

ECOLE NATIONALE SUPERIEURE AGRONOMIQUE DE MONTPELLIER

**THESE
Présentée
Pour obtenir le grade de**

**DOCTEUR
DE
L'ECOLE NATIONALE SUPERIEURE AGRONOMIQUE DE MONTPELLIER**

ECOLE DOCTORALE: BIOLOGIE INTEGRATIVE

Par

Louis Marie Raboin

**Génétique de la résistance au charbon de la canne à sucre causé par
Ustilago scitaminea syd.: caractérisation de la diversité génétique du
pathogène, cartographie de QTL dans un croisement bi-parental et étude
d'associations dans une population de cultivars modernes.**

Soutenue le 22 juin 2005 devant la commission d'examen :

V. LEFEBVRE, Directrice de recherche à l'INRA
G. CHARMET, Directeur de recherche à l'INRA
J.L. NOTTEGHEM, Professeur à l'ENSA
F. QUETIER, Professeur à l'université de Paris XI
A. D'HONT, CIRAD

Rapporteur
Rapporteur
Examinateur
Examinateur
Directrice de thèse

Remerciements

Je remercie au premier chef l'équipe génétique canne à sucre à la Réunion (3P/PVBMT) composée de : Hugues Télismart, Iréné Promi et Michel Hoarau (qui profite maintenant d'une « LOOM » bien méritée). C'est grâce à leur travail que l'on a pu mettre en place et évaluer de nombreux essais sur le terrain. Ce qui n'est pas une mince affaire!

Je tiens aussi à remercier tout particulièrement Danièle Roques qui a créé la population sur laquelle est basée le travail de cartographie génétique.

Je remercie tout particulièrement et j'exprime toute mon amitié à l'ensemble de l'équipe canne à sucre de l'UMR PIA (ex BIOTROP) : Angélique D'Hont, Florence Paulet, Olivier Garsmeur, Loic Lecunff, Caroline Catalayud, Selvi Athiappan, Laurent Grivet, Jean Yves Hoarau, Jérôme Pauquet et Jean Christophe Glaszmann, directeur de l'UMR PIA. Cette thèse s'inscrit dans la dynamique scientifique qu'ils ont créée depuis maintenant dix ans sur la canne à sucre, une plante aussi «monstrueuse» qu'intéressante sur le plan génétique. Sans cette interaction, ce travail n'aurait tout simplement pas pu être réalisé. Je remercie aussi chaleureusement le canal historique de «Biotrop» Ange Marie Risterucci, Maguy Rodier-Gout, Jean Louis Noyer, Brigitte Courtois pour leur aide et leurs conseils précieux .

Je remercie aussi chaleureusement mes collègues de l'UMR PVBMT (3P). En premier lieu, je remercie Laurent Costet pour son aide lors des inoculations artificielles des différents essais et les différents conseils prodigués. Je le remercie donc pour son temps précieux compte tenu de ses responsabilités familiales qui semblent s'étendre chaque année... Je remercie Samuel Nibouche pour ses compétences précieuses avec «SAS». Je remercie également Magalie Payet et Emmanuelle Chapier pour leur aide. Je remercie bien sûr Jacques Dintinger, chef de l'axe génétique, pour ses conseils mais aussi pour ses ruggissements qui m'ont permis de rester en éveil devant mes pages blanches. Je remercie aussi Bernard Reynaud en tant que directeur de l'UMR PVBMT et du pôle de protection des plantes (3P) dans le cadre duquel s'est déroulé une grande partie des travaux de cette thèse.

Je remercie les sélectionneurs du CERF, Jacques Payet et Jean Yves Hoarau, pour les discussions que nous avons eues sur la canne à sucre et sa sélection. Je remercie aussi le CERF, et plus particulièrement Jean Pierre Diorflar, pour avoir conduit sur la station du Gol des essais d'évaluation pour la résistance au charbon. Je remercie aussi les autres chefs de station Jacques Bernard, Gilles Chassagne, Pascal Langlois et le directeur du CERF Michel Hellmann. A Jean-Yves, je dis un très grand merci pour son aide lors de la rédaction de cette thèse mais aussi pour la collaboration continue que nous avons eue sur le terrain au cours de ces 7 (Déjà!) dernières années.

A l'UMR BGPI, je remercie Jean Carlier pour son aide cruciale concernant l'étude de la diversité du charbon de la canne à sucre. Je remercie grandement, Jean Claude Girard et Philippe Rott, pour leurs informations et leur expertise en pathologie de la canne.

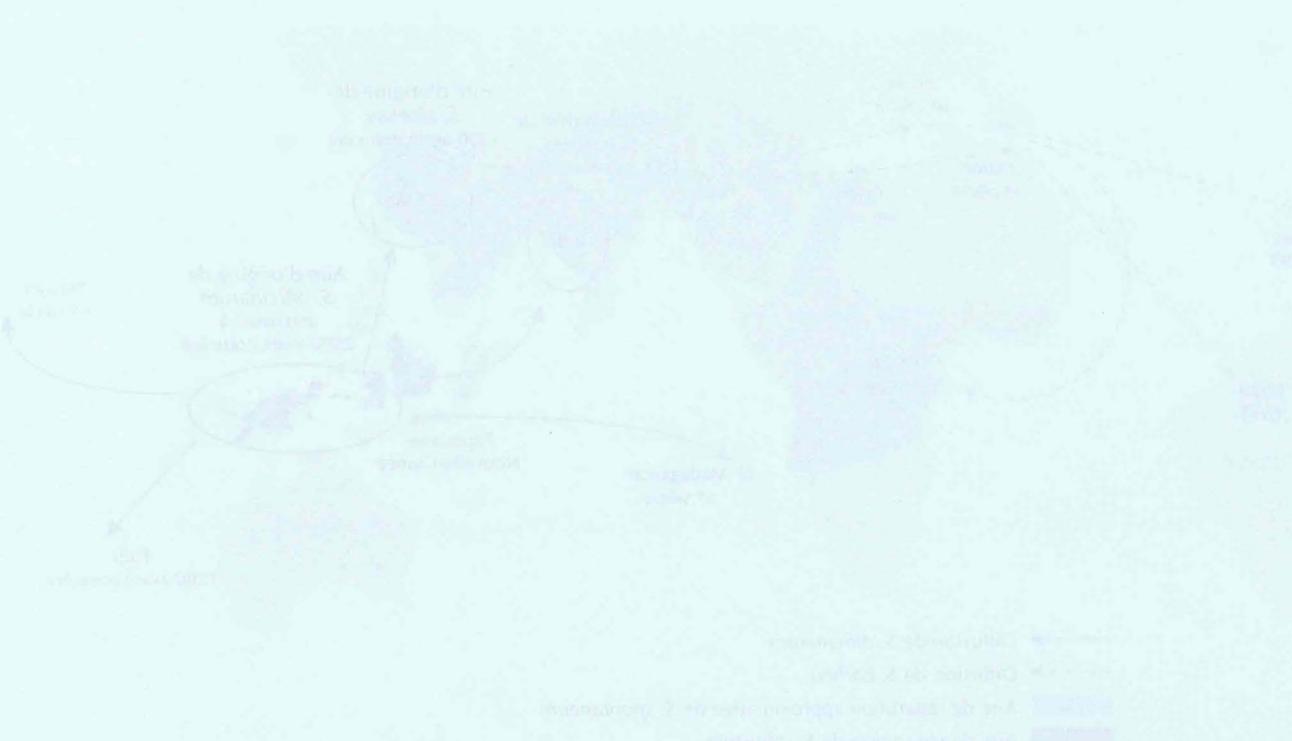
A tout le monde, je dis à bientôt mais sur une autre île, un peu plus grande que la Réunion, «Madagascar», et pour l'étude d'une autre herbe, un peu plus petite que la canne à sucre, «le Riz »...

TABLES DES MATIERES

I. INTRODUCTION	1
I.1 Histoire et géographie	1
I.1.1 Histoire de la diffusion de la culture de la canne à sucre (D'après Glaszmann et al. 2003, Daniels et Roach 1987 et Meyer 1989)	1
I.1.2 Importance économique de la canne à sucre	1
I.2 Taxonomie, évolution et domestication de la canne à sucre	2
I.2.1 Le «complexe Saccharum»	2
I.2.2 Le genre Saccharum	3
I.2.3 Scénario pour la domestication de la canne à sucre (Grivet et al. 2004)	6
I.2.4 Diversité génétique au sein du genre Saccharum (fig. 4)	6
I.3 La sélection de la canne à sucre	7
I.3.1 Début de l'amélioration génétique	7
I.3.2 La nobilisation et les premiers hybrides interspécifiques	8
I.3.3 Maîtrise de la floraison (revue par Moore et Nuss 1987)	9
I.3.4 Les méthodes de sélection	10
I.4 Etude du génome des cultivars modernes de canne sucre	11
I.4.1 Structure du génome	11
I.4.2 Cartographie génétique	12
I.4.3 L'ère de la génomique	13
I.5 Pathogènes et ravageurs de la canne à sucre	14
I.5.2 Les principaux ravageurs	14
I.5.2 Les principales maladies	16
I.5.3 Le charbon de la canne à sucre	17
I.6 Introduction au sujet de thèse	20
II Diversité génétique de <i>Ustilago scitaminea</i> agent du charbon de la canne à sucre	22
II.1 Introduction	24
II.2 Material and Methods	26
Fungal isolates	26
DNA extraction	26
Construction of an enriched genomic library and definition of microsatellite primers	27
Microsatellite analysis	27
Data analysis	28
II.3 Results	28
II.4 Discussion	30
Selfing—the predominant reproduction mode	30
An Asian center of origin	31
Global dispersal process	32
Implications for studies on pathogenic variation	33
Implications for resistance breeding and management	33
III Cartographie génétique dans le croisement R 570 x MQ 76/53	39
III.2 Introduction	41
III.3 Materials and Methods	43
Plant and DNA material:	43
Field evaluation of rust resistance and stalk colour	43

AFLP markers	44
Microsatellites markers:	44
RFLP markers:	44
Map construction:	45
Use of double dose (DD) markers of bi-parental origin	45
III.4 Results	46
Parental linkage maps	46
Alignment of different R570 AFLP-based maps	47
Tentative identification of common chromosome segments between parental genomes	47
Identification and mapping of a putative major rust resistance gene in MQ76-53	48
Identification and mapping of a major gene controlling stalk colour in MQ76-53	49
III.5 Discussion	49
IV Etude de la structure du déséquilibre de liaison dans une population non structurée de cultivars de canne à sucre à l'aide de marqueurs AFLP	58
IV.1 Introduction	59
IV.2 Materials and methods	61
Plant material:	61
AFLP protocol	61
Statistical analyses	61
IV.3 Results	63
An overview of the raw AFLP data	63
Population structure and correlation between coefficient of parentage and Genetic dissimilarity	64
‘True’ associations (i.e. associations between markers genetically linked on a same chromosome segment) versus ‘false associations’: Where is the threshold?	65
The extent of Linkage Disequilibrium in sugarcane	66
Linkage disequilibrium among all the AFLP markers produced	67
IV.4 Discussion	68
V Cartographie génétique de la résistance de la canne à sucre au charbon dans une descendance bi-parentale et par association dans une population de cultivars.	76
V.1 Introduction	77
V.2 Material and Methods:	79
Plant material	79
Inoculum used in the greenhouse and field trials	80
Comparison of different inoculation methods under greenhouse:	80
Greenhouse evaluation of the progeny population	81
Field evaluation of the progeny population	81
Analysis of quantitative traits	81
Available molecular data for the progeny population:	82
Available molecular data for the cultivar population:	83
QTL analysis:	83
Analysis of marker-smut resistance associations in the cultivar population	84
V.3 Results	84
Comparison of different inoculation methods in Greenhouse trials:	84
Smut resistance evaluation of the progeny population	85
QTL analysis	86
Analysis of marker-smut resistance associations in the cultivar population	87
V.4 Discussion	87

VI CONCLUSIONS ET PERSPECTIVES	93
Diversité génétique de <i>Ustilago scitaminea</i> , agent du charbon de la canne à sucre	93
Cartographie génétique dans le croisement biparental ‘R570’ x ‘MQ 76/53’	95
Structure du déséquilibre de liaison chez la canne à sucre	97
Etude du déterminisme génétique de la résistance au charbon de la canne à sucre et des autres caractères étudiés concomitamment.	100
Perspectives pour la création variétale	101
REFERENCES BIBLIOGRAPHIQUES	106
ANNEXES	



I. INTRODUCTION

Recent technological advances have enabled the study of the gut microbiome and its interactions with the central nervous system (CNS) to become feasible. The gut microbiome consists of trillions of microorganisms that colonize the gut, primarily in the large intestine. These microorganisms are involved in various physiological processes, including nutrient metabolism, energy production, and immune system regulation. The gut microbiome has been implicated in a variety of diseases, including inflammatory bowel disease, obesity, and mental health disorders. The CNS, on the other hand, is a complex network of neurons and glial cells that regulate various bodily functions. The gut microbiome and the CNS are interconnected through various pathways, including the vagus nerve, the enteric nervous system, and the release of hormones and neurotransmitters. This review will discuss the current understanding of the gut microbiome and its interactions with the CNS, focusing on the role of the gut microbiome in mental health disorders.

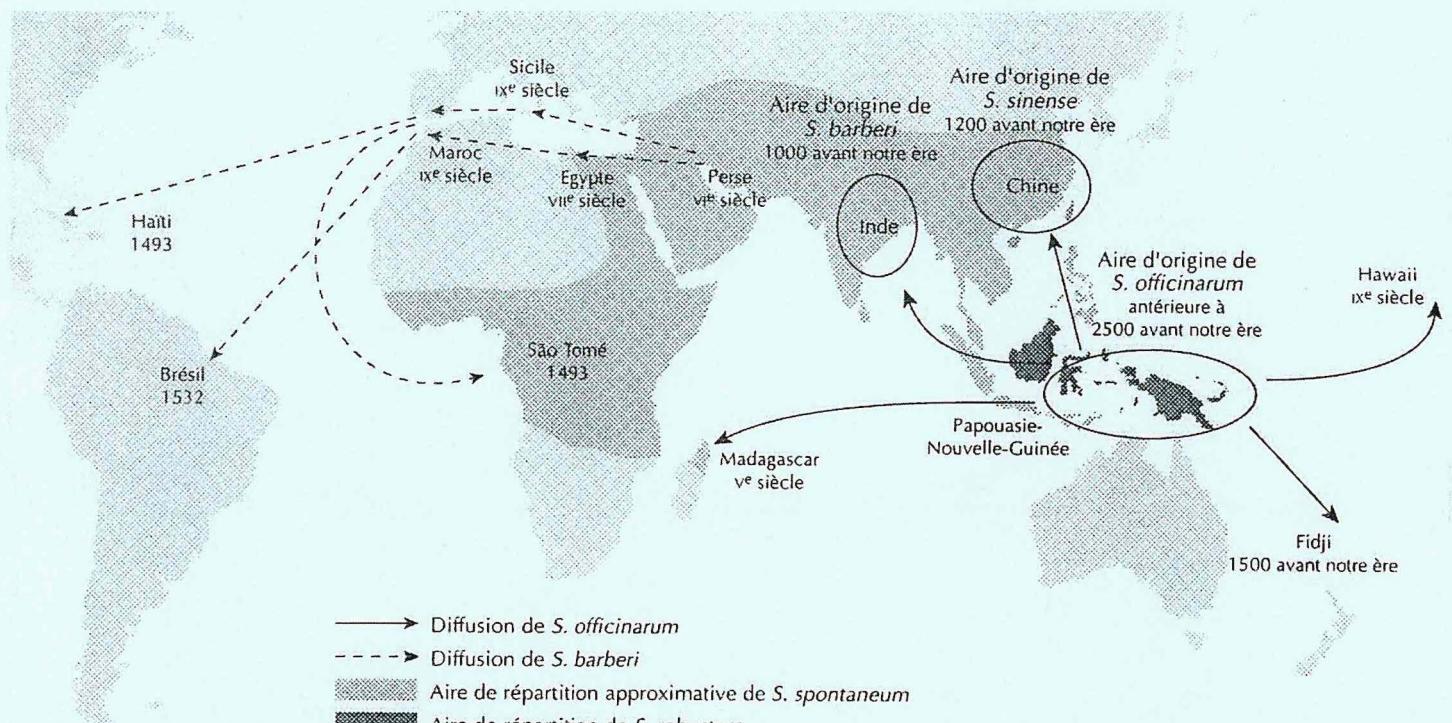


Figure 1 : Aire d'origine des trois espèces de canne à sucre domestiquées, *S. officinarum*, *S. barberti* et *S. sinense*, et dispersion jusqu'au début du XVI^e siècle, d'après Blume (1985), Daniels et Roach (1987) et Meyer (1989)

I. INTRODUCTION

I.1 Histoire et géographie

I.1.1 Histoire de la diffusion de la culture de la canne à sucre (D'après Glaszmann *et al.* 2003, Daniels et Roach 1987 et Meyer 1989)

La domestication de la canne à sucre serait survenue en Papouasie Nouvelle Guinée, où existe une grande diversité morphologique de l'espèce domestiquée *Saccharum officinarum*. La domestication de la canne à sucre serait antérieure à 2500 avant notre ère. La canne à sucre y était cultivée dans les jardins pour être mâchée. La canne à sucre se serait ensuite répandue dans les îles du Pacifique sud, en Inde et en Chine au cours des migrations austroasiatiques qui se déroulèrent aux alentours de 1500 à 1000 ans avant notre ère (Fig. 1). C'est vraisemblablement en Inde et/ou en Chine que l'industrie de l'extraction du sucre est née. En 500, la canne à sucre était cultivée en Perse. A partir de là, les Arabes étendirent sa culture en Afrique du nord et aux îles méditerranéennes. Par la suite, les Portugais, puis les espagnols propagèrent sa culture dans les îles de l'Atlantique (Madère, Canaries, îles du Cap-vert, Sao Tomé). Christophe Colomb, à l'occasion de son deuxième voyage, emmena la canne à sucre vers les Amériques et c'est à Haïti (Hispaniola) que la canne à sucre a été cultivée en premier.

La canne à sucre aurait été introduite à la Réunion, appelée île Bourbon à l'époque, vers 1665 à partir de Madagascar où elle serait arrivée à l'occasion des migrations malaises. L'arrivée de la canne à sucre est donc presque simultanée avec le premier peuplement permanent de l'île par deux français et dix malgaches débarqués en 1663 en baie de Saint-Paul. Pendant très longtemps, la canne à sucre ne servit qu'à la préparation d'un jus fermenté appelé «Fangourin» dont la distillation permettait de fabriquer «l'Arak». Ce terme, «rak», est encore utilisé aujourd'hui pour désigner le rhum. La première fabrication de sucre à la Réunion daterait de 1784 mais ne fut convenablement maîtrisée qu'à partir de 1815 par Charles Desbassyns (Trouette (1983), Chaudenson (1980), Lavaux (1975) cités par Girard et Payet, 1997)

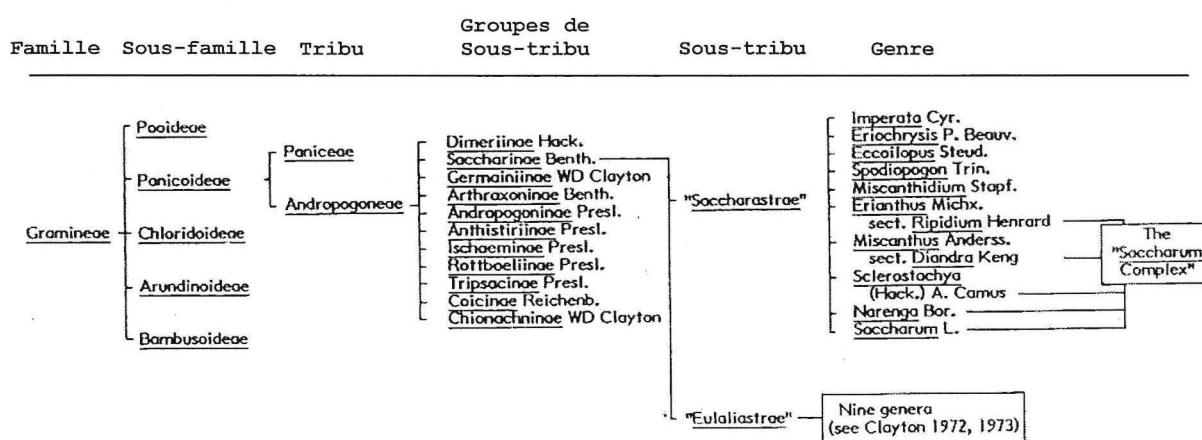
I.1.2 Importance économique de la canne à sucre

La canne à sucre est une monoculture industrielle semi-pérenne et récoltée annuellement. Elle est cultivée dans plus de 115 pays, tropicaux pour l'essentiel d'entre eux puisque la plante ne tolère pas le gel. Les principaux producteurs mondiaux sont le Brésil, l'Inde et la Chine. En

Tableau 1: Statistiques de la FAO concernant la production de la canne à sucre et les surfaces consacrées à sa culture (données FAO STAT, 2004)

Pays	Surfaces (ha)	Production (MT)
Brésil	5 336 980	389.8
Inde	4 607 600	281.6
Chine	1 432 960	92.0
Thaïlande	1 121 000	78.2
Pakistan	1 086 000	52.0
Mexique	639 061	45.1
Australie	423 000	37.9
Colombie	435 000	37.0
USA	402 420	30.7
Philippines	391 095	25.9
Indonésie	340 000	24.5
Cuba	654 331	22.9
Afrique du sud	325 000	20.4
Réunion	29 000	2.0
World	20 698 803	1 351.9

Tableau 2: Place de Saccharum dans la famille des graminées (Daniels et Roach 1987)



2004, 1.351 milliards de tonnes de canne à sucre ont été produites dans le monde sur une surface de 20.698 millions d'hectares (tableau 1). La consommation mondiale de sucre continue de croître d'environ 2,5% par an pour atteindre près de 143 millions de tonnes en 2004. Cette augmentation a été initiée par la croissance des marchés asiatiques alors que la consommation dans les pays occidentaux est restée stable ou a légèrement baissé. Plus des deux tiers du saccharose produit au niveau mondial est issu de la canne à sucre. Le tiers restant est apporté par la culture de la betterave sucrière. La proportion de la production mondiale de sucre provenant de la canne à sucre est en constante augmentation.

A la Réunion, la canne à sucre est plantée sur environ 27 000 ha ce qui représente presque la moitié de la surface agricole utile. La production annuelle moyenne oscille autour de 1,8 millions de tonnes de canne dont on tire environ 200 000 tonnes de saccharose. Il s'agit d'une filière industrielle intégrée sucre/rhum/bagasse. Deux complexes industriels, l'un au nord et l'autre au sud de l'île, traitent l'ensemble de la production de l'île. Ces usines sont couplées avec des centrales thermiques qui brûlent la bagasse, résidu de l'extraction du sucre, pendant la campagne sucrière et du charbon pendant le reste de l'année pour produire de l'électricité. La mélasse, autre sous-produit de l'extraction du sucre, permet la production d'un rhum industriel. On peut aussi parler de la contribution de la filière canne à la filière élevage par l'intermédiaire de la paille de canne qui est utilisée comme fourrage. En dépit de cela, de nombreuses ombres planent sur l'avenir de la filière canne à la Réunion. D'une part, la canne est en concurrence forte avec d'autres spéculations (la première étant l'immobilier) pour l'utilisation du foncier disponible. D'autre part, le soutien de l'union européenne est amené à diminuer et enfin la canne à sucre représente une part de plus en plus marginale de la richesse produite à la Réunion (Rochoux, 2002).

I.2 Taxonomie, évolution et domestication de la canne à sucre

I.2.1 Le «complexe *Saccharum*»

La canne à sucre est une grande plante herbacée qui appartient à la famille des graminées, à la tribu des andropogonées et au genre *Saccharum* (tableau 2). Des échanges génétiques sont possibles entre *Saccharum* et plusieurs genres voisins : *Erianthus*, *Misanthus*, *Sclerotaschya* ou *Narenga*. Ces genres s'hybrident plus ou moins facilement avec le genre *Saccharum* et les descendances obtenues sont plus ou moins fertiles (Grassl 1972; Chen *et al.* 1983; Sreenivasan *et al.* 1987). Ces genres ont longtemps été considérés comme étroitement liés à *Saccharum* et comme ayant participé à l'origine de la canne à sucre. En réalité, il semble que ce concept de «complexe *Saccharum*» ait conduit à surestimer la contribution des genres apparentés à l'émergence de la canne à sucre cultivée. Un ensemble d'éléments moléculaires

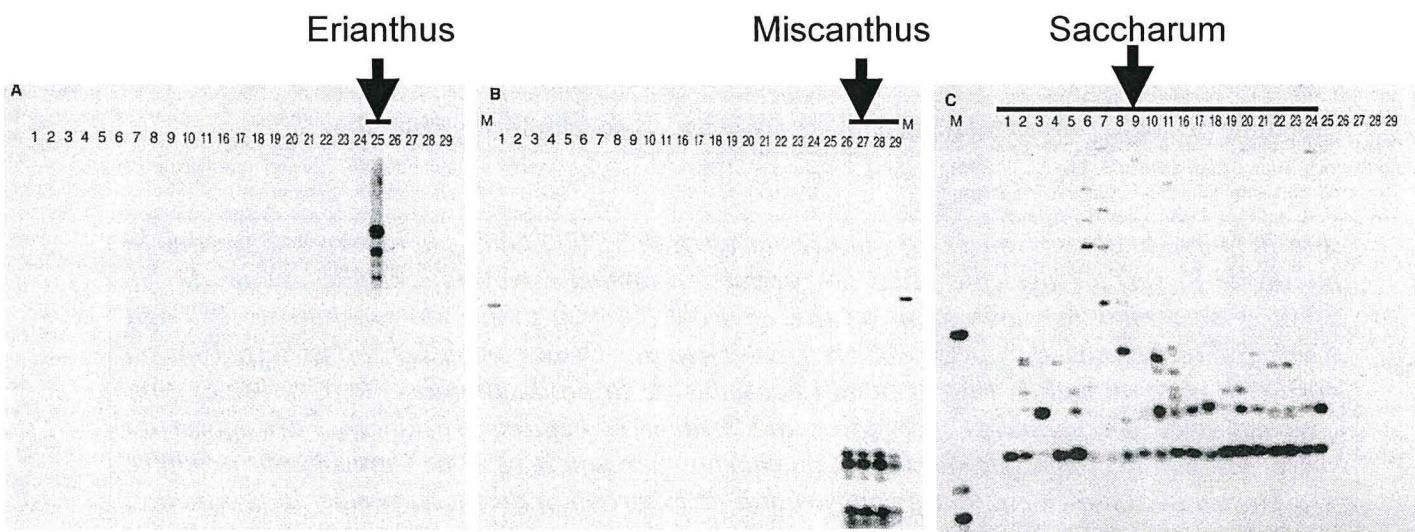


Figure 2: hybridations de séquences répétées de type Alu, des genres *Erianthus*, *Miscanthus* et *Saccharum* sur l'ADN de représentants de chacun de ces genres (Alix et al., 1998 et 1999)

démontre que la canne à sucre est issue spécifiquement du phylum *Saccharum* (Besse et al. 1997 ; Grivet *et al.* 2004). A titre d'exemple, l'absence d'hybridation de sondes répétées, de type Alu like ou satellite, spécifiques des genres *Miscanthus* et *Erianthus* sur l'ADN de différents cultivars de canne à sucre est un élément qui étaye clairement cette thèse (Alix *et al.* 1998 et 1999) (Fig. 2). Dans la pratique, la réussite des croisements inter-génériques semble souvent délicate en raison d'une compatibilité souvent médiocre en hybridation. L'observation de caractères agro-morphologiques ne constitue pas un critère fiable permettant de distinguer les descendants légitimes de ceux qui ne le sont pas, comme des travaux moléculaires ont pu le montrer sur des descendances fondées sur des croisements entre *Saccharum* et *Erianthus* (D'Hont *et al.* 1995; Piperidis *et al.* 2000).

I.2.2 Le genre *Saccharum*

La canne à sucre a évolué à partir de six espèces de base appartenant au genre *Saccharum*, parmi lesquelles on distingue deux catégories:

- les espèces sauvages
- les espèces domestiquées

Les espèces sauvages

S. spontaneum est une espèce très rustique que l'on retrouve dans des écologies très diverses. Elle peut coloniser des régions extrêmement contrastées de part leurs températures, leur humidité ou leurs types de sols. En conséquence *S. spontaneum* possède une étendue géographique très vaste qui s'étale du Japon jusqu'à l'Afrique de l'Est en passant par l'Asie du sud-est, le continent indien, le moyen orient et le bassin méditerranéen (Brandes *et al.* 1939). La reproduction se fait par voie sexuée et végétative. La plante produit des rhizomes. Il existe des écotypes annuels, semi-pérennes ou pérennes. *S. spontaneum* présente généralement un faciès herbacé caractérisé par des tiges grêles qui tallent abondamment. La hauteur de la plante peut être très variable selon les écotypes, depuis des formes très courtes jusqu'à 5 m de hauteur. Il existe une certaine variabilité génétique des caractères agronomiques, mais beaucoup d'entre eux sont assez typés par rapport aux espèces sucrières (Tai *et al.* 1995) : les feuilles sont étroites, les tiges sont relativement fines, elles ne contiennent pas ou pratiquement pas de sucre, mais sont riches en fibre.

Le nombre de chromosomes varie de $2n = 40$ à $2n = 128$. Les cytotypes les plus fréquents correspondent à un multiple de huit (Panje et Babu 1960; Sreenivasan *et al.* 1987, Burner 1991). D'Hont *et al.* (1998) ont montré par hybridation *in-situ* des ADN ribosomaux, que le nombre chromosomique de base est $x = 8$, et que le niveau de ploïdie varie entre 8 et 12 selon les cytotypes étudiés.

S. robustum est la deuxième espèce sauvage du genre *Saccharum*. Elle se distingue de *S. spontaneum* par l'absence de rhizomes, une inflorescence plus grande, des épillets plus petits, une biomasse bien plus abondante, qui résulte de touffes compactes, formées de tiges généralement longues (qui peuvent atteindre jusqu'à 10 mètres), relativement grosses et particulièrement ligneuses (Stevenson 1965). La Nouvelle-Guinée et les îles de Mélanésie constituent le centre d'origine de *S. robustum* qui est resté endémique à cette région (Daniels et Roach 1987). Son habitat naturel est situé sur des berges de rivières.

On trouve deux cytotypes euploïdes: $2n = 60$ ou bien $2n = 80$ et des aneuploïdes variant de $2n = 63$ à 205. Ces clones aneuploïdes pourraient correspondre à des hybrides naturels entre clones de *S. robustum* ou entre *S. robustum* et d'autres espèces (Price 1957 et 1965).

Les espèces domestiquées

S. officinarum est l'espèce sucrière par excellence, que l'on appelle encore l'espèce «noble». Elle correspond aux formes domestiquées, dont les plus intéressantes ont été propagées en culture pendant des siècles jusqu'à l'intervention de l'hybridation interspécifique. L'espèce présente certaines qualités agronomiques et industrielles correspondant aux critères de sélection les plus importants : grosses tiges avec un port érigé, forte richesse en saccharose, et faible teneur en fibre et en amidon (Bremer 1961; Roach 1986). Bien que considéré initialement comme très homogène, *S. officinarum* présente des caractéristiques agronomiques qui sont en réalité assez variables suivant les clones (Daniels et Roach 1987), comme l'étaye par ailleurs une récente étude de diversité génétique menée à l'aide de marqueurs moléculaires (Jannoo *et al.* 1999a).

S. officinarum est une espèce euploïde à $2n=80$ chromosomes (Bremer 1930; Li et Price 1967; Price et Daniels 1968). La méiose est généralement normale avec formation régulière de bivalents. *S. officinarum* est une espèce octoploïde; elle présente un nombre chromosomique de base de $x=10$ (D'Hont *et al.* 1998).

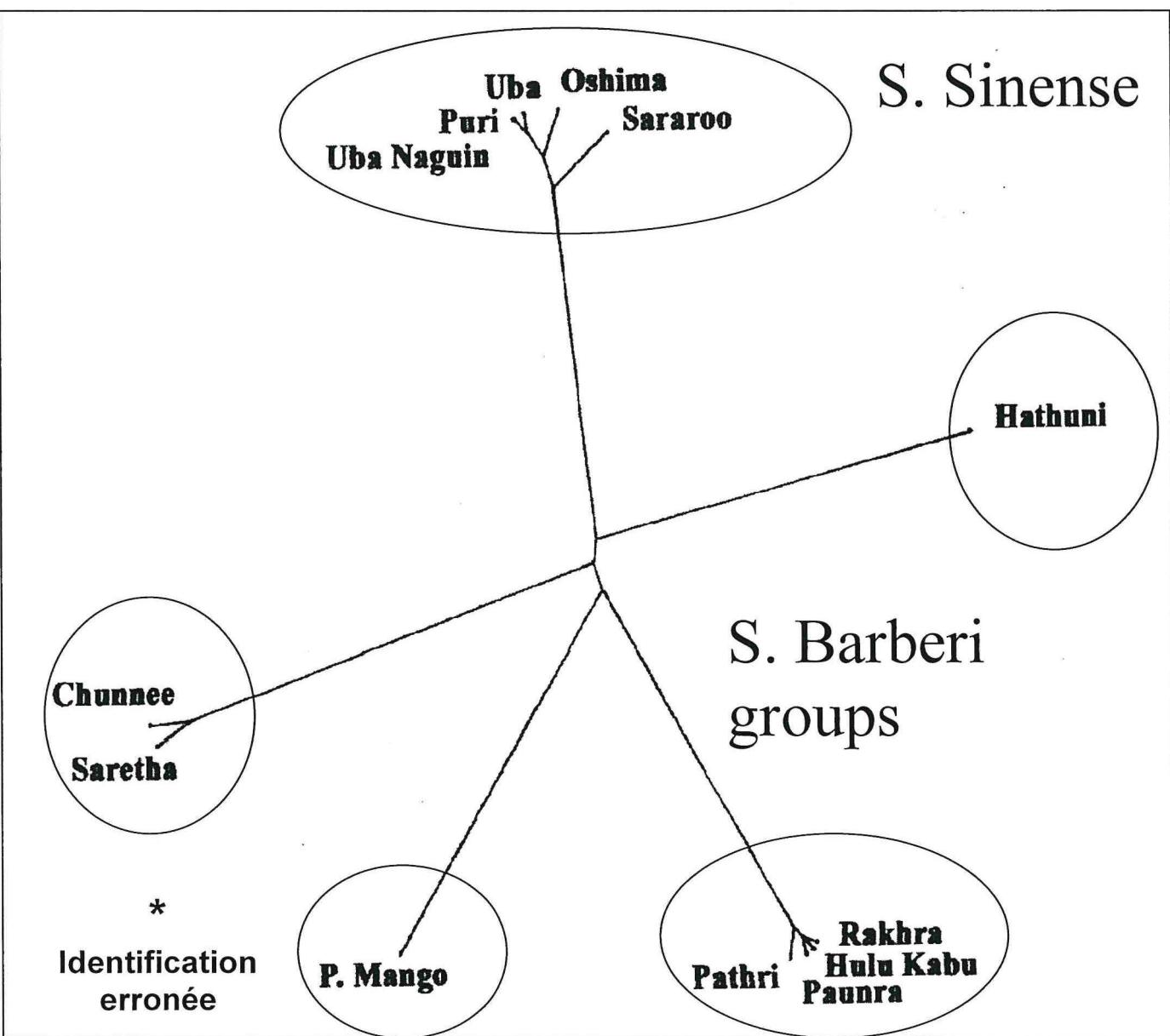


Figure 3: Diversité génétique dans les espèces *S. sinense* et *S. barberi* (D'hont et al, 2002)

S. barberi et *S. sinense* : depuis des millénaires, ces deux espèces sucrées ont été respectivement cultivées en Inde du Nord (qui est devenu depuis l'Est du Pakistan) et en Chine du Sud (Stevenson 1965). Elles se distinguent des clones de *S. officinarum* par leurs caractères floraux, leur haute teneur en fibre et leur grande rusticité qui est due à leur meilleure tolérance aux stress environnementaux. De nombreux clones sont résistants aux principales maladies de la canne (Daniels et Roach 1987). Des clones de ces deux espèces ont été utilisés en croisement avec *S. officinarum* dans les premières étapes de l'amélioration variétale de la canne à sucre Le nombre chromosomique de ces deux espèces varie entre $2n = 82$ et 124 . Les observations cytologiques montrent des méioses extrêmement irrégulières. Ces deux espèces possèdent une origine interspécifique qui impliquerait *S. officinarum* et *S. spontaneum* (Brandes 1958). Ce sont vraisemblablement des hybrides de première génération car très peu d'échanges interspécifiques entre les chromosomes ont été observés par Hybridation *in-situ* (D'Hont *et al.* 1996). Par ailleurs, l'étude de la diversité génétique entre différents groupes de clones appartenant à ces deux espèces montre qu'ils devraient être plutôt considérés comme des groupes horticoles plutôt que comme des espèces distinctes (Fig. 3). En effet la dissimilarité entre les différents groupes de *S. barberi* est équivalente à la dissimilarité entre *S. barberi* et *S. sinense* (D'Hont *et al.* 2002).

S. edule : cette espèce stérile comprend un petit groupe de clones, circonscrits à la Nouvelle Guinée et aux îles voisines. Elle est cultivée pour ses inflorescences avortées comestibles dans les jardins villageois de Nouvelle Guinée et de Fidji. Les clones constituent une série polyploïde avec $2n = 60$, 70 ou 80 (Roach 1972). On la retrouve sous le nom de 'Naviso' au Vanuatu ou 'Pitpit' en Papouasie Nouvelle Guinée. Les clones de cette espèce sont probablement issus de l'espèce *S. robustum* bien qu'une origine interspécifique avec *S. robustum* comme donneur du gamète femelle ne puisse pas être écartée (Grivet *et al.* 2004). Les profils moléculaires mitochondriaux et chloroplastiques des clones de *S. edule* sont identiques au profil le plus fréquemment observé dans l'espèce *S. robustum* (D'Hont *et al.* 1993; Sobral *et al.* 1994)

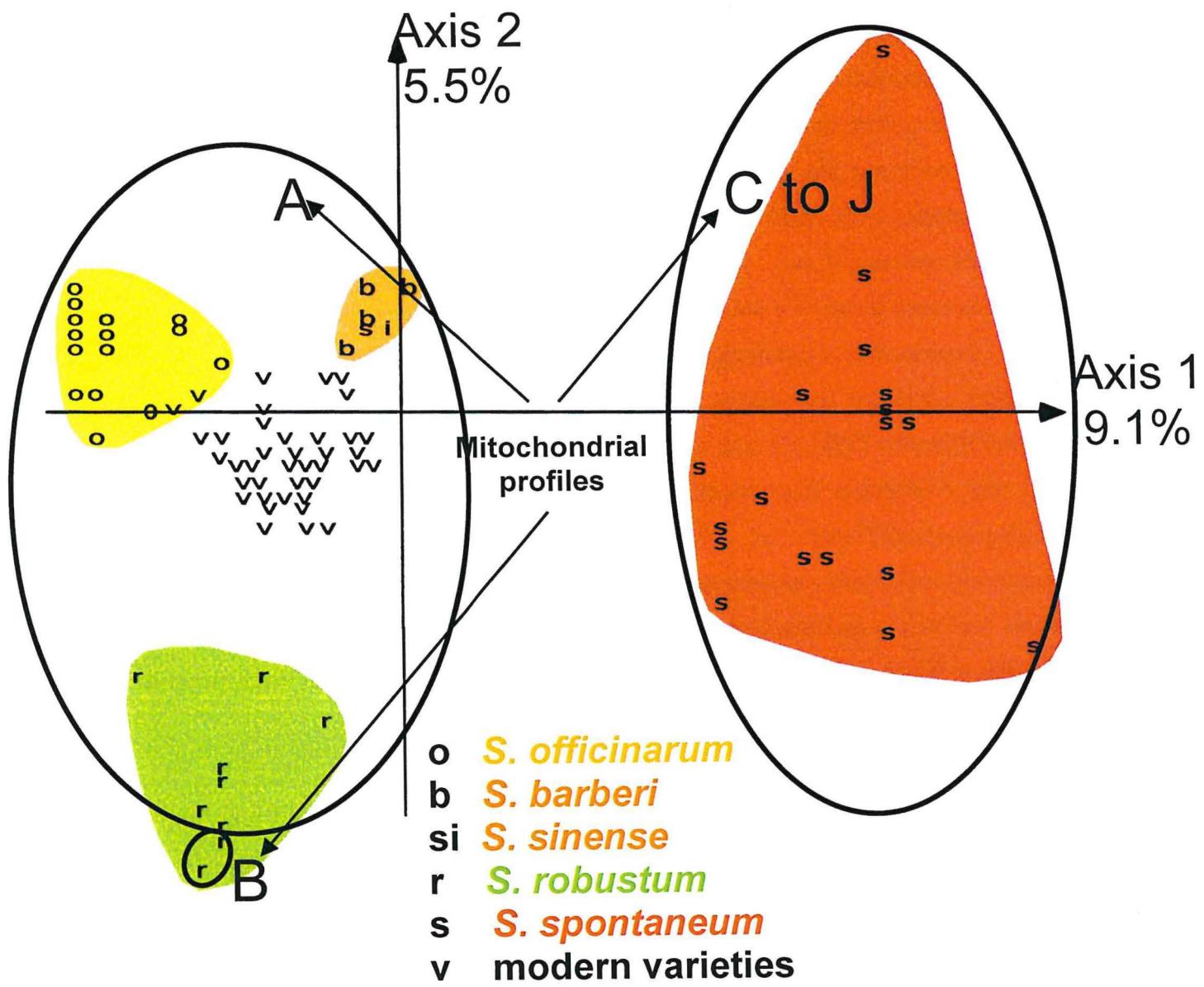


Figure 4: Diversité génétique dans le genre *Saccharum* (D'Hont et al. 1993, Lu et al. 1994, Jannoo et al, 1999a)

I.2.3 Scénario pour la domestication de la canne à sucre (Grivet *et al.* 2004)

La ligne de Wallace sépare deux zones du continent asiatique, la partie Sunda (Sud-est asiatique continental, Java, Sumatra et Kalimantan) et la partie Sahul (Papouasie Nouvelle Guinée et les îles avoisinantes). Ces deux zones sont restées séparées par les eaux pendant l'essentiel des trois millions d'années de l'ère quaternaire. Néanmoins, il y a eu des opportunités d'échange génétique entre elles au moment des périodes glaciaires grâce à la baisse du niveau des eaux. Il est donc proposé que les deux espèces sauvages du phylum *Saccharum* se soient différenciées de part et d'autre de cette ligne de séparation: *S. spontaneum* du côté Sunda et *S. robustum* du côté Sahul. La domestication de *Saccharum officinarum* s'est très vraisemblablement déroulée en Papouasie Nouvelle Guinée et dans les îles avoisinantes à partir de l'espèce *S. robustum*. C'est en effet dans cette région que l'on trouve la plus grande diversité morphologique des clones de *S. officinarum* (Daniels et Roach 1987). Les clones nobles (clones de l'espèce domestiquée *S. officinarum*) auraient été ensuite diffusés vers Sunda et aussi vers les îles du Pacifique au cours de migrations humaines. Des hybridations interspécifiques entre clones nobles et *S. spontaneum* seraient alors devenues possibles. Elles auraient permis l'apparition naturelle d'hybrides interspécifiques tels que les clones des espèces *S. barberi* (en Inde) et *S. sinense* (en Chine)

I.2.4 Diversité génétique au sein du genre *Saccharum* (Fig. 4)

La diversité cytoplasmique a été étudiée à l'aide de sondes chloroplastiques et mitochondrielles hétérologues (D'Hont *et al.* 1993). Les sondes chloroplastiques n'ont révélé aucune variabilité. En revanche, les sondes mitochondrielles ont permis de définir 8 profils différents au sein du genre *Saccharum*. Deux profils ont été révélés pour l'espèce *S. robustum*. Les espèces *S. officinarum*, *S. barberi*, *S. sinense* ainsi que les cultivars modernes de canne à sucre présentent le même profil mitochondrial qui est aussi identique au profil dominant de l'espèce *S. robustum*. Ces observations indiquent vraisemblablement que les clones nobles ont été impliqués comme femelles lors des hybridations naturelles qui ont fondé les espèces *S. barberi*, *S. sinense* et les cultivars modernes de canne à sucre. Chez *S. spontaneum*, la diversité cytoplasmique est plus importante. Six profils mitochondriaux différents y ont été identifiés.

L'étude de la diversité nucléaire apporte aussi des informations intéressantes (Lu *et al.* 1994a et 1994b; Jannoo *et al.* 1999a). La diversité génétique de l'espèce *S. spontaneum* est beaucoup plus grande que celle des autres espèces. L'essentiel de la diversité de l'espèce *S. officinarum* est retrouvée dans le pool des cultivars modernes (seulement 15% des allèles de *S. officinarum* n'ont pas été retrouvés dans les 109 cultivars étudiés). L'apparente séparation entre l'espèce *S. robustum* et l'espèce *S. officinarum* révélée par ces études s'explique vraisemblablement par le manque de représentativité de l'échantillonnage des clones *S. robustum* étudiés mais aussi par un effet de fondation chez l'espèce *S. officinarum* comme semble le suggérer la relativement faible diversité de cette dernière.

Ces données sont compatibles avec l'hypothèse d'une évolution de *S. officinarum* à partir de *S. robustum*. Elles confirment l'origine interspécifique de *S. barberi*, *S. sinense* et des cultivars modernes. Elles indiquent aussi qu'il y a peu à gagner dans la mise en place de programmes d'élargissement de la base génétique à partir de l'espèce *S. officinarum*.

