UNIVERSITE MONTPELLIER II SCIENCES ET TECHNIQUES DU LANGUEDOC

THESE

Présentée pour l'obtention du titre de

DOCTEUR EN SCIENCES DE L'UNIVERSITE MONTPELLIER II

Formation Doctorale : Biologie, Diversité et Adaptation des Plantes Cultivées *Ecole Doctorale* : Biologie Intégrative

par

Sandra NOIR

Diversité des gènes de résistance au sein du génome des caféiers (Coffea L.)

Analyse génétique de la résistance au nématode à galles,

Meloidogyne exigua chez C. arabica

soutenue publiquement le 19 décembre 2002 devant le jury composé de,

A.-F. ADAM-BLONDON, Chargé de Recherche, INRA Evry
A. CHARRIER, Professeur, ENSA Montpellier
T. LANGIN, Directeur de Recherche, CNRS Orsay
P. LASHERMES, Directeur de Recherche, IRD Montpellier
M. LEBRUN, Professeur, Université Montpellier II
A. PALLOIX, Directeur de Recherche, INRA Avignon

Examinatrice Examinateur Rapporteur Directeur de Thèse Président Rapporteur

Sommaire

AVANT-PROPOS		1
CHAPITRE 1	INTRODUCTION BIBLIOGRAPHIQUE	
1. LE CAFEIER		5
1. 1 QUELQUES GE	NERALI TES	5
1. 1. 1 Class	sification botanique et mode de reproduction	5
1. 1. 2 Phylo	ogénie des caféiers du genre <i>Coffea</i>	5
1. 2 Coffea arabi	ca, ORI GI NE ET CARACTERI STI QUES	6
1. 3 I MPORTANCE	ECONOMI QUE	8
2. CONTEXTE DE	E L'AMELIORATION GENETIQUE DE C. arabica	8
2. 1 OBJECTI FS DE	E SELECTI ON CHEZ LE CAFEI ER <i>C. arabica</i>	8
$2.\ 2\ R\ \text{essources}$	GENETI QUES	9
2. 3 AMELIORATIO	ON GENETIQUE DU CAFEIER <i>C. arabica</i>	11
3. L'INTERACTI	ON SPECIFIQUE PLANTE - AGENT PATHOGENE : MECANISMES	>
DE RESISTANCE	ET EVOLUTION DES GENES DE RESISTANCE	11
3. 1 LA RESISTAN	E DES PLANTES	11
3. 2 LE CONCEPT D	E RESI STANCE GENE-A-GENE	13
3. 2. 1 Les r	nécanismes de défense : HR, LAR et SAR !	13
3. 2. 2 Cara	ctérisation des gènes de résistance	14
	Architecture des produits de gène <i>R</i> , la structure à la rencontre de la fonction ?	
	Classification des gènes R	
3. 2. 3 De la	reconnaissance aux mécanismes de défense	
	Modeles de reconnaissance specifique entre les produits des genes <i>R</i> et <i>avr</i> definis	10
	autour du concept géne-a-géne	
	Complexite des signaux de transduction aboutissant aux mecanismes de detense	
3. 3 EVOLUTION L	JES GENES DE RESISIANCE	
3. 3. TOrga	nisation moleculaire des genes R et mecanismes genetiques impliques	22
	Comployité dos locus P	22
	Logue simple à multiples spécificités de résistance	22 22
	Locus simple a maniples specificaes de resistance	20 23
	Co-iocuisation de genes de resistance à afferents agents painogenes Multiples copies de séquences analogues organisées en clusters	23
	Mécanismes génétiques impliqués dans l'évolution des gènes R	20
	Substitutions	24
	Différents types de recombinaison	
	Eléments transposables	
	Produits tronqués et pseudogènes	

3. 3. 2 Evolution des gènes <i>R</i> : des modèles d'évolution	27
Notion de pression de sélection réciproque ou co-évolution	27
Déterminisme de la spécificité des gènes de la classe NBS-LRR	27
Compromis entre diversification et conservation	28
4. UNE CONTRAINTE AGRONOMIQUE MAJEURE, LES NEMATODES	31
4. 1 L'AGENT PHYTOPATHOGENE	31
4. 1. 1 Les nématodes phytoparasites	31
4. 1. 2 Les nématodes du genre <i>Meloidogyne</i>	32
Généralités	32
Mode de reproduction	32
Cycle biologique	32
Symptomatologie	33
4. 2 LES GENES DE RESI STANCE SPECIFIQUE AUX NEMATODES	34
4. 2. 1 Les gènes de résistance aux nématodes clonés	34
4. 2. 2 L'exemple du gène <i>Mi</i> de la tomate	35
4. 3 LES NEMATODES DU GENRE Meloidogyne PARASI TES DU CAFEI ER C. arabica	36
4. 3. 1 Importance économique	
4. 3. 2 Meloidogyne exigua	
Symptômes associés à <i>M. exiqua</i> chez le caféier	
Sources de résistance à <i>M. exigua</i> chez le caféier	
5. LES OBJECTIFS DE THESE	38
5 1 E MATERIEL D'ETUDE	39

5.	1. I	LE MATERI EL D'ETUDE	9
5.	2 L	_ES OBJECTIFS4	0

CHAPITRE 2 DIVERSITE ET EVOLUTION DES GENES DE RESISTANCE DE LA CLASSE NBS-LRR, AU SEIN DU GENOME DU CAFEIER

Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (<i>Coffea</i> L .)	43
Abstract	44
Introduction	44
Materials and Methods	46
Plant material	46
Primers and PCR conditions	46
Cloning and sequencing of PCR products	47
Sequence analysis	47
Results	48
PCR amplification of RGAs	48
Analysis of coffee RGA	49
Comparison analysis	50

Discussion	51
Degenerate primer approach in coffee	51
Coffee RGA diversity versus R-genes diversity	52
Variation and evolution of NBS-encoding sequences in coffee trees	53
Functional inference from NBS domain evolution	54

CHAPITRE 3 IDENTIFICATION ET CARTOGRAPHIE GENETIQUE D'UN LOCUS DE RESISTANCE AU NEMATODE A GALLES, *MELOIDOGYNE EXIGUA*, CHEZ *COFFEA ARABICA*

I dentification of a major gene (*Mex-1*) from *Coffea canephora* conferring resistance to *Meloidogyne exigua* in *Coffea arabica*.....**57**

Abstract	58
Introduction	58
Materials and methods	59
Plant material	59
Nematode resistance evaluation	60
AFLP analysis	60
Analysis of data	61
Results	61
Inheritance of <i>M. exiqua</i> resistance	61
Identification of molecular markers associated with the resistance to <i>M. exiqua</i>	62
Linkage analysis and construction of a localized linkage map	62
Consistency of <i>Mex-1</i> in different Timor Hybrid-derived Arabica lines	
Discussion	

