

Présentation des projets financés au titre de l'édition 2010 du  
 Programme « Génomique, Biotechnologies végétales »

<b>ACRONYME et titre du projet</b>	<b>Page</b>
<b>BrassiNAM</b> – Développement d'une population NAM pour l'analyse de caractères complexes chez Brassica napus	3
<b>CartoSeq</b> – Identification en masse des variants génétiques influençant les caractères d'élevage chez les trois principales races laitières françaises	5
<b>Chloro-types</b> – Adaptation du chloroplaste aux stress abiotiques : utilisation de la protéomique pour révéler les phénotypes moléculaires	7
<b>CNV-MAIZE</b> – Étude d'association sur génome entier entre variation structurale, variation des caractères d'intérêt agronomique et hétérosis chez le maïs	9
<b>DURAVITIS</b> – Bases développementales, moléculaires et génétiques de l'adaptation de la vigne à la contrainte thermique	11
<b>EFG-MIG</b> – Génomique évolutive et fonctionnelle de biosynthèse des glucosinolates indoles modifiés	13
<b>GEMBAL</b> – Génomique multi-race des Bovins Allaitants et Laitiers	15
<b>GENAMIBE</b> – Comprendre la pathogénicité de Entamoeba par l'utilisation de la transcriptomique et la phylogénétique comparatives	17
<b>GnpAsso</b> – Un outil générique pour la gestion et l'exploitation des résultats de génétique d'association utilisant les données de génotypage et de phénotypage hauts débits	19
<b>Immunit-Ae</b> – Diversité des composants génétiques et des mécanismes de résistance à Aphanomyces euteiches chez les légumineuses	21
<b>PHENOBLE</b> – Mise au point et valorisation d'outils de phénotypage nouvelle génération pour l'analyse des déterminants génétiques de l'efficacité d'utilisation des engrais azotés chez le blé tendre	23
<b>PhyloFish</b> – Analyse phylogénomique des duplications géniques chez les poissons téléostéens: une approche par RNA-Seq	25

<b>RepliColScope</b> – Séquençage et génomique comparative des plasmides de Escherichia coli : apport à la compréhension du succès évolutif des $\beta$ -lactamases à spectre étendu	27
<b>SOLAR</b> – Développer et tester de nouveaux outils biotechnologiques pour accroître la teneur en huile de la graine : une clé pour la production de chaînes hydrocarbonées chez les oléagineux	29
<b>SUS_FLORA</b> – Contribution du microbiote intestinal à l'homéostasie du système immunitaire chez le porc: approches génétiques et génomiques	31
<b>UtOpIGe</b> – Vers une Utilisation Optimale de l'Information Génomique dans les schémas pyramidaux	33
<b>WALLARRAY-II</b> – Biopuces à polysaccharides pariétaux en génomique fonctionnelle	34
<b>XANTHOMIX</b> – Etude comparative des génomes et des transcriptomes de Xanthomonas phytopathogènes	36

**Titre du projet**

**BrassiNAM** – Développement d’une population NAM pour l’analyse de caractères complexes chez *Brassica napus*

**Résumé**

Le colza est la culture oléagineuse la plus cultivée en Europe. La surface cultivée ne cesse de croître depuis 2003 dû au fort développement de la demande en huile de colza, à la fois pour l’alimentation humaine et pour des utilisations non alimentaires. Récemment la demande sur les produits agricoles, et sur l’huile végétale en particulier, s’est fortement accrue. Ceci est surtout lié à la demande alimentaire croissante des pays émergents ainsi qu’au développement des bioénergies. Il devient clair maintenant que pour satisfaire à l’augmentation de la demande, la productivité des cultures, et notamment celle des cultures oléagineuses, doit s’accroître fortement. Dans le même temps les intrants (pesticides, fertilisants,...) doivent être réduits à la fois pour des raisons écologiques et économiques évidentes. Dans ce contexte, l’amélioration génétique est une réponse majeure permettant le développement d’une agriculture durable respectueuse de l’environnement. Au cours des dernières années, un véritable saut technologique s’est produit dans le domaine de la sélection génétique lié notamment au développement des technologies de séquençage et de génotypage à haut débit. Parallèlement des résultats très prometteurs ont été obtenus récemment dans le domaine de la génétique d’association. L’association de ces nouvelles technologies et méthodes d’analyse permet maintenant d’envisager avec succès la dissection de caractères complexes majeurs et leur utilisation en amélioration variétale. Dans cet objectif, un nouveau type de population a été très récemment développé avec succès chez le maïs et est en cours de développement chez le blé. Cette nouvelle ressource permet de rassembler les avantages des deux principales méthodes d’analyses permettant de disséquer les caractères quantitatifs : les analyses de liaisons et les analyses de déséquilibre de liaison. Cette population appelée NAM, pour Nested Association Mapping, constitue un outil performant pour valider rapidement et en grande quantité des gènes candidats responsables des caractères observés et pour caractériser finement des QTL chez les plantes. Le projet OSR\_NAM vise à développer chez le colza une population NAM qui fournira aux

scientifiques et aux sélectionneurs impliqués dans l'amélioration du colza, un outil majeur de cartographie génétique à haute résolution. Cette ressource sera développée par 2 laboratoires leaders en génétique, génomique et en génotypage, au sein d'un projet regroupant à la fois la recherche publique et privée. Cette association garantit l'exploitation de cette ressource à la fois pour une recherche fondamentale mais aussi pour une recherche appliquée, visant à mettre sur le marché de nouvelles variétés aux performances agronomiques accrues.

**Partenaires** Biogemma  
UMR\_APBV

**Coordinateur** BRUNO GREZES-BESSET - Biogemma  
[bruno.grezes-besset@biogemma.com](mailto:bruno.grezes-besset@biogemma.com)

**Aide de l'ANR** 477 k€

**Début et durée** Janvier 2011 - 48 mois

**Référence** ANR-10-GENM-001

**Label pôle** AGRIMIP INNOVATION

Titre du projet

**CartoSeq – Identification en masse des variants génétiques influençant les caractères d'élevage chez les trois principales races laitières françaises**

Résumé

L'ANR a financé ces dernières années plusieurs projets de génomique bovine, dont le projet CartoFine. Le but de ce projet était la localisation fine de QTLs pour des caractères laitiers chez les trois principales races laitières présentes en France, Holstein, Normande et Montbéliarde. Chacun des 3200 taureaux d'insémination artificielle choisis pour ce projet, a été caractérisé par les performances d'une centaine de filles, sur 25 caractères de production et d'adaptation (par exemple la production et composition du lait, la fertilité ou la résistance aux mammites). Tous ces taureaux ont été génotypés pour 54001 SNPs à l'aide de la puce bovine SNP50K d'Illumina. Une analyse pan-génomique combinant analyse de liaison et analyse du déséquilibre de liaison a permis de détecter et de localiser finement 305 régions QTLs. La plupart des QTLs identifiés ne sont pas partagés entre races. La sélection génomique permet désormais d'améliorer la sélection génétique d'animaux laitiers, cependant comme la plupart des SNPs présents sur la puce Illumina, ne sont pas dans des gènes et à cause du déséquilibre de liaison, les SNPs associés avec les caractères agronomiques d'intérêt, ne sont pas les variants génétiques directement responsables de la variabilité de ces caractères. L'identification des variants directement impliqués dans les phénotypes d'intérêt, reste une tâche difficile. Il est donc important de développer des stratégies pour cibler rapidement les variants génétiques responsables de ces phénotypes d'intérêt. L'identification des variants génétiques causaux, aussi appelés QTNs (quantitative trait nucleotides) implique la localisation des QTLs, la découverte de nouveaux marqueurs génétiques dans ces régions QTLs, la localisation fine de ces QTLs puis le re-séquençage de gènes candidats. Ce processus itératif prenait jusqu'à récemment beaucoup de temps. Grâce à la disponibilité d'un très grand nombre de SNPs et de méthodes de génotypage pan-génomique relativement peu chères, la localisation fine des QTLs est maintenant grandement accélérée. L'apparition de nouvelles méthodes de séquençage à très haut débit permettant de séquencer des génomes entiers de mammifères à un coût relativement abordable, offre désormais de nouvelles opportunités pour la découverte des QTNs. Par exemple, Eck et

collaborateurs (2009) ont récemment produit la séquence pan-génomique d'un taureau Fleckvieh. Ils ont généré 24 Gb de séquences en utilisant des courtes lectures de 36 bases, avec une couverture d'environ 7,4 fois. Cette couverture était suffisante pour identifier 2,44 millions de SNPs, dont 82% étaient nouveaux, et 115000 petites insertions/délétions. Une comparaison entre les génotypes obtenus pour le même animal, avec la puce bovine SNP50K d'Illumina, révèle que 74% et 30% des génotypes homozygotes et hétérozygotes pouvaient être détectés, respectivement. Le but du projet de recherche proposé est le développement d'une approche à grande échelle basée sur l'utilisation du séquençage nouvelle génération, pour identifier les QTNs influençant les caractères laitiers chez les trois principales races laitières présentes en France (Holstein, Normande et Montbéliarde). Le projet proposé est la suite du projet CartoFine, précédemment financé par l'ANR. A notre connaissance, ce projet est la première étude pan-génomique d'association à grande échelle basée sur le re-séquençage, chez les bovins. Les méthodes et outils qui résulteront de ce projet, pourront être utilisés dans de futures analyses d'association pour cibler les variants génétiques responsables de caractères d'intérêt et pourront être facilement adaptés à l'identification de QTNs pour d'autres caractères ou d'autres espèces. Le projet proposé bénéficiera donc non seulement à la communauté scientifique travaillant sur la production laitière ou le génome bovin, mais également potentiellement aussi à d'autres scientifiques.

