

**UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE I**

**HABILITATION A DIRIGER DES RECHERCHES**

**Guillaume BESNARD**

**Génétique et évolution des plantes  
en milieux méditerranéen et tropical**



**Soutenu publiquement le 13 Janvier 2009 à l'Université de Lille I**

**Jury :**

Joël CUGUEN

Jacques DAVID

Nicole GALLAND

Sylvain JEANDROZ

Pierre SAUMITOU-LAPRADE

Professeur, Université de Lille I

Professeur, SupAgro Montpellier

Professeure, Université de Lausanne

Maître de conférences, ENESAD Dijon

Chargé de recherches, CNRS, Université de Lille I

Directeur

Rapporteur

Rapporteur

Examineur

Rapporteur

# SOMMAIRE

<b>1. Présentation de mes activités de recherche</b>	<b>3</b>
<b>2. Caractérisation des ressources génétiques de l'Olivier</b>	<b>4</b>
2.1. Introduction	4
2.2. Clarification de la taxonomie et de la diversification du genre <i>Olea</i> en Afrique	6
2.3. Phylogéographie du complexe de l'Olivier ( <i>Olea europaea</i> )	10
2.4. Domestication multilocale de l'Olivier méditerranéen	15
2.5. Apport de l'outil moléculaire dans la conservation des ressources génétiques cultivées <i>ex situ</i>	16
2.6. Conservation des ressources génétiques sauvages <i>in situ</i> : le cas de l'Olivier de Laperrine ( <i>O. europaea</i> subsp. <i>laperrinei</i> ) du Sahara Central	17
2.7. Conclusions	19
<b>3. Bases moléculaires de l'évolution de la photosynthèse C<sub>4</sub> chez les Poales</b>	<b>21</b>
3.1. La photosynthèse C <sub>4</sub> et les facteurs environnementaux favorisant son évolution	21
3.2. Taxonomie, datation moléculaire et origines multiples de la photosynthèse C <sub>4</sub> chez les Poales	25
3.3. Evidence de sélection positive parallèle sur <i>rbcL</i> suite à l'acquisition du système C <sub>4</sub>	27
3.4. Caractérisation de la famille multigénique codant les PEPC : évidence de convergences moléculaires	28
3.5. La photosynthèse C <sub>4</sub> et les changements globaux : quelles conséquences pour l'amélioration des plantes ?	33
<b>4. Perspectives de recherche</b>	<b>35</b>
4.1. Origines et domestication d'une espèce ligneuse méditerranéenne : l'olivier comme cas d'étude	36
4.2. Génétique de la conservation des sous-espèces menacées du complexe <i>O. europaea</i>	40
4.3. Transfert et valorisation	40
<b>Conclusion</b>	<b>41</b>
<b>Références</b>	<b>42</b>

## 1. Présentation de mes activités de recherche

Depuis septembre 1994, j'ai été impliqué dans divers projets de recherche ayant pour principaux objectifs l'utilisation et la gestion des ressources génétiques de plantes cultivées (Olivier, Platane, Tournesol) ainsi que des projets plus fondamentaux visant en particulier à mieux comprendre l'évolution de caractères adaptatifs. Durant mon cursus, j'ai travaillé dans plusieurs laboratoires en Europe ainsi qu'en milieu tropical. J'ai effectué tout d'abord un DEA (Bases de la Production Végétale) et une thèse (en Biologie de l'Evolution et Ecologie) à l'INRA de Montpellier sous la direction du Dr André BERVILLE, suivis d'un ATER à l'Université de la Réunion et un post-doc à l'Université de Nancy. Par la suite, j'ai passé cinq ans à l'université de Lausanne (Suisse) en tant que premier assistant, ce qui m'a permis notamment d'encadrer plusieurs étudiants en thèse et en master. J'ai enfin obtenu une bourse Marie Curie (de « Research Associate ») pour travailler à l'Imperial College (Silwood Park, Londres, UK) avec Vincent SAVOLAINEN durant deux ans à partir de septembre 2008. Durant ces différentes activités de recherche, je me suis particulièrement spécialisé dans le développement de divers outils de génétique moléculaire, et des problématiques de recherche très variées ont ainsi été abordées :

- La **taxonomie moléculaire** et l'étude de la diversification de différentes familles de plantes, en particulier les Poacées, les Cypéracées et les Oléacées ;
- L'histoire évolutive de plusieurs taxons (i.e. *Biscutella laevigata*, *Heracleum mantegazzianum*, *Olea europaea*, *Picea abies*, *Prunus serotina* et *Thlaspi caerulescens*) à l'aide d'une approche de **phylogéographie**. A cette occasion, les origines d'espèces envahissantes et de la polyploïdie ont été étudiées ;
- La **conservation** de populations menacées d'extinction (Olivier de Laperrine) ;
- La caractérisation de formes sauvages et cultivées d'Olivier et de Platane pour comprendre leur **domestication/sélection** et optimiser la **gestion de leurs ressources génétiques** ;
- L'étude des **bases moléculaires** de l'évolution de **caractères adaptatifs**, en particulier la photosynthèse C<sub>4</sub> et l'hyperaccumulation des métaux lourds ;
- La construction de **cartes génétiques** (*Olea europaea*, *Picea abies*) pour la **détection de QTLs** (*Helianthus* spp.) ou de zones affectées par des effets de sélection prézygotiques (*P. abies*).

Dans ce document, je vais me focaliser sur mes travaux les plus significatifs. Dans une première partie, je ferai le point sur l'étude des ressources génétiques et de la domestication d'une espèce ligneuse typiquement méditerranéenne (olivier), puis dans une seconde partie, je présenterai mes travaux sur l'évolution de la photosynthèse C<sub>4</sub> chez les Poales et tout particulièrement dans la famille des Poacées. A travers ces deux grandes thématiques, je vais notamment discuter de l'importance des recherches effectuées et montrer comment j'ai pu acquérir différentes compétences de chercheur tant au niveau technique que dans la gestion de ces activités. Mes différentes expériences d'encadrement d'étudiants et de conduite de nouvelles études en collaboration avec des laboratoires étrangers seront également mentionnées. Pour finir, mes projets de recherche pour les années à venir seront présentés.

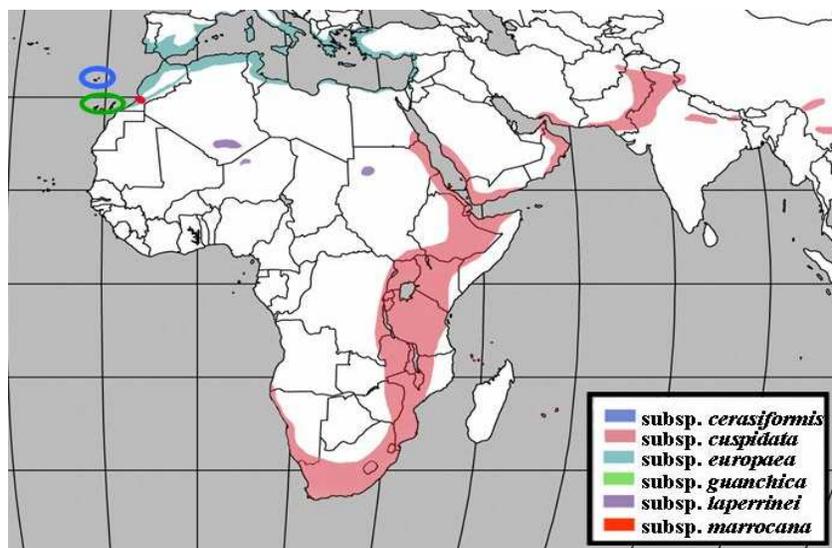
## 2. Caractérisation des ressources génétiques de l'Olivier

Mes études sur les ressources génétiques de l'olivier ont débuté durant ma thèse, il y a environ 12 ans, en Avril 1996. Ces premiers travaux s'inscrivaient dans le projet européen FAIR CT95-0689, intitulé « Evaluation of olive genetic diversity and identification of molecular markers related to the aspects of food products, oil and olives, their mapping for marker assisted breeding ». Durant ce travail, j'ai notamment collaboré avec Bouchaïb KHADARI et Philippe BARADAT. J'ai soutenu ma thèse en Octobre 1999. Mes activités sur cette problématique ont repris à l'occasion de collaborations avec Pablo VARGAS et Rafael RUBIO DE CASAS au Jardin Botanique de Madrid en Février 2002, puis surtout en Mars 2004 (dans le cadre de projets BIODIBERIA A82 et SYNTHESYS ES-TAF-244). Par ailleurs, Djamel BAALI-CHERIF de l'université d'Alger puis Fabien ANTHLEME de l'IRD de Montpellier m'ont contacté indépendamment en 2004, pour débiter un travail sur la conservation de l'olivier de Laperrine. En parallèle, plusieurs projets d'initiation à la recherche ont également été proposés à des étudiants de Master pour étudier des populations envahissantes d'olivier en Australie et à Hawaii (Philippe HENRY et Lukas WILLE) et mesurer l'importance des flux de gènes dans des populations isolées d'olivier de Laperrine à l'aide de tests de paternité (Sandra BETTINELLI-RICARDI et David PARIETTI). Récemment, j'ai rejoint le groupe de Vincent SAVOLAINEN (Imperial College London) pour approfondir mes travaux sur la phylogéographie de l'olivier méditerranéen.

Au cours de ces dix années d'investigations sur l'olivier, j'ai acquis des compétences en génétique moléculaire, phylogénie et génétique des populations. En parallèle, j'ai également approfondi mes connaissances sur cette espèce cultivée et ma vision sur la caractérisation de ses ressources génétiques a aussi fortement évolué au cours du temps. Dans cette synthèse, je commencerais par une brève introduction sur l'olivier comme espèce d'étude, puis l'ensemble des travaux sur ce taxon auxquels j'ai participé ces dernières années sera présenté.

### 2.1. Introduction

L'olivier (*Olea europaea* L. subsp. *europaea*, Oléacées) est l'arbre emblématique de la zone méditerranéenne. Ce taxon a une longévité remarquable et des individus supposés plurimillénaires sont répertoriés un peu partout dans le Bassin Méditerranéen. De plus, il est capable de se régénérer par des rejets de souche, une caractéristique biologique assez fréquente en milieu méditerranéen (ex. *Quercus ilex* et *Q. coccifera*) permettant une rapide recolonisation après un feu de forêt. Comme de nombreux arbres, l'olivier est une espèce allogame avec une floraison abondante et la dispersion de son pollen est essentiellement assurée par le vent. Son système d'autoincompatibilité serait de type gamétophytique (Cuevas & Polito 1997 ; C. Breton, pers. com.), mais il existe aussi des cas de stérilités mâles (Villemur *et al.* 1984). La dissémination de ses fruits (drupes) est essentiellement assurée par les oiseaux (Spennemann & Allen 2000) mais des petits mammifères (rongeurs) y participent probablement. L'homme a évidemment aussi fortement contribué à sa dispersion. L'olivier appartient à un complexe de six sous-espèces (supposées inter-fertiles) présent depuis l'Afrique du Sud jusqu'en Chine, en Méditerranée, en Macaronésie et enfin dans les montagnes sahariennes (Fig. 1 ; Green & Wickens 1989). Des membres de ce complexe sont également devenus envahissants en Australie et dans plusieurs archipels de l'Océan Pacifique (Spennemann & Allen 2000 ; Green 2002).



**Fig. 1.** Distribution naturelle du complexe *Olea europaea* dans le monde (d'après Rubio de Casas *et al.* 2006)

L'olivier fut l'un des premiers arbres cultivés avec le figuier et le dattier (Zohary & Spiegel-Roy 1975). Il est clairement démontré que son exploitation a fortement augmenté à l'Holocène, il y a environ 8000 ans (Zohary & Hopf 2000 ; Terral *et al.* 2004a). Sa domestication, par le biais de la multiplication végétative des individus les plus intéressants d'un point de vue agronomique et l'établissement de vergers constitués des premiers clones, aurait débuté il y a environ 6000 ans principalement au Levant, à l'Est de la Méditerranée (Zohary & Spiegel-Roy 1975 ; Galili *et al.* 1997). Par la suite, l'olivier a été fortement disséminé par l'homme dans le Bassin Méditerranéen, au gré des principales civilisations telles que les Phéniciens, les Romains ou encore les Arabes (Terral 1997 ; Zohary & Hopf 2000). Plus récemment, il y a environ deux siècles, sa culture a été introduite dans des pays du nouveau monde (e.g. Australie, Californie). Actuellement, l'olivier est une importante ressource de lipides (environ 3 millions de tonnes/an) et il y a eu une forte augmentation de sa production (d'environ 40%) durant la dernière décennie en réponse à une demande grandissante (CNUCED 2006).

Si l'histoire récente de l'olivier est clairement établie, il reste de nombreuses incertitudes sur ses origines précises. Plusieurs hypothèses ont notamment été proposées telles qu'une introduction de formes cultivées depuis l'Afrique de l'Est ou du Moyen Orient (Chevalier 1948 ; Turrill 1951), ou au contraire une origine autochtone en Méditerranée (Zohary & Spiegel-Roy 1975). De plus, les connaissances sur les ressources génétiques sauvages de l'olivier (complexe *O. europaea* ; Green & Wickens 1989) restaient très limitées au début de mon travail de thèse. En particulier, les relations phylogénétiques et les compatibilités de croisements entre ces taxons n'avaient jamais été étudiées. Seuls quelques hybrides supposés entre les sous-espèces *cuspidata* et *europaea* étaient en collection à l'INRA de Montpellier ou encore au CNR de Perugia (Italie). Pour cerner l'histoire de cet arbre cultivé, la confrontation de données archéologiques, biogéographiques et génétiques s'avérait indispensables (e.g. Zeder *et al.* 2006).

Dans l'objectif de mieux comprendre les origines de l'olivier et pour une meilleure gestion de sa diversité, il s'avérait tout d'abord nécessaire de faire un inventaire exhaustif des ressources

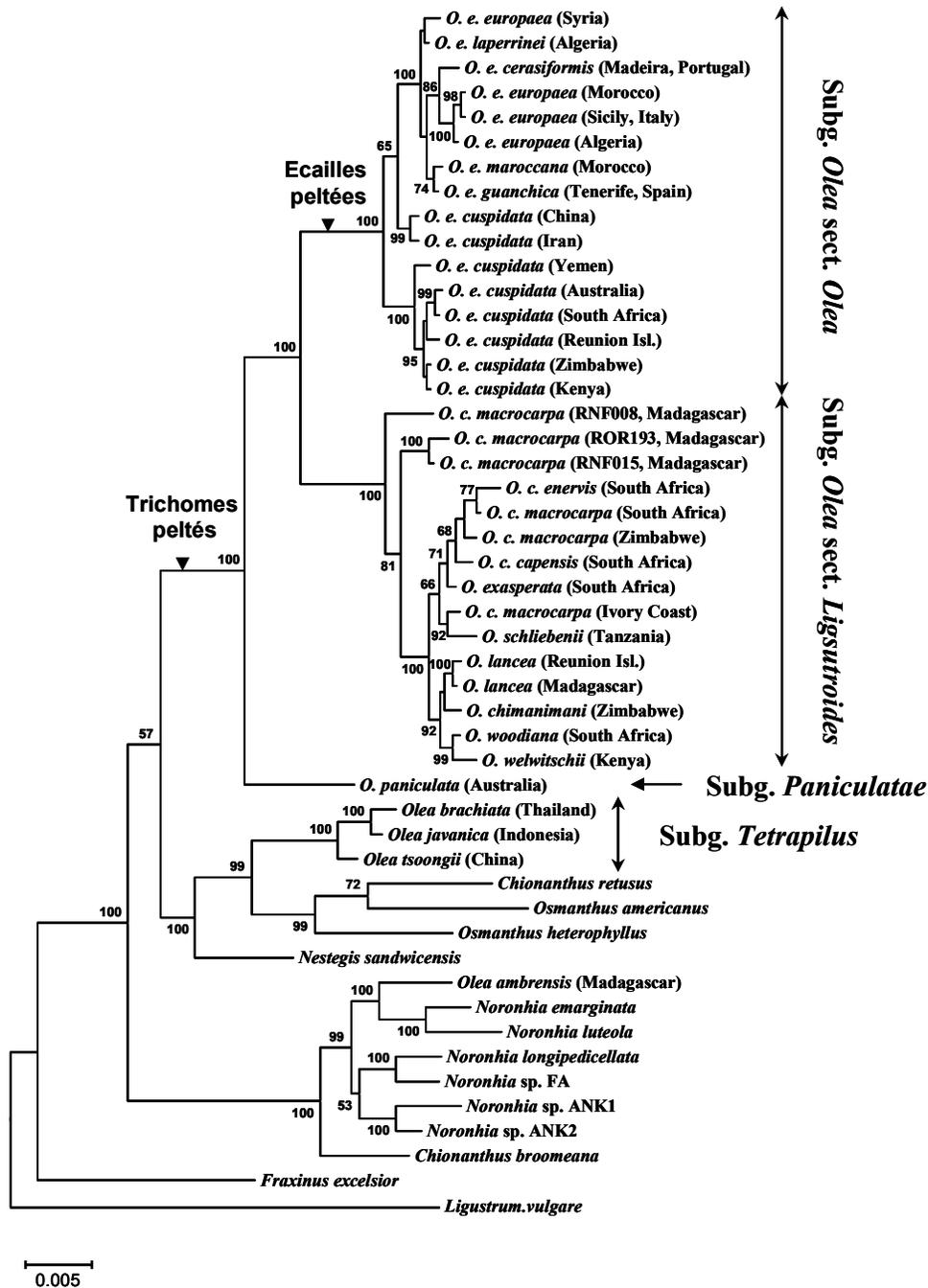
génétiques d'*Olea europaea*. Une caractérisation phylogénétique du genre *Olea* devait permettre de mieux connaître sa position systématique dans la sous-tribu Oleinae ainsi que de tester certaines hypothèses sur la diversification du genre (i.e. polyphylie, datation des principaux événements de divergence, importance de la polyploïdie). Par la suite, l'histoire de l'olivier sauvage a été étudiée à l'aide d'une approche de phylogéographie sur l'ensemble de son aire de répartition (Fig. 1). La domestication de l'olivier ainsi que l'origine de ses formes envahissantes ont été reconsidérées sur les bases de données moléculaires. Les marqueurs moléculaires ont été aussi utilisés pour l'identification variétale et la gestion de la collection française d'oliviers cultivés (au CBNM Porquerolles). Enfin, la conservation *in situ* de taxons menacés a été abordée, en particulier au Sahara Central, sur la base d'une étude de génétique des populations et d'analyses de paternité (flux de gènes polliniques).

## 2.2. Clarification de la taxonomie et de la diversification du genre *Olea* en Afrique

### 2.2.1. Utilisation de l'outil moléculaire dans la classification du genre *Olea*

Durant ma thèse, j'ai collecté puis caractérisé à l'aide de marqueurs RAPD plusieurs espèces appartenant aux différentes sous-familles et sections du genre *Olea* (Besnard *et al.* 2002a). Ces investigations préliminaires sont cependant restées très limitées du fait de l'utilisation de marqueurs RAPD qui présentent une trop grande variation et sont donc très peu informatifs pour établir les liens entre sous-familles et sections. Par la suite, avec mes collègues de Madrid, nous avons obtenu l'autorisation de caractériser plusieurs spécimens d'herbier en collection aux jardins botaniques de Madrid, Harvard, Genève et Kew. Avec cette opportunité d'établir un échantillonnage beaucoup plus représentatif du genre *Olea*, il était désormais possible de reconstruire une phylogénie quasi-exhaustive du genre et de reconsidérer certaines hypothèses sur sa diversification. Ainsi, quatre fragments d'ADN chloroplastique ainsi que l'espaceur interne transcrit 1 de l'ADN ribosomique nucléaire (ITS-1 ; provenant de gènes ribosomiques fonctionnels) ont été séquencés sur nos échantillons d'*Olea* et des représentants de genres apparentés (principalement de la sous-tribu Oleinae ; Besnard *et al.* 2009c). Ces données ont permis des reconstructions phylogénétiques et des inférences sur l'évolution de ce groupe (datation moléculaire). De plus, sur un échantillonnage représentatif du genre *Olea*, des observations microscopiques d'herbier nous ont permis de caractériser l'indument de la face abaxiale des feuilles, un critère taxonomique important dans l'identification des sections et sous-familles (Green 2002).

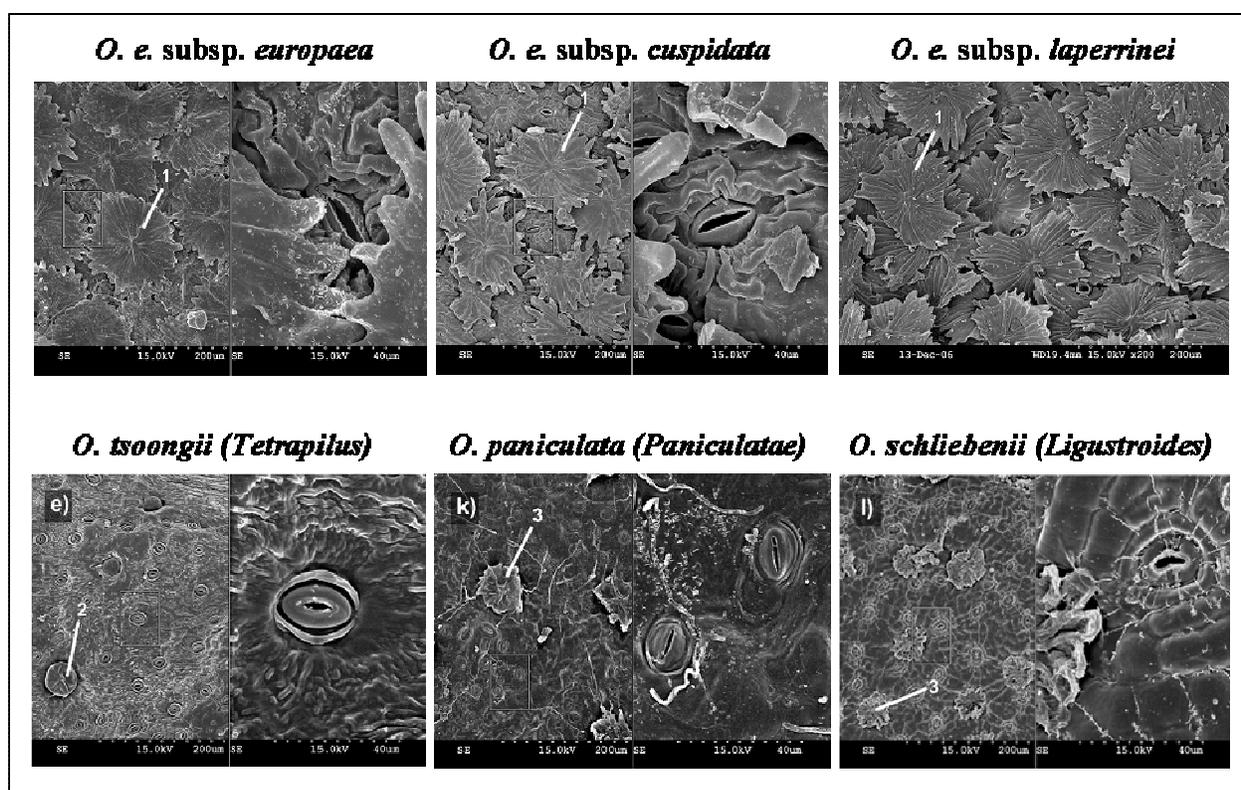
Les quatre grandes lignées d'*Olea* décrites dans une récente monographie du genre *Olea* (Green 2002) étaient fortement soutenues par nos données moléculaires (Fig. 2). Toutefois, *Olea ambrensis* (Madagascar) n'était pas apparenté à la section *Ligustroides* mais au genre *Noronhia* (endémique de Madagascar). La diagnose de cette espèce très rare étant incomplète (Perrier de la Bâthie 1952), sa position systématique avait été mal considérée. Par ailleurs, le genre *Olea* est apparu comme polyphylétique du fait que le sous-genre *Tetrapilus* est relié aux genres *Chionanthus*, *Osmanthus*, *Phillyrea* et *Nestegis* (à la fois sur la base des séquences de l'ADN chloroplastique et nucléaire). Cette étude a notamment démontré la nécessité de réviser toute la sous-tribu Oleinae à l'aide de l'outil moléculaire, et selon nos résultats (voir aussi Johnson 1957), le genre *Olea* devrait être restreint aux sous-genres *Olea* et *Paniculatae* (neuf espèces : *O. europaea*, *O. capensis*, *O. chimanimani*, *O. exasperata*, *O. lancea*, *O. schliebenii*, *O. welwitschii*, *O. woodiana* et *O. paniculata*).