CHAPITRE 4 CONSTRUCTION D'UNE BANQUE BAC DE C. ARABICA

Construction and	characterisation	of	а	BAC	library	of	coffee	tree	
(Coffea arabica L.).					-				67
````									
Abstract									68
Introduction									68
Materials and metho	ods								69
Plant material									69
BAC library cor	struction								69
Isolation of	of high molecular weig	ht DN	VA						69
Hind///-dig	gestion of high molecu	ılar w	eigh	t DNA					70
Size sele	ction of partially restric	cted F	IMN	/-DNA					70
Ligation a	and construction of the	BAC	libra	ary					70
Estimation of in	sert sizes of BAC clon	ies							71
BAC filter and S	Southern hybridisation								71

Results	72
Construction of the Arabica line BAC library	72
Insert size estimation and organellar DNA contamination control	72
BAC library screening	73
Discussion	73

# CHAPITRE 5 DEVELOPPEMENT D'UNE CARTE PHYSIQUE DU LOCUS *MEX*-1 CONFERANT LA RESISTANCE A *M. EXIGUA*

Physical mapping of the *Mex*-1 region and identification of BAC clones containing putative resistance genes in coffee (*Coffea arabica*) ......**77** 

hstract 7
atroduction 7
laterials and methods
Material80
BAC library screening80
BAC clone characterisation8
BAC-end sequencing and mapping8
Search of disease-resistance gene analogues82
esults
BAC library screening
Development of BAC contig
Identification of BACs containing RGAs83
iscussion
Development of contigs
Nature of the introgression carrying Mex-184
BAC library screening
The locus Mex-1 located on a R gene-cluster

# CHAPITRE 6 DISCUSSION GENERALE

1. SYNTHESE DES PRINCIPAUX RESULTATS	88
1. 1 Etude de la diversite et de l'evolution des genes de resistance de la classe	
NBS-LRR, AU SEIN DU GENOME DU CAFEIER	88
1. 2 IDENTIFICATION ET CARTOGRAPHIE GENETIQUE D'UN LOCUS DE RESISTANCE A	
Meloidogyne exigua, chez Coffea arabica	90
1. 3 DEVELOPPEMENT D'UNE CARTE PHYSIQUE DU LOCUS MEX-1 CONFERANT LA RESISTANCE	
А <i>М. ЕХІ GUA</i>	91
1.3.1 Construction et caractérisation d'une banque BAC, à partir du génome	
C. arabica	91
1. 3. 2 Carte physique de la région portant le locus Mex-1 et analyses préliminaires	92

2. PERSPECTIVES	93
2. 1 VERS UNE CONNAISSANCE PLUS APPROFONDIE DE L'ORGANISATION ET DE L'EVOLUTION	
DES GENES DE RESI STANCE AU SEI N DU GENOME DES CAFEI ERS	93
2. 1. 1 Distribution et localisation des RGA au sein du génome des caféiers	93
2. 1. 2 Relation avec la structure polyploïde du génome	94
2. 2 VERS UNE CONNAISSANCE PLUS APPROFONDIE DU CLUSTER DE RESISTANCE	
COMPRENANT LE LOCUS <i>MEX-1</i>	95
2. 3 VERS UNE CONNAISSANCE PLUS APPROFONDIE DE LA STRUCTURE ET DE LA FONCTION	
DU GENE <i>MEX-1</i>	97
2. 3. 1 Isolement du gène <i>Mex-1</i> et validation fonctionnelle	97
Détermination de la distance physique du locus Mex-1	97
Stratégies de clonage envisagées	98
Validation fonctionnelle	99
2. 3. 2 Etude du gène <i>Mex-1</i>	100
2. 3. 3 Sélection assistée	100
D D	
REFERENCES BIBLIOGRAPHIQUES	103
ANNEXES	120
Abréviations	121
Figures et tableaux par chapitre (en version électronique)	122
Résumé	167

Avant-propos

La pression sanitaire, exercée par les parasites et les ravageurs sur les plantes cultivées, représente l'une des contraintes majeures de la production végétale à l'échelle mondiale. Tout en contrôlant et limitant le développement de ces infections, les méthodes de lutte développées tentent de satisfaire le souci croissant du respect de l'environnement, et les exigences des consommateurs pour des produits alimentaires dépourvus de résidus de traitement. Dans le cadre d'une lutte intégrée, l'utilisation, quand elles sont disponibles, de variétés résistantes aux maladies est une alternative efficace à l'emploi de produits phytosanitaires.

Le café est une boisson qui doit son succès à son goût unique et agréable mais également à l'effet stimulant qu'il procure. Parmi la centaine d'espèces de caféiers recensées à ce jour, deux sont cultivées à l'échelle mondiale, *Coffea arabica* (75% de la consommation mondiale) et *C. canephora* (25%). La culture des caféiers est un enjeu économique majeur pour les 15 à 20 millions d'exploitations situées dans les pays producteurs de toute la zone intertropicale humide. Alors que le marché mondial du café est déjà en crise, la pression parasitaire importante au niveau des zones d'exploitation constitue une contrainte sérieuse en caféiculture. Cette pression s'exerce tout particulièrement sur les variétés *C. arabica* dont les plus cultivées se sont avérées sensibles à la plupart des maladies et parasites du caféier, tels que la rouille orangée, l'anthracnose des fruits, les nématodes ou encore le scolyte des baies. Les nématodes phytopathogènes sont notamment responsables d'importants dommages dans les caféières de la plupart des pays d'Amérique Latine. Les dégâts provoqués vont de la baisse sensible de la production de café, jusqu'à la mort des plantes et l'abandon de la caféiculture.

Le développement de lignées résistantes, notamment à la rouille, a fait l'objet de programmes de sélection généalogique qui ont duré plus de quarante années. Or, en comparaison des variétés traditionnelles, la majorité des lignées sélectionnées présentent une proportion importante de défauts des grains et une baisse de la qualité à la tasse. L'analyse des résultats de cette sélection montre qu'il est difficile de parvenir à conserver le ou les caractères introgressés tout en réduisant les effets défavorables de l'introgression. Le transfert de nouvelles résistances dans les variétés cultivées rend nécessaire le contrôle des fragments chromosomiques hérités du parent sauvage pour réduire l'effet négatif de l'introgression sur la qualité du café produit.

2

A ces contraintes de sélection, s'ajoute une difficulté supplémentaire. La diversité des agents pathogènes du caféier nécessite le développement de variétés cumulant de multiples résistances. Dans cet objectif, le besoin de connaissances relatives à l'ensemble des caractères de résistance chez les caféiers devient toujours plus important.

Dans ce contexte, l'analyse globale et la connaissance des gènes de résistance au sein du genre *Coffea*, et, l'identification au sein d'espèces sauvages de caféiers, de nouvelles sources de résistance, en particulier, vis-à-vis des nématodes, constituent les objectifs majeurs des programmes d'amélioration de la principale espèce d'intérêt agronomique, *C. arabica.* Une connaissance de l'ensemble des gènes de résistance et l'identification de marqueurs moléculaires associés à la résistance de ravageurs des caféiers permettront d'optimiser les méthodes de gestion et d'utilisation des ressources génétiques des caféiers.

Chapitre 1

# Introduction Bibliographique

# **1.LE CAFEIER**

## 1. 1 QUELQUES GENERALITES

#### 1. 1. 1 Classification botanique et mode de reproduction

Les caféiers sont des arbres ou arbustes aux fleurs odorantes, appartenant à la famille des *Rubiaceae* et constituant la tribu des *Coffeeae*. Dans les classifications taxonomiques les plus récentes, on distingue sur la base de caractéristiques florales, deux genres, *Coffea* L. et *Psilanthus* Hook (Leroy, 1980 ; Bridson et Verdcourt, 1988 ; Bridson, 1994). Le genre *Coffea* est subdivisé en deux sous-genres, *Coffea* et *Baracoffea*, sur la base de la position des inflorescences et du mode de développement (Bridson et Verdcourt, 1988).

Plus de 80 taxons, dont les deux espèces principalement cultivées, *C. arabica* et *C. canephora*, ont été recensés au sein du genre *Coffea* sous-genre *Coffea*. Ces caféiers sont endémiques des forêts intertropicales humides d'Afrique (au moins 30 taxons), de Madagascar (au moins 50 taxons), de l'archipel des Comores et des îles Mascareignes.

Les espèces appartenant au genre *Coffea* sont toutes diploïdes (2n = 2x = 22) à l'exception de l'espèce *C. arabica* qui est tétraploïde (2n = 4x = 44). L'espèce *C. arabica* est préférentiellement autogame alors que les caféiers diploïdes sont allogames stricts à l'exception de *C. heterocalix* et *C. sp* Moloundou (Anthony, 1992). Le système d'auto-incompatibilité, étudié chez l'espèce *C. canephora*, est de type gamétophytique et apparaît contrôlé par un seul locus présentant plusieurs allèles (Berthaud, 1986 ; Lashermes *et al.*, 1996). Les espèces diploïdes ont un génome de base commun peu différencié comprenant 11 chromosomes (Bouharmont, 1963 ; Charrier, 1978). Le génome diploïde a été cartographié chez l'espèce *C. canephora*. La carte génétique obtenue présente 160 locus, et distingue 11 groupes de liaison correspondant potentiellement aux 11 chromosomes du génome de *C. canephora* (Lashermes *et al.*, 2001a).

# 1. 1. 2 Phylogénie des caféiers du genre Coffea

Plusieurs analyses moléculaires ont été conduites afin de préciser l'organisation des espèces du genre *Coffea*.

Dans un premier temps, une transmission strictement maternelle de l'ADN chloroplastique (ADNcp) a été démontrée chez l'espèce *C. canephora* et les hybrides interspécifiques entre *C. arabica* et *C. canephora* (Lashermes *et al.*, 1996). Par la suite, les variations de l'ADN chloroplastique de 27 taxons représentatifs du genre *Coffea* sous-genre *Coffea* ont été étudiées par plusieurs approches incluant une analyse RFLP (Restriction Fragment Length Polymorphism) à l'aide de sondes hétérologues (*Lactuca sativa*) et le séquençage de la région intergénique *trnL-trn*F (Cros *et al.*, 1998).

Le génome nucléaire a aussi été étudié afin de pallier le faible polymorphisme présenté par l'ADNcp des caféiers et de comparer les phylogénies obtenues par l'analyse des génomes chloroplastique et nucléaire. Plus particulièrement, le fragment ITS2 (Internal Transcribed Spacers) de la grande sous-unité nucléaire codant pour l'ARN ribosomique a été séquencé pour 40 génotypes représentant 28 taxons de *Coffea* et 4 espèces de *Psilanthus* (Lashermes *et al.*, 1997).

Les analyses phylogénétiques de ces données moléculaires suggèrent une explosion radiative récente (très largement postérieure à la dislocation du Gondwana) et l'existence d'au moins 5 groupes phylogénétiques relativement peu différenciés. Ces groupes d'espèces de caféiers (*Coffea* sous-genre *Coffea*) correspondent à de vastes ensembles biogéographiques : zone guinéo-congolaise (deux groupes), Afrique de l'Est, et Madagascar et les îles Mascareignes (Fig. 1. 1) Des résultats remarquablement convergents sont obtenus à partir des deux génomes analysés (i. e. nucléaire et chloroplastique). Toutefois, quelques divergences ont pu être observées suggérant la possibilité d'hybridations interspécifiques et d'introgressions en relation avec le processus de spéciation. Par ailleurs, la distribution actuelle des différentes espèces de caféiers semble associée d'une part à la formation de la vallée du Grand Rift africain, et d'autre part, aux variations climatiques et l'histoire des refuges forestiers tropicaux du quaternaire (Berthaud, 1986 ; Cros, 1994).

# 1. 2 Coffea arabica, ORIGINE ET CARACTERISTIQUES

*C. arabica* est un arbuste originaire des forêts tropicales d'altitude d'Ethiopie, du Sud du Soudan et du Nord du Kenya (Anthony *et al.*, 1987). Plusieurs approches complémentaires ont été conduites afin de caractériser et de déterminer l'origine du génome de *C. arabica*. Aucune différence entre l'ADNcp de *C. arabica* et celui de certaines accessions diploïdes (*C. eugenioides*, *C. sp.* Moloundou) n'a été mise en évidence (Lashermes *et al.*, 1996). De plus, la séquence ITS2 analysée de *C. arabica* apparaît proche de celles des espèces du groupe des

"canephoroïdes" (*C. brevipes*, *C. canephora*, *C. congensis*...) (Lashermes *et al.*, 1997). Le polymorphisme nucléaire a été analysé par RFLP en utilisant des sondes spécifiques de locus (Lashermes *et al.*, 1999a). Les allèles présents chez *C. arabica* ont été recherchés au sein des espèces diploïdes. Par ailleurs, l'hybridation génomique *in situ* (GISH) qui consiste en l'hybridation des chromosomes métaphasiques d'une espèce donnée par l'ADN total d'une autre espèce, a été utilisée (Lashermes *et al.*, 1999a). Le génome de *C. arabica* apparaît constitué par l'association de deux génomes diploïdes. Cette espèce résulterait d'une hybridation récente entre deux espèces diploïdes de caféier proches des espèces actuelles *C. eugenioides* (parent femelle, génome E^a) et *C. canephora* (parent mâle, génome C^a), en zones de sympatrie.

Le caractère polyploïde du caféier *C. arabica* pourrait être à l'origine de la capacité de cette espèce à coloniser des milieux marginaux d'altitude auxquels les espèces diploïdes ne sont pas adaptées. En effet, la croissance de *C. arabica* est optimale à des conditions de température comprises entre 17 et 25°C et, à raison de précipitations annuelles de 1200 à 2000 mm (Campos *et al.*, 1990).

L'espèce *C. arabica* présente un comportement méiotique comparable à celui des espèces diploïdes avec la formation prépondérante de bivalents (Grassias et Kammacher, 1975). Par l'analyse, à plusieurs locus RFLP, des ségrégations au sein d'une population  $F_2$ , un mode d'hérédité disomique a été établi chez *C. arabica* (Lashermes *et al.*, 2000a). A la méiose, les chromosomes homologues (i. e. appartenant au même sous-génome) s'apparieraient de façon systématique pour former des bivalents.

Les principales variétés de *C. arabica* (Caturra, Catuai, Mundo Novo) sont hautement productives et donnent un café de bonne qualité. Elles furent sélectionnées à partir de l'étroite base génétique dispersée au début du XVIII^e siècle (Fig. 1. 2) par les puissances coloniales (Pays-Bas, France, Angleterre). Cependant, en conséquence du nombre extrêmement réduit d'individus fondateurs à l'origine des caféiers en plantation à travers le monde, ainsi que du mode de spéciation de *C. arabica*, les variétés cultivées présentent une très faible diversité génétique (Anthony *et al.*, 2002). Cette caractéristique rend la caféiculture d'Arabica particulièrement vulnérable aux aléas. En effet, les caféiers cultivés se sont révélés sensibles à la plupart des maladies et ravageurs qui attaquent le caféier, notamment la rouille orangée due à *Hemileia vastatrix*, l'anthracnose des baies causée par *Colletotrichum kahawae*, les nématodes des genres *Meloidogyne* et *Pratylenchus*, la mineuse des feuilles *Perileucoptera coffeella* ou encore le scolyte des baies *Hypothenemus hampei* (Bertrand *et al.*, 1999).

7

# 1. 3 IMPORTANCE ECONOMIQUE

Bien que l'expansion de la caféiculture soit un phénomène récent lié à l'histoire coloniale des nations européennes, la culture du caféier constitue de nos jours une source de revenus vitale pour l'économie de nombreux pays tropicaux.

Deux espèces de caféiers sont principalement cultivées. Le café, de qualité très appréciée, produit par l'espèce C. arabica L. représente 75% de la consommation mondiale et est cultivé dans plus de 60 pays principalement d'Amérique Latine et d'Afrique de l'Est. La culture de C. canephora (environ 25% de la consommation mondiale) qui donne le café Robusta, se concentre plus particulièrement en Asie, Afrique de l'Ouest, et Brésil (Campos et al., 1990). Ces zones de culture du caféier recouvrent plus de 12 millions d'hectares de terres. Au niveau mondial, le café occupe l'un des premiers marchés de matière première en valeur, après celui du pétrole, et génère un chiffre d'affaires annuel estimé entre 7 et 12 milliards de dollars selon les cours mondiaux (Anonyme, 1998). Au cours de l'année 2001-2002, la production mondiale s'est élevée à 113 millions de sacs de 60 kg (indicateur ICO, 2002a). Le café fait ainsi l'objet d'enjeux économiques importants. A titre d'exemple, les exportations de café représentent entre 20 et 25% de la valeur totale des exportations, selon les années, pour les pays producteurs d'Amérique Centrale. Outre le rôle évident du caféier dans le développement économique de ces pays, sa culture constitue un important facteur de stabilité sociale en raison des emplois qu'elle génère. Le nombre de personnes directement concernées par la production est estimé à 125 millions (indicateur ICO 2002b ; Pare, 2002).

# 2. CONTEXTE DE L'AMELIORATION GENETIQUE DE Coffea arabica

# 2. 1 OBJECTIFS DE SELECTION CHEZ LE CAFEIER C. arabica

Les objectifs de sélection propres à l'arabica-culture sont directement liés aux problèmes biologiques associés aux systèmes de culture, aux contraintes économiques ainsi qu'aux exigences commerciales.

La culture de C. arabica présente différentes caractéristiques :

### - Caractéristiques relatives aux zones de culture

D'une part, l'importante diversité des zones de culture (reliefs et microclimats) aboutit à une grande diversité des terroirs. D'autre part, la faible disponibilité en terres agricoles nécessite une intensification de la culture et donc la recherche de gains de productivité. Dans cette optique, les producteurs souhaitent globalement multiplier les variétés de port-nain (i. e. cv Caturra) afin de cultiver à haute densité.

### - Caractéristiques relatives au système de production

L'instabilité des cours mondiaux du café a pour conséquence des pratiques agronomiques soumises à ces fluctuations. La durabilité des systèmes de production est alors régulièrement écourtée par la recherche de baisse de coûts de production (en intrants et en main d'œuvre).

### - Caractéristiques relatives aux exigences des consommateurs

En effet, la demande en cafés fins ou encore cafés "terroirs" se fait croissante. Par ailleurs, les consommateurs soutenus par les producteurs, exercent une forte pression commerciale par l'aspiration à une caféiculture "plus biologique" : un café exempt de résidus de produits phytosanitaires, une teneur minimum en myco-toxines et une réduction des intrants au niveau des plantations.

Ces contraintes multiples soulignent la nécessité de développer de nouvelles variétés, présentant la ou les qualités suivantes :

- Une résistance aux maladies, existante ou potentielle
- Une productivité au moins égale à celle des meilleures lignées actuelles
- Une qualité organoleptique comparable ou supérieure à celle de la variété Caturra
- Une adaptation aisée à la réduction des intrants, c'est-à-dire une bonne stabilité de production
- Une fertilité comparable à celle de la variété Caturra.

# 2. 2 RESSOURCES GENETIQUES

Les caféiers subspontanés du centre primaire de diversité de *C. arabica* constituent une importante ressource génétique. Toutefois, la diversité génétique présente au sein de ce matériel végétal reste limitée, et l'exploitation des ressources génétiques que représentent les

espèces diploïdes de caféiers apparaît incontournable. Ces deux principales ressources représentent un réservoir de diversité génétique essentiel pour l'amélioration de *C. arabica* dans une perspective de développement durable (Tableau 2. 1) (Carvalho, 1988 ; Anthony *et al.*, 1999).

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 (Van der Vossen, 1985 ; Carvalho, 1988 ; Charrier et Eskes, 1997 ; Couturon *et al.*, 1998). 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'introgressions (Lashermes *et al.*, 2000b ; Herrera *et al.*, 2002 ; Prakash *et al.*, 2002).

La diversité génétique des espèces diploïdes représente ainsi un intérêt agronomique majeur (Berthaud et Charrier, 1988). Notamment, l'espèce *C. canephora* dont la mise en culture reste récente (i. e. il y a un siècle environ), est particulièrement intéressante. Les variétés de cette espèce, en dépit d'une qualité à la tasse moindre comparée à celle du café Arabica, se caractérisent par une bonne adaptation à basse altitude, une importante vigueur, une productivité élevée et la résistance vis à vis de divers agents pathogènes du caféier.

Un autre matériel génétique important est constitué par les hybrides spontanés qui sont apparus lorsque des populations de *C. arabica* et de caféiers d'espèces diploïdes ont été cultivées ensembles. L'Hybride de Timor (HdT), notamment, a été identifié au sein d'une plantation de *C. arabica* (plantée en 1927) sur l'île de Timor en Indonésie (Bettencourt, 1973). Les analyses génétique et moléculaire, de cet hybride et de ses descendants, ont démontré que l'HdT est issu d'un croisement interspécifique spontané entre *C. arabica* et *C. canephora* (Bettencourt, 1973 ; Goncalves et Rodrigues, 1976 ; Lashermes *et al.* 2000b). Les descendants de l'HdT, largement diffusés à travers le monde, ont été très exploités dans les programmes d'amélioration des caféiers ces 50 dernières années. Ils constituent la principale source de résistance aux maladies et ravageurs du caféier, en particulier, vis-à-vis de la rouille orangée, de *Colletotrichum kahawae* et des nématodes à galles du genre *Meloidogyne* (Bertrand *et al.*, 2001 ; Van der Vossen, 2001).

# 2. 3 AMELIORATION GENETIQUE DU CAFEIER C. arabica

L'amélioration génétique de l'espèce *C. arabica* est basée sur l'exploitation des ressources génétiques des caféiers sauvages. Elle consiste principalement à introduire plus de variabilité génétique au sein de l'espèce. Le caféier *C. arabica* est sélectionné essentiellement par la méthode généalogique et par rétro-croisements.

Bien que la sélection généalogique nécessite plus de trente années pour l'obtention de lignées chez le caféier (i. e. un cycle de sélection dure au moins 5 ans), elle s'est avérée efficace pour améliorer la production, la vigueur et la résistance de *C. arabica*, tout en conservant les caractéristiques du café Arabica (Charrier et Eskes, 1997). Ainsi, les hybrides interspécifiques, artificiel ou naturel, ont servis de relais pour le transfert de caractères agronomiques par rétro-croisements avec *C. arabica*. De nombreuses variétés lignées (e. g. Catimor, Sarchimor, ICATU, variété Columbia) ont ainsi été créées.

Au cours des dernières années, une stratégie de valorisation plus rapide de ces lignées sélectionnées a été de créer des variétés hybrides (Van der Vossen, 2001). Les premières variétés hybrides ont été obtenues par croisement en conditions de pollinisation contrôlée (manuelle) entre 2 lignées ou 2 pools de lignées de caféiers (e. g. Ruiru 11). Plus récemment, des variétés hybrides F₁, diffusées par multiplication végétative *in vitro* (embryogenèse somatique) d'arbres hybrides, ont été proposées (Bertrand *et al.*, 1999 ; Etienne *et al.*, 2002).

# 3. L'INTERACTION SPECIFIQUE PLANTE-AGENT PATHOGENE : MECANISMES DE RESISTANCE ET EVOLUTION DES GENES DE RESISTANCE

# **3. 1** LA RESISTANCE DES PLANTES

A chaque instant, les plantes doivent faire face à l'agression de micro-organismes tels que les champignons, les bactéries, les virus ou encore les nématodes. Ne pouvant fuir devant l'agresseur, elles ont élaboré au cours de leur évolution des mécanismes de résistance efficaces contre les divers agents pathogènes. La résistance correspond à la capacité de la plante à retarder ou supprimer l'activité et le développement du parasite. Selon la mise en place et la nature des mécanismes impliqués en réponse à une attaque phytopathogène, différents

Introduction Bibliographique

types de résistance sont distingués : la résistance non-hôte à déterminisme passif et, la résistance hôte, non-spécifique ou spécifique, à déterminisme actif.

Quel que soit leur environnement, les espèces végétales sont protégées contre la plupart des agents pathogènes, on parle de résistance non-hôte pour laquelle deux niveaux ont été distingués (Heath, 2000a). Le plus souvent, les micro-organismes sont incapables de franchir les barrières physiques telles que les cires cuticulaires de la surface des feuilles ou, la taille étroite des stomates. Il s'agit d'une résistance passive qui a lieu indépendamment de la présence du micro-organisme. Le second niveau de la résistance non-hôte correspond à la situation où l'agent pathogène amorce le processus de pénétration de la plante. Toutefois, le parasite ne trouvant pas nécessairement les conditions physiologiques requises pour son développement, l'infection est stoppée.

Dans le cas de la résistance hôte non-spécifique, l'agresseur, par la sécrétion de toxines et/ou d'enzymes hydrolytiques (cutinases, protéases, pectinases...), a la capacité de contourner les barrières mécaniques de la plante, d'y pénétrer et quelquefois, d'entreprendre son développement. Cette résistance est induite par la reconnaissance d'éliciteurs généraux (e. g. polysaccharides, glycoprotéines, protéines fongique ou bactérienne). Les mécanismes de défense mis en place modifient les tissus infectés et stoppent la progression du parasite. Selon le degré d'efficacité des mécanismes activés, le site d'infection peut devenir impropre à la dissémination et à la survie de l'agent pathogène.

Dans le cas d'une résistance hôte-spécifique, les mécanismes de défense exprimés par la plante reposent sur une spécificité parasitaire beaucoup plus étroite. Cette résistance active confère à la plante la capacité à confiner -et parfois à tuer- l'agresseur en son site de pénétration ou d'attaque. Une telle résistance chez la plante est associée à une espèce donnée de parasite, et le plus souvent même, l'efficacité de cette résistance est restreinte à une race donnée de l'espèce pathogène, on parle de résistance hôte à spécificité race/cultivar.

D'un point de vue physiologique, certains mécanismes induits lors de la résistance hôtespécifique ne se distinguent pas fondamentalement de ceux mis en place lors d'une résistance hôte non-spécifique. Globalement, la mise en place de la résistance hôte-spécifique s'effectue en 3 phases (Fig. 3. 1). La première étape, la reconnaissance de l'agent pathogène, est essentielle. L'issue de l'interaction plante/agent pathogène est en grande partie déterminée par la rapidité de la plante à mettre en place les mécanismes de défense active. Sans la reconnaissance de l'agent pathogène, les étapes ultérieures ne se déclenchent pas. La deuxième phase est celle de l'activation d'une cascade de signaux au niveau des cellules attaquées et de la transmission des signaux d'alerte aux cellules environnantes, puis à la plante entière. Enfin, au cours de la troisième étape, les réactions de défense s'expriment.

# 3. 2 LE CONCEPT DE RESISTANCE GENE-A-GENE

La résistance fondée sur une relation race/cultivar a atteint le niveau le plus spécifique et le plus raffiné des interactions plantes/parasites. Cette résistance, largement étudiée aux niveaux moléculaire et génétique, est associée au concept de résistance dit gène-à-gène, particulièrement développé ces dernières décennies (Flor, 1971). Selon ce concept, lorsque la plante détecte la présence de l'agent pathogène et parvient à le neutraliser, la plante est dite résistante et l'agent pathogène est considéré comme avirulent, on parle d'interaction incompatible. Alternativement, si l'agent pathogène parvient à coloniser la plante hôte et que la maladie se développe, la plante est dite sensible et l'agent pathogène virulent, il s'agit d'une interaction compatible (Tableau 3. 1).

Lors d'une résistance gène-à-gène, la spécificité race/cultivar observée apparaît liée à l'événement précoce de reconnaissance, directe ou indirecte, entre le produit d'un gène de la plante hôte, le gène de résistance (gène *R*) et le produit d'un gène de l'agent pathogène, le gène d'avirulence (gène *avr*). En aval de cette étape de reconnaissance, se distinguent trois niveaux d'expression de la résistance : la réaction d'hypersensibilité (HR), la résistance locale acquise (LAR) et, la résistance systémique acquise (SAR) (Blumwald *et al.*, 1998).

# 3. 2. 1 Les mécanismes de défense : HR, LAR et SAR !

Dans la majorité des cas, lors d'une interaction incompatible, toute une série d'événements se met en place à l'issue de la perception de l'agent pathogène par la plante (Blumwald *et al.*, 1998). L'ensemble du métabolisme des cellules infectées est sollicité et fait intervenir des protéines G, des flux ioniques, notamment de Ca²⁺, H⁺, K⁺ et Cl⁻ (Grant et Mansfield, 1999), des formes activées d'oxygènes (ou ROS pour, *Reactive Oxygen Species*) et toute une cascade de phosphorylation/déphosphorylation de protéines telles que les MAPK (*Mitogen-Activated-Protein-Kinase*) (Zhang et Klessig, 2001).

Le plus souvent, ce déferlement signalétique se traduit phénotypiquement par la mort rapide des cellules infectées par l'agent pathogène et la formation de lésions nécrotiques localisées autour du site de pénétration du parasite. Cette réaction à déterminisme génétique est dite réaction d'hypersensibilité (ou HR). A ce jour, la HR est comparée à une forme de mort cellulaire programmée (Heath, 2000b). Les modifications métaboliques importantes permettant de confiner l'agresseur en son site de pénétration (Goodman et Novacky, 1994) sont le résultat de l'expression de nombreux gènes de défense au niveau de la zone infectée. La HR génère également des signaux qui activent la LAR, au contact des lésions, et la SAR, à distance du site d'infection (Fig. 3. 1).

Lors de l'établissement de la LAR, les réponses de défenses sont intenses et tout particulièrement localisées au niveau de l'anneau de cellules entourant la zone HR (Dorey *et al.*, 1997). Ces réponses incluent la synthèse d'un large spectre de composés anti-microbiens tels que les protéines PR (*Pathogenesis Related Protein*), certains métabolites secondaires aux propriétés antibiotiques, par exemple les phytoalexines (Fritig *et al.*, 1998), ainsi que l'accumulation de molécules intervenant dans les voies de signalisation, notamment les ROS et diverses hormones dont l'acide salicylique (Dorey *et al.*, 1997).

Ces deux niveaux de résistance locale (HR et LAR) s'accompagnent d'une résistance établie à l'échelle de la plante, la SAR, dont l'expression, plus tardive et moins intense, peut persister plusieurs semaines après l'infection. La SAR maintient la plante en un état de veille qui lui permet de résister non seulement à l'agresseur d'origine, mais aussi à une large gamme d'autres parasites pouvant intervenir ultérieurement (Maleck et Dietrich, 1999). La mise en place de cette résistance systémique, au niveau de tissus non infectés, dépend d'un réseau élaboré de communication intercellulaire. Des signaux libérés par les cellules exprimant la HR diffusent et atteignent d'autres cellules qui à leur tour déclenchent une réponse spécifique. L'acide salicylique, l'éthylène, l'acide jasmonique ou encore la systémine ont été reconnus depuis plusieurs années comme des messagers chimiques. Ces messagers alertent les cellules non infectées et dirigent leur métabolisme vers la mise en place de réponses de défenses, notamment, le renforcement de la paroi cellulaire, la stimulation des voies du métabolisme secondaire (i. e. enzymes du métabolisme des phénylpropanoïdes, de la voie de biosynthèse de l'éthylène et de l'acide jasmonique) et l'accumulation de protéines PR.

### 3. 2. 2 Caractérisation des gènes de résistance

Au cours des dernières années, le clonage de plusieurs gènes de résistance et la caractérisation des produits déduits de leur séquence respective ont permis des avancées considérables dans la connaissance des bases moléculaires de la résistance hôte-spécifique des plantes.

14

Ces gènes *R* ont été isolés d'espèces végétales variées (e. g. la tomate, le lin, le riz, le tabac, *Arabidospsis*, la betterave sucrière) (Hammond-Kosack et Jones, 1997 ; Dangl et Jones, 2001). Malgré l'importante diversité des parasites auxquels ils confèrent la résistance (champignons, bactéries, virus ou encore nématodes), leur comparaison révèle une forte homologie de séquences ainsi que la conservation de motifs structuraux.

### Architecture des produits de gène R, la structure à la rencontre de la fonction ?

La majorité des protéines déduites des gènes *R* isolés de plantes, partage différents éléments structuraux (Fig. 3. 2) caractérisés au sein de domaines protéiques intervenant dans les mécanismes de défense chez la levure, la drosophile ou encore les vertébrés (Michelmore et Meyers, 1998 ; Cohn *et al.*, 2001 ; Dangl et Jones, 2001 ; Staskawicz *et al.*, 2001 ; Nürnberger et Brunner, 2002). Par analogie, la plupart des produits R combinerait un domaine récepteur et un domaine effecteur (Hammond-Kosack et Jones, 1997) assurant respectivement deux fonctions majeures : la reconnaissance de molécules élicitrices par des mécanismes d'interactions protéine-protéine et l'activation directe ou indirecte de signaux de transduction (Blumwald *et al.*, 1998). Cinq principaux domaines structuraux conservés ont été distingués. La figure 3. 3 illustre ces différents domaines et leur localisation (Dangl et Jones, 2001).

La majorité des produits déduits des séquences de gènes R clonés à ce jour présente en position C-terminale un domaine LRR (Leucine-Rich Repeats) intra- ou extra- cytoplasmique. Les domaines LRR correspondent à la répétition d'un motif de taille variable comprenant des leucines. Ils interviendraient dans les mécanismes d'interactions protéine-protéine et protéinepolysaccharide (Jones et Jones, 1997). Le domaine NBS (Nucleotide Binding Site) ou NB, associé au domaine LRR (NBS-LRR), est largement distribué au sein des gènes R clonés. Ce domaine, composé notamment de différents motifs conservés de type kinase, correspond à un site de fixation et d'hydrolyse des nucléotides triphosphates ATP et GTP. Il présente des analogies avec des protéines animales potentiellement impliquées dans les phénomènes de mort cellulaire apoptotique, notamment APAF-1 identifiée chez l'homme et CED-4 identifiée chez Caenorhabditis elegans. L'ensemble de ces domaines se regroupe sous la terminologie NB-ARC (Van der Biezen et Jones, 1998). Deux autres domaines, dits TIR et CC (ou encore LZ), moins fréquents au sein des gènes R identifiés, peuvent être associés, en position Nterminale, aux produits NBS-LRR. Le domaine TIR (Toll Interleukin Receptor) présente d'importantes homologies de séguences avec les domaines intracellulaires de récepteurs protéiques isolés de la drosophile (récepteur Toll) et de l'homme (récepteur Interleukin-1)

(Hammond-Kosack et Jones, 1997). Sur la base de ces analogies, un rôle dans la cascade de signalisation cellulaire est attribué au domaine TIR. Le domaine CC (*Coiled-Coil*) ou LZ (*Leucine-Zipper*) qui peut varier en taille et en position est connu pour assurer un rôle dans l'homo- ou l'hétérodimérisation des protéines.

Enfin, le domaine Sérine/Thréonine kinase se présente seul (e. g. produit du gène *Pto*) ou, associé à un domaine LRR (e. g. produit du gène *Xa21*). Il serait impliqué dans des réactions de phosphorylation lors de cascades de signalisation.

#### Classification des gènes R

Les différentes classes de gènes R sont définies sur la base de l'association de domaines conservés au sein des produits de gènes R (e. g. LRR, NBS, CC, TIR, Ser/Thr kinase). Plus de 10 classes de gènes peuvent être distinguées à ce jour (Tableau 3. 2). Quelques classes regroupent un grand nombre de représentants, toutefois, la majorité des classes de gènes R, ne présente qu'un unique membre.

Une première classe de gènes *R* coderait des protéines cytoplasmiques de type récepteur. Ces gènes présentent à la fois, un domaine LRR à l'extrémité 3' et un domaine NBS à l'extrémité 5'. Ces gènes NBS-LRR constituent la principale classe de gènes *R* clonés jusqu'à présent. En fonction de la présence en position N-terminale, au sein des produits NBS-LRR, du domaine TIR ou du domaine CC, 2 sous-classes de gènes *R* sont distinguées, TIR-NBS-LRR et Non-TIR-NBS-LRR (Meyers *et al.*, 1999).

Une deuxième classe regroupe les gènes *Cf-2* (Dixon *et al.*, 1996) et *Cf-9* (Jones *et al.*, 1994) de la tomate et le gène  $HS1^{pro-1}$  de la betterave sucrière (Cai *et al.*, 1997). Les produits codés par ces gènes présentent un domaine LRR extracellulaire, un domaine transmembranaire et une courte région cytoplasmique.

Les autres classes de gènes *R* ne sont constituées, pour l'instant, que d'un seul gène-membre mis en évidence. Certaines classes présentent une structure nouvellement identifiée dans le contexte de résistances spécifiques chez les plantes.

L'une de ces classes est représentée par le gène *Pto* isolé de la tomate (Martin *et al.*, 1993). Ce gène code une protéine kinase de type Sérine/Thréonine. Bien qu'aucune étude n'ait permis de montrer une association directe entre les deux protéines, les produits des gènes *Pto* et *Prf* (de classe NBS-LRR) sont tous les deux exigés pour qu'il y ait expression de la

16

résistance après infection par *Pseudomonas syringae* exprimant le gène *avrPto* ou *avrPtoB* (Salmeron *et al.*, 1994 ; Kim *et al.*, 2002).

Le gène *Xa21* isolé du riz constitue une quatrième classe (Song *et al.*, 1995). Un domaine LRR extracellulaire, une région trans-membranaire et un domaine kinase Sérine/Thréonine cytoplasmique ont été identifiés au sein de la protéine Xa21. La structure de cette protéine suggère, d'un point de vue évolutif, une relation entre les protéines de type LRR (*Cf*) et celles de type kinase (*Pto*).

Deux gènes, récemment isolés, présentent un domaine de type kinase. Le gène *Pbs1*, identifié chez *Arabidopsis thaliana* confère la résistance à *Pseudomonas syringae* (Swiderski *et al.*, 2001). Le produit de ce gène présente un domaine putatif Sérine/Thréonine kinase, différent de celui identifié au sein du produit *Pto*, ainsi que de courtes extensions N- et C-terminales. Par ailleurs, la fonctionnalité de ce gène *Pbs1* est indissociable de celle du gène *RPS5* (Warren *et al.*, 1999).

Chez l'orge, le produit de type kinase codé par le gène de résistance à l'agent de la rouille *Puccinia graminis*, *Rpg1*, se compose de deux domaines kinases en tandem, sans domaine récepteur ou domaine d'ancrage à la membrane identifiables. Par homologie de séquences, le gène *Rpg1* est plus proche du domaine kinase du gène *Xa21*, et d'un point de vue structural, le produit Rpg1 s'apparente mieux au produit Pto (Brueggemen *et al.*, 2002).

Les gènes de résistance à *Verticillium* chez la tomate illustrent également un nouveau type de récepteur (Kawchuk *et al.*, 2001). En effet, les produit de gènes *Ve* associent un domaine LRR, une séquence-signal potentielle d'endocytose, ainsi qu'un domaine LZ ou une séquence PEST (Pro-Glu-Ser-Thr) connue pour intervenir dans l'ubiquitinisation, la compartimentation et la dégradation des protéines.

Quant au gène de résistance *Rpw8*, conférant la résistance à plusieurs isolats du genre *Erysiphe* agent de l'oïdium, chez *A. thaliana*, il est constitué d'un domaine CC, sans homologie significative avec toute autre protéine caractérisée, associé en position N-terminale à un domaine trans-membranaire ou un peptide signal potentiel (Xiao *et al.*, 2001).

Il convient de mentionner que certains gènes de résistance, admis comme tels, ont la particularité d'être actifs à l'état récessif.

L'allèle, du gène *mlo*, membre de la famille de gènes *Mlo* codant des protéines membranaires (Devoto *et al.*, 1999), conférant la résistance à l'ensemble des isolats connus d'oïdium chez l'orge, est fonctionnel à l'état récessif.

17

De même, l'allèle récessif *Rrs1-R*, composé en position N-terminale des motifs identifiés TIR, NBS, LRR et en position C-terminale d'un signal potentiel de localisation nucléaire et d'un domaine WRKY caractéristique de facteur de transcription chez les plantes, confère chez *A. thaliana* la résistance à la bactérie *Ralstonia solanacearum* (Deslandes *et al.*, 2002 ; Lahaye, 2002).

Enfin, le gène *Hm1* chez le maïs fonctionne bien à l'état dominant mais pour coder une toxineréductase capable d'inactiver la toxine du champignon pathogène *Cochliobolus carbonum* race 1. *Hm1* se différencie des gènes *R* précédemment cités par le fait qu'aucun composé Avr n'est impliqué dans la dégradation de la toxine par l'activité réductase (Johal *et al.*, 1992).

# 3. 2. 3 De la reconnaissance aux mécanismes de défense

Initialement, dans le cadre de la résistance gène-à-gène, les gènes R ont été définis génétiquement comme une unité polymorphe entre les génotypes résistant et sensible d'une plante. A l'origine du concept, tout composant non-fonctionnel du dispositif de reconnaissance du parasite par la plante-hôte, au sein de la variété sensible, pouvait être considéré comme un gène R au sein de la variété résistante correspondante (Flor, 1971). A ce jour, différents exemples expérimentaux récents confirment que la présence et l'activité d'un gène R donné est indispensable à l'expression de la résistance qu'elle détermine. De plus, différents composés regroupés sous le terme de résistosome (Lahaye et Bonas, 2001), intervenant dès la phase de reconnaissance mais aussi en aval de cette étape, apparaissent indispensables à l'expression de la résistance.

# Modèles de reconnaissance spécifique entre les produits des gènes *R* et *avr* définis autour du concept gène-à-gène

La localisation de la plupart des agents phytopathogènes pourrait laisser supposer que les protéines R correspondent à des produits extracellulaires de type récepteurs. Si certaines protéines R (e. g. protéines Xa21 et Cf) apparaissent effectivement extracellulaires, la majorité des protéines codées par les gènes *R* étudiés semble être intracellulaire. Des systèmes de sécrétion, dits de type III, au sein de bactéries phytopathogènes ont été découverts (Lahaye et Bonas, 2001). Ils permettraient l'injection de protéines Avr au sein même de la cellule végétale, offrant la possibilité d'une interaction directe et/ou intracellulaire de ces dernières avec les produits R. Différents modèles d'interaction, non-exclusifs entre eux, ont été proposés. Ils

reposent sur la fonction de reconnaissance du produit R et suggèrent : soit une interaction directe entre les produits R et Avr, soit la présence d'un composé tiers nécessaire à la formation d'un complexe reconnu par R, soit encore, une étape de maturation du produit Avr pour être reconnu.

Le premier modèle, dit "récepteur/éliciteur", suggère une interaction directe entre les produits R et Avr (Keen, 1982) (Fig. 3. 4a). Une telle interaction a été démontrée au sein de deux pathosystèmes distincts, d'une part entre les produits Pto et AvrPto (Scofield *et al.*, 1996 ; Tang *et al.*, 1996) et tout dernièrement entre Pto et AvrPtoB (Kim *et al.*, 2002), d'autre part, entre les produits Pi-ta et Avr-Pita (Jia *et al.*, 2000), intervenant respectivement lors des réactions incompatibles tomate/*P. syringae* pv *tomato* et riz/*Magnaporthe grisea*.

Plus récemment, des alternatives au modèle ligand/récepteur, impliquant une interaction indirecte entre les produits R et Avr, ont été proposées (Bonas et Lahaye, 2002).

Notamment, le *guard model* ou "modèle de garde" (Van der Biezen et Jones, 1998) (Fig. 3. 4b) suppose l'interaction de la molécule effectrice Avr et d'une protéine de la plante, désignée comme "cible de la pathogénie" (*pathogenicity target*). Ces protéines-cibles agiraient comme des molécules de liaison entre les produits complémentaires R et Avr (Bonas et Van den Ackerveken, 1999). Les produits R seraient, les gardiens de ces protéines-cibles et, par reconnaissance du complexe produit Avr/cible de la pathogénie, ils initieraient les réponses de défense au sein de la plante. Par exemple, le gène de résistance *Prf*, de la classe NBS-LRR, dont l'activité est indispensable à la fonction du gène *Pto* (Salmeron *et al.*, 1996), pourrait agir comme une protéine R capable de reconnaître le complexe de pathogénie Pto/AvrPto, pour activer le mécanisme de résistance (Van der Biezen et Jones, 1998). Selon ce modèle, le produit Pto jouerait le rôle de cible de la pathogénie (Dangl et Jones, 2001).

Le modèle de garde permet d'expliquer différentes observations expérimentales :

- Certains produits R ont la capacité de "reconnaître" deux types d'agents pathogènes. C'est le cas des produits des gènes *Rpm1* (Grant *et al.*, 1995) et *Mi-1* (Vos *et al.*, 1998). Ces protéines pourraient "garder" le même produit, cible de produits Avr différents (Dixon *et al.*, 2000).

- Pour la majorité des couples R/Avr connus, aucune interaction physique directe n'a pu être mise en évidence (Dangl et Jones, 2001).

- L'absence de la cible d'assemblage adéquate expliquerait qu'un nombre important de gènes *R* ne sont fonctionnels qu'au sein de plantes relativement proches de l'espèce donneuse

de la résistance considérée, phénomène désigné sous le terme de "fonctionnalité restreinte au taxon" (*restricted taxonomic functionnality*) (Tai *et al.*, 1999a).

Par ailleurs, l'étude de certains pathosystèmes a révélé que plusieurs gènes *avr* de virus (Mestre *et al.*, 2000), de champignons (Orbach *et al.*, 2000) ou encore de bactéries (Orth *et al.*, 2000) codaient des protéines de type protéases. De plus, la mutation des sites catalytiques potentiels de ces protéines inhibent toute réponse de défense. Or, chez la drosophile, le récepteur Toll reconnaît le ligand Spätzle issu d'un processus de maturation protéolytique (Van Eeden et St Johnston, 1999). Par analogie, la possibilité de réactions protéolytiques, en amont de la reconnaissance de l'éliciteur Avr, a été proposée. Les éliciteurs des mécanismes de défense, reconnus (directement ou non) par la protéine R, pourraient être des produits de la plante hôte, devenus matures par protéolyse (Fig. 3. 4c). L'hypothèse évolutive que des protéines R aient "développé" des sites de clivage, afin de devenir des éliciteurs de défense, lors d'attaques d'origine protéolytique, a donc été proposée (Bonas et Lahaye, 2002).