I.3 La sélection de la canne à sucre

I.3.1 Début de l'amélioration génétique

Jusqu'à la fin du XIXème siècle la culture de la canne a reposé sur la propagation végétative d'un nombre très réduit de clones provenant des trois espèces sucrées (*S. officinarum*, *S. sinense*, *S. barberi*). Il fallut attendre la fin du XIXème siècle que la fertilité de la canne soit découverte car les clones avec lesquels les planteurs étaient le plus familiers étaient souvent mâles stériles. La canne à sucre était donc considérée comme une plante totalement stérile jusqu'à la découverte de plantules issues de semences naturelles à Barbades en 1858 (Stevenson 1965). Cette aptitude ne fut utilisée délibérément pour produire de nouvelles variétés que bien plus tard en 1888, d'abord à Java, puis la même année à Barbades. Pendant les vingt années suivantes, de nombreux croisements furent réalisés par les pionniers de la sélection à Java, en Inde, à Maurice et à Hawaï (Bremer 1961). Dans un premier temps, les programmes d'amélioration ont été fondés sur des croisements entre clones «nobles» (*S. officinarum*). Ces croisements intra-spécifiques remportèrent un succès indéniable. Les variétés POJ100, 247B, EK28 ainsi obtenues à Java ou encore les clones BH10/12 ou B726 à Barbades furent successivement cultivés sur de grandes surfaces dans ces deux îles pendant

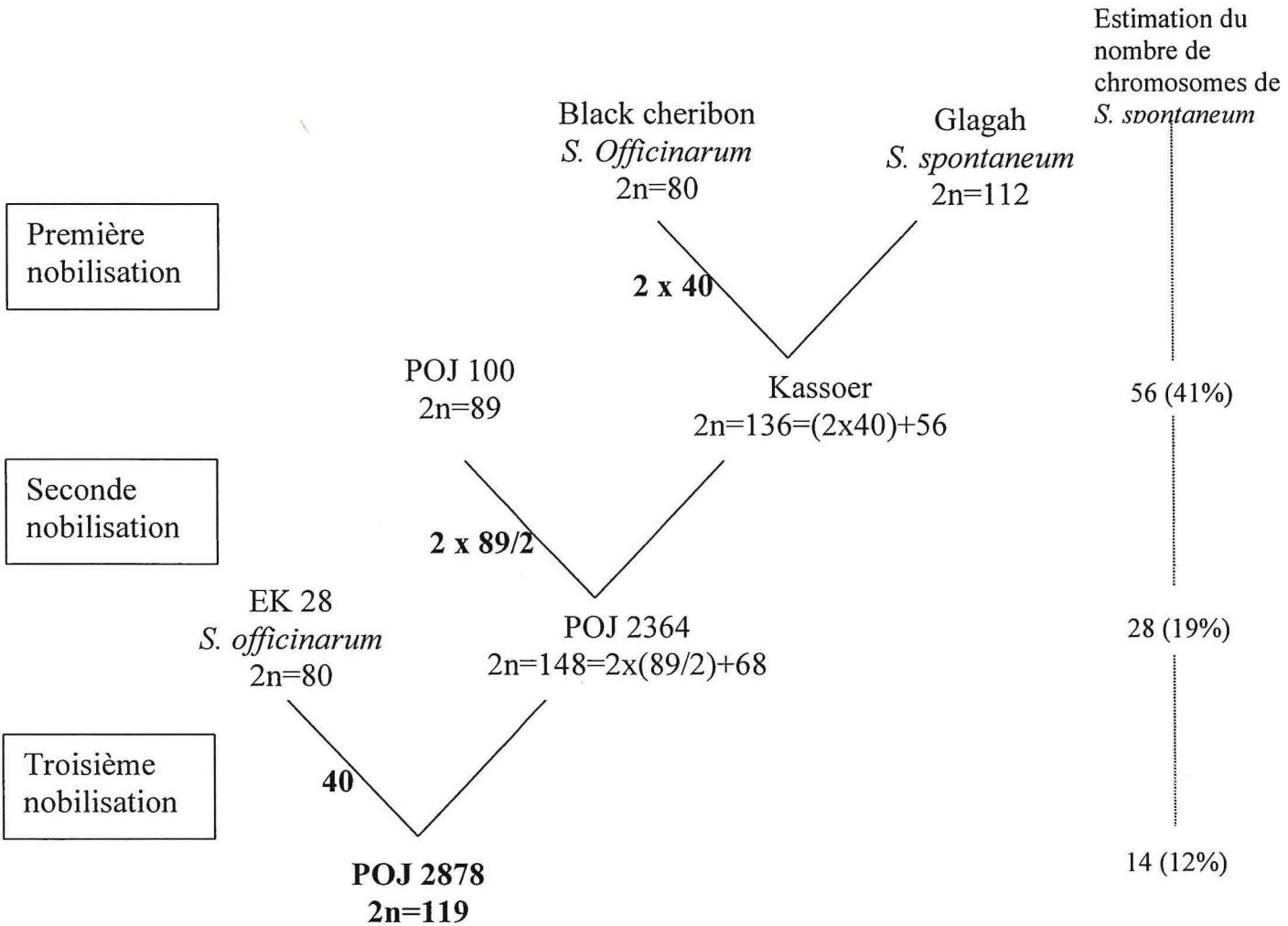


Figure 5: Exemple de nobilisation ayant amené à la création du cultivar POJ 2878. Au cours des deux premières générations de nobilisation le clone *S. officinarum* utilisé comme femelle transmet son nombre somatique de chromosomes ($2n$). Il est à noter toutefois que POJ 100 n'est pas un pur *S. officinarum* (son nombre chromosomique est 89 et non 80). A la troisième nobilisation EK 28 a transmis un gamète réduit normal (Bremer, 1961)

les deux premières décennies du siècle (Stevenson 1965). Ces travaux visaient à maîtriser les principales maladies de l'époque qu'étaient le sereh (Rands et Abbotts 1964) et la mosaïque, deux maladies virales ainsi qu'une bactériose, la gommosé (Hughes 1964).

A la Réunion, une forte épidémie de mosaïque s'est développée entre 1920 et 1930 provoquant des pertes considérables essentiellement dans la zone 'sous le vent' (ouest de l'île) sur les cannes nobles cultivées à l'époque (Lousier, Port Mackay , Bigtanna rayée, Bigtanna blanche, Batavia, DK 74, MP 55, MP 33, MP 131...). Pour répondre à cette crise, la station d'essai de la Bretagne fut créée en 1929 avec pour mission de rechercher des variétés de canne à sucre résistantes . Elle deviendra ensuite le CERF (Centre d'essais, de recherche et de formation). Le premier sélectionneur de l'île, Donald d'Emmerez de Charmoy, part alors se former en Inde sur la station de Coimbatore où historiquement se sont déroulés les premiers travaux d'hybridation de la canne à sucre parallèlement aux travaux entrepris par les Hollandais sur l'île de Java. Dès son retour, les cannes nobles sont progressivement remplacées par des hybrides étrangers venus de Java (POJ 2878...) ou de Coimbatore Co 214, Co 281, Co 290...) puis par des hybrides réunionnais tel que R 337(POJ 2878 x Uba Marot) qui a occupé jusqu'à 60 % des surfaces en canne après guerre. La création variétale va alors accompagner un véritable boom sucrier à la Réunion qui verra la production de sucre passer de 50000 tonnes en 1930 à 260 000 tonnes en 1960 alors que les surfaces consacrées à la canne n'ont pratiquement pas évoluées (Hoarau et al. 2002).

I.1.3.2 La nobilisation et les premiers hybrides interspécifiques

Les premiers travaux dans ce domaine débutèrent à la station de sélection de Proefstation Ost Java créée par des Hollandais. Ces derniers désignèrent par le terme de «nobilisation» l'action consistant à croiser un clone «noble» (*S. officinarum*) riche en sucre par un clone d'une espèce apparentée, vigoureux ou résistant à une maladie, puis à rétro-croiser éventuellement plusieurs fois l'hybride obtenu avec le même parent «noble» ou un autre, de façon à récupérer un phénotype cultivable devenu plus rustique et plus vigoureux (Fig. 5). Ces premiers hybrides interspécifiques ont permis un progrès génétique considérable.

Au cours de la nobilisation, le clone *S. officinarum*, utilisé comme femelle, transmet la totalité de son nombre somatique de chromosomes ($2n=80$ chromosomes) lors du premier croisement F1 et du premier rétro-croisement. Le mécanisme exact de la transmission $2n+n$ chromosomes n'est pas parfaitement établi (Bhat et Gill, 1985). Plusieurs hypothèses ont été avancées pour expliquer ce phénomène telle que le doublement chromosomique (endoduplication) ou encore la fusion de deux nuclei après la seconde division de meiose. Dans tous les cas, ce phénomène

provoque une réduction rapide du nombre relatif de chromosomes de *S. spontaneum* à la meïose (Sreenivasan, 1987).

Les premiers hybrides interspécifiques issus de ces travaux de nobilisation ont permis un progrès génétique rapide et ont été par la suite utilisés massivement en croisement dans tous les programmes de sélection à travers le monde. Il s'agit notamment des clones POJ 2878, POJ2725, POJ213 de Java ou les clones Co281, Co290 de Coimbatore.

I.3.3 Maîtrise de la floraison (revue par Moore et Nuss 1987)

La synchronisation des floraisons est une préoccupation essentielle du sélectionneur qui souhaite réaliser les combinaisons parentales de son choix. Ce qui est très difficile lorsqu'il s'agit de croiser des clones à floraison tardive avec des clones à floraison précoce. La floraison est un phénomène saisonnier. Dans l'hémisphère nord les floraisons commencent fin septembre pour les zones équatoriales ou seulement au mois de décembre si on s'éloigne vers 30° de latitude nord. Dans l'hémisphère sud on observe le même type de progression (de mars à juin).

Les facteurs qui influencent la floraison sont nombreux. Un des plus important est la photopériode. L'induction florale peut être avancée en soumettant artificiellement les plantes à une diminution de la photopériode et cela avant que cette diminution ne se produise naturellement. En zone équatoriale (ou la durée du jour est quasi constante et ne dépasse pas 12,5 heures), il est aussi possible d'avancer la floraison en illuminant les plantes 20 minutes avant le levée du soleil puis en réduisant graduellement cet apport de lumière. Inversement, pour retarder la floraison, il faut apporter artificiellement un complément de lumière qui compense la diminution de la photopériode (qui induirait la floraison).

La floraison peut pénaliser le rendement. En effet, la transition du méristème apical vers la production d'une inflorescence entraîne un arrêt de la croissance végétative. La floraison est associée à un effet de maturation (accumulation de saccharose dans la tige). En général, la teneur en sucre va augmenter au moment de la floraison pour diminuer ensuite (ce qui intensifie encore la perte de rendement due à l'arrêt de la croissance). La floraison sera moins pénalisante pour des clones de début de campagne (qui mûrissent précocement). Il y a donc

une certaine antinomie entre la nécessité d'avoir des fleurs pour la création variétale et la nécessaire contre-sélection des clones qui fleurissent beaucoup.

La fertilité pollinique de la canne à sucre est très variable. Un certain nombre de clones sont mâles stériles et doivent donc obligatoirement être utilisés comme femelle dans les croisements. La castration des clones mâles fertiles, qui permet un contrôle total sur les combinaisons de parents (traitement à l'eau chaude à 50°C par exemple) peut être utilisée dans certains cas. Le contrôle de la fertilité pollinique se fait par observation des anthères ou par coloration des grains de pollen pour contrôler leur viabilité.

I.3.4 Les méthodes de sélection

L'objectif de l'amélioration variétale est l'augmentation de la rentabilité économique de la culture. C'est à dire de produire le maximum de sucre à l'hectare. Les critères de sélection sont donc essentiellement le rendement, la richesse en sucre. La résistance aux maladies et ravageurs sont aussi des critères recherchés puisqu'ils permettent de sécuriser le rendement.

La plupart des schémas de sélection reposent sur une démarche récurrente dans laquelle les meilleurs clones issus de la sélection locale, dits élites, sont utilisés comme géniteurs. Néanmoins les ressources génétiques utilisées en hybridation sont continuellement enrichies par des variétés étrangères dans le cadre d'échanges internationaux.

Les interactions génotype x environnement sont fortes. En conséquence, la sélection s'effectue le plus tôt possible dans l'environnement ciblé.

A la Réunion, par exemple, la sélection s'opère dans cinq stations correspondant à cinq environnements contrastés: Vue-Belle (zone de moyenne altitude sèche de l'ouest à 600-700m et hauts de l'ouest pour la station de Vue Belle II située à 1000m d'altitude), Etang salé (zone sèche du littoral sud), Sainte Marie (zone littorale sud-est), Saint Benoit (zone très humide de l'est), Saint Philippe (zone volcanique très arrosée >4m d'eau dans le sud-est)

Le choix des géniteurs est raisonné en fonction de leur potentiel génétique. Celui-ci est défini sur la base de différents critères:

- la valeur propre des géniteurs
- la valeur générale en croisement des géniteurs ('proven parent')

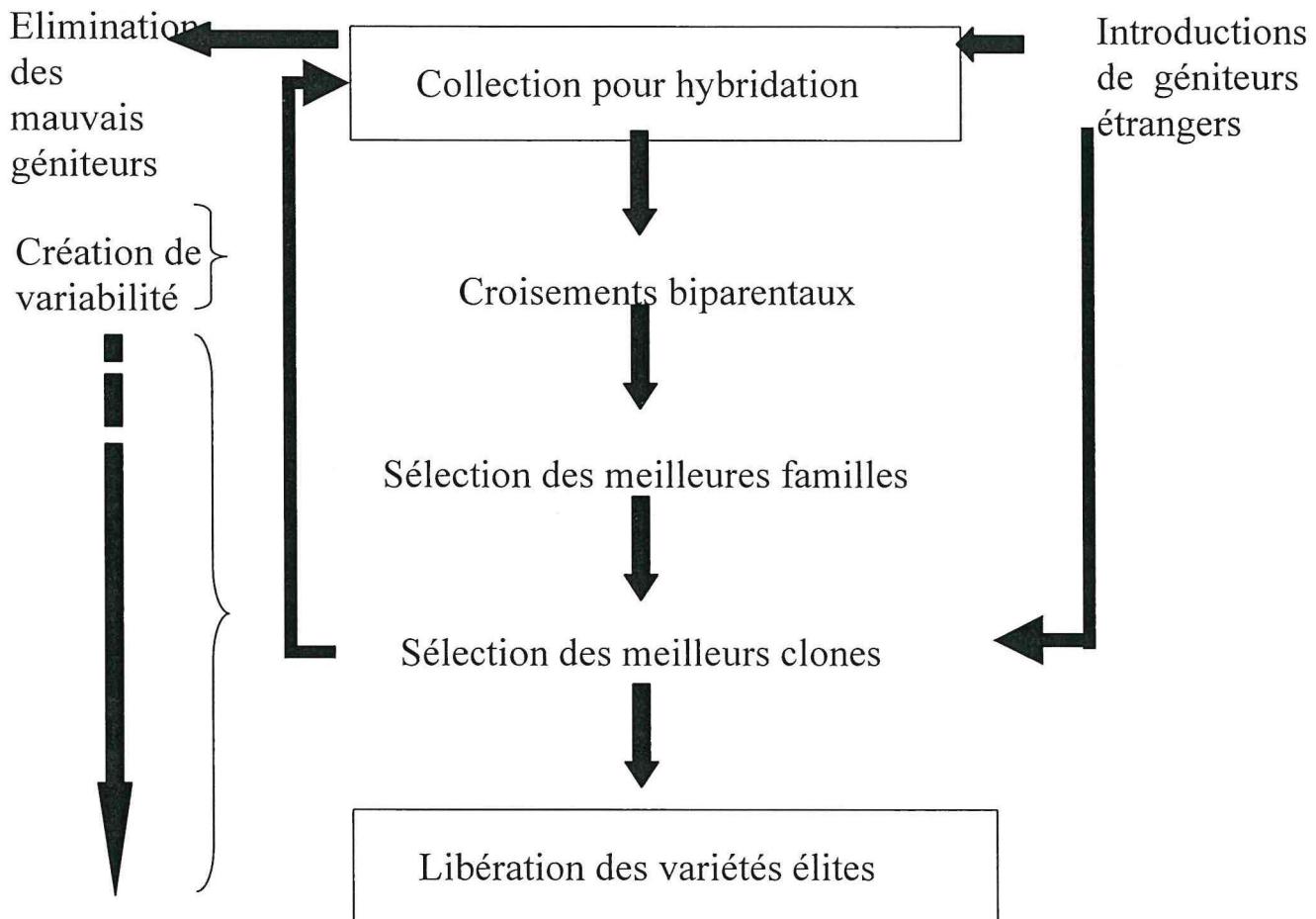


Figure 6: Schémas de sélection en Australie (Cox et al., 2000)

Tableau 3: Schémas de sélection à la Réunion du CERF (Centre d'Essai de Recherche et de Formation)

Année	0	1	2 3	4 5	6 7	8 9 10	11 12 13 14	13 14 15
Stade de sélection	Hybridation Semis	Pépinières	"1 touffe"	"1 ligne"	"2 lignes"	"3 lignes"	Essais régionaux multilocaux	Pré-libération Libération
Nombre de variétés	± 1 200 croisements 1 à 2 Millions	200 000	80000	5000	350	80 à 100	15 à 30	0 à 2
Parcelle élémentaire	1 graine	1 plantule	1 touffe	1 ligne de 5 mètres	2 lignes de 5 mètres 3 répétitions	3 lignes de 10 mètres 3 répétitions -1 essai récolté en début de saison -1 essai récolté en fin de saison	3 lignes de 10 mètres 5 répétitions	1 à 2 Ha
Sites de sélection	La Bretagne	Sur 3 stations CERF : La Mare, Etang Salé, Vue Belle	Sur 4 stations CERF : La Mare, Saint Benoit, Etang Salé, Saint Philippe	Sur 5 stations CERF : La Mare, Saint Benoit, Etang Salé, Saint Philippe, Vue belle	Sur 5 stations CERF : La Mare, Saint Benoit, Etang Salé, Saint Philippe, Vue belle	Sur 5 stations CERF : La Mare, Saint Benoit, Etang Salé, Saint Philippe, Vue belle	Essais chez les agriculteurs et sur les 5 stations du CERF	Echelle semi-industrielle sur les grosses exploitations + Diffusion pépinière CTICS
Critères de sélection	-Programme d'hybridation suivant étude des descendants et pool génétique à disposition	-Sélection naturelle : compétition entre les semis (Aspect végétatif, maladies)	-Sélection en repousse (Aspect végétatif, maladies et BRIX)	-Observation et sélection en repousse (Aspect végétatif, maladies, et BRIX)	-Sélection sur vierge et 1 ^{re} repousse (Aspect végétatif, maladies, rendement, richesse et indice de rendement économique)	-Sélection sur vierge et 2 repousses (Aspect végétatif, maladies qualités agronomiques, rendement, richesse et indice de rendement économique)	-Sélection sur vierge et 3 repousses (Adaptation aux diverses zones pédoclimatiques, aspect végétatif, maladies qualités agronomiques, rendement, richesse et indice de rendement économique)	-Thermotherapie sur bouture -Appréciation des planteurs -Confirmation des qualités industrielles

-la valeur spécifique de certaines combinaisons de géniteurs ('proven cross') accessible uniquement dans les croisements bi-parentaux contrôlés. Dans le cas de pollinisations libres ou de type polycross (croisement contrôlé mais en utilisant différents géniteurs mâles polliniseurs) on accède uniquement à la valeur générale des géniteurs femelles testés.

Les géniteurs les plus intéressants sont ceux dont le maximum de descendants atteindront le stade de sélection le plus avancé dans le cadre d'une sélection massale. En revanche, dans le cadre d'une sélection familiale on compare directement la valeur des familles (Skinner *et al.* 1987). Deux exemples de schémas de sélection sont présentés (tableau 3, Fig. 6): la sélection massale réalisée par le Centre d'Essai, de Recherche et de Formation (CERF) à la Réunion et la sélection sur une base familiale pratiquée au BSES en Australie (Bureau of Sugar Experiment Stations).

I.4 Etude du génome des cultivars modernes de canne sucre

I.4.1 Structure du génome

La canne à sucre a un génome de grande taille principalement à cause de son haut niveau de polyploidie. La taille du génome de R 570 a été estimée à 10 000 Mb et la taille de la carte génétique saturée de R 570 est estimée autour de 17 000 cM. Néanmoins, la taille du génome monoploïde de la canne à sucre est sensiblement équivalente à celle du sorgho et de l'ordre de deux fois plus grande que celle du riz. Le génome monoploïde de *S. officinarum* comprend 930 Méga paires de bases (0.96 pg) et celui de *S. spontaneum* 750 Méga paires de bases (0.78 pg) (D'Hont et Glaszmann 2001). Le génome des cultivars modernes de canne à sucre est caractérisé par son haut niveau de ploidie ($2n = 100\text{-}130$ chromosomes) et sa fréquente aneuploidie. Chaque chromosome de base est potentiellement présent en dix à douze exemplaires homo(eo)logues. L'hybridation *in-situ* (GISH) a permis de mettre en évidence la double structure du génome des cultivars. Pour cela l'ADN total de chacune des espèces ancestrales (*S. spontaneum* et *S. officinarum*) est hybride sur les préparations chromosomiques des hybrides mais révélé chacun par un fluorochrome différent. Chacune des sondes marque l'ensemble des chromosomes mais avec une intensité variable. C'est la superposition des deux marquages qui permet de repérer l'origine spécifique des chromosomes. Le génome du cultivar moderne 'R 570', par exemple, est constitué de 80% de

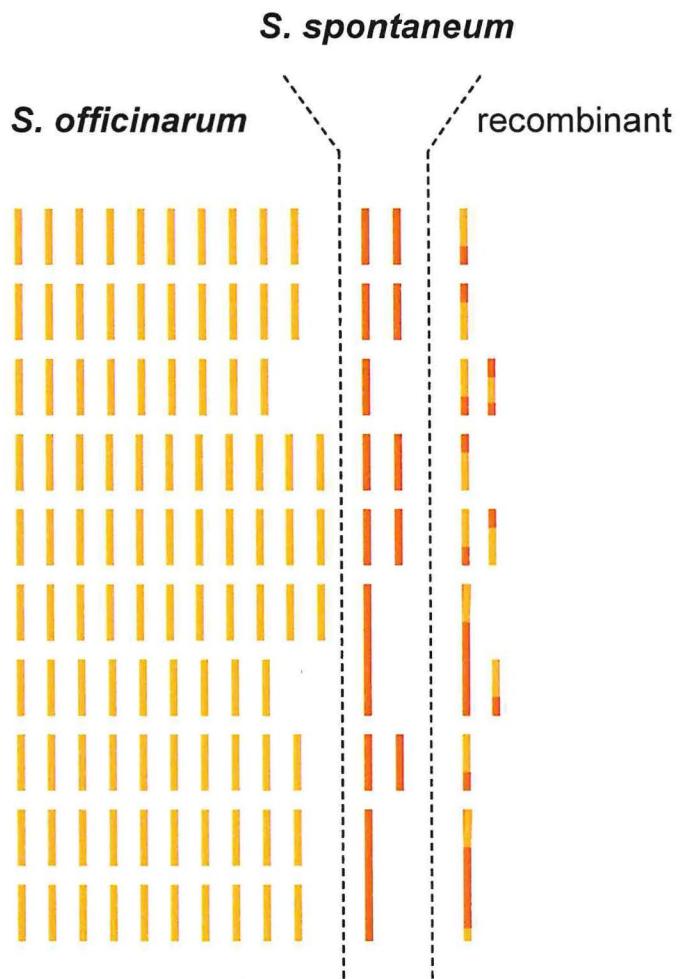


Figure 7: Représentation schématique du génome d'un cultivar moderne de canne à sucre d'après les études réalisées par hybridation in-situ (FISH et GISH). Les barres jaunes et oranges correspondent aux chromosomes ou aux segments de chromosomes hérités de *S. officinarum* et *S. spontaneum* respectivement (D'Hont, 2005).

chromosomes issus de *S. officinarum*, de 10% de chromosomes hérités de *S. spontaneum* et 10% de chromosomes issus de recombinaisons entre chromosomes des deux espèces ancestrales (D'Hont *et al.* 1996) (Fig. 7).

I.4.2 Cartographie génétique

Des cartes génétiques partielles ont été réalisées sur les espèces ancestrales des cultivars modernes: *S. spontaneum* (Ming *et al.* 1998 et 2002; Al Janabi *et al.* 1993 ; Da Silva *et al.* 1995) et *S. officinarum* (Mudge *et al.* 1996 ; Ming *et al.* 1998 et 2002 ; Guimaraes *et al.* 1997,1999). Des travaux de cartographie génétique ont aussi été menés sur les cultivars modernes (D'Hont *et al.* 1994, Grivet *et al.* 1996, Hoarau *et al.* 2001, Rossi *et al.* 2003 et Aitken *et al.* 2005). Néanmoins, il n'existe aucune carte génétique saturée du génome d'une variété moderne à ce jour. La carte la plus complète est celle du cultivar 'R 570' qui compte 128 groupes de coségrégation rassemblés en 7 groupes d'homologie (Rossi *et al.* 2003). La longueur cumulée de ces groupes représente 7800 cM, à comparer avec la taille estimée du génome de 17 000 cM. De plus, le marquage du génome est toujours très inégal. Les chromosomes d'origine *S. spontaneum* sont couverts plus densément que les chromosomes d'origine *S. officinarum* (Grivet et Arruda 2001). La cartographie génétique confirme la double structure du génome de la canne à sucre. La part du génome héritée de *S. spontaneum* est d'environ 20% et le nombre de chromosomes non recombinés issus de *S. spontaneum* est faible dans les cultivars modernes (peut être 6 chez 'R 570' selon Hoarau *et al.* 2001) Par ailleurs, un groupe d'homologie (VIII) chez 'R 570' est constitué de deux grands chromosomes issus de *S. spontaneum* homologues à deux séries de petits chromosomes issus de *S. officinarum* (Grivet *et al.* 1996; Rossi *et al.* 2003). C'est une conséquence de la différence de structure entre les génomes des deux espèces ancestrales: 8 chromosomes de base pour *S. spontaneum* et 10 chromosomes de base pour *S. officinarum* (D'Hont *et al.* 1998 ; Ha *et al.* 1999). Par ailleurs, les travaux de cartographie génétique ont révélés des appariements préférentiels entre certains chromosomes à la meïose avec des fréquences d'appariements très variables et des sous groupes de chromosomes montrant des affinités préférentielles. La meïose d'un cultivar moderne ne s'effectue donc pas selon un schéma tranché tel que la disomie ou la polysomie (Grivet *et al.* 1996; Hoarau *et al.* 2001; Jannoo *et al.* 2004).

Parallèlement à ces travaux de cartographie, des tentatives ont été faites pour identifier les QTL impliqués dans des caractéristiques agronomiques importantes. Pour les espèces ancestrales, on peut citer, les travaux de Mudge et al. (1996) sur la résistance au ‘eyespot’, de Guimaraes et al. (1997) sur la floraison ou ceux de Ming et al. (1998) sur la richesse en sucre. Chez les cultivars modernes, une étude a été réalisée sur les composantes du rendement (Hoarau et al. 2002) et une autre sur le tallage (Jordan et al., 2004). Ces études révèlent un grand nombre de QTL à effets faibles (<10% de la variation phénotypique totale) et instables entre les environnements. Malgré les difficultés que pose la polyploidie pour ce type d’approche, un gène majeur de résistance à la rouille a pu être identifié (Daugrois et al., 1996) et récemment une équipe mauricienne a identifié un QTL majeur pour la résistance au « Yellow spot ».

I.4.3 L’ère de la génomique

Malgré la taille du génome de la canne à sucre, l’analyse détaillée du génome de la canne à sucre a été entreprise. Le projet brésilien SUCEST (<http://sucest.lad.ic.unicamp.br/en/>) a permis de construire une base de données de 250 000 Expressed Sequence Tags (ESTs) provenant de différentes banques construites dans différentes conditions environnementales et à partir de différents organes et tissus. Ces ESTs correspondent à la séquence des extrémités 5’ ou 3’ de 230 000 ADNc clonés (Vettore *et al.* 2003). Ces informations faciliteront l’identification de gènes et l’analyse de leur fonction. Elles permettent l’analyse comparée du génome de la canne à sucre avec les génomes d’autres espèces, à l’intérieur des graminées ou à l’extérieur (Vincentz *et al.* 2004). Cette banque de données est aussi une source pour la découverte de polymorphismes de type Single Nucleotide Polymorphism (SNP) ou Insertion-Deletion (INDEL) (Grivet *et al.* 2003) et microsatellites (Pinto *et al.* 2004) qui peuvent permettre de développer des marqueurs utiles à l’analyse du génome de la canne à sucre

Le projet de clonage du gène de résistance à la rouille cartographié chez le cultivar ‘R 570’ (Daugrois *et al.* 1996) illustre parfaitement la synergie qui est créée par l’exploitation des cartes génétiques, des banques de ressources et de données génomiques constituées pour la canne à sucre mais aussi pour d’autres graminées comme le sorgho et le riz (plante modèle). Cette démarche est favorisée par l’existence de relations synténiques fortes entre la canne à sucre et le sorgho et dans une moindre mesure entre la canne à sucre et le riz (Paterson *et al.* 1995; Dufour *et al.* 1996 and 1997; Glaszmann *et al.* 1997; Guimaraes *et al.* 1997; Ming *et al.*

1998). Avant d'entreprendre la cartographie physique et la marche chromosomique vers le gène, une cartographie fine du gène a été réalisée en plusieurs étapes. La synténie entre les Poacées (Maïs, sorgho et riz) a été exploitée pour saturer le groupe de coségrégation (i.e. le chromosome) portant le gène de résistance avec des sondes RFLP hétérologues (Asnaghi *et al.* 2000). Cela n'a pas permis d'encadrer le gène qui est situé sur une extrémité du chromosome. Un travail de cartographie complémentaire à l'aide marqueurs AFLP ainsi qu'une approche BSA ont par la suite permis de placer des marqueurs de part et d'autre du gène à 2.2 cM en position proximale et 2.5 cM et 1.9 cM en position distale (Hoarau *et al.* 2001; Asnaghi *et al.* 2004). Ces nouveaux marqueurs après clonage et utilisation comme sonde ont permis d'identifier une région homéologue localement inversée sur la carte du Sorgho. Réciproquement, les marqueurs compris dans cette région chez le sorgho ont été cartographiés chez R 570 ce qui a permis de placer des marqueurs à 0.3 cM et 0.6 cM en position distale et proximale respectivement par rapport au gène. Ces marqueurs ont aussi permis d'identifier les contigs de la banque BAC sorgho et de la banque BAC riz couvrant la zone ciblée. La comparaison de la séquence du riz avec les séquences contenues dans la base de données SUCEST a permis d'identifier un certain nombre de marqueurs ou gènes candidats. La population utilisée pour réaliser la cartographie a été élargie de façon à avoir une résolution de 0.14 cM par recombinant. Les marqueurs BAC-ends identifiés chez le sorgho et les marqueurs EST identifiés par comparaison avec le riz ont été cartographiés et ont permis de trouver des marqueurs qui co-localisent avec le gène résistance à la rouille. La marche chromosomique peut donc commencer en utilisant la banque BAC construite à partir du cultivar ‘R 570’ (Tomkins *et al.* 1999). La difficulté chez le haut polyploïde qu'est la canne à sucre est d'obtenir un contig complet de l'haplotype contenant le gène cible (D'Hont *et al.* 2001; Lecunff *et al.* 2004).

I.5 Pathogènes et ravageurs de la canne à sucre

I.5.2 Les principaux ravageurs

Trois grands groupes d'insectes peuvent provoquer des dégâts importants sur la canne à sucre (Pour revue, Williams *et al.* 1969 ou Fauconnier et Bassereau 1970) :

Les lépidoptères foreurs de tige:

Ce groupe d'insecte dont la chenille se développe à l'intérieur des tiges, peut occasionner des pertes économiques importantes à travers le monde. On peut citer les espèces *Chilo sacchariphagus* (d'origine asiatique mais qui est présent dans l'océan indien et qui se repand sur le continent africain, notamment au Mozambique), *Eldana saccharina* (en Afrique) et *Diatrea* ssp. (continent américain).

Les piqueurs suceurs:

Ils sont favorisés par des conditions climatiques particulières (ensoleillement et irrigation). On peut citer notamment par ordre d'importance::

- Perkinsiella saccharicida* (Hemiptera, Delphacidae) qui est le vecteur de la maladie de Fidji.
- Les pucerons vecteurs des virus de la mosaïque et du syndrome de la feuille jaune (*Rophalosiphum maidis*, *Melanaphis sacchari...*)
- Les cochenilles telles que *Aulacaspis tegalensis* qui peut entraîner des encroûtements spectaculaires et le dessèchement des tiges

Les ravageurs du système racinaire:

Ils appartiennent principalement à l'ordre des coléoptères. En Australie, où on en recense plus de dix espèces, et en Amérique du nord ils posent des problèmes de façon plus ou moins localisée.

*A la Réunion, les trois groupes d'insectes sont présents. Le foreur ponctué, *Chilo sacchariphagus*, introduit d'Asie via l'île Maurice est signalé pour la première fois à la Réunion en 1855. Il peut provoquer des pertes significatives en fonction de la sensibilité des variétés cultivées. Ainsi le remplacement des variétés 'S 17' et 'M 134/32' par 'R 570', résistante, a grandement diminué l'impact de ce ravageur. Mais récemment, ce problème a repris de l'ampleur avec l'augmentation des surfaces plantées avec la variété sensible 'R 579'. Une lutte biologique est donc envisagée contre ce foreur reposant sur l'élevage intensif et l'épandage d'une guêpe parasitoïde, *Trichogramma chilonis* (Soula, 2003). Le ver blanc (*Hoplochelus marginalis*) a causé d'énormes problèmes dans les années 1980, suite à son introduction accidentelle de Madagascar. Une lutte biologique utilisant un champignon entomo-pathogène (*Beauveria Brongnartii*) a été mise en place. C'est un des rares exemples de succès d'une lutte micro-biologique de ce type (Neuvéglise et al., 1994).*

I.5.2 Les principales maladies

De nombreuses maladies ont été recensées pour la canne à sucre à travers le monde (Pour revues, «A guide to sugarcane diseases», Rott P., Bailey R. A., Comstock J.C., Croft B.J., Salem Saumtally A. Eds. 2000 et «Diseases of sugarcane», Ricaud C., Egan B.T., Gillaspie A.G., Hugues C.G. Eds 1989) qui peuvent toutes dans certaines conditions environnementales provoquer des dégâts sur des cultivars sensibles. Néanmoins, les maladies citées ci-dessous représentent un problème plus général.

Pour les maladies fongiques:

- Le charbon de la canne à sucre qui est évoqué plus en détail dans le chapitre suivant.
- Le mildiou (downy mildew) dû à *Peronosclerospora sacchari* et *P. philippinensis* (oomycètes qui désormais ne sont plus classés avec les champignons mais avec un nouveau groupe d'organismes appelé Straménopiles. Les straménopiles constituent un des 5 règnes eukaryotes) est une maladie importante dans le Pacifique et en Asie du sud-est. Les symptômes sont des stries jaunâtres et la production d'un duvet blanchâtre à la face inférieure des feuilles.
- La pourriture rouge ou morve rouge (red rot) due à *Glomerella tucumanensis* (ascomycètes) est très répandue. Elle est caractérisée par des nervures médianes qui se colorent en rouge et par la coloration en rouge des tissus internes de la tige.
- Les rouilles brunes et orangées (brown rust et orange rust) provoquées respectivement par *Puccinia melanocephala* et par *P. Kuehnii* (Basidiomycètes)

Pour les maladies bactériennes:

- Le rabougrissement des repousses (ratoon stunting disease) dû à *Leifsonia xyli* (anciennement *Clavibacter xyli*) qui est une bactérie ‘fastidieuse’. Les symptômes classiques sont une coloration rouge des faisceaux vasculaires provoquant un ralentissement de la croissance de la canne.
- L'échaudure des feuilles (Leaf scald disease) due à *Xanthomonas albilineans* qui est caractérisée par l'apparition de lignes blanches bien dessinées sur les feuilles. Ces stries peuvent évoluer dans les cas plus graves vers un blanchiment du limbe et même jusqu'à la nécrose des feuilles. La tige prend alors un aspect échaudé particulier avec les feuilles nécrosées qui se recourbent vers l'intérieur. Les faisceaux vasculaires peuvent présenter une couleur rouge.



Figure 8: symptômes dûs au charbon de la canne à sucre

1 touffes de canne à sucre saines (variété 'R579')

2 touffes de canne à sucre infestées par le charbon

3 « buisson » de fouets charbonneux

4 détail d'un fouet charbonneux

5 prolifération de bourgeons axillaires formant un « balai de sorcière » (symptôme atypique)

-La gombose (gumming disease) due à *Xanthomonas axonopodis* pv *vasculorum*. qui est caractérisée par l'exudation de gommes à partir des tiges coupées. Les symptômes foliaires sont des lignes crèmes tachetées de rouge qui peuvent évoluer vers des nécroses. Lorsque la bactérie est installée de façon systémique dans la plante on peut aussi observer des chloroses du feuillage comme dans le cas de l'échaudure.

Pour les maladies virales:

-La mosaïque due à un Potyvirus qui provoque un faciès particulier sur les feuilles avec des variations entre des zones vertes et des zones jaunes. Les zones cholorotiques ont généralement des bordures diffuses.

-Le syndrome de la feuille jaune (Yellow leaf syndrom) du à un lutéovirus, parfois en association avec un phytoplasme, est caractérisée par le jaunissement plus ou moins important de la nervure centrale des feuilles.

-La maladie de Fidji due à un Fijivirus est présente essentiellement en Océanie, dans le Pacifique et en Asie du sud-est. Elle est caractérisée par la formation de galles foliaires.

*A la Réunion, la plupart des maladies sont présentes mais la plupart d'entre elles se cantonnent aux parcelles de sélection du CERF car les variétés cultivées à la Réunion (R 570, R 579, R 575 ou R 777) y sont résistantes. Deux maladies fongiques importantes ne sont pas présentes à la Réunion, la rouille orangée et le mildiou, ainsi qu'une maladie virale, la maladie de Fiji. Dans le cas de cette dernière, l'insecte vecteur du virus est présent à la Réunion et R 570 est sensible à la maladie. Un contrôle sanitaire strict doit donc être opérée aux frontières avec des échanges de matériel qui doivent impérativement se faire via les quarantaines. L'évolution de deux maladies doit être surveillée à la Réunion: celle du syndrome de la feuille jaune qui est présent sur toutes les variétés cultivées et celle du Yellow spot (*Mycovellosiella koepkei*), un champignon qui se développe dans des conditions particulières d'humidité et de forte nébulosité et auquel R 570 et R 579 sont assez sensibles.*

I.5.3 Le charbon de la canne à sucre

Histoire et distribution:

Le charbon de la canne à sucre est une des maladies les plus importantes au monde pour cette culture. Tous les grands pays producteurs ont eu à faire face à ce problème. Jusque dans les années cinquante, le charbon ne concernait que les pays d'Asie et en Afrique essentiellement dans le Natal et en Afrique de l'est. La maladie ne s'est ensuite répandue en Afrique de l'ouest qu'au cours des années soixante-dix. Hawaii, les Caraïbes et l'ensemble de l'Amérique ont

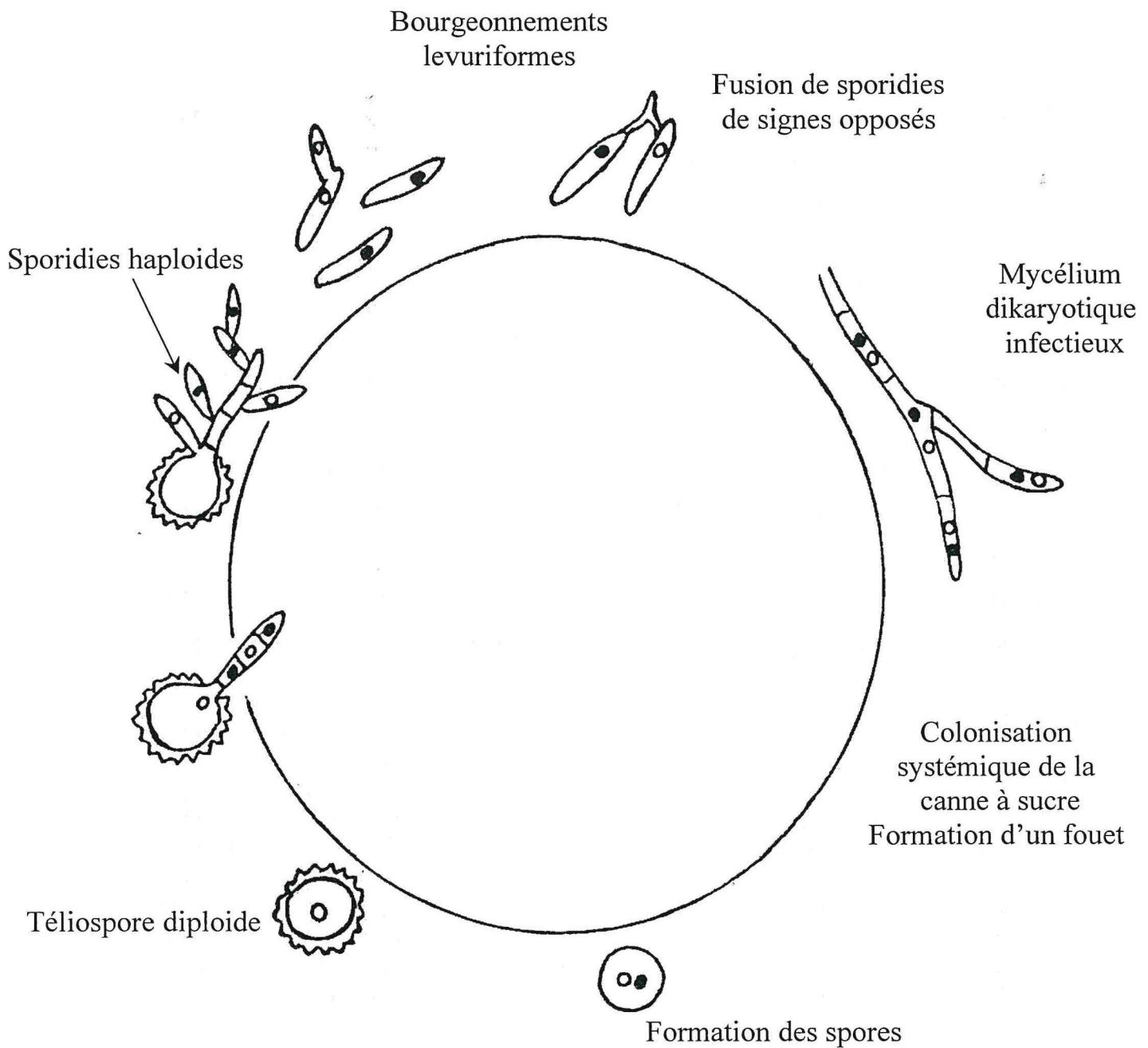


Figure 9: Cycle biologique de *Ustilago scitaminea*

ensuite été progressivement contaminés à partir de la fin des années 70. Un foyer de charbon avait toutefois éclaté dès 1948 en Argentine, contaminant de proche en proche le sud du Brésil, le Paraguay et la Bolivie mais sans toutefois s'étendre au reste du continent (Antoine 1961; Lee-lowick 1978; Comstock 2000). Seule l'Australie, en dehors d'une zone marginale à l'ouest 'l'Ord river district', Fidji et la Papouasie Nouvelle Guinée sont indemnes de la maladie.

Le charbon est présent à la Réunion depuis longtemps. Il est déjà cité par de Sornay en 1936.

Symptômes:

Le symptôme caractéristique de la maladie est le fouet charbonneux qui se développe à l'extrémité des tiges malades (Fig. 8). Le fouet mesure quelques cm à plus de 1.5 m pouvant alors dominer le couvert végétal. Le sore qui se dégage du faisceau foliaire croît pendant plusieurs semaines et libère de très grandes quantités de spores, 10^8 à 10^9 selon Waller (1969), après que la membrane argentée qui les protégeaient se soit déchirée. Les souches de canne à sucre atteinte par la maladie peuvent avoir un aspect buissonnant avec très peu de tiges développées usinables mais un tallage excessif. D'autres symptômes plus occasionnels peuvent se rencontrer comme l'apparition de galles ou de cals sur les tiges, la prolifération de bourgeons axillaires ou encore des distorsions de la tige (Byther et Steiner , 1974).

Cycle biologique du pathogène et cycle infectieux:

Le charbon de la canne à sucre, *Ustilago scitaminea* Sydow, est un basidiomycète appartenant à l'ordre des Urédinales. Il produit des téliospores diploïdes qui permettent la dissémination de la maladie par le vent. La méiose se produit au moment de la germination de la téliospore qui aboutit à la formation d'un promycélium comportant quatre cellules haploïdes. Chacune de ces cellules forme une sporidie (basidiospore). *U. scitaminea* a un système de reproduction bipolaire diallelique (Alexander et Srinivasan, 1966. Leu, 1978) dans lequel seules les sporidies de signe sexuel opposé peuvent conjuguer. La fusion de sporidies compatibles aboutit à la formation d'un mycélium cloisonné dicaryotique infectieux capable d'infecter les bourgeons de la canne à sucre (Fig. 9).

Les téliospores germent rapidement dans des conditions favorables (humidité et température optimale de 25-30°C (Saxena et Khan 1963)). Au contact de la canne à sucre le promycélium donne directement un hyphe infectieux. Il est à noter toutefois que sur milieux artificiels, les sporidies peuvent adopter un mode de croissance par bourgeonnement de type levure. Le

mycélium infectieux pénètre à travers les écailles des bourgeons et s'installe dans les zones méristématiques. Le pathogène suit le développement de la plante de façon systémique en s'installant préférentiellement dans chacun des méristèmes latéraux formés. La formation du fouet est précédée par une perturbation du mode de croissance des tiges qui produisent alors des entre nœuds plus longs et plus fins. Une prolifération du mycélium est observée autour du méristème apical qui change de mode de fonctionnement et devient intercalaire permettant la croissance continue du fouet (Waller 1969; Lee-Lovick 1978). Le pathogène se développe particulièrement dans les tissus du fouet où les hyphes évoluent physiologiquement et cytologiquement. Les hyphes qui se développent dans les couches superficielles du fouet sont uninucléées (Il y a donc eu caryogamie). Les cellules des hyphes s'agrègent et finissent par former une matrice «gélantineuse» à mesure que les parois des cellules s'hydrolisent. Les spores se forment à partir de cette matrice en agglomérats et ne sont pas connectées directement à un hyphe (Trione 1980). Les téliospores uni-nucléées matures prennent une couleur marron et sont libérées lorsque la membrane extérieure qui couvre le fouet se rompt.

Génétique de la résistance au charbon

Composantes de la résistance

Les caractéristiques morphologiques des bourgeons peuvent contribuer à la résistance en ralentissant l'accès du champignon aux zones meristématiques (Waller, 1970). D'autre part des substances glycosidiques, présentes dans les écailles des bourgeons, identifiables et quantifiables par chromatographie liquide haute performance (Lloyd et Naïdo, 1983), semblent jouer un rôle dans la résistance. La permet de séparer et quantifier ces substances. Il y aurait donc au minimum deux mécanismes intervenant dans l'expression de la maladie : la résistance à la germination et à la pénétration des spores dans les bourgeons et la résistance à la colonisation des tissus par le champignon (Walker, 1987). Certaines méthodes d'inoculation, telle que l'injection d'une suspension de spores dans les jeunes pousses, permettent d'éliminer la première barrière de résistance qui est le franchissement des écailles des bourgeons et d'évaluer la résistance "physiologique" (Burner, 1993).

Déterminisme génétique de la résistance

Les études de génétique quantitative ont montré que l'héritabilité de la résistance au charbon au sens large est élevée et l'héritabilité au sens strict modérée et qu'il est donc possible de sélectionner avec succès pour ce caractère (Wu *et al.* 1988 ; Chao *et al.* 1990 ; Balance et al.

1996). Ceci d'autant plus qu'il n'existe aucune corrélation négative entre la résistance au charbon et les paramètres du rendement : diamètre des tiges, nombre de tiges, longueur des tiges ou Brix (Wu *et al.*, 1983). L'élimination du matériel sensible dans les descendances ne devrait donc pas limiter les gains génétiques sur le rendement. En revanche, l'évaluation systématique du matériel végétal dans les premiers stades de sélection est lourde. Les tests à mettre en œuvre exigent un grand nombre de répétitions par clone (Whittle et Walker, 1982). Récemment, une étude d'expression différentielle (techniques cDNA –AFLP et SSH) a permis de révéler un certain nombre de gènes exprimés spécifiquement dans le cas d'une interaction entre un cultivar résistant et le charbon. Ces gènes seront utilisés ensuite comme marqueurs (Heinze *et al.* 2001).