**Fig. 2.** Arbre phylogénétique consensus du genre *Olea* basé sur quatre séquences d'ADN chloroplastiques (*trnL-trnF*, *trnT-trnL*, *trnS-trnG* et *matK*) et l'ITS-1 et reconstruit à l'aide d'une inférence Bayésienne. Les probabilités postérieures (> 50) sont indiquées sur chaque branche. Les apparitions des trichomes peltés puis d'un indument constitué de nombreuses écaillés peltées (cf. Fig. 3) sont indiquées sur l'arbre (Besnard *et al.* 2009c).

En considérant les fossiles d'Oléacées disponibles dans la littérature, une datation moléculaire relaxée (Thorne *et al.* 1998 ; Thorne & Kishino 2001) a permis d'estimer les temps de divergence entre les principales lignées d'*Olea*. Selon nos résultats les sous-genres *Olea* et *Paniculatae* ont divergé il y a  $28.0 \pm 4.4$  Millions d'années (Ma). La séparation des sections

*Ligustroides* et *Olea* serait apparue il y a  $20.5 \pm 5.5$  Ma. La diversification de ces deux sections (qui a suivi leur divergence) semble donc concomitante avec le début de l'aridification de l'Afrique continentale ( $\approx 16$  Ma ; Retallack 1992). Ainsi dans le complexe *O. europaea*, la divergence de la sous-espèce *cuspidata* est estimée à  $7.5 \pm 3.0$  Ma, probablement en liaison avec l'apparition du Sahara ( $\approx 7$  Ma ; Sepulchre *et al.* 2006). Par ailleurs, dans la section *Olea*, l'indument de la face abaxiale des feuilles est constitué d'écailles peltées chevauchantes (Fig. 3). Ces larges trichomes sont considérés comme étant impliqués dans l'adaptation à la sécheresse chez les membres de cette section car en protégeant les stomates ils créent des conditions favorables pour les échanges gazeux (Bongi *et al.* 1987). En particulier chez l'olivier de Laperrine du Sud Algérien, les trichomes sont si nombreux qu'il est impossible d'observer des stomates (Fig. 3). Nos datations suggèrent donc que l'apparition des feuilles densément recouvertes de trichomes (dans l'ancêtre de la section *Olea*) devrait être concomitante avec l'apparition de conditions xériques sur l'Afrique à partir du Miocène Moyen (Besnard *et al.* 2009c).



**Fig. 3.** Observations en microscopie électronique à balayage de la surface abaxiale de feuilles de trois membres du complexe *Olea europaea* [sous-espèces *europaea* (Espagne), *cuspidata* (Afrique du Sud) et *laperrinei* (sud Algérie)], *O. tsoongii* (subgen. *Tetrapilus*), *O. paniculata* (subgen. *Paniculatae*) et *O. schliebenii* (sect. *Ligustroides*). Ces photos illustrent la présence d'écailles peltés (1) dans la section *Olea*, de trichomes glandulaires (2) dans le sous-genre *Tetrapilus*, et de trichomes peltés (3) dans le sous-genre *Paniculatae* et la section *Ligustroides*. Des détails de stomates (à gauche) sont également donnés, excepté pour la sous-espèce *laperrinei*, où ils ne sont pas visibles. Echelles des photographies : 200  $\mu$ m et 40  $\mu$ m (détails des stomates). Photos. R. RUBIO DE CASAS.

### 2.2.2. Mise en évidence de polyploïdes dans le genre *Olea* à l'aide de différentes approches

A l'université de Lausanne, lors des premiers tests de marqueurs microsatellites nucléaires sur des individus appartenant aux six sous-espèces d'*O. europaea*, j'ai observé des patterns très

complexes (avec un maximum de six allèles par locus) sur tous les individus analysés appartenant aux sous-espèces *maroccana* (Haut Atlas marocain) et *cerasiformis* (Madère). Ces résultats étaient très inattendus puisque la polyploïdie n'avait pas été reportée dans le genre *Olea* (Green & Wickens 1989). Cependant, le comptage des chromosomes par les méthodes classiques s'est avéré fort compliqué car les chromosomes d'*O. europaea* sont très petits et leur nombre potentiellement très grand (e.g. 46 pour un diploïde, 92 pour un tétraploïde, 138 pour un hexaploïde...). Le comptage des chromosomes d'individus diploïdes étant déjà une gageure, il s'est avéré quasi-impossible d'estimer de façon précise le nombre de chromosomes sur les sous-espèces *maroccana* et *cerasiformis* par les méthodes classiques. Toutefois, la combinaison de marqueurs microsatellites très variables avec une méthode de cytométrie a permis de déterminer sans équivoque que les sous-espèces *maroccana* et *cerasiformis* sont respectivement hexaploïde et tétraploïde, tandis que les quatre autres sous-espèces d'*O. europaea* sont diploïdes (Besnard *et al.* 2008). Par ailleurs, sur la base d'analyses moléculaire et de cytométrie, j'ai récemment mis en évidence dans une population d'olivier de Laperrine que trois individus (sur un total de 102 génotypes ; soit env. 3%) étaient triploïdes (Besnard & Baali-Cherif 2009). De récentes investigations avec différents types de marqueurs (i.e. microsatellites, AFLP, séquences de l'intron du gène FAD6) ont permis de supporter l'hypothèse d'une origine allopolyploïde (segmentaire) des sous-espèces *maroccana* et *cerasiformis* (García-Verdugo *et al.* 2009 ; Besnard *et al.* en prep.). Les ancêtres des sous-espèces *europaea* et *guanchica* sont très probablement impliqués. En revanche, ces mêmes investigations ont fortement soutenu une origine autopolyploïde pour les trois individus triploïdes de l'olivier de Laperrine. Comparée aux diploïdes, la forte diversité haplotypique observée sur l'intron d'un gène nucléaire (FAD6) chez les polyploïdes *maroccana* et *cerasiformis* (Besnard *et al.* en prep.) supporte le fait que la polyploïde joue un rôle essentiel dans le maintien de la diversité génétique dans ces populations. Pour finir, sur la base de marqueurs multialléliques, aucune preuve de polyploïde n'a été observée dans la section *Ligustroides* (*O. woodiana*, *O. welwitschii*, *O. c. capensis*, *O. c. macrocarpa*, *O. exasperata* et *O. lancea* ; huit individus) et le sous-genre *Paniculatae* (*O. paniculata* ; un individu). Ceci n'exclut cependant pas que des populations polyploïdes existent aussi dans ces groupes taxonomiques.

Cette preuve de polyploïdie dans le complexe de l'olivier s'est avérée très intéressante sous plusieurs aspects. Tout d'abord, le statut polyploïde des sous-espèces *maroccana* et *cerasiformis* soutient le traitement taxonomique récemment proposé par Vargas *et al.* (2001) et Green (2002). Par ailleurs, on peut se demander pourquoi les deux sous-espèces macaronésiennes sont polyploïdes. En fait, il s'agit de deux taxa avec une aire de répartition extrêmement réduite et la sous-espèce *maroccana* est un taxon relique qui est considéré comme l'un des arbres les plus menacés de Méditerranée (Médail *et al.* 2001). Ceci suggère que le succès des oliviers polyploïdes est limité et que leur avenir est très incertain. Ils pourraient donc représenter un cul de sac évolutif. La polyploïdie pourrait avoir été initialement favorisée dans des populations très réduites permettant ainsi de limiter des effets de consanguinité (Soltis *et al.* 2004). Ensuite, certains désavantages liés à la polyploïde (comme un gros génome engendrant des difficultés de méiose et de mitose ; Comai 2005) ont probablement affectés leur succès. Actuellement, ces populations se reproduisent probablement très peu par voie sexuée (car des fleurs et des fruits ont très rarement été observés sur la sous-espèce *maroccana* ; Maire 1933 ; Médail *et al.* 2001) mais elles privilégient un fort développement végétatif des individus existants comme il a été observé chez l'olivier de Laperrine (Baali-Cherif & Besnard 2005). Ceci devrait permettre leur maintien sur le long terme (Médail *et al.* 2001). Par ailleurs, la forte diversité haplotypique observée chez

les polyploïdes suggère que les sous-espèces *maroccana* et *cerasiformis* devraient être des ressources génétiques importantes à conserver de toute urgence.

### 2.3. Phylogéographie du complexe de l'Olivier (*Olea europaea*)

Nos études ont ensuite été focalisées sur le complexe *O. europaea* dans l'objectif de mieux comprendre les origines de l'olivier méditerranéen. Si il est bien établi que le genre *Olea* est présent au nord-ouest de l'Afrique et le sud de l'Europe depuis la fin du Paléogène (environ 24 Ma ; pour une revue voir Terral *et al.* 2004b), il est également admis que les oliviers ont fortement régressé face à des changements climatiques du Tertiaire et particulièrement durant les glaciations du Pléistocène (e.g. Médail *et al.* 2001 ; Terral *et al.* 2004b). Une étude phylogéographique d'*O. europaea* s'avérait donc nécessaire pour mieux connaître les relations entre sous-espèces, ainsi que pour identifier les différentes lignées d'olivier et leurs voies de recolonisation en Méditerranée durant la période postglaciaire.

Pour cela, j'ai tout d'abord prospecté des échantillons sauvages des six sous-espèces, et en particulier des oléastres (terme utilisé pour désigner les formes sauvages méditerranéennes). J'ai pu en multiplier une partie par voie végétative, qui constitue aujourd'hui une collection unique (à l'INRA de Montpellier) d'une bonne centaine de représentants des sous-espèces *europaea* (Syrie, Turquie, Libye, Algérie, Maroc, Espagne et France), *cerasiformis* (Madère), *maroccana* (Haut Atlas), *laperrinei* (Hoggar) et *cuspidata* (Afrique du Sud, Réunion, Kenya, Yémen, Iran et Australie) plus quelques hybrides diploïdes récemment obtenus et caractérisés [i.e. *O. e. cuspidata* (Iran) x *europaea* (?), *O. e. cuspidata* (Afrique du Sud) x *europaea* (?) et *O. e. laperrinei* (Hoggar) x *cuspidata* (Iran) ; données non publiées]. Par ailleurs, pour inclure dans notre étude quelques arbres de régions montagneuses particulièrement difficiles d'accès (e.g. Erythrée, Sud Egypte, Nord Est Soudan, Somalie), j'ai de nouveau obtenu les autorisations de caractériser quelques spécimens en collection aux herbiers de Montpellier, Genève et Kew. L'ensemble de nos échantillons ont été analysés avec différents types d'outils moléculaires et notamment des marqueurs RAPD (Besnard *et al.* 2001), AFLP (Rubio de Casas *et al.* 2006), microsatellites nucléaires (Breton *et al.* 2006 ; Besnard *et al.* 2007b), et cytoplasmiques (RFLP de l'ADN mitochondrial, RFLP, PCR-RFLP et microsatellites de l'ADN chloroplastique ; Besnard *et al.* 2002b, 2007c) ou encore des séquences d'ITS-1 (Besnard *et al.* 2007c). A cette occasion, j'ai développé différents outils au laboratoire en particulier les marqueurs RFLP des ADN cytoplasmiques (Besnard *et al.* 2000) et les marqueurs PCR sur l'ADN chloroplastique (ADNcp) pour analyser des ADN dégradés (i.e. ADN d'herbiers ; Besnard *et al.* 2003b).

#### 2.3.1. Investigations dans l'aire native

**Apports des marqueurs « fingerprinting » :** Les marqueurs de type RAPD et AFLP ont tout d'abord permis la distinction des six sous-espèces comme proposée par Green (2002). Ces marqueurs ont également clairement indiqué que les populations de la sous-espèce *cuspidata* sont bien différenciées des cinq autres sous-espèces (Besnard *et al.* 2001 ; Rubio de Casas *et al.* 2006) probablement du fait que la barrière du Sahara a grandement limité les flux de gènes entre la Méditerranée et l'Afrique tropicale (cf. partie 2.2.1). En Méditerranée (subsp. *europaea*), les populations d'oléastres de l'Est étaient clairement différenciées des populations de l'Ouest. Cependant, nos populations prospectées au Centre de la Méditerranée (Libye, Tunisie, France) étaient aussi génétiquement intermédiaires entre ces deux pools « Est » et « Ouest » suggérant

l'existence de flux de gènes (Besnard *et al.* 2001 ; Rubio de Casas *et al.* 2006). Un pattern similaire a été récemment observé sur des données microsatellites nucléaires (Breton *et al.* 2006 ; Besnard *et al.* 2007b). De façon très étrange, des populations d'oléastres relativement éloignées (« Ostricone » en Corse et « Tamanar » au Maroc) n'étaient pas significativement différenciées alors que des populations très proches (« Cap des Mèdes » et « La Repentance » à l'île de Porquerolles) pouvaient être très différentes (Besnard *et al.* 2001 ; 2002b). La distribution très fragmentée de l'oléastre en Méditerranée, l'influence humaine jouant un rôle dans le déplacement de cultivars ou de « landraces » à plus ou moins longue distance ainsi que l'existence de flux de gènes depuis les formes cultivées vers les formes sauvages sont très probablement responsables de ce type de pattern. Il apparaît donc que l'interprétation de ces résultats est très délicate car l'impact anthropique dans le Bassin Méditerranéen est extrêmement important depuis l'Holocène.

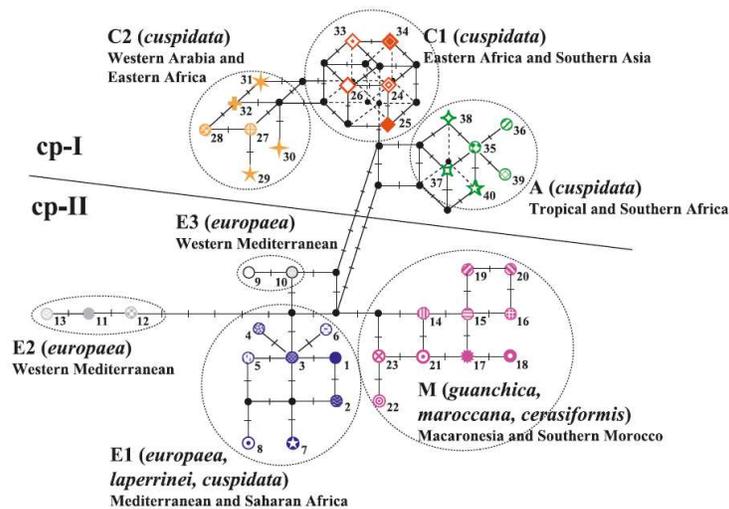
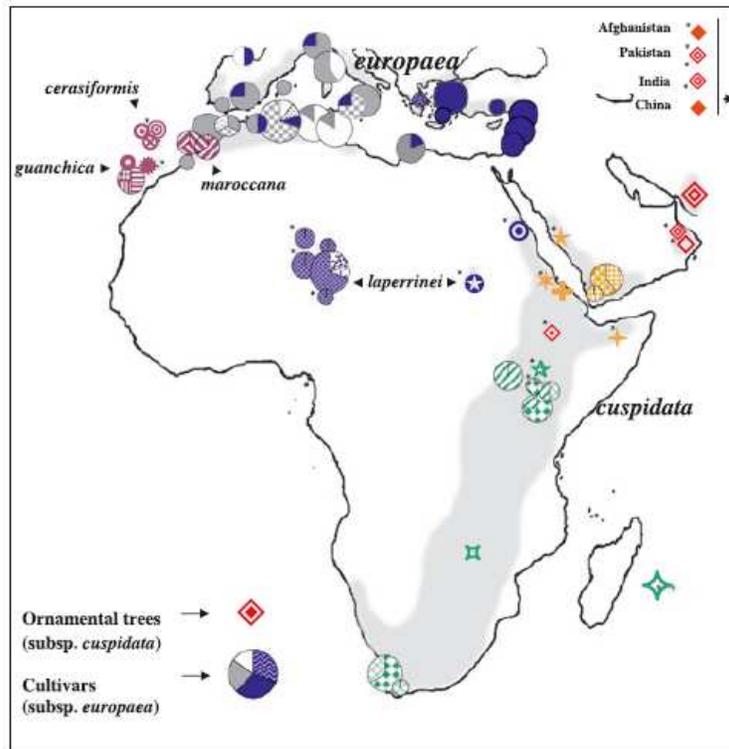


Fig. 4. Réseau des haplotypes chloroplastiques du complexe de l'olivier (reconstruit selon un principe de Parcimonie ; Besnard *et al.* 2007c)

**Caractérisation du polymorphisme des génomes cytoplasmiques :** En parallèle, nous avons caractérisé le polymorphisme des ADN cytoplasmiques. Leur transmission étant maternelle (Besnard *et al.* 2000 ; 2002b), leur utilisation permet de caractériser les flux de gènes par graines. De plus, comparée au pollen, la dissémination des graines étant beaucoup plus limitée et les phénomènes de dérive génétique plus marqués (Petit *et al.* 1993), une très forte structure géographique de la diversité génétique des marqueurs cytoplasmiques est attendue. L'absence supposée de recombinaison facilite également la reconstruction de réseau d'haplotypes. Pour ces raisons, les marqueurs cytoplasmiques sont des marqueurs de choix pour la phylogéographie. Cependant, la variation des séquences d'ADNcp d'olivier est extrêmement faible (similarité de 99.7% entre les haplotypes les plus distants ; Fig. 2). Il a donc été nécessaire de développer des marqueurs ADNcp de type PCR-RFLPs et microsatellites (Besnard *et al.* 2003b). Un échantillonnage de 185 arbres bien répartis sur l'ensemble de l'aire de répartition du complexe *O. europaea* a ensuite été caractérisé avec 20 loci. Quarante profils d'ADNcp ont pu être mis en évidence et un réseau a été reconstruit sur un principe de maximum de parcimonie (Fig. 4). L'existence de sept lignées d'haplotypes a été mise à jour. Une forte structuration géographique de la diversité est également observée puisque chaque lignée présente une distribution bien délimitée (Fig. 5). Ainsi, les trois lignées du clade cp-I sont uniquement détectées dans la sous-

espèce *cuspidata* (Fig. 4 & 5), tandis que les quatre lignées E1, E2, E3 et M (clade cp-II) sont restreintes au Nord de l'Afrique. La lignée E1 est partagée par les sous-espèces *cuspidata* (Sud Égypte), *laperrinei* et *europaea* suggérant de possibles échanges entre l'Afrique tropicale et la Méditerranée dans les populations de l'Afrique Saharienne. Les sous-espèces *guanchica*, *maroccana* et *cerasiformis* présentent une lignée spécifique d'haplotypes (M) suggérant un ancêtre maternel commun pour ces trois taxa. Enfin, la distinction entre les populations méditerranéennes de l'Est et de l'Ouest (déjà précédemment observée sur la base de marqueurs AFLP et RAPD) est également bien soutenue par les marqueurs cytoplasmiques (Fig. 5 ; voir aussi Besnard *et al.* 2002b).



**Fig. 5.** Répartition géographique des haplotypes chloroplastiques dans l'aire native du complexe *Olea europaea* (Besnard *et al.* 2007c)

**Utilisation de séquences de l'ADN ribosomique :** Un sous-échantillonnage d'arbres (38) appartenant aux différentes sous-espèces a été caractérisé avec des séquences d'ITS-1 en utilisant un couple d'amorces différent (Hess *et al.* 2000 ; Besnard *et al.* 2007c) de celui utilisé dans la partie taxonomique (ce travail étant antérieur à celui présenté en partie 2.2.1). Les gènes ribosomiques sont constitués de plusieurs milliers de copies en tandem dans un génome, et de ce fait, les marqueurs de ces gènes sont assez particuliers à utiliser du fait qu'ils peuvent être constitués d'un mélange de séquences différentes plus ou moins soumises à l'évolution concertée. Cependant, leur utilisation présente certains avantages et en particulier ils peuvent être facilement amplifiés même sur du matériel très dégradé comme des herbiers ou des restes archéologiques (e.g. Elbaum *et al.* 2006). Les séquences que nous avons générés se sont avérés correspondre à des unités ribosomiques non fonctionnelles (pseudogènes ; notamment caractérisés par une faible teneur en GC, un haut niveau de substitution, une structure secondaire instable ou encore une

taille relativement courte par rapport à des ITS-1 d'unités fonctionnelles ; Besnard *et al.* 2007c). Toutefois, l'analyse de ces données a révélé une très forte structuration géographique de la diversité très comparable à celle observée sur l'ADNcp. Le mélange d'unités phylogénétiquement très divergentes dans plusieurs individus d'Afrique Centrale et du Nord Ouest suggère encore l'hypothèse d'une évolution réticulée dans le complexe de l'olivier. En particulier, ces données soutiennent le fait que la sous-espèce *laperrinei* du Sahara central est issue de l'hybridation entre des individus d'Afrique tropicale et du Nord de l'Afrique (Angiolillo *et al.* 1999 ; Besnard *et al.* 2007c).

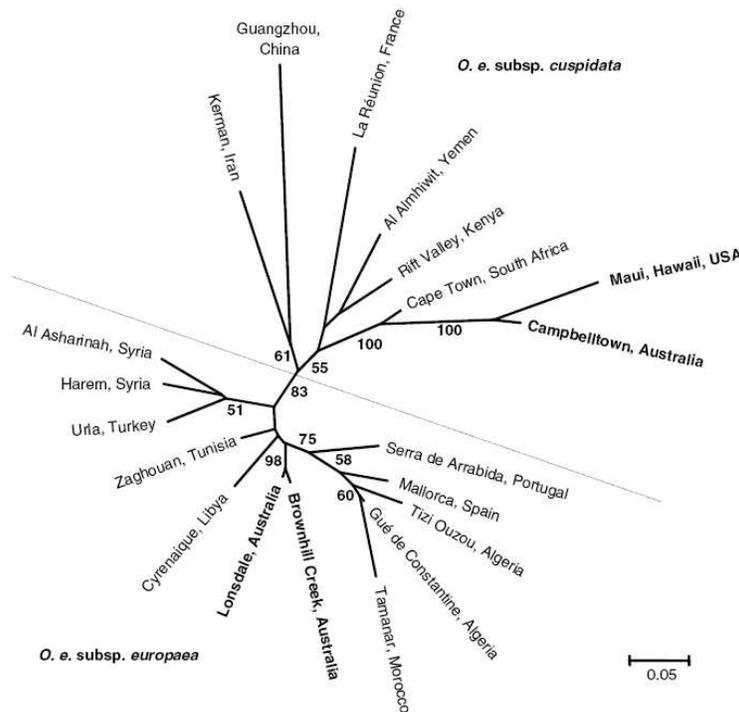
Il est bien évident que cette étude s'appuie sur un relativement faible nombre d'individus, mais grâce à l'utilisation d'herbiers, nous avons eu accès à un échantillonnage assez représentatif de l'aire de répartition naturelle de l'olivier. Bien que des désaccords aient pu être observés selon le type de marqueurs utilisés, en particulier dans les massifs du Sahara ou en Macaronésie (Madère), des lignées d'olivier bien différenciées ont clairement été identifiées notamment en Afrique Tropicale, Asie du Sud, Canaries et Méditerranée (« Est » et « Ouest »). Notre travail constitue désormais une base de travail pour des études plus approfondies sur des populations bien ciblées, mais aussi pour des études sur matériel archéologique (e.g. ADNcp et pseudogènes d'ITS-1 ; Elbaum *et al.* 2006).

