

LE SELECTIONNEUR FRANÇAIS

**BULLETIN DE L'ASSOCIATION
DES SÉLECTIONNEURS FRANÇAIS**

“ Apport des marqueurs moléculaires et de la génomique à l'utilisation des ressources génétiques pour l'amélioration des plantes ”

La diversité génétique des plantes cultivées à l'heure de l'information génétique à haut débit : génomique comparative de la domestication Jacques David, Sylvain Glémin & coll.	3
Pré-sélection stratégique pour l'amélioration du blé Graham Moore	15
Re-synthèse d'espèces : l'exemple du colza Anne-Marie Chèvre, Alexandre Pelé, Sophie Paillard, Mathieu Rousseau-Gueutin	21
L'apport des marqueurs moléculaires et de la génomique pour l'utilisation des ressources génétiques chez la tomate Christopher Sauvage, Guillaume Bauchet et Mathilde Causse	29
Des espèces sauvages à la sélection génomique : l'exemple de la betterave Karine Bounan-Henry et Bruno Desprez	41
Gestion et utilisation des ressources génétique chez le maïs Alain Charcosset, Brigitte Gouesnard, Laurence Moreau, Stéphane Nicolas, Anne Zanetto, Jacques Laborde, Carine Palaffre et Cyril Bauland	57
En guise de conclusion : vers une nouvelle conception du pré-breeding Michel Renard.....	69

LA DIVERSITE GENETIQUE DES PLANTES CULTIVEES A L’HEURE DE L’INFORMATION GENETIQUE A HAUT DEBIT : GENOMIQUE COMPARATIVE DE LA DOMESTICATION

Jacques DAVID¹, Sylvain GLEMIN² &

Gautier Sarah¹, Benoît Nabholz², François Sabot³, Yan Holtz¹, Felix Homa¹, Stephanie Pointet, Sandy Contreras¹, Laure Sauné, Morgane Ardisson¹, Roberto Bacilieri¹, Bouchaib Khadari⁴, Claire Lanaud¹, David Pot¹, Christopher Sauvage⁵, Nora Scarcelli³, James Tregear³, Yves Vigouroux³, Nabila Yahiaoui¹, Manuel Ruiz¹, Sylvain Santoni¹, Jean-Pierre Labouisse¹, Jean-Louis Pham³

¹ UMR AGAP, Montpellier Supagro-INRA-CIRAD, ² UMR ISEM-CNRS, ³ UMR DIADE-IRD,

⁴ Conservatoire Botanique National, ⁵ UR GAFL, INRA

Email : jacques.david@supagro.fr

1 - LES RESSOURCES GENETIQUES ET L’HISTOIRE DES ESPECES CULTIVEES

Les ressources génétiques ont une acception très large. Pour le sélectionneur, elles sont vues comme la diversité nécessaire à l’introduction de la variabilité génétique indispensable pour améliorer ses variétés, leur donner une plus grande rusticité ou des tolérances à différents stress, développer des produits d’usage spécifique ; la gamme des mises en œuvre de la diversité des plantes dans les variétés est très large.

Pour autant, la diversité mesurée pour une espèce à un instant de son histoire n’est pas figée, elle résulte d’une histoire évolutive complexe, au cours de laquelle ont interagi les pratiques de culture, de sélection, d’adaptation aux territoires, le hasard, les échanges de toutes sortes et ce pendant plusieurs millénaires depuis la domestication.

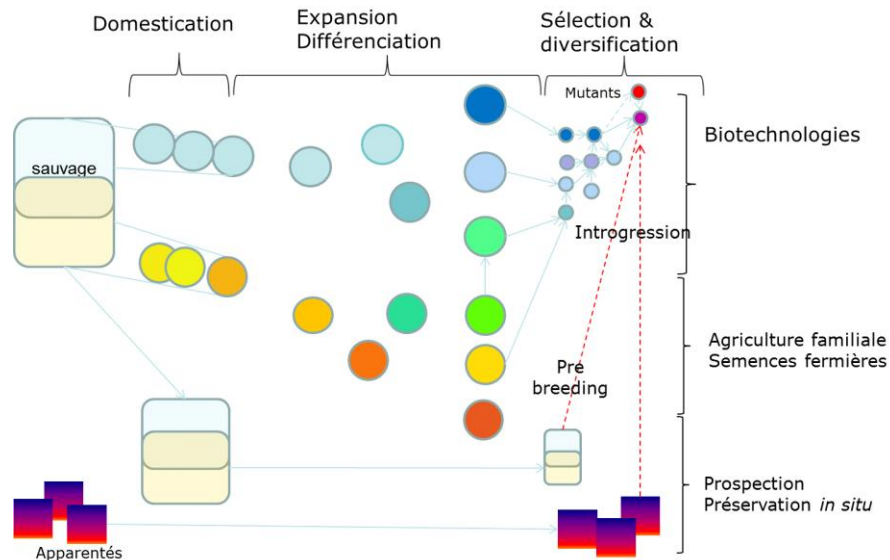


Figure 1 : Représentation schématique de l’évolution des compartiments cultivés et sauvages d’une espèce domestique. La diversité évolue en fonction des événements démographiques et sélectifs qui jalonnent l’histoire du compartiment cultivé. Ce schéma peut être modifié assez profondément en fonction de l’espèce considérée, notamment pour les espèces pérennes ou récemment domestiquées.

Nombre d’auteurs se sont penchés sur cette histoire et les relations qu’ont entretenues l’homme et les plantes cultivées, voire plus généralement les organismes domestiqués (Darwin, 1868 ; Candolle, 1883 ; Harlan, 1992). De manière générale, les ressources génétiques sont souvent caractérisées par leur éloignement du compartiment élite cultivé aujourd’hui (Buckler *et al.*, 2001). Elles sont les descendants actuels de formes plus anciennement cultivées, mais aussi de formes sauvages présentes au moment de la domestication. Un schéma très simplifié qui conviendrait à poser un cadre historique pour un grand nombre d’espèces, surtout de grande culture, est donné à la figure 1, bien qu’il existe de nombreuses déviations à cette représentation, notamment pour les espèces pérennes, ou les espèces à domestication quasi immédiate telles que des espèces polyploïdes comme le bananier (Perrier *et al.*, 2011). Un compartiment sauvage soumis à une agriculture pré-domestique (Tanno and Willcox, 2006) durant quelques siècles, dans une ou plusieurs zones d’origine, donne les premières formes cultivées qui vont ensuite se disperser et se différencier dans un espace géographique et temporel plus ou moins large selon les espèces. Certaines vont se déplacer à des latitudes homogènes comme le blé tandis que d’autres vont partir des tropiques et changer de latitude comme le maïs ou la pomme de terre.

Ces deux événements, domestication et expansion, fondent en général la base de la diversité utilisée par la sélection actuelle. La domestication, parce qu’elle est l’étape à travers laquelle s’est transmise la diversité initialement présente dans la forme sauvage, et l’expansion, car elle a créé une large diversité de formes dans l’immense espace (en termes de durée et de superficie) de l’agriculture traditionnelle. La mise en relation du Nouveau et de l’Ancien Monde a aussi créé de grandes modifications dans les patrons de diversité. Enfin les pratiques de sélection moderne, délibérées, basées à la fois sur une exploitation efficace de la diversité intra-variétale des populations traditionnelles par l’intercroisement réfléchi et contrôlé des meilleurs parents, et sur l’expérimentation et la diffusion de variétés beaucoup plus homogènes que les formes traditionnelles vont grandement modifier les patrons de diversité des formes actuelles.

Lors de ces grandes transitions, les génomes ont subi des pressions de sélection visant à adapter les variétés aux besoins des cultures. Elles ont pu perdre nombre d’aptitudes importantes pour une plante sauvage mais gênantes en agriculture comme la dissémination spontanée, ou la dormance (Nave *et al.*, 2016; Purugganan and Fuller, 2009 ; Glemin and Bataillon 2009 ; Ross-Ibarra *et al.*, 2007) et gagner des adaptations favorables à l’intensification des pratiques, comme des statures demi-naines pour les céréales à paille ou une accumulation de résistance aux maladies. Ces pressions de sélection ont été responsables d’effets de balayage sélectif plus ou moins prononcés : la sélection des allèles favorables à un locus pouvant avoir un effet d’entraînement sur les autres locus situés à proximité, neutres ou dont la valeur sélective était moins importante localement (Smith and Haigh, 1974). La sélection interagit alors avec la dérive génétique et vient réduire localement la diversité de gènes non touchés directement par la sélection. Par exemple, en séquençant le génome de la téosinte et du maïs domestique, des premières estimations pan-génomiques des transitions démographiques lors de la domestication montrent que la réduction de la taille efficace a été très intense au départ : seuls 5 % de la population de la téosinte a été mise à contribution pour la fondation du compartiment domestique tandis qu’une rapide et intense expansion démographique du maïs a suivi, permettant d’atteindre des tailles efficaces bien supérieures à celles connues par son ancêtre téosinte (Beissinger *et al.*, 2016).

Ainsi la littérature rapporte souvent que les plantes cultivées actuelles montrent moins de diversité moléculaire que les formes ancestrales (Glemin and Bataillon, 2009) mais une compensation s’observe fréquemment entre niveau de diversité (ou originalité génétique) et performances agronomiques.

2 - UTILISATION DE LA DIVERSITE GENETIQUE POUR DISSEQUER LES CARACTERES

2.1. Les panels de génétique d’association

L’accès au génome par les outils de séquençage a relancé l’intérêt des collections de ressources

génétiques. Des panels pour mener des études de génétique d’association en ont été sortis pour un très grand nombre d’espèces (Glaszmann *et al.*, 2010). La recherche sur l’architecture génétique des caractères a été révolutionnée par cette approche et l’utilisation des techniques modernes (Rafalski, 2002 ; Kilian and Graner, 2012). Ces panels apportent plusieurs allèles possibles par locus et l’histoire de leur recombinaison est beaucoup plus riche et complexe que celle découlant d’un croisement biparental simple suivi d’une fixation. Ces panels sont ainsi devenus de formidables plateformes pour la recherche d’associations entre polymorphismes et caractères quantitatifs. La plus grande difficulté étant d’assembler un panel adapté à la question posée (Glaszmann *et al.*, 2010). Nous ne rentrerons pas ici dans le détail des possibilités permises par les panels de génétique d’association qui a fait l’objet d’un grand nombre de travaux. Cependant, si la diversité présente au sein des panels est un atout, la complexité de l’histoire et de l’origine du matériel nécessite de prendre de multiples précautions dans l’interprétation des associations trouvées positives dans la mesure où elles peuvent être le fait d’une structure cachée du matériel (Korte and Farlow, 2013).

2.2. Détection des zones chromosomiques impliquées dans l’adaptation des formes cultivées

Avec le développement de la connaissance sur les génomes, de nombreuses statistiques de diversité et de neutralité par locus peuvent être calculées et visualisées le long des chromosomes pour réaliser de véritables scans génomiques (Siol *et al.*, 2010). Les empreintes laissées dans les génomes par les épisodes sélectifs majeurs se caractérisent, parmi d’autres statistiques pouvant être calculées à partir de données de polymorphismes, par trois propriétés essentielles : des pertes de diversité locale (mesurée par la diversité de Nei ou bien la diversité nucléotidique) plus importantes que celles subies par le génome en moyenne lors des épisodes de goulots démographiques, une augmentation du déséquilibre de liaison qui signe que la rapidité de l’évènement n’a pas laissé le temps à la recombinaison de rétablir l’équilibre des fréquences haplotypiques à proximité des gènes ayant subi la sélection, et enfin, une différenciation accentuée par rapport à la forme sauvage par la mesure de critères comme le *Fst* par exemple. Sur le soja, ces trois signaux concomitants ont permis d’identifier des zones soumises à sélection de manière très nette (Li *et al.*, 2013).

Ces « scans » génomiques donnent accès à des informations sans *a priori* sur les caractères concernés et permettent d’identifier les zones des génomes ayant connu le plus grand nombre de modifications (Vigouroux *et al.*, 2002 ; Wright *et al.*, 2005). Par exemple, sur le maïs, des estimations donnent deux à quatre pour cent des locus sous sélection lors de la transition sauvage vers domestique (Wright *et al.*, 2005). Très récemment, il a été suggéré que l’impact de ces sélections de quelques gènes à effets forts reste difficile à montrer chez le maïs (Beissinger *et al.*, 2016). D’autres travaux sur le haricot ont aussi montré que des échanges peuvent s’instaurer entre sauvage et cultivé et qu’une sélection différentielle, humaine d’un côté et naturelle de l’autre, crée des îlots de différenciation qui peuvent aider à pointer les zones dans lesquelles les gènes de domestication se trouvent (Papa and Gepts, 2003).

2.3. Interprétation du polymorphisme au niveau moléculaire

Si de nombreuses avancées ont eu lieu grâce à la production de marqueurs bon marché, nombreux et faciles à utiliser, elles sont souvent basées sur des polymorphismes déjà identifiés dans des panels de référence. L’information complète de séquence, via les méthodes modernes de séquençage dites de nouvelle génération, permet d’aller plus loin dans l’interprétation des patrons de polymorphisme entre les formes sauvages et cultivées. En accédant aux séquences codantes des gènes, il est possible de différencier la dynamique d’évolution des fréquences des mutations qui modifient les protéines (mutations non synonymes) et dont l’effet attendu est plutôt une altération de la fonction de la protéine (effet légèrement délétère), de celle des mutations qui ne modifient pas la protéine (mutations synonymes) dont on attend qu’elles soient *a priori* neutres en première approximation (Figure 2).

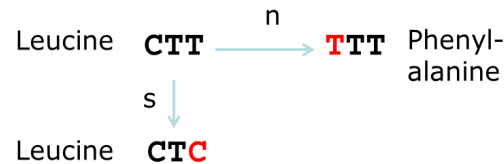


Figure 2 :Mutations synonymes (s) et non synonymes (n). Le codon initial CTT code pour une leucine. En raison du code génétique dégénéré, une mutation qui touche la troisième base du codon, (ici T mute vers C) ne change pas l’acide aminé. Cette mutation est dite synonyme (s). En revanche, un mutant C vers T à la première base du codon modifie l’acide aminé codé et donc est non synonymes (n) vis-à-vis de la protéine.

Ainsi les polymorphismes établis dans les gènes peuvent être utilisés pour calculer deux indices de polymorphisme, la diversité nucléotidique synonyme (π_s) et la diversité non synonyme (π_n). Sous l’hypothèse que la mutation fait entrer dans la population des mutations des deux formes, une façon de mesurer l’efficacité de la sélection est de comparer ces deux niveaux de polymorphisme en posant comme hypothèses que la plupart des mutations non synonymes sont légèrement délétères et que la sélection maintient leur fréquence à de faible niveau. Si la sélection est globalement attendue plus efficace dans une population que dans une autre, alors le ratio des polymorphismes π_n/π_s doit être plus faible dans la population où la sélection est la plus efficace, car elle maintient les mutations délétères à des fréquences plus faibles.

Il est évident que, de temps à autre, des mutations non synonymes peuvent se révéler avantageuses et qu’elles permettent l’adaptation. Dans ce cas, si elles échappent à la dérive lorsqu’elles sont encore rares, leur destinée n’est pas de rester polymorphes mais au contraire de se fixer rapidement en fonction de leur gain de valeur sélective. Reste alors le cas où ces mutations sont bénéfiques dans un environnement et pas dans un autre, ou celui où elles sont maintenues par des mécanismes de fréquence-dépendance. Malgré ces possibilités de sélection positive ou stabilisante, plusieurs études sont venues démontrer qu’utiliser le ratio π_n/π_s est un bon indicateur de l’efficacité de la sélection (Ellegren and Galtier, 2016).

Lorsque la sélection n’est plus assez efficace, la dérive peut entraîner l’augmentation en fréquence de mutations légèrement défavorables, voire permettre leur fixation. Ce phénomène est une des explications données de la vigueur hybride lorsque sont croisés des parents ayant divergé de la même source mais ayant accumulé des mutations différentes. Si chaque mutation a un petit effet, les effets se cumulent sur l’ensemble du génome et finissent par peser sur la valeur sélective de la population, c’est le fardeau génétique.

Cette sélection contre les allèles faiblement défavorables à effets cumulatifs dépend essentiellement de la taille efficace de la population (Ne). Plus la taille efficace est élevée, plus la sélection va pouvoir maintenir à faible fréquence des allèles à effets délétères faibles. Une autre façon de voir les choses est de regarder la valeur seuil de n pour laquelle la sélection reste efficace. Si Ne est petit, la dérive agit et fait augmenter en fréquence des mutations dont l’effet défavorable est plus grand que lorsque Ne est élevé. Conséquemment, les fréquences des allèles délétères vont pouvoir prendre des valeurs plus élevées dans les petites populations que dans les grandes, et ainsi il est attendu que π_n/π_s augmente lorsque Ne diminue.

Une transition de ce type est donc attendue lors de la domestication et dans la suite de l’histoire du matériel cultivé puisqu’elle se manifeste souvent par une réduction de l’effectif efficace au début de la domestication. Ceci aboutit à émettre l’hypothèse d’un coût de la domestication en terme d’accumulation de fardeau génétique. Il serait logique que cet effet « d’autostop génétique » de ces allèles délétères soit plus fort chez les espèces autogames qui recombinent moins (Nabholz *et al.*, 2014). Etudier ce phénomène en comparant différents couples d’espèces sauvages et cultivées

permettrait d'en donner la valeur générale et de mieux saisir l'impact des facteurs évolutifs sur le polymorphisme au-delà des gènes de domestication à fort effets, détectés par les analyses QTL ou en génétique d'association.

A l'heure actuelle, les effets de la domestication ont été évalués avec des données moléculaires sur un petit nombre d'espèces, majoritairement des espèces de grande culture. Seules quelques espèces pérennes comme le pommier (Cornille *et al.*, 2012) ont été étudiées et donnent des patrons relativement différents de ce qui a été observé sur les espèces annuelles. La longévité, la capacité de s'hybrider avec le compartiment sauvage, le faible nombre de générations séparant domestique et sauvage pourraient expliquer l'absence de goulot d'étranglement. Il est intéressant de prendre du recul et de documenter les effets de la domestication sur un nombre plus large d'espèces, si possible en suivant le même protocole et en les analysant de la même manière. C'est l'objet du projet ARCAD.

3 - ARCAD : UNE ETUDE COMPARATIVE DE LA DOMESTICATION

La communauté montpelliéraine comprend des équipes travaillant sur la diversité génétique de nombreuses plantes cultivées, méditerranéennes et tropicales. Elle possédait donc les ressources idéales pour effectuer une analyse comparative de la domestication en suivant un protocole standardisé et en utilisant les technologies de séquençage NGS. Ce projet a été soumis à la Fondation Agropolis qui l'a soutenu (<http://www.arcad-project.org/>). Comme beaucoup d'espèces ciblées ne possédaient pas de génome de référence au démarrage du projet, une approche par le RNAseq a été privilégiée pour obtenir les données sur une fraction réduite et répétable du génome. Elle permettait également en accédant à l'espace codant de pouvoir documenter les polymorphismes affectant les gènes.

Parmi les questions posées par cette étude figuraient les suivantes : la diversité du compartiment sauvage est-elle comparable entre espèces annuelles, pérennes, autogames ou allogames ? Les pertes de diversité lors du goulot de domestication varient-elle en fonction des traits d'histoire de vie des espèces (autogamie vs. allogamie, pérennité vs. annualité) ? Peut-on tester l'hypothèse du coût de la domestication au travers d'une accélération du taux de fixation des mutations non synonymes ?

3.1. Le projet

Tableau 1 : Espèces prises en compte dans le projet ARCAD, réparties selon des critères taxonomiques, de traits d'histoire de vie et de régime de reproduction. Les espèces marquées d'un * ont bénéficié de l'infrastructure ARCAD pour la production et l'analyse bio-informatique des données mais ont été financées par d'autres voies que le projet ARCAD pour la production des échantillons et le séquençage.

Espèce	Taxonomie	Statut	Régime de reproduction
Bananier	Monocot	Pérenne	Allogame
Palmier	Monocot	Pérenne	Allogame
Igname	Monocot	Annuelle	Allogame/Clonale
Engrain	Monocot	Annuelle	Autogame
Mil	Monocot	Annuelle	Allogame
Riz Africain	Monocot	Annuelle	Autogame
Sorgho	Monocot	Annuelle	Autogame
Fonio*	Monocot	Annuelle	Autogame
Cacaoyer	Dicot	Pérenne	Allogame
Caféier	Dicot	Pérenne	Allogame
Olivier*	Dicot	Pérenne	Allogame
Vigne	Dicot	Pérenne	Allogame
Tomate	Dicot	Annuelle	Autogame

Pour chacune des espèces, une dizaine de représentants sauvages et autant de cultivés ont été échantillonnés sur plusieurs tissus (fleurs et feuilles lorsque cela était possible) et autant de banques d’ADNc ont été élaborées et séquencées (Nabholz *et al.*, 2014). A ces paires de taxons étaient associées au moins deux espèces phylogénétiquement distantes afin de pouvoir disposer de groupes externes pour l’orientation des SNP en état ancestral et dérivé.

Le collectif mobilisé est large et recoupe plusieurs institutions de recherche créant de ce fait une animation scientifique et méthodologique importante. Au total, ce sont 332 banques de cDNA qui ont été produites sur 45 espèces. On a ainsi produit plus de 13 milliards de fragments de cDNA séquencés à leurs extrémités sur 100bp par la technologie Illumina (read pairs). Leur analyse a nécessité l’écriture de pipe-lines bio-informatiques spécifiques (Gautier *et al.*, sous presse ; Nabholz *et al.*, 2014).

Ces fragments ont été alignés, soit sur les génomes de référence lorsqu’ils étaient disponibles, soit sur des transcriptomes assemblés *de novo* (Gautier *et al.*, *in press*). Le pipe-line informatique a permis de produire les résultats figurant dans le tableau 2. La stratégie RNAseq s’est révélée pertinente avec les progrès réalisés ces dernières années sur les NGS. Plusieurs milliers de contigs ont pu être reconstruits pour chacune des espèces et la diversité a pu être ainsi caractérisée pour l’ensemble des espèces

3.2. Polymorphismes détectés

Tableau 2 : Polymorphismes détectés au sein des différentes espèces du dispositif ARCAD dans leurs compartiments sauvage et domestique. (N : effectif ; # Contigs : nombre de contigs sur lesquels les reads ont pu être cartographiés et des polymorphismes détectés ; SNP : Single Nucleotid polymorphism).

Espèce	Compartiment sauvage			Compartiment cultivé		
	N	# Contig	SNP	N.	# Contigs	SNP
<i>Coffea canephora</i>	12	14 581	148 508	12	15 722	49 790
<i>Dioscorea rotundata</i>	10	15 981	139 368	9	16 708	80 335
<i>Eleais guineensis</i>	10	16 975	54 238	10	17 220	51 722
<i>Musa acuminata</i>	10	21 422	286 784	10	277 146	274 179
<i>Olea europa</i>	10	19 549	209 402	10	21 168	168 372
<i>Oryza glaberrima</i>	9	12 169	23 578	7	12 169	7 597
<i>Pennisetum glaucum</i>	10	15 714	143 297	10	17 586	94 058
<i>Solanum lycopersicum</i>	10	18 887	73 778	10	16 396	40 189
<i>Sorghum bicolor</i>	10	11 397	46 298	10	13 231	41 266
<i>Theobroma cacao</i>	10	13 901	83 613	10	13 899	76 114
<i>Triticum monococcum</i>	10	5 847	11 148	10	14 066	20 776
<i>Vitis sylvestris</i>	12	14 397	133 697	12	16 222	195 603

Parmi les espèces dont les représentants sauvages étaient les plus polymorphes figurent le bananier diploïde (*Musa acuminata*) et l’olivier (*Olea europa*) tandis que le riz africain (*Oryza glaberrima*) et l’engrain (*Triticum monococcum*) ont les plus faibles nombres de SNP identifiés. Cela peut pour partie être aussi du au moment auquel ont été produites les données, les plus tardives ayant bénéficié des progrès des séquenceurs et ayant donc bénéficié d’une couverture plus importante.

3.3. Taille efficace et fardeau de mutation chez les formes sauvages

Comme nous l’avons introduit plus haut, des populations de grande taille efficace devraient avoir une meilleure efficacité vis-à-vis de la sélection et notamment vis-à-vis de l’élimination des mutations ayant des effets défavorables. Il est par ailleurs attendu que dans de grandes populations, la diversité nucléotidique synonyme soit plus élevée que dans de petites populations. Si le taux de mutation par paire de base est comparable entre espèces, alors comparer les taux de diversité revient à comparer les tailles efficaces. Ainsi, il est possible de tester l’attendu théorique que les populations à fort taux de polymorphisme (grand N_e) aient aussi des ratios π_n/π_s plus faibles (sélection efficace) que des populations à faible N_e .

Pour l’ensemble des espèces sauvages d’ARCAD la relation est donnée figure 3. La relation négative entre les deux paramètres apparaît très nettement entre π_s et le ratio π_n/π_s . Ainsi, le niveau et la composition du polymorphisme de départ des différentes espèces n’est pas comparable. Dans les espèces pour lesquelles il y a peu de polymorphisme, il est probable qu’une partie importante des variants non synonymes soient défavorables et qu’ils aient augmenté en raison de la dérive, au moins dans les conditions écologiques de la forme sauvage, un allèle délétère dans la forme sauvage pouvant devenir favorable dans la forme cultivée. La situation particulière du blé diploïde sauvage pourrait être due à une couverture moins bonne des transcriptomes ou bien à une évolution récente de sa taille efficace.

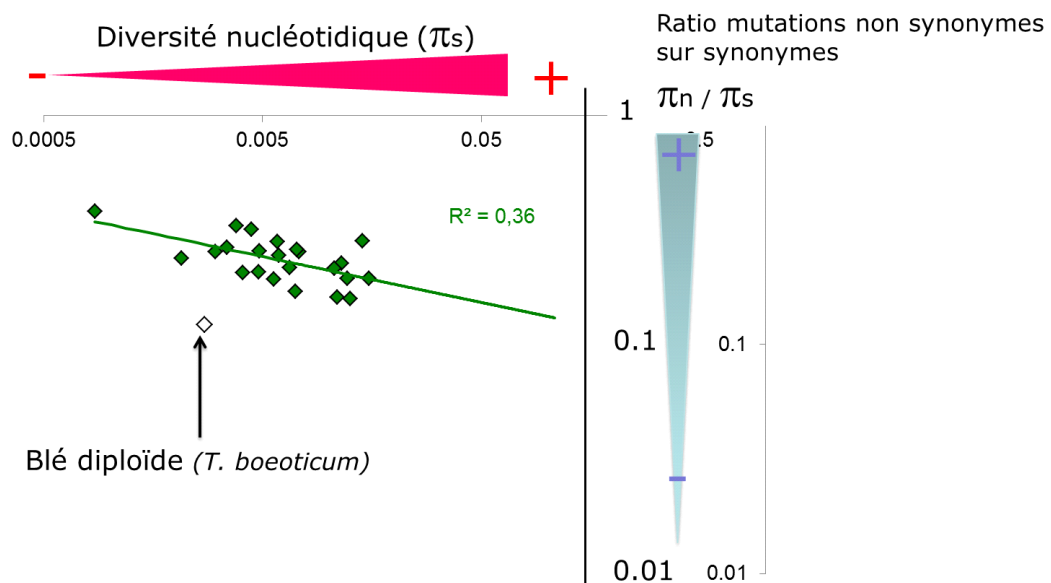


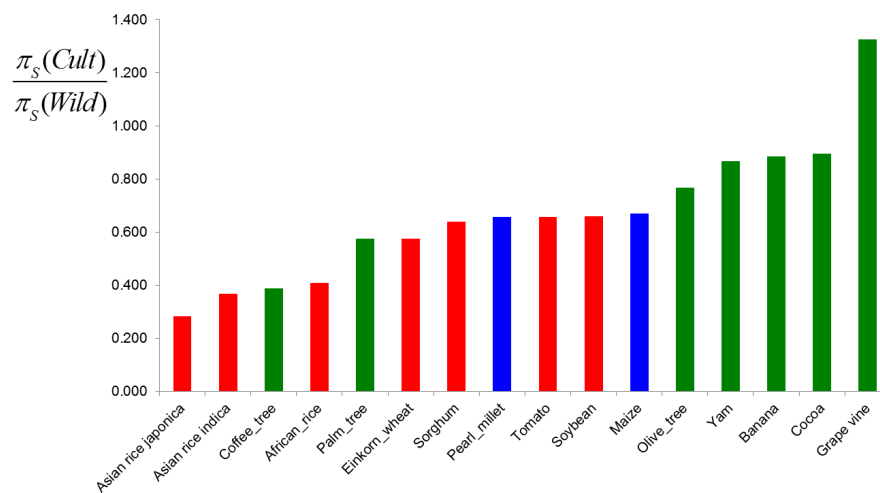
Figure 3 : Impact de la taille efficace (approximée par le taux de variation nucléotidique synonyme π_s) sur l’efficacité de la sélection des variants défavorables (approximé par le ratio π_n/π_s). Chaque point est un échantillon du dispositif Arcad.

3.4. Goulots d’étranglement

Les pertes de diversité lors du passage de la forme sauvage à la forme cultivée sont très variables (Figure 4). Elles sont de l’ordre de 60 à 70 % pour les riz asiatiques (*Oryza sativa ssp. japonica* et *indica*, *Oryza glaberrima*) et pour le café diploïde (*Coffea canephora*). Elles sont très réduites pour le cacaoyer (*Theobroma cacao*), le bananier diploïde (*Musa acuminata*), pour l’igname (*Dioscorea rotundata*) et l’olivier pour lesquels la perte de diversité est de 20 %. Les autres espèces se situent dans une zone intermédiaire de 40 % de perte environ. La vigne a un statut particulier. Pour notre

échantillon, la forme cultivée est plus diverse que les lambrusques. Cela interroge sur la représentativité de notre échantillon ou sur les processus qui ont mené à la création du pool domestique. L’Est du bassin méditerranéen et le Caucase n’ont probablement pas été bien échantillonnés ou il est aussi possible que des introgressions secondaires à la domestication aient eu lieu avec une espèce non encore identifiée (R. Bacilieri, com. pers.).

A l’instar de ce qui avait déjà été décrit chez le pommier (Cornille *et al.*, 2012) il est remarquable de constater que les espèces pérennes ont globalement moins perdu de diversité que les espèces annuelles si on excepte les situations particulières du café et du palmier à huile. Les formes domestiques de ces deux espèces sont probablement plus récentes en comparaison des autres espèces. Il est possible que pour ces espèces la domestication soit plutôt à assimiler à des épisodes de sélection consciente assez drastique, consistant à un échantillonnage conscient de plantes rapidement isolées génétiquement de leur compartiment sauvage et ne pouvant bénéficier de réintroduction régulière de diversité via des événements d’interfécondations.



* (Caicedo et al. 2007), ** (Lam et al. 2010), *** (Wright et al. 2005)

Figure 4 : Perte de diversité de plusieurs plantes domestiques lors de la transition sauvage-cultivé mesurée dans le dispositif Arcad. Les données du maïs, du soja et des riz asiatiques prises dans la littérature ont été rajoutées. L’intensité des goulots d’étranglement est mesurée par le ratio de la diversité nucléotidique synonyme de la forme cultivée rapportée à celle du compartiment sauvage. En rouge figurent les espèces annuelles, en vert les espèces pérennes.

3.5. Coût de la domestication

Pour les espèces pour lesquelles le goulot est fort, nous avons cherché à savoir si le taux de fixation de mutations non synonymes avait été accéléré dans le compartiment domestique en comparaison du compartiment sauvage. Sur le riz africain, pour lequel la diversité nucléotidique synonyme ($\pi_s = 0.0006$ per site) apparaît une des plus faible jamais rencontrée, le goulot a été extrêmement fort (Nabholz *et al.*, 2014). A partir des séquences des deux espèces externes (*Oryza meridionalis* et *Oryza sativa*), il a pu être prédit si les allèles observés chez le sauvage ou le cultivé avait un statut ancestral (car identique à celui des espèces externes) ou dérivé (car divergent de celui des espèces externes). En cumulant l’information sur l’ensemble des sites polymorphes, deux taux de divergence ont été calculés pour chacune des deux formes de riz africain, cultivé et sauvage : le taux non-synonyme (Dn) et le taux synonymes (Ds). Leur ratio Dn/Ds peut donner une indication sur la vitesse de fixation des mutations synonymes et non synonymes dans le compartiment cultivé comparativement au compartiment sauvage. Cela a été fait pour deux fractions de gènes, faiblement ou

fortement exprimés. Les résultats (Figure 5) montrent clairement que le taux de fixation de mutation non synonymes rapporté au taux de fixation synonymes (Dn/Ds) est beaucoup plus élevé dans la forme cultivée que dans la forme sauvage et ce notamment pour les gènes fortement exprimés. L’hypothèse d’un coût de domestication est donc très fortement soutenue par les données.

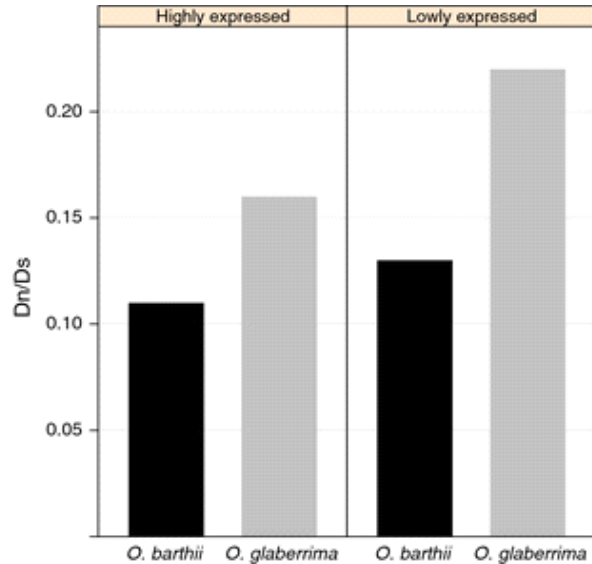


Figure 5 : Illustration du coût de la domestication chez le riz africain, *Oryza glaberrima*. Sur la forme sauvage (*O. barthii*), le taux de fixation de mutation non synonymes est très inférieur à celui observé sur la forme cultivée (d’après Nabholz *et al.*, 2014).

Nous avons effectué ce même calcul pour six autres espèces : trois annuelles (sorgho, blé et mil) et trois pérennes (bananier, olivier, cacaoyer). L’hypothèse de coût de la domestication est significative sur le blé et le mil mais n’est pas décelable sur les trois pérennes.

4 - CONCLUSIONS ET PERSPECTIVES

L’étude comparée de la domestication apporte de précieux éléments pour la compréhension de la dynamique de la diversité des plantes cultivées. Suivant leurs traits d’histoire de vie, l’ancienneté de leur domestication, leur régime de reproduction... la diversité génétique est répartie différemment et l’appel aux ressources des plantes sauvages doit être discutée au cas par cas, notamment lorsque la diversité du compartiment domestique n’a pas été réduite de manière drastique. La domestication ne découle pas d’un modèle unique et n’a pas eu les mêmes conséquences, la signification des polymorphismes n’est pas identique. Les ratios des mutations synonymes et non synonymes montrent que la sélection joue un rôle important dans leur répartition en interaction avec la dérive génétique.