**Partenaires**

G2B  
UNCEIA  
PlaGe  
SIGENAE

**Coordinateur**

Dominique Rocha - G2B  
[dominique.rocha@unilim.fr](mailto:dominique.rocha@unilim.fr)

**Aide de l'ANR**

412 k€

**Début et durée**

Janvier 2011 - 36 mois

**Référence**

ANR-10-GENM-018

**Label pôle**

**Titre du projet**

**Chloro-types – Adaptation du chloroplaste aux stress abiotiques : utilisation de la protéomique pour révéler les phénotypes moléculaires**

**Résumé**

L'objectif du projet Chloro-types est d'étendre notre connaissance sur le lien qui existe entre la régulation du métabolisme du chloroplaste, son adaptation aux conditions environnementales et la dynamique de la biogénèse de cet organite. Ainsi nous proposons de déchiffrer les mécanismes qui soutendent la régulation de la dynamique du protéome chloroplastique et d'étendre la compréhension des processus qui permettent l'adaptation du chloroplaste à différentes conditions de stress. Dans le cadre d'un projet collaboratif (ANR Génoplante 2007-2010; projet "GlycoChloroplast") réalisé entre nos deux laboratoires (laboratoires LPCV and EDyP, associés dans le projet Chloro-types), nous avons produit des outils importants pour analyser la dynamique du protéome du chloroplaste à l'échelle d'un organite entier. L'un de ces outils, encore unique, est la première base de données AMT consacrée au chloroplaste, la base de données AT-CHLORO. Cette base de données protéomiques a été générée de telle façon qu'elle peut être utilisée pour des analyses quantitatives (par exemple la comparaison de mutants ou l'impact de conditions variables de culture) en reposant sur la stratégie AMT (accurate mass and time tags). Une autre avancée récente, issue du travail collaboratif des laboratoires PCV et EDyP (autres programmes Génoplante), est l'identification d'un mécanisme de régulation original pour l'import des protéines dans les chloroplastes. Dans ce processus, les protéines adressées au chloroplaste sont dans un premier temps séquestrées dans le cytosol au travers d'une interaction avec un partenaire protéique, puis, libérées dans le cytosol par un signal spécifique avant d'être importées dans le chloroplaste. Des données préliminaires indiquent que ce mécanisme de régulation pourrait contrôler différentes voies métaboliques essentielles du chloroplaste et pourraient être sous contrôle de conditions environnementales particulières. L'objectif de ce projet est 1) d'identifier les conditions de stress abiotique qui ont un impact sur ces protéines du chloroplaste qui sont adressées de manière spécifique au chloroplaste, 2) de déchiffrer la régulation des mécanismes induits par le stress et qui

contrôlent l'adressage des protéines au chloroplaste et 3) de comprendre la signification physiologique de ces mécanismes.

**Partenaires** CNRS-LPCV  
INSERM-LEDyP

**Coordinateur** Norbert ROLLAND - CNRS - LPCV  
[norbert.rolland@cea.fr](mailto:norbert.rolland@cea.fr)

**Aide de l'ANR** 371 k€

**Début et durée** Janvier 2011 - 48 mois

**Référence** ANR-10-GENM-002

**Label pôle**



**Titre du projet**

**CNV-MAIZE** – Étude d'association sur génome entier entre variation structurale, variation des caractères d'intérêt agronomique et hétérosis chez le maïs

**Résumé**

Les variations structurales (SV) intraspécifiques des génomes, telles que les insertions, délétions, duplications, inversions et translocations, ont été jusqu'ici considérées comme des évènements rares impactant peu les variations phénotypiques. De récentes études montrent que ces SV, en particulier chez le maïs, concernent une part du génome plus importante qu'attendu. La compréhension de la structure, de l'évolution et de la variabilité du génome du maïs nécessite une étude approfondie de ces SV. Le maïs étant l'une des céréales majeures, la découverte de SV particulières affectant le nombre de copie (CNV) ou la présence/absence (PAV) de séquences géniques est importante pour évaluer leur contribution aux variations des caractères complexes et à l'hétérosis. Les ressources génétiques du maïs comprennent des lignées très contrastées présentant vraisemblablement d'importants SV. Nous proposons dans ce projet une stratégie originale combinant les techniques de séquençage de nouvelle génération et l'hybridation génomique comparative sur puce d'ADN (aCGH) pour la caractérisation des SV du maïs à l'échelle du génome. Ce projet sera conduit par trois équipes françaises leaders en génomique et génétique du maïs ainsi qu'en statistique pour l'étude des génomes, ce qui permettra de rassembler des compétences complémentaires en bioinformatique, biologie moléculaire et statistiques. Le projet bénéficiera des appuis décisifs offerts par une équipe américaine pionnière du domaine, qui a contribué au séquençage du génome du maïs, d'une société de biotechnologie leader dans la conception et la production des aCGH et du Génoscope. Dans un premier temps, nous identifierons à l'aide de puces aCGH les CNV présents dans une «core collection» de lignées américaines et européennes. La séquence du génome de la lignée B73 a été achevée et une puce aCGH est d'ores et déjà disponible. Pour mettre en évidence le plus de CNV possibles, notamment ceux qui pourraient être spécifiques des maïs européens, nous allons enrichir cette puce aCGH avec les régions du génome spécifiques de la lignée française F2, largement utilisée dans

les programmes de sélection européens. L'hybridation des ADN génomiques de notre core collection sur cette nouvelle puce CGH pangénomique permettra de caractériser l'étendue, l'organisation et la nature des CNV chez le maïs. Dans un second temps, les CNV les plus pertinents parmi ceux découverts dans la phase précédente seront utilisés pour construire une puce aCGH innovante destinée au génotypage d'un large panel d'association. Pour déterminer si l'information apportée par les SV peut être complémentaire de celle issue des marqueurs conventionnels (SNP) pour des études de génétique d'association, nous caractériserons la nature et la prévalence de chaque SV d'intérêt ainsi que sa liaison avec les autres types de polymorphismes. Des approches de LD mapping («Linkage Disequilibrium») que nous aurons spécialement adaptées pour l'utilisation des SV permettront d'étudier les associations génotype-phénotype pour des caractères d'intérêt agronomique au niveau de l'ensemble du génome (GWAS). Nous essaierons en particulier de déterminer la valeur prédictive des SV au regard de l'intensité de l'hétérosis. Nous développerons également les méthodologies nécessaires pour la prise en compte de ces nouvelles informations dans les programmes de sélection variétale. Finalement, à partir des SV les plus significatifs en GWAS, nous construirons une boîte à outils de marqueurs CNV au profit de la communauté maïs. Ces nouveaux marqueurs seront très utiles aux futures études sur la recombinaison, la plasticité des chromosomes et l'hétérosis. Le développement de marqueurs spécifiques des lignées européennes offrira un apport crucial pour les travaux de sélection européens et, de façon générale, contribuera significativement aux futurs programmes de recherche sur le maïs en Europe.