Dans le cadre de l'interaction *Cladosporium fulvum*/tomate, il a été montré que la maturation des produits Avr9 (Van den Ackerveken *et al.*, 1993) et Avr4 (Joosten *et al.*, 1997) nécessite à la fois une activité protéase chez le champignon mais aussi chez la plante. Plus récemment, il a été montré que le produit du gène *Rcr3* correspond à une cystéine protéase secrétée et que celle-ci joue un rôle spécifique en amont de la résistance conférée par *Cf-2* (Krüger *et al.*, 2002). Sans exclure la possibilité de la formation d'un complexe entre le produit d'avirulence Avr2 et le produit Rcr3, complexe qui serait alors reconnu par le produit Cf-2, ces auteurs suggèrent que Rcr3 pourrait agir dans la maturation du produit d'avirulence Avr2 pour générer un ligand mature reconnu par Cf-2.

#### Complexité des signaux de transduction aboutissant aux mécanismes de défense

Comme cela a pu être décrit auparavant, les produits des gènes R interagiraient en aval de l'étape de reconnaissance avec des composés de cascades signalétiques aboutissant à la mise en place des mécanismes de résistance. Or, malgré l'importante diversité des modes d'infection phytopathogène observée, l'analyse des produits de gènes R a révélé l'intervention d'un nombre limité de motifs structuraux. Une telle organisation suggère non seulement une convergence des mécanismes de reconnaissance, mais également une restriction de la diversité des voies de signalisation activées en aval de cette étape (Feys et Parker, 2000). Cette hypothèse est confortée par différentes données expérimentales.

Il a été montré que certains gènes impliqués dans des voies de signalisation, sont nécessaires à la fonction de différents gènes *R*. Ainsi, la mutation *rar1* chez l'orge compromet la fonction de 9 gènes distincts de la famille *Mla*, conférant la résistance à différents isolats de *E. graminis f. sp. hordei* (oïdium). Ceci permet de positionner le gène *Rar1* au niveau d'un point de convergence entre les voies de transduction sous-jacentes à la série de gènes *Mla-X* de l'orge (Schulze-Lefert et Vogel, 2000).

La convergence des voies de signalisation aboutissant aux mécanismes de défense a été observée, également, dans le contexte de résistance à des agents pathogènes très variés. Les mutants, *eds1* (*enhanced disease susceptibility*) (Parker *et al.*, 1996) et *ndr1* (*non-race-specific disease resistance*) (Century *et al.*, 1995), d'*Arabidopsis* sont affectés dans leur résistance vis-à-vis de bactéries et de champignons. Chez le tabac, les gènes *Rar1* et *Eds1* interviennent dans le cadre de la résistance au virus de la mosaïque du tabac (Liu *et al.*, 2002).

La comparaison de ces analyses de mutants suggère que l'activation d'une cascade de signalisation est déterminée par la structure du gène *R* intervenant, et non pas, par la nature de l'agent pathogène. Globalement, les mutations au sein des gènes *Eds1* impliqueraient la perte de la résistance conférée par les gènes *R* de la classe TIR-NBS-LRR, tandis que les mutants *ndr1* supprimeraient la résistance assurée par les gènes *R* de la classe non-TIR-NBS-LRR (Aarts *et al.*, 1998 ; Liu *et al.*, 2002) (Fig. 3. 5). Toutefois, chez *Arabidopsis*, ni la mutation *eds1*, ni la mutation *ndr1* ne suppriment significativement la résistance à *Peronospora parasitica*, conférée par le gène *Rpp8* de la classe non-TIR-NBS-LRR (McDowell *et al.*, 1998 ; Aarts *et al.*, 1998). L'existence d'autres voies de signalisation non encore identifiées est fortement suggérée. De plus, il a été mis en évidence que le gène *Rar1* est indispensable à la résistance conférée par les gènes *Mla*, de la classe non-TIR-NBS-LRR, chez l'orge (Schulze-Lefert et Vogel, 2000). Or, *Rar1* (i. e. son orthologue), est également nécessaire à l'activité du gène *N* chez le tabac (Liu *et al.*, 2002). Ce gène constitue le premier exemple de point de convergence dans les voies de signalisation de résistance race-spécifique conférée par des gènes *R* de classes TIR-NBS-LRR et non-TIR-NBS-LRR.

Par ailleurs, le criblage de mutants altérés dans une résistance race-spécifique a permis l'identification de gènes spécifiquement requis pour la fonction d'un gène *R* donné, soulignant l'existence d'association protéine R/composé auxiliaire, hautement discriminatoire.

Ainsi, chez la tomate, bien que les gènes *Cf-2* et *Cf-5* présentent 93% de similarités, la mutation *rcr3* supprime spécifiquement la résistance conférée par le gène *Cf-2*, sans altérer celle liée au gène *Cf-5* (Dixon *et al.*, 2000). De même, la mutation du gène *Pbs1*, identifié chez *Arabidopsis* 

21

(Swiderski et Innes, 2001), s'avère supprimer la résistance liée à *RPS5* et n'affecte aucune autre résistance spécifique testée (Warren *et al.*, 1999).

Enfin, des études récentes mettent en évidence la convergence de voies impliquées dans l'expression de gènes *R* spécifiques et dans l'accumulation d'acide salicylique (Feys *et al.*, 2001 ; Rustérucci *et al.*, 2001 ; Van der Biezen *et al.*, 2002). Ceci illustre bien la complexité et la convergence, non seulement entre les voies de signalisation à l'origine de l'activation de gènes *R* distincts, mais aussi entre les voies de signalisation conduisant aux mécanismes de la HR et de la SAR.

# 3.3 EVOLUTION DES GENES DE RESISTANCE

Les gènes *R* majoritairement identifiés à ce jour appartiennent aux classes TIR et non-TIR NBS-LRR. A partir du génome entièrement séquencé de la plante modèle dicotylédone *Arabidopsis thaliana* écotype Col-O, 166 gènes de type NBS-LRR (dont 65% de la classe TIR-NBS-LRR) ont été répertoriés (Richly *et al.*, 2002). Le génome du riz, plante modèle monocotylédone, présente environ 600 gènes de type NBS-LRR. Contrairement à *Arabidopsis*, seuls 5 gènes potentiels de type NBS-LRR seraient associés, en position N-terminale, à une région présentant de faibles homologies avec les domaines TIR (Goff *et al.*, 2002). L'ensemble des travaux traitant de l'évolution des gènes *R* s'appuie essentiellement sur les connaissances relatives à ces classes de gènes.

# 3. 3. 1 Organisation moléculaire des gènes *R* et mécanismes génétiques impliqués dans leur évolution

## Complexité des locus R

La survie d'un organisme est fonction de la présence de systèmes génétiques spécifiques capables de maintenir un certain niveau de diversité des mécanismes de défense de l'organisme considéré, face à un environnement variable. Tout comme au sein du complexe majeur d'histocompatibilité chez l'homme (Trowsdale, 2002), des études génétiques et moléculaires ont révélé une organisation génomique particulière des gènes de résistance chez les plantes. La plupart d'entre eux sont génétiquement étroitement associés à d'autres gènes

ou séquences homologues, et constituent des locus complexes ou clusters (Pryor et Ellis, 1993).

# Locus simple à multiples spécificités de résistance

La majorité des locus aujourd'hui répertoriés présente de multiples spécificités. Ces locus présentent de nombreux allèles qui confèrent la résistance à une race spécifique de l'agent pathogène considéré. Chez le lin, par exemple, la résistance à la rouille *Melampsora lini* est assurée par au moins 30 allèles regroupés en 5 locus polymorphes (K, L, M, N et P) présentant des organisations variées (Fig. 3. 6a). Le locus *L*, notamment, correspondrait à un gène unique constituant une série allélique d'au moins 13 allèles (*L*, *L1* à *L11* et *LH*) relatifs à 13 spécificités (Ellis *et al.*, 1999). De même le locus *Rpp13*, conférant la résistance à *P. parasitica* chez *A. thaliana* correspond à un locus simple présentant 5 allèles distincts fonctionnels (Bittner-Eddy *et al.*, 2000). Dans plusieurs autres cas de locus simple de gène *R*, tels que *Rpm1* (Grant *et al.*, 1996) et *Rps2* (Caicedo *et al.*, 1999) chez *Arabidopsis*, un allèle résistant a été caractérisé et la spécificité d'autres membres de la série allélique n'a pas été identifiée.

# Co-localisation de gènes de résistance à différents agents pathogènes

Chez diverses espèces végétales, des locus de résistance, conférant chacun une résistance spécifique dirigée contre des agents pathogènes différents, sont co-localisés sur un même fragment chromosomique (Jones *et al.*, 1993 ; Witsenboer *et al.*, 1995 ; Spielmeyer *et al.*, 1998). Ainsi, à l'extrémité du bras court du chromosome VI de la tomate, 6 spécificités de résistance sont regroupées sur une région d'environ 20 cM : deux spécificités de résistance à *C. fulvum* (*Cf-2* et *Cf-5*) ; le gène *Mi/Meu1*, qui confére la résistance au nématode *Meloidogyne incognita* et au puceron *Macrosiphum euphorbiae* (Dickinson *et al.*, 1993 ; Kaloshian *et al.*, 1995) ; le gène de résistance *Ol-1* à *Oïdium lycopersicum* (Vanderbeeck *et al.*, 1994) et le gène *Ty-1*, gène de tolérance au TYLCV (*Tomato Yellow Leaf Curl Virus*) (Zamir *et al.*, 1994).

# Multiples copies de séquences analogues organisées en clusters

Un grand nombre de séquences présentant des similarités avec les gènes *R* clonés ont été mises en évidence au sein des génomes de plantes : on parle d'analogues de gènes de résistance (ou RGA, *Resistance Gene Analogs*). Ces RGA ont été le plus souvent obtenus et caractérisés suite à des amplifications *in vitro* (PCR) utilisant des amorces oligonucléotidiques dégénérées correspondant à des domaines conservés des gènes de résistance (Leister *et al.*, 1996). La plupart de ces RGA sont organisés en clusters et co-localisés avec des gènes de

résistance connus. Le locus M (Fig. 3. 6b), notamment est composé d'environ 15 gènes en cluster, distribués sur moins d'1 Mb et assurant 7 spécificités connues (Ellis *et al.*, 1997). Le locus *Xa21* chez le riz est constitué d'au moins 8 membres répartis sur 230 kb (Williams *et al.*, 1996) et le locus *Cf-4/9* de la tomate présente 5 membres étroitement liés (Parniske *et al.*, 1997). Une telle organisation correspond à un ensemble de séquences répétées en tandem, toutes ces séquences paralogues ne constituent pas forcément des gènes fonctionnels (*Cf* § Produits tronqués et pseudogènes). Les gènes de résistance de types NBS-LRR ou LRR/kinase et les séquences qui leur sont apparentées, constituent ainsi des familles multigéniques majeures au sein du génome des plantes.

Les clusters homogènes constitués d'un gène R et de séquences paralogues de la même famille se distinguent des clusters hétérogènes. Ces derniers associent un gène R à des séquences RGA ou à un autre gène R de familles différentes. Toutefois, aucun cluster hétérogène présentant à la fois un gène de classe TIR-NBS-LRR et un gène de classe non-TIR-NBS-LRR n'a été identifié (Richly *et al.*, 2002).

L'organisation en clusters des gènes *R* observée à l'échelle génomique a abouti aux termes de super- ou mega-clusters de gènes *R*. Ces groupes de gènes s'étendent sur plusieurs millions de paires de bases et peuvent présenter quelques douzaines de gènes (Young, 2000). Par exemple, le chromosome VI d'*Arabidopsis* présente de multiples séquences NBS-LRR co-localisées, elles constituent un mega-cluster qui s'étend sur 4,6 Mb (Meyers *et al.*, 1999; Young, 2000).

## Mécanismes génétiques impliqués dans l'évolution des gènes R

L'organisation et l'instabilité des locus de gènes *R* ont été largement étudiées. Les mécanismes moléculaires impliqués dans leur évolution et dans l'apparition de nouvelles spécificités de résistance, en réponse à la pression parasitaire, ont été mis en évidence (Leister *et al.*, 1998 ; McDowell *et al.*, 1998 ; Parniske *et al.*, 1999). Au sein des clusters de gènes *R* analysés, l'accumulation de substitutions (mutation d'une seule base), les recombinaisons, les délétions ou encore les transpositions, sont autant de mécanismes génétiques à l'origine de nouvelles séquences fonctionnelles pouvant présenter de nouvelles spécificités de résistance (Parniske *et al.*, 1997 ; Song *et al.*, 1997 ; Leister *et al.*, 1998 ; McDowell *et al.*, 1998).

### Substitutions

Les mutations ponctuelles de nucléotides (transversions et transitions) sont une des sources de variation des molécules d'ADN. En raison de la dégénérescence du code génétique, ces mutations n'impliquent pas nécessairement une modification de la séquence correspondante en acide aminé : ce sont les substitutions synonymes (ou silencieuses), par opposition aux substitutions non-synonymes (ou non-silencieuses). L'analyse comparative des taux de substitutions synonymes (K_s) et non-synonymes (K_a) a été effectuée au sein de domaines LRR de gènes *R* (Michelmore et Meyers, 1998, Bergelson *et al.*, 2001). Les taux importants de substitutions non-synonymes observés suggèrent que ce mécanisme est impliqué dans la diversification des domaines potentiels de reconnaissance, contribuant à leur adaptation en tant que récepteurs spécifiques (Parniske *et al.*, 1997 ; McDowell *et al.*, 1998 ; Bergelson *et al.*, 2001).

#### Différents types de recombinaison

Il est apparu que le processus classique de recombinaison génétique (Kover et Caicedo, 2001) et les phénomènes de conversion génique (i. e. transfert unidirectionnel d'information génétique) (Bendahmane *et al.*, 2000) constituent une source majeure de diversification des gènes de résistance (Fig. 3. 7). De même, des événements de crossing-over inégal, résultant d'un mauvais appariement entre des séquences apparentées répétées en tandem, interviennent dans l'évolution de gènes *R*. Ces recombinaisons inégales peuvent être intergéniques (entre séquences séparant les gènes), ou intragéniques (entre séquences de gènes).

Les recombinaisons inégales intergéniques seraient à l'origine de la variation en nombre des copies de gènes au niveau des clusters de gènes de résistance. La présence de patchwork de séquences homologues entre différents gènes *R* serait le résultat de crossing-over inégaux intragéniques. Cette possibilité a été suggérée à partir des analyses génétiques menées au niveau du locus *Rp1*, conférant la résistance à la rouille chez le maïs (Sudupak *et al.*, 1993). Depuis, les études comparatives des familles de gènes *Cf-4/9* (Parniske *et al.*, 1997 ; Van der Hoorn *et al.*, 2001) et *Xa21* (Song *et al.*, 1997) ont confirmé que les séquences homologues, observées entre divers membres de ces familles, étaient en partie issues d'événements de recombinaison entre séquences de séries de gènes en tandem (i. e. paralogues). Ce mécanisme de création de "gènes-hybrides" concernerait tout particulièrement la région LRR qui présente de nombreuses répétitions. Soulignons toutefois, que ces mécanismes, au sein d'un même cluster, ne sont pas exclusifs. L'étude du cluster de résistance *Dm3* chez la laitue

montre que ces deux types de recombinaisons inégales interviennent au sein d'un même locus (Chin *et al.*, 2001).

#### Eléments transposables

Les éléments transposables sont des composants majeurs de la plupart des génomes de plantes. Certains clusters de gènes R présentent les séquences d'éléments transposables et différents rôles leur sont attribués dans l'évolution des locus R. La mutation et la délétion d'allèle de résistance ont été associées à l'insertion d'éléments transposables au sein de gènes R (Johal *et al.*, 1992 ; Henk *et al.*, 1999). D'autres observations montrent que les éléments transposables favoriseraient les événements de crossing-over inégal à l'origine, par exemple, de la duplication des gènes R (Wessler *et al.*, 1995). Les éléments transposables constitueraient ainsi un outil efficace pour la diversification de locus R.

Notamment, 15 éléments transposables potentiels ont été identifiés au sein de la famille *Xa21* (Song *et al.*,1997). Deux d'entre eux sont localisés au niveau de régions codantes et seraient à l'origine de la création de 2 nouveaux cadres de lecture codant chacun une protéine tronquée (Richter et Ronald, 2000).

#### Produits tronqués et pseudogènes

Comme rapporté précédemment, les gènes *R* sont associés à de nombreuses copies de RGA au sein de clusters. Ces séquences paralogues, plus ou moins diversifiées, peuvent correspondre à des pseudogènes (i. e. séquences d'ADN, apparentées à des gènes fonctionnels, non-exprimées ou inactives du fait de mutations) ou encore à des gènes de fonction non-identifiée.

Le gène tronqué *Xa21D*, membre de la famille *Xa21*, présente un domaine LRR sans domaine trans-membranaire ni domaine kinase. Il confère toutefois, une résistance partielle selon le même spectre de résistance que le gène *Xa21* (Song *et al.*, 1997). D'autres gènes tronqués, de type NBS-LRR, par exemple, au sein des locus *Rpp5* et *Rp1*, coderaient des protéines fonctionnelles, mais aucune activité ne leur a été associée (Parker *et al.*, 1997 ; Ayliffe *et al.*, 1999 ; Collins *et al.*, 1999). Ces séquences paralogues ne présentent généralement qu'un unique événement de mutation, sans accumulation d'autres mutations. Certaines auraient (ou auraient eu) la capacité d'interagir avec des éliciteurs (Noel *et al.*, 1999 ; Sun *et al.*, 2001). Sous l'effet de mécanismes génétiques permettant de garder intacte leur région codante, ces séquences pseudogènes, ou non, semblent constituer "un réservoir de séquences" contribuant à l'évolution des familles de gènes *R* (Michelmore et Meyers, 1998 ; Hulbert *et al.*, 2001).

# 3. 3. 2 Evolution des gènes R : des modèles d'évolution...

## Notion de pression de sélection réciproque ou co-évolution

La très haute spécificité des relations gène-à-gène trouve son origine dans la co-évolution des espèces de plantes et de micro-organismes au sein d'un même biotope (Rausher, 1996). Etant donnée la non-mobilité des plantes, on conçoit que leur survie requiert l'apparition de nouvelles résistances. La génération active de nouvelles spécificités confère alors un avantage sélectif (*fitness*) à la plante-hôte menacée (Geffroy *et al.*, 1999). Cependant, bien qu'il s'agisse de caractères transmis génétiquement, l'efficacité des gènes R apparaît transitoire. L'expression d'une résistance induit une pression de sélection, sur la population d'agents pathogènes, qui favoriserait le développement de souches virulentes : le gène R est contourné par la modification du caractère avirulent du micro-organisme qui devient alors virulent. La perte de cette fonction d'avirulence peut être occasionnée par de simples mutations chez l'agent pathogène. De nouveau, la population de parasites exerce, alors, une forte sélection sur les plantes induisant la génération de gènes R présentant de nouvelles spécificités.

Dans ce processus de co-évolution permanent, la courte durée de génération et la taille importante de certaines populations d'agents pathogènes constituent des paramètres déterminants (Bennetzen *et al.*, 1994 ; Kalt et Shykoff, 1998). Toutefois, le fait même que les gènes d'avirulence soient présents chez les micro-organismes alors qu'ils sont néfastes pour l'état parasitaire laisse supposer qu'ils sont nécessaires à leur adaptation ou leur développement. Dans ces conditions, il pourrait être désavantageux pour l'agent pathogène de perdre une telle fonction. Des analyses par mutation ont démontré, dans plusieurs cas, que les gènes *avr* présentaient un avantage sélectif pour l'agent pathogène en absence du gène *R* correspondant (Staskawicz *et al.*, 2001). Il est également postulé que les gènes *avr* agiraient à l'origine comme des facteurs de virulence, lesquels au cours de la co-évolution plante/parasite, seraient devenus la cible des gènes *R* des plantes.

# Déterminisme de la spécificité des gènes de la classe NBS-LRR

La connaissance des bases moléculaires de la spécificité de reconnaissance intervenant dans le système gène-à-gène reste encore un objectif important.

L'analyse comparative de gènes *R* semble montrer que la spécificité de résistance réside largement dans la séquence des domaines LRR. Classiquement, de façon à maintenir la fonction d'un gène, la majorité des gènes montre une fréquence plus importante des substitutions synonymes (K_s) par rapport aux substitutions non-synonymes (K_a). Or, dans le cas des gènes *R*, il a été observé pour la première fois au sein des séquences membres de la famille *Cf-4/9* que le taux de substitutions non-synonymes était majoritaire (Parniske *et al.*, 1997). Cette observation a été étendue aux domaines LRR des différents membres de la famille *Xa21* (Wang *et al.*, 1998) ainsi qu'aux domaines LRR de plusieurs locus NBS-LRR dont *Rpp8* (McDowell *et al.*, 1998), *Rpp5* (Noel *et al.*, 1999), *Rpp1* (Botella *et al.*, 1998), *Dm3* (Meyers *et al.*, 1998) et *Rp1* (Sun *et al.*, 2001). Ces régions LRR, connues pour intervenir dans les interactions protéines-protéines, apparaissent soumises à une pression de diversification favorisant la variabilité des acides aminés supposés assurer un contact de surface. Ces données sont en accord avec l'implication potentielle des régions LRR dans les processus de reconnaissance spécifique.

Par ailleurs, l'étude du locus *L*, conférant la résistance à la rouille chez le lin, a montré que la région LRR ne contrôle pas exclusivement la spécificité de reconnaissance (Ellis *et al.*, 1999). En effet, bien que les allèles *L6* et *L7* codent des spécificités de résistances distinctes, leurs produits présentent une région LRR identique et ne se différencient qu'au niveau de leur domaine TIR. De plus, non seulement une pression de sélection exercée sur les résidus acido-aminés des domaines TIR des allèles *L* a été mise en évidence, mais une co-évolution entre ces résidus et ceux du domaine LRR semble nécessaire pour assurer la spécificité du produit R (Luck *et al.*, 2000).

### Compromis entre diversification et conservation

La dynamique d'évolution des familles de gènes *R* est déterminée par deux forces opposées. D'une part, le besoin de générer de nouvelles spécificités en réponse à l'évolution des agents pathogènes se traduit par une diversification des domaines potentiels de reconnaissance (i. e. LRR). Ceci est révélé notamment par l'analyse des taux de substitutions synonymes et nonsynonymes observés au sein des domaines LRR (i. e. K_a:K_s>1 ; pour synthèse, Bergelson *et al.*, 2001). D'autre part, les gènes fonctionnels doivent être conservés en absence, ou lors de faible pression parasitaire, et protégés contre les forces d'homogénéisation invariablement associées aux mécanismes d'échanges de séquences. Les gènes *R* au sein d'une population végétale naturelle seraient généralement soumis à une sélection dite fréquence-dépendante (Van der Hoorn *et al.*, 2002). En effet, différentes observations, la longévité (Stahl *et al.*, 1999 ; Riely et Martin, 2001) et l'évolution lente de certains gènes *R* liée à l'absence de recombinaison (Wei *et al.*, 1999 ; Chin *et al.*, 2001 ; Van der Hoorn *et al.*, 2001), mais aussi l'importante diversité allélique des locus *R* observés (Caicedo *et al.*, 1999) ou encore la co-existence au sein de populations naturelles de plantes capables et non-capables de reconnaître un facteur Avr donné (Stahl *et al.*, 1999 ; Van der Hoorn, 2001), confortent cette idée. Cette pression de sélection s'exercerait à l'échelle de "l'éco-biologie" de l'interaction plante/micro-organisme considérée (e. g. systèmes de reproduction de la plante, tailles et durées de génération respectives des populations de l'hôte végétal et du parasite, organes infectés aérien ou racinaire).

Différentes théories quant à l'évolution des gènes *R* ont été proposées. L'hypothèse *Arms Race* (ou "course à l'armement") impliquant un polymorphisme transitoire associé au renouvellement rapide des gènes *R*, a rapidement été écartée (Stahl et Bishop, 2000). D'autres théories ont été développées notamment celle du *Trench Warfare* ("guerre des tranchées") (Stahl *et al.*, 1999) selon laquelle la prévalence d'un gène *R* serait relative à l'avantage sélectif que celui-ci confère à la plante-hôte, ou encore, la théorie *Birth-and-Death* où les gènes *R* seraient soumis à une sélection de diversification menant à la génération de nouvelles spécificités (Michelmore et Meyers, 1998).

Toutefois, l'acquisition récente de données expérimentales souligne la variabilité considérable des modes d'évolution des locus de résistance chez les plantes. Les dogmes actuellement proposés sont adaptés à la structure (i. e. taille, complexité) et/ou à l'âge d'un nombre restreint de gènes de résistance, voir à seulement une région donnée de ces gènes. En effet, il apparaît qu'un modèle unique d'évolution des gènes R ne puisse répondre à la diversité des modes d'évolution des gènes R observés (e. g. importance relative et fréquence, à court- et long- termes des mécanismes génétique et moléculaire impliqués, variabilité considérable des interactions plante-parasites relatives aux gènes R et Avr, mais aussi aux caractéristiques de la plante-hôte et de l'agent pathogène).

Les informations issues d'expérimentations récentes et l'analyse comparative entre les séquences de locus *R*, orthologues et paralogues, ont permis de distinguer deux situations "types" de gènes *R* répondant à deux systèmes extrêmes d'évolution (Michelmore, 2002). Entre ces deux situations types, il existerait toute une gamme de modes évolutifs intermédiaires. Chacun serait configuré par les caractéristiques génétique et biologique de l'interaction phytopathogène considérée.

Le premier modèle expérimental d'évolution des gènes R suggère que les crossing-over et les conversions de gènes sont les mécanismes génétiques majeurs menant à la génération rapide de nouvelles spécificités. Ces clusters de gènes R sont supposés être particulièrement dynamiques et constitués de séries instables de séquences homologues (Ellis et al. 1997; Hulbert, 1997). Ces mécanismes ont effectivement été corrélés à la création de nouvelles spécificités de résistance au sein du locus Rp1 chez le maïs (Sudupak et al., 1993; Richter et al., 1995). De même, les recombinaisons inégales intragéniques au sein de régions LRR se sont montrées impliquées dans la perte de résistance au niveau des locus M et Rpp5 chez le lin et chez Arabidopsis, respectivement (Anderson et al., 1997; Parker et al., 1997). De plus, des analyses de séquences ont montré que les recombinaisons et les conversions géniques étaient responsables de la formation de gènes chimériques au niveau des clusters Cf-4/9 et Cf-2/5 chez la tomate (Parniske et al., 1997; Dixon et al., 1998; Parniskes et Jones, 1999), Xa21 chez le riz (Song et al., 1997), Dm3 chez la laitue (Meyers et al., 1998) ou encore Rpp5 et Rpp8 chez Arabidopsis (McDowell et al., 1998 ; Noel et al., 1999). Enfin, comme cité précédemment, il peut être rappelé que les ratios K_a/K_s, observés au sein des domaines LRR des gènes des clusters Cf-4/9, Xa21, Rpp1, Rpp8, Rpp5 et Dm3 sont tous supérieurs à 1 (Parniske et al., 1997; Botella et al., 1998; McDowell et al., 1998; Michelmore et Meyers, 1998; Wang et al., 1998; Noel et al., 1999). Ce type de locus R, fortement soumis à une sélection de diversification, serait donc caractérisé par une évolution rapide et la génération d'un nombre important d'allèles, chaque allèle étant peu fréquent au sein de la population hôte.

Par opposition à ce premier modèle d'évolution, l'analyse comparative de différents locus suggère que les gènes *R* puissent évoluer lentement et que la structure de certains clusters de résistance serait relativement stable. Ainsi, le séquençage complet des locus *Pto* à partir d'haplotypes/génotypes, résistant et sensible, chez la tomate a mis en évidence une structure conservée de ces locus présentant des relations évidentes entre orthologues (Chin *et al.*, 2001). Au sein des clusters *Dm*3, *Cf*, *Xa*21 et *Pto*, il a été montré que les séquences orthologues partagent entre elles plus de similarités qu'elles n'en partagent avec leurs séquences paralogues respectives. De plus, contrairement à ce qui est attendu sous l'effet de fréquents crossing-over et de conversions de gènes, ces séquences paralogues ne présentent pas de signes d'homogénéisation particulière (Parniske *et al.*, 1997 ; Meyers *et al.*, 1998 ; Chin *et al.*, 2001). L'ensemble de ces informations suggère qu'une sélection de purification plutôt que des échanges importants de séquences entre paralogues, agirait au sein des ces clusters. Ceux-ci seraient caractérisés par une évolution lente et la génération d'un nombre limité d'allèles, mais de fréquence importante au sein de la population hôte.
# 4. UNE CONTRAINTE AGRONOMIQUE MAJEURE, LES NEMATODES

Les nématodes (ou "vers ronds") appartiennent à l'embranchement des némathelminthes. Ce groupe zoologique, présent dans presque tous les milieux, affiche une grande diversité (plus de 25 000 espèces décrites) (Maggenti, 1991). Il comprend des parasites de vertébrés, d'insectes, et de plantes ainsi que des espèces bactériophage ou saprophage (*Caenorhabditis elegans*).

#### 4. 1 L'AGENT PHYTOPATHOGENE

#### 4. 1. 1 Les nématodes phytoparasites

Les nématodes phytoparasites sont des parasites obligatoires biotrophes. Présents en abondance dans tous les sols arables, ils se nourrissent des cellules végétales vivantes. A ce jour, environ 3 000 espèces phytophages sont décrites et se répartissent en deux ordres, les *Tylenchidae* et les *Dorylaimidae*. L'ordre des *Tylenchidae* regroupe environ 2 200 espèces parasitant un grand nombre de cultures dans toutes les zones agricoles du monde, il constitue une contrainte économique importante (Luc *et al.*, 1990).

Selon la relation établie entre le nématode et la plante-hôte, différents groupes de nématodes phytoparasites sont distingués (Sijmons *et al.*, 1994). Les ectoparasites (migrateurs et sédentaires) vivent libres dans le sol, ils ne pénètrent jamais à l'intérieur de la racine et pour se nourrir, ils se servent de leur stylet. Ces nématodes sont généralement responsables de dommages limités. Les nématodes endoparasites pénètrent entièrement dans la racine. Les endoparasites migrateurs restent mobiles à tous les stades de leur développement (e. g. espèces du genre *Pratylenchus*). Ils sont très polyphages et provoquent des lésions tissulaires massives (Wyss, 1997). Les endoparasites sédentaires, appartenant à la famille des *Heteroderidae*, sont responsables des dommages les plus importants au niveau mondial. Ils établissent une interaction complexe avec la plante-hôte. Au sein de cette famille se distinguent, les nématodes à kystes (*Heterodera* spp., *Globodera* spp.) qui n'attaquent qu'un nombre limité d'espèces végétales, et les nématodes à galles (*Meloidogyne* spp.) particulièrement polyphages (Hussey et Williamson, 1998).

### 4. 1. 2 Les nématodes du genre Meloidogyne

#### Généralités

Les nématodes à galles se distinguent, notamment, par une vaste distribution géographique, aussi bien dans les régions tempérées que tropicales, et par un très large spectre d'hôtes comprenant plus de 3 000 espèces de plantes (Trugdill et Block, 2001 ; Jepson, 1987). Ces ravageurs s'attaquent aussi bien aux grandes cultures qu'aux cultures maraîchères, fruitières et florales. Dans les régions tropicales et tempérées chaudes, favorables à leur développement, les nématodes à galles constituent l'un des principaux ennemis des plantations et cultures (e. g. caféier, cotonnier, bananier, ananas, maïs et sorgho) ainsi que des cultures maraîchères cultivées sous serre (Netscher et Sikora, 1990). A eux seuls, les nématodes à galles du genre *Meloidogyne* seraient responsables de 5% des pertes agricoles à l'échelle mondiale.

#### Mode de reproduction

Les espèces appartenant au genre *Meloidogyne* présentent différents modes de reproduction. Certaines sont amphimictiques (*M. carolinensis*, *M. megatyla*...), d'autres se reproduisent de façon parthénogénétique méiotique facultative (*M. hapla*, *M. chitwoodi*, *M. exigua*...) ou parthénogénétique mitotique stricte (*M. arenaria*, *M. incognita*, *M. javanica*, *M. paranensis*) (Triantaphyllou, 1985 ; Trugdill et Block, 2001)

La parthénogenèse méiotique facultative remplace la reproduction sexuée en absence de mâles. Dans le cas de la reproduction par parthénogenèse mitotique stricte, le mâle n'intervient pas dans la reproduction. Une seule division non-réductionnelle se produit et le nombre de chromosomes est maintenu durant toute la maturation de l'ovocyte (Dalmasso et Bergé, 1975).

#### Cycle biologique

La durée du cycle biologique des *Meloidogyne* est fonction principalement de la température et de la nature même de la plante-hôte. Pour des températures variant entre 20 et 30°C, le stade adulte du nématode peut être atteint 16 à 30 jours après pénétration des larves infestantes. A des températures inférieures à 15°C et supérieures à 33°C, le cycle biologique des *Meloidogyne* ne peut être complété (Noe *et al.*, 1991).

Les nématodes à galles du genre *Meloidogyne* présente un parasitisme composé d'une phase mobile et d'une phase sédentaire (Fig. 4. 1).

Lors de la phase mobile, la larve infestante, dite juvénile de deuxième stade (J2), après exploration de la surface racinaire, pénètre dans la racine, soit au niveau de la zone d'élongation, soit au niveau de petites lésions déjà existantes. La larve progresse de façon intercellulaire le long du cylindre central de la racine et se sédentarise dans la zone de différenciation.

Lors de la phase sédentaire, la larve induit la formation d'une structure complexe, le site nourricier. Cette structure est constituée de 5 à 7 cellules végétales, multinuclées au métabolisme hyperactif et appelées cellules géantes (Wyss *et al.*, 1992). La formation de ce site résulte de la différenciation des cellules procambiales en réponse à des signaux émis par le nématode (e. g. des sécrétions salivaires). La formation des cellules géantes s'accompagne de l'hypertrophie puis de l'hyperplasie des cellules corticales environnantes, ce qui conduit à la formation d'une galle au niveau de la zone parasitée (Lowenberg *et al.*, 1960). Ces galles sont généralement visibles 2 à 3 jours après la pénétration des J2.

Après induction des cellules nourricières, les juvéniles J2 s'alimentent et réalisent 3 mues (juvéniles de 3^{ème} et 4^{ème} stade - adulte) à l'issue desquelles a lieu la différenciation sexuelle. Cette différenciation aboutit d'une part à la formation de mâles, filiformes et mobiles, lesquels quittent les tissus de l'hôte, d'autre part, à la formation de femelles renflées, sédentaires qui continuent de s'alimenter durant plusieurs semaines. Chez les nématodes à galles, la fréquence des mâles augmentent lorsque les conditions sont défavorables. Les femelles adultes de *Meloidogyne*, à maturité sexuelle, pondent plusieurs centaines d'œufs au sein d'une matrice protectrice gélatineuse formant une masse intra ou extra racinaire. Après éclosion, les femelles meurent, les cellules géantes et les tissus de la galle dégénèrent. Quant aux larves J2 disséminées dans le sol, elles peuvent soit pénétrer aussitôt dans la racine et reprendre un cycle de reproduction, soit attendre jusqu'à plusieurs mois en cas d'absence d'hôte ou de conditions défavorables.

#### Symptomatologie

Les premiers symptômes de la maladie sont observés sur la partie aérienne de la plante, il ne s'agit pas de symptômes spécifiques. Globalement, la croissance est retardée, les feuilles présentent des signes de carence (i. e. chlorose). Plus tardivement, la partie aérienne présente un aspect chétif, la floraison et la fructification peuvent être fortement diminuées et, dans les

33

cas les plus graves, une défoliation pouvant aller jusqu'à la mort de la plante infectée, peut être observée.

Au niveau racinaire, les déformations dues aux galles constituent le symptôme le plus caractéristique de l'attaque des *Meloidogyne*. Les conséquences d'une forte infestation sont la réduction du chevelu racinaire, une spoliation de la plante au profit du nématode qui détourne les nutriments synthétisés par le végétal et une forte perturbation de la structure vasculaire de la racine infectée (De Guiran et Netsher, 1970).

La baisse de l'alimentation hydrique et minérale de la plante et, une résistance à la sécheresse diminuée, aboutissent au dépérissement de la plante. Cette déficience globale peut être accélérée et aggravée lors des saisons sèches. De plus, l'affaiblissement généralisé de la plante favorise l'attaque d'autres agents pathogènes (Williamson et Hussey, 1996).

#### 4. 2 LES GENES DE RESISTANCE SPECIFIQUE AUX NEMATODES

#### 4. 2. 1 Les gènes de résistance aux nématodes clonés

La relation étroite entre le nématode endoparasite et la plante infectée a favorisé l'émergence de gènes de résistance chez les nombreuses espèces végétales soumises à cette pression parasitaire. La majorité de ces gènes de résistance, identifiés au sein d'espèces végétales variées, interviennent lors de résistance spécifique de type gène-à-gène (Flor, 1971). A ce jour, une quinzaine de gènes de résistance spécifiques aux nématodes endoparasites ont été identifiés et cartographiés (Tableau 4. 1). Trois de ces gènes ont été clonés et caractérisés. A l'exception du gène *Hs1*^{pro-1}, les données actuelles tendent à montrer que les gènes de résistance de la classe NBS-LRR sont fréquemment rencontrés au cours des interactions incompatibles plantes-nématodes (Williamson, 1999).

En effet, le gène de résistance *Hs1^{pro-1}*, cloné chez la betterave à sucre, confère la résistance aux nématodes *H. schachtii* (Cai *et al.*, 1997). La séquence et la structure de *Hs1^{pro-1}* (i. e. une région transmembranaire associée à une région riche en leucines) n'est semblable à aucune autre protéine connue.

Le gène *Mi* -présenté ultérieurement- isolé d'une espèce sauvage de tomate, *L. peruvianum*, confère, entre autre, la résistance à plusieurs espèces de nématodes du genre *Meloidogyne* sp. (Williamson, 1998). Le gène *Gpa2* qui confère la résistance aux nématodes *Globodera pallida*, a été cloné chez la pomme de terre (Van der Vossen *et al.*, 2000). Ces deux gènes, *Mi* et *Gpa2*, sont membres de la famille des gènes *R* de type NBS-LRR. De plus, des études préliminaires laissent à penser que les gènes *Gro1* (Leister *et al.*, 1996) et *Cre3* (Lagudah *et al.*, 1997), lesquels confèrent respectivement la résistance à *G. rostochiensis* chez la pomme de terre et la résistance à *H. avenae* chez le blé, seraient également des gènes *R* de la classe NBS-LRR.

#### 4. 2. 2 L'exemple du gène *Mi* de la tomate

La résistance de la tomate à *M. incognita* a été mise en évidence chez l'espèce sauvage *Lycopersicon peruvanium* (Bayley, 1941). Dans les années 1940, cette résistance fut introduite au sein de l'espèce cultivée, *L. esculentum*, suite à l'obtention par sauvetage d'embryons d'un hybride interspécifique (Smith, 1944). Cet hybride constitue l'unique source de résistance aux nématodes à galles des variétés de tomates actuellement commercialisées. A ce jour, la majorité des tomates cultivées présente un gène unique, dominant, le gène *Mi.* Contrairement à la plupart des gènes de résistance spécifique, *Mi* confère la résistance à plusieurs isolats de 4 espèces de nématodes du genre *Meloidogyne* : *M. incognita* à l'origine du nom de ce gène, *M. javanica, M. arenaria* et *M. arabicida* (Quénéhervé *et al.*, 2002).

Le gène *Mi* a été isolé par une approche de clonage positionnel, après localisation du locus sur le bras court du chromosome VI de la tomate (Milligan et al., 1998). Les efforts entrepris pour localiser le gène ont longtemps été retardés du fait d'un important phénomène de répression de la recombinaison au niveau du locus *Mi* introgressé au sein des lignées *L. esculentum* (Ho et al., 1992; Liharska et al., 1996). A l'issue du clonage et du séguencage du locus Mi, trois cadres de lecture ouverts ont été identifiés : Mi-1.1 et Mi-1.2 correspondent à deux gènes intacts présentant des similarités de séquences avec les gènes de résistance connus, ils flanquent le troisième cadre de lecture ouvert. Ce troisième ORF, tronqué et présentant plusieurs codons-stop, correspondrait à un pseudogène. Des travaux de complémentation à partir d'espèces de tomate sensible ont mis en évidence que, contrairement à Mi-1.1, Mi-1.2 seul confère la résistance à la lignée transformée (Milligan et al., 1998). Par ailleurs, une étude antérieure rapportait la présence d'un gène Meu-1 conférant la résistance au puceron de la pomme de terre (M. euphorbiae) au sein des tomates introgressées avec Mi et mettait en évidence la liaison génétique indissociable de ces deux gènes (Kaloshian et al., 1995). La disponibilité des clones *Mi* et les travaux de complémentation ont permis de montrer que *Mi* et Meu-1 sont en fait le même gène (Rossi et al., 1998). Le gène Mi constitue ainsi le premier exemple de gène chez les plantes conférant une résistance à deux bio-agresseurs de phylums très éloignés, un nématode et un insecte.

35

L'analyse de la structure de *Mi* a montré que ce gène est membre de la classe NBS-LRR des gènes de résistance (Milligan *et al.*, 1998). Comme de nombreux gènes de résistance (cf § Complexité des locus *R*), la région génomique présentant le gène *Mi* est très complexe. Six gènes et QTL de résistance aux maladies sont regroupés sur le bras court du chromosome VI de la tomate. De plus, 6 séquences hautement similaires à *Mi* sont présentes au sein des plantes résistantes *L. peruvianum* et la majorité d'entre elles forment un cluster (Milligan *et al.*, 1998).

Enfin, le mécanisme par lequel le produit du gène *Mi* confère la résistance à la fois à des nématodes et à un puceron, reste encore mal connu. Toutefois, il a été mis en évidence que la fonction du locus *Rme1* est indispensable à l'expression des résistances conférées par *Mi* (de Llarduya *et al.,* 2001). Des mutants *rme1* révèlent une totale absence de résistance aux nématodes ainsi qu'au puceron. Par contre, ces mêmes mutants ne sont pas affectés dans leur résistance à *Fusarium oxysporum f.sp lycopersici* race 2 conférée par le gène, membre de la classe NBS-LRR, *I-2* de la tomate (Simons *et al.,* 1998). Le locus *Rme1* pourrait intervenir lors d'étapes précoces dans la cascade de signalisation relative à *Mi* (de Llarduya *et al.,* 2001).

# 4.3 LES NEMATODES DU GENRE *Meloidogyn*e parasites du cafeier *C. arabica*

#### 4.3.1 Importance économique

De nombreux genres et espèces de nématodes sont associés aux cultures de caféiers à travers le monde. A l'heure actuelle, 17 espèces de nématodes à galles du genre *Meloidogyne* ont été répertoriées au sein de systèmes racinaires de caféiers (Tableau 4 .2) (Campos *et al.*, 1990 ; Carneiro *et al.*, 2000). Ces nématodes présentent différents niveaux d'agressivité. Notamment, le cas du nématode *M. paranaensis*, capable de provoquer la mort des caféiers, est très préoccupant du fait de l'étendue de son aire de distribution (Fig. 4. 2) (Carneiro *et al.*, 2000). D'autres nématodes tels que *M. arenaria*, *M. arabicida* et *M. konaensis*, provoquent des dégâts également importants, mais relativement localisés.

Globalement, les dommages provoqués par les *Meloidogyne* sur la caféiculture en Amérique Centrale sont d'une grande importance économique au champ tout comme en pépinière (Bertrand *et al.*, 1995).

#### 4.3.2 Meloidogyne exigua

En Amérique Latine, le nématode à galles présentant la plus large distribution est *M. exigua*. Ce nématode a été identifié dans la majorité des cultures de caféier d'Amérique Centrale et d'Amérique du Sud (Tableau 4.2). Malgré l'agressivité réduite de ce nématode, comparativement aux espèces précédemment citées, il est responsable d'importantes pertes de production. Par exemple, au Costa Rica, les baisses de production dues aux attaques de *M. exigua* ont été estimées de 10 à 20% (Bertrand *et al.*, 1997).

L'interaction parasitaire *M. exigua*/caféier est actuellement le seul modèle ayant fait l'objet d'études préliminaires (Fig 4. 3).

#### Symptômes associés à *M. exigua* chez le caféier

Chez le caféier sensible (e. g. les variétés *C. arabica*), *M. exigua* provoque la formation de galles rondes majoritairement au niveau des jeunes racines. Ces galles se distribuent de façon isolées ou en groupes, multipliant jusqu'à 8 fois le diamètre de la racine infectée (Fig. 4. 4a). Leur couleur, blanche ou jaunâtre, tourne au brun lorsque les racines vieillissent (Di Vito *et al.*, 2000). Des zones de nécrose, observées au niveau des racines modifiées, peuvent être ultérieurement le site d'infections secondaires (Campos *et al.*, 1990).

Les modifications tissulaires des racines peuvent être observées dès 3 jours après infection (Rodrigues *et al.*, 2000). La dissection aléatoire des galles met en évidence, plusieurs femelles, au sein d'une même galle. La tête des femelles, profondément implantées au niveau du système vasculaire central, est entourée de 3 à 8 cellules géantes bien formées constituant le site nourricier (Di Vito *et al.*, 2000). Ces femelles sont associées à de très larges masses d'œufs (Bertrand *et al.*, 2001).

Généralement, le caféier sensible infecté présente au champ une croissance réduite, des feuilles chlorosées et une chute prématurée de ces feuilles. Dans le cas de jeunes plants, ces symptômes peuvent entraîner la mort (Campos *et al.*, 1990).

Dans le cas de caféiers résistants (e. g. variétés de *C. canephora*), infectés par *M. exigua*, leur croissance ne semble pas modifiée et la reproduction du nématode est très réduite voir nulle (Rodrigues *et al.*, 2000).

