



Polyploïdie et adaptation des plantes Étude de l'expression des gènes homéologues chez le caféier (Coffea arabica)

Marie-Christine Combes Gavalda

► **To cite this version:**

Marie-Christine Combes Gavalda. Polyploïdie et adaptation des plantes Étude de l'expression des gènes homéologues chez le caféier (Coffea arabica) . Biochimie, Biologie Moléculaire. 2011. <hal-01461992>

HAL Id: hal-01461992

<https://hal-ephe.archives-ouvertes.fr/hal-01461992>

Submitted on 8 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
ÉCOLE PRATIQUE DES HAUTES ÉTUDES

Sciences de la Vie et de la Terre

MÉMOIRE

Présenté par

Marie-Christine Combes Gavalda

Pour l'obtention du diplôme de l'École Pratique des Hautes Études

POLYPLOÏDIE ET ADAPTATION DES PLANTES
ETUDE DE L'EXPRESSION DE GENES
HOMEOLOGUES CHEZ LE CAFEIER
(*Coffea arabica*)

Soutenu le 7 octobre 2011 devant le jury suivant :

Claudine Montgelard – Président
Philippe Lashermes – Tuteur scientifique
Thierry Dupressoir – Tuteur pédagogique
Alain Rival – Rapporteur
Juan-Carlos Herrera – Examineur

Mémoire préparé sous la direction de :

Philippe Lashermes

Centre IRD-Montpellier

UMR Résistance des Plantes aux Bioagresseurs

Equipe DivA Diversité Amélioration Adaptation

Directeur : *Michel Nicole*

Philippe.Lashermes@ird.fr

Thierry Dupressoir

Laboratoire de Pathologie comparée des invertébrés

Diversité, Génomes et Interactions Microorganismes Insectes INRA UMR 1333-DGIMI

EPHE (Sciences de la Vie et de la Terre)

Thierry.Dupressoir@univ-montp2.fr

ÉCOLE PRATIQUE DES HAUTES ÉTUDES
Sciences de la Vie et de la Terre

**POLYPLOÏDIE ET ADAPTATION DES PLANTES ETUDE DE L'EXPRESSION DE
GENES HOMEOLOGUES CHEZ LE CAFEIER (*Coffea arabica*)**

Marie-Christine Combes Gavalda

Soutenu le 7 octobre 2011

La polyplœidie joue un rôle fondamental dans l'évolution des plantes. En effet, elle est considérée comme un facteur majeur de spéciation, de diversification et d'adaptation écologique des plantes. En particulier au sein des allopolyploïdes, l'expression des gènes dupliqués (gènes homœologues) permettrait une plasticité fonctionnelle et serait à l'origine de capacités adaptatives. Parmi les espèces polyplœides, la contribution relative des gènes homœologues au transcriptome peut varier selon les organes de la plante et selon les conditions de culture.

Le caféier *C. arabica* est une espèce allotétraploïde présentant deux sous-génomes homœologues peu différenciés. Cette espèce provient de l'hybridation récente de deux espèces diploïdes de caféiers proches des espèces actuelles de *C. eugenioides* (génome E^a de *C. arabica*) et de *C. canephora* (génome C^a de *C. arabica*). Sur le plan adaptatif, *C. arabica* présente la caractéristique de tolérer de plus grande amplitude de température que les deux espèces parentales diploïdes, l'une *C. canephora* étant plus adaptée à des températures chaudes (22 à 26 °C) et l'autre *C. eugenioides* à des températures froides (18 à 22°C).

Pour comprendre la spécificité adaptative des allopolyploïdes, l'influence de la température sur l'expression des gènes homœologues de *C. arabica*, a été étudiée en estimant l'expression relative de 13 paires de gènes homœologues de la même région génomique sur différents organes. La quantification du ratio de SNP estimée à partir de chromatogrammes de séquençage de type Sanger a permis de mesurer l'expression relative des gènes homœologues.

Pour les 130 combinaisons analysées, les deux sous-génomes ont été exprimés même si les conditions de culture conviennent mieux à l'une ou à l'autre des espèces parentales, aucun cas de « silencing » n'a été décelé. L'expression relative des gènes homœologues dépend des gènes, des organes et des conditions de culture. De plus l'expression relative des gènes homœologues adjacents apparaît indépendante et les profils d'expression sur la région étudiée sont relativement stables au travers des organes et des conditions de culture.

Ces résultats suggèrent que *C. arabica* n'a pas subi de phénomènes de diploïdisation notables. Ces résultats indiquent que l'expression des gènes homœologues au sein de *C. arabica* est régulée par un mécanisme de trans-régulation partagé qui n'est pas spécifique du sous-génome. Par contre les biais d'expression relative des gènes homœologues résulteraient de facteurs de régulation à l'échelle du gène. La régulation de l'expression des gènes homœologues semble intervenir sur les capacités adaptatives particulières de cette espèce.

Mots clés : polyplœidie, adaptation, gènes homœologues, *Coffea arabica*, transcription, introgression, expression gènes dupliqués

TABLES DES MATIERES

INTRODUCTION	5
CHAPITRE I : SYNTHESE BIBLIOGRAPHIQUE	
1/ LA POLYPLOÏDIE	
1-1/ Définition	7
1-2/ Fréquence de la polyploïdie et paléopolyploïdie	7
1-3/ Formation des polyploïdes	9
1-4/ Etude de la polyploïdie	10
2/ LES EFFETS DE LA POLYPLOÏDIE	
2-1/ Les réarrangements génomiques	11
2-1-1/ Les réarrangements de la structure des génomes	11
2-1-2/ L'activation des transposons et méthylation de l'ADN	12
2-2/ La régulation de l'expression des gènes	12
2-2-1/ Estimation de l'expression des gènes	13
2-2-2/ Mécanismes de régulation de l'expression des gènes	13
3/ CONSEQUENCES EVOLUTIVES DE LA POLYPLOÏDIE	
3-1/ Evolution des gènes dupliqués	15
3-2/ Polyploïdie et adaptation des plantes	17
4/ LE CAFEIER <i>COFFEA ARABICA</i>	
4-1/ Phylogènie et origine des caféiers	18
4-2/ Caractéristiques phénotypiques des caféiers	20
4-3/ Culture des caféiers, amélioration de <i>C. arabica</i>	20
5/ LES OBJECTIFS DE L'ETUDE	21
BIBLIOGRAPHIQUE	23

Liste des abréviations

ADN	Acide DésoxyriboNucléique
ADNc	Acide DésoxyriboNucléique complémentaire
ADNg	Acide DésoxyriboNucléique génomique
AFLP	Polymorphisme de longueur de fragments amplifiés (Amplified Fragment-Length Polymorphism)
ANOVA	Analyse de Variance (Analysis Of Variance)
ARNi	Acide RiboNucléique interférent
ARNm	Acide RiboNucléique messenger
BAC	Chromosome Artificiel Bactérien (Bacterial Artificial Chromosom)
Bp	Paire de base
DNase	DésoxyriboNucléase
dNTP	Désoxyribonucléotide Tri Phosphate
EST	Expressed Sequence Tag
Kb	Kilobase
Ks	Nombre de changements synonymes / nombre de sites synonymes
MgCl ₂	Chlorure de Magnésium
Mb	Mégabase (1 million de paires de bases)
mg	Milligramme
nm	Nanomètre
PCR	Réaction de polymérase en chaîne (Polymerase Chain Reaction)
pg	Picogramme
Poly-A	Poly-Adénosine
RNase	RiboNucléase
SNP	Polymorphisme d'une seule base (Single Nucleotide Polymorphism)
Taq polymerase	Taq DNA polymerase
U/ μ l	Unité/ μ l
°C	Degré Celsius
μ g/ μ l	microgramme / microlitre
μ l	microlitre
μ M	micromolaire
μ Em ⁻² s ⁻¹	microEinstein par mètre carré par seconde, énergie lumineuse

INTRODUCTION

Parmi les plantes cultivées, la polyploidie est reconnue d'intérêt agronomique, les plantes polyploïdes présentent en général des caractéristiques phénotypiques intéressantes pour l'agriculture. L'amélioration génétique de telles espèces nécessite l'approfondissement des connaissances concernant l'aptitude des polyploïdes à disposer de l'information génétique redondante.

Depuis une quinzaine d'années, de nombreuses études sont investies autant sur des polyploïdes établis depuis quelques milliers voire quelques millions d'années que sur des polyploïdes synthétisés *de novo* pour identifier les mécanismes à l'origine des avantages adaptatifs. Au cours de ces études, l'organisation génomique, et l'expression des gènes sont tout particulièrement considérées. Le contexte génétique et évolutif des espèces parentales à l'origine des allopolyploïdes, leur processus de formation et d'évolution depuis l'évènement d'hybridation font partie des nombreux paramètres qui influent sur l'état actuel des allopolyploïdes. Multiplier le nombre de modèles étudiés semble nécessaire pour observer les diverses configurations génétiques et transcriptomiques qui se déroulent au sein des polyploïdes.