**Partenaires**

UMR320  
Biogemma  
UMR518

**Coordinateur**

Stéphane Nicolas - UMR320  
[stephane.nicolas@versailles.inra.fr](mailto:stephane.nicolas@versailles.inra.fr)

**Aide de l'ANR**

232 k€

**Début et durée**

Novembre 2010 - 36 mois

**Référence**

ANR-10-GENM-003

**Label pôle**

Céréales Vallée

**Titre du projet**

**DURAVITIS – Bases développementales, moléculaires et génétiques de l'adaptation de la vigne à la contrainte thermique**

**Résumé**

Comme d'autres plantes cultivées, la vigne est confrontée à des changements climatiques, notamment à une augmentation de la température, peu rassurants pour ses perspectives d'évolution. Les températures élevées sont connues pour agir défavorablement sur certaines composantes du rendement (ex : initiation des boutons floraux, anthèse, avortement des fleurs et jeunes baies), sur la qualité des raisins, ainsi que sur les qualités organoleptiques des vins qui en sont issus (ex : acidité (fraîcheur), aspect (couleur), agrément aromatique (précurseurs)). La qualité alimentaire des vins est également impactée par les températures élevées qui réduisent leurs propriétés anti-oxydantes et augmentent leur toxicité (éthanol). En permettant la sélection de cultivars plus adaptés, la création variétale par hybridation, combinée aux biotechnologies, peut permettre une adaptation durable de la viticulture aux contraintes thermiques annoncées. Mais, pour pouvoir mettre en oeuvre des programmes de sélection efficaces, les mécanismes de réponse de la vigne aux températures élevées doivent être connus et les caractères d'adaptation identifiés. C'est la problématique centrale du projet DURAVITIS. Il s'agit de réévaluer l'impact de la température sur les développements végétatifs et reproducteurs de la vigne selon différents angles d'analyses, comprenant le niveau plante entière et le micro-environnement. Le but est de développer un cadre d'analyse de l'organogenèse, de la morphogénèse, et des composantes de la qualité des baies, pour comprendre les effets de la température sur la production. Une autre finalité du projet DURAVITIS est d'identifier les régions impliquées dans la tolérance de la vigne à la température en étudiant des ressources génétiques contrastées pour ces caractères. Compte tenu des objectifs variés de ce projet et de la complexité des cadres d'analyses envisagés, ce projet est conçu comme une collaboration multi-disciplinaire associant écophysiolistes, physiologistes moléculaires et généticiens. Nous pensons que l'accès à des outils de génomique haut-débit combiné avec le développement d'approches

expérimentales innovantes (boutures fructifères, microvignes) aidera à la dissection des mécanismes écophysiologiques et moléculaires de la réponse adaptative de la vigne aux fortes températures, et facilitera l'identification de caractères génétiques de tolérance. Plusieurs types de retombées sont attendues, en particulier : i) Retombées scientifiques - DURAVITIS devrait fournir des informations originales sur ces aspects et conduire à une avancée significative de la compréhension des mécanismes de réponses aux contraintes environnementales chez les plantes à fruits. ii) Développements technologiques - Le projet DURAVITIS propose un nouveau format méthodologique basé sur l'utilisation de boutures fructifères et de microvignes pour accélérer et améliorer à la fois les études d'écophysiologie et de génétique. En contournant plusieurs difficultés expérimentales associées aux vignes classiques, ce projet offre la possibilité d'établir un cadre d'analyse phénotypique qui pourrait ensuite servir à traiter tout un ensemble de questions variées en rapport avec l'adaptation de la vigne à un environnement contraignant. iii) Opportunités économiques - La sélection de nouvelles variétés capables de faire face aux changements climatiques, notamment à l'augmentation de la température, tout en assurant une production de qualité semble l'approche la plus efficace dans une perspective de long terme. L'étude proposée dans DURAVITIS est un pré-requis pour établir les principes de l'analyse phénotypique et la sélection génotypes tolérants aux fortes températures.

**Partenaires**

UMR DIAPC  
UMR LEPSE  
UMR SPO  
UMR EGFV

**Coordinateur**

Laurent Torregrosa - UMR DIAPC  
[laurent.torregrosa@supagro.inra.fr](mailto:laurent.torregrosa@supagro.inra.fr)

**Aide de l'ANR**

547 k€

**Début et durée**

Janvier 2011 - 48 mois

**Référence**

ANR-10-GENM-004

**Label pôle**

<b>Titre du projet</b>	<b>EFG-MIG – Génomique évolutive et fonctionnelle de biosynthèse des glucosinolates indoles modifiés</b>
<b>Résumé</b>	<p>Le système glucosinolate-myrosinase est un mécanisme de défense inductible chez la plante modèle <i>Arabidopsis thaliana</i> et chez les espèces proches de l'ordre Brassicales. Ce système protège les plantes efficacement contre la plupart des insectes herbivores et d'autres ravageurs, en générant au moment de l'attaque des molécules effectrices toxiques à partir de précurseurs biologiquement inactifs. Le principe fonctionnel de base de ce mécanisme de défense inductible est simple, mais le système lui-même est extrêmement complexe, et comporte une très grande variabilité de structure et de régulation, à la fois aux niveaux inter et intra-espèce. Les recherches précédentes se sont en grande partie concentrées sur les glucosinolates dérivés de méthionine, la classe la plus abondante et la plus diversifiée structurellement parmi les glucosinolates chez <i>Arabidopsis</i>. Ce n'est que très récemment que l'énorme importance des glucosinolates indoles pour les interactions écologiques entre les plantes et leurs ravageurs a été reconnue. L'architecture génétique qui contrôle la variation des structures des glucosinolates indoles est complexe, et nous avons cloné seulement très récemment le premier de nombreux QTL. Ce projet au cadre de recherche fondamentale vise à comprendre l'architecture génétique qui contrôle la variation structurale et quantitative de cette classe importante de métabolites secondaires chez la plante modèle <i>Arabidopsis</i>. Par ailleurs, ce projet vise à comprendre les facteurs qui déterminent la variabilité dans la biosynthèse des glucosinolates indoles et sa régulation à différents niveaux taxonomiques, par une étude comparative des gènes importants au sein d'<i>Arabidopsis thaliana</i>, mais aussi chez des espèces apparentées plus ou moins proches, et entre les familles de Brassicaceae, de Cleomaceae, de Capparaceae et de Resedaceae, de l'ordre Brassicales.</p>
<b>Partenaires</b>	UPS/ESE IBP
<b>Coordinateur</b>	Juergen Kroymann - UPS/ESE <a href="mailto:juergen.kroymann@u-psud.fr">juergen.kroymann@u-psud.fr</a>

**Aide de l'ANR** 481 k€

**Début et durée** Janvier 2011 - 48 mois

**Référence** ANR-10-GENM-005

**Label pôle**

Titre du projet

**GEMBAL – Génomique multi-race des Bovins Allaitants et Laitiers**

Résumé

Adapter les outils et les objectifs de sélection pour gérer efficacement les productions françaises de lait et viande bovine est un enjeu majeur des prochaines années. La sélection génomique fournit une occasion prodigieuse de réorienter la sélection bovine pour un élevage plus durable. Le projet Gembal a pour objet le développement d'une évaluation génomique multi- raciale afin d'étendre la sélection génomique à toutes les races allaitantes et laitières, y compris celles à faibles effectifs. L'accent sera mis sur les caractères fonctionnels et maternels : facilité de vêlage, fertilité et longévité des vaches tant laitières qu'allaitantes. A l'échelle nationale, ce projet devrait être le socle commun de tous les programmes de sélection génomique, évitant ainsi un foisonnement de petites initiatives inefficaces. Le coeur du projet est la constitution de la base technique nécessaire au développement d'une sélection génomique multi- raciale chez les bovins laitiers et allaitants. L'idée centrale est qu'un échantillon – appelé population d'imputation – sera génotypé avec une puce haute densité dans chaque race, alors que la plupart des individus sera génotypée à plus faible coût sur une puce moyenne densité. La condition requise pour constituer les populations d'imputation est l'utilisation massive d'un nouvel outil moléculaire, une puce haute densité avec 800 000 SNP développée par Illumina au sein d'un consortium incluant l'INRA et l'UNCEIA. La tâche 2 est dédiée à cette partie technique du projet. Les ressources génétiques bovines ainsi générées en tâche 2 serviront de matériel biologique pour les recherches académiques de la tâche 3 visant à caractériser la diversité génétique entre races et l'histoire de chaque population soumise à son propre contexte de dérive et de sélection. La tâche 4 correspond à l'imputation, à savoir la procédure statistique permettant d'inférer les génotypes manquants de la plupart des individus à partir d'une information génotypique complète dans une population d'imputation de taille limitée. La qualité de l'imputation sera étudiée en fonction des tailles effectives de population et d'échantillon d'imputation de chaque race. Des algorithmes de calcul efficace seront développés car l'imputation va induire beaucoup de contraintes informatiques

avec le développement rapide de la sélection génomique. Puis, un modèle de prédiction génomique, utilisant le déséquilibre de liaison entre races, sera développé en tâche 5. Les défis méthodologiques sont le développement d'approches statistiques robustes et puissantes ainsi que de programmes informatiques pour la prédiction dans un contexte multi-racial, en particulier pour des caractères fonctionnels avec des effets génétiques directs et maternels corrélés. Les applications correspondantes seront conduites en tâches 6 et 7, respectivement pour les races laitières et allaitantes. En tâche 6, la disponibilité dans les trois principales races françaises de populations de référence de taille conséquente permettra des comparaisons précises d'évaluation génomique intra- et entre-races, révélant probablement ainsi les conditions nécessaires à une implémentation réussie d'une sélection génomique multi-raciale. Une stratégie alternative consistera à vérifier si les fragments chromosomiques conservés correspondant à des haplotypes favorables détectés dans une grande race ségrègent aussi dans les plus petites races. Alors une évaluation génomique basée sur ces haplotypes pourrait être proposée pour ces dernières races. En tâche 7, la population de référence multi-raciale sera constituée des 2 300 taureaux qui constitueront aussi les populations allaitantes d'imputation. Si un nombre suffisant de QTL est détecté comme commun aux races allaitantes et laitières, une détection de QTL et un calcul d'équation de prédiction génomique entre l'ensemble des populations de référence laitières et allaitantes sera effectué pour les caractères maternels.