La démarche décrite pour le projet ARCAD s’est révélée efficace, l’utilisation du RNAseq pour accéder à une sous partie du génome, même pour des espèces complexes, est envisageable et ARCAD a produit de nombreuses ressources bio-informatiques pour pouvoir utiliser relativement facilement ce type de données. Les données ARCAD sont publiques ; elles constituent d’importantes ressources pour l’utilisation des ressources génétiques comme les bases de données de polymorphismes ou les transcriptomes de référence assemblés *de novo*. Elles peuvent d’ores et déjà servir à établir des outils de génotypage à haut débit. Il est possible de procéder à l’exploration des séries alléliques de gènes candidats ou à la caractérisation de l’évolution des familles multigéniques.

Les données très riches sont encore à exploiter pour analyser les conditions dans lesquelles il sera possible de détecter les empreintes de sélection ; nos résultats sur le riz africain suggèrent qu’un trop fort goulot d’étranglement gêne considérablement leur détection.

Pour les sélectionneurs, il est important de prendre en compte l’hypothèse du coût de la domestication sur leur matériel et donc de s’intéresser à la nature des polymorphismes qui sont détectés et utilisés. L’apport de l’évolution moléculaire est précieux et devrait être beaucoup plus systématiquement pris en compte.

L’étude de l’évolution des patrons d’expression des gènes est également permise par les données RNAseq et les données s’accumulent pour suggérer que la diversité des patrons d’expression a elle aussi été touchée par les phénomènes de sélection pour les formes domestiques (Bellucci *et al.*, 2014).

“Journée ASF du 4 février 2016”
“ Apport des marqueurs moléculaires et de la génomique à l’utilisation
des ressources génétiques pour l’amélioration des plantes ”

REFERENCES BIBLIOGRAPHIQUES

- Beissinger TM, Wang L, Crosby K, Durvasula A, Hufford MB, Ross-Ibarra J (2016) Recent demography drives changes in linked selection across the maize genome. *Nature plants* 2:16084. doi:10.1038/nplants.2016.84
- Bellucci E, Bitocchi E, Ferrarini A, Benazzo A, Biagetti E, Klie S, Minio A, Rau D, Rodriguez M, Panziera A, Venturini L, Attene G, Albertini E, Jackson SA, Nanni L, Fernie AR, Nikoloski Z, Bertorelle G, Delledonne M, Papa R (2014) Decreased Nucleotide and Expression Diversity and Modified Coexpression Patterns Characterize Domestication in the Common Bean. *The Plant cell* 26 (5):1901-1912. doi:10.1105/tpc.114.124040
- Buckler ESt, Thornsberry JM, Kresovich S (2001) Molecular diversity, structure and domestication of grasses. *Genetical research* 77 (3):213-218
- Candolle Ad (1883) *Origine des plantes cultivées*. Bibliothèque scientifique internationale,, vol XLIII. G. Baillièrre et cie, Paris,
- Cornille A, Gladieux P, Smulders MJ, Roldan-Ruiz I, Laurens F, Le Cam B, Nersesyan A, Clavel J, Olonova M, Feugey L, Gabrielyan I, Zhang XG, Tenaillon MI, Giraud T (2012) New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS genetics* 8 (5):e1002703. doi:10.1371/journal.pgen.1002703
- Darwin C (1868) *The variation of animals and plants under domestication*. Authorized edn. O. Judd & company, New York,
- Ellegren H, Galtier N (2016) Determinants of genetic diversity. *Nature reviews Genetics* 17 (7):422-433. doi:10.1038/nrg.2016.58
- Glaszmann JC, Kilian B, Upadhyaya HD, Varshney RK (2010) Accessing genetic diversity for crop improvement. *Current opinion in plant biology* 13 (2):167-173. doi:10.1016/j.pbi.2010.01.004
- Glemin S, Bataillon T (2009) A comparative view of the evolution of grasses under domestication. *The New phytologist* 183 (2):273-290. doi:10.1111/j.1469-8137.2009.02884.x
- Harlan JR (1992) *Crops & man*. 2nd edn. American Society of Agronomy : Crop Science Society of America, Madison, Wis., USA
- Kilian B, Graner A (2012) NGS technologies for analyzing germplasm diversity in genebanks. *Briefings in functional genomics* 11 (1):38-50. doi:10.1093/bfpg/elr046
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant methods* 9:29. doi:10.1186/1746-4811-9-29
- Li YH, Zhao SC, Ma JX, Li D, Yan L, Li J, Qi XT, Guo XS, Zhang L, He WM, Chang RZ, Liang QS, Guo Y, Ye C, Wang XB, Tao Y, Guan RX, Wang JY, Liu YL, Jin LG, Zhang XQ, Liu ZX, Zhang LJ, Chen J, Wang KJ, Nielsen R, Li RQ, Chen PY, Li WB, Reif JC, Purugganan M, Wang J, Zhang MC, Wang J, Qiu LJ (2013) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. *BMC genomics* 14:579. doi:10.1186/1471-2164-14-579

- Nabholz B, Sarah G, Sabot F, Ruiz M, Adam H, Nidelet S, Ghesquiere A, Santoni S, David J, Glemin S (2014) Transcriptome population genomics reveals severe bottleneck and domestication cost in the African rice (*Oryza glaberrima*). *Molecular ecology* 23 (9):2210-2227. doi:10.1111/mec.12738
- Nave M, Avni R, Ben-Zvi B, Hale I, Distelfeld A (2016) QTLs for uniform grain dimensions and germination selected during wheat domestication are co-located on chromosome 4B. *Theoretical and applied genetics* 129 (7):1303-1315. doi:10.1007/s00122-016-2704-4
- Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theoretical and applied genetics* 106 (2):239-250. doi:10.1007/s00122-002-1085-z
- Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, Bakry F, Carreel F, Hippolyte I, Horry JP, Jenny C, Lebot V, Risterucci AM, Tomekpe K, Doutrelepont H, Ball T, Manwaring J, de Maret P, Denham T (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences of the United States of America* 108 (28):11311-11318. doi:10.1073/pnas.1102001108
- Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature* 457 (7231):843-848. doi:10.1038/nature07895
- Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. *Current opinion in plant biology* 5 (2):94-100
- Ross-Ibarra J, Morrell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences of the United States of America* 104 Suppl 1:8641-8648. doi:10.1073/pnas.0700643104
- Siol M, Wright SI, Barrett SC (2010) The population genomics of plant adaptation. *The New phytologist* 188 (2):313-332. doi:10.1111/j.1469-8137.2010.03401.x
- Smith JM, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetical research* 23 (1):23-35
- Tanno K, Willcox G (2006) How fast was wild wheat domesticated? *Science* 311 (5769):1886. doi:10.1126/science.1124635
- Vigouroux Y, McMullen M, Hittinger CT, Houchins K, Schulz L, Kresovich S, Matsuoka Y, Doebley J (2002) Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proceedings of the National Academy of Sciences of the United States of America* 99 (15):9650-9655. doi:10.1073/pnas.112324299
- Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS (2005) The effects of artificial selection on the maize genome. *Science* 308 (5726):1310-1314. doi:10.1126/science.1107891

PRE-SELECTION STRATEGIQUE POUR L'AMELIORATION DU BLE¹

Graham MOORE

Co-ordinator of the Wheat Improvement Strategic Programme

John Innes Centre

Colney Lane, Norwich, Royaume-Uni

E-mail : graham.moore@jic.ac.uk

RESUME

Augmenter le rendement des cultures nécessite de collecter puis d'exploiter toute la diversité génétique conservée, bien plus large que celle existante au sein de la gamme étroite des variétés couramment cultivées. Une telle entreprise nécessite un partenariat entre la recherche académique et l'industrie.

INTRODUCTION

Le blé est l'une des trois céréales les plus cultivées dans le monde avec une récolte de 713 millions de tonnes en 2013. La demande en blé est telle que la production cumulée des 50 prochaines années devra dépasser le total de ce qui a été produit depuis les débuts de l'agriculture il y a 10.000 ans. Cependant, les rendements en blé progressent actuellement à des taux annuels inférieurs à ceux observés dans les décennies précédentes. L'exploitation de l'ensemble de la diversité génétique disponible peut aider à développer les nouvelles variétés de blé qui permettront d'atteindre cet objectif alors que les cultivars commerciaux actuels n'exploitent environ que 10 % de cette ressource. Le Royaume-Uni, par exemple, a pratiquement cessé de produire des croisements expérimentaux avec le blé sauvage et les autres graminées (c'est-à-dire du matériel génétique de pré-breeding) du fait de la fermeture en 1987 du Plant Breeding Institute (PBI) basé à Cambridge. Or c'est précisément ces hybridations avec des apparentés plus ou moins éloignés qui permettent d'augmenter la diversité génétique du blé par le transfert de caractères agronomiques potentiellement très intéressants. Mais un travail de sélection complémentaire est ensuite nécessaire pour permettre la création de variétés élites hautement performantes.

En 2004, une étude sur le secteur de la sélection végétale conduite par le Biotechnology and Biological Sciences Research Council (BBSRC) du Royaume-Uni identifiait une discordance entre les résultats produits par la recherche publique et les besoins de la sélection privée en matière de renouvellement du matériel génétique (Gilligan *et al.*, 2004). C'est ainsi que, quelque vingt ans après la fermeture du PBI, le BBSRC a relancé au Royaume-Uni un programme de pré-sélection du blé tendre avec le Wheat Improvement Strategic Programme (WISP, www.wheatisp.org). Rassemblant un consortium de chercheurs universitaires (Tableau 1) et de sélectionneurs du privé, WISP peut servir de modèle quant à la manière d'améliorer le rendement des cultures à l'échelle nationale, voire au niveau mondial. Cependant, l'équilibre entre les objectifs parfois contradictoires d'une agence gouvernementale, des universitaires du secteur public et le secteur privé ne va pas sans difficultés.

¹ Traduction de l'article original paru dans NATURE PLANTS, vol. 1, March 2015 et reproduite ici avec l'aimable autorisation de l'auteur.

CONSTRUCTION DU CONSORTIUM

WISP se concentre sur le rassemblement de la diversité utile issue des ressources “exotiques” qui a, soit été perdue durant le processus de la sélection moderne, soit encore jamais exploitée. Trois groupes indépendants se consacrent aux différentes sources de matériel génétique exotique : variétés locales, blés « synthétiques » nouvellement créés et espèces apparentées sauvages (Figure 1). Le matériel génétique est phénotypé à Rothamsted et à l'Université de Nottingham. Puis il est génotypé sur les plates-formes mises au point par l'Université de Bristol.

Les sélectionneurs du secteur privé ont été impliqués à chacune des étapes du programme WISP, depuis sa conception initiale au moyen de propositions de financement, jusqu'à la livraison des résultats finaux. Au cours des réunions préliminaires, les universitaires et les sélectionneurs du privé ont convenu : *i*) du choix des croisements ; *ii*) de la nécessité de baser les croisements sur le même fond génétique élite ; *iii*) d'utiliser comme parents élites Paragon en blé de printemps et Robigus en blé d'hiver ; *iv*) des caractères prioritaires à cibler (rendement/biomasse, efficacité de l'utilisation de l'azote, résistance aux pucerons et au piétin-échaudage).

Le projet WISP a tout d'abord été soumis au comité d'examen du BBSRC selon la procédure classique et représente la plus importante subvention accordée par cette voie par le BBSRC. Le coût total du programme évalué à ce jour dépasse les 15 millions de £. Mon rôle en tant que coordinateur du WISP a été d'agir comme un “honnête courtier” dans les discussions entre le BBSRC, les universitaires du secteur public et les sélectionneurs du secteur privé. Que mes activités de recherche propres n'aient rien à voir avec le projet WISP m'a été infiniment précieux pour tenir ce rôle ; le fait qu'aucune décision de WISP ne pouvait affecter mon propre financement m'a permis d'être, ce qui est exceptionnel, un acteur engagé mais impartial.

WISP dispose également d'un gestionnaire de projet, Daniela Tomescu, dont la première tâche a été d'élaborer un accord de consortium afin de clarifier les activités et les responsabilités des partenaires, et d'établir un accord de transfert de matériel couvrant l'ensemble des produits générés par le consortium. Ainsi, tout le matériel génétique développé au sein du programme est exempt de tout droit de propriété intellectuelle et à la disposition de tous ceux qui en font la demande. Les partenaires et les sélectionneurs du secteur privé se réunissent tous les 6 mois, dans le cadre de journées portes ouvertes qui leur permettent d'observer le matériel génétique au champ.

Tableau 1 : *Partenaires du consortium WISP.*

Groupe	Chercheur principal	Localisation
Variétés locales	Simon Griffiths	John Innes Centre, Norwich
Blés synthétiques	Andy Greenland Alison Bentley Phil Howell	NIAB, Cambridge NIAB, Cambridge NIAB, Cambridge
Espèces sauvages apparentées	Ian King Julie King	Université de Nottingham Université de Nottingham
Caractérisation du matériel génétique	Malcolm Hawkesford Peter Shewry John Foulkes	Rothamsted Rothamsted Université de Nottingham
Génotypage	Keith Edwards Gary Barker	Université de Bristol Université de Bristol

Pour accéder aux ressources créées dans ce programme, contacter directement l'un des partenaires ci-dessus.
NIAB : National Institute of Agricultural Botany.

Les sélectionneurs, à titre individuel, peuvent également rencontrer chacun des partenaires pour faire part de leurs besoins en matière de matériels ou de ressources génétiques. De fait, le succès du programme WISP repose sur la fluidité des échanges entre tous les partenaires, libres de tout droit de propriété intellectuelle, quant au matériel génétique, aux données de phénotypage et à l’information relative aux croisements à effectuer. Certains avaient prédit que cet exercice serait difficile au sein du secteur public, et plus encore avec une grande diversité d’organismes. Mais jusqu’à ce jour ce n’est pas ce qui s’est passé ; le processus s’est parfaitement bien déroulé.

RESULTATS DEJA ACQUIS

Durant les décades 1920 et 1930, un fonctionnaire britannique avait pris l’initiative de faire envoyer des semences collectées sur les marchés locaux dans le monde entier par le biais des ambassades du Royaume-Uni à la Chambre de commerce de Londres. Nous disposons encore de plus de 800 entrées de cette collection de variétés locales constituée par A.E. Watkins. Le génotypage a permis de réduire ce nombre à 120 variétés locales qui retiennent l’essentiel de la diversité de départ (Wingen L. *et al.*, 2014). Chacune de ces lignées a été croisée avec la variété Paragon et les populations en disjonction correspondantes ont été produites. Ces populations, composées de plus de 9000 individus originaux ont été génotypées pour produire plus de trois millions de données élémentaires qui, combinées aux données phénotypiques recueillies relatives à l’utilisation de l’azote et à la production de la biomasse, ont permis de localiser sur les chromosomes 130 loci, contrôlant une variabilité génétique nouvelle et utile. La sélection assistée par marqueurs permettra d’introduire ces loci dans les programmes de sélection des variétés de blé élites. Un “kit” de 40 lignées dérivées de ces croisements variété locale x Paragon, et englobant toute la diversité de ces caractères, a été fourni aux sélectionneurs et est disponible sans contrepartie sur simple demande. Nous prévoyons que cet ensemble sera porté à environ 100 lignées d’ici à la fin du programme.



Figure 1 : Apparentés sauvages du blé cultivés en serre.

Le blé tendre moderne est hexaploïde et contient trois génomes (*A*, *B* et *D*) résultant de plusieurs hybridations naturelles qui se sont produites au hasard il y a quelques 10.000 ans entre le blé amidonnier tétraploïde (*Triticum dicoccoides* ; génomes *A* et *B*) et une graminée sauvage diploïde (*Aegilops tauschii* ; génome *D*). Du fait que ces hybridations ont eu lieu assez récemment, le blé moderne affiche une relativement faible variabilité du génome *D* comparée à celle des génomes *A* et *B*.

Le programme “Blés synthétiques” crée de nouveaux blés hexaploïdes en hybridant le blé dur moderne tétraploïde (*Triticum turgidum*) avec *Aegilops tauschii* (Oliveira H. R., 2014 ; Jones H. *et al.*, 2013). Un deuxième volet du programme consiste à introduire une nouvelle diversité des génomes A et B par croisement avec le blé dur et le blé amidonnier tétraploïde. Les résultats de génotypage et de phénotypage ont permis d’identifier un certain nombre d’accessions intéressantes pour divers caractères tels que la tolérance à la chaleur et à la sécheresse. Plus de cinquante nouveaux génomes hexaploïdes D et 50 géniteurs de blé dur et de blé amidonnier ont été croisés avec les variétés élités Paragon et Robigus. Environ 100 lignées originales de pré-breeding ont été dérivées de chacun de ces croisements, qui offrent des possibilités considérables pour l’amélioration génétique du blé. Ces matériels font actuellement l’objet d’une évaluation approfondie afin d’identifier les meilleures lignées, beaucoup d’entre elles ayant déjà été transmises aux sélectionneurs pour exploitation.

Les génomes entiers de toute une gamme d’espèces, incluant *Secale cereale*, *Aegilops speltoides*, *Amblyopyrum muticum*, *Triticum urartu*, *Thinopyrum bessarabicum*, *Th. elongatum*, *Th. intermedium*, *Th. Ponticum* et *Th. timopheevii* ont été transférés dans le blé sous forme de petits segments chromosomiques qui se chevauchent, selon une approche déjà développée pour les graminées fourragères (King J. *et al.*, 2013). Une série de rétrocroisements est ensuite effectuée sur la variété Paragon, jusqu’à l’obtention de lignées d’introgession portant un seul segment chromosomique. Plus de 20.000 croisements ont été réalisés au cours des trois dernières années afin d’obtenir des descendances de backcross présentant une fertilité normale. Les introgressions sont détectées en utilisant une gamme de près de 450.000 marqueurs SNP, mise au point par l’Université de Bristol et Affymetrix, qui permet de distinguer les apparentés plus ou moins éloignés des blés élités. Bien que ce travail soit encore à un stade préliminaire, du matériel introgressé est déjà en cours d’exploitation par le CIMMYT (Mexique), l’ICARDA (Maroc), le Directorate of Wheat Research (Inde), l’Université de Sydney (Australie) et la Chinese Academy of Science (Pékin).

ANALYSE DES CARACTERES

Le grand nombre de lignées générées à partir des trois principales sources de germplasm ont nécessité la mise au point de nouvelles méthodes d’analyse à haut débit, tant au niveau du champ que du laboratoire, aux stades récolte et post-récolte. L’évaluation au champ du rendement en grain et de la biomasse a été réalisée sur plusieurs années et sur deux sites, Rothamsted et Nottingham (Sikder S. *et al.*, 2015). En outre, afin d’évaluer un caractère clé pour la durabilité des systèmes agricoles, les essais sont conduits avec deux niveaux d’azote (haut et bas) ce qui permet d’évaluer la réponse du matériel aux pratiques modernes de la fertilisation azotée. Le matériel est également évalué pour la teneur en micronutriments (fer et zinc) ; ainsi, si les lignées modernes demi-naines présentent généralement une faible concentration en raison de la dilution due à un rendement élevé, quelques lignées anciennes prometteuses associent une biomasse élevée à une forte aptitude à l’absorption des micronutriments dans le feuillage et à leur transfert dans les grains. En outre, un certain nombre de lignées ont été évaluées pour leur teneur en fibres alimentaires ainsi que pour leur résistance au piétin-échaudage (*Gaeumannomyces graminis*), aux pucerons et à la mouche grise des céréales. En particulier, les lignées de Watkins se sont avérées être une source prometteuse de variation avec une plus grande diversité pour la teneur et la composition en fibres alimentaires que celles existantes dans les variétés commerciales modernes. Cette variabilité porte à la fois sur la quantité totale de fibres et sur sa proportion soluble, cette dernière étant impliquée dans le ralentissement du taux de digestion et d’absorption des nutriments dans le tractus intestinal (conférant aux produits un index glycémique plus bas et un effet hypolipémiant).

Jusqu’à tout récemment, la disponibilité des marqueurs SNP associée à l’accès à des plateformes de génotypage capables de traiter les génomes polyploïdes, était l’un des principaux goulots d’étranglement auxquels étaient confrontés les sélectionneurs de blé. L’exploitation du matériel génétique dans le cadre du programme WISP a nécessité la création d’une plate-forme de génotypage du blé publique et libre de tout droit de propriété intellectuelle. A cette fin, le premier ensemble complet de marqueurs SNP cartographiés a été basé sur la plate-forme Kompetitive Allele Specific

PCR ou KASP (Allen A. M. *et al.*, 2012). Cela a été suivi, début 2014, par la publication d'un ensemble de données génotypiques basées sur les 90.000 SNP de la puce iSelect (Wang S. *et al.*, 2014), et plus récemment par un autre ensemble de données basé sur les 35 000-SNP du Axiom Wheat Breeders Array et les 820.000 SNP du Axiom High Density Array (www.cerealsdb.uk.net). Au total, il s'agit de plus de 400 millions de données élémentaires, constituant l'ensemble de données de génotypage du blé le plus complet mis librement à la disposition des chercheurs blé au niveau mondial. L'intérêt suscité a été tel que, par exemple, l'ensemble des 35.000 données SNP Axiom a été téléchargé plus de 2600 fois au cours des trois premières semaines suivant sa publication. Ainsi, grâce au programme WISP, la communauté blé a librement accès à près d'un million de marqueurs SNP validés et caractérisés.

LES ENSEIGNEMENTS A RETENIR

Au-delà de ses résultats immédiats, l'expérience du programme de WISP peut aider à la mise en place de programmes de sélection pour d'autres cultures. Pour établir une collaboration à cette échelle, il faut avoir clairement identifié les besoins auxquels on veut répondre, quels sont les partenaires à impliquer pour y parvenir et pourquoi. Cela permet d'expliquer beaucoup plus facilement à ceux qui pourraient se sentir potentiellement concernés les raisons pour lesquelles ils ne sont pas indispensables. Pour les partenaires retenus, ils doivent s'interroger dès le départ sur les sorties qu'ils attendent du programme, de façon à bien équilibrer leurs besoins avec ceux de l'ensemble du programme. Les négociations sur le financement et entre partenaires individuels sont grandement facilitées si le leader lui-même ne reçoit pas d'aide financière du programme. De plus, un comité de pilotage des parties prenantes (dans le cas de WISP, les sélectionneurs) est irremplaçable pour résoudre les problèmes et établir les priorités entre les partenaires. Enfin, un programme tel que WISP n'a pas pour objet principal la production de publications scientifiques, ce qui pourrait conduire à une sous-appréciation des contributions des chercheurs universitaires qui y participent. Quoiqu'il en soit, le programme WISP a une claire vision quant aux objectifs qu'il s'est donné et aux moyens de les atteindre, de telle sorte que sa réussite ne fait aucun doute.

“Journée ASF du 4 février 2016”

“Apport des marqueurs moléculaires et de la génomique à l'utilisation des ressources génétiques pour l'amélioration des plantes”

REFERENCES BIBLIOGRAPHIQUES

- Allen, A. M. *et al.*, 2012. Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotech. J.* 9, 1086–1099.
- Gilligan, C. *et al.*, 2004. *Review of BBSRC-funded Research Relevant to Crop Science* (BBSRC, 2004) ; <http://go.nature.com/4ye7RF>.
- Jones, H. *et al.*, 2013. Strategy for exploiting exotic germplasm using genetic, morphological, and environmental diversity: the *Aegilops tauschii* Coss. example. *Theor Appl Genet.* 126, 1793–1808.
- King, J. *et al.*, 2013. Exploitation of interspecific diversity for monocot crop improvement. *Heredity* 110, 475–483 (2013).
- Oliveira, H. R., 2014. Wheat in the Mediterranean revisited – tetraploid wheat landraces assessed with elite bread wheat Single Nucleotide Polymorphism markers. *BMC Genetics* 15, 54.
- Sikder, S. *et al.*, 2015. Evaluation of photosynthetic potential of wheat genotypes under drought condition. *Photosynthetica* 53, 47–54 (2015).
- Wang, S. *et al.*, 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotech. J.* 12, 787–796.
- Wingen, L. *et al.*, 2014. Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat. *Theor. Appl. Genet.* 127, 1831–1842.

RE-SYNTHESE D'ESPECES : L'EXEMPLE DU COLZA

**Anne-Marie CHEVRE, Alexandre PELE,
Sophie PAILLARD, Mathieu ROUSSEAU-GUEUTIN**
INRA, UMR1349 Institut de Génétique, Environnement et Protection des Plantes
BP35327, 35653 Le Rheu cedex
E-mail : anne-marie.chevre@rennes.inra.fr

RESUME

Le colza (*Brassica napus*, AACC, $2n=4x=38$) est une espèce allopolyploïde d'intérêt agronomique et économique majeur qui constitue un excellent modèle pour identifier la stratégie optimale afin d'introduire de la diversité génétique provenant de ses deux espèces constitutives, la navette (*Brassica rapa*, AA, $2n=2x=20$) et le chou (*Brassica oleracea*, CC, $2n=2x=18$). La première stratégie est de produire des formes resynthétisées par croisement des espèces diploïdes. Chez ce matériel, la première méiose est à l'origine d'une explosion de recombinaisons homéologues entre les génomes A et C qui entraînent de nombreux réarrangements, une instabilité méiotique dans la descendance et une faible fertilité. Cependant, le croisement avec le colza naturel peut permettre de modifier le dosage de certaines régions génomiques d'intérêt et de cumuler ainsi des QTL à effets forts. Pour la seconde stratégie, correspondant au croisement direct entre le colza et l'une de ses espèces constitutives, nous avons montré qu'il est possible d'augmenter le nombre et de modifier la position des crossovers entre chromosomes homologues. Cette stratégie a été exploitée pour introduire la diversité génétique présente dans les core collections des deux espèces diploïdes dans une même variété de colza. Ces deux stratégies permettent de réduire le goulot d'étranglement subi au cours de la sélection, ouvrant de nouvelles perspectives pour l'amélioration du colza.

Mots-clés : diversité génétique, recombinaison homologue, recombinaison homéologue, polyplôïdie, *Brassica napus*.

1 - INTRODUCTION

Le colza (*Brassica napus*, AACC, $2n=4x=38$) est un hybride naturel issu d'un croisement entre la navette (*Brassica rapa*, AA, $2n=2x=20$) et le chou (*Brassica oleracea*, CC, $2n=2x=18$) (U, 1935). Sa culture a connu une véritable explosion à partir de la moitié du XX^e siècle avec deux défis qui ont rapidement été surmontés par la sélection pour la qualité de la graine : la suppression de l'acide érucique et la réduction de la teneur en glucosinolates. Cependant, ces contraintes d'amélioration ont entraîné une forte réduction de la variabilité génétique (Qian *et al.*, 2006). Cette difficulté nécessite d'exploiter la très large diversité de ses deux espèces progénitrices.

Ces travaux d'amélioration sont facilités par les récentes données génomiques acquises ces dernières années faisant du colza un modèle pour l'étude de la polyplôïdie et de l'impact des duplications génomiques. Cette espèce, qui n'existe pas sous forme spontanée, a une origine encore spéculative : 500 à 600 ans pour les botanistes (Gomez-Campo, 1980) alors que les datations moléculaires suggèrent une origine plus ancienne d'environ 7500 ans (Chalhoub *et al.*, 2014). Les données de cartographie génétique (Parkin, 2011), de séquençage du colza et de ses deux progéniteurs

(Wang *et al.*, 2011 ; Chalhoub *et al.*, 2014 ; Liu *et al.*, 2014 ; Parkin *et al.*, 2014) indiquent que les génomes constitutifs du colza sont bien conservés par comparaison avec les formes modernes diploïdes. Nous savons également que tous les génomes diploïdes de Brassicacées dérivent d’un ancêtre commun hexaploïde et que chaque région génomique est tripliquée (Parkin *et al.*, 2011 ; Schranz *et al.*, 2006). Ces régions montrent un fractionnement différentiel, présentant plus ou moins de gènes conservés (Tang *et al.*, 2012 ; Cheng *et al.*, 2012, 2013 ; Murat *et al.*, 2015). En dépit de ce haut niveau d’homéologie entre ses génomes, le colza naturel présente une méiose régulière et une hérédité disomique. Cette similarité de structure doit cependant permettre d’introduire de façon optimale de la variabilité à partir de ses espèces parentales diploïdes par recombinaison.

La recombinaison, par l’intermédiaire des crossovers (COs), est le principal mécanisme pour générer de la diversité dans les descendances. La recombinaison homologue est un mécanisme strictement contrôlé reposant sur un CO obligatoire par paire de chromosomes homologues et parfois deux ou trois (Mézard *et al.*, 2007). Par ailleurs, la localisation de ces COs dépend de la structure du chromosome, les COs étant fréquemment localisés dans les régions télomériques et absents des zones centromériques (e.g. Drouaud *et al.*, 2007 ; Saintenac *et al.*, 2009). Dans le cas d’une espèce allopolyploïde récente telle que le colza, il est possible de la croiser avec l’une de ses espèces diploïdes progénitrices afin d’introduire de la diversité génétique par recombinaison homologue. Par ailleurs, Jenczewski *et al.* (2003) a démontré qu’il existe un contrôle génétique de l’appariement homéologue qui promeut ou non l’appariement entre les génomes A et C chez des haploïdes AC de colza. Seuls deux types d’appariements sont observés, haut ou faible niveau d’appariement (Cifuentes *et al.*, 2010) ; ces auteurs ont montré que ce caractère peut être relié aux deux origines phylogéniques du colza. L’appariement homéologue est sous contrôle polygénique, dont un locus majeur, *PrBn* pour « Pairing regulator in *B. napus* » (Liu *et al.*, 2006 ; Nicolas *et al.*, 2009, 2012).

Nous proposons de décrire les deux principales stratégies (Figure 1) pour exploiter la diversité génétique disponible chez les deux parents diploïdes pour l’amélioration du colza. Il est possible, soit de produire des colzas resynthétisés par croisement entre les deux espèces parentales, soit de croiser directement le colza avec ses progéniteurs actuels. L’intérêt de chacune de ces stratégies sera analysé au regard des taux de recombinaison et du contrôle génétique des caractères que l’on souhaite introduire.

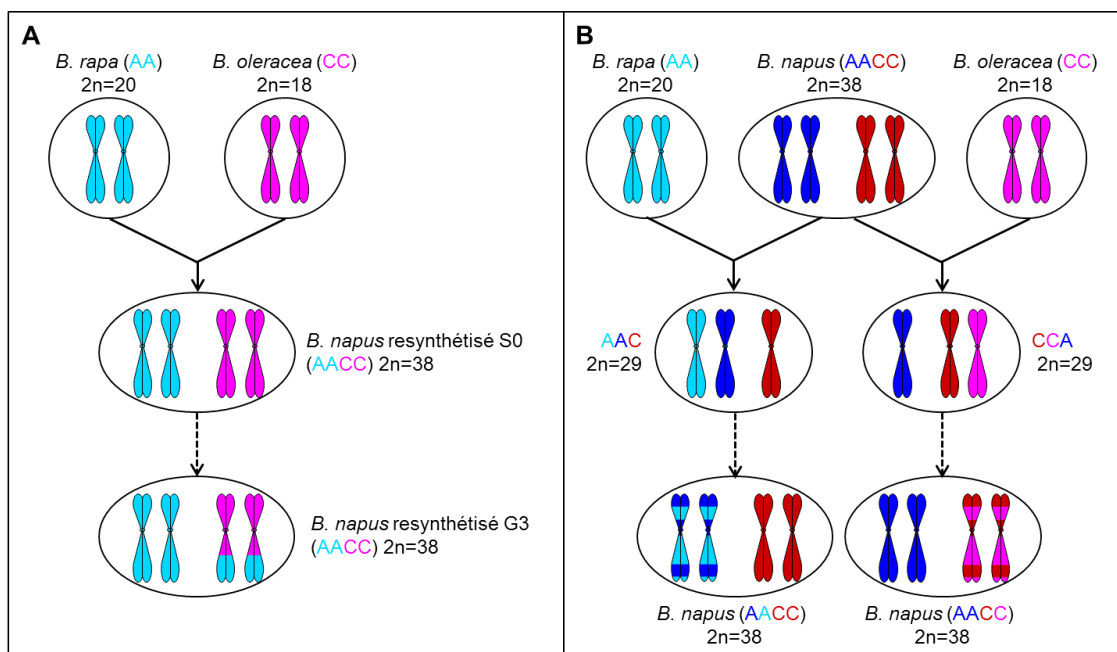


Figure 1 : Présentation des différentes stratégies pour introduire de la variabilité génétique dans le colza à partir de ses espèces diploïdes progénitrices.

2 - INTRODUCTION DE DIVERSITE GENETIQUE VIA LA CREATION DE COLZAS RESYNTHETISES

L'amélioration du colza naturel via la production de colzas resynthétisés (Figure 1A) a pour avantage de capter toute la diversité présente au sein des variétés de navette et de chou choisies. Cependant, la première méiose se traduit toujours par une explosion de recombinaison entre les génomes homéologues A et C, que ce soit lors de la formation de gamètes non réduits (Szadkowski *et al.*, 2011) ou après doublement chromosomique (Szadkowski *et al.*, 2010) de l'hybride F1 de génome AC. La principale différence entre ces deux voies d'obtention de la plante AACC resynthétisée (S0) est la fréquence et la taille des échanges entre les deux génomes homéologues. Cette instabilité méiotique initiale conduit à la formation d'aneuploïdes et à de nombreux réarrangements, à l'origine d'une très faible fertilité par rapport aux colzas naturels (Gaeta *et al.*, 2007 ; Gaeta & Pires, 2010 ; Xiong *et al.*, 2011).

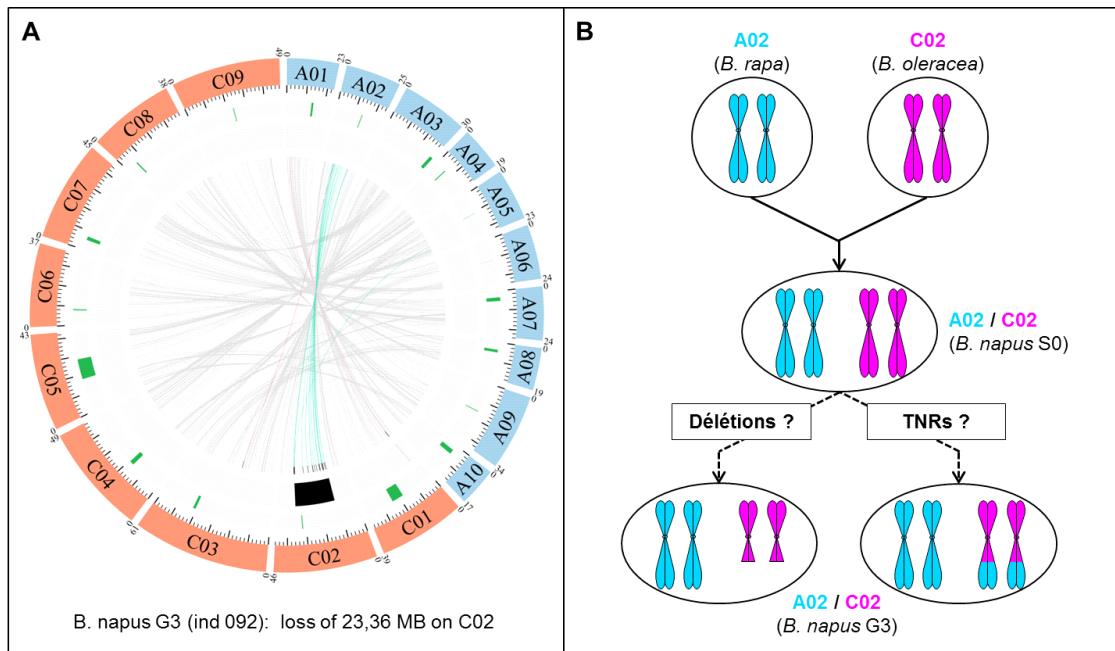


Figure 2 : Exemple d'une déletion de 23,36Mb observée sur le chromosome C02 dans une plante de colza resynthétisé de la troisième génération. (A) Les 10 chromosomes du génome A sont représentés en bleu et ceux du génome C en rouge, la taille physique des chromosomes étant indiquée. Les rectangles verts représentent la position des centromères. Les liens indiquent les localisations physiques (relation d'homéologie) des marqueurs SNP polymorphes entre les espèces diploïdes parentales ; ceux qui sont gris présentent toujours les deux allèles parentaux chez la plante synthétique alors que les liens de couleur indiquent la perte d'un allèle parental. Le bloc noir correspond à la région déléetée. (B) La perte de marqueurs peut correspondre à une déletion ou une translocation non réciproque (TNR) à l'état homozygote.