La conciliation de plusieurs sous-génomes dans une même cellule entraîne des phénomènes de régulation de l'expression des gènes. La contribution relative des gènes de chaque sous-génome au transcriptome du polyploïde est un des aspects de l'expression des gènes qui a été étudié auprès de certains allopolyploïdes mais pas au sein d'un allopolyploïde récent provenant de l'hybridation de deux espèces diploïdes peu divergentes.

Le café fait partie des principaux produits du commerce mondial. Plus de 80 millions de personnes travaillent dans la culture, le traitement et le commerce du café. La production commerciale de café repose principalement sur deux espèces : *Coffea arabica* et *Coffea canephora*, représentant respectivement 65% et 35% de la production mondiale de café. *Coffea arabica* est une espèce allopolyploïde ($4n = 4x = 44$), la seule du genre *coffea*, elle provient d'un évènement d'hybridation récent entre deux génomes provenant d'espèces diploïdes ($2n = 2x = 22$) peu divergentes, *Coffea eugenoides* et *Coffea canephora*. *C. arabica* présente des capacités adaptatives à des variations environnementales dont ne bénéficient pas les espèces diploïdes du genre *Coffea*. Par ailleurs l'amélioration génétique de *C. arabica* est basée principalement sur l'exploitation des ressources génétiques des caféiers diploïdes. Elle consiste à introduire au sein des espèces cultivées, des gènes d'intérêt à partir des espèces diploïdes en croisant *C. arabica* avec certaines d'entre elles. L'étude de l'expression relative des gènes homéologues de *C. arabica* décrira un nouveau modèle d'allopolyploidie et viendra augmenter les connaissances concernant cet aspect de l'étude de l'expression des gènes au sein des polyploïdes. Cette étude informera en outre de l'influence de l'expression relative des sous-génomes sur les capacités adaptatives de *C. arabica* même dans le cas d'introgession.

SYNTHESE BIBLIOGRAPHIQUE

1 La polyplœidie

1.1 Définition

Le terme de polyplœide qualifie les individus ou les cellules qui présentent plus de deux lots de chromosomes. Plusieurs types de polyplœides sont distingués selon les événements qui sont à leur origine (Chen & Ni 2006), (Chen 2007). Les autopolyplœides au sein desquels des génomes provenant d'une même espèce sont juxtaposés sont composés de plusieurs jeux de chromosomes homologues. Chez les végétaux, de nombreuses espèces autopolyplœides sont répertoriées: des espèces cultivées comme la Pomme de terre, la Luzerne, le Soja et aussi des espèces sauvages ou ornementales comme les *Lantana*.

Les allopolyplœides qui résultent d'une hybridation interspécifique associée à un doublement de chromosomes présentent à la fois des paires de chromosomes homologues mais aussi des jeux de chromosomes apparentés (homéologues), plus ou moins divergents, provenant des espèces parentales. Des espèces cultivées comme le Coton, le Blé, le Colza, le Caféier, le Tabac, le Citronnier, le Concombre, l'Arachide mais aussi de nombreuses espèces sauvages comme la Spartine (*Spartina*), *Tragopogon* et *Senecio* sont des allopolyplœides.

1.2 Fréquence de la polyplœidie et paléopolyplœidie

La polyplœidie se rencontre chez de nombreux Eucaryotes, mais sa prévalence varie grandement. Elle se produit chez les Eucaryotes unicellulaires comme les levures (Scannell *et al.* 2006), mais aussi parmi les insectes et les vertébrés (Otto 2007), néanmoins ces événements sont rares: un seul cas d'organisme polyplœide a été observé chez les mammifères (Golden Vizcacha Rat *Pipanacoctomys aureus*) et aucun chez les oiseaux. La sensibilité du développement embryonnaire aux variations de dose de copies de gènes, notamment ceux liés au déterminisme sexuel expliquerait la fréquence faible d'événements de polyplœidisation au sein des animaux (Chen & Ni 2006).

L'avancée des techniques de marquage moléculaire et plus récemment des données de séquençage des génomes a permis de mettre en évidence des duplications localisées qui proviennent d'événements anciens de polyplœidie (Cui *et al.* 2006), (Blanc & Wolfe 2004). Des traces de duplication ont été détectées dans de nombreux génomes (Zhou *et al.* 2010). Les anciens polyplœides auraient subi au cours de l'évolution des remaniements chromosomiques, des mutations et des pertes de gènes. Ainsi chez les végétaux, après avoir longtemps considéré que 70% à 80% des Angiospermes étaient des polyplœides (Masterson 1994), on estime maintenant que l'ensemble des Angiospermes ont une origine polyplœide. L'existence d'événements anciens de polyplœidie (*i.e.* paléopolyplœidie) a été démontrée par différentes approches expérimentales :

- L'analyse de blocs de synténie présents en plusieurs copies et répartis dans un même génome : par cartographie, analyse de génome séquencé, ou par hybridation *in situ*, des blocs de synténie se dégagent entre espèces et au sein même des génomes (Salse *et al.* 2008).
- L'analyse du nombre de paires de gènes paralogues qui présentent un même taux de substitution synonyme par site (K_s) : il est possible de détecter des périodes évolutives au cours desquelles de nombreux gènes se trouvent dupliqués simultanément (Blanc & Wolfe 2004), (Cui *et al.* 2006). L'exemple de la figure 1 illustre comment détecter des événements anciens et récents de polyploïdie à partir d'EST (Expressed Sequence Tag). Si un événement de polyploïdie survient, un grand nombre de gènes sont dupliqués en même temps, et la probabilité de trouver des paires avec un K_s faible est très importante. Ceux-ci auront un K_s proche de 0 dans le cas d'autopolyploïdie, ou un K_s de l'ordre de celui existant entre les copies de gènes des espèces progénitrices s'il s'agit d'un allopolyploïde. Il en résultera un pic de paires de gènes autour du K_s moyen de ces copies dupliquées. Mais de la même façon, si un événement de duplication ancien est survenu, les gènes dupliqués conservés vont diverger (donc auront un $K_s > 0$) a priori avec une vitesse similaire. Sur le graphe, il s'agira d'un deuxième pic, plus réduit, et décalé vers les K_s plus élevés selon l'âge de l'événement ancien de polyploïdie. L'étude effectuée sur le coton et présentée dans la figure 1 illustre bien un événement récent de polyploïdie, et un second événement plus ancien.

En utilisant les bases de données d'EST, Cui *et al.* (Cui *et al.* 2006) ont montré que la paléopolyploïdie est extrêmement fréquente chez les Angiospermes, puisque presque toutes les espèces étudiées (à l'exception d'*Amborella*) présentent des traces de paléopolyploïdie dans leurs génomes.

L'étude de gènes dupliqués partagés par divers Eucaryotes (animaux, champignons, plantes) indique que les événements de duplication se sont produits dès le début de l'histoire évolutive des Eucaryotes (Zhou *et al.* 2010). Pour les Angiospermes, plusieurs événements de polyploïdisation sont rapportés au cours de leur histoire évolutive : à la base des lignées des Angiospermes, à l'origine d'un paléohexaploïde proche de la divergence des Eudicotes (Soltis *et al.* 2009), (Cui *et al.* 2006) et enfin plusieurs événements survenus sur différentes branches des Angiospermes (Fawcett *et al.* 2009), (Jiao *et al.* 2011).

1.3 La formation des polyplœides

La probabilité d'hybridation mais aussi la divergence entre les génomes de deux espèces conditionnent fortement la formation d'un allopolyploïde (Buggs *et al.* 2009), mais également la transmission de diversité des espèces parentales vers les polyplœides (Soltis & Soltis 1999). Certaines espèces ont des origines polyphylétiques (multiparentales), issues de croisements multiples et récurrents, d'autres polyplœides sont issus de croisements réciproques (populations avec les deux cytoplasmes parentaux possibles), c'est le cas pour *Tragopogon* et *Brassica*. Au contraire, certaines espèces polyplœides comme le coton, le blé ou *Arabidopsis* sont issues de peu de croisements à leur origine (Chen & Ni 2006).