L'observation microscopique de racines de caféier résistant infectées par *M. exigua*, révèle l'absence totale de galle ou la présence de très petites galles (Fig. 4. 4b). L'examen de ces petites galles lorsqu'elles ne sont pas vides (hypertrophie seule), met en évidence un faible

niveau de développement du nématode : une seule femelle associée à une masse d'œufs faiblement ou anormalement développée. L'altération de la masse d'œufs serait liée au développement incomplet des cellules géantes (Rodrigues *et al.*, 2000 ; Bertrand *et al.*, 2001). Une réaction de résistance peut être observée au niveau de la stèle, 6 jours après infection. Les cellules géantes avant de mourir passent par une étape de dégradation des membranes et la formation de vacuoles autophagiques. La présence de ces cellules nécrotiques, autour et au sein du site nourricier altéré, pourrait être associée aux symptômes de la résistance génétique de type HR (Rodrigues *et al.*, 2000).

#### Sources de résistance à M. exigua chez le caféier

Plusieurs sources de résistance spécifique à l'encontre des nématodes à galles du genre *Meloidogyne* spp., ont été identifiées (Tableau 2. 1).

En ce qui concerne l'espèce *M. exigua*, tous les génotypes de *C. arabica*, sauvages ou cultivés, sont sensibles (Bertrand *et al.*, 2001). La résistance vis-à-vis de ce nématode a été identifiée au sein de l'espèce diploïde *C. canephora*. Ce caractère a été transféré par hybridation et rétrocroisement à l'espèce *C. arabica*. Différentes lignées dérivées d'hybrides interspécifiques, soit l'Hybride de Timor (spontané), soit l'hybride contrôlé ICATU développé au Brésil, présentent une résistance à *M. exigua* similaire à celle conférée par *C. canephora* (Silvarolla *et al.*, 1998).

Les premières études de plantes  $F_1$  et populations  $F_2$  provenant de l'HdT, en ségrégation pour la résistance à *M. exigua*, indiquent un déterminisme oligogénique de ce caractère (Bertrand *et al.*, 2001).

# 5. LES OBJECTIFS DE THESE

La pression phytoparasitaire exercée par les ravageurs sur les plantes cultivées représente l'une des contraintes majeures de l'agriculture actuelle. L'utilisation répandue de la lutte intégrée marque l'émergence d'une ère nouvelle dans le domaine de la protection des cultures tout en respectant l'environnement et les exigences, de qualité et de soucis sanitaires, des consommateurs. La lutte intégrée inclut notamment le développement et l'utilisation de variétés résistantes pour combattre les phytoparasites. La connaissance des interactions spécifiques entre les espèces végétales cultivées et les micro-organismes qui leur sont pathogènes, est devenue indispensable pour la sélection et l'amélioration variétale de ces plantes. Différentes stratégies, en phytopathologie et en génétique, sont mises en œuvre pour comprendre les mécanismes de résistance spécifique des plantes vis-à-vis de parasites.

Les travaux de recherche effectués au cours de cette thèse s'inscrivent dans le cadre du projet scientifique de l'équipe IRD-CIRAD "Physiologie et Génétique de la Résistance des Caféiers aux Parasites" mené à l'IRD-Montpellier. Les objectifs généraux entrepris au sein de cette équipe reposent sur deux modèles principaux d'interactions spécifiques caféier/micro-organisme, *C. arabica/Meloidogyne* spp. et *C. arabica/Hemileia vastatrix*. Ces objectifs sont les suivants :

- 1. identifier de nouvelles sources de résistance à ces maladies afin de les introduire au sein des variétés cultivées de *C. arabica*;
- 2. identifier et cartographier certains gènes gouvernant ces résistances, en vue de leur clonage et de l'analyse de leur fonctionnement ;
- 3. analyser les mécanismes physiopathologiques ;
- 4. optimiser les méthodes de gestion et d'utilisation de ces gènes en sélection, en relation avec les phénomènes d'évolution et d'introgression propres à *C. arabica*.

Après une présentation synthétique du matériel végétal développé et exploité au cours de ces travaux, les objectifs de cette thèse seront exposés.

# 5. 1 LE MATERIEL D'ETUDE

Le matériel végétal étudié appartient aux espèces *Coffea arabica, C. canephora* et *C. eugenioides*. Les accessions utilisées sont soit entretenues en serre à l'IRD (Montpellier) ou proviennent des collections situées au CATIE et au CICAFE (Costa Rica).

Dans les expérimentations effectuées, les lignées Arabica introgressées utilisées sont issues de croisements entre des variétés de *C. arabica* et des descendants de l'hybride interspécifique spontané (*C. arabica* x *C. canephora*), l'Hybride de Timor (Fig. 5. 1). Les fragments introgressés au sein des lignées Arabica sont issus du génome de *C. canephora*. Parmi les lignées Arabica introgressées, les cultivars IAPAR 59 et T5296 ont été sélectionnés pour leur résistance à *M. exigua* (Bertrand *et al.*, 2001).

Une autre série de matériel végétal a été utilisée : les descendances d'hybrides F1 entre des variétés cultivées, incluant les lignées introgressées, et des caféiers sauvages d'espèces

Arabica. Ces descendances  $F_2$  permettent d'étudier la ségrégation des caractères génétiques tels que la résistance aux nématodes. Plusieurs populations  $F_2$  (75 à 500 individus) en ségrégation pour la résistance à *M. exigua* [lignée introgressée résistante x *C. arabica* sensible] ont été développées au Costa Rica. L'évaluation de ces individus pour la résistance testée est basée sur l'indice de galles observé 3 mois après inoculation du nématode (Anthony *et al.*, 2002b).

### 5. 2 LES OBJECTIFS DE CETTE THESE

Dans le contexte de l'interaction *C. arabica/M. exigua*, les objectifs de cette thèse consistent, d'une part, en l'analyse globale et en l'amélioration des connaissances des gènes de résistance au sein du genre *Coffea*, et d'autre part, en l'identification et la cartographie du gène de résistance au nématode à galles, *M. exigua*, chez le caféier. Les travaux de recherche effectués s'organisent autour de quatre analyses principales. Ces quatre études font chacune l'objet d'un chapitre de résultats rédigé sous la forme d'article scientifique.

- La première partie de ces résultats (chapitre 2) s'intéresse à la diversité et à l'évolution des gènes de résistance chez le caféier. Cette analyse s'appuie sur la mise en évidence de séquences analogues de gènes de résistance (RGA) de type NBS au sein du génome des caféiers.

- Parallèlement à cette approche visant l'ensemble des gènes *R* de type NBS chez le caféier, l'analyse du modèle d'interaction spécifique étudié au laboratoire, *C. arabica/M. exigua*, constitue une seconde partie de résultats (chapitre 3). Le déterminisme génétique de cette résistance a été précisé et la cartographie génétique du locus de résistance au nématode *M. exigua*, *Mex-1*, a été entreprise.

- La troisième partie consacrée aux résultats de ces travaux de thèse (chapitre 4) présente la construction et la caractérisation d'une banque BAC (*Bacterial Artificial Chromosome*), à partir de la lignée introgressée Arabica, IAPAR 59. La construction de cette banque génomique est apparue indispensable tant pour l'étude du locus de résistance *Mex-1* que pour l'analyse de l'ensemble des gènes *R* chez le caféier.

- Enfin, le développement d'une carte physique du locus *Mex*-1 conférant la résistance à *M. exigua* a été entreprise. Les premiers résultats de construction des contigs associés à *Mex-1* font l'objet de la dernière partie de résultats (chapitre 5). La mise en évidence de séquences RGA au niveau de ces contigs est rapportée.

A l'issue de ces travaux, une synthèse des résultats obtenus (chapitre 6) est présentée et des perspectives de recherches sont exposées.

# CHAPITRE 2

# Diversité et évolution des gènes de résistance

de la classe NBS-LRR, au sein du génome du caféier

Molecular Genetics and Genomics (2001) 265: 654-662

# Origin, diversity and evolution of NBS-type diseaseresistance gene homologues in coffee trees (*Coffea* L.)

S. Noir, M.-C. Combes, F. Anthony, P. Lashermes

### Abstract

The majority of plant disease-resistance genes (R-genes) isolated so far encode a predicted nucleotide-binding site (NBS) domain. NBS domains related to R-genes show a highly conserved backbone of amino acid motifs offering the possibility of isolating resistance gene analogous sequences (RGAs) by polymerase chain reaction (PCR) with degenerate primers. Multiple combinations of primers with low degeneracy designed from two conserved motifs in the NBS regions of R-genes of various plants were used in coffee trees, an important perennial tropical crop. Nine distinct classes of NBS-like RGAs representing a large diversity were isolated from Coffea arabica and C. canephora species. The analysis of one coffee RGA family suggested point mutations as the primary source of diversity. With one exception, coffee RGA families appeared closely related by sequence to at least one cloned R-gene. In addition, deduced amino acid sequences of coffee RGAs were identified showing strong sequence similarity with almost all known non-TIR type R-genes. The high similarity between particular coffee RGAs and R-genes isolated from other angiosperm species such as Arabidopsis, tomato and rice indicated an ancestral relationship and the existence of common ancestors. As revealed in coffee trees, the evolution of NBS-encoding sequences seems to involve accumulation and slow divergence mechanisms within distinct R-gene families rather than a fast-evolving process. Functional inference of the suggested NBS domain type of evolution is also discussed.

### Introduction

In recent years, a growing number of disease-resistance genes (R-genes) conferring resistance to a diverse spectrum of pathogens have been isolated from a wide range of plant species (Richter and Ronald 2000). Sequence comparisons among these genes revealed remarkable similarities in general structure, and conservation of specific domains that participate in proteinprotein interactions and signal transduction (Staskawicz *et al.*, 1995). These R-genes have been grouped into several classes on the basis of their predicted protein products. The most prevalent class encodes a predicted nucleotide-binding site (NBS) attached to a C-terminal leucine-rich repeat (LRR) of variable length. So far, the only demonstrated role for NBS-LRRencoding genes in plants is in disease or pest resistance (Michelmore 2000). The NBS domain is a common protein element essential for the catalytic activity of various prokaryotic and eukaryotic proteins. This functional domain, which occurs in several related structural forms, is required for ATP- and GTP-binding (Saraste *et al.*, 1990). The primary sequence of the NBS is so distinct that protein sequences can be assigned to separate subgroups based on the conserved motifs found within the domain (Traut 1994). Phylogenetic analysis (Meyers *et al.*, 1999) indicated that plant NBS domains of R-genes can be categorized into two major types. One type, including *N*, *L*6, *RPP5*, *M* and *RPP1*, encodes proteins containing a Toll/Interleukin-1 receptor homology region (TIR) N-terminal to the NBS. On the other hand, the second type does not encode a TIR domain and commonly has a predicted leucine zipper motif in the N-terminal region (Pan *et al.*, 2000a). The non-TIR class includes proteins encoded by the resistance genes *RPS2* and *RPM1* from *Arabidopsis*, and *I2*, *Mi*, and *Prf* from tomato as well as *Dm3* from lettuce. While TIR and non-TIR sequences have been isolated from dicot species, TIR-type genes have not been detected in genomic or expressed sequence tag (EST) sequences from any grass species (Pan *et al.*, 2000a; Meyers *et al.*, 1999).

All plant NBS sequences of R-genes, both TIR and non-TIR, have a highly conserved backbone of amino acid motifs. Eight major conserved motifs have been identified in the NBS region, some of which are specific to the non-TIR class of predicted proteins (Meyers *et al.,* 1999). The overall sequence homology among members of the NBS-LRR genes is low, and not sufficient to be detected by cross-hybridization. However, the existence of conserved motifs offers opportunities for the design of degenerate primers and the isolation of disease-resistance gene analogous sequences (RGAs) by PCR from plant genomes. This approach was successfully applied to isolate NBS-LRR genes from several monocot and dicot species (Kanazin *et al.,* 1996; Leister *et al.,* 1996; Shen *et al.,* 1998; Yu *et al.,* 1998).

Many plant pathogens exhibit a high mutation rate that renders obsolete the effectiveness of individual R-genes. Therefore, the ability of plant species to survive over evolutionary time might depend on their ability to maintain and generate useful diversity at resistance loci (Hammond-Kosack and Jones 1997). R-genes are often members of tightly linked multigene families, which can be functionally diversified (Pryor 1987; Pryor and Ellis 1993; Martin *et al.*, 1994; Hulbert 1997). It has long been speculated that DNA rearrangements play a key role in evolution of the R-gene loci, thus allowing plants to generate novel resistance to match the changing pattern of pathogen virulence. The evolution of R-genes remains largely unexplored, but information has recently been gained from molecular genetic analysis (Michelmore and Meyers 1998; Meyers *et al.*, 1999; Pan *et al.*, 2000a). However, these analysis

45

concern plants with short life cycle while data from perennial plants have retained so far little attention.

Coffee trees belong to the genus *Coffea* in the family Rubiaceae. The subgenus *coffea* consists of approximately 100 taxa so far identified (Bridson and Verdcourt 1988). All species are woody and have a relatively long life-cycle since generation times have been estimated to be between 20 and 30 years (Berthaud 1996). *Coffea arabica* L. and *Coffea canephora* Pierre are the only cultivated species of economic importance. While *C. canephora* is a diploid species (2n=2x=22), *C. arabica* is an allotetraploid (2n=4x=44) involving two genomes that originated from two different diploid progenitors, *C. canephora* and *C. eugenioides* (Lashermes *et al.,* 1999).

In the present study, a number of RGAs were obtained from two coffee species, *C. arabica* and *C. canephora*, using degenerate primers based on conserved motifs of the NBS domain. The sequence characterization and diversity analysis of these RGAs are reported as well as their relationships with the NBS sequences of known R-genes from other plant species. In so doing, we sought to gain insights into the origin, diversification and evolution of NBS-LRR resistance genes in perennial species like coffee trees.

# Materials and Methods

#### Plant material

Plant material involved two accessions of *C. arabica* (i. e. Caturra, Et30) and four accessions of *C. canephora* (i. e. IF181, IF200, IF133, 02722). Genomic DNA was isolated from lyophilized leaves collected from a single plant through a nuclei isolation step as described by Agwanda *et al.* (1997).

#### **Primers and PCR conditions**

A large set of degenerate and non-degenerate primers was designed to amplify NBS-containing sequences from coffee trees (Table 1). Degenerate oligonucleotide primers were designed based on conserved motifs in the aligned amino acid sequences of known NBS-containing R-genes. These primers were designed to have low degeneracy, particularly at the 3' end. Six degenerate primers were designed to correspond to the P-loop motif (i. e. GGVGKTT) in the

sense direction while four degenerate primers were made corresponding to the GLPLAL motif in the anti-sense direction. On the basis of first coffee RGAs isolated from IF181 (*C. canephora*) with the Ploop1/GLPL1 primer combination (Fig. 2), two pairs of non-degenerate primers were designed corresponding to either the GVGKTT and DGLPLAL amino acid motifs (i. e. Ploop-Cof and GLPL-Cof primers) or the internal coffee-specific deduced amino acid motifs, GIKSWVC and AGEEVP (i. e. Int-Cof1 and Int-Cof2 primers).

PCR amplification was performed in a 50  $\mu$ l reaction volume containing the following reagents: 250 ng genomic DNA, 8  $\mu$ M of each degenerate primer or 1  $\mu$ M of each nondegenerate primer, 50  $\mu$ M dNTPs, 1X PCR buffer (Promega), 1.5 mM MgCl₂ (Promega) and 0.5 U *Taq* polymerase (Promega). An MJ research PTC-200 thermocycler was used for the PCR amplification. After denaturation of DNA template at 94°C for 5 min, PCR cycle conditions were 35 cycles each of denaturation at 94°C for 30 s, annealing at 42°C (with degenerate primers) or at 58°C (with no degenerate primers) for 45 s and elongation at 72°C for 1.5 min. The last round of elongation was for 6 min at 72°C. In the case of degenerate primers, the annealing temperature was selected after testing temperatures from 38 to 60°C. PCR products were separated on 1.5% agarose gel, and DNA fragments of the appropriate size were extracted.

#### **Cloning and sequencing of PCR products**

Cloning was performed with the TOPO TA Cloning® kit (Invitrogen) and plasmid DNA was purified with the WIZARD Plus SV Miniprep kit (Promega), according to the supplier's instructions. Each insert of the appropriate size was sequenced from both sides using universal primers T7 and SP6 (Genome-Express laboratories, France).

Nucleotide and amino acid sequences of coffee RGAs have been submitted to EMBL as accession numbers AJ298883, AJ298884, AJ298885, AJ298886, AJ298887, AJ298888, AJ298889, AJ298890, AJ298891, AJ298892, AJ298893, AJ298894, AJ298895, AJ298896.

#### Sequence analysis

Phylogenetic analyses were performed on the putative coffee RGA nucleotide sequences as well as on their deduced protein sequences. The mutations putatively introduced by the PCR amplifications were estimated by comparison of related sequences obtained from 8 distinct PCR amplifications. The rate of mutation-errors was estimated to 0.1% and appeared negligible in our analyses. Search for introns was done using the NetPlantGene program (Hebsgaard *et al.,* 

1996). Similarity searches were performed with the BLAST program (Altschul *et al.*, 1990) in the GenBank database. Pairwise comparisons and multiple alignment were performed using the ALIGN program (Myers and Miller 1988) and the ClustalW version 1.8 (Thompson *et al.*, 1994), respectively. A CLUSTALW DNA matrix and a BLOSUM protein matrix were used. Neighbor-Joining trees (Saitou and Nei 1987) were generated from sequence alignment using the TRECONW version 1.3b package (Van der Peer and De Wachter 1994). The bootstrap method (Felsenstein 1985) was employed to evaluate the reliability of tree branching. Kimura's correction (1980) was applied, and insertions and deletions were taken into account. Non-synonymous (K_A) and synonymous (K_S) substitutions were calculated using the SNAP method (Nei and Gojobori 1986). NBS sequences of the following R-genes were used in our analysis: *N* (U15605), *L6* (U27081), *M* (U73916), *Prf* (U65391), *RPM1* (X87851), *RPS2* (U12860), *RPS5* (AF074916), *RPP1* (AF098962), *RPP5* (U97106), *RPP8* (AF089710), *I2C-1* (AF004878), *Mi1-2* (AF039682), *RP1-D* (AF107294), *Pi-B* (AB013448), *Dm3* (AF113948), *Xa-1* (AB002266), *GPA2* (AF195939), *Bs2* (AF202179).

### Results

#### PCR amplification of RGAs

#### Amplification with degenerate primers

Twenty-four combinations of degenerate primers (Table 1) were tested on genomic DNA samples of two accessions, Caturra of *C. arabica* and IF181 of *C. canephora*. As expected considering the NBS domain size of known R-genes, major amplification products were in the 520-bp size range. However, PCR products of non-expected sizes (larger, > 600 bp, or shorter, < 500 bp) were also observed and cloned. When sequenced, these fragments revealed no characteristic motif of cloned R-genes. Some of them showed identity to retro-transposon-like elements as determined through BLAST searches of GenBank. Therefore, only DNA fragments of the appropriate size (480 to 600 bp) were considered further. A total of 120 clones were sequenced. In all cases, amplified bands contained a mixture of sequences. One hundred presented no resemblance to NBS domains of R-genes while 19 appeared related to the NBSs of known R-genes. These 19 RGA sequences presented an uninterrupted open reading frame. Their deduced amino acid sequences showed not only the two motifs targeted by the primers used but also the internal motifs characteristic of NBS-LRR gene class (Kin-2, and RNBS-B;

Meyers *et al.*, 1999). These sequences resulted from 8 different Ploop/GLPL primer combinations (Table 2). The number of distinct isolated NBS-like sequences per primer combination varied from 1 to 4. No putative splicing site was detected in these clones.

#### Amplification with non-degenerate primers

On the basis of one coffee RGA sequence (CrgaA1), two pairs of non-degenerate primers were defined (Table 1). The pair matching the P-loop and GLPL motifs was tested on genomic DNA samples of different accessions of both species, *C. arabica* (Caturra, ET30) and *C. canephora* (IF181, IF133, IF200, 02722). From all templates, an amplified band of the expected size was observed. Indeed the cloning and the sequencing of these PCR products allowed isolation of 20 NBS-RGAs of 524 to 527 bp sizes in which no putative splicing site was detected. The pair targeting the two motifs, GIKSWVC and AGEEVP, was tested on Caturra (*C. arabica*) and IF181 (*C. canephora*). Following amplification and cloning, 4 sequences of 399 to 403 bp long, corresponding to NBS-RGAs, were identified.

#### Analysis of coffee RGA

#### Diversity analysis of coffee RGAs

A diversity analysis of coffee NBS-RGAs, isolated from one accession of *C. arabica* (Caturra) and one accession of *C. canephora* (IF181), was performed. Twenty-seven nucleotide sequences obtained from either degenerate (19) or non-degenerate (8) Ploop/GLPL primer combinations were included in the analysis.

A Neighbor-Joining tree (Saitou and Nei 1987) is presented in figure 1. Nine coffee RGA families (A to I) were distinguished. A parsimony analysis produced a similar tree (data not shown) which confirmed the tree is robust. Three families consisted of only one RGA (i. e. B, C and D), while other families were composed of several members. Furthermore, families composed of several members are supported by very high bootstrap values (Fig. 1). Identities between RGA members of a given family range between 72 and 99.5%. Inter-family identities ranged from 50 to 60% depending on the families compared and the RGAs considered. The highest levels of inter-family identity are observed between members of the C, D and E families. The ratio of non-synonymous to synonymous substitution (K_A:K_S ratio; Nei and Gojobori 1986) of the nucleotides encoding amino acid sequences of coffee RGAs were calculated for each family. They ranged from 0.2 to 0.57.

The diversity of identified RGAs appeared to be related to the primer combination used (Table 2). For a given primer combination, sequences belonging to only either one or two

families were isolated. The non-degenerate primer pairs defined on the basis of the A family sequences allowed isolation of the RGAs of the same family. The RGA association did not appear to be related to the species of origin. Indeed, sequences of both species were found within 5 of these families (A, F, G, H, I). For example, Argal1 (*C. arabica*) and Crgal3 (*C. canephora*), in the I family, share up to 99% identity.

#### Nucleotide analysis of the A family

Considering all coffee RGAs isolated, a total of 23 RGAs, derived from 2 accessions of C. arabica and 4 accessions of C. canephora, were grouped in the A family. Based on Neighbor-Joining analysis, 2 subfamilies were distinguished. One subfamily (Aa) was composed of 2 RGAs isolated with degenerate primers, while the second subfamily (Ab) comprised 21 RGAs which were all isolated with non-degenerate primers, except for CrgaA3. A fine analysis was under taken from the majority consensus of each subfamily (Fig. 2). Mutational events observed among the A family corresponded to numerous nucleotide substitutions and one insertiondeletion mutation. RGAs of both subfamilies showed several autapomorphies, the total amount being related to the number of analyzed sequences. The Ab subfamily presented 56 autapomorphies while 12 were observed in the Aa subfamily. These mutations appeared evenly distributed along the sequences. Within the codons, these autapomorphies affected the third base more frequently (44% and 46%, for Aa and Ab families, respectively). Moreover, a very significant number (i. e. 79) of substitutions were shared (i. e. synapomorphies) by RGAs within each subfamily. In particular, 63 of those supported the distinction between the two subfamilies. In addition, the 2 subfamilies were distinguished by the presence of a short deletion (3 bases) within Ab subfamily RGAs. Furthermore, some mutations are shared between RGAs not belonging to the same subfamily. In particular, four of these mutations shared between Aa subfamily RGAs and the CrgaA18 RGA, belonging to the Ab subfamily, are co-localized in the sub-terminal position of the sequence alignment. CrgaA18 could derive therefore from an intersubfamilies recombination event.

#### **Comparison analysis**

Phylogenetic relationships between deduced amino acid coffee RGA sequences and known Rgene products were investigated. Almost all R-genes of the NBS-LRR class, listed in the GenBank database at the time of this study, were included in the analysis (*Mi1-2*, *I2C-1* and *Prf* from tomato, *Bs2* from pepper, *GPA2* from potato, *RPM1*, *RPS2*, *RPS5*, *RPP1* and *RPP8* from *Arabidopsis*, *RP1-D* from maize, *Dm3* from lettuce, *Pib* and *Xa1* from rice, *M* and *L6* from flax, and *N* from tobacco). The Neighbor-Joining phylogenetic tree constructed from the amino acid alignment of NBS domains of these R-genes and coffee RGAs (1 to 4 members per family) is shown in figure 3. The tree has long branch lengths and closely clustered nodes, reflecting a high level of sequence divergence. The distinction of the 9 coffee RGA families is clearly confirmed. However few divergences with the nucleotide analysis are observed in the relationships between coffee RGA families. All these families (except the A family) are associated with at least one of the NBS domains of cloned R-genes. According to the previously defined distinction between the TIR class and the non-TIR class, all isolated coffee RGAs seemed to belong to the non-TIR class type of R-genes. Moreover, all considered NBS domains of the non-TIR class (except *Dm3*) were associated with one coffee family.

In addition to this phylogenetic analysis, searches of GenBank using BLAST (BLASTN and BLASTP) algorithms confirmed that coffee RGA families are closely related to some NBS domains of cloned R-genes. In particular, the highest BLAST scores were observed between the coffee families, H and I, and the NBS domains of *I2C-1* and *Mi1-2* genes, respectively (Fig. 4). For instance, ArgaH3 and *I2C-1* share up to 48.6% identity and 68.3% similarity, while CrgaI3 and *Mi1-2*, showed 44.2% identity and 62.7% similarity. In comparison, coffee RGAs of different families share on average 35% identity and 55% similarity. The 2 respective alignments, H family RGAs with *I2C-1* NBS domain and I family RGAs with *Mi1-2* NBS domain, revealed that similarities were shared beyond conserved NBS motifs and supplementary specific consensus motifs were identified (Fig. 4).

#### Discussion

#### Degenerate primer approach in coffee

The results presented here show the great potential of applying heterologous PCR approaches to the cloning and study of R-gene in a perennial plant like coffee tree. Using numerous degenerated primers matching two conserved motifs among the NBS regions of R-genes of different species, we were able to amplify and to clone from coffee species nine distinct classes of NBS-like sequences representing a large diversity. As previously suggested (Shen *et al.,* 1998), the use of multiple combinations of primers with low degeneracy rather than few primers with greater degeneracy (Kanazin *et al.,* 1996; Leister *et al.,* 1996) could be more efficient for isolating highly different NBS sequences. The amplification of retro-transposon related

sequences, in addition to the NBS sequences observed in this study, is likely to result from the interference of the high copy number of such sequences in plant genome. For instance, it has been shown that retro-transposons constitute at least 50% of the maize genome (SanMiguel *et al.,* 1996).

The PCR-derived coffee NBS sequences were identified as RGAs or portions of Rgenes based on several features. These sequences contained uninterrupted open reading frames. Moreover, in preliminary transcription analysis based on RT-PCR experiments using primer combinations specific to the coffee RGA classes, cDNAs corresponding to 8 distinct NBS-like sequence classes were detected in coffee leaves of both species, *C. arabica* and *C. canephora*. In addition, all sequences contained the characteristic conserved motifs of NBS Rgenes (Meyers *et al.*, 1999). Most coffee RGAs were also closely related by sequence to at least one known R-gene and consequently some of them may encode resistance gene products of as yet unknown specificity.

#### Coffee RGA diversity versus R-genes diversity

The NBS-encoding RGA isolated from coffee trees showed considerable sequence variation. Moreover, 8 of the 9 coffee-RGA families identified appeared to be closely related to one or several known plant R-genes. It is noteworthy that no TIR-type NBS sequences were isolated from the two coffee species surveyed, and that coffee RGAs were identified showing strong sequence similarity with almost all known non-TIR-type R-genes. The only exception is *Dm3* from lettuce for which no equivalent coffee RGA has been isolated so far.

The failure to detect non-TIR encoding sequences in coffee suggests that this R-gene family is absent or has diverged beyond recognition in coffee trees. Similarly, TIR-type sequences are absent in cereal genomes (Meyers *et al.*, 1999; Pan *et al.*, 2000a). On the other hand, one TIR-type sequence has been identified in *Pinus radiata*, suggesting that the differentiation of TIR and non-TIR types pre-dates the divergence of angiosperms and gymnosperms (Meyers *et al.*, 1999; Pan *et al.*, 2000a). The very frequent sequence similarity between particular coffee RGAs and R-genes isolated from other angiosperm species such as *Arabidopsis*, tomato and rice suggests an ancestral relationship and the existence of common ancestors. This also supports an origin and primary diversification of NBS-LRR R-genes more ancient than the divergence of monocots and dicots which occurred 130-150 million years ago (Crane *et al.*, 1995). A possible scenario would be that modern NBS-LRR R-gene families derive from only a few ancestral genes arising from divergent gene duplication. Pan *et al.*, *and* and *et al.*, *and* and *et al.*, *and* and *et al.*, *and*, *and* 

(2000a) suggested the presence of ancestral loci in low-copy number and a differential gene family expansion between monocot and dicot species to explain the absence of non-TIR-type sequences in cereal species. Such a scenario can also be adduced to explain the possible absence of TIR-type sequences in coffee trees. However, in the case of coffee species, this event would have occurred after the monocot/dicot divergence.

#### Variation and evolution of NBS-encoding sequences in coffee trees

NBS sequences are abundant in plant genomes. For instance, the *Arabidopsis* genome is estimated to contain approximately 200 NBS-encoding genes (150 of the TIR type and 50 of the non-TIR type; Meyers *et al.*, 1999). Obviously, other plants may have significantly greater numbers of R-genes. Using a non-exhaustive degenerate primer approach, up to 18 different coffee RGAs were identified from a single accession of coffee trees suggesting the presence of a large number of R-genes in the coffee genome.

Molecular data are increasingly consistent with plants having R-genes arrayed in complex clusters. Several studies have shown that within a disease-resistance cluster, there are often genes that show high similarity to resistance genes but do not encode a functional product (Song et al., 1997; Ori et al., 1997). Other data revealed that clusters of R-genes may contain sequences related in function but not in sequence (Milligan et al., 1998; Dixon et al., 1996). Indeed, clustering of R-genes and homologous sequences is considered as an interesting gene organization for generating diversity and new specificities of resistance. Various genetic mechanisms, including point mutation, recombination, unequal crossing over, and (or) gene conversion, have been proposed to ensure the evolution of the R-gene (Michelmore and Meyers 1998; Noel et al., 1999). Coffee NBS sequences appeared to be grouped in distinct families of ancestral origin. While the RGAs belonging to the same family showed high identities, sequence variation between RGA families was notably important suggesting an independent evolution of these different families. Furthermore, C. canephora is one of the progenitor species of *C. arabica* (Lashermes *et al.*, 1999). The identification, in both analyzed coffee tree species, of closely related RGAs was therefore expected. Interestingly, the observation of some highly identical RGA sequences is consistent with the postulated recent speciation of C. arabica and would indicate a slow evolution of the coffee NBS RGAs.

The fine analysis of one coffee RGA family suggested that point mutations are the primary source of diversity. Nevertheless, sequence exchange can not be excluded as one putative recombination event was also detected. Otherwise, the estimated ratios of nonsynonymous to synonymous substitutions ( $K_A:K_S$ ) for all coffee RGA families are of the same order as those for NBS domains of known R-genes (i. e. 0.37 to 0.44; Michelmore and Meyers 1998). The NBS region therefore seems to be under purifying selection, which is consistent with its proposed but unproven effector function.

In conclusion, the evolution of NBS-encoding sequences in coffee trees appeared to involve accumulation and slow divergence mechanisms within distinct R-gene families rather than a fast-evolving process. Associated with concerted evolution, fast-evolving mechanisms such as unequal crossing-over and gene conversion have been adduced as a major way of diversification in several studies (Leister *et al.*, 1998; McDowell *et al.*, 1998). However, concerning coffee NBS sequences, this kind of evolution would be not preponderant. As suggested by Michelmore and Meyers (1998) and Stahl *et al.*, (1999), R-genes might not be evolving rapidly in order to keep pace with changes in the pathogen, but rather are evolving fairly slowly to provide resistance against pathogen populations that are heterogeneous in space and time populations. This could be particularly relevant in the case of perennial plants like coffee trees.

#### Functional inference from NBS domain evolution

Plant disease R-genes are assumed to encode proteins that both determine recognition of specific pathogen-derived avirulence proteins and initiate signal transduction pathways leading to complex defense responses (Grant and Mansfield 1999). Several studies have proposed that the NBS domain would ensure this signaling function. Remarkably, at least 9 distinct RGA families of ancestral origin have been maintained in the coffee tree genome. The concurrent maintenance of different families suggests an evolutionary inference. Availability of R-genes involving various NBS domains might result in fitness superiority.

Sequence analysis of different loss-of-function alleles in R-genes *RPM1* (Grant *et al.*, 1995), *RPS2* (Mindrinos *et al.*, 1994), and *Prf* (Salmeron *et al.*, 1996) demonstrated the functional requirements for conservation in the consensus motifs characterizing the NBS region of plant R-genes (Pan *et al.*, 2000a). Specific signature was revealed when comparing coffee RGAs and the different closely related R-genes isolated from monocot and/or dicot species distantly-related to coffee species from a phylogenetic point of view. Within families constituted by the NBS sequences of hypothesized common ancestral origin, consensus motifs appeared extended and short additional conserved blocks were observed. Since those genes or RGAs result from an ancient divergence, this conservation suggests a functional family-specificity. It

would be possible that some of the differences between the NBS sequences encoded by Rgenes influence the specificity of interactions either with a pathogen elicitor or with proteins downstream in the signal transduction cascade. For instance, studies of R-gene-mediated responses in *Arabidopsis* disease resistance/susceptibility mutants suggest that there may be an association between signaling pathways and structural features of the R-genes (Aarts *et al.*, 1998).

Further genetic and biochemical analyses are required to define better the structural and functional roles of particular motifs and of diversity within the NBS domain. Elucidating the evolution of R-genes is of particular importance for the understanding of how plants maintain and adapt their defenses to pathogens (Pan *et al.*, 2000a). Since the highly contrasted situations occurring in plants, investigations of various species, including tree-species plants, appear particularly necessary besides plant models such as *Arabidopsis* and rice. In so doing, coffee trees could constitute an attractive model for perennial plants.

# CHAPITRE 3

# Identification et cartographie génétique

d'un locus de résistance au nématode à galles,

Meloidogyne exigua, chez Coffea arabica

Plant Pathology (2003) 52: 97-103

# Identification of a major gene (*Mex*-1) from *Coffea canephora* conferring resistance to *Meloidogyne exigua* in *Coffea arabica*

Short title: Nematode resistance gene in coffee

S. Noir, F. Anthony, B. Bertrand, M.-C. Combes and P. Lashermes

### Abstract

Among the most damaging root-knot nematode species, Meloidogyne exigua is especially common in Latin America and constitutes a major agronomic constraint in all major coffeegrowing (Coffea arabica L.) areas. Growing nematode-resistant coffee trees represents the most promising option to control the pest. The present study aimed at determining the mode of inheritance of the *M. exigua* resistance transferred into *C. arabica* from a relative species *C.* canephora, and attempting the identification of associated molecular markers. Segregation data analysis of a F₂ progeny derived from a cross between the resistant introgression line T5296 and the susceptible accession Et6 showed that the resistance to *M. exigua* is controlled by a simply inherited major gene (designated the Mex-1 locus). The gall index distribution exhibited by the F₂ individuals suggested an incomplete dominant expression. Fourteen AFLP markers were found associated with the resistance to *M. exigua* and a localised genetic map of the chromosome segment carrying Mex-1 was constructed. Furthermore, the association of the identified AFLP markers with Mex-1 was confirmed by the analysis of a set genotypes involving 28 introgression Arabica lines either resistant or susceptible to *M. exigua* in field conditions. These results represent an important starting point to enhance backcross-breeding programmes and to perform an early selection of resistant seedlings.

### Introduction

Root-knot nematodes (*Meloidogyne* spp.) have become a major threat in all major coffeegrowing (*Coffea arabica* L.) areas throughout the world. So far, more than fifteen species of *Meloidogyne* have been reported as pathogens of coffee (Carneiro *et al.*, 1996; Campos *et al.*, 1990). However, yield losses due to root-knot nematode infections are very dependent on the region and species involved (Campos *et al.*, 1990). Among the most damaging species, *M. exigua* is especially common in Latin America, where it constitutes a major agronomic constraint. Nematicide treatments are only effective in the short term, and they are also expensive and environmentally hazardous (Campos *et al.*, 1990). Therefore, growing nematode-resistant coffee trees constitutes so far the most promising option to control the pest.

The cultivated coffee *C. arabica* L. (2n = 4x = 44) is the only polyploid species in the *Coffea* genus. Recent investigations established that *C. arabica* is an amphidiploid formed by natural hybridisation between two diploid species, *C. eugenioides* and *C. canephora* 

(Lashermes *et al.*, 1999). *C. arabica* is characterised by low genetic diversity (Anthony *et al.*, 2002). In contrast, a considerable variability was reported among diploid coffee species (Berthaud & Charrier 1988). Therefore, the transfer of desired characters, in particular disease resistance, from the diploid relative species into cultivars of *C. arabica*, has been a continuous priority in coffee breeding (Van der Vossen, 2001).

While the standard Arabica coffee varieties are highly susceptible to *M. exigua*, a high level of resistance was identified in several accessions of *C. canephora* (Curi *et al.*, 1970). Moreover, several lines derived from a spontaneous interspecific hybrid (i. e. Timor Hybrid) between *C. arabica* and *C. canephora* exhibited a resistance against *M. exigua* similar to that observed in *C. canephora* (Bertrand *et al.*, 1997; Gonçalves & Pereira, 1998; Bertrand *et al.*, 2001). Progenies of the Timor Hybrid have been distributed world wide, and constitute the main source of pest and disease resistances used so far in coffee breeding (Lashermes *et al.*, 2000b).

The objective of the study presented here was to determine the mode of inheritance of the M. exigua resistance introgressed into C. arabica and to attempt identification of associated molecular markers. The amplified fragment length polymorphism (AFLP) approach (Vos *et al.*, 1995) which allows simultaneous analysis of a large number of marker-loci throughout the genome, was investigated. The identification of markers linked to the resistance would represent an invaluable tool for the resistance genetic analysis. In addition, they would represent an essential step towards both marker-assisted selection and map-based gene cloning.

#### Materials and methods

#### **Plant material**

The plant material employed in the present study consisted of i) accessions of *C. arabica* (i. e. Catuaí, Caturra, Et6, T5296) and *C. canephora* (i. e. IF200), ii) 28 introgression Arabica lines in generation  $F_4$  to  $F_7$  derived from different hybrids between individuals of various Timor Hybrids progenies (832-1, 832-2, and 1343) and commercial Arabica cultivars, and iii) a  $F_2$  population of 93 individuals derived from a cross between the *M. exigua* resistant introgression line T5296 and the susceptible accession Et6. A single  $F_1$  tree was self-pollinated to produce the  $F_2$  individuals.

#### Chapitre 3 -

#### Nematode resistance evaluation

Resistance evaluations have been performed either on seedlings in greenhouse or on 10 years old trees in field conditions. Since the conditions of infestation are very dissimilar, two different ways of evaluation have been used.

A single egg-mass isolate of *M. exigua* that originated from a severely infested plot at the CICAFE experimental station (Barva de Heredia, Costa Rica) was used for nematode resistance evaluation under greenhouse conditions at CATIE (Costa Rica). The initial population was characterised by electrophoretic patterns for four enzyme systems (Hernández *et al.*, 1996). The nematode population was maintained and multiplied on tomato roots (*Lycopersicon esculentum* cv. Nainemor). Nematode resistance was evaluated as previously described (Bertrand *et al.*, 2001) for individual coffee seedlings cultivated in pots containing a sterilised mixture (2:1) of leaf mould and sand. The inoculum (i. e.  $1000 \pm 100 \text{ J2}$  juveniles per pot) was applied around the collar of each seedling using a pipette when the plants were about four months old (two pairs of leaves). Observations were carried out three months after inoculation. The root system was carefully washed free of soil and the seedlings classified according to a gall index (GI), corresponding to a scale proposed by Kinloch (1990), where GI = 0, 0 gall, GI = 1, 1-2 very small galls, GI = 2, 3-10 very small galls, GI = 3, 11-30 small and large galls, GI = 4, 31-100 small and large galls, and GI = 5, over 100 small and large galls.

Field assessment of cultivars and introgression Arabica lines was carried out on trees that were more than 10 years old. These trees were assessed on two occassions, in November 1996 and October 1999, in trials at the CICAFE experimental station (Barva de Heredia, Costa Rica), under conditions of severe infestation as previously reported (Bertrand *et al.*, 2001). Trees for which no galls were observed were scored as resistant (R), and those for which at least one gall was found were scored susceptible (S).