L'existence de deux pools génétiques en Méditerranée (aussi détectés avec des isozymes par Lumaret & Ouazzani 2001) constitue un point essentiel pour la suite de nos investigations sur l'origine de formes cultivées. D'autres espèces, telles que la vigne sauvage (Arroyo-Garcia *et al.* 2006), présentent d'ailleurs un pattern très similaire. Nous avons interprété cette différenciation « Est - Ouest » comme une conséquence d'un isolement des populations d'Oléastres à l'Est et l'Ouest de la Méditerranée durant le Pléistocène, notamment dû à des barrières désertiques (Egypte-Libye) et maritimes (e.g. mer adriatique). La recolonisation relativement récente de la Méditerranée par l'olivier durant l'Holocène est également attestée à l'Est comme à l'Ouest par les données palynologiques et anthracologiques et notamment en relation avec les activités humaines (e.g. Carrión & Dupré 1996 ; Yll *et al.* 1997 ; Mudie *et al.* 2007). Il apparaît que l'un des points critiques de notre étude est le choix de nos populations car il existe des populations ensauvagées un peu partout en Méditerranée (Besnard *et al.* 2001c ; Belaj *et al.* 2007) et les flux de gènes entre cultivars et oléastres pourraient être localement non négligeables. Pour approfondir ce travail, l'échantillonnage devrait être mieux constitué mais la démarche reste encore à définir. Faudra t-il faire un choix de populations difficiles d'accès et donc *a priori* moins affectées par des flux de gènes depuis les formes cultivées ou à l'inverse faudra t-il faire une prospection systématique des populations dans n'importe quelle situation ? J'envisagerai actuellement plutôt de suivre cette deuxième option, mais il sera aussi nécessaire de bien caractériser l'environnement socio-écologique dans lequel chaque population a été prospectée pour ensuite tester son influence sur la structuration de la diversité génétique intra- et inter-populations.

### 2.3.2. Origines des oliviers envahissants du nouveau monde

L'ensemble de nos données sur les populations natives a récemment été utilisé pour déterminer les origines de populations envahissantes (Besnard *et al.* 2007b). Quatre populations naturalisées ont été échantillonnées en Australie et à Hawaii (par D. COOKE, F. & K. STARR) puis caractérisées avec divers marqueurs (ITS-1, ADNcp, microsatellites nucléaires ; Stage MSc I de

L. WILLE & P. HENRY). Nous avons très clairement démontré que les populations de l'Est de l'Australie (Campbelltown) et d'Hawaii sont fortement apparentées et probablement originaires du Sud de l'Afrique (subsp. *cuspidata*; Fig. 6). En revanche les populations du Sud de l'Australie (Lonsdale et Brownill Creek) étaient fortement apparentées aux cultivars méditerranéens et à des populations du centre de la Méditerranée (Tunisie, Libye ; subsp. *europaea* ; Fig. 6). Un hybride naturel (probablement de première génération) entre les sous-espèces *europaea* et *cuspidata* a aussi été mis en évidence dans une population de l'Est de l'Australie bien que les taxa envahissants ne soient pas à l'heure actuelle en contact. Toutefois, l'olivier méditerranéen est fréquemment planté dans les jardins à l'Est de l'Australie où la sous-espèce *cuspidata* est envahissante, permettant ainsi ces hybridations. Par ailleurs, l'histoire d'invasion de ces deux taxa semble très différente puisque les populations de *cuspidata* ont subi une forte érosion génétique probablement liée à l'introduction de quelques individus et des goulots d'étranglement successifs (cf. Hawaii), tandis que les populations d'Australie du Sud présentent une forte diversité génétique comparable aux populations méditerranéennes ce qui est probablement dû à l'introduction multiple de cultivars et/ou de portes greffes (Besnard *et al.* 2007b). Ces populations dont les origines datent tout au plus du début du 19<sup>ème</sup> siècle sont d'un intérêt tout particulier car elles devraient être considérées (en comparaison des populations d'oléastres) pour mieux comprendre la recolonisation de l'olivier en Méditerranée qui a aussi sans doute été relativement rapide si on considère certaines données palynologiques de la fin de l'Holocène (e.g. Kelly & Huntley 1991 ; Yll *et al.* 1997 ; Mudie *et al.* 2007).



**Fig. 6.** Arbre phénétique non enraciné illustrant les similitudes génétiques entre populations invasives et natives (basées sur huit loci microsatellites nucléaires). L'arbre est reconstruit à partir de distances inter-populations (« minimum genetic distances » ; Nei, 1987) à l'aide de l'algorithme du « neighbor-joining » (Nei, 1987). Les valeurs de Bootstrap supérieures à 50% sont indiquées sur chaque branche correspondante. Les deux groupes de populations appartenant aux sous-espèces *europaea* et *cuspidata* sont délimités par un trait (Besnard *et al.* 2007b).

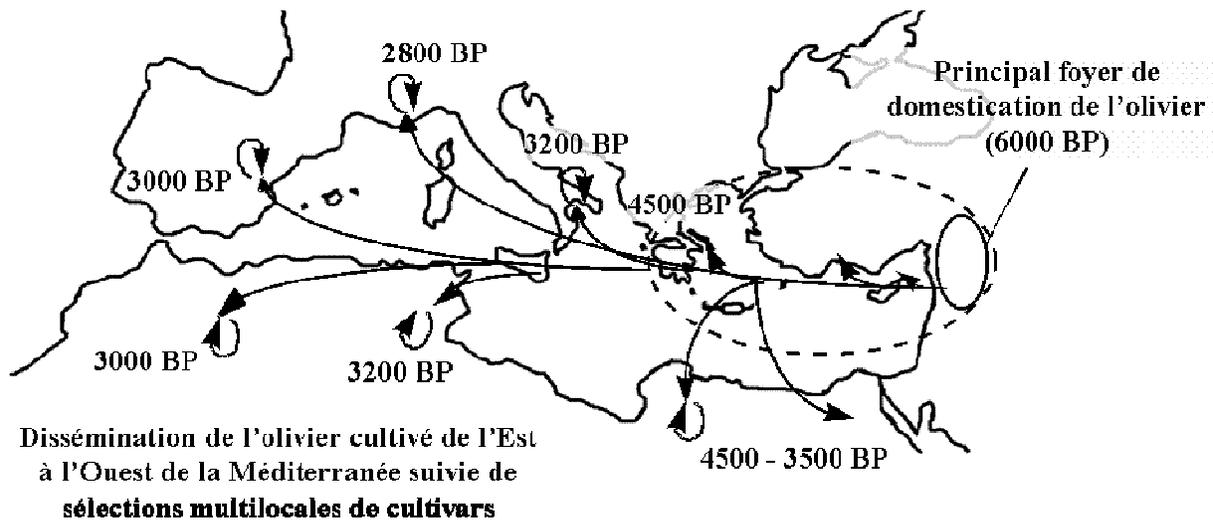
## 2.4. Domestication multilocale de l'Olivier méditerranéen

La comparaison de la diversité génétique des formes sauvages et cultivées a été très utilisée pour inférer les origines de plantes cultivées (e.g. Heun *et al.* 1997 ; Salamini *et al.* 2002 ; Zeder *et al.* 2006). Pour suivre cet objectif, un échantillonnage d'une centaine de cultivars de l'ensemble de la Méditerranée a été rassemblé durant ma thèse, notamment avec l'aide des collègues espagnols de l'IRO de Cordoue (en particulier avec L. RALLO) où la collection internationale d'olivier a été établie. Ces arbres ont été caractérisés à la fois avec des marqueurs cytoplasmiques et RAPD, et plus récemment à l'aide de marqueurs microsatellites (Breton *et al.* 2008). Ces données étaient ensuite analysées et notamment comparées à celles des formes sauvages. Avec Ph. BARADAT, nous avons tout d'abord utilisé des analyses discriminantes basées sur les données RAPD pour tester l'existence d'une structuration génétique liée à l'utilisation des olives et à l'origine géographique des cultivars. Nous avons ensuite mesuré les similitudes de chaque cultivar avec les deux principaux pools génétiques de la Méditerranée (« Est » et « Ouest »). Cette analyse s'apparente à un test d'assignement d'un individu à des populations prédéfinies.

Notre étude a établi que les cultivars utilisés pour l'huile et ceux utilisés pour les olives de tables étaient significativement différenciés sur la base de marqueurs RAPD. De même, les variétés d'Andalousie, d'Algérie et de Corse étaient très nettement différenciées des autres variétés analysées. Ces observations sont probablement liées au fait que certaines variétés de l'Ouest utilisées pour l'huile présentent globalement des profils moléculaires atypiques (Besnard *et al.* 2001a). Par ailleurs, les marqueurs cytoplasmiques ont attesté que les oliviers cultivés ont des haplotypes communs avec les deux principaux pools génétiques méditerranéens (« Est » et « Ouest »). Toutefois, une origine principale à l'Est était suggérée puisque presque 85% des cultivars caractérisés ont les chlorotypes de la lignée E1 (no 1 et 2) qui est prépondérante dans les formes sauvages de l'Est (Fig. 5 ; Besnard *et al.* 2001b). Les chlorotypes de l'Ouest no 9, 11, 12 et 13 (lignées E2 et E3 ; Fig. 4) étaient uniquement détectés dans des cultivars de l'Ouest. Basé sur les marqueurs RAPD, les cultivars sont aussi globalement génétiquement plus proches des formes sauvages de l'Est (Besnard *et al.* 2001b). Toutefois, les variétés avec un chlorotype spécifique de l'Ouest de la Méditerranée (lignées E2 et E3) présentent souvent des profils RAPD similaires aux formes sauvages de l'Ouest de la Méditerranée (e.g. 'Sabina', 'Chemlal' ou 'Lechin de Sevilla'; Besnard *et al.* 2001b). De nombreuses variétés de l'« Ouest » présentaient aussi des profils intermédiaires suggérant des mélanges entre les deux pools.

Nos résultats supportent l'idée d'une domestication multi-locale, comme attestée par des données archéologiques (e.g. Terral & Arnold-Simard 1996 ; Galili *et al.* 1997), mais il est clair que le pattern a été probablement fortement obscurci par les déplacements de cultivars à longue distance et les hybridations entre des arbres d'origine différente (Fig. 7). Récemment, Breton *et al.* (2008) ont essayé d'approfondir ce travail en reprenant une démarche assez semblable pour assigner les cultivars à des pools génétiques panmictiques sur la base données microsatellites (12 loci). Cette idée semble *a priori* très intéressante, mais je suis resté toutefois très septique sur l'utilité d'une telle approche en particulier pour inférer précisément l'origine géographique d'un cultivar. Tout d'abord, comme précédemment évoqué, la pureté des populations d'oléastres reste très discutable et incertaine. La détection de groupes panmictiques dans les populations sauvages d'olivier reste aussi problématique. En effet, l'analyse Bayésienne effectuée avec le logiciel STRUCTURE (Pritchard *et al.* 2000) par Breton *et al.* (2006) montre que plus d'un tiers (36%) des oléastres ne peuvent être attribués à des groupes panmictiques. De plus, sur les sept groupes

identifiés, deux n'étaient constitués que de trois individus. A mon avis, il est difficile d'interpréter ces résultats et il n'est pas évident que tous les groupes d'oléastres observés représentent différentes populations reliques du Tertiaire comme proposé par Breton *et al.* (2006). En effet, une différenciation inter-population peut rapidement apparaître lors de la recolonisation et tout particulièrement dans un habitat très fragmenté (par ex entre les îles Méditerranéennes), comme observé entre les deux populations envahissantes de la sous-espèce *cuspidata* (Besnard *et al.* 2007d). De plus, les clones d'olivier cultivé (qui fleurissent chaque année et produisent une énorme quantité de pollen comparée à des formes sauvages) pourraient avoir fortement « pollué » l'intégrité des populations d'oléastres sur des millénaires, même si ce phénomène pourrait être relativement limité à l'échelle d'une génération (cf. cas de la vigne sauvage ; di Vecchi-Staraz *et al.* 2009). A partir de ce constat, l'assignement d'un cultivar (clone) à un de ces groupes ne reflète pas forcément son origine dans ce groupe mais potentiellement le fait qu'il ait secondairement échangé des gènes avec lui. Une autre approche intégrant la coalescence des gènes devrait être plus informative sur la dynamique des populations et apporter de nouvelles informations (cf. partie 4).



**Fig. 7.** Scénario de dissémination de l'oléiculture dans le Bassin Méditerranéen d'après Besnard *et al.* (2001b). Un foyer principal de domestication, attesté par des données archéologiques (e.g. Galili *et al.* 1997 ; Zohary & Hopf 2000), est représenté au Levant, tandis que la sélection multilocale de l'olivier suite à des introductions depuis l'Est (mise en évidence sur les données moléculaires) est également datée d'après des données archéologiques (selon Terral 1997).

## 2.5. Apport de l'outil moléculaire dans la conservation des ressources génétiques cultivées *ex situ*

Chaque pays du pourtour Méditerranéen est responsable de gérer ses propres ressources génétiques d'olivier cultivé. Des collections nationales ont ainsi été mises en place, chacune généralement constituée de plusieurs centaines d'arbres (497 représentant plus de 100 dénominations dans la collection française de Porquerolles en 2003). Cependant, l'identification variétale basée sur des critères morphologiques n'est pas toujours aisée et il n'est pas rare de trouver dans les collections des génotypes en plusieurs exemplaires sous des noms différents (synonymies) ou à l'inverse des génotypes différents sous le même nom (homonymie). Des

mutations vont également apparaître dans les différents clones d'olivier, pouvant mener à la distinction phénotypique de nouvelles variétés. Des erreurs d'étiquetage sont aussi relativement fréquentes. La gestion de ces ressources génétiques pose donc certains problèmes notamment sur l'identification précise des génotypes mais aussi sur la définition des priorités de conservation. Pour aider la gestion des collections, les marqueurs génétiques ont donc aussi été utilisés notamment dans le but d'identifier les variétés d'olivier.

L'utilisation des marqueurs RAPD a tout d'abord été testée sur un échantillonnage exhaustif du Bassin Méditerranéen (Besnard *et al.* 2001c). Ces marqueurs se sont avérés très discriminants comparés à des marqueurs de type isozymes ou cytoplasmiques, la combinaison de trois amorces permettant de distinguer tous les génotypes étudiés. Ces marqueurs ont ensuite été utilisés en combinaison avec quelques loci microsatellites pour caractériser la collection de Porquerolles (Khadari *et al.* 2003). Le multiallélisme des marqueurs microsatellites s'est avéré également très discriminant (e.g. Tessier *et al.* 1999). Parmi les résultats notoires de ces travaux, on peut mentionner que certaines dénominations ont été détectées comme polyclonales. C'est le cas de 'Sabina' ou 'Ghjermana' en Corse. A l'inverse des cas de synonymies ont également été observés comme pour les variétés 'Picholine Marocaine' (Maroc), 'Cañivano Blanco' (Espagne), Sigoise (Algérie) et 'Shimlali' (Israël), ou les clones de 'Ghjermana' qui correspondent aux variétés italiennes 'Moraiolo' et 'Frantoio'. Ces observations attestent clairement de déplacements plus ou moins importants de variétés entre différents pays de la Méditerranée. Ceci mène potentiellement à de nombreuses redondances entre les collections nationales. Cependant, les vieilles variétés locales, souvent considérées agronomiquement moins intéressantes, représentent certainement un patrimoine génétique plus riche mais leur prospection n'est sans doute pas toujours optimale et nécessiterait encore quelques investigations (e.g. Khadari *et al.* 2008). La mise en place d'une banque de données génétiques représente désormais un outil très important pour caractériser de nouvelles prospections à inclure dans la collection de Porquerolles (Khadari *et al.* 2003). Par ailleurs, la caractérisation des formes cultivées avec les marqueurs cytoplasmiques nous a permis de démontrer sans équivoque que le déterminisme génétique d'un type de stérilité mâle de l'olivier est cytoplasmique (Besnard *et al.* 2000).

## 2.6. Conservation des ressources génétiques sauvages *in situ* : le cas de l'Olivier de Laperrine (*O. europaea* subsp. *laperrinei*) du Sahara Central

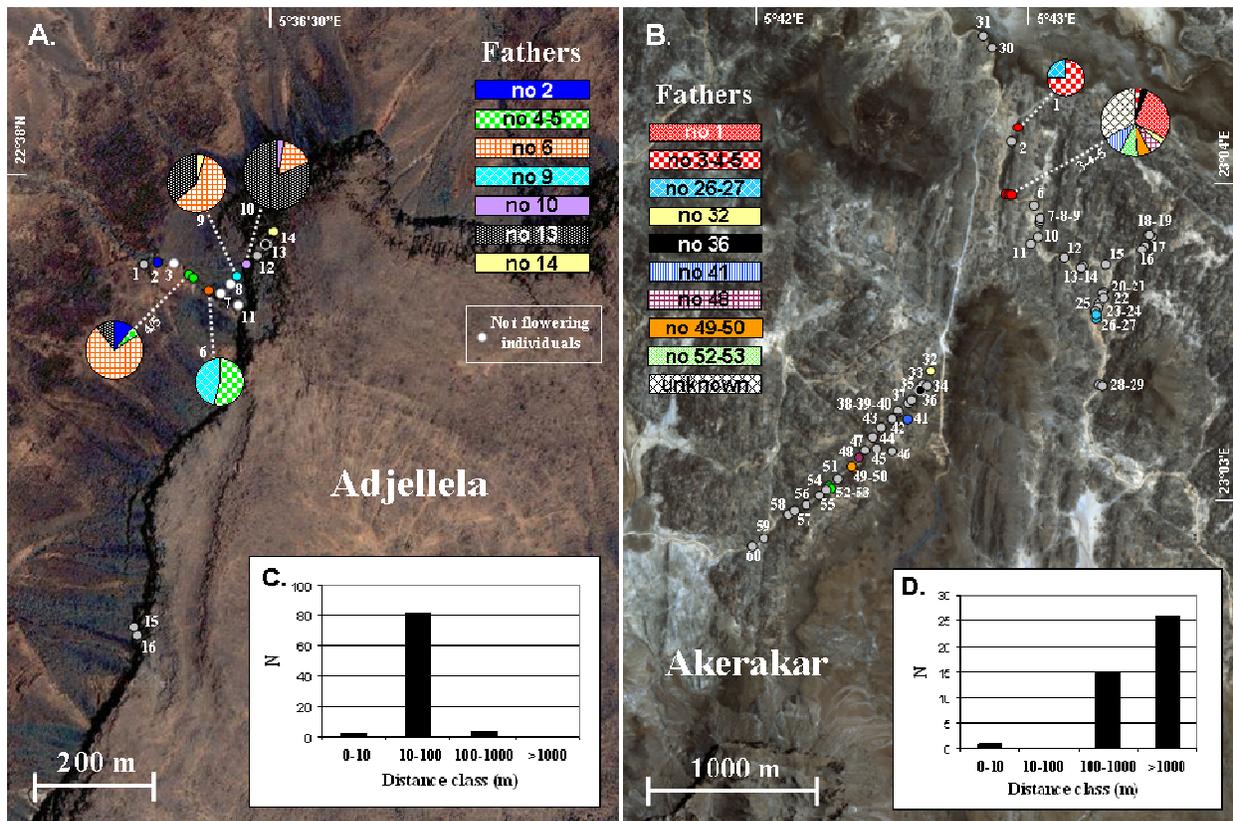
Les populations d'oliviers de Laperrine des massifs montagneux du Sahara central (Nord Niger et Sud Algérie) ont finalement attiré mon attention durant ces quatre dernières années car elles sont de plus en plus menacées en raison de changements climatiques et des activités humaines grandissantes. Ces arbres représentent pourtant un élément fondamental de l'écosystème des montagnes du Hoggar et de l'Aïr et leur préservation est un challenge pour les générations futures, notamment pour lutter contre la désertification. L'olivier de Laperrine est généralement présent au bord des lits d'oued ou dans les falaises (parfois très difficiles d'accès) et chaque station est très souvent constituée de moins de 100 individus. Les principales espèces arbustives présentes dans ces milieux sont *Rhus tripartitum*, *Acacia ehrenbergiana*, *Ficus salicifolia*, *Maerua crassifolia*, *Myrtus nivellei*, *Ziziphus lotus* et *Pistacia atlantica* (Quézel 1965). On le trouve également avec l'emblématique *Cupressus dupreziana* dans le Tassili n'Ajjer. Dans son milieu naturel, cet olivier représente une source de fourrage notamment pour le bétail et les chameaux, mais surtout pour les animaux sauvages en particulier le mouflon à manchette (*Ammotragus lervia*), une espèce menacée sur la liste rouge. Par ailleurs, certains vieux arbres

remarquables (parfois sacrés) peuvent également servir à promouvoir le tourisme dans ces régions. Enfin, ce taxon est intéressant pour sa résistance à la sécheresse potentiellement utile dans l'amélioration de l'olivier cultivé ou pour son utilisation comme porte greffe en milieu aride.

Afin de mieux connaître les stratégies de reproduction de l'olivier de Laperrine, avec D. BAALI-CHERIF (USTHB d'Alger), nous avons entrepris une première étude de génétique moléculaire (à l'aide de loci microsatellites nucléaires et chloroplastiques) sur des prospections du Hoggar en juin 2004. Par la suite, avec la collaboration de F. ANTHELEME (IRD Montpellier) et A. ABDOULKADER (Université de Niamey), nous avons étendu notre étude aux massifs du Niger et du Tassili n'Ajjer (dans le sud Est de l'Algérie). L'organisation de la diversité génétique entre massifs montagneux et à l'intérieur de chaque massif a été étudiée. Pour finir, nous avons eu l'opportunité de récolter un assez grand nombre de graines (fait assez exceptionnel) en 2007 dans le massif du Hoggar. Afin, de mesurer l'impact d'une multiplication par graines de l'olivier de Laperrine, nous avons caractérisé environ 200 graines avec huit loci microsatellite pour une étude de paternité (stage MSc d'initiation à la recherche de S. BETTINELLI-RICARDI et D. PARIETTI).

Nos premières investigations ont abouti à la conclusion que les souches d'olivier de Laperrine sont probablement très anciennes car environ un quart des individus analysés présentent plusieurs troncs (avec les mêmes profils génétiques) apparemment non connectés au dessus du sol. De plus, la reproduction par voie sexuée semble inexistante. La croissance clonale à partir de la souche est assez remarquable et nous avons estimé que certaines souches avoisinent les 100 m<sup>2</sup> alors que certains individus présentent plus de vingt pieds (Baali-Cherif & Besnard 2005 ; Anthelme *et al.* 2008). A l'aide d'une analyse Bayésienne (Pritchard *et al.* 2000), nous avons pu montrer l'existence de quatre groupes d'individus (« populations ») dans notre échantillonnage (Besnard *et al.* 2007a). Ces groupes correspondent approximativement aux quatre massifs prospectés bien que la différenciation génétique entre le Hoggar et le Tassili n'Ajjer soit très modérée ( $F_{ST} = 2.1\%$ ). Ce résultat démontre que les barrières désertiques ont considérablement limité les flux de gènes inter-massifs. Nous avons pu observer une forte diversité génétique dans le Hoggar comparée aux trois autres populations, et ce à la fois sur les loci nucléaires et chloroplastiques. Ceci suggère que le Hoggar est un important réservoir de gènes pour notre taxon et peut être son centre d'origine. Par ailleurs, des analyses d'auto-corrélation spatiale (Hardy & Vekemans 2002) dans chaque massif indiquent une relation significative entre les distances géographiques et les différenciations génétiques uniquement dans les massifs du Niger (Bagzane et Tamgak). Ce résultat surprenant a été interprété comme une conséquence d'une colonisation plus récente de ces massifs et/ou d'un habitat beaucoup plus fragmenté dans l'Aïr (du fait d'un espace altitudinal très limité pour l'olivier de Laperrine) limitant les flux de gènes entre patches d'individus et créant ce pattern d'isolement par la distance. Pour finir, une récente étude effectuée sur des lots de graines provenant de dix pieds mères de quatre patches différents dans le Hoggar a permis de faire une première analyse de paternité (Besnard *et al.* 2009a). Il est évidemment apparu que, plus le patch est petit, plus le nombre de pères est restreint. Ainsi, les 86 graines analysées provenant de trois génotypes maternels dans un patch isolé et constitué de sept individus (Tonget) étaient toutes issues d'une pollinisation à très courte distance (moins de 50 mètres) et l'un des parents était impliqué comme mère ou père dans toutes les descendance. Une situation assez similaire (avec des flux de pollens n'excédant pas 140 m) était observée sur la station d'Adjellela (Fig. 8a & 8c), ceci étant bien sûr grandement lié au fait que le patch est très isolé et que peu d'arbres (11 genets) étaient en

floraison dans la zone d'étude. Toutefois, des incompatibilités de croisement entre certains géotypes étaient aussi fortement suspectées (e.g. entre Tonget 2 et 3, Adjellela 9 et 10, Adjellela 6 et 13). Par ailleurs, la situation était assez différente dans une autre station constituée d'un grand nombre d'individus (> 150) dispersés dans un réseau d'Oueds au pied du mont Akerakar. En effet, les 42 graines analysées de deux pieds mères étaient issues d'au moins 15 pères différents, certains localisés à plusieurs milliers de mètres des pieds mères (dans un oued parallèle ; Fig. 8b & 8d).