**Partenaires**

Inra - GABI  
Institut de l'Elevage  
UNCEIA  
Races de France

**Coordinateur**

Florence PHOCAS - GABI  
[Florence.phocas@jouy.inra.fr](mailto:Florence.phocas@jouy.inra.fr)

**Aide de l'ANR**

662 k€

**Début et durée**

Janvier 2011 - 36 mois

**Référence**

ANR-10-GENM-014

**Label pôle**



**Titre du projet**

**GENAMIBE – Comprendre la pathogénicité de Entamoeba par l'utilisation de la transcriptomique et la phylogénétique comparatives**

**Résumé**

Entamoeba histolytica est une amibe pathogène, parasite protozoaire dépourvu de mitochondries. C'est l'agent étiologique de l'amibiase qui se caractérise chez l'homme par une dysenterie et des abcès hépatiques. Près de 50 millions de cas cliniques et 100 000 décès sont recensés annuellement. La virulence de l'amibe varie selon les isolats: la souche E. histolytica HM1:IMSS (isolée d'un patient souffrant de dysenterie) est virulente alors que la souche E. histolytica Rhaman (isolée d'un patient asymptomatique) est atténuée. Deux espèces très proches, E. dispar et E. moshkovskii sont non-pathogènes; la première est très répandue chez les humains alors que la seconde (isolée du sol) est rarement retrouvée chez des patients. Bien que ces diverses amibes soient très proches dans leur morphologie et phylogénie, leur issue clinique est dramatiquement différente, suggérant l'existence des facteurs indispensables au développement de la maladie et à l'adaptation du parasite à son hôte. Nous voulons caractériser les différences phénotypiques entre espèces/souches de Entamoeba par la comparaison de leurs génomes et de leurs transcriptomes. Nos objectifs majeurs sont : 1) déterminer précisément les différences transcriptomiques et phylogénomiques entre les espèces/souches de Entamoeba ; 2) annoter d'une manière fonctionnelle les familles protéiques et découvrir l'ensemble des gènes responsables des différences phénotypiques entre ces parasites. En première approche, nous utiliserons les nouvelles technologies de séquençage (dites NGS) pour caractériser le profil transcriptionnel des parasites. Nous cartographierons très précisément les transcrits codants et non codants, les petits ARN et ARN antisens. Les ARNs seront extraits d'amibes provenant de trois conditions expérimentales : en culture axénique, en présence de monoxyde d'azote (marqueur de l'inflammation pendant l'amibiase) et en contact avec des explants de colon humain. La comparaison des profils d'expression des petits ARN et des ARN non codants des différentes espèces d'amibes nous donnera des informations précieuses quant au rôle des ARN non-codants dans la régulation du transcriptome. Les

différences phénotypiques entre souches pathogènes et non-pathogènes sont le résultat d'une sélection naturelle des mutations affectant des régions précises du génome (loci). L'identification de ces loci nous apportera des renseignements sur les bases génomiques responsables des différences phénotypiques. En deuxième approche, nous réaliserons donc l'exploration globale du génome à la recherche de ces mutations intéressantes pour l'évolution ; parmi celles-ci, nous identifierons les sites d'évolution adaptative et les mutations ponctuelles qui ont tous deux des impacts fonctionnels. Cette étude nous apportera une vision claire des spécificités génomiques des espèces de *Entamoeba* et contribuera à une meilleure compréhension de la phylogénie des organismes eucaryotes unicellulaires. En troisième approche, nous proposons d'annoter les protéines prédites par les séquences génomiques codantes; en effet, les génomes des amibes sont en général peu annotés. Les données seront hébergées dans une base simple permettant aux chercheurs d'analyser et comparer leurs résultats de transcriptomique. Ce projet représente la première étude complète s'intéressant aux génomes des eucaryotes unicellulaires. Elle se veut compréhensive par : (i) le type de transcrits que nous serons capables de capturer : codants, petits ARN, ARN antisens, ARNs longs non codants ; (ii) la cartographie des sites d'initiation et de terminaison de la transcription, des jonctions introns-exons, des sites de polyadénylation ; (iii) les différentes comparaisons entre espèces, intra-espèce et entre différents environnements. Le but est d'appliquer des méthodes modernes de transcriptomique et de phylogénomique pour caractériser l'ensemble des gènes responsables du caractère pathogène de *E. histolytica*.

**Partenaires** Institut Pasteur

**Coordinateur** NANCY GUILLEN – Institut Pasteur  
[nguillen@pasteur.fr](mailto:nguillen@pasteur.fr)

**Aide de l'ANR** 642 k€

**Début et durée** Janvier 2011 - 36 mois

**Référence** ANR-10-GENM-011

**Label pôle**

**Titre du projet**

**GnpAsso – Un outil générique pour la gestion et l'exploitation des résultats de génétique d'association utilisant les données de génotypage et de phénotypage hauts débits**

**Résumé**

Depuis plusieurs années, de nombreux progrès ont permis l'obtention de données génomiques et génétiques à l'échelle des génomes pour les espèces de grandes cultures. Les développements en bio-informatique visant à la fois l'intégration des données et leur analyse sont de plus en plus nécessaires pour l'exploitation de ces informations. Aux niveaux international et national, plusieurs bases de données ont été élaborées pour répondre aux attentes des chercheurs, malheureusement ces outils sont souvent dédiés à une espèce en particulier ou limités à un certain type de données. La génétique d'association est une méthode d'analyse basée sur l'exploitation du déséquilibre de liaison (LD) dont la puissance permet entre autres la dissection des traits complexes. La pertinence de cette approche en fait un outil largement utilisé pour l'analyse de nombreuses espèces végétales. En effet, la génétique d'association a montré son utilité pour l'analyse du déterminisme génétique des caractères quantitatifs et a permis de découvrir ou confirmer de nombreux gènes impliqués dans la variation des caractères d'intérêt agronomique. Avec l'augmentation des projets de séquençage de génomes entiers de plantes ou de re-séquençage et la diminution rapide des coûts de typage moléculaire, associées à la mise en place de méthodes de génotypage haut-débit, il est maintenant possible, pour un nombre croissant d'espèces, de génotyper de nombreux polymorphismes SNP pour effectuer des études d'association exhaustives à l'échelle du génome entier "Whole genome association mapping", ou dans les régions où des QTL ont été détectés par études de liaison. Les efforts ne se concentrent donc plus uniquement sur les gènes candidats putativement impliqués dans la variation des caractères d'intérêt. Combinée à des approches traditionnelles comme la cartographie de QTL, cette méthode permet de renforcer de nombreuses activités telles que i) la cartographie fine de QTL, ii) l'identification accélérée de nouveaux marqueurs utiles en sélection, iii) la recherche d'allèles d'intérêt dans les collections de ressources génétiques et leur utilisation pour la

création de matériel dédié à la sélection. L'objectif du projet GnpAsso est de mettre à disposition une nouvelle base de données bio-informatique dédiée à la gestion et l'exploitation des résultats issus d'études d'association. Ces résultats seront en lien avec des données de cartographie génétique, des données génomiques et des données issues des plateformes de génotypage haut-débit et des données de phénotypage. La ressource GnpAsso permettra également d'avoir des vues, à la fois synthétiques et détaillées, de toutes les données liées aux résultats de ces études d'association (ressources génétiques, géotypes, SNP, trait, allèles, QTL, annotation de gènes, expression génique, ...) via des liens vers d'autres ressources existantes au sein du système GnpIS. Pour atteindre ces objectifs, nous nous appuyons sur le système d'information GnpIS URGI existant, qui comporte déjà les bases suivantes : GnpMap cartographie, GnpSNP polymorphisme, Siregal ressource génétique et GnpGenome annotation des génomes. L'extension proposée avec GnpAsso offrira aux chercheurs et sélectionneurs une ressource complète qui intégrera les données d'association à l'échelle du gène comme du génome.