Cependant les travaux initiaux ont été conduits par Single Seed Descent (SSD) et l'intercroisement en fécondation libre de ces colzas resynthétisés permet l'élimination naturelle des aneuploïdes peu fertiles, en dépit de variations structurales (Figure 2A) (Rousseau-Gueutin *et al.*, soumis). Ces remaniements identifiés via la perte de marqueurs SNP adjacents, physiquement ancrés sur le génome, peuvent correspondre soit à des déléetions, soit à des translocations non réciproques homozygotes (TNRs) pour lesquelles la déléetion est remplacée par la région homéologue correspondante (Figure 2B). Le mécanisme impliqué peut être identifié, soit par séquençage NGS, la zone dupliquée étant deux fois plus représentée qu'attendu dans le cas de TNRs homozygotes, soit par hybridation *in situ* à l'aide de BACs spécifiques de la région remaniée. De même que chez les colzas

resynthétisés, des TNRs de plus petite taille ont été mises en évidence chez des colzas naturels, 17 étant présentes chez la variété séquencée ‘Darmor’ (Chalhoub *et al.*, 2014). Ainsi, la recombinaison homéologue permet de modifier le nombre de doses d’une région génomique portant des gènes d’intérêt. En effet, Pires *et al.* (2004) ont montré qu’en changeant le nombre de doses du locus *FLC*, contrôlant la précocité de floraison, il est possible d’obtenir des plantes plus ou moins tardives.

L’instabilité de ce matériel resynthétisé ne permet pas une exploitation directe en culture. Seule une variété utilisée comme légume a été proposée (Fujii *et al.*, 2011). En revanche, les hybrides produits par croisement de ces formes resynthétisées avec des colzas naturels peuvent présenter un bon niveau d’hétérosis (Seyis *et al.*, 2006 ; Jesske *et al.*, 2013) et l’étude de descendants en ségrégation permet d’améliorer des composantes du rendement (Radoev *et al.*, 2008).

3 - INTRODUCTION DE DIVERSITE GENETIQUE VIA LE CROISEMENT ENTRE LE COLZA ET L’UN DE CES PROGENITEURS DIPLOÏDES

Le croisement direct entre un colza naturel et l’une de ses espèces parentales, la navette ou le chou, permet de produire respectivement des structures allotriploïdes AAC ou ACC (Figure 1B). Nous avons montré que la structure AAC, présentant une méiose proche du comportement méiotique attendu (10 bivalents A et 9 univalents C en métaphase I de méiose) (Leflon *et al.*, 2006), favorise la recombinaison homéologue entre les chromosomes A. En effet, une augmentation d’un facteur 4 à 6 chez les hybrides AAC a été observée par comparaison à un hybride AA (Leflon *et al.*, 2010). L’utilisation de marqueurs SNP ancrés sur l’ensemble du génome de la navette séquencée (Wang *et al.*, 2011) et régulièrement répartis (un SNP pour 1,25 Mpb) a permis de mettre en évidence que la distribution des COs est également affectée. En effet, à partir d’une descendance de 528 plantes issues d’un hybride AA et de 142 plantes issues d’un hybride AAC, la formation de COs a été observée dans la totalité des intervalles entre SNPs chez les hybrides AAC, y compris dans les régions péricentromériques dépourvues de COs chez les hybrides AA (Pelé *et al.*, en prép.).

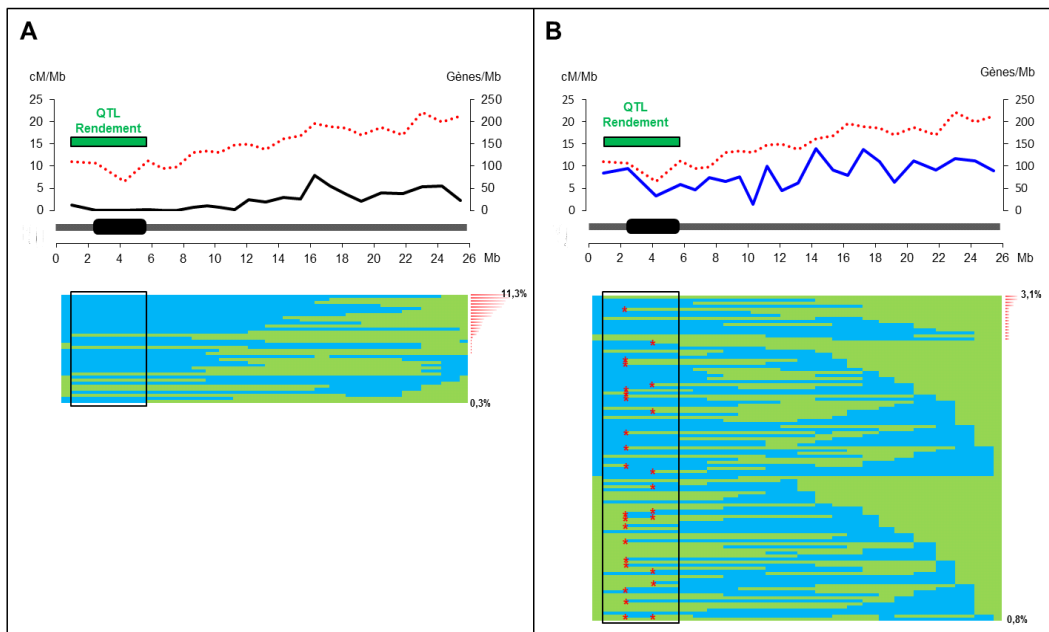


Figure 3 : Fréquence de recombinaison sur le chromosome A07 exprimée en cM/Mb (A) à partir d’un hybride diploïde AA (courbe noire) et (B) d’un hybride AAC (courbe bleue). La densité en gènes correspond à la courbe en pointillés rouges exprimée en gènes/Mb. Les différentes structures génétiques observées dans les descendance d’un hybride (A) AA ou (B) AAC sont représentées sous chaque graphique. La fréquence de chaque structure génétique est indiquée par l’histogramme à droite de chaque graphique.

Des résultats similaires ayant été observés entre l'hybride AA et le colza (AACC), l'exploitation d'hybrides AAC présente deux intérêts majeurs : favoriser le brassage génique et réduire la taille de certains QTL d'intérêt. Près de 8,4 % des gènes totaux du génome A (3100 gènes) sont identifiés dans des régions non recombinantes chez les hybrides AA et AACC. L'emploi d'hybrides AAC permettrait pour la première fois de faire ségréger dans la descendance ces différents locus. De plus, en se basant sur la détection de QTL impliqués dans les composantes du rendement réalisée par Bouchet *et al.* (soumis), nous avons pu observer qu'il semble possible de réduire la taille d'un QTL identifié sur le chromosome A07 via l'exploitation d'hybride AAC. En effet, dans cette région génomique d'environ 5 Mpb, aucun CO n'est observé parmi les 528 plantes produites à partir du diploïde AA alors que 29 COs sont identifiés avec seulement 142 plantes obtenues à partir de l'hybride AAC (Figure 3). Ces résultats ouvrent ainsi de nouvelles perspectives pour l'introduction de caractères d'intérêt provenant de la navette à partir d'hybrides AAC.

Compte tenu de ce résultat, nous avons exploité la voie allotripléide pour introgresser de la diversité génétique présente chez la navette et le chou dans le colza. Des accessions représentatives de la diversité de chacune des deux espèces diploïdes (9 à 10 accessions par espèce) ont été croisées avec une même variété de colza mâle stérile sous cages. Environ 80 hybrides F1 AAC ont été produits par accession alors que les hybrides CCA ont été plus difficiles à obtenir. Ces allotripléides ont été pollinisés sous cages par la même variété de colza (hétérozygote pour le gène de restauration de la fertilité mâle). Seuls les gamètes AC de chaque hybride allotripléide permettent de produire une plante AACC de même structure que le colza. La fréquence des gamètes AC n'étant que de 5 % environ (Leflon *et al.*, 2006 ; Namai *et al.*, 1987), environ 32 000 plantes ont été triées par cytométrie en flux afin d'obtenir ~80 plantes à $2n=38$ par accession d'origine. En théorie, ces colzas présentent 25 % du génome de l'accession diploïde parentale (Paillard *et al.*, en prép.). Les graines de plus de 1200 plantes de colza ont été distribuées aux partenaires du projet (Promosol : projet 'ProBiodiv'). La première analyse de ce matériel à partir de puces SNP révèle que des allèles des accessions diploïdes d'origine ont été introduits de façon homogène sur tous les chromosomes.

4 - CONCLUSION

Les deux stratégies analysées mettent en évidence qu'il est possible d'introduire une importante diversité génétique dans le colza allotétraploïde, par recombinaison homéologue entre les génomes A et C ou par recombinaison homologue. La première possibilité à partir de formes resynthétisées s'avère particulièrement intéressante dans le cas de caractères quantitatifs pour lesquels les QTL sont soumis à des effets de dosage. En effet, les analyses prenant en compte la structure du génome révèlent que les QTL sont souvent portés par les chromosomes homéologues (e.g. Fomeju *et al.*, 2014) et cette stratégie permettrait de dupliquer le QTL présentant le plus fort effet sur son chromosome homéologue. Il reste à élucider comment rétablir le rendement et contrôler les appariements homéologues lors de la formation d'un colza resynthétisé.

La seconde possibilité par le croisement direct d'un colza naturel et une de ses espèces constitutives a pour avantage de modifier le contrôle de la recombinaison homologue. La très forte fréquence de recombinaison entre les chromosomes A dans des hybrides AAC couplée à une distribution des COs tout le long des chromosomes permet de limiter la taille des régions génomiques introgressées aux seules zones d'intérêt. Par ailleurs, cette stratégie pourrait s'avérer pertinente pour limiter l'intervalle de confiance des QTL et ainsi faciliter le clonage positionnel. Les premières observations indiquent que la fréquence de recombinaison pourrait être également élevée dans les hybrides ACC entre les chromosomes du génome C (travaux en cours). Cette importante variabilité génétique introduite dans le colza via la voie allotripléide a déjà permis de produire des populations à base génétique très large qui sont en cours d'exploitation dans les programmes de sélection.

“Journée ASF du 4 février 2016”

“ Apport des marqueurs moléculaires et de la génomique à l'utilisation des ressources génétiques pour l'amélioration des plantes ”

REFERENCES BIBLIOGRAPHIQUES

- Chalhoub B., Denoeud F., Liu S.Y. *et al.*, 2014. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345 (6199):950-953.
- Cheng F., Mandakova T., Wu J., Xie Q., Lysak M.A., Wang X.W., 2013. Deciphering the diploid ancestral genome of the mesohehexaploid *Brassica rapa*. *Plant Cell* 25 (5):1541-1554.
- Cheng F., Wu J., X. Wang X., 2014. Genome triplication drove the diversification of *Brassica* plants. *Hort. Res.*1:14024.
- Cifuentes M., Eber F., Lucas M.O., Lode M., Chevre A.M., Jenczewski E., 2010. Repeated polyploidy drove different levels of crossover suppression between homoeologous chromosomes in *Brassica napus* allohaploids. *Plant Cell* 22 (7):2265-2276.
- Drouaud J., Mercier R.I., Chelysheva L. *et al.*, 2007. Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. *Plos Genetics* 3(6): 1096-1107.
- Fomeju B. F., Falentin C., Lassalle G., Manzanares-Dauleux M.J., Delourme R.. 2014. Homoeologous duplicated regions are involved in quantitative resistance of *Brassica napus* to stem canker. *BMC Genomics* 15:498
- Fujii K., Ohmido N., 2011. Stable progeny production of the amphidiploid resynthesized *Brassica napus* cv. Hanakkori, a newly bred vegetable. *Theor. Appl. Genet.* 123 (8):1433-1443.
- Gaeta R. T., Pires J.C., Iniguez-Luy F., Leon E., Osborn T.C., 2007. Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell* 19 (11):3403-3417.
- Gaeta R. T., Pires J.C., 2010. Homoeologous recombination in allopolyploids: the polyploid ratchet. *New Phyt.* 186:18-28.
- Gomez-Campo C., 1980. Morphology and morphotaxonomy of the tribe Brassiceae. In *Brassica crops and wild allies: Biology and breeding*, edited by S. Tsunoda, K. Hinata, and C. Gomez-Campo. Tokyo: Japan Scientific Soc. Press.
- Jenczewski E., Eber F., Grimaud A. *et al.* 2003. *PrBn*, a major gene controlling homeologous pairing in oilseed rape (*Brassica napus*) haploids. *Genetics* 164 (2):645-653.
- Jeske T., Olberg B., Schierholt A., Becker H.C., 2013. Resynthesized lines from domesticated and wild *Brassica* taxa and their hybrids with *B. napus* L.: genetic diversity and hybrid yield. *Theor. Appl. Genet.* 126:1053–1065.
- Leflon M., Eber F., Letanneur J.C. *et al.*, 2006. Pairing and recombination at meiosis of *Brassica rapa* (AA) × *Brassica napus* (AACC) hybrids. *Theor. Appl. Genet.* 113:1467-1480.
- Leflon M., Grandont L., Eber F. *et al.*, 2010. Crossovers get a boost in *Brassica* allotriploid and allotetraploid hybrids. *Plant Cell* 22 (7):2253-64.
- Liu S. Y., Liu Y.M., Yang X.H. *et al.*, 2014. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Comm.* 5:3930.
- Liu Z., Adamczyk K., Manzanares-Dauleux M. *et al.*, 2006. Mapping *PrBn* and other quantitative trait loci responsible for the control of homeologous chromosome pairing in oilseed rape (*Brassica napus* L.) haploids. *Genetics* 174 (3):1583-1596.
- Mézard C., Vignard J., Drouaud J., Mercier R., 2007. The road to crossovers: plants have their say. *Trends Genet.* 23 (2):91-99.
- Murat F., Louis A., Maumus F. *et al.*, 2015. Understanding Brassicaceae evolution through ancestral genome reconstruction. *Genome Biol.* 16:262.
- Namai H., 1987. Inducing cytogenetical alterations by means of interspecific and intergeneric hybridization in *Brassica* crops. *Gamma Field Symp.* 26:41-89.
- Nicolas S. D., Monod H., Eber F., Chèvre A.M. Jenczewski E., 2012. Non-random distribution of extensive chromosome rearrangements in *Brassica napus* depends on genome organization. *Plant J.*70 (4):691-703.
- Nicolas S. D., Leflon M., Monod H. *et al.*, 2009. Genetic regulation of meiotic cross-overs between related genomes in *Brassica napus* haploids and hybrids. *Plant Cell* 21 (2):373-385.

- Parkin I.A.P., 2011. Chasing Ghosts: Comparative Mapping in the Brassicaceae. In *Genetics and Genomics of the Brassicaceae*, edited by R. Schmidt, and I. Bancroft: Springer New York.
- Parkin I.A.P., Koh C., Tang H. *et al.*, 2014. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea* *Genome Biol.* 15 (6):R77.
- Pires J. C., Zhao J., Schranz M.E. *et al.*, 2004. Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (*Brassicaceae*). *Biological Journal of the Linnean Society* 82 (4):675-688.
- Qian W., Meng J., Li M. *et al.*, 2006. Introgression of genomic components from Chinese *Brassica rapa* contributes to widening the genetic diversity in rapeseed (*B. napus* L.), with emphasis on the evolution of Chinese rapeseed. *Theor. Appl. Genet.* 113 (1):49-54.
- Radoev M., Becker H.C., Ecke W., 2008. Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by quantitative trait locus mapping. *Genetics* 179:1547–1558.
- Saintenac C., Falque M., Martin O.C., Paux E., Feuillet C., Sourdille P., 2009. Detailed Recombination Studies Along Chromosome 3B Provide New Insights on Crossover Distribution in Wheat (*Triticum aestivum* L.). *Genetics* 181: 393–403
- Schranz M.E., Lysak M.A., Mitchell-Olds T., 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends Plant Sci.* 11 (11):535-542.
- Seyis F., Friedt W., Luhs W., 2006. Yield of *Brassica napus* L. hybrids developed using resynthesized rapeseed material sown at different locations. *Field Crops Res.* 96 (1):176-180.
- Szadkowski E., Eber F., Huteau V. *et al.*, 2010. The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phyt.* 186 (1):102-112.
- Szadkowski E., Eber F., Huteau V. *et al.*, 2011. Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus*. *New Phyt.* 191 (3):884-894.
- Tang H. B., Woodhouse M.R., Cheng F. *et al.*, 2012. Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy. *Genetics* 190 (4):1563-1574.
- U N., 1935. Genome-analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jap. J. Bot.* 7:389-452.
- Wang X. W., Wang H.Z., Wang J. *et al.*, 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genet.* 43 (10):1035-1039.
- Xiong Z. Y., Gaeta R.T., Pires J.C., 2011. Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proc. Natl. Acad. Sci. U. S. A.* 108 (19):7908-7913.

L'APPORT DES MARQUEURS MOLECULAIRES ET DE LA GENOMIQUE POUR L'UTILISATION DES RESSOURCES GENETIQUES CHEZ LA TOMATE

Christopher SAUVAGE¹, Guillaume BAUCHET^{1,2} et Mathilde CAUSSE¹

¹ - INRA UR1052, GAFL, Centre de recherche PACA, domaine Saint Maurice,
67 Allée des chênes, CS 60094, 84143, Montfavet Cedex, France

² - Boyce Thompson Institute, Université Cornell, Etats-Unis

E-mail : sauvage@avignon.inra.fr

RESUME

La tomate est l'un des légumes les plus produits dans le monde et cette espèce est également un modèle dans le domaine scientifique pour l'étude du développement du fruit. Au cours des 60 dernières années, de nombreuses campagnes de collecte en Amérique Centrale et du Sud (zone d'origine), mais aussi dans le monde entier, ont permis de conserver plus de 83.000 accessions chez cette espèce. En parallèle, l'apport de la biologie moléculaire au travers du développement de marqueurs moléculaires a permis de cartographier de nombreux QTL pour des caractères d'intérêt tels que la qualité du fruit ou la résistance à des stress biotiques et abiotiques. Depuis 2012, date de la publication du génome séquencé de la tomate, la combinaison de ces importantes ressources génétiques et de la génomique haut-débit permet d'étendre les travaux à l'échelle du génome afin de mieux comprendre sa nature, d'identifier un grand nombre de sites polymorphes mais aussi de revisiter des questions scientifiques, telles que l'histoire évolutive de cette espèce et l'étude de l'architecture génétique de caractères quantitatifs par le biais de nouvelles approches comme la génétique d'association ou la sélection génomique. Les ressources génétiques disponibles chez cette espèce sont primordiales pour la recherche scientifique fondamentale mais aussi pour l'amélioration génétique de cette espèce.

Mots-clés : *Solanum lycopersicum*, génomique, marqueurs moléculaires, ressources génétiques, étude d'association, génétique quantitative

1 - INTRODUCTION

La tomate, premier légume produit dans le monde, représente 14 % de la production mondiale de légumes (plus de 140 millions de tonnes/an pour une valeur marchande de 1,6 milliard de dollars, Food and Agriculture Organisation [FAO] 2010). La tomate est une source importante en micronutriments pour l'alimentation humaine. Cette espèce est également une espèce modèle pour l'étude du développement des fruits et de l'accumulation de métabolites. Les principaux objectifs visés par les producteurs de tomates tendent à une plus grande productivité, une meilleure tolérance aux stress biotiques et abiotiques et à l'augmentation de la valeur sensorielle et nutritionnelle du fruit. L'amélioration de ces caractères nécessite la caractérisation et la gestion de la diversité génétique de la tomate présente dans les ressources génétiques. En raison de son origine latino-américaine et de l'histoire de sa domestication, la tomate cultivée a fait face à plusieurs goulots d'étranglement au cours du temps. Cela a conduit à une réduction drastique de sa diversité génétique. L'étude du centre d'origine de la tomate a permis des avancées majeures dans la caractérisation de sa diversité naturelle. En parallèle, les initiatives de conservation des plantes *ex situ* fleurissaient, assurant la collecte et la conservation des variétés locales et des espèces sauvages à travers le développement de banques de semences.

Depuis le milieu du XX^e siècle, de nouvelles méthodes telles que l'hybridation contrôlée ont permis de croiser les tomates sauvages avec les tomates cultivées. La génétique moderne et l'amélioration des méthodes de culture ont contribué à une meilleure compréhension du contrôle génétique des caractères d'intérêt, mais aussi accentué les progrès génétiques sur ces caractères. L'introgession de caractères d'intérêt des accessions sauvages dans la tomate cultivée n'a pas toujours été simple, notamment en raison de liaison entre les effets «favorables» et «défavorables» des fragments introgressés.

L'avènement de la biologie moléculaire dans les années 80 a suscité de grands espoirs en termes de caractérisation de la diversité génétique présente dans les deux compartiments sauvages et cultivés. En outre, de grandes attentes ont émergé depuis le développement des techniques moléculaires pour "repérer" des régions génomiques impliquées dans des caractères d'intérêt. L'étude du contrôle génétique des caractères complexes, en utilisant des techniques *ad hoc* de la génétique quantitative, est devenue possible, conduisant à l'identification des allèles clés impliqués dans divers caractères agronomiques, provenant de plusieurs espèces sauvages apparentées. Le génome de la tomate a été entièrement séquencé (TGC, 2012). Cet apport constitue une nouvelle étape dans la connaissance de la diversité de la tomate en lien avec les disciplines de type “omiques” et l'apport constant des techniques de séquençage de nouvelle génération (NGS). Ces technologies et l'analyse des données associées, permettent d'obtenir une vision intégrée du génome et d'autres niveaux de régulation et d'information tels que le transcriptome, le protéome ou le métabolome, obtenus également à haut débit.

Parmi toutes ces nouvelles approches, qui émergent suite à la publication du génome de la tomate, les approches de cartographie de QTL, aussi connues sous le terme de génétique d'association à l'échelle du génome (GWAS pour Genome-Wide Association Study) tirent profit des populations naturelles, et facilitent la caractérisation génétique des caractères complexes et la gestion du matériel génétique pour les variétés de tomates sauvages et cultivées.

2 - RESSOURCES GENETIQUES ET DIVERSITE

La tomate appartient à la famille des Solanacées qui comprend plus de trois mille espèces. Parmi celles-ci, les plus cultivées aujourd'hui sont originaires du Vieux Monde (aubergine) et du Nouveau Monde (piment, pomme de terre, tabac, tomate). Le clade *Lycopersicon* (Figure 1) contient la tomate domestiquée (*Solanum lycopersicum*) et ses 12 espèces sauvages apparentées (Peralta et Spooner 2005) que l'on peut désigner sous le terme de "crop wild relative" (CWR). L'âge du clade de la tomate a été estimé entre 2,7 et 7,8 millions d'années par des comparaisons de séquences nucléotidiques entre *Solanum lycopersicum* et *S. pennellii* (Nesbitt et Tanksley, 2002). Les premières études détaillées sur les CWR ont été réalisées par Charles M. Rick et ses collègues depuis les années 40. Les espèces du clade de la tomate sont originaires des Andes, notamment du Pérou, de la Bolivie, et l'Equateur, de la Colombie et du Chili. Leurs habitats naturels sont très variés, depuis le niveau de la mer jusqu'à 3300 m d'altitude, et des climats arides à équatoriaux (*Solanum cheesmaniae*; *S. galapagense*). Cette large variation est également observée aux niveaux morphologique, physiologique, reproductif et moléculaire (Peralta et Spooner, 2005). La figure 1, présente la variabilité de la couleur et de la forme du fruit dans le clade de la tomate.

Plusieurs classifications phylogénétiques ont été proposées et plusieurs ajustements ont eu lieu. La première classification a été fondée sur des critères morphologiques. L'apport des marqueurs moléculaires a tout d'abord permis de revoir cette classification en sous-taxons (Peralta *et al.*, 2005). Le clade de la tomate est un exemple intéressant pour la recherche sur la biodiversité végétale, notamment, sur l'histoire évolutive (macro-échelle), l'adaptation, la domestication par l'homme (micro-échelle) (Peralta et Spooner, 2007). De nos jours, à travers l'Amérique du Sud, les populations de tomates sauvages sont sévèrement réduites. Leurs habitats naturels sont en diminution en raison du développement urbain, de l'agriculture intensive ainsi que de l'élevage de chèvres dans les montagnes, comme l'a récemment documenté une expédition botanique au Pérou (Grandillo *et al.*, 2011).

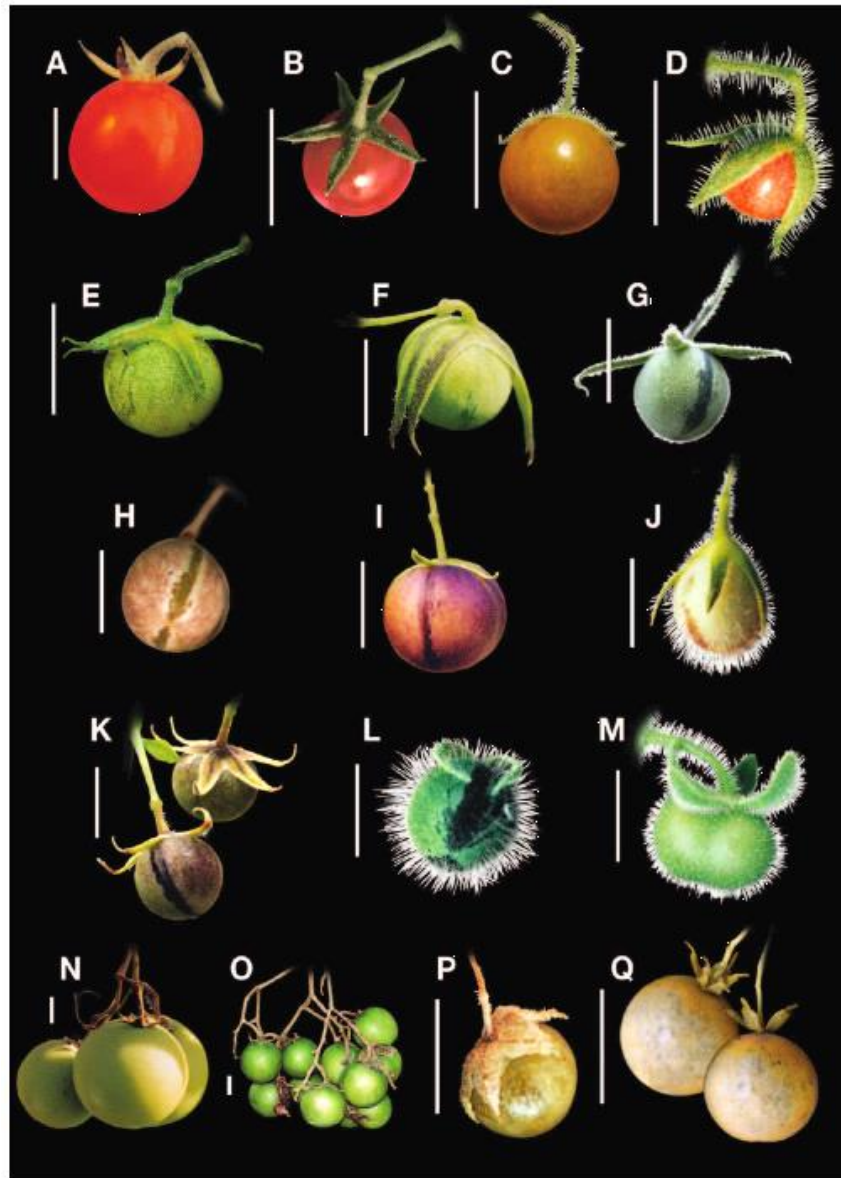


Plate 3. Fruits of *Solanum* sect. *Lycopersicon*, sect. *Juglandifolia*, and sect. *Lycopersicoides*. A. *S. lycopersicum* (LA1673). B. *S. pimpinellifolium* (LA1237). C. *S. cheesmaniae* (LA1450). D. *S. galapagense* (LA317). E. *S. neorickii* (LA2200). F. *S. chmielewskii* (LA2663). G. *S. arcanum* (LA2548). H. *S. huaylasense* (LA1365). I. *S. peruvianum* (LA153). J. *S. corneliomulleri* (LA1609). K. *S. chilense* (LA1930). L. *S. habrochaites* (LA1353). M. *S. pennellii* (LA716). N. *S. ochranthum* (Spooner et al. 5000). O. *S. juglandifolium* (Castillo et al. 1206). P. *S. lycopersicoides* (LA2772). Q. *S. sitiens* (LA2876). Scale bars = 1 cm. LA numbers from the C. M. Rick Tomato Genetic Resources Center.

Figure 1 : Représentation de la diversité de forme et de couleur du fruit au sein du clade de la tomate. Illustration tirée de Peralta, Spooner et Knapp, 2005.

Légende : A. *Solanum lycopersicum* (LA1673). B. *S. pimpinellifolium* (LA1237). C. *S. cheesmaniae* (LA1450). D. *S. galapagense* (LA317). E. *S. neorickii* (LA2200). F. *S. chmielewskii* (LA2663). G. *S. arcanum* (LA2548). H. *S. huaylasense* (LA1365). I. *S. peruvianum* (LA153). J. *S. corneliomulleri* (LA1609). K. *S. chilense* (LA1930). L. *S. habrochaites* (LA1353). M. *S. pennellii* (LA716). N. *S. ochranthum* (Spooner et al. 5000). O. *S. juglandifolium* (Castillo et al. 1206). P. *S. lycopersicoides* (LA2772). Q. *S. sitiens* (LA2876). Scale bars = 1 cm. LA numbers from the C. M. Rick Tomato Genetic Resources Center.

Aujourd'hui, plus de 83.000 accessions de tomates sont stockées dans des banques de semences dans le monde entier (FAO 2010). Les principales collections dans le monde sont : aux Etats-Unis, le Centre des ressources génétiques de tomate en Californie (TGRC) (www.tgrc.ucdavis.edu) et la collection de l'USDA (www.ars.usda.gov), le World Vegetable Center à Taiwan (www.avrdc.org) et plusieurs collections européennes (voir ci-dessous). La mise en place des collections de ressources de tomates a permis de mieux comprendre la répartition de sa diversité à travers le monde (Blanca *et al.*, 2015). Néanmoins, le manque de coordination et les données passeport imprécises sont une limite pour une gestion du matériel génétique. Des efforts sont actuellement déployés pour coordonner les initiatives nationales et internationales. Depuis 2001, le programme de coopération pour les ressources phytogénétiques (www.ecpgr.cgiar.org) organise la collaboration à l'échelle du continent européen pour la conservation à long terme et l'utilisation des ressources génétiques des Solanacées. Ce projet est basé sur un réseau de centres nationaux pour les ressources génétiques de Solanacées, dont la tomate, comme le COMAV (Espagne), le CGN (Pays-Bas), l'INRA (France), l'IPK (Allemagne), l'Institut Vavilov (Russie) entre autres. Ces institutions partagent leurs informations sur le matériel génétique grâce à une base de données centrale (<http://documents.plant.wur.nl/cgn/pgr/tomato/>). Plus récemment, dans le cadre d'un projet de l'Union Européenne (EU-SOL, www.eu-sol.wur.nl), une collection de plus de 6.000 accessions de tomates cultivées a été établie, phénotypée et accompagnée d'une base de données *ad hoc*. En outre, la tomate est une espèce faisant partie du projet de conservation à long terme des espèces végétales, lancé par l'initiative Svalbard Global Seed Vault (Food, 2008).

3 - MARQUEURS MOLECULAIRES ET APPLICATIONS POUR L'AMELIORATION

Les marqueurs moléculaires ont permis la construction de cartes génétiques du génome de la tomate à des densités augmentant avec le progrès dans les technologies et la disponibilité des marqueurs moléculaires. La première du genre fut proposée par Tanksley *et al.* dès 1992. Cela a permis la dissection des caractères quantitatifs en QTL (Quantitative Trait Loci) (Tanksley *et al.*, 1992). Cette stratégie a également ouvert la voie vers la cartographie physique et le clonage moléculaire des facteurs génétiques sous-jacents de quelques caractères quantitatifs (Paterson *et al.* 1991). En outre, les génomes de la tomate et de ses CWR sont diploïdes et colinéaires, ce qui rend la dissection génétique simple. Le premier gène cloné par clonage positionnel a été le gène *Pto*, conférant une résistance à *Pseudomonas syringae* (Martin, Brommonschenkel *et al.*, 1993). Depuis, des croisements interspécifiques avec chacune des espèces sauvages apparentées ont été réalisés. En raison de la faible diversité génétique au sein du compartiment cultivé (Miller et Tanksley, 1990), la plupart des populations de cartographie sont basées sur des croisements interspécifiques entre un cultivar et les espèces sauvages apparentées du groupe *Lycopersicon* (revues de Foolad, 2007 ; Labate *et al.*, 2007) ou à partir de *S. lycopersicoides* et le groupe *juglandifolia*. Cependant, des cartes génétiques basées sur des croisements intra-spécifiques ont également été réalisées et ont prouvé leur intérêt notamment sur les aspects liés à la qualité des fruits (Saliba-Colombani *et al.*, 2001). Toutes ces populations ont permis de caractériser une myriade de gènes majeurs impliqués dans divers traits, notamment liés à la taille et la forme du fruit (Rodriguez *et al.*, 2011).

Rapidement, les stratégies de sélection moléculaire ont été mises en œuvre pour accumuler ou «pyramider» des gènes d'intérêt pour les caractères agronomiques, notamment en utilisant la méthode de rétrocroisement avancé de QTL (AB-QTL) (Tanksley *et al.*, 1996). En utilisant cette approche basée sur un croisement interspécifique entre *Solanum lycopersicum* et *S. pimpinellifolium*, dans lequel des allèles favorables ont été détectés, Tanksley et ses collègues ont montré comment une espèce sauvage pouvait contribuer à améliorer la tomate cultivée (Tanksley *et al.*, 1996). Des lignées d'introgession ou Introgression Lines (IL) provenant de croisements interspécifiques ont également permis de disséquer l'effet des fragments de chromosomes provenant d'un donneur (généralement un parent sauvage) introgressé dans une lignée élite récurrente. Les IL offrent la possibilité d'évaluer les performances agronomiques d'un ensemble spécifique de QTL (Paran *et al.*, 1995). Ces IL ont été utilisées comme base pour la cartographie fine et le clonage positionnel de plusieurs gènes et QTL d'intérêt. La première population d'IL a été développée entre *Solanum pennellii* et *S. lycopersicum*

(Eshed et Zamir, 1995; 2001).