Les espèces ayant une origine unique pourraient avoir moins de diversité à diffuser que celles qui ont une origine multiple et qui favorisent l'introduction d'une plus grande part de la diversité des espèces parentales. Cependant les pressions de

sélection (anthropiques et naturelles) peuvent éroder la variabilité transférée initialement. De plus d'autres mécanismes concernant la régulation de l'expression des gènes dupliqués chez les allopolyploïdes vont moduler ce phénomène et générer de la diversité aussi bien génétique que morphologique.

Plusieurs voies permettent d'aboutir à un allopolyploïde contenant des paires de chromosomes homologues et au moins deux jeux de chromosomes homéologues (Ramsey & Schemske 1998). Plusieurs modèles sont proposés. En deux étapes, l'allopolyploïde est formé par l'hybridation de deux espèces diploïdes suivie d'un doublement chromosomique de l'hybride F1. En une étape, la formation de l'allopolyploïde se ferait par la fusion de gamètes mâles et femelles non réduits de deux espèces diploïdes ou par l'hybridation directe entre deux autotétraploïdes. La formation d'autopolyploïdes pourrait suivre les mêmes modèles de formation à l'exception près qu'une seule espèce est impliquée dans les différents modèles.

Sachant que de nombreuses espèces végétales présentent des anomalies de la méiose, la voie de formation d'allopolyploïdes par la fusion de gamètes non réduits pourrait être le modèle prédominant (Chen & Ni 2006). Le déroulement classique de la méiose conduit à la formation de cellules à n chromosomes à partir d'un organisme à $2n$ chromosomes. Pour cela, deux divisions successives font suite à une seule étape de réplication de l'ADN. Lors de la première division, dite réductionnelle, les chromosomes homologues sont séparés, alors que ce sont les chromatides sœurs qui sont séparées lors de la seconde (division équationnelle). Le déroulement de la méiose peut être perturbé et peut conduire à la formation de gamètes possédant $2n$ chromosomes, gamètes dits « non-réduits ». On distingue globalement deux types de gamètes non réduits. Les gamètes du type FDR (pour « First-Division Restitution ») qui hériteront d'une chromatide (recombinée ou non) de chaque chromosome présent chez le parent. Le produit final est ainsi équivalent à une absence de première division méiotique. D'un point de vue mécanistique, plusieurs aberrations différentes du déroulement de la méiose peuvent conduire à la formation de ce type de gamètes (Bretagnolle & Thompson 1995). Le second type de gamètes non réduits sont les gamètes de type SDR pour « Second-Division Restitution », ils contiennent au contraire des paires de chromatides sœurs, résultant de problèmes ou d'une absence complète de deuxième division.

1.4 Etude de la polyplœidie

Depuis une quinzaine d'années, les conséquences de la polyplœidie sur les gènes, l'expression des gènes et l'évolution des génomes sont particulièrement étudiées. Selon les polyplœidies étudiées, les enseignements sont distincts et complémentaires (Yang *et al.* 2010). En effet des phénomènes différents de la polyplœidie peuvent être examinés selon le type de polyplœidie des plantes considérés : le doublement génomique dans le cas de l'autopolyploïdie, la fusion des génomes divergents dans le cas de l'allopolyploïdie. Les mécanismes de régulation de l'expression des gènes développés dans chacune des deux situations pourraient être variables (Wang *et al.* 2004). L'estimation de la période de l'évènement de polyplœidisation est aussi considérée, elle informe sur l'importance et le type des phénomènes d'évolution que certains polyplœidies ont subi depuis le doublement génomique ou l'hybridation des génomes à leur origine. Enfin dans le cas des allopolyploïdes, la divergence entre les espèces diploïdes qui sont à leur origine influerait sur les phénomènes qui permettraient de concilier les génomes homéologues dans le même noyau (Jackson & Chen 2010).

Pour certaines plantes, des polyplœidies artificiels ont été créés. Ces polyplœidies de synthèse ou néopolyploïdes présentent l'avantage de ne pas avoir subi les phénomènes d'évolution des gènes dupliqués, ils permettent d'étudier les premiers

effets au niveau génomique et transcriptomique des étapes suivant la fusion et/ou le doublement du génome en les comparant à des polyploïdes naturels. Ils facilitent l'étude de la polyploïdie dans des contextes plus simples en dissociant les effets du doublement génomique et de la fusion de génomes divergents. Ils permettent également d'étudier les effets du doublement génomique et de la fusion génomique sans l'impact de l'évolution du génome du polyploïde depuis l'évènement d'hybridation. Mais il faut noter que cette approche n'est pas toujours possible, pour certains polyploïdes, les plantes à l'origine de l'hybridation ne sont plus toujours disponibles ou ne sont pas connues (Peng *et al.* 2008).

2 Les effets de la polyploïdie

Les polyploïdes sont étudiés dans de nombreux contextes : génétique, génomique fonctionnelle, génomique comparative, génomique évolutive, phénotypique pour la sélection et l'amélioration... Dans les paragraphes suivants, seront présentés les effets de la polyploïdie qui se produisent avant tout au niveau du génome, les nombreux réarrangements génomiques intervenant après l'hybridation de deux génomes et enfin la régulation de l'expression des gènes dupliqués.

2.1 Les réarrangements génomiques

Les réarrangements génomiques sont les effets de plusieurs types d'évènements au sein du génome : des mouvements (inversions, translocations), des duplications, des insertions, ou des délétions. Ils concernent des fragments génomiques suffisamment grands pour créer des changements détectables dans l'ordre, la taille ou la position de marqueurs moléculaires. Parmi les Eucaryotes polyploïdes, ce sont essentiellement les plantes qui présentent ces modifications du génome (Hufton & Panopoulou 2009).

2.1.1 Changements rapides et dynamiques dans la structure des génomes

Lors de l'évènement de polyploïdisation, au moment de l'hybridation des génomes, de fortes contraintes se produisent pour l'organisation du nouveau génome. Des études de marquage moléculaire effectuées sur des polyploïdes de synthèse chez *Brassica* ont permis de mettre en évidence des changements rapides de l'organisation du génome, et de détecter des transmissions de fragments génomiques non-additifs au sein des nouveaux allopolyploïdes. Ces changements se traduisent par la disparition de fragments génomiques parentaux, et l'apparition de nouveaux fragments qui sont absents chez les parents de l'hybride. Lors des premières générations après l'hybridation, peu de changements génomiques sont constatés, les génomes homéologues apparaissent quiescents, ce n'est que vers la troisième génération que surviennent de nombreux changements génomiques (Gaeta *et al.* 2007), (Udall *et al.* 2005), (Osborn *et al.* 2003). A l'inverse de ces observations, les allopolyploïdes de synthèse de blé montrent dès l'hybridation de nombreux changements génomiques (5 à 14% du génome total du polyploïde) principalement des délétions spécifiques des génomes. Ces pertes de séquences spécifiques permettraient la reconnaissance des chromosomes homéologues facilitant ainsi la formation de bivalents pendant la méiose. A partir de la troisième génération, peu de variations sont observées dans la descendance (Feldman *et al.* 1997). Levy et Feldman

(Levy & Feldman 2002) ont suggéré en 2002 que les allopolyploïdes subissent une « phase de révolution » de changements génétiques et épigénétiques immédiatement après la formation de l'allopolyploïde suivie d'une « phase d'évolution » correspondant à des modifications et des réarrangements plus lents des séquences d'ADN. Contrairement aux changements rapides qui surviennent chez *Brassica* et chez le blé, les allopolyploïdes de synthèse chez le coton présentent peu de changements génomiques, Liu *et al* en 2001 (Liu *et al.* 2001) ont montré que presque tous les fragments AFLP observés présentaient un profil d'additivité. Des analyses phylogénétiques moléculaires ont indiqué que les gènes orthologues des espèces parentales évoluaient indépendamment à des taux similaires des gènes homéologues combinés dans les allopolyploïdes (Cronn *et al.* 1999). De la même façon, les changements génomiques dans les polyploïdes de *Spartina* se produisent à une fréquence faible (Salmon *et al.* 2005). Le coton et *Spartina* tolèrerait plus le doublement du génome et l'hybridation interspécifique que *Brassica* et le blé, ces exemples illustrent la diversité des phénomènes d'évolution moléculaire et de réaction des génomes à l'hybridation.