**Fig. 8.** Mesures de flux de gènes polliniques dans deux stations d'oliviers de Laperrine du Hoggar (Besnard *et al.* 2009a). Les pères identifiés pour chaque pied-mère sont représentés dans les camemberts. **A.** Population d'Adjellela (tous les individus présents ont été échantillonnés et caractérisés) ; **B.** Population d'Akerakar (tous les individus matures présents dans un rayon de 1500 m autour des pieds-mères ont été échantillonnés). Pour Akerakar 3-4-5, le père n'était pas identifié pour onze graines. Un minimum de six pères étaient impliqués dans ces onze fécondations ; **C-D.** Diagramme des flux polliniques en fonction de la distance pour les deux populations.

A partir de cette étude, quelques enseignements peuvent être tirés sur les stratégies à adopter pour une meilleure persévérance de l'olivier de Laperrine. Dans l'objectif de multiplier ce taxon, deux méthodes peuvent être utilisées. La première est la multiplication végétative d'arbres existants et adaptés aux conditions locales. Cette stratégie nécessite de bien maîtriser le bouturage sur un maximum de géotypes différents. La seconde est une multiplication par voie sexuée. Dans cette stratégie alternative, il sera important de prospecter des graines sur un grand nombre d'arbres de préférence dans de relativement grandes populations et lors d'années très favorables pour la floraison et la fructification afin de maximiser le nombre d'arbres impliqués dans la reproduction (et donc d'éviter une érosion génétique ; Schoen & Brown 2001). Le

renforcement des populations actuelles avec ces plantules reste évidemment un autre problème et nécessitera de sensibiliser au préalable les populations humaines locales (pour protéger les jeunes plants du broutement par le bétail et les chameaux). Pour finir, des études sur d'autres organismes (et en particulier sur des essences arbustives menacées de disparaître) sont évidemment nécessaires pour une préservation optimale de tout l'écosystème.

## 2.7. Conclusions

L'ensemble de ces travaux a permis d'affiner nos connaissances sur la diversification de l'espèce *O. europaea* et sur les origines de l'olivier cultivé. Ils ont notamment contribué à mieux délimiter les ressources génétiques de cet arbre et auront un impact non négligeable dans leur gestion (e.g. collections, orientation des croisements dans d'éventuels programme de sélection) et leur conservation (*in situ* ou *ex situ*). Par ailleurs, des marqueurs moléculaires ont été mis au point et certains (e.g. ADNcp) devraient constituer des outils très importants pour la suite des investigations sur les ressources génétiques de cet arbre. Toutefois, il reste encore d'importantes questions sur les origines de l'olivier, et tout particulièrement sur le statut des formes sauvages méditerranéennes. On peut en effet encore se demander si la plupart des populations méditerranéennes sont véritablement sauvages ou tout simplement ensauvagées. L'importance de cette question réside dans les problèmes de conservation qui sont associés. La localisation des refuges glaciaires serait évidemment très utile pour définir des stratégies de protection des populations sauvages et potentiellement maximiser la conservation de la diversité génétique. Pour répondre à cette question, une approche assez différente de celle utilisée jusqu'ici devra être utilisée (cf. partie 4). Par ailleurs, il serait certainement intéressant de s'attarder à la domestication de l'olivier pour mieux comprendre les processus de sélection des formes cultivées. Avec les avancées technologiques actuelles, des approches de génomique fonctionnelle couplées à des études de génétiques quantitatives devraient ouvrir de nouvelles perspectives dans ce domaine.

### 3. Bases moléculaires de l'évolution de la photosynthèse C<sub>4</sub> chez les Poales

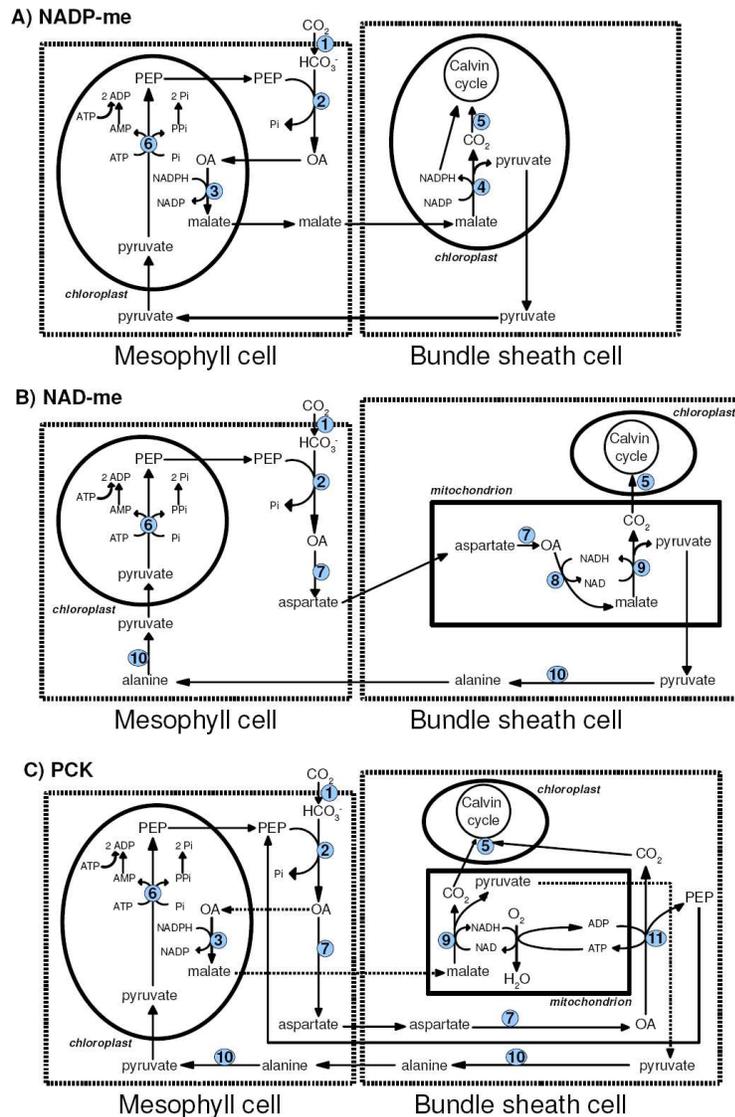
En parallèle de mes travaux sur l'olivier, j'ai développé un second axe de recherche sur l'évolution de la photosynthèse C<sub>4</sub>. J'ai eu cette opportunité après ma soutenance de thèse, lorsque j'ai été engagé en tant qu'ATER au laboratoire de Biochimie et de Biologie Moléculaire (LBGM) de l'université de la Réunion (entre 2000 et 2001). Il m'a été proposé de travailler sur les gènes codant les phosphoenolpyruvate carboxylases (PEPC) de la canne à sucre (*Saccharum* spp.) sur lesquelles une thèse avait été soutenue quelques années plus tôt dans l'unité de recherche (Barret 1997). En particulier, je devais développer des outils pour étudier la transcription de ces gènes. Les objectifs scientifiques de ce travail n'étaient toutefois pas clairement définis, me laissant une certaine autonomie sur la direction à donner. Après quelques investigations préliminaires, j'ai commencé une étude sur les bases moléculaires de l'évolution de la photosynthèse C<sub>4</sub>, un trait adaptatif supposé être apparu au moins dix fois chez les Graminées (Kellogg 2001). J'ai tout d'abord caractérisé des gènes codant les PEPC dans cette famille de plantes, puis avec Ph. RONDEAU (LBGM), nous avons analysé les gènes codant les malate déshydrogénases à NADP (Rondeau *et al.* 2005). L'ensemble de ces travaux a ensuite été abandonné quelques années avant d'être repris à l'université de Lausanne dans un contexte beaucoup plus favorable. J'ai en effet eu l'opportunité de monter un projet sur ce sujet, qui a été financé par la fondation nationale pour la recherche suisse (FNRS) : « Evolution of C<sub>4</sub> photosynthesis: Phylogenomic approach with phosphoenolpyruvate carboxylase-encoded genes in the grass family (Poaceae) » (no 3100AO-105886). Un étudiant en thèse (P.A. CHRISTIN) a commencé à travailler sur ce sujet en mars 2005 et soutenu son travail en Septembre 2008. Des collègues travaillant en Suisse (tout particulièrement N. SALAMIN), en Angleterre (V. SAVOLAINEN), en Irlande (T.R. HODCKINSON), en Afrique du Sud (A.M. MUASYA) et aux Etats Unis (M.R. DUVALL, E.H. ROALSON & E.A. KELLOGG) ont été associés à ces travaux dans le cadre d'échanges d'échantillons mais surtout pour l'analyse et l'interprétation des données. Ces collaborations ont été fructueuses, ayant déjà mené à d'importantes publications dans *Current Biology*, *Molecular Biology and Evolution*, *Genome Biology and Evolution*, *Plant Physiology* et *American Journal of Botany*.

Après un bref aperçu de l'importance écologique et économique de la photosynthèse C<sub>4</sub>, je vais présenter mes différents travaux entrepris sur ce caractère chez les familles des Graminées et des Cypéracées (incluant une partie des résultats de thèse de P.A. CHRISTIN). Je commencerai par présenter nos travaux sur la systématique moléculaire de l'ordre des Poales qui comprend environ 80% des espèces C<sub>4</sub>. A l'aide de fossiles, les datations moléculaires ont permis de reconsidérer le facteur temps et de tester l'hypothèse selon laquelle la baisse du CO<sub>2</sub> atmosphérique durant le Tertiaire aurait induit l'évolution du trait C<sub>4</sub> (Ehleringer *et al.* 1991 ; Pagani *et al.* 2005). Ensuite, nous avons caractérisé les bases moléculaires de l'évolution d'enzymes photosynthétiques suite à l'acquisition du caractère C<sub>4</sub> (tout particulièrement sur le gène *rbcL* et les gènes codant les PEPC). L'apport de l'ensemble de ces travaux dans la compréhension de l'évolution d'un caractère adaptatif primordial chez les plantes sera enfin discuté, notamment face aux changements globaux prédis dans les prochaines décennies.

#### 3.1. Quelques généralités sur l'évolution de la photosynthèse C<sub>4</sub>

La photosynthèse C<sub>4</sub> consiste en des modifications morphologiques (anatomie de Kranz) et

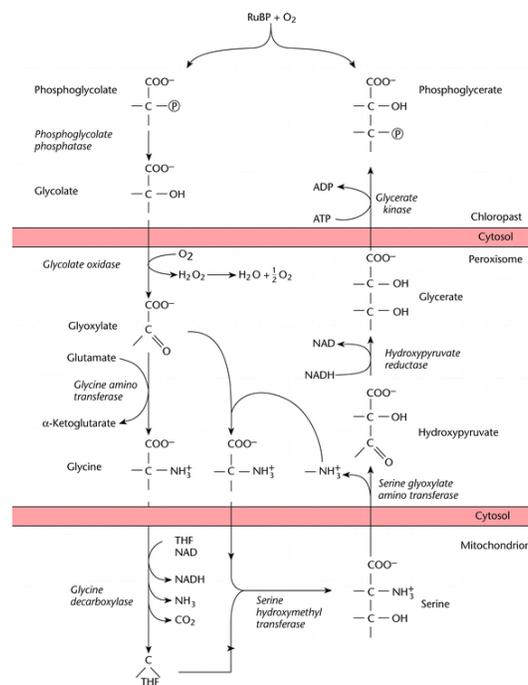
biochimiques au niveau de la feuille qui permettent de séparer les fonctions de fixation du dioxyde de carbone atmosphérique et de photosynthèse (cycle de Calvin ou photosynthèse dite en  $C_3$ ). Ainsi, le  $CO_2$  est tout d'abord fixé dans les cellules du mésophile par la PEPC puis une voie biochimique permet de le transporter et le libérer dans les chloroplastes des cellules de la gaine péri-vasculaire où il entre dans le cycle de Calvin (Fig. 9 ; von Caemmerer & Furbank 2003).



**Fig. 9.** Différentes voies biochimiques de la photosynthèse  $C_4$  (schémas d'après Christin 2008). Trois sous-types sont reconnus : A) NADP-me, B) NAD-me, et C) PCK. Les enzymes sont numérotées comme suit : 1 = anhydrase carbonique (CA) ; 2 = phosphoenolpyruvate carboxylase (PEPC) ; 3 = malate déshydrogénase à NADP (NADP-MDH) ; 4 = enzyme malique à NADP (NADP-ME) ; 5 = Rubisco ; 6 = pyruvate phosphate dikinase (PPDK) ; 7 = aspartate aminotransférase (AA) ; 8 = NAD-malate déshydrogénase (NAD-MDH) ; 9 = enzyme malique à NAD (NAD-ME) ; 10 = alanine aminotransférase ; et 11 = phosphoenolpyruvate carboxykinase (PCK).

La photosynthèse  $C_4$  est considérée comme un caractère d'adaptation à une faible teneur en  $CO_2$  dans l'atmosphère qui est apparue durant le Tertiaire (Elheringer *et al.* 1991 ; Pagani *et al.* 2005). Pour comprendre l'origine de ce trait, il est indispensable de considérer les propriétés

catalytiques et l'histoire de la Rubisco, la protéine la plus abondante de tout le règne végétal et dont l'importance capitale dans la fixation initiale du carbone durant la synthèse des composés organiques (cycle de Calvin) est bien connue. Cette enzyme présente une affinité pour le dioxyde de carbone mais également pour le dioxygène, car ces deux molécules sont très semblables (Tcherkez *et al.* 2006). Cependant, la fixation d'O<sub>2</sub> mène au phénomène de photorespiration (Fig. 10) qui est considérée comme une perte d'énergie que la plante doit limiter. Lorsque la Rubisco est apparue, il y a environ 2,7-2,9 milliards d'années (Nisbet *et al.* 2007), la composition de l'atmosphère était très différente de celle d'aujourd'hui. Le rapport des concentrations en CO<sub>2</sub> et O<sub>2</sub> était particulièrement élevé. Suite à l'apparition des organismes photosynthétiques, l'O<sub>2</sub> a progressivement augmenté dans l'atmosphère. Le rapport CO<sub>2</sub>/O<sub>2</sub> a ainsi été considérablement abaissé et l'affinité de la Rubisco pour le CO<sub>2</sub> a dû être améliorée pour éviter une trop grande fixation d'O<sub>2</sub>. Cependant, il a fallu résoudre un compromis entre l'affinité pour le CO<sub>2</sub> et l'activité catalytique de l'enzyme devait (Tcherkez *et al.* 2006). En conséquence, la plupart des plantes ont dû sur-exprimer la Rubisco pour palier au problème de sa faible activité catalytique. Chez certaines espèces C<sub>3</sub>, la Rubisco représente ainsi jusqu'à un tiers des protéines solubles exprimées par la plante, représentant un investissement en azote extrêmement important (Ghannoum *et al.* 2005). Par ailleurs, d'autres facteurs favorisent la photorespiration. En particulier, avec l'augmentation de température, l'affinité relative de la Rubisco pour l'O<sub>2</sub> par rapport au CO<sub>2</sub> augmente. De même, les conditions xériques vont également affecter la conductance des gazes (du fait de la fermeture des stomates en conditions de sécheresse) et contribuer à l'augmentation de la photorespiration. Pour faire face à ce phénomène, les plantes ont développé des systèmes de concentration du CO<sub>2</sub> autour de la Rubisco, réduisant ainsi la possibilité de fixation d'O<sub>2</sub> par cette enzyme. Le plus connu de ces systèmes est la photosynthèse C<sub>4</sub>, mais il faut citer également la photosynthèse de type CAM (Crassulean Acid Metabolism).



**Fig. 10.** La photorespiration : réactions biochimiques et réactions annexes conduisant à la formation de la glycine dans le peroxysoxe, puis de la sérine dans la mitochondrie. Notons que la régénération du phosphoglycérate nécessite ensuite une molécule d'ATP et de NADH ; selon Sharkey (2001).

Le système  $C_4$  est bien sûr coûteux en énergie pour la plante (Fig. 9), mais il apporte un avantage dans les conditions où la photorespiration est importante, c'est-à-dire en condition de forte température, de sécheresse et/ou de faible teneur en  $CO_2$ . Pour ces raisons, les plantes  $C_4$  sont principalement distribuées dans les milieux ouverts en zone tropicale (i.e. savanes, milieux semi-désertiques). Son importance écologique est également indéniable si on considère que 3% des angiospermes utilisent ce système mais représentent presque 20% de la productivité des plantes sur terre (Elheringer *et al.* 1997). Les espèces  $C_4$  dominent de nombreux écosystèmes tropicaux et représentent une source de fourrage extrêmement importante pour des animaux sauvages (i.e. savanes). De même, leur importance économique est majeure en condition tropicale avec des espèces à forte biomasse comme le maïs, le sorgho ou encore la canne à sucre. On trouve également des espèces tolérantes de fortes sécheresses et cultivées de façon extensive (en particulier le mil). Actuellement, on dénombre environ 7500 espèces  $C_4$  qui appartiennent à 18 familles de plantes (Sage 2004). Les plus importantes sont les Graminées ( $\approx 4600$  sp.), les Cypéracées ( $\approx 1300$  sp.) et les Amarantacées ( $\approx 800$  sp.). Seules les Graminées et les Cypéracées  $C_4$  dominent certains écosystèmes tropicaux, alors que les dicots  $C_4$  sont généralement caractéristiques de milieux ouverts fortement perturbés (e.g. sites cultivés, bords de route) ou de milieux extrêmement arides (Muhaidat *et al.* 2007).

La photosynthèse  $C_4$  est apparue plusieurs fois dans certaines familles de plantes et en particulier chez les Graminées (au moins dix fois ; Kellogg 2001), les Cypéracées (au moins quatre fois ; Soros & Bruhl 2000) et les Amarantacées (au moins 13 fois ; Kadereit *et al.* 2003 ; Sage *et al.* 2007). Ces phénomènes de convergence évolutive, qui comptent parmi les plus remarquables (Conway-Morris 2006), sont particulièrement intrigants du fait de la complexité du système  $C_4$  (Fig. 10). La multiplicité de l'apparition du caractère dans certaines familles laisse également supposer que certaines lignées de plantes étaient prédisposées à l'acquérir, d'un point de vue génétique et/ou écologique. Les taxa  $C_4$  sont presque exclusivement des plantes herbacées et très souvent des espèces à temps de génération court qui peuvent ainsi évoluer relativement rapidement (Sage 2001). Les gènes codant les enzymes  $C_4$  appartiennent presque tous à des familles multigéniques, et Monson (2003) a proposé que des gènes dupliqués ont été recrutés pour le système  $C_4$  via néofonctionnalisation. Les lignées de plantes ayant subi des événements récurrents de polyploïdisation (telles les Graminées ; Levy & Feldman 2002) auraient ainsi augmenté leur probabilité d'acquérir ce système. Par ailleurs, les caractérisations morphologique et biochimique d'espèces intermédiaires «  $C_3$ - $C_4$  » représentatives de plusieurs lignées  $C_4$  relativement récentes (e.g. *Flaveria*) indiquent que les modifications requises pour acquérir le système  $C_4$  suivraient un ordre bien précis, et plusieurs stades intermédiaires semblent évolutivement stables (Sage 2004). L'incrément du système aurait ainsi suivi des voies évolutives similaires chez les différentes lignées  $C_4$ .

L'apparition indépendante de la photosynthèse  $C_4$  chez de nombreuses lignées a également abouti à une relativement grande variation du système tant aux niveaux morphologique que biochimique. Ainsi plus de 15 types différents d'anatomie de Kranz ont été reportées dans la littérature (e.g. Soros & Dengler 2001 ; Kadereit *et al.* 2003) et même trois espèces  $C_4$  sans anatomie de Kranz ont été décrites (Edwards *et al.* 2004). Au niveau biochimique, trois principaux sous-types  $C_4$  sont actuellement reconnus selon l'enzyme impliquée dans la décarboxylation (libération du  $CO_2$  dans les chloroplastes des cellules de la gaine péri-vasculaire ; Fig. 9) : NADP-ME (le plus fréquent), NAD-ME et PCK (le plus rare et uniquement reporté chez deux sous-familles de Graminées). Cependant, des changements de sous-type  $C_4$

sont supposés avoir eu lieu durant l'évolution, et certaines espèces sont capables d'utiliser plusieurs sous-types en même temps (Sinha & Kellogg 1996 ; Giussani *et al.* 2001 ; Muhaidat *et al.* 2007). La valeur adaptative de ces variations n'est pas bien comprise à l'heure actuelle. Pour finir, on peut également mentionner le cas de certains *Eleocharis* (Cypéracées) qui sont capables d'exprimer ou non le système C<sub>4</sub> selon l'environnement (Ueno 2001).

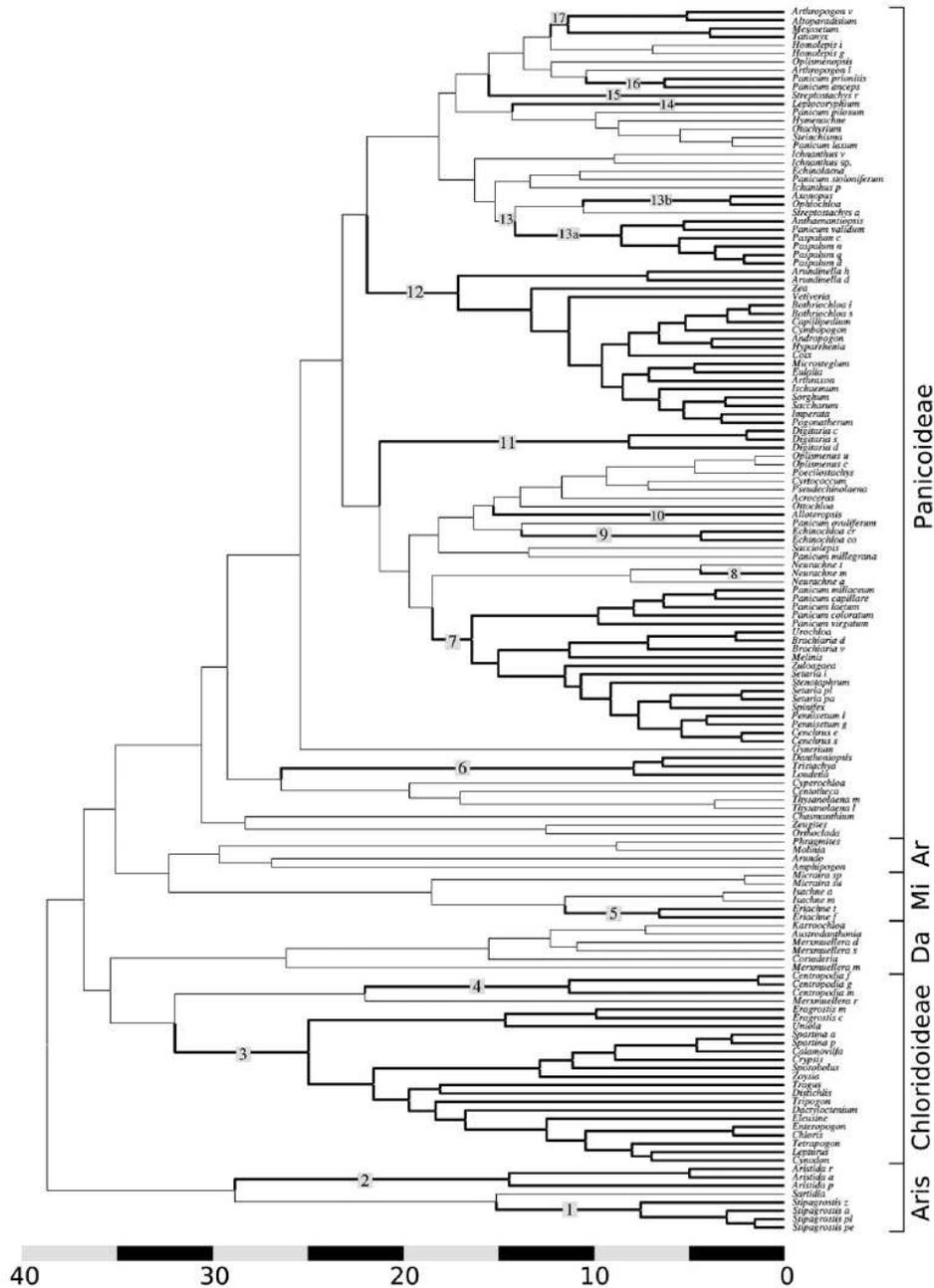
Depuis plus de 40 ans, la photosynthèse C<sub>4</sub> est un sujet de recherche très intense, et en particulier de nombreux auteurs se sont activement intéressés à des aspects physiologiques et biochimiques du système C<sub>4</sub>, à l'évolution moléculaire des gènes codant les enzymes C<sub>4</sub> ou encore à l'importance des espèces C<sub>4</sub> dans divers écosystèmes. De nombreuses études moléculaires se sont focalisées sur un genre d'Astéracées, *Flaveria* (e.g. Svensson *et al.* 2003 ; Westhoff & Gowik 2004), qui aurait acquis son système C<sub>4</sub> relativement récemment (moins de cinq Ma ; Sage 2004). Les Graminées et Cypéracées ont curieusement été quelque peu délaissées, car certains auteurs ont considéré que l'ancienneté du système C<sub>4</sub> dans ces familles est un désavantage pour des études d'évolution moléculaire. Toutefois, la multiplicité de l'apparition du système dans ces familles peut permettre de nombreuses comparaisons entre des lignées relativement apparentées. Toutes ces raisons m'ont poussé à monter un projet de recherche sur l'évolution de la photosynthèse C<sub>4</sub> chez les Poales lorsque je suis arrivé à l'université de Lausanne pour un poste de moyenne durée (cinq ans ; en tant que 1<sup>er</sup> assistant).