#### **AFLP** analysis

DNA was extracted from lyophilised leaves and the AFLP procedure was performed as described elsewhere (Lashermes *et al.*, 2000b). An aliquot of 500 ng genomic DNA was digested with restriction enzymes *Eco*RI and *Mse*I. Restriction fragments were ligated with double-strand *Eco*RI and *Mse*I-adapters. A preamplification was done using the appropriate primers combination with one added selective nucleotide: E+A or C combined with M+A or C. The code following E or M corresponds to the selective nucleotides at the 3'-end of the *Eco*RI and *Mse*I primers, respectively. The reaction mix was diluted 1/30, and 10 µI was used for the

final amplification with two primers each having three selective nucleotides. A total of 342 primer combinations derived from 16 *Eco*RI primers and 22 *Mse*I primers were employed in this study. Those 16 *Eco*RI primers included 11 E+Axx (x represents one of the four bases A, T, G and C) and 5 E+Cxx. Likewise, the 22 *Mse*I primers included 10 M+Axx and 12 M+Cxx.

#### Analysis of data

The segregation in the F₂ population of each marker associated with the *M. exigua* resistance was tested for goodness-of-fit to the expected Mendelian segregation ratio by chi-square analysis. Markers were compared with *M. exigua* response score (GI) by one-way analysis of variance (ANOVA). Linkage analysis and map construction were performed using the computer program Mapmaker version 3.0b (Lander *et al.*, 1987). The most-likely order of the markers was obtained using exhaustive multipoint analyses. The Kosambi function (Kosambi, 1944) was used for converting recombination frequencies into map distances or centiMorgan (cM) values.

# Results

#### Inheritance of *M. exigua* resistance

The resistance against *M. exigua* of the susceptible line Et6, the resistant line T5296, and 93  $F_2$  individuals derived from a cross between T5296 and Et6 was evaluated under artificial inoculation in greenhouse conditions. The responses of the two parental lines (Fig. 1a, b) were clearly differentiated using the 0-5 gall index (GI). Seedlings of Et6 showed a gall index of 3 or higher (the mean being 4.6) while the seedlings of T5296 exhibited a gall index of 2 or lower (the mean being 0.6). In contrast, the gall index exhibited by the  $F_2$  individuals (T5296 x Et6) varied from 0 to 5, with a maximum frequency of plants showing a GI of 2 (Fig. 1c). The observed phenotype segregation was compared with Mendelian expectations. The plants were considered as resistant when they showed a GI inferior to 3. Under this assumption, the  $F_2$  segregated 65 resistant and 28 susceptible plants, which is consistent with the presence of one dominant gene ( $\chi^2 = 1.29$ ; P= 0.25). Most of the resistant plants showed a GI higher than the mean value of the resistant parent T5296 suggesting an incomplete dominant expression. Though other more complex hypotheses could not be excluded, the most simple "one incompletely-dominant gene hypothesis" was further considered ; the symbol *Mex-*1 being used to designate the locus for resistance to *M. exigua*.

# Identification of molecular markers associated with the resistance to *M. exigua*

A total of 342 AFLP primer combinations (Table 1) were screened on DNA samples representing either 2 resistant (T5296 and T17925) versus 2 susceptible (T17928 and T18137) Timor hybrid-derived genotypes or 2 resistant versus 2 susceptible individuals from the  $F_2$  (T5296 x Et6) population (Fig. 2). On average, about 40 clear amplified fragments were detected per primer combination for a total of 13680 AFLP products. Out of the 564 polymorphic fragments identified, 30 appeared as potentially linked with either the resistance or the susceptibility and were further screened on an expanded set of DNA samples including 5 resistant and 5 susceptible  $F_2$  individuals. Finally, 14 AFLP markers were found associated with the resistance to *M. exigua* and, presumably, with the locus *Mex*-1. No markers associated with the susceptibility were identified.

#### Linkage analysis and construction of a localized linkage map

Further linkage analysis was performed using the  $F_2$  (T5296 x ET6) population. The 14 candidate makers (Table 2) were tested on the 93 individuals. The segregation ratios of all 14 AFLP markers in the  $F_2$  population fit the 3:1 inheritance mode expected for a dominant marker (data not presented). All makers appeared to be tightly linked to each other (LOD score > 10) and mapped within a 8.2 cM interval (Fig. 3). The one-way analysis of variance showed a highly significant linkage (P < 0.001) of all markers with *M. exigua* resistance as estimated by GI.

Mapping of the locus *Mex-1* was attempted based on the interpretation of the disease evaluation data in terms of resistance or susceptibility. Since any misinterpretations would decrease the accuracy of the map position, the 11 plants presenting a GI of 3 were not considered in the analysis. In addition, two  $F_2$  plants classified as "susceptible" (GI of 4 and 5, respectively) and showing all markers as well as two  $F_2$  plants classified as "resistant" (GI of 0) and exhibiting none markers, were considered as ambiguously evaluated and were discarded from the analysis. Considering the 78 remaining  $F_2$  individuals, the resistance trait was found to co-segregate perfectly with the AFLP marker Exi-11. Consequently, the location of locus *Mex-1* was deduced to be in the vicinity of Exi-11 (Fig. 3).

#### Consistency of Mex-1 in different Timor Hybrid-derived Arabica lines

To confirm the features of *Mex-*1, the identified AFLP markers were analysed in a set of genotypes involving 28 introgression Arabica lines derived from different progenies of the Timor Hybrid, and cultivars of both *C. arabica* and *C. canephora* (Table 3). All markers showed a highly significant association ( $\chi^2$  values with *P* < 0.001) with *M. exigua* resistance evaluated in field conditions. Three recombinant introgression lines were observed which confirmed the order of markers previously determined and the linkage of *Mex-*1 with Exi-11. Whatever the Timor Hybrid progenitor used, an association between the marker and *M. exigua* resistance could be detected. All markers, with the only exception of Exi-10 located in the distal position, were also detected in the *C. canephora* accession and not in the *C. arabica* cultivars confirming the introgressed origin from *C. canephora* of the resistance to *M. exigua*.

# Discussion

Segregation data analysis of the  $F_2$  progeny derived from the cross between the resistant introgression line T5296 and the susceptible accession Et6 has provided convincing evidence about the inheritance of resistance to *M. exigua*. Most of the variation in the  $F_2$  could be explained by a single-locus model. Furthermore, despite the ambiguities in the interpretation of the disease evaluation data, the gene mapping approach employed proved to be very efficient. The resistance to *M. exigua* in introgression Arabica lines appeared clearly controlled by the action of one (or several) gene(s) located at a single genomic region (designated the *Mex-1* locus). In addition, our results indicated that the resistance exhibited by various progenies of the Timor Hybrid derived introgression lines is due to an unique introgressed fragment carrying *Mex-1*. Furthermore, the molecular data confirmed the *C. canephora* origin of the resistance to *M. exigua* exhibited in the Timor Hybrid derivative material. The simple inheritance of the *M. exigua* resistance observed in the present study conformed with most genetic studies of plantroot knot nematode interactions (Trudgill and Blok, 2001). Nematode resistance conferred by simply inherited major genes have been identified in a large number of crop species including cereals, cowpea, tomato, potato, and soybeans (Young, 1998).

The gall index distribution exhibited by the  $F_2$  individuals provided supporting evidence for an incomplete dominant expression ; the resistance as estimated by the GI being expressed at a slightly lower level in the heterozygous condition. However, it is not possible to discount the existence of minor genes with modulating effects (Roberts *et al.*, 1998). Further analyses are

Chapitre 3 -

necessary to clarify this question. In particular, it would be interesting to vegetatively propagate the  $F_1$  hybrid and to evaluate its resistance behaviour.

As previously reported (Silvarolla *et al.*, 1998 ; Bertrand *et al.*, 2001), the *M. exigua* resistance screening procedure used in the present study was very effective. The amount of root galling appeared to be a reliable indicator of the difference of nematode reproduction among genotypes. Furthermore, the resistance phenotype associated with *Mex-*1 was displayed by the screening procedure as well as in field conditions. Nevertheless, a few  $F_2$  individuals showed an ambiguous classification. The observed resistant plants lacking the *Mex-*1 introgressed fragment are likely to result from infection escape. Most of these plants presented a low vigour and an underdeveloped root system. On the other hand, the two plants that appeared susceptible although showing the introgressed fragment carrying *Mex-*1 should be evaluated again. The deficiency in expression of the resistance is intriguing and could have various origins such as the physiological status of the plant as well as error in the inoculum dosage.

In this study, the AFLP technique has proven to be a powerful tool for identifying tightly linked markers. Out of the 14 identified markers, the marker Exi-11 is probably the most-closely-linked marker to the *Mex-1* locus. As anticipated (Bertrand *et al.*, 2001) the genetic divergence between the resistant and susceptible genotypes was very low explaining *a posteriori* the non-performance of a bulked segregant analysis (Michelmore *et al.*, 1991). The relatively high proportion of polymorphic fragments associated with *Mex-1* (i. e. 14 among 564) is likely to be due to the introgressed origin of the resistance. As a matter of fact, most of the genetic diversity revealed by molecular markers in such plant material was attributed to the introgression process (Lashermes *et al.*, 2000b). Furthermore, it is noticeable that despite the alien origin of the DNA fragment carrying *Mex-1*, the associated markers did not exhibited segregation distortion. This result conforms with data indicating that introgression of genes into *C. arabica* from *C. canephora* appears not limited by differences either in sequence homology or in the chromosomal structure (Herrera *et al.*, 2002).

Efficient use of the genetic resources available in wild relative species is essential to the continued improvement of coffee varieties (Van der Vossen, 2001). However, breeding programmes face considerable difficulties in doing so. In particular, strong limitations are due to the long generation time of coffee-trees, the high cost of field trials, and the lack of accuracy of the current strategy. The identification of molecular markers linked to *M. exigua* resistance represents an important starting point to enhance backcross-breeding programmes and to perform an early selection of resistant seedlings. *Mex*-1 is the first gene for nematode

resistance in coffee to be identified. However, several complementary sources of resistance against *Meloidogyne* spp. have been reported (Gonçalves *et al.*, 1996; Anzueto *et al.*, 2001). Thus, the *Mex-*1 markers may also be used in constructing coffee resistant genotypes by pyramiding the different sources of resistance. Efforts will be done to convert the identified AFLP markers into PCR-specific markers suitable for marker-assisted selection. These sequence-specific PCR-based markers would be also very valuable to proceeding with map-based gene cloning.

# CHAPITRE 4

# Construction d'une banque BAC de *C. arabica*
Chapitre rédigé sous la forme d'un article à soumettre à la revue Genome.

La construction de la banque BAC a été effectuée au sein de l'équipe de B. Chalhoub au laboratoire INRA-URGV à Evry.

#### Construction and characterisation of a BAC library of coffee tree (*Coffea arabica* L.)

#### Abstract

In order to promote genome research on coffee tree, one of the most important tropical crop, a bacterial artificial chromosome (BAC) library of the coffee allotetraploid species, *Coffea arabica*, has been constructed. The variety IAPAR 59 widely distributed in Latin America and exhibiting a fair level of resistance to several pathogens, was chosen. High-efficiency BAC cloning was achieved through the use of a two-step selection for the high-molecular-weight genomic DNA partially digested by *Hind*III. In total, the library contains 88 813 clones with an average insert size of 130 kb, and represents approximately 8 *C. arabica* dihaploid genome equivalents. Characterisation of the library showed that less than 4.5% of the clones consisted of organelle DNA. Furthermore, good genome coverage and representation of the library was established by hybridisation screening of high density-filters using a number of nuclear probes distributed across the genome. This Arabica BAC library, the first large-insert DNA library so far constructed in the genus *Coffea*, appeared well-suited for many applications in genome research, including physical mapping, map-based cloning as well as functional and comparative genomics.

#### Introduction

Coffee is one of the most important agricultural export commodity in the world. *Coffea arabica* (Arabica coffee) accounts for about 70% of the total coffee production. It is the only polyploid species (2n = 4x = 44) in the genus, while others species are diploid (2n = 22). Recent investigations established that *C. arabica* is an amphidiploid formed by hybridisation of two closely related diploid species, *C. eugenioides* and *C. canephora* (Lashermes *et al.,* 1999). In addition, the low divergence between the two diploid constitutive sub-genomes of *C. arabica*, namely C^a and E^a, and those of its modern-day progenitor species, C (*C. canephora*) and E (*C. eugenioides*) respectively, suggests *C. arabica* results from a very recent speciation (Lashermes *et al.,* 1999; Herrera *et al.,* 2002).

Availability of a large-insert genomic DNA library is crucial for physical mapping, mapbased gene cloning, and analysis of gene structure and function in most organisms including plant species. In recent years, bacterial artificial chromosome (BAC) vectors have emerged as the cloning system of choice (Choi and Wing, 2000). The easy handling and propagation of the clones, together with their relatively high stability and low degree of chimerism compared with yeast artificial chromosome (YAC) vectors, make BACs an invaluable tool (Osoegawa *et al.,* 2000). In plant, numerous BAC libraries have been constructed and successfully used for a variety of applications including positional cloning and genome sequencing (Shizuya *et al.,* 1992; Martin *et al.,* 1993; Sasaki and Burr, 2000). So far, no large-insert genomic DNA library has been constructed in coffee. Availability of a BAC library would therefore constitute an important breakthrough in coffee genome research.

In this paper, we report the construction and the characterisation of a 8 genomeequivalent BAC library from the Arabica cultivar IAPAR 59. We demonstrate that this BAC library is a valuable resource for genome analysis in *C. arabica* by carrying out hybridisation screens using mitochondrial and chloroplast DNA probes as well as nuclear probes distributed across the genome.

#### Materials and methods

#### **Plant material**

The coffee material, consisted of plants of the commercial variety IAPAR 59, was grown in greenhouse conditions until aged of 8 months. Before sampling plants were placed in dark for 4 days. Young leaves were collected and stored in liquid nitrogen until needed.

#### **BAC library construction**

Isolation of high molecular weight DNA, digestion and ligation were done according to protocols fully described in Chalhoub *et al.* (in preparation).

#### Isolation of high molecular weight DNA (HMW-DNA)

Nuclei were isolated from 25 g of leaf tissue as reported elsewhere (Peterson *et al.*, 2000; Chalhoub *et al.*, in preparation). One ml of isolated nuclei solution was embedded in 1 ml of 1.5% low melting point agarose (InCert® Agarose Bio-Rad) and poured into plug molds (Bio-Rad). Solidified plugs were incubated under low agitation at 45°C for 48 h in lysis buffer. While some of these plugs were stored at 4°C in this buffer for future use, 11 lysed plugs were washed in successive baths to destroy residual proteinase K and to remove high concentrations of EDTA from the plugs.

#### Chapitre 4 -

#### Hind/II-digestion of high molecular weight DNA

Partial digestion step was performed according to the strategy of Chalhoub *et al.* (In preparation). Before HMW-DNA restriction, the 11 plugs were equilibrated for 1 h at 4°C under agitation in the appropriate *Hind*III restriction buffer according to manufacturer's instructions (Promega). For partial digestion, a range of ten *Hind*III restriction enzyme concentration solutions was prepared. Each plug was sliced into fine pieces (sides about of 1 mm) and respectively transferred into one of the scale concentrated restriction solution. After incubation for 1 h on ice, all reactions were transferred at 37°C for a fixed time of 20 min. Digestion was stopped by adding 150 µl of 0.5 M EDTA pH 8 per tube. One plug was incubated without *Hind*III enzyme as a DNA degradation control.

#### Size selection of partially restricted HMW-DNA

Restriction reactions were collected together and loaded on a 1% agarose gel (Gold Seakem) in 0.25X TBE. Restriction fragments were resolved in a CHEF-DRIII apparatus (Bio-Rad) by pulsed field gel electrophoresis (PFGE) using the appropriated parameters. After electrophoresis, the sides of the gel containing marker lanes were stained with ethidium bromide and aligned with the unstained portion of the gel. DNA fragments corresponding to 100-150, 150-200, 200-250 and 250-350 kb were subjected to a second size selection on a 1% low melting agarose (Sea Plaque GTG) in 0.25X TBE by PFGE. The gel was stained and aligned as before, and slices of agarose containing DNA ranging between 100-150, 150-200, 200-250 and 250-350 kb were subjected. The DNA concentration was estimated by gel migration and comparison with Lambda DNA standards.

#### Ligation and construction of the BAC library

The pIndigo BAC vector was prepared for the ligation step. Ligation reactions were carried out at 16°C overnight using the four different size selected partially digested DNA samples. Ligated products issue from the selected ranges 100-150, 150-200, 200-250 and 250-350 kb were noted ligation 1, ligation 2, ligation 3 and ligation 4, respectively. After dialysis, 15 µl of each ligation mixture were transformed by electroporation in 110 µl of competent *Escherichia coli* DH10B cells (Gibco-BRL). Transformants were selected on LB-X-gal-IPTG plates containing 12.5 µg/ml chloramphenicol. White colonies were picked using a Genetix Q-Bot and stored in 384-well microtiter plates (Genetix) at -80°C.

#### Estimation of insert sizes of BAC clones

BAC clones were randomly selected from each ligation and grown for 48 h at 37°C in 1.5 ml 2YT medium containing chloramphenicol (12.5  $\mu$ g/ml). The BAC DNA was extracted using an alkaline lysis procedure (Sambrook *et al.,* 1989) and digested with *Not*I (Biolabs). Digested products were separated on a 1% agarose gel (Gold Seakem) in 0.5X TBE, in a CHEF-DRIII apparatus (Bio-Rad) by PFGE.

#### **BAC filter and Southern hybridisation**

High-density colony filters were prepared using the Genetix Q-Bot. BAC clones were gridded in double spots onto Hybond-N⁺ filters (Amersham-Pharmacia) using a 4 x 4 pattern as described by Tomkins *et al.* (1999). Clones from the ligation reaction of the expected 250-300 and 300-350 kb fragments were preferably patterned. In total, 27 648 clones (i. e. 44 plates of the ligation 4, 27 plates of the ligation 3 and 1 plate of the ligation 2) of the BAC library were spotted onto a 22.5 x 22.5 cm filter. The inoculated filters were incubated for 14 h at 37°C. Colony DNA transfer on the filter was performed according to standard procedure (G. Gyapay, personal communication).

The filters were hybridised with 11 low-copy number nuclear probes from *C. arabica* that were selected to represent the 11 different linkage groups of the genome of *C. canephora* (Lashermes *et al.,* 2001). In addition, *CoxIII* and *NAD3/rps12* mitochondrial clones from *Oenothera* (Schuster *et al.,* 1990) were used while screening for chloroplast DNA was performed using Arabica PCR-generated probes corresponding to the *trnD/trnT*, *trnK/mat*K and *trnQ/Rps16* fragments (Taberlet *et al.,* 1991; Johnson and Soltis, 1994).

Probes were labelled with [ 32 P]-dCTP according to the manufacturer's recommendations (Megaprime DNA labelling systems kit, Amersham) and added to the hybridisation solution. Hybridisation was carried out overnight at 60°C in 5X SSC, 0.5% SDS, 5X Denhardt's, 25 mM potassium phosphate pH 6.5, 0.1mg/ml haring sperm DNA and filters were washed at 60°C successively for 15 min in 2X SSC – 0.1% SDS, in 1X SSC – 0.1% SDS and in 0.5X SSC – 0.1% SDS.

#### Results

#### **Construction of the Arabica line BAC library**

Isolated megabase DNA from the *C. arabica* IAPAR 59 variety was partially digested and separated by PFGE. After two size selection steps, the library was generated from the 4 selected size DNA fractions (100-150, 150-200, 200-250 and 250-350 kb). Each size fraction, corresponding to four sublibraries, was ligated and selected transformed colonies were picked in 230 microtiter plates of 384 wells. Collected BAC clone number for each ligation is noticed in the Table 1, they represent in total 88 320 BAC clones.

#### Insert size estimation and organellar DNA contamination control

The estimation of mean insert size for each ligation was based on PFGE gel analysis of 120 randomly chosen clones digested with the *Not*l restriction enzyme. Examples of *Not*l digestions representing the four different sublibraries are illustrated in Figure 1. Of the 120 clones analysed, 5 (4%) had no inserts suggesting that 84787 (96%) clones of the library have DNA inserts. Insert size distribution of BAC clones derived from the ligation reactions 3 and 4 (i. e. sublibraries 3 and 4) is shown in Figure 2. The estimated mean insert sizes for each ligation reaction are indicated in Table 1, and appeared lower than anticipated based on the apparent gel-size-selection procedure. More than 89% had inserts larger than 100 kb. Taking into account the relative number of picked clones from the four ligation reactions, the average insert size of the library was estimated at 130 kb. Furthermore, it can be noted that only one insert band per clone was generally observed after *Not*l digestion of BAC DNA (Fig. 1).

The percentage of clones in the library containing organelle DNA was determined by colony hybridisation of BAC clone-spotted filters. Filters of 27 648 clones were screened with three probes derived from different region (*trnD/trn*T, *trnK/mat*K and the *trnQ/Rps*16) of the chloroplast DNA. Depending on the probe used, the number of positive clones ranged from 1160 to 1190, 1150 being detected by the three probes. Altogether in the fraction of the library screened (i. e. 27 648 clones), 1237 clones contained chloroplast DNA indicating that roughly 4.5% of the BAC in the library had inserts that originated from the chloroplast genome.

Screening of the same BAC filters with two mitochondrial DNA probes, *Cox* III and *NAD3/rps*12, detected 8 positive clones for each probe, 4 being in common. Thus, 12 different positive clones were observed suggesting that about 0.0004 % of the BAC clones are derived

from the mitochondrial geneome. In total, less than 4.5% of the BAC clones were derived from the organellar DNA.

#### **BAC library screening**

To test the distribution of BAC clones, RFLP markers anchored to the 11 Canephora linkage groups that putatively correspond to the 11 gametic chromosomes, were used to screen BAC filters (Table 2). These RFLP mapping probes identified one locus in *C. canephora* with the exception of the cR167 that revealed two loci. Furthermore, five of them were reported to reveal two loci in *C. arabica*, one on each of the two diploid sub-genomes (C^a and E^a) associated in *C. arabica* (Lashermes *et al.*, 2000). The remaining probes exhibited only one constant allele in *C. arabica*. In consequence, the number of loci in *C. arabica* remained ambiguous since this figure could reflect the presence of either one locus or two loci (one on each sub-genome) exhibiting the same allele (Lashermes *et al.*, 1999).

For each of the eleven probes, clones were identified by hybridisation screening of BAC filters containing 27 648 clones (Fig. 3). The number of hits per probe ranged from 2 to 12 (6.3 on average). Taking into account that the fraction of clones analysed represents approximately half (44%) of the overall BAC library size, the number of positive clones per probe in the whole library was estimated (Table 2). Depending on the probe considered, the number of positive clones varied between 5 and 27, the average being fourteen. Assuming that the probes analysed present two loci in *C. arabica* (one on each sub-genome), the mean number of copies per di-haploid genome equivalents would be as 7.

#### Discussion

An Arabica BAC library using the cultivar IAPAR 59 was successfully constructed. This introgressed variety, derived from the Timor Hybrid (Lashermes *et al.*, 2000), is widely distributed in Latin America. A number of such Arabica coffee introgression lines exhibiting a fair level of resistance to several pathogens have been developed (Bertrand *et al.*, 1997; Bertrand *et al.*, 2001). The cultivar IAPAR 59 was selected for resistance to leaf rust and root-knot nematodes (Sera, 2001) and it presents the introgressed DNA fragment carrying the *Mex-1* locus that confers the resistance to *M. exigua* (Noir *et al.*, 2003).

The construction of a BAC library depends on both a high transformation efficiency and high proportion of desired BAC insert sizes. For both aspects, the protocol adopted in the

present study appeared highly satisfactory. The current library consists of 88 320 clones stored in 384-well microtiter plates. Taking into consideration false-positive (i. e. 4 %) and organelle clones (4.5%), it can be considered that the library is composed of 80 813 BAC clones of Arabica nuclear DNA. The di-haploid genome size of *C. arabica* has been estimated at 1 300 Mb (Cros *et al.*, 1995). Thus, considering an average BAC insert size of 130 kb, the entire library corresponds to 10 506 Mb of cloned nuclear DNA and would represent approximately 8 *C. arabica* genome equivalents.

Assuming an average insert size (I) of 130 kb, a total (N) of 80 813 clones in the library and a genome size (GS) of 1 300 Mb, the probability (P) to have a particular nuclear sequence represented at least once in this library would be  $P = 1 - e^{N[ln(1-l/GS)]} = 0.9997$  (Clarke and Carbon, 1976). Furthermore, positive clones were identified by hybridisation screening with eleven distinct probes representing all linkage groups. The number of positive clones in each case was roughly consistent with the allotetraploid structure and the estimated eight-times genome coverage of the library. The slight difference between the deduced mean number of copies per genome equivalents (i. e. 7) and the expected number of copies (i. e. 8) would result from either a small under-representation of particular sequences or difficulties in interpreting the data associated with the allotetraploid structure (e. g. estimation of RFLP loci number). Altogether, these evaluations indicated that the Arabica genome is well covered and represented by the BAC library.

Although nuclear isolation and purification steps were performed, screening results showed that approximately 4.5% of the library sequences derived from the organellar DNA. Contamination with the chloroplast DNA appeared much more serious than with the mitochondrial genome suggesting that the coffee chloroplast DNA is easier to spin down along with nuclei than mitochondria. Unintentionally, some of these BACs could correspond to the entire chloroplast genome and could be particularly useful when studying the coffee chloroplast genome (Cros *et al.*, 1998). Alternative hypothesis to explain part of the chloroplast or mitochondrial DNA-containing BAC clones could be suggested. The exchange of DNA sequences between mitochondria, chloroplasts, and nucleus is not a rare event. For instance, large insertions of mitochondrial DNA have been observed in *Arabidopsis thaliana* (Stupar *et al.*, 2001). Studies have shown that nuclear genomes of *Beta vulgaris* and related Chenopodiaceae contain large amounts of chloroplast DNA (Ayliffe *et al.*, 1998). Similarly, it has been estimated that the nuclear genome of spinach contains 5-10 copies of the chloroplast genome (Scott and Timmis, 1984). Therefore, it remains open whether the organelle DNA sequences found in the

Arabica library originate all from chloroplasts and mitochondria or partially from integrated nuclear copies.

Beside an average BAC insert size of 130 kb, the constructed Arabica library contained only a low number clones with insert DNA smaller than 75 kb. In addition, a high frequency (i. e. 55%) of the BAC clones deriving the larger size DNA fraction (i. e. ligation reaction 3 and 4) contained inserts with a size superior to 200 kb. Two-step size selection procedure as performed in the present study have been reported to effectively reduce the number of small insert clones in the library (Woo *et al.*, 1994; Zhang *et al.*, 1996; Xu *et al.*, 2001). Nevertheless, the average insert size for each of the 4 size-selected DNA fractions of the Arabica library appears to be inferior to the range of molecular weights selected.

The analysis of BAC clones by *Not*I digestion followed by PFGE showed that the majority of Arabica DNA inserts were typically present as single *Not*I fragment inserts. Thus, the Arabica genome apparently contains few *Not*I sites, a feature commonly observed with the genomes of other dicot species (Danesh *et al.*, 1998; Nam *et al.*, 1999; Deng *et al.*, 2001) and contrary to the results obtained with monocot species (Wang *et al.*, 1995; Yang *et al.*, 1997).

In conclusion, the Arabica BAC library described here represents the first large-insert DNA library for the genus *Coffea*. It contains large inserts (130 kb on average), has sufficient genome coverage (8 X) and contains a low fraction of clones with mitochondrial or chloroplast DNA. These features indicate that the BAC library is suitable for many applications in genome research, including physical mapping, map-based cloning, and functional and comparative genomics.

### CHAPITRE 5

### Développement d'une carte physique du locus Mex-1

### conférant la résistance à M. exigua

Chapitre rédigé sous la forme d'un article à soumettre à la revue Theoretical and Applied Genetics

# Physical mapping of the *Mex*-1 region and identification of BAC clones containing putative resistance genes in coffee (*Coffea arabica*)

S. Noir, L. Diniz, M.-C. Combes, P. Lashermes

#### Abstract

The root-knot nematode, *Meloidogyne exigua*, can be a redoubtable parasite especially in coffee-growing (*Coffea arabica* L.) areas of Latin America. A number of Arabica coffee introgression lines exhibiting a fair level of resistance to *M. exigua* have been selected. Previous studies showed that the resistance to *M. exigua*, introgressed from *C. canephora*, is controlled by a simply inherited major gene designated the *Mex-1* locus. Using cloned AFLP markers spanning the *Mex-1* locus, the construction of a physical map of the *Mex-1* region was initiated. Eighteen BAC clones corresponding to molecular markers flanking *Mex-1* were assembled in 3 contig regions. As expected from the allotetraploid structure of *C. arabica*, homoeologous containing resistance gene analogs (RGAs) was investigated by both PCR amplification with degenerate primers and hybridisation with coffee RGA probes. Positive BAC clones were identified for several RGA families. The *Mex-1* region appeared therefore to form a complex resistant-genes cluster.

#### Introduction

Coffee is one of the most important agricultural export commodity in the world. *Coffea arabica* (Arabica coffee) accounts for about 70% of the total coffee production. It is the only polyploid species (2n = 4x = 44) in the genus, while others species are diploid (2n = 22). Recent investigations established that *C. arabica* is an allotetraploid derived from the combination of the diploid *C. eugenioides* and *C. canephora* genomes (Lashermes *et al.*, 1999).

*C. arabica* is characterised by low genetic diversity (Anthony *et al.*, 2002a) and the transfer of desired traits from diploid relative species has been a continuous priority in coffee breeding (Carvalho, 1988; Van der Vossen, 2001). In particular, progenies from a spontaneous interspecific hybrid (i.e. Timor Hybrid) between *C. arabica* and *C. canephora*, have been used intensively as the main source of resistance to pests and diseases including coffee leaf rust (*Hemileia vastatrix*), Coffee Berry Disease (CBD) caused by *Colletotrichum kahawae*, and resistance to root-knot nematode (*Meloidogyne exigua*). A number of Arabica coffee introgression lines exhibiting a fair level of resistance to several pathogens have been developed (Bertrand *et al.*, 1997; Bertrand *et al.*, 2001).

There are now more than 25 resistance genes (R) of known function cloned from multiple plant genomes and structurally analysed (Hulbert *et al.*, 2001). A major class of cloned plant R genes contains nucleotide-binding site (NBS) and leucine-rich repeat (LRR) domains (Hammond-Kosack and Jones, 1997). The well-conserved regions of the NBS domains have been used to design degenerate primers enabling the PCR-amplification of resistance gene analogs (RGAs). Many of the RGAs are associated with known gene loci conferring resistance to viruses, bacteria, fungi or nematodes (Kanazin *et al.*, 1996; Leister *et al.*, 1996, Yu *et al.*, 1996). Indeed, previous studies showed that R genes and RGAs tend to cluster in plant genomes (Meyers *et al.*, 1998).

To understand the organisation of *R* genes or RGA clusters, large DNA fragments like BAC inserts or BAC contigs are necessary. Plant BAC libraries have been used for a number of structural genomic applications. In *Arabidopsis* and rice, BAC libraries have and are being used to develop physical maps allowing examination and comparison of large genomic regions (Lysak *et al.*, 2001; Liu *et al.*, 2001). In addition, some BAC libraries have been developed to perform positional cloning of genes of interest such as disease resistance genes (e. g. Song *et al.*, 1995a).

Segregation data analysis of  $F_2$  progenies showed that the resistance to the root-knot nematode *M. exigua,* introgressed from *C. canephora,* is controlled by a simply inherited major gene designated the *Mex-1* locus (Noir *et al.,* 2003). Fourteen AFLP markers were found tightly associated with the resistance to *M. exigua* and a localised genetic map spanning the *Mex-1* locus was constructed. Furthermore, a BAC library of the *M. exigua* resistant introgressed variety IAPAR 59 has been recently established (Noir *et al.,* in preparation), offering the possibility to envisage the map-based cloning of *Mex-1*. This paper describes the first efforts towards the construction of a physical map of the *Mex-1* region. Several BAC contigs corresponding to molecular markers flanking *Mex-1* were assembled. In addition, the presence within the *Mex-1* region of BAC clones containing RGAs was investigated by means of both PCR amplification with degenerate primers and hybridisation with coffee RGA probes.

#### Materials and methods

#### Material

The BAC library constructed from the genomic DNA of the Arabica introgressed line IAPAR 59 was used (Noir *et al.*, in preparation). This variety, derived from the Timor Hybrid (Lashermes, *et al.* 2000), is widely distributed in Latin America. It was selected for resistance to leaf rust and root-knot nematodes (Sera, 2001) and it presents the introgressed DNA fragment carrying the *Mex-1* locus that confers the resistance to *M. exigua* (Noir *et al.*, 2003).

#### **BAC Library screening**

The identification of BAC clones corresponding to the AFLP molecular markers (Exi-2, Exi-3 and Exi-4) flanking the *Mex-1* resistance gene was performed by both hybridisation and PCR-based screening. The Exi-2, Exi-3 and Exi-4 marker bands were cloned and sequence-characterised amplified region (SCAR) markers were developed (Noir *et al.*, 2003). When used as probes, the Exi-2 and Exi-3 markers are non repeated-copy sequences in the *C. arabica* genome. In contrast, the Exi-4 probe revealed complex profile (i.e. repeated sequence) and was therefore considered as useless (data not shown).

High-density colony filters screened comprise of 27 648 clones of the IAPAR 59 BAC library. This fraction of clones analysed represents approximately 3.5 *C. arabica* genome equivalents. The labelling of Exi-2 and Exi-3 probes and filter hybridisation conditions were performed as described in Noir *et al.* (in preparation).

The screening of the IAPAR 59 BAC library carried out by PCR amplification was performed from 36 864 BAC clones stored in 96 plates of 384 wells. These BAC clones include the 27 648 colony-filter spotted clones and represent approximately 4.2 *C. arabica* genome equivalents. From each plate, one pool was prepared by mixing of the 384 clones. From cell culture of each pool, *E. coli* cells were collected by centrifugation and dissolved in 10 mM Tris, 10 mM EDTA pH 8, 5% glycerol. BAC DNA of the 96 pools was extracted by four repeats of heating until ebullition followed by centrifugation 5 min at 5000 rpm. Diluted supernatants containing BAC DNA were used as template DNA for screening. The 96 pools of BAC DNA were subjected to the PCR-based screening using the 3 SCAR markers designated Exi-2, Exi-3 and Exi-4 SCAR-primers, respectively.

#### **BAC clone characterisation**

Hit BAC clones were grown for 48 h at 37°C in 1.5 ml 2YT medium containing chloramphenicol (12.5  $\mu$ g/ml). The BAC DNA was extracted using an alkaline lysis procedure (Sambrook *et al.*, 1989). To check the size of each BAC clone, plasmid DNAs were digested by *Not*I (Biolabs) and separated on a 1% agarose gel (Gold Seakem) in 0.5X TBE, in a CHEF-DRIII apparatus (Bio-Rad) with following PFGE parameters: pulse ramping from 1 to 40 s, angle 120°, current 6 V/cm, run time 11 h at 12°C.

BAC clone plasmids were digested by *Hind*III (Promega) at 37°C for 5 h. Digested DNA fragments were loaded onto a 1% agarose gel (Sea Plaque) and electrophoresed at 100 V for 22 h in a refrigerated 1X TAE running buffer (12°C). After electrophoresis, gels were stained with SYBR® Green and visualised under an appropriate transilluminator. Recovered *Hind*III-digested DNA fragments of BAC clones were used to BAC contig assembling.

#### BAC-end sequencing and mapping

BAC-end of clones associated with Exi-2, Exi-3 or Exi-4 markers were sequenced from one or the two insert sides using universal primers T7 and SP6 (Genome-Express laboratories, France). Primer combinations specific of BAC-end sequences were designed. PCR amplification was performed in 50 µl reaction volume containing the following reagents: 25 ng plasmid DNA, 1 µM of each BAC-end primer, 50 µM dNTPs, 1X PCR buffer (Promega), 1.5 mM MgCl₂ (Promega) and 0.5U *Taq* polymerase (Promega). A MJ research PTC-200 thermocycler was used for the PCR amplification. After a denaturation step at 94°C for 5 min, PCR cycles conditions were 35 cycles each of denaturation at 94°C for 30 s, annealing at 58°C for 45 s and elongation at 72°C for 1 min. The last round of elongation was for 6 min at 72°C. PCR products were separated on 1.5% agarose gel.

Plasmid DNAs of BAC clones related to Exi-2, Exi-3 or Exi-4 markers were digested with *Hind*III and *Eco*RI, respectively, for Southern blotting. Digested fragments were separated on a 1% agarose gel. The separated DNA fragments were transferred to a Hybond-N⁺ nylon membrane. BAC-end primers detecting putative overlapping BAC clones were used to amplified PCR-products from corresponding BAC plasmids. PCR-products were purified, cloned and used as probes (BAC-end probes). Labelling and hybridisation were carried out as above for library screening.

#### Search of disease-resistance gene analogues

The research of resistance gene analogue (RGA) sequences in the *Mex-1* region was carried out by PCR-amplification and Southern blotting. Twelve degenerate primer combinations (Ploop 1, 4, 5 or 6 versus GLPL 1, 3, 4) defined in Noir *et al.* (2001) were used in the same conditions. These combinations were tested with plasmid DNAs of BAC clones positives with either the Exi-2, Exi-3 or Exi-4 markers.

Previously, nine distinct classes of RGAs of the NBS-like type, representing a highly diverse sample, were isolated from *C. arabica* and *C. canephora* species (Noir *et al.*, 2001). To facilitate identification of RGA-containing BAC, one representative clone was selected from each of the 9 coffee RGA families and used as probes on Southern blot. EMBL accession numbers of the 9 coffee RGA used as probes are: AJ298883, AJ298885, AJ298886, AJ298887, AJ298888, AJ298889, AJ298891, AJ298893, AJ298895.

#### Results

#### **BAC library screening**

The identification of BAC clones corresponding to molecular markers linked to the *Mex-1* resistance gene was performed for 3 different AFLP markers. The AFLP marker bands, Exi-2, Exi-3 and Exi-4, were cloned and the screening of the BAC library was carried out by both hybridisation and PCR-based methods. The Exi-2 and Exi-3 markers are located on one side of the *Mex-1* locus while the Exi-4 marker is situated on the other side. The distance between these markers and the *Mex-1* locus was estimated at approximately 2 cM (Noir *et al.*, in press).

High-density colony filters were screened with 2 single-copy probes, Exi-2 and Exi-3 probes (Fig. 1). Probes Exi-2 and Exi-3 detected on the filter 5 and 11 clones, respectively. Each clone was identified and isolated (Table 1). Their insert sizes ranged from 130 to 250 kb. The two probes did not detected common clone.

PCR-based screening on pooled BAC DNA with Exi-2, Exi-3 and Exi-4 SCAR primers identified 2, 6 and 2 positive pooled BAC DNA plates, respectively. In the case of Exi-2 and Exi-3 markers, no pools were detected by PCR amplification with no corresponding hybridised clones. In order to isolate BAC clones identified with the Exi-4 marker, the two positive pooled

BAC DNA plates were individualised and subjected to screening using the same Exi-4 SCAR primers until the two single positive BAC clone were identified (Table 1).

#### **Development of BAC contigs**

BAC clones identified by screening using the Exi-2, Exi-3 and Exi-4 markers were analysed by fingerprinting. The banding patterns of clones were compared pairwise (Fig. 2). Then, the clone DNAs were assayed by Southern hybridisation and PCR-amplification with SCAR (i. e. Exi-2 and Exi-3) and 9 BAC-end primer combinations (i.e. 3A-T7, 3B-T7, 3B-SP6, 3G-T7, 3H-T7, 3K-T7, 4A-T7, 4A-SP6 and 4B-T7). Positive BAC clones relative to Exi-2, Exi-3 and Exi-4 markers were assembled in 5 contigs.

A first contig appears composed of the two BAC clones positive with the Exi-4 marker. Their fingerprinting-patterns showed strong similarities and they presented the same RFLP profile with the 4B-T7 probe.

From the 5 BAC clones identified with the Exi-2 marker, two contigs designed Exi-2 I and Exi-2 II were distinguished. The contig Exi-2 I consisted in the BACs 2E and 2B, while the BACs 2A, 2C and 2D formed the contig Exi-2 II. The Exi-2 SCAR primers enabled only amplification of the expected product from BAC plasmids of contig Exi-2 I. On the contrary, the Exi-2 probe hybridised to the BAC clones of the two contigs but revealed two distinct RFLP profiles.

The Exi-3 positive BAC clones are arranged in two distinct contigs designed Exi-3 I and Exi-3 II (Fig. 3). In total, 9 PCR or restriction sites corresponding to 7 physical markers (i. e. sequences identified either by hybridation or PCR) were located on the Exi-3 region (Fig. 4). Based on the 6 physical markers shared by the two contigs, the colinearity of contigs was observed. There was no discernible difference in marker positions between the two Exi-3 contigs. Nevertheless, the two homoeologous contigs showed notable differences. The 3 probes studied revealed distinct RFLP profiles between the contigs, and 1 of the 6 examined primers combinations amplified exclusively from the BACs belonging the Exi-3 I contig.

The 3 chromosomic regions corresponding to the Exi-4 contig, the two Exi-2 contigs and the two Exi-3 contigs, respectively, revealed no overlapped relationships.