**Partenaires**

INRA URGI  
Biogemma  
UMR Génétique végétale du Moulon  
UMR LEG  
UMR DIA PC  
UMR RPB

**Coordinateur**

Delphine Steinbach - INRA URGI  
[delphine.steinbach@versailles.inra.fr](mailto:delphine.steinbach@versailles.inra.fr)

**Aide de l'ANR**

382 k€

**Début et durée**

Février 2011 - 36 mois

**Référence**

ANR-10-GENM-006

**Label pôle**

Céréales Vallée

**Titre du projet**

**Immunit-Ae – Diversité des composants génétiques et des mécanismes de résistance à *Aphanomyces euteiches* chez les légumineuses.**

**Résumé**

*Aphanomyces euteiches*, un oomycète racinaire des légumineuses, est le principal facteur biotique limitant de la production de pois en France et en Europe. Pour accélérer l'amélioration génétique de la résistance à *A. euteiches* chez cette espèce cultivée, nous nous proposons d'utiliser les ressources développées chez la légumineuse modèle *M. truncatula* et d'exploiter la découverte récente d'une région génomique impliquée dans la résistance à *A. euteiches*. Ce locus, identifié à l'aide de 2 sources de résistance différentes est impliqué dans la résistance à plusieurs isolats du parasite. Dans la lignée séquencée A17, ce locus, appelé prAe1, confère une résistance partielle récessive et a été réduit à un intervalle de 135 kb, qui ne contient pas de gènes de résistance classiques (de type NBS-LRR). Chez la lignée DZA45.5, le gène AER1, situé dans la même région génomique confère une résistance complète dominante et est impliqué dans des interactions épistatiques avec d'autres loci pour la résistance vis-à-vis d'autres isolats. Partant de ces résultats, le premier objectif de ce projet est d'identifier le(s) séquence(s) impliquée(s) dans la résistance, située(s) dans le locus prAe1 / AER1. Après obtention des séquences de ce locus chez la lignée DZA45.5 et chez la lignée sensible F83005.5, la comparaison des données acquises avec celle de la lignée A17 permettront d'entreprendre une cartographie fine et d'identifier les séquences impliquées dans la résistance. Ces dernières seront alors validées fonctionnellement dans des approches de complémentation, après obtention de racines transformées avec ces séquences, analyses de mutants disponibles et/ou analyses de luzernes transgéniques. Pour analyser les mécanismes de résistance associées à prAe1, des lignées quasi-isogéniques (NIL) pour ce locus seront analysées après inoculation à l'aide de puces Affymétrie. Le transcriptome de cellules de péricycle en division, mécanisme spécifique présent uniquement chez les NIL résistantes sera également analysé en utilisant les ARN de ces cellules obtenues par microdissection laser. Le deuxième objectif est d'analyser la diversité des loci de résistance de *M. truncatula* impliquées dans la résistance à

différentes souches d'A. euteiches grâce à une approche de génétique d'association sur 192 lignées de M. truncatula, rendue possible par l'accès privilégié aux données génomiques générées par le projet américain « Medicago Hapmap ». Cette approche sera complétée par l'analyse génétique de trois nouvelles populations de RIL connectée, dont les parents font également partie des lignées reséquencées. Les résultats complémentaires obtenus par les deux techniques seront comparés et permettront d'augmenter la résolution des QTL détectés. Le troisième volet de ce projet a pour objectif d'étudier la conservation des gènes, loci et mécanismes de résistances à A. euteiches entre M. truncatula et les principales légumineuses cultivées, en particulier le pois. Ce dernier point sera facilité par l'utilisation des connaissances et des outils génétiques déjà générés (marqueurs, séquences, QTL) chez cette espèce. Dans un premier temps, la syntenie des loci de résistance détectés chez Medicago et le pois sera étudiée grâce à la production de marqueurs ponts entre les deux espèces dans les régions identifiées dans les deux génomes. Des gènes orthologues des gènes candidats identifiés chez M. truncatula seront ensuite recherchés chez plusieurs légumineuses cultivées (notamment le pois) et des marqueurs de ces gènes seront développés. Pour le pois, les meilleurs gènes candidats (localisés dans des régions synténiques de QTL) seront analysés dans une « core-collection » de pois et leur rôle fonctionnel validé dans des mutants de TILLING. Enfin, des analyses cytologiques seront également entreprises pour comparer les mécanismes de résistance observés chez M. truncatula et ceux de plusieurs légumineuses cultivées.

**Partenaires**

UPS - SCSV  
UMR INRA-Agrocampus Ouest-Université Rennes I APBV  
INRA DIAPC Montpellier  
URGV Evry

**Coordinateur**

Christophe jacquet - UPS - SCSV  
[jacquet@scsv.ups-tlse.fr](mailto:jacquet@scsv.ups-tlse.fr)

**Aide de l'ANR**

686 k€

**Début et durée**

Janvier 2011 - 36 mois

**Référence**

ANR-10-GENM-007

**Label pôle**

**Titre du projet**

**PHENOBLE** – Mise au point et valorisation d’outils de phénotypage nouvelle génération pour l’analyse des déterminants génétiques de l’efficacité d’utilisation des engrais azotés chez le blé tendre

**Résumé**

Afin de pouvoir répondre à la demande mondiale en céréales qui augmentera constamment dans les 50 prochaines années, le rendement, notamment du blé tendre, devra être amélioré de 2% chaque année. Ce niveau ne pourra être atteint qu’après avoir optimisé la sélection variétale et les pratiques agronomiques. L’optimisation de la productivité passe par l’identification des facteurs limitants et le développement de stratégies adaptées, en particulier celles liées à l’amélioration génétique. Deux domaines majeurs doivent être travaillés: le génotypage et le phénotypage. Des efforts notables ont été réalisés ces dernières années pour développer et donner accès à la communauté scientifique, à de nouvelles technologies permettant de générer des marqueurs moléculaires peu coûteux, en grand nombre, et même plus récemment de séquencer des génomes entiers. En dépit de cette avancée notable dans le domaine de la biologie moléculaire et de la génétique, peu d’efforts ont encore été réalisés dans le domaine du phénotypage et particulièrement en ce qui concerne le phénotypage au champ qui est nécessaire lorsqu’on s’adresse à des caractères agronomiques complexes. Quelques avancées récentes concentrées dans un champ de recherche nommé « phénomique », ont cependant permis de voir apparaître des technologies : -Permettant de caractériser la performance des plantes, leur rythme de croissance et certaines de leurs fonctions -Permettant de tester des plantes dans des scénarios agronomiques contrôlés à l’aide d’outils haut débit, en particulier, des outils de proximité-détection. Notre capacité à disposer de phénotypes clairs et précis, est d’une importance majeure pour les associer au génotype et identifier les gènes qui aideront à mieux comprendre et maîtriser les caractères complexes. Cependant, la plupart des outils disponibles ont été développés pour des environnements contrôlés comme des serres, relativement éloignés des réalités agronomiques. Dans le cadre du projet PHENOBLE, nous proposons de construire une plateforme dédiée au phénotypage fin et haut débit en

plein champ. Cette plateforme sera d'abord dédiée au blé tendre, et se focalisera sur l'identification des facteurs génétiques impliqués dans les interactions génotype x environnement liés à l'assimilation de l'azote. Les outils évalués seront testés sur un dispositif multi local. PHENOBLE vise à valider et à adapter des outils disponibles actuellement à travers l'évaluation sous différentes contraintes azotées d'un panel de lignées de blés commercialisés en France. Le projet focalisera ses travaux sur l'azote en raison de son importance aussi bien économique, environnementale, qu'agronomique. L'azote a une influence connue sur de nombreux caractères, comme la surface foliaire, la sénescence, l'activité chlorophyllienne ou la teneur en métabolites azotés. Ces caractères paraissent accessibles et mesurables par la nouvelle génération d'outils que nous nous proposons d'évaluer. Les livrables principaux de ce projet seront doubles : générer une liste d'outils et de méthodes validés et utilisables dans des programmes de génétique et de sélection variétale, mais aussi de produire durant le projet une liste de phénotypes héréditaires qui serviront à conduire des études de génétique d'association. Ces études devraient permettre de préciser et d'identifier les mécanismes clefs liés à l'utilisation efficace de l'azote et ses implications sur le rendement. Les outils et les marqueurs moléculaires constitueront tous deux des avancées majeures pour les sélectionneurs et la communauté scientifique.