I.6 Introduction au sujet de thèse

La création variétale a permis un essor considérable de la culture de la canne à sucre grâce à la création des premiers hybrides interspécifiques entre *S. officinarum* et *S. spontaneum* puis à leur utilisation en croisement. Cependant ces travaux se font de manière tout à fait empirique sans la moindre connaissance des gènes (et de leur diversité) impliqués dans les caractères qui font l'objet de la sélection. La génétique moléculaire devrait donc pouvoir dégager des marges de progrès supplémentaires en contribuant à identifier les gènes importants sur le plan agronomique, permettant ainsi de cibler la sélection. C'est dans ce cadre général que se placent l'ensemble des travaux, menés par l'équipe «Connaissance du génome de la canne à sucre» du CIRAD, concernant l'exploration de la diversité génétique, la cartographie du génome ou l'étude la structure de ce génome haut-polyplioïde. La résistance de la canne à sucre aux maladies est une des préoccupations majeures des sélectionneurs. Cependant, on ne connaît pratiquement rien sur le déterminisme génétique des résistances aux maladies, à l'exception notable du gène majeur de résistance à la rouille identifié chez le cultivar 'R570' (Daugrois *et al.* 1996; Asnaghi *et al.* 2004), et du QTL majeur de résistance au yellow spot identifié récemment chez le cultivar mauricien «M 134/75» (Al-Janabi, 2005). Le charbon de la canne à sucre étant une des maladies les plus importantes de la canne à sucre au plan mondial, nous nous y sommes intéressés avec l'objectif de progresser dans la compréhension du déterminisme génétique de la résistance. La démarche expérimentale que nous avons adoptée a été tout d'abord d'étudier la diversité du pathogène puis d'entreprendre à la fois i) un travail de cartographie génétique de QTL, à la Réunion, sur un croisement entre un cultivar

résistant et un clone très sensible au charbon ii) une étude d'association, exploitant le déséquilibre de liaison, sur une population de 74 cultivars caractérisés pour leur résistance au charbon dans des essais menés par la SOSUCO au Burkina Faso.

**II DIVERSITE GENETIQUE DE
USTILAGO SCITAMINEA**

II Diversité génétique de *Ustilago scitaminea* agent du charbon de la canne à sucre

Les évaluations pour la résistance au charbon que ce soit au Burkina Faso et à la Réunion ont été réalisées à partir d'isolats locaux non contrôlés sur le plan génétique. Nous avons donc voulu connaître quelle était la variabilité locale du champignon à l'intérieur et entre ces deux sites. D'une manière générale, la diversité génétique du champignon a été très peu décrite en dépit de quelques tentatives utilisant des marqueurs morphologiques et physiologiques (Lambat et al. 1962; Amire et al. 1982) ou à travers l'étude de la variabilité du pouvoir pathogène (Péros 1981). Seule une étude préliminaire récente à l'aide de marqueurs AFLP a permis de révéler un très faible niveau de variabilité génétique global à l'exception des quelques isolats provenant d'Asie du sud-est (Braithwaite et al. 2004). Nous avons donc profité d'un appel d'offre du Génoscope pour réaliser une banque enrichie en microsatellites à partir de l'ADN génomique de *Ustilago scitaminea* et la faire séquencer. Les marqueurs microsatellites développés nous ont permis de réaliser une étude de la diversité mondiale en y intégrant un échantillonnage d'isolats collectés au Burkina Faso et à la Réunion.

Ce travail est présenté sous la forme d'un article, soumis à la revue *Fungal Genetics and Biology*, intitulé « **Evidence of selfing and rare dispersal events from an Asian centre of diversity in the sugarcane smut fungus, *Ustilago scitaminea*** ».

Evidence of selfing and rare dispersal events from an Asian centre of diversity in the sugarcane smut fungus, *Ustilago scitaminea*

Louis-Marie Raboin^{1*}, Athiappan Selvi^{2*}, Karine Miranda Oliveira³, Florence Paulet⁴, Caroline Calatayud⁴, Marie-Françoise Zapater⁵, Philippe Brottier⁶, Olivier Garsmeur⁴, Jean Carlier⁵ and Angélique D'Hont⁴

*The first two authors contributed equally to this paper

¹.CIRAD (Centre de coopération internationale en recherche agronomique pour le développement) - UMR PVBMT, Pôle de Protection des Plantes, Ligne Paradis, 97410 Saint-Pierre, Réunion, France

². Sugarcane Breeding Institute, Coimbatore, India

³. Universidade Estadual de Campinas, Brazil

⁴. CIRAD - UMR 1096, PIA, TA40/03, Avenue Agropolis, 34398 Montpellier cedex 5, France

⁵. CIRAD - UMR BGPI, TA 41 / K, Campus International de Baillarguet

34398 Montpellier Cedex 5, France.

⁶. Genoscope, Centre National de Séquençage, 2 rue Gaston Crémieux, 91057 Evry, France.

Abstract

Ustilago scitaminea, which causes sugarcane smut disease, has been spreading throughout Africa and America since the 1950s. The overall genetic structure of this fungus was investigated. A total of 142 single-teliospore isolates from 15 countries worldwide were analysed using 17 polymorphic microsatellite loci. All isolates but one were homozygous for all loci, indicating that selfing is the highly preferential reproductive mode of *U. scitaminea*. In America and Africa, genetic diversity was found to be extremely low and all isolates belonged to a single inbred lineage. This inbred lineage was also found in some parts of the Asian continent where most *U. scitaminea* genetic diversity was detected. These observations support the hypothesis that the fungus originated in Asia. The strong founder effect observed in the global genetic structure of *U. scitaminea* suggests that the fungus migrated from Asia to other continents on rare occasions.

Index descriptors:

Sugarcane, smut, *Ustilago scitaminea*, population genetic structure, inbreeding, fungus dispersal, microsatellites

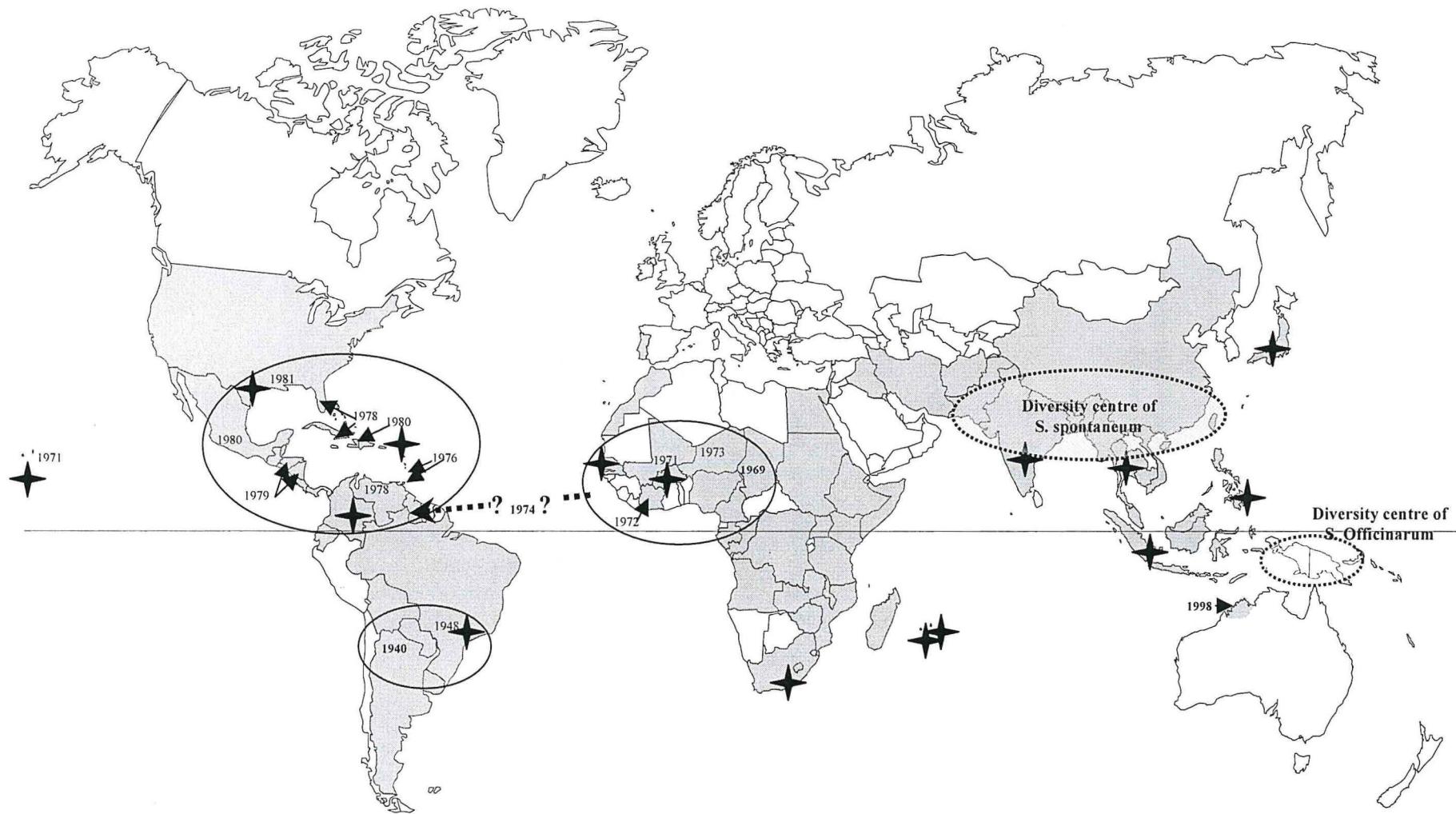


Figure 1: Worldwide distribution of *Ustilago scitaminea* (in gray) according to Antoine (1961), the Commonwealth Mycological Institute (1977) and Comstock (2000). In West Africa, the disease was first detected in 1969 at Banda, near Fort Archambault (Chad). From there, cuttings were sent to Upper Volta (now Burkina Faso), triggering a new epidemic (Barat, 1971). Later, smut was found in Mali and Côte d'Ivoire in 1972, Cameroon and Niger in 1973 (Delassus, 1975). The disease first occurred in South America in 1940, in Tucuman province, Argentina (Hirschhorn, 1949), and rapidly spread throughout the sugarcane-growing areas. The disease later was observed in Paraguay in 1944 (James, 1978) and in São Paulo state of Brazil in 1948 (Veiga, 1972). Smut was also present in Bolivia in 1957, but no further extension of the epidemic was observed on the American continent until being detected in the Caribbean state of Guyana in 1974 (Bates, 1975), and it then successively appeared throughout the West Indies, Central America and mainland USA: Martinique (1976), Trinidad (1976), Jamaica (1976), Guadeloupe (1978), Belize (1978), Venezuela (1978), St Kitts (1978), Cuba (1978), Florida (1978), Honduras (1979), Nicaragua (1979), Barbados (1979), Haiti (1980), Mexico close to the Belize border (1980), Puerto Rico (1981) and Louisiana (1981) (James, 1976; Flores, 1978; Flores, 1980; Liu, 1981). Before that, smut had also been reported in 1971 on the Hawaiian island of Oahu (Ferreira and Comstock, 1989). Australia was recently hit, in 1998, probably due to wind blown spores originating from Indonesia (Riley *et al.*, 1999). ♦ Black stars indicate locations from which *U. scitaminea* teliospores were collected for analysis. Areas where recent outbreaks of the disease occurred in America and West Africa are circled. The approximate diversity centers of *Saccharum spontaneum* and *S. officinarum* are circled with dotted lines.

II.1 Introduction

Smut, caused by the fungus *Ustilago scitaminea* Sydow, is a major sugarcane disease that can cause considerable yield loss. Smut is easily identified by the characteristic whip-like sorus that forms from the apex of infected stalks. The disease now occurs in nearly all sugarcane-producing countries. Prophylactic measures such as roguing diseased stools or planting healthy material have proved insufficient in preventing epidemics. The use of systemic fungicides is not cost-effective, except for seed treatment, but in such cases the crop protection efficacy decreases over successive ratoons (review by Lee-Lovick, 1978). In the end, the use of resistant cultivars has proved to be the only efficient way to control the disease. Smut resistance is a highly heritable trait in the broad sense and a fairly heritable trait in the narrow sense (Chao *et al*, 1990; Wu *et al*, 1988).

Sugarcane smut was first discovered in Natal in 1877 and identified as *Ustilago sacchari* Rabenh by the Mycological Institute at Kew, and later renamed *Ustilago scitaminea* Sydow. Then it was reported in many countries of Asia and East Africa in the late 19th and early 20th centuries (review by Antoine 1961). The distribution and most recent outbreaks of the disease on the American continent and in West Africa are presented in Figure 1. Papua New Guinea and Fiji are the only sugarcane-producing countries still free of the disease. In Australia, the disease was first reported in 1998 in a marginal production area on the west side but the main producing area (east coast) remains free of the disease. Historically, typical sugarcane smut patterns have involved severe epidemic outbreaks, sometimes followed by long periods during which the disease has remained latent. Smut recurred in Taiwan in 1964 after a few decades (Leu and Teng, 1972) and also on Okinawa island in Japan in 1972 after a lapse of 40 years (Yamauch, 1978).

Smut is a basidiomycetes fungus belonging to the Ustilaginales order, like other smut fungus plant pathogens. Ustilaginales (Martinez-Espinosa *et al*, 2002) produce diploid spores or teliospores. When it germinates, the teliospore undergoes meiosis and gives rise to a septate promycelium bearing four haploid sporidia (basidiospores). *Ustilago scitaminea*, like most parasitic heterobasidiomycetes (Bakkeren and Kronstad, 1994), has a diallelic bipolar mating system (Alexander and Srinivasan, 1966; Leu, 1978) in which only sporidia of opposite

mating types conjugate. Of the four initial sporidia or basidiospores from each teliospore, two have a positive mating allele and two have a negative mating allele. *U. scitaminea* can thus both self and outcross, but the natural selfing *versus* outcrossing frequency is unknown. A dikaryotic mycelium develops after fusion of compatible sporidia. This dikaryotic mycelium is infectious, penetrates behind bud scales and invades the meristematic zone of the bud. When the bud is germinating, the mycelium systematically colonizes the plant by growing in association with each developing bud primordium (Ferreira and Comstock, 1989). Finally, the apical meristem of smut-infected cane produces a long whip-like structure bearing billions of teliospores (i.e. sorus). The whip is covered by a thin silvery membrane which soon ruptures and the spores disperse via the wind.

Knowledge on the genetic structure and evolutionary potential of pathogens could be useful for drawing up guidelines for breeding programs and management of plant resistance (Leung *et al.*, 1993, McDonald and Linde, 2002). Level and distribution of pathogen diversity can be estimated through analyses of population structures at different geographical scales using molecular markers. These analyses, combined with studies on pathogenic variation, should allow more accurate assessment of germplasm resistance. Population structure analysis could also help identify potential sources of resistance, which are expected in areas where genetic diversity of both pathogen and host is maximal. Finally, these analyses may be used to evaluate the relative importance of factors such as genetic recombination and gene flow during the evolution of these pathogens. These factors, related to the reproduction system and dispersal processes, may condition the choice of resistance breeding and deployment strategies (McDonald and Linde, 2002). Ustilaginales have been poorly investigated in this respect, although they are important plant pests and considered as a model system (Martinez-Espinoza *et al.*, 2002).

In *U. scitaminea*, the lack of reliable morphological and physiological markers has hampered description of its genetic diversity, despite a few attempts (Lambat *et al.*, 1962; Amire *et al.*, 1982). Only recently, a preliminary study using molecular markers (AFLP) gave the first insight into the worldwide diversity of *U. scitaminea* (Braithwaite *et al.*, 2004). Thus, in addition to the description of the global population structure of *U. scitaminea* on the basis of a larger number of isolates, the objective of the present study was to infer the reproduction system, the centre of origin and dispersal processes of this fungus on a global scale.

Table 1: Origin of the 142 *Ustilago scitaminea* isolates studied.

Country	Locality	Variety	Contact	Nº isolate	n ^a
America					
Brazil	Piracicaba	IAC 66/6	Copersucar Technologie Center / M. V. Casagrande	031	1
Brazil	Piracicaba	NA 56/79	Copersucar Technologie Center / M. V. Casagrande	033	1
Brazil	Piracicaba	SP 70/11343	Copersucar Technologie Center / M. V. Casagrande	030	1
Brazil	Piracicaba	SP 80/185	Copersucar Technologie Center / M. V. Casagrande	032	1
Brazil	Jau	Pool	Copersucar Technologie Center / M. V. Casagrande	034	1
Colombia	Cenicana	CC 84/75	CENICANA / J.C. Angel	046	1
Colombia	Cenicana	CO 421	CENICANA / J.C. Angel	047	1
Colombia	Cenicana	CP 57/603	CENICANA / J.C. Angel	048	1
Colombia	Central sicarare	B 61	CENICANA / J.C. Angel	050	1
Colombia	Central sicarare	CC 89/452	CENICANA / J.C. Angel	049	1
Colombia	Central sicarare	CC89/43	CENICANA / J.C. Angel	052	1
Colombia	Central sicarare	PH 5333	CENICANA / J.C. Angel	051	1
Colombia	Cucuta	B 49/191	CENICANA / J.C. Angel	054	1
Colombia	Cucuta	CO 421	CENICANA / J.C. Angel	055	1
Colombia	Cucuta	RD 75/11	CENICANA / J.C. Angel	053	1
Guadeloupe	Roujol	HJ 57/41	CIRAD / J.H. Daugrois	007	1
Hawai	Kauai	H 77/4643	Hawaii agriculture Research Center / S. Schenck	004	1
Hawai	Maui Island	H 78/7750	Hawaii agriculture Research Center / S. Schenck	003	1
Texas	Weslaco	NCo310	Texas Agricultural Experiment Station / J. Da Silva	001	1
					64
Asia					
India	Gorakhpur	Unknown	U.P. Council of Sugarcane Research / G.P Rao	041	1
Indonesia	3rd division	RGM 94/137	GPM Sugar group / S. Lamadji	079	3
Indonesia	5th division	GM 23	GPM Sugar group / S. Lamadji	082	3
Indonesia	5th division	RGM 94/122	GPM Sugar group / S. Lamadji	078	3
Indonesia	ILP factory	Berat	GPM Sugar group / S. Lamadji	015	2
Japan	Nago	Ni 9	JIRCAS / M. Matsuoka	021	2
Japan	Naha	RK 97/35	JIRCAS / M. Matsuoka	020	2
Japan	Naha	RK 97/7020	JIRCAS / M. Matsuoka	017	2
Japan	Naha	RK 97/7032	JIRCAS / M. Matsuoka	019	2
Japan	Naha	RK 97/7065	JIRCAS / M. Matsuoka	018	2
Philippines	Brgy Luy a	84524	Philippine Sugar Research Inst. Foundation / R. Cu	045	2
Philippines	Brgy Luy a	86550	Philippine Sugar Research Inst. Foundation / R. Cu	042	2
Philippines	Brgy Luy a	HOCP 91/555	Philippine Sugar Research Inst. Foundation / R. Luzaran	097	3
Philippines	Brgy Luy a	VMC 95/100	Philippine Sugar Research Inst. Foundation / R. Luzaran	099	3
Philippines	Brgy Luy a	VMC 95/104	Philippine Sugar Research Inst. Foundation / R. Luzaran	098	2
Philippines	Brgy Luy a	VMC 95/262	Philippine Sugar Research Inst. Foundation / R. Luzaran	096	3
Philippines	Hda Cansilayan	Phil 74/64	Philippine Sugar Research Inst. Foundation / R. Luzaran	093	3
Philippines	Hda Florencia	VMC 84/524	Philippine Sugar Research Inst. Foundation / R. Luzaran	085	3
Philippines	Hda Florencia	VMC 86/551	Philippine Sugar Research Inst. Foundation / R. Luzaran	084	2
Philippines	Hda Linay	Phil 74/64	Philippine Sugar Research Inst. Foundation / R. Luzaran	090	2
Philippines	Hda Linay	VMC 86/550	Philippine Sugar Research Inst. Foundation / R. Luzaran	092	1
Philippines	Hda Santa Cruz	VMC 88/354	Philippine Sugar Research Inst. Foundation / R. Luzaran	087	2
Philippines	Hda Tinibahan	Phil 74/64	Philippine Sugar Research Inst. Foundation / R. Luzaran	088	3
Philippines	La Granja	Phil 80/13	Philippine Sugar Research Inst. Foundation / R. Luzaran	091	2
Philippines	MDDC Nursery	VMC 86/550	Philippine Sugar Research Inst. Foundation / R. Luzaran	094	1
Thailand	Ban Bua	K88-92	Mitr Phol Sugarcane Research Center / M. S. Rao	027	2
Thailand	Ban Bua	Marcos	Mitr Phol Sugarcane Research Center / M. S. Rao	025	2
Thailand	Ban Kud Jork	K86-161	Mitr Phol Sugarcane Research Center / M. S. Rao	028	2
Thailand	Nong Kon thai	K84-200	Mitr Phol Sugarcane Research Center / M. S. Rao	026	2
					59
Africa					
Burkina Faso	plot BN2	SP 70/1006	SN SOSUCO / Balaya	114	3
Burkina Faso	plot BS2	Co 64/15	SN SOSUCO / Balaya	112	3
Burkina Faso	plot BS5	PR 1007	SN SOSUCO / Balaya	115	3
Burkina Faso	plot diaraba	Nco 310	SN SOSUCO / Balaya	104	1
Burkina Faso	plot NN12	SP 70/1006	SN SOSUCO / Balaya	110	3
Burkina Faso	plot NS16	R 570	SN SOSUCO / Balaya	113	3
Burkina Faso	plot NS28	Co 64/15	SN SOSUCO / Balaya	111	3
Burkina Faso	plot PK1 SNSo	Co 64/15	SN SOSUCO / Balaya	109	1
Burkina Faso	plot PL4	R 570	SN SOSUCO / Balaya	108	3
Burkina Faso	plot PL4 SnSo	PR 1007	SN SOSUCO / Balaya	105	3
Burkina Faso	plot PS1	PR 1007	SN SOSUCO / Balaya	106	3
Mauritius	Belle vue	M 1030/71	MSIRI / A. Dookun-Saumtally	118	1
Mauritius	Pamplemousse	M 376/64	MSIRI / A. Dookun-Saumtally	119	3
Mauritius	Petite rivière	M 1030/71	MSIRI / A. Dookun-Saumtally	117	3
Réunion	La Mare	Unknown	CERF / J.Y. Hoarau	056	1
Réunion	La Mare	Unknown	CERF / J.Y. Hoarau	057	1
Réunion	Gol 343	R 00/122	CERF / J.Y. Hoarau	070	2
Réunion	La Bretagne	R 85/1238	CERF / J.Y. Hoarau	065	2
Réunion	Saint Benoit 111	R 88/215	CERF / J.Y. Hoarau	058	2
Réunion	Ligne Paradis	Unknown	CERF / J.Y. Hoarau	075	3
Réunion	Saint Benoit 112	R 90/4143	CERF / J.Y. Hoarau	067	2
Réunion	ST Phillippe 415	R 83/267	CERF / J.Y. Hoarau	074	2
Réunion	ST Phillippe 415	R 83/267	CERF / J.Y. Hoarau	076	1
Réunion	ST Phillippe 415	R 83/644	CERF / J.Y. Hoarau	076	1
Réunion	Vue Belle 205	R 91/4297	CERF / J.Y. Hoarau	060	2
Senegal	CSS A4	N 14	Compagnie Sucrière Sénégalaise / R. Toll	013	1
Senegal	CSS A7	N 14	Compagnie Sucrière Sénégalaise / R. Toll	014	1
Senegal	CSS A5a	N 14	Compagnie Sucrière Sénégalaise / R. Toll	011	1
South Africa	Pongola	unknown	SASEX / I. Antwerp	024	1

^a number of single spore isolates from the same whip that have been genotyped.

Microsatellite markers were developed for *U. scitaminea* and used to analyse a sample of single-teliospore isolates from various sugarcane-producing countries around the world.

II.2 Material and Methods

Fungal isolates

We surveyed 142 single spore isolates of *Ustilago scitaminea* for genetic diversity (Table 1). The fungal samples were teliospores from 77 single whips collected on various cultivars and at different locations in 15 sugarcane-growing countries throughout the world. Four areas were more represented than others: Réunion (11 whips/7 localities), the Philippines (15 whips/9 localities), Colombia (10 whips/3 localities) and Burkina Faso (11 whips from various plots in the SOSUCO plantation). We were especially interested in the Burkina Faso and Réunion isolates as we are involved in field-screening trials for smut resistance in these regions as part of a project aimed at identifying the genetic factors determining resistance. The Philippines were sampled more intensively because preliminary analyses had shown high diversity among isolates.

Teliospores were germinated by serial dilution and plating on solid agar medium. Single germinating teliospores were isolated and grown on potato dextrose agar at 26°C for 1 week. The white mycelial colonies were further multiplied in order to produce sufficient fungal material. The mycelium was dried overnight at 60°C, stored at -20°C and later used for DNA extraction. The white mycelial-type colonies are made up of dikaryotic hyphae (Trione, 1990). DNA extracts thus represent the diploid stage of the pathogenic fungus.

DNA extraction

Genomic DNA was extracted using a modification of the method described by Rogers and Bendich (1988). Approximately 350 mg of dried mycelium was ground with sand, dispersed in 800 µl of cetyltrimethylammonium bromide (CTAB) extraction buffer (1% CTAB, 0.7 M NaCl, 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1% β-mercaptoethanol, 100 µg/ml

Table 2: Characteristics of the 22 *Ustilago scitaminea* SSR loci analysed
(Ta: annealing temperature, n: number of alleles detected in this study)

Embl accession number	Name	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Repeat motif	Ta(°C)	Product size	n
	mUsCIR1	GTCGCTCGTTCACTT	CTACCTCGCTATCCCTT	(GT) ₁₈	50	175	4
	mUsCIR2	ATAGAAATGGGCTGAAGA	GATTGATACCGAGAGGG	(CA) ₃ A ₅ (CA) ₆	51	192	1
	mUsCIR3	ACTGTAGAACCTGGTGCCT	TTTCTTTTCCCTCTCCC	(GT) ₆	50	309	2
	mUsCIR4	ACAAACCACATCCCCCT	CGTGGGTATTCTACTGCT	(GT) ₆ (GA) ₂	51	196	1
	mUsCIR5	CGATACCATCCACACC	AGTGCCTTAGGTTTTTC	(TGC) ₅	50	309	1
	mUsCIR6	ATCTTGCCTCTCTGTATC	CGGTCAAGGTGTTTC	(CA) ₇	50	183	1
	mUsCIR7	TCGTTCATCCCTCATG	GAGAAAGAGCAGTAACCC	(CT) ₂ G(CT) ₇	49	287	3
	mUsCIR8	TCTCAGGTTATTGAT	AACATACCCCTTTCAC	(GT) ₆ GC(GT) ₃ GC(GT) ₄ AAA(GT) ₃ GC(GT) ₃ GC(GT) ₄ AAA(GT) ₃ GC(GT) ₂	45	204	2
	mUsCIR9	CTCTTGGGTTTCTTTTC	ACATAGACCCACACCTTC	(CT) ₂₃	50	269	9
	mUsCIR10	AGAGCTTAGGAGGCAAC	GCTACTCCCCACATCA	(GT) ₇	50	164	2
	mUsCIR12	AGCCACCACTACTCATT	ACGGTCTTACCCATTCT	(CA) ₂₈	50	205	4
	mUsCIR13	ATTGGAAAGAAAGGAGG	TTTCCGACTGATTACTGA	(GA) ₁₈	50	168	4
	mUsCIR15	CACCACTCTACTTGCT	GCTCGTTGAGGCTAC	(GA) ₆ (GT) ₇	50	121	2
	mUsCIR18	GCTTCTCCTTTTGCTT	ATCCAGCGTTCTCTCTC	(GT) ₁₅	50	270	3
	mUsCIR19	AACCTCGCACTTGATT	TGAAAGGTTGTATTGGG	(CA) ₄ CG(CA) ₆ AA(CA) ₄	49	158	1
	mUsCIR21	ACTACTTAGGCGAGGG	TCCGTGAATGCTTATT	(GT) ₁₇	50	144	4
	mUsCIR29	CATTTCACTCGTCTCTT	ATTTGGTCTGTGTGGTG	(CA) ₁₄	50	203	2
	mUsCIR30	AAGGTGGCGAACAGAG	AATGTAGGGCGAAAGG	(GA) ₇ (GT) ₈ (GA) ₁₂	52	109	3
	mUsCIR32	TCGTTAGGGAAAGGTCT	CGGTTAGGTTGAGTGTG	(GT) ₁₁	50	265	2
	mUsCIR33	GCTCGTTGAGGCTAC	CACCACTCTACTTGCT	(CA) ₆ (CT) ₇	50	121	2
	mUsCIR34	CGATTATTTGTCTATGCTC	TGAGGGTTGTGGTTT	(CA) ₁₁	50	278	3
	mUsCIR38	TTCTTGACTGGCGATT	GATGACTTACCTCTGGCT	(CA) ₅ TACATA(CA) ₉	50	273	2

RNase) and incubated at 65°C for 60 min with continuous gentle rocking. Two successive chloroform:isoamyl alcohol (24:1) purifications were performed. The DNA was precipitated by adding 0.8 v of isopropanol and the DNA pellet was washed with 2 ml of 70% ethanol. The DNA pellet was air dried for 20 min and then resuspended in 200 µl of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0). The extractions yielded from 3 to 10 µg of DNA.

Construction of an enriched genomic library and definition of microsatellite primers

A genomic library enriched with GA and GT repeated microsatellite sequences was produced according to the protocol previously described by Billotte *et al.* (1999) with the following modifications. Total DNA of *Ustilago scitaminea* was restricted with the endonuclease *Rsa*I. The microsatellite sequences were selected using two biotin-labelled oligoprobe in a mix for simultaneous selection of CA and GA repeats: 5'-I*IIIITCTCTCTCTCTC-3' and 5'-I*IIIITGTGTGTGTG-3', where I* is a biotinylated inosine. After ligation, transformation was made in the Epicurian XL1-blue *Escherichia coli* strain. A total of 480 clones were transferred to microplates containing 150 µl of LB/ampicillin solution. Insert sizes were estimated by agarose gel electrophoresis of insert PCR products using M13 universal primers. Gels were alkaline Southern transferred to Hybond N+ nylon membranes (Amersham) and hybridized with $\gamma^{32}\text{P}$ radiolabeled (GA)₁₅ and (GT)₁₅ probes. Inserts of 200 clones that gave a satisfactory positive signal with a size over 500 bp were sequenced by the Centre National de Séquençage (CNS, Génoscope, Evry, France). Primer pairs were designed using the online Primer 3 package (Whitehead Institute for Biomedical Research) for 38 inserts containing more than 5 repeated dinucleotides motifs. A total of 22 microsatellite loci were used for the present study (Table 2).

Microsatellite analysis

The forward primers were end-labeled with [$\text{c-}33\text{P}$]ATP, and amplification was performed in an MJ Research PTC 100 Thermal Cycler in 20 µl reaction mixtures containing 10 ng of template DNA, 200 µM dNTP mix, 1.5 mM MgCl₂, 50 mM KCl, 10 mM TRIS-HCl (pH 8.3), each primer at 0.2 µM, and 1 U of Taq polymerase. The samples were denatured at 94°C for 4 min and subjected to 35 cycles of 94°C for 30 s, 50°C for 1 min, and 72°C for 1 min.

After the addition of 20 µl of loading buffer (98% formamide, 10 mM EDTA, bromophenol blue, xylene cyanol), the amplified products were denatured at 94°C for 3 min, and 5 µl of each sample was loaded onto a 5% polyacrylamide gel with 7.5 M urea and electrophoresed in 1% TBE buffer at 60 W for 1 h 40 min. The gels were dried for 30 min at 80°C and exposed for 4 days to X-ray film (Fuji RX).

Data analysis

Genetic dissimilarities were computed from the microsatellite data using the shared allele distance (Chakraborty and Jin, 1993) in the Darwin software package (Perrier *et al.*, 2003). A pairwise dissimilarities matrix was used to construct a neighbour joining (N-J) tree. A bootstrapping procedure was performed over 1000 replications. Unbiased estimates of expected heterozygosity under Hardy Weinberg assumptions, the observed heterozygosity, the percentage of polymorphic loci when the most frequent allele did not exceed 95% and the mean number of alleles per locus were computed with the Genetix version 4.03 software package (Belkhir K. *et al.*, Centre National de Recherche Scientifique, UMR 5000, Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier, France).

II.3 Results

Seventeen of the 22 primer pairs used to analyse 142 single spore isolates from 77 whips collected in 15 different countries revealed polymorphism. The total number of alleles amplified by these 17 primer pairs was 53, with a mean of 3.1 (Table 2). Although the DNA samples had been extracted from diploid mycelium cultures, the 2114 combinations tested (142 single spore isolates x 17 polymorphic microsatellites) revealed (with one exception) only one SSR allele at a particular locus. This indicates that the fungus genotypes were homozygous at all loci, except for a sample collected in the Philippines that showed a heterozygotic pattern with microsatellite Usc8. For 45 whips, collected in 7 different countries (including 13 whips from the Philippines), two to three single spores per whip were analysed and no polymorphism among whips was observed, except in two cases. In one case, whip 42

Philippines 93
Philippines
Philippines

Philippines

Philippines

Philippines

Philippines

Philippines

Japan

Japan

Japan

Japan

Japan

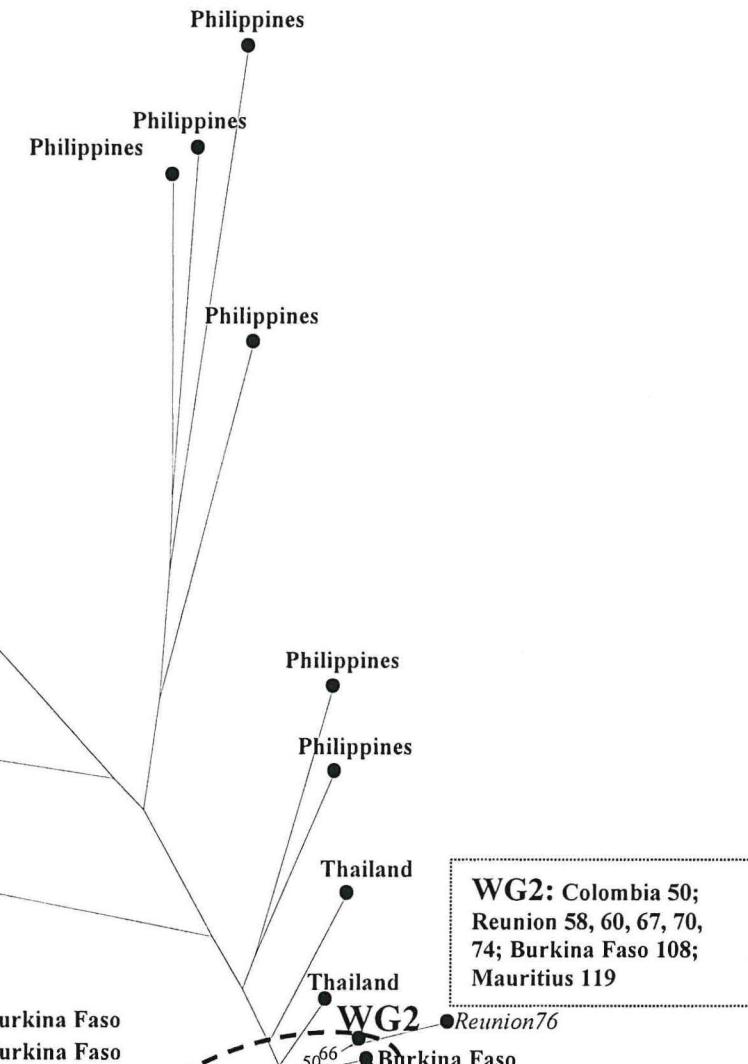
Philippines

India

0 0.2

Figure 2: Neighbour joining tree of the 79 genotypes isolated from 77 distinct whips collected in 15 different countries. WG1 is a genotype consisting of 39 isolates. WG2 is a genotype consisting of 8 isolates. The worldwide lineage (WL) encompasses WG1, WG2, isolates 104,105,106,113,115 from Burkina Faso and isolate 119 from Mauritius. Only bootstrap values over 50% are given in the graph.

Worldwide Lineage (WL)



from the Philippines, the two isolates presented very different genotypes (13 loci out of 17 differed). In the second case (whip 119 from Mauritius), the two single spores differed at only one locus (Usc9) out of 17. Therefore, further analyses were conducted with one isolate per homogenous whip (75 isolates) and the four isolates discussed above, resulting in a total of 79 isolates.

The Neighbour-joining tree of the 79 isolates (Figure 2) showed a very contrasted situation with high variation in the Asian samples, particularly those from the Philippines, and very low diversity in America and Africa. Overall, 22 distinct genotypes were observed. A total of 39 isolates from five African countries (South Africa, Senegal, Burkina Faso and the Mascarene islands of Réunion and Mauritius), five countries of the Americas (Colombia, Brazil, Texas, French West Indies or Guadeloupe), Hawaii (Pacific region), Indonesia, and two isolates (out of 4) from Thailand displayed exactly the same genotype at the 17 studied microsatellite loci. They belong to worldwide genotype 1 (WG1). From this main group, 15 closely related isolates were identified as they differed only at one locus out of the 17 studied. Eight of them, pooled under worldwide genotype 2 (WG2), carried allele 3 of microsatellite Usc9 instead of allele 4 for WG1. The size difference between allele 3 and 4 of Usc9 corresponded to only one repeat motif slippage. Isolates Mauritius119 and Burkina Faso104 carried allele 2 and 5 for microsatellite Usc9, respectively. Isolates 106/113/115/105 from Burkina Faso carried allele 4 for microsatellite Usc1 instead of allele 3 for WG1. Again, the size difference between allele 3 and 4 of Usc1 corresponded to only one repeat motif slippage. Thus, it is likely that all the above isolates belong to the same worldwide inbred lineage, referred to as WL in Figure 1. One isolate from Réunion (76) differed at two loci from WG1 and one locus from WG2. In the Philippines, substantial diversity was observed with 12 distinct genotypes for the 16 isolates analysed. The only isolate from India differed from the other Asian isolates. All isolates from the two locations in Japan had an identical genotype which differed from all the other isolates.

Genetic diversity statistics were computed for countries where the sample size was large enough (Table 3): Colombia, the Philippines, Réunion and Burkina Faso. For Colombia, the Philippines and Réunion, samples were collected from locations tens of kilometers apart and therefore represented diversity on a regional scale. The Burkina samples were collected from various plots within the same commercial plantation. Unbiased gene diversity ($H_{n.b.}$) was

Table 3: Diversity statistics for *U. scitaminea* populations from Burkina Faso, Colombia, Réunion and the Philippines.

	size	Hn.b.	Hobs	P(0.95)	A
Burkina Faso	11	0.0509 (0.1419)	0.00 (0.00)	0.125	1.1875
Colombia	10	0.0118 (0.0474)	0.00 (0.00)	0.0625	1.0625
Philippines	16	0.573 (0.1045)	0.0039 (0.0156)	1.00	2.8125
Réunion	10	0.0434 (0.1319)	0.00 (0.00)	0.125	1.125

Hn.b.= unbiased gene diversity (Nei,1978)

Hobs = observed heterozygosity

P(0.95) = proportion of polymorphic loci when most frequent allele does not exceed 95%

A = mean number of alleles per locus

Standard deviations are indicated in parenthesis

much higher in the Philippines than in Burkina Faso, Réunion and Colombia. In these latter three countries, it was null for most of the studied microsatellite loci, as confirmed by the very low mean number of alleles per locus. Microsatellite Usc9 was the only one to present polymorphism in the four populations. The near absence of heterozygotes (Hobs) in the four populations was also striking. The three populations from Burkina Faso, Colombia and Réunion were genetically identical and presented the same allele set. The Philippine population was completely different, mainly because of its high level of diversity. It encompasses most of the alleles (94%) from the Burkina Faso, Réunion and Colombia populations and it displays variability within some of the localities. For example, in “Brgy Luy a”, four distinct whips were collected, which revealed four distinct homozygous genotypes. These four genotypes encompassed 91% of the alleles revealed in the overall population (41/45 alleles at 17 loci). The two whips from “Hacienda Linay” also revealed two distinct homozygous genotypes encompassing 62% of the alleles revealed in the overall population (28/45 alleles at 17 loci). In contrast, the two whips collected from two different varieties in “Hacienda Florencia” were found to have the same genotype.

II.4 Discussion

Selfing—the predominant reproduction mode

Since *U. scitaminea* isolates were cloned from diploid teliospores it was possible—using the co-dominant microsatellites markers developed in this study—to infer the reproduction mode of this fungus. All isolates but one analysed in this study were homozygous for all 17 polymorphic microsatellite loci (Table 3). Moreover, almost no polymorphism was detected between teliospores from a single whip, even in areas where polymorphism does exist. It is also noteworthy that very different homozygous genotypes could be present at the same location, possibly even in the same plant, as observed in the Philippines. This indicates that selfing is the most predominant reproductive mode of *Ustilago scitaminea*, and populations of this fungus appeared to be mainly composed of inbred lineages. In this respect, *U. scitaminea* can be compared to *Ustilago hordei* in which outcrossing is considered uncommon (Abdennadher and Mills, 2000), and to the *Microbotryum violaceum* model (formerly *Ustilago violaceum*) where fungal populations displayed unexpectedly high frequencies of

homozygotes. The fungi are transmitted by insect pollinators which prefer healthy over diseased plants, thus reducing the probability of spore mixtures being deposited on the same flower (Delmotte *et al.*, 1999). Since outcrossing opportunities are likely rare in this fungus, the evolutionary trend may have favoured selfing as a mechanism to ensure reproductive success (Kaltz and Shykoff, 1999). Two mechanisms have been described in *Microbotryum violaceum* that may foster inbreeding: preferential conjugations between sporidia in selfing rather than in outcrossing combinations (Kaltz and Shykoff, 1999) and intratetrad mating (i.e. conjugation between haploid cells originating from the same teliospore; Hood and Antonovics, 2000). Intratetrad mating may also explain the predominance of selfing in *Ustilago scitaminea* since Leu (1972) and Bock (1964) observed a similar mechanism in *Ustilago scitaminea*. This mechanism might ensure reproductive success because it is quite unlikely that two teliospores from distinct genetic sources could germinate simultaneously and closely enough on the same sugarcane bud to allow outcrossing conjugations. As for *M. violaceum*, selfing might be an appropriate strategy for initiating infection. A population structure analysis should be conducted at the field and plant scale in areas where genetic diversity is maintained in order to further test this hypothesis.

An Asian center of origin

Genetic diversity in *U. scitaminea* was only noted in Asia (Figure 2). Globally, all isolates from America and Africa were found to be closely related genetically and belong to the same worldwide inbred lineage. This lineage was also detected in some Asian countries (Indonesia and Thailand). These observations strongly suggest an Asian origin of *U. scitaminea*. On the basis of our findings, it was possible to start circumscribing the diversity centre of *U. scitaminea* within the Asian area. We detected polymorphism in the Philippines and Thailand and Braithwaite *et al.* (2004) revealed polymorphism among isolates from Taiwan. Westward, our sampling was too weak to draw any conclusions, but the only isolate from Indias differed from all the other samples, including the worldwide lineage. Northward, in Japan, only one genotype, which differed from all the others, was detected at two different localities 66 km apart, suggesting an independant introduction from the centre of diversity. Polymorphism was no longer detected southward in Indonesia, where all isolates collected belonged to the worldwide lineage. Southeastward, in Papua New Guinea, the disease may never have occurred (Antoine, 1961), although it is the center of origin of *Saccharum officinarum*. As

modern sugarcane cultivars are derived from interspecific hybrids between the domesticated species *S. officinarum* and the wild relative *S. spontaneum* (Daniels and Roach, 1987), *S. spontaneum* could be the natural host of *U. scitaminea*. Moreover, there is little evidence indicating that alternative hosts other than *Saccharum* species and their interspecific hybrid derivatives (*S. sinense* and *S. barbieri*) exist naturally (Ferreira and Comstock, 1989). *S. spontaneum* has a large distribution from West Africa to Japan, but it is strongly suggested that the Asian continent is a center of diversity (Panje and Babu, 1960). The diversity centres of *U. scitaminea* and *S. spontaneum* overlap. This indicates that *U. scitaminea* and *S. spontaneum* could have coevolved. Further investigations on the genetic structure of *U. scitaminea* should be conducted in countries such as India, Pakistan, China, Malaysia or Cambodia, where the two species coexist and for which there is still a dramatic lack of information. This should help in determining the centre of diversity of *U. scitaminea* and in localising the origin of the worldwide lineage.

Global dispersal process

The genetic diversity observed in *U. scitaminea* outside Asia and even in some Asian countries was extremely low, with the presence of the same inbred lineage in most countries around the world. This suggests that events leading to the introduction of the fungus in Africa, America and some parts of Asia were linked. The history of sugarcane has involved successive cultivar substitutions on a worldwide scale. In particular, sugarcane hybrids produced by the very first breeding programs in Java (Indonesia) and Coimbatore (India) were very successful and disseminated throughout the world during the 1920s and 1930s. The worldwide spread of the fungus might have followed the release of contaminated cuttings. Long distance aerial dispersal of teliospores might also have been involved as this is a common invasion strategy among plant pathogenic fungi (Brown and Hovmöller, 2002). For example, it has been proposed that teliospores could have been transported like sand and camel dung particles by trade winds entering the Caribbean area (James, 1976; Prospero *et al.*, 1970). This could explain the sudden smut outbreak in the West Indies during the 1970s following an epidemic that spread throughout West Africa in the late 1960s and early 1970s. Nevertheless, the strong founder effects observed in the global population structure of *U. scitaminea* suggest that the fungus seldomly migrated between continents. Since *U. scitaminea* produces teliospores with wind dispersal potential, the degree of gene flow is

expected to be higher at a lower geographic scale and should be investigated in Asian regions where genetic diversity exists.

Implications for studies on pathogenic variation

Race-cultivar specialisation controlled by a gene-for-gene interaction was demonstrated in *Ustilago hordei* (Linning *et al.*, 2004). In *Ustilago scitaminea*, race-cultivar specialisation through controlled artificial inoculations has only been clearly demonstrated in Taiwan (Leu and Teng, 1972, Leu, 1978, Lee *et al.*, 1999). Interestingly, the pathogenicity variations observed in Taiwan coincide with the genetic variation revealed with AFLP markers among Taiwanese strains (Braithwaite *et al.*, 2004). There have been other reports of race-cultivar specialisation in Pakistan, Brazil and Hawaii (Muhammed and Kausar, 1962; Toffano, 1976, Comstock and Heinz, 1977). Nonetheless, an international project involving Argentina, Barbados, Colombia, India, South Africa, Sudan, Taiwan, USA and Zimbabwe revealed no evidence of race-cultivar specialisation except in Taiwan (Grisham, 2001). Similarly, Pérès and Baudin (1983) observed no differences in pathogenicity and no race-cultivar specialisation among isolates from Réunion, Côte d'Ivoire, Mali, Guadeloupe, Kenya and Burkina Faso, which is congruent with the near absence, in our study, of genetic variability among isolates from Africa and America. Thus, we hypothesize that pathogenicity variations may be greater in parts of Asia where a high level of genetic variation in *U. scitaminea* has been found. To further test the above hypothesis, pathogenicity variations should be analysed using appropriate samples (e.g. consisting of isolates collected from multiple cultivars in the same locations and characterized with molecular markers), with pathogenicity tested under controlled conditions on a reference set of sugarcane cultivars.