#### Identification of BACs containing RGAs

The presence within the *Mex-1* region of BAC clones containing RGAs has been investigated. The 18 BAC clone plasmids identified with the Exi-markers were screened by PCR-amplification with 12 degenerate primer combinations known to target RGA sequences. Three combinations allowed the amplification of products with an expected size (Table 2). Putative RGA PCRproducts amplified were cloned and sequenced. In addition, Southern blot of the positive clones was hybridised using as probe one coffee RGA of the 9 identified RGA families. Seven familyspecific RGA probes hybridised none of the BAC clones identified with the Exi-markers. In contrast, the RGA F probe hybridised with the BAC clones 2A and 2B (Fig. 5), and the RGA-B probe with the BAC clone 4B (Table 2).

By these two approaches, RGA sequences have been detected within the 3 chromosomic regions related to the Exi-2, Exi-3 and Exi-4 markers, respectively. The RGAs present in the *Mex-1* resistance region belong to at least 3 distinct RGA families.

#### Discussion

Many investigations have demonstrated that BAC libraries are well-suited for map-based cloning and physical mapping. The results presented here demonstrate the feasibility of such approaches in an allotetraploid species, *C. arabica*.

#### **Development of contigs**

The coffee BAC library was screened with 3 *Mex-1*-linked markers. From the Exi-2, Exi-3 and Exi-4 markers, 5, 11 and 2 positive clones respectively were identified. The number of positive clones in each case was roughly consistent with the estimated genome coverage of the fraction of BAC clones analysed (about 4 X) and the fact that both sub-genomes are hybridised in southern analysis. The eighteen positives clones associated to the target *Mex-1* region, have been isolated. Their insert sizes range from 130 to 250 kb. Fingerprinting of the positive clone plasmids allowed the construction of preliminary contigs related to the 3 regions, Exi-2, Exi-3 and Exi-4, respectively. By PCR and Southern analyses, no link between the 3 chromosomic regions were detected. In the objective of the physical mapping and the cloning of the *Mex-1* gene, ours efforts will be bent towards extending the 3 Exi-contig regions to each other.

#### Nature of the introgression carrying Mex-1

The Exi-2 and Exi-3 BAC clones, respectively, were discriminated in 2 homoeologous contigs. This is consistent with the allotetraploid structure of the *C. arabica* species. Each homoeologous contig would be carried by one of the 2 sub-genomes of *C. arabica*. Moreover, the distinction of only 2 homoeologous contigs suggests that the introgressed fragment carrying *Mex-1* is the

result of an homologous recombination between *C. canephora* genome and one of the two subgenomes of *C. arabica*. Furthermore, an important level of polymorphism was observed by RFLP analysis between the homoeologous contigs suggesting that the *C. canephora* fragment carrying the *Mex-1* gene was introgressed into the C^a sub-genome of *C. arabica*.

#### **BAC library screening**

There are two basic methods to screen a BAC library which are hybridisation and the PCR. The two procedures adopted here have proven to be very efficient and complementary.

In the case of the identification of BAC clones related to the Exi-4 marker, only a PCRscreening was possible. After intermediate pooling steps, 2 positive BAC clones were assembly in a unique contig. This approach has allowed the allele-specific identification of positive BAC clones. The homoeologous BAC clones have not been detected. Screening by PCR allows the highly specific isolation of clones even when the marker sequence is a repetitive nature or has dispersed homologous copies (e. g. gene families) (O'Sullivan *et al.*, 2001).

After analysis, the Exi-2 and Exi-3 clones detected by hybridisation are appeared distributed in two homoeologous chromosomic regions, respectively. Even if hybridisation screening allows the direct determination of detected clone co-ordinates, in standard stringency conditions, this approach detects the interest allele clone and its homologous sequences without discrimination. This characteristic is nevertheless a powerful way of isolating all members of a gene family (Wang *et al.*, 2001). Likewise, by manipulation of the stringency of hybridisation, this screening method allows the use of heterologous probes (O'Sullivan *et al.*, 2001).

The combination of these two screening methods seems particularly complementary for polyploid genomes ; notably by allowing the discrimination between homologous and homoeologous sequences.

#### The locus *Mex-1* located on a *R* gene-cluster

The investigation of NBS-like RGA sequences was performed on BAC clones of the *Mex*-1 region. BAC clones containing RGA sequences were detected in the 3 chromosomic regions Exi-2, Exi-3 and Exi-4. At least 3 distinct RGA families were identified among clones constitutive of the *Mex-1* resistance region. This suggests that the *Mex-1* region contain clustered disease resistance gene analogs, and that the *Mex-1* gene could be a NBS-type *R* gene. In plant genome, complex and heterogeneous cluster of resistance, presenting distinct RGA sequences,

have been recently reported (Richly *et al.*, 2002). Resistance gene clusters have been postulated to have arisen from recombination-duplication and rearrangements events, to produce novel specificity for different pathogens (Hulbert *et al.*, 2001). In this context, the presence of disease resistance genes in the *Mex-1* region might provide further insight into disease resistance gene organisation and evolution in the coffee genome.

Since resistance genes tend to be clustered, a BAC contig that spans one *R* gene may also span others (Martin *et al.*, 1994). This study is a first example of clustered resistance gene analogs in the coffee genome. Coupled with an approach of using conserved domain to search for putative resistance genes (Deng *et al.*, 2001; Yu *et al.*, 2000), BAC library provide a valuable tool for plant-resistance candidate gene isolation.

### CHAPITRE 6

### Discussion Générale

#### 1. SYNTHESE DES PRINCIPAUX RESULTATS

# 1. 1 ETUDE DE LA DIVERSITE ET DE L'EVOLUTION DES GENES DE RESISTANCE DE LA CLASSE NBS-LRR, AU SEIN DU GENOME DU CAFEIER

Dans le contexte de la sélection et de l'amélioration de la résistance aux ravageurs chez le caféier, nous nous sommes proposés d'analyser l'origine, la diversité et le mode d'évolution des gènes de résistance spécifique au sein du génome de cette plante cultivée.

Actuellement, la classe de gènes *R* la plus largement étudiée correspond aux gènes de type NBS-LRR. La conservation de motifs peptidiques au sein des domaines NBS de ces gènes offre, en effet, une opportunité remarquable pour isoler les membres de cette classe. Des expérimentations utilisant une gamme d'amorces dégénérées, ciblant 2 motifs conservés des domaines NBS de gènes *R* clonés, ont été effectuées à partir d'ADN génomique de caféiers. Les séquences analogues de gènes de résistance (RGA) de type NBS ainsi isolées, constituent les 40 premières séquences caractérisées relatives à des gènes de résistance identifiées au sein du génome des caféiers. La nature de ces séquences a été confirmée, par ailleurs, grâce à l'isolement et l'analyse d'une séquence complémentaire d'ADN pleine longueur de la classe NBS-LRR (Fig. 1.1). Cette séquence a été isolée d'une banque phagique d'ADNc de *C. arabica*, criblée à partir d'une sonde RGA de caféier.

Bien que l'approche utilisée pour isoler des RGA de type NBS du génome de caféier ne soit pas exhaustive (i. e. biais liés aux motifs peptidiques ciblés et à l'amplification par PCR), l'analyse des séquences de RGA de caféier a mis en évidence leur grande diversité. Leur classification a permis de distinguer 9 familles (A à I). Par ailleurs, une analyse transcriptionnelle préliminaire de ces familles de RGA a été effectuée. Les séquences ADNc correspondant à 8 des familles de RGA de caféiers ont été détectées, par RT-PCR, chez *C. arabica* et *C. canephora*, indiquant que des membres de ces familles sont transcrits et pourraient coder des protéines de gènes *R*.

La présence et le nombre de séquences RGA de type NBS ont été analysés au sein de différents génotypes de caféiers. Des séquences des 9 familles identifiées sont observées au sein d'un même individu. De plus, le nombre de membres au sein des 9 familles a été estimé

par RFLP. Globalement, les 9 familles de RGA de caféiers étudiées apparaissent faiblement répétées. La présence des différentes familles de RGA identifiées au sein d'un même génotype de caféier et leur niveau de répétition suggèrent la présence de centaines de gènes *R* de type NBS au sein du génome de caféier. Par comparaison, les avancées dans le séquençage des génomes d'*A. thaliana* écotype Col-0 et, d'*Oryza sativa* L. ssp. *japonica*, ont permis d'identifier respectivement, 166 et au moins 600 gènes présentant un domaine NBS (Goff *et al.*, 2002 ; Richly *et al.*, 2002).

Les analyses phylogénétiques effectuées à partir des séquences peptidiques déduites des RGA de caféiers et des domaine NBS de gènes *R* clonés ont montré que les séquences de RGA de caféiers sont étroitement proches de celles de domaines NBS de gènes *R*. La similarité entre certains RGA de caféiers et les gènes *R* isolés d'autres angiospermes est en accord avec l'origine ancestrale de ces séquences et l'existence d'ancêtres communs (Pan *et al.*, 2000; Cannon *et al.*, 2002). Par ailleurs, l'analyse des séquences de RGA de caféiers a montré une évolution lente et intra-famille des domaines NBS chez le caféier.

L'analyse de diversité de plus de 800 séquences de la classe NBS-LRR (TIR et non-TIR) parmi 30 genres et 9 familles de plantes, a mis en évidence la distinction de 6 clades principales, 5 composées de séquences non-TIR-NBS-LRR et une seule regroupant les séquences TIR-NBS-LRR plus homogènes (Pan *et al.*, 2000 ; Cannon *et al.*, 2002). Chez *A. thaliana*, 40 des 166 gènes NBS-LRR identifiés correspondent au groupe non-TIR-NBS-LRR et 108 au groupe TIR-NBS-LRR (Richly *et al.*, 2002). Chez le caféier, tous les RGA isolés appartiennent à la classe non-TIR de gènes *R*. Toutefois, ils présentent une importante diversité : les principaux domaines NBS de type non-TIR illustrant la variabilité des domaines NBS de gènes *R* connus, sont associés à l'une des familles de RGA de caféiers isolés.

La fonction des domaines NBS serait d'initier l'activation de voies signalétiques aboutissant à la mise en place des réponses de défense de la plante (Baker *et al.*, 1997). Le maintien simultané de ces différentes formes de domaines NBS issus de divergence ancienne suggère une spécificité de fonction, relative à chaque famille conservée. Les caractéristiques structurelles des familles de RGA impliqueraient la spécificité d'une interaction, soit avec un agent éliciteur, soit plus probablement avec des produits de signalisation intervenant en aval des produits R.

# 1. 2 IDENTIFICATION ET CARTOGRAPHIE GENETIQUE D'UN LOCUS DE RESISTANCE A MELOIDOGYNE EXIGUA, CHEZ COFFEA ARABICA

Parallèlement à l'étude globale des gènes *R* de la classe NBS-LRR et de leur RGA, nous nous sommes intéressés au développement des connaissances relatives à un locus de résistance chez le caféier. Dans cette optique, l'étude d'un modèle d'interaction spécifique, tel que le pathosystème, caféier *C. arabica*/nématode *M. exigua*, est apparu incontournable. Le nématode à galles, *M. exigua*, est responsable de pertes économiques considérables en caféiculture, particulièrement dans les pays d'Amérique Latine. Dans le contexte du parasitisme des caféiers par les nématodes, l'interaction *M. exigua*/caféier est actuellement le seul modèle ayant fait l'objet d'analyses histologiques. De plus, des études génétiques préliminaires relatives à la résistance à *M. exigua* chez les caféiers suggèreraient un déterminisme oligogénique de ce caractère (Bertrand *et al.*, 2001).

Afin de préciser le déterminisme génétique de la résistance à *M. exigua* chez les caféiers, des marqueurs AFLP associés à cette résistance ont été recherchés par analyse de lignées Arabica introgressées résistantes (R) ou sensibles (S) et, analyse d'individus F₂ en ségrégation pour la résistance testée (F₂ [T5296 x Et6]). Cette analyse a montré que la résistance observée au sein de la population en ségrégation est conférée par un gène majeur, appelé *Mex-*1.

Une cartographie génétique de ce locus a été entreprise. Quatorze marqueurs AFLP, étroitement liés au locus *Mex-1* ont été identifiés. L'ensemble de ces marqueurs se distribue sur un fragment d'une taille d'environ 8,2 cM. Aucune distorsion de ségrégation n'a été observée. Quelques marqueurs SCAR ont été développés à partir de plusieurs marqueurs AFLP. L'association de ces marqueurs à la résistance à *M. exigua* a été confirmée, d'une part, à partir d'un lot de 28 lignées d'introgression Arabica (12 S et 16 R), et d'autre part, par le criblage de nouvelles populations d'individus F₂ en ségrégation pour la résistance à *M. exigua* (Tableau 1. 1). Au total, plus de 700 individus ont été génotypés et évalués pour cette résistance. Néanmoins, lors du criblage de ces nouvelles populations F₂, aucun nouvel événement de recombinaison n'a été détecté entre les marqueurs SCAR flanquant le gène *Mex-1*.

Dès lors, dans le contexte de la sélection de variétés de caféiers résistantes à *M. exigua*, les marqueurs SCAR encadrant *Mex-1* offrent la possibilité d'une sélection assistée par marqueurs.

Ces marqueurs constituent également, les données de base aux analyses de cartographie physique visant à cloner le gène *Mex-1* et/ou à étudier l'introgression portant ce gène.

# 1. 3. Developpement d'une carte physique du locus Mex-1 conferant la resistance a M. exigua

# 1.3.1 Construction et caractérisation d'une banque BAC, à partir du génome *C. arabica.*

Afin de permettre la cartographie physique du gène *Mex-1* et de poursuivre l'analyse des gènes *R* sur l'ensemble du génome de *C. arabica*, une banque génomique de grands fragments a été construite et caractérisée.

Le génome retenu pour la construction de cette banque est celui du cultivar IAPAR 59, lignée introgressée de l'espèce *C. arabica*. Cette lignée a été sélectionnée pour ses résistances à la rouille et aux nématodes à galles, notamment à *M. exigua* (Sera, 2001). La lignée Arabica, IAPAR 59 présente le fragment d'introgression issu de *C. canephora*, portant le gène *Mex-1*. Cette banque constitue la première banque BAC issue d'un génome du genre *Coffea*.

L'estimation des clones faux-positifs (4%) et des contaminations (4.5%) chloroplastique et mitochondriale a été effectuée. La banque est composée de 88 813 clones d'ADN nucléaire de *C. arabica*. La taille moyenne des inserts des clones BAC a été estimée à 130 kb. Considérant la taille du génome dihaploïde de *C. arabica* évaluée à 1 300 Mb (Cros *et al.*, 1995), l'ensemble de la banque BAC est constitué de 10 506 Mb d'ADN rucléaire, ce qui représente approximativement, 8 équivalents-génome de *C. arabica*. Par ailleurs, le criblage de la banque à partir de sondes caractérisant chacun des groupes de liaison du génome, a révélé une bonne représentativité de l'ensemble du génome du génome de *C. arabica*.

Les caractéristiques mises en évidence montrent que cette banque BAC constitue un outil de qualité pour des analyses de cartographie physique de gènes d'intérêt et l'étude des sousgénomes constitutifs de *C. arabica*.

# 1. 3. 2 Carte physique de la région portant le locus Mex-1 et analyses préliminaires

L'étude de l'organisation et de la composition de la région portant le locus *Mex-1* permettrait d'envisager les mécanismes évolutifs à l'origine de ce locus de résistance. Ainsi, afin de développer une carte physique de cette région, trois marqueurs génétiques étroitement associés à *Mex-1* et encadrant ce gène ont été utilisés. Ces marqueurs ont permis le criblage de la banque BAC de IAPAR 59, selon 2 approches complémentaires, l'hybridation de filtres BAC et l'amplifiaction par PCR.

Dix-huit clones BAC correspondant à des marqueurs (Exi-2, Exi-3 ou Exi-4) encadrant le locus *Mex-1* ont été détectés. L'analyse par *fingerprint* de ces clones BAC positifs a permis la construction de 5 contigs. L'un de ces contigs est composé de 2 clones BAC identifiés par le marqueur Exi-4, 2 contigs sont composés de 5 clones BAC identifiés par le marqueur Exi-2 et, 2 contigs sont composés de 11 clones BAC identifiés par le marqueur Exi-3. Les 3 régions relatives à chacun des marqueurs Exi-2, Exi-3 et Exi-4 respectivement, ne sont pas contigues. En outre, il a été montré que les 2 contigs composant respectivement les régions Exi-2 et Exi-3, sont des contigs homéologues (i. e. chacun de ces contigs est porté par l'un des sous-génomes constitutifs de *C. arabica*).

Par ailleurs, la recherche de séquences RGA au sein des clones BAC recouvrant la région *Mex-1*, a été effectuée. A partir des 3 régions distinguées Exi-2, Exi-3 et Exi-4, des clones BAC présentant des séquences RGA ont été identifiés. Ainsi la région portant le locus *Mex-1* semble correspondre à un cluster de gènes de résistance. De plus, au moins 3 familles distinctes de RGA de type NBS ont été détectées. La co-localisation de ces 3 familles distinctes de RGA au sein des contigs relatifs à *Mex-1* souligne la complexité et l'hétérogénéité de cette région (Richly *et al.*, 2002).

#### 2. PERSPECTIVES

Etant donnée la dimension économique que représente la culture du café et l'importance des pertes de production liées à la pression phytosanitaire exercée sur les caféiers, les connaissances relatives à la résistance aux parasites chez cette plante sont et resteront une priorité des programmes de recherche. Ces programmes bénéficient d'une part, de la collecte de ressources génétiques considérables, et d'autre part, de l'importante création de matériel végétal depuis la seconde moitié du XX^e siècle. De plus, des outils, biotechnologiques (e. g. micropropagation, embryogenèse somatique et transformation génétique *in vitro*) et génomiques (cartes génétiques, banques BAC, séquençage en masse d'EST), sont déjà largement maîtrisés. La combinaison de ces facteurs permet d'envisager de nombreux travaux de recherche, et font du caféier un modèle remarquable pour l'étude des interactions hôte/parasite chez une plante pérenne.

#### 2. 1 VERS UNE CONNAISSANCE PLUS APPROFONDIE DE L'ORGANISATION ET DE L'EVOLUTION DES GENES DE RESISTANCE AU SEIN DU GENOME DES CAFEIERS

#### 2.1.1 Distribution et localisation des RGA au sein du génome des caféiers

Différentes études de locus de résistance ont souligné l'agencement en clusters des gènes R et d'analogues de gènes R (RGA), au sein des génomes de plantes (Kanazin *et al.*, 1996 ; Leister *et al.*, 1996 ; Yu *et al.*, 1996). Le nombre et la distribution des séquences RGA et/ou des clusters constitués de RGA pourront être déterminés au sein du génome des caféiers. La localisation de ces séquences constituera une indication sur la localisation d'un grand nombre de gènes de résistance au sein du génome des caféiers. Ces données, associées aux programmes en cours d'analyse de la résistance des caféiers aux nématodes, à la rouille ou encore à l'anthracnose des baies, constitueront des bases solides pour faciliter la localisation et, éventuellement, le clonage des gènes R impliqués.

La cartographie physique de clusters de résistance nécessite l'analyse de larges fragments d'ADN tels que les inserts de type YAC et BAC. Le criblage de banques génomiques par des sondes spécifiques type RGA a permis d'identifier et d'organiser en contigs des clones présentant des séquences caractéristiques de gènes *R*. Chez *Arabidopsis*, le criblage d'une banque YAC a permis de mettre en évidence 21 séquences présentant les caractéristiques de RGA dont 12 sont associées à des locus *R* identifiés (Speulman *et al.*, 1998). Au sein des génomes de *Citrus* et de *Hordeum vulgare* L., 25 et 20 contigs respectivement, constitués de clones BAC associés à des séquences type NBS-LRR ont été identifiés (Deng *et al.*, 2001 ; Yu *et al.*, 2000).

Chez le caféier, la construction de contigs présentant des séquences RGA a été entreprise en utilisant les familles de RGA de type NBS identifiées et la banque BAC de *C. arabica.* Ces expérimentations pourront être complétées et étendues à d'autres familles de RGA de type NBS et à d'autres séquences caractéristiques de gènes R (e. g. domaines kinases, domaines trans-membranaires). Le nombre et la composition (i. e. en RGA de familles distinctes) des différents clusters de gènes de résistance (ou "clusters de résistance") au sein du génome de caféier seront ainsi déterminés.

La localisation des contigs portant des séquences RGA, au sein du génome de caféier, est envisagée à partir d'une approche génétique et d'une approche cytologique. La cartographie génétique des différents clusters de résistance permettra de mettre en évidence l'organisation génomique de ces gènes. Il n'existe pas actuellement de carte génétique de *C. arabica*. Par contre, la carte génétique de *C. canephora* qui distingue 11 groupes de liaison correspondant potentiellement aux 11 chromosomes du génome de *C. canephora* (Lashermes *et al.*, 2001), permettra la cartographie des différents clusters de résistance au sein du génome de *C. canephora*. La distribution et le regroupement des gènes de résistance au sein du génome du caféier permettront de souligner l'implication particulière de groupes de liaison dans des mécanismes de résistance (Fourmann *et al.*, 2001). En complément de cette approche, une approche cytologique (e. g. BAC-FISH) permettra de localiser physiquement les dusters de résistance identifiés au sein du génome de *C. canephora*.

#### 2.1.2 Relation avec la structure polyploïde du génome

Dans plus de 70% des cas, l'histoire évolutive des plantes à fleurs (sauvages ou cultivées) comporte des stades de polyploïdie. La large distribution de ce caractère suggère un avantage adaptatif de ces génomes complexes (pour revue, Wendel, 2000). Notamment, la duplication de gènes peut engendrer de nouvelles fonctions génétiques. Toutefois, l'analyse d'espèces allotétraploïdes spontanées et la création de plantes allotétraploïdes synthétiques montrent que

ces génomes, au moins lors de leur construction, sont plus ou moins sujets à des phénomènes de restructuration et d'instabilité génétique (e. g. recombinaison homéologue, activation de transposons, extinction de gènes, variation des niveaux de méthylation) (Song *et al.*, 1995b; Feldman *et al.*, 1997).

L'allotétraploïdie de *C. arabica* a été analysée à l'échelle génomique en termes de comportement méiotique (Lashermes *et al.*, 2000a) et de niveaux de recombinaison entre les génomes constitutifs de *C. arabica* (E^a et C^a) (Herrera *et al.*, 2002). Les clusters de résistance identifiés pourront être employés comme des marqueurs chromosomiques du génome des caféiers. Des BAC des contigs associés aux gènes *R* seront utilisés comme sondes cytologiques. La comparaison des localisations de ces marqueurs, au niveau des génomes de *C. arabica* et des espèces diploïdes parentales, *C. canephora* et *C. eugenioides*, permettra de vérifier la synténie entre les génomes des différentes espèces, et éventuellement, de mettre en évidence des remaniements chromosomiques associés à la spéciation de *C. arabica*.

# 2. 2 VERS UNE CONNAISSANCE PLUS APPROFONDIE DU CLUSTER DE RESISTANCE COMPRENANT LE LOCUS *MEX-1*

En complément de l'approche génomique qui aura permis de déterminer la distribution des clusters de résistance au sein du génome du caféier, il serait particulièrement intéressant d'analyser finement un cluster de résistance. Nous nous proposons ainsi, d'étudier l'organisation et l'évolution du cluster de résistance comprenant le gène *Mex-1*, chez le caféier. La mise en évidence des mécanismes moléculaires à l'origine de la structure de ce cluster apportera de nouvelles données quant à la dynamique évolutive des gènes *R* chez le caféier.

De nombreux clusters de résistance, constitués notamment de RGA de la classe NBS-LRR, ont été décrits chez différentes espèces végétales. L'analyse et la comparaison de ces clusters montrent d'importantes variations au niveau du nombre et de la distribution des RGA qu'ils présentent. Par exemple, le cluster *Dm3* chez la laitue, est composé de la famille de RGA, *RGC2*. Chacun des membres de cette famille sont distants d'au moins 120 kb (Meyers *et al.*, 1998). Par contre, jusqu'à 11 RGA, membres de 3 familles distinctes de la classe NBS-LRR, ont été identifiés au sein des 240 kb constituant la région *Mla* chez l'orge (Wei *et al.*, 1999). La diversité des structures des clusters de résistance est généralement le résultat des

mécanismes de pression de sélection exercés au niveau de ces séquences (Hulbert *et al.*, 2001).

La localisation de plusieurs RGA de familles distinctes au sein du contig présentant *Mex-1* souligne la complexité et l'hétérogénéité de ce cluster. L'étude de l'organisation (e. g. séquences en tandem) et de la diversité de ces séquences (e. g. degrés d'homologies) permettra d'envisager les mécanismes évolutifs à l'origine de ce cluster de résistance (Richly *et al.*, 2002).

La structure tétraploïde du génome de *C. arabica* ajoute un niveau de complexité à l'organisation et l'évolution des gènes au sein de cette espèce. Peu de travaux proposent le suivi évolutif de séquences dupliquées après polyploïdisation d'une espèce. Ces études ont, toutefois, confirmé l'existence de deux modes d'évolution possibles, d'une part, l'évolution concertée, et de ce fait, l'homogénéisation des locus homéologues considérés (Wendel *et al.*, 1995), et d'autre part, l'évolution indépendante des gènes dupliqués après polyploïdie (Cronn *et al.*, 1999 ; Feuillet *et al.*, 2001).

La comparaison des contigs homéologues associés à la région chromosomique portant *Mex-1* au sein de la variété IAPAR 59, permettrait la confirmation et l'affinement des mécanismes évolutifs liés à la formation de ce cluster de résistance.

Par ailleurs, il est envisagé d'analyser l'évolution de certains gènes composant le cluster de résistance *Mex-1* depuis la formation de *C. arabica.* Il s'agira alors d'isoler l'ensemble des RGA membres de différentes familles, associés à *Mex-1*, à partir du génome allotétraploïde de *C. arabica* et des génomes diploïdes proches des espèces parentales de *C. arabica, C. canephora* et *C. eugenioides.* Les clones BAC comprenant les RGA du cluster étudié permettront de repérer génétiquement les régions flanquantes (non transcrites) de ces séquences. L'isolement et l'analyse phylogénétique des RGA de cette famille (Feuillet *et al.*, 2001). Le mode évolutif des séquences homéologues au sein de *C. arabica* (i. e. évolution concertée ou indépendante) et la perte éventuelle de certains membres de cette famille de RGA au sein des espèces de caféier analysées pourront être discutés.

96

# 2. 3 VERS UNE CONNAISSANCE PLUS APPROFONDIE DE LA STRUCTURE ET DE LA FONCTION DU GENE MEX-1

A terme, la caractérisation du gène *Mex-1* constitue une étape incontournable dans la volonté de mieux comprendre les mécanismes de résistance associés à ce gène. Ces données constitueront également une ouverture vers le mode de fonctionnement et l'identification de nouvelles résistances chez les caféiers.

#### 2. 3. 1 Isolement du gène Mex-1 et validation fonctionnelle

La cartographie génétique du locus *Mex-1* permet le positionnement du (ou des) gène(s) correspondant(s) dans un intervalle de moins de 4 cM. L'état d'avancée de la carte physique de ce même locus ne permet pas pour l'instant de donner une estimation de la distance physique entre le locus *Mex-1* et les marqueurs génétiques les plus proches de ce locus.

#### Détermination de la distance physique du locus Mex-1

Les stratégies de clonage positionnel de gènes nécessitent la mise en évidence de marqueurs génétiques physiquement proches du gène d'intérêt (Tanksley *et al.*, 1995 ; Patocchi *et al.*, 1999). Or, il a été montré chez différentes espèces végétales que le niveau de recombinaison génétique est loin d'être homogène sur l'ensemble d'un génome (Werner *et al.*, 1992). Ces taux de recombinaison sont fonction, notamment, du niveau de ploïdie de l'espèce considérée, et, de la structure (e. g. euchromatine ou hétérochromatine) et de la région (e. g. centromérique, télomérique) chromosomiques étudiées (Moore, 2000). Plus particulièrement, différents exemples font référence à des phénomènes de répression de la recombinaison notamment au sein d'hybrides interspécifiques et de fragments introgressés, impliquant une forte sous estimation du ratio distance génétique/distance physique (cM/kb) (Van Daelen *et al.*, 1993 ; Wei *et al.*, 1999 ; Chin *et al.*, 2001).

Dans la perspective de cloner le gène *Mex-1*, la distance physique entre les marqueurs génétiques encadrant directement ce gène doit être précisée. En effet, la région chromosomique portant le locus *Mex-1* pourrait être le siège de restriction de recombinaison. Toutefois, l'analyse des relations génétiques entre lignées dérivées de l'Hybride de Timor (Lashermes *et al.*, 2000 ; Noir *et al.*, 2003) et l'estimation des fréquences de recombinaison au

sein d'hybrides tétraploïdes (*C. arabica* x *C. canephora* 4x) (Herrera *et al.*, 2002) montrent que les phénomènes d'introgressions et les recombinaisons entre chromosomes homéologues ne sont pas limités au sein de ce matériel végétal. Ces résultats sont sans doute à associer à la faible différenciation génomique observée entre les sous-génomes C^a et E^a de *C. arabica*, et entre ces sous-génomes et le génome de *C. canephora*.

Par une approche cytologique de type BAC-FISH, la distance physique entre les marqueurs flanquant le gène *Mex-1* sera estimée (Tör *et al.*, 2002) et, le positionnement ainsi que l'orientation des contigs associés à *Mex-1* seront déterminés.

#### Stratégies de clonage envisagées

Plusieurs gènes R ont été isolés avec succès, selon la stratégie de "marche chromosomique" (e. g. Martin et al., 1993; Mindrinos et al., 1994; Cai et al., 1997; Milligan et al., 1998; Tai et al., 1999b). Toutefois, la mise en œuvre de cette approche reste techniquement lourde, tout particulièrement dans le cas de génomes complexes et de grande taille (e. g. margueurs génétiques physiquement proches). Quelques approches modifiées ont été développées sur la base du clonage positionnel. Par exemple, le clonage du gène Ve chez la tomate associe le clonage positionnel et le criblage de banque d'ADNc (Kawchuk et al., 2001). La localisation physique du locus de résistance Lr10 du génome hexaploïde de Triticum aestivum L. a été réalisée sur la base de la colinéarité des génomes diploïdes et hexaploïde du blé (Stein et al., 2000). Plus généralement, les gènes R étant caractérisés par la présence de motifs très conservés, l'identification de RGA a été développée chez de nombreuses plantes (e. g. Kanasin et al., 1996 ; Leister et al., 1996 ; Aarts et al., 1998b). Ces RGA, par leur structure et/ou par leur position, constituent de bonnes séquences dans le contexte d'une approche "gène-candidat" (pour revue, Pflieger et al., 2001). Récemment, la combinaison de la marche chromosomique et de l'approche gène-candidat tend à se développer pour le clonage de gènes R (Bendahmane et al., 1999; Van der Vossen et al., 2000; Ballvora et al., 2002).

Les RGA identifiés au sein du cluster de résistance à *M. exigua* chez le caféier constituent de bons candidats pour le clonage du gène *Mex-1*, à la fois par leur nature et leur position. La combinaison des marqueurs génétiques étroitement liés à *Mex-1* et l'inventaire, dans un premier temps, des gènes types RGA, assignés aux clones BAC associés à *Mex-1*, permettront d'établir une liste de gènes correspondant potentiellement à *Mex-1*. Par la suite, le séquençage complet d'un BAC portant le gène *Mex-1* permettra le repérage de toutes les séquences de cadres de lecture ouverts de cette région.

#### Validation fonctionnelle

La co-localisation génétique entre un gène candidat et un locus d'intérêt ne permet pas de conclure définitivement que ce gène est responsable du phénotype observé. De même, les analyses d'expression (RT-PCR, *Northern*) fournissent des arguments en faveur ou contre un candidat potentiel, mais n'apportent pas de preuve irréfutable.

La transformation génétique est un outil bien adapté à l'étude de l'expression de gènes à caractères monogéniques tels que les gènes *R*. La transformation de la plante étudiée, d'espèces apparentées ou de plante modèle (système hétérologue) à partir de bactéries du genre *Agrobacterium* s'est particulièrement développée. Les transformations transitoires par complémentation d'écotypes sensibles, de mutants de la résistance ou encore par co-expression des gènes *R* et *Avr* (Milligan *et al.*, 1998 ; Tai *et al.*, 1999a ; Bendahmane *et al.*, 2000 ; Swiderski et Innes, 2001), ainsi que les transformations stables (i. e. régénération de plantes transformées) (Tai *et al.*, 1999a ; Kawchuk et al., 2001 ; Swiderski et Innes, 2001; Deslandes *et al.*, 2002), ont été largement pratiquées à partir d'*A. tumefaciens*. L'utilisation d'*A. rhizogenes* a entraîné chez plusieurs espèces végétales le développement de phénotypes anormaux (des parties racinaires, caractère *hairy root,* et/ou aériennes de la plante) (Christey, 2001). Toutefois, lorsque le phénotype normal est conservé, cet outil s'avère particulièrement adapté à l'étude de gènes *R* conférant la résistance aux ravageurs des racines végétales (Cai *et al.*, 1997 ; Kifle *et al.*, 1999).

A ce jour, le développement de racines de caféiers transformées par des souches sauvages d'*A. rhizogenes* est maîtrisé au sein du laboratoire (H. Etienne, en préparation). Cette approche permet l'obtention rapide (3 semaines) et efficace (60 à 80% d'explants transformés) de racines, de phénotype normal, constituant un stock aseptique de matériel végétal de cultivars sensible (i. e. Caturra) et résistant (i. e. IAPAR59) à *M. exigua*. La mise au point de la transformation de ces racines à partir de vecteurs binaires (e. g. exprimant le gène rapporteur GUS) est en cours.

Ces expérimentations permettront par complémentation de racines sensibles à partir des candidats de *Mex-1* retenus (i. e. transformation *sens*), ou par extinction de la résistance au sein de racines transformées résistantes (i. e. transformation *antisens*) de valider l'identification du gène *Mex-1*.

#### 2. 3. 2 Etude du gène Mex-1

A plus longs termes, et une fois le gène *Mex-1* identifié et cloné, une caractérisation physiologique et fonctionnelle de ce gène pourra être envisagée. Dans ce cadre, la maîtrise de l'outil de transformation *in vitro* de racines de caféiers et l'accès à des ressources génétiques importantes seront mis à profit, pour analyser :

#### - L'expression du gène Mex-1

Le niveau des transcrits du gène *Mex-1* (i. e. constitutifs ou induits lors d'une infection par le nématode) sera observé dans les racines (organe-cible du nématode) et dans les feuilles. A partir de racines de caféiers transformées par *A. rhizogenes* développées *in vitro*, l'étude transcriptionnelle de *Mex-1* pourra être effectuée (e. g. localisation de l'expression, analyses des séquences régulatrices de *Mex-1*).

#### - La diversité génétique du gène Mex-1 au sein des caféiers

L'étude comparative entre les séquences du gène *Mex-1*, issues de caféiers sensibles et résistants à *M. exigua*, permettra de mettre en évidence les variations moléculaires liées aux caractères résistant ou sensible du génotype considéré.

#### 2.3.3 Sélection assistée

Dans le contexte de l'amélioration du caféier *C. arabica*, la sélection de variétés résistantes à plusieurs bio-agresseurs est un objectif majeur. Des sources de résistances vis-à-vis de divers agents pathogènes ont été identifiées au sein de caféiers sauvages (cf § Chapitre 1- Tableau 2. 1). Toutefois, la création de variétés combinant plusieurs sources de résistance est impossible dans un laps de temps acceptable, par les techniques traditionnelles d'amélioration (Lashermes *et al.*, 2000c).

L'identification de marqueurs moléculaires associés à ces résistances est une étape essentielle dans les programmes d'amélioration de la résistance aux maladies chez le caféier. De tels marqueurs peuvent être utilisés précocement sur les plantes en pépinière, permettant ainsi la réduction de la taille des essais aux champs. De plus, cette sélection précoce peut être entreprise simultanément vis-à-vis de plusieurs ravageurs, tout en s'astreignant des contraintes bio-géographiques d'une évaluation traditionnelle (e. g. dissémination de l'agent pathogène, sélection par anticipation, lieu de l'évaluation indépendant du site de propagation du parasite...). Les marqueurs associés au gène *Mex-1* permettent d'ores et déjà la sélection

assistée de variétés de caféier résistantes au nématode *M. exigua*. Cette approche sera d'autant plus performante que de nouveaux marqueurs moléculaires associés à d'autres résistances chez le caféier auront été identifiés.

### Références Bibliographiques
- Aarts M. G. M., Metz M., Holub E., Staskawicz B. J., Daniels M. J., et Parker J. E. (1998a). Different requirements for *EDS1* and *NDR1* by disease resistance genes define at least two *R* gene-mediated signaling pathways in *Arabidopsis*. *Proc Natl Acad Sci USA* 95: 10306-10311.
- Aarts M. G. M., Lintel Hekkert B., Hdub E. B., Beynon J., Stieckema W. J., et Pereira A (1998b). Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Mol Plant-Microbe Interact* 11: 251-258.
- Agrios G. N. (1997). "Plant pathology" Academic Press, San Diego.
- Agwanda C., Lashermes P., Trouslot P., Combes M. C., et Charrier A. (1997). Identification of RAPD markers for resistance to Coffee Berry Disease, Colletotrichum kahawae, in Arabica coffee. *Euphytica* 97: 241-248.
- Altschul S. F., Gish W., Miller W., Myers E. W., et Lipman D. J. (1990). Basic local alignement search tool. *J Mol Biol* 215: 403.
- Anderson P. A., Lawrence G. J., Morrish B. C., Ayliffe M. A., Finnegan E. J., et Ellis J. G. (1997). Inactivation of the flax rust resistance *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell* 9: 641-651.
- Anonyme (1998). Commodity market review 1997-1998. *In* "Fao commodities and trade division", Rome, Italie, pp. 91.
- Anthony F. (1992). "Les ressources génétiques des caféiers.," ORSTOM Editions.
- Anthony F., Astorga C., et Berthaud J. (1999). Los recursos genéticos: las bases de una solucion genética a los problemas de la caficultura latinoamericana. *In* "Desafios de la caficultura centroamericana" (B. Bertrand, et B. Rapidel, Eds.), IICA-CIRAD, San José, Costa Rica, pp. 369-406.
- Anthony F., Berthaud J., Guillaumet J.-L., et Lourd M. (1987). Collecting wild *Coffea* species in Kenya and Tanzania. *Plant Genetic Resources Newsletter* 69: 23-29.
- Anthony F., Bertrand B., Quiros O., Wilches A., Lashermes P., Berthaud J., et Charrier A. (2001). Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. *Euphytica* 118: 53-65.
- Anthony F., Combes M. C., Astorga C., Bertrand B., Graziosi G., et Lashermes P. (2002a). The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theor Appl Genet* 104: 894-900.
- Anthony F., Topart P., Anzueto F., Astorga C., et Bertrand B. (2002b). La resistencia genética del café (*Coffea* spp.) a los nematodos (*Meloidigyne* spp.): identification y utilizacion para la caficultura latinoamericana. *Manejo integradode plagas y Agroecoloia CATIE*, sous presse.
- Anzueto F., Bertrand B., Sarah J. L., Eskes A. B., et Decazy B. (2001). Resistance to *Meloidogyne incognita* in Ethiopian Coffea arabica accessions. *Euphytica* 118: 1-8.
- Ayliffe M. A., Frost D. V., Finnegan E. J., Lawrence G. J., Anderson P. A., et Ellis J. G. (1999). Analysis of alternative transcripts of the flax *L6* rust resistance gene. *Plant J* 17: 287-292.
- Bailey D. M. (1941). The seeding test method for root-knot nematode resistance. *Proc Am Soc Hortic Sci* 38: 573-575.
- Baker B., Zambryski P., Staskawicz B. J., et Dinesh-Kumar S. P. (1997). Signaling in plantmicrobe interactions. *Science* 276: 726-733.
- Ballvora A., Ercolano M. R., Weiss J., Meksem K., Bormann C. A., Oberhagemann P., Salamini F., et Gebhardt C. (2002). The *R1* gene for potato resistance to late blight (*Phytophtora infestans*) belongs to the leucine zipper/ NBS/LRR class of plant resistance genes. *Plant J* 30: 361-371.

- Ballvora A., Hesselbach J., Niewohner J., Leister D., Salamini F., et Gebhardt C. (1995). Marker enrichment and high-resolution map of the segment of potato chromosome VII harbouring the nematode resistance gene *Gro1*. *Mol Gen Genet* 249: 82-90.
- Bendahmane A., Kanyuka K., et Baulcombe D. C. (1999). The *Rx* gene from potato controls separate virus resistance and cell death responses. *Plant Cell* 11: 781-792.
- **Bendahmane A., Querci M., Kanyuka K., et Baulcombe D.** (2000). *Agrobacterium* transient expression system as a tool for the isolation of disease resistance genes: application to the Rx2 locus in potato. *Plant J* 21: 73-81.
- Bennetzen J., Richter T., Hu G. (1994). Organisation and hyperevolution of rust resistance genes in maize. *In* "Advances in molecular genetics of plant-microbe interactions." (M. J. Daniels, Ed.), Kluwer Academic Publishers, Netherlands.
- Bent A. F., Kundel B. N., Dahlbeck D., Brown K. L., Schmidt R., Giraudat J., Leung J., et Staskawicz B. J. (1994). *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* 265: 1856-1860.
- Bergelson J., Kreitman M., Stahl E. A., et Tian D. (2001). Evolutionary dynamics of plant *R*-genes. *Science* 292: 2281-2285.
- **Berthaud J.** (1986). Les ressources génétiques pour l'amélioration des caféiers africains diploïdes. *In* "Evaluation de la richesse génétique des populations sylvestres et de ses mécanismes organisateurs. Conséquences pour l'application", ORSTOM, Paris, pp. 379.
- Berthaud J., et Charrier A (1988). Genetic resources of *Coffea. In* "Coffee vol 4: Agronomy" (Clarke R. J, et Macrae R., Eds), London, pp. 1-40.
- Bertrand B., Aguilar G., Bompard E., et Rafinon A. (1997). Comportement agronomique et résistance aux principaux déprédateurs des lignées Sarchimor et Catimor au Costa Rica. *Plantations, Recherche, Développement*, 312-318.
- Bertrand B., Aguilar G., Santacreo R., et Anzueto F. (1999). El mejoramiento genético en América Central. *In* "Desafios de la caficultura centroamericana" (Bertrand B., et Rapidel, B., Eds), IICA-CIRAD, San José, Costa Rica, pp. 405-456.
- Bertrand B., Anthony F., et Lashermes P. (2001). Breeding for resitance to *Meloidogyne* exigua of *Coffea arabica* by introgression of resistance genes of *Coffea canephora*. *Plant Pathol* 50: 637-643.
- Bertrand B., Vasquez N., et Decazy B. (1995). Nature of coffee resistance to two costa rican *Meloidogyne* populations, ASIC, 16e colloque, Kyoto.
- **Bettencourt A.** (1973). Cosideradoçoes gerais sobre o "Hybrido de Timor", Instituto Agronomico de Campinas, Campinas, Brasil, pp. 256.
- Bittner-Eddy D., Crute I. R., Holub E. B., et Beynon J. L. (2000). *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *Plant J* 21: 177-188.
- Blumwald E., Aharon G. S., et Lam B. C. H. (1998). Early signal transduction pathways in plant-pathogen interactions. *Trds Plant Sci* 3: 342-346.
- Bonas U., et Lahaye T. (2002). Plant disease resistance triggered by pathogen-derived molecules: refined models of specific recognition. *Cur Opin Microbiol* 5: 44-50.
- Bonas U., et Van den Ackerveken G. (1999). Gene-for-gene interactions: bacterial avirulence proteins specify plant disease resistance. *Cur Opin Microbiol* 2: 94-98.
- Botella M. A., Parker J. E., Frost L. N., Bittner-Eddy P. D., Beynon J. L., Daniels M. J., Holub E. B., et Jones J. D. (1998). Three genes of the *Arabidopsis RPP1* complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. *Plant Cell* 10: 1848-1860.
- Bouharmont J. (1963). Somatic chromosomes of some Coffea species. Euphytica 12: 254-257.

- Bridson D. M. (1994). Additional note on *Coffea* (*Rubiaceae*) from tropical east Africa. *Kew Bulletin* 49: 331-342.
- Bridson D. M., et Verdcourt B. (1988). "Flora of tropical east Africa *Rubiacae*." (Polhill R. M., Ed) Part 2, 727 pp.
- Brommonschenkel S. H., Frary A., Frary A., et Tanksley S. D. (2000). The broad-spectrum tospovirus resistance gene *Sw*-5 of tomato is a homolog of the root-knot nematode resistance gene *Mi. Mol Plant-Microbe Interact* 13: 1130-1138.
- **Brown C. R., Yang C. P., Mojtahedi H., Santo G., et Masuelli R.** (1996). RFLP analysis of resistance to Clumbia root-knot nematode derived from *Solanum bulbocastanum* in a BC₂ population. *Theor Appl Genet* 92: 572-576.
- Brueggeman R., Rostoks N., Kudrna D., Kilian A., Han F., Chen J., Druka A., Steffenson B., et Kleinhofs A. (2002). The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proc Natl Acad Sci USA* 99: 9328-9333.
- Cai D., Kleine M., Kifle S., Harloff H. J., Sandal N. N., Marcker K. A., Klein-Lankhorst R. M., Salentijn E. M. J., Lange W., Stiekema W. J., Wyss U., Grundler F. M. W., et Jung C. (1997). Positional cloning of a gene for nematode resistance in sugar beet. *Science* 275: 832-834.
- Caicedo A. L., Schaal B. A., et Kunkel B. N. (1999). Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 96: 302-306.