**Partenaires**

ARVALIS  
BIOGEMMA  
UMR GDEC  
INRA-Bordeaux  
INRA-EMMAH  
UMR DIAPC  
UMR AGIR

**Coordinateur**

Katia BEAUCHENE - ARVALIS  
[k.beauchene@arvalisinstitutduvegetal.fr](mailto:k.beauchene@arvalisinstitutduvegetal.fr)

**Aide de l'ANR**

365 k€

**Début et durée**

Janvier 2011 - 48 mois

**Référence**

ANR-10-GENM-008

**Label pôle**

Céréales Vallée



**Titre du projet**

**PhyloFish** – Analyse phylogénomique des duplications géniques chez les poissons téléostéens: une approche par RNA-Seq

**Résumé**

Les duplications génomiques sont au cœur du processus d'évolution des Vertébrés. Ainsi à titre d'exemple beaucoup de familles protéiques présentes chez l'homme sont le résultat des deux duplications génomiques qui se sont succédé à l'origine de la radiation évolutive des vertébrés. Le groupe des poissons téléostéens est de ce point de vue particulièrement intéressant car suite à une duplication génomique surnuméraire à la base de leur radiation évolutive, de nombreux gènes présents en copie simple dans le génome humain sont en copies doubles chez les poissons téléostéens. Ces gènes « surnuméraires » sont en principe disponibles pour l'évolution de nouvelles fonctions et ces innovations évolutives sont une des causes principales à l'origine de la diversification de la vie sur Terre. L'étude de l'évolution des gènes en relation avec les duplications génomiques reste donc une question majeure de la biologie évolutive moderne car elle permet de mieux comprendre les mécanismes par lesquels les génomes évoluent et permettent ainsi la mise en place d'un développement et d'une physiologie propre à chaque espèce de Vertébrés. Malheureusement, nous n'avons toujours pas à l'heure actuelle une connaissance suffisante des génomes de Vertébrés pour pouvoir répondre complètement à cette question. Ce projet contribuera à apporter des réponses à cette question en utilisant comme modèle d'étude la (les) duplication(s) génomique(s) spécifique(s) des poissons téléostéens. Grâce à l'utilisation des nouvelles générations de séquençage à très haut débit, nous proposons de produire de nouvelles ressources de séquences des gènes transcrits chez différentes espèces de poissons choisies sur la base de leur position taxonomique originale par rapport à l'évolution des poissons téléostéens. Ce jeu de données original et pertinent d'un point de vue évolutif, sera utilisé comme base de travail pour la mise en place d'une analyse automatique à haut débit combinant à la fois des analyses phylogénétiques, des recherches de synténies conservées et des corrélations d'expression géniques tissulaires. Les résultats de ce projet permettront

d'apporter des réponses à l'échelle du génome à des questions telles que : combien de fois ces gènes dupliqués sont perdus ou retenus de façon indépendante dans différents groupes de poissons et est-ce que ces changements groupes spécifiques, soit du répertoire de ces gènes dupliqués, soit de leur profils d'expression, peuvent être des facteurs évolutifs explicatifs de l'extraordinaire diversité des espèces observée au sein du groupe des poissons téléostéens. De plus, comme cette complexité supplémentaire liée aux duplications génomiques a un impact majeur sur la qualité de l'annotation des gènes chez les poissons téléostéens, notre projet proposera une amélioration de cette annotation basée sur les résultats de notre analyse évolutive. Cela permettra d'établir des liens évolutifs gènes à gènes entre différentes espèces de poissons et de Vertébrés permettant ainsi de transférer de façon rigoureuse des informations fonctionnelles sur tel ou tel gène d'intérêt depuis une ou plusieurs espèces modèles vers des espèces d'intérêt socio-économique. Pour mener à bien ces objectifs, ce projet réunit cinq équipes de recherche possédant une expertise reconnue dans différents domaines complémentaires de la génomique comme : les techniques de séquençage nouvelle génération, la transcriptomique, la bio-informatique, et la phylogénomique. Ce projet intègre aussi une collaboration formelle avec une équipe américaine dont l'expertise dans le domaine de la recherche sur les duplications des génomes de poissons, est internationalement reconnue. Enfin, ce projet fournira aussi une masse d'informations actuellement sans précédent sur les répertoires des gènes transcrits chez plusieurs espèces de poissons d'intérêt socio-économique.

**Partenaires**

INRA SCRIBE  
UOregon  
MGX  
INRA SIGENAE  
U Aix Mars 1

**Coordinateur**

Julien BOBE - INRA SCRIBE  
[Julien.Bobe@rennes.inra.fr](mailto:Julien.Bobe@rennes.inra.fr)

**Aide de l'ANR**

481 k€

**Début et durée**

Février 2011 - 48 mois

**Référence**

ANR-10-GENM-017

**Label pôle**

Titre du projet

**RepliColScope – Séquençage et génomique comparative des plasmides de Escherichia coli : apport à la compréhension du succès évolutif des  $\beta$ -lactamases à spectre étendu**

Résumé

La résistance aux antibiotiques est devenue un enjeu mondial de santé publique. Une illustration en est l'explosion de la résistance acquise aux céphalosporines de troisième génération (G3G), couramment utilisées pour traiter les infections à Gram négatif. Cette résistance est principalement due à la production de  $\beta$ -lactamases à spectre étendu (BLSE), enzymes qui hydrolysent les  $\beta$ -lactamines à l'exception des carbapénèmes et des céphamycines. Les gènes codant les BLSE, essentiellement portés par des plasmides, diffusent actuellement par plusieurs mécanismes non exclusifs : (i) transferts horizontaux de gènes entre plasmides, (ii) transferts horizontaux de plasmides parmi des clones bactériens non reliés et (iii) diffusion de clones épidémiques. Pendant les années 1990, les principales BLSE étaient dérivées des enzymes de type TEM et SHV. Depuis les années 2000, les enzymes de type CTX-M ont mondialement diffusé et sont devenues majoritaires. *Escherichia coli*, un des principaux pathogènes opportunistes de l'homme et des animaux qui était sensible aux  $\beta$ -lactamines, est devenu l'hôte principal des BLSE. Les *E. coli* producteurs de BLSE sont retrouvés de plus en plus fréquemment chez l'homme, dans les infections communautaires, hospitalières et en portage, et chez les animaux d'élevage, mais aussi chez les animaux sauvages. L'émergence soudaine de souches de *E. coli* productrices de BLSE, et notamment de CTX-M, ne peut être expliquée uniquement par la pression de sélection exercée par l'utilisation des G3G. D'autres facteurs écologiques et/ou moléculaires pourraient influencer cette évolution. L'objectif du projet RepliColScope est de comprendre le succès évolutif des BLSE dans l'espèce *E. coli*, et en particulier l'émergence des enzymes de type CTX-M par rapport aux enzymes de type SHV et TEM. Pour atteindre nos objectifs, nous proposons : (i) de séquencer 80 plasmides de souches de *E. coli* codant des BLSE des 3 principaux types (70 plasmides de souches humaines et 10 de souches animales) ainsi que des plasmides de souches de *E. coli* présents avant l'utilisation des G3G (17 plasmides de souches la collection ECOR et 3 plasmides de

souches de E. coli de la collection Murray datant de l'ère pré-antibiotique), (ii) d'effectuer des comparaisons génomiques des données générées et des données déjà disponibles et (iii) de tester in vitro des hypothèses générées à partir des données obtenues in silico. Les plasmides seront séquencés en utilisant une stratégie incluant du pyroséquençage à haute densité (A 454 Titanium). Des procédures d'annotations automatiques structurelles et fonctionnelles seront développées et les résultats intégrés dans une base de données spécifique permettant de faire des annotations expertes et d'utiliser les données. Des analyses de génomique comparative seront effectuées pour reconstruire l'histoire évolutive des gènes plasmidiques en identifiant ceux ayant le contenu génétique le plus similaire. Ces données seront également analysées à la lumière du fond génétique chromosomique, représenté par l'histoire phylogénétique des souches et par les CRISPRs. Nous rechercherons in vitro des communications possibles entre des modules portés par des plasmides et ceux trouvés sur des chromosomes de E. coli, qui peuvent jouer un rôle clé dans la stabilisation de ces éléments extra chromosomiques. Par la quantité et la qualité des informations obtenues, ce projet devrait apporter de nouveaux éclairages tant fondamentaux que médicaux sur l'évolution des plasmides et sur la résistance aux antibiotiques et permettre de trouver de nouvelles stratégies pour prévenir la dissémination des plasmides de résistance. RepliColScope associe des équipes qui ont une expertise dans les domaines de la biologie évolutive de E. coli (U722), de l'épidémiologie des BLSE chez les souches animales et humaines (ER8 et UR1282), de la plasticité du génome bactérien (UPGB) et le séquençage des génomes et leur annotation (Genoscope).

**Partenaires**

U722  
Genoscope  
ER8  
UR1282  
UPGB

**Coordinateur**

Catherine Branger - U722  
[catherine.branger@lmr.aphp.fr](mailto:catherine.branger@lmr.aphp.fr)

**Aide de l'ANR**

380 k€

**Début et durée**

Janvier 2011 - 36 mois

**Référence**

ANR-10-GENM-012

**Label pôle**

**Titre du projet**

**SOLAR** – Développer et tester de nouveaux outils biotechnologiques pour accroître la teneur en huile de la graine : une clé pour la production de chaînes hydrocarbonées chez les oléagineux.