### 3.2. Taxonomie, datation moléculaire et origines multiples de la photosynthèse C<sub>4</sub> chez les Poales

Tout d'abord, nous avons reconstruit une phylogénie de l'ordre des Poales à l'aide de deux marqueurs plastidiques [*ndhF* (≈ 2200 bp) et *rbcL* (≈ 1400 bp)]. Sur la base de la littérature, un échantillonnage de Graminées et de Cypéracées le plus exhaustif possible a été tout d'abord constitué (avec l'aide de nombreux collègues). Nous avons inclus toutes les lignées C<sub>4</sub> connues et un maximum de leurs espèces sœurs C<sub>3</sub>. Une phylogénie a été reconstruite à l'aide d'une méthode Bayésienne (Ronquist & Huelsenbeck 2003) puis la topologie obtenue a été utilisée pour une datation moléculaire (Thorne *et al.* 1998 ; Thorne & Kishino 2001). Une fois l'âge des lignées C<sub>4</sub> obtenu, l'impact des concentrations passées en CO<sub>2</sub> atmosphérique (Pagani *et al.* 2005) sur l'induction du caractère C<sub>4</sub> a été testé à l'aide d'une méthode de maximum de vraisemblance développée par N. SALAMIN.

La topologie de l'arbre obtenu était tout à fait cohérente par rapport à des phylogénies récemment publiés sur les Graminées (Duvall *et al.* 2007) ou les Cypéracées (Simpson *et al.* 2007). Un minimum de 17 apparitions indépendantes du système C<sub>4</sub> était observé chez les Graminées (dans le clade des PACMAD, Fig. 11 ; Christin *et al.* 2008a) tandis qu'au moins six lignées C<sub>4</sub> étaient révélées chez les Cypéracées (Besnard *et al.* 2009b). Les lignées les plus anciennes appartiennent à la famille des Graminées avec entre autres la sous-famille Chloridoideae (32-25 Millions d'années ; Ma), le genre *Aristida* (29-14 Ma), la tribu Andropogoneae (22-17 Ma) et un clade de la tribu Paniceae (18.5-16.5 Ma). Les lignées de Cypéracées sont globalement plus récentes et la plus ancienne correspond au genre *Bulbostylis* (20-10 Ma). Basé sur les estimations de l'âge des lignées C<sub>4</sub> de Graminées, nous avons démontré que la probabilité de transition entre C<sub>3</sub> et C<sub>4</sub> a fortement augmenté après une baisse importante du CO<sub>2</sub> dans l'atmosphère il y a environ 30 Ma (Pagani *et al.* 2005). Ce résultat soutient que la faible teneur en CO<sub>2</sub> serait un des principaux facteurs favorisant l'apparition du système C<sub>4</sub>. Toutefois, nos résultats attestent également que la majorité des lignées (17/24) serait apparu bien

plus tard, après 15 Ma. Ceci suggère que bien d'autres facteurs sont impliqués (Osborne 2008). En particulier, l'augmentation de température et la perturbation grandissante des espaces ouverts tropicaux, du fait de sécheresses saisonnières, de feux récurrents empêchant la régénération des forêts, ainsi que l'apparition des grands herbivores, devraient avoir mené à la dominance des écosystèmes C<sub>4</sub> vers 5-8 Ma (i.e. Cerling *et al.* 1997), longtemps après les premières inductions du syndrome C<sub>4</sub>.



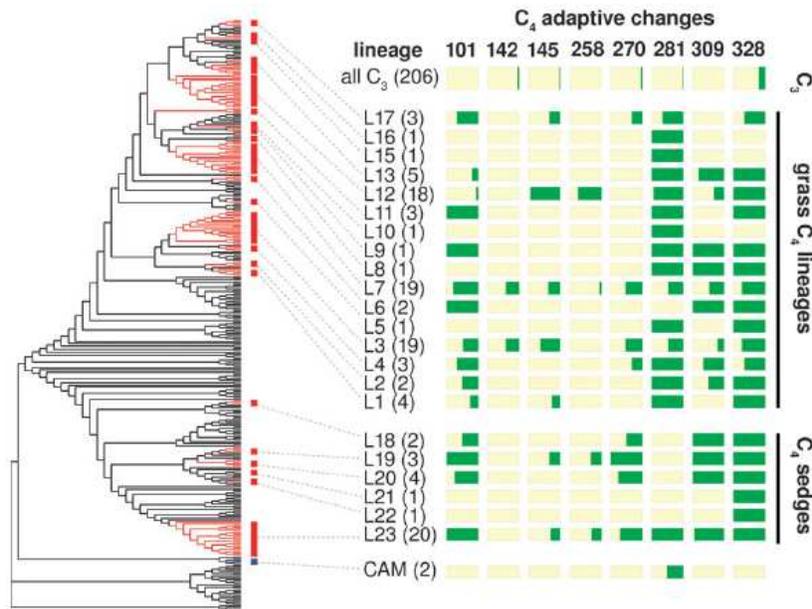
**Fig. 11.** Arbre phylogénétique calibré du clade des PACMAD [Paniceae, Arundinoideae (= Ar), Chloridoideae, Micrairoideae (= Mi), Aristidoideae (=Aris) & Danthonioideae (= Da)]. La longueur des branches est proportionnelle au temps et l'échelle est donnée en millions d'années. Les branches des 17 lignées C<sub>4</sub> sont en gras.

### 3.3. Evidence de sélection positive parallèle sur *rbcL* suite à l'acquisition du système C<sub>4</sub>

Basé sur nos données précédentes, nous avons ensuite testé l'action de la sélection positive sur *rbcL*, le gène qui code pour la grande sous-unité de la Rubisco. En effet, plusieurs éléments suggéraient que ce gène a subi de nouvelles pressions de sélection après l'acquisition du système C<sub>4</sub>. Tout d'abord, il est clairement établi que les Rubiscos d'espèces C<sub>4</sub> ont une activité catalytique supérieure et une affinité pour le CO<sub>2</sub> plus faible que les Rubiscos C<sub>3</sub> (e.g. Yeoh *et al.* 1980, 1981). Ceci peut s'expliquer par le fait que la teneur en CO<sub>2</sub> autour de la Rubisco est plus élevée chez les espèces C<sub>4</sub> offrant la possibilité d'augmenter l'activité catalytique de l'enzyme au dépend d'un abaissement de l'affinité pour le CO<sub>2</sub> (Tcherkez *et al.* 2006). De plus, il a été montré que ces caractéristiques C<sub>4</sub> sont maternellement transmises dans des croisements réciproques C<sub>3</sub> x C<sub>4</sub> (chez *Flaveria* et *Atriplex* ; Hudson *et al.* 1990) suggérant que le déterminisme génétique est plastidique. Le gène plastidique *rbcL* peut donc être légitimement considéré comme responsable de ces différences.

Basé sur ces constats, nous avons testé que le gène *rbcL* a subi de nouvelles pressions sélectives après l'acquisition du système C<sub>4</sub>. Différents modèles basés sur le ratio entre substitutions non-synonymes et synonymes ont été utilisés ( $d_N/d_S$ ; Yang *et al.* 2000 ; Yang & Nielsen 2002 ; Yang *et al.* 2005 ; Zhang *et al.* 2005). Le modèle de sélection positive ( $d_N/d_S > 1$ ) dans les branches C<sub>4</sub> était le meilleur modèle pour expliquer l'évolution du gène *rbcL* dans notre échantillon de plantes. Par ailleurs, huit codons sous sélection positive étaient identifiés avec des probabilités postérieures (pp) supérieures à 0.999. Cinq d'entre eux (e.g. 101, 258, 270, 281 et 309) étaient sous sélection purifiante dans les branches C<sub>3</sub> tandis qu'ils étaient sous sélection positive dans les branches C<sub>4</sub>. Des changements parallèles à ces huit codons étaient ainsi mis en évidence dans les différentes lignées C<sub>4</sub> étudiées (Fig. 12 ; Christin *et al.* 2008b). Considérant ces résultats, on peut aussi se demander si le fait que le gène *rbcL* soit sous sélection positive pourrait être un problème pour reconstruire un arbre des espèces (cf. partie précédente 3.2). En éliminant les huit codons sous sélection, les reconstructions phylogénétiques étaient refaites et ont mené à des topologies quasi identiques. Ce faible biais est notamment dû au fait que *rbcL* était combiné avec *ndhF*, qui est plus long et beaucoup plus variable (donc plus informatif).

Les sites de *rbcL* sous sélection positive codent probablement pour des acides aminés responsables des différences catalytiques observées entre Rubiscos C<sub>3</sub> et C<sub>4</sub>. Les propriétés biochimiques de ces changements mis en évidence devraient être caractérisées afin de le confirmer. Des applications biotechnologiques pourraient découler de ces résultats. En effet, les changements observés peuvent être considérés comme une adaptation à une plus forte concentration en CO<sub>2</sub> autour de la Rubisco. Dans cette optique, puisque les conditions atmosphériques futures devraient être caractérisées par une relativement forte concentration en CO<sub>2</sub> (de 2 à 3 fois plus élevées que celles d'avant la révolution industrielle ; Ainsworth *et al.* 2008), les pressions de sélection sur les Rubiscos C<sub>3</sub> pourraient être sensiblement modifiées donnant l'opportunité d'augmenter l'activité catalytique de l'enzyme au dépend d'une plus faible affinité pour le CO<sub>2</sub> (comme mis en évidence chez les plantes C<sub>4</sub>).



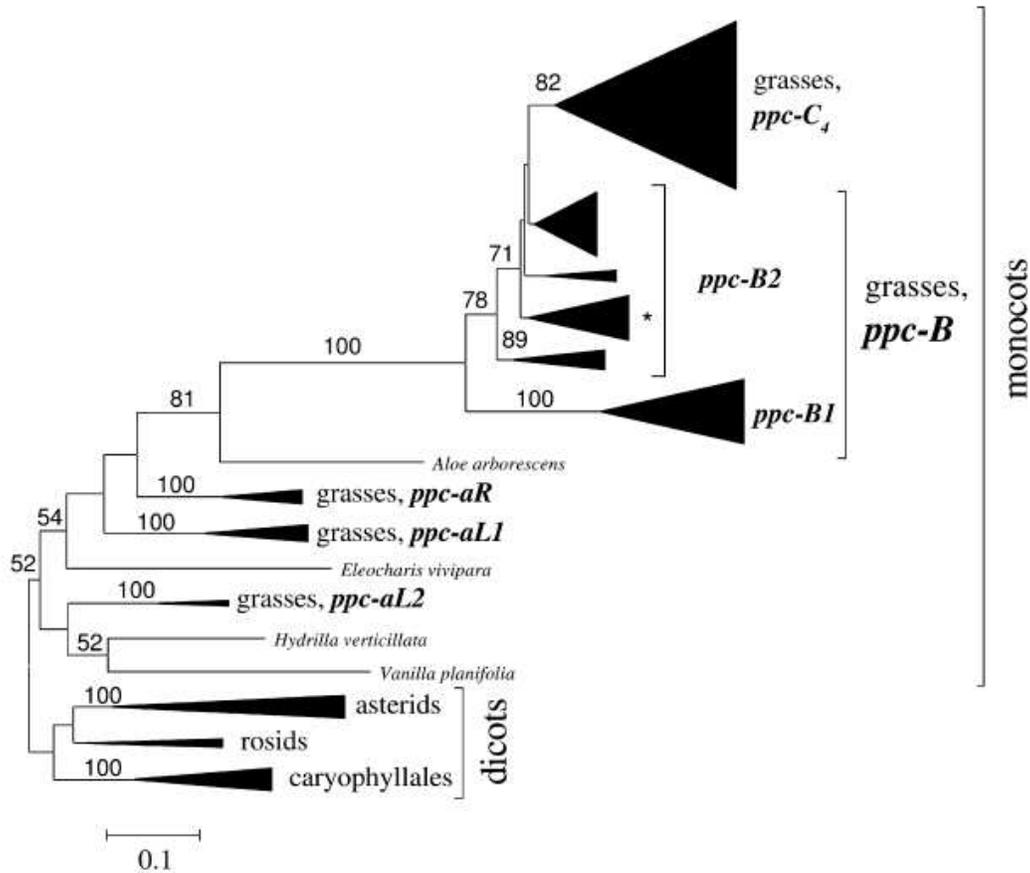
**Fig 12.** Codons du gène *rbcL* sous sélection positive dans les différentes lignées C<sub>4</sub> (selon Christin *et al.* 2008b). L'arbre phylogénétique des Poales (à gauche) est basé sur les gènes plastidiques *rbcL* et *ndhF* et reconstruit à l'aide d'une inférence Bayésienne. Les branches de 23 lignées C<sub>4</sub> sont représentées en rouge. Les changements d'acides aminés « caractéristiques » des lignées C<sub>4</sub> sont représentés à droite en vert. La fréquence de chaque changement est représentée pour chaque lignée. Le nombre d'espèces considérées pour chaque lignée est donné entre parenthèse.

### 3.4. Caractérisation de la famille multigénique codant les PEPC : évidence de convergences moléculaires

Toutes les enzymes recrutées dans le système C<sub>4</sub> préexistent dans les espèces C<sub>3</sub> (Monson 2003). Cependant, des modifications du promoteur sont nécessaires pour une forte transcription dans les cellules requises (cellules du mésophile ou de la gaine périvasculaire). De plus, des changements vont également affecter les propriétés catalytiques des enzymes C<sub>4</sub> car l'environnement biochimique est considérablement modifié (Svensson *et al.* 2003). Afin d'étudier des bases moléculaires de l'évolution de la photosynthèse C<sub>4</sub>, et en particulier les changements opérants au niveau de la séquence codante, nous avons caractérisé plusieurs gènes nucléaires codant des enzymes C<sub>4</sub>. Ces travaux ont concerné les familles multigéniques codant les PEPC, les NADP-MDH, les NADP-ME et les PCK (Fig. 9 ; Christin 2008). Des reconstructions phylogénétiques suivies de test de sélection ont été utilisées dans chaque cas. Je ne présenterai ici que les résultats obtenus sur la famille de gènes qui codent les PEPC, mais des résultats assez similaires ont été observés pour les différentes familles de gènes étudiées (Rondeau *et al.* 2005 ; Christin 2008).

Mes premières études sur les gènes de PEPC de Graminées ont notamment permis de montrer que le gène *ppc-C<sub>4</sub>* (qui code la PEPC C<sub>4</sub>) a accumulé 6 à 12 fois plus de substitutions non synonymes que les autres membres de cette famille multigénique (Besnard *et al.* 2002). Ceci pourrait être dû à une relaxation de la sélection purifiante mais aussi à une sélection positive sur le gène *ppc-C<sub>4</sub>*. Par ailleurs, ces travaux préliminaires n'avaient permis d'isoler que des fragments d'ADN complémentaires (ADNc) codant des PEPC C<sub>4</sub>. A Lausanne, nous avons donc

entrepris d'isoler des fragments d'ADN génomique de différents gènes *ppc* à la fois chez les espèces C<sub>3</sub> et C<sub>4</sub>. Avec ce jeu de données, nous avons ensuite utilisé des reconstructions phylogénétiques ainsi que des tests de sélection positive pour inférer l'histoire évolutive de cette famille de gènes en relation avec l'acquisition du système C<sub>4</sub>. Ce travail était effectué à la fois sur les Graminées et les Cypéracées. Les gènes codant des enzymes C<sub>4</sub> pouvaient être identifiés *a posteriori* par la présence d'un acide aminé particulier (Ser<sup>780</sup>) responsable d'une grande partie des propriétés C<sub>4</sub> de l'enzyme, chez des lignées de plantes très divergentes appartenant aux dicots et aux monocots (Svensson *et al.* 2003).

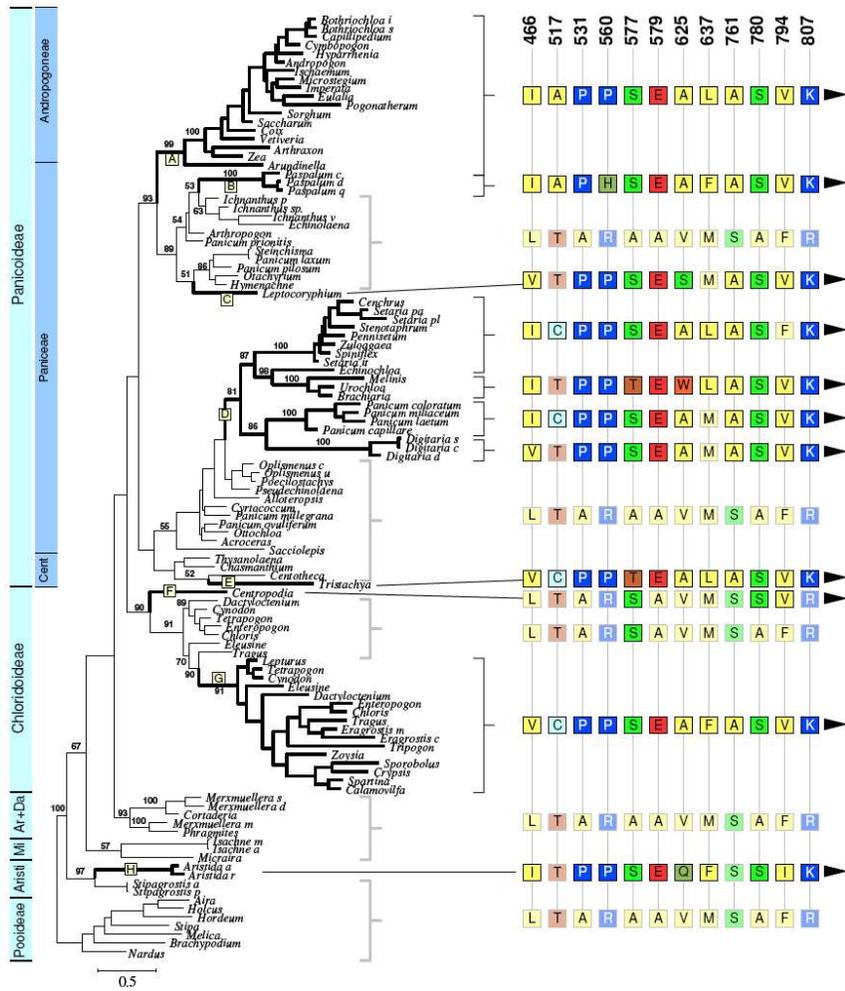


**Fig. 13.** Arbre phylogénétique (basé sur la séquence codante) de la famille multigénique codant les PEPC d'Angiospermes reconstruit à l'aide d'une méthode maximum de vraisemblance (Christin *et al.* 2007b). Tous les gènes de *ppc-C<sub>4</sub>* de Graminées forment un clade monophylétique (excepté le gène de *Centropodia forskalii* dont la position est indiquée par \*).

Chez les Graminées, cinq gènes codant des PEPC ont été trouvés dans le génome du riz (*ppc-B*, *ppc-aL1a*, *ppc-aL1b*, *ppc-aL2*, *ppc-aR* ; Besnard *et al.* 2003a ; Christin *et al.* 2007a). Un sixième gène (appelé *ppc-B2* ; absent chez le riz) était mis en évidence durant nos investigations chez un grand nombre d'espèces C<sub>3</sub> de Graminées (Christin *et al.* 2007b). Nos analyses phylogénétiques suggèrent une duplication du gène *ppc-B* (générant les gènes *ppc-B1* et *ppc-B2*) avant la diversification des Graminées. Certaines espèces, en particulier, les Pooideae ou certaines PACMAD C<sub>3</sub>, possèdent encore ces deux gènes tandis qu'un grand nombre semble en avoir perdu une copie. Un gène *ppc-C<sub>4</sub>* (i.e. Ser<sup>780</sup>) était identifié dans toutes les espèces C<sub>4</sub>

étudiées. Nos analyses phylogénétiques ont montré que la plupart de ces gènes ont probablement dérivé de *ppc-B2* (Fig. 13). En fait, une seule exception était récemment mise en évidence. Le gène de la PEPC  $C_4$  du genre *Stipagrostis* a dérivé du gène *ppc-aLIB* (Christin & Besnard 2010), mais je ne m'attarderai pas sur ce cas bien qu'il soit vraiment intéressant (et une preuve irréfutable de deux apparitions indépendantes de la photosynthèse  $C_4$  dans la sous-famille Aristidoideae). Dans l'arbre phylogénétique basé sur la séquence codante (Fig. 13), un clade renfermant presque tous les gènes *ppc-C4* (sauf celui de *Centropodia*) était observé. Les relations phylogénétiques entre tribus et sous-familles étaient particulièrement inattendues et assez incohérentes par rapport aux arbres d'espèces connus des Graminées (Christin *et al.* 2007b). Nos analyses phylogénétiques ont donc ensuite été approfondies et ont focalisé sur le clade *ppc-B2/ppc-C4*. Nous avons reconstruit un arbre sur les introns et les 3<sup>ème</sup> position des codons (car supposés moins soumis à sélection). Une topologie très différente était alors obtenue (Fig. 14). Cette fois les relations phylogénétiques entre tribus et sous-familles de Graminées étaient tout à fait celles attendue *a priori*. Des apparitions multiples du gène *ppc-C4* à partir de *ppc-B2* sont désormais fortement suggérées par cette topologie. Huit clades  $C_4$  sont ainsi présents [appartenant à dix lignées  $C_4$  sur la base des données chloroplastiques (Christin *et al.* 2007b)]. Les tests basés sur le ratio entre substitutions non-synonymes et synonymes ( $d_N/d_S$ ) supportent le fait que des événements parallèles de sélection positive ont eu lieu à la base de ces huit clades. Des codons sous sélection positive étaient identifiées (21 pour  $pp > 0.95$ , et 12 pour  $pp > 0.999$ ) et des changements convergents étaient mis en évidence (Fig. 14). Cette forte sélection positive sur le gène *ppc-C4* explique le biais phylogénétique sur la topologie présentée en Fig. 13.