- Campos V. P., Sivapalan P., et Gnanapragasam C. (1990). Nematode parasites of coffee, cocoa and tea. *In* "Plant parasitic nematodes in subtropical and tropical agriculture" (Luc M., Sikora R. A., et Bridge J., Eds), pp. 387-400.
- Cannon S. B., Zhu H., Baumgarten A. M., Spangler R., May G., Cook D. R., et Young N. D. (2002). Diversity, distribution and ancient taxonomic relationships within the TIR and Non-TIR NBS-LRR resistance gene subfamilies. *J Mol Evol* 54: 548-562.
- Carneiro R. M. D. G., Almeida M. R. A., et Quénéhervé P. (2000). Enzymes phenotypes of *Meloidogyne* spp. populations. *Nematology* 2: 645-654.
- Carneiro R. M. D. G., Carneiro R. G., Abrantes I. M. O., Santos M. S. N., et Almeida M. R. A. (1996). *Meloidogyne paranaensis* sp. (*Nemata: Meloidogynidae*) a root-knot nematode parazitizing coffee from Brazil. *J Nematology* 28: 177-189.
- **Carvalho A.** (1988). Principles and practices of coffee plant breeding for productivity and quality factors: *Coffea arbica. In* "Coffee vol 4: Agronomy", (Clarke R. J., et Macrae R., Eds), London, pp. 129-165.
- Century K. S., Holub E. B., et Staskawicz B. J. (1995). *NDR1*, a locus of *Arabidospis thaliana* that is required for disease resistance to both a bacterial and a fungal pathogen. *Proc Natl Acad Sci USA* 92: 6597-6601.
- **Charrier A.** (1978). Etude de la structure et de la variabilité génétique des caféiers : résultats des études et des expérimentations réalisées au Cameroun, en Côte d'Ivoire et à Madagascar sur l'espèce *Coffea arabica* L. collectée en Ethioie par une mission ORSTOM en 1966. Bulletin IFCC n°14, Paris, France, 100p.
- Charrier A., et Eskes A. B. (1997). "Les caféiers" (Charrier A., Jacquot M., Hamon, S. et Nicolas D., Eds) Paris.
- Chin D. B., Arroyo-Garcia R., Ochoa O. E., Kesseli R. V., Lavelle D. O., et Michelmore R.
   W. (2001). Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (*Lactuca sativa*). *Genetics* 157: 831-849.
- **Christey M. C.** (2001). Use of *Ri*-mediated transformation for production of transgenic plants. In vitro *Cell Dev Biol-Plant* 37 : 687-700.
- Cohn J., Sessa G., et Martin G. B. (2001). Innate immunity in plants. *Cur Opin Immunol* 13: 55-62.

- Collins N., Drake J., Ayliffe M. A., Sun Q., Ellis J. G., Hulbert S., et Pryor T. (1999). Molecular characterisation of the maize *Rp1-D* rust resistance haplotype and its mutants. *Plant Cell* 11: 1365-1376.
- Cooley M. B., Pathirana S., Wu H. J., Kachroo P., et Klessig D. F. (2000). Members of the *Arabidopsis* HRT/*RPP8* family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell* 12: 663-676.
- **Couturon E., Lashermes P., et Charrier A.** (1998). First intergeneric hybrids (*Psilanthus ebracteolatus* H. x *Coffea arabica*) in coffee trees. *Can J of Bot* 76: 542-546.
- Crane P. R., Friis E. M., et Pederson K. J. (1995). The origin and early diversification of angiosperms. *Nature* 374: 27-33.
- Creusot F., Macadré C., Ferrier Cana E., Riou C., Geffroy V., Sévignac M., Dron M., et Langin T. (1999). Cloning and molecular characterization of three members of the NBS-LRR subfamily located in the vicinity of the *Co-2* locus for anthracnose resistance in *Phaseolus vulgaris*. *Genome* 42: 254-264.
- Cronn R. C., Small R. L., et Wendel J. F. (1999). Duplicated genes evolve independantly after polyploid formation in cotton. *Proc Natl Acad Sci USA* 96: 14406-14411.
- **Cros J.** (1994). Implications phylogénétiques des variations de l'ADN chloroplastique chez les caféiers (genres *Coffea* L. et *Psilanthus* Hook. f.), *Thèse Doctorale*, Univerité Montpellier II, pp. 160.
- Cros J., Combes M. C., Trouslot P., Anthony F., Hamon S., Charrier A., et Lashermes P. (1998). Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. *Mol Phylogenet Evol* 9: 109-117.
- Curi S. M., Carvalho A., Moraes F. P., Monaco L. C., et Arruda H. V. (1970). Novas fontes de resistencia genética de *Coffea* no controle de nematoide do cafeeiro *Meloidogyne exigua*. *Biologico* 36: 293-295.
- Dangl J. L., et Jones J. D. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411: 826-833.
- **De Guiran, et Netsher** (1970). Les nématodes du genre *Meloidogyne*, parasites de cultures tropicales. *In* "Cahiers de l'ORSTOM", pp. 150-185.
- de Ilarduya O. M., Moore A. E., et Kaloshian I. (2001). The tomato *Rme1* locus is required for *Mi-1*-mediated resistance to root-knot nematodes and the potato aphid. *Plant J* 27: 417-25.
- Deng Z., Tao Q., Chang Y. L., Huang S., Ling P., Yu C., Chen C., Gmitter Jr F. G., et Zhang H. B. (2001). Construction of a bacterial artificial chromosome (BAC) library for citrus and identification of BAC contigs containing resistance gene candidates. *Theor Appl Genet* 102: 1177-1184.
- Deslandes L., Olivier J., Theulières F., Hirsch J., Feng D. X., Bittner-Eddy P., Beynon J., et Marco Y. (2002). Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by te recessive *RRS1-R* gene, a member of a novel family of resistance genes. *Proc Natl Acad Sci USA* 99: 2404-2409.
- Devoto A., Piffanelli P., Nilsson I., Wallin E., Panstruga R., von Heijne G., et Schulze-Lefert P. (1999). Topology, subcellular localization, and sequence diversity of the Mlo family in plants. *J Biol Chem* 274: 34993-5004.
- Di Vito M., Grozzoli R., et Vovlas N. (2000). Pathogenicity of *Meloidogyne exigua* on coffee (*Coffea arabica* L.) in pots. *Nematropica* 30: 55-61.
- Dickinson M. J., Jones D. A., et Jones J. D. (1993). Close linkage between the Cf-2/Cf-5 and Mi resistance loci in tomato. *Mol Plant-Microbe Interac* 6: 341-347.
- Dixon M. S., Golstein C., Thomas C. M., van der Biezen E. A., et Jones J. D. G. (2000). Genetic complexity of pathogen perception by plants: the example of *Rcr3*, a tomato gene required specifically by *Cf-2*. *Proc Natl Acad Sci USA* 97: 8807-8814.

- Dixon M. S., Hatzixanthis K., Jones D. A., Harrison K., et Jones J. D. (1998). The tomato *Cf*-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. *Plant Cell*. 10: 1915-1925.
- **Dixon M. S., Jones D. A., Keddie J. S., Thomas C. M., Harrison K., et Jones J. D. J.** (1996). The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* 84: 451-459.
- Djian-Caporalino C., Pijarowski L., Fazari A., Samson M., Gaveau L., O'Byrne C., Lefebvre V., Caranta C., A. P., et Abad P. (2001). High-resolution genetic mapping of the pepper (*Capsicum annuum* L.) resistance loci *Me3* and *Me4* conferring heat-stable resistance to root-knot nematodes (*Meloidogyne* spp.). *Theor Appl Genet* 103: 592-600.
- **Dodds P. N., Lawrence G. J., et Ellis J. G.** (2000). Six amino acid changes confined to the Leucine-Rich Repeat b-strand/b-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell* 13: 163-178.
- **Dorey S., Baillieul F., Pierrel M. A., Saindrenan F., Fritig B., et Kauffmann S.** (1997). Spatial et temporal induction of cell death, defense genes and accumulation of salicylic acid in tobacco leaves reacting hypersensitively to a fungal glycoprotein elicitor. *Mol Plant-Microbe Interac*10: 646-655.
- **Dufour M., Anthony F., Bertrand B., et Eskes A. B.** (1997). Identification de caféiers mâlesstériles de *Coffea arabica* au CATIE, Costa Rica. *Plantations, Recherche, Développement .* 4: 401-407.
- Eastwood R. F., Lagudah E. S., Appels R., Hannah M., et Kollmorgen J. F. (1991). *Triticium tauschii*: a novel source of resistance to cereal cyst nematode (*Heterodera avanae*). *Aust J Agri Res* 42: 69-77.
- **Eisenback J. D.** (1998). Morphology and systematics. *In* "Plant nematode interactions" (Barker K. R., Pedreson G. A., et Windham G. L., Eds), American society of agronomy, Crop science society of America, Soil science society of America, Madison, USA, pp. 37-63.
- Ellis J. G., Lawrence G. J., Ayliffe M. A., Anderson P., Collins N., Finnegan J., Frost D., Luck J., et Pryor T. (1997). Advances in the molecular genetic analysis of the flax-flax rust interaction. *Ann Rev Phytopathol* 35: 271-291.
- Ellis J. G., Lawrence G. J., Luck J. E., et Dodds P. N. (1999). Identification of regions in alleles of the flax rust resistance gene *L* that determine differnces in gene-for-gene specificity. *Plant Cell* 11: 495-506.
- Ellis P. R., et Smith M. J. W. (1971). Inheritance of resistance to potato cyst-eelworm (*Heterodera rostochinsis* Woll.) in the genus *Lycopersiscon*. *Euphytica* 20: 93-101.
- **Eskes A. B.** (1989). Resistance. *In* "Coffee rust: Epidemiology, resistance and management", (Kushalappa A. C., et Eskes A. B., Eds), Boca Raton, Florida, pp. 171-291.
- Etienne H., Anthony F., Dussert S., Fernandez D., Lashermes P., et Bertrand B. (2002). Biotechnological applications for the improvement of coffee (*Coffea arabica*). In vitro *Cell Dev Biol-Plant* 38: 129-138.
- Feldman M., Liu B., Segal G., Abbo S., Levy A. A., et Vega J. M. (1997). Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes. *Genetics* 147: 1381-1387.
- **Felsenstein J.** (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.
- Feuillet C., Penger A., Gellner K., Mast A., et Keller B. (2001). Molecular evolution of receptor-like kinase genes in hexaploid wheat. Independant evolution of orthologs after polyploidization and mechanisms of local rearrangements at paralogous loci. *Plant Physiol* 125: 1304-1313.
- Feys B. J., Moisan L. J., Newman M.-A., et Parker J. E. (2001). Direct interaction between the *Arabidopsis* disease resistance signaling proteins, *EDS1* and *PAD4*. *EMBO J* 20: 5400-5411.

- **Feys B. J., et Parker J. E.** (2000). Interplay of signaling pathways in plant disease resistance. *Trds Genet* 16: 449-455.
- Flor H. H. (1971). Current status of the gene-for-gene concept. Ann Rev Phytopathol 9: 275-296.
- **Fourmann M., Charlot F., Froger N., Delourme R., et Brunel D.** (2001). Expression, mapping and genetic variability of *Brassica napus* disease resistance gene analogues. *Genome* 44: 1083-1099.
- Fritig B., Heitz T., et Legrand M. (1998). Antimicrobial proteins in induced plant defense. *Cur Opin Immunology* 10: 16-22.
- Ganal M. W., Simon R., Brommonschenkel S., Arndt M., Phillips M. S., Tanksley S. D., et Kumar A. (1995). Genetic mapping of a wide spectrum nematode resistance gene (*Hero*) against *Globodera rostochiensis* in tomato. *Mol Plant Microbe Interact* 8: 886-891.
- **Gassman W., Hinsch M. E., et Staskawicz B. J.** (1999). The *Arabidopsis RPS4* bacterialresistance gene is a memeber of the TIR-NBS-LRR family of disease-resistance genes. *Plant J* 20: 265-277.
- Geffroy V., Sicard D., de Oliveira J. C. F., Sévignac M., Cohen S., Gepts P., Neema C., Langin T., et Dron M. (1999). Identification of an ancestral resistance gene cluster involved in the coevolution process between *Phaseolus vulgaris* and its fungal pathogen *Colletotrichum lindemuthianum*. *Mol Plant-Microbe Interac* 12: 774-784.
- **Glazebrook J.** (2001). Genes controlling expression of defense responses in Arabidopsis 2001 status. *Curr Opin Plant Biol* 4: 301-308.
- Goff S. A., Ricke D., Lan T. H., Presting G., Wang R., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H., Hadley D., Hutchison D., Martin C., Katagiri F., Lange B. M., Moughamer T., Xia Y., Budworth P., Zhong J., Miguel T., Paszkowski U., Zhang S., Colbert M., Sun W. L., Chen L., Cooper B., Park S., Wood T. C., Mao L., Quail P., Wing R., Dean R., Yu Y., Zharkikh A., Shen R., Sahasrabudhe S., Thomas A., Cannings R., Gutin A., Pruss D., Reid J., Tavtigian S., Mitchell J., Eldredge G., Scholl T., Miller R. M., Bhatnagar S., Adey N., Rubano T., Tusneem N., Robinson R., Feldhaus J., Macalma T., Oliphant A., et Briggs S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92-100.
- Gonçalves W., Ferraz L. C. C. B., Lima M. M. A., et Silvarolla M. B. (1996). Reações do cafeeiro às raças 1, 2 e 3 de *Meloidigyne incognita*. *Summa Phytopathologica* 22: 172-177.
- **Gonçalves W., et Pereira A. A.** (1998). Resistência do cafeeiro a nematóïdes. IV. Reacção de cafeeiros derivados do Híbrido de Timor a *Meloidogyne exigua*. *Nematologia Brasileira* 22: 39-50.
- **Gonçalves M. M., et Rodrigues, M. L.** (1976). Estudos sobre o café de Timor. II. Nota sobre as posibilidades de producao do hibrido de Timor no seu habitat natural. *In* "Missao de Estudos Agronomicos do Ultramar", Lisboa, Portugal, Comunicacoes 86: 31-72.
- Goodman R. N., et Novacky A. J. (1994). "The hypersensitive reaction in plants to pathogens," APS Press, St Paul.
- Grant M., et Mansfield J. (1999). Early events in host-patogen interactions. *Cur Opin Plant Biol* 2: 312-319.
- Grant M. R., Godiard L., Straube E., Ashfield T., Lewald J. (1996). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* 269: 843-846.
- Grant M. R., Godiard L., Straube E., Ashfield T., Lewald J., Sattler A., Innes R. W., et Dangl J. L. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* 269: 843-846.
- **Grassias M., et Kammacher P.** (1975). Observations sur la conjugaison chromosomique de *Coffea arabica* L. *Café Cacao Thé* 19: 177-190.

- Hammond-Kosack K. E., et Jones J. D. G. (1997). Plant disease resistance genes. Ann Rev Plant Physiol Plant Mol Biol 48: 575-607.
- Heath M. C. (2000a). Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology* 3: 315-319.
- Heath M. C. (2000b). Hypersensitive response-related death. Plant Mol Biol 44: 321-334.
- Hebsgaard S. M., Korning P. G., Tolstrup N., Engelbrecht J., Rouze P., et Brunak S. (1996). NetPlantGene WWW Server, Splice Site Prediction in *Arabidopsis thaliana* DNA by combining local and global sequence information. *Nuc Acids Res* 17: 3439-3452.
- Heller R., Schondelmaier J., Steinrücken G., et Jung C. (1996). Genetic localisation of four genes for nematodes (*Heterodera schachtii* Schm.) resistance in sugar beet (*Beta vulgaris* L.). *Theor Applied Genet* 92: 991-997.
- Henk A. D., Warren R. F., et Innes R. W. (1999). A new Ac-like transposon of *Arabidopsis* is associated with a deletion of the RPS5 disease resistance gene. *Genetics* 151: 1581-1589.
- Hernández A. M., Fargette V., Molinier H., Ramanason B., Decazy B., et Sarah J. L. (1996). Enzymatic characterization and reproductive fitness on coffee of root-knot nematode populations from Central America. *Nematropica* 26: 274.
- Herrera J. C., Combes M. C., Anthony F., Charrier A., et 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* x *C. canephora*). Theor Appl Genet 104: 661-668.
- Ho J. Y., Weide R., Ma H. M., Van Wordragen M., Lambert K. N., Koornneef M., Zabel P., et Williamson V. M. (1992). The root-knot nematode resistance gene (*Mi*) in tomato: construction of a molecular linkage map and identification of dominant cDNA markers in resistant genotypes. *Plant J* 2: 971-982.
- Hulbert S., Webb C. A., Smith S. M., et Sun Q. (2001). Resistance gene complexes: evolution and utilization. *Ann Rev Phytopathol* 39: 285-312.
- Hulbert S. H. (1997). Structure and evolution of the rp1 complex conferring rust resistance in maize. *Ann Rev Phytopathol* 35: 293-310.
- Hussey R. S., et Williamson V. M. (1998). Physiological and molecular aspects of nematodes parasitism. *In* "Plant and nematode interactions" (Barker K. R., Pederson G. A., et Windham G. L., Eds), Madison, Wisconsin, USA, pp. 87-109.
- ICO (2002a). Coffee market report, International Coffee Organization.
- **ICO** (2002b). The global crisis: A threat to sustainable development, *International Coffee Oraganization.*
- Janssen G. J. H., Bakker J., et Gommers F. J. (1991). Mendelian proof for a gene-for-gene relationship between virulence of *Globodera rostochiensis* and the *H1* resistance gene in *Solanum tuberosum* spp. *andigena* CPC 1673. *Rev Nématol* 14: 207-211.
- Jepson S. B. (1987). "Identification of root-knot nematodees (*Meloidogyne* species)," CAB International, Walingford, United Kingdom.
- Jia Y., Mc Adams S. A., Bryan G. T., Hershey H. P., et Valent B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO J* 19: 4004-4014.
- Johal G. S., et Briggs S. P. (1992). Reductase activity encoded by the *Hm1* disease resistance gene in maize. *Science* 258: 985-987.
- Jones D. A., et Jones J. D. G. (1997). The role of leucine-rich repeat proteins in plant defenses. *Adv Bot Res* 24: 89-167.
- Jones D. A., Thomas C. M., Hammond-Kosack K. E., Balint-Kurti P. J., et Jones J. D. (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266: 789-793.

- Jones J. D. G., Thomas C. M., Hammond-Kosack K. E., Balint-Kurti P. J. et Jones D. A. (1993). Two complex resistance loci revealed in tomato by classical and RFLP mapping of the *Cf-2*, *Cf-4*, *Cf-5* and *Cf-9* genes for resistance to *Cladosporium fulvum*. *Mol Plant-Microbe Interact* 6: 348-357.
- Joosten M. H. A. J., Vogelsang R., Cozijnsen T. J., Verberne M. C., et de Wit P. J. G. M. (1997). The biotrophic fungus *Cladosporium fulvum* circumvents *Cf-4*-mediated resistance by producing unstable AVR4 elicitors. *Plant Cell* 9: 367-379.
- Kaloshian I., Lange W. H., et Williamson V. M. (1995). An aphid-resistance locus is tightly linked to the nematode-ressitance gene, *Mi*, in tomato. *Proc Natl Acad Sci USA* 92: 622-625.
- Kaltz O., et Shykoff J. A. (1998). Local adaptation in host-parasite systems. *Heredity* 81: 361-370.
- Kanazin V., Marek L. F. M., et Shoemaker R. C. (1996). Resistance gene analogs are conserved and clustererd in soybean. *Proc Natl Acad Sci USA* 93: 11746-11750.
- Kawchuk L. M., Hachey J., Lynch D. R., Kulcsar F., van Rooijen G., Waterer D. R., Robertson A., Kokko E., Byers R., Howard R. J., Fischer R., et Prüfer D. (2001). Tomato Ve disease resistance genes encode cell surface-like receptors. Proc Natl Acad Sci USA 98: 6511-6515.
- Keen N. T. (1982). Specific recognition in gene-for-gene host-parasite systems. Adv Plant Pathol 1: 35-81.
- Kifle S., Shao M., Jung C., et Cai D. (1999). An improved transformation protocol for studying gene expression in hairy roots of sugar beet (*Beta vulgaris* L.). *Plant Cell Report* 18: 514-519.
- Kim Y. J., Lin N. C., et Martin G. B. (2002). Two distinct *Pseudomonas* effector proteins interact with the *Pto* kinase and activate plant immunity. *Cell* 109: 598-598.
- **Kimura M.** (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16: 111-120.
- Kinloch R. A. (1990). Screening for resistance to root-knot nematodes. *In* "Methods for evaluating plant species for resistance to plant-parasitic nematodes" (Starr J. L., Ed), Society of Nematologists, Hyattsville, MD-USA, pp. 16-23.
- Kosambi D. D. (1944). The estimation of map distances from recombination values. *Annals of Eugenics* 12: 172-175.
- Kover P. X., et Caicedo A. L. (2001). The genetic architecture of disease resistance in plants and the maintenance of recombination by parasites. *Mol Ecol* 10: 1-16.
- Kreike C. M., de Koning J. R. A., Vinke J. H., Van Ooijen J. W., et Stiekema W. J. (1994). Quantitatively-inherited resistance to *Globodera pallida* is dominated by one major locus in *Solanum spegazzinii. Theor Appl Genet* 88: 764-769.
- Krüger J., Thomas C. M., Golstein C., Dixon M. S., Smoker M., Tang S., Mulder L., et Jones J. D. G. (2002). A tomato cysteine protease required for *Cf-2* dependent disease resistance and suppression of autonecrosis. *Science* 296: 744-747.
- Lagudah E. S., Moullet O., et Appels R. (1997a). Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. *Genome* 40: 659-665.
- Lahaye T. (2002). The *Arabidopsis RRS1-R* disease resistance gene-uncovering the plant's nucleus as the new battlefield of plant defense? *Trds Plant Sci* 7: 425-427.
- Lahaye T., et Bonas U. (2001). Molecular secrets of bacterial type III effector proteins. *Trds Plant Sci* 6: 479-485.
- Lander E. S., Green P., Abrahamson J., Bralow A., Daly M. J., Lincoln S. E., et Newburg L. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174-181.

- Lashermes P., Andrzejewski S., Bertrand B., Combes M. C., Dussert S., Graziosi G., Trouslot P., et Anthony F. (2000b). Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theor Appl Genet* 100: 139-146.
- Lashermes P., Bertrand B., et Anthony F. (2000c). Current status of molecular research on coffee with special reference to marker assisted breeding scope and limitations. *In* "Proceedings of the International Scientific Symposium on Coffee", (Surya Prakash, N., Raghuramulu, Y et Devasia, J. Eds), Bangalore, India, pp. 31-39.
- Lashermes P., Combes M. C., Prakash N. S., Trouslot P., Lorieux M., et Charrier A. (2001). Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome* 44: 589-596.
- Lashermes P., Combes M. C., Robert J., Trouslot P., D'Hont A., Anthony F., et Charrier A. (1999a). Molecular characterisation and origin of the *Coffea arabica* L. genome. *Mol Gen Genet*261: 259-266.
- Lashermes P., Combes M. C., Trouslot P., et Charrier A. (1997). Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. *Theor Appl Genet* 94: 947-955.
- Lashermes P., Couturon E., Moreau N., Paillard M., et Louarn J. (1996). Inheritance and genetic mapping of self-incompatibility in *Coffea canephora* Pierre. *Theor Appl Genet* 93: 458-462.
- Lashermes P., Paczek V., Trouslot P., Combes M. C., Couturon E., et Charrier A. (2000a). Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interspecific hybrid *C arabica x C. canephora. J Heredity* 91: 81-85.
- Lawrence G. J., Finnegan E. J., Ayliffe M. A., et Ellis J. G. (1995). The *L6* gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N. Plant Cell* 7: 1195-1206.
- Lecouls A. C., Rubio-Cabetas M. J., Minot J. C., Voisin R., A. B., Salesses G., Dirlewanger E., et Esmenjaud D. (1999). RAPD and SCAR markers linked to the Ma1 root-knot nematode resistance gene in Myrobalan plum (*Prunus cerasifera* Ehr). *Theor Appl Genet* 99: 328-335.
- Leister D., Ballvora A., Salamini F., et Gehardt C. (1996). A PCR based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nat Genet* 14: 421-429.
- **Leister D., et Katagiri F.** (2000). A resistance gene product of the nucleotide binding site-leucin rich repeats class can form a complex with bacterial avirulence proteins *in vivo*. *Plant J* 22: 345-354.
- Leister D., Kurth J., Laurie D. A., Yano M., Sasaki T., Devos K., Graner A., et Schulze-Lefert P. (1998). Rapid reorganisation of resistance gene homologues in cereal genomes. *Proc Natl Acad Sci USA* 95: 370-375.
- Leroy J. F. (1980). Evolution et taxogenèse chez les caféiers (*Coffea* L., *Psilanthus* Hook. f. et *Nostolachma* Durand). Hypothèse sur leur origine. *In* "Comptes rendus de l'Académie des Sciences", Paris, pp. 593-596.
- Lewers K. S., Nilmalgoda S. D., Warner A. L., Knap H. T., et Matthews, B. F. (2001). Physical mapping of resistant and susceptible soybean genomes near the soybean cyst nematode resistance gene *Rhg4*. *Genome* 44: 1057-1064.
- Liharska T., Koornneef M., van Wordragen M., van Kammen A., et Zabel P. (1996). Tomato chromosome 6: effect of alien chromosomal segments on recombinant frequencies. *Genome* 39: 485-491.
- Liu H., Sachidanandam R., et Stein L. (2001). Comparative genomics between rice and *Arabidopsis* shows scant collinearity in gene order. *Genome Res* 11: 2020-2026.

- Liu Y., Schiff M., Marathe R., et Dinesh-Kumar S. P. (2002). Tobacco *Rar1*, *EDS1* and *NPR1/NIM1* like genes are required for N-mediated resistance to tobacco mosaic virus. *Plant J* 30: 415-429.
- Lowenberg J. R., Sullivan T., et Schuster M. L. (1960). Gall induction by *Meloidogyne incognita* by surface feeding and factors affecting the behaviour pattern of the second-stage larvae. *Phytopathology* 50: 322-333.
- Luc M., Hunt D. J., et Machon E. J. (1990). Morphology, anatomy and biology of plant parasitic nematodes - A synopsis. *In* "Plant parasitic nematodes in subtropical and tropical agriculture" (Luc, M., Sikora, R. A. et Bridge, J. Eds.), CAB International, Walingford, United Kingdom, pp. 1-44.
- Luck J. E., Lawrence G. J., Dodds P. N., Shepherd K. W., et Ellis J. G. (2000). Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *Plant Cell* 12: 1367-1377.
- Lysak M. A., Fransz P. F., Ali H. B., et Schubert I. (2001). Chromosome painting in *Arabidopsis* thaliana. *Plant J* 28: 689-697.
- Maggenti A. R. (1991). Nemata: higher classification. *In* "Manual of agricultural nematology" (Nickle W. R., Ed), Dekker, New York, pp. 147-187.
- Maleck K., et Dietrich R. A. (1999). Defense on multiple fronts: how do plants cope with diverse enemies? *Trds Plant Sci* 4: 215-219.
- Martin G. B., Brommonschenkel S. H., Chunwongse J., Frary A., Ganal M. W., Spivey R., Wu T. E., E. D., et Tanksley S. D. (1993). Map-based-cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262: 1432-1436.
- Martin G. B., Frary A., Wu T., Brommonschenkel S., Chunwongse J., Earle E. D., et Tanksley S. D. (1994). A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. *Plant Cell* 6: 1543-52.
- Mazzafera P., et Carvalho, A. (1992). Breeding for low seed caffeine content of coffee (*Coffea* L.) by interspecific hybridization. *Euphytica* 59: 55-60
- McDowell J. M., Dhandaydham M., Long T. A., Aarts M. G. M., Goff S., Holub E. B., et Dangl J. L. (1998). Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis*. *Plant Cell* 10: 1861-1874.
- **Mestre P., Brigneti G., et Baulcombe D. C.** (2000). An *Ry*-mediated response in potato requires the intact active site of the NIa proteinase from potato virus Y. *Plant J* 23: 653-661.
- Meyers B. C., Chin D. B., Shen K. A., Sivaramakrishnan S., Lavelle D. O., Zhang Z., et Michelmore R. W. (1998). The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *Plant Cell* 10: 1817-1832.
- Meyers B. C., Dickerman A. W., Michelmore R. W., Sivaramakrishnan S., Sobral B. W., et Young N. D. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant JI* 20: 317-332.
- **Michelmore R.** (2002). Disease resistance in plant pathgology. 6th conference of European Foundation for Plant Pathology. Prague, République Tchèque, 8-14 septembre 2002.
- Michelmore R. (2000). Genomic approaches to plant disease resistance. *Curr Opin Plant Biol* 3: 125-31.
- Michelmore R. W., et Meyers B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and birth-and-death process. *Genome Res* 8: 1113-1130.
- Michelmore R. W., Paran I., et Kesseli R. V. (1991). Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genome regions by using segregating populations. *Proc Natl Acad Sci USA* 88: 9828-9832.

- Milligan S. B., Bodeau J., Yaghoobi J., Kaloshian I., Zabel P., et Williamson V. M. (1998). The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. *Plant Cell* 10: 1307-1319.
- Mindrinos M., Katagiri, F., Yu G. L., et Ausubel F. M. (1994). The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* 78: 1089-99.
- Moore G. (2000). Cereal chromosome structure, evolution, and pairing. Ann Rev Plant Physiol Plant Mol Biol 51: 195-222.
- Myers E. W., et Miller W. (1988). ALIGN: Optimal alignments in linear space. *Comput Appl Biosci* 4: 11-17.
- **Nei M., et Gojobori T.** (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418-26.
- Netscher C., et Sikora R. A. (1990). Nematode parasites of vegetable in plant-parasitic nematodes in subtropical agriculture. *In* "Plant parasitic nematodes in subtropical agriculture" (Luc M., Sikora R. A., et Bridge J., Eds), CAB International, London.
- Noe J. P., Sasser J. N., et Imbriani J. L. (1991). Maximising the potential of cropping systems for nematode management. *J Nematol* 23: 353-362.
- Noel L., Moores T. L., Van der Biezen E. A., Parniske M., Daniels M. J., Parker J. E., et Jones J. D. (1999). Pronouced intraspecific haplotype divergence at the *RPP5* complex disease resitance locus of *Arabidopsis*. *Plant Cell* 11: 2099-112.
- Noir S., Anthony F., Bertrand B., Combes M. C., et Lashermes P. (2003). Identification of a major gene (*Mex-1*) from *Coffea canephora* conferring resistance to *Meloidogyne exigua* in *Coffea arabica. Plant Pathol* 52: 97-103.
- Noir S., Combes M. C., Anthony F., et Lashermes P. (2001). Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (*Coffea* L.). *Mol Genet Genomics* 265: 654-662.
- Noir S., et Lashermes P. (2000). Organisation et évolution des gènes de résistance chez les plantes. *Cahiers Agricultures* 8: 301-309.
- Nürnberger T., et Brunner F. (2002). Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-assiociated molecular patterns. *Cur Opin Plant Biol* 5: 1-7.
- Orbach M. J., Farral L., Sweigaard J. A., Chumley F. G., et Valent B. (2000). A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. *Plant Cell* 12: 2019-2032.
- Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S., Zamir D., et Fluhr R. (1997). The *I2C1* family from the wilt disease resistance locus *I2* belong to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* 9: 521-532.
- Orth K., Xu Z., Mudgett M. B., Bao Z. Q., Palmer L. E., Bliska J. B., Mangel W. F., Staskawicz B. J., et Dixon J. E. (2000). Disruption of signaling by *Yersinia* effector YopJ, a ubiquitin-like protein protease. *Science* 290: 1594-1597.
- O'Sullivan D.M., Ripoll P.J., Rodgers M., et Edwards K.J. (2001). A maize bacterial artificial chromosome (BAC) library from the European flint inbred line F2. *Theor Appl Genet* **103**: 425-432.
- **Pan Q., Wendel J., et Fluhr R.** (2000a). Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *J Mol Evol* 50: 203-13.
- **Pare S.** (2002). Max Havelaar, expériences du commerce équitable du café. *In* "Recherche et caféiculture", CIRAD-CP, Montpellier, France, pp. 22-29.
- Parker J. E., Coleman M. J., Szabo V., Frost L. N., Schmidt R., van der Biezen E. A., Moores T., Dean C., Daniels M. J., et Jones J. D. (1997). The *Arabidopsis* downy mildew

resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with *N* and *L6*. *Plant Cell* 9: 879-894.

- Parker J. E., Holub E., Frost L. N., Falk A., Gunn N. D., et Daniels M. J. (1996). Characterisation of *eds1*, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* genes. *Plant Cell* 8: 2033-2046.
- Parniske M., Hammond-Kosack K. E., Golstein C., Thomas C. M., Jones D. A., Harrison K., Wulf B. B. H., et Jones J. D. G. (1997). Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* 91: 821-832.
- Parniske M., et Jones J. D. G. (1999). Recombination between diverged clusters of the tomato *Cf-9* plant disease resistance gene family. *Proc Natl Acad Sci USA* 96: 5850-5855.
- Patocchi A., Gianfranceschi L., et Gessler C. (1999). Towards the map-based cloning of *Vf*: fine and physical mapping of the *Vf* region. *Theor Appl Genet* 99: 1012-1017.
- Pflieger S., Lefebvre V., et Causse M. (2001). The candidate gene approach in plant genetics: a review. *Mol Breed* 7: 275-291.
- Prakash N., Combes, M. C., Sommana, N., et Lashermes, P. (2002). AFLP analysis of introgression in coffee cultivars (*Coffea arabica* L.) derived from a natural interspecific hybrid. *Euphytica* 124: 265-271.
- **Pryor T.** (1987). The origin and structure of fungal disease resistance genes in plants. *Trds Genet* 3: 157-61.
- **Pryor T., et Ellis J. G.** (1993). The genetic complexity of fungal resistance genes in plants. *Adv Plant Pathol* 10: 282-305.
- Quénéhervé P., Anthony F., et Topart P. (2002). *Meloidogyne arabicida*: another *Meloidogyne* species controlled by the *Mi* gene. *Nematropica*, sous-presse.
- **Rausher M. D.** (1996). Genetic analysis of coevolution between plants and their natural enemies. *Trends in Genetics* 12: 212-217.
- Richly E., Kurth J., et Leister D. (2002). Mode of amplification and reorganisation of resistance genes during recent *Arabidopsis thaliana* evolution. *Mol Biol Evol* 19: 76-84.
- Richter T. E., Prior T. J., Bennetzen J. L., et Hulbert S. H. (1995). New rust resistance specificities associated with recombination in the *Rp1* complex in maize. *Genetics* 141: 373-381.
- Richter T. E., et Ronald P. C. (2000). The evolution of disease resistance genes. *Plant Mol Biol* 42: 195-204.
- Riely B. K., et Martin G. B. (2001). Ancient origin of pathogen recognition specificity conferred by the tomato disease resistance gene *Pto. Proc Natl Acad Sci USA* 98: 2059-2064.
- Roberts P. A., Matthews W. C., et Veremis J. C. (1998). Genetic mechanisms of host-plant resistance to nematodes. *In* "Plant and Nematode Interactions. Agronomy" (Barker K. R., Pederson G. A., et Windham G. L., Eds), American Society of Agronomy, Madison, Wisconsin, USA, pp. 209-38.
- Rodrigues A. C. F. d. O., Abrantes I. M. d. O., Melillo M. T., et Bleve-Zacheo T. (2000). Ultrastructural response of coffee roots to root-knot nematodes, *Meloidogyne exigua* and *M. megadora. Nematropica* 30: 201-210.
- Ross H. (1986). Potato breeding-problems and perspectives. Z Planzenzüchtung.
- Rossi M., Goggin F. L., Milligan S. B., Kaloshian I., Ullman D. E., et Williamson V. M. (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proc Natl Acad Sci USA* 95: 9750-9754.
- Rouppe Van der Voort J. N. A. M., Wolters P., Folkertsma R., Hutten R., van Zanvoort P., Vinke H., Kanyuka K., Bendahmane A., Jacobsen E., Janssen G. J. H., et Bakker J.

(1997). Mapping of the cyst nematode resistance locus *Gpa2* in potato using a strategy based on co-migrating AFLP markers. *Theor Appl Genet* 95: 874-880.

- Rustérucci C., Aviv D. H., Holt III B. F., Dangl J. L., et Parker J. E. (2001). The disease resistance signaling components *EDS1* and *PAD4* are essential regulators of the cell death pathway controlled by *LSD1* in *Arabidopsis*. *Plant Cell* 13: 2211-2224.
- Saitou N., et Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425.
- Salmeron J. M., Barker S. J., Carland F. M., Mehta A. Y., et Staskawicz B. J. (1994). Tomato mutants altered in bacterial disease resistance provide evidence for a new locus controlling pathogen recognition. *Plant Cell* 6: 511-520.
- Salmeron J. M., Oldroyd G. E. D., Rommens C. M. T., Scofield S. R., Kim H. S., Lavelle D. T., Dahlbeck D., et Staskawicz B. J. (1996). Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* 86: 123-133.
- Sambrook J., Fritsch E.F., et Maniatis T. (1989). Molecular Cloning A Laboratory Manual, 2nd Edition. Cold Spring Habour Laboratory Press, New-York.
- San Miguel P., Tikhonov A., Jin Y. K., Motchoulskaia N., Zakharov D., Melake-Berhan A., Springer P. S., Edwards K. J., Lee M., Avramova Z., et Bennetzen J. L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274: 765-68.
- Saraste M., Sibbald P. R., et Wittinghofer A. (1990). The P-loop: A common motif in ATP- and GTP-binding proteins. *Trds Biochem Sci* 15: 430-434.
- Sasaki T. et Burr B. (2000). International rice genome sequencing project: the effort to completely sequence the rice genome. *Cur Opin Plant Biol* 3: 138-141.
- Schulze-Lefert P., et Vogel J. (2000). Closing the ranks to attack by powdery mildew. *Trds Plant Sci* 5: 343-348.
- Scofield S. R., Tobias C. M., Rathjen J. P., Chang J. H., Lavelle D. T., Michelmore R. W., et Staskawicz B. J. (1996). Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. *Science* 274: 2063-2065.
- Sera, T. (2001). Coffee genetic breeding at IAPAR. Crop Breed Appl Biotechnol 1: 179-199.
- Shen K. A., Meyers B. C., Nurul Islam Faridi M., Chin D. B., Stelly D. M., et Michelmore R.
   W. (1998). Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. *Mol Plant-Microbe Interact* 11: 815-823.
- Sijmons P. C., Atkinson H. J., et Wyss U. (1994). Parasitic strategies of root nematodes and associated host cell responses. *Ann Rev Phytopathol* 32: 235-259.
- Silvarolla M. B., Gonçalves W., et Lima M. M. A. (1998). Resistência do cafeeiro a nematoides V Reprodução de *Meloidogyne exigua* em cafeeiros derivados da hibridação de *Coffea arabica* com *C. canephora. Nematologia Brasileira* 22: 51-59.
- Simons G., Groenendijk J., Wijbrandi J., Reijans M., Groenen J., Diergaarde P., van der Lee T., Bleeker M., Onstenk J., de Both M., Haring M., Mes J., Conelissen B., Zabeau M., et Vos P. (1998). Dissection of the *Fusarium I2* gene cluster in tomato reveals six homologs and one active gene copy. *Plant Cell* 10: 1055-1068.
- Smith P. G. (1944). Embryo culture of a tomato species hybrid. *Proc Am Soc Hortic Sci* 44: 413-416.
- **Song K., Lu P., Tang K., et Osborn T. C.** (1995b). Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc Natl Acad Sci U S A* 92: 7719-7723.
- Song W. Y., Pi L. Y., Wang G. L., Gardner J., Holsten T., et Ronald P. C. (1997). Evolution of the rice *Xa21* disease resistance gene family. *Plant Cell* 9: 1279-1287.

- Song W. Y., Wang G. L., Chen L. L., Kim H. S., Pi L. Y., Holsten T., Gardner J., Wang B., Zhai W. X., Zhu L. H., Fauquet C., et Ronald, P.C. (1995a). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 270: 1804-1806.
- **Speulman E., Bouchez D., Holub E., et Beynon J. L.** (1998). Disease resistance gene homologs correlate with disease resistance loci of *Arabidopsis* thaliana. *Plant J* 14: 467-474.
- Spielmeyer W., Robertson M., Collins N., Leister D., Schulze-Lefert P., Seah S., Moullet O., et Lagudah E. S. (1998). A superfamily of disease resistance gene analogs is located on all homoeologous chromosome groups of wheat (*Triticum aestivum*). *Genome* 41: 782-788.
- Stahl E. A., et Bishop J. G. (2000). Plant-pathogen arms races at the molecular level. Cur Opin Plant Biol 3: 299-304.
- Stahl E. A., Dwyer G., Mauricio R., Kreitman M., et Bergelson J. (1999). Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* 400: 667-671.
- Staskawicz B. J., Ausubel F. M., Baker B. J., Ellis J. G., et Jones J. D. G. (1995). Molecular genetics of plant disease resistance. *Science* 268: 661-667.
- Staskawicz B. J., Mudgett, M. B., Dangl J. L., et Galan J. E. (2001). Common, and contrasting themes of plant and animal diseases. *Science* 292: 2285-2289.
- Stein N., Feuillet C., Wicker T., Schlagenhauf E., et Keller B. (2000). Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proc Natl Acad Sci USA* 97: 13436-13441.
- Sudupak M. A., Bennetzen J. L., et Hulbert S. H. (1993). Unequal exchange and meiotic instability of disease-resistance genes in *the Rp1* region of maize. *Genetics* 133: 119-125.
- Sun Q., Collins N., Ayliffe M., Smith S., Drake J., Pryor T., et Hulbert S. (2001). Recombination between paralogues at the *Rp1* rust resistance locus in maize. *Genetics* 158: 423-438.
- Swiderski M. R., et Innes R. W. (2001). The *Arabidopsis PBS1* resistance gene encodes a member of a novel protein kinase subfamily. *Plant J* 26: 101-112.
- Tai T., Dahlbeck D., Stall R. E., Peleman J., et Staskawicz B. J. (1999b). High-resolution genetic and physical mapping of the region containing the *Bs2* resistance gene of pepper. *Theor Appl Genet* 99: 1201-1206.
- Tai T. H., Dahlbeck D., Clark E. T., Gajiwala P., Pasion R., Whalen M. C., Stall R. E., et Staskawicz B. J. (1999a). Expression of the Bs2 pepper gene confers resistance to bacterial spot in tomato. Proc Natl Acad Sci USA 96: 14153-14158.
- Takken F. L., Thomas C. M., Joosten M. H. A. J., Golstein C., Westerink N., Hille N., Nijkamp H. J. J., de Wit P.J.G.M., et Jones, J. D. G. (1999). A second gene at the tomato *Cf-4* locus confers resistance to *Cladosporium fulvum* through recognition of a novel avirulence determinant. *Plant J* 20: 279-288.
- Tang X., Frederick R. D., Zhou J., Halterman D. A., Jia Y., et Martin G. B. (1996). Initiation of plant disease resistance by physical interaction of *AvrPto* and *Pto* kinase. *Science* 274: 2060-2063.
- Tanksley S. D., Ganal M. W., et Martin G. B. (1995). Chromosome landing: a paradigm for map-based cloning in plants with large genomes. *Trds Genet* 11: 63-68.
- Thomas C. M., Jones D. A., Parniske M., Harrison K., Balint-Kurti P. J., Hatzixanthis K., et Jones J. D. (1997). Characterization of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. *Plant Cell* 9: 2209-2224.
- **Thompson J. D., Higgins D. G., et Gibson T. J.** (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment trough sequence weighting, positions-specific gap penalties and weight matrix choice. *Nuc Acids Res* 22: 4673-4680.

- Tör M., Manning K., King G. J., Thompson A. J., Jones G. H., Seymour G. B., et Armstrong S. J. (2002). Genetic analysis and FISH mapping of the *Colourless non-ripening* locus of tomato. *Theor Appl Genet* 104: 165-170.
- **Traut T. W.** (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide binding-sites. *Eur J Bioch.* 222: 9-19.
- Triantaphyllou A. C. (1985). Cytogenetics, cytotaxonomy and phylogeny of root-knot nematodes. *In* "An advance treatise on *Meloidogyne*. Volume I: biology and control" (Sasser J. N., et Carter C. C., Eds), University Graphics, Raleigh, North Carolina State, pp. 113-126.
- **Trowsdale J.** (2002). The gentle art of gene arrangement: the meaning of gene clusters. *Genome Biology* 3: 2002.1-2002.5.
- **Trudgill D. L., et Block V. C.** (2001). Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. *Ann Rev Phytopathol* 39: 53-77.
- Van Daelen R. A. J. J., Gerbens F., van Ruissen F., Aarts J., Hontelez J., et Zabel P. (1993). Long-range physical maps of two loci (Aps-1 and GP79) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. *Plant Mol Biol* 23: 185-192.
- Van den Ackerveken G. F. J. M., Vossen P., et de Wit P. J. G. M. (1993). The AVR9 racespecific elicitor of *Cladosporium fulvum* is processed by engenous and plant proteases. *Plant Physiol* 103: 91-96.
- Van der Beek J. G., Pet G., et Lindhout P. (1994). Resistance to powdery mildew (*Oidium lycopersicum*) in *Lycopersicon hirsutum* is controlled by an incompletely-dominant gene *Ol-1* on chromosome 6. *Theor Appl Genet* 89: 467-473.
- Van der Biezen E. A., Freddie C. T., Kahn K., Parker J. E., et Jones J. D. G. (2002). *Arabidopsis RPP4* is a member of the *RPP5* multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. *Plant J* 29: 439-451.
- Van der Biezen E. A., et Jones J. D. G. (1998). The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Cur Biol* 8: R226-R227.
- Van der Hoorn R. A. L., de Wit P. J. G. M., et Joosten M. H. A. J. (2002). Balancing selection favors guarding resistance proteins. *Trds Plant Sci*7: 67-71.
- Van der Hoorn R. A. L., Kruij M., Roth R., Brandwagt B. F., Joosten M. H. A. J., et de Wit P. J. G. M. (2001). Intragenic recombination generated two distinct *Cf* genes that mediate AVR9 recognition in the natural population of *Lycopersicon pimpinellifolium*. *Proc Natl Acad Sci USA* 98: 10493-10498.
- Van der Peer Y., et de Wachter R. (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Applic Biosci* 10: 569-570.
- Van der Vossen E. A. G., Van der Voort J. R., Kanyuka K., Bendahmane A., Sandbrink, H., Baulcombe D., Bakker J., Stiekema W. J., et Klein-Lankhorst R. M. (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J* 23: 567-576.
- Van der Vossen H. A. M. (1985). Coffee selection breeding. *In* "Coffee: Botany, biochemistry and production of beans and beverage" (Clifford M. N., et Wilson K. C., Eds), Croom Helm, London, Sydney, pp. 48-97.
- Van der Vossen H. A. M. (1997). Quality aspects in arabica coffee breeding programmes in Africa. *In* "Proceedings of the 17th ASIC Colloquium", pp. 430-438, Nairobi.

- Van der Vossen H. A. M. (2001). Agronomy I: Coffee Breeding Practices. In "Coffee, Recent Developments" (Clarke R. J., et Vitzthum O. G., Eds), Backwell Science, Ames, Iowa, USA pp. 184-201.
- Van Eeden F., et St Johnston D. (1999). The polarisation of the anterior-posterior and dorsalventral axes during *Drosophila oogenesis*. *Cur Opin Genet Devt* 9: 396-404.
- Van Wordragen M. F., Weide R., Liharska T., Van der Steen A., Koornneef M., et Zabel P. (1994). Genetic and molecular organisation of the short arm and pericentrometric region of tomato chromosome 6. *Euphytica* 79: 169-174.
- Villain L., Baujard P., Anzueto F., Hernandez A., et Sarah J. L. (2002). Protection intégrée des caféières d'Amérique Centrale contre les nématodes. *In* "Recherche et Caféiculture", CIRAD, Montpellier, France, pp. 118-127.
- Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., et Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting. *Nuc Acids Res* 23: 4407-4414.
- Vos P., Simons G., Jesse T., J. W., Heinen L., Hogers R., Frijters A., Groenendijk J., Diergaarde P., Reijans M., Fierens-Onstenk J., de Both M., Peleman J., Liharska T., Hontelez J., et Zabeau M. (1998). The tomato *Mi-1* gene confers resistance to both rootknot nematodes and potato aphids. *Nat Biotechnol* 16: 1365-1369.
- Wang G. L., Ruan D. L., Song W. Y., Sideris S., Chen L., Pi L. Y., Zhang S., Zhang Z., Fauquet C., Gaut B. S., Whalen M. C., et Ronald P. C. (1998). Xa21D encodes a receptorlike molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. *Plant Cell* 10: 765-779.
- Wang Q., Zhang K., Qu X., Jia J., Shi J., Jin D., et Wang B. (2001). Construction and characterisation of a bacterial artificial chromosome library of peach. *Theor Appl Genet* **103**: 1174-1179.
- Wang Z. X., Yano M., Yamanouchi U., Iwamoto M., Monna L., Hayasaka H., Katayose Y., et Sasaki T. (1999). The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucin-rich repeat class of plant disease resistance genes. *Plant J* 19: 55-64.
- Warren R. F., Henk A., Mowery P., Holub E., et Innes R. W. (1998). A mutation within the leucine-rich repeat domain of the *Arabidopsis* disease resistance gene *RPS5* partially suppresses multiple bacterial resistance genes. *Plant Cell* 10: 1439-1452.
- Warren R. F., Meritt P. M., Holub E., et Innes R. W. (1999). Identification of three putative signal transduction genes involved in *R* gene-specified disease resistance in *Arabidopsis*. *Genet Soc America* 152: 401-12.
- Wei F., Gobelman-Werner K., Moroll S. M., Kurth J., Mao L., Wing R., Leister D., Schulze-Lefert P., et Wise R. P. (1999). The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR genes familes and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* 153: 1929-1948.