2.1.2 L'activation de transposons et la méthylation de l'ADN

Les éléments transposables représentent 50 à 80 % du génome des plantes (Chen & Ni 2006). Bien que la plupart soient inactifs dans leur génome respectif, certains peuvent être activés en réponse à un stress ou à un « genomic shock », par exemple confronter deux génomes divergents dans un même noyau après un événement de polyploïdisation. Ainsi la réactivation de transposons à ADN et de rétrotransposons est observée chez le blé, dans des allopolyploïdes d'*Arabidopsis* et dans des allopolyploïdes récents de *Spartina* (Kashkush *et al.* 2002), (Madlung *et al.* 2004), (Salmon *et al.* 2005). L'activation de transposons peut avoir un effet de régulateur cis ou trans pour les gènes voisins de la transposition (Wittkopp *et al.* 2004). Chez d'autres espèces allopolyploïdes, les phénomènes d'activation des transposons sont indissociables de mécanismes de régulation de l'expression des gènes et de mécanismes épigénétiques. Par exemple, chez *Spartina* (Salmon *et al.* 2005) et chez des allopolyploïdes de synthèse d'*Arabidopsis* (Madlung *et al.* 2004), l'état de méthylation de l'ADN est corrélé à l'activation des transposons, l'allopolyploïdisation perturberait l'organisation de la chromatine, levant temporairement le niveau de méthylation de l'ADN et permettant ainsi les transpositions des éléments mobiles dans le génome, ensuite un retour à un niveau de méthylation plus important réprimerait les transposons, assurant une certaine stabilité génomique.

2.2 Régulation de l'expression des gènes

Les variations phénotypiques observées chez les allopolyploïdes en les comparant aux diploïdes, ont amené les chercheurs à étudier l'expression et la régulation des gènes au sein des allopolyploïdes.

2.2.1 Estimation de l'expression des gènes

Dans les diverses études effectuées sur les allopolyploïdes, l'expression des gènes est analysée par deux démarches différentes : la première consiste à comparer le niveau d'expression des gènes de l'allopolyploïde à l'expression des mêmes gènes des parents diploïdes, la seconde méthode vise à estimer la contribution relative des

gènes homéologues (appartenant à chaque sous-génome de l'allopolyploïde) dans l'allopolyploïde.

Dans la première démarche, le niveau d'expression des gènes de l'allopolyploïde peut être classifié en deux catégories :

- additivité : elle est définie par certains auteurs comme la somme du niveau d'expression des parents (quantitativement et proportionnellement), alors que d'autres auteurs prennent comme valeur de l'additivité la moyenne du niveau d'expression des parents (« mid-parent value »).
- Non-additivité : elle reflète tout écart à l'additivité, le niveau d'expression du gène chez le polyploïde est supérieur ou inférieur au niveau d'expression du gène chez les parents diploïdes (Chen 2010).

Pour la seconde démarche, différentes méthodes sont utilisées pour estimer l'expression relative des gènes homéologues dans le transcriptome de l'allopolyploïde. Le rapport des quantités des transcrits homéologues est exprimé en pourcentage ou en proportion de l'expression totale. Un tétrapolyploïde (comme le coton, figure 7) dont le génome est composé de deux sous-génomes, peut présenter des niveaux d'expression relative des gènes homéologues variables, de la présence équivalente des deux sous-génomes jusqu'au biais de l'expression vers l'un ou l'autre des sous-génomes et parfois jusqu'à la disparition de l'expression de l'un d'eux.

2.2.2 Mécanismes de régulation de l'expression

L'hybridation de deux génomes va confronter les gènes et les facteurs de régulation ayant évolué indépendamment dans les espèces progénitrices. Il en résulte que l'expression des gènes homéologues dans le contexte polyploïde pourra être modifiée par des mécanismes génétiques ou épigénétiques. En 2003, Osborn résume par la figure 8 les principaux phénomènes qui agissent sur l'expression des gènes au sein des polyploïdes (Osborn *et al.* 2003).

Les réarrangements génomiques et les effets « dose »

Les mécanismes génétiques intervenant dans la régulation des gènes homéologues, sont les réarrangements génomiques présentés précédemment et leurs effets dose-dépendants sur les gènes (nombre de copies du gène considéré). Les réarrangements, par délétion ou duplication de régions génomiques, vont entraîner la présence d'une même région génomique en un nombre variable de copies dans un polyploïde, ce qui peut entraîner pour certains gènes un déséquilibre de l'expression. Les gènes sous contrôle de régulateurs vont avoir un niveau final d'expression qui va différer entre les espèces diploïdes parentales et l'allopolyploïde. Birchler *et al.* (Birchler *et al.* 2005) démontrent que les translocations non réciproques ou les délétions déséquilibrent les gènes « dose-dépendant » dans ces régions. Les réarrangements génomiques modifient ainsi l'expression des copies homéologues avec soit une surexpression (en cas de duplication du gène ou de ses séquences régulatrices), soit un « silencing » définitif (en cas de délétion ou mutation ponctuelle non synonyme délétère : par exemple mutation dans le site d'épissage de l'intron conduisant à sa conservation, codon stop prématuré conduisant à une protéine tronquée non fonctionnelle).

Les insertions d'éléments transposables

L'insertion d'élément transposable est aussi un mécanisme génétique qui va modifier la régulation de l'expression des gènes homéologues. La conséquence peut être un « silencing » définitif si l'élément transposable de type I (copier-coller) s'insère dans la séquence promotrice, entre la séquence promotrice et le gène, ou dans un exon. La conséquence peut aussi être réversible pour une insertion d'un élément transposable

de type II (couper-coller) (Parisod *et al.* 2010). De plus, les promoteurs d'éléments transposables insérés à côté de gènes peuvent modifier le niveau d'expression de celui-ci, comme illustré chez la souris où l'insertion d'un élément transposable à proximité d'un gène de la couleur de pelage conduit à sa surexpression (Rakyan *et al.* 2001). De telles situations sont envisageables chez les polyploïdes, pour lesquels des nouvelles insertions d'éléments transposables sont observées.

Les modifications transcriptionnelles par méthylation

Dans l'exemple précédent des souris, l'élément transposable inséré a des niveaux variables de méthylation selon les descendants, ce qui conduit à faire varier le niveau d'expression du gène de couleur du pelage. Dans le cas des polyploïdes, les éléments transposables sont souvent la cible de méthylation et de compaction de la chromatine par les histones (Parisod *et al.* 2010) mais leur expression peut être transitoirement déréprimée lors de l'hybridation.

Les modifications transcriptionnelles ou post transcriptionnelles par ARNi.

La découverte des ARN interférents a été une des grandes révélations de la fin du vingtième siècle. Ces petits ARN possèdent une séquence complémentaire d'ARN de gènes qui en s'hybridant à l'ARN messager du gène va conduire à sa dégradation. La rencontre des ARNi des deux génomes peut perturber l'expression d'un gène, que ce soit en réprimant l'expression des gènes directement (régulation post transcriptionnelle) ou via sa cascade de gènes régulateurs (transcriptionnelle), (Chen & Ni 2006). Etant donné que les espèces parentales ont évolué indépendamment, l'importance de la régulation par ces ARNi peut être variable et perturber l'expression des copies homéologues en contexte polyploïde (Chen *et al.* 2008).

Les interactions des séquences régulatrices : régulation en cis et en trans

L'hybridation va combiner dans le même génome les gènes homéologues et les gènes des facteurs de transcription. Les espèces hybridées étant plus ou moins divergentes, des interactions peuvent se produire entre les facteurs de transcription, qui peuvent agir mutuellement sur les gènes homéologues. Selon les niveaux d'interaction, plusieurs combinaisons sont envisageables (Chen & Ni 2006).

Modification des rapports stoechiométriques

Selon l'hypothèse « dosage balance hypothesis », pour un gène donné à nombre de copies constant, la variation de l'équilibre stoechiométrique des copies de gènes en amont qui ont un rôle dans sa transcription va moduler son expression. Birchler *et al.* ont proposé un modèle pour représenter ce phénomène lors de l'hybridation entre deux espèces (Birchler *et al.* 2005).

Lister individuellement les mécanismes de régulation donne une vision simplifiée de la régulation des gènes homéologues. Il est évident que certains des phénomènes de régulation interagissent et sont parfois indissociables (par exemple : les éléments transposables et l'état de méthylation de l'ADN), la plupart de ces phénomènes ne s'excluent pas les uns des autres et régulent les gènes homéologues de façon combinée. Le schéma de Jackson 2010 est un des schémas les plus complets qui illustre parfaitement les interactions des différents mécanismes observés chez les allopolyploïdes.

3 Conséquences évolutives de la polyploïdie

3.1 Evolution des gènes dupliqués

Chaque espèce possède *a priori* au moins une copie de chaque gène rencontré dans les autres espèces (gènes orthologues). Lors de l'hybridation, ces copies se retrouvent confrontées, et le devenir de ces copies a été au cœur des réflexions sur l'évolution des génomes dès Ohno (1970) qui leur suppose différentes possibilités d'évolution à court, moyen et long terme. Hahn interprète ces trois possibilités d'évolution en considérant les séquences codantes et les séquences régulatrices des gènes dupliqués (Hahn 2009).