Implications for resistance breeding and management

Gene flow between continents appeared to be very low, thus greatly reducing the risk of new pathotypes being introduced from the Asian centre of diversity to other sugarcane-producing areas. However, since teliospores can be easily disseminated by wind, the degree of gene flow might be high on a local scale. We demonstrated that the preferential reproductive mode of *U. scitaminea* is selfing with a very low recombination level. According to the risk model of McDonald and Linde (2002), such pathogens may have an intermediate evolutionary

potential. In this setting, a breeding strategy that would optimise the combination of different major resistance genes within the polyploid sugarcane genome might be durable since it could require a long period for accumulation of multiple virulence mutations in inbred lineages. However, frequency and evolutionary impact under selection pressure of rare recombination events should be further investigated. Better characterisation and genetic analysis of resistant germplasm is also needed to identify resistance genes within modern sugarcane cultivar germplasm. Studies on pathogenic variations based on the population structure drawn in this study should allow more accurate characterisation of resistant germplasm. If this approach turns out to be unsuccessful, the exploration could be broadened to wild ancestral germplasm. *S. spontaneum*—based on the assumption that it coevolved with *U. scitaminea*—could be an interesting source of new resistance genes to be assessed. Indeed, numerous resistant clones have been found within *S. spontaneum*, but unfortunately the two main *S. spontaneum* clones used in early interspecific hybridizations from which modern sugarcane cultivars are derived are susceptible to smut (Srinivasan and Alexander, 1971).

Acknowledgements

We gratefully acknowledge Genoscope (Centre National de Séquençage, Evry, France) for sequencing the sugarcane smut SSR clones and all the institutions listed in Table 2 that kindly provided us with smut samples. We also thank Georges Piperidis, Laurent Costet, Fabien Halkett, Didier Tharreau and Jean Pierre Péros for critical reading of the manuscript.

References:

- Abdennadher, M., Mills, D., 2000. Telomere-associated RFLPs and electrophoretic karyotyping reveal lineage relationships among race-specific strains of *Ustilago hordei*. Curr. Genet. 38, 141-147.
- Alexander, K.C., Srinivasan, K.V., 1966. Sexuality in *Ustilago Scitaminea* Syd. Curr. Sci. 23, 603-604.
- Amire, O.A., Trione, E.J., Schmitt, R.A., 1982. Characterization of pathogenic races of the sugarcane smut fungus by neutron activation analysis. Journal of Radioanalytical Chemistry 75, 195-203.
- Antoine, R., 1961. Smut. In: Martin, J.P., Abbott, E.V., Hughes, C.G. (Eds), Sugar Cane Disease of the World. Elsevier Publishing Co., Amsterdam, pp 327-345.
- Barat, H., 1971. Etat Phytosanitaire des cannes à sucre de Côte d'Ivoire. Dangers résultant de la présence du charbon en Haute-Volta. Rapport IRAT.
- Bates, J.F., 1975. Smut disease in Guyana. Sugarcane Pathologists' Newsletter 13/14, 8.
- Bakkeren, G., Kronstad, J.W., 1994. Linkage of mating type loci distinguishes bipolar from tetrapolar mating in Basidiomycetous smut fungi. Proc. Natl. Acad. Sci. USA 91, 7085-7089.
- Billotte, N., Lagoda, P.J.L., Risterucci, A.M., Baurens, F.C., 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54 (4), 277-288.
- Bock, K.R., 1964. Studies on sugar-cane smut (*Ustilago scitaminea*) in Kenya. Trans. Brit. Mycol. Soc. 47 (3), 403-417.
- Braithwaite, K.S., Bakkeren, G., Croft, B.J., Brumbley, S.M., 2004. Genetic variation in a worldwide collection of the sugarcane smut fungus Ustilago scitaminea. Proc. Aust. Soc. Sugar cane Technol. 26, abstract p 48.
- Brown, J.K.M., Hovmöller, M.S., 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science 297, 537-541.
- Chakraborty, R., Jin, L., 1993. A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and populations distances. In: Pena,

- S.D.J., Chakraborty, R., Epplen, J.T., Jeffreys, A.J. (Eds), DNA fingerprinting: State of the science. Birkhäuser Verlag, Basel, pp 153-175.
- Chao, C.P., Hoy, J.W., Saxton, A.M., Martin, F.A., 1990. Heritability of resistance and repeatability of clone reactions to sugarcane smut in Louisiana. *Phytopathology* 80, 622-626.
- Commonwealth Mycological Institute, 1977. Distribution Maps of Plant Diseases. N°79, 4th ed.
- Comstock, J.C., 2000. Smut. In: Rott, P., Bailey, R.A., Comstock, J.C., Croft, B.J., Salem Saumtally, A. (Eds), A guide to sugarcane diseases. CIRAD and ISSCT, pp. 181-185.
- Comstock, J. C., Heinz, D.J., 1977. A new race of culmicolous smut of sugarcane in Hawaii. *Sugarcane Pathologists' Newsletter* 19, 24-25.
- Daniels , J., Roach, B.T., 1987. Taxonomy and evolution. In: Heinz, D.J. (Ed.), Sugarcane Improvement through Breeding, Elsevier, Amsterdam, pp. 7-84.
- Delassus, M., 1975. Rapport sur le charbon de la canne à sucre. Situation au Niger et méthodes de lutte à envisager. Rapport de l'IRAT.
- Delmotte, F., E., Bucheli, Shykoff, J.A., 1999. Host and parasite population structure in a natural plant-pathogen system. *Heredity* 82, 300-308.
- Ferreira, S.A., Comstock, J.C., 1989. Smut. In: Ricaud, C., Egan, B.T., Gillaspie, A.G., Hughes, C.G. (Ed.), Diseases of sugarcane. Elsevier, Amsterdam, pp. 211-229.
- Flores, S.C., 1978. Smut in Florida and Belize - A serioius threat to the mexican sugar industry. *Sugarcane Pathologists' Newsletter* 21, 1-2.
- Flores, S.C., 1980. Sugarcane smut in Mexico. *Sugarcane Pathologists' Newsletter* 24: 8-10.
- Genetix version 4.0. Logiciel sous Windows TM pour la génétique des populations.
Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II,
Montpellier, France.
- Grisham, M.P., 2001. An international project on sugarcane variability within sugarcane smut. *Proc. Int. Soc. Sugar Cane Technol.* 24 (2), 459-461.
- Hirshhorn, E., 1949. Un nuevo metodo de infección artificial con el carbon de la caña de azucar. *Revista de Investigaciones Agricolas*. III (4), 335-344.
- Hood, M.E., Antonovics, J., 2000. Intratetrad mating, heterozygosity, and the maintenace of deleterious alleles in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity* 85, 231-241.

- James, G.L., 1976. Sugarcane smut infection in Guyana and Martinique. The Sugar Journal May, 17.
- James, N.I., 1978. Sugarcane smut - An eminent threat to the U.S. mainland. Proc. Amer. Soc. Sugar Cane Technol. 7, 63-66.
- Kaltz, O., Shykoff, J.A., 1999. Selfing versus outcrossing propensity of the fungal pathogen *microbotryum violaceum* across *Silene latifolia* host plants. J. Evol. Biol. 12, 340-349.
- Lambat, A.K., Chenulu, V.V., Chona, B.L., 1968. Morphological variation in the sugarcane smut *Ustilago scitaminea* Syd. Mycopathologia and Mycologia Applicata 36, 301-304.
- Lee, C.S., Yuan, C.H., Liang, Y.G., 1999. Occurrence of a new pathogenic race of culmicolous smut of sugarcane in Taiwan. Proc. Int. Soc. Sugar Cane Technol. 23(2) : 406-407
- Lee-Lovick, G., 1978. Smut of sugarcane - *Ustilago scitaminea*. Review of Plant Pathology of the Commonwealth Mycological Institute 57 (5), 181-187.
- Leu, L.S., 1972. Culmicolous smut of sugar cane in Taiwan III. Germination and storage of teliospores, and compatibility of *Ustilago scitaminea* Sydow. Report of Taïwan sugarcane experimental station 56, 37-45
- Leu, L.S., 1978. Culmicolous smut of sugar cane in Taiwan VI. New pathogenic strain obtained by artificial hybridization and further studies on compatibility of *Ustilago scitaminea* Sydow. Ann. Phytopath. Soc. Japan 44, 321-324.
- Leu, L.S., Teng, W.S., 1972. Pathogenic Strains of *Ustilago scitaminea* Sydow. Sugarcane Pathologists' Newsletter 8, 12-13.
- Leung, H., Nelson, R.J., Leach, J.E., 1993. Population structure of plant pathogenic fungi and bacteria. Adv. Plant Pathol. 10, 157-205.
- Linning, R., Lin, D., Lee, N., Abdennadher, M., Gaudet, D., Thomas, P., Mills, D., Kronstad, J.W., Bakkeren, G., 2004. Marker-based cloning of the region containing the UhAvr1 avirulence gene from the basidiomycete barley pathogen *Ustilago hordei*. Genetics 166, 99-111.
- Liu, L.-J., 1981. Sugarcane smut in Puerto Rico. Sugarcane Pathologists' Newsletter 26, 52.
- McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential , and durable resistance. Annu. Rev. Phytopathol. 40, 349-379.
- Martinez-Espinoza, A.D., Garcia-Pedrajas, M.D., Gold, S.E., 2002. The Ustilaginales as plant pests and model system. Fungal Genet. Biol. 35, 1-20.

- Muhammad; S.; Kausar; A.G.; 1962. Preliminary studies on the genetics of sugarcane smut, *Ustilago scitaminea* Sydow. Biologia 8, 65-74.
- Panje; R.R.; Babu; C.N.; 1960. Studies in *Saccharum spontaneum*. Distribution and Geographical association of chromosome numbers. Cytologia 25, 152-172.
- Péros, J.P., Baudin, P., 1983. Etude de la variabilité d'*Ustilago scitaminea* Syd., agent du charbon de la canne à sucre. L'Agronomie Tropicale 38 (3), 234-239.
- Perrier, X., Flori, A. , Bonnot, F., 2003. Data analysis methods. In : Hamon, P., Seguin, M., Perrier, X. , Glaszmann, J. C. (Ed.), Genetic diversity of cultivated tropical plants. Enfield, Science Publishers. Montpellier, pp. 43-76.
- Prospero, J.M., Bonatti, E., Schubert, C., Carlson, T.N., 1970. Dust in the Caribbean atmosphere traced to an african dust storm. Earth and Planetary Science letters 9, 287-293.
- Riley, I.T., Jubb T.F., Egan B.T., Croft, B.J., 1999. First outbreak of sugarcane smut in Australia. Proc. Int. Soc. Sugar Cane Technol. 23 (2), 333-336.
- Rogers, S, Bendich, A, 1988. Extraction of DNA from plant tissues. In: Gelvin, S.B., Schilperoort, R.A. (Eds), Plant Molecular Biology Manual A6. Kluwer Academic Publishers, Dordrecht, pp 1-10.
- Srinivasan, K.V., Alexander, K.C., 1971. Sources of resistance to red rot and smut in the species of *Saccharum*. Sugarcane pathologist's newsletter 6, 6-7.
- Toffano, W.B., 1976. Estudos sobre polaridade sexual e possíveis raças fisiológicas de *Ustilago scitaminea* Syd., no estado de São Paulo. Arq. Inst. Biol., São Paulo 43, 65-79.
- Trione, E.J., 1990. Growth and sporulation of *Ustilago scitaminea*, *in vivo* and *in vitro*. Mycol. Res. 94 (4), 489-493.
- Veiga, F.M., 1972. Smut in Brazil. Sugarcane Pathologists' Newsletter 9 (25), 17.
- Wu, K.K., Heinz, D.J., Hogarth, D.M., 1988. Association and heritability of sugarcane smut resistance to races A and B in Hawaii. Theor. Appl. Genet. 75, 754-760.
- Yamauch, S., 1978. Epidemic of smut (*Ustilago scitaminea* Syd.) In Okinawa and its control. Sugarcane Pathologists' Newsletter 21, 4-5.

III CARTOGRAPHIE GENETIQUE DANS UN CROISEMENT BI-PARENTAL

III Cartographie génétique dans le croisement R 570 x MQ 76/53

Le cultivar réunionnais ‘R570’ a fait l’objet d’un investissement continu en cartographie génétique depuis plus de 10 ans. Une première carte a été construite à l’aide de marqueurs RFLP (Grivet et al. 1996) à partir d’une descendance, de taille limitée (80 individus), issue de l’autofécondation de ‘R570’. Une deuxième carte a été construite à l’aide de marqueurs AFLP à partir d’une nouvelle descendance autofécondée de taille plus importante (300 individus) (Hoarau et al., 2001). Cette dernière carte a ensuite été complétée à l’aide de marqueurs microsatellites et de sondes RFLP « analogues de gènes de résistance » (Rossi et al., 2003). Pour mettre en évidence les facteurs génétiques mendéliens impliqués dans la résistance au charbon du cultivar ‘R 570’ nous avons choisi d’utiliser un nouveau croisement, bi-parental cette fois, entre ‘R 570’ et ‘MQ 76/53’ un clone extrêmement sensible utilisé comme testeur. Il faut indiquer ici que la descendance issue de l’autofécondation de ‘R570’ ségrège très mal pour la résistance au charbon dans les conditions de la Réunion (Hoarau comm. pers.). De plus, la comparaison dans un essai préliminaire de deux croisements bi-parentaux, ‘R570’ x ‘MQ 76/53’ et ‘Nco 376’ (sensible au charbon) x ‘R 570’, a montré que le premier présentait une meilleure ségrégation que le second. Un nouvel investissement en cartographie génétique a donc été réalisé sur 200 descendants du croisement entre ‘R 570’ et ‘MQ 76/53’, en utilisant principalement le potentiel des marqueurs AFLP pour couvrir rapidement le génome haut polyplioïde des deux parents du croisement.

Ce travail est présenté sous la forme d’un article, soumis à la revue *Theoretical and Applied Genetics*, intitulé « **Genetic mapping in the high polyploid sugarcane using a bi-parental progeny; identification of a gene controlling stalk colour and a new rust resistance gene** »

La carte complète du clone MQ 76-53 est montrée en Annexe 1

Genetic mapping in the high polyploid sugarcane using a bi-parental progeny; identification of a gene controlling stalk colour and a new rust resistance gene

L.-M. Raboin (1), K.M. Oliveira (2), L. Lecunff (3), H. Telismart (1), D. Roques (1), M. Butterfield (4), J.-Y. Hoarau (5), A. D'Hont (3)

(1)CIRAD (Centre de coopération internationale en recherche agronomique pour le développement) - UMR PVBMT, Ile de la Réunion, France

(2)Universidade Estadual de Campinas, 13083-970 Campinas, SP, Brazil

(3) CIRAD, UMR 1096, TA 40/03, Avenue Agropolis 34 398 Montpellier cedex 5, France

(4) SASRI (South African Sugarcane Research Institute), Private bag X02, Mount Edgecombe, 4300 South Africa

(5) CERF (Centre d'essai, de recherche et de formation), BP 315 97494 Ste-Clotilde, Ile de la Réunion, France

Submitted to TAG with 4 tables and 3 figures

Abstract

Two genetic maps were constructed in the high polyploid sugarcane ($2n=100-130$) using a population of 198 progeny derived from a cross between R570, a commercial sugarcane cultivar, and MQ76-53 an old Australian sugarcane clone (Trojan \times *Saccharum spontaneum* SES528). A total of 1666 polymorphic markers were produced using 37 AFLP primer pairs combinations, 46 SSRs and 9 RFLP probes. Linkage analysis allowed the construction of 86 cosegregation groups for R570 and 105 cosegregation groups for MQ76-53 encompassing 424 and 536 simple dose markers respectively. The cumulative length of the R570 map was 3144 cM. The cumulative length of the MQ76-53 map was 4329 cM. Double dose markers common to R 570 and MQ76-53 were used apart to identify double dose cosegregation groups that may correspond to chromosome segments conserved between the two parents. A synthesis of the mapping information gathered on R570 from this study and a previous map based on a selfed population of R570 is presented. Two new genes controlling mendelian traits were localised on the MQ76-53 map: a gene controlling the red colour of the stalks found linked at 6.5 cM to an AFLP marker and a new brown rust resistance gene found linked at 23 cM to an AFLP marker. Beside another brown rust resistance gene (*Bru1*) previously identified, these two genes are the only major genes identified in sugarcane so far.

Key words: sugarcane, genetic mapping, AFLP, rust resistance gene, colour gene

III.2 Introduction

Modern sugarcane cultivars possess one of the most complex genomes of commercially cultivated plants (Grivet and Arruda 2001). They are polyploid and aneuploid hybrid derivatives from two highly polyploid species, *Saccharum officinarum* ($x=10$, $2n=8x=80$) the domesticated sugar-producing species, and *Saccharum spontaneum* ($x=8$, $2n=40-128$) a wild species. The first interspecific hybrids, produced in the early 20th century, were backcrossed to *S. officinarum* leading to modern sugarcane cultivars. In both interspecific F1 and BC1 crosses, *S. officinarum* transmitted its somatic chromosome number (2n) to the progeny (Bhat and Gill 1985). As a result, modern cultivars have between 100 and 130 chromosomes. Despite this complicated picture, the meiosis of modern sugarcane cultivars involves mainly bivalent pairing (Price 1963; Burner and Legendre 1993 and 1994) and thus make it possible to address genetic mapping on the basis of simplex markers (Wu et al. 1992).

Genetic maps have been produced for both ancestral species, *S. spontaneum* (Al Janabi et al. 1993; Da Silva et al. 1993 and 1995; Ming et al. 1998 and 2002b) and *S. officinarum* (Mudge et al. 1996; Guimaraes et al. 1997; Ming et al. 1998 and 2002b). Genetic maps have also been constructed for modern sugarcane cultivars, representative of *S. officinarum* introgressed germplasm exploited in modern breeding programs. A map has recently been published on Australian cultivar Q165 (Aitken et al. 2005) and continuing efforts have also been achieved to build a map of cultivar R570 from Reunion. A first RFLP map has been constructed using a selfed population derived from R570 (Grivet et al. 1996). Later, using a larger number of selfed progeny of R570, AFLPs (Hoarau et al. 2001), SSRs and Resistance Gene Analogs (Rossi et al. 2003) were mapped. This AFLP-based map encompassed more than 1100 markers and reached a total length of 7800 cM, which represents coverage of around 46% of the anticipated genome size (17000 cM). The species origin of markers revealed that about 10% of the chromosomes of R570 were inherited from *S. spontaneum* and that another 10% were derived from recombination between *S. spontaneum* and *S. officinarum*. These ancestral contributions were congruent with previous cytogenetic studies based on in situ hybridization (D'Hont et al. 1996). At the moment marker coverage in the map of the interspecific cv. R570 is uneven, with *S. spontaneum* chromosomes being covered more densely than those of *S. officinarum*. This discrepancy results from a lower polymorphism rate in the *S. officinarum*

component of modern sugarcane cultivar genome as compared to *S. spontaneum* (D'Hont et al. 1994, Grivet et al. 1996).

Developing a saturated map is a prerequisite to be able to efficiently localize major genes or Mendelian factors involved in Quantitative Trait Loci (QTL). QTL mapping is expected to be a real challenge in sugarcane where many alleles (potentially up to 12) coexist at each locus due to the high ploidy. Indeed, at a particular locus, the effect of an allele should be perceptible only if it exceeds the average effect of all the other segregating alleles in the background (up to 11), and not only if its effect exceeds the effect of a single alternative allele as in a diploid (D'Hont and Glaszmann 2001; Hoarau et al. 2002). Ming et al. (2001 and 2002a) investigated the genetic basis of traits related to sugar content, plant height and flowering in interspecific crosses between *S. officinarum* and *S. spontaneum*. Numerous QTLs were detected that could be localised in a few genomic regions suggesting that the actual number of genes involved in the genetic control of these traits may be much less. Within an inter-specific cross, wide segregation of phenotype can provide a favourable context for QTL detection. By comparison, a study of yield components (plant height, stalk diameter, stalk number and Brix) in the selfed progeny of the modern cultivar R570 revealed numerous QTLs with smaller individual effects (Hoarau et al 2002). Similarly, Jordan et al. (2004) detected numerous small QTLs for stalk number in sugarcane and found that the same RFLP probes, when used in Sorghum, were also associated with tillering. So far, the only major gene localized in the sugarcane genome is a rust resistance gene (Daugrois et al. 1996). This gene (called *Bru1* for brown rust) is currently the focus of a map-based cloning project (D'Hont et al. 2001; Asnaghi et al. 2004).

In this paper we describe the identification and mapping of two other major genes, a new putative rust resistance gene and a gene controlling stalk colour. These genes were identified in the sugarcane clone MQ76-53 thanks to an extensive genetic mapping study of a bi-parental cross derived from the cultivar R570 and MQ76-53. Common markers between the two parents allowed us to identify putative conserved chromosome segments. We discuss the significance of these common segments and the constraints and advantages of working in a bi-parental cross in comparison with a selfed progeny population.

III.3 Materials and Methods

Plant and DNA material:

The mapping population consisted of 198 progeny derived from a cross between R570 and MQ76-53. R570 is a modern cultivar developed at CERF [Centre d'Essai de Recherche et de Formation, Réunion] and is derived from a cross between two modern cultivars H32-8560 and R445. MQ76-53 is an old Australian sugarcane clone that comes from a cross between Trojan (Co 270 x *S. officinarum*), an old cultivar, and SES528, a *S. spontaneum* clone. Its genetic structure should therefore be close to that of an F1 interspecific hybrid. R570 is resistant to rust and has green stalks. MQ76-53 is also resistant to rust but has red stalks. Total genomic DNA of mapping progenies was extracted from fresh leaves according to the method described by Hoisington (1992).

In addition, 133 progeny from a cross between the rust susceptible female tester B63-758 and the rust resistant male MQ76-53 has been evaluated for rust resistance.

Field evaluation of rust resistance and stalk colour

A subset of 166 out of the 198 mapping progenies was evaluated for rust resistance in a randomized complete block design with three replicates in Ligne Paradis CIRAD experimental station (Saint Pierre, Reunion). Each plot was composed of 4 distinct stools. Rust resistance was scored on each plot on a 1 (the most resistant:) to 9 (the most susceptible) scale according to Tai et al. (1981). A score of 1 indicated the absence of sporulating pustules (Uredosores). Susceptible plants may have scores ranging from 2 (a few sporulating pustules) to 9 (many pustules even on younger leaves and necrosis of older ones). Segregation analysis was performed on the mean score computed over replications for each progeny. Progeny from the cross B63-758xMQ76-53 were studied in the field without any particular statistical design, with seedlings planted 50 cm apart. In this cross, rust resistance was scored for each individual seedling on the presence/absence of sporulations. Seedlings bearing sporulating pustules were classified as susceptible, otherwise they were classified as resistant.

Stalk colour of the 198 mapping progenies was scored in three different trials on the basis of the discrete classification red versus non-red colour (mostly green or yellow stalks).

Table 1: Description of the candidate genes used as RFLP probes provided by SASRI

Probe name	Putative gene function
ADS1	Thaumatin
ADS2	Flavonoid pathway transcription factor
ADS3	Pto ser/thr protein kinase
ADS4	Pathogen induced nucleotide binding site
ADS5	Phosphoprotein phosphatase
ADS6	Isoflavone reductase
ADS7	G protein receptor
ADS8	Cellulose synthase
ADS9	Cell wall associated kinase

AFLP markers

AFLP analysis (Vos et al. 1995) was performed using Gibco BRL kit genome I following manufacturer recommendations except for slight modifications as in Hoarau et al. (2001). Each AFLP marker was identified by the primer combination consisting of 6 letters plus a band number indicated as a suffix. The first three letters represent *EcoRI* selective nucleotides and the last three *MseI* selective nucleotides. We used the same 37 combinations as used to build R570 map of Hoarau et al. (2001) and Rossi et al. (2003) except the three *EcoRI/MseI* primer combinations aag/caa, aag/cat and aag/cta. Instead, the combinations aag/ctg, aac/ctg and acc/cta have been used. Therefore, most of the bands inherited from R570 could be labelled as in the previously published maps.

Microsatellites markers:

Thirty-three SSRs developed at CIRAD in collaboration with Génoscope (Evry, France) in addition to 13 SSRs mined from the Brazilian sugarcane EST database (Pinto et al. 2004 and in prep) were analysed following the protocol described in Rossi et al. (2003). CIRAD SSR bands inherited from R570 were coded according to this previously published map using the following nomenclature: m (microsatellite) followed by the number of the SSR and then the letter m (as for marker) followed by a number identifying the band. CIRAD SSRs (complete nomenclature = mSSCIRxx) are described at <http://tropgenedb.cirad.fr/en/sugarcane.html>. The nomenclature of EST SSR was m (microsatellite) followed by a letter and a two digit code number and then the letter m (as for marker) followed by a number identifying the band. SSR bands were scored as dominant markers (presence vs absence).

RFLP markers:

Nine candidate genes differentially expressed in response to challenge by smut (Heinze et al. 2001) were kindly provided by the South African Sugarcane Research Institute (SASRI) and used as RFLP probes (Table 1). The nomenclature used for those RFLP markers was ‘ADS’ followed by a figure between 1 and 9, and then a code for the restriction enzyme used (Eco for EcoRV and Hin for HindIII) followed by a number identifying the band. DNA

preparation, Southern blotting and hybridizations were performed as described previously by Grivet et al. (1996). RFLP bands were scored as dominant markers (presence vs. Absence).

Map construction:

The genetic maps were constructed using linkages in coupling phase between simple dose markers only (Wu et al. 1992). In order to distinguish simplex markers (1:1) from double dose (DD) markers of bi-parental origin (3:1), we considered a segregation ratio threshold of 1.73:1 since this ratio gives equal χ^2 value for both 1:1 and 3:1 hypotheses (Mather 1957). Two-point analyses between single dose markers were performed at a LODscore threshold of 5 and a recombination fraction threshold of 0.35 using Mapmaker 3.0 (Lander et al. 1987). Markers within each cosegregating group (CG) were then ordered by multipoint analysis using the Haldane mapping function. CGs were gathered into the same Homology Groups (HG) either if (1) they had at least two RFLP probes or SSR in common or if (2) they were linked in repulsion. Moreover, some R570 CGs could be assigned to HG on the basis of anchor markers (AFLP markers or SSR markers) they had in common with CGs from a previous R570 RFLP-based map (Grivet et al. 1996 and unpublished results). Chromosome pairing behaviour was investigated following the procedure described in Hoarau et al (2001).

Use of double dose (DD) markers of bi-parental origin

Double dose (DD) markers of bi-parental origin (3:1) were distinguished from triple dose markers or higher multiple dose markers (ratio >7:1) by considering the segregation ratio threshold of 4.58:1 which gives equal χ^2 statistical value for both 3:1 and 7:1 hypotheses (Mather, 1957). DD markers of bi-parental origin were tentatively used to detect chromosomal segments that may be common between the two cultivars. For this objective, CGs of DD markers of bi-parental origin were constructed by performing a two-point analyses (Mapmaker LOD = 5 and $\theta= 0.2$). The risk of homoplasy (i.e. confusion of two distinct single dose markers with a double dose marker because of comigration of amplification products) is assumed to be low when considering groups of cosegregating DD markers.

In a second step, those putative common chromosomal segments revealed by DD marker CGs were tentatively assigned to the respective simplex markers CGs of R570 and MQ76-53 to which they belong. The assignment to a R570 simplex markers CG was straightforward since most of DD markers of bi-parental origin had been previously mapped in R 570 (Rossi et al., 2003). The assignment to a simplex CG of MQ 76-53 map was determined on the basis of a systematic comparison of the considered DD marker with all possible virtual DD markers resulting from the combination of each simplex marker of MQ76-53 and each simplex marker of R570. A genetic distance smaller than 10 cM between the actual DD segment and one virtual DD marker was retained to assign with a good level of confidence the DD segment to the corresponding simplex CG of MQ 76-53.

III.4 Results

Parental linkage maps

A total of 1666 polymorphic markers were produced in the progeny using 37 AFLP primer pair combinations, 46 SSRs and 9 RFLP probes. Among these markers, 1057 were simplex markers and 168 were DD markers of bi-parental origin. The remaining markers (duplex or monoparental origin, triplex or higher multiplex markers) were not exploited in this work. Out of the 1057 simplex markers, 584 (55%) were specific to MQ76-53 and 473 (45%) were specific to R570. Linkage analysis of all these simplex markers resulted in 191 CGs encompassing a total of 960 markers while 97 simplex markers remained unlinked: 424 markers specific of R570 formed 86 CGs and 536 markers specific of MQ76-53 formed 105 CGs. The length of the CGs ranged from 1.6 cM to 179.4 cM in MQ 76-53 and from 1.1 cM to 158.3 cM in R570. The cumulative length of MQ76-53 map was 4329 cM. The cumulative length of R570 map was 3144 cM. The detailed maps are presented at the web site <http://tropgenedb.cirad.fr/en/sugarcane.html> (Ruiz et al. 2004). Only 17 out of the 105 MQ76-53 CGs could be assigned to the Homology Group framework defined for the R570 map on the basis of at least two RFLP probes or SSR loci in common. Out of the 86 CGs of the R 570 map, 60 CG (70%) could be assigned to a HG. No preferential pairing between MQ76-53 CGs was detected at LOD=3. In contrast, preferential pairing was observed between 18 pairs of CGs in R570, 14 of which had been previously detected by Hoarau et al. (2001).

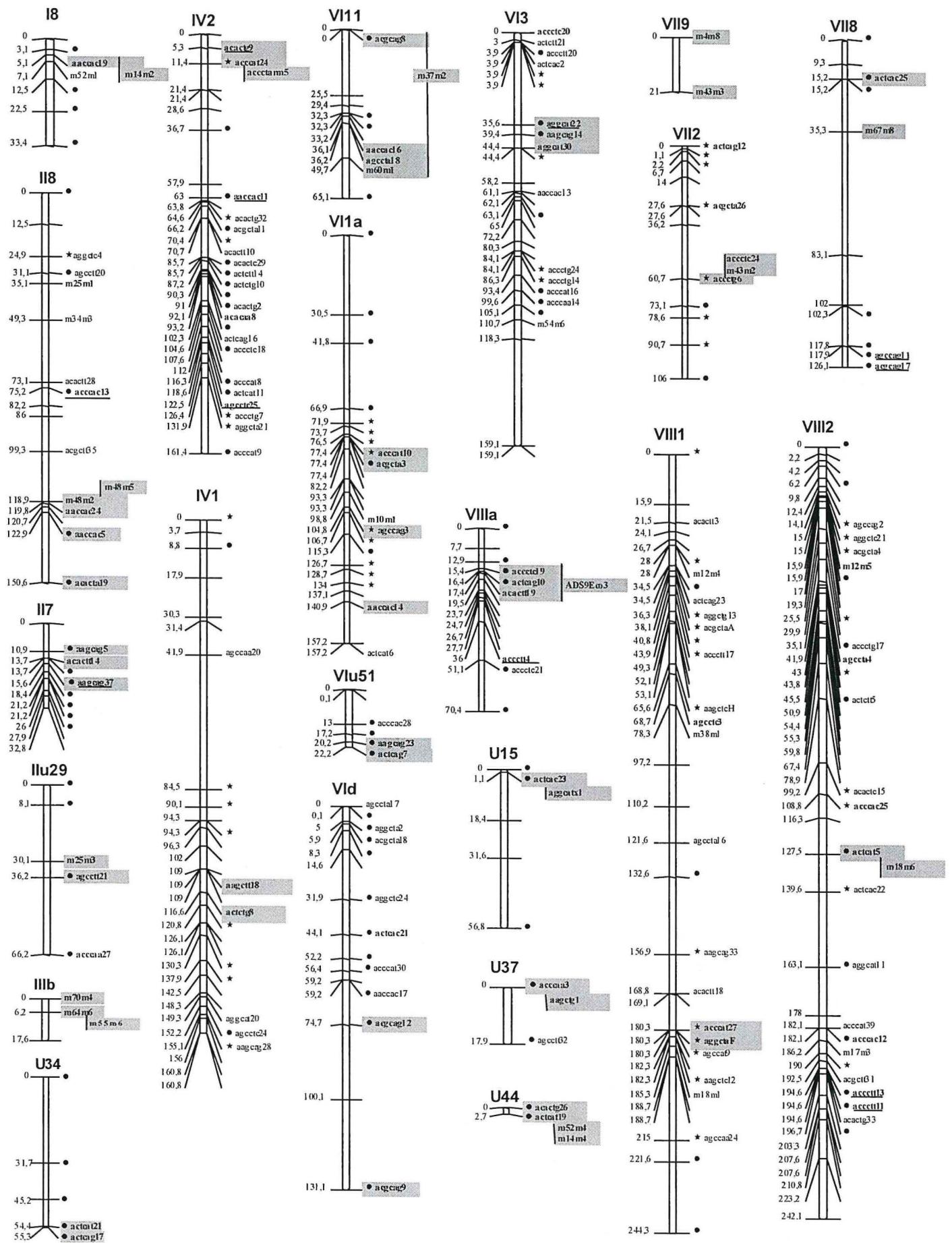


Figure 1: Distribution on the simple-dose-marker-based map of R570 (Rossi et al. 2003) of the double doses (DD) markers involved in a DD cosegregation group in the present study and visualisation of putative common haplotype between R570 and MQ76-53. Markers in normal and bold types were scored in the present study whereas markers without a name were not available (unscorable or not produced). Markers in bold types were double dose markers of bi-parental origin. Underlined markers were multiplex markers (≥ 3) of bi-parental origin. Groups of markers in bold types highlighted in grey represent putative common haplotypes between the two genitors. New markers belonging to those groups that had not been previously mapped by Rossi et al. (2003) are represented in a right-shifted position. Multiplex markers (≥ 3) linked to any double dose haplotypes were also highlighted. The species origin of markers, as determined by Hoarau et al. (2001), is indicated with: * for a putative *S. spontaneum* origin or • for a putative *S. officinarum* origin.

Alignment of different R570 AFLP-based maps

The present map of R570 was aligned with the former AFLP-based map of the same cultivar (Hoarau et al. 2001 and Rossi et al. 2003). This was possible because most of the single dose AFLP markers specific to R570 in the present map (based on a bi-parental progeny derived from the cross R570xMQ76-53) were common with the former map (based on a selfed progeny of R570). The localisation of 32 new locus-specific markers (19 SSRs and 13 RFLPs) and the detection of 4 new pairs of CGs in repulsion improved the assignment of the CGs of the previously published map to the Homology Group (HG) framework. This percentage reached 72% (92 over 128 CGs) compared to 52% before Rossi et al. (2003). When we integrated the information of all three studies, we obtained a tentative consensus map organized in 7 HGs which encompassed a total of 90 assigned CGs. The number of CGs per HG ranged from 2 (HG IV) to 22 (HG VI). The present mapping work also helped refine CGs of previous published map (Table 2). In particular, the basal part of CG VI3 in Rossi et al. (2003) may have been erroneously attached to the upper part of this CG, assigned to HG VI. Indeed, this basal part, which corresponds to CG 59 in this study carries two locus specific markers (m40 and R149Eco) that suggest it belongs to HG III. In three cases, the new data permitted the coalescence of two former distinct CGs into a single CG (CG30 = Ib + U62 (Rossi et al. 2003) ; CG7 = VI3+VI8 (Rossi et al. 2003) ; CG111 = I6 + U3 (Rossi et al. 2003)). In addition to the CGs presented in Table 2, the R570 genetic map is also composed of 30 small CGs (average size= 21.8 cM) that do not bear any locus specific markers. Therefore, those CGs can not yet be assigned to Homology Groups.

Tentative identification of common chromosome segments between parental genomes

A total of 168 double dose (DD) markers fitting a 3:1 ratio had a bi-parental origin (i.e. MQ76-53 and R570 contributed one dose each). A linkage analysis (LOD=5, $\theta=0.2$) with these DD markers allowed the identification of 22 CGs containing between 2 and 5 markers and encompassing a total of 57 bi-parental DD markers. The distribution of those markers on the R570 selfed map of Rossi et al. (2003) is shown in Figure 1. Most of the identified CGs of DD markers are clusters of tightly linked markers but in a few instances they extend over tens of centiMorgans. These cosegregating DD markers indicate that fragments of homologous CGs (i. e. chromosomes) may have the same haplotype in R570 and MQ76/53 - or in other words, chromosome segments may be common between R570 and MQ76-53. Nearly all of

Table 2 : Synthesis of mapping information on R570: alignment of the present map with maps of Hoarau et al. (2001) and Rossi et al. (2003). Locus-specific markers (SSRs and RFLPs) highlighted in bold characters represent newly localised markers. CGs pairing preferentially at meiosis are associated within boxes (bold boxes indicate newly detected repulsion linkages). When a CG is involved in repulsion linkage with two CGs a vertical line is drawn between each pair of CGs in repulsion). When part of a former CG was presumably wrongly linked in coupling phase to its remaining part, it appears crossed out in bold characters. When two former CGs could be linked together to form a new CG, they appear in bold characters with a + sign. Detailed information on the different genetic maps of R570 is available at <http://tropgenedb.cirad.fr/en/sugarcane.html>.

HG ⁴	Cons ⁵	Rossi ¹	JYH ²	LMR ³	SSR loci	RFLP loci
I					m2, m13, m14, m19, m27 m42*, m52, m53	R024Sst ³ , R057hin, R083Sst, R125Sst, R183Sst, R313eco, R371Dra
I	I a	I a	51	54	m52m8	R057Hinc, R024Sstf
I	I b	I b+U 62	52+38	30	m14m3, 27m1	R057Hinf
I	I c	I c	105	29	m27m2	
I	I 4	I 4	1	43	m19m1, m42m7, m52m3, m53m2	R083Sstc
I	I 5	I 5	3	154	m2m2, m14m6	R083Sste, R125Ssta
I	I 6	I 6+U 3	37+72	111	m13m2, m27m3, m52m6	R024Sstc, R024Sstg, R183Sstg
I	I 7	I 7	74	40	m19m5, m53m9	R024Ssta, R183Ssta, R313Ecob, R313Ecoe, R371Draa
I	I 8	I 8	68	-	m52m1	R024Sste, R083Sstd
I	I 9	I 9	2	108	m13m1, m14m1	R125Sstc
I	I 10	I 10	10	42	m19m2, m53m1	R057Hine, R057Hind, R125Sstb, R313Ecoa, R313Ecod
I	I 17	U 8	82	114+143	m52m2	
II					mA03, mA16*, mB01, m25, m34, m35, m39, m41, m48, m50, m56, m69*	R012Eco, R145Eco, R169Hin, R258Dra, R275Eco, R366Hin, R386Dra
II	II a	II a	90	128		R012Ecof, R145Ecoe, R275Ecoe
II	II b	II b	71	166		R012Ecoc, R275Ecoc
II	II 4	U 35	84	-	m39m1	
II	II u21	U 21	85	107		
II	II 6	II 6	92+5	21	mA03m5, m25m2, m34m1	R012Ecoa, R386Drac
II	II u29	U 29	88	-	m25m3	
II	II 7	II 7	62	-	m39m2	R275Ecof
II	II 8	II 8	59	23	m25m1, m34m3, m48m2, m50m1, m69m2	R169Hina, R386Draa
II	II 9	II 9	4	69	mA03m6, mB01m1, m34m2, m35m1, m41m1	R169Hinc, R258Draa, R275Ecod, R366Hina
II	II 10	II 10	7	148	mA16m1, m35m2, m41m2	R169Hinf, R258Drab, R366Hinc
II	II 11	II 11	-	-	m48m3, m56m4	
II	II 12	II 12	-	-	m35m3	R366Hind
II	II 14	U 61	6	115		
III					mC03, m33, m40, m42*, m55, m64, m70	R149Eco, R523Eco
III	III a	III a	-	189	m40m6, m64m2	R523Ecoc
III	III b	III b	-	-	m64m6, m70m4	R523Ecoa
III	III 2a	U 57	24	-		
III	III 3	III 3	23	53	m42m4, m55m2, m70m2	
III	III 2	III 2	24	126	m55m8, m70m6	R149Ecoa
III	III 4	III 4	22	68+191	m40m1	
III	III 7	III 7	83	82	m40m4, m55m9, m64m8, m70m7	
III	III 8	III 8	107	27	m33m5, m55m5, m64m5, m70m3	
III	III u	VI 3	29	59	mC03m3, m40m2	R149Ecoe
III	III u54	U54	34	-	m33m2, m42m1	
III	III 10	III 10	114	57	m64m1, m70m1	
IV					mA06, m74	
IV	IV 1	IV 1	8	81+109	mA06m4	R142Eco*, R184Eco, R372Eco, R441Sst, R526Eco*, R533Hin
IV	IV 2	IV 2	9	106	m74m2, mA06m2	R142Ecod, R184Ecog, R184Ecof, R372Ecoe, R372Ecoa, R441Sstd, R441Ssta, R526Ecob, R533Hinb
VI					m7, m9, m10, m23, m31, m37, m47, m54, m57, m58, m60, m68, m73*	R441Sstf, R441Sste, R533Hina
VI	VI 1a	VI 1a	26	-	m10m1	
VI	VI 1b	VI 1b	33	-	m47m4, m60m4	
VI	VI 2	VI 2	28	8+173	m10m7, m54m7	
VI	VI 3	VI 3+VI 8	29+70	7	m10m3, m31m6, m37m1, m54m6, m60m3, m73m3	ADS1Eco, ADS3Hin, ADS4Eco, ADS6Hin, ADS8Hin, R087Hin, R152Hin, R162Eco, R196Eco, R272Eco, R367Dra, R396Hin, R482Eco
VI	VI 4	VI 4	31	125	m9m2, m68m1	R196Ecoc, R196Ecod
VI	VI 9	VI 9	32	32	m37m5, m47m5, m57m3	ADS3Hin10, ADS6Hin4, ADS8Hin3b, R272Ecoa, R482Ecod
VI	VI 5	VI 5	94	-	m10m2, m58m2	ADS6Hin8, ADS8Hin3h, R272Ecod
VI	VI 10	VI 10	25	66	m37m4, m57m1, m60m2	
VI	VI 11	VI 11	27	-	m60m1, m73m1	
VI	VI 12	VI 12	30	113	m37m3, m57m2	
VI	VI 15	VI 15	109	169	m10m5	R 367Draa

Table 2 (continued)

HG ⁴	Cons ⁵	Rossi ¹	JYH ²	LMR ³	SSR	RFLP
VI	VI 20	U 11	77+104	164+178	m54m8	R087Hinb, R152 Hinb, R162Ecoa, R162Ecob, R162Ecoc
VI	VI a	VI a	-	-	m10m6	R196Ecof, R272Ecob
VI	VI b	VI b	-	-	m58m1	R196Ecoa, R272Ecog
VI	VI c	VI c	57	11		ADS4Eco6, R396Hinb
VI	VI u52	U 52	56	12	m23m2, m54m2, m7m1	ADS4Eco5
VI	VI d	VI d	61	65+110		R367Dra, R396Hinc
VI	VI u19	U 19	73	19		ADS3Hin4, ADS8Hin7
VI	VI u36	U 36	95	16		ADS1Eco4, R272Ecoc
VI	VI u60	U 60	75	72	m23m1, m54m1	
VI	VI u51	U 51	54	-		R152Hinc, R396Hina
VI	VI u45	U 45	117	-	m31m1	
VII					mB41, m4, m8, m21, m29, m36, m43, m67, m69*	R019Sst, R137Sst
VII	VII a	VII a	80		m4m4	R019Sstc, R 137Ssta, R 137Sstb, R 137Sstf
VII	VII la	VII la	60	146	m4m5, m43m1	R137Std, (Rust resistance gene)
VII	VII lb	VII lb	13	48	m8m3, m67m1, m69m1	
VII	VII 2	U 12	53	-		
VII	VII 3	U 14	11	-		
VII	VII 4	U 50	67	49	m8m4	
VII	VII 5	U 13	12	-		
VII	VII 8	VII 8	63	-	m21m1, m36m2, m67m8	
VII	VII 9	VII 9	-	-	m4m8, m43m3	
VII	VII 14	VII 14	-	-	m4m3	R019Sste
VII	VII 18	U 2	-	-	m21m3, m29m3	
VII	VII 20	U 40	39	129	mB41m2, m29m1	
VIII					mA16*, m1, m12, m17, m18, m26, m28, m38, m44, m46, m61, m65, m73*	ADS7Hin, ADS9Eco, R039Sst, R082Dra, R118Eco, R131Eco, R142Eco*, R 173Sst, R185Sst, R267Sst, R281Hin, R326Sst, R335Eco, R342Sst, R406Sst, R408Sst, R488Hin, R526Eco*, R542Hin
VIII	VIII a	VIII a	108	184	mA16m3	R118Ecoa, R185Ssth, R281Hind, R281Hing, R326Ssta, R326Sstf
VIII	VIII b	VIII b	65	55		R185Sstb, R267Sste, R326Sstb
VIII	VIII d	VIII d	55	156	m61m2	R542Hinb
VIII	VIII u16	U 16	58	-		
VIII	VIII e	VIII e	78	71		R542Hina
VIII	VIII f	VIII f	45	-	m17m4, m46m1	
VIII	VIII g	VIII g	102	-		R039Sstf, R488Hinc
VIII	VIII u58	U 58	101	-		R488Hind
VIII	VIII i	VIII i	14	67	m1m2, m12m4, m18m1, m38m1	R082Draa, R082Drab, R185Ssti, R281Hinb, R326Ssti
VIII	VIII 2	VIII 2	15	5	m1m1, m12m5, m17m3, m28m6, m44m1, m46m2, m65m3	ADS7Hin2, R039Sstd, R118Eco, R131Eco, R173Sst, R185Sst, R267Sst, R281Hine, R326Sste, R335Ecog, R342Sstb, R406Ssta
VIII	VIII 11	VIII 11	16	56	m17m2, m26m3, m65m2	R039Ssta, R335Ecoa, R335Ecob, R342Sste, R342Sstd, R488Hinb
VIII	VIII u49	U 49	18	-		
VIII	VIII 3	VIII 3	-	-	m12m1	R142Eco, R526Ecoc
VIII	VIII 4	VIII 4	43	155	m12m8, m28m2	R142Ecoc, R526Ecoa
VIII	VIII 5	VIII 5	20	1	m38m2, m73m2	ADS9Eco2, R118Ecob, R 173Sstf, R173Sstd, R173Ssta, R185Sstf, R267Sstc, R281Hina, R326Sstd
VIII	VIII 16	VIII 16	21	-	m61m3	R173Sste
VIII	VIII 8	U 53	17	-		
VIII	VIII 10	VIII 10	19	-	m18m3	
VIII	VIII 15	VIII 15	106	-	m38m4	R118Ecoe, R118Ecob, R173Sstb, R173Sstg, R281Hinf
-	U 1	U 1	99+110	-	m49m6	
-	U 4	U 4	69	172	m23m3	
-	U 5	U 5	-	-		R016Drae
-	U 15	U 15	-	-		R482Ecob
-	U 27	U 27	81	176		R149Ecod
-	U 42	U 42	120	181	m4m2	R012Ecob
-	U 43	U 43	-	-		R031Ecoa
-	U 46	U 46	-	-		R129Ecob, R371Drab
-	U 47	U 47	100	112		R088Ecoa
-	U 55	U 55	118	-	R129Ecoc	
-	-	-	-	175	m54m4	
-	-	-	-	186	m54m5	
-	-	-	-	46	m16m1	
-	-	-	-	58	m4m6	

¹Rossi *et al.* (2003); ²Hoarau *et al.* (2001), ³This study, ⁴Proposed Homology Groups, ⁵Consensus map integrating informations from ^{1,2,3}

* indicates Probe/enzyme combinations or SSR primers producing alleles (bands) attributed to CGs belonging to different HGs

^{\$}Resistance gene analogs mapped by Rossi *et al.* (2003)

Table 3: Observed ratios of resistant and susceptible clones in three different populations and the Chi-square test of two hypotheses: one major gene is segregating in the population (1:1 ratio is expected) or two major genes are segregating (3:1 ratio is expected).