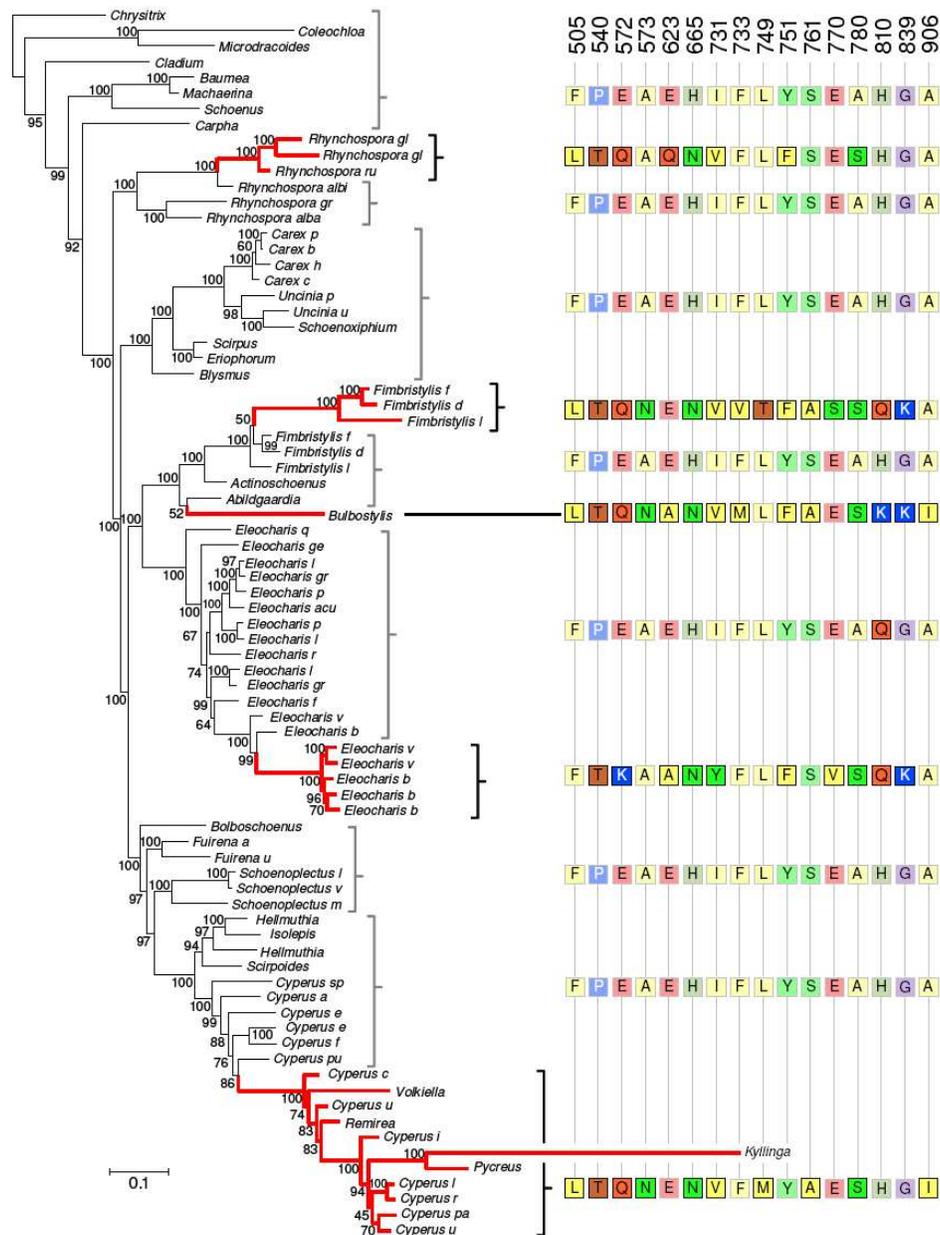
Chez les Cyperacées, le cadre était assez différent car seulement une séquence codante de PEPC (PEPC  $C_4$  d'*Eleocharis vivipara*) était connue au début de ce travail. Avec un étudiant de master (Flavien RUSSIER), nous avons toutefois été capables de résoudre ce problème et générer un segment de gène de PEPC pour un échantillonnage exhaustif d'espèces  $C_3$  et  $C_4$  de la famille des Cyperacées (Besnard *et al.* 2009b). Là encore, au moins un fragment de gène  $C_4$  était isolé chez toutes les espèces  $C_4$  étudiées. Chez certaines espèces, plusieurs séquences de gènes de PEPC étaient isolées (notamment dans les genres *Eleocharis*, *Fimbristylis* ou certaines Cyperaceae). Environ 80 séquences étaient isolées et formaient un clade monophylétique au sein des Poales (Besnard *et al.* 2009b). Sur ce jeu de données, la topologie obtenue (Fig. 15) était très cohérente par rapport à celle observée sur l'ADNcp suggérant peu de biais phylogénétiques dus à la sélection. Dans l'arbre, cinq lignées de gènes  $C_4$  étaient observées et la grande différence avec l'arbre ADNcp réside dans le fait que *Eleocharis vivipara* et *E. baldwinii* possèdent un gène de PEPC  $C_4$  très similaire (ce qui est aussi fortement soutenu par les introns) alors qu'ils sont très clairement séparés dans les arbres basés sur l'ADNcp ou l'ITS (Roalson and Friar 2000). Ceci suggère que l'une de ces deux espèces pourrait avoir acquis ce gène  $C_4$  par hybridation. Ces deux espèces sont d'ailleurs en sympatrie en Floride, et *E. vivipara* présente une grande diversité dans l'expression de son caractère  $C_4$  (Murphy *et al.* 2007). Des études plus approfondies sont nécessaires pour mieux comprendre ce phénomène. Par ailleurs, les tests basés sur le ratio  $d_N/d_S$  montrent que le gène a évolué sous sélection positive à la base des cinq clades correspondant aux gènes de PEPC  $C_4$ . Seize codons sous sélection positive étaient identifiés ( $pp > 0.999$  ; Fig. 15). Cinq d'entre eux sont également sous sélection chez les Graminées, et les mêmes changements d'acides aminés sont observés sur les isoformes  $C_4$  (Tableau 1). Plusieurs de ces changements étaient aussi observés chez les Dicots et sur le gène de PEPC  $C_4$  de *Stipagrostis* (Tableau 1).



**Fig. 14.** Arbre phylogénétique des gènes *ppc-B2* (branches normales ; Ala<sup>780</sup>) et *ppc-C4* (branches en gras ; Ser<sup>780</sup>) basé sur les 3<sup>ème</sup> bases des codons et les introns, et reconstruit à l'aide d'une inférence Bayésienne (Christin *et al.* 2007b). A droite, les acides aminés résultants des codons détectés sous sélection positive (pp > 0.999) sont représentés. Abrév. : Arist = Aristidoideae ; Mi = Micrairoideae ; Ar = Arundinoideae ; Da = Danthonioideae ; Cent = Centothecaceae.

**Tableau 1.** Changements d'acides aminés codés par les cinq codons sous sélection positive (pp > 0.95) chez les Graminées et les Cypéracées. Le nombre d'événements pour chaque changement est indiqué pour chaque groupe et chaque codon. N = nombre de lignées C<sub>4</sub> considérées.

Groupe	N	No codon				
		572	665	733	761	780
Graminées	8	E → Q: 7	H → N: 6 H → R: 1	F → V: 3 F → M: 2	S → A: 6	A → S: 8
Cypéracées	5	E → Q: 4 E → K: 1	H → N: 5	F → V: 1 F → M: 1	S → A: 3	A → S: 5
Dicots	5	E → Q: 1 E → K: 1	H → N: 3		S → A: 3	A → S: 5
<i>Stipagrostis</i>	1	E → Q: 1	H → N: 1			A → S: 1



**Fig. 15.** Arbre phylogénétique basé sur la séquence codante des gènes de PEPC isolées chez les Cypéracées et reconstruit à l'aide d'une inférence Bayésienne (Besnard *et al.* 2009b). Les branches des gènes non C<sub>4</sub> (Ala<sup>780</sup>) sont en noire tandis que celles des gènes C<sub>4</sub> (Ser<sup>780</sup>) sont en rouge. A droite, les acides aminés résultants des codons détectés sous sélection positive (pp > 0.999) sont représentés.

L'ensemble de ces investigations a très clairement apporté des évidences moléculaires à l'apparition multiple du système C<sub>4</sub> chez les Poales. De plus, nous avons développé une méthodologie assez originale (i.e. test de sélection appliqué sur plusieurs phylums indépendants) pour identifier les codons sous sélection, qui ont probablement une grande importance dans l'acquisition des propriétés catalytiques des enzymes C<sub>4</sub>. Leur identification ouvre certes d'importantes perspectives biotechnologiques, mais il reste à déterminer quelles propriétés

biochimiques sont apportées par ces changements d'acides aminés, ce qui pour le moment dépasse le cadre de mes compétences. Chez les espèces exprimant des niveaux variables du trait  $C_4$  (par exemples *Eleocharis vivipara* ou *Alloteropsis semialata*), il serait également très intéressant de tester si il existe une variation aux sites sous sélection, et le cas échéant, si il est possible de la relier à des variables environnementales dans les populations naturelles d'origine.

Par ailleurs, il est intéressant de noter qu'un même gène peut être indépendamment recruté pour la photosynthèse  $C_4$  de nombreuses fois, que ce soit chez les Cypéracées ou les Graminées (en particulier ici pour les gènes de PEPC  $C_4$ ). Nous pouvons proposer que la fonction originale du gène, ainsi que son mode de transcription et la localisation cellulaire de l'enzyme codée, sont probablement des facteurs prépondérants dans la possibilité d'acquérir ou non une fonction  $C_4$ . Des duplications géniques associées au recrutement d'un gène pour la fonction  $C_4$  ont aussi été révélées (e.g. pour les gènes de PEPC chez *Eleocharis*, *Fimbristylis* et les Chloridoideae) confirmant que ce phénomène joue probablement un rôle important (Monson 2003). Par ailleurs, de fortes convergences ont été observées sur l'évolution des gènes codant les enzymes  $C_4$ , et tout particulièrement lorsque les copies recrutées sont à l'origine très semblables (e.g. appartenant à une même lignée génique comme *ppc-B2*). Il est très vraisemblable qu'il existe des contraintes évolutives pour améliorer une protéine (Weinreich *et al.* 2006). Dans une famille de plantes donnée, le recrutement récurrent d'un même gène expliquerait ainsi les fortes convergences que nous avons mises en évidence. Toutefois, lorsque les gènes recrutés sont à l'origine divergents et probablement impliqués dans des fonctions assez différentes (e.g. le cas des gènes de PEPC des Cypéracées, *Stipagrostis* et autres Graminées), les phénomènes de convergence sont beaucoup moins marqués.

### 3.5. La photosynthèse $C_4$ et les changements globaux : quelles conséquences pour l'amélioration des plantes ?

L'importante augmentation de la teneur en  $CO_2$  atmosphérique attendue durant les prochaines décennies est l'un des facteurs anthropiques actuels les plus préoccupants, car il devrait induire des changements environnementaux majeurs par le biais, notamment, d'une forte augmentation de la température moyenne sur terre (e.g. Williams *et al.* 2007). Par ailleurs, il devrait également minimiser l'importance de la photorespiration, et des études récentes suggèrent que les savanes africaines ou certaines prairies nord américaines (principalement constituées de graminées  $C_4$ ) pourraient être recolonisées par des espèces arbustives  $C_3$  (Morgan *et al.* 2008). Cela signifie que de fortes perturbations dans les chaînes trophiques de ces écosystèmes peuvent être attendues dans les prochaines décennies. Toutefois, tout cela ne signifie par pour autant que les plantes  $C_4$  seront inévitablement moins compétitives. En effet, les plantes  $C_4$  sont particulièrement fréquentes en milieux ouverts où la photorespiration est accentuée, et les activités humaines contribuent largement à leur succès par la création de tels espaces par le biais de la déforestation en milieu tropical (Sage & Kubien 2003). Les plantes  $C_4$  à forte biomasse (c.a.d. forte surface d'évapotranspiration) resteront compétitives dans des conditions bien particulières et tout particulièrement lorsqu'elles sont cultivées de façon intensive. Ainsi des cultures tropicales, telles la canne à sucre ou le maïs, devraient rester performantes car leur système  $C_4$  leur confère une meilleure efficacité de l'utilisation de l'eau mais aussi de l'azote (notamment par des besoins fortement réduits en Rubisco) par rapport aux espèces  $C_3$  (e.g. Ghannoum *et al.* 2005). Toutefois, les plantes  $C_3$  sont naturellement capables de réduire très significativement l'expression de la Rubisco en présence d'une plus forte concentration en  $CO_2$ , ce qui leur permet d'économiser une

partie de l'azote disponible (jusqu'à 6% pour les concentrations en CO<sub>2</sub> attendues en 2050 ; Ainsworth *et al.* 2008).

Nos résultats sur les gènes codant la large sous-unité de la Rubisco et certaines enzymes C<sub>4</sub> pourraient avoir des applications biotechnologiques dans le futur afin d'augmenter les performances des cultures, en particulier pour transformer les espèces C<sub>3</sub> (Ainsworth *et al.* 2008 ; Hibberd *et al.* 2008). Toutefois, il y a encore une grande controverse sur ce sujet car il n'est pas évident que ce genre de travail soit toujours justifié et les résultats obtenus jusqu'ici ce sont avérés très décevants. Transformer une plante C<sub>3</sub> en C<sub>4</sub> n'est pas encore possible, car manipuler un caractère aussi complexe (avec notamment le problème de la compartimentation des fonctions de fixation du CO<sub>2</sub> et de photosynthèse) n'est pas encore maîtrisée (Hibberd *et al.* 2008 ; Christin *et al.* 2009b). Les résultats accumulés dans nos études ont toutefois apporté des connaissances potentiellement très importantes qui devraient à terme permettre de modifier les enzymes, par exemple afin de moduler leur activité catalytique et/ou l'affinité pour leur substrat. Toutefois, comme précédemment mentionné, des études biochimiques préalables sont encore nécessaires pour bien déterminer le rôle des acides aminés sous sélection positive dans les espèces C<sub>4</sub>. Chez les espèces C<sub>3</sub>, l'obtention d'une Rubisco avec une activité adaptée à une forte concentration en CO<sub>2</sub> (en s'inspirant de nos résultats) devrait permettre en outre d'encore réduire les besoins en Rubisco des plantes, induisant une économie d'azote et donc potentiellement une augmentation de productivité. Ce type de mutations devrait également apparaître naturellement dans les populations (de façon stochastique). L'impact d'un tel phénomène sur la compétitivité interspécifique et notamment la perturbation de l'équilibre des écosystèmes mériterait aussi d'être étudié dans les prochaines décennies.

Pour finir, dans de prochaines investigations, il serait très intéressant de s'attarder sur les différents sous-types biochimiques C<sub>4</sub> (Christin *et al.* 2009a, b, c). Le sous-type PCK semble le plus intéressant si l'on considère son coût énergétique (Muhaidat *et al.* 2007), et il serait primordial de comprendre pourquoi il est si peu représenté dans les lignées C<sub>4</sub> actuelles. Certaines espèces sont capables de l'utiliser pour, semble t'il, compléter les sous-types NAD-ME ou NADP-ME (Muhaidat *et al.* 2007). Des désavantages ou des contraintes évolutives pourraient expliquer pourquoi le sous-type PCK n'est prépondérant que dans certaines espèces de deux sous-familles de Graminées, comme par exemples une sensibilité accrue à la sécheresse ou tout simplement une plus grande difficulté à mettre en place le système au niveau cellulaire (Fig. 9 ; Muhaidat *et al.* 2007). La signification écologique de ces différents sous-types C<sub>4</sub> est également très mal comprise mais elle pourrait être liée à des facteurs environnementaux tels la disponibilité en eau (Taub 2000 ; Carmo-Silva *et al.* 2007). Leur compréhension pourrait aussi permettre de nouvelles avancées dans l'amélioration du système C<sub>4</sub> de plantes cultivées (principalement de type NADP-ME) comme le maïs, le sorgho ou la canne à sucre.

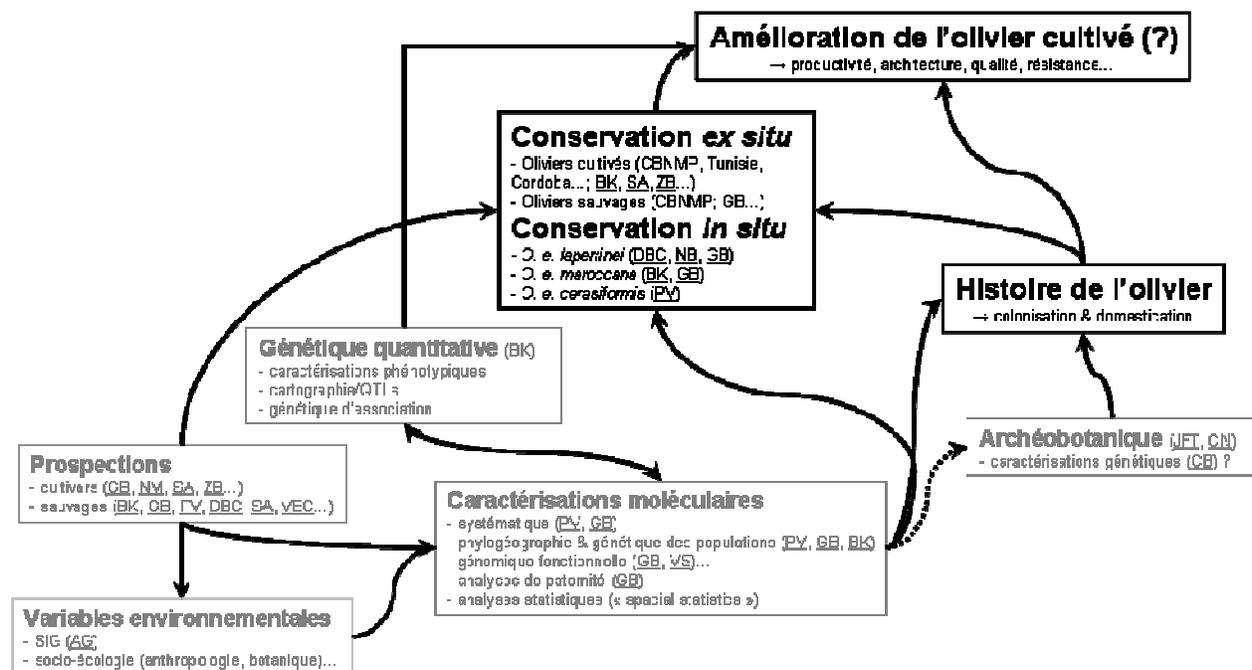
## 4. Perspectives de recherche

Mes différentes expériences dans le domaine de la recherche m'ont tout d'abord permis de travailler avec de nombreux collègues (e.g. j'ai cosigné des publications dans des journaux internationaux avec une centaine de personnes) et j'ai aussi pris beaucoup de plaisir à encadrer des étudiants en thèse et master. Cela m'a également donné plusieurs opportunités d'écrire et de mener des projets, impliquant des collaborations internationales, notamment avec des pays émergents (Algérie, Niger). Mes différents travaux entrepris ont généralement mené à plusieurs publications, ce qui démontre que j'ai aussi acquis de solides compétences pour gérer et/ou orienter les recherches sur des thèmes d'intérêt scientifique. Pendant mes activités, des sujets très variés utilisant des données génétiques ont été abordés, comme l'histoire de l'olivier et l'évolution de la photosynthèse  $C_4$ , et des avancées très significatives ont été obtenues sur ces deux sujets en particulier. L'ensemble de ces travaux ont bien sûr fortement contribué à m'ouvrir l'esprit sur des problématiques de recherche de plus en plus larges.

Après mes cinq années passées à Lausanne, je suis dans une phase de transition où il m'est clairement nécessaire de redéfinir mes objectifs de travail. Mes investigations sur des problématiques très différentes ont abouti à plus de 45 publications, mais il me semble désormais plus judicieux de recadrer mes recherches afin de me focaliser sur un projet mieux ciblé. J'ai choisi de travailler sur la génétique de l'olivier, ayant cette opportunité de reprendre ce sujet dans le cadre d'une bourse intra-européenne Marie-Curie à l'Imperial College dans l'équipe de V. SAVOLAINEN. Cette démarche a aussi pour but de me resituer sur une problématique de génomique sur une plante cultivée en milieu méditerranéen, un thème de recherche que je voudrais faire valoir dans l'optique de me stabiliser dans les années à venir. A cette occasion, j'aimerais combiner l'ensemble de mes compétences et mettre à profit celles de mes collaborateurs pour définir un projet ambitieux sur les ressources génétiques de l'olivier. A partir des résultats de nos précédentes investigations et l'évolution des connaissances sur cette espèce, il est désormais nécessaire de redéfinir les questions auxquelles je souhaiterais répondre durant les prochaines années :

- Tout d'abord, peut-on mieux comprendre les origines de l'olivier méditerranéen ? En particulier que représentent aujourd'hui les formes sauvages (i.e. oléastres) et sont elles un patrimoine génétique important en regard de l'olivier cultivé ? Peut-on identifier des populations reliques des glaciations passées ?
- Que doit-on conserver en collection en priorité : des cultivars ou également des formes sauvages ?
- Dans l'objectif de caractériser les collections, quels caractères phénotypiques d'intérêt faut-il étudier en priorité ?
- Quelles sont les bases génétiques de la domestication de l'olivier ? Comment faut-il étudier la variation des gènes d'intérêt potentiellement impliqués dans la valeur agronomique des individus ?
- Enfin, pour la conservation *ex-situ* de populations d'oliviers sauvages (en particulier pour subspp. *maroccana*, *cerasiformis* et *laperrinei*), quelle démarche faudra-il adopter dans les prochaines décennies ?

A travers ces questions, l'un des enjeux pour la prochaine décennie sera aussi d'intégrer les approches de génomique et de transcriptomique. Des collègues, en particulier dans le sud de l'Europe (Italie et Espagne), devraient évidemment développer et utiliser ce type d'approche. Il sera donc très important de réunir autour d'un objectif commun différents chercheurs aux compétences complémentaires, en génétique, en physiologie, en agronomie, en biologie de la conservation ou encore en écologie, afin d'optimiser l'impact de ce type d'étude (Fig. 16). Des collaborations internationales devront être initiées au travers de projets européens mais aussi d'autres projets internationaux tout particulièrement avec nos collègues maghrébins et d'autres pays émergents.



**Fig. 16.** Organigramme des différentes activités proposées sur la caractérisation des ressources génétiques de l'olivier. L'expertise de chaque collaborateur pour les différentes approches impliquées est indiquée. Abrév. des collaborateurs : AG = A. Guisan (UNIL), BK = B. Khadari (INRAM), DBL = D. Baali-Cherif (Univ. Alger), CN = C. Newton (Univ. Nottingham), GB = G. Besnard, JFT = J.F. Terral (Univ. Montpellier); NB = N. Bouguédoura (Univ. Alger), PV = P. Vargas (RJB Madrid), SA = S. Abdelhamid (Institut de l'Olivier, Sfax), VBC = V. Bronzini de Caraffa (Univ. Corte), VS = V. Savolainen (Imperial College), ZB = Z. Bouchera (Univ. Corte).

#### 4.1. Origines et domestication d'une espèce ligneuse méditerranéenne : l'olivier comme cas d'étude

##### 4.1.1. Analyses génétiques et génomiques des populations d'oléastres

Pour mieux comprendre les origines de l'olivier cultivé, il est tout d'abord nécessaire de bien cerner l'origine des populations d'oléastres. L'approche phylogéographique développée durant et après ma thèse ne nous a pas permis de trancher avec une absolue certitude sur le statut

réellement sauvage des populations d'oléastres alors prospectées. Il m'apparaît aujourd'hui possible de comparer les populations dans différentes situations et d'apporter de nouveaux éléments pour répondre à cette question.

D'un point de vue méthodologique, les démarches à suivre devront bien intégrer nos connaissances sur l'olivier, et en premier point l'échantillonnage devra être constitué de façon très méthodique. De nouvelles campagnes de prospection à l'ouest de la Méditerranée sont prochainement envisagées. Un échantillonnage conséquent devrait être établi avec nos partenaires (P. VARGAS en Espagne, B. KHADARI au Maroc, D. BAALI-CHERIF en Algérie, S. ABDELHAMID en Tunisie, V. BRUNINI en Corse et Sardaigne...). Comme précédemment abordé (partie 2.3.1), j'envisage de bien caractériser le milieu dans lequel chaque population d'oléastre sera (ou a été) collectée. Il faudra notamment énumérer les espèces associées (relevés de végétation) et prendre en compte la proximité des cultures. Des indices d'anthropisation pourraient être ainsi définis sur ces données. Les variables environnementales (SIG) devront être également considérées. L'ensemble de ces informations pourra ensuite être intégré dans certaines analyses des données génétiques (par exemple, par le biais d'une Analyse Canonique des Correspondances).