- Pseudogénéisation : les deux gènes peuvent conserver tous deux leur fonction d'origine dans les premières générations, mais une copie peut ne pas être transcrite (de façon aléatoire ou stochastique) sur le moyen ou long terme, perte qui est compensée par l'autre copie homéologue. Plusieurs causes sont possibles, soit en raison de la dégradation d'une copie en gène non fonctionnel accumulant des mutations délétères à sa fonction (pseudogène), ou par recombinaison homéologue (translocation ou conversion génique). La recombinaison illégitime peut également conduire directement à la perte de séquence. De la même façon, un gène mis sous silence pourra être sujet à pseudogénéisation.
- Néofonctionnalisation : sous cette notion coexistent la répartition par organe de l'expression d'un gène (court terme) avec l'acquisition d'une nouvelle fonction (sur le long terme). Ainsi l'expression d'un gène dans des organes nouveaux par rapport aux espèces parentales est considérée comme de la néofonctionnalisation. Sous pression de sélection, une copie de gène peut accumuler des mutations non synonymes qui modifient la protéine et lui donnent une nouvelle fonction (nouvelle activité enzymatique en modifiant le site réactionnel par exemple).
- Sous-fonctionnalisation : Chez les polyploïdes, le « silencing » de l'expression entre copies homéologues dépendant de l'organe est considéré comme une marque de sous fonctionnalisation. En effet, si ce phénomène perdure, les deux copies seront conservées car elles se spécialisent selon l'organe ou le tissu cellulaire et partagent la fonction originelle du gène. Ce modèle est appelé DDC Duplication-Dégénération-Complémentation. La principale différence avec la néofonctionnalisation réside dans le fait que les deux copies n'ont fait que fragmenter la fonction ancestrale en deux (les tissus concernés par l'expression des copies ne sont pas nouveaux dans le polyploïde par rapport aux espèces parentales).

Selon les plantes polyploïdes et la date de l'évènement de polyploïdisation, plusieurs états d'évolution des gènes dupliqués sont proposés. L'allopolyloïde fixerait en quelque sorte un état hétérozygote qui lui serait avantageux parce qu'il disposerait d'un panel de gènes homéologues. Les gènes maintenus dupliqués fourniraient à l'allopolyloïde de nombreuses possibilités d'évolution le conduisant vers la spéciation en lui facilitant les mécanismes d'adaptation face à des conditions environnementales variables (Flagel *et al* 2009), (Rapp & Wendel 2005).

Ha *et al* (Ha *et al.* 2007) ont mis en évidence que la divergence entre les gènes homéologues est plus forte sous l'action de stress environnementaux que sous l'action de changements développementaux internes. La divergence rapide entre les gènes dupliqués en réponse à un stress abiotique et biotique faciliterait la sous-fonctionnalisation, la néofonctionnalisation et le développement de mécanismes pour s'adapter à des changements de l'environnement. Mais l'allopolyloïde disposerait

d'un temps limité pour ces évolutions, il perdrait progressivement, à l'échelle de temps de l'évolution, les copies des gènes dupliqués en accumulant des mutations, l'allopolyploïde reviendrait à l'état de gènes « singleton », il s'agirait de la diploïdisation. Un cycle de polyploïdisation et diploïdisation pourrait alors se produire à nouveau pour permettre à l'allopolyploïde de disposer de nouvelles sources de diversité (Doyle *et al.* 2008). Dans le schéma proposé, Doyle *et al.* font une synthèse des mécanismes génétiques et épigénétiques qui agissent sur l'expression des gènes homéologues, des phénomènes d'évolution des gènes dupliqués qui vont provoquer des caractéristiques phénotypiques (qui peuvent être avantageuses) et enfin sur la récurrence des événements de polyploïdie. En effet, la polyploïdisation est considérée comme un facteur de diversification, et spéciation, de nombreuses Angiospermes présentant plusieurs événements de polyploïdisation supporteraient l'idée d'un tel mécanisme d'évolution.

Des situations illustrant ces mécanismes d'évolution ont été observées sur plusieurs polyploïdes anciens, chez le coton et le blé (Flagel *et al.* 2009), (Liu & Adams 2007), (Stamati *et al.* 2009). Dans ces études, des gènes homéologues étudiés ont présentés des profils d'expression de sous-fonctionnalisation, de « silencing » spécifiques d'organe sont décrits. En revanche des néopolyploïdes, ou des polyploïdes provenant d'évènements récents permettent d'évaluer la vitesse de mise en place des phénomènes génétiques et épigénétiques, par ailleurs ils ont permis de constater que l'hybridation de génomes lors de la polyploïdisation a plus d'effets sur l'expression des gènes que le doublement du génome (Adams 2007).

3.2 Polyploïdie et adaptation

On estime qu'entre 2% et 15% des événements de spéciation des Angiospermes seraient causés par un événement de polyploïdisation (Wood *et al.* 2009). Certains auteurs supposent que la polyploïdie serait également un facteur clé de la diversification des espèces végétales (Soltis & Soltis 1999), en leur permettant de mieux s'adapter à des environnements plus extrêmes que ceux tolérés par leurs espèces parentales. L'observation des allopolyploïdes et leur comparaison avec les espèces parentales diploïdes mettent en évidence des caractères phénotypiques différents et parfois nouveaux, qui jouent sans doute un rôle important dans l'adaptation des polyploïdes à de nouveaux environnements ou à de nouvelles niches écologiques. Ainsi un niveau élevé de tolérance à la sécheresse, la résistance à des parasites, l'apomixie, une variation de la période de floraison... ont permis aux polyploïdes de mieux survivre que leurs parents diploïdes à des environnements défavorables : hautes altitudes, climat froid... (Chen 2007).

Récemment Fawcett *et al.* (Fawcett *et al.* 2009) ont montré que la datation des événements anciens de polyploïdie présents chez de nombreuses angiospermes correspondrait à celle de la crise Crétacé-tertiaire. Ces événements indépendants sont survenus dans une période de grands bouleversements climatiques et ont accompagné l'établissement des Angiospermes et le déclin des Gymnospermes, ce qui laisse entrevoir l'importance de la polyploïdie dans l'adaptation des espèces aux changements climatiques.

Mais la polyploïdie ne semble pas seulement un avantage dans les environnements défavorables, en effet de nombreuses espèces domestiquées sont polyploïdes. L'augmentation de la taille des organes pourrait être un des facteurs qui favorise les espèces polyploïdes parmi les espèces cultivées (Mizukami 2001), mais aussi leur plasticité phénotypique notamment dans les premières générations (Jackson & Chen 2010), l'état de vigueur hybride qui semble fixé et enfin le mode de régulation des gènes par effet dose (Chen 2007) participeraient à la domestication par

la sélection. Mais certaines caractéristiques des polyploïdes peuvent être aussi des désavantages, ainsi l'augmentation de la taille des génomes qui induit un volume cellulaire plus important entrainerait un métabolisme cellulaire et un développement plus lent (Otto 2007).

4 Le caféier

4.1 Phylogénie, et origine des caféiers

Les caféiers appartiennent à la famille des *Rubiaceae* qui compte plus de 6000 espèces. Ce sont des plantes pérennes présentant une taille très variable qui peut aller de petits arbustes jusqu'à des arbres d'environ dix mètres. Dans cette famille, les caféiers constituent la tribu *Coffea* caractérisée par une placentation dite « cofféenne », où la partie ventrale de l'albumen des graines forme une invagination. Aujourd'hui deux genres sont considérés : le genre *Coffea* L. subdivisé en deux sous-genre *Coffea* et *Baracoffea*, et le genre *Psilanthus* spp. lui même divisé en deux sous-genre *Psilanthus* et *Afrocoffea*. Des différences au niveau de la position des fleurs et du mode de développement sont à l'origine de cette séparation (Leroy 1980, Bridson 1987, Bridson et Verdcourt 1988). Le genre *Coffea* sous genre *Coffea* comprend plus de 95 taxons (Davis et 2006), et notamment les deux principales espèces cultivées : *Coffea arabica* et *Coffea canephora*.