**Résumé**

Objectifs du projet : Les huiles végétales constituent une denrée alimentaire précieuse dont la consommation augmente pour cause de croissance démographique mondiale soutenue. De plus, les acides gras végétaux stockés dans les graines sous forme de triglycérides (TAGs) ont une structure comparable à celle des chaînes carbonées des hydrocarbures et représentent une alternative potentiellement compétitive aux produits dérivés des hydrocarbures pour la chimie verte (fabrication de produits industriels tels que les détergents, peintures, plastiques et autres lubrifiants). La demande croissante d'huiles végétales pour des applications tant industrielles que nutritionnelles met en lumière un besoin urgent de développer de nouvelles méthodologies susceptibles d'accroître encore le contenu en huile des graines quand les programmes d'amélioration variétale classique s'essouffent, limités par la variabilité naturelle connue et/ou utilisable. L'obtention d'espèces oléagineuses à fort rendement en huile sera facilitée par l'élucidation complète des mécanismes contrôlant la production des acides gras et leur acylation. La voie de biosynthèse des TAGs se compose de deux parties : un bloc A de réactions enzymatiques plastidiales assure la synthèse d'acides gras, un bloc B de réactions microsomales assure l'assemblage des TAGs. Jusqu'ici, l'immense majorité des approches biotechnologiques destinées à accroître la teneur en huile des graines s'est focalisée sur le bloc B. Ces approches reposaient sur l'idée aujourd'hui contestée qu'un petit nombre de réactions du bloc B constituait le principal goulot d'étranglement du réseau de biosynthèse des huiles végétales. Récemment, des analyses de contrôle métabolique ont révélé que les étapes limitant la production des TAGs sont en fait nombreuses, disséminées au sein des blocs A et B. Il apparaît donc urgent d'étudier la régulation du bloc A pour trouver des outils biotechnologiques originaux capables de stimuler de manière efficace la synthèse des acides gras. Les données disponibles indiquent que cette synthèse est fortement régulée au niveau transcriptionnel et qu'une co-

activation d'une majorité des gènes codant les enzymes du bloc A est requise pour la stimuler efficacement. Nous avons par conséquent décidé de nous focaliser sur la régulation transcriptionnelle du bloc A dans ce projet. Plan de travail : Un premier régulateur transcriptionnel du bloc A, WRINKLED1, a été isolé et étudié en détail. Néanmoins, des données récentes obtenues dans notre équipe indiquent qu'il n'agit pas seul pour induire la lipogenèse. - L'objectif premier de ce projet est d'identifier de nouveaux facteurs de transcription impliqués dans la régulation de la biosynthèse des acides gras chez l'espèce modèle Arabidopsis. - Le second objectif du projet est d'étudier un certain nombre des candidats ainsi isolés pour aboutir à une connaissance approfondie du mécanisme contrôlant l'expression des gènes impliqués dans de la synthèse des acides gras. - Le troisième objectif est d'exploiter les connaissances acquises pour mettre au point et tester des stratégies biotechnologiques efficaces pour stimuler la production d'acides gras dans la graine d'Arabidopsis. Résultats attendus : - ce projet nous apportera d'abord des connaissances nouvelles relatives à l'activation transcriptionnelle de la voie de biosynthèse des acides gras chez les plantes. - ce projet permettra de tester des approches biotechnologiques susceptibles d'être utilisées pour accroître la teneur en huile des graines.

**Partenaires**

INRA IJPB SDQ  
INRA IJPB PLHS

**Coordinateur**

Sébastien BAUD - INRA IJPB SDQ  
[sbaud@versailles.inra.fr](mailto:sbaud@versailles.inra.fr)

**Aide de l'ANR**

320 k€

**Début et durée**

Février 2011 - 48 mois

**Référence**

ANR-10-GENM-009

**Label pôle**

<b>Titre du projet</b>	<b>SUS_FLORA – Contribution du microbiote intestinal à l’homéostasie du système immunitaire chez le porc: approches génétiques et génomiques</b>
<b>Résumé</b>	<p>Une meilleure compréhension des mécanismes de résistance/susceptibilité aux agents pathogènes devient un axe de recherche prioritaire dans une perspective européenne de rationalisation des élevages, de sécurité alimentaire et sanitaire et de diminution des additifs et traitements médicamenteux. En lien avec des sélectionneurs dans l’espèce porcine, un programme d’étude du contrôle génétique de la réponse immunitaire innée et adaptative chez le porc Large White a été démarré avec le projet IMMOPIG (ANR 2007-2009). Nos résultats montrent que la majorité des paramètres mesurés présente une héritabilité moyenne à forte, ce qui suggère une part génétique importante de la réponse immunitaire, en accord avec des travaux publiés par d’autres équipes. Les animaux étudiés sont en cours de génotypage avec la puce SNP porcine 60K iSelect pour des études d’association. Une sélection divergente pour un index de quatre paramètres héréditaires est engagée et la première génération sera disponible en 2010. Les bactéries commensales sont présentes sur toutes les surfaces corporelles exposées à l’environnement et l’intestin est considéré comme l’un des meilleurs exemples de cohabitation entre hôte et bactéries (microbiote). Le microbiote intestinal se développe comme un paramètre spécifique de l’hôte et se stabilise au cours de la vie. L’étude des interactions entre le système immunitaire de l’hôte et le microbiote intestinal, avec des hypothèses de coévolution de l’hôte et des bactéries du microbiote est un domaine en plein essor. Dans nos projets sur la caractérisation de l’immunocompétence des porcs, il nous semble novateur et prospectif d’envisager le microbiote intestinal comme un phénotype à part entière à intégrer dans notre analyse des paramètres de la réponse immunitaire. Le but majeur du projet consistera à caractériser le microbiote intestinal d’animaux pour lesquels des paramètres classiques de la réponse immunitaire innée et adaptative sont mesurés. Le projet inclut, d’une part une étude globale des animaux Sus_Flora avec du phénotypage et des analyses génétiques et, d’autre part, une étude des interactions locales entre le microbiote et l’épithélium intestinal pour un sous-groupe de ces animaux bien caractérisés pour leurs phénotypes et génotypes. Quatre-vingt-dix familles seront spécifiquement</p>

produites pour ce projet et quatre descendants par famille seront étudiés. Nous caractériserons le microbiote des porcelets à 60 jours et celui de leurs mères après la mise-bas. Nous étudierons les variations du microbiote des porcelets entre la naissance et la période de croissance pour un sous-groupe de 30 animaux. Nous calculerons l'héritabilité du caractère microbiote intestinal et les corrélations avec les paramètres de la réponse immunitaire. Les 360 porcelets et leurs 90 pères seront génotypés avec la puce SNP porcine 60K iSelect, afin d'accumuler des premières données pour entamer des études d'associations génétiques. L'haplotype des animaux au locus CMH sera reconstruit pour analyser d'éventuelles associations entre la composition du microbiote et le CMH dont le polymorphisme guide la capacité de présentation antigénique. Nous inclurons également la recherche du polymorphisme des récepteurs Toll-like. L'étude des interactions entre le microbiote intestinal et l'épithélium sera menée sur le sous-groupe de porcelets suivis pour les variations du microbiote et nous aurons la possibilité d'inclure des animaux en sélection divergente. L'expression différentielle des gènes de l'hôte dans cinq sites du tube digestif (duodénum, jéjunum, iléon, colon et plaques de Peyer) sera analysée par une approche transcriptome. Trois fonctions clés seront étudiées: l'histocompatibilité, avec l'expression des gènes de classe I non classiques (SLA-Ib, MIC-2, CD1, MR1), la reconnaissance de motifs à la surface des cellules (TLRs, NODs) et l'expansion des IgA dans les muqueuses. Les compartiments cellulaires T périphériques et intestinaux seront comparés.

**Partenaires**

INRA-GABI  
CEA-DSV/IRCM/SREIT/LREG  
INRA - MICALIS  
INRA-LPT  
INRA-VIM  
INRA-GEPA  
BIOPORC  
INRA-IASP

**Coordinateur**

Claire Rogel-Gaillard - INRA-GABI  
[claire.rogel-gaillard@jouy.inra.fr](mailto:claire.rogel-gaillard@jouy.inra.fr)

**Aide de l'ANR**

766 k€

**Début et durée**

Mars 2011 - 36 mois

**Référence**

ANR-10-GENM-016

**Label pôle**



# Programme « Génomique, Biotechnologies végétales »

Edition 2010

Titre du projet	<b>UtOpIGe – Vers une Utilisation Optimale de l'Information Génomique dans les schémas pyramidaux</b>
<b>Résumé</b>	<p>Ce projet vise à fournir les informations nécessaires à la mise en place d'une sélection génomique dans les schémas de sélection pyramidaux caractérisés par : un étage de sélection en race pure pour produire des terminaux croisés; un milieu de contrôle en sélection de haut niveau sanitaire en comparaison du milieu dans lequel la production a lieu. Une approche expérimentale sera entreprise dans les deux espèces porc et poulet. Des populations de référence, de races pures et croisées, élevées dans des milieux différents, seront créées et contrôlées pour un grand nombre de caractères. Les génotypages seront réalisés avec une puce SNP haut débit (actuellement 64K disponibles). Différentes méthodes d'évaluation génomique seront comparées quant à leur capacités à prédire précisément la valeur génétique des reproducteurs dans ces différentes situations (type génétique, environnement). L'ensemble des données obtenues sera également utilisé pour réaliser l'optimisation des schémas de sélection. Les résultats seront présentés à l'ensemble de la communauté des sélectionneurs des espèces concernées.</p>
<b>Partenaires</b>	GARen GABI LGC SAGA UE TP IFIP SYSAAF Novogen
<b>Coordinateur</b>	Pascale LE ROY - GARen <a href="mailto:pascale.leroy@rennes.inra.fr">pascale.leroy@rennes.inra.fr</a>
<b>Aide de l'ANR</b>	968 k€
<b>Début et durée</b>	Janvier 2011 - 48 mois
<b>Référence</b>	ANR-10-GENM-015
<b>Label pôle</b>	VALORIAL - l'Aliment de demain