Studied populations	Parental rust resistance status		Number of progenies			χ^2	
	Female (R/S)	Male (R/S)	total	Resistant	Susceptible	2 doses ^a	1 dose ^a
Full population of clones evaluated for rust resistance	R570 (R)	MQ76-53 (R)	166	105	61	11.66 *** ↔	12.22 ***
Subpopulation of clones without markers flanking R570 resistance gene <i>Bru1</i>	R570 (R)	MQ76-53 (R)	90	36	54	36.75 ***	3.60 ns
Control cross:	B63-758 (S)	MQ76-53 (R)	133	68	65	6.61 ***	0.07 ns

^a $\chi^2=3.84$ at the 5% level

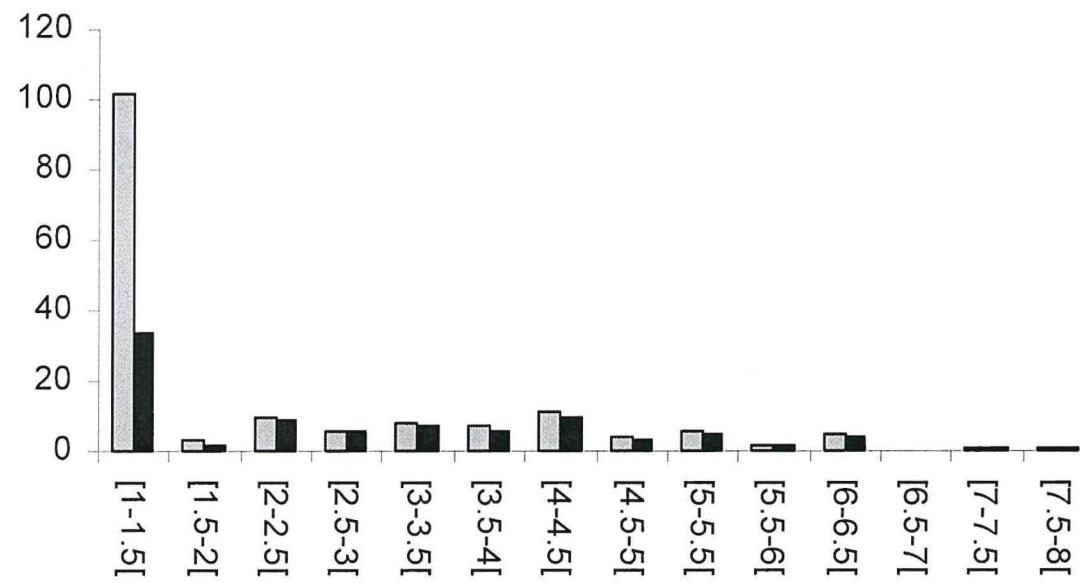


Figure 2: Distribution of rust resistance mean scores. In gray (166 unselected clones of the progeny) and in black (90 clones of the progeny selected because they do not have the AFLP markers flanking R570 rust resistance gene)

these segments appear to have an homogeneous species origin: 15 segments showed a putative *S.officinarum* origin, 4 segments a putative *S. spontaneum* origin, 1 segment a recombined origin and two segments had an unknown origin. The cumulative length of the common segments identified reached 263 cM which represents less than 10 percent of the cumulative length of R570 map in our study. Out of the 22 putative common chromosomal segments revealed by DD markers, 6 cases could be assigned to a homologous simplex-marker CG of MQ76-53. These 6 CGs, CG187, CG61, CG133, CG 103, CG36 and CG 6 are homologous to CGs I8, IV1, VI^d, U37 VIII1 and VIII2 of R570 respectively. On the basis of this homology, CG187, CG61, CG133, CG36 and CG6 were respectively assigned to Homology groups I, , IV, VI, VIII, VIII. For CG187 and CG36 this is congruent with their previous assignment on the basis of locus specific markers.

Identification and mapping of a putative major rust resistance gene in MQ76-53

The distribution of rust resistance scores of 166 progeny clones is presented on figure 2. It showed a clear segregation between resistant clones (mean score < 2) and susceptible clones (mean score ≥ 2) with a segregation ratio between 1:1 and 1:3 (Table 3). We already know that R570 possess one copy of a rust resistance gene (*Bru1*) that has been mapped, flanked with AFLP markers (Asnaghi et al. 2004). Two of these AFLP markers, aaccac6 mapped at 1 cM on the distal side of the gene, and actctg9R at 2.2 cM on the proximal side of the gene, have also been mapped in the present study. Therefore, we had the opportunity to select from the 166 clones evaluated for rust those that do not carry the *Bru1* gene. To do this, we retained only clones with simultaneous absence of the two AFLP markers flanking *Bru1*. This left us with a subpopulation of 90 clones that do not bear the *Bru1* gene (Here we assume that the probability of having a double recombination (one recombination on each side of *Bru1*) within the 3.2 cM defined by the flanking AFLP markers is almost null). The distribution of rust resistance scores of these 90 clones presented on figure 2 shows a clear segregation between resistant clones (mean score <2) and susceptible clones (mean score ≥ 2). The segregating ratio is skewed but is not statistically different from 1:1(Table 3). This lead us to hypothesize the presence of a second rust resistance gene inherited from MQ76-53. This hypothesis is confirmed by the 1:1 segregation ratio (68 resistants: 65 susceptibles) observed within 133 progenies derived from a cross between the cultivar B63-758 (susceptible to rust) and MQ76-53 (Table 3).

Table 4: Significant associations between markers and rust resistance according to a Fisher's exact test at $P < 3 \times 10^{-5}$

marker	P<F	HG	CG	Position
actctg7	1.08×10^{-5}	VIII	3	83.7
putative gene		VIII	3	48.5
acgcta16	6.98×10^{-9}	VIII	3	25.4
agccta16	5.37×10^{-8}	VIII	3	23.5
aaggag16	3.77×10^{-7}	VIII	3	21.1
accctc4	3.77×10^{-7}	VIII	3	18.5
acccac2	2.36×10^{-6}	VIII	3	16.9
acactg6	8.30×10^{-6}	VIII	3	13.2
accctt10	1.47×10^{-5}	VIII	3	8
<i>actctt26</i>	1.73×10^{-5}			<i>unlinked</i>

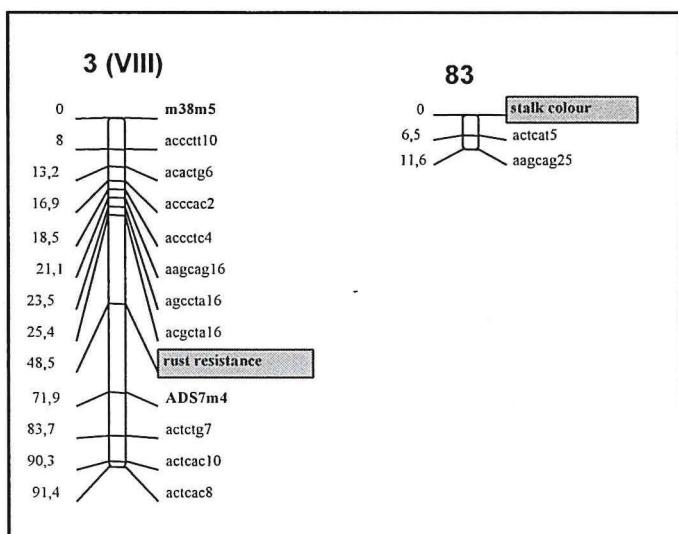


Figure 3: Representation of cosegregation groups of MQ76-53 genetic map bearing the putative new rust resistance gene (CG 3, HG VIII) and the putative gene controlling stalk colour (CG 83, unassigned to an HG). The detailed genetic map of MQ76-53 is available at <http://tropgenedb.cirad.fr/en/sugarcane.html>.

For mapping purposes, rust resistance, in the subpopulation of 90 clones not bearing *Brul*, was used as a morphological marker coded 1 for resistant clones and 0 for susceptible clones. This marker was localized on CG3 of MQ76-53 map (Figure 3) which was assigned to HGVI thanks to two locus specific markers (ADS7 and m38). A Fisher's exact test (Mehta and Patel 1983; SAS institute 1990, FREQ procedure) performed between rust resistance and each of the 1666 segregating markers at the conservative threshold of $P = 3 \times 10^{-5}$ ($P=0.05/1666$) revealed that the only markers significantly associated to the rust resistance gene belonged to CG3, except an unmapped DD marker (Table 4). This test which confirmed the localisation of the putative rust resistance gene on CG3 showed that the most tightly associated marker is acgcta16 (Table 4) located at 23.1 cM from the gene (Figure 3).

Identification and mapping of a major gene controlling stalk colour in MQ76-53

Stalk colour of 198 progeny clones has been regularly checked in various trials and in the collection after stripping off leaves from a few stalks. A clear segregation was observed between 91 red stalk clones and 89 non-red stalk clones (this subpopulation of clones could have green stalks or yellow stalks). Eighteen clones were not classified because of difficulties to decide between the two categories. For example, some clones appeared lightly red in some plots and non red in others and some clones presented a peculiar striped colour pattern. These few ambiguous clones were removed from the analysis. The resulting observed segregation ratio was not statistically different from 1:1 ($\chi^2 = 0.02$), suggesting that red stalk colour is under the control of a major gene inherited from MQ76-53. Stalk colour was therefore used as a morphological marker coded 1 for red stalk clones and 0 for non-red stalk clones for mapping purposes. The putative major gene was localized on cosegregation group83 of MQ76-53 map closely linked at 6.5 cM to an AFLP marker (actcat5 see figure 3).

III.5 Discussion

The cumulative lengths of the genetic maps developed from our bi-parental population were 3144 cM and 4329 cM for R570 and MQ76-53 respectively. For R570, the map coverage (3144 cM) is less than that of the map built from a selfed progeny of R570 (5849 cM in Hoarau et al. (2001)) although the number of AFLP primer combinations used was nearly the same. As expected, the same genotyping effort yielded less informative markers for the

mapping of a given parent when comparing a bi-parental population to a selfed population. This is because it is not possible to exploit markers common to both parents for mapping. In our study, 168 double dose (DD) markers common to R570 and MQ76-53 (identical alleles or homoplasic band level) were not usable for mapping whereas most of them had been mapped in the selfed progeny of R570. These DD markers of bi-parental origin lost for mapping represent at least 26 % of the markers in single dose in R570 genome in this study. Moreover, this loss of markers is underestimated since it does not take into account single dose markers of R570 not coded due to confounding effect with multiple doses brought by MQ76-53.

MQ76-53 is a F1 hybrid between an old cultivar and a wild *S. spontaneum* clone (SES 528) that is not found in the recorded pedigree of any international cultivars (Machado 2001). Therefore this clone is presumably rather genetically distant from R570 and from other present-day cultivars. This was confirmed by an AFLP-based diversity study that included R570, MQ 76-53 and 72 modern cultivars from various breeding programs. This study revealed that the proportion of single dose AFLP bands mapped in R570 and common to R570 and the other cultivars ranged from 40% up to 65%(unpublished data). These figures give an interesting estimation of the potential loss of informative markers in a bi-parental mapping project. The minimum proportion of 40% was observed between R570 and MQ 76-53. Consequently, we can expect the loss of informative markers for mapping to be even higher in a cross between two classic modern cultivars. Despite the lower yield of useful markers for mapping, a bi-parental mapping population seems more suitable than a selfed population when the final objective is QTL detection especially if the two parents have highly contrasted phenotypes for the trait of interest. Indeed, in a bi-parental progeny (1) the two populations to be compared (the one with the marker to be tested and the one without it) are of much more balanced size (1:1) than in a selfed progeny context (3:1) (2) the buffering effect of the background of alternative alleles should be less strong because half of this background is inherited from the “contrasted tester”. The detection power of QTLs is therefore expected to be stronger in a bi-parental context other parameters being equal (population size, type1-error...) and should help detect with better accuracy at least the most favorable alleles.

The alignment of the different R570 AFLP-based maps allowed the cosegregation groups of R570 to be grouped into 7 HGs (Homology groups) containing between 2 (HG IV) and 22 (HG VI) CGs. Structural differences between *S. spontaneum* and *S. officinarum* are expected

since they have different basic number $x=8$ and $x=10$ respectively (D'Hont et al. 1998). These differences may result from simple fusion or fission events as suggested for HG VIII in which two sets of homologous *S. officinarum* chromosomes are assigned to the same HG (VIII) because of their homology to the same *S. spontaneum* CG (Grivet et al. (1996); D'Hont et al. (1996) and this study). Thus, considering the basic chromosome number of *S. spontaneum*, we may have expected the constitution of 8 HGs instead of the 7 found up to now. Tentative explanations can be proposed for this situation: (1) genome coverage with locus specific markers is still insufficient in order to identify all HGs (some unassigned CGs may belong to the missing HGs) (2) duplication of some SSR and RFLP locus within the basic chromosome set may have led to wrongly assemble sets of homologous CGs into the same HG. (Butterfield et al. 2001) (3) complex structural differences between the basic chromosomes set of *S. officinarum* and *S. spontaneum* may also have led to wrongly assemble sets of homologous or homeologous CGs. Mapping additional locus specific markers (SSRs and RFLPs) should help to resolve this question.

No linkage in repulsion was detected between CGs of MQ76-53 while preferential pairing was observed for 18 pairs of CG in R570. This contrasting situation may be in part explained by the differences of genome constitution between these two clones. MQ76-53 is a clone originating from a cross between a cultivar (Trojan) and a *S. spontaneum* clone (SES 528) which has $2n=64$ chromosomes (Panje and Babu 1960). The *S. spontaneum* component of the MQ76-53 genome is therefore composed of at least 4 copies of each *S. spontaneum* basic chromosome inherited from SES 528 plus the *S. spontaneum* chromosomes inherited from Trojan. By comparison, R570 is composed of only one or two *S. spontaneum* chromosomes per homologous class. Thus, the *S. spontaneum* chromosome component of MQ76-53 is roughly of equal size to that of *S. officinarum* and no preferential pairing, could be detected within this component at LOD=3. This picture is in agreement with the lack of preferential pairing observed in *S. spontaneum* SES 208 that suggests polysomic inheritance and autopolyploidy in this species (Al Janabi et al. 1993; Ming et al. 1998). However, the absence of preferential pairing among MQ76-53 *S. officinarum* chromosomes would be a departure from the observation of incomplete polysomy in *S. officinarum* (Mudge et al. 1996; Guimaraes et al. 1997). That may be due to a dramatic deficiency in the coverage of the *S. officinarum* component of the MQ76-53 genome that renders it almost invisible. In R570, chromosome pairing behaviour is complex. Preferential pairing involves CGs of *S. officinarum* origin as well as CGs recombinant between *S. spontaneum* and *S. officinarum* and

CGs of sole *S. spontaneum* origin (Hoarau et al. 2001 and Grivet et al. 1996). The fine analysis of all pairing frequencies between homo(homeo-)logous chromosomes of a single Homology group (HGI) in R570 revealed pairing affinities between chromosomes ranging from 0 to 40 %, chromosome affinities being only partly explained by their species origin (Jannoo et al. 2004).

Segments of the genome are common between R570 and MQ76-53. Their cumulative length represents around 8% of R570 map. MQ76-53 is an atypical sugarcane clone, rather genetically distant from R570 and derived from the hybridisation between an old cultivar and a *S. spontaneum* (SES528) that has not been used elsewhere in breeding programs. On the contrary, modern cultivars are usually derived from a limited number of founding interspecific crosses through less than ten meiosis. Thus we hypothesize that in other bi-parental mapping projects involving classic modern cultivars, the same approach should lead to the identification of a higher proportion of common chromosome segments. The identification of common chromosome segments between R570 and MQ76-53 illustrate the linkage disequilibrium existing in sugarcane (Jannoo et al. 1999). Associations between alleles are thus expected to be conserved throughout sugarcane germplasm. This is a feature that should help trace alleles of agronomic interest.

A Mendelian factor governing stalk colour has been identified in the genome of MQ76-53. This gene is linked to two AFLP markers in a cosegregation group (CG 83). Although cane colour varies considerably according to age of the stalk and to the amount of sunlight received (Stevenson 1965), a rather clear 1:1 segregation ratio between red stalk and non-red stalks cane could be observed. This finding may not be relevant from an agronomic viewpoint but is another indication that even in the complex polyploid background of sugarcane it is possible to identify traits controlled by a single gene, and that genetic mapping efforts should continue in sugarcane to identify new major genes. In addition, a putative new rust resistance gene has been located in the MQ76-53 genome. This gene is clearly different from the R570 rust resistance gene *Bru1* (Asnagi et al. 2004). Indeed, none of the AFLP markers flanking *Bru1* in R570 is present in MQ76-53 indicating that this new putative resistance gene belongs to a different haplotype. Moreover, this gene is located on CG3 which should belong to the Homology group VIII whereas *Bru1* is located on a R570 CG that belong to Homology group VII. Rossi et al. (2003) identified three Resistance Gene Analogs clusters, including RGA of the NBS-LRR and S/T Kinase types, in homology group VIII. These RGAs could be used as

candidate genes and mapped on our population in order to check if they are located in the vicinity of this new rust resistance gene. This approach could help refine the location of the gene more rapidly (Pfleiger et al. 2001). In addition, it would be interesting to trace the origin of this resistance gene. If it is inherited from the *S. spontaneum* SES528 it may constitute an alternative source of resistance to rust. In this case, its occurrence in the genome pool of modern sugarcane cultivars could be surveyed in order to determine if introgressing this new haplotype from SES 528 into commercial breeding programs would be beneficial.

Acknowledgements:

We thank Iréné Promi for his help in the fieldwork. Experiments presented in this paper comply with the current laws of France

REFERENCES

- Aitken KS, Jackson PA, McIntyre CL (2005) A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. *Theor Appl Genet* 110:789-801
- Al-Janabi SM, Honeycutt RJ, McClelland M, Sobral BWS (1993) A genetic linkage map of *Saccharum spontaneum* L. ‘SES 208’. *Genetics* 134:1249–1260
- Asnaghi C, Roques D, Ruffel S, Kaye C, Hoarau JY, Telismart H, Girard JC, Raboin LM, Risterucci AM, Grivet L, D’Hont A (2004) Targeted mapping of a sugarcane rust resistance gene (*Bru1*) using bulked segregant analysis and AFLP markers. *Theor Appl Genet* 108:759-764
- Bhat SR, Gill BS (1985) The implication of 2n egg gametes in nobilisation and breeding of sugarcane. *Euphytica* 34:377–384
- Burner DM, Legendre BL (1993) Chromosome transmission and meiotic stability of sugarcane (*Saccharum* spp.) hybrid derivatives. *Crop Sci* 33:600-606
- Burner DM, Legendre BL (1994) Cytogenetic and fertility characteristics of elite sugarcane clones. *Sugar Cane* 1:6-10.
- Butterfield MK, D’Hont A, Berding N (2001) The sugarcane genome : a synthesis of current understanding, and lessons for breeding and biotechnology. *Proc S Afr Sug Technol Ass* 75:1-5
- Da Silva JAG, Burnquist WL, Tanksley SD (1993) RFLP linkage map and genome analysis of *Saccharum spontaneum*. *Genome* 36:782–791
- Da Silva J, Honeycutt RJ, Burnquist W, Al-Janabi SM, Sorrells ME, Tanksley SD, Sobral BWS (1995) *Saccharum spontaneum* L. ‘SES 208’ genetic linkage map combining RFLP- and PCR-based markers. *Mol Breed* 1:165–179
- Daugrois JH, Grivet L, Roques D, Hoarau JY, Lombard H, Glaszmann JC, D’Hont A (1996) A putative major gene for rust resistance linked with an RFLP marker in sugarcane cultivar R570. *Theor Appl Genet* 92:1059–1064
- D’Hont A, Lu YH, González de Leòn D, Grivet L, Feldmann P, Lanaud C, Glaszmann JC (1994) A molecular approach to unraveling the genetics of sugarcane, a complex polyploid of the *Andropogoneae* tribe. *Genome* 37 : 222-230
- D’Hont A, Grivet L, Feldmann P, Rao S, Berding N, Glaszmann JC (1996) Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol Gen Genet* 250:405–413

- D'Hont A, Ison D, Alix K, Roux C, Glaszmann JC (1998) Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome* 41:221-225
- D'Hont A, Garsmeur O, Raboin LM, Paulet F, Begum D, Wing R, Glaszmann JC (2001) Chromosome walking toward a major resistance gene for common rust of sugarcane. *Proc Int Soc Sugarcane Technol* 24: 315-317
- D'hont A, Glaszmann JC (2001) Sugarcane genome analysis with molecular markers, a first decade of research. *Proc Int Soc Sugarcane Technol* 24: 556-559
- Grivet L, Arruda P (2001) Sugarcane genomics: depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology* 5:122-127
- Grivet L, D'Hont A, Roques D, Feldmann P, Lanaud C, Glaszmann JC (1996) RFLP mapping in a highly polyploid and aneuploid interspecific hybrid. *Genetics* 142:987-1000
- Guimaraes CT, Sills GR, Sobral BWS (1997) Comparative mapping of Andropogoneae: *Saccharum* L. (sugarcane) and its relation to sorghum and maize. *Proc Natl Acad Sci USA* 94:14261–14266
- Heinze BS, Thokoane LN, Williams NJ, Barnes JM, Rutherford RS (2001). The smut-sugarcane interaction as a model system for the integration of marker discovery and gene isolation. *Proc S Afr Sug Technol Ass* 75: 88-93
- Hoarau JY, Offmann B, D'Hont A, Risterucci AM, Roques D, Glaszmann JC, Grivet L (2001) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp). I. Genome mapping with AFLP markers. *Theor Appl Genet* 103:84–97
- Hoarau JY, Grivet L, Offmann B, Raboin LM, Diorflar JP, Payet J, Hellmann M, D'Hont A, Glaszmann JC (2002) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp). II. Detection of QTLs for yield components. *Theor Appl Genet* 105:1027-1037
- Hoisington D (1992) Laboratory protocols. CIMMYT applied molecular genetics laboratory, CIMMYT, Mexico, D.F.
- Jannoo N, Grivet L, Dookun A, D'Hont A, Glaszmann JC (1999). Linkage disequilibrium among modern sugarcane cultivars. *Theor Appl Genet* 99: 1053-1060
- Jannoo N, Grivet L, D'Hont A, Glaszmann JC (2004) Differential chromosome pairing affinities at meiosis in polyploid sugarcane revealed by molecular markers. *Heredity* 93:460-467
- Jordan DR, Casu RE, Besse P, Carroll BC, Berding N, McIntyre CL (2004) Markers associated with stalk number and suckering in sugarcane collocate with tillering and rhizomatousness QTLs in sorghum. *Genome* 47:988-993

- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Machado GR (2001). Sugarcane variety notes – An international directory 7th revision, Piracicaba, Int. Soc. Sugar Cane Techn. Februray 2001, 132 p
- Mather K (1957) The measurement of linkage in heredity. Methuen, London.
- Mehta CR, Patel NR (1983) A network algorithm for performing Fisher's exact test in rxc contingency tables. *J Am Statist Assoc* 78:427-434
- Ming R, Liu S-C, Lin Y-R, Da Silva J, Wilson W, Braga D, Van Deynze A, Wenslaff TF, Wu KK, Moore PH, Burnquist W, Sorrells ME, Irvine JE, Paterson AH (1998) Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. *Genetics* 150:1663–1682
- Ming R, Liu SC, Moore PH, Irvine JE, Paterson AH (2001) QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. *Genome Res* 11:2075–2084
- Ming R, Del Monte TA, Hernandez E, Moore PH, Irvine JE, Paterson AH (2002a) Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genomes. *Genome* 45:794-803
- Ming R, Liu SC, Bowers JE, Moore PH, Irvine JE, Paterson AH (2002b) Construction of a *Saccharum* consensus genetic map from two interspecific crosses. *Crop Sci* 42:570-583
- Mudge J, Andersen WR, Kehler RL, Fairbanks DJ (1996) A RAPD genetic map of *Saccharum officinarum*. *Crop Sci* 36: 1362–1363
- Panje RR, Babu CN (1960). Studies in *Saccharum spontaneum*. Distribution and Geographical association of chromosome numbers. *Cytologia* 25:152-172.
- Pfleiger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. *Mol Breed* 7:275-291.
- Pinto LR, Oliveira KM, Ulian EC, Garcia AA, de Souza AP (2004) Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome* 47(5):795-804
- Price S (1963) Cytogenetics of modern sugar canes. *Economic Botany* 17: 97–105
- Rossi M, Araujo PG, Paulet F, Garsmeur O, Dias VM, Chen H, Van Sluys MA, D'hont A (2003) Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. *Mol Gen Genet* 269:406-419
- Ruiz M, Rouard M, Raboin LM, Lartaud M, Lagoda P, Courtois B (2004) Tropgene-DB, a multitropical crop information system. *Nucleic Acids Res* 32:D364-D367

- Stevenson GC (1965) Genetics and Breeding of Sugar cane. Longmans, London pp284
- SAS Institute (1990)SAS procedures guide, version 6, third edition. SAS Institute Inc, Cary, N.C.
- Tai PYP, Miller JD, Dean JL (1981) Inheritance of resistance to rust in the sugarcane field. Crops Res 4:261–268
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
- Wu KK, Burnquist W, Sorrels ME, Tew TL, Moore PH, Tanksley SD (1992) The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor Appl Genet 83:294–300

IV STRUCTURE DU DESEQUILIBRE DE LIAISON

IV Etude de la structure du déséquilibre de liaison dans une population non structurée de cultivars de canne à sucre à l'aide de marqueurs AFLP

Le déséquilibre de liaison (DL) est la mesure de l'association non aléatoire entre allèles à des locus distincts. Chez les plantes cultivées, le déséquilibre de liaison est plus ou moins fort selon l'histoire de leur domestication par l'homme et leur régime de reproduction. De l'intensité du déséquilibre de liaison va dépendre, la résolution avec laquelle on pourra localiser les gènes par association avec des marqueurs et aussi l'effort de génotypage qu'il faudra fournir pour atteindre ce résultat (Flint-Garcia et al. 2003; Gupta et al. 2005). Chez la canne à sucre en particulier, les variétés modernes résultent de quelques hybridations interspécifiques réalisées au début du siècle dernier en Inde et en Indonésie, dont les produits ont ensuite été entrecroisés moins d'une dizaine de fois pour aboutir aux cultivars modernes de canne à sucre. Il y a donc eu un fort effet de fondation créateur de DL. Ce déséquilibre de liaison a été conservé dans la population des cultivars modernes de canne à sucre comme cela a été démontré par Jannoo et al. (1999). Ainsi, des associations ont été détectées entre locus RFLP distants de plus de 20 cM. Le déséquilibre de liaison chez la canne devrait donc conférer à la diversité moléculaire en un point du génome une valeur prédictive pour la diversité au niveau de gènes utiles présents dans le voisinage.

La population de 74 cultivars que nous avons constitué pour réaliser l'étude d'association a fait l'objet d'un effort de génotypage important par AFLP. Le codage des marqueurs a été effectué en référence au codage des marqueurs AFLP produits pour la construction des cartes de 'R 570' et 'MQ 76/53' (40 couples d'amorces sur 42 sont communs aux deux études). En effet, ces deux clones encadrent systématiquement les cultivars de la population sur les gels de migration. L'objectif ici, en plus de l'objectif initial qui est de repérer des associations entre marqueurs et résistance au charbon, est de tenter d'affiner la compréhension de la structure du déséquilibre de liaison chez la canne à sucre. Notamment, il devrait nous être possible grâce aux cartes génétiques (celles construites dans le cadre de cette thèse et celle de Hoarau et al. 2001) d'établir une relation assez fine entre déséquilibre de liaison et distance génétique sur les cartes.

Ce travail est présenté sous la forme d'un projet d'article intitulé « **Genome wide analysis of Linkage Disequilibrium in the high polyploid sugarcane using AFLP markers** »

Genome wide analysis of Linkage Disequilibrium in the high polyploid sugarcane using AFLP markers

Louis-Marie Raboin, Jérôme Pauquet, Mike Butterfield, Angélique D'Hont and Jean-Christophe Glaszmann

IV.1 Introduction

Sugarcane represents more than two thirds of the world sugar production, a share that is continuously growing. Sugarcane breeding was initiated at the end of the 19th century and greatly contributed to improve the productivity of this industry. Sugarcane cultivars derive from interspecific hybridization between the domesticated sugar producing species *Saccharum officinarum* ($x=10$, $2n=8x=80$) and the wild species *Saccharum spontaneum* ($x=8$, $2n=40-128$) followed by several backcrosses to *S. officinarum*. Subsequently, sugarcane breeding has relied on the reccurent intercrossing of elite hybrid cultivars and selection among the progeny. Modern sugarcane cultivars have a complex aneuploid and polyploid genome constituted of 100 to 130 chromosomes representing an estimated 10^9 base pairs. Genomic In Situ Hybridisation (GISH) revealed that 80% of R 570 (a typical modern cultivar) chromosomes were inherited from *S. officinarum*, 10% from *S. spontaneum* and 10% were derived from recombinations between the two ancestral species (Dhont et al. 1996). These features hamper the efforts made toward the identification and localization of loci that contribute to the variation of important agronomic traits. The best genetic maps are far from saturated mainly because the highly polyploid and redundant *S. officinarum* compartment of the genome is poorly covered. The effects of the numerous detected QTLs are small (Grivet and Arruda 2001). Mining alleles relevant for breeding is a real challenge in sugarcane. It will require the combined implementation of new high-throughput marker technologies, synteny with related grasses (sorghum and rice), continuous efforts in genetic mapping and the exploration of sugarcane germplasm thanks to linkage disequilibrium based approaches (association mapping and candidate gene association studies).

Linkage disequilibrium (LD) which is the non random association of alleles at distinct loci has been exploited to map human disease genes for some times (Cardon and Bell 2001) but

has only recently become the focus of interest in plant genetics. Unlike linkage analysis, association studies do not require the constitution of segregating populations of known parentage. They are based on existing populations or germplasm collections which may present various advantages: (1) individuals in such collections may be more distantly related and may have accumulated recombination events over many generations allowing high resolution mapping (2) individuals in such collections may be well characterized for a range of interesting traits (3) individuals in such collections may encompass a significant proportion of useful alleles (Morgante and Salamini 2003, Rafalski and Morgante 2004). The extent of LD determines whether genome scans or candidate gene association approaches can be used (Flint-Garcia et al. 2003). *Arabidopsis*, for which LD decays within 250kb, or 1 cM is well suited for genome-wide association mapping (Nordborg et al. 2002). Maize on the other hand is characterized by a rapid decline of LD within a few hundred base pairs (Tenailleon et al. 2001) which makes genome-wide association mapping nearly impractical because it would necessitate hundreds thousands of markers. Very few attempts have been made so far to identify marker-trait associations through whole-genome scan as reported in sugar beet for mapping the bolting gene (Hansen et al. 2001) or in Barley for mapping yield, a complex quantitative trait (Kraakman et al. 2004). More results have been published on associations with markers of targeted genomic regions or directly with candidate genes. In *Arabidopsis*, LD mapping has been used to examine haplotype variation in the targeted region of the CRY2 gene associated to flowering time variation. One of the identified early flowering haplotypes could even be associated with a unique serine substitution in the cry2 gene (Olsen et al. 2004). In tetraploid potato, LD has been successfully used to find associations between disease resistance and candidate gene markers (Simko et al. 2004).

Sugarcane breeding history is recent and is based on the use of a limited number of founder parents. Only a few *S. officinarum* and *S. spontaneum* clones, either directly or through *S. barbieri* (probable natural hybrid between *S. officinarum* and *S. spontaneum*) were used to create the first interspecific hybrids (Arceneaux 1965). Sugarcane breeding is a recurrent process in which each breeding cycle takes between 10 and 15 years. Consequently, only a few generations separate modern cultivars from the first interspecific hybrids and only a few meioses have had the opportunity to recombine chromosomes inherited from founder sugarcane clones. Therefore, a strong linkage disequilibrium should have been conserved among modern sugarcane cultivars. This assumption was confirmed in a pioneer work which revealed significant associations between RFLP markers, previously mapped in cultivar 'R

Table 1: Identification of the 74 sugarcane clones along with their parentage and breeding station

Clones	Parent1	Parent2	Breeding station	Clones	Parent1	Parent2	Breeding station
R 570	H 32/8560	R 445	Reunion	FR 84/344	KWT 56/26		FWI
MQ 76/53	TROJAN	SES 528	Australia	FR 84/387	NA 63/90		FWI
B 47/258	B 39/254	B 34/104	Barbados	FR 84/166	B 80/574		FWI
B 51/129	B 45/170	B 41/227	Barbados	H 32/8560	Co 213	POJ 2878	Hawaii
B 63/119	B 49/6	B 49/119	Barbados	H 39/3633	H 32/8560		Hawaii
B 66/23	M 147/44	B 49/119	Barbados	H 39/7028	H 32/8560		Hawaii
B 75/524			Barbados	H 44/3098	H 32/8560		Hawaii
B 80/66	B 75/738		Barbados	H 49/5	H 41/3340	H 37/1933	Hawaii
B 80/6	B 73/348	B 74/172	Barbados	H 50/7209	H 44/3098		Hawaii
B 82/288	B 74/142	B 73/428	Barbados	H 61/1721	H 49/3533		Hawaii
B 85/356	WI 73/48	BJ 63/132	Barbados	IAC 64/257	Co 419	IAC 49/131	Brazil
B 86/406	F 146	BR 62/49	Barbados	J 59/3			Jamaica
B 86/409	BJ 74/59		Barbados	Ja 64/19	Ja 55/663	Ja 54/309	
B 86/839	B 66/210	CR 68/188	Barbados	LF 53/4789			Fiji
B 87/1172	WI 80/703	CR 63/100	Barbados	LF 53/4825			Fiji
B 77/84	CP 52/43	HJ 57/41	Barbados	M 202/46	Co 281	M 63/39	Mauritius
BJ 78/128			Barbados	Mex 68/P23	Mex 59/89		Mexico
BR 71/48	B 50/135	B 49/119	Barbados	MY 53/53	B 42/231	Co 453	Cuba
BR 75/48	B 63/118	B 56/95	Barbados	MY 55/14	CP 34/79	B 45/181	Cuba
BR 79/4	L 60/14		Barbados	N 12	NCo 376	Co 331	South Africa
BT 72/344			Barbados	N 17	NCo 376	CB 38/22	South Africa
C 227/59	EK 2	POJ 2878	Cuba	NA 56/62	Co 290	CP 43/74	Argentina
CB 56/171	POJ 2961		Brazil	NCo 310	Co 421	Co 312	India
Co 1157	Co 419		India	NCo 376	Co 421	Co 312	India
Co 1177	Co 677	POJ 2961	India	Phil 56/226	POJ 2878	CP 36/105	Phillipines
Co 1186	Co 312	Co 617	India	Phil 66/7	Phil 54/60	Co 440	Phillipines
Co 1208	Co 312	CoL 9	India	PR 61/632	S 56/287	M 336	Puerto Rico
Co 449	POJ 2878	Co 331	India	Q 84	TROJAN	Co 475	Australia
Co 462	Co 421	Co 313	India	R 526	POJ 2878	R 397	Reunion
Co 842	Co 464	Co 617	India	R 574	H 39/3633	R 567	Reunion
CP 61/37	CP 48/103	CP 55/38	USA	RB 70/96	CB 36/14		Brazil
CP 66/315	CP 52/68	CP 53/17	USA	RB 72/5828	NA 56/79		Brazil
CP 70/1133	CP 56/63	67 P 6	USA	SP 70/1005			Brazil
D 172			Guyana	SP 70/1284	CB 41/76		Brazil
DB 73/419	B 67/128	B 63/118	Barbados	SP 70/3225			Brazil
DB 80/104	B 73/405	WI 73/14	Barbados	SP 70/1423			Brazil
F 160	NCo 310	F 141	Taiwan	SP 71/6113	Co 775		Brazil

570' (Grivet et al. 1996), distant of up to 10 cM (Jannoo et al. 1999a). Here we report on the extensive analysis of LD in a population of 74 sugarcane cultivars using the potential of the AFLP technique to efficiently cover the large polyploid genome of sugarcane with markers. Hundreds of AFLP markers have already been mapped in cultivar 'R 570' (Hoarau et al. 2001) or in sugarcane clone 'MQ 76/53' (Raboin et al. in prep). This valuable mapping information was used in order to investigate the genome wide structure and the extent of LD in sugarcane. The objective was to provide a guideline for the practical use of association studies to trace alleles of breeding value in sugarcane germplasm.

IV.2 Materials and methods

Plant material:

The studied population consisted of 74 sugarcane clones from various breeding stations around the world (Table 1). Clones from Barbados are the most numerous. Two of these clones that have been the focus of genetic mapping with AFLP markers were used as repeated controls in the DNA analysis: 'R 570' ('H 32/8560' crossed with 'R 445') a modern cultivar from Reunion and 'MQ 76/53' ('Trojan' crossed with a *S. spontaneum* clone 'SES 528') an old Australian sugarcane clone whose genetic structure should be close to that of an F1 interspecific hybrid.

AFLP protocol

Genomic DNA was extracted and prepared for AFLP analysis according to Hoarau et al. (2001). AFLP analysis (Vos et al. 1995) was performed using the invitrogen AFLP analysis system I. Gibco as recommended by the manufacturer except for slight modifications as in Hoarau et al. (2001). The 74 clones have been genotyped using 42 primer pairs of which 40 had been previously used for the genetic mapping of either 'R 570' or 'MQ 76/53'. 'R 570' and 'MQ 76/53' have been used on the acrylamide gels as systematic controls flanking each of the other 72 samples. Therefore we were able to identify unambiguously many of the AFLP bands previously mapped.

Statistical analyses

Genetic structure of the population

Artifactual or spurious associations (i.e. associations between unlinked markers or between phenotypes and markers at non causative genome regions) can arise because of population admixture. The population homogeneity should therefore be checked before performing an association study (Lander and Schork 1994).

Genetic similarities between all pairwise combinations of clones were calculated using the Dice index (Nei and Li 1979) according to the following formula:

$$GS_{ij}=2N_{ij}/(N_i+N_j)$$

where GS_{ij} is the measure of the genetic similarity between sugarcane clone i and j , N_{ij} is the number of bands in common between i and j , N_i and N_j are the number of bands in i and j respectively. The matrix of pairwise similarities was then used to build a Neighbour Joining (N-J) tree using the Darwin software (Perrier et al. 2003).

Coefficient of parentage

A subset of 30 clones, out of the 74 in the population, for which genealogical records were available up to the great-grandfather generation have been retained. We also collected all available genealogical information connected to the great-great grand father generation; up to the founding interspecific crosses. The coefficient of parentage among remote ancestors for which parentage was unknown, was assumed equal to 0. The coefficient of parentage has been calculated for all pairwise combinations among the 30 selected clones (435 combinations) using the INBREED procedure (SAS institute Inc. 1996). The Pearson's correlation between coefficient of parentage and genetic similarity has been calculated.

Testing for significant associations between markers

All the classic measures of LD (D' , r^2 , d^2) are related to the standard χ^2 statistics for a 2x2 contingency table (Nordborg and Tavare, 2002) but exploit allele frequency or haplotype

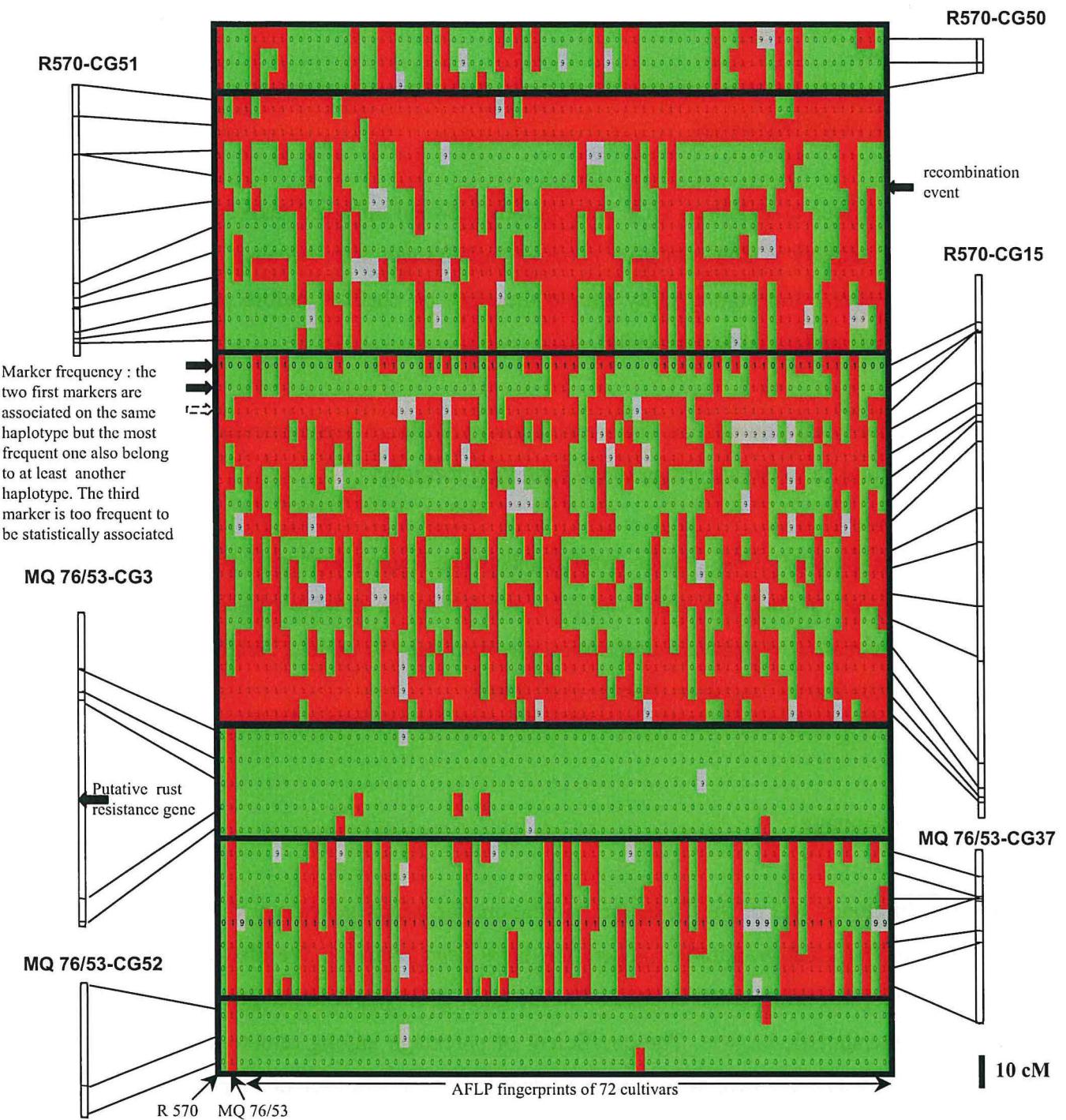


Figure 1: Excerpts from the raw AFLP data file. The 74 studied sugarcane genotypes are presented in column, the first two column correspond to the controls R 570(1st column) and MQ 76/53 (2nd column). Markers are presented in lines and were sorted according to their positions on the genetic maps. The upper part of the figure correspond to markers mapped in 'R 570' and the lower part to markers mapped in 'MQ 76/53'. Presence of a marker, coded 1, is highlighted in red whereas absence of the marker ,coded 0, is highlighted in green and missing data, coded 9, were highlighted in grey.

frequency. Because of the high polyploidy of sugarcane and because of the dominant nature of the markers used, allele frequency or haplotype frequency could not be calculated. We therefore used the Fisher exact probability to test for associations between markers.

For each pair of markers a 2x2 contingency table (presence versus absence) the Fisher probability and the correlation coefficient Phi were computed (Mehta and Patel, 1983; SAS Institute Inc., 1990).

To define a threshold that will distinguish between ‘true’ associations (i.e. associations between markers that are genetically linked on the same chromosome segment and therefore belong to a same haplotype) and ‘false’ (i.e. random associations between markers occurring because of the important number of pairwise comparisons performed) we used two methods. The first one was to apply the Bonferroni procedure by dividing the significance threshold of 5% by the number of comparisons performed. The second method took advantage of our knowledge of the position of part of the AFLP markers produced on the genetic map of ‘R 570’ (Hoarau et al. 2001; Rossi et al. 2003; Raboin et al. submitted). A great part of those mapped AFLP markers have been assigned to Homology groups. Therefore using the subset of markers for which the full mapping information is available (i.e. position on a cosegregation group as well as assignation to a homology group) we were able to define three classes of associations: intra cosegregation group associations, inter cosegregation group but intra homology group associations and finally inter homology group associations. The objective was to empirically set the threshold so that inter homology group associations (this kind of association should not correspond to haplotypes and were therefore considered as ‘False’ associations) would not represent over 5 percent of the total number of putative significant associations.