Le développement des outils d'analyse moléculaire a permis de structurer les espèces du sous genre *Coffea*. Les variations de l'ADN chloroplastique et nucléaire de taxons représentatifs du genre *Coffea* sous-genre *Coffea* ont été étudiées. Les analyses phylogénétiques des données moléculaires à partir des deux génomes analysés (e.g. nucléaire et chloroplastique) suggèrent l'existence d'au moins 4 groupes phylogénétiques relativement peu différenciés. Ces groupes d'espèces de caféiers (*Coffea* sous genre *Coffea*) recouvrent les trois ensembles biogéographiques : zone guinéo-congolaise (1 groupe), Afrique de l'Est et centrale (2 groupes), et Madagascar, l'archipel des Comores et les îles Mascareignes (1 groupe) (Lashermes *et al.* 1997) (figure 17). Une étude complémentaire récente de séquences intergéniques du génome chloroplastique indique que la diversification des espèces au sein du genre *Coffea* s'est produite par radiation rapide et récente 100 000 à 460 000 ans selon la région chloroplastique considérée (Anthony *et al.* 2010). L'analyse comparative d'une région génomique de *Coffea* (ordre des *Gentianaceae* dans les *Asterideae*) et des régions orthologues d'espèces du groupe des *Rosideae* (*Vitis vinifera*, *Populus Trichocarpa* et *Arabidopsis thaliana*) a montré que ces deux groupes proviennent du même génome ancestral hexaploïde pour lequel le processus de diploïdisation était bien avancé au moment de la divergence entre les deux groupes et qu'aucun événement supplémentaire de polyploïdisation ne s'est produit au cours de l'évolution du genre *Coffea* (Cenci *et al.* 2010)

Les espèces du sous genre *Coffea* sont toutes diploïdes ($2n = 2x = 22$) à l'exception de l'espèce *C. arabica* qui est tétraploïde ($2n = 4x = 44$). L'analyse des caryotypes des 11 chromosomes du génome de caféier, rapportée par Sybenga (1959), met en évidence l'existence d'un génome de base composé de 11 chromosomes ($n = 11$) présentant une taille moyenne réduite (1 à 3 μm) et peu variable. En revanche la quantité d'ADN nucléaire est très variable entre les différentes espèces de caféier, avec des valeurs comprises entre 0.95 et 1.78 pg d'ADN par noyau, jusqu'à 2.61 pg d'ADN par noyau pour *C. arabica* qui possède deux fois plus de chromosomes (Cros *et al.* 1995). La taille

du génome de *C. arabica* est de 1300 Mb approximativement et de 710 Mb pour *C. canephora* (Anthony & Lashermes 2005).

C. arabica est une espèce amphidiploïde (Grassias & Kammacher 1975). Par des approches moléculaires et notamment l'utilisation de marqueurs chloroplastiques et ribosomiques (Cros *et al.* 1998), (Lashermes *et al.* 1995) il a été possible de vérifier une forte affinité entre *C. arabica* et des espèces du groupe des diploïdes originaires d'Afrique centrale et de l'ouest (*C. eugenioides*, *C. canephora*, *C. congensis*, *C. brevipes*, et *Coffea sp* « Mouloundou »). Une étude plus approfondie appuyée sur des analyses RFLP (en utilisant des sondes spécifiques de locus) et cytogénétiques (par hybridation de l'ADN génomique des espèces potentielles avec les chromosomes de *C. arabica*) a permis de démontrer la présence de deux lots de chromosomes E^a et C^a, qui proviendraient respectivement de formes ancestrales des espèces actuelles *C. eugenioides* (parent femelle, génome E^a) et *C. canephora* (parent mâle, génome C^a). Plus récemment des données provenant de séquences ont permis de confirmer cette hypothèse.

Ainsi l'hybride ancestral tétraploïde pourrait résulter de croisements interspécifiques faisant intervenir des diplo-gamètes (2n) dont la formation aurait été favorisée par des conditions climatiques particulières (figure 18). Il est aussi possible que la polyploïdisation ait eu lieu après la formation d'un hybride interspécifique diploïde suivie d'un doublement chromosomique. Les variations climatiques importantes pendant le quaternaire et leurs implications sur la distribution des différentes espèces de caféiers rendent difficiles l'interprétation de l'origine géographique de *C. arabica*. Toutefois, Lashermes *et al.* (Lashermes *et al.* 1999) ont suggéré le versant nord-ouest du rift africain comme zone d'origine de *C. arabica*, zone géographique que les espèces parentales *C. canephora* et *C. eugenioides* auraient occupé simultanément. Par la suite, l'espèce *C. arabica* aurait colonisé la région comprenant le sud du Soudan, le Nord du Kenya et le sud-ouest de l'Ethiopie qui est considéré comme son centre de diversité. L'analyse des fréquences de substitution de régions génomiques orthologues des espèces *C. arabica*, *C. canephora* et *C. eugenioides* a révélé que cet événement d'hybridation entre les deux espèces diploïdes ne s'est pas produit avant 10 000 à 50 000 ans. La faible divergence de ces mêmes régions génomiques, la présence d'éléments transposables partagés et une synténie des gènes presque parfaite indiquent que les génomes diploïdes à l'origine de *C. arabica* sont très peu divergents (Cenci *et al.* article soumis). Malgré cette faible divergence entre ses sous-génomes constitutifs, *C. arabica* se comporte sur le plan méiotique comme un diploïde, c'est-à-dire que seuls les chromosomes homologues s'hybrident sous l'action de facteurs de régulation de l'appariement (Lashermes 2000).

4.2 Caractéristiques phénotypiques des caféiers

De façon globale, les caféiers se répartissent dans les zones équatoriales et inter-tropicales, ils sont sensibles aux facteurs écologiques liés à la température et à l'humidité. Les plus fortes limitations climatiques sont le gel et la sécheresse. Le gel qui peut se produire en altitude, provoque des dommages létaux pour les arbres, la sécheresse qui peut être plus fréquente est considérée comme le principal stress environnemental pour la culture de caféiers. Les effets de la sécheresse sont considérables sur le rendement des parcelles productives (DaMatta & Ramalho 2006). Mais hormis ces conditions extrêmes, les caractéristiques écologiques des caféiers et les capacités adaptatives pour résister à des environnements stressants diffèrent selon les espèces de caféier considérées. *C. canephora* peut être cultivé sur des terres à faible altitude (du niveau de la mer jusqu'à 800 m) dans des régions à climat chaud et

humide ayant une température annuelle moyenne de 22 et 26 °C. *C. eugenoides* qui n'est pas une espèce cultivée est endémique de régions présentant une température annuelle moyenne de 18 à 22°C. Ces deux espèces tolèrent mal les fortes fluctuations de température. En revanche *C. arabica* dont la température optimale annuelle moyenne est de 20 à 24 °C peut s'adapter à des augmentations de température allant jusqu'à 25 ou 26 °C et à l'inverse à des baisses de température allant jusqu'à 18 °C. Par conséquent *C. arabica* peut être cultivé entre 1000 et 2000 m d'altitude dans les régions équatoriales et à des altitudes plus basses (400 à 1200 m) dans les régions plus éloignées de l'équateur (DaMatta & Ramalho 2006).

4.3 Culture des caféiers, amélioration de *C. arabica*

Le café est l'une des marchandises agricoles d'exportation les plus précieuses au monde. Pour les pays situés dans les régions tropicales et sub-tropicales, l'exportation du café a une grande importance sociale et économique (International Coffee Organization, <http://www.ico.org/>). En 2008, en Amérique Latine, en Afrique et en Asie, la production de café était la principale source de revenus pour de nombreux pays, et un moyen de subsistance pour 80 millions de personnes.

Deux espèces sont principalement cultivées : *C. arabica* pour 65 % de la production mondiale et *C. canephora* connu sous le nom commercial de Robusta pour environ 35 %. Les principales variétés de *C. arabica* cultivées en Amérique Latine (Caturra, Catuai, Mundo Novo) hautement productives donnent un café de bonne qualité aux arômes fins et à faible teneur en caféine alors que *C. canephora* fournit un café plus fort avec une teneur en caféine plus importante. Les variétés de *C. arabica* cultivées furent sélectionnées à partir d'une base génétique étroite dispersée au début du XVIII^{ème} siècle (figure 19) par les puissances coloniales.

Cependant, en conséquence du nombre réduit d'individus fondateurs à l'origine des caféiers en plantation à travers le monde, ainsi que du mode de reproduction de *C. arabica*, les variétés cultivées présentent une très faible diversité génétique (Anthony *et al.* 2002) rendant la culture de cette espèce particulièrement vulnérable aux parasites et ravageurs.