A partir de données génétiques, je propose d'utiliser des approches de coalescence pour affiner nos connaissances sur la recolonisation post-glaciaire de l'oléastre dans le bassin méditerranéen. Ce type d'approche s'appuyant sur des modèles de démographie devrait être une méthode alternative (et bien évidemment complémentaire) à l'analyse de structuration (e.g. inférence de populations panmictiques). De plus, cette démarche devrait potentiellement être généralisable à d'autres espèces cultivées méditerranéennes telles que la vigne ou le figuier sur lesquels certains de mes collaborateurs ou collègues travaillent (e.g. B. KHADARI, P. THIS ; INRA Montpellier). Des données des génomes cytoplasmiques (maternellement transmis) et nucléaires (transmission bi-parentale) devront être utilisées :

- Durant les dernières années, moins de 5% du génome chloroplastique (ADNcp) de l'olivier a été étudié pour rechercher du polymorphisme. Avec les nouvelles séquences d'ADNcp d'Oléacées (en particulier du Jasmin mais aussi d'olivier), disponibles dans les banques de données publiques, nous avons entrepris le séquençage complet de plusieurs génomes chloroplastiques d'olivier (à Londres) pour développer de nouveaux marqueurs (notamment en collaboration avec A. FOREST et P. VARGAS du RJB Madrid). Avec ces nouveaux outils, nous espérons pouvoir détecter de nombreux haplotypes permettant une mesure plus précise de la diversité génétique intra-population pour chaque lignée chloroplastique. Ceci permettrait d'inférer l'histoire démographique récente des populations : les populations ayant eu une récente expansion devraient être caractérisées par peu de diversité (mais potentiellement par des haplotypes phylogénétiquement divergents comme dans les populations australiennes, qui pourrait être maintenue par la présence de CMS ; Besnard *et al.* 2007b), ou au contraire, les populations plus stables et anciennes devraient avoir une plus forte diversité (avec des haplotypes phylogénétiquement apparentés). L'analyse d'un échantillonnage exhaustif de populations, en particulier de l'Ouest de la Méditerranée, permettra de caractériser l'organisation de la diversité sur une échelle géographique relativement large.

- Par ailleurs, l'utilisation de séquences de l'ADN nucléaire (beaucoup plus variable que l'ADNcp) est également envisageable. Cependant, la démarche est compliquée du fait de l'hétérozygotie. Il est nécessaire *a priori* d'isoler les haplotypes par clonage ce qui rend ce travail extrêmement fastidieux (e.g. gène FAD6 ; Besnard *et al.* en prep). Des investigations

préliminaires sont donc nécessaires. Par ailleurs, avec le développement du pyroséquençage (Margulies *et al.* 2005 ; Hudson 2008), il est fort probable que de nouvelles techniques pourront être développées d'ici quelques années. A Londres, nous envisageons sérieusement d'initier cette technologie afin de générer un grand nombre de données pour une nouvelle approche de génomique des populations. L'un des problèmes de cette technique reste encore son coût très élevé. Une méthodologie appropriée devra être également développée. La diversité nucléotidique pourra être mesurée sur plusieurs loci (au moins une 10-15<sup>aine</sup>) et la structure du déséquilibre de liaison entre substitutions seront également étudiées à chaque locus. Dans un premier temps, on pourra comparer l'organisation de la diversité sur des populations d'origines et de situations différentes (e.g. en conditions anthropisées ou non), afin de tester et de mesurer l'importance des effets anthropiques sur la structuration de la variation génétique.

Les patterns observés sur l'ADNcp et l'ADN nucléaire seront comparés afin d'évaluer la concordance de ces deux types d'information. Ensuite, à partir de ces données, des modèles estimant l'importance des flux de gènes devraient être utilisés (modèle ME, « isolement avec migration » ; Hey 2006) notamment pour tenter de mesurer l'importance des flux de gènes depuis les formes cultivées de l'Est dans les populations sauvages de l'Ouest de la Méditerranée. Des collaborations avec des spécialistes de génétique des populations (e.g. R. RUBIO DE CASAS, RJB Madrid) devront être mises en place. On pourra également utiliser des modèles de démographie pour tester différentes hypothèses de refuge ou non refuge sur les populations étudiées (e.g. Perron *et al.* 2000 ; Heuertz *et al.* 2006). Ceci devrait nous permettre de résoudre une partie de l'histoire post-glaciaire de l'oléastre en Méditerranée.

#### *4.1.2. Génétique de la domestication de l'olivier*

Sur le pool cultivé, une première étude relativement facile à mettre en œuvre (d'un point de vue méthodologique) est l'analyse de parenté entre cultivars. Sur la base de marqueurs microsatellites, d'anciens croisements entre clones cultivés ayant mené à de nouvelles variétés peuvent être mis en évidence (e.g. Bowers *et al.* 1999). Ces analyses devraient aussi permettre de repérer des cultivars qui ont un lien de parenté. Quelques exemples sont déjà répertoriés dans la littérature pour l'olivier (e.g. La Mantia *et al.* 2005). Ce type d'information permettra de mieux comprendre la sélection des cultivars mais aussi leur dissémination par l'homme à une échelle régionale (e.g. Vouillamoz *et al.* 2003). En outre, ce type d'étude pourra également avoir certaines applications pour orienter des croisements intéressants dans d'éventuels programmes de sélection. J'ai déjà généré quelques données en collaboration avec B. ZITOUN de l'université de Corte (12 loci microsatellites sur 72 cultivars provenant en particulier de Tunisie) mais ceci reste anecdotique. Avec des partenaires des différents pays méditerranéens, il serait intéressant de constituer une base de données sur un large échantillonnage provenant d'une vaste aire géographique. Des données de qualité irréprochable (ce qui n'est malheureusement pour le moment pas le cas ; e.g. Doveri *et al.* 2008 ; Rekik *et al.* 2008) ont besoin d'être générés pour un échantillonnage le plus exhaustif possible et un nombre de loci suffisant (au moins une 30<sup>aine</sup>). Le problème des synonymies et homonymies doit être aussi mentionné car il ne facilite pas ce type d'étude en apportant une certaine confusion dans les nomenclatures (e.g. Besnard *et al.* 2001d ; Khadari *et al.* 2003 ; La Mantia *et al.* 2005 ; Rekik *et al.* 2008). La relative fréquente apparition de mutations sur des allèles de type microsatellite (e.g. Baali-Cherif & Besnard 2005) devra aussi être prise en compte. Il y a déjà beaucoup de collègues sur cette thématique, et pour le moment je n'envisage de m'impliquer dans ce domaine que par le biais de collaborations (e.g. B. ZITOUN,

S. ABDELHAMID, B. KHADARI). Toutefois, les relations généalogiques entre cultivars devraient être des informations potentiellement très importantes pour la suite des investigations sur la domestication de l'olivier, et cet aspect ne devra donc pas être négligé.

Le syndrome de domestication de l'olivier (et plus généralement des espèces ligneuses multipliées par voie végétative) est à appréhender différemment de celui des espèces annuelles telles que le maïs, le tournesol ou encore le blé (Doebley *et al.* 2006), notamment parce que le nombre de générations impliquées est probablement très faible et les formes cultivées ne sont pas facilement distinguables des formes sauvages que ce soit sur la base de critères morphologiques ou génétiques. La domestication de l'olivier s'est basée sur la multiplication des individus les plus intéressants par voie végétative, ce qui permet de fixer un génotype donné sans avoir ensuite recours à la voie sexuée. Les caractères agronomiques sélectionnés par l'homme sont généralement assez complexes et à déterminisme multigénique ce qui rend difficile l'identification des facteurs génétiques impliqués. Les méthodes de génétique d'association (Gupta *et al.* 2005) peuvent permettre l'identification de zones du génome sous sélection, mais la finesse de la détection va dépendre du déséquilibre de liaison (LD ; ou association entre les polymorphismes physiquement liés sur le génome). L'olivier est anémophile et strictement allogame permettant un mélange très important des gènes (comparé à une espèce autogame par exemple). En conséquence, du fait de la recombinaison, on peut s'attendre à ce que le LD décroisse relativement rapidement avec la distance séparant les polymorphismes sur le génome (e.g. Flint-Garcia *et al.* 2003). Cependant, des phénomènes d'admixture relativement récents (e.g. Besnard *et al.* 2001 ; Baldoni *et al.* 2006 ; Breton *et al.* 2006 ; Belaj *et al.* 2007) devraient avoir contribué à une augmentation significative du LD moyen dans certaines populations.

Dans le but d'identifier des polymorphismes liés à la valeur agronomique des individus, il nous sera tout d'abord nécessaire de caractériser le phénotype (pour différents traits d'intérêt) sur un échantillonnage exhaustif de cultivars et d'oléastres, et ceci dans les mêmes conditions environnementales (« common garden »). Chaque individu sera également typé génétiquement. Des associations entre certains marqueurs et certains traits seront ensuite recherchées. Des gènes candidats potentiellement soumis à sélection durant la domestication (codant pour des protéines impliquées dans l'architecture, la qualité de l'huile, la taille du fruit, la teneur en huile...) devraient être étudiés dans ce type d'approche en combinaison avec des marqueurs considérés comme « neutres ». Un des problèmes auxquels nous serons confrontés chez l'olivier est l'existence d'effets confondants liés à l'histoire des populations. Notamment il existe deux pools génétiques hautement différenciés en Méditerranée et il y a eu de nombreuses hybridations entre ces deux pools. Cet aspect devra être absolument pris en compte. Plusieurs possibilités peuvent être envisagées pour y remédier. On pourrait constituer des échantillons d'individus appartenant aux deux pools méditerranéens selon les données génétiques (*a priori* « neutres ») accumulées sur différents génomes et faire les analyses indépendamment sur chacun de ces pools. L'autre possibilité est de considérer la structuration de la diversité génétique basé sur un grand nombre de loci, et tester si l'association trait marqueur reste significative une fois ce facteur pris en compte (e.g. Thornsberry *et al.* 2001). Les polymorphismes mis en évidence comme potentiellement sous sélection pourront être validés par des approches QTLs (B. KHADARI ; INRA Montpellier) ou de transcriptomique.

Enfin, une étude très prometteuse serait la caractérisation moléculaire de matériels archéologiques (e.g. Elbaum *et al.* 2006). Cette démarche permettrait d'introduire une composante spatio-temporelle dans nos investigations. Toutefois, dans l'état actuel des choses,

nous avons encore besoin de bien définir la démarche à suivre. Cette analyse nécessite de détruire des échantillons archéologiques ce qui pose un problème éthique, et il est donc nécessaire qu'elle génère des données indispensables à la compréhension de l'histoire de l'olivier. Je suis actuellement en contact avec deux archéobotanistes spécialisés sur l'histoire de l'olivier [J.F. TERRAL (Université de Montpellier II) et C. NEWTON (Université de Nottingham)] et nous devons encore définir des objectifs clairs. Dans un avenir relativement proche, les techniques de pyroséquençage (Millar *et al.* 2008) pourraient également révolutionner le domaine des analyses de matériel archéologiques, mais actuellement le coût de ce type d'approche reste encore prohibitif.

#### 4.2. Génétique de la conservation des sous-espèces menacées du complexe *O. europaea*

Mes investigations sur la conservation de l'olivier de Laperrine au Hoggar devraient se poursuivre au moins quelques années. En particulier, il reste à mieux étudier certains aspects sur la dynamique des flux de gènes dans les populations. J'aimerais approfondir l'étude basée sur les analyses de paternité avec un design mieux construit. En particulier, l'effet de la taille du patch ou de la topographie du paysage (et même de l'année de floraison) devrait être pris en compte, car selon nos études préliminaires, ils devraient avoir une influence relativement forte sur la dissémination effective du pollen. Par ailleurs, il serait également judicieux de s'attarder aux problèmes d'incompatibilité entre génotypes (e.g. Mookerjee *et al.* 2005). La caractérisation des gènes d'auto-incompatibilité serait certainement très informative (e.g. Charlesworth *et al.* 2005; C. BRETON, Comm. Pers.). Ces différents types d'information devraient s'avérer primordiaux pour la conservation de ce taxon et notamment pour sa multiplication par voie sexuée. De même, les performances des individus issus de graines devraient être suivies durant les premiers mois de croissance pour mesurer si une dépression de consanguinité peut être observée sur les descendances obtenues dans des conditions très particulières (patch réduit à quelques individus potentiellement apparentés ; cf. Tonget). Dans les deux prochaines années, j'aimerais monter un projet sur ces problématiques avec mes partenaires algériens (D. BAALI-CHERIF & N. BOUGUEDOURA ; USTHB d'Alger). Par ailleurs, il serait fortement judicieux d'ouvrir la thématique sur la préservation d'autres essences ligneuses présentes dans cet écosystème (e.g. *Acacia*, *Ficus*, *Pistacia*...), qui sont parfois en fort déclin. L'étude de la structuration de la diversité génétique sur différentes espèces, combinée à une meilleure caractérisation de l'habitat, devraient à terme permettre de définir des stratégies adaptées de préservation de ces espèces ligneuses des montagnes sahariennes.

En parallèle, d'autres taxons d'oliviers fortement menacés (subsp. *maroccana* et *cerasiformis*) pourraient aussi faire l'objet d'investigations. Ces sous-espèces d'oliviers polyploïdes sont d'un intérêt tout particulier car ils semblent retenir une diversité génétique très importante (García-Verdugo *et al.* 2009 ; Besnard *et al.* en prep.). En particulier, mon expertise dans le marquage moléculaire pourrait être mise à contribution dans des projets annexes sur ces taxa (avec respectivement, B. KHADARI et P. VARGAS).

#### 4.3. Transfert et valorisation

L'histoire de l'olivier, qui représente l'un des premiers arbres domestiqués avec le figuier et le dattier (Zohary & Spiegel-Roy 1975), est évidemment liée à l'histoire des populations humaines depuis plusieurs millénaires. Actuellement, l'étude de la domestication de cette espèce répond à

une réelle demande sociale, en particulier dans les pays du pourtour Méditerranéen où l'importance culturelle de l'olivier est extrêmement grande. De plus, une meilleure connaissance de la diversité et l'origine des cultivars devrait apporter des arguments pour une meilleure valorisation des produits du terroir (comme par exemple pour la mise en place d'AOC). J'ai quelques contacts avec le monde de l'oléiculture (et notamment avec l'Afidol, Nyons) ce qui me permet de transmettre le résultat de mes recherches à la profession, mais des échanges plus importants seraient à l'avenir nécessaires pour un meilleur transfert de ces nouvelles données.

Par ailleurs, la conservation *in situ* et la mise en collection d'oliviers cultivés et sauvages est sans doute un enjeu important pour préserver la diversité de l'olivier. L'étude et la mise en place de stratégies de préservation de taxa menacés (tels que les subspp. *laperrinei*, *guanchica*, *cerasiformis* ou *maroccana*) est également primordiale dans l'optique de protéger des écosystèmes, particulièrement face aux changements climatiques futurs et aux pressions anthropiques. Par ailleurs, des caractères d'intérêt (e.g. résistance à des pathogènes ou à des conditions d'extrême aridité, qualité d'huile) dans les populations sauvages (et notamment dans les sous-espèces apparentées) pourront être plus facilement identifiés dans notre collection. Des croisements pourront éventuellement être effectués afin d'introduire de tels caractères dans des formes cultivées ou pour simplement mieux comprendre le déterminisme génétique de tels traits. Cette collection pourrait donc avoir un intérêt dans l'amélioration de l'olivier mais ceci représente évidemment un challenge sur le très long terme dans la vie d'un chercheur.

## Conclusion

Durant les deux années à venir, j'envisage donc très clairement de développer de nouveaux outils de génomique (e.g. cpDNA, pyroséquençage) pour de nouvelles approches sur la compréhension des origines de l'olivier et de la conservation de sa diversité. Je vais aussi élargir mon réseau de relations (notamment en Italie, Grèce et Proche-Orient), afin de pouvoir initier ou m'impliquer activement dans de nouvelles recherches sur l'olivier. Au terme de ces deux années, l'objectif est de proposer un ou plusieurs projets afin de financer un programme ambitieux sur les ressources génétiques de cette espèce. Ceci pourra clairement être établi lorsque j'aurai obtenu une position stable dans un institut de recherche ou une université. Je pense que mes expériences dans le domaine de la Génétique des plantes et en Botanique devraient me permettre d'atteindre cet objectif assez rapidement. Par la suite, je souhaite m'impliquer activement dans l'animation d'une équipe de recherche notamment par l'encadrement d'étudiants de thèse et de master, et pourquoi pas élargir mes investigations à de nouvelles thématiques. A travers ces activités, j'espère pouvoir transmettre à de jeunes chercheurs mon enthousiasme pour la recherche mais aussi mes connaissances, et ainsi contribuer à dynamiser mon domaine de recherche. Des collaborations avec d'autres chercheurs impliqués dans l'origine d'espèces ligneuses cultivées en Méditerranée (en particulier avec B. KHADARI sur le figuier ou P. THIS sur la vigne ; INRA de Montpellier) seraient également une opportunité pour élargir nos investigations à des questions plus générales sur la domestication.

## Références

- Angiolillo A, Mencuccini M, Baldoni L (1999) Olive genetic diversity assessed using amplified polymorphic fragment length polymorphisms. *Theoretical and Applied Genetics* **98** : 411-421.
- Ainsworth EA, Rogers A, Leakey ADB (2008) Targets for crop biotechnology in a future high-CO<sub>2</sub> and high-O<sub>3</sub> world. *Plant Physiology* **147** : 13-19.
- Anthelme F, Abdoukader A, Besnard G (2008) Distribution, shape and clonal growth of the rare endemic tree *Olea europaea* subsp. *laperrinei* (Oleaceae) in the Saharan mountains of Niger. *Plant Ecology* **198** : 73-87.
- Arroyo-García R, et al. (2006) Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Molecular Ecology* **15** : 3707-3714.
- Baali-Cherif D, Besnard G (2005) High genetic diversity and clonal growth in relict populations of *Olea europaea* subsp. *laperrinei* (Oleaceae) from Hoggar, Algeria. *Annals of Botany (London)* **96** : 823-830.
- Barret P (1997) Contribution à l'étude du cycle photosynthétique de Hatch et Slack des plantes en C<sub>4</sub>. Mise en évidence d'un complexe multi-enzymatique. Thèse Université Saint Denis, La Réunion, 211 p.
- Belaj A, Muñoz-Díez C, Baldoni L, Porceddu A, Barranco D, Satovic Z (2007) Genetic diversity and population structure of wild olives from the north-western Mediterranean assessed by SSR markers. *Annals of Botany (London)* **100** : 449-458.
- Besnard G, Khadari B, Villemur P, Bervillé A (2000) Cytoplasmic male sterility in the olive (*Olea europaea* L.). *Theoretical and Applied Genetics* **100** : 1018-1024.
- Besnard G, Baradat P, Bervillé A (2001a) Genetic relationships in the olive (*Olea europaea* L.) reflect multilocal selection of cultivars. *Theoretical and Applied Genetics* **102** : 251-258.
- Besnard G, Baradat P, Breton C, Khadari B, Bervillé A (2001b) Olive domestication from structure of oleasters and cultivars using nuclear RAPDs and mitochondrial RFLPs. *Genetics, Selection, Evolution* **33** : S251-S268.
- Besnard G, Baradat P, Chevalier D, Tagmount A, Bervillé A (2001c) Genetic differentiation in the olive complex (*Olea europaea*) revealed by RAPDs and RFLPs in the rRNA genes. *Genetic Resources and Crop Evolution* **48** : 165-182.
- Besnard G, Breton C, Baradat P, Khadari B, Bervillé A (2001d) Cultivar identification in olive based on RAPD markers. *Journal of the American Society for Horticultural Science* **126** : 668-675.
- Besnard G, Green PS, Bervillé A (2002a) The genus *Olea*: molecular approaches of its structure and relationships to other Oleaceae. *Acta Botanica Gallica* **149** : 49-66.
- Besnard G, Khadari B, Baradat P, Bervillé A (2002b) Combination of chloroplast and mitochondrial DNA polymorphisms to study cytoplasmic genetic differentiation in the olive complex (*Olea europaea* L.). *Theoretical and Applied Genetics* **105** : 139-144.
- Besnard G, Offmann B, Robert C, Rouch C, Cadet F (2002c) Assessment of the C<sub>4</sub> phosphoenolpyruvate carboxylase gene diversity in grasses (Poaceae). *Theoretical and Applied Genetics* **105** : 404-412.
- Besnard G, Pinçon G, D'Hont A, Hoarau JY, Cadet F, Offmann B (2003a) Characterisation of the phosphoenolpyruvate carboxylase gene family in sugarcane (*Saccharum* spp.). *Theoretical and Applied Genetics* **107** : 470-478.
- Besnard G, Rubio de Casas R, Vargas P (2003b) A set of primers for length and nucleotide-substitution polymorphism in chloroplast DNA of *Olea europaea* L. (Oleaceae). *Molecular Ecology Notes* **3** : 651-653.
- Besnard G, Christin PA, Baali-Cherif D, Bouguedoura N, Anthelme F (2007a) Spatial genetic structure in the Laperrine's olive (*Olea europaea* subsp. *laperrinei*), a long-living tree from central-Saharan mountains. *Heredity* **99** : 649-657.
- Besnard G, Henry P, Wille L, Cooke D, Chapuis E (2007b) On the origin of the invasive olives (*Olea europaea* L., Oleaceae). *Heredity* **99** : 608-619.
- Besnard G, Rubio de Casas R, Vargas P (2007c) Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea* L.). *Journal of Biogeography* **34** : 736-752.
- Besnard G, Wille L, Henry P, Chapuis E, Christin PA (2007d) Can microsatellite data allow identification of oleaster Plio-Pleistocene refuge zones in the Mediterranean Basin? *Journal of Biogeography* **34** : 559-560.
- Besnard G, García-Verdugo C, Rubio de Casas R, Treier UA, Galland N, Vargas P (2008) Polyploidy in the olive complex (*Olea europaea* L.): evidence from flow cytometry and nuclear microsatellite analyses. *Annals of Botany (London)* **101** : 25-30.
- Besnard G, Baali-Cherif D (2009) Coexistence of diploids and triploids in a Saharan relict olive: evidence from nuclear microsatellite and flow cytometry analyses. *Comptes Rendus Biologies* **332** : 1115-1120.
- Besnard G, Bettinelli-Riccardi S, Parietti D, Bouguedoura N, Baali-Cherif D (2009a) Pollen gene flow in a highly fragmented landscape: consequences for defining a conservation strategy of the relict Laperrine's olive. *Comptes Rendus Biologies* **332** : 662-672.
- Besnard G, Muasya AM, Russier F, Roalson EH, Salamin N, Christin PA (2009b) Phylogenomics of C<sub>4</sub> photosynthesis convergent evolutions in Cyperaceae. *Molecular Biology and Evolution* **26** : 1909-1919.
- Besnard G, Rubio de Casas R, Christin PA, Vargas P (2009c) Phylogeny of *Olea* (Oleaceae) based on plastid and nuclear-ribosomal DNA sequences: Tertiary climatic shifts and lineage divergence times. *Annals of Botany (London)* **104** : 143-160.
- Bongi G, Mencuccini M, Fontanazza G (1987) Photosynthesis of olive leaves: Effect of light flux density, leaf age, temperature, peltates, and H<sub>2</sub>O vapor pressure deficit on gas exchange. *Journal of the American Society for Horticultural Science* **112** : 143-148.
- Bowers J, Boursiquot JM, This P, Chu K, Johansson H, Meredith C (1999) Historical genetics: The parentage of Chardonnay, Gamay, and other wine grapes of northeastern France. *Science* **285** : 1562-1565.
- Breton C, Tersac M, Bervillé A (2006) Genetic diversity and gene flow between the wild olive (oleaster, *Olea europaea* L.) and the olive: several Plio-Pleistocene refuge zones in the Mediterranean basin suggested by Simple Sequence Repeats analysis. *Journal of Biogeography* **33** : 1916-1928.
- Breton C, Pinatel C, Médail F, Bonhomme F, Bervillé A (2008) Comparison between classical and Bayesian methods to investigate the history of olive cultivars using SSR-polymorphisms. *Plant Science* **175** : 524-532.
- Carmo-Silva AE, Soares AS, Marques da Silva J, Bernardes da Silva A, Keys AJ, Arrabaça MC (2007) Photosynthetic responses of three C<sub>4</sub> grasses of different metabolic subtypes to water deficit. *Functional Plant Biology* **34** : 204-213.
- Carrión JS, Dupré M (1996) Late Quaternary vegetational history at Navarrés, Eastern Spain. A two core approach. *New Phytologist* **134** : 177-91.
- Cerling TE, Harris JM, MacFadden BJ, Leakey MG, Quade J, Eisenmann V, Ehleringer JR (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389** : 153-158.
- Charlesworth D, Vekemans X, Castric V, Glémin S (2005) Plant self-incompatibility systems: a molecular evolutionary perspective. *New Phytologist* **168** : 61-69.
- Chevalier A (1948) L'origine de l'olivier cultivé et ses variations. *Revue Internationale de Botanique Appliquée et d'Agriculture Tropicale* **28** : 1-