La puissance de détection de QTL a été accrue par rapport à des populations bi-alléliques et a de nouveau été améliorée par la constitution de sous-IL portant des fragments introgressés de plus petite taille. Ce type de descendance a permis d'identifier des QTL pour des caractères liés aux fruits (Causse *et al.* 2004), à la teneur en vitamine C (Stevens *et al.*, 2007), etc. L'introgession d'un QTL identifié dans ces IL a permis aux sélectionneurs d'augmenter le pourcentage de matière sèche soluble (Brix) dans les variétés commerciales ainsi que le rendement (Fridman *et al.*, 2004). D'autres populations ont ainsi été conçues par croisement avec plusieurs autres espèces sauvages, notamment *Solanum pimpinellifolium*, *S. habrochaites* (Doganlar *et al.*, 2002 ; Monforte et Tanksley, 2000; Finkers *et al.*, 2007) et *S. lycopersicoides*. Des lignées d'introgession ont également été utilisées pour disséquer les bases génétiques de l'hétérosis.

4 - CAS PARTICULIERS DES GENES DE LA FORME DU FRUIT

Le processus de domestication et la diversification ultérieure des types de fruits ont conduit à une grande diversité morphologique des fruits de tomate (Figure 2) : taille (de petite à grande) ; forme (ronde, polyédrique, allongée, en poire) ; avec des couleurs allant du rouge au vert, blanc, noir, rose, orange ou jaune. Au contraire, les espèces de tomates sauvages portent de petits fruits rouges ou verts, ronds, avec une faible diversité phénotypique intra-spécifique. Ces observations ont attiré l'attention des scientifiques sur l'hérédité et le modèle de développement de la taille et la forme des fruits chez la tomate. L'influence du chromosome 2 sur la morphologie des fruits a été mise en évidence très rapidement (Butler, 1952). En utilisant des techniques moléculaires, les QTL sous jacents à ce caractère ont été disséqués (Grandillo *et al.*, 1999 ; Lippman et Tanksley, 2001).

Le premier QTL, *fw2.2*, contrôlant la variation du poids du fruit a été cloné (Frary *et al.*, 2000). Il a été suggéré que la diversité de la forme du fruit dans le matériel génétique cultivé peut être largement expliquée par quatre gènes (Rodriguez *et al.*, 2011). L'étude a établi un modèle pour l'évolution de la forme du fruit de la tomate. Ce modèle comprend quatre grandes mutations récemment identifiées (Cong *et al.*, 2008) : *FAS* qui augmente le nombre de loges, la fasciation et la taille du fruit ; *LC*, qui augmente le nombre de loges et la taille des fruits (Muños *et al.*, 2011) ; *Ovate*, qui donne la forme du fruit ovoïde (Liu, van Eck *et al.*, 2002) et *SUN*, qui donne une forme de fruit allongée (van der Knaap *et al.* 2002 ; Xiao *et al.*, 2008) ou la forme de type cordiforme lorsqu'elle est associée à *LC* et *FAS*. La distribution allélique des quatre gènes a été associée à la morphologie, la répartition géographique et des données historiques dans une collection d'accessions cultivées diverses.

Cette étude a établi que la sélection a eu lieu selon cet ordre : *LC* a surgi en premier dès la domestication, suivi par *Ovate* chez *Solanum lycopersicum* puis chez *S. l. cerasiforme* mais dans des populations distinctes. *FAS* a surgi plus tard dans un fond génétique présentant la mutation *LC*. La présence de ces trois mutations dans le matériel génétique d'Amérique latine suggère des mutations précolombiennes. Combinés avec *fw2.2*, ces loci ont fortement contribué à l'augmentation de la taille des fruits lors de la domestication de la tomate. Au contraire, la mutation *SUN* (fruits allongés) n'est pas portée par l'ensemble du matériel originaire d'Amérique latine testé dans cette étude, ce qui suggère que la mutation *SUN* est apparue après la diversification en zone Européenne (probablement en Italie). “L'ironie de tout cela”, dit Steve Tanksley (généticien à l'Université Cornell (Etats-Unis) et précurseur de toutes ces études), "est que toute la diversité des variétés anciennes peut être expliquée par une poignée de gènes. Il n'y a probablement pas plus de 10 gènes mutants qui créent la diversité des variétés que vous voyez " (Borrell, 2009). La domestication et la sélection chez la tomate ainsi que sa diffusion dans le monde entier ont conduit à l'immense diversité des variétés qui caractérisent de nombreuses espèces de plantes domestiquées (Purugganan et Fuller, 2009).

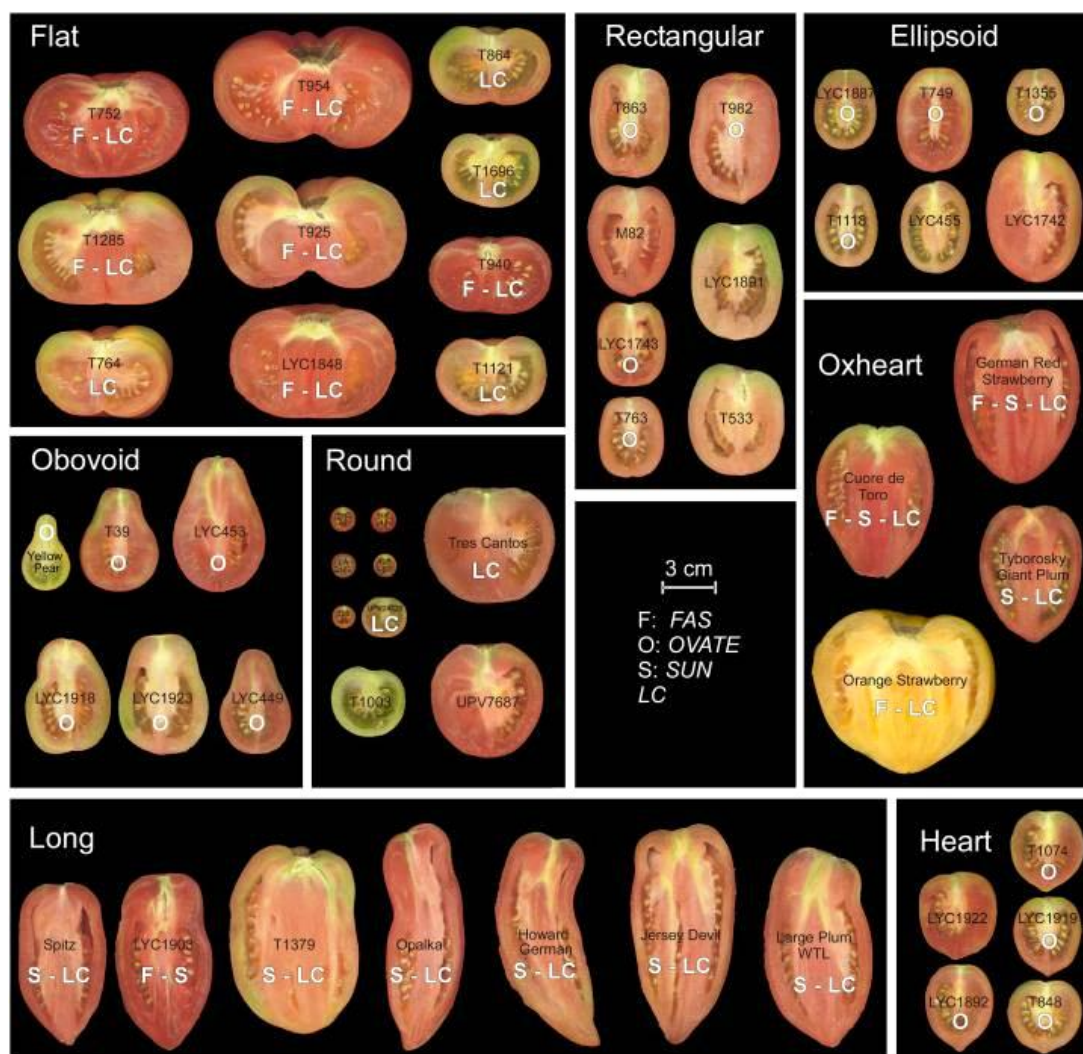


Figure 2 : Catégories de formes de fruits de tomate adaptés de l'UPOV (2001) et l'IPGRI (1996). Chaque fruit est caractérisé (informations tirées de <http://solgenomics.net/>) par la présence des mutations SUN, ovate, LC et/ou des gènes FAS (abrégé en S, O, LC, et F, respectivement).

5 - HISTOIRE EVOLUTIVE, GENETIQUE D'ASSOCIATION ET SELECTION GENOMIQUE : LES NOUVELLES VALORISATIONS DE LA DIVERSITE GENETIQUE A L'ERE DU HAUT DEBIT

Suite au séquençage du génome de la tomate en 2012, les progrès récents de la biologie moléculaire et de la puissance de calcul ont permis d'aborder ou de revisiter de nouvelles questions de recherches notamment liées à l'histoire évolutive de l'espèce, à l'architecture génétique de caractères d'intérêt agronomiques à l'échelle du génome et de proposer de nouvelles pistes d'amélioration génétique telle que la sélection génomique. Aujourd'hui, la masse de données génomiques accumulée chez la tomate est relativement importante avec plus de 500 génomes différents séquencés qui sont libres d'utilisation (Causse *et al.*, 2013 ; Aflitos *et al.*, 2014 ; Lin *et al.*, 2014) mais aussi avec les génomes des CWR, telle que *Solanum pennellii* (Bolger *et al.*, 2014), espèce largement utilisée pour l'introgession de gènes d'adaptation aux stress biotiques et abiotiques. D'autres espèces sauvages sont en cours de séquençage telles que *Solanum arcanum* et *S. chilense* (Cf. projet SOL100 - <https://solgenomics.net/organism/sol100/view>). Ces données génomiques ont également permis le développement d'outils de génotypage tel qu'une puce de génotypage (ou Array) portant plus de 7720 marqueurs de type SNP, identifiés à partir de séquences d'EST. Cette puce, développée par le consortium SolCAP, est largement utilisée dans des études de déséquilibre de liaison (Sim *et al.*,

2012), de diversité génétique (Blanca *et al.*, 2015) ou de génétique d'association (Sauvage *et al.*, 2014).

Jusqu'à présent, l'histoire évolutive de la tomate a été étudiée à l'échelle du génome au travers de deux études (Koenig *et al.*, 2013 ; Lin *et al.*, 2014). Ces études ont notamment permis de mettre en évidence, d'une part, une perte sévère de diversité nucléotidique à l'échelle du génome due au goulot d'étranglement de la domestication, et d'autre part, la détection de la sélection négative ou stabilisante, qui permet de purger les allèles délétères à l'échelle du génome à l'exception d'une cinquantaine de gènes (Koenig *et al.*, 2013). De plus, la domestication n'a pas seulement modifié les patrons de diversité nucléotidique mais aussi les patrons de régulation de gènes avec l'identification de gènes différentiellement exprimés entre le compartiment cultivé et les CWR. Ces données sont également confirmées par les résultats obtenus dans le cadre du projet de recherche 'Arcad' (<http://www.arcad-project.org/>) visant à étudier les conséquences de la domestication au niveau moléculaire par la comparaison des patrons de diversité et d'expression de milliers de gènes entre des accessions cultivées (*Solanum lycopersicum*) et des accessions sauvages (*S. pimpinellifolium*). Dans le cadre de ce projet, les modélisations du scénario de domestication le plus probable chez la tomate ont notamment démontré que le dernier goulot d'étranglement majeur subi par cette espèce date de 420 ans environ et a réduit la diversité nucléotidique génique synonyme de plus de 34 % à l'échelle du génome, malgré des événements de migrations récurrents depuis le compartiment sauvage vers le compartiment cultivé (Sauvage *et al.* en préparation). Enfin, de nombreux balayages sélectifs (ou fixation très rapide d'un haplotype portant une mutation d'intérêt) ont ponctué l'histoire évolutive de la tomate amenant à la fixation d'allèles et/ou d'haplotypes au sein du compartiment cultivé de l'espèce. Ces événements ont eu pour conséquence de permettre la transmission d'allèles d'intérêt pour la sélection moderne chez cette espèce, illustrée notamment pour la couleur du fruit (Lin *et al.*, 2014).

La cartographie d'association tire parti des événements de recombinaison historiques et de la diversité génétique naturelle. En utilisant un grand nombre de lignées et de marqueurs moléculaires répartis sur l'ensemble du génome, la résolution obtenue pour une étude de GWAS est beaucoup plus élevée que dans les populations dites classiques telles que les populations biparentales. Une telle approche nécessite une estimation précise de la structure génétique de l'échantillon étudié et du déséquilibre de liaison (LD) qui s'étend le long des chromosomes. Plus le LD est court, plus il faudra de marqueurs moléculaires pour tester la liaison statistique entre le marqueur moléculaire et le caractère phénotypique étudié. Afin de prendre en compte ces facteurs (structure de population et apparentement), Yu *et al.* (2005) ont proposé un modèle mixte tenant compte de la structure génétique de l'échantillon, basé sur l'analyse des marqueurs moléculaires. Chez la tomate, la première étude d'association a été réalisée par Nesbitt et Tanksley (2002) pour identifier le SNP responsable de la variation du QTL *fw2.2* qu'ils avaient cloné. Ils n'ont pas trouvé de lien entre la taille des fruits et la séquence génomique de la région de *fw2.2* dans une collection de 39 accessions de tomates cerises. Ranc et ses collègues (2010) y ont identifié une association significative dans la région promotrice, grâce à un échantillon représentatif plus large. Du point de vue des caractères de reproduction, la cartographie par croisements interspécifiques entre la tomate cultivée et les CWR est une méthode de choix pour l'identification d'allèles dans le matériel génétique semi-sauvage. Muñoz *et al.* (2011) ont utilisé cette approche pour identifier le polymorphisme causal du QTL contrôlant le nombre de loges sur le chromosome 2.

Depuis, de nombreuses études, principalement menées à l'INRA d'Avignon, ont identifié des locus impliqués dans la qualité du fruit chez la tomate par génétique d'association en augmentant successivement le nombre de marqueurs moléculaires utilisés et en utilisant des modèles linéaires mixtes toujours plus sophistiqués (Ranc *et al.*, 2012 ; Xu *et al.*, 2013 ; Bauchet *et al.* 2014 ; Sauvage *et al.*, 2014). L'ensemble de ces travaux a notamment permis de constituer et de caractériser des core-collections regroupant un maximum de la diversité génétique disponible dans les ressources génétiques conservées au sein de l'INRA. Ces études ont abouti à la caractérisation fine de la structure populationnelle de cette core-collection et du niveau moyen d'apparentement (ou 'kinship'). Ces caractéristiques populationnelles ont alors été incluses dans des modèles linéaires mixtes de génétique

d’association et ont permis d’identifier des marqueurs associés à des caractères de la qualité du fruit telle que la teneur en acide malique, pour lequel deux déterminants majeurs ont été repérés sur les chromosomes 2 et 6. La figure 3 représente les résultats obtenus pour la teneur en acide malique du fruit de tomate tel que décrit dans Sauvage *et al.*, 2014.

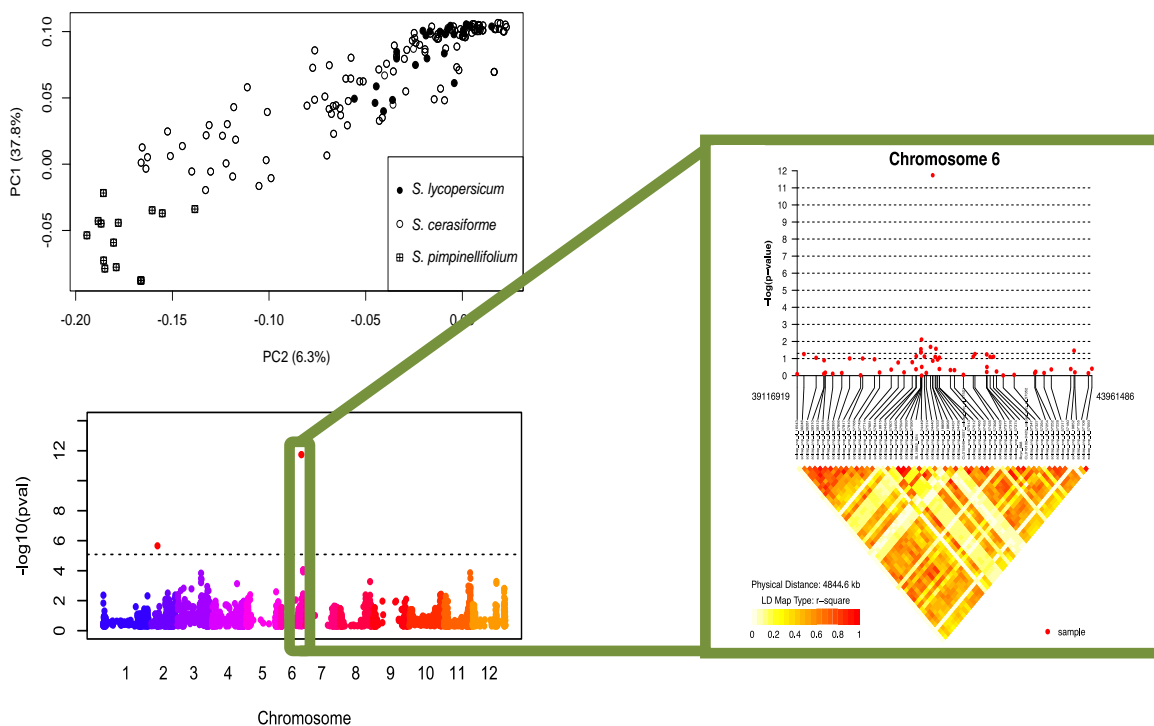


Figure 3 : En haut à gauche, représentation de la structure populationnelle de la population calculée par l’approche ACP (Analyse en composantes principales). En bas à gauche, représentation en ‘Manhattan plot’ (axe X : position chromosomique le long des 12 chromosomes du génome de la tomate ; axe Y, $-\log_{10}$ de la p-valeur du test d’association) identifiant deux loci associés à la teneur en acide malique situés sur les chromosomes 2 et 6. A droite, représentation du déséquilibre de liaison (demi-matrice) autour du locus candidat identifié sur le chromosome 6. D’après Sauvage *et al.*, 2014.

Brièvement, la structuration de la population a été étudiée, par analyse en composantes principales (ACP, sur les marqueurs moléculaires) afin de prendre en compte ce paramètre (Figure 3, haut). Puis un modèle linéaire mixte multilocus (MLMM) a été employé afin d’identifier des associations génotype-phénotype (ici pour la teneur en acide malique) et représenter selon la visualisation graphique de type ‘Manhattan plot’ (Figure 3, bas). Enfin, afin d’identifier des gènes candidats en déséquilibre de liaison avec le SNP identifié par le modèle mixte, les patrons de déséquilibre de liaison de la région chromosomique (Figure 3, droite, demi-matrice rouge et jaune) ont été superposés au manhattan plot de la région étudiée pour répertorier les gènes contenu dans cette région (Figure 3, droite).

En prolongement de la génétique d’association, la sélection génomique a émergé il y a quelques années (Jonas and de Koning, 2013) chez les plantes cultivées et donc chez la tomate. Le principe général de cette approche vise à utiliser les connaissances acquises au sein d’une population (pedigree ou core-collection par exemple) sur les liens génotype-phénotype pour prédire le phénotype d’individus issus d’autres croisements, uniquement sur la base des marqueurs moléculaires. Cela a notamment pour objectifs potentiels de réduire les coûts de phénotypage, d’accélérer le gain génétique pour un caractère polygénique d’intérêt par augmentation de la pression de sélection, diminution du temps de génération et sélection précoce des meilleurs géniteurs ou lignées pour soutenir l’innovation

variétale. Pour ce faire, des équations mathématiques permettent de prédire la valeur phénotypique des individus d'une population dite de 'test' pour lesquels seul le génotype est connu à partir des informations génotypiques et phénotypiques connus chez des individus de la population dite 'd'entraînement'. Chez les bovins laitiers, cette approche a été couronnée de succès avec un gain génétique multiplié par deux. Chez les plantes, ce schéma de sélection n'est pas encore pleinement utilisé. Les travaux récents visent à déterminer la faisabilité de cette approche chez les espèces de grande culture (maïs, blé, palmier à huile), notamment au travers de la prise en compte des interactions génotype x environnement (spécifique des végétaux).

Chez la tomate, des travaux de validation croisée ont été entrepris pour tester la faisabilité de cette approche pour des caractères complexes liés à la qualité du fruit. Les populations d'entraînement et de test reposent sur la collection de ressources génétiques utilisées dans le cadre des travaux de génétique d'association telles qu'utilisées dans Sauvage *et al.* (2014). L'étape de validation croisée repose sur le calcul de corrélations entre le phénotype prédit par les équations mathématiques et la valeur phénotypique réelle mesurée en serre. Brièvement, Duangjit *et al.* (2016) ont démontré que cette corrélation est très dépendante de la nature et de l'héritabilité du caractère considéré, de l'étendue du déséquilibre de liaison, du nombre de marqueurs moléculaires utilisés mais aussi de la composition et de l'apparentement génétique des populations d'entraînement et de test. Cependant, beaucoup d'inconnues restent à résoudre notamment autour de la place de la sélection génomique dans le paysage des filières de production de la tomate et plus largement des plantes potagères et de grande culture.

6 - CONCLUSIONS ET PERSPECTIVES

Au cours des trois dernières décennies, l'apport des marqueurs moléculaires a largement contribué à poursuivre la sélection génétique pour l'amélioration de la tomate au travers de l'identification de QTL liés à des caractères de qualité du fruit ou de résistance à des stress biotiques et abiotiques. Depuis 2012 et l'accès à la séquence complète du génome, des efforts de re-séquençage de lignées cultivées et des espèces sauvages apparentées et les connaissances sur la nature même du génome de la tomate ont fait un bond en avant permettant d'explorer des questions scientifiques à l'échelle du génome, notamment sur son histoire évolutive. Les outils de la génomique moderne ont également permis d'identifier de nouveaux locus impliqués dans des caractères polygéniques par le biais de la génétique d'association. Cette approche tire bénéfice des larges ressources génétiques (banques de graines) disponibles. L'amélioration des plantes repose désormais sur la génétique mendélienne, quantitative et des populations, deux disciplines qui sont restées relativement séparées par le passé, mais qui, en combinaison, permettent de poursuivre le progrès génétique des plantes cultivées.

Enfin, dans un futur proche, de nouveaux types de marqueurs moléculaires, les marqueurs épigénétiques (méthylation de l'ADN), devraient devenir utilisables en amélioration des plantes. En effet, les connaissances scientifiques sur ce type de marqueurs deviennent de plus en plus précises, notamment autour de leur mode de transmission verticale entre générations et de leur rôle dans la régulation de l'expression des gènes. Pour preuve, le locus épigénétique *Cnr*, identifié chez la tomate, est responsable du défaut de coloration du fruit (Manning *et al.*, 2006). L'intégration de ce nouveau type de marqueurs rendra l'amélioration des plantes un peu plus ardue.

“Journée ASF du 4 février 2016”
“ Apport des marqueurs moléculaires et de la génomique à l'utilisation
des ressources génétiques pour l'amélioration des plantes ”

RÉFÉRENCES BIBLIOGRAPHIQUES

- Aflitos, S., E. Schijlen, H. De Jong, D. De ridder, S. Smit *et al.*, 2014 Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *The Plant Journal* 80: 136-148.
- Bauchet, G., M., Causse, 2012. Genetic Diversity in Tomato (*Solanum lycopersicum*) and Its Wild Relatives. *Genetic Diversity in Plants* 1–30 (2012). Edited by Mahmut Çalişkan, ISBN 978-953-51-0185-7.
- Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E. *et al.*, 2015 Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16: 257.
- Bolger, a., f. Scossa, M. E. Bolger, C. Lanz, F. Maumus *et al.*, 2014 The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nat Genet* 46: 1034-1038.
- Borrell, B. (2009). How to Grow a Better Tomato: The Case against Heirloom Tomatoes. *Scientific American* <http://www.scientificamerican.com/article.cfm?id=case-against-heirloom-tomatoes>
- Causse, M. Duffe, P., *et al.* (2004). A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. *Journal of Experimental Botany* 55(403): 1671- 1685.
- Causse, M., N. Desplat, L. Pascual-banuls, M.-C. Le Paslier, C. Sauvage *et al.*, 2013 Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genomics* 14: 791.
- Cong, B., L. S. Barrero, *et al.* (2008). Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nature Genetics* 40(6): 800-804.
- Doganlar, S., A. Frary, *et al.* (2002). Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). *Genome* 45(6): 1189-1202.
- Duangjit, J., Causse, M. & Sauvage, C. Efficiency of genomic selection for tomato fruit quality. *Mol Breeding* 1–16 (2016).doi:10.1007/s11032-016-0453-3
- FAO (2010). Plant genetic resource for food and agriculture. Rome, Food and Agriculture Organisation of the United Nations
- Finkers, R., A. van Heusden, *et al.* (2007). The construction of a *Solanum habrochaites* LYC4 introgression line population and the identification of QTLs for resistance to *Botrytis cinerea*. *Theoretical and Applied Genetics* 114(6): 1071-1080.
- Food, Royal Norwegian Ministry of Agriculture (2008) Svalbard global seedvault - more about the constuction from the directorate of public construction and property. 27.
- Foolad, M. R. (2007). Genome Mapping and Molecular Breeding of Tomato. *International Journal of Plant Genomics* 2007: 1-52.
- Frary, A., Nesbitt *et al* (2000). fw2.2: A quantitative trait locus key to the evolution of tomato fruit size. *Science* 289(5476): 85-88.
- Fridman, E., F. Carrari, *et al.* (2004). Zooming on a quantitative trait for tomato yield using interspecific introgressions. *Science* 305(5691): 1786-1789.
- Grandillo, S., H. M. Ku, *et al.* (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics* 99(6): 978-987.
- Grandillo, S., R. Chetelat, *et al.* (2011). *Solanum* sect. *Lycopersicon* Wild Crop Relatives: Genomic and Breeding Resources. C. Kole, Springer Berlin Heidelberg: 129-215.
- Jonas, E. & de Koning, D.-J. (2013). Does genomic selection have a future in plant breeding? *Trends Biotechnol* 31, 497–504 (2013).
- Koenig, D., J. M. Jimenez-gomez, S. Kimura, D. Fulop, D. H. Chitwood *et al.*, 2013 Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc Natl Acad Sci U S A* 110: E2655-2662.
- Labate, J. A., S. Grandillo, *et al.* (2007). Tomato. Genome mapping and molecular breeding in plants. C. Kole. NY, Springer Publishing. 5: 1-125.
- Lin, T., G. Zhu, J. Zhang, X. Xu, Q. Yu *et al.*, 2014 Genomic analyses provide insights into the history of tomato breeding. *Nat Genet* 46: 1220-1226.

- Lippman, Z. and S. D. Tanksley (2001). Dissecting the Genetic Pathway to Extreme Fruit Size in Tomato Using a Cross Between the Small-Fruited Wild Species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom. *Genetics* 158(1): 413-422.
- Liu, J., van Eck et al (2002). A new class of regulatory genes underlying the cause of pear- shaped tomato fruit. *Proceedings of the National Academy of Sciences* 99(20): 13302- 13306.
- Manning, K., M. Tor, M. Poole, Y. Hong, A. J. Thompson *et al.*, 2006 A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38: 948-952.
- Martin, G., S. Brommonschenkel, *et al.* (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262(5138): 1432-1436.
- Monforte, A. J. and S. D. Tanksley (2000). Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. *Genome* 43(5): 803-813.
- Muños, S., N. Ranc, *et al.* (2011). Increase in tomato locule number is controlled by two SNPs located near WUSCHEL. *Plant Physiology*. 156(4): 2244-2254.
- Nesbitt, T. C. and S. D. Tanksley (2002). Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* 162(1): 365-379.
- Paran, I., I. Goldman, *et al.* (1995). Recombinant inbred lines for genetic mapping in tomato. *Theoretical and Applied Genetics* 90(3): 542-548.
- Paterson, A., S. Damon, *et al.* (1991). Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics* 127: 181-187.
- Peralta, I. E. and D. M. Spooner (2005). Morphological characterization and relationships of wild tomatoes (*Solanum* L. Section *Lycopersicon*). A Festschrift for William G. D'Arcy. T. B. Croat, V. C. Hollowell and R. C. Keating, Missouri Botanical Garden Press. 104: 227-257.
- Peralta, I. E. and D. Spooner (2007). History, origin and early cultivation of tomato (*Solanaceae*). Genetic improvement of *Solanaceous* crops. M. K. Razdan and A. K. Mattoo. Enfield (NH), Science Publisher. 2: 1-24.
- Peralta, I. E., D. Spooner, *et al.* (2007). Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; *Solanaceae*). *Systematic Botany Monographs* 84.
- Peralta, I. E., S. Knapp and D. M. Spooner, 2005 New Species of Wild Tomatoes (*Solanum* Section *Lycopersicon* : *Solanaceae*) from Northern Peru. *Systematic Botany* 30: 424-434.
- Purugganan, M. D. and D. Q. Fuller (2009). The nature of selection during plant domestication. *Nature* 457(7231): 843-848.
- Ranc, N. (2010). Analyse du polymorphisme de gènes de composantes de la qualité des fruits dans les ressources génétiques sauvages et cultivées de tomate ; recherche d'association gènes/QTL. Montpellier, SupAgro. Thèse de doctorat: 348.
- Ranc, N., S. Munos, J. Xu, M.-C. Le paslier, A. Chauveau *et al.*, 2012 Genome-wide association mapping in tomato (*Solanum lycopersicum*) is possible using genome admixture of *Solanum lycopersicum* var. *cerasiforme*. *Genes Genomes Genetics* 2: 853-864.
- Rodriguez, G. R., S. Munos, *et al.* (2011). Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. *Plant Physiology* 156(1): 275-285.
- Saliba-Colombani, V., M. Causse, *et al.* (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics* V102(2): 259-272.
- Sauvage, C., V. Segura, G. Bauchet, R. Stevens, P. T. Do *et al.*, 2014 Genome-Wide Association in Tomato Reveals 44 Candidate Loci for Fruit Metabolic Traits. *Plant Physiol* 165: 1120-1132.
- Sim, S.-C. *et al.* High-Density SNP Genotyping of Tomato (*Solanum lycopersicum* L.) Reveals Patterns of Genetic Variation Due to Breeding. *PLoS ONE* 7, e45520 (2012).
- Stevens, R., M. Buret, *et al.* (2007). Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. *Plant Physiology* 143(4): 1943- 1953.
- Tanksley, S. D., M. W. Ganai, *et al.* (1992). High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132(4): 1141-1160.

Tanksley, S. D., S. Grandillo, *et al.* (1996). Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics* 92(2): 213-224.

The Tomato Genome Consortium, 2012 The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485: 635-641.

van der Knaap, E., Z. B. Lippman, *et al.* (2002). Extremely elongated tomato fruit controlled by four quantitative trait loci with epistatic interactions. *Theoretical and Applied Genetics* 104(2): 241-247.

Xiao, H., N. Jiang, *et al.* (2008). A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit. *Science* 319(5869): 1527-1530.

Xu, J., N. Ranc, S. Munos, S. Rolland, J.-P. Bouchet *et al.*, 2013. Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species. *Theoretical and Applied Genetics* 126: 567-581.

Yu, J., G. Pressoir, *et al.* (2005). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38(2): 203-208.

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nature Review Genetics* 2(12): 983 - 989.

DES ESPECES SAUVAGES A LA SELECTION GENOMIQUE : L'EXEMPLE DE LA BETTERAVE

Karine BOUNAN-HENRY & BRUNO DESPREZ¹

Florimond Desprez Veuve & Fils SAS
3, Rue Florimond Desprez, BP 41, F-59242 CAPPELLE-EN-PEVELE
E-mail : bruno.desprez@florimond-desprez.fr

RESUME

La betterave à sucre est cultivée depuis Napoléon et a considérablement progressé en termes de rendement sucre par hectare (x 20), mais aussi sur de nombreux caractères d'adaptation aux stress biotique et abiotique, à la nutrition azotée, etc. Malgré un taux quasi constant d'amélioration de 2 % de sucre/ha/an, le risque d'une perte de compétitivité est bien présent. La forte pression de sélection et le nombre important de caractères à accumuler dans le matériel élite invitent le sélectionneur à la plus grande prudence, en particulier au regard d'une perte importante de variabilité. Le programme AKER qui est ici décrit tente de sécuriser les processus d'amélioration en explorant une stratégie originale de l'apport et de l'exploitation des ressources génétiques. Après une présentation des principes et du schéma général d'AKER, les premiers résultats sont ici décrits. Les analyses réalisées dans ce cadre montrent que 40 plantes suffisent à représenter la variabilité allélique totale disponible du genre *Beta*, section *Beta*. De même, 15 plantes permettent la conservation de toute la variabilité complémentaire de celle disponible au sein du compartiment des betteraves à sucre cultivées. Les 15 plantes de la collection de référence AKER, une fois séquencées, ont été croisées et recroisées pour introgresser leur génome par morceaux dans du matériel élite, à raison de 200 plantes par descendance. Ainsi 3.000 plantes ont été obtenues et feront l'objet d'une analyse phénotypique. L'analyse informatique devrait permettre d'associer les compositions génétiques aux phénotypes et ainsi révéler la variabilité vraiment utile tout en fournissant les outils nécessaires à son exploitation. Plusieurs collaborations, exploitations de ce programme sont également évoquées en conclusion.

Mots-clés : ressources génétiques, marqueurs moléculaires, *Beta*, génomique, diversité allélique, collection de référence, génomique du paysage

1 - DE LA GESTION DE LA VARIABILITE AUX RESSOURCES GENETIQUES

Deux cents ans d'amélioration de la betterave à sucre ont permis de passer d'une production de sucre à l'hectare de 700 kg du temps de Napoléon 1er, à l'origine de la culture, à une production aujourd'hui de plus de 20 fois supérieure (avec une teneur en sucre de 18 % au lieu de 7 %). Ces progrès considérables continuent à un rythme soutenu de 2 % par an. La betterave d'aujourd'hui résiste à de nombreuses maladies, s'adapte à des environnements variés, réagissant parfaitement aux changements climatiques et ceci en consommant de moins en moins d'engrais (moins 45 % par tonne produite, en 20 ans). Le bilan environnemental de la betterave est particulièrement remarquable. Depuis 1983, on a ainsi observé une baisse de 50 % des quantités de produits phytosanitaires utilisés

¹ Karine Bounan-Henry est Responsable de la sélection Betterave sucrière et Fourragère et Chicorée, Leader du Groupe de Travail WP2 et Coordinatrice scientifique du Programme AKER ; Bruno Desprez est Président de Florimond-Desprez et Président du Comité de Coordination du Programme AKER

et la tendance continue. Il faut aussi compter sur les améliorations au niveau du procédé industriel, avec des extractions d'un sucre de qualité bien supérieure. La consommation énergétique à la tonne de betterave a été presque divisée par deux en quarante ans. De même, l'importante quantité d'eau contenue dans la betterave (80 %) fait qu'une sucrerie est globalement excédentaire en eau. Celle-ci est recyclée en étant de plus en plus utilisée dans différentes étapes industrielles (lavage, diffusion...). C'est l'eau même contenue dans la betterave qui va servir au procédé industriel d'extraction et de purification du sucre.