L'amélioration génétique de l'espèce *C. arabica* est basée essentiellement sur l'exploitation des ressources génétiques des caféiers diploïdes. Elle consiste à introduire plus de variabilité génétique au sein des variétés cultivées de l'espèce, par introgression de fragments génomiques des espèces diploïdes. Le croisement et l'obtention d'hybrides interspécifiques entre l'espèce tétraploïde *C. arabica* et les espèces diploïdes ne constituent pas un obstacle au transfert de gènes par voie sexuée. Après doublement chromosomique du parent diploïde, des hybrides interspécifiques, et intergénériques, tétraploïdes relativement fertiles peuvent être obtenus (Couturon *et al.* 1998), (Charrier & Eskes 1988). L'analyse génétique de ces hybrides et de leurs descendants a montré une fréquence importante de recombinaisons génétiques entre génomes et un niveau élevé d'introgression (Lashermes *et al.* 2000), (Herrera *et al.* 2002), (Prakash *et al.* 2002). Les collections de caféiers diploïdes, mais aussi des hybrides interspécifiques naturels (exemple : Hybride de Timor) représentent des réservoirs de diversité qui sont exploités dans les programmes d'amélioration

5 Les objectifs de l'étude

C. arabica, se caractérise par une faible diversité génétique mais aussi par une grande variabilité phénotypique. La polyploïdie conférerait à *C. arabica* une grande capacité adaptative. Dans l'étude des capacités adaptatives des polyploïdes et des

mécanismes liés à la polyploïdie, le contexte génétique et évolutif des espèces polyploïdes est particulièrement important. Jusqu'à présent l'étude de l'expression des gènes homéologues n'a pas été effectuée sur des polyploïdes récents et provenant de l'hybridation de deux espèces peu divergentes. Notre étude a donc pour objectif d'examiner un nouvel exemple de polyploïde, *C. arabica*, allopolyploïde récent entre deux espèces diploïdes peu divergentes.

Contrairement aux espèces diploïdes du genre *Coffea* qui ne tolèrent pas de grandes variations des conditions environnementales, *C. arabica* est moins sensible aux variations de température. Cette particularité pourrait être attribuée au caractère polyploïde de *C. arabica*. La compréhension des mécanismes propres à l'état de polyploïdie paraît un point particulièrement important à appréhender et à comprendre pour l'amélioration génétique de l'espèce. Il semblerait que plusieurs phénomènes complexes (expression et régulation de l'expression des gènes dupliqués, évolution de ces gènes) soient à l'origine des avantages adaptatifs liés à la polyploïdie. Dans le programme de recherche « Etude des modifications génomiques structurales et fonctionnelles associées à l'allopolyploïdie », l'étude de l'expression des gènes est effectuée par deux approches différentes comme dans la plupart des études portant sur ce sujet pour d'autres espèces polyploïdes. L'expression des gènes de l'allopolyploïde peut être étudiée d'une part en comparant le transcriptome de l'allopolyploïde à celui des deux parents diploïdes et d'autre part en identifiant la contribution relative des deux sous-génomes dans le transcriptome de l'allopolyploïde. Dans l'étude présentée, il s'agit d'étudier la contribution relative de l'expression des deux sous-génomes dans *C. arabica*. Comment *C. arabica* concilie-t-il ses deux sous-génomes ? Cette conciliation participe-t-elle aux capacités adaptatives de *C. arabica* ? *C. arabica* a-t-il développé des mécanismes particuliers de néo-fonctionnalisation ou sous-fonctionnalisation pour concilier les deux sous-génomes ? Dans le cas de *C. arabica* introgressé, en raison de l'importance de cette voie d'amélioration, comment l'expression des deux sous-génomes (dont l'un est nouvellement intégré au génome de la plante) se déroule-t-elle ? Pour répondre à ces nombreuses questions, deux expérimentations ont été développées. L'expression relative des sous-génomes de *C. arabica* provenant des parents diploïdes proches des caféiers actuels *C. eugenoides* (E^a) et *C. canephora* (C^a) a été étudiée dans différents organes de plantes cultivées dans deux conditions de culture contrastées correspondant chacune aux conditions favorables au développement des espèces diploïdes parentales. Ensuite l'expression relative des sous-génomes dans *C. arabica* introgressé a été analysée dans différents organes de la plante. L'expression relative des sous-génomes a été mesurée sur une région génomique pour laquelle on dispose de l'information permettant de distinguer les gènes homéologues. La mesure de l'expression relative des gènes homéologues a été effectuée par l'analyse du rapport de SNP sur des chromatogrammes de réactions de séquence de type Sanger.

BIBLIOGRAPHIE

- Adams KL, & Wendel JF. 2005. Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*. 8(2): 135-141.
- Adams KL, Percifield R, Wendel JF. 2004. Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. *Genetics*. 168(4): 2217-2226.
- Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences of the United States of America*. 100(8): 4649-4654.
- Adams KL. 2007. Evolution of duplicate gene expression in polyploid and hybrid plants. *Journal of heredity*. 98(2): 136-1
- Anthony F, Topart P, Anzueto F, Astorga C, Bertrand B, Graziosi G, Lashermes P. 2002. The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theoretical and Applied genetics*. 104: 894-900
- Anthony F, Diniz LEC, Combes MC, Lashermes P. 2010. Adaptive radiation in *Coffea* subgenus *Coffea* L. (Rubiaceae) in Africa and Madagascar. *Plant Systematics and Evolution*. 285(1-2): 51-64.
- Birchler JA, Riddle NC, Auger DL, Veitia RA. 2005. Dosage balance in gene regulation: biological implications. *Trends in Genetic*. 21(4): 219-226.
- Blanc G, Wolfe KH. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *The Plant Cell*. 16(7): 1667-1678.
- Bretagnolle F, Thompson JD. 1995. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist*. 129(1): 1-22.
- Buggs RJA, Elliott NM, Zhang L, Koh J, Viccini LF, Soltis DE, Soltis PS. 2010. Tissue-specific silencing of homoeologs in natural populations of the recent allopolyploid *Tragopogon mirus*. *The New Phytologist*. 186(1): 175-183.
- Buggs RJA, Soltis PS, Soltis DE, 2009. Does hybridization between divergent progenitors drive whole-genome duplication? *Molecular Ecology*. 18(16): 3334-3339.
- Cenci A, Combes MC, Lashermes P. 2010. Comparative sequence analyses indicate that *Coffea* (Asterids) and *Vitis* (Rosids) derive from the same paleo-hexaploid ancestral genome. *Molecular Genetics and Genomics*. 283(5): 493-501.
- Charrier A, Eskes AB. 1988. Les caféiers. L'amélioration des plantes tropicales. coédition Cirad-Orstom. 171-196.

- Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, Springer NM, Wendel JF. 2009. Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). *Genetics*. 182(2): 503-517.
- Chen M, Ha M, Lackey E, Wang J, Chen ZJ. 2008. RNAi of *met1* reduces DNA methylation and induces genome-specific changes in gene expression and centromeric small RNA accumulation in *Arabidopsis* allopolyploids. *Genetics*. 178(4): 1845-1858.
- Chen ZJ. 2010. Molecular mechanisms of polyploidy and hybrid vigor. *Trends in Plant Science*. 15(2): 57-71.
- Chen ZJ. 2007. Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids. *Annual review of plant biology*. 58: 377-406.
- Chen ZJ, Ni Z. 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *BioEssays* 28(3): 240-252.
- Couturon E, Lashermes P, Charrier A. 1998. First intergeneric hybrids (*Psilanthusebracteolatus* Hiern × *Coffea arabica* L.) in coffee trees. *Canadian Journal of Botany*. 76: 542-546.
- Cronn RC, Small RL, Wendel JF. 1999. Duplicated genes evolve independently after polyploid formation in cotton. *Proceedings of the National Academy of Sciences of the United States of America*. 96(25): 14406-14411.
- Cros J, Combes MC, Chabrilange N, Duperray C, Monnot des Angles A, Hamon S. 1995. Nuclear DNA content in the subgenus *Coffea* (Rubiaceae): inter- and intra-specific variation in African species. *Canadian Journal of Botany*, 73: 14-20.
- Cros J, Combes MC, Trouslot P, Anthony F, Hamon S, Charrier A, Lashermes P. 1998. Phylogenetic Analysis of Chloroplast DNA Variation in *Coffea* L. *Molecular Phylogenetics and Evolution*. 9(1): 109-117.
- Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, Albert VA, Ma H, dePamphilis CW. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research*. 16(6): 738-749.
- DaMatta FM, Ramalho JDC, 2006. Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology*. 18(1): 55-81.
- Dong S, Adams KL. 2011. Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. *New Phytologist*. 190(4): 1045-1057.
- Doyle JJ, Flagel LE, Andrew HP, Rapp RA, Soltis DE, Soltis PS, Wendel JF. 2008. Evolutionary Genetics of Genome Merger and Doubling in Plants. *Annual Review of Genetics*. 42(1): 443-461.