- Wendel J., Schnabel A., et Seelanan T. (1995). Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc Natl Acad Sci USA* 92: 280-284.
- Wendel J. F. (2000). Genome evolution in polyploids. Plant Mol Biol 42: 225-249.
- Werner J. E., Endo T. R., et Gill B. S. (1992). Toward a cytogenetically based physical map of the wheat genome. *Proc Natl Acad Sci USA* 89: 11307-11311.
- Wessler S. R., Bureau T. E., et White S. E. (1995). LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr Opin Genet Dev* 5: 814-821.
- Whitham S., McCormick S., et Baker B. (1996). The *N* gene of tobacco masaic virus in transgenic tomato. *Proc Natl Acad Sci USA* 93: 8776-8781.

- Williams C. E., Wang B., Holsten T. E., Scambray J., Dasilva F. D. G., et Ronald P. C. (1996). Markers for selection of the rice *Xa21* disease resistance gene. *Theor Appl Genet* 93: 1119-1122.
- Williamson V. M. (1998). Root-knot nematode resistance genes in tomato and their potential for future use. *Ann Rev Phytopathol* 36: 277-293.
- Williamson V. M. (1999). Plant nematode resistance genes. Curr Opin Plant Biol 2: 327-331.
- Williamson V. M., et Hussey R. S. (1996). Nematode pathogenesis and resistance in plants. *Plant Cell* 8: 1735-1745.
- Witsenboer H., Kesseli R., Fortin M., Stanghellini M., et Michelmore R. W. (1995). Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. *Theor Appl Genet* 91: 178-188.
- **Wyss U.** (1997). Root parasitic nematodes: an overview. *In* "Cellular and molecular aspects of plant-nematode interactions" (Fenoll C., Grundler F. M. W., et Ohl S., Eds), Kluwer Academic Publishers, Dordrecht, pp. 5-22.
- Wyss U., Grundler F. M. W., et Münch A. (1992). The parasitic behaviour of second stage juveniles of *Meloidogyne incognita* in roots of *Arabidopsis thaliana*. *Nematologica* 38: 98-111.
- Xiao S., Ellwood S., Calis O., Patrick E., Li T., Coleman M., et Turner J. G. (2001). Broadspectrum mildew resistance in *Arabidopsis thaliana* mediated by *RPW8*. *Science* 291: 118-120.
- Yaghoobi J., Kaloshian I., Wen Y., et Williamson V. M. (1995). Mapping of a new nematode resistance locus in *Lycopersicon peruvianum*. *Theor Appl Genet* 91: 457-464.
- Yoshimura S., Yamanouchi U., Katayose Y., Toki S., Wang Z.-I., Kono I., Kurata N., Yano M., Iwata N., et Sasaki T. (1998). Expression of Xa1, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. Proc Natl Acad sci USA 95: 1663-1668.
- Young L. D. (1998). Breeding for nematode resistance and tolerance. *In* "Plant and Nematode Interactions. Agronomy" (Barker K. R., Pederson G. A., et Windham G. L., Eds), American Society of Agronomy, Madison, Wisconsin, USA pp. 187-207.
- Young N. D. (2000). The genetic architecture of resistance. Curr Opin Plant Biol 3: 285-290.
- Yu Y., Tomkins J. P., Waugh R., Frisch D. A., Kudrna D., Kleinhofs A., Brueggeman R. S., Muehlbauer G. J., Wise R. P., et Wing R. A. (2000). A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor Appl Genet* 101: 1093-1099.
- Yu Y. G., Buss G. R., et Saghai Maroof M. A. (1996). Isolation of a superfamily of candidate disease-resistance genes in soybean based on conserved nucleotide-binding site. *Proc Natl Acad sci USA* 93: 11751-11756.
- Zamir D., Ekstein-Michelson I., Zakay Y., Navot N., Zeidan M., Sarfatti M., Eshed Y., Harel E., Pleban T., Van Hoss H., Kedar N., Rabinowithch H. D., et Czosneck H. (1994). Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, *Ty-1*. *Theor Appl Genet* 88: 141-146.
- Zhang S., et Klessig D. F. (2001). MAPK cascades in plant defense signaling. *Trds Plant Sci* 6: 520-527.
- Zhou F., Kurth J., Wei F., Elliott C., Vale G., Yahiaoui, N., Keller, B., Somerville, S., Wise, R. et Schulze-Lefert, P. (2000). Cell-autonomous expression of barley *Mla1* confers race specific resistance to the powdery mildew fungus via *Rar1* independent signaling pathway. *Plant Cell* 13: 337-350.

## Annexes

## Abréviations

ADN	Acide Désoxyribo-Nucléique
ADNc	Acide Désoxyribo-Nucléique complémentaire
AFLP	Amplification Fragment Length Polymorphism
ARNm	Acide Ribonucléique messager
<i>avr</i> /Avr	avirulence
BAC	Bacterial Artificial Chromosome
CATIE	Centro Agronomico Tropical de Investigacion y Enseñanza
CC	Coil-Coiled domain
CICAFE	Institut national de recherche sur le caféier au Costa Rica
сM	centimorgans
FISH	Fluorescence in situ Hybridisation
GISH	Genomic fluorescence in situ Hybridisation
HR	Hypersensitive Response
IRD	Institut de Recherche pour le Développement
LRR	Leucine-Rich Repeat
LZ	Leucine-Zipper
ORF	Open Reading Frame
NBS	Nucleotide Binding Site
pb/kb	paires de bases/kilo-paires de bases
PCR	Polymerase Chain Reaction
<i>R</i> /R	résistance
RFLP	Restriction Fragment Length Polymorphism
RGA	Resistance Gene Analogues
RT-PCR	Reverse Transcriptase PCR
SAR	Systemic Acquired Resistance
SCAR	Sequence-Characterised Amplified Region
TIR	Toll-Interleukin 2 Receptor
ТМ	Trans-Membranaire

## Figures et Tableaux

Par chapitre et ordre d'apparition dans le texte :

## **CHAPITRE 1**

- **Figure 1.1** Classification phylogénétique de 32 taxons de caféiers et distribution géographique des principaux groupes de la classification.
- Figure 1.2 Base génétique étroite des variétés de caféiers Arabica.
- **Figure 3.1** Représentation schématique de la mise en place de la résistance active selon 3 phases : reconnaissance, activation d'une cascade de signaux, expression des réactions de défense.
- Figure 3.2 Représentation synthétique des produits intervenant dans les voies de signalisation menant à l'activation des réponses de défense chez les mammifères (a), les insectes (b) et les plantes (c).
- **Figure 3.3** Représentation schématique de la localisation et de la structure de quelques principaux produits de gènes *R*.
- Figure 3.4 Modèles de reconnaissance spécifique entre les produits des gènes *R* et *avr.*
- **Figure 3.5** Représentation schématique de quelques points de convergence entre les voies de signalisation identifiées et impliquées dans l'expression de la résistance spécifique chez les plantes.
- **Figure 3.6** Représentation schématique des structures génétiques des locus *L* et *M* de résistance à la rouille chez le lin.
- **Figure 3.7** Différents types de recombinaison génétique à l'origine de la multiplication et de la diversification des gènes de résistance.
- Figure 4.1 Cycle de développement des nématodes du genre *Meloidogyne*.
- Figure 4.2 Plantation de caféiers *C. arabica* au Guatemala, dévastée par *M. paranaensis*.
- **Figure 4.3** Racines de caféiers infectées par *Meloidogyne exigua*, (a) plante sensible, (b) plante résistante.
- **Figure 4.4** Observations cytologiques de racines de caféier d'espèce *C. arabica* infectées par *M. exigua.*
- Figure 5.1 Lignées d'introgression *C. arabica*.
- Tableau 2.1
   Exemples de caractères d'intérêts agronomiques identifiés au sein d'espèces diploïdes de caféiers ou de formes sauvages d'Arabica.
- **Tableau 3.1**Interactions plante-hôte/agent pathogene, compatible et incompatible, selon le<br/>concept de la résistance gène-à-gène (Agrios, 1997).
- **Tableau 3. 2**Classification structurale des gènes R caractérisés.
- Tableau 4.1
   Locus de résistance spécifiques aux nématodes endoparasites des plantes cartographiés
- Tableau 4. 2
   Distribution géographique des nématodes du genre
   Meloidogyne parasites du caféier C. arabica



**Figure 1. 1** Classification phylogénétique de 32 taxons de caféiers, sur la base du polymorphisme de l'ADN nucléaire ribosomal et distribution géographique des principaux groupes de la classification. Les valeurs de *bootstrap* (%) sont indiquées sur les principales branches de l'arbre (Lashermes *et al.*, 1997).



**Figure 1.2** Représenation de la base génétique étroite des variétés cultivées de caféiers Arabica.

Caractères d'intérêt agronomique	Sources de diversité	Références	
Sources de résistance aux maladies			
- Perileucoptera coffeela	C. racemosa	Carvalho <i>et al.,</i> 1988	
- Hemileia vastatrix	C. canephora, C. liberica	Eskes, 1989	
- Colletotrichum kahawae	C. canephora	Van der Vossen, 1997	
- Meloidogyne exigua	C. canephora, C. racemosa	Bertrand <i>et al.</i> , 1999 ; Anthony <i>et al.</i> , 2002b	
- Meloidogyne paranaensis	Formes sauvages d'Arabica, <i>C. canephora</i>	Anthony et al., 2002b	
- Meloidogyne incognita	Formes sauvages d'Arabica, <i>C. canephora</i>	Anzueto <i>et al.</i> , 2001	
- Meloidogyne arabicida	Formes sauvages d'Arabica, <i>C. canephora</i>	Anthony <i>et al.</i> , 2002b	
Stérilité mâle	Formes sauvages d'Arabica	Dufour <i>et al.</i> , 1997	
Durée du cycle de maturation des fruits	C. congensis, C. canephora	Données non publiées	
Composition biochimique des graines	C. pseudozanguebariae	Mazzafera et Carvalho, 1992	

**Tableau 2. 1**Exemples de caractères d'intérêt agronomique identifiés au sein d'espècesdiploïdes de caféiers ou de formes sauvages d'Arabica.



**Figure 3. 1** Représentation schématique de la mise en place de la résistance active selon 3 phases : reconnaissance, activation d'une cascade de signaux, expression des réactions de défense. HR, réaction d'hypersensibilité ; LAR, résistance locale acquise ; SAR, résistance systémique acquise ; *avr*, avirulence ; *R*, résistance.

Nature de		Génes de la plante-hôte ²		
l'interaction		R	r	
Génes de	avr A	incompatible	compatible	
pathogène ¹	<i>avr</i> a	compatible	compatible	

¹ gène d'avirulence *avr* A, gène de virulence *avr* a. ² gènes de résistance dominant (R) ou récessif (r).

Tableau 3. 1 Interactions plante-hôte/agent pathogène, compatible et incompatible, selon le concept de la résistance gène-à-gène (Agrios, 1997).



**Figure 3. 2** Représentation synthétique des produits intervenant dans les voies de signalisation menant à l'activation des réponses de défense chez les mammifères (a), les insectes (b) et les plantes (c). Les voies de signalisation, représentées par les flèches, partagent des composés de structure commune, leurs noms respectifs sont précisés en italique (d'après Cohn *et al.*, 2001). NBS, nucleotide binding site ; LRR, leucine-rich repeat ; TIR, domaine homologue aux domaines intracellulaires des récepteurs protéiques isolés de la drosophile (récepteur Toll) et de l'homme (récepteur Interleukin-1) ; PK, protéine kinase.



**Figure 3. 3** Représentation schématique de la localisation et de la structure de quelques principaux produits de gènes R (d'après Dangl et Jones, 2001). Les principales régions conservées entre ces produits sont précisées sur la figure ; NBS, nucleotide binding site ; LRR, leucine-rich repeat ; TIR, domaine homologue aux domaines intracellulaires des récepteurs protéiques isolés de la drosophile (récepteur Toll) et de l'homme (récepteur Interleukin-1). Les produits de la classe NBS-LRR sont présumés être cytoplasmiques, toutefois, il n'est pas écarté que ceux-ci puissent être associés à la membrane. Certains gènes R caractéristiques des classes représentées sont précisés sur la figure. Tous les produits représentés sont orientés selon les extrémités N- et C- terminales précisées sur la figure.

Gène / famille de gènes	Interaction hôte/parasite	Structure protéique prédite	Références	
			Lowropco of al. 1005	
	Lin/ Melampsora lini		Anderson <i>et al.</i> , 1995	
			Whitham $et al$ 1996	
	Lin/Melamosora lini		Dodds <i>et al.</i> , 2000	
F RDD1	Arabidonsis/Peronospora		Botella <i>et al.</i> , 1998	
RPP5	Arabidopsis/Peronospora		Parker et al., 1997	
RPS4	Arabidopsis/Peronospora	TIR-NBS-LRR	Gassman et al., 1999	
Bs2	Poivron/Xanthomonas	NBS-LRR	Tai <i>et al.</i> , 1999b	
Dm3	Laitue/Bremia	NBS-LRR	Meyers <i>et al.</i> , 1998	
Gpa2/Rx1	Pomme de terre /Globodera	NBS-LRR	Van der Vossen et al.,	
0,000,000	Pomme de terre/PVX		2000 ;Bendahmane et al., 1999	
12	Tomate/Fusarium	NBS-LRR	Ori <i>et al.</i> , 1997 ; Simons <i>et al.</i> , 1998	
Mi	Tomate/ <i>Meloidogyne</i> /Macrosiphum	NBS-LRR	Milligan <i>et al.</i> , 1998	
Mla	Orge/ <i>Blumeria</i>	NBS-LRR	Zhou <i>et al.</i> , 2000	
Pib	Riz/Magnaporthe	NBS-LRR	Wang <i>et al.</i> , 1999	
Pi-ta	Riz/Magnaporthe	NBS-LRR	Bryan <i>et al.</i> , 2000	
Prf	Tomate/Pseudomonas	NBS-LRR	Salmeron <i>et al.</i> , 1996	
Rp1	Maïs/Puccinia	NBS-LRR	Collins <i>et al.</i> , 1999	
RPM1	Arabidopsis/Pseudomonas	NBS-LRR	Grant <i>et al.</i> , 1996	
RPP8/HRT	Arabidopsis/Pernospora	NBS-LRR	McDowell et al., 1998 ; Cooley	
	Arabidopsis/TCV (HRT)		<i>et al.</i> , 2000	
RPP13	Arabidopsis/Peronospora	NBS-LRR	Bittner-Eddy et al., 2000	
RPS2	Arabidopsis/Pseudomonas	NBS-LRR	Bent <i>et al.</i> , 1994 ; Mindrinos <i>et al.</i> , 1994	
RPS5	Arabidopsis/Pseudomonas	NBS-LRR	Warren <i>et al.</i> , 1998	
Rx2	Pomme de terre/PVX	NBS-LRR	Bendahmane <i>et al.</i> , 1999	
R1	Pomme de terre/Phytophtora	NBS-LRR	Ballvora <i>et al.</i> , 2002	
Sw-5	Tomate/ <i>Topovirus</i>	NBS-LRR	Brommonschenkel <i>et al.</i> , 2000	
Xa-1	Riz/Xanthomonas	NBS-LRR	Yoshimura et al., 1998	
Cf-2/5	Tomate/ Cladosporium	LRR-TM	Dixon <i>et al.</i> , 1998	
Cf-4/9	Tomate/ Cladosporium	LRR-TM	Jones <i>et al.</i> , 1994 ; Takken <i>et al.</i> , 1999 ; Thomas <i>et al.</i> , 1997	
HS1 ^{pro1}	Betterave/Heterodera	LRR ² -TM	Cai <i>et al.</i> , 1997	
Pto	Tomate/Pseudomonas	kinase	Martin <i>et al.</i> ,1993	
Xa21	Riz/Xanthomonas	LRR-TM-kinase	Song <i>et al.</i> , 1995	
Pbs1	Arabidopsis/Pseudomonas	Kinase ²	Swiderski et Innes, 2001	
Rpg1	Orge/ <i>Puccinia</i>	Kinase ²	Brueggeman et al., 2002	
Ve	Tomate/Verticillum	unique	Kawchuk, <i>et al.</i> , 2001	
Rpw8	Arabidopsis/Erysiphe	unique	Xiao <i>et al.</i> , 2001	
MIo ¹	Orge/Blumeria	Protéine membranaire	Devoto et al., 1997	
Rrs1-R ¹	Arabidopsis/Ralstonia	unique	Deslandes et al., 2002	
Hm1	Maïs/Cochliobolus	Toxine réductase	Johal et Briggs, 1992	

NBS, nucleotide binding site ; LRR, leucine-rich repeat ; TIR, domaine homologue aux domaines intracellulaires des récepteurs protéiques isolés de la drosophile (récepteur Toll) et de l'homme (récepteur Interleukin-1) ; TM, domaine trans-membranaire

¹*mlo* et *Rrs1-R* confèrent une résistance à l'état récessif. ²Ces domaines présentent des structures potentielles semblables mais ne sont pas homologues.

**Tableau 3.2** Classification structurale des gènes *R* caractérisés.



**Figure 3. 4** Modèles de reconnaissance spécifique entre les produits des gènes *R* et *avr* (d'après Bonas et Lahaye, 2002). (a) modèle récepteurligand, (b) modèle de garde, (c) résistance protéolyse-dépendante. R, produit du gène R ; Avr, produit du gène avr ; P, "cible de la pathogénie" ; X, cible de l'activité protéase.



**Figure 3. 5** Représentation schématique de quelques points de convergence entre les voies de signalisation identifiées et impliquées dans l'expression de la résistance spécifique chez les plantes (Glazebrook, 2001; Schulze-Lefert et Vogel, 2000; Feys et Parker, 2000). En vert, se distinguent les produits de gènes associés à la résistance, mis en évidence uniquement chez *Arabidopsis*, en bleu, chez l'orge et en rose, chez le tabac. SA, acide salicylique.



Figure 3.6 Représentation schématique des structures génétiques des locus L et M de résistance à la rouille chez le lin. Les rectangles correspondent aux régions codantes apparentées (Ellis et al., 1997). (a) Le locus L. Les 3 allèles, ici présentés L2, L6 et L10, ont été respectivement identifiés au sein des lignées de lin Stewart, Birio et

Bolley Golden. (b) Le locus M. 1. La spécificité M est présente au sein de lignées du lin Forge, 2. la spécificité M1 au sein de lignée du lin Williston Brown. Les positions relatives de M et M1 sont arbitraires.





**Figure 3. 7** Différents types de recombinaison génétique à l'origine de la multiplication et de la diversification des gènes de résistance (R).



- 1. Attraction et pénétration dans la racine des larves J2 (**a**).
- 2. Larves de 2^{ème} stade ayant rejoint leur site définitif.
- Début de formation de la galle et développement des larves en femelles (b) ou en mâles (c).
- Galle contenant une jeune femelle (d) et un mâle (e) au stade J4.
- 5. Galle contenant une femelle adulte (f) associée à une gangue gélatineuse (g) contenant les œufs pondus. De ces œufs, sortiront les larves J2 (a). Le mâle adulte (h) est libéré.

Figure 4.1 Cycle de développement des nématodes du genre *Meloidogyne*.

Chez ces nématodes, la taille moyenne des juvéniles (J2, stade infestant) est de 500  $\mu$ m de longueur et 15  $\mu$ m de largeur. Au stade adulte, les mâles sont filiformes et mesurent entre 1 et 2 mm de long et environ 20  $\mu$ m de large. Les femelles, sont quant à elles piriformes, elles mesurent environ 700  $\mu$ m de long et 400  $\mu$ m de large (Eisenback, 1998).

	Plante	Espèce végétale	Locus <i>R</i> (chromosome/ groupe de liaison) ¹	Nématodes	Références
Nématodes à galles	Tomate	Lycopersicum peruvianum	<i>Mi</i> (∨I)	Meloidogyne incognita M. javanica M. arenaria M. arabicida	Ho <i>et al.</i> , 1992 ; Van Wordragen <i>et al.</i> , 1994 ; Quénhérvé <i>et al.</i> , 2002
			<b>Mi3</b> (XII)	M. incognita M. javanica	Yaghoobi <i>et al.</i> , 1995
	Piment	Capsicuum annuum	<i>Me3</i> (LG ₈ P) <i>Me4</i>	M. incognita M. javanica M. arenaria	Djian-Caporalino <i>et</i> a <i>l.</i> , 2001
	Pomme de terre	Solanum bulbocastanum	<b>R</b> _{Mc1} (XI)	M. chitwoodi	Brown <i>et al.</i> , 1996
	Prunier	Prunus cerasifera	<b>Ma1</b> (7)	M. incognita M. javanica M. arenaria	Esmanjaud <i>et al.</i> , 1996 ; Lecouls <i>et al.</i> , 1999
Nématodes	Tomate	L. pimpinellifolium	<i>Hero</i> (IV)	Globodera rostochiensis	Ellis et Smith, 1971 ; Ganal <i>et al.</i> , 1995
à kystes	Pomme de terre	S. tuberosum spp. andigena	<i>H1</i> (∨)	G. rostochiensis patothypes Ro1 et Ro4	Janssen <i>et al.</i> , 1991
		S. spegazzinii	<i>Gro1</i> (∀II)	<i>G. rostochiensis</i> patothypes <i>Ro1</i> et <i>Ro5</i>	Barone <i>et al</i> , 1990 ; Ballvora <i>et al.</i> , 1995
E			Gpa (V) Gpa2 (XII)	<i>G. pallida</i> pathotypes <i>Pa</i> 2 et <i>Pa</i> 3	Kreike <i>et al.</i> , 1994 Rouppe van der Voort <i>et al</i> , 1997
		S. vernei	<i>GroVI</i> (∨)	<i>G. rostochiensis</i> patothype <i>Ro1</i>	Ross, 1986
	Betterave	Beta pastellaris B. procumbens	Hs1 ^{pat-1} (I) Hs1 ^{pro-1} (I)	Heterodera schachtii	Heller <i>et al.</i> , 1996 Cai <i>et al.</i> , 1997
	Soja	Glycine max	<b>Rhg₄</b> (A)	H. glycines race3	Lewers et al., 2001
	Blé	Triticium aestivum	Cre (IIB) Cre3 (IID)	H. avenae	Eastwood <i>et al.</i> , 1991 ; Lagudah <i>et</i> <i>al.</i> , 1997

¹Les localisations au niveau de chromosomes sont indiquées en chiffres romains.

**Tableau 4.1** Locus de résistance spécifiques aux nématodes endoparasites des plantes cartographiés.

Pays répertoriés
Brésil, Guatemala, République Dominicaine, Colombie, Salvador, Venezuela, Nicaragua, Bolivie, Pérou, Honduras, Costa Rica
Brésil, Tanzanie, Jamaïque, Venezuela, Guatemala, Côte d'Ivoire, Inde
Brésil, Tanzanie, République Démocratique du Congo, Inde, Guatemala, Salvador
Brésil, Tanzanie, Salvador, République Démocratique du Congo, Inde,
Costa Rica
Jamaïque, Salvador
Brésil, Guatemala
Hawaii
Cuba
Brésil
Kenya, République Démocratique du Congo
Tanzanie, Sao Tome
Tanzanie
Angola, Ouganda
Guatemala
République Démocratique du Congo
Inde

**Tableau 4. 2** Distribution géographique des nématodes du genre *Meloidogyne* parasites du caféier *C. arabica* (pour revues, Campos *et al.*, 1990 ; Carneiro *et al.*, 2000 ; Villain *et al.*, 2002).



**Figure 4. 2** Plantation de caféiers *C. arabica* au Guatemala, dévastée par *M. paranaensis* (photo., B. Bertrand).


**Figure 4.3** Observations cytologiques de racines de caféier d'espèce *C. arabica* infectées par *M. exigua* (grossissement X50, coloration à la fuschine acide). (a) Larve de deuxième stade, la tête figée au niveau du cylindre central d'une racine de Caturra. (b) Femelle sphérique (photo., A. Martinez et F. Anthony).



**Figure 4. 4** Racines de caféiers infectées par *Meloidogyne exigua*. (a) Plante sensible. (b) Plante résistante.



Figure 5. 1 Lignées d'introgression *C. arabica*.

- **Figure 1.** Phylogenetic tree for nucleotide RGA sequences isolated from Caturra (*C. arabica*) and IF181 (*C. canephora*) accessions.
- Figure 2. Majority consensus of Aa and Ab subfamily RGAs alignment.
- **Figure 3.** Neighbor-Joining tree based on alignment of amino acid sequences of representative coffee RGAs (1 to 4 members per family) and NBS domains of cloned R-genes.
- **Figure 4.** Multiple amino acid sequence alignment of coffee RGAs of a given family and the NBS domain of closely related R-genes.
- **Table 1.**Oligonucleotide primers used to amplify resistance gene analogs.
- **Table 2.**Characteristics of RGAs isolated from coffee.

	Conserved amino-	Primer name	Forward/	Primer sequence (5'-3') ^a				
	acid motif		Reverse					
		Ploop1	F	GGI GGI GTI GGI AAR ACN AC				
	P-loop	Ploop2	F	GGN GGN RTN GGN AAA ACA AC				
	(G G V/I/M G K T T)	Ploop3	F	GGN GGN RTN GGN AAG ACG AC				
Dogonorato		Ploop4	F	GGN GGN RTN GGN AAR ACT AC				
primer		Ploop5	F	GGN GGN RTN GGN AAR ACC AC				
printer		Ploop6	F	GGN GGN RTN GGN AAR ACA AC				
	Hydrophobic Domain (G L/F P L/F A L/V)	GLPL1 GLPL2 GLPL3 GLPL4	R R R R	IAR IGC IAR IGG IAR NCC CAH HGC NAA HGG HAA HCC CAA NGC CAA NGG CAA NCC CAG NGC NAG NGG NAG NCC				
	"GVGKTT" (P-loop)	Ploop-Cof	F	G GGG GTG GGG AAG ACG ACT C				
No- degenerate primer ^b	"DGLPLAL" (GLPL)	GLPL-Cof	R	AG GGC GAG GGG GAG GCC ATC				
•	"GIKSWVC"	Int-Cof1	F	T GGC ATA AAG AGT TGG GTT TG				
	"AGEEVP"	Int-Cof2	R	TGG TAC TTC TTC CCC TGC				

^a I=deoxy-inosine, R=A/G, H=A/C/T, N=A/T/G/C ^b Designed on the basis of first isolated coffee RGAs.

**Table 1.** Oligonucleotide primers used to amplify resistance gene analogs.

Primer	Number of	Diversity families of
combinations	isolated RGAs	isolated RGAs
Ploop1/GLPL1	2	А
Ploop2/GLPL2	4	Ε, Ι
Ploop1/GLPL3	4	G, H
Ploop2/GLPL1	2	A, D
Ploop1/GLPL4	2	B, C
Ploop4/GLPL3	2	E, H
Ploop5/GLP4	2	F
Ploop6/GLPL4	1	F
Ploop-Cof/GLPL-Cof	20	А
Int-Cof1/Int-Cof2	4	А

 Table 2. Characteristics of RGAs isolated from coffee.



**Figure 1.** Phylogenetic tree for nucleotide RGA sequences isolated from Caturra (*C. arabica*) and IF181 (*C. canephora*) accessions. This tree was constructed by the Neighbor-Joining method and Kimura's correction was applied. Numbers (%) on the branches represented the bootstrap values (for 500 iterations). Branches corresponding to RGA groups of high sequence identity are noted A to I.



**Figure 3.** Neighbor-Joining tree based on alignment of amino acid sequences of representative coffee RGAs (1 to 4 members per family) and NBS domains of cloned R-genes. Kimura's correction was applied. Coffee RGAs are marked by gray shaded. The dotted line distinguishes TIR and non-TIR type sequences.

### A- H family versus I2C-1

CrgaHl	GVGKTTLAQMVYEDLGVEVSLPTRALVCISEEYDPTRITKEILRQLGISFGES	DNLLSLQVKLRGGLTEKKFLLVLDDVWN
CrgaH2	GVGKTTLARMIYKDSRVDVSFPTRAWVCVSEGYDATRITKELLRELNISFVDS	DNLFSLQVKLQGGLTQKKFLLVLEDVWN
ArgaH3	GMGKTTLAQLVYNDKRVNNHFTTKAWVCVSEAYDATRITKELLRELEISFSDSG	ESLNSLQLKLQLGLTEKRFLLVLDDVWN
I2C-1	GMGKTTLAKAVYNDERVQKHFGLTAWFCVSEAYDAFRITKGLLQEIGSTDLKADDNLNQLQVKL	KADDNLNQLQVKLKEKLNGKR <mark>FLVVLDDVW</mark> N
		· · · · · · · · · · · · · · · · · · ·
CrgaHl	DNYNDWDNLRTPFKGGSRGSKIIVTTRNQQVARMMAEERSIHHLDPMLEEDCRSLFKKHAFENR	DGNENAELEEIGNKIVTKCRGLPLAL
CrgaH2	NNYNQWDNLRSPFNVGSCGSKIIITTRDQNVARMMAKERSIHHLNFIQEEDCRSLFKKHAFENK	DGNENAKLEEIGKKIVKKCGGLPLAL
ArgaH3	RDYDDWDKLKMMLKGGSEGSKIIVTTRDNRIALMSIHHLDLISEEDSWSLFEKHAFRDK	DNENWRELEVIGKKIVNKCEGLPLDF
I2C-1	DNYPEWDDLRNLFLQGDI <mark>GSKIIVTTRK</mark> ESVALMMDSGAIYMG <u>ILSSEDSWALFKRHSLEHK</u>	DPKEHPEFEEVGKQIADK <u>CKGLPLAL</u>

#### **B**- I family versus *Mi1-2*

CrgaI3 ArgaI1 ArgaI2 Mi1-2	GVGKTTLAKK GVGKTTLAKK GYRETTLAKK <u>GSGKTT</u> LAYK :	VYNDSS VYNDSS VYNDSS VYNDKS	VICNFI VICNFI IIYNF( VSR <u>HFI</u> : :	HIRLWC HIRLWC QIRLWC DLRAWC .:	TVSPE TVSPE TLSQA TVDOG	FNTKSI FNTKSI FNMKN YDDKKI	LLIQI LLIQI VLLQI LLDTI :	LCSNGK LCSNGK LCSEGK FSQVS- :	QSRMD QSRMD HSRMD G	EELKN EELKN DELKN SDSNL	LNEHE LNEHE LDEHV SENII ::	CLLHKI CLLHKI VLLEKI OVADKI : .	YQRLH YQRLH YQKLH RKQLH ::	CTKRY CTKRY CENRY GKRY :	LVVF LVVF LVVF LIVI : :	DDVW DDVW DDVW DDVW	DIKVV DIKVV DIKVV DTTTI	INELRI INELRI INELRI JDELTR :
CrgaI3	SFPDEKKGSR	IIFTSR	SSNVAS	SQVEFG	GKPHN	LVPLSI	EKESFI	ELLLKK	VFGNE	DCPQA	LHGLO	MEIAK	KCRGI	<b>JPFAM</b>				
ArgaIl	SFPDEKKGSR	IIFTSR	SSNVAS	SQVEFG	GKPHN	LVPLSI	EKESFI	ELLLKK	VFGNE	DCPQA	LHGLO	GMEIAK	KCRGI	PFAM				
ArgaI2	SFPDDKKGSR	IIFTSR	SSNVAS	SQVEYG	GKSHN	LRPLSI	EKESFI	ELLLKK	VFGKG	DCPQA	LHGLO	GMEIAK	KCMGI	PLAM				
Mil-2	PFPEAKKGSR	IILTTR	EKEVA	LHGKLN	TDPLD	LRLLRI	PDESWI	ELLDKR	TFGNE	SCPDE	LLDVG	GKEIAE	NCKGI	'PLVA				
	. :	: :	:	: :	:		. :	:	. :	. :	.:	:	:	: :.				

**Figure 4.** Multiple amino acid sequence alignment of coffee RGAs of a same family and the NBS domain of closely related R-genes. Strict consensus residues are shown by shading. Residues sharing high (;) or low (.) physico-chemical properties are specified. Sequence blocks marked with boxes correspond to conserved motifs of NBS-encoding region of related R-gene sequences defined by Meyers et al. (1999).

ConsAb ConsAa	G G <b>G</b> G GG	G G <b>T</b> C G GTC	g GGG g GGG	B <b>A</b> AG B AAG	ас <b>о</b> Ас <b>л</b>	ACI ACI	<b>CT</b> CT	r GCI r GCI	' AAA ' AAA	TCA TCA	GTC GTC	TAC TAC	<b>A</b> AC AAC		ACA CNA	A AAA A AAA	A ATT A ATT	GAI GAI	r gag r gag	F AAT F AAT	TTT TTT	GGC GGC	C ATA	AAG AAG	AGT ; AGT	TGG TGG	GTT GTT	TG <b>T</b> TGT
				Ploo	p-Co	f																		Int-	Cof1			
ConsAb	G <b>T</b> G	G <b>C</b> T	AG <b>A</b>	GAA	ATT	AAA	<b>a</b> t <b>a</b>	GTG	GA <b>G</b>	CTG	TTC	AAA	<b>C</b> TC	ATT	TTA	GAA	TCG	TTG	ACA	AG <b>A</b>	ACA	AAG	GTT	GAA	G <b>T</b> G	GAT (	GGT Z	AGG
ConsAa	GTG	GCT	AGA	AAA	ATT	GAT	Aca	GTG	GGG	CTG	TTC	AAA	CTC	ATT	TTA	GAA	TCG	TTG	ACA	AAA	ACA	AAG	GTT	GAA	GTG	GAT (		AGG
ConsAb	GAT	GCC	АТА	<b>G</b> TT	C <b>AA</b>	GAA	ATT	CGA	GGA	<b>a</b> aa	C <b>T</b> T	GGG	GA <b>A</b>	ааа	<b>A</b> GA	<b>T</b> TT	CTC	CTT	GTT	<b>C</b> T <b>T</b>	GA <b>T</b>	GAT	GTG	TGG	aat	TGT (	GAA (	CAA
ConsAa	GAA	GCC	АТА	GTT	CAA	GAA	<b>A</b> TT	CGA	GGA	aag	CTC	AAG	GGA	Саа	AGA	TAT	TTC	CTT	GTT	CTT	GAT	GAT	GTG	TGG	gat	CAT (	GA <b>T</b> (	CAA
ConsAb	ga <b>a</b>	IT <b>C</b>	TGG	AGT	GAC	TTT	TTC	ACC	acg	TTG	TI <b>G</b>	GGA	CTC	AG <b>C</b>	A <b>C</b> A	<b>a</b> ct	ААА	GGA	AGC	TGG	TGT	GCT	CTC	ACT	AC <mark>G</mark>	CGT (	CT <b>A (</b>	<b>G</b> AA
ConsAa	gga	TCG	TGG	GA <b>T</b>	GAC	TAT	TTC	AAC	act	TTG	ATG	GGA	CTC	AAC	GAA	acc	ААА	GGA	AGC	TGG	TGT	CTT	CTC	ACT	ACT	CGT (	CTA (	GAA
ConsAb	CCA	GTG	GCT	аа <b>т</b>	GCT	GTG	CC <b>T</b>	AGA	C <b>A</b> T	<b>CTG</b>	CAA	<b>A</b> TG	AAT	GAT	<b>G</b> GT	CCT	ТАТ	<b>TT</b> C	CT <b>A</b>	GGA	AAG	C <b>TA</b>	TCA	GA <b>T</b>	GAT	GCA :	rgc '	TGG
ConsAa	<b>TCA</b>	GTG	GCT	аат	GCT	GTG	CCT	AGA	CAT	TTG	CAA	ATG	AAT	GAT	CGT	CCT	ТАТ	TTC	TTA	GGA	AAG	CTA	TCA	GGT	GAT	GAG :	rgc '	TGG
ConsAb ConsAa	TCT TCC	AT <b>C</b> ATC	СТ <b>А</b> АТА		<b>G</b> GA GGA	AAA AAG	CT <b>G</b> GTG	<b>G</b> TA ATG	GTT AGT	GCA GCA	GGG GGG	G <b>A</b> A GAA Int-Co	GAA GAA of2	GTA GTA	CCA CCA	aa <b>t</b> gaa	GAA GAA	TTG TTG	GAA GAA	GCA GC A	TTG CTG	AAG AAG	AAG GAG	C <b>A</b> A CAA	A <b>TT</b> ATT	TT <b>A</b> 2 TTA 2 *	A <b>AA</b> AGG	A <b>A</b> A AGA *
ConsAb ConsAa	TG <b>T</b> TGC *	GAT GAT	GGC GGC <b>G</b>	CTC CTC LPL-C	ccc ccc	CTC NTC	GCC GCC	<b>C</b> T CT																				

**Figure 2.** Majority consensus of Aa and Ab subfamily RGAs alignment. Autapomorphies are indicated in bold type, homoplasies by shading and synapomorphies by framing. Mutations shared between RGAs not belonging to the same subfamily are specified by an asterisk. Non-degenerate primer pairs are indicated by arrows.

- **Figure 1.** Distribution of seedlings from: a- the susceptible line Et6 (34 individuals), b- the resistant line T5296 (16 individuals) and, c- the  $F_2$  progeny (93 individuals) derived from a T5296 x Et6 cross according to their gall index (GI) in response to *M. exigua* inoculation.
- **Figure 2.** Example of AFLP markers (Exi-7 and Exi-10) found to be linked with the resistance to *M. exigua* based on  $F_2$  (T5296 x Et6) individuals analysis.
- **Figure 3.** Genetic map of the chromosome region containing the *Mex-*1 gene for resistance to *M. exigua* based on the  $F_2$  (T5296 x Et6) population.
- **Table 1.** Screening of AFLP markers associated with the resistance to *M. exigua*.
- Table 2.
   List of AFLP markers associated with the resistance to *M. exigua*.
- Table 3.Survey of AFLP markers in some advanced introgressed lines of *C. arabica*<br/>derived from the Timor Hybrid (TH) either resistance (R) or susceptible (S) to<br/>*Meloidogyne exigua* in field conditions.



**Figure 1.** Distribution of seedlings from: a- the susceptible line Et6 (34 individuals), b- the resistant line T5296 (16 individuals) and, c- the  $F_2$  progeny (93 individuals) derived from a T5296 x Et6 cross according to their gall index (GI) in response to *M. exigua* inoculation.

Screening	No. of primer	No. of polymorphic	No. of AFLP markers
Material ^a	combinations	AFLP bands	associated with the
			resistance to <i>M. exigua</i>
Timor hybrid-	232	403	10
derived			
genotypes			
F2 individuals	110	161	4
Total	342	564	14

^a Screening material included either 2 resistant versus 2 susceptible Timor hybrid-derived genotypes or 2 resistant versus 2 susceptible  $F_2$  (T5296 x Et6) individuals.

Table 1. Screening of AFLP markers as	ssociated with the	resistance to M.	. exigua.
---------------------------------------	--------------------	------------------	-----------



**Figure 2.** Example of AFLP markers (Exi-7 and Exi-10) found to be linked with the resistance to *M. exigua* based on  $F_2$  (T5296 x Et6) individuals analysis.

Code	Primer	Size of the
	combinations	informative band (bp)
Exi-1	E-AGC/M-CAA	201
Exi-2	E-AGG/M-CTT	141
Exi-3	E-CAT/M-CTT	109
Exi-4	E-CAT/M-AAC	124
Exi-5	E-ACG/M-ACC	247
Exi-6	E-AGC/M-AGT	100
Exi-7	E-AGC/M-ACA	77
Exi-8	E-AAG/M-ACG	110
Exi-9	E-AAG/M-CGT	280
Exi-10	E-ACT/M-AAC	110
Exi-11	E-CAC/M-CTA	194
Exi-12	E-CAA/M-AGG	120
Exi-13	E-CAC/M-CAT	114
Exi-14	E-CTA/M-ACA	41

**Table 2.** List of AFLP markers associated with the resistance to *M. exigua*.



**Figure 3.** Genetic map of the chromosome region containing the *Mex*-1 gene for resistance to *M. exigua* based on the  $F_2$  (T5296 x Et6) population.

N°	Code	TH Parent ^a	R/S					C	rde	red	mar	ker	s ^b				
				Exi-	Exi-1	Exi-4	Exi-5	Exi-6	Exi-7	Exi-8	Exi-9	Exi-	Exi-	Exi-	Exi-	Exi-3	Exi-2-
1	CR 95 cv	1	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	T17928	3	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	T17929	3	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	T18121	1	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	T18122	1	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	T18123	3	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	T18127	1	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8	T18137	2	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	T18140	2	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	T5175	1	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11	T8666	1	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12	T18139	2	S	-	-	-	-	-	-	-	-	-	-	-	-	+	+
13	T17935	3	R	-	-	-	-	-	-	-	-	-	-	-	+	-	-
14	T17940	3	R	-	-	-	-	-	-	-	-	-	-	-	+	-	-
15	IAPAR cv	2	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16	T16784	2	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
17	T17924	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
18	T17925	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
19	T17926	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
20	T17927	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
21	T17930	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
22	T17933	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
23	T17936	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
24	T17937	3	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
25	T18126	3	S	+	+	+	+	+	+	+	+	+	+	+	+	+	+
26	T18130	1	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
27	T18138	2	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
28	T5296	2	R	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Con	trol :																
29	IF 200 cv	C. canephora	R	-	+	+	+	+	+	+	+	+	+	+	+	+	+
30	Caturra cv	C. arabica	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	Catuaí cv	C. arabica	S	+	-	-	-	-	-	-	-	-	-	-	-	-	-

^a When known the Timor Hybrid-derived progenitor used in the introgression programme is indicated with  $1 = CIFC832/1 \ 2 = CIFC832/2 \ and \ 3 = CIFC1343$ 

^b Markers order deduced from the F₂ (T5296 x Et6) analysis; "+" indicate presence of AFLP markers while "-" indicate absence of AFLP markers

**Table 3.** Survey of AFLP markers in some advanced introgressed lines of *C. arabica* derived from the Timor Hybrid (TH) either resistance (R) or susceptible (S) to *Meloidogyne exigua* in field conditions.

- Figure 1. Pulsed-field gel electrophoresis (PFGE) analysis of BAC clones.
- **Figure 2.** Insert size distribution of Arabica BAC clones.
- Figure 3. Colony hybridisation of BAC filters.
- Table 1.
   Composition of the Arabica BAC library
- **Table 2.**BAC library filter-hybridisation results using 11 low-copy Arabica RFLP mapping<br/>probes anchored to the 11 linkage groups of the genome of *C. canephora*.

Ligation	Size selection	Number of	Mean insert	Proportion (%) of the
reaction	range (kb)	clones picked	size (kb)	whole library size
1	100-150	34 944	90	26 %
2	150-200	23 040	130	25 %
3	200-250	13 824	170	20 %
4	250-350	16 896	200	29 %

Table 1. Composition of the Arabica BAC library



**Figure 1.** Pulsed-field gel electrophoresis (PFGE) analysis of BAC clones. Ethidium bromide stained gel of BAC DNA digested with *Not*I and separated in a 1% agarose gel. Lanes 1-8, BAC DNA from the ligation 1, lanes 9-15 BAC DNA from the ligation 2, lanes 16-23 BAC DNA from the ligation 3, lanes 24-31 BAC DNA from the ligation 4.  $\lambda$  is a standard ladder (PFG marker, Biolabs).



**Figure 2.** Insert size distribution of Arabica BAC clones. The data are based on 23 and 24 randomly selected clones derived from the ligation reaction 3 (in white) and 4 (in grey), respectively

Probe	Linkage group	No. of loci in	No. of loci in	No. of hits among	Estimated number of
11000	in <i>C. canephora</i>	C. canephora	C. arabica ^a	27648 BACs	positive clones in the
				tested	whole library ^b
gA 67	1	1	2	8	18
gR19	2	1	1 or 2	2	5
gA 14	3	1	1 or 2	8	18
cR167	4	2	2	2	5
gA12	5	1	2	8	7
gA 11	6	1	1 or 2	3	18
gA 61	7	1	1 or 2	12	27
gA 6	8	1	2	9	20
gA 1	9	1	2	8	18
gA 13	10	1	1 or 2	6	14
g	11	1	1 or 2	3	7
R 1210					

^a See text for more information.

^b Taking into account that the fraction (i.e. 27648) of BACs analysed represents approximately 44% of the overall BAC library size.

**Table 2.** BAC library filter-hybridisation results using 11 low-copy Arabica RFLP mapping probes anchored to the 11 linkage groups of the genome of *C. canephora*.



**Figure 3.** Colony hybridisation of BAC filters: 27648 clones were double spotted onto nylon membrane and hybridised with nuclear single-copy probe gA 6.

- Figure 1. High-density filter screened with the Exi-3 probe.
- Figure 2. *Hind*III-fingerprinting gel of digested Exi-3 hybridising BAC clones.
- Figure 3. Distribution of Exi-3 positive BAC clones in 2 contigs, namely Exi-3 I and Exi-3 II.
- Figure 4. Collinearity between the two Exi-3 homoeologous contigs.
- Figure 5. Example of BAC clones constituting the *Mex-1* region, containing RGAs.
- Table 1.Positive BAC clones identified by either BAC filter hybridisation or BAC pool<br/>PCR-based screening.
- **Table 2.**Exi-2, Exi-3 and Exi-4 BAC clones containing RGA sequences.

Markers	Screening method	Number of positive clones	Identification (Exi code-clone co-ordinates)
Exi-2	- Hybridisation	5	2A-39C8; 2B-135L12; 2C-49B24; 2D-66A19; 2E-54L11
	- PCR		
Exi-3	- Hybridisation	11	3A-65K19; 3B-137O10; 3C-139G14;3D-127J15; 3E-
	- PCR		49P15; 3F-49P15; 3G-36A15; 3H-70O24; 3I-122F16; 3J-36L24; 3K-51P2
Exi-4	- PCR	2	4A-41D15; 4B-66P9

**Table 1.** Positive BAC clones identified by either BAC filter hybridisation or BAC pool PCR-based screening. The probes and SCAR primers were derived from cloned AFLP fragments corresponding to the Exi-2, Exi-3 and Exi-4 markers.



**Figure 1.** High-density filter screened with the Exi-3 probe. Eleven positive clones were detected among the 27 648 doublespotted clones. Arrows show the positive clones.



**Figure 2.** *Hind*III-fingerprinting gel of digested Exi-3 hybridising BAC clones. The BAC clone is identified at the top of each lane. M, DNA size marker (*Raoul*TM, Q-BIOgene).



**Figure 3.** Distribution of Exi-3 positive BAC clones in 2 contigs, namely Exi-3 I and Exi-3 II. BAC DNAs of Exi-3 positive clones were digested with *Hind*III for Southern blotting and hybridised with the Exi-3 probe.



**Figure 4.** Colinearity between the two Exi-3 homeologous contigs. The BAC clones were assembled into contigs based on BAC DNA fingerprinting. They are drawn to the scale. PCR sites of BAC-end and Exi-3 primers are indicated as full vertical bars, while RFLP markers are indicated by dotted vertical bars. Horizontal bracket over marker designations indicates markers with undetermined position relative to each other.

	BAC clones associated to <i>Mex-1</i> carrying RGA sequences
PCR-based research	
5'-3' primer sequence (reverse/forward)	
GGNGGNRTNGGNAARACTAC / CAANGCCAANGGCAANCC	2C-49B24
GGNGGNRTNGGNAARACTAC / CAGNGCNAGNGGNAGNCC	3B-137O10; 3D-127J15; 3F-49P15; 3J-36L24
GGNGGNRTNGGNAARACCAC / IARIGCIARIGGIARNCC	2B-135L12; 3A-65K19; 3B-137O10
Southern blotting	
(Coffee family RGA probes)	
RGA F	2A-39C8; 2B-135L12
RGA B	4B-66P9

Table 2. Exi-2, Exi-3 and Exi-4 BAC clones containing RGA sequences.

a <i>- Hi</i>	ndIII		b- EcoRI									
2A	2B	ЗП	9 0 0	4A	4B		2A	2B	3Е	30	4A	4B
		100	1999				137					-
												1
100												
				•								

**Figure 5.** Example of BAC clones constituting the *Mex-1* region, containing RGAs. BAC DNAs of Exi-2, Exi-3 and Exi-4 positive clones were digested with *Hind*III (a) and *EcoR*I (b) for Southern blotting and hybridised with a RGA sequence (AJ298889) member of the F coffee RGA family.

- **Figure 1.1** Séquence peptidique déduite de la séquence NBS-LRR pleine longueur isolée par criblage d'une banque d'ADNc de *C. arabica* à partir d'une sonde RGA de caféier, spécifique de la famille E.
- **Tableau 1.1**Populations  $F_2$  en ségrégation pour la résistance à *M. exigua* développées au<br/>CATIE (Costa Rica).

MAETVLSFVLDQLSIFLREEGRLLGGLHQEVQLISDELGHMRAFVGVAETIEEGADPRI	LQ.	60	1
EWIKQVREAAYDTEDVLDEFVACFAHHHATGFYGSVRKIFNSIKTLRARRKVAAQIQS	ΓK	120	
ARVKNISEGHQRYQSEFGGTTQAAESLAAVNNTTWRYSRDDALLVEEAELVGIDHPKQQ	)L	180	
${\tt ISQLLERDDSQLKVVSVV} \textbf{GMGGLGKTTL} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt QPCNFLE} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt QPCNFLE} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt QPCNFLE} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt QPCNFLE} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt QPCNFLE} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt VS} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt PVRAFVT} \textbf{VS} {\tt VS} {\tt VKKVHEDLAIRRH} \textbf{VS} {\tt VKKVHEDLAIRRH} \textbf{F} {\tt VKKVHEDLAIRRH} \textbf{VS} {\tt VS} {\tt$	240	NBS	
KDLTRQLHNELKKPVPESIEAMTAYQLKLCVKDFLQQPVRYAIVFDDVWDVEFWNEIR	300		
LPENGY GNRVMLTTRK TDVASASCNKSQDYVFKMAPLSFEDSWTLFCSKIFKENGCPA	360		
IDVAKGILGK <b>CEGLPLAV</b> LAISGLLALKDLNIAEEWEMVRRSLGGELEGCGMLDRVKK	420		
<b>SLSYNDL</b> PCH <b>LKTC</b> LLYLSIYPEDFEIRCGRLVQL <b>W</b> SAERFVGKKEGMTVEDVGYN		476	
YLRE LVNRS LIQVTQI FYEGI PYAC		501	1
RIHDLVREIVLSKAREQNMIAITTGQ		527	
YTKWLYEKVRRLVVHSSSNNTEQHQES		542	
QCYS FNHLRSF ITIE	557		
SMNPLVSRLLSEV		570	
LKSGRLLKVLD		581	1
LCDEKTLEEL		591	
PNEIFNLFRLRHLNLCRTGVKAV	614		
PKFIGKLRKLEYLNLGETQVKEL	637		
PVEILKLQKLEHLIAYKKVDSSDYSQGYHG	667		
FKAPSKLGGLLALQILDTIDASSGSVV		694	LRR
VKGIGKLTQLRRLQISNLRREDG	717		
KELCSSLATLTRLWELNIASIRNDDADYEVMDLNYHDQQQHSHSSS		763	
MPSTTFLQSLRMLILYGRLEKM	785		
PQWIAHLESLVRIALCWSRL		805	
RVEEDP LAP LHHLPNLVTIQF IGSYQGEGLCFK		838	
AGGFPKLKDLCLEKLEKLKWLKVE		862	
DGVLPNLQELYPDTLPLLEELPLG		886	
IQH <i>SRNLRKLYLS</i> ELSSQLMEKLENLNEETEDYRKIAHI		925	
SEVVTVLWTDEEGWRLHRLWGKKM		949	

**Figure 1. 1** Séquence peptique déduite d'un gène de la classe NBS-LRR isolé de *C. arabica.* La séquence nucléique pleine longueur a été isolée par criblage d'une banque d'ADNc de *C. arabica* à partir d'une sonde RGA de caféier, spécifique de la famille E. Les motifs conservés caractéristiques des domaines NBS sont précisés en gras et ceux caractéristiques des séquences non-TIR NBS, en gras-soulignés. Le domaine LRR potentiel présente 14 répétitions imparfaites du motif consensus xLxxLxxLx, surligné en gris.

Population F ₂	Parent résistant		Parent sensible
(nombre d'individus évalués)	Lignée introgressée Arabica		C. arabica
50	T5296	х	Et5
50	T5296	Х	Et32b
530	IAPAR59	Х	Java

**Tableau 1.1** Populations  $F_2$  en ségrégation pour la résistance à *M. exigua* développées au CATIE (Costa Rica).

#### Diversité des gènes de résistance au sein du génome des caféiers (Coffea L.) Analyse génétique de la résistance au nématode à galles, Meloidogyne exigua, chez C. arabica

**Résumé** La pression phytoparasitaire exercée par les ravageurs sur les plantes cultivées à l'échelle mondiale, représente l'une des contraintes majeures de l'agriculture actuelle. Le développement et l'utilisation de variétés résistantes pour combattre les phytoparasites sont devenus incontournables pour la protection des cultures tout en respectant l'environnement et les exigences des consommateurs. Dans le cadre de la caféiculture, cette pression s'exerce tout particulièrement sur les variétés *C. arabica* dont les plus cultivées se sont avérées sensibles à la plupart des parasites du caféier. Les nématodes phytopathogènes sont notamment responsables d'importants dommages dans les caféières de la plupart des pays d'Amérique Latine. Les objectifs de ce présent travail consistent, d'une part, en l'analyse globale et en l'amélioration des connaissances des gènes de résistance au sein du genre *Coffea*, et d'autre part, dans le contexte de l'interaction *C. arabica/M. exigua*, en l'identification et la cartographie du gène de résistance au nématode à galles, *M. exigua*, chez le caféier.

La diversité et le mode d'évolution des gènes de résistance chez le caféier ont été analysés, à partir de séquences analogues de gènes de résistance (RGA) de type NBS identifiées au sein du génome des caféiers. Parallèlement à cette approche visant l'ensemble des gènes *R* chez le caféier, les connaissances relatives au modèle d'interaction spécifique *C. arabical M. exigua*, ont été développées. Le déterminisme génétique de cette résistance a été précisé et la cartographie génétique du locus de résistance au nématode *M. exigua, Mex-1*, a été entreprise. La construction d'une banque génomique de grands fragments est apparue indispensable tant pour l'étude du locus de résistance *Mex-1* que pour l'analyse de l'ensemble des gènes *R* chez le caféier. A partir de la lignée introgressée Arabica, IAPAR 59, résistante à *M. exigua*, une banque BAC a été construite et caractérisée. Cet outil a permis d'entreprendre le développement d'une carte physique du locus *Mex-1*. Les premières constructions de contigs associés à *Mex-1* et la mise en évidence de séquences RGA au niveau de ces contigs sont rapportées.

**Mots-clefs** caféier (*Coffea* spp.), gènes de résistance, marqueurs moléculaire, banque BAC, cartographie physique, introgression, polyploïdie.

#### Resistance gene diversity in the coffee tree genome (*Coffea* L.) Genetic analysis of the resistance to the root-knot nematode, *M. exigua* in *C. arabica*

**Abstract** The world parasite pressure on crops is one of the main drawbacks in the agriculture. In the fight against phytopathogens, the development of resistant varieties and their use are become an obligatory strategy for a sustainable crop protection and in the same time, the environment respect and the consumer's satisfaction. In the coffee-culture, this pressure is particularly important on the *C. arabica* species. The most cultivated Arabica coffee trees are susceptible to the majority of coffee tree pathogens. Nematodes are namely responsible for great damages in the growing coffee-tree areas in Latin America. The aims of the present work are, on one hand, the whole analysis of resistance genes in the *Coffea* genus, and on the other hand, in the study of the specific interaction *C. arabica/M. exigua*, the identification and the mapping of the M. exigua root-knot nematode resistance gene in the coffee tree.

In the coffee tree, diversity and evolution of resistance genes have been studied on the base of NBS-like resistance gene analogs (RGA) isolated from the coffee genome. In addition to this global approach, the *C. arabica/M. exigua* interaction model has been studied. The genetic determinism of the analysed resistance has been precise and the genetic mapping of the *M. exigua* resistance locus, *Mex-1*, has been performed. To carry on these both approaches, the construction of a genomic library presenting large DNA fragments is become necessary. From the *M. exigua* resistant Arabica introgressed line IAPAR 59, a BAC library has been constructed and characterised. Using this BAC library, the preliminary physical map of *Mex-1* has been developed. The contig assembly related to the *Mex-1* locus and the display of these contigs containing RGA sequences are reported.

**Key-words** coffee tree (*Coffea* spp.), specific resistance, resistance genes, molecular markers, BAC library, physical mapping, introgression, polyploidy.