L'amélioration génétique demeure un élément essentiel dans toute cette progression, même si elle n'est pas le seul. Elle ne peut cependant se poursuivre que si une variabilité suffisante reste disponible. Examinons tout d'abord comment s'est déroulée la gestion des ressources génétiques au cours de l'amélioration de la betterave à sucre.

Historiquement la betterave sucrière a été sélectionnée au début du XIX^e siècle à partir d'une population de betteraves que nous pourrions qualifier de pré-sucrière, dénommée la betterave de Silésie et de différents types de betteraves fourragères. Philippe André de Vilmorin (1776-1862) avait d'ailleurs prospecté et collectionné différentes ressources génétiques, y compris des betteraves sauvages. Il s'était attaché à essayer de comprendre l'évolution de la teneur en sucre, collectant des betteraves sauvages maritimes mais aussi des variétés fourragères de toutes formes et de toutes couleurs. Jusqu'à la fin de la première guerre mondiale, la variabilité disponible dans les betteraves cultivées était suffisante pour améliorer par sélection massale puis par sélection sur descendance ces différents types de populations. L'intégration systématique des ressources génétiques autres que celles des betteraves dites cultivées était très faible, voire inexistante.

Avec le chercheur italien Muneratti en 1930, l'exploitation des ressources génétiques devient une véritable stratégie d'amélioration de la betterave. Il fut en effet le premier à prospecter systématiquement les betteraves sauvages de la côte adriatique, des betteraves maritimes ou *Beta maritima*, avec pour objectif de trouver des sources de résistance aux maladies foliaires dont le *Cercospora*, une maladie fongique fort répandue dans le contexte méditerranéen et dans les milieux à climats continentaux. Il le fera avec succès. La collaboration de la station de Rovigo dans la vallée du Pô avec les chercheurs américains fera école et cette approche deviendra dès lors quasi systématique pour les maladies de la betterave.

L'utilisation des ressources génétiques était ainsi relativement simple et efficace. Elle était et est toujours réalisée suivant le même principe : les ressources génétiques sont dans un premier temps phénotypées, criblées pour la tolérance ou la résistance aux maladies, puis s'en suit une série de rétrocroisements pour introduire celle-ci dans du matériel élite. Cette méthode a pour avantages la simplicité et sa grande efficacité pour les caractères monogéniques, ce qui est souvent le cas des résistances aux maladies. Mais elle a aussi pour inconvénient d'entraîner un phénotypage long et très coûteux qu'il faut appliquer « au hasard » sur le plus grand nombre possible d'accessions. Le phénotypage étant très onéreux, son application à grande échelle pose ainsi des problèmes de débit, de coût, mais aussi de précision. Cette méthode reste également très difficile, voire non utilisable, pour les caractères polygéniques tels que le rendement. Enfin, et ce n'est pas anodin, le phénotypage peut être biaisé par les effets du génome d'origine. Il est fréquent d'observer une « résistance », par exemple sur du matériel sauvage, et s'apercevoir ensuite que cette caractéristique n'est finalement pas transmise lors de rétrocroisements dans du matériel cultivé élite. La plante peut « éviter », « échapper » à la maladie du fait par exemple de sa conformation racinaire chevelue ou de ses feuilles épaisses ou de son port prostré. Dès l'arrivée dans un contexte avec une racine plus « pivotante », la maladie se déclare...

Les marqueurs moléculaires peuvent bien sûr aider au transfert, encore faut-il qu'ils soient disponibles, mais c'est par essence rarement le cas, du fait qu'ils sont en effet mis au point après détection et croisements... Avec cette approche en sélection assistée par marqueurs, l'approche QTL (*Quantitative Trait Loci*), plus élaborée, ne permet cependant de ne capturer qu'une fraction de la variance génétique, e.g. 10 QTL ~ 50 % de la variance génétique. Dans cette approche, que l'on peut

qualifier de « traditionnelle », on constate une réelle difficulté dans l’optimisation du phénotypage et du génotypage. Les coûts très élevés du phénotypage limitent ainsi fortement l’apport des ressources génétiques.

La trop forte pression de sélection devenue souvent nécessaire par l’accumulation des traits de performance est, quant à elle, très réductrice de cette variabilité déjà difficile à alimenter. La variabilité exploitable du matériel élite décroît fortement, aboutissant à des croisements de plus en plus consanguins. L’effet « entonnoir » est ainsi notable. Tout cela nous invite à la prudence sur la capacité des sélectionneurs à conserver cette croissance soutenue de 2 %.

Le programme d’investissements d’Avenir dénommé AKER (2012-2020), s’est ainsi donné pour objectif de conserver, et même d’essayer de doubler à terme, le taux de progrès annuel. L’objectif étant donc de passer de 2 à 4 % de progrès par an. Le levier génétique a de suite été identifié comme le principal levier à actionner et c’est pourquoi nous avons essayé de proposer une nouvelle approche pour réincorporer de la variabilité (utile si possible, bien sûr) pour éviter le fameux effet « entonnoir », tout cela en incorporant les progrès réalisés dans le génotypage et le phénotypage. C’est ce programme de 8 ans (2012-2019), doté d’un budget total de plus de 18,5 millions d’euros, réalisé en collaboration avec 11 partenaires français, qui est maintenant détaillé avec ses premiers résultats.

2 - LA STRATEGIE AKER DANS L’EXPLOITATION DES RESSOURCES GENETIQUES

L’introduction de nouvelles sources de diversité génétique au sein du germplasm élite est une des approches les plus prometteuses pour l’amélioration génétique des variétés cultivées, et celle de la betterave en particulier, pour les caractères d’intérêt que sont le rendement en sucre ou la résistance aux stress biotiques ou abiotiques. Le projet AKER propose d’élargir la base génétique des variétés actuelles grâce à la découverte de nouveaux allèles favorables aux caractères spécialement ciblés dans AKER.

En comparaison avec les techniques actuelles de sélection, le projet AKER se veut un projet innovant et ambitieux. Contrairement à l’approche traditionnelle, AKER se propose de génotyper les plantes avant de les phénotyper. Le phénotypage étant beaucoup plus cher que le génotypage, il faut maximiser son intervention, en particulier en n’intervenant que sur un nombre choisi de plantes et le plus tard possible dans le processus pour éviter la découverte de faux positifs. Le processus qui consiste à croiser les plantes exotiques avec les plantes élites sera donc largement entamé lorsque l’on appliquera le phénotypage. Réaliser le génotypage en premier permettra ainsi de sélectionner des allèles favorables cachés ne s’exprimant pas dans le fond génétique exotique. Les allèles non-héréditaires dans le contexte élite seront aussi évités.

AKER a aussi pour objectif de maximiser la connaissance du génome exotique et de ses apports, au cours du processus d’introduction du matériel sauvage. Pour l’approche retenue, il faut donc améliorer en parallèle la qualité du diagnostic et du phénotypage, notamment sa précision et son débit, ceci pour exploiter au maximum les données moléculaires qui seront générées dès le début du processus. Ce phénotypage n’interviendra alors qu’à la fin du processus, de manière optimale à partir de techniques à haut-débit si possible et non destructrices.

La betterave appartient au genre *Beta* qui comprend quatre sections, dont trois exclusivement composées de betteraves sauvages/exotiques. La première section *B. vulgaris* quant à elle comprend deux groupes : d’une part, les betteraves cultivées (sucrières, fourragères, potagères...) et d’autre part les betteraves maritimes (*B. maritima*) présentes sur les côtes européennes et les betteraves *B. adanensis* et *B. macrocarpa*. C’est ce que l’on appelle le Groupe I.

Nous n’avons donc considéré dans le programme AKER que les espèces et sous-espèces du Groupe I, dans lequel nous allons chercher à exploiter une variabilité qui se vaudra différente de celle exploitée jusqu’à présent. Le schéma général du programme est représenté ci-dessous (Figure 1).

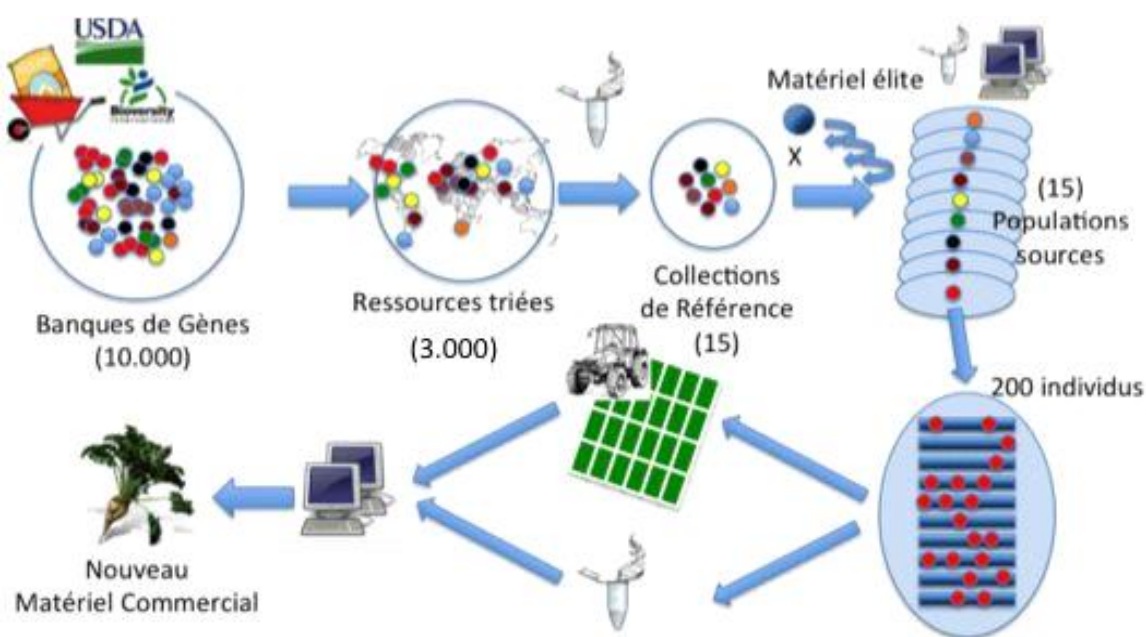


Figure 1 : Schéma général du programme AKER.

La première étape du programme AKER consiste à recenser toutes les semences de betteraves du Groupe I présentes dans les 50 banques de gènes identifiées. Nous avons recensé 20.000 sachets ou lots de semences. Après nettoyage de ces collections (retrait des doublons, des lots de semences non-germantes, des sachets non suffisamment renseignés par ce que l'on appelle les données passeport (latitude et longitude des lieux d'origine), etc.), nous sommes parvenus à un total de 10.000 accessions ou lots de semences.

La seconde étape consiste à placer les accessions sur une carte géographique et en sélectionner 1/3 (soit environ 3.000 accessions) en respectant la distribution géographique et la représentativité des différentes origines (sauvage, cultivée, rouge de table, carde, fourragère, sucrière...). Après une étude génétique réalisée sur ces 3.000 accessions « mondiales » à l'aide de très nombreux marqueurs, nous avons essayé de représenter le maximum de variabilité avec le nombre minimum de betteraves. Nous avons initialement calculé qu'en théorie une trentaine de plantes devrait suffire à représenter la totalité de la variabilité allélique des accessions « mondiales ». De même, nous avons estimé que 15 plantes permettent la conservation de toute la variabilité complémentaire de celle disponible au sein du compartiment des betteraves à sucre cultivées. L'objectif est alors de séquencer entièrement ces 15 plantes constituant ce que l'on a appelé la collection de référence AKER.

La troisième étape du programme AKER consiste à prendre chacune de ces 15 plantes et à les croiser séparément avec du matériel élite. Après croisements et recroisements, l'objectif est de pulvériser le génome de chacune des 15 plantes exotiques d'origine en une multitude de morceaux, les plus petits possibles et répartis dans 200 plantes descendantes. L'ensemble du génome de chacune des plantes exotiques est ainsi totalement représenté mais distribué dans différentes plantes avec un même fond génétique élite. Un marquage moléculaire dense, ainsi que le séquençage des accessions, a permis la constitution de cette collection de référence et permettra également de suivre de manière précise les introgressions dans le matériel élite.

L'étape ultime consiste à phénotyper les 3.000 plantes ainsi obtenues et par la connaissance précise de leur génome, à combiner les données pour assigner à chaque morceau une valeur pour un phénotype donné. Le phénotypage n'intervient ainsi qu'à la fin du programme, de matière optimale, à partir de techniques haut-débit invasives ou non destructives. La variabilité sera ainsi mieux

caractérisée et la construction de nouvelles variétés élite pourra se réaliser par combinaisons de ces « bons morceaux » évitant les mauvais... La création variétale revient ainsi au coeur du processus de l'amélioration variétale.

Nous nous proposons, après une rapide présentation du matériel génétique utilisé dans AKER, d'examiner plus précisément à la fois la méthodologie de chacune de ses étapes mais aussi les premiers résultats obtenus.

3 - LE CHOIX DU MATERIEL GENETIQUE

La betterave sucrière appartient à la famille des Amaranthacées. Le genre *Beta*, aujourd'hui divisé en 4 sections inclut des formes cultivées (fourragères, potagères, sucrières et cardes), comme des formes sauvages. La section *Beta* contient trois espèces et trois sous-espèces. *Beta macrocarpa* est une espèce annuelle, autofertile, distribuée sur l'ensemble du bassin méditerranéen et dans les îles Canaries. *Beta patula* est une espèce pérenne, autofertile, endémique d'une petite île au large de Madère. *Beta vulgaris* est subdivisée en 3 sous-espèces : *Beta vulgaris* subsp. *maritima* est une espèce pérenne (parfois annuelle), largement auto-incompatible avec une aire large de distribution allant des côtes atlantiques d'Europe de l'Ouest (Scandinavie aux Açores), aux côtes méditerranéennes, et jusqu'au Moyen-Orient et l'Asie. *Beta vulgaris* subsp. *Adanensis* est une espèce annuelle, autofertile, distribuée essentiellement autour de la mer Egée (Chypre, Turquie, Grèce, Israël...). *Beta vulgaris* subsp. *Vulgaris* regroupe l'ensemble des espèces cultivées.

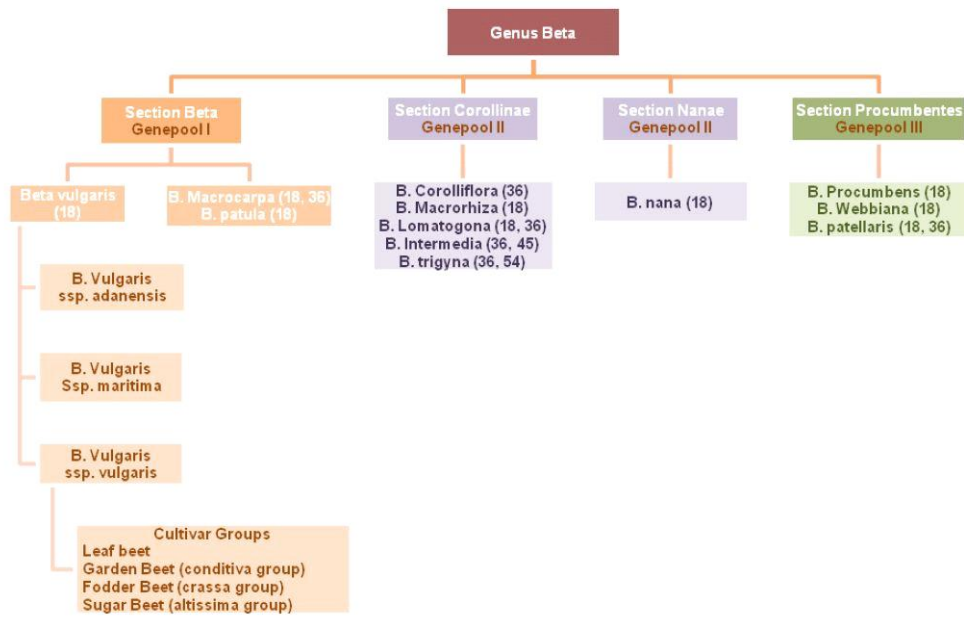


Figure 2 : Description taxonomique du genre *Beta*.

La première partie du programme a consisté à recueillir toutes les informations provenant des bases de données européenne (Eurisco) et américaine (GRIN-ARS), pour constituer une première liste de 10.000 accessions. Sur la base des informations éco-géographiques (données passeport : latitude, longitude, données taxonomiques) contenues dans cette liste, 3.000 accessions ont été dans un premier temps sélectionnées de manière à représenter l'ensemble des espèces disponibles (sauvages et cultivées) et l'ensemble des zones géographiques possibles. Nous avons essayé autant que possible de couvrir toutes les zones d'origine de la betterave sauvage ainsi que les zones de culture actuelles de la betterave sucrière. Pour les betteraves exotiques, nous avons pu collecter des accessions originaires

des côtes atlantiques (du Danemark au Maroc), des côtes méditerranéennes (du Maroc jusqu’à Israël et la Turquie). Pour les betteraves cultivées, la plupart des régions cultivées sont représentées : Europe (Ouest, Est, Sud, Centre), Asie, Proche-Orient et USA.

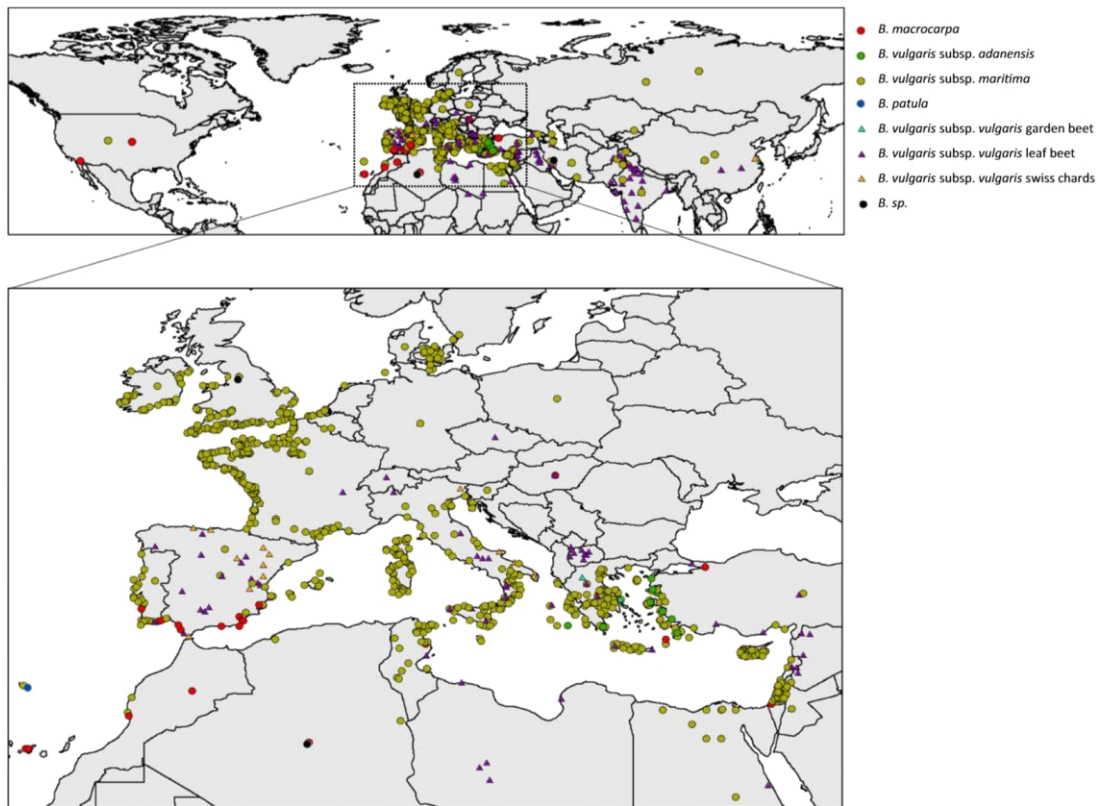


Figure 3. Distribution géographique des 3.000 accessions sélectionnées.

A partir de l’ADN des accessions reçues, des marqueurs DArTs (*Diversity Arrays Technology*) ont été générés selon la technique décrite par Kilian *et al.* (2012). Au total 2.680 ADN extraits, représentant le matériel analysé, ont été utilisés pour construire les bibliothèques avec un total de 15.360 clones DArT. Nous avons utilisé l’ensemble des accessions pour la construction des bibliothèques pour éliminer toute forme de biais (le problème majeur de la plupart des marqueurs SNPs (*Single Nucleotide Polymorphism*) est qu’ils ont été développés à partir d’un petit échantillon de matériel, essentiellement élite et ne sont donc pas forcément représentatifs du pool exotique). Au final, des données de génotypage ont été générées pour 2.680 accessions et 4.497 marqueurs DArTs. En parallèle, les accessions ont également été analysées avec un set interne de 378 marqueurs SNP (technologie Kaspar).

4 - METHODES D’ANALYSE DE LA DIVERSITE

La diversité génétique au sein des accessions sélectionnées a été analysée à 2 niveaux : la variabilité inter-accessions et la variabilité intra-groupe d’espèces. Le but de cette approche était de caractériser à l’aide d’un ensemble de paramètres, la variabilité de la population totale et la proximité génétique des populations. Pour décrire la diversité génétique intra-groupe, divers paramètres ont été calculés à l’aide du logiciel *GenAlex* : le taux de polymorphisme, le nombre moyen d’allèles par locus, le taux moyen d’hétérozygotie observé et attendu, l’indice de diversité de Shannon...

Nous avons utilisé deux approches complémentaires pour caractériser la structure génétique des accessions analysées et mettre en évidence de potentiels groupes d’accessions : (i) une analyse discriminante en composantes principales (Jombard *et al.*) ; (ii) une approche de type clustering Bayésien avec le logiciel *Structure*. Les deux analyses ont été utilisées pour vérifier la consistance du clustering obtenu.

Nous avons ensuite complété les analyses en estimant la diversité génétique entre taxa puis entre les groupes identifiés par l’analyse précédente. La phylogénie a été reconstruite en utilisant l’approche *Neighbour Joining* (NJ) à l’aide du logiciel *Darwin* version 5.0.155 (Perrier *et al.*, 2006). Nous avons également utilisé le logiciel *DIVA-GIS* (X. Scheldeman, M. van Zonneveld, 2012) pour réaliser une analyse spatiale de la diversité ; une méthode de raréfaction permettant de compenser la distribution inégale des accessions, a été utilisée pour l’analyse spatiale de la richesse allélique. La méthode de raréfaction recalcule la richesse allélique mesurée dans les différentes cellules géographiques, comme si le même nombre standard d’observations avaient été effectué.

Finalement, l’ensemble de ces données a été compilé et utilisé *a posteriori* pour choisir la collection de référence de 15 accessions. En effet, les données DArTs et SNP ont été introduites dans le logiciel *CoreHunter* (H. De Beukelaer, G. Davenport, 2012). Les algorithmes disponibles permettent la constitution d’une collection de référence à partir de sept mesures de diversité : deux mesures de distance génétiques, trois indices de diversité allélique et deux mesures auxiliaires. Nous avons dans un premier temps utilisé l’ensemble des données de génotypage pour estimer le nombre d’accessions nécessaires pour couvrir 100 % de la diversité allélique totale, à l’aide de l’indice CV (indice représentant le pourcentage d’allèles conservés entre la collection initiale et la collection de référence). En parallèle, les mêmes analyses ont été réalisées à partir de la liste des allèles complémentaires du matériel élite. Plusieurs jeux de collections de référence représentant 100 % de la diversité allélique complémentaires du pool élite, ont ainsi été générés : la collection finale a été choisie parmi ces différents jeux, de manière à respecter la diversité taxonomique et la diversité géographique.

5 - METHODES D’INTROGRESSION DE LA COLLECTION DE REFERENCE DANS LE MATERIEL ELITE

Les 15 accessions sélectionnées ont été introduites dans le matériel élite de manière à pulvériser le génome de chaque accession en plusieurs fragments distribués sur environ 200 plantes. L’objectif est de créer des populations de type *AB-QTL* avec les propriétés suivantes : bonne puissance de détection de QTL, résolution maximale, 100 % de couverture du génome sauvage, taille de segment exotique minimale. Pour ce faire, des simulations ont été réalisées de manière à optimiser les nombres de plantes utilisées à chaque génération.

Suite à la création de la collection de référence, chaque accession est croisée avec un génotype récurrent élite, jusqu’au développement de 15 populations de BC2S1 (Figure 4). Le génotype élite, une ligne HD développée avant AKER, contient la stérilité mâle génique pour faciliter chaque cycle de croisements. La sélection à chaque génération des plantes BC1, BC2 et BC2S1, s’appuie sur les données de génotypage afin d’obtenir des plantes ayant des profils différents de gènes exotiques. Ainsi, chaque génotype BC2S1 comprendra un même patrimoine génétique de l’élite et des segments d’un génome sauvage donneur tels que le génome de l’exotique soit représenté intégralement dans la population de BC2S1 (à l’exception de certains gènes sauvages qu’il faut enlever afin de mesurer les caractéristiques d’une betterave cultivée). Un test-cross sera réalisé en fin de cycle à partir des BC2S1 et les 3.000 hybrides obtenus seront évalués au champ.

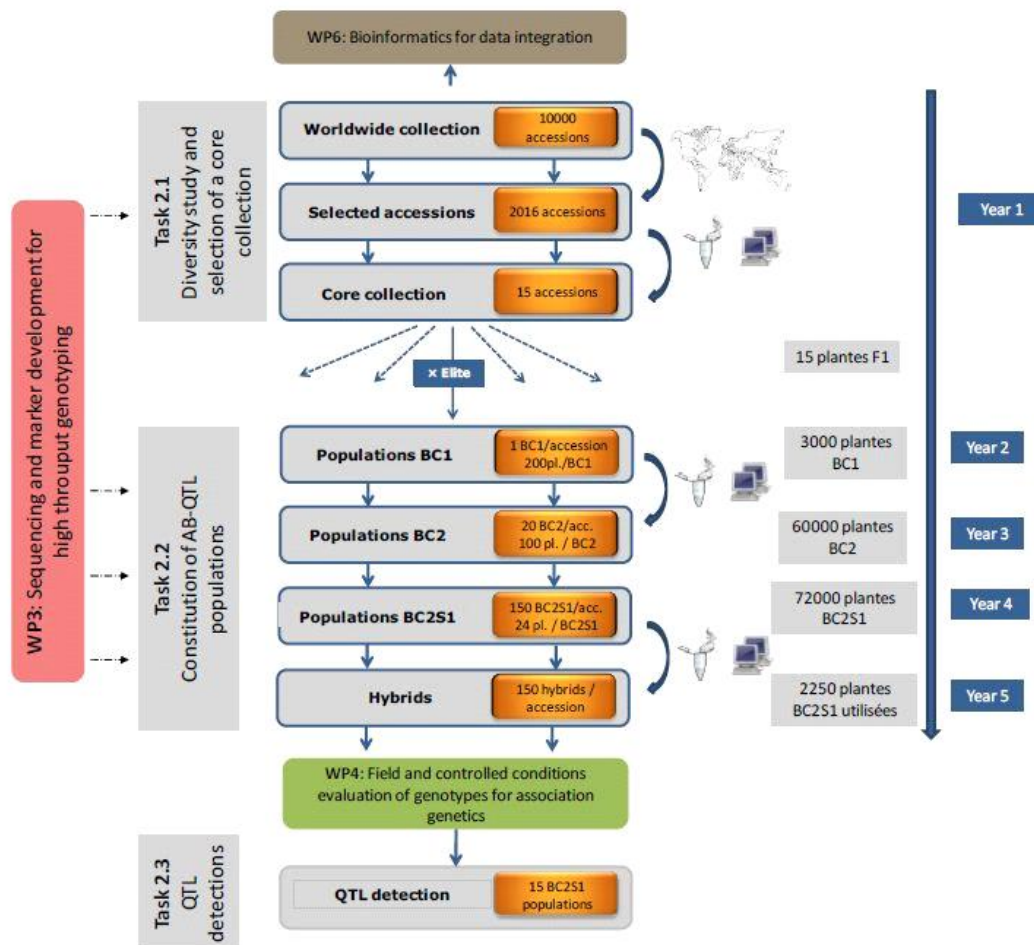


Figure 4 : Schéma d’introgression de la collection de référence dans le matériel élite.

6 - PRINCIPAUX RESULTATS

La question qui nous intéresse au sein d’AKER et à laquelle l’analyse de diversité doit permettre de répondre est : « Est-ce que les ressources génétiques sélectionnées sont des sources de diversité allélique ou pas, par rapport notamment aux populations cultivées élites ? ». Nous nous sommes donc attachés à :

- premièrement, décrire la diversité entre populations à divers degrés, en prenant en compte les sous-groupes tels que les espèces/sous-espèces, les pays... ;
- deuxièmement, calculer les relations deux à deux entre les populations à partir de calculs de distance génétique, puis exprimer ces relations avec des méthodes de classification ou d’ordination ;
- troisièmement, combiner les résultats des données moléculaires aux données géographiques, afin de mettre en évidence des points chauds de richesse allélique ;
- Finalement, combiner toutes ces informations afin de constituer une collection de référence. Il y a différentes façons de constituer une collection de référence : différents algorithmes, différents critères de constitution, clustering *a priori* ou *a posteriori*...

Dans AKER, nous avons fait le choix d’utiliser comme critère principal, la richesse allélique complémentaire de la richesse allélique élite. Plusieurs répétitions ont été réalisées, le choix final s’est fait en combinant la fréquence d’apparition des accessions dans les répétitions, et les données

passport et biologiques pour éviter à richesse allélique égale de n’avoir que des *maritima* ou cinq accessions de la même région géographique.

Les analyses génétiques réalisées sur les 2.600 accessions ont ainsi permis une étude très fine de leur diversité, que nous présentons ici.

6.1. Structure génétique et spatiale

Les différentes analyses effectuées mettent en évidence 4 à 10 clusters ainsi qu’un nombre élevé de populations admixés (33 %). Les analyses réalisées à partir des données DArTs mettent en évidence un nombre plus élevé de clusters (9 clusters *versus* 4 clusters pour les données SNP) et permettent une meilleure discrimination des différents taxons.

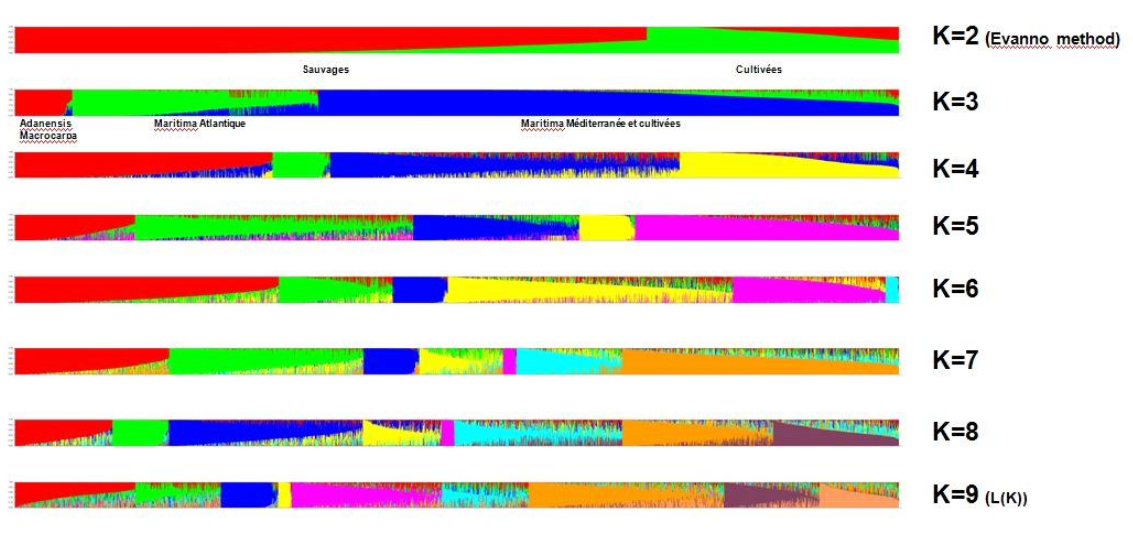


Figure 5 : Représentation graphique des résultats obtenus avec Structure pour k=2 à 9.

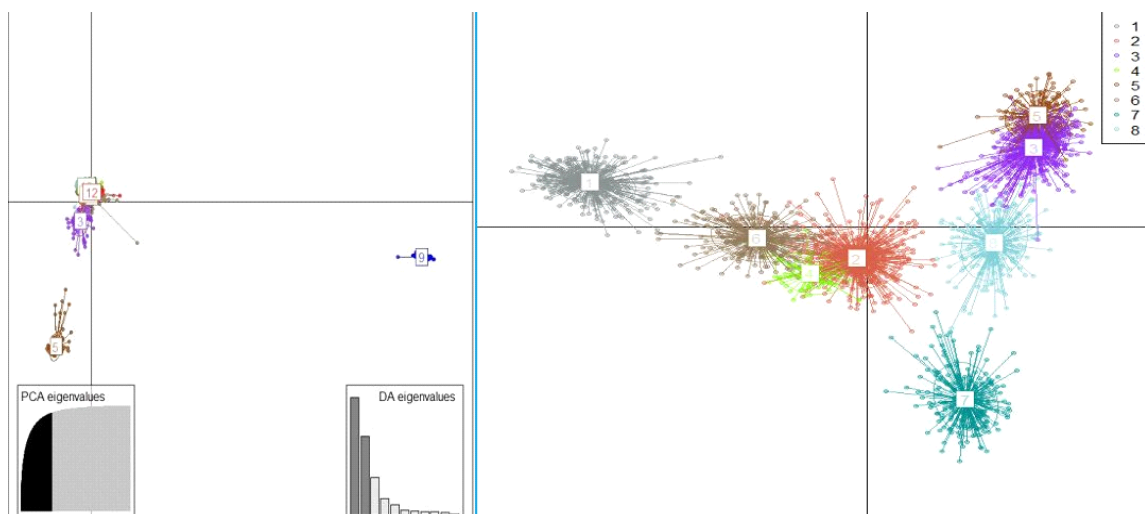


Figure 6 : DAPC (Adegenet) avec k=10 (à gauche), k=8 (après avoir enlevé les groupes 9 et 10 très distants).

Le cluster 9 est principalement composé d’accessions *Beta macrocarpa* et de quelques accessions *Beta vulgaris* subsp. *maritima*, probablement mal répertoriées. Le cluster 10 est composé par la majorité des accessions *Beta vulgaris* subsp. *adanensis*. Les autres clusters sont séparés principalement en deux sous-groupes : le premier inclut les clusters 1, 2, 4 et 6, correspondant principalement aux accessions *Beta vulgaris* subsp. *maritima* d’origine atlantique et le deuxième, les clusters 3, 5, 7 et 8, correspondant aux accessions cultivées (fourragères, potagères, sucrières et cardes) et *Beta vulgaris* subsp. *maritima*, d’origine méditerranéenne.

L’arbre phylogénétique reconstruit à partir de la méthode *Neighbour Joining*, donne une image très similaire (cf. figure 7).