- Fawcett JA, Maere S, Van de Peer Y. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proceedings of the National Academy of Sciences*. 106(14): 5737-5742.
- Feldman M, Liu B, Segal G, Abbo S, Levy AA, Vega JM. 1997. Rapid Elimination of Low-Copy DNA Sequences in Polyploid Wheat: A Possible Mechanism for Differentiation of Homoeologous Chromosomes. *Genetics*. 147(3): 1381-1387.
- Flagel LE, Chen L, Chaudhary B, Wendel JF. 2009. Coordinated and fine-scale control of homoeologous gene expression in allotetraploid cotton. *The Journal of Heredity*. 100(4): 487-490.
- Flagel LE, Wendel JF. 2009. Gene duplication and evolutionary novelty in plants. *New Phytologist* 183(3): 557-564.
- Flagel L, Udall J, Nettleton D, Wendel J. 2008. Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. *BMC Biology*. 6: 1-16.
- Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC. 2007. Genomic Changes in Resynthesized *Brassica napus* and Their Effect on Gene Expression and Phenotype. *The Plant Cell*. 19(11): 3403-3417.
- Ha M, Li WH, Chen ZJ. 2007. External factors accelerate expression divergence between duplicate genes. *Trends in Genetic*. 23(4): 162-166.
- Hahn MW. 2009. Distinguishing among evolutionary models for the maintenance of gene duplicates. *The Journal of Heredity*. 100(5): 605-617.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*. 41: 95-98.
- Herrera JC, Combes MC, Anthony F, Charrier A, Lashermes P. 2002. Introgression into the allotetraploid coffee (*Coffea arabica* L.): segregation and recombination of the *C. canephora* genome in the tetraploid interspecific hybrid (*C. arabica* × *C. canephora*). *Theoretical and Applied Genetics*. 104(4): 661-668.
- Hufton AL, Panopoulou G. 2009. Polyploidy and genome restructuring: a variety of outcomes. *Current Opinion in Genetics & Development*. 19(6): 600-606.
- Jackson S, Chen ZJ. 2010. Genomic and expression plasticity of polyploidy. *Current Opinion in Plant Biology*. 13(2): 153-159.
- Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y., Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature*. 473(7345): 97-100.
- Kashkush K, Feldman M, Levy A. 2002. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nature Genetics*. 33(1): 102-106.

- Lashermes P, Combes MC, Cros J, Trouslot P, Anthony F, Charrier A. (1995). Origin and genetic diversity of *Coffea arabica* L. based on DNA molecular markers. In : 16ème Colloque Scientifique International sur le Café (Kyoto, Japon, 9-14 Avril 1995) Association Scientifique Internationale du Café, Paris, France. Volume II: 528-536,
- Lashermes P. 2000. Brief communication. Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interspecific Hybrid *C. arabica* X *C. canephora*. *Journal of Heredity*. 91(1): 81-85.
- Lashermes P, Andrzejewski S, Bertrand B, Combes MC, Dussert S, Graziosi G, Trouslot P, Anthony F. 2000. Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theoretical and Applied Genetics*, 100(1): 139-146.
- Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A. 1999. Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics*. 261(2): 259-266.
- Lashermes P, Combes MC, Trouslot P, Charrier A. 1997. Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. *Theoretical and Applied Genetics*. (94): 947-955.
- Lashermes P, Combes MC, Ribas A, cenci A, Mahé L, Etienne H. 2010. Genetic and physical mapping of the SH3 region that confers resistance to leaf rust in coffee tree (*Coffea arabica* L.). *Tree Genetics & Genomes*. 6(6): 973-980.
- Levy AA, Feldman M. 2002. The Impact of Polyploidy on Grass Genome Evolution. *Plant Physiology* 130(4): 1587-1593.
- Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF. 2001. Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome*. 44(3): 321-330.
- Liu Z, Adams KL. 2007. Expression Partitioning between Genes Duplicated by Polyploidy under Abiotic Stress and during Organ Development. *Current Biology*. 17(19): 1669-1674.
- Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154(1): 459-473.
- Madlung A, Tyagi AP, Watson B, Jiang H, Kagochi T, Doerge RW, Martienssen R, Comai L. 2004. Genomic changes in synthetic *Arabidopsis* polyploids. *The Plant Journal*. 41(2): 221-230.
- Mizukami Y. 2001. A matter of size: developmental control of organ size in plants. *Current Opinion in Plant Biology*. 4(6): 533-539.
- Mochida K, Yamazaki Y, Ogihara Y. 2003. Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Molecular Genetics and Genomics*. 270(5): 371-377.
- Osborn TC, Butrulle DV, Sharpe AG, Pickering KJ, Parkin IA, Parker JS, Lydiate DJ. 2003. Detection and effects of a homeologous reciprocal transposition in *Brassica napus*. *Genetics*. 165(3): 1569-1577.

- Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee HS, Comai L, Madlung A, Doerge RW, Colot V, Martienssen RA. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics*. 19(3): 141-147.
- Otto SP. 2007. The Evolutionary Consequences of Polyploidy. *Cell*. 131(3): 452-462.
- Parisod C, Alix K, Lust J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhou B, Grandbastion MA. 2010. Impact of transposable elements on the organization and function of allopolyploid genomes. *The New Phytologist*. 186(1): 37-45.
- Peng H, Zhang J, Wu X. 2008. The ploidy effects in plant gene expression: progress, problems and prospects. *Science in China. Series C, Life Sciences / Chinese Academy of Sciences*. 51(4): 295-301.
- Pignatta D, Comai L. 2009. Parental squabbles and genome expression: lessons from the polyploids. *Journal of Biology*. 8(4): 43.
- Prakash NS, Combes MC, Somanna N, Lashermes P. 2002. AFLP analysis of introgression in coffee cultivars (*Coffea arabica* L.) derived from a natural interspecific hybrid. *Euphytica* (124): 265-271.
- Rakyan VK, Preis J, Morgan HD, Whitelaw E. 2001. The marks, mechanisms and memory of epigenetic states in mammals. *Biochemical Journal*. 356(1): 1-10.
- Ramsey J, Schemske DW. 1998. Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants. *Annual Review of Ecology and Systematics* 29: 467-501.
- Rapp RA, Wendel JF. 2005. Epigenetics and plant evolution. *The New Phytologist*. 168(1): 81-91.
- Rauscher JT, Doyle JJ, Brown AHD. 2002. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. *Molecular Ecology*. 11(12): 2691-2702.
- Riddle NC, Birchler JA. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. *Trends in Genetics*. 19(11): 597-600.
- Rozen S, Skaletsky HJ. 2000. Primer 3 on the WWW for general users and for biologist programmers. *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ. 365-386.
- Salmon A, Ainouche ML, Wendel JF. 2005. Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Molecular Ecology*. 14(4): 1163-1175.
- Salse J, Bolot S, Throude M, Jouffe V, Piegue B, Umar Masoosd Q, Calcagno T, Cooke R, Delseny M, Feuillet C. 2008. Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution. *The Plant Cell*. 20(1): 11-24.
- Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. 2006. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature*. 440(7082): 341-345.

- Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution*. 14(9): 348-352.
- Soltis DE, Albert Va, leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of botany*. 96(1): 336-348.
- Stamati K, Mackay I, Powell W. 2009. A quantitative genomic imbalance gene expression assay in a hexaploid species: wheat (*Triticum aestivum*). *Genome*. 52(1): 89-94.
- Udall JA, Quijada PA, Osborn TC. 2005. Detection of Chromosomal Rearrangements Derived From Homeologous Recombination in Four Mapping Populations of *Brassica napus* L. *Genetics*. 169(2): 967-979.
- Vidal RO, Costa Mondego JM, Pot D, Ambrosio AB, Carvalho Andrade A, Protasio Pereira LP, Colombo CA, Esteves Vieira LG, Carazolle Falsarella M, Pereira GAG. 2010. A High-Throughput Data Mining of Single Nucleotide Polymorphisms in *Coffea* Species Expressed Sequence Tags Suggests Differential Homeologous Gene Expression in the Allotetraploid *Coffea arabica*. *Plant Physiology*. 154(3): 1053-1066.
- Wang J, Tian L, Madlung A, Lee HS, Chen M, Lee JJ, Watson B, kagochi T, Comai L, Chen ZJ. 2004. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics*. 167(4): 1961-1973.
- Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation. *Nature* 430(6995): 85-88.
- Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the national Academy of sciences*. 106(33): 13875-13879.
- Yang X, Ye CY, Cheng ZM, Tschaplinski TJ, Wullschleger SD, Yin W, Tuskan GA. 2010. Genomic aspects of research involving polyploid plants. *Plant Cell, Tissue and Organ Culture*. 104(3): 387-397.
- Zhou X, Lin Z, Ma H. 2010. Phylogenetic detection of numerous gene duplications shared by animals, fungi and plants. *genome biology*. 11(4): R38.