Titre du projet

**WALLARRAY-II – Biopuces à polysaccharides pariétaux en génomique fonctionnelle**

Résumé

Le projet WallArray a été soumis à l'appel d'offres ANR-Génomique en avril 2008. Dans ce cadre, le projet a été financé pour une étude de faisabilité sur un an. Les résultats acquis confortent les différents partenaires du projet dans les objectifs initialement proposés. Le nouveau projet s'appuie sur les points forts de l'étude de faisabilité, et développe de nouvelles solutions techniques et de nouvelles propositions dans la réalisation de ces objectifs. La soumission du projet WallArray 2 à l'appel d'offres Génomique, Biotechnologies Végétales s'inscrit donc dans la continuité du précédent projet. Il rassemble les mêmes équipes de recherche de plusieurs instituts publics CEA, CNRS et INRA, des universités de Grenoble et Toulouse et de la société privée Genoptics. Ces équipes développent des recherches dans les domaines de la physico-chimie, la chimie, la biologie structurale et la biologie végétale. Les parois cellulaires sont des structures composites naturelles, renfermant essentiellement des polymères de fortes masses moléculaires (polysaccharides et lignines) et des protéines interagissant avec ces polymères. Les parois cellulaires sont des structures dynamiques impliquées dans la division, l'expansion, la différenciation et l'adhérence cellulaire contrôlant ainsi la morphologie de la plante. Les parois sont également la source de signaux oligosaccharidiques: leur reconnaissance moléculaire contrôle des processus liés au développement et même l'issue d'interactions entre plantes et microorganismes. Ces dernières années, des progrès décisifs en génomique, protéomique, glycomique et bio-informatique ont considérablement augmenté nos connaissances des gènes et des protéines liés à la paroi, en particulier les gènes et les protéines interagissant, ou encore modifiant les polysaccharides pariétaux. Toutefois, très peu d'entre eux ont une fonction expérimentalement démontrée. Ainsi, toutes les données structurales ou fonctionnelles concernant l'interaction du glycome pariétal (défini comme l'ensemble de structures glucidiques de la paroi) et du protéome pariétal (l'ensemble des protéines) sont d'un grand intérêt afin de comprendre la dynamique et la signalisation pariétale. Le but du projet est de développer des biopuces à polysaccharides

pariétaux comme outils innovants pour sélectionner des protéines capables d'interactions et évaluer leur spécificité. Les données fonctionnelles et structurales, qui seront acquises, compléteront les banques de données déjà en place et ouvertes à la communauté scientifique (glyco3d). Ces données serviront de support expérimental au développement de logiciels pour la prédiction de structure de protéines et de polysaccharides, voire à l'élaboration de logiciels pour la prédiction d'interactions protéines/polysaccharides pariétaux. Le projet rassemble des nouvelles technologies éprouvées et déjà mises en œuvre dans notre étude de faisabilité: de la constitution de collections de polysaccharides pariétaux et leur fixation à une surface, à la production de protéines pariétales et l'analyse de leurs interactions avec des polysaccharides par imagerie en résonance plasmonique de surface. La compréhension des interactions protéines/polysaccharides au niveau des surfaces cellulaires complète les données de protéomique et de glycomique, et représente une étape essentielle en post-génomique. Par ailleurs, les parois des plantes sont les constituants les plus abondants de la biomasse végétale. Les matériaux de la paroi (cellulose, pectines) ont été largement exploités par les industries alimentaires et non-alimentaires. Avec la reconnaissance actuelle des molécules glucidiques comme vecteurs de propriétés biologiques, les sociétés de glyco-biotechnologie sont apparues. Les biopuces à polysaccharides seront un outil utile pour l'élaboration de nouveaux matériaux, la découverte ou la modification d'enzymes intervenant sur les polysaccharides pariétaux, l'amélioration des espèces végétales.

**Partenaires** UPS - SCSV  
INRA  
INAC/SPrAM  
GENOPTICS  
CNRS

**Coordinateur** Hervé CANUT - UPS - SCSV  
[canut@scsv.ups-tlse.fr](mailto:canut@scsv.ups-tlse.fr)

**Aide de l'ANR** 568 k€

**Début et durée** Janvier 2011 - 36 mois

**Référence** ANR-10-GENM-010

**Label pôle** AGRIMIP INNOVATION

Titre du projet

**XANTHOMIX** – Etude comparative des génomes et des transcriptomes de *Xanthomonas* phytopathogènes

Résumé

L'objectif du projet XANTHOMIX est de fournir des données de génomique et de transcriptomique pour une meilleure connaissance des mécanismes d'infection, de la spécialisation d'hôte et la spécificité tissulaire, des mécanismes évolutifs et de l'épidémiologie de cinq espèces de *Xanthomonas*. Les Xanthomonadacées sont des bactéries à GRAM négatif responsables de plus de 400 maladies sur des plantes monocotylédones comme dicotylédones, plusieurs d'entre elles d'importance économique majeure (par exemple orge, seigle, blé, riz, haricot, chou, manioc, agrumes et coton). Les Xanthomonadacées possèdent un haut degré de spécialisation tant pour leurs hôtes que pour les tissus qu'ils infectent. Le genre *Xanthomonas* est donc un très bon modèle pour étudier les aspects moléculaires, évolutifs et épidémiologiques de l'interaction hôte-bactérie. Nous souhaitons séquencer 22 souches de *Xanthomonas* qui sont pour nos laboratoires des modèles d'étude de toutes les phases du processus infectieux. Nous avons aussi choisi des souches de *Xanthomonas* qui sont d'importance majeure en tant qu'organismes de quarantaine car présentant une menace importante pour l'agriculture et la sécurité alimentaire. Toutes les souches choisies appartiennent à des lignées pour lesquelles aucune donnée génomique ou transcriptomique n'est disponible pour l'instant. Ce choix comprend notamment deux espèces et 14 pathovars jamais séquencés. La proximité évolutive des souches sélectionnées doit cependant permettre d'identifier des déterminants potentiels pour la spécificité tissulaire et/ou la spécialisation d'hôte par le biais d'analyses comparatives. La technologie de séquençage « Next-Generation » sera utilisée. Plus spécifiquement le séquençage d'ADN génomique par la technique 454 sera combiné au séquençage des ADNc par la technique Illumina. Ceci doit permettre de produire d'une part des séquences génomiques ébauches de haute qualité et d'autre part d'étudier le transcriptome sous différentes conditions. Les ADNc seront générés de bactéries cultivées en conditions standard et en conditions correspondant à des étapes importantes de l'infection. La comparaison de ces données doit permettre d'émettre des

hypothèses concernant les régulons qui pourront être rapidement testées à l'aide des séquences génomiques disponibles. La synthèse des ADNc sera effectuée à l'aide de protocoles spécifiques permettant l'identification de petits ARN non codants et d'ARN antisens en plus des ARNm. La connaissance des sites de démarrage de la transcription permettra de caractériser d'éventuels « Riboswitches » régulateurs agissant en cis et les codons d'initiation de traduction. Ces informations permettront l'obtention d'une annotation de très haute qualité des génomes concernés ainsi que de ceux déjà disponibles. L'annotation fonctionnelle experte des gènes impliqués dans l'infection sera une ressource très précieuse pour le reste de la communauté scientifique internationale. Les hypothèses issues des données de génomique et de transcriptomique concernant le rôle de nouveaux déterminants de l'infection seront complétées par des analyses fonctionnelles. Les nouvelles ressources génomiques ainsi produites seront également utilisées pour le développement d'outils de typage moléculaire à haute résolution applicables à l'épidémiologie-surveillance de ces agents pathogènes majeurs. Le projet implique quatre partenaires français d'instituts hautement actifs dans la recherche fondamentale et appliquée. Les chercheurs de ce projet, appartenant à l'IRD, à l'INRA et au CIRAD sont des experts reconnus dans leurs domaines (microbiologie, pathologie végétale, phylogénie, génétique, génomique, épidémiologie) avec une forte expertise en génétique, génomique et épidémiologie des bactéries phytopathogènes en général et du genre *Xanthomonas* en particulier.

**Partenaires**

RPB  
LIPM  
PaVé  
PVBMT

**Coordinateur**

Ralf KOEBNIK – RPB  
[Ralf.Koebnik@ird.fr](mailto:Ralf.Koebnik@ird.fr)

**Aide de l'ANR**

887 k€

**Début et durée**

Janvier 2011 - 30 mois

**Référence**

ANR-10-GENM-013

**Label pôle**