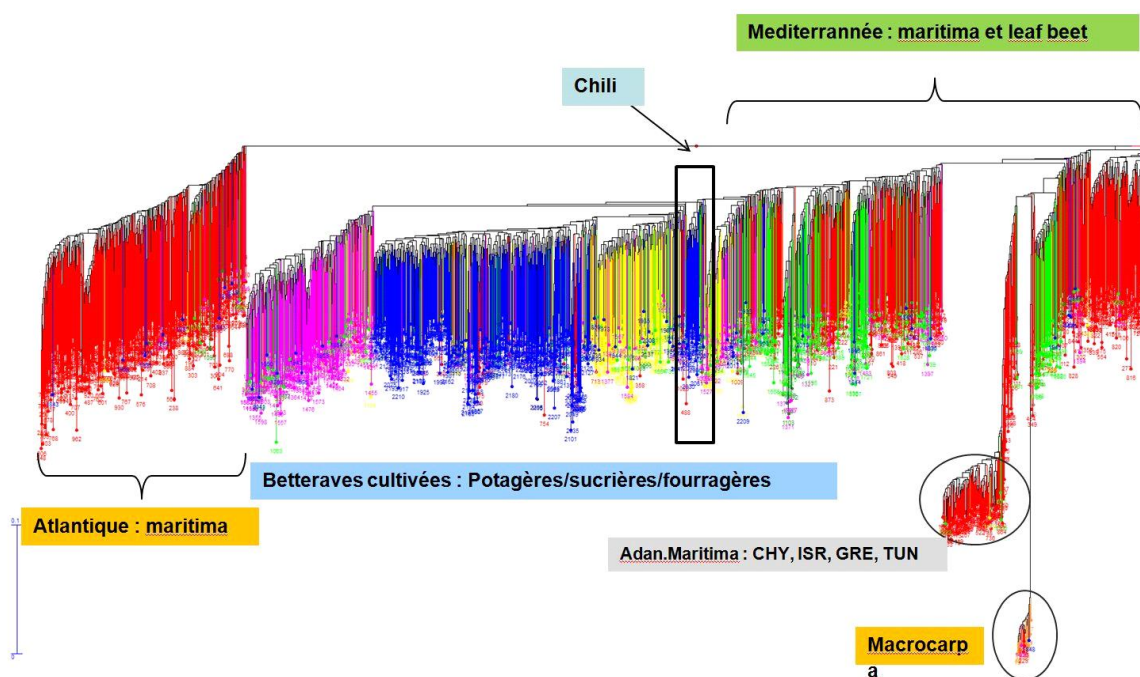


Figure 7. Arbre phylogénétique établi à partir des marqueurs DArTs.

Les taxons *Beta macrocarpa* et *Beta vulgaris* subsp. *adanensis* sont les plus éloignés de la betterave cultivée avec un indice de diversité intra-groupe relativement faible dû à leur mécanisme de reproduction exclusivement autogame. Le taxon *Beta vulgaris* subsp. *maritima* est le plus proche du pool élite.

La représentation géographique des clusters identifiés dans l’analyse discriminante pour *Beta vulgaris* subsp. *Maritima* (10 clusters) montre un gradient de différenciation spatiale très net. La distribution des clusters dans le plan défini par les 2 premières fonctions discriminantes, reproduit parfaitement leur distribution géographique. Le cluster le plus différencié est celui incluant les accessions asiatiques.

La différenciation génétique des accessions du genre *Beta* section *Beta* repose donc essentiellement sur la distinction entre *macrocarpa*, *adanensis* et les autres taxons. *Beta macrocarpa* est le taxon le plus éloigné. *Beta adanensis* est également très éloigné du reste mais montre une distance plus proche du groupe méditerranée contrairement à *Beta macrocarpa* qui est à égale distance des groupes méditerranée et atlantique. Les autres taxa forment un groupe modérément différencié qui inclut *Beta vulgaris maritima*, *Beta patula* et les formes cultivées.

Au sein du groupe des betteraves *maritima* et cultivées, la variation génétique est plus structurée selon la distribution spatiale que taxonomique avec l’émergence de deux groupes distincts : méditerranée et atlantique (Tableau 1 ; Figure 8). Ce gradient de variation génétique spatiale témoigne de la combinaison d’effets de dérives génétiques locales et de flux de gènes limités. La structure génétique sur de grandes distances n’a pas pu être corrélée avec les courants marins mais résulterait plutôt d’un modèle d’isolation par distance sans discontinuité.

Tableau 1. Distance génétique entre taxons.

Species	Macrocarpa	Patula	Adanensis	Maritima	Elite	Fodder beet	Garden beet	Leaf beet
Patula	0.309	0.000						
Adanensis	0.356	0.264	0.000					
Maritima	0.228	0.143	0.116	0.000				
ELITE	0.277	0.139	0.209	0.080	0.000			
Fodder beet	0.313	0.117	0.203	0.069	0.046	0.000		
Garden beet	0.292	0.137	0.210	0.066	0.053	0.019	0.000	
Leaf beet	0.269	0.127	0.154	0.032	0.051	0.024	0.025	0.000
Sugar beet	0.312	0.115	0.219	0.075	0.029	0.015	0.031	0.036

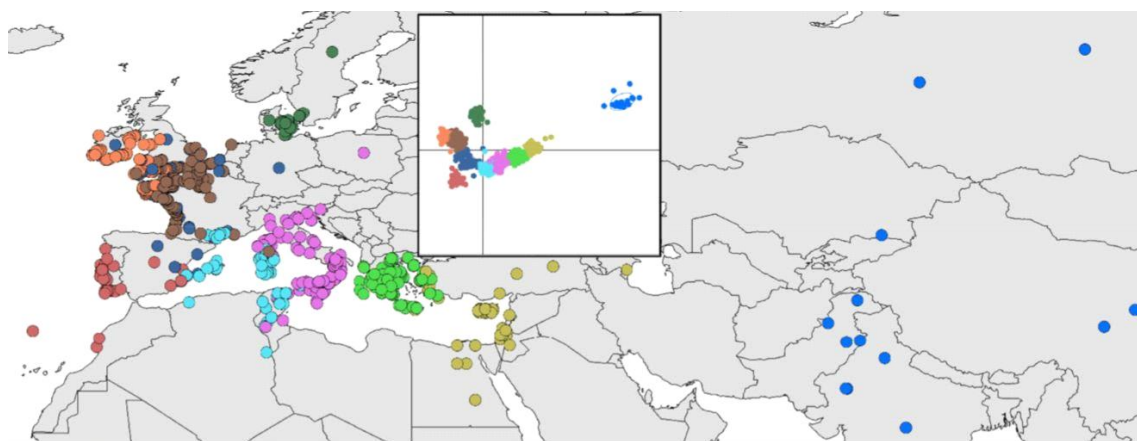


Figure 8 : Structure génétique des 923 accessions *Beta vulgaris* subsp. *maritima* estimée par DAPC (k=10 clusters).

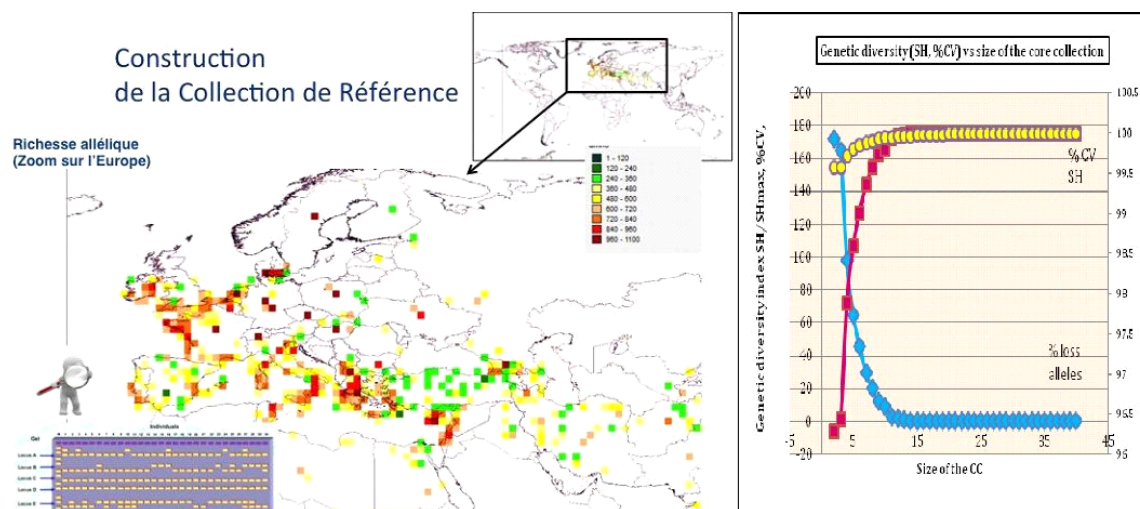
6.2. Introgression de la collection de référence dans le matériel élite et évaluation phénotypique

D’après les analyses réalisées, 40 accessions permettent de conserver toute la variabilité allélique disponible au sein du genre *Beta* section *Beta*. De même 15 accessions permettent de conserver toute la variabilité allélique complémentaire de celle disponible au sein de notre pool cultivé (Figures 9 et 10). Ainsi, l’analyse de plusieurs millions de données moléculaires nous a permis de vérifier que l’hypothèse initiale d’AKER était la bonne. Cette découverte est déjà en soit très importante puisqu’elle démontre, au moins chez la betterave, que très peu d’individus sont nécessaires pour représenter la diversité allélique potentiellement utile.

Cette collection de référence de 15 plantes a depuis été entièrement séquencée et alignée sur une séquence de référence. Nous disposons en effet déjà d’un génome de base de betterave, ceci depuis 3 ans. Il est quatre fois plus petit que celui de l’homme avec 750 Mio de bases (contre 3 milliards pour l’homme). Suite à ce séquençage des 15 plantes, nous avons pu mettre en évidence 40 millions de différences, de SNPs, réparties tout au long du génome. Cela veut dire en d’autres termes qu’en moyenne, environ une base sur 20 est variable, les autres non. Cette variabilité est inégalement

répartie sur les chromosomes. Nous avons, pour les besoins de la sélection, mis au point une puce avec 30.000 SNP répartis de façon régulière sur l'ensemble des chromosomes, ajoutant des SNPs connus pour leur proximité de gènes d'intérêts voire des gènes eux-mêmes. La difficulté a quand même été de bien recouvrir le génome avec cette puce, sans trop de distance non marquée par endroit. Nous avons aussi dû éviter la proximité trop grande de SNPs. Celle-ci peut en effet compromettre leur bonne lecture sur les systèmes de puces actuels. Indirectement cet outil va constituer également une ressource très intéressante pour la sélection génomique, voire d'autres programmes de sélection.

Les 15 plantes exotiques ont alors été croisées et recroisées avec des plantes dites élites, pour enfin assigner cette nouvelle variabilité allélique. Pour les rétrocroisements et leur suivi nous avons aussi constitué une collection de référence de marqueurs et ainsi pu suivre les recombinaisons. La dernière étape combinera le génotypage systématique déjà réalisé à un phénotypage le plus divers et précis possible.

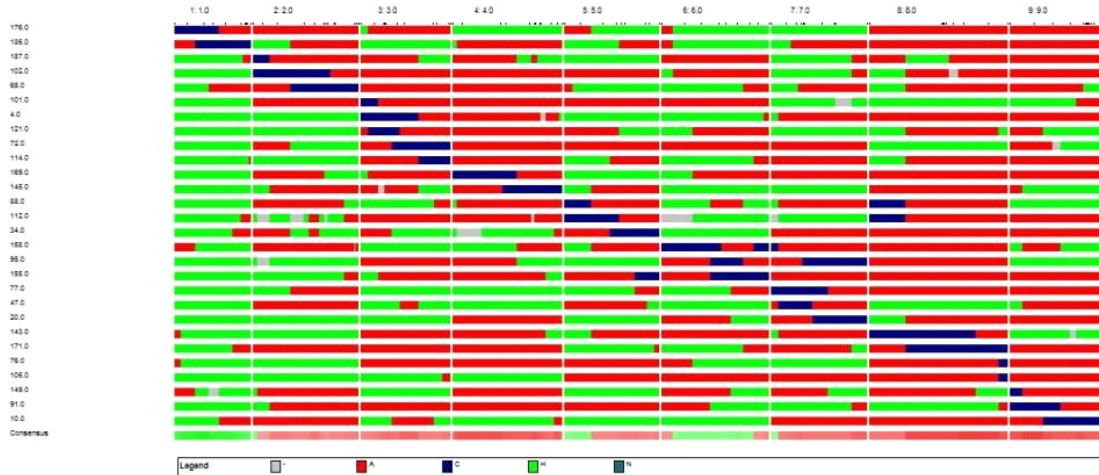


Figures 9 et 10 : Répartition de la richesse allélique en Europe (vert pauvre – jaune moyen – rouge riche) ; Taille de la collection de référence et représentativité de la richesse allélique (en bleu pourcentage de perte allélique, en rouge gain allélique).

Des simulations ont été réalisées pour déterminer comment introgresser au mieux cette collection de référence de manière à obtenir en deux générations de croisements, 200 lignées BC2S1 contenant chacune un fragment différent du génome de l'exotique de départ. Ainsi 188 plantes BC1 par accession de départ, ont été analysées avec un set spécifique de 100 marqueurs SNP, ce qui nous a permis d'établir une carte génétique précise pour chaque accession. Parmi les descendances BC2 obtenues, 6.400 plantes ont été génotypées avec un set de 100 marqueurs idéalement répartis sur l'ensemble du génome, de manière à fragmenter le génome de la plante exotique en fragments d'environ 10-15 cM et avoir un taux de retour vers l'élite d'environ 90 %. L'utilisation d'un nombre important de descendants a permis d'obtenir les recombinaisons nécessaires pour cibler le fragment souhaité dans un fond génétique élite.

Il ne faut pas oublier que la partie moléculaire d'AKER est indissociable du phénotypage. Ce phénotypage représente d'ailleurs 60 % du budget (des 20 millions d'euros). La spécificité d'AKER est un phénotypage si possible à haut débit, ceci pour évaluer un très grand nombre d'individus, précis pour détecter les parties du génome correspondantes, mais aussi intégrant une notion de dynamique et donc en privilégiant le « non destructif ». Ainsi, par exemple, pour le rendement, nous allons nous attacher non pas à une évaluation d'un rendement par le simple poids d'une récolte à un moment donné, mais essayer de trouver des méthodologies non destructives de la racine nous permettant de

tracer une courbe de croissance, une « élaboration du rendement » au cours du temps, afin de détecter des parties du génome responsables de la vitesse de mise en place du rendement. Sprinteurs ou marathoniens...et ainsi « disséquer » aux mieux les traits désirés comme le rendement ou la teneur en sucre ou une maladie par exemple.



Figures 11 et 12 : Représentation de l’introgression du génome exotique dans un contexte receveur élite. (en rouge le génome élite, en vert le génome exotique). Chaque ligne représente un individu de la descendance, les colonnes représentent les 9 chromosomes. Le BC1 et les BC2 S1 sont représentés Figure 11 et Figure 12 respectivement.

Dans le domaine des semences, nous nous sommes aussi attachés à étudier la morphologie des différents compartiments grâce à des outils extrêmement précis comme la tomographie rayons X (mesure des volumes de chaque compartiment, des surfaces de contact entre l’embryon et les réserves, épaisseur de l’enveloppe de la semence, etc.) et bien d’autres caractères encore, en utilisant des techniques ou des dispositifs encore plus innovants comme l’analyse par laser des surfaces pour déterminer la viabilité d’une graine pleine, les bancs de germination contrôlés par caméras, classiques

ou thermiques, les cinétiques de germination dans des conditions simulées et contrôlées, stressantes ou non, etc. Bien sûr, le phénotypage au champ est aussi un élément essentiel, comme la mesure de la teneur en sucre non destructive par sonde ou par patches, le suivi des parcelles, de la croissance et de la couverture foliaire par drones équipés de différents types de caméras incluant des mesures corrélées avec les teneurs en azote, à l'échelle de la feuille ou la mesure de réponse à un stress... L'utilisation de tests en serre enfin est une approche importante, permettant de contrôler l'environnement, comme par exemple pour l'inoculation de maladies racinaires comme le champignon *Rhizoctonia*, et les mesures en continu, assistée par caméras, de l'évolution des dégâts après inoculations artificielles.

7 - CONCLUSION ET PERSPECTIVES

En parallèle de la colonne vertébrale du programme AKER, nous nous sommes également attachés à la mise en place de la sélection génomique. En effet, nous allons disposer à la fin du programme AKER (début 2020), d'un matériel particulièrement intéressant entièrement séquencé et comportant une richesse allélique particulièrement originale. Lui appliquer la sélection génomique et ainsi mettre au point des algorithmes prédictifs de la valeur génétique nous apparaît comme une approche essentielle, à tester dès à présent.

Nous avons ainsi cumulé pendant trois ans les résultats phénotypiques d'un grand nombre de « candidats variétés » et nous les avons tous génotypés. Nous nous sommes servis des deux premières années pour modéliser les algorithmes de prédiction de la valeur génétique. Nous avons alors appliqué ces algorithmes sur la génétique connue des candidats testés en dernière année, ceci bien sûr sans utiliser leur valeur phénotypique aux champs. Enfin nous avons pu comparer ces valeurs prédictives aux valeurs réellement obtenues aux champs. Les résultats ont été particulièrement intéressants : ainsi, pour le rendement en sucre mesuré et le rendement en sucre prédit, on constate une corrélation de 0,78 entre les deux ce qui prouve la très bonne qualité de la prédiction. Pour le rendement, caractère complexe, la corrélation est tout de même de 0,65. S'agissant de la teneur en sucre, caractère très héritable, la corrélation est de 0,85. Les autres critères, comme la teneur en potassium et en sodium, présentent des corrélations intermédiaires. Ces résultats encourageants seront bien sûr à confirmer avec un matériel beaucoup plus variable et malgré tout plus « éloigné » du matériel élite.

Après AKER, nous aurons à réaliser des constructions de type pyramidal ou à gérer la variabilité par l'introduction de ce matériel en populations récurrentes en utilisant au maximum les marqueurs moléculaires et de façon plus précise l'ensemble du génome. Cela constituera sans doute un beau défi ! Bien sûr, il faut être également critique sur l'approche AKER. L'une des principales limites est qu'AKER considère la variabilité génétique au sens séquence génomique. Bien sûr, et nous l'avons vu dans une AG de l'ASF, l'épigénétique est à considérer. Il serait ainsi possible d'aborder la même démarche avec des marqueurs de méthylation. Au-delà de l'ADN, c'est certainement les approches ARN, protéomique et métabolomique qui nous font défaut. Néanmoins, on pourra assez vite les intégrer, dès que les techniques de haut débit dans ces domaines seront plus abordables. Déjà les recherches de voies métaboliques sont plus faciles et la combinaison des connaissances avec celles des génomes pourra sans nul doute créer une synergie permettant une meilleure distinction des apports dus à la variabilité utile et à celle qui ne l'est pas.

Les NBT (Nouvelles Techniques de Sélection), dont le statut légal est en train d'être discuté au niveau européen, pourront sans nul doute permettre de mieux exploiter les connaissances issues d'AKER. La construction d'algorithmes prédictifs et par conséquent la mise en évidence de l'impact plus ou moins forts de certains SNPs sur les phénotypes devraient fournir des cibles idéales pour les technologies comme les CRISPR capables de modifier une seule base. Inversement, la modification de certaines bases choisies pourrait préciser l'impact de certains SNPs et ainsi donner plus de poids pour une validation d'algorithmes. A terme, certaines procédures de rétrocroisements pourraient aussi être évitées : la base responsable d'un phénotype désiré et repérée dans un contexte exotique serait ainsi modifiée directement sur le génotype élite. Le fait d'être une variation allélique exotique réellement

observée (et donc « naturelle ») renforcerait aussi son « image » de modification acceptable par la société...

Nous voudrions aussi terminer sur de nouvelles perspectives : les données générées par les marqueurs moléculaires sont très nombreuses et précises et déjà nous permettent d'envisager de nouvelles études. Une « nouvelle science » appelée la *landscape genomic* ou « génomique du paysage » exploite ainsi nos données génétiques et les données passeport de nos accessions. Son objet (projet ADAPT, collaboration avec le CNRS et l'Université de Marseille) est d'étudier la corrélation entre la présence de certains allèles et l'adaptation de la plante à des conditions environnementales particulières. Il peut s'agir d'allèles liés à la longueur du jour, à l'altitude, à tel ou tel type de sol, au climat... par exemple. Elles nous permettront à terme d'approcher une variabilité « utile » liée à l'adaptation. Un certain nombre de régions génomiques pourront alors être mises en évidence, introduites dans les matériels élites destinés à certains marchés spécifiques. Notons que nous avons ici la chance de bénéficier de données passeport précises associées à de très nombreux paramètres descriptifs.

Nous avons également un autre projet dénommé BE DOMINO qui étudie la betterave et sa domestication... Son approche est simple : rechercher dans le compartiment exotique des allèles de résistance déjà utilisés par les sélectionneurs. Prenons par exemple l'allèle *Rz2*, allèle utilisé pour la résistance à la Rhizomanie, une maladie virale très grave et mondialement répandue, dont l'agent infectieux est un virus transmis par un champignon du sol *Polymyxa betae*. C'est un allèle d'un autre gène que celui utilisé historiquement : le *Rz1*. Nous l'avons retrouvé dans des accessions provenant de plusieurs zones géographiques très différentes. Une approche moléculaire plus précise nous a montré que la ressource actuellement utilisée par un de nos concurrents dans ses variétés de betterave provenait d'une *Beta maritima* poussant sur les côtes danoises. D'autres allèles différents (démontrés avec plusieurs marqueurs internes au gène) et aussi différemment « entourés » (démontrés à l'aide de marqueurs flanquants) sont donc présents dans d'autres zones géographiques. Nous projetons de vérifier qu'ils correspondent bien à des résistances valides qui, si c'est le cas, constituent une nouvelle source de variabilité intéressante, non pas uniquement par l'allèle légèrement différent (ou pas) mais aussi par l'environnement génétique proche, qu'il nous serait difficile de séparer, sauf à disposer d'effectifs colossaux et de chance pour obtenir des recombinaisons proches...

“Journée ASF du 4 février 2016”

“Apport des marqueurs moléculaires et de la génomique à l'utilisation des ressources génétiques pour l'amélioration des plantes”

REMERCIEMENTS

Les auteurs tiennent à remercier chaleureusement pour leurs contributions tous les partenaires et collaborateurs du Programme AKER.

This work received supports from the French Government supervised by the Agence Nationale de la Recherche in the framework of the program Investissements d'Avenir under reference ANR-11-BTBR-0007 (AKER program).

Nous invitons le lecteur à visiter le site AKER (<http://www.aker-betterave.fr/fr/>). Il lui donnera plus de précisions sur nos partenaires et sur le projet lui même qui dépasse l'utilisation des ressources génétiques.

REFERENCES BIBLIOGRAPHIQUES

- Beukelaer (de) H. , Smykal P., Davenport G.F., Fack V., 2012. CoreHunterII:fast core subset selection based on multiple genetic diversity measures using Mixed Replica search. *BMC Bioinformatics*, 13, 1, 312.
- Desprez, B., 1996. Towards a French network for beet genetic resources. In International *Beta* Genetic Resources Network, L.P. L. Frese, H.M. Srivastava and W. Lange, ed (Izmir, Turquie: International Crop Network Series), 82.
- Desprez, M., Desprez B., 1993. Evolution des méthodes de sélection de la betterave sucrière des origines à nos jours, C.R. Acad. Agric. Fr. 79, N°6, 71-84.
- Desprez, M., Desprez B., 2015. De la sélection phénotypique à la sélection génotypique : l'exemple de la betterave, in *Evolution des méthodes de sélection de Louis de Vilmorin à aujourd'hui*, Journée du conseil scientifique de la SNHF, Édition 2015, p30-34.
- Frese, L., Desprez B., 1999. Utilization of Beta genetic resources, In Implementation of the Global Plan of Action in Europe, *Proceedings of the European Symposium*, 30 Jun-3 Jul 1998, L.F. T. Gass, F. Begemann and E. Lipman, ed (Braunschweig, Germany: IPGRI), 172-180.
- Frese, L., desprez, B., Ziegler, D., 2001. Potential of genetic resources and breeding strategies for base broadening in *Beta*. In *Broadening the Genetic Base of Crop production*, C.S.a.T.H. eds H.D. Cooper, ed (IPGRI-FAO), 295-309.
- Jombart T., Ahmed I., 2011. Adegnet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* btr521. doi: 10.1093/bioinformatics/btr521
- Jombart T., Devillard S., Balloux F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94
- Kilian A., Wenzl P., Huttner E. *et al*, 2012. Diversity arrays technology: a generic genome profiling technology on open platforms. In: Pompanon F, Bonin A (eds) *Data production and analysis in population genomics*. Humana Press, New York, pp 67–91.
- Mangin B., Sandron F., Henry K., Devaux B., Willems G., Devaux P., Desprez B., Goudemand E., 2015. Breeding patterns and cultivated beets origins by genetic diversity and linkage disequilibrium analyses. Online. DOI: 10.1007/s00122-015-2582-1.
- Munerati, O., 1932. Sull'incrocio della barbabietola coltivata con la beta selvaggia della costa adriatica, *L'Industria Saccarifera Italiana*, **25**, 303-304.
- Perrier X., Flori A., Bonnot F., 2003. Data analysis methods. In: Hamon, P., Seguin, M., Perrier, X., Glaszmann, J. C. Ed., *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43 - 76.
- Perrier X., Jacquemoud-Collet, J.P., 2006. DARwin software <http://darwin.cirad.fr/>
- van Zonneveld M., Scheldeman X., Escribano P., Viruel M.A., van Damme P., Garcia W., Tapia C., Romero J., Siguenas J.I., 2010. Mapping genetic diversity of cherimoya (*Annona cherimola* Mill.) - Application of spatial analysis for conservation and use of plant genetic resources. *Training manual on spatial analysis of plant diversity and distribution*. Author: Scheldeman, X.; van Zonneveld, M. Corporate Author: Bioversity International, Rome (Italy), 179 pp.

GESTION ET UTILISATION DES RESSOURCES GENETIQUES CHEZ LE MAÏS

Alain CHARCOSSET¹, Brigitte GOUESNARD², Laurence MOREAU¹,
Stéphane NICOLAS¹, Anne ZANETTO³, Jacques LABORDE⁴,
Carine PALAFFRE⁴ & Cyril BAULAND¹

¹Génétique Quantitative et Évolution - Le Moulon,
INRA - Université Paris-Sud - CNRS – AgroParisTech, Ferme du Moulon, 91190 Gif-sur-Yvette,
France

²UMR Amélioration Génétique et Adaptation des Plantes méditerranéennes et tropicales (AGAP UMR
1334), INRA – Montpellier, 2 place Pierre Viala bat21, 34060 Montpellier Cedex 1, France

³DIASCOPE, INRA Domaine de Melgueil, Chemin de Mezouls, 34130 Mauguio, France

⁴INRA St Martin de Hinx, 2297 route INRA, 40390 Saint Martin de Hinx, France

E-mail : alain.charcosset@moulon.inra.fr

RESUME

De nombreuses ressources génétiques sont disponibles chez le maïs, notamment des variétés populations traditionnelles et des lignées anciennes. Les premières présentent une très forte diversité intra-population, qui s’explique par le mode de fécondation allogame de l’espèce et la contre-sélection des plantes consanguines qui présentent une vigueur inférieure. Du fait des difficultés d’obtention des premières lignées utilisées pour créer des hybrides, une partie seulement de cette diversité a contribué aux programmes de sélection modernes. Le marquage moléculaire permet de comprendre l’organisation de la diversité au sein des ressources génétiques, d’identifier des sources de diversité originales et facilite le choix des croisements destinés à évaluer le potentiel de celles-ci. Compte-tenu d’un progrès génétique très important réalisé depuis le début de la sélection moderne, une ou plusieurs générations de back-cross avec le matériel élite apparaissent nécessaires pour valoriser les ressources génétiques et permettre d’obtenir des transgressions prometteuses. L’efficacité de ces approches est appelée à être améliorée par des approches spécifiques de sélection génomique.

Mots-clés : diversité intra-population, ressources génétiques, maïs, dépression de consanguinité, hétérosis, groupes hétérotiques, back-cross, génétique d’association.

1 - INTRODUCTION

Le maïs a été domestiqué au Mexique puis s'est répandu sur le continent américain avec succès, en s'adaptant à des climats très divers, notamment le climat tempéré (pour une synthèse cf Tenaillon et Charcosset, 2011). A l'époque de la découverte du Nouveau Monde, à la fin du XV^e siècle et dans les décennies qui ont suivi, il a été adopté avec enthousiasme dans de nombreuses régions du monde (Mir *et al.*, 2013). Jusqu'au début du XX^e siècle, la sélection du maïs a été réalisée par sélection massale, c'est-à-dire en sélectionnant les plantes les plus « belles » ou les mieux adaptées pour produire les semences de la génération suivante. La sélection « moderne » du maïs a démarré au début du XX^e siècle avec une innovation de rupture due notamment à Shull (Shull, 1908), le concept de variété hybride (Charcosset, 2002). En permettant de multiplier à l'identique des individus hétérozygotes, celle-ci a permis une sélection efficace malgré la dépression de consanguinité très forte liée au mode de reproduction allogame de l'espèce. Après leur développement aux USA dans le deuxième quart du XX^e siècle, les hybrides de maïs se sont développés dans de nombreuses régions du monde, et notamment en Europe après la seconde guerre mondiale.

L'allogamie et la dépression de consanguinité très fortes chez le maïs ont des conséquences très importantes sur l'organisation de la diversité et la nature des ressources génétiques de l'espèce ; (i) du

fait d'une sélection pour la vigueur et la valeur reproductive qui tend à favoriser le maintien de l'hétérozygotie, les variétés traditionnelles sont des populations génétiquement hétérogènes, et (ii) la création de lignées homozygotes à partir des variétés populations est en général difficile, ce qui a amené à un nombre de fondateurs assez faible pour les programmes de sélection modernes. Les variétés traditionnelles contiennent donc *a priori* une diversité très importante qui n'a pas encore contribué aux programmes de sélection modernes.

La dépression de consanguinité a régressé au fur et à mesure des générations de sélection. Des travaux sur les variétés de la société Pioneer ont montré que le progrès de productivité des lignées a été approximativement parallèle à celui des hybrides (Duvick, 2001). En d'autres termes, l'hétérosis, c'est-à-dire l'écart de productivité entre l'hybride et ses lignées parentales est resté constant mais a décru en valeur relative. Bien que la cible principale était le rendement hybride, la sélection directe sur la capacité des lignées à produire des semences et/ou une réponse indirecte liée à la sélection pour la performance hybride ont entraîné une amélioration de la performance des lignées, d'un facteur deux environ pour le matériel de Pioneer entre 1930 et 1990. Il est donc beaucoup plus facile aujourd'hui de créer une nouvelle lignée en croisant deux lignées modernes que de revenir aux variétés traditionnelles. De plus, les nouvelles lignées issues de croisements entre lignées élites conduisent en moyenne à de meilleurs hybrides que celles qui sont issues de variétés populations.

Alors, quelles ressources génétiques et comment les utiliser ? Il est important de noter que la notion de ressources génétiques est relative à un programme de sélection donné, car un matériel directement cultivable dans certaines conditions environnementales et d'utilisation peut être une ressource génétique pour d'autres. Au sens large, tout ensemble d'individus susceptible d'apporter une diversité nouvelle et favorable par rapport à celle du programme de sélection en cours peut être considéré comme une ressource génétique. Chez le maïs, cette notion de programme de sélection intègre les zones de culture visées et l'organisation de la diversité en groupes d'aptitudes à combinaison (Charcosset and Gallais, 2009).

La notion de ressources génétiques implique aussi souvent qu'une procédure de sélection spécifique sera nécessaire pour les exploiter. Nous nous concentrerons ici sur la caractérisation et la valorisation de ressources génétiques constituées par des variétés populations anciennes et des lignées de première génération ou anciennes. Du fait des progrès déjà réalisés au sein des programmes de sélection depuis la mise en œuvre de la sélection hybride, ces ressources génétiques ne peuvent pas conduire par fixation directe à des parents de variétés hybrides performantes, ce qui suppose la mise en œuvre de stratégies de sélection spécifiques. Nous envisageons dans ce qui suit, trois sujets complémentaires : (i) l'apport des marqueurs pour la description de la diversité et l'identification de ressources originales, (ii) des éléments sur la variation phénotypique au sein de panels de ressources génétiques sélectionnées pour leur représentativité, et enfin (iii) des résultats d'expériences visant à utiliser ces ressources dans les programmes de sélection élite.

2 - APPORT DES MARQUEURS POUR LA DESCRIPTION DE LA DIVERSITE ET L'IDENTIFICATION DE RESSOURCES ORIGINALES

Du fait de l'allogamie et de la dépression de consanguinité évoquées ci-dessus, les populations traditionnelles de maïs conservées dans les collections présentent chacune une diversité importante. La prise de conscience de cette diversité et de son importance par les chercheurs INRA et les sélectionneurs de Promaïs a été à l'origine de l'action menée à partir de 1983 sur la collection de populations de maïs conservée depuis lors à l'INRA de Montpellier (Gallais et Monod, 1998). Des travaux (Dubreuil and Charcosset, 1998) conduits avec des marqueurs RFLP sur 10 variétés traditionnelles illustrent la coexistence d'une différenciation des populations et d'une forte variabilité intra-population. Au sein de cette collection, cette diversité « intra-population » peut représenter de l'ordre de 75 % de la diversité totale si on l'évalue avec l'indicateur classique de (Nei, 1973). Cet ordre de grandeur a été confirmé avec de nombreuses techniques de marquage plus récentes (isozymes, microsatellites et SNP notamment). La présence ou l'absence d'allèles donne avec les marqueurs RFLP une image beaucoup plus contrastée des populations avec en moyenne pour les

différents locus étudiés 2,7 allèles par population sur un total de 13. Il apparaît donc important de comprendre la différenciation des populations, tout en gardant en tête que chacune présente une certaine diversité intrinsèque.

Ces caractéristiques nous ont amené à étudier les collections de populations avec des approches de génotypage de mélanges d'ADN visant à identifier les allèles présents et à estimer leurs fréquences au travers d'un seul échantillon d'ADN issus d'un échantillon de plantes représentatif de l'ensemble de la population. Cela nous a permis d'obtenir des résultats robustes sur l'organisation des collections de populations françaises et européennes (Gauthier *et al.*, 2002) et sur l'origine des populations européennes (Rebourg *et al.*, 2001 ; Rebourg *et al.*, 2003). Plus récemment, un travail conduit en collaboration avec le CIMMYT et plusieurs instituts nationaux de pays asiatiques et africains a permis de caractériser 800 populations représentatives de l'ensemble de la diversité mondiale avec des marqueurs microsatellites (Mir *et al.*, 2013). Ce travail a permis de révéler des profils de similarité entre origines géographiques qui, confrontés à des éléments historiques, suggèrent un premier portrait de la diffusion mondiale de l'espèce. Des éléments nouveaux ont pu être obtenus sur l'introduction en Europe, l'existence de deux voies principales de diffusion en Asie et des échanges multiples avec le continent Africain, liés aux principales voies de commerce des esclaves. Au-delà de la meilleure connaissance de la diversité et des processus qui ont pu la structurer, ces outils peuvent aussi être très intéressants pour échantillonner dans les collections des sous collections représentatives en vue d'une analyse plus poussée, notamment au niveau phénotypique (cf. section 2).