- Christin PA, Salamin N, Savolainen V, Besnard G (2007a) A phylogenetic study of the phosphoenolpyruvate carboxylase multigene family in *Poaceae*: understanding the molecular changes linked to C<sub>4</sub> photosynthesis evolution. *Kew Bulletin* **62** : 455-462.
- Christin PA, Salamin N, Savolainen V, Duvall MR, Besnard G (2007b) C<sub>4</sub> photosynthesis evolved in grasses via parallel adaptive genetic changes. *Current Biology* **17** : 1241-1247.
- Christin PA (2008) Phylogenetic analyses of C<sub>4</sub> photosynthesis evolution in grasses: genetics of convergence and convergence of genes. PhD thesis, Lausanne University, CH, 156 p.
- Christin PA, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, Savolainen V, Salamin N (2008a) Oligocene CO<sub>2</sub> decline promoted C<sub>4</sub> photosynthesis in grasses. *Current Biology* **18** : 37-43.
- Christin PA, Salamin N, Muasya AM, Roalson EH, Russier F, Besnard G (2008b) Evolutionary switch and genetic convergence on *rbcL* following the evolution of C<sub>4</sub> photosynthesis. *Molecular Biology and Evolution* **25** : 2361-2368.
- Christin PA, Besnard G (2009) Two independent C<sub>4</sub> origins in Aristidoideae (Poaceae) revealed by the recruitment of distinct phosphoenolpyruvate carboxylase genes. *American Journal of Botany* **96** : 2234-2239.
- Christin PA, Petitpierre B, Salamin N, Büchi L, Besnard G (2009a) Evolution of C<sub>4</sub>-related phosphoenolpyruvate carboxykinase in grasses, from genotype to phenotype. *Molecular Biology and Evolution* **26** : 357-365.
- Christin PA, Salamin N, Kellogg EA, Vicentini A, Besnard G (2009b) Integrating phylogeny into studies of C<sub>4</sub> variation in the grasses. *Plant Physiology* **149** : 82-87.
- Christin PA, Samaritani E, Petitpierre B, Salamin N, Besnard G (2009c) Evolutionary insights on C<sub>4</sub> photosynthetic biochemical subtypes in grasses from genomics and phylogenetics. *Genome Biology and Evolution* **1** : 221-230.
- CNUCED (2006) Information des marchés dans le secteur des produits de base: Huile d'olive. <http://www.unctad.org/infocomm/francais/olive/marche.htm>
- Comai L (2005) The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* **6** : 836-846.
- Conway-Morris S (2006) Evolutionary convergence. *Current Biology* **16** : R826-R827.
- Cuevas J, Polito VS (1997) Compatibility relationships in 'Manzanillo' olive. *HortScience* **32** : 1056-1058.
- di Vecchi-Staraz M, Laucou V, Bruno G, Lacombe T, Gerber S, Bourse T, Boselli M, This P (2009) Low level of pollen-mediated gene flow from cultivated to wild grapevine: Consequences for the evolution of the endangered subspecies *Vitis vinifera* L. subsp. *silvestris*. *Journal of Heredity* **100** : sous presse.
- Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. *Cell* **127** : 1309-1321.
- Doveri S, Gil FS, Diaz A, Reale S, Busconi M, Machado AD, Martin A, Fogher C, Donini P, Lee D (2008) Standardization of a set of microsatellite markers for use in cultivar identification studies in olive (*Olea europaea* L.). *Scientia Horticulturae* **116** : 367-373.
- Duvall MR, Davis JI, Clark LG, Noll JD, Goldman DH, Sánchez-Ken JG (2007) Phylogeny of the grasses (Poaceae) revisited. *Aliso* **23** : 237-247.
- Elbaum R, Melamed-Bessudo C, Boaretto E, Galili E, Lev-Yadun S, Levy AA, Weiner S (2006) Ancient olive DNA in pits: preservation, amplification and sequence analysis. *Journal of Archaeological Science* **33** : 77-88.
- Edwards GE, Franceschi VR, Voznesenskaya EV (2004) Single-cell C<sub>4</sub> photosynthesis versus the dual-cell (Kranz) paradigm. *Annual Review of Plant Biology* **55** : 173-196.
- Ehleringer JR, Sage RF, Flanagan LB, Pearcy RW (1991) Climate change and the evolution of C<sub>4</sub> photosynthesis. *Trends in Ecology and Evolution* **6** : 95-99.
- Ehleringer JR, Cerling TE, Helliker BR (1997) C<sub>4</sub> photosynthesis, atmospheric CO<sub>2</sub>, and climate. *Oecologia* **112** : 285-299.
- Flint-Garcia SA, Thornsberry JM, Buckler IV ES (2003) Structure of linkage disequilibrium in plants. *Annual Reviews of Plant Biology* **54** : 357-374.
- Galili E, Stanley DJ, Sharvit J, Weinstein-Evron M (1997) Evidence for earliest olive-oil production in submerged settlements off the Carmel coast, Israel. *Journal of Archaeological Science* **24** : 1141-1150.
- García-Verdugo C, Fay MF, Granado-Yela C, Rubio de Casas R, Balaguer L, Besnard G, Vargas P (2009) Genetic differentiation under geographical isolation in a ploidy series of *Olea europaea*: A multiscale approach from subspecies to islands. *Molecular Ecology* **18** : sous presse.
- Ghannoum O, Evans JR, Chow WS, Andrews TJ, Conroy JP, von Caemmerer S (2005) Faster Rubisco is the key to superior nitrogen-use efficiency in NADP-malic enzyme relative to NAD-malic enzyme C<sub>4</sub> grasses. *Plant Physiology* **137** : 638-650.
- Giussani L, Cota-Sánchez JH, Zuloaga F, Kellogg EA. 2001. A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C<sub>4</sub> photosynthesis. *American Journal of Botany* **88** : 1993-2012.
- Green PS (2002) A revision of *Olea* L. *Kew Bulletin* **57** : 91-140.
- Green PS, Wickens GE (1989) The *Olea europaea* complex. The Davis & Hedge Festschrift, ed. Kit Tan, pp. 287-299. Edinburgh University Press.
- Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Molecular Biology* **57** : 461-485.
- Hey J (2006) Recent advances in assessing gene flow between divergent populations and species. *Current Opinion in Genetics & Development* **16** : 592-596.
- Hess J, Kadereit JW, Vargas P (2000) The colonization history of *Olea europaea* L. in Macaronesia based on internal transcribed spacer 1 (ITS-1) sequences, randomly amplified polymorphic DNAs (RAPD), and intersimple sequence repeats (ISSR). *Molecular Ecology* **9** : 857-868.
- Heuertz M, de Paoli E, Källman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. *Genetics* **174** : 2095-2105.
- Heun M, SchaferPregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* **278** : 1312-1314.
- Hibberd JM, Sheehy JE, Langdale JA (2008) Using C<sub>4</sub> photosynthesis to increase the yield of rice: rationale and feasibility. *Current Opinion in Plant Biology* **11** : 228-231.
- Hudson GS, Mahon JD, Anderson PA, Gibbs MJ, Badger MR, Andrews TJ, Whitfield PR (1990) Comparisons of *rbcL* genes for the large subunit of Ribulose-bisphosphate carboxylase from closely related C<sub>3</sub> and C<sub>4</sub> plant species. *Journal of Biological Chemistry* **265** : 808-814.
- Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources* **8** : 3-17.
- Johnson LAS (1957) A review of the family Oleaceae. *Contributions from the New South Wales National Herbarium* **2** : 397-418.
- Kadereit G, Borsch T, Weising K, Freitag H (2003) Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C<sub>4</sub> photosynthesis. *International Journal of Plant Science* **164** : 959-986.

- Kellogg EA (2001) Evolutionary history of the grasses. *Plant Physiology* **125** : 1198-1205.
- Kelly MG, Huntley B (1991) An 11 000-year record of vegetation and environment from Lago di Martignano, Latium, Italy. *Journal of Quaternary Science* **6** : 209-224.
- Khadari B, Breton C, Moutier N, Roger JP, Besnard G, Bervillé A, Dosba F (2003) The use of molecular markers for germplasm management in a French olive collection. *Theoretical and Applied Genetics* **106** : 521-529.
- Khadari B, Charafi J, Moukhli A, Ater M (2008) Substantial genetic diversity in cultivated Moroccan olive despite a single major cultivar: a paradoxical situation evidenced by the use of SSR loci. *Tree Genetics & Genomes* **4** : 213-221.
- Kislev ME, Nadel D, Carmi I (1992) Epipalaeolithic (19 000 B.P.) cereal and fruit diet at Ohalo II, Sea of Galilee, Israel. *Review of Palaeobotany and Palynology* **73** : 161-166.
- La Mantia M, Lain O, Caruso T, Testolin R (2005) SSR-based DNA fingerprints reveal the genetic diversity of Sicilian olive (*Olea europaea* L.) germplasm. *Journal of Horticultural Science & Biotechnology* **80** : 628-632.
- Levy AA, Feldman M (2002) The impact of polyploidy on grass genome evolution. *Plant Physiology* **130** : 1587-1593.
- Loussert R, Brousse G (1978) *L'olivier : Techniques agricoles et productions méditerranéennes*. G.P. Maisonneuve et Larose, Paris.
- Lumaret R, Ouazzani N (2001) Ancient wild olives in Mediterranean forests. *Nature* **413** : 700.
- Margulies M, Egholm M, Alman WE, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437** : 376-380.
- Médail F, Quézel P, Besnard G, Khadari B (2001) Systematics, ecology and phylogeographic significance of *Olea europaea* L. subsp. *maroccana* (Greuter & Burdet) P. Vargas et al., a relictual olive tree in south-west Morocco. *Botanical Journal of the Linnean Society* **137** : 249-266.
- Millar CD, Huynen L, Subramanian S, Mohandesan E, Lambert DM (2008) New developments in ancient genomics. *Trends in Ecology and Evolution* **23** : 386-393.
- Monson RK (2003) Gene duplication, neofunctionalization, and the evolution of C<sub>4</sub> photosynthesis. *International Journal of Plant Science* **164** : S43-S54.
- Mookerjee S, Guerin J, Collins G, Ford C, Sedgley M (2005) Paternity analysis using microsatellite markers to identify pollen donors in an olive grove. *Theoretical and Applied Genetics* **111** : 1174-1182.
- Morgan JA, Milchunas DG, LeCain DR, West M, Mosier AR (2007) Carbon dioxide enrichment alters plant community structure and accelerates shrub growth in the shortgrass steppe. *Proceedings of the National Academy of Sciences, USA* **104** : 14724-14729.
- Mudie PJ, Marret F, Aksu AE, Hiscott RN, Gillespie H (2007) Palynological evidence for climatic change, anthropogenic activity and outflow of Black Sea water during the late Pleistocene and Holocene: Centennial- to decadal-scale records from the Black and Marmara Seas. *Quaternary International* **167/168** : 73-90.
- Muhaidat R, Sage RF, Dengler NG (2007) Diversity of Kranz anatomy and biochemistry in C<sub>4</sub> eudicots. *American Journal of Botany* **94** : 362-81.
- Murphy LR, Barroca J, Franceschi VR, Lee R, Roalson EH, Edwards GE, Ku MSB (2007) Diversity and plasticity of C<sub>4</sub> photosynthesis in *Eleocharis* (Cyperaceae). *Functional Plant Biology* **34** : 571-580.
- Nisbet EG, Grassineau NV, Howe CJ, Abell PI, Regelous M, Nisbet RER (2007) The age of Rubisco: the evolution of oxygenic photosynthesis. *Geobiology* **5** : 311-335.
- Osborne CP (2008) Atmosphere, ecology and evolution: what drove the Miocene expansion of C<sub>4</sub> grasslands? *Journal of Ecology* **96** : 35-45.
- Pagani M, Zachos JC, Freeman KH, Tipler B, Bohaty S (2005) Marked decline in atmospheric carbon dioxide concentrations during the Paleogene. *Science* **309** : 600-603.
- Perrier de la Bâthie H (1952) Oléacées (*Oleaceae*). In : *Flore de Madagascar et des Comores*. H Humert (ed.), Paris, Gouvernement Général de Madagascar. 166<sup>e</sup> famille, pp. 1-89.
- Perron M, Perry DJ, Andalo C, Bousquet J (2000) Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. *Proceedings of the National Academy of Sciences of the United States of America* **97** : 11331-11336.
- Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plants. *Heredity* **71** : 630-641.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure from multilocus genotype data. *Genetics* **155** : 945-959.
- Quézel P (1965) *La végétation du Sahara. Du Tchad à la Mauritanie*. Stuttgart: Gustav Fisher Verlag.
- Rekik I, Salimonti A, Grati Kamoun N, Muzzalupo I, Lepais O, Gerber S, Perri E, Rebai A (2008) Characterization and identification of Tunisian olive tree varieties by microsatellite markers. *HortScience* **43** : 1371-1376.
- Roalson EH, Friar EA (2000) Infrageneric classification of *Eleocharis* (Cyperaceae) revisited: Evidence from the internal transcribed spacer (ITS) region of nuclear ribosomal DNA. *Systematic Botany* **25** : 323-336.
- Rondeau P, Rouch C, Besnard G (2005) NADP-malate dehydrogenase gene evolution in Andropogoneae (Poaceae): gene duplication followed by sub-functionalization. *Annals of Botany (London)* **96** : 1307-1314.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19** : 1572-1574.
- Rubio de Casas R, Besnard G, Schoenswetter P, Balaguer L, Vargas P (2006) Extensive gene flow blurs phylogeographic but not phylogenetic signal in *Olea europaea* L. *Theoretical and Applied Genetics* **113** : 575-583.
- Rutschmann F (2006) Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. *Diversity & Distribution* **12** : 35-48.
- Sage RF (2001) Environmental and evolutionary preconditions for the origin and diversification of the C<sub>4</sub> photosynthetic syndrome. *Plant Biology* **3** : 202-213.
- Sage RF, Kubien DS (2003) Quo vadis C<sub>4</sub>? An ecophysiological perspective on global change and the future of C<sub>4</sub> plants. *Photosynthesis Research* **77** : 209-225.
- Sage RF (2004) The evolution of C<sub>4</sub> photosynthesis. *New Phytologist* **161** : 341-370.
- Sage RF, Sage TL, Percy RW, Borsch T (2007) The taxonomic distribution of C<sub>4</sub> photosynthesis in Amaranthaceae sensu stricto. *American Journal of Botany* **94** : 1992-2003.
- Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. *Nature Reviews Genetics* **3** : 429-441.
- Schoen DJ, Brown AHD (2001) The conservation of wild plant species in seed banks. *BioScience* **51** : 960-966.
- Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M (2006) Tectonic uplift and Eastern Africa aridification. *Science* **313** : 1419-1423.
- Sharkey DS (2001) Photorespiration. *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd. doi: 10.1038/npg.els.0001292.
- Simpson DA, Muasya AM, Alves MV, Bruhl JJ, Dhooge S, Chase MW, Furness CA, Ghamkhar K, Goetghebeur P, Hodkinson TR, Marchant AD, Reznicek AA, Nieuwborg R, Roalson EH, Smets E, Starr JR, Thomas WW, Wislon KL, Zhang X (2007) Phylogeny of Cyperaceae based on DNA sequence data – A new *rbcL* analysis. *Aliso* **23** : 72-83.

- Sinha NR, Kellogg EA (1996) Parallelism and diversity in multiple origins of C<sub>4</sub> photosynthesis in the grass family. *American Journal of Botany* **83** : 1458-1470.
- Soltis DE, Soltis PS, Tate JA (2004) Advances in the study of polyploidy since Plant speciation. *New Phytologist* **161** : 173-191.
- Soros CL, Bruhl JJ (2000) Multiple evolutionary origins of C<sub>4</sub> photosynthesis in the Cyperaceae. In: *Monocots: Systematics and Evolution*, Vol. 1. KL Wilson & DA Morrison, editors. CSIRO Publishing, Collingwood, Victoria, Australia, p. 629-636.
- Soros CL, Dengler NG (2001) Ontogenic deviation and cell differentiation in photosynthetic tissues of C<sub>3</sub> and C<sub>4</sub> Cyperaceae. *American Journal of Botany* **88** : 992-1005.
- Spennemann DHR, Allen LR (2000). Feral olives (*Olea europaea*) as future woody weeds in Australia: a review. *Australian Journal of Experimental Agriculture* **40** : 889-901.
- Svensson P, Bläsing OE, Westhoff P (2003) Evolution of C<sub>4</sub> phosphoenolpyruvate carboxylase. *Archives of Biochemistry and Biophysics* **414** : 180-188.
- Taub DR (2000) Climate and the US distribution of C<sub>4</sub> grass subfamilies and decarboxylation variants of C<sub>4</sub> photosynthesis. *American Journal of Botany* **87** : 1211-1215.
- Taylor H (1945) Cyto-taxonomy and phylogeny of the *Oleaceae*. *Brittonia* **5** : 337-367.
- Tcherkez GBB, Farquhar GD, Andrews TJ (2006) Despite slow catalysis and confused substrate specificity, all ribulose biphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences, USA* **103** : 7246-7251.
- Terral JF, Arnold-Simard G (1996) Beginnings of olive cultivation in eastern Spain in relation to Holocene bioclimatic changes. *Quaternary Research* **46** : 176-185.
- Terral JF (1997) La domestication de l'olivier (*Olea europaea* L.) en Méditerranée nord-occidentale : Approche morphométrique et implications paléoclimatiques. Thèse, Université Montpellier II, France.
- Terral JF, Alonso N, Buxó i Capdevila R, Chatti N, Fabre L, Fiorentino G, Marinval P, Pérez Jordá G, Pradat B, Rovira N, Alibert P (2004a) Historical biogeography of olive domestication (*Olea europaea* L.) as revealed by geometrical morphometry applied to biological and archaeological material. *Journal of Biogeography* **31** : 63-77.
- Terral JF, Badal E, Heinz C, Roiron P, Thiebault S, Figueiral I (2004b) A hydraulic conductivity model points to post-Neogene survival of the Mediterranean olive. *Ecology* **85** : 3158-3165.
- Tessier C, David J, This P, Boursiquot JM, Charrier A (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *Theoretical and Applied Genetics* **98** : 171-177.
- Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* **15** : 1647-1657.
- Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* **51** : 689-702.
- Thornberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES IV (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nature Genetics* **28** : 286-289.
- Turrill WB (1951) Wild and cultivated olives. *Kew Bulletin* **3** : 437-442.
- Ueno O (2001) Environmental regulation of C<sub>3</sub> and C<sub>4</sub> differentiation in the amphibious sedge *Eleocharis vivipara*. *Plant Physiology* **127** : 1524.
- Vargas P, Munoz Garmendia F, Hess J, Kadereit JW (2001) *Olea europaea* ssp *guanchica* and ssp *maroccana* (*Oleaceae*), two new names for olive tree relatives. *Anales del Jardín Botánico de Madrid* **58** : 360-361.
- Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population. *Molecular Ecology Notes* **2** : 618-620.
- Villemur P, Musho US, Delmas JM, Maamar M, Ouksili A (1984) Contribution to the study of flower biology of olive tree (*Olea europaea* L.): male-sterility, pollen flow and effective pollination period. *Fruits* **39** : 467-473.
- von Caemmerer S, Furbank RT (2003) The C<sub>4</sub> pathway: an efficient CO<sub>2</sub> pump. *Photosynthesis Research* **77** : 191-207.
- Vouillamoz J, Maigre D, Meredith CP (2003) Microsatellite analysis of ancient alpine grape cultivars: pedigree reconstruction of *Vitis vinifera* L. 'Cornalin du Valais'. *Theoretical and Applied Genetics* **107** : 448-454.
- Wallander E, Albert VA (2000) Phylogeny and classification of Oleaceae based on *rps16* and *trnL-F* sequence data. *American Journal of Botany* **87** : 1827-1841.
- Weinreich DM, Delaney NF, DePristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312** : 111-114.
- Westhoff P, Gowik U (2004) Evolution of C<sub>4</sub> phosphoenolpyruvate carboxylase. Genes and proteins: a case study with the genus *Flaveria*. *Annals of Botany (London)* **94** : 13-23.
- Williams JW, Jackson ST, Kutzbach JE (2007) Projected distributions of novel and disappearing climates by 2100AD. *Proceedings of the National Academy of Sciences, USA* **104** : 5738-5742.
- Yang ZH, Nielsen R, Goldman N, Krabbe Pedersen AM (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155** : 431-449.
- Yang ZH, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19** : 908-917.
- Yang ZH, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acids sites under positive selection. *Molecular Biology and Evolution* **22** : 1107-1118.
- Yeoh HH, Badger MR, Watson L (1980) Variations in K<sub>m</sub>(CO<sub>2</sub>) of ribulose-1,5-bisphosphate carboxylase among grasses. *Plant Physiology* **66** : 1110-1112.
- Yeoh HH, Badger MR, Watson L (1981) Variations in kinetic properties of ribulose-1,5-bisphosphate carboxylase among plants. *Plant Physiology* **67** : 1151-1155.
- Yll EI, Perez-Obiol R, Pantaleon-Cano J, Roure JM (1997) Palynological evidence for climatic change and human activity during Holocene on Minorca (Balearic Islands). *Quaternary Research* **48** : 339-347.
- Zeder MA, Emshwiller E, Smith BD, Bradley DG (2006) Documenting domestication: the inter-section of genetics and archaeology. *Trends in Genetics* **22** : 139-155.
- Zhang JZ, Nielsen R, Yang ZH (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution* **22** : 2472-2479.
- Zohary D, Hopf M (2000) *Domestication of plants in the Old World*. Third edition, Oxford Clarendon Press.
- Zohary D, Spiegel-Roy P (1975) Beginnings of fruit growing in the Old World. *Science* **187** : 319-327.

## Résumé

Depuis septembre 1994, j'ai été impliqué dans divers projets de recherche ayant pour principaux objectifs l'utilisation et la gestion des ressources génétiques de plantes cultivées ainsi que des projets plus fondamentaux visant à mieux comprendre l'évolution de caractères adaptatifs. Durant mon cursus, j'ai été amené à travailler dans plusieurs laboratoires en Europe ainsi qu'en milieu tropical. Je me suis particulièrement spécialisé dans le développement d'outils de génétique moléculaire pour des approches de génétique des populations et de phylogénie. Dans mon document de synthèse pour l'habilitation à diriger de recherche, mes travaux les plus significatifs sont présentés. Dans une première partie, l'ensemble de mes travaux sur les ressources génétiques et la domestication de l'olivier est abordé. Parmi les résultats importants, on peut mentionner la mise en évidence d'une domestication multilocale de l'olivier en Méditerranée ainsi que la reconstruction d'une première phylogéographie exhaustive du genre *Olea* et du complexe de l'olivier. Dans une seconde partie, je présente mes travaux sur les bases moléculaires des apparitions multiples de la photosynthèse C<sub>4</sub>, un caractère d'adaptation aux conditions favorisant une forte photorespiration (tout particulièrement en milieu tropical ouvert). Au moins 22 lignées C<sub>4</sub> ont pu être mises en évidence chez les Poales à l'aide d'une phylogénie de l'ADN chloroplastique. De très fortes évidences de convergence ont ensuite été obtenues sur les gènes impliqués dans le système C<sub>4</sub>, particulièrement sur ceux codant la phosphoenolpyruvate carboxylase (qui fixe initialement le CO<sub>2</sub>) mais aussi la grande sous-unité de la Rubisco, l'enzyme la plus abondante du règne végétale. A travers ces deux grandes thématiques, je discute des intérêts fondamentaux et appliqués de mes recherches, mais j'insiste aussi sur les différentes compétences de chercheur que j'ai acquises tant au niveau technique que dans la gestion des activités. Mes différentes expériences d'encadrement d'étudiants et de conduite de nouveaux projets en collaboration avec des laboratoires étrangers sont ainsi mentionnées. Pour finir, mes projets de recherche pour les prochaines années sur la phylogéographie et la domestication de l'olivier sont brièvement présentés.

**Mots-clés :** Convergence évolutive, Domestication, Génétique des populations, Photosynthèse C<sub>4</sub>, Phylogénie.