IV.3 Results

An overview of the raw AFLP data (Fig. 1)

A total of 1626 polymorphic markers were scored from the 42 AFLP primer pairs used. Among those markers, 737 belong to a cosegregation group of either the ‘MQ 76/53’ (258 markers) or ‘R 570’ (479 markers) AFLP genetic maps. The two controls, ‘R 570’ and ‘MQ

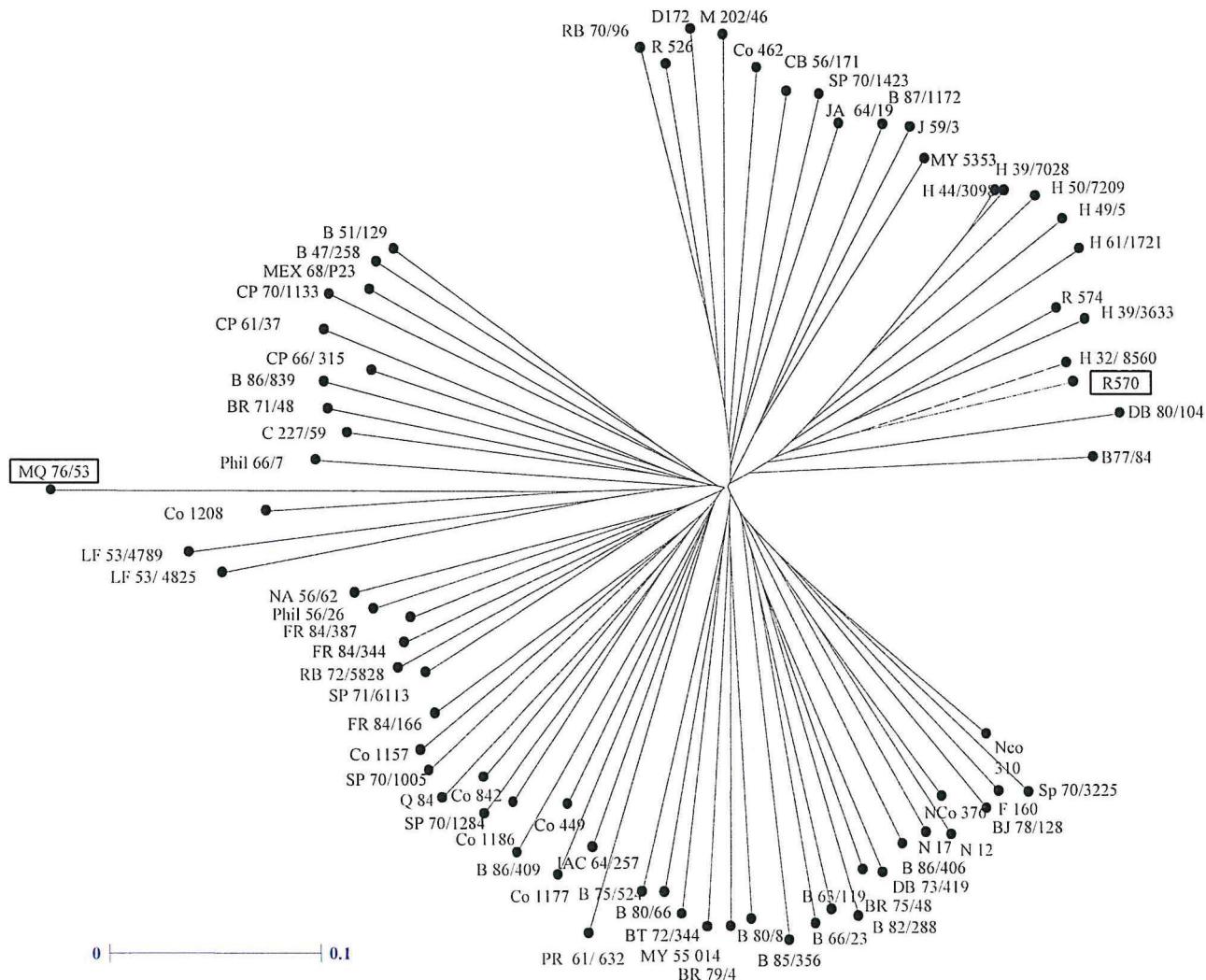


Figure 2: Neighbour-joining tree, based on the Dice dissimilarity index calculated from AFLP data (1626 polymorphic markers), presenting the 74 sugarcane genotypes. The two mapped cultivars used as controls ('R 570' and 'MQ 76/53') are presented in a box. Dotted lines have been drawn between genealogically closely related clones.

76/53', encompass 1254 of the 1626 markers scored in the population (77.1%). Markers of the controls are slightly over-represented as they encompass 49.6% and 50.1 % of the total number of scored markers for 'R 570' and 'MQ 76/53' respectively whereas the other cultivars encompass 44% of the total number of scored markers on average. This is a consequence of the systematic repetition of the controls on the gels. Each of the controls was repeated 36 times whereas the other cultivars were present only once on the gels which resulted in higher risk of artefacts and scoring errors and more missing data for those latter.

Ordering the markers of the AFLP data set according to their position in the cosegregation groups allowed the direct visualisation of linkage disequilibrium in sugarcane (Fig.1). Indeed it is possible to identify sequence of associated markers (putative haplotypes) that cosegregate in the studied population. Recent recombination events could even be identified because of ruptures in the visual pattern of LD. In these cases, the sequence of markers that is observed in a particular cosegregation group of cultivar 'R 570' results from the recombination of two distinct ancestral haplotypes that are still perceptible in the overall population of cultivars. AFLP markers show a great range of frequency, with 'MQ 76/53' clearly appearing as different from the other cultivars. Some of the 'MQ 76/53' markers (sometimes all markers of a cosegregation group) appeared to be specific to this clone. Most probably, those specific cosegregation groups or chromosomes were inherited from 'SES 528', the *S. spontaneum* parent of 'MQ 76/53'. Such markers are useless in an analysis of LD. It is noteworthy though that a putative new rust resistance gene was identified on CG 3 of 'MQ 76/53' (Raboin et al., submitted) whose markers are almost completely absent from the other cultivars. This gene may therefore constitute a new source of rust resistance, of putative *S. spontaneum* origin, worth introgressing into commercial breeding programs.

Population structure and correlation between coefficient of parentage and Genetic dissimilarity

No particular structuration was observed in the population of sugarcane clones (Fig. 2). Each sugarcane clone is roughly genetically equidistant from every other clone of the population which gave a star-like aspect to the N-J tree. Nevertheless, the few closely related clones (i. e. father-son relationship) are easily spotted: 'R 570' with 'H 32/8560', 'Nco 310' with 'SP 70/3225' or 'F 160' and 'Nco 376' with 'N 12' or 'N17' for example (Table 1 for declared

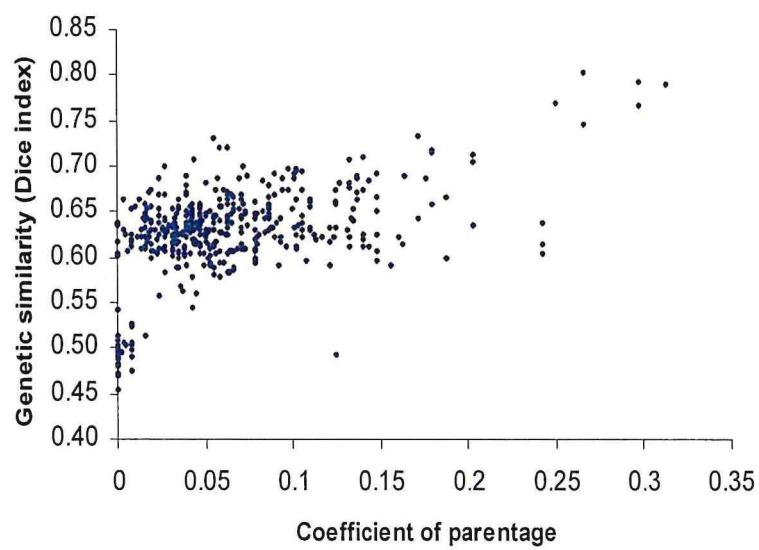


Figure 3: Plot of AFLP-based genetic similarity (Dice index) and coefficient of parentage for 435 pairwise combinations of sugarcane clones.

genealogical records). The control clone ‘MQ 76/53’ clearly distinguished itself from the other clones in agreement with its particular *S. spontaneum* origin. Two clones , ‘H 44/3098’ and ‘H 39/7028’, appeared surprisingly very closely related genetically. 65 differences (out of a total of 1465 markers with no missing data) have been observed between the two AFLP fingerprints. We propose two distinct explanations for such a result (1) both cultivars derive from the selfing of H 32/8560 (although in such a case we would have expected more differences between the two cultivars) (2) these two clones may have been mislabelled and could be identical and therefore the differences between the AFLP profiles may correspond to genotyping errors. In this latter case, it would set the error rate at 4.4%.

The coefficient of parentage was poorly correlated ($r=0.5$) to genetic similarity (Fig. 3). If we excluded the 29 pairwise comparisons involving MQ76/53 which correspond to particularly low genetic similarities (below 0.55) the correlation between the coefficient of parentage and the genetic similarity was equal to 0.44. This figure is close to that ($r=0.42$) calculated from 3043 pairs of sugarcane cultivars by Lima et al (2002).

‘True’ associations (i.e. associations between markers genetically linked on a same chromosome segment) versus ‘false associations’: Where is the threshold?

From the data set, we used 382 markers for which we had full information on their position in the cosegregation groups of the ‘R 570’ genetic map (Hoarau et al. 2001). A total of 72771 independent 2x2 Fisher exact test were performed corresponding to all possible pairwise combinations between these markers. Among these associations, 32177 had a negative PHI coefficient (i.e. the presence of one marker was associated with the absence of the other). These have been excluded from our interpretation scheme explained below. The remaining 40594 associations were divided between 3 groups: the group of associations between markers belonging to the same cosegregation group (group intra-CG containing 1082 associations), the group of associations between markers belonging to the same homology group but distinct cosegregation groups (group intra-HG containing 5759 associations) and finally the group of associations between markers belonging to different homology groups (group inter-HG containing 33753 associations). The analysis of the distribution of associations according to their Fisher probability within each of the groups defined above revealed that putative ‘true’ associations (within the intra-CG group) are rapidly overwhelmed

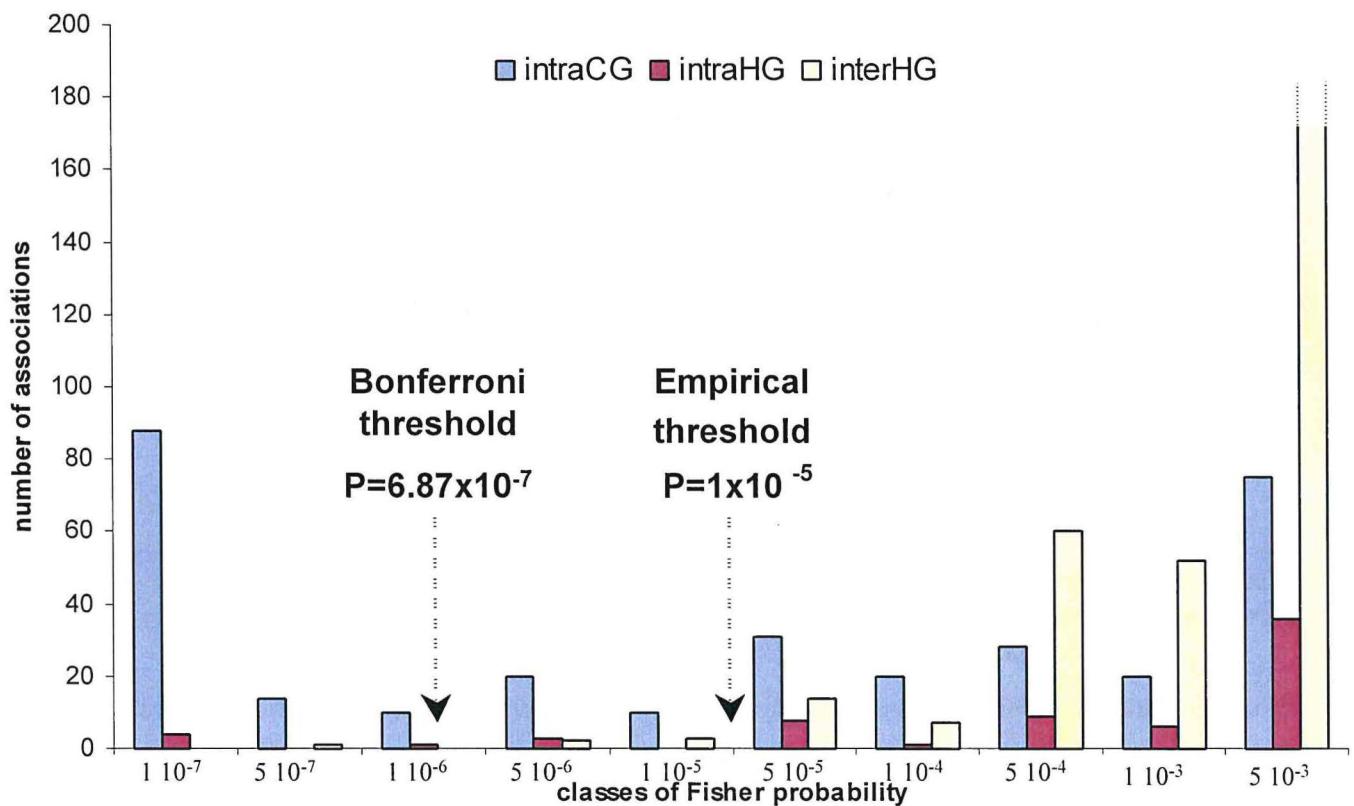


Figure 4: Distribution of pairwise associations of markers according to the class of Fisher probability to which they belong (Classes of Fisher probability are not shown for $P>10^{-3}$). Pairwise associations of markers are divided in three categories: intra-CG, intra-HG and inter-HG (see material and methods for details).

by associations detected within the inter-HG group (Fig. 4). In order to keep the proportion of ‘wrong’ associations (i.e. associations detected within the inter-HG group) below the 5% limit, an empirical threshold of $P=1\times 10^{-5}$ was set. Under these conditions, 142 associations detected within the intra-CG group could be considered as genuine and 6 associations detected within the inter-HG group could be considered as false. Only 8 associations from the intra-HG group were detected using this threshold and given their low number they were excluded as it is not clear whether they are ‘true’ or ‘false’ associations. It should also be noted that no associations having a negative PHI coefficient value would have been detected at $P=1\times 10^{-5}$. When using the Bonferroni correction, only association tests that yielded a probability threshold inferior to $P=6.87\times 10^{-7}$ were considered as significant (Fig. 4). Under these conditions, 102 ‘true’ associations of markers were retained.

The statistical threshold set using the Bonferroni procedure appeared too selective (the risk to fail to detect ‘true’ associations was rather high). Nevertheless, this threshold should not be relaxed too much because of the numerous artifactual associations that could be retained. In our example, below $P=1\times 10^{-3}$ approximately half of the associations would putatively be artifactual.

The extent of Linkage Disequilibrium in sugarcane

All the 142 associations detected using the empirical threshold ($P=1\times 10^{-5}$) were highlighted on the framework of ‘R 570’ AFLP genetic map (Figure 5¹). These associations are distributed in clusters of associated markers scattered over the entire sugarcane genome. Some of those clusters span over large genetic distances (over 50 cM). These clusters of markers in Linkage Disequilibrium correspond to haplotype blocks segregating in the population of sugarcane cultivars. Although the map coverage with markers scored in our LD study remains low (we could score 479 markers out of 887 positioned on the map), the results clearly indicated that LD occurs over a large range. The pattern of LD along the map of R 570 appeared heterogeneous though, which probably reveals the distinct history of the different cosegregation groups or chromosomes.

Among the 142 associations between pairs of markers, 55 involved two markers of *S. officinarum* origin, 27 involved two markers of *S. spontaneum* origin, and 20 involved markers of different origin. If we consider the distinct haplotype blocks (i.e. genomic regions

¹ Figure présentée en vis-à-vis des pages 70 et 71

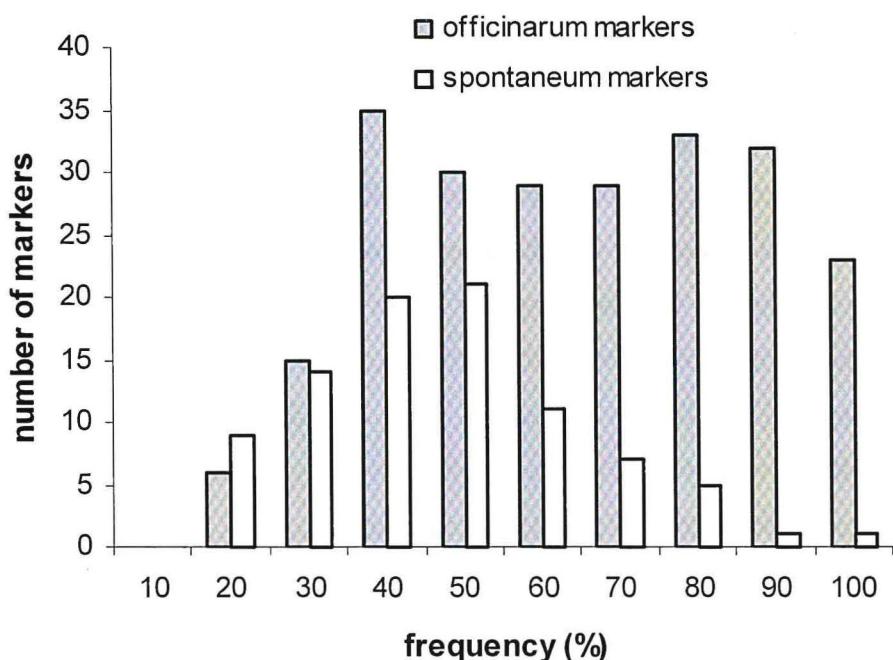


Figure 6: Distribution of the frequency of the markers according to their putative origin

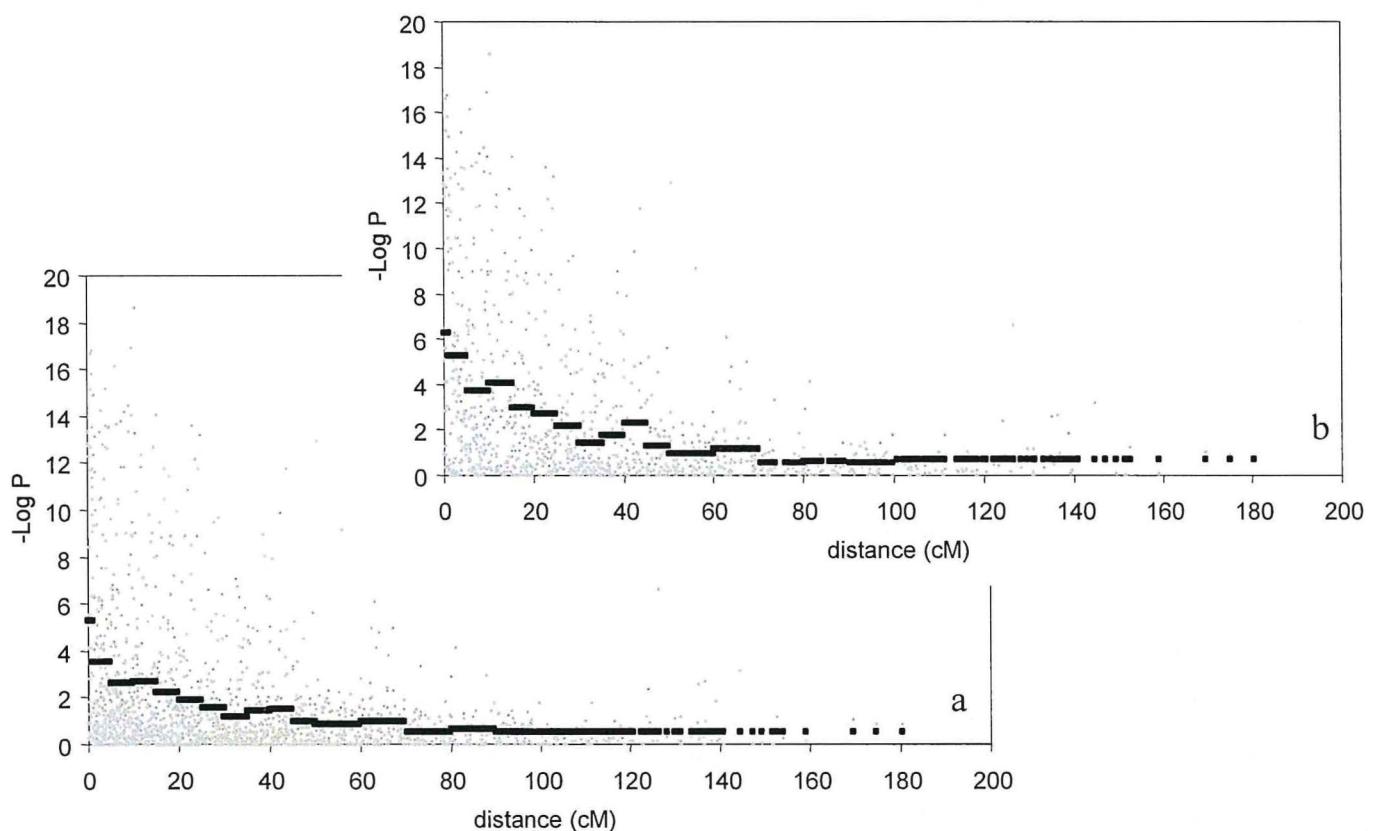


Figure 7: LD as a function of genetic distance in centimorgan (cM). LD was measured as the Logarythm transformation of the Fisher exact test Pvalue (1745 pairwise combinations of markers are plotted). Figure 7b was constructed using only markers whose frequency was >0.1 and <0.85 . The bold lines correspond to average LD values for marker pairs in 5 cM bins (5-50cM), 10 cM bins (50-100) and then over the totality of the remaining 100-200 cM. In the first 5 cM bin, the average was calculated within the first cM, and then the next four cM.

encompassing groups of markers significantly associated with one another), 31 had a homogeneous *S. officinarum* constitution, 8 had a homogeneous *S. spontaneum* constitution and 9 were recombinant. It is noteworthy that those chromosomal regions recombined between the ancestral species correspond to regions revealing extensive LD. Markers of *S. officinarum* origin have a wider frequency range in the 74 cultivars than markers of *S. spontaneum* origin (Fig. 6). Many of the *S. officinarum* markers are present at high frequency and therefore become statistically invisible as far as LD is concerned.

LD decay plots were constructed with all the associations between pairs of markers genetically linked on either R 570 map or MQ 76/53 map (1745 pairwise combinations in total). The Logarithm of the Fisher probability was used as a measure of LD and plotted as a function of the genetic distance between markers (Fig. 7a and b). Figure 7b was constructed using only markers whose frequency was >0.1 and <0.85. So doing, much of the background noise due to low-probability associations between genetically linked markers was removed. Both graphs revealed that LD extends over distances of up to 30 cM but decays sharply along the first five cM.

Linkage disequilibrium among all the AFLP markers produced

A total of 1 321 125 independent 2x2 Fisher exact tests, corresponding to all the possible pairwise comparisons between the 1626 polymorphic markers, were performed. To achieve an overall significant threshold of 5%, a nominal significance threshold of $P=3.78 \times 10^{-8}$ must be applied to each test according to the Bonferroni procedure. Under this conservative threshold, 297 associations encompassing 293 distinct markers, were considered as significant. The distribution of the frequency of the markers involved in Linkage disequilibrium is compared to the distribution of the frequency of all the polymorphic markers (Fig. 8). A vast majority of markers involved in LD were present in more than 10 percent of the clones and in less than 75% of the clones. The distribution of the frequency of markers involved in LD is centered around 35 percent, a figure which is similar to that observed with RFLP markers involved in LD (Jannoo et al 1999a). It is noteworthy that 208 significant associations out of the total of 297 (70%) significant associations between markers corresponded to combinations of markers present in ‘R 570’ (165/297) or combinations of markers present in ‘MQ 76/53’ (43/297). Only 64 significant associations (21%) corresponded to combinations of markers absent from

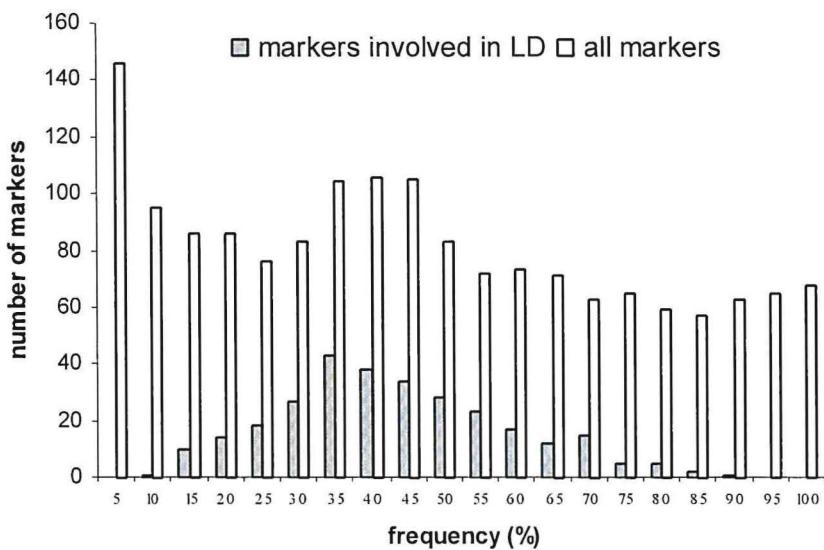


Figure 8: Distribution of markers according to their frequency among the 74 sugarcane genotypes (grey bars for markers involved in Linkage disequilibrium at $P = 3.78 \times 10^{-6}$ and white bars for the full set of markers).

Table 2: Number and size (number of constituting markers) of the putative haplotypes identified at a given threshold.

P	Nb of Markers	Nb of haplotypes	Number of haplotypes per class of haplotype size															
			2	3	4	5	6	7	8	9	10	11	12	13	15	16	
10^{-4}	810	143	81	27	12	6	8	1	3	3	0	0	0	1	0	0	0 1
10^{-5}	538	151	86	34	6	9	3	5	0	2	1	0	1	0	1	2	1 0
10^{-6}	396	128	77	23	10	6	2	3	2	2	1	1	0	0	1	0	0 0
10^{-7}	313	106	61	22	8	7	3	2	2	0	1	0	0	0	0	0	0 0
^b 3.78×10^{-8}	293	102	61	19	11	5	2	1	2	0	1	0	0	0	0	0	0 0

^aHaplotype size = number of markers, associated by transitivity, in the same haplotype

^b Threshold according to the bonferroni procedure

both control cultivars. The remaining 23 significant associations corresponded to combinations of markers between a marker absent from both control cultivars and markers of either ‘R 570’ or ‘MQ76/53’ cultivar. This result is parallel to the observation that markers scored in ‘R 570’ and ‘MQ 76/53’ encompassed 77% of the total number of polymorphic markers.

At $P=3.78 \times 10^{-8}$, 101 haplotypes could be constructed by grouping all significant pairwise combinations of markers by transitivity. Those haplotypes were composed of from two markers (61/101) up to 10 markers (Table 2). When the threshold for significance was relaxed, the number and the size of the constructed haplotypes increased. But when the threshold was too permissive, many markers were gathered together in a large haplotype containing markers that should not be associated (markers of different HGs for example). Thus, the threshold should rather be set around $P=1 \times 10^{-6}$.

IV.4 Discussion

Our results confirmed that linkage disequilibrium is extensive among sugarcane cultivars. Significant LD has been detected between AFLP markers up to 30 centimorgans apart. Nevertheless, an important drop in LD intensity was observed for distances over 5 cM. Unlike Jannoo et al. (1999a) LD was observed mainly between markers of putative *S. officinarum* origin although *S. spontaneum* markers involved in LD were over-represented if we consider that only 20% of the genome of modern sugarcane cultivars is inherited from this species.

A major drawback in the LD structure observed in sugarcane could be its “heterogeneity”. The strong linkage disequilibrium revealed among modern sugarcane cultivars may have been created by a two steps foundation effect. First of all, a few interspecific hybridizations involving a limited number of founding *S. officinarum* cultivars (also called noble clones) and *S. spontaneum* clones constituted a first genetic bottleneck. Secondarily, a few superior interspecific hybrids have been used extensively in all breeding stations around the world. Indeed, the analysis of the genealogy (Fig. 9) of the subpopulation of 30 clones, for which genealogical records were available, revealed a few preferential genealogical pathways through what we could call secondary genetic bottlenecks. ‘POJ 2878’, a well known example for sugarcane specialists, is found 5 times as parent/10 times as grandparent/8 times as great

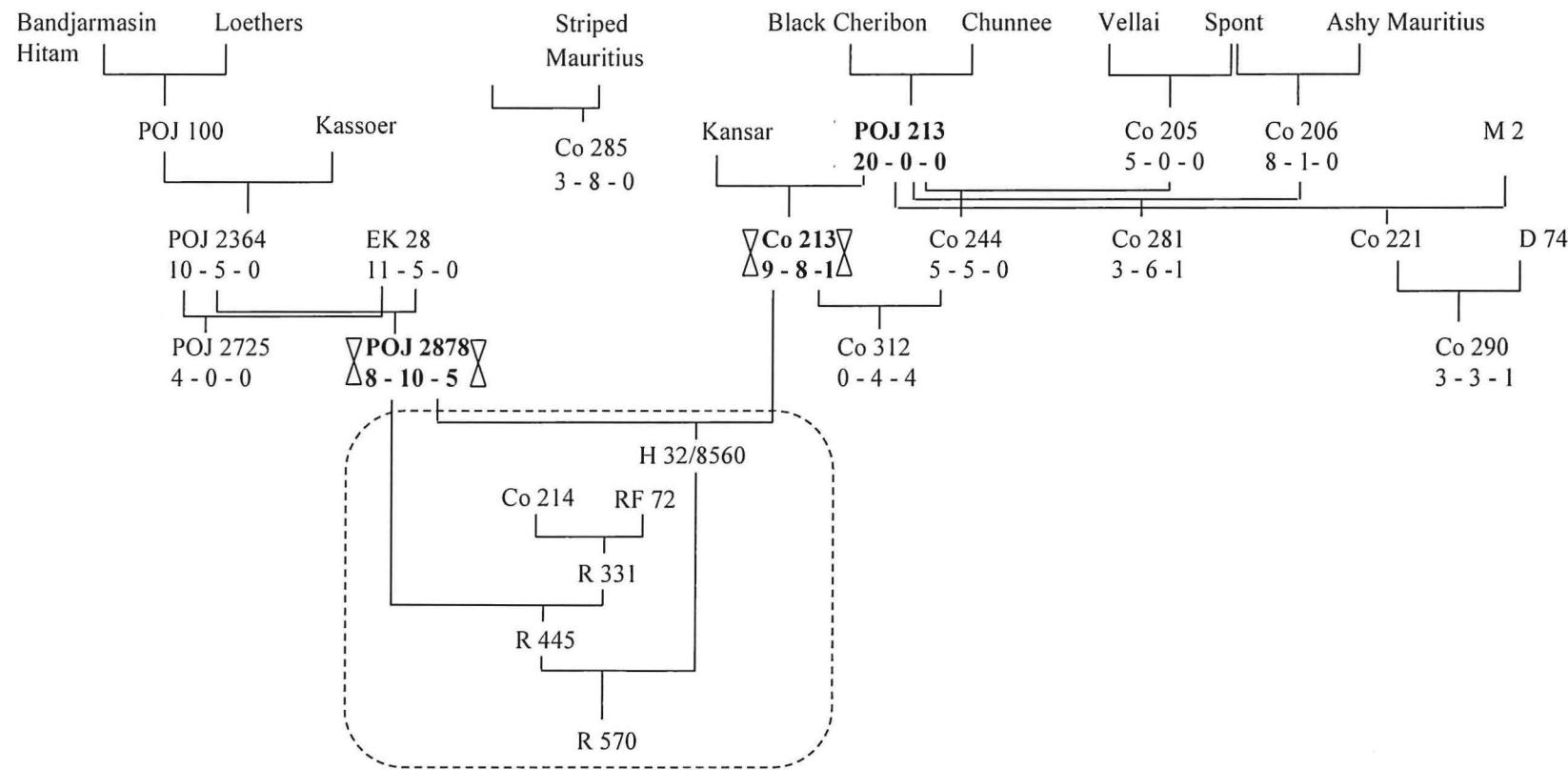


Figure 9: Description of the genealogy of 30 cultivars (out of 74 involved in this study). The main genealogical pathways (cultivars most frequently found as ancestor of the studied set of cultivars) are indicated with a funnel-like symbol. For the most frequent ancestors, we indicated below the cultivar name, the number of occurrence of the given cultivar as a great grandparent/grandparent/parent of the subset of 30 clones.

grandparent in the genealogy of those cultivars. Another major pathway is revealed through cultivars ‘POJ 213’ and ‘Co 213’. Cultivar ‘R 570’ is connected 3 times to those pathways (‘POJ 2878’ is found two times as a grandparent and ‘Co 213’ is found one time as a grandparent of ‘R 570’). Thus, we hypothesize that most of the haplotype blocks visualized on ‘R 570’ map should correspond to haplotypes inherited from either ‘Co 213’ or ‘POJ 2878’. When the total set of AFLP markers was considered, only around 400 markers (out of 1626) were statistically involved in haplotypes at $P=1\times 10^{-6}$. More than half of those haplotypes were constituted of markers present in ‘R 570’. Most probably, markers of the “secondary bottleneck ancestors”, some of them identified above, were over-represented in this subset of markers. This assumption should be checked using the most frequent ancestors as controls in future experiments focussed on LD analysis after a detailed analysis of the declared pedigree of the studied cultivars. Haplotypes present at too low frequency in the sugarcane germplasm will be difficult to detect. This difficulty might be partly circumvented by working on larger populations of cultivars and also by carefully choosing the cultivars used in an LD-based study. Different subpopulations of cultivars should be designed, based on the available genealogical records, in order to fully exploit all the genetic base of modern sugarcane germplasm. The main limitation here is that the accuracy of these genealogical records is sometimes questionable. On this subject, the pedigree of new elite cultivars extensively used as genitors in the breeding programs should be systematically checked now that routine genotyping methods are available. Another point is that most of the cultivars used in this work were bred thirty or more years ago. Current elite cultivars involved in the recurrent breeding schemes throughout the world are a few generations ahead and have accumulated more recombination events. Therefore, there is an opportunity in sugarcane to fit the population to the mapping objectives. If a better mapping resolution is required, more advanced cultivars should be chosen.

A further complication is that in a polyploid like sugarcane, linked markers may have very different frequencies in the population of cultivars. Markers at low frequency belonging to distinct haplotypes may appear linked to a same high frequency marker (Butterfield 2005). In that respect, markers of *S. officinarum* origin may cause the most problems. Indeed, diversity analysis has revealed a rather high genetic similarity (0.56) among *S. officinarum* clones involved in the genealogy of modern sugarcane cultivars (Jannoo et al. 1999b). Thus, markers at high frequency in the *S. officinarum* “genealogic germplasm” involved in distinct haplotype

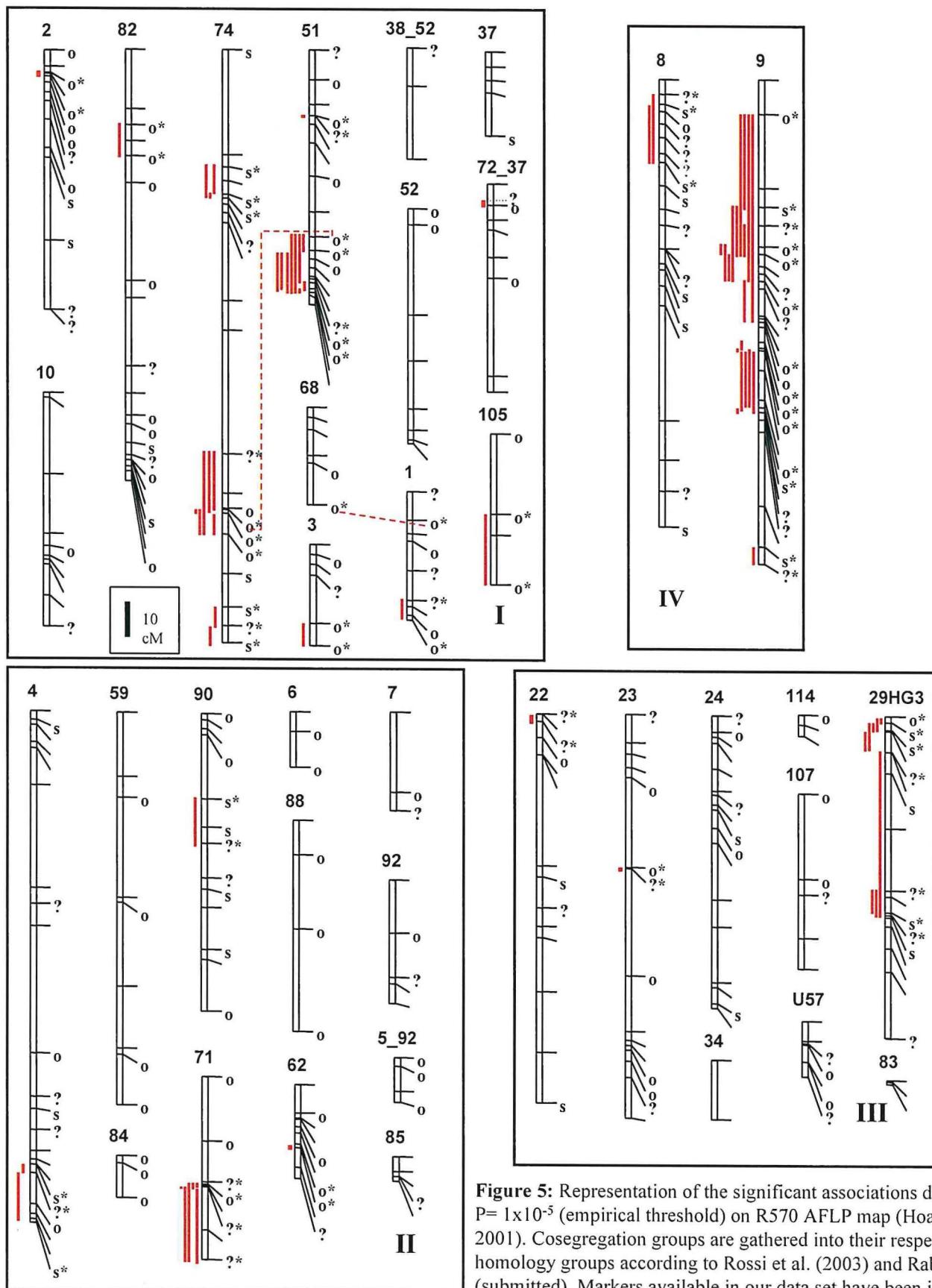


Figure 5: Representation of the significant associations detected at $P = 1 \times 10^{-5}$ (empirical threshold) on R570 AFLP map (Hoarau et al. 2001). Cosegregation groups are gathered into their respective homology groups according to Rossi et al. (2003) and Raboin et al. (submitted). Markers available in our data set have been identified with a reference mark labelled with a letter code corresponding to their putative origin: o for *S. officinarum*, s for *S. spontaneum* and ? for unknown origin. When a marker of the published map was not available in our data set, its position was identified with an unlabelled reference mark. Markers involved in significant associations are also identified with *.

structures (i.e. distinct homo(eo)logous chromosome segments constituted of distinct sequences of markers/alleles) could have been transmitted into the modern sugarcane germplasm. This kind of situation should be less frequent with markers of *S. spontaneum* origin because the genetic similarity among *S. spontaneum* clones is much lower (0.31 among *S. spontaneum* clones compared to 0.66 among *S. officinarum* clones) according to Lu et al. (1994b). Indeed, if a marker is common, it is likely to have been introduced by multiple founders and in that case the extent of LD depend on the LD structure in its population of origin (Kruglyak, 1999). The comparison of markers frequency (Figure 6) illustrates this situation where markers of *S. officinarum* origin have a much wider frequency range than markers of *S. spontaneum* origin.

Most of *S. officinarum* diversity is still found among modern sugarcane cultivars (Jannoo et al. 1999b). In contrast, only part of *S. spontaneum* high diversity may have been transmitted to sugarcane cultivars (Lu et al. 1994a). The genetic bottleneck due to the limited number of founding interspecific crosses from which modern sugarcane germplasm was derived lead to loss of diversity among *S. spontaneum* alleles but not among *S. officinarum*. This is illustrated, in our data set (Figure 1), by the absence of many haplotypes of clone ‘MQ 76/53’ from the population of cutivars we studied. These haplotypes have most probably been inherited from *S. spontaneum* ‘SES 528’, the father clone of ‘MQ 76/53’. ‘SES 528’ is almost absent from the genealogy of modern sugarcane (<http://tropgenedb.cirad.fr/>; Ruiz et al. 2004). The loss of *S. spontaneum* alleles may have gotten worse because of the rapid elimination of *S. spontaneum* chromosomes due to the 2n (somatic number of chromosomes) transmission of *S. officinarum* chromosomes occuring in the first interspecific crosses and in the first back cross with *S. officinarum* (Bhat and Gill 1985).

LD observed in this study was extensive but dropped sharply when markers were separated by 5 cM and more. The order of magnitude of LD in sugarcane resembles that described in cattle (Farnir et al. 2000). A situation that should make genome-wide LD mapping achievable in sugarcane provided that the difficulties expressed above are controlled. The minimum number of locus-specific-markers required to achieve a density of one or two markers every five cM would lie between 300 and 600 (the size of the haploid sugarcane genome is about 1500 cM). To date, only 260 microsatellites markers have been characterized in sugarcane. Although more SSRs should be provided by mining EST databases (Pinto et al. 2004), this is far from sufficient to achieve an even and dense enough coverage. AFLPs have a fair throughput

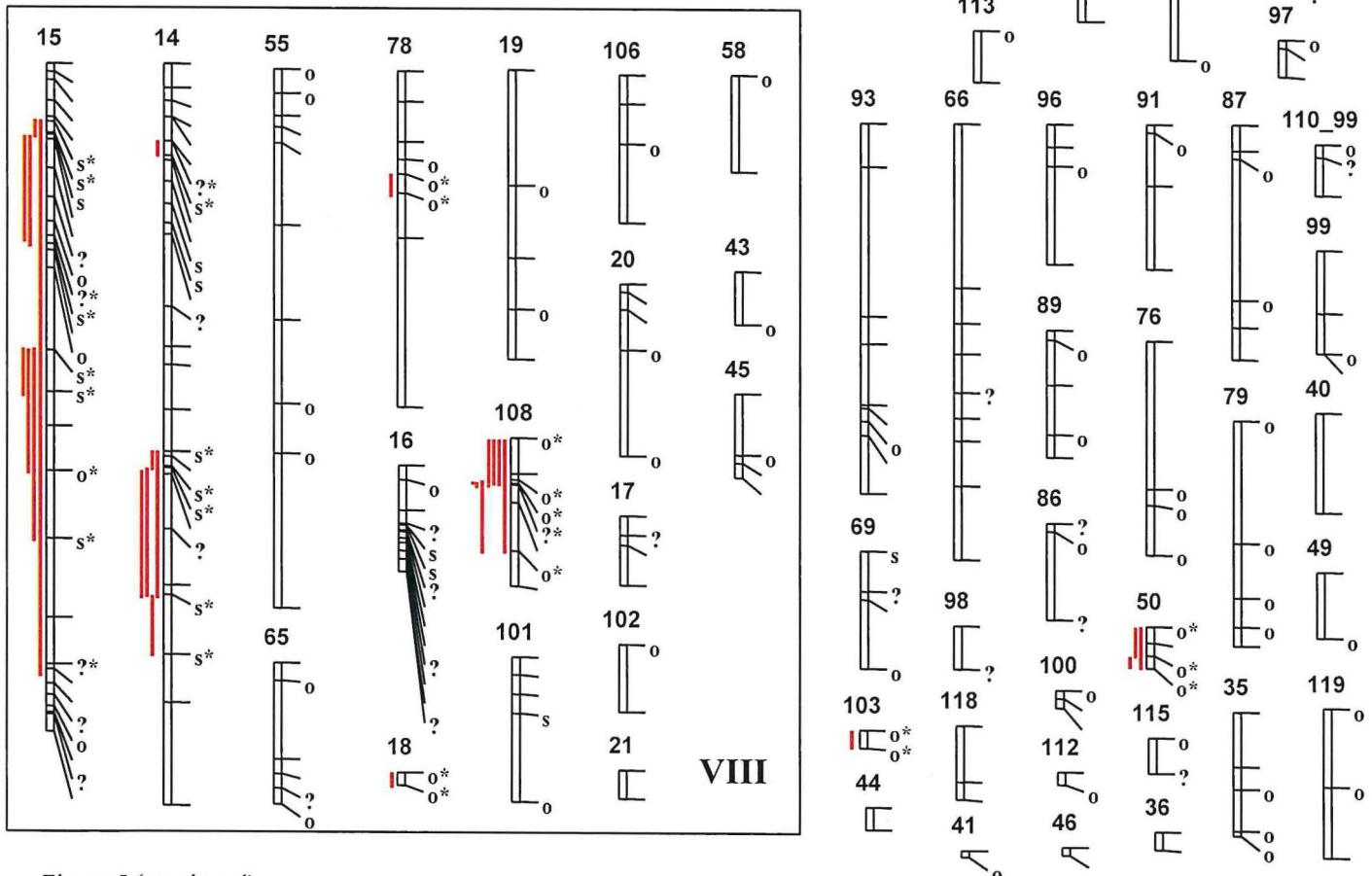
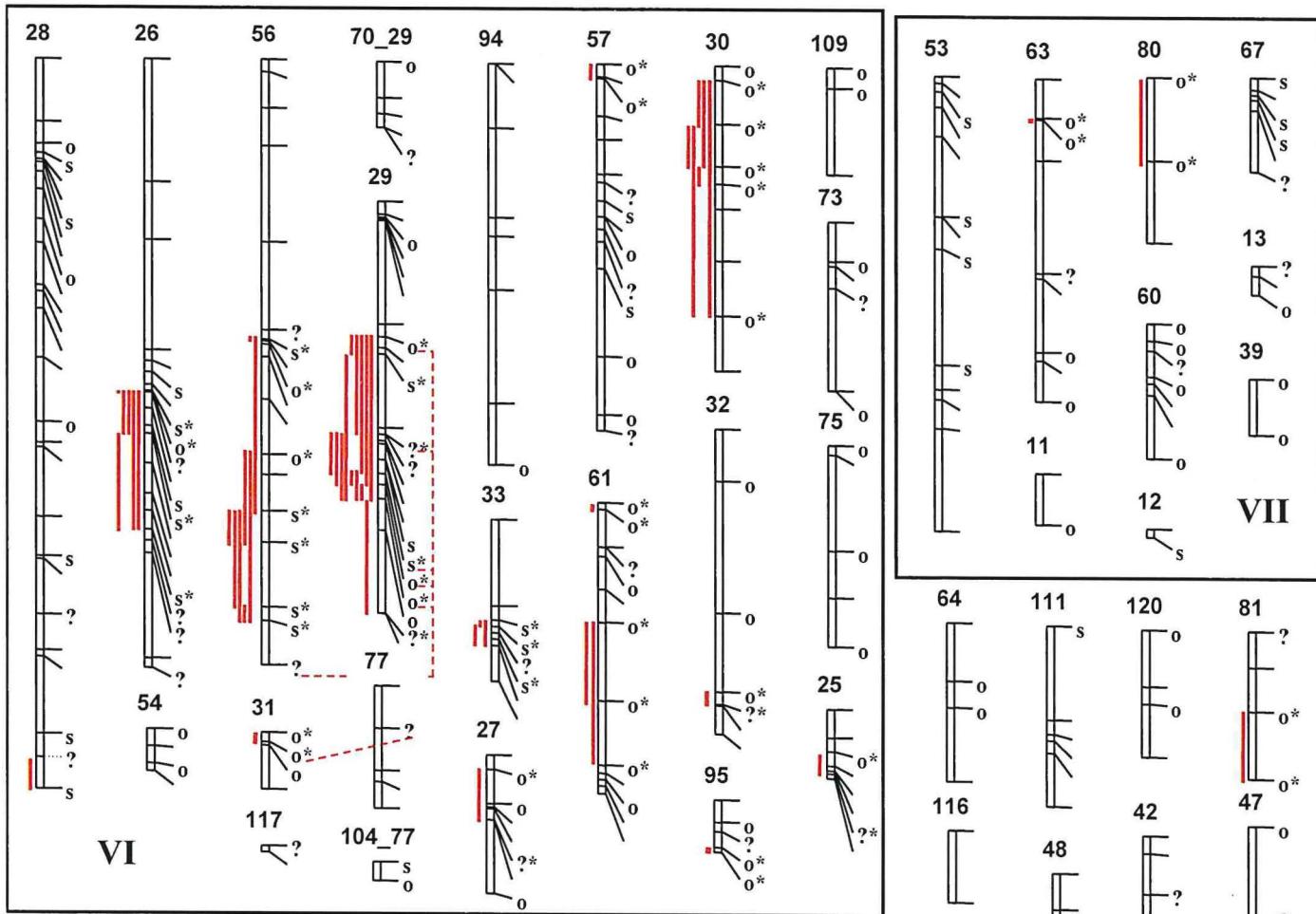


Figure 5 (continued)

potential, but they also are tedious to score. In this study, we used 42 primer pairs (out of the 64 commercially available) in order to produce 1626 markers, a figure that revealed clearly insufficient to apprehend the totality of the haplotypes diversity in the sugarcane germplasm. RFLPs are too time consuming to produce and tedious to score. Moreover, when it comes to the question of the capitalisation of data from distinct experiments, disconnected in time and space, none of these markers, comprising microsatellites, are the markers of choice. SNPs are very expensive to develop especially in the high polyploid sugarcane where the identification of useful polymorphic SNPs shall be more difficult (Grivet et al. 2003) although EST resources are available in sugarcane (<http://sucest.lad.ic.unicamp.br/en/>). The application of molecular genotyping to sugarcane breeding clearly requires the switch from analytical separation techniques to high throughput hybridization techniques based on the use of DNA arrays. Diversity Arrays Technology (DArT) appears particularly promising as it can detect polymorphism at several hundred loci simultaneously without relying on sequence informations (Wenzl et al. 2004). This technique may prove useful to address large and complex polyploid genomes like in sugarcane. Arrays could be specifically designed for applications based on LD for example targeting markers of the appropriate frequency among sugarcane cultivars (around 0.35 as suggested from RFLP data (Jannoo et al. 1999a) as well as from AFLP data presented in this study). This type of high throughput automated technology will open the way to the construction of large databases gathering informations at the whole genome scale on a growing number of genotypes. Thus, it should soon be possible to correlate these genotypes with phenotypes measured in a wide range of environments (under different pathogen pressure, different abiotic stresses....) with the objective to identify haplotypes conferring a superior agronomic value.