Dans le cadre d'une valorisation des ressources génétiques, il est aussi particulièrement important de comprendre les relations entre la diversité des variétés traditionnelles et celle du matériel élite contemporain. Les travaux de (Camus-Kulandaivelu *et al.*, 2006) comparent dans ce but trois niveaux de diversité : (i) les populations traditionnelles conservées dans les collections, (ii) les principales lignées « de premier cycle » créées à partir de ces populations, notamment dans les premières phases des programmes de sélection hybride aux USA et en Europe et (iii) des lignées issues de programmes de sélection plus avancés ayant « recyclé » les premières lignées créées. Ils illustrent une filiation entre les groupes observés à ces trois niveaux mais aussi l'apparition d'un nouveau groupe très important, les « Stiff Stalk » façonné par la sélection hybride nord-américaine. Ce processus de création de groupes génétiques nouveaux liés à la sélection moderne conduite par différents acteurs publics et privés dans le monde s'est poursuivi depuis. Des résultats de Rincent *et al.* (2014) illustrent notamment la place très importante prise par le groupe denté Iodent dans la sélection européenne. La diversité évolue donc en permanence avec la sélection de variétés hybrides, qui façonne de nouveaux groupes génétiques.

La valorisation des ressources génétiques telle que nous l'avons définie précédemment doit donc commencer par une étude comparative des collections de ressources génétiques candidates et du matériel de sélection plus avancé pour identifier des sources de diversité originales pour enrichir celui-ci. Dans le cadre du projet Amaizing nous avons donc, récemment, utilisé une puce de génotypage Illumina 50k qui comporte 50.000 SNP (Ganal *et al.*, 2011) pour la caractérisation conjointe de populations traditionnelles et de lignées (Arca *et al.*, *in prep.*). Les résultats illustrent tout d'abord une bonne capacité de la méthode à déterminer à l'aveugle la population d'origine d'une lignée de premier cycle. Si l'on calcule la distance génétique entre une population donnée et chacune des lignées d'une collection de référence, on observe qu'au sein de populations présentant une même diversité intra-population, certaines possèdent des lignées proches et d'autres non. Ces dernières peuvent donc être considérées comme des cibles prioritaires pour enrichir la diversité des programmes de sélection. L'échantillonnage de sources de diversité potentiellement intéressantes doit donc intégrer à la fois l'originalité des variétés populations les unes par rapport aux autres et par rapport aux programmes de sélection en cours.

3 - VARIATION PHENOTYPIQUE

Pour caractériser la variation phénotypique des ressources génétiques, il est possible d'avoir une stratégie en cascade pour réserver les évaluations les plus coûteuses à une collection plus réduite de

populations qui représente au mieux la diversité disponible. Dans le cadre du projet européen RESGEN88, qui a porté sur l’analyse des collections françaises, espagnoles, portugaises, italiennes, allemandes et grecques, de l’ordre de 3000 populations européennes ont été décrites par des données de passeport (origines géographiques notamment) qui ont permis de sélectionner un sous-ensemble de 400 populations. Celles-ci ont été génotypées avec des marqueurs RFLP (Gauthier *et al.*, 2002). Les données de génotypage ont été analysées avec le logiciel MStrat (Gouesnard *et al.*, 2001) pour sélectionner une collection « noyau » finale de 100 populations représentant 93 % des allèles observés. Cette collection représentative a été évaluée pour des caractères d’intérêt : tolérance à des ravageurs et à la sécheresse, digestibilité de la plante entière et qualité du grain. Des populations présentant une tolérance à la sécheresse intéressante ont pu être identifiées et leur caractérisation a été poursuivie dans le cadre d’expériences complémentaires (Gouesnard *et al.*, 2016). De la même façon, une collection représentative de 88 populations du sud-ouest de la France a pu être définie et caractérisée pour des caractères de valeur alimentaire. Plus de la moitié des populations apparaissent supérieures au témoin commercial Anjou 285 pour le critère « Dinag » et les plus digestibles ont été utilisées en croisement avec du matériel élite (cf. partie 3). Les données sont accessibles dans la base données Multicrop (<http://bioweb.supagro.inra.fr/multicrop/>).

Cette caractérisation directe des populations est pertinente pour un certain nombre de caractères à déterminisme essentiellement additif, comme c’est le cas pour la digestibilité. Pour des caractères présentant de la dominance, une évaluation directe peut conduire à sous évaluer l’intérêt de populations présentant une consanguinité plus forte que les autres. Rebourg *et al.* (2001) ont montré que, dans une collection de populations européennes, les populations présentant la diversité la plus faible pour des marqueurs RFLP avaient des performances particulièrement faibles pour le rendement et ses composantes. Pour de telles populations, un rendement faible en valeur propre n’indique pas nécessairement un manque de potentiel pour créer des parents de variétés hybrides. Il est donc important de compléter ces évaluations en valeur propre par des évaluations du potentiel en croisement, en phénotypant une descendance hybride. L’approche la plus simple est de croiser les populations avec des testeurs représentatifs des groupes hétérotiques travaillés au sein du programme de sélection. Les marqueurs moléculaires peuvent aider dans une certaine mesure dans le choix du ou des testeurs les plus appropriés pour une population donnée. (Rebourg 1996) a évalué 65 populations pour leurs valeurs en croisement avec trois testeurs représentatifs des groupes corné précoce, denté précoce et denté demi-tardif (respectivement F2xF283, F252xF272, MBS847) et décomposé la variation en aptitudes générale et spécifique à la combinaison. Ces mêmes populations ont été caractérisées avec des marqueurs RFLP qui ont permis d’obtenir une classification en cinq groupes principaux. Cette structuration en groupe explique 42,7 % de la variation observée pour l’aptitude spécifique à la combinaison (ASC). En particulier, les populations classées dans un groupe proche du groupe auquel appartient un testeur donné ont en général une ASC négative avec celui-ci. Une classification moléculaire peut donc permettre de cibler le choix de testeurs *a priori* favorables pour révéler le potentiel d’une population.

Les résultats ci-dessus concernent la valeur phénotypique moyenne des populations. Il est important aussi de considérer que, comme au niveau moléculaire, les populations peuvent présenter une variation phénotypique importante au niveau intra-population. Dubreuil (1996) a étudié la variation inter et intra-population (30 individus par population échantillonnée) de cinq populations cornées et trois populations dentées pour le rendement de leur descendance en croisement avec un testeur denté (MBS847), en comparaison avec des témoins représentant différentes époques. La figure 1 illustre tout d’abord la moindre performance des populations dentées (US) par rapport aux populations européennes (EU) en croisement avec ce testeur. Pour ces dernières, on observe une forte variation intra-population par rapport aux différences de moyennes entre populations. Cette variation a de plus une ampleur variable selon les populations. Les toutes meilleures familles peuvent être observées au sein de populations présentant une moyenne assez forte (ex. « EU-RC ») mais aussi une moyenne plus faible si celle-ci est compensée par une variation importante (« EU-Lac »). Il est d’ailleurs intéressant de noter que, si l’on considère un index qui combine rendement et précocité, les meilleures familles sont observées pour la population Lacaune, à l’origine des lignées F2 et F7 qui ont été à l’origine du succès des hybrides de maïs précoces en France. De plus, ces résultats montrent que la lignée F2, qui a eu un impact exceptionnel, était au niveau des toutes meilleures familles étudiées

ici. Le niveau de performance de cette lignée est donc remarquable compte-tenu des moyens expérimentaux limités employés à l'époque de sa création. *A contrario*, la lignée F7 apparaît plus proche de la moyenne de la population. Ces résultats illustrent aussi que les meilleurs hybrides obtenus entre familles issues des populations et le testeur MBS847 sont en deçà de la performance du témoin élite de l'époque (Fanion). Même si cette expérience est de taille limitée, elle illustre que le progrès génétique creuse régulièrement l'écart entre le matériel élite et les meilleures lignées qui peuvent être créées par extraction directe des populations. D'autres stratégies de sélection apparaissent donc nécessaires pour valoriser les ressources génétiques.

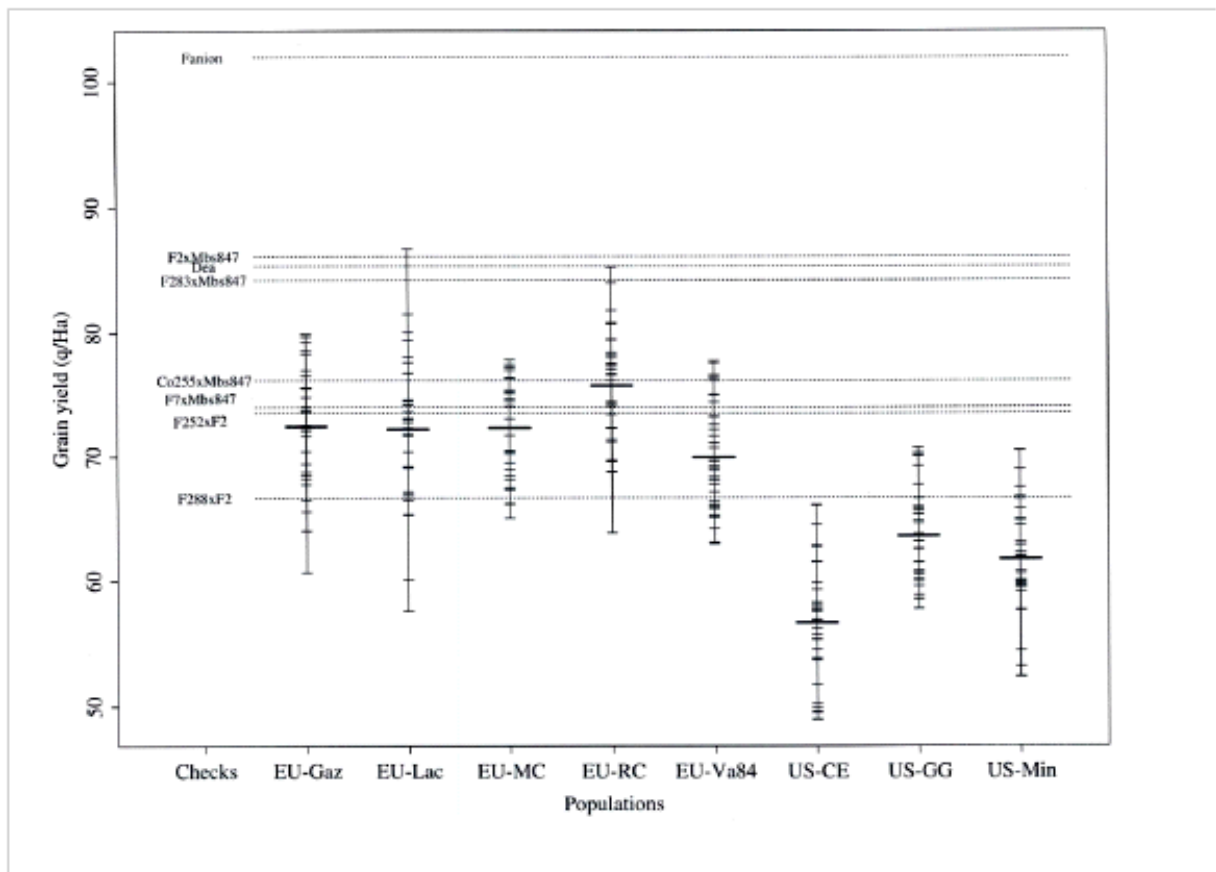


Figure 1 : Evaluation de 240 familles S1 issues de 8 populations traditionnelles pour leur performance en croisement avec le testeur denté MBS847 (rendement grain évalué dans 4 essais conduits en 1994 et 1995). Les traits horizontaux indiquent la performance de témoins de référence. Le préfixe EU désigne les populations d'origine européenne, le préfixe US les populations d'origine américaine.

4 - QUELLES STRATEGIES D'UTILISATION DANS LES PROGRAMMES DE SELECTION ELITE ?

Différentes stratégies d'utilisation des ressources génétiques ont été tentées chez le maïs, avec pour but d'enrichir la diversité génétique des populations de sélection. Dans le cadre du programme INRA-Promaïs « populations sources », démarré en 1983, André Gallais et collaborateurs ont évalué la performance en croisement avec des testeurs de plus de mille populations. Les résultats ont été distribués aux sélectionneurs participant au programme en vue d'une utilisation en interne des populations les plus intéressantes. Les populations présentant les meilleurs résultats avec chaque testeur ont de plus été inter-croisées pour produire des pools. Chacun a été décliné en une version « de base » et une version « améliorée » par les meilleures lignées INRA de l'époque présentant un profil en croisement comparable. Ces différentes versions des pools ont été distribués aux partenaires du programme pour une utilisation interne. Cette stratégie de « pré-breeding » était justifiée dans le contexte de l'époque où les programmes privés restaient relativement proches génétiquement du

matériel public à partir duquel ils avaient été initiés. Elle présente toutefois l'inconvénient de « marquer » le matériel créé par les lignées élites utilisées, ce qui peut poser des problèmes de compatibilité avec l'organisation en groupes hétérotiques d'un sélectionneur particulier.

Une alternative est donc de croiser les ressources génétiques identifiées comme potentiellement intéressantes avec les meilleures lignées élites du moment, en choisissant des couples donneurs-receveurs spécifiques. Le matériel créé peut avoir une double finalité : la création directe de matériel élite, ou la création d'un matériel intermédiaire pas encore utilisable directement en création variétale mais qui va être utilisé comme parent de futures populations. Cela pose plusieurs questions de stratégie comme le nombre de croisements à réaliser avec la lignée élite. Un programme pilote a été mis en place pour éclairer ces questions (Gouesnard *et al.*, 2005). Cinq populations cornées ont été retenues pour améliorer la lignée F2. Nous avons considéré comme source de diversité la population elle-même ou la meilleure famille S1 sélectionnée parmi 30 sur la base de la valeur en croisement avec le testeur denté MBS847 (cf. expérience mentionnée précédemment). Chaque donneur a été croisé au receveur une ou deux fois, conduisant à un total de quatre modalités pour chaque population. Pour chaque modalité de chaque population, 30 familles ont été évaluées pour leur performance en croisement avec le testeur MBS847. Les résultats illustrent une forte variation des performances entre les différentes populations et modalités. On observe une grande variation intra-modalité et l'existence de familles supérieures à la lignée élite receveuse. Pour approfondir l'évaluation de l'effet du nombre de croisement à la lignée élite, nous avons calculé le critère d'utilité (Schnell, 1983) de chaque modalité, à savoir la valeur moyenne attendue des familles après sélection. Ce critère présente l'intérêt de prendre en compte simultanément la moyenne et la variance des populations de sélection. Ce point est particulièrement important dans le contexte de la valorisation des ressources génétiques où, aux extrêmes, (i) une part trop importante de génome issu de la ressource génétique peut dégrader tellement la performance moyenne qu'il ne sera plus possible de remonter au niveau visé malgré une variation très forte, et *a contrario* (ii) une part trop importante de génome élite limiterait tellement la variation qu'il ne serait plus possible de sélectionner des familles intéressantes malgré une moyenne très élevée. Cette analyse, réalisée avec une intensité de sélection théorique de 10 % illustre clairement que, pour cette étude pilote, la stratégie de rétrocroisement est supérieure au croisement pour toutes les populations et toutes les modalités. Un dosage à 25 % de la part de ressources génétiques apparaît plus favorable qu'un dosage à 50 %.

Ce projet, conduit à une échelle limitée, a servi de base à la mise en place d'un programme pilote plus ambitieux de valorisation des ressources génétiques dans le cadre d'un partenariat privé-public associant 8 partenaires privés membres de Promaïs et l'INRA, le projet « piémont pyrénéen ». Son organisation présente des points communs avec le programme présenté par Bruno Desprez dans le cadre de cette journée, avec une nuance importante qui est que chaque partenaire privé apportait sa propre lignée élite. 22 familles S1 issues de 7 populations présentant une très bonne valeur alimentaire ensilage dans une étude préliminaire ont été considérées comme donneurs. Elles ont été sélectionnées parmi 800 familles S1 par observation phénotypique en pépinière dans deux lieux. L'objectif de cette sélection en valeur propre était d'éliminer les familles dont une augmentation de l'homozygotie laissait s'exprimer des caractères défavorables qui compromettraient leur utilisation dans du matériel élite (verse, sensibilité aux maladies, faibles aptitudes à produire des graines ou du pollen). Les semences de chacune des 22 familles ont été partagées entre les participants, ainsi que quatre témoins donneurs, dont la lignée F4 présentant une très forte digestibilité mais aussi des problèmes agronomiques importants. Ces donneurs ont été croisés à la lignée élite de chaque société pour créer trois familles BC1-S1 par couple donneur receveur. Ceux-ci ont été évalués avec un testeur commun, F1890, sur trois lieux et deux années d'expérimentation.

Si l'on regarde tout d'abord le critère de digestibilité des parois (Figure 2a), on voit que les lignées élites les moins digestibles (à gauche du graphe) peuvent être améliorées de façon notable par les donneurs utilisés, ce qui confirme leur intérêt pour ce caractère. En vue d'une utilisation directe du matériel créé en création variétale, il est toutefois important d'évaluer l'impact des introgressions sur d'autres caractères et notamment le rendement.

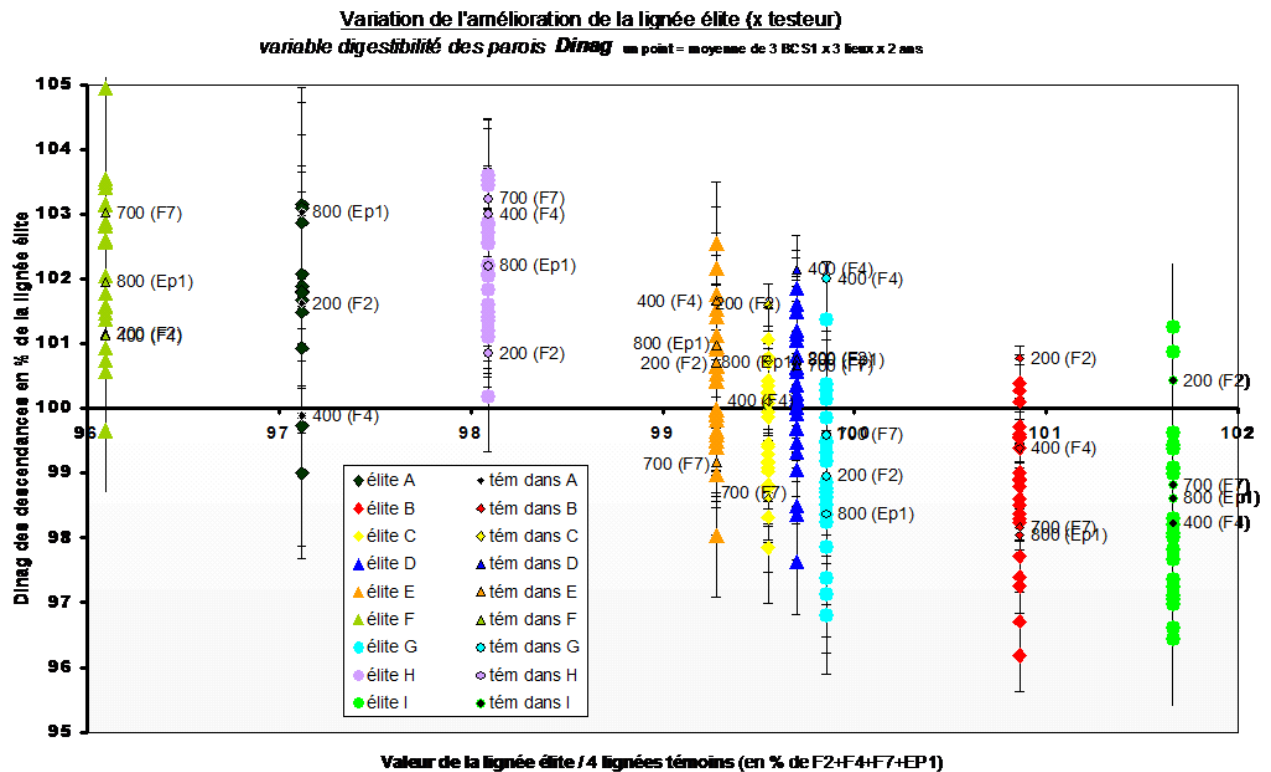


Figure 2a : Résultats d’essais d’hybrides BC1S1*testeur (3 lieux * 2 années : 2007 et 2008) : 9 lignées élites ont été introgressées en BC1S1 par 22 S1 de populations du Piedmont Pyrénéen et 4 lignées de référence (F2, F4, F7 et Ep1). Trois familles ont été développées pour chaque couple donneur – receveur.

Cas du caractère de qualité des parois végétales (*Dinag*, Barrière *et al.*, 1995) pour lequel les populations ont été choisies. Une couleur indique les descendance obtenues avec une élite donnée. Sur l’axe des abscisses, les descendance sont classées par lignée élite receveuse, suivant la valeur relative de la lignée élite par rapport à la moyenne des 4 lignées de référence. La moins bonne lignée élite en qualité *Dinag* (en vert foncé, à gauche) est à 96 % des lignées de référence. La meilleure lignée élite (en vert fluo à droite) est à 102 % des lignées de référence. Sur l’axe vertical, chaque point représente la moyenne des 3 BC1S1 issues de la même S1 (ou de la lignée de référence, servant comme donneur) exprimée en relatif par rapport à la lignée élite receveuse. La lignée élite est donc à 100 % de l’axe vertical. La moins bonne lignée élite (vert foncé à gauche) est améliorée par toutes les S1 sauf une (en-dessous de 100 %). La meilleure lignée élite (vert fluo à droite) n’est améliorée que par 3 S1. L’introgression peut faire baisser le BC1 jusqu’à 96 % de la valeur du receveur mais il y a toujours une S1 pour améliorer la lignée élite entre 101 et 105 %.

La figure 2b présente les résultats obtenus pour l’index Rendement Fourrage Pondéré Utile, qui détermine une valeur économique globale intégrant valeur alimentaire, productivité et précocité. Cet index est utilisé par le CTPS pour les épreuves officielles d’inscription au catalogue. Comme attendu, les résultats sont globalement décalés vers le bas. Les nouvelles familles sont en moyenne inférieures à la lignée qualité de référence mais, à l’exception de la meilleure lignée élite, il existe toujours des familles transgressives présentant une amélioration globale par rapport à l’élite de référence. L’approche apparaît donc encourageante. Pour faciliter sa mise en œuvre et les partages d’information entre programmes de sélection, il serait souhaitable de la mettre en place avec des donneurs homozygotes pouvant être multipliés à l’identique.

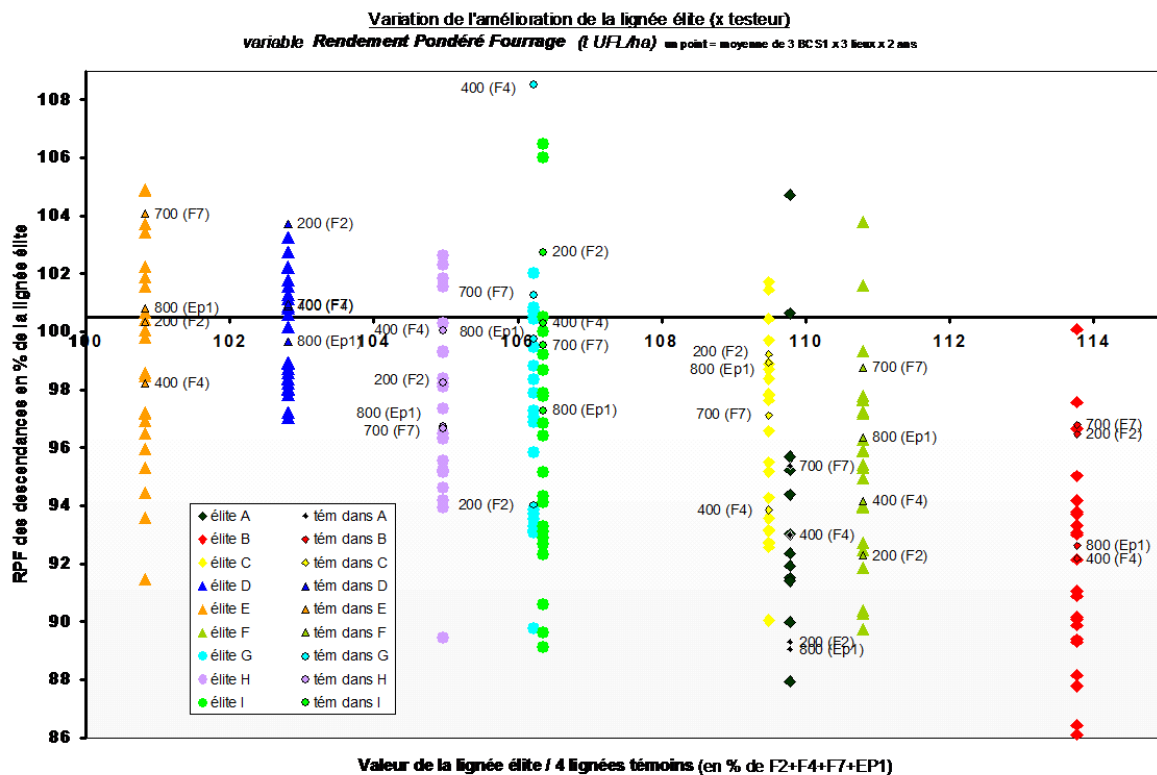


Figure 2b : Résultats d’essais d’hybrides BC1S1*testeur (3 lieux * 2 années : 2007 et 2008) : 9 lignées élites ont été introgressées en BC1S1 par 22 S1 de populations du Piedmont Pyrénéen et 4 lignées de référence (F2, F4, F7 et Ep1). Trois familles ont été développées pour chaque couple donneur – receveur.

Cas de l’index synthétique des épreuves d’inscription maïs fourrage: rendement pondéré UFL (tonnage hectare, précocité et valeur énergétique fourrage). Le positionnement des receveurs élites illustre le progrès réalisé par les sociétés semencières : toutes les lignées élites sont supérieures aux lignées de référence (toutes supérieures à 100 % sur l’axe horizontal) malgré l’utilisation d’un testeur commun. La moins bonne lignée élite sur l’index (en orange, à gauche) est à 101 % des lignées de référence. La meilleure lignée élite (en rouge à droite) est à 114 %. L’introgression peut faire baisser le BC1 jusqu’à 86 % de la valeur du receveur. Sauf pour la meilleure lignée élite (en rouge à droite), il y a toujours une S1 pour améliorer la lignée élite (entre 101 et 106 %) sur l’index utilisé pour l’inscription au catalogue.

Des efforts ont donc été entrepris pour mieux caractériser les collections de lignées existantes et créer de nouvelles lignées à partir de populations. Celles-ci peuvent de plus constituer une voie originale de maintien à long terme de la diversité des populations, dans la mesure où il serait possible de « reconstituer » une population perdue par la création d’une synthétique entre les lignées qui en ont été dérivées. A titre d’illustration Nicolas *et al.* (pers. comm.) ont analysé par marquage SSR 102 lignées créées par HD et SSD à partir de 55 populations dans le cadre du programme ProMaïs Diversité Zea. Celles-ci apportent un total de 26 nouveaux allèles (+0.48 allèle / marqueurs) qui n’avaient pas été observés dans la collection de lignées analysée par (Camus-Kulandaivelu *et al.*, 2006), ce qui illustre l’aptitude des populations à apporter une diversité nouvelle par rapport aux principaux fondateurs des programmes de sélection actuels.

Une voie complémentaire de valorisation des ressources génétiques par rapport aux approches présentées précédemment est d’identifier des allèles favorables au sein des collections pour ensuite gérer leur introduction dans les pools élites. Une première approche est de développer des panels de génétique d’association incluant des lignées originales. Nous avons analysé dans ce but dans le cadre du projet Amaizing l’ensemble des lignées présentant un type de grain corné dans la collection de l’INRA, les 60 lignées cornées obtenues par HD et SSD du programme ProMaïs Diversité Zea et les avons comparées aux lignées du panel corné analysé par Rincant *et al.* (2014), soit un total de 1191 lignées. Pour limiter les coûts de génotypage, ces lignées ont été analysées avec une approche de génotypage par séquençage, une technique très efficace mise en place par l’université de Cornell (Elshire *et al.*, 2011) permettant d’obtenir de l’ordre de 500.000 marqueurs SNP polymorphes à un coût modéré. Les résultats de génétique d’association permettent de retrouver sans ambiguïté des gènes ayant des effets très forts sur la coloration du grain, ou des effets plus modérés sur la précocité de floraison. Une des limites intrinsèques de cette approche est toutefois de manquer de puissance pour détecter l’effet de locus dont les fréquences alléliques sont très déséquilibrées, ou très liées à la structure de population. Ces limites seront toutefois levées en partie par une augmentation de la taille des populations mais, lorsque les allèles d’intérêt sont *a priori* très rares et/ou portés par des donneurs posant de gros problèmes d’adaptation aux conditions d’expérimentation, il peut être pertinent de développer des populations de back-cross à partir de donneurs choisis judicieusement. Plusieurs populations ont été développées à partir de donneurs tropicaux tolérants aux basses températures (lignées issues de populations cultivées à haute altitude) ou au stress hydrique et leur étude a commencé dans le cadre du projet Amaizing.

5 - CONCLUSION-PERSPECTIVES

En conclusion, le maïs est caractérisé par une diversité génétique et une variation phénotypique très fortes. La dépression de consanguinité a limité au départ la part de la diversité qui a été utilisée comme base des programmes de sélection moderne. Sans même parler du pool sauvage des téosintes complètement interfertiles avec le maïs, les variétés tropicales et les variétés populations anciennes cultivées dans le climat tempéré sont susceptibles d’enrichir la diversité des programmes de sélection élites actuels. Différentes voies de valorisation sont possibles pour cette diversité et doivent prendre en compte que la différence de performance se creuse avec le matériel élite, qui de plus est fortement structuré en groupes hétérotiques spécifiques à chaque entreprise de sélection. La mise en place de stratégies de back-cross avec le matériel élite apparaît aujourd’hui indispensable pour gérer ce décalage. Des projets pilotes illustrent l’intérêt de départ de type BC1 pour des caractères ciblés tels que la valeur alimentaire ensilage. Des backcross plus avancés peuvent être une bonne stratégie pour des allèles à effets forts apportés par du matériel inadapté. Il reste toutefois des challenges importants pour mieux valoriser les ressources génétiques pour des caractères complexes contrôlés par de nombreux locus. Trois questions liées restent ouvertes et particulièrement importantes à notre sens : (i) comment gérer l’introduction de diversité dans les pools de sélection, (ii) est-ce que la génétique d’association peut être pertinente pour identifier des allèles originaux, avec quels types de populations et (iii) est-ce que la sélection génomique peut aider à atteindre ces objectifs ? Différentes approches de la sélection génomique peuvent être envisagées dans ce contexte, telle celle proposée récemment par le CIMMYT dans le cadre du projet Seeds of Discovery avec une notion de « Bridge germplasm » (Gorjanc *et al.*, 2016). Au niveau français un nouveau projet a été mis en place dans le cadre de ProMaïs, qui reprend les principes du projet « piémont pyrénéen » (cf. Figure 2) avec des lignées élites privées et des donneurs issus de populations, mais avec un marquage systématique du matériel créé pour conduire une sélection génomique et/ou détecter des allèles à effets forts.

“Journée ASF du 4 février 2016”
“ Apport des marqueurs moléculaires et de la génomique à l’utilisation
des ressources génétiques pour l’amélioration des plantes ”

REMERCIEMENTS

Merci à André Gallais pour cette invitation à contribuer à cette réflexion sur la gestion et l'utilisation des ressources génétiques, et pour avoir initié en 1983 ces projets de maintien, caractérisation et valorisation des ressources génétiques, en collaboration avec Jean-Pierre Monod, qui a joué un rôle structurant au niveau du partenariat privé sur ces sujets.

Merci à Jean Beigbeder pour de nombreuses discussions très intéressantes sur ces sujets et son rôle actif de promotion de la diversité génétique du maïs.

Nous remercions les gestionnaires des collections à l'INRA à Montpellier et Saint-Martin de Hinx, les collègues impliqués dans le développement du matériel expérimental, son phénotypage et son génotypage, les jeunes scientifiques qui ont contribué à ces études : Pierre Dubreuil, Cécile Rebourg, Bastien Lemoine, Perrine Gauthier, Letizia Camus-Kulandaivelu, Renaud Rincent, Mariangela Arca, et enfin nos partenaires privés dans les projets conduits dans le cadre de Promaïs (<http://promais.org/>) ou Amaizing (<http://www.amaizing.fr/>) avec la contribution de l'ANR.

REFERENCES BIBLIOGRAPHIQUES

- Barrière, Y., J. C. Emile, O. Argillier and Y. Hebert (1995). "Effets du génotype de maïs ensilage sur les performances zootechniques de vaches laitières. ." *Inra Production Animale* 8(5): 315-320.
- Camus-Kulandaivelu, L., J. B. Veyrieras, D. Madur, V. Combes, M. Fourmann, S. Barraud, P. Dubreuil, B. Gouesnard, D. Manicacci and A. Charcosset (2006). "Maize adaptation to temperate climate: Relationship between population structure and polymorphism in the Dwarf8 gene." *Genetics* 172(4): 2449-2463.
- Charcosset, A. (2002). Le fait hybride, conditions de l'innovation et choix stratégiques. *L'amélioration des plantes, continuités et ruptures*. Montpellier.
- Charcosset, A. and A. Gallais (2009). "Emergence et développement du concept de variétés hybrides chez le maïs." *Le Sélectionneur Français* 60: 21-30.
- Dubreuil, P. (1996). *Etude de l'apport de marqueurs RFLP pour l'analyse de la diversité génétique et sa structuration chez le maïs (Zea mays L.). Relations avec les caractéristiques agronomiques de populations traditionnelles*. Thèse de doctorat, Université Paris XI Orsay.
- Dubreuil, P. and A. Charcosset (1998). "Genetic diversity within and among maize populations : a comparison between isozyme and nuclear RFLP loci." *Theoretical and Applied Genetics* 96(5): 577-587.
- Duvick, N. D. (2001). "Biotechnology in the 1930s: the development of hybrid maize." *Nature Reviews Genetics* 2: 69-74.
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler and S. E. Mitchell (2011). "A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species." *PLoS One* 6(5): e19379.
- Gallais A., Monod J.P., 1998. La gestion des ressources génétiques du maïs en France : de leur caractérisation jusqu'aux premiers stades de leur valorisation. *C. R. Acad. Agric. Fr.*, 84, 173-181.
- Ganal, M. W., G. Durstewitz, A. Polley, A. Berard, E. S. Buckler, A. Charcosset, J. D. Clarke, E. M. Graner, M. Hansen, J. Joets, M. C. Le Paslier, M. D. McMullen, P. Montalent, M. Rose, C. C. Schon, Q. Sun, H. Walter, O. C. Martin and M. Falque (2011). "A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome." *PLoS One* 6(12): e28334.
- Gauthier, P., B. Gouesnard, J. Dallard, R. Redaelli, C. Rebourg, A. Charcosset and A. Boyat (2002). "RFLP diversity and relationships among traditional European maize populations." *Theoretical and Applied Genetics* 105(1): 91-99.
- Gorjanc, G., J. Jenko, S. J. Hearne and J. M. Hickey (2016). "Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations." *Bmc Genomics* 17.
- Gouesnard, B., T. M. Bataillon, G. Decoux, C. Rozale, D. J. Schoen and J. L. David (2001). "MSTRAT: An algorithm for building germ plasm core collections by maximizing allelic or phenotypic richness." *Journal of Heredity* 92(1): 93-94.
- Gouesnard, B., J. Dallard, P. Bertin, A. Boyat and A. Charcosset (2005). "European

maize landraces: genetic diversity, core collection definition and methodology of use." *Maydica* **50** (50th anniversary issue): 225-234.

Gouesnard, B., A. Zanetto and C. Welcker (2016). "Identification of adaptation traits to drought in collections of maize landraces from southern Europe and temperate regions." *Euphytica* **209**: 565-584.

Mir, C., T. Zerjal, V. Combes, F. Dumas, D. Madur, C. Bedoya, S. Dreisigacker, J. Franco, P. Grudloyma, P. X. Hao, S. Hearne, C. Jampatong, D. Laloe, Z. Muthamia, T. Nguyen, B. M. Prasanna, S. Taba, C. X. Xie, M. Yunus, S. Zhang, M. L. Warburton and A. Charcosset (2013). "Out of America: tracing the genetic footprints of the global diffusion of maize." *Theor Appl Genet* **126**(11): 2671-2682.

Nei, M. (1973). "Analysis of gene diversity in subdivided populations." *Proceedings of the National Academy of Sciences of the United States of America* **70**(12): 3321-3323.

Rebourg, C. (1996). *Analyse de la diversité de populations de maïs à l'aide de marqueurs moléculaires. Application à la gestion des ressources génétiques*. DEA Université Pierre et Marie Curie. Orsay.

Rebourg, C., M. Chastanet, B. Gouesnard, C. Welcker, P. Dubreuil and A. Charcosset (2003). "Maize introduction into Europe: the history reviewed in the light of molecular data." *Theoretical and Applied Genetics* **106**(5): 895-903.

Rebourg, C., B. Gouesnard and A. Charcosset (2001). "Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation." *Heredity* **86**: 574-587.

Rincint, R., S. Nicolas, S. Bouchet, T. Altmann, D. Brunel, P. Revilla, R. A. Malvar, J. Moreno-Gonzalez, L. Campo, A. E. Melchinger, W. Schipprack, E. Bauer, C. C. Schoen, N. Meyer, M. Ouzunova, P. Dubreuil, C. Giauffret, D. Madur, V. Combes, F. Dumas, C. Bauland, P. Jamin, J. Laborde, P. Flament, L. Moreau and A. Charcosset (2014). "Dent and Flint maize diversity panels reveal important genetic potential for increasing biomass production." *Theoretical and Applied Genetics* **127**(11): 2313-2331.

Schnell, F. (1983). Probleme der elternwahl - ein überblick. Ber Arbeitstag 1983 Arbeitsgem Saatzuchtleiter, Gumpenstein, Austria.

Shull, G. H. (1908). "The composition of a field of maize." *American Breeder's Association Reports* **4**: 296-301.

Tenaillon, M. I. and A. Charcosset (2011). "A European perspective on maize history." *Comptes Rendus Biologies* **334**(3): 221-228.

VERS UNE NOUVELLE CONCEPTION DU PRE-BREEDING ? QUEL PARTENARIAT PUBLIC-PRIVE ?

Michel RENARD¹, Laurence GARMENDIA², Rémy CAILLIATTE³, Carole CARANTA⁴

¹UMR IGEPP INRA, Agrocampus Ouest, Université Rennes 1, Domaine de la Motte, BP 35327,
F-35653 Le Rheu

²INRA, Adjointe Département BAP- Partenariat, Domaine de la Grande Ferrade, CS 20032,
F-33882 Villenave d'Ornon Cedex

³INRA, Adjoint Département BAP - Innovation variétale, Domaine St Maurice, CS 60094,
F-84143 Montfavet cedex

⁴INRA, Chef du Département BAP, Route de Saint Cyr,
F-78026 Versailles

E-mail : renard@rennes.inra.fr

RESUME

Après avoir défini le pré-breeding, son rôle et son positionnement dans les schémas de sélection, une analyse de l'organisation du partenariat public-privé est présentée en s'appuyant sur des exemples. L'impact des marqueurs moléculaires sur l'efficacité du transfert de caractères innovants dans des pools élites est également souligné par des exemples. Des éléments d'évolution du contexte de mise en œuvre du pré-breeding sont déclinés. Des interrogations sont formulées quant à l'évolution de la répartition des missions entre la recherche publique et la recherche/sélection privée, les moyens d'assurer le transfert, la valorisation des connaissances produites par la recherche publique, et le rôle de celle-ci concernant les espèces mineures, voire orphelines de recherche.

Mots-clés : pré-breeding, définition, rôle, partenariat public-privé.

1 – INTRODUCTION

En 2015, l'INRA, dans le cadre de son Département Biologie et Amélioration des Plantes- (DBAP) a mené une réflexion sur le pré-breeding (PB) en prévision de sa participation au colloque de Montpellier ayant pour thème le Partenariat Public-Privé en Pré-breeding (International Workshop Public-Private Partnerships on Pre-breeding, 2-4 février 2015, Agropolis, Montpellier). Celle-ci nous a amenés à préciser la notion de pré-breeding, son positionnement dans un schéma de sélection, son rôle, ainsi que l'impact des outils moléculaires dans la conception du PB et l'organisation des activités de PB en partenariat. Elle a reposé, en partie, sur le résultat d'une enquête adressée à un certain nombre de responsables de programmes d'amélioration variétale.

2 – QUELLE DEFINITION DU PRE-BREEDING ?

Il existe diverses conceptions du pré-breeding, selon qu'on le considère plus ou moins en amont de l'activité de création variétale. Pour la suite de notre réflexion, nous avons retenu au départ la définition suivante :

« Pre-breeding (PB) refers to all activities designed to identify desirable characteristics and/or genes from unadapted materials that cannot be used directly in breeding populations and to transfer these traits to an intermediate set of materials that breeders can use further in producing new varieties for farmers. It is a necessary first step in the use of diversity arising from wild relatives and other unimproved materials.

These activities are a collaboration between the germplasm curator and the plant breeder who need to work together to understand the scope and value of germplasm collections and how new traits from these collections can be bred into new varieties. The adoption of pre-breeding facilitates the efficiency and effectiveness of crop improvement programmes by enabling increased access to, and use of, genetic variations conserved in genebanks. »
(Source: Pre-breeding for Effective Use of Plant Genetic Resources, GIPB e-learning course)

Cette définition met en avant tout particulièrement le processus de transfert de caractères d'un matériel non adapté dans du matériel directement exploitable par le sélectionneur pour la création de variétés commerciales, tout en soulignant la contribution du PB à une meilleure exploitation de la diversité génétique présente dans les ressources génétiques.

3 – QUAND COMMENCE – QUAND FINIT LE PREBREEDING ?

3.1. Place du PB dans le processus de sélection

Le PB se positionne après le criblage du germplasma et l'identification de progéniteurs intéressants, sur des critères de diversité génétique. Il commence réellement avec l'introduction de variabilité génétique, de nouveaux caractères, dans du matériel adapté. Le PB s'arrête donc au moment de l'application de méthodes de création variétale proprement dite à un pool génétique introgressé afin de créer des variétés. Ainsi, de notre point de vue, le PB se situe en aval des activités de collecte et de maintien des ressources génétiques ainsi que de leur caractérisation phénotypique et génotypique, mais se situe en amont de la sélection de variétés commerciales (Figure 1). Le PB peut cependant, dans certains cas, conduire à la production directe de variétés exploitables commercialement.

Les caractères introduits dans du matériel adapté ont des origines diverses. Ils peuvent être issus :

- des ressources génétiques plus ou moins exotiques de l'espèce sélectionnée ;
- des espèces apparentées sauvages ou cultivées ;
- des biotechnologies vertes : variants issus de mutagénèse ou des NPBT (New Plant Breeding Techniques), organismes génétiquement modifiés issus de la transgénèse.

L'efficacité du PB dépend donc à l'amont des conditions d'accès à la diversité génétique et à des données de caractérisation sur des caractères pertinents de ces ressources. Le schéma de la figure 2 explicite la complexité croissante de l'accès aux ressources génétiques dans le contexte du Traité de Nagoya (www.cbd.int). Dans le cadre de ce traité, applicable sur le territoire de l'Union Européenne depuis octobre 2015 (règlement européen 511/2014), tout utilisateur de ressources génétiques doit s'assurer préalablement à leur utilisation du consentement du fournisseur (consentement matérialisé par un PIC ou Prior Informed Consent), puis il doit établir avec le fournisseur un accord pour un partage équitable des avantages issus de l'utilisation de ces ressources (matérialisé sous forme d'un MAT ou Mutually Agreed Terms). L'ensemble de ce processus est *in fine* validé par l'autorité compétente du pays fournisseur de la ressource génétique (matérialisé sous forme d'un CCIR ou Certificat de Conformité Internationalement Reconnu). Pour prendre en compte les besoins spécifiques liés à la sélection variétale pour l'agriculture et l'alimentation, et en particulier pour éviter de multiples négociations bilatérales de PIC et MAT, la FAO a proposé une implémentation de ce traité dans le cadre du Traité International sur les Ressources Génétiques Liées à l'Agriculture et l'Alimentation (TIRPAA, www.planttreaty.org/fr).

Dans ce cadre, également reconnu par l'Union Européenne, un MTA unique (SMTA) est associé aux échanges et prévoit les modalités de partage des avantages (versement à un fond global). Les difficultés actuelles rencontrées par les utilisateurs de ressources génétiques sont à la fois liées au fait que ce processus d'accès est plus complexe qu'auparavant, mais également à l'hétérogénéité de son application à travers le monde avec des pays qui signent ou non ces différents traités puis qui les implémentent de façon hétérogène dans le temps et dans les modes de gouvernance.

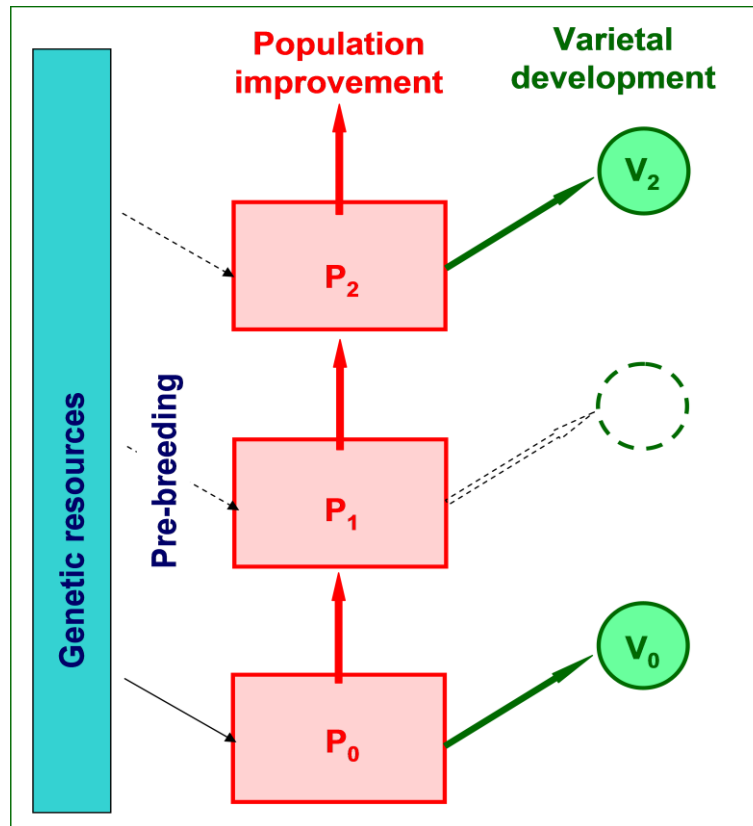
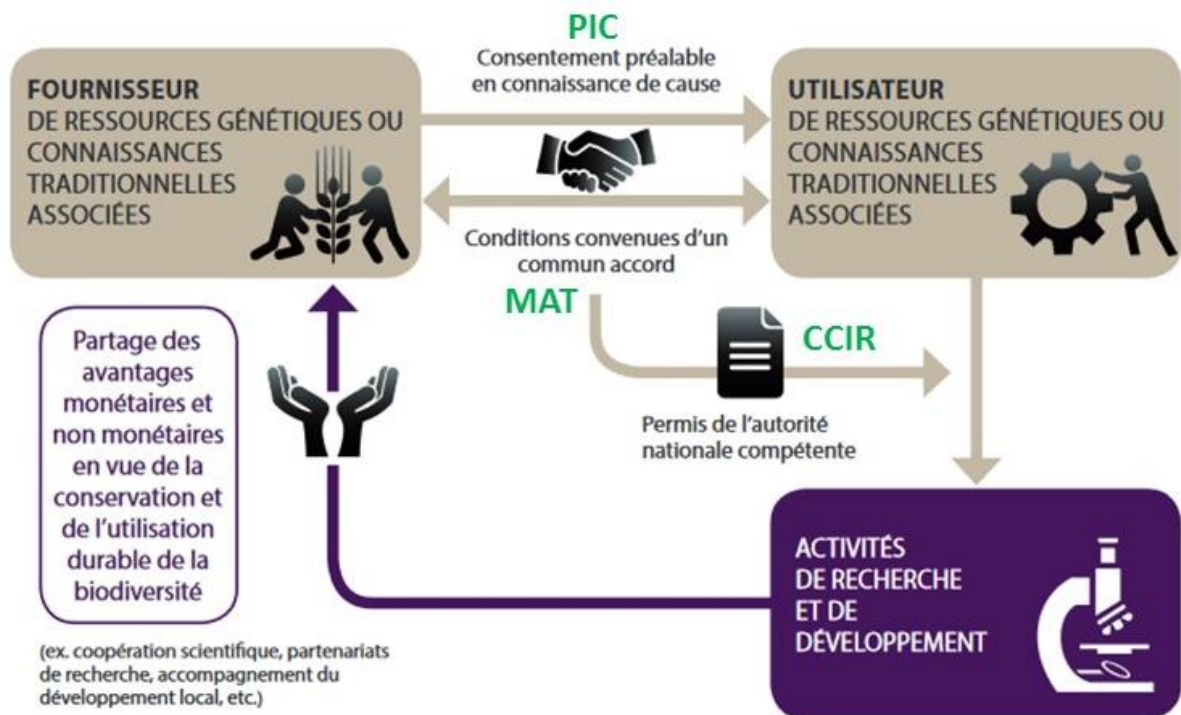


Figure 1 : Place du pré-breeding dans la stratégie de sélection et de création variétale (A. Gallais, Académie Agricole).



Source: FRB 2015

Figure 2 : Processus à mettre en œuvre pour échanger des ressources génétiques dans le cadre du Protocole de Nagoya.
(Source : http://www.fondationbiodiversite.fr/images/documents/APA/Note_FRB_reglement_EU.pdf)

3.2. Rôle du PB

L’objectif du PB est de produire un matériel donneur (lignée, clone...), une population source, voire un pool génétique élite, avec un ou des nouveaux caractères totalement ou partiellement fixés. L’approche de PB se focalise en général sur des caractères monogéniques ou oligogéniques mais peut/pourra également être appliquée à des caractères génétiquement plus complexes. Il s’agit donc :

- d’introduire de la diversité primaire (e.g. nécessité dans le cas du triticales créé par l’homme), dans les programmes de sélection sans baisse drastique de la valeur agronomique ;
- d’adapter la diversité avant de l’utiliser dans des programmes de création variétale ;
- de contribuer progressivement à mieux comprendre l’architecture génétique des caractères introgressés ;
- d’inclure une approche intégrative multi-caractères ;
- de développer, dans le cadre d’un partenariat public-privé, une approche prospective et collaborative de l’introgession de nouveaux caractères d’intérêt, en intégrant parfois une phase de validation ;
- de stimuler le transfert d’une innovation génétique (caractère nouveau) au niveau de l’exploitation agricole, en assurant le transfert de l’innovation dans du matériel élite, en traduisant les connaissances académiques en outils opérationnels et/ou en matériels adaptés.

En quelque sorte, le PB constitue un maillon intermédiaire entre la diversité primaire et le matériel élite.

4 – ORGANISATION DU PRE-BREEDING

4.1. Articulation entre recherche publique et acteurs privés

L’activité de PB peut être portée, soit par des acteurs publics, d’ailleurs souvent en charge également de programmes de sélection, soit directement par des acteurs privés. Cependant, elle s’inscrit le plus souvent dans le cadre d’un partenariat public-privé. Par essence, elle nécessite des compétences en sélection variétale.

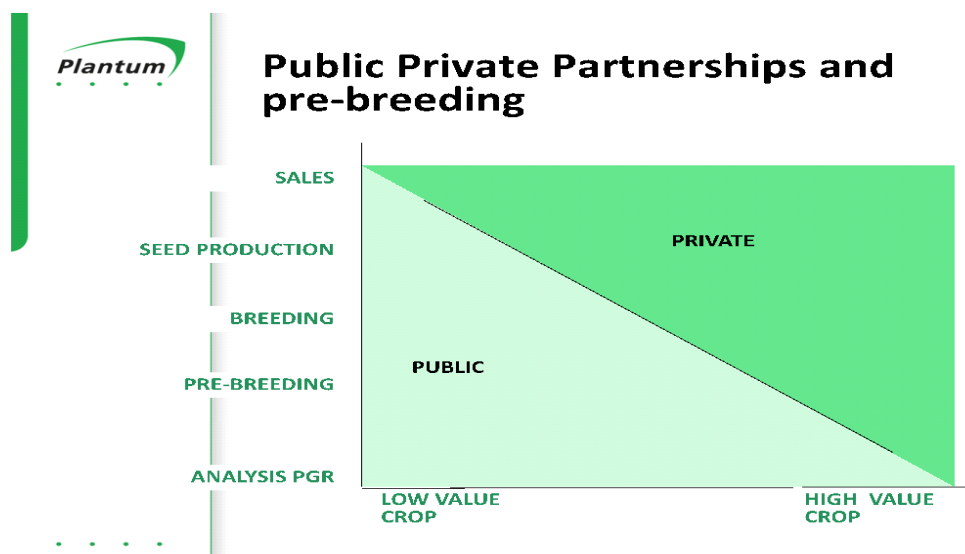


Figure 3 : Répartition possible du rôle de la recherche publique et privée dans le processus de recherche, de sélection et de commercialisation de variétés végétales en fonction de la valeur économique de l’espèce (source : Anke Vandenhurk – Plantum, Février 2016).

L'intérêt des acteurs privés dépend de la valeur économique de l'espèce, ce qui impacte la répartition des investissements entre recherche publique et recherche privée. Pour ce qui concerne les espèces stratégiques sur le plan économique, l'activité de PB, voire de recherche, reste internalisée dans le but de maximiser les marges et de garder un avantage concurrentiel. Ainsi le partenariat public-privé pour le PB dépend du statut économique de l'espèce (lui-même fonction des modalités de valorisation de cette espèce), de l'opportunité du projet de recherche et de l'intérêt pour le partenaire privé à y contribuer, financièrement et/ou en s'impliquant directement dans le projet. Globalement, le secteur privé tend à déployer une approche sectorisée dans laquelle chaque espèce travaillée est source de rentabilité économique. Dans ce contexte, le secteur public se doit de trouver un équilibre entre un partenariat public-privé portant sur les espèces majeures, avec une activité de recherche amont positionnée à l'international, et un continuum recherche-PB-sélection pour des espèces dites mineures, en réponse à la demande sociale et en cohérence avec les politiques publiques nationales et européennes.

4.2. Modalités et exemples de partenariat

Dans le domaine du PB, le partenariat de l'INRA se construit autour des principaux critères suivants :

- la liberté d'utilisation des résultats (Freedom To Operate) ce qui implique un partenariat non exclusif dans la valorisation des résultats;
- un partenariat privilégié à destination des entreprises françaises et des PME/TPE tout en restant en cohérence avec les dynamiques internationales ;
- un partenariat associant les organisations professionnelles, en lien avec les programmes de sélection de l'INRA dans le but de préparer les conditions d'adoption et d'utilisation les plus pertinentes des innovations issues des travaux collaboratifs.

Des programmes de PB sont mis en place sur le long terme par l'INRA avec des consortia larges (e.g., Promosol pour les oléagineux, Promaïs, UFS-Plantes légumières, Terres Inovia et Terres Univia pour les légumineuses à graines), des consortia plus restreints (e.g., GIE Colza), des partenariats bilatéraux (avec IFV pour la vigne et CEP Innovation pour les arbres fruitiers) ou encore sans partenaire privé, dans le cas de projets de création variétale propres à l'INRA. Ces exemples témoignent de la nécessité d'adapter les modalités de partenariat aux espèces travaillées et à la structuration des filières impliquées. Dans ce contexte, l'INRA privilégie une approche au cas par cas prenant en compte les spécificités de chaque situation.

Des exemples de résultats issus de ces partenariats INRA-privés sont donnés ci-dessous.

a) Partenariat fédéré

Avec ce type de partenariat, ont été créées :

- avec Promosol (Association loi 1901 : UFS, Terres Inovia, Terres Univia et INRA sur les oléagineux dont le colza) : des lignées MS (cybrides) mais également des lignées restauratrices (gène *Rfo*) ainsi que des lignées naines issues de mutagenèse (gène *Bzh*) en tant que géniteurs pour l'utilisation en sélection ; plus récemment, de la variabilité génétique des progéniteurs *Brassica oleracea* et *B. rapa* a été introgressée dans une population de colza, en stimulant la recombinaison homologue ;
- avec le GIE blé dur, des populations exotiques ;
- avec le GIE Triticale des populations améliorées (Figure 4) ;
- avec le GIE Promaïs, une organisation pour le maintien des populations françaises et des populations sources à base génétique large ;
- avec l'UNIP/Terres Inovia, des lignées de pois d'hiver résistantes à certaines maladies.

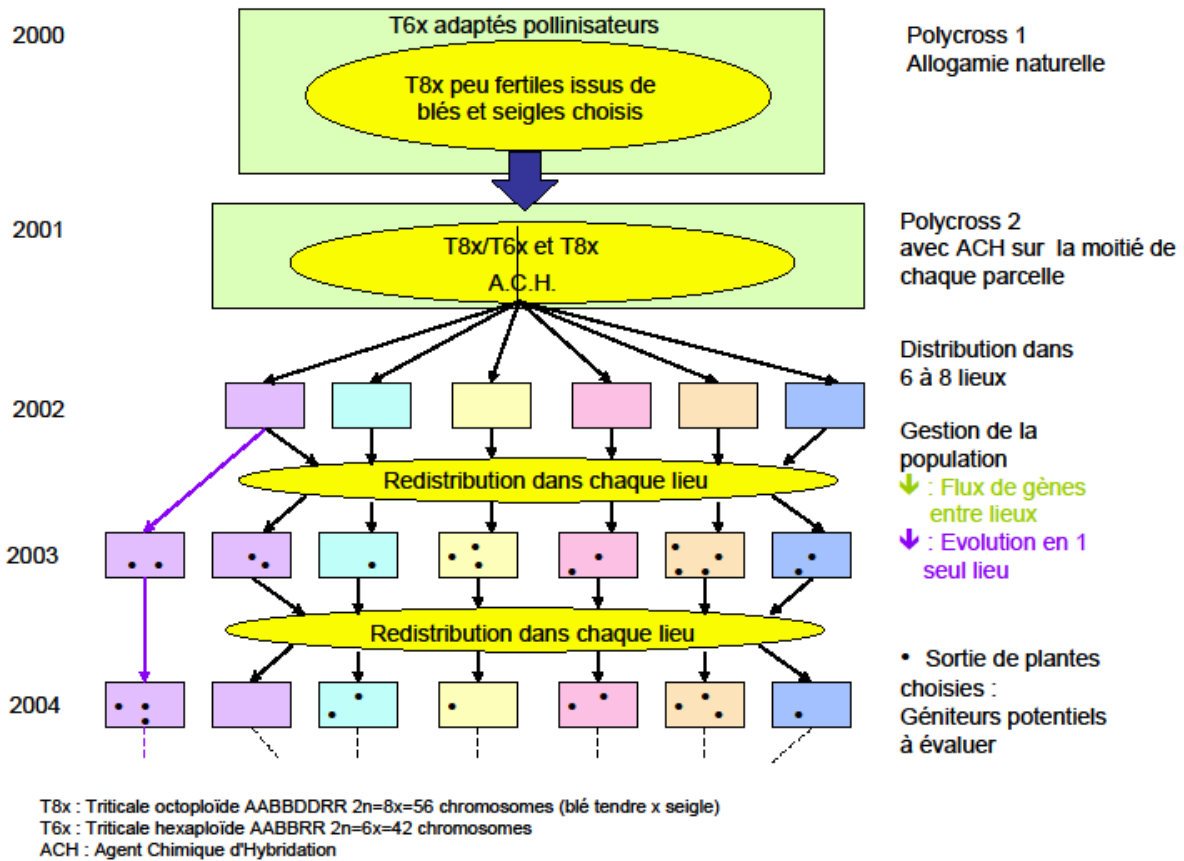


Figure 4 : Exemple de schéma de pré-breeding pour la sélection du triticale (A. Bouguennec, UMR GDEC, Clermont-Ferrand).

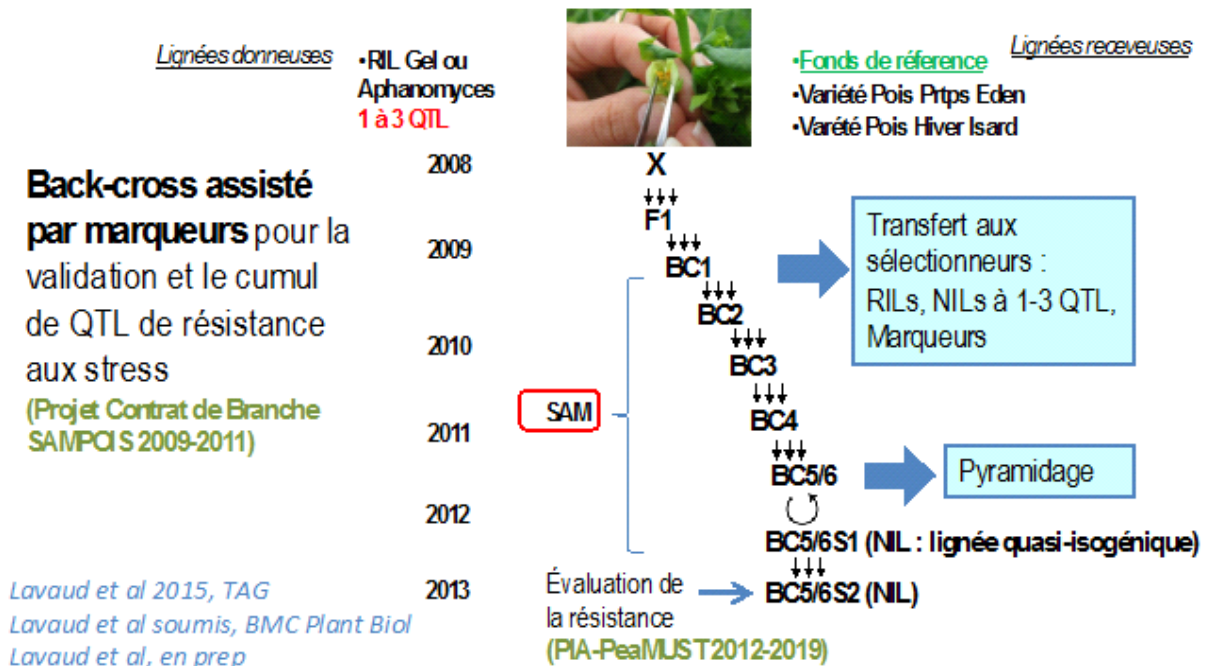


Figure 5 : Utilisation de marqueurs moléculaires pour le pyramidage de QTL de résistance à *Aphanomyces* dans des lignées élites de pois (M.L. Pilet, UMR IGEPP, Le Rheu).

Différentes combinaisons de QTL de résistance à *Aphanomyces* ont ainsi été créées par *back-cross* assistés de marqueurs (Figure 5), ce qui a débouché sur la production de géniteurs présentant un meilleur niveau de résistance. Ce matériel végétal ainsi que les marqueurs associés sont utilisés par les partenaires pour la sélection de nouvelles variétés.

b) Partenariat plus restreint

Avec le GIE Colza, ont été produites des populations élites pour modifier la composition en acides gras ou convertir un pool printemps en type hiver, à partir de matériel élite des partenaires privés.

Le transfert du caractère haute teneur en acide oléique et faible teneur en acide linoléique, contrôlé par quatre gènes indépendants de désaturase dans une population source élite a été réalisé par *back-cross* assistés de marqueurs en utilisant des lignées à bonne valeur agronomique ainsi que le système d'hybridation ogu-INRA pour brasser le matériel. L'utilisation de marqueurs a permis de sélectionner le caractère recherché à chaque génération sans avoir à passer par deux générations d'autofécondation nécessaires à un phénotypage robuste. Ce programme a abouti à la production d'une population à bonne valeur en lignées exploitable par les différents partenaires pour la sélection de nouvelles variétés lignées ou hybrides.

c) Partenariat bilatéral

Ont été sélectionnés ou sont en cours de sélection :

- avec le CEP-Innovation, des fruitiers résistants à certaines maladies ;
- avec l'IFV, des vignes résistants à l'oïdium et au mildiou.

L'efficacité des programmes de PB dépend de plus en plus de l'identification de marqueurs moléculaires du caractère, exploitables en sélection. Il est souvent préférable d'attendre la disponibilité de ces marqueurs avant de se lancer dans un programme de PB reposant uniquement sur du phénotypage.

5 - PERSPECTIVES

Les activités de PB seront impactées à court terme par :

- le développement d'approches de génomique et post-génomique, lesquelles devraient booster le transfert de gènes et la valorisation du germplasm ;
- le développement d'approches translationnelles permettant de faire bénéficier les espèces « mineures » des avancées des espèces dites « majeures », palliant ainsi le déficit de données de séquençage du génome, de techniques de régénération... de l'espèce d'intérêt ;
- la plus grande sensibilisation aux enjeux de durabilité et de diversification qui tend à encourager les travaux sur des combinaisons originales de traits et sur des espèces jusqu'à présent peu travaillées (moutarde, bleuet...) ;
- le développement et l'accès de plus en plus routinier au séquençage et au génotypage haut-débit ;
- les NPBT (pour muter spécifiquement un gène plutôt que de transférer l'allèle original accompagné des gènes environnants qui lui sont liés) ;
- l'utilisation de “*short cycling genes*” chez les espèces pérennes.

Dans le cas des fruitiers, il est ainsi envisagé d'utiliser un transgène de précocité afin de pouvoir réaliser une génération par an dans le transfert, à partir d'espèces apparentées, de gènes de résistance à la tavelure pour le pommier (Figure 6). L'utilisation dans le futur de tels schémas en sélection dépendra fortement du statut qui sera donné aux ségréants négatifs. Cette même approche est en cours d'étude chez la vigne, mais en utilisant simplement une mutation naturelle récessive.

Pseudo-backcross scheme done at Agroscope using the early flowering approach to develop pre-breeding material carrying the fir b light resistance gene of Evereste



Slide kindly provided by A. Patocchi
Agroscope, Switzerland.

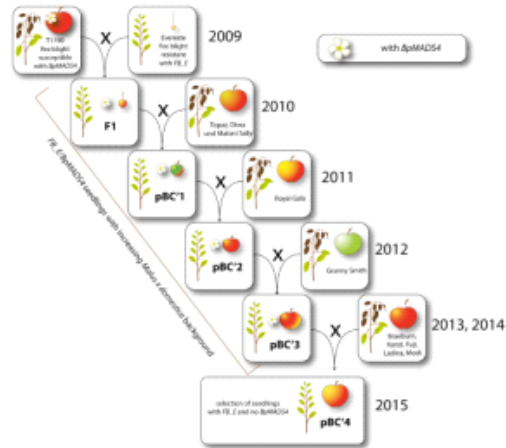


Figure 6 : Schéma d’utilisation en pré-breeding d’un gène de floraison précoce pour l’accélération des cycles de sélection du pommier (A. Patocchi, Agroscope, Suisse, Janvier 2016).

7 - QUESTIONS A DEBATTRE SUR LE PRE-BREEDING

Il ressort de cette première analyse la nécessité d’approfondir la réflexion plus particulièrement sur certaines questions clés :

- Comment améliorer l’efficacité des schémas de PB ? Les approches de PB peuvent induire des questions de méthodologie de la sélection, de gestion dynamique des pools génétiques ; des questions génériques portant sur l’augmentation et le pilotage de la recombinaison...
- Comment soutenir financièrement les activités de PB ?
- Comment soutenir en particulier le PB d’espèces mineures voire orphelines de recherche? Faut-il envisager de mutualiser une partie des bénéfices de l’ensemble de l’activité de sélection nationale afin de dégager une enveloppe dédiée spécifiquement à ces filières de diversification ?
- Quel est l’impact de la taille de la firme de sélection/du partenaire privé sur la demande réelle auprès de la recherche publique ?
- Quel sera l’impact des biotechnologies vertes sur la demande future des partenaires privés ?

8 – CONCLUSION

Le PB s’inscrit dans un continuum recherche-PB-Sélection. Le PB a longtemps mis en œuvre des schémas de rétrocroisements reposant le plus souvent sur une sélection phénotypique du caractère à introgresser. Les nouvelles technologies devraient fortement améliorer l’efficacité de ces schémas, voire s’y substituer par l’induction de mutations ciblées, appliquées directement à du matériel élite plus directement exploitable en création variétale.

Le PB participe fondamentalement au processus de transfert et de valorisation de caractères innovants. La valorisation des activités de la recherche publique en génétique végétale dépend ainsi des ressources allouées aux activités de PB et de l'organisation de son partenariat avec la recherche privée. La recherche publique doit s'assurer d'un partenariat public-privé diversifié tout en l'adaptant au contexte des filières concernées et en équilibrant ses investissements entre les espèces majeures et les espèces dites mineures voire orphelines, en réponse aux attentes sociétales.

« Journée ASF du 4 février 2016 »
« Apport des marqueurs moléculaires et de la génomique
à l'utilisation des ressources génétiques pour l'amélioration des plantes »

REMERCIEMENTS

Anne-Françoise Adam-Blondon (Ressources Génétiques, DBAP, Versailles)
Jean-Marc Audergon (Fruitiers, GAFL, Avignon)
Annaïg Bouguennec (Triticale, GDEC, Clermont-Ferrand)
Mathilde Causse (Tomate et Fruitiers, UAFL, Avignon)
Alain Charcosset (Maïs, GQE, Le Moulon)
Anne-Marie Chèvre (Colza, IGEPP, Le Rheu)
Jacques David (Blé dur, AGAP, Montpellier)
André Gallais (Schémas de sélection, Académie d'Agriculture de France)
Marc Ghesquière (Fourrages, PPPF, Lusignan)
Joseph Jahier (Blé, IGEPP, Le Rheu)
Alain Label (Réglementation, IGEPP, Ploudaniel)
François Laurens (Fruitiers, IRHS, Angers)
Marie-Laure Pilet-Nayel (Pois, IGEPP, Le Rheu)