References

- Arceneaux, G. 1965. Cultivated sugarcanes of the world and their botanical derivation. Proc. Int. Soc. Sugar Cane Technol. 12: 844-854.
- Bhat, S.R. and B.S. Gill. 1985. The implications of $2n$ egg gametes in nobilization and breeding of sugarcane. Euphytica 34: 377-384.
- Butterfield M (2005) Molecular breeding of sugarcane using DNA markers and linkage disequilibrium mapping. . Proc. Int. Soc. Sugar Cane Technol , 25 :545-548.
- Cardon, L.R. and J.I. Bell. 2001. Association study designs for complex diseases. Nature Reviews Genetics 2: 91-99.
- D'Hont, A., L. Grivet, P. Feldmann, P.S. Rao, N. Berding, and J.C. Glaszmann. 1996. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. Mol. Gen. Genet. 250: 405-413.
- Farnir, F., W. Coppieters, J.-J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar, and M. Georges. 2000. Extensive Genome-wide Linkage Disequilibrium in Cattle. Genome Research 10: 220-227.
- Flint-Garcia, S.A., T.J. M., and E.S. Buckler IV. 2003. Structure of Linkage Disequilibrium in plants. Annu. Rev. Plant Biol. 54: 357-374.
- Grivet, L., A. D'Hont, D. Roques, P. Feldmann, C. Lanaud, and J.C. Glaszmann. 1996. RFLP mapping in cultivated sugarcane (*Saccharum* spp.): Genome organization in a highly polyploid and aneuploid interspecific hybrid. Genetics 142: 987-1000.
- Grivet, L. and P. Arruda. 2001. Sugarcane genomics: depicting the complex genome of an important tropical crop. Curr. Opin. Plant Biol. 5: 122-127.
- Grivet, L., Glaszmann J.C., Vincentz M.A., da Silva F., and Arruda P.(2003). ESTs as a source for sequence polymorphism discovery in sugarcane: example of the Adh genes. Theor. Appl. Genet. 106: 190-197.
- Hansen, M., T. Kraft, S. Ganestam, T. Säll, and N.-o. Nilsson. 2001. Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. Genet. Res. Camb. 77: 61-66.
- Hoarau, J.Y., B. Offmann, A. D'Hont, A.-M. Risterucci, D. Roques, J.C. Glaszmann, and L. Grivet. 2001. Genetic dissection of a modern cultivar (*Saccharum* spp.) I. genome mapping with AFLP markers. Theor. Appl. Genet. 103: 84-97.

- Jannoo, N., L. Grivet, A. Dookun, A. D'Hont, and J.C. Glaszmann. 1999a. Linkage disequilibrium among modern sugarcane cultivars. *Theor. Appl. Genet.* 99: 1053-1060.
- Jannoo, N., L. Grivet, M. Seguin, F. Paulet, R. Domaingue, P.S. Rao, A. Dookun, A. D'Hont, and J.C. Glaszmann. 1999b. Molecular investigation of the genetic base of sugarcane cultivars. *Theor. Appl. Genet.* 99: 171-184.
- Kraakman, A.T.W., R.E. Niks, P.M.M.M. Van den Berg, P. Stam, and F.A. Van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168: 435-446.
- Kruglyak, L. 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics* 22: 139-144.
- Lander, E.S. and N.J. Schork. 1994. Genetic dissection of complex traits. *Science* 265: 2037-2048.
- Lima, M.L.A., A.A.F. Garcia, K.M. Oliveira, S. Matsuoka, H. Arizono, C.L. de Souza, and A.P. de Souza. 2002. Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugarcane (*Saccharum* spp.). *Theor. Appl. Genet.* 104: 30-38.
- Lu, Y.H., A. D'Hont, F. Paulet, L. Grivet, M. Arnaud, and J.C. Glaszmann. 1994. Molecular diversity and genome structure in modern sugarcane varieties. *Euphytica* 78: 217-226.
- Lu, Y.H., A. D'Hont, D.I.T. Walker, P.S. Rao, P. Feldmann, and J.C. Glaszmann. 1994. Relationships among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. *Euphytica* 78: 7-18.
- Mehta CR, Patel NR (1983) A network algorithm for performing Fisher's exact test in rxc contingency tables. *J Am Statist Assoc* 78:427-434
- Morgante, M. and F. Salamini. 2003. From plant genomics to breeding practice. *Curr. Opin. Biotech.* 14: 214-219.
- Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76 : 427-434
- Nordborg, M., J.O. Borevitz, J. Bergelson, C.C. Berry, J. Chory, J. Hagenblad, M. Kreitman, J.N. Maloof, T. Noyes, P.J. Oefner, E.A. Stahl, and D. Weigel. 2002. The extent of linkage disequilibrium in *Arabidopsis Thaliana*. *Nature Genetics* 30: 190-193.
- Nordborg, M. and S. Tavaré. 2002. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 18: 83-90.

Olsen, K.M., S.S. Halldorsdottir, J.R. Stinchcombe, C. Weinig, J. Schmitt, and M.D. Purugganan. 2004. Linkage disequilibrium mapping of *Arabidopsis* CRY2 flowering time alleles. *Genetics* 167: 1361-1369.

Perrier, X., Flori, A. , Bonnot, F. (2003). Data analysis methods. In : Hamon, P., Seguin, M., Perrier, X. ,Glaszmann, J. C. Ed., Genetic diversity of cultivated tropical plants. Enfield, Science Publishers. Montpellier. pp 43 - 76.

Pinto, L.R., K.M. Oliveira, E.C. Ulian, A.A.F. Garcia, and A.P. de Souza. 2004. survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome* 47: 795-804.

Raboin L-M, Oliveira K-M, Lecunff L, Telismart H, Roques D, Butterfield M, Hoarau J-Y, D'Hont A (submitted to TAG). Genetic mapping in the high polyploid sugarcane using a biparental progeny; identification of a gene controlling stalk colour and a new rust resistance gene.

Rafalski, A. and M. Morgante. 2004. Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* 20: 103-111.

Rossi, M., P.G. Araujo, F. Paulet, O. Garsmeur, V.M. Dias, H. Chen, V.S. M.-A, and A. D'Hont. 2003. Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. *Mol. Gen. Genet.* 269: 406-419.

Ruiz M, Rouard M, Raboin LM, Lartaud M, Lagoda P, Courtois B (2004) Tropgene-DB, a multitropical crop information system. *Nucleic Acids Research* 32:D364-D367

SAS Institute Inc. (1996) SAS/STAT® Software: Changes and enhancements through release 6.11, Cary , NC:SAS Institute Inc. 1104 pp.

SAS Institute (1990)SASprocedures guide, version 6, third edition. SAS Institute Inc, Cary, N.C.

Simko, I. (2004). "One potato, two potato: haplotype association mapping in autotetraploids." *Trends Plant Sci.* 9(9): 441-448.

Tenaillon, M.I., M.C. Sawkins, A.D. Long, R.L. Gaut, J.F. Doebley, and B.S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci.* 98: 9161-9166.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wenzl, P., J. Carling, D. Kudrna, D. Jaccoud, E. Huttner, A. Kleinhofs, and A. kilian. 2004. Diversity arrays technology (DArT) for whole-genome profiling of barley. Proc. Natl. Acad. Sci. 101: 9915-9920.

V MARQUAGE DE LA RESISTANCE AU CHARBON

V Cartographie génétique de la résistance de la canne à sucre au charbon dans une descendance bi-parentale et par association dans une population de cultivars.

Finalement, les travaux de génotypage réalisés d'une part pour construire les cartes génétiques de ‘R 570’ et de ‘MQ 76/53’ (voir chapitre III) et d'autre part pour caractériser le déséquilibre de liaison dans la population de 74 cultivars de canne à sucre (voir chapitre IV) ont été mis à profit pour tenter de marquer les zones du génome impliquées dans la résistance au charbon. La descendance de 198 clones issue du croisement ‘R 570’ et ‘MQ 76/53’ a été évaluée pour son comportement vis à vis du charbon dans différents essais en serre et au champ. Une détection de QTL a donc pu être réalisée. Par ailleurs, la population de 74 cultivars de canne à sucre a été constituée à partir deux sous ensembles de cultivars, l'un extrêmement résistant et l'autre extrêmement sensible au charbon. Ces sous ensembles de clones ont été sélectionnés sur la base des résultats de nombreux essais de sélection menés au Burkina Faso en conditions d'infestation naturelle forte par le charbon. Une analyse des associations entre marqueurs et résistance au charbon a ainsi pu être réalisée.

Ce travail est présenté sous la forme d'un projet d'article intitulé « **Genetic mapping of sugarcane resistance to smut through bi-parental segregation and associations among modern cultivars** »

Genetic mapping of sugarcane resistance to smut through bi-parental segregation and associations among modern cultivars

Louis-Marie Raboin, Nibouche Samuel, Pauquet Jérôme, Telismart Hugues, Jacques Dintinger, Jean-Yves Hoarau, Laurent Costet, Angélique D'Hont

V.1 Introduction

Smut caused by *Ustilago scitaminea* Syd. is one of the most important diseases of sugarcane that has a world-wide geographical distribution (Comstock 2000). Only Papua New Guinea, Fiji and East Australia remain free of the disease. After infection of the buds by the germinating teliospores, the fungus invades the meristematic region. Then, the systemic infection occurs along with the colonisation of each bud primordium produced by the growing stalk. Eventually, the apical meristem function is modified and the apex produces a typical whip-like sorus bearing billions of teliospores (Ferreira and Comstock 1989). Smut is a major concern for breeders as it can cause severe losses on susceptible cultivars (Antoine (1961) reported yield reductions over 70% and more recently Hoy et al. (1985) demonstrated that cane yield is negatively correlated with smut incidence and estimated losses up to 50%). Quantitative genetic studies indicated that broad sense heritability was high and narrow sense heritability was moderate for resistance to smut (Wu et al. 1983 and 1988; Chao et al. 1990). That's why breeding and selecting for smut resistance revealed efficient so far. Nevertheless, little is known about the complexity (i.e. polygenic or oligogenic) of sugarcane resistance to smut. Reports of the possible existence of race-cultivar specialisation in Hawaii (Comstock and Heinz 1977), Brazil (Toffano 1976), Pakistan (Muhammed and Kausar 1962) and Taiwan (Leu and Teng 1972) even raise the question of the possible involvement of major resistance genes. Those cases need further clarification and moreover, quantitative trait loci (QTL) as well, may be responsible for race-cultivar specialisation (Leonards-Schippers et al. 1994).

QTL mapping is a powerful tool to dissect complex disease resistance into simple mendelian components (Michelmore 1995; Young 1996). In sugarcane, the challenge is to achieve a good coverage of the genome suitable for a genome scan QTL approach. Sugarcane is characterized by its high polyploidy (typical cultivars have between 100 and 130 chromosomes) as it is derived from interspecific hybridization between *S. spontaneum* ($x=8$, $2n=40-128$) and *S. officinarum* ($x=10$,

$2n=80$) (D'Hont et al. 1996). This feature dictates particular constraints for genetic mapping (Grivet and Arruda 2001). The effects of detected QTLs are often small mainly because of a complex pattern of dominant and epistatic interactions at and between loci presenting up to 12 distinct alleles. Thus, very few attempts have been made to unravel the molecular basis of disease resistance in sugarcane. The only major resistance gene mapped in sugarcane is the resistance gene to brown rust *brul* identified in cultivar 'R 570' (Daugrois et al. 1996). This gene is at present the focus of a map-based cloning project (D'Hont et al. 2001; Asnaghi et al. 2004). Recently a major QTL (explaining over 20% of the phenotypic variation) for resistance to yellow spot, a disease caused by the fungus *Mycovellossiella koepkei*, was identified in a Mauritian cultivar 'M 134/75' (Al-Janabi et al. 2005).

An alternative approach to QTL mapping, based on the analysis of segregating progeny, for the identification of mendelian factors underlying disease resistance is association mapping, based on the exploitation of Linkage Disequilibrium (LD) in (preferably) unstructured populations (Neale and Savolainen, 2004). When LD extends over large distances on chromosomes, a whole-genome scan association study is feasible. Conversely, when LD decays rapidly, a candidate-gene association study is more appropriate (Flint-Garcia 2003). These strategies are only starting to be applied in plants. A few examples are available from litterature on studies aimed at identifying loci involved in disease resistance. The tomato Ve1 gene, a candidate gene for potato resistance to *Verticillium Dahliae*, and potato homologues to this gene were used with success to identify a resistance gene in potato using LD mapping (Simko 2004). In rice, LD has been characterized in a targeted region (70 kb) where the recessive gene xa5, conferring race-specific resistance to bacterial blight, has been mapped (Garris et al. 2003). This analysis revealed genetic heterogeneity. Distinct haplotypes and loci may confer the resistant phenotype. A genome-wide association study in barley, revealed effective to identify SSR markers associated to powdery mildew resistance (Ivandic et al. 2003).

Here we report on the implementation of both approaches, QTL mapping in a bi-parental progeny and association mapping in a population of cultivars, in an exploratory tentative to identify loci and haplotypes involved in sugarcane resistance to smut. LD is extensive among sugarcane cultivars. Significant LD was detected between markers up to 30 centimorgans apart (Jannoo et al. 1999) although it dropped sharply between markers separated by 5 cM and more (Raboin et al. in prep¹).

¹ Chapitre IV de la thèse

Table 1: list of cultivars of the LD mapping population. The smut rating has been obtained from trials performed in Burkina Faso. The set of highly susceptible cultivars was supplemented with cultivars identified as extremely susceptible in Reunion island. These clones are identified with the symbol *.

subset of resistant clones						subset of susceptible clones					
clone	origin	rating ¹	clone	origin	rating ¹	clone	origin	rating ¹	clone	origin	rating ¹
B 80/66	Barbados	1	Co 1157	India	2	F 160	Taiwan	5*	SP 71/6113	Brazil	8
B 85/356	Barbados	1	Co 449	India	2	R 574	Reunion	5*	FR 84/387	FWI	9
BJ 78/128	Barbados	1	CP 61/37	USA	2	B 63/119	Barbados	7	H 32/8560	Hawaii	9*
Co 462	India	1	CP 66/315	USA	2	B 66/23	Barbados	7	H 39/3635	Hawaii	9*
D 172	Guyana	1	CP 70/1133	USA	2	B 75/524	Barbados	7	H 39/7028	Hawaii	9*
FR 84/166	FWI	1	DB 80/104	Barbados	2	B 86/409	Barbados	7	H 44/3098	Hawaii	9*
M 202/46	Mauritius	1	FR 84/344	FWI	2	BR 79/4	Barbados	7	H 49/5	Hawaii	9*
SP 70/1284	Brazil	1	IAC 64/257	Brazil	2	Co 1177	India	7	H 50/7209	Hawaii	9*
B 47/258	Barbados	2	J 59/3	Jamaica	2	Co 842	India	7	H 61/1721	Hawaii	9*
B 51/129	Barbados	2	Ja 64/19	Cuba	2	DB 73/419	Barbados	7	LF 53/4789	Fiji	9*
B 80/8	Barbados	2	Mex 68/P23	Mexico	2	N 12	South Africa	7	LF 53/4825	Fiji	9*
B 82/288	Barbados	2	MY 53/53	Cuba	2	NA 56/62	Argentina	7	MQ 76/53	Australia	9*
B 86/839	Barbados	2	MY 55/14	Cuba	2	NCo 310	India	7	PR 61/632	Puerto Rico	9
B 86/406	Barbados	2	N 17	South Africa	2	NCo 376	India	7	Q 84	Australia	9
B 87/1172	Barbados	2	R 526	Reunion	2	Phil 56/226	Phillipines	7	RB 70/96	Brazil	9
BR 71/48	Barbados	2	SP 70/1005	Brazil	2	Phil 66/7	Phillipines	7			
BR 75/48	Barbados	2	SP 70/1423	Brazil	2	RB 72/5828	Brazil	7			
BT 72/344	Barbados	2	SP 70/3225	Brazil	2	B 77/84	Barbados	8			
C 227/59	Cuba	2	R 570	Reunion	3	Co 1186	India	8			
CB 56/171	Brazil	2				Co 1208	India	8			

¹ rating based on a 1(the most resistant) to 9(the most susceptible) scale established according to a set of standards (see Material and Methods).

This order of magnitude makes genome-wide LD mapping achievable in sugarcane. The population used by Raboin et al. (in prep.) to characterize LD among sugarcane cultivars has been extensively genotyped with AFLP and involved cultivars well characterized for smut resistance. Thus, in the present study, we exploited this population to search for associations between resistance to smut and AFLP markers on a whole-genome scale. In addition, for the purpose of QTL detection, we characterized a progeny, from a cross between a resistant cultivar ‘R 570’ and a sugarcane clone highly susceptible to smut ‘MQ 76/53’, for resistance to smut through various trials in the field or under greenhouse. The genetic map derived from this progeny has been described by Raboin et al. (submitted²). The results obtained with these two approaches are compared.

V.2 Material and Methods:

Plant material

The progeny used for the QTL study consisted of 198 clones derived from a cross between ‘R 570’ (used as a female after emasculation by hot water treatment) and ‘MQ 76/53’. ‘R 570’ (H 32/8560 x R445) is a modern cultivar developed by CERF (Centre d’Essai de Recherche et de Formation, Réunion) and is considered as resistant to smut. ‘MQ 76/53’ (Trojan x SES 528) is an Australian clone and is highly susceptible to smut.

The population used in the association study consisted of 74 cultivars released from various breeding programs around the world (Table 1). For most of these cultivars, rating for smut resistance was established in trials conducted in Burkina Faso (SOSUCO, Société sucrière de la Comoé, Burkina Faso)) using randomized block designs of 6 replicates (one plot consisted of a 10 m row). Evaluations were performed under natural pressure of the pathogen. The rating was established in second ratoon according to the same set of standards (‘Co 449’ rated 2, ‘B 46/364’ rated 5, ‘Nco 376’ rated 7 and ‘Nco 310’ rated 8). Data of 17 distinct trials (carried out since 1986), provided by SOSUCO, involving a total of 254 cultivars were analysed. In order to constitute our association mapping population, we only selected the 39 most resistant cultivars (more or equally resistant than ‘Co 449’) and the 23 most susceptible cultivars (more or equally susceptible than ‘Nco 376’). The subset of highly susceptible cultivars had to be supplemented with cultivars unambiguously rated as highly susceptible in Reunion island. They were granted a rating of 9 (‘MQ

² chapitre III de la thèse

76/53' was part of them). 'R 570' rated 3 in Burkina Faso and two moderately susceptible clones in Reunion 'F 160' and 'R 574' granted a rating of 5 were also part of the studied population. We are confident that the smut rating of the supplemented clones from Reunion should be homogenous with the smut rating they would obtain in Burkina Faso because smut isolates from Reunion island are genetically very close to smut isolates from Burkina Faso. They both belong to a worldwide lineage according to a microsatellite based diversity study (Raboin et al. submitted³). Moreover, Péros and Baudin (1983) observed no differences in pathogenicity and no race-cultivar specialisation among isolates from Réunion, and Burkina Faso.

Inoculum used in the greenhouse and field trials

As mentionned above, a diversity study based on microsatellites (Raboin et al., submitted) demonstrated the near absence of variability among isolates collected in different regions of Reunion island and we have no indications of pathogenicity variation in Reunion island. Therefore spores used for artificial inoculations were mixed from whips collected on different cultivars in different plots of Reunion island. Before each inoculation, the viability of spores (stored in glass vacuum bell jars with silicagel) was checked on agarose in Petri dishes (over 80% germination).

Comparison of different inoculation methods under greenhouse:

We conducted two preliminary trials under greenhouse in order to compare three inoculation methods. Single-bud cuttings were inoculated either by dipping in a spore suspension (5×10^6 spores per ml), puncturing the bud with a needle previously covered with smut spores (first the needle is dipped in sterile water and then dipped in a tube containing spores) or by injecting a spore suspension (5×10^6 spores per ml) in germinating shoot of 5 to 10 cm (review of inoculation techniques by Lee-Lovick 1978). The behaviour of two susceptible clones 'MQ 76/53' and 'B 34/104' was compared to that of resistant clone 'R 570'. Inoculated cuttings (dipping and puncturing method) and un-inoculated one-bud cuttings (controls and injection method) were incubated during 24 hours under a humid floorcloth and then transplanted in small pots containing a mixture of half soil and half pouzzolane. Each week, the newly diseased plants (i.e. plants presenting newly emerged whips) were counted and removed from the greenhouse. The analysed variables were the maximum percentage of plants having developed a whip and the Area Under

³ chapitre II de la thèse

Disease Progress Curve (AUDPC) of the cumulated number of whips (i.e. integration over time of the cumulated number of whips).

Greenhouse evaluation of the progeny population

We conducted two trials in early may 2003 and mid april 2004 at Ligne-Paradis CIRAD experimental station in south Reunion island (20°S). The experimental design was a randomized block design with four replicates in the first trial and three replicates in the second trial. Plots consisted of 12 small pots each containing one single-bud inoculated cutting. Inoculation was performed by puncturing the buds with a needle previously covered with spores. Inoculated cuttings were left for germination in trays under a humid floorcloth. After one week, only growing buds (most of them with roots or roots primordiums) were transferred into small pots containing a mixture of half pouzzolane and half soil. The emergence of whips was assessed every week over a period of 9 months.

Field evaluation of the progeny population

We planted a first trial in September 2000 at Ligne-Paradis CIRAD experimental station and a second one in september 2001 at Le Gol CERF experimental station in Reunion island (20°S). For both trials the experimental design was a 20x10 alpha-lattice design (Patterson and Williams, 1976) with three replications (consisting of 20 blocks of 10 basic plots). Each plot consisted of 2.5 m row planted with 8 inoculated cuttings. Along with the progeny, the two parental clones ('R570' and 'MQ 76/53') were also evaluated. Moreover, single stools of the highly susceptible clone 'MQ 76/53' were interspersed every two plots in order to facilitate the epidemic. Inoculation was performed by dipping three-bud cuttings for 20 minutes in a spore suspension of 5×10^6 spores per ml. Inoculated cuttings were incubated for 24 hours under a humid floorcloth before planting. Smut incidence was measured by counting newly emerged whips every 15 days. Old whips were marked with paint to avoid counting them twice. This assessment was performed in plant cane and two ratoons. Before harvesting, the number of millable stalks was counted over the entire plot.

Analysis of quantitative traits

In the field trials, analyses of variance were performed for each variable within each of the two trials separately using SAS Mixed procedure (SAS version 6.12, SAS Institute, In., NC, USA). The analysed variables, calculated on a plot basis, were the total number of whips counted over the

campaign and the final number of millable stalks. Block and genotype (clone) factors were considered as random, and replication as fixed. The following model was used:

$Y_{ijk} = \mu + G_i + Rep_j + (Block/Rep)_{jk} + e_{ijk}$, where Y_{ijk} is the variable measured on the i^{th} genotype in the k^{th} block of the j^{th} replication, μ is the general mean, G_i the effect of the i^{th} genotype, Rep_j the effect of the j^{th} replication, $(Block/Rep)_{jk}$ the effect of the k^{th} block of the j^{th} replication and e_{ijk} the residual effect.

In the greenhouse trials, analyses of variance were performed within each of the two trials separately using SAS Mixed procedure. The analysed variable calculated on a plot basis, was the AUDPC (integration over time of the cumulated number of whips). Genotype (clone) factor was considered as random and replication as fixed. The following model was used:

$Y_{ij} = \mu + G_i + Rep_j + e_{ij}$, where Y_{ij} is the variable measured on the i^{th} genotype in the j^{th} replication, μ is the general mean, G_i the effect of the i^{th} genotype, Rep_j the effect of the j^{th} replication and e_{ij} the residual effect.

In both greenhouse and field trials, Best Linear Unbiased Predictors (BLUPs) (Henderson 1975) were obtained by adding the general mean of the trial to the solution of the random genotype effect. Broad-sense heritabilities were calculated at the experimental design level as $\sigma_p^2 = \sigma_g^2 + \sigma_e^2/j$, where σ_g^2 is the genetic variance, σ_p^2 the phenotypic variance, σ_e^2 the error variance and j the number of replications.

Available molecular data for the progeny population:

A total of 1666 polymorphic markers are available. They were produced using 37 AFLP primer pairs, 46 SSRs and 9 RFLP probes (Raboin et al. submitted⁴). The nine RFLP probes are candidate genes differentially expressed in response to challenge by smut provided by the South African Sugarcane Research Institute (Heinze et al. 2001). 424 markers specific of the parent 'R570' formed 86 cosegregation groups (CGs) and 536 markers specific of the parent 'MQ76-53' formed 105 CGs. The detailed maps are presented at the web site <http://tropgenedb.cirad.fr/en/sugarcane.html> (Ruiz et al. 2004). Moreover, 111 double-dose (DD) markers formed 39 double dose CGs containing between 2 and 7 markers (24 CGs constituted of DD markers common to 'R 570' and 'MQ 76/53', 9 CGs constituted of markers specific of 'R 570' and 6 CGs constituted of markers specific of 'MQ 76/53'). Nevertheless, the totality of the markers (i.e. linked or unlinked) were used for the

⁴ chapitre 3 de la thèse

detection of marker-trait associations in the segregating progeny population. A great part of the markers of 'R 570' had also been previously mapped using a selfed progeny of 'R 570' (Hoarau et al. 2001; Rossi et al. 2003).

Available molecular data for the cultivar population:

A total of 1626 polymorphic markers are available. They were produced using 42 AFLP primer pairs. Among those markers, 396 could be assigned to 128 haplotypes composed of from two markers to 15 markers (Raboin et al. in prep⁵). Moreover, 258 of those 1626 markers have known positions on 'MQ 76/53' genetic map (raboin et al. submitted) and 479 of those markers have known positions on 'R 570' genetic maps (Hoarau et al. 2001; Rossi et al. 2003 and Raboin et al. Submitted). 'R 570' and 'MQ 76/53' were used as controls for the genotyping of the cultivar population. Thus, markers were labeled as follows: the marker name presented the suffix (1) “_r” if the marker was present in the pattern of the control 'R 570' only (2) “_m” if the marker was present in the pattern of the control 'MQ 76/53' only (3) “_rm” if the marker was present in both controls (4) “_N” if the marker was absent from both controls. The suffixes “_rN”, “_mN”, “_rmN” correspond to markers present in the pattern of one or both controls but were not coded in the previous genetic mapping works. All the markers were used for the detection of marker-trait associations in the population of cultivars.

QTL analysis:

One way ANOVA have been performed in order to detect QTLs (i.e. to test the difference between genotypes (clones) that present and those that do not present a given marker). These analyses were performed using the totality of available markers and BLUPs values calculated for each genotype (clone) of the progeny population as described above. Empirical threshold were calculated by permuting the trait data with respect to the genetic marker data. 10 000 permutations were performed to allow the estimation of an empirical Fstatistic giving a type I error of 0.005 for each individual test (Churchill and Doerge 1994; software by Offmann (2000)). All markers identified at P=0.005 have been considered for the analysis of results. The proportion of the total phenotypic variance explained by a marker (R^2) was calculated using the sums of square provided by the ANOVA.

⁵ chapitre IV de la thèse

Table 2: Analysis of variance in the two trials comparing three inoculation techniques

Source of variation	ddl	AUDPC	Percentage of diseased plants
Trial(1)	1	ns	ns
Clone(2)	2	****	****
Inoculation(3)	2	****	****
Bloc(trial)	6	ns	ns
(1)*(2)	2	ns	ns
(2)*(3)	4	****	****
(1)*(3)	2	ns	ns
(1)*(2)*(3)	4	ns	ns

****(P<0.0001), ns (P> 0.05)

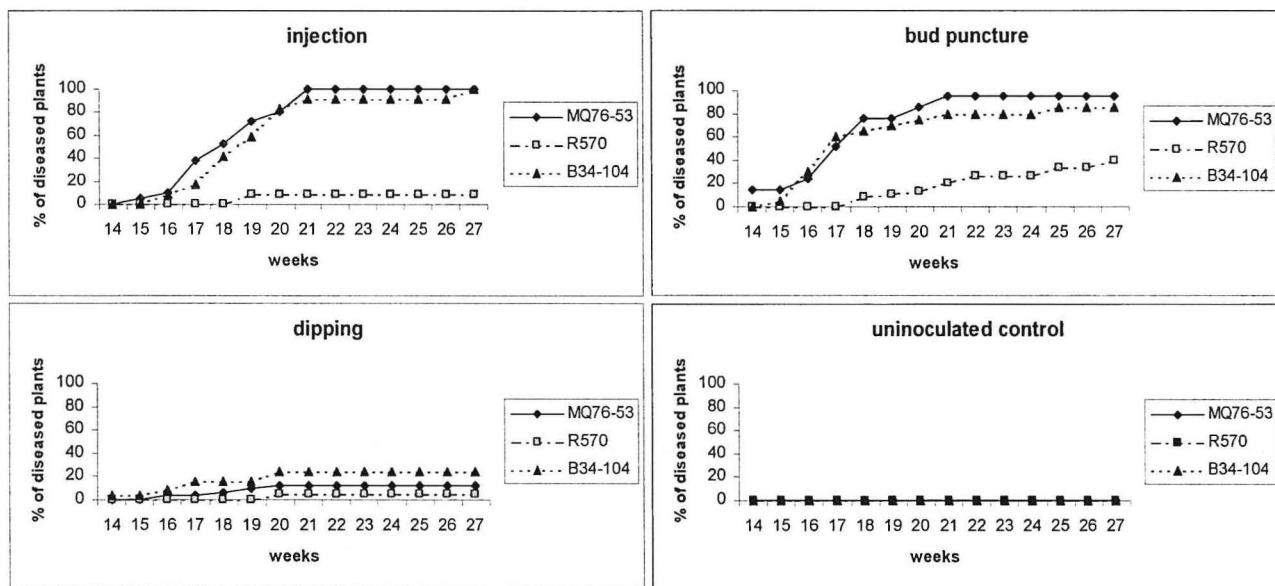


Fig. 1: Comparison of the different inoculation methods. Evolution across weeks of the cumulated number of diseased plants expressed as the percentage of the total number of inoculated plants.

Analysis of marker-smut resistance associations in the cultivar population

The association between markers and smut rating (table1) was tested by using a one-way ANOVA following the method described in the above section. In addition, given that our cultivar population has been constituted from extreme phenotypes for resistance to smut (table 1), we converted the smut rating of the clones into a binary rating. Resistant clones from the population were coded 1 whereas susceptible clones were coded 0. The independance between this binary smut rating and the 1626 available markers was therefore tested using a Fisher exact test (Mehta and Patel 1983; SAS institute 1990, FREQ procedure). We retained all marker-trait associations giving either a Fisher probability inferior to 0.005 or presenting a F statistic giving a $P < 0.005$ for the ANOVA.

V.3 Results

Comparison of different inoculation methods in Greenhouse trials:

Both greenhouse trials comparing different inoculation methods led to similar results (Table 2). Clones had contrasted responses to inoculation by smut. However, this response depended on the type of inoculation performed as indicated by a highly significant inoculation method effect as well as a highly significant interaction effect between clone and inoculation method. When the dipping method was performed, clones could not be satisfactorily distinguished for their resistance to smut. On the other hand, both injection and bud-puncture clearly distinguished the resistant clone from the susceptible clones (Fig 1). Inoculation by injection of a spore suspension appeared more difficult to standardize. It induced a high mortality of the inoculated buds in the first trial (75%). Therefore, we decided to use the bud puncture method for the greenhouse evaluation of the mapping population for resistance to smut. The parameters AUDPC and ‘percentage of diseased plants’ led to similar results.

Table 3: Statistical parameters(Means, Variance components, Broad sense heritabilities at the experimental design level, pearson correlations between traits) for each variablextrials combination

	Field (cumulated nb of whips per plot)					Greenhouse (AUDPC)	
	LP2001 plant cane	LP2002 1st ratoon	LP2003 2nd ratoon	LP2004 3rd ratoon	Gol2004 2nd ratoon	2003	2004
Means							
R 570	0.0	0.0	0.3	0.0	0.0	0.0	0.3
MQ 76/53	46.3	92.0	75.0	26.3	46.3	69.3	51.9
Progeny	3.8	8.0	8.8	4.8	4.4	7.2	6.6
variance components							
σ_g^2	15.9 ± 2.5	39.4 ± 5.5	63.4 ± 8.2	17.9 ± 2.9	24.7 ± 3.2	34.8 ± 4.8	53.4 ± 6.4
σ_e^2	22.2 ± 1.7	40.7 ± 3.1	48.1 ± 3.7	29.1 ± 2.2	19.3 ± 1.4	35.8 ± 2.6	28.7 ± 2.1
broad sense heritability							
h^2	0.68	0.74	0.80	0.65	0.79	0.74	0.85
interaction with nb of stalks							
pearson (Prob)	0.017 (0.81)	-0.117 (0.10)	0.052 (0.47)	0.022 (0.76)	-0.013 (0.86)	-	-
r_g	0.06	-0.12	0.07	0.009	0.003	-	-
effect as covariable	none	none	none	none	none	-	-
QTL analysis							
	no	no	yes	no	yes	yes	yes
Pearson correlations (probability)							
LP2001	-	0.62 (<0.0001)	0.44 (<0.0001)	0.33 (<0.0001)	0.28 (<0.0001)	0.39 (<0.0001)	0.48 (<0.0001)
LP2002	-	-	0.82 (<0.0001)	0.67 (<0.0001)	0.58 (<0.0001)	0.39 (<0.0001)	0.52 (<0.0001)
LP2003	-	-	-	0.83 (<0.0001)	0.68 (<0.0001)	0.43 (<0.0001)	0.52 (<0.0001)
LP2004	-	-	-	-	0.69 (<0.0001)	0.36 (<0.0001)	0.47 (<0.0001)
Gol2004	-	-	-	-	-	0.36 (<0.0001)	0.53 (<0.0001)
AUDPC2003	-	-	-	-	-	-	0.65 (<0.0001)
AUDPC2004	-	-	-	-	-	-	-

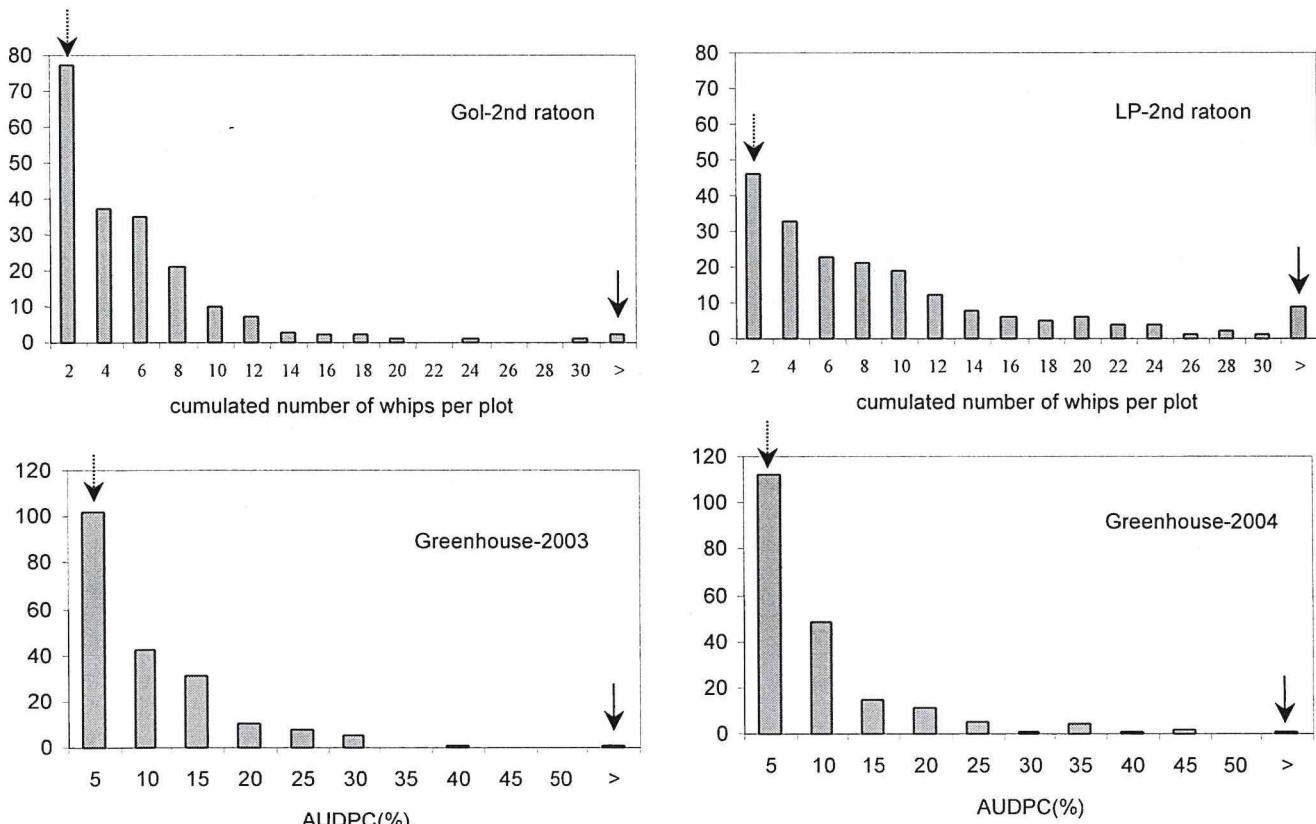


Fig. 2: Distribution of the smut resistance parameters measured in field trials (cumulated number of whips in Ligne-Paradis and Gol trials in 2nd ratoon) and greenhouse trials (AUDPC in 2003 and 2004 trials). The dotted arrows indicate the position of the resistant parent 'R570' and the plain arrows indicate the susceptible parent 'MQ 76/53'.

Smut resistance evaluation of the progeny population

In all field and greenhouse trials, the genotypic variance components were highly significant ($P<0.0001$) for the cumulated number of whips per plot in each crop cycle (field trials) and the AUDPC (greenhouse trials) (Table 3). In Ligne Paradis field trial (LP), the disease was assessed in plant cane and over three successive ratoons. The cumulated number of whips per plot increased over years to reach a maximum in the second ratoon. It decreased in the third ratoon. The broad sense heritability of the number of whips per plot behaved in the same way. For that reason, in the second trial (Gol) the disease was assessed only in second ratoon. For both trials, QTL analysis were performed on data from the second ratoon ($h^2= 0.80$ in LP2002 and $h^2= 0.79$ in Gol2004). In our conditions, the cumulated number of whips per plot do not interact with the number of millable stalks per plot measured before harvest. Pearson correlations and genetic correlations between these two parameters were almost null. Moreover, when the cumulated number of stalks was used as a covariate in the analysis of variance model, its effect was not significant. As a consequence, we used the raw cumulated number of whips per plot in further QTL analysis and not the ratio between number of whips and number of stalks. The Pearson coefficient of correlations between the different evaluations of smut resistance were always highly significant. In Ligne Paradis field trial, high correlations were observed between successive crop cycles. This was expected because when a stool is diseased, most of the times it remained diseased during the following crop-cycle. The cumulated number of whips measured in second ratoon of Gol trial was fairly correlated to that observed in second ratoon of Ligne Paradis trial ($r=0.69$). The AUDPC measured in the two greenhouse trials were also fairly correlated ($r=0.65$). By comparison, the correlations between field evaluations and greenhouse evaluations were lower but they still were significant.

Distributions, in the progeny population, of both cumulated number of whips per plot in the field trials and AUDPC in the greenhouse trials were L-shaped in a continuing pattern (Fig. 2). This suggests a polygenic determinism of the resistance. The two parents were always positioned at each extremity of the distributions: 'R570' always being one of the most resistant clones under evaluation and 'MQ 76/53' always being the most susceptible clone under evaluation. A great majority of the clones from the progeny were rated rather resistant (on the resistant 'R 570' side) and conversely a very small number of clones were rated as susceptibles (on the susceptible 'MQ 76/53' side). This situation suggests a dominant mode of inheritance of some of the multiple alleles involved in R 570 resistance.

marker ¹	type	CG ²	R/S ³	cM	HG	Rossi ²	JYH ²	n(+) ⁴	n(-) ⁴	AUDPC 2003			AUDPC 2004			LP 2003 (2nd ratoon)			Gol 2004 (2nd ratoon)														
										avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P												
R570 single dose markers																																	
acacaa_rx1	AFLP	group53	s	46.8	III			95	98								10.019	7.098	0.042 ***														
aggctc_r12	AFLP	group53	s	48	III	III-3	23	92	103								10.057	7.058	0.045 ***														
agccat_r9	AFLP	group67	r	31.4	VIII	VIII-1	14	92	78										3.678	5.595	0.042 ***												
acactt_r18	AFLP	group67	r	46	VIII	VIII-1	14	96	61								7.227	10.756	0.057 ***	3.487	5.538	0.049 ***											
m12_r4	SSR	group67	r	111.3	VIII	VIII-1		103	91					5.130	8.002	0.049 ***	7.194	9.882	0.036 **	3.631	5.198	0.031 **											
m38_r1	SSR	group67	r	119.2	VIII	VIII-1		105	89					5.192	7.994	0.047 ***	7.126	9.990	0.041 ***														
ADS8_r3h	RFLP	group7	r	106.1	VI			45	42								6.440	10.436	0.098 ***														
acccaa_r14	AFLP	group7	r	121.9	VI	VI-3	29	92	100								6.703	10.049	0.055 ***														
acgcag_r15	AFLP	group7	r	122.9	VI					91	104						6.790	9.945	0.049 ***														
accat_r16	AFLP	group7	r	125.6	VI	VI-3	29	94	91								6.864	10.191	0.053 ***														
acccac_rx2	AFLP	group7	r	127.7	VI					94	95						6.941	9.945	0.046 ***														
accctg_r24	AFLP	group7	r	129.7	VI	VI-3	29	95	98								6.977	10.049	0.047 ***														
acacla_rx1	AFLP	group7	r	131.8	VI					94	92						7.033	10.289	0.051 ***														
aaccac_r13	AFLP	group7	r	142.8	VI	VI-3	29	91	103								6.872	9.950	0.047 ***														
m54_r6	SSR	group7	r	151.3	VI	VI-3				99	97						6.808	10.138	0.055 ***														
ADS6_r8	RFLP	group7	r	158.3	VI					96	91						7.095	10.138	0.045 ***														
m54_r7	SSR	group8	s	37.6	VI	VI-2		86	110								10.189	7.101	0.047 ***	5.349	3.615	0.038 ***											
acacaa_r7	AFLP	group40	r	15.3	I	I-7	74	88	105										3.312	5.245	0.047 ***												
m4m_r6	SSR	group58	s	0						79	117						10.218	7.266	0.042 ***														
agccaa_r5	AFLP	group130	r	23						96	71								3.519	5.671	0.054 ***												
acactt_r10	AFLP	group106	r	51.7	IV	IV-2	9	89	89										3.434	5.188	0.038 ***												
m4m_r3	SSR	r			VII	VII-14				95	101						6.911	9.909	0.045 ***	3.392	5.281	0.045 ***											
ADS2_r2	RFLP		r							25	20								3.548	5.898	0.160 ***												
acacaa_r14	AFLP		r		VIII	U-53	17	108	85								7.167	10.275	0.047 ***	3.455	5.553	0.055 ***											
MQ76/53 single dose markers																																	
aagcac_m8	AFLP	group10	r	71.2				87	100										3.319	5.407	0.054 ***												
agccit_m17	AFLP	group10	r	77.9				64	71										3.377	5.682	0.054 ***												
actctg_m3	AFLP	group119	s	35.9				91	101										5.537	3.330	0.062 ***												
aaccgt_m12	AFLP	group119	s	40.2						98	92								5.264	3.452	0.041 ***												
acactc_m3	AFLP	group119	s	48.6				96	99										5.253	3.531	0.038 ***												
m38_m5	SSR	group3	r	0	VIII			98	96								6.951	9.960	0.045 ***	3.323	5.429	0.057 ***											
accctt_m10	AFLP	group3	r	9.6	VIII			65	71					5.245	8.528	0.059 ***				3.310	5.641	0.056 ***											
aagcag_m16	AFLP	group3	r	12.6	VIII			91	103								6.878	9.883	0.045 ***	3.233	5.376	0.058 ***											
acgcta_m16	AFLP	group3	r	16.9	VIII			88	95	5.959	8.066	0.043 **	5.019	7.759	0.046 ***			6.726	9.869	0.051 ***	3.189	5.426	0.063 ***										
agccta_m16	AFLP	group3	r	18.7	VIII			91	100								7.000	9.889	0.041 **	3.297	5.347	0.053 ***											
accctc_m4	AFLP	group3	r	20.3	VIII			92	101								7.105	9.844	0.037 **	3.302	5.401	0.056 ***											
acccac_m2	AFLP	group3	r	21.9	VIII			93	99										3.336	5.311	0.050 ***												
acactg_m6	AFLP	group3	r	25.7	VIII			96	98										3.398	5.343	0.048 ***												
actctg_m7	AFLP	group3	r	66.8	VIII			93	100										3.436	5.244	0.042 ***												
actcac_m10	AFLP	group3	r	73.5	VIII			96	97										3.448	5.262	0.042 ***												
actcac_m8	AFLP	group3	r	74.6	VIII			95	99										3.466	5.252	0.041 ***												
m54_m15	SSR	group63	r	0	VI			96	100								6.917	9.933	0.045 ***														
m10_m9	SSR	group20	r	0	VI			94	101								6.686	9.996	0.055 ***	3.555	5.095	0.030 **											
aggcat_m1	AFLP	group80	s	0				110	84								9.681	6.852	0.039 **	5.125	3.288	0.043 ***											
acactg_m13	AFLP	group87	s	9.9				87	104	8.102	6.047	0.042 ***																					
multiple doses markers																																	
aagcag_rm23	AFLP	DDgroup10	r		VI	U51		149	46										3.856	6.060	0.045 ***												
acttag_rm7	AFLP	DDgroup10	r		VI	U51		147	46										3.860	6.135	0.048 ***												
m67_rm8	SSR	DDgroup36	r		VII	VII8		121	67	6.289	8.627	0.047 ***																					
agccat_rm15	AFLP		r		VI	U-36	95	131	59									7.412	10.822	0.049 ***													
aagcgt_rm4	AFLP		r							149	24			5.822	11.258	0.079 ***			7.774	13.182	0.070 ***												
m12_rm1	SSR		r		VIII	VIII-3	141	53						5.742	8.433	0.053 ***																	
m38_rm4	SSR		r		VIII	VIII-15	186	8						6.151	14.058	0.059 ***																	
m62_rm6	SSR		s		I	I-6	134	58						7.313	4.414	0.042 ***																	
agccag_r5	AFLP	DDgroupR28	r					141	50					5.555	8.994	0.055 ***																	
acacaa_x2	AFLP		r					189	4	6.865	15.635	0.060 ***			6.147	23.773	0.151 ***																
agccag_rx1	AFLP		r					189	3	6.830	19.670	0.097 ***			6.067	31.360	0.236 ***																
aagcag_m14	AFLP		r					131	61									7.465	10.739	0.046 ***													
actctg_m6	AFLP		r					183	10																								
aaccgt_m10	AFLP		r					130	59																								
acactc_m4	AFLP		r					142	42	6.454	9.099	0.046 ***																					
m33_m1	SSR		r					133	56					</td																			

QTL analysis

The marker-trait associations detected at $P=0.005$ are presented in Table 4. Considering both greenhouse trials and both field trials altogether, 60 markers (out of 1666) were involved in those associations (44 markers were single dose markers and the other 16 were multiple doses ones). Among those markers, a great majority (50 of 60) presented an effect increasing resistance although an equivalent number of markers were specific of the susceptible parent ‘MQ 76/53’ and the resistant parent ‘R 570’ (27 markers were specific of ‘R 570’, 25 markers were specific of ‘MQ 76/53’ and 8 markers were common to both parents). Marker-trait associations involving closely linked markers were considered to reflect the effect of the same Quantitative Trait Allele (QTA). When cosegregating markers linked to the same trait were separated by more than 30 cM they were considered as linked to distinct QTAs. This may lead to an under-estimation of the genuine QTA number since several QTAs might be involved over a segment of 30 cM. Nevertheless, following this method, 12 QTAs were identified in ‘R 570’ (9 for resistance and 3 for susceptibility) and 9 QTAs were identified in ‘MQ 76/53’ (6 for resistance and 3 for susceptibility). Moreover, 16 multiple-doses markers were also found in association with smut resistance rating in one of the four evaluations performed. Two double-dose markers belonged to a same double-dose cosegregation group. Interestingly, alleles of the RFLP probes ADS2 (Flavonoid transcription factor), ADS6 (Isoflavone reductase) and ADS8 (Cellulose synthase), that correspond to loci suggested by Heinze et al. (2001) to be associated with smut resistance, have been found linked to a QTL of resistance in our study.

The proportion of the total phenotypic variation explained by an individual QTA was generally small and ranged from 3% to 6.3 % except for two markers, with a lot of missing values, which may have resulted in an over-estimation of the marker effect (9.8% for ADS8_3h and 16% for ADS2_r2). The proportion of the total phenotypic variation explained by a given multiple doses marker was also small except for a few markers with highly unbalanced segregation most probably inducing an over-estimation of the marker effect. Only a few QTAs were detected across field and greenhouse trials involving SSR alleles m12_r4 and m38_r1 (HG VIII and CG67), AFLP markers accctt_m10 and acgcta_m16 (HG VIII and CG3) and double dose markers m12_rm1 (HG VIII) and aagctg_rm4. Interestingly, at two SSR loci (m38 and m54 assigned to HG VIII and HG VI respectively) three distinct QTAs were detected. Similarly, at two other SSR loci (m4 and m12 assigned to HG VII and HG VIII respectively) two distinct QTAs were detected.

Table 5: Associations between AFLP markers and smut resistance rating in the LD population

marker ¹	freq ²	R ³		S ³		ANOVA ⁵				Fisher prob	haplotype ⁶	mapping information			
		+	-	+	-	avg(+) ⁴	avg(-) ⁴	F	F5pm	R ²		HG ⁷	CG ⁷	position ⁷	QTL ⁸
Resistance															
acacta_rX1	0.56	27	12	11	18	3.42	5.53	8.96	8.84	0.12	0.013932	acgctt_rm20	VI		X
aaccac_r13	0.56	28	11	12	21	3.48	5.97	13.10	8.84	0.16	0.004079	acgctt_rm20	VI	JYH_29	55.7 X
accctg_r14	0.53	27	12	11	22	3.42	5.79	12.18	8.47	0.15	0.004168	acgctt_rm20	VI	JYH_29	64.9
accctat_r16	0.54	27	10	12	23	3.56	6.09	13.67	9.06	0.16	0.001886	acgctt_rm20	VI	JYH_29	67.6 X
acccaa_r14	0.50	26	13	11	24	3.46	5.84	12.39	8.34	0.15	0.004858	acgctt_rm20	VI	JYH_29	71.1 X
aagcaa_rm20	0.72	30	4	19	15	4.02	6.89	13.55	9.13	0.17	0.006006		VI	JYH_29	98.3
aagctt_r21	0.52	23	10	11	21	3.59	5.94	10.86	8.81	0.15	0.006303	acgctt_rm20	VI	JYH_56	144.6
aacctg_N17	0.32	19	20	5	30	2.83	5.52	14.14	8.30	0.16	0.002502	agccaa_N14			
acccac_N4	0.33	19	20	5	29	2.83	5.49	13.60	8.27	0.16	0.002631	agccaa_N14			
acccta_N10	0.49	25	14	10	23	3.51	5.59	8.74	8.64	0.11	0.005129	agccaa_N14			
accctg_N10	0.35	19	20	6	27	3.00	5.36	10.78	8.23	0.13	0.012206	agccaa_N14			
agccat_N6	0.27	17	22	3	31	2.65	5.32	12.46	8.16	0.15	0.001273	agccat_N6			
acacaa_r7	0.46	21	13	10	23	3.61	5.78	8.64	8.54	0.12	0.014277	aggcat_N9	I	JYH_74	174.7 X
acccac_rX3	0.61	27	9	13	17	3.65	5.96	9.49	8.80	0.13	0.011884	acccac_rX3			
accctg_rR	0.36	17	15	4	23	3.05	5.42	8.58	8.60	0.13	0.002772	actctg_rm0	VI	JYH_56	130.4
accctt_N1	0.12	9	30	0	34	2.00	4.98	7.83	7.76	0.10	0.002723				
acactc_N6	0.64	34	5	13	22	3.38	6.85	29.33	8.28	0.29	0.000009				
aacctg_N3	0.25	15	24	3	30	2.67	5.22	10.20	8.53	0.13	0.005710				
agccag_N1	0.51	27	12	11	24	3.29	6.08	18.27	8.52	0.20	0.002236				
aaccac_r29	0.79	35	3	22	12	4.07	6.87	10.70	8.00	0.13	0.007543				
acccta_rN17	0.86	37	1	25	9	4.19	7.50	10.87	8.28	0.13	0.004843				
acactg_r15	0.57	27	12	15	20	3.69	5.91	10.29	8.13	0.13	0.034056		I	JYH_1	39.2
acacta_rX3	0.49	26	13	10	25	3.36	5.87	14.02	8.17	0.16	0.001276				
Susceptibility															
accctat_m11	0.18	2	37	11	24	7.23	4.10	12.46	7.95	0.15	0.004665	accctat_m11	MQ_100	0	
aagctg_m16	0.21	3	36	12	21	6.80	3.96	11.24	8.31	0.14	0.003655	accctat_m11	MQ_182	34.7	
acccaa_N1	0.14	0	37	10	24	7.90	4.13	15.09	8.18	0.18	0.000284	acgcag_N14			
acactg_N5	0.16	1	38	11	23	7.75	3.97	18.52	8.30	0.21	0.000849	acgcag_N14			
acacctt_N3	0.13	1	37	8	26	7.33	4.25	8.63	8.34	0.11	0.010638	acgcag_N14			
acactg_m16	0.09	0	39	7	28	8.14	4.28	10.99	7.42	0.13	0.003737	actcac_N4	VIII	MQ_50	90.9
actcac_N4	0.09	0	39	7	28	8.14	4.28	10.99	7.42	0.13	0.003737	actcac_N4			
aagcta_mN6	0.14	1	38	9	24	7.60	4.03	13.60	8.00	0.16	0.004163	actcac_N4			
acccac_m7	0.04	0	39	3	31	9.00	4.43	6.60	6.39	0.09	0.096212		VIII	MQ_123	1.1
accctg_m9	0.03	0	39	2	31	9.00	4.41	4.47	4.30	0.06	0.206573		VI	MQ_20	9.8
accctat_m5	0.07	0	39	5	30	8.60	4.36	9.58	6.64	0.12	0.020153			MQ_34	6.1
acgctt_m6	0.05	0	39	4	30	8.50	4.39	7.04	6.79	0.09	0.042608		VIII	MQ_36	80.7
acactg_m17	0.40	9	29	20	15	5.93	3.86	8.41	8.26	0.11	0.004428		VIII	MQ_84	24.9
accctc_m6	0.15	2	37	9	26	7.27	4.19	10.27	8.21	0.12	0.020054		VIII	MQ_90	0
aagcaa_mN2	0.08	0	39	6	29	8.00	4.35	8.26	7.37	0.10	0.008762				
actcat_m4	0.07	0	38	5	29	8.20	4.42	7.29	7.18	0.09	0.019887				
aagcac_m9	0.07	0	39	5	30	8.60	4.36	9.58	7.55	0.12	0.020153				
accctat_mN4	0.07	0	39	5	30	8.40	4.38	8.52	6.64	0.11	0.020153				
accctc_m20	0.07	0	39	5	30	8.40	4.38	8.52	7.55	0.11	0.020153				
acgcag_N15	0.07	0	39	5	30	8.20	4.39	7.55	6.64	0.09	0.020153				
accctt_N3	0.07	0	33	4	24	9.00	4.30	9.06	7.20	0.13	0.039235				
aggctc_N19	0.05	0	39	4	31	9.00	4.40	9.11	6.95	0.11	0.045506				
acacta_mN8	0.31	8	31	14	19	6.09	3.90	8.25	8.24	0.11	0.071316				
accctg_m6	0.03	0	39	2	31	9.00	4.41	4.47	4.30	0.06	0.206573				
acccaa_m16	0.09	0	39	7	28	8.43	4.25	13.20	8.17	0.15	0.003737				
actctg_N13	0.09	0	39	7	28	8.14	4.28	10.99	7.42	0.13	0.003737				
aggctg_mN1	0.72	21	16	30	4	5.33	3.10	8.04	8.68	0.10	0.003852				
acactc_m5	0.19	2	37	12	23	7.29	4.03	14.58	8.40	0.17	0.002144				
aagcaa_mN4	0.45	10	27	20	10	5.97	3.30	14.57	9.09	0.18	0.001491				
aagctc_rD	0.53	14	25	25	10	5.54	3.66	7.26	8.52	0.09	0.002727				
acgcat_r24	0.40	9	29	20	15	5.76	3.98	6.06	8.38	0.08	0.004428		VIII	JYH_47	0
accctat_r11	0.24	5	34	13	22	6.44	4.07	8.68	7.73	0.11	0.028243		II	JYH_85	6.1
acactg_rm23	0.28	6	33	15	20	6.48	3.92	11.47	8.35	0.14	0.010878	acccac_rm10	VI	JYH_33	23.9
acccat_rm18	0.66	15	17	25	4	5.70	2.81	13.64	9.03	0.19	0.001425		VIII	JYH_86	0
agccat_rmX1	0.68	16	17	26	3	5.60	2.65	13.86	8.76	0.19	0.000846				
acccat_rm41	0.60	14	19	22	5	5.61	2.92	12.22	8.96	0.17	0.003303				

¹ the suffix of the marker name indicates if the marker is present in R 570 (_r), in MQ 76/53 (_m), in both of them (_rm) or unknown (_N)² frequency of the marker in the LD population³ number of clones presenting the marker (+) or not (-) within the resistant group of clones (R)and the susceptible group of clones (S)⁴ average smut rating among clones presenting the marker (+) or not (-)⁵ analysis of variance between the groups of clones presenting the marker or not. F5pm empirical threshold giving a type I error of P=0.005⁶ haplotypes identified in Raboin et al. (in prep)⁷ mapping information from (Hoarau et al. 2001; Rossi et al. 2003; Raboin et al. submitted). HG homology group; CG cosegregation group⁸ markers involved in a QTL for smut resistance (this paper)

Analysis of marker-smut resistance associations in the cultivar population

The most significant associations between markers and smut rating are presented in Table 5. In total, 59 markers (of 1626) were involved in those associations among which 24 were associated with resistance and 35 with susceptibility. Twenty three of those markers could be assigned to 10 previously identified haplotypes (Raboin et al, in prep). For five of those haplotypes, between 2 and 6 markers were found associated with smut resistance or susceptibility. Five markers involved in QTLs were also found significantly associated with smut resistance in the cultivar population (4 of them cosegregated in CG29 and the remaining one belonged to CG74 (Hoarau et al, 2001)). In both QTL mapping and association mapping approaches, the effect of these 5 markers was similar and enhanced the level of resistance. An interesting discussion point is that markers of ‘MQ 76/53’, the highly susceptible parent of the progeny population, were detected in association with susceptibility. Conversely most markers of ‘R 570’, the resistant parent of the progeny population, were detected in association with resistance. Four markers common to ‘R 570’ and ‘MQ 76/53’ were associated with susceptibility and one with resistance.

V.4 Discussion

The biparental cross between ‘R 570’ and a highly susceptible tester clone ‘MQ 76/53’ disappointingly revealed unsuited to unravel the genetic architecture of ‘R 570’ resistance to smut through a QTL mapping approach. The progeny did not segregate appropriately for a QTL analysis either in the field trials or in the greenhouse trials. In our conditions, most of the clones from the biparental progeny appeared rather resistant. Thus, although broad-sense heritabilities in the different evaluations were fairly good, this lack of variability was not favorable for the statistical detection of loci involved in the resistance. For single dose markers in particular, the difference for resistance traits means (AUDPC, cumulated number of whips...) between genotypes presenting a given marker and genotypes not presenting this marker was never important. That’s probably why in the two greenhouse trials single dose markers accounted for only 8% and 10% of the phenotypic variation. Besides, in those same greenhouse trials, very few markers specific of the resistant parent ‘R 570’ have been found linked to a putative QTL of resistance. In the field trials, more QTLs have been detected explaining together 46% and 58% of the phenotypic variation in Ligne Paradis and Gol trials respectively. Nevertheless, like in greenhouse evaluations clones were distributed in their grand majority toward the resistant parent. The disease never reached a significative impact, except

for a very few number of clones, as suggested by the absence of interaction of the cumulated number of whips with the final number of millable stalks. The QTLs identified had small effects and lacked of repeatability, a common drawback in QTL approaches (Mackay, 2001) and even more common in the polyploid context of sugarcane (Hoarau et al, 2002). In R 570, resistance probably result from multiple dominant factors but not much progress has been made here toward their identification. It still is not clear whether these multiple factors belong to a limited or a great number of loci.

There are key differences between greenhouse and field trials evaluations. The bud puncture method of inoculation used in greenhouse evaluations facilitate the penetration of the fungus by eliminating the natural barrier of the bud scales. Moreover under greenhouse we looked at the percentage of inoculated buds that convert themselves into whips whereas in the field trials we measured the number of whips that developed from an unknown number of diseased stools (originating from a “successful” contamination event by the fungus). The sugarcane-smut fungus interaction may schematically be divided into three distinct phases (1) the fungus is successful (or not) to break through the barrier (mechanical and/or chemical barrier (Lloyd and Naidoo 1983)) of the bud scales (2) the fungus is successful (or not) to become established in the meristematic area of the targeted sugarcane bud (3) the established fungus triggers (or not) the production of a lot of diseased suckers that will evolve in secondary whips. Distinct resistance mechanisms may be involved for each phase. In our greenhouse trials we should have been able to assess Phase II resistance only whereas the clone behaviour in the field trials resulted from interactions at the three Phases. Our hypothesis is that on one hand we probably missed most of loci involved in the “Phase II” resistance mechanism but that on the other hand we may have identified interesting loci involved in the “pre-infection” or “post-infection” reaction of sugarcane plants to the fungus.

The association mapping approach gave promising results although the validity of the markers revealed by the analysis has to be confirmed. The fact that markers, present in the highly susceptible parent ‘MQ 76/53’, were associated exclusively with susceptibility and that markers, present in the resistant ‘R 570’, were mainly associated with resistance give credit to the biological reality of those associations. Moreover, two of the haplotypes associated with resistance had also been identified as QTAs of resistance in the progeny population. The appropriate number of evenly positionned locus-specific markers required for a whole-genome LD-based association study has been estimated around 600 (300 at the minimum corresponding to one marker every 5 cM of the

estimated 1500 cM of the basic monoploid genome of sugarcane) (Raboin et al. in prep). It is difficult to apply this mode of calculation to dominant AFLP markers but the number of markers analysed in this study (1626) was far from sufficient to identify the totality of the haplotype diversity present in the cultivar population.

An interesting feature, in both approaches, is that a great part of the markers associated with smut resistance or susceptibility belong to Homology Group (HG) VI or VIII. Interestingly, Rossi et al. (2003) localised three clusters of Resistance Gene Analogs on HG VIII of a total of eight clusters identified in R 570. Moreover, the candidate defense genes, differentially expressed in response to challenge by smut (Heinze et al. 2001), used as RFLP probes in the QTL approach have all been located on HG VI (ADS1, ADS3, ADS4, ADS6 and ADS8) and HG VIII (ADS7 and ADS9) (Raboin et al submitted⁶).

In conclusion, the understanding of the genetic control of sugarcane resistance to smut will require the screening of multiple bi-parental combinations between resistant cultivars and highly susceptible testers so as to identify crosses presenting appropriate segregations for QTL detection. The resistant parents could be chosen within the cultivar population from the association mapping approach. This would be a way to validate candidate markers identified in the association mapping approach. Association mapping is promising but must be implemented on a larger scale: larger populations and probably a minimum of 4 times more markers.

Acknowledgments: we thank SOSUCO (Société sucrière de la Comoé au Burkina Faso) for sharing informations regarding the smut resistance rating of most of the cultivars used in the present study. We thank Iréné Promi for his help in the field work. We thank Jean Pierre Diorflar for the management of Gol field trial (CERF experimental station).

⁶ chapitre 3 de la thèse

References :

- Al-Janabi S.M., Parmessur, Kross H., Dhayan S., Saumtally S. and Dookun-Saumtally A. (2005) Characterization of quantitative trait loci (QTLs) contributing to field resistance to yellow spot disease in sugar cane. Proc. Int. Soc. Sugarcane Technol.
- Asnaghi C., Roques D., Ruffel S., Kaye C., Hoarau J.Y., Télismart H., Girard J.C., Raboin L.M., Risterucci A.M., Grivet L., D'Hont A. 2004. Targeted mapping of a sugarcane rust resistance gene (*Bru1*) using bulked segregant analysis and AFLP markers. *Theor. Appl. Genet.* 108:759-764.
- Antoine, R., 1961. Smut. In: Martin, J.P., Abbott, E.V., Hughes, C.G. (Eds), Sugar Cane Disease of the World. Elsevier Publishing Co., Amsterdam, pp 327-345.
- Chao, CP, Hoy, JW, Saxton, AM and Martin, FA (1990). Heritability of resistance and repeatability of clone reactions to sugarcane smut in Louisiana. *Phytopathology* 80: 622-626.
- Churchill, G.A. and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963-971.
- Comstock, J. C., Heinz, D.J., 1977. A new race of culmicolous smut of sugarcane in Hawaii. *Sugarcane Pathologists' Newsletter* 19, 24-25.
- Comstock, J.C., 2000. Smut. In: Rott, P., Bailey, R.A., Comstock, J.C., Croft, B.J., Salem Saumtally, A. (Eds), A guide to sugarcane diseases. CIRAD and ISSCT, pp. 181-185.
- Daugrois JH, Grivet L, Roques D, Hoarau JY, Lombard H, Glaszmann JC, D'Hont A (1996) A putative major gene for rust resistance linked with an RFLP marker in sugarcane cultivar R570. *Theor Appl Genet* 92:1059–1064
- D'hont A., Grivet L., Feldmann P., Seshagiri Rao, Berding N. et Glaszmann J.C. 1996. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum spp.*) by molecular cytogenetics. *Mol. Gen. Genet.* 250:405-413.
- D'Hont A, Garsmeur O, Raboin LM, Paulet F, Begum D, Wing R, Glaszmann JC (2001) Chromosome walking toward a major resistance gene for common rust of sugarcane. *Proc Int Soc Sugarcane Technol* 24: 315-317
- Ferreira, S.A., Comstock, J.C., 1989. Smut. In: Ricaud, C., Egan, B.T., Gillaspie, A.G., Hughes, C.G. (Ed.), Diseases of sugarcane. Elsevier, Amsterdam, pp. 211-229.
- Flint-Garcia, S.A., T.J. M., and E.S. Buckler IV. 2003. Structure of Linkage Disequilibrium in plants. *Annu. Rev. Plant Biol.* 54: 357-374.
- Garris, A.J., S.R. McCouch, and S. Kresovich. 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus in rice (*Oryza sativa L.*). *Genetics* 165: 759-769.
- Grivet L, Arruda P (2001) Sugarcane genomics: depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology* 5:122-127
- Heinze BS, Thokoane LN, Williams NJ, Barnes JM, Rutherford RS (2001). The smut-sugarcane interaction as a model system for the integration of marker discovery and gene isolation. *Proc S Afr Sug Technol Ass* 75: 88-93

Henderson C (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31 : 423-447.

Hoarau JY, Offmann B, D'Hont A, Risterucci AM, Roques D, Glaszmann JC, Grivet L (2001) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp). I. Genome mapping with AFLP markers. Theor Appl Genet 103:84-97

Hoarau JY, Grivet L, Offmann B, Raboin LM, Diorflar JP, Payet J, Hellmann M, D'Hont A, Glaszmann JC (2002) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp). II. Detection of QTLs for yield components. Theor Appl Genet 105:1027-1037

Hoy, J.W., Hollier C.A., Fontenot D.B., Grelen L.B. 1986. Incidence of sugarcane smut in Louisiana and its effect on yield. Plant Disease 70: 59-60.

Ivandic, V., W.T.B. Thomas, E. Nevo, Z. Zhang, and B.P. Forster. 2003. Associations of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. Plant Breeding 122: 300-304.

Jannoo, N., L. Grivet, A. Dookun, A. D'Hont, and J.C. Glaszmann. 1999a. Linkage disequilibrium among modern sugarcane cultivars. Theor. Appl. Genet. 99: 1053-1060.

Lee-Lovick, G., 1978. Smut of sugarcane - *Ustilago scitaminea*. Review of Plant Pathology of the Commonwealth Mycological Institute 57 (5), 181-187.

Leonards-Schippers C, Gieffers W, Schafer-Pregl R, Ritter E, Knapp Sj, Salamini F and Gebhardt C (1994) Quantitative resistance to *Phytophtora infestans* in potato: a case study for mapping in an allogamous plant species. Genetics 137:67-77.

Leu, L.S., Teng, W.S., 1972. Pathogenic Strains of *Ustilago scitaminea* Sydow. Sugarcane Pathologists' Newsletter 8, 12-13.

Lloyd H.L. and Naidoo G. 1983. Chemical assay potentially suitable for determination of smut resistance of sugarcane cultivars. Plant disease. 67:1103-1105.

Mackay, T.F.C. 2001. The genetic architecture of quantitative traits. Annu. Rev. Genet. 35: 303-339.

Mehta CR, Patel NR (1983) A network algorithm for performing Fisher's exact test in rxc contingency tables. J Am Statist Assoc 78:427-434

Michelmore R., 1995. Molecular approaches to manipulation of disease resistance genes. Annu. Rev. Phytopathol. 15: 393-427.

Muhammad; S.; Kausar; A.G.; 1962. Preliminary studies on the genetics of sugarcane smut, *Ustilago scitaminea* Sydow. Biologia 8, 65-74.

Neale, D.B. and O. Savolainen. 2004. Association genetics of complex traits in conifers. Trends Plant Sci. 9: 325-330.

Offmann B. 2000. Caractérisation et analyse génétique de la résistance de la canne à sucre *Xanthomonas albilineans*. Thèse de doctorat en sciences, Université de la Réunion, 116 p.

Patterson HD, Williams ER (1976) A new class of resolvable incomplete block designs. *Biometrika* 63:83-92.

Péros, J.P., Baudin, P., 1983. Etude de la variabilité d'*Ustilago scitaminea* Syd., agent du charbon de la canne à sucre. *L'Agronomie Tropicale* 38 (3), 234-239.

Rossi M, Araujo PG, Paulet F, Garsmeur O, Dias VM, Chen H, Van Sluys MA, D'hont A (2003) Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. *Mol Gen Genet* 269:406-419

Ruiz M, Rouard M, Raboin LM, Lartaud M, Lagoda P, Courtois B (2004) Tropgene-DB, a multitropical crop information system. *Nucleic Acids Res* 32:D364-D367

Simko, I., S. Costanzo, K. Haynes, B. Christ, and R. Jones. 2004. Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. *Theor. Appl. Genet.* 108: 217-224.

Toffano, W.B., 1976. Estudos sobre polaridade sexual e possíveis raças fisiológicas de *Ustilago scitaminea* Syd., no estado de São Paulo. *Arq. Inst. Biol.*, São Paulo 43, 65-79.

Wu, KK, Heinz, DJ and Meyer, HK (1983). Heritability of sugarcane smut resistance and correlation between smut grade and yield components. *Crop Science* 23: 54-56.

Wu K.K., Heinz D.J. and hogarth D.M., 1988. Association and heritability of sugarcane smut resistance to races A and B in Hawaii. *Theor.Appl.Genet.* 75: 754-760.

Young, N.D. 1996. QTL mapping and quantitative disease resistance in plants. *Annu. Rev. Phytopathol.* 34: 479-501.

VI CONCLUSIONS ET PERSPECTIVES

VI CONCLUSIONS ET PERSPECTIVES

Le fil conducteur et l'objectif de départ de ce travail de thèse étaient de progresser dans la compréhension du déterminisme génétique de la résistance de la canne à sucre au charbon. En préalable, une étude de la diversité génétique du pathogène *Ustilago scitaminea* a été conduite pour déterminer i) en particulier, quelle variabilité existait dans les deux zones géographiques (la Réunion et le Burkina Faso) où les évaluations de résistance au charbon concernant ce travail de thèse ont été menées ii) plus généralement, comment se structurait la diversité du pathogène au niveau mondial. Pour tenter de localiser les locus impliqués dans la résistance au charbon, deux stratégies ont été mises en œuvre i) la cartographie génétique et la détection de QTL dans un croisement (évalué à la Réunion pour la résistance au charbon) entre un géniteur résistant et un géniteur testeur extrêmement sensible ii) l'exploitation du déséquilibre de liaison pour identifier des associations entre marqueurs et résistance au charbon (évaluée au Burkina Faso) dans une population non-structurée de cultivars. A ce stade, peu de progrès ont été réalisés dans la compréhension génétique de la résistance au charbon même si l'analyse comparée des approches QTL et DL a mis en évidence des zones probablement impliquées dans la résistance. En revanche, des résultats «collatéraux» très intéressants ont été obtenus concernant i) la diversité génétique et la biologie du champignon pathogène ii) la cartographie génétique chez la canne à sucre et l'identification de nouveaux caractères morpho-agronomiques à hérédité mendélienne iii) la structure du déséquilibre de liaison chez la canne à sucre.

Diversité génétique de *Ustilago scitaminea*, agent du charbon de la canne à sucre

Nouvelles connaissances apportées

- Pour la première fois chez ce champignon, 38 locus microsatellites ont été caractérisés et des amores définies pour chacun de part et d'autre du motif répété. Vingt deux de ces locus ont été utilisés pour étudier la diversité génétique du champignon.
- Le mode de reproduction du champignon est très préférentiellement l'autofécondation. Les populations de cet agent pathogène sont constituées de lignées consanguines. Tous les échantillons diploïdes analysés, à une exception près, étaient homozygotes.
- Tous les échantillons provenant d'Amérique et d'Afrique qui ont été analysés appartiennent à une même filiation consanguine mondiale provenant d'Asie. L'Asie est le

continent qui présente l'essentiel de la diversité du pathogène. Il y a donc eu de très rares événements de migration du champignon de l'Asie vers les autres continents.

Quelle suite à donner?

Hors d'Asie, il semble bien qu'un génotype unique se soit répandu. Cette hypothèse forte devrait néanmoins être vérifiée, particulièrement en ce qui concerne des zones de contaminations relativement anciennes par le charbon comme l'Afrique de l'Est (Kenya, Tanzanie, Mozambique) ou l'Argentine (contaminée dès 1940) et pour lesquelles nous ne disposons pas d'échantillons. Le centre de diversité du pathogène est clairement situé en Asie mais il est néanmoins nécessaire d'entreprendre une étude approfondie des populations du pathogène sur ce continent pour en préciser la localisation. L'Inde, la Chine, le Pakistan en particulier doivent être explorés.

L'existence de spécialisations race-cultivar a été évoquée par différents auteurs pour le couple charbon-canne mais n'est solidement étayée que dans le cas de Taiwan (Leu 1978; Lee et al. 1999) et dans ce cas les différentes races du champignon ont pu être génétiquement différenciées à l'aide de marqueurs AFLP (Braithwaite et al. 2004). Cette situation rapprocherait notre cas de modèles plus classiques permettant d'introduire la spécificité dans l'analyse génétique de la résistance, voire d'entreprendre la cartographie génétique des facteurs d'avirulence chez le champignon à l'instar de *Ustilago hordei* (Linning et al., 2004). Nous espérons que les marqueurs que nous avons développés pourront être rapidement utilisés pour tester les souches caractérisées à Taiwan.

Les cultivars modernes de canne à sucre sont issus de l'hybridation interspécifique entre *S. officinarum* (espèce domestiquée riche en sucre) et *S. spontaneum* (espèce sauvage). Il semble que *S. officinarum* originaire de Papouasie Nouvelle-Guinée n'ait jamais été en contact avec le pathogène. Par contre, le centre d'origine de *S. spontaneum* recouvre celui du champignon. L'hypothèse d'une co-évolution entre *Ustilago scitaminea* et *S. spontaneum* peut donc être formulée. Les ressources génétiques de cette espèce devraient donc être exploitées pour y rechercher de nouvelles sources de résistance. D'après Srinivasan et Alexander (1971) les clones de *S. spontaneum* qui ont fondé les populations modernes de cultivars de canne à sucre étaient sensibles au charbon, mais il existe de nombreux clones résistants au sein de cette espèce.

Cartographie génétique dans le croisement biparental ‘R570’ x ‘MQ 76/53’

Nouvelles connaissances apportées

- La cartographie de nouveaux marqueurs locus spécifiques (microsatellites découverts dans les banques EST (Pinto et al. 2004), sondes RFLPs) ainsi que la mise en évidence de nouveaux appariements préférentiels entre chromosomes ont permis de progresser dans la connaissance des relations d'homo(éo)logie entre groupes de co-ségrégation de la carte génétique du cultivar ‘R 570’ depuis les travaux de Hoarau et al. (2001) et Rossi et al. (2003). La carte compte maintenant 90 groupes de co-ségrégation affectés à 7 groupes d'homologie.
- Une nouvelle carte génétique a été construite pour le clone ‘MQ 76/53’, constituée de 105 groupes de co-ségrégations regroupant 535 marqueurs simple dose.
- Un gène majeur responsable de la couleur rouge des tiges, hérité du parent ‘MQ 76/53’, a été identifié et localisé à 6.5 cM d'un marqueur AFLP.
- Un nouveau gène majeur de résistance à la rouille, distinct du gène Bru1 de ‘R 570’ cartographié sur le groupe de coségrégation 1a du groupe d'homologieVII (Asnaghi et al, 2004), a été identifié chez le parent ‘MQ 76/53’. Il a été localisé sur le groupe de coségrégation 3 du groupe d'homologie VIII à 23.1 cM d'un marqueur AFLP. Ce gène de résistance à la rouille appartient à un groupe d'homologie (VIII) différent de celui du gène Bru1 (VII). Ce gène correspond donc vraisemblablement à un nouveau locus. Les résultats apportés par le génotypage de la population de 73 cultivars (article3 de la thèse), montrent que l'ensemble des marqueurs du groupe de coségrégation portant le gène sont absents ou extrêmement rares dans le reste de la population. Etant donné que ‘MQ 76/53’ est issu du croisement entre un cultivar et un clone *S. spontaneum* ‘SES 528’, il est fort probable que ce gène putatif soit hérité de ‘SES 528’ et constitue une nouvelle source de résistance intéressante à introgresser dans le matériel génétique utilisé en sélection.
- La cartographie génétique dans un croisement bi-parental présente des contraintes particulières. Un grand nombre de marqueurs qui sont «simple dose» chez chacun des parents mais en commun entre les deux parents du croisement ne sont pas exploitables. Cette proportion de marqueurs communs peut être très importante. La comparaison des marqueurs AFLP simple-dose cartographiés chez ‘R 570’ avec les marqueurs AFLP révélés dans chacun des 73 autres cultivars de la population (cf. étude du DL de l'article 3) a montré que 40% à 65% des marqueurs étaient communs. C'est un paramètre à prendre en compte avant de choisir une combinaison parentale pour réaliser un travail de cartographie génétique.

Quelle suite à donner ?

L'organisation du génome des cultivars modernes de canne à sucre est encore loin d'être totalement élucidée. Ainsi, alors que le nombre de chromosomes de base de *S. officinarum* est de 10 et celui de *S. spontaneum* est de 8 (D'Hont et al. 1998), seulement 7 groupes d'homologie ont pu être établis pour la carte de 'R 570', dont l'un n'est constitué que de deux chromosomes. Une telle situation peut être expliquée par i) une couverture du génome avec des marqueurs locus spécifiques, insuffisante pour identifier tous les groupes d'homologie (il reste en effet une quarantaine de petits groupes de co-ségrégations qui ne sont pas affectés à un groupe d'homologie) ii) la duplication de certains marqueurs locus spécifiques, qui entraîne des erreurs lors du regroupement des groupes de co-ségrégations en groupe d'homologie iii) l'existence de différences structurelles importantes entre les chromosomes des deux espèces qui compliquent la définition du canevas des groupes d'homologie. On ne pourra répondre sur ce dernier point que si des travaux de cartographie génétique comparée poussés, ou de cartographie cytogénétique par BAC-FISH, sont entrepris sur les deux espèces ancestrales de la canne à sucre. Dans un premier temps, un effort supplémentaire de cartographie du cultivar 'R 570' doit être fait en utilisant d'autres marqueurs locus spécifiques, à commencer par les 120 microsatellites disponibles développés dans le cadre du consortium international pour les microsatellites de la canne à sucre (Cordeiro et al. 2000). Cela constituerait aussi une bonne base de comparaison avec les autres cartes développées par d'autres équipes, en particulier en Australie (Aitken et al. 2005).

Le nouveau gène de résistance putatif à la rouille identifié chez 'MQ 76/53' appartient au groupe d'homologie VIII, dans lequel trois clusters de RGA ont été mis en évidence (Rossi et al. 2003). Ces analogues de gènes de résistance pourraient être utilisés comme marqueurs candidats pour vérifier s'ils se cartographient à proximité du gène de résistance et en préciser plus rapidement la position (Pflieger et al. 2001). Le contexte du croisement bi-parental R 570 x MQ 76/53 n'est pas le plus favorable pour cartographier ce nouveau gène compte tenu de la ségrégation de l'autre allèle de résistance hérité de 'R 570'. Néanmoins, la population disponible en collection à la Réunion est de 300 clones dont seuls 198 ont été utilisés pour le travail de cartographie présenté ici. Il est donc possible d'augmenter la taille de la population utile pour cartographier ce nouveau gène putatif jusqu'à 150 individus à peu près, une moitié des 300 individus portant le gène Bru1 et devant être éliminés de l'analyse.

Structure du déséquilibre de liaison chez la canne à sucre

Nouvelles connaissances apportées

- Le déséquilibre de liaison chez la canne à sucre est fort puisque des associations significatives entre marqueurs peuvent être détectées jusqu'à des distances de plus de 20 cM. Cela confirme la situation déjà décrite par Jannoo et al. (1999).
- Compte tenu du grand nombre de tests statistiques réalisés il est nécessaire de travailler avec des seuils statistiques «stringents» de façon à ne pas être «noyé» par un grand nombre d'associations déclarées vrais par le test alors qu'elles impliquent des marqueurs qui ne sont pas liés physiquement sur un même haplotype.
- L'intensité du déséquilibre de liaison chute fortement lorsque la distance génétique entre les marqueurs est supérieure à 5 cM.
- Les marqueurs détectés en déséquilibre de liaison appartiennent à une gamme de fréquence centrée autour de 35%. Les marqueurs d'origine *S. spontaneum* ont une gamme de fréquence autour de ces valeurs (plutôt 40-50%) en revanche les marqueurs d'origine *S. officinarum* ont une gamme de fréquence beaucoup plus étendue avec un grand nombre de marqueurs très fréquents (jusqu'à 100%).
- Le déséquilibre de liaison existe entre marqueurs d'origine *S. officinarum* comme entre marqueurs d'origine *S. spontaneum* même si les associations entre marqueurs *S. spontaneum* sont sur-représentées par rapport à la proportion de marqueurs *S. spontaneum* dans l'ensemble des marqueurs.

Quelle suite à donner ?

La structure du déséquilibre de liaison chez la canne à sucre permet d'envisager des études d'associations à l'échelle du génome. Cependant, l'exploitation du déséquilibre de liaison chez la canne à sucre doit tenir compte d'un certain nombre de contraintes liées à l'histoire de la sélection. Un nombre restreint de croisements interspécifiques, impliquant un nombre limité d'ancêtres chez *S. officinarum* et *S. spontaneum*, ont fondé les populations de cultivars modernes. Cet effet de fondation a été générateur de déséquilibre de liaison. Dans un deuxième temps, un nombre très limité d'hybrides interspécifiques très performants, issus des

travaux de nobilisation du début du 20^{ème} siècle, ont été massivement utilisés comme géniteurs dans l'ensemble des programmes de sélection à travers le monde. Par conséquent, les haplotypes les plus fréquents et ceux qui ont la plus grande probabilité d'être détectés sont vraisemblablement hérités de ces goulots d'étranglement génétique. Donc pour avoir une chance d'exploiter la plus grande diversité possible des haplotypes, y compris les haplotypes rares, il sera nécessaire de travailler sur des populations de grands effectifs dont on contrôlera dans la mesure du possible l'histoire généalogique (dans le but d'éviter la sur-représentation des haplotypes issus de ces goulets génétiques). Par ailleurs, il est apparu clairement que le nombre de marqueurs utilisé, même si il peut paraître important (1626), était toutefois loin d'être suffisant pour couvrir l'ensemble de la diversité des haplotypes de la population. Pour cela il faudrait placer au minimum un marqueur tous les 5 cM. Cela nécessiterait l'utilisation d'un minimum de 300 marqueurs locus spécifiques (à condition qu'ils soient choisis en fonction de leur distance respective sur les cartes, or un tel niveau d'information cartographique n'est pas encore disponible chez la canne à sucre) ou plusieurs milliers de marqueurs de type AFLP (6000, 8000 voir plus...). L'ajustement du nombre de marqueurs peut se faire en cours d'étude en fonction des résultats obtenus.

Un projet impliquant un plus grand nombre de cultivars est déjà en cours dans le cadre d'un projet financé par le consortium international pour les biotechnologies chez la canne à sucre (ICSB) en partenariat entre le CIRAD et la SASRI (Afrique du sud). Ce projet implique le génotypage par AFLP de 180 génotypes dont 20 cultivars ancestraux et vise à identifier les haplotypes impliqués dans des caractères complexes comme le rendement ou la richesse en sucre. A la Réunion, un projet de même ampleur est prêt à démarrer, dont l'objectif serait d'identifier des haplotypes conférant la résistance aux maladies et aux ravageurs de la canne à sucre dans une population de cultivars modernes utilisés comme géniteurs dans le programme de création variétale du CERF (organisme obtenteur des variétés de canne à sucre à la Réunion). Dans le contexte réunionnais, il est en effet possible d'étudier dans de très bonnes conditions des maladies comme la rouille brune, le yellow spot et un ravageur, le foreur ponctué. Par ailleurs, l'ensemble des maladies les plus importantes de la canne à sucre au plan mondial est présent à la Réunion et peut également y être étudié: le charbon, l'échaudure des feuilles, la gombose ou le syndrome de la feuille jaune. Si ce projet est mis en œuvre, il faudra s'assurer de travailler avec la même gamme de témoin utilisée pour le projet SASRI-CIRAD et il serait judicieux d'inclure de la même manière les 20 géniteurs ancestraux qui ont été retenus par ce projet. C'est une perspective intéressante pour la Réunion car elle peut

fédérer les intérêts des sélectionneurs, qui bénéficieront d'une meilleure caractérisation de leurs ressources génétiques pour la résistance aux maladies, et ceux des généticiens, qui pourront valoriser ces informations grâce au travail de génotypage massif qui devra être entrepris en parallèle.

Nous disposons, avec le gène de résistance Bru1 du cultivar 'R 570' (Asnaghi et al. 2004), d'un modèle intéressant pour valider l'approche «Déséquilibre de liaison» comme outil permettant d'identifier les haplotypes associés à des caractéristiques agronomiques supérieures chez la canne à sucre. En effet, dans le cadre du travail en cours visant au clonage de ce gène, une carte génétique fine de la zone portant le gène a été réalisée et la construction d'un contig de BAC canne à sucre couvrant la zone cible (cartographie physique) est en cours (Lecunff et al. 2004). Un grand nombre de marqueurs est donc disponible autour du gène de résistance (en attendant d'avoir accès à la séquence du gène lui-même) pour entreprendre une étude ciblée du déséquilibre de liaison. Cette étude a débuté, basée sur un travail important mené sur le terrain à la Réunion pour caractériser la résistance à la rouille de plus de 120 cultivars de canne à sucre utilisés dans le programme de création variétale du CERF. Cette étude permettra de savoir si la résistance à la rouille (dans le contexte de la Réunion) est associée à un haplotype unique (celui correspondant au gène Bru1) ou au contraire si des haplotypes distincts peuvent conférer la résistance indiquant la présence d'autres gènes de résistance dans la population étudiée. Par ailleurs, en utilisant des marqueurs de plus en plus éloignés du gène, on devrait pouvoir préciser la distance entre marqueurs et gène de résistance à partir de laquelle le déséquilibre de liaison s'estompe et ne permet plus d'identifier le gène dans une population non structurée de cultivars.

Il est nécessaire d'insister sur un point de détail qui a son importance: l'intérêt qu'il y a de pouvoir s'appuyer sur des données généalogiques fiables lorsque l'on doit faire des interprétations génétiques de résultats de recherche (notamment lorsque l'on s'intéresse au déséquilibre de liaison) ou de sélection. Maintenant que les techniques de génotypage sont largement accessibles, la biologie moléculaire devrait au moins être mise à contribution pour la mise en place de «passeports moléculaires» pour les variétés ainsi que pour la vérification systématique des descendants (père et mère) des cultivars élites qui sont utilisés en croisements par les sélectionneurs.

Etude du déterminisme génétique de la résistance au charbon de la canne à sucre et des autres caractères étudiés concomitamment.

Nouvelles connaissances apportées

- Un certain nombre de QTL ont été mis en évidence pour la résistance au charbon. Ils sont caractérisés par des effets faibles et par une instabilité entre les différents environnements.
- Deux QTL de résistance au charbon ont aussi été détectés par l'approche déséquilibre de liaison.

Quelle suite à donner ?

La couverture de la carte du cultivar 'R 570' est probablement trop faible pour avoir pu détecter l'ensemble des zones génomiques impliquées dans la variation phénotypique observée. D'autre part, les conditions d'évaluation de la maladie et/ou le niveau de résistance qui s'exprime dans le croisement n'ont peut-être pas permis de révéler une variabilité suffisante pour la résistance au charbon permettant une détection statistique efficace des QTL impliqués dans la résistance. Il pourrait donc être utile de poursuivre le projet i) en complétant le génotypage de la descendance ii) en insistant sur la caractérisation de la résistance du croisement bi-parental étudié en menant des essais en serre en période cyclonique plus chaude (mais avec le risque inhérent à la saison) dans le but d'augmenter l'expression de la maladie et décaler les ségrégations observées. De la même manière, on peut envisager de mettre au champ des essais inoculés par piqûre dans lesquels des boutures 1 œil seraient suivies individuellement (ce type d'essai suppose un effort d'entretien des parcelles très supérieur à une conduite normale). Une autre solution est de rechercher d'autres sources de résistance (le projet DL évoqué dans le paragraphe précédent devrait être très intéressant à cet égard) et de créer de nouvelles descendances dont il faudrait vérifier la réponse à la maladie avant d'engager le génotypage.

Parallèlement à l'évaluation de la résistance au charbon de la canne à sucre, plusieurs paramètres agronomiques ont été mesurés sur les descendants (brix, nombre de tiges

usinables, diamètre des tiges et floraison, pour lesquels les résultats des détection de QTL sont présentés en annexe 2). Le taux de floraison est apparu particulièrement intéressant à étudier dans ce croisement car c'est un paramètre hautement héritable et la plupart des QTL détectés pour ce caractère se sont révélés répétables d'une année sur l'autre (à l'exception du cycle de culture en vierge). Il semble donc qu'il soit possible de bien appréhender le contrôle génétique de ce caractère. Pour cela, il faudrait affiner l'évaluation de la floraison en évaluant la précocité de floraison des clones en plus de l'intensité de floraison. Il faudrait aussi continuer à saturer les cartes génétiques des deux parents du croisement. Dans ce cadre là, il serait utile de cartographier un certain nombre de gènes candidats impliqués dans le contrôle photopériodique de la floraison et la régulation des rythmes circadiens chez d'autres Poacées. Il s'agit notamment des gènes Hd1 (Yano et al. 2000), Hd3a (Kojima et al. 2002) et Hd6 (Takashi et al. 2004) du riz, respectivement similaires aux gènes de la famille des gènes contans (CO), flowering locus T (FT), sous-unité alpha casein kinase 2 (CK2) de *Arabidopsis* (Hayama et Coupland 2004; Griffiths et al. 2003). Il est intéressant de souligner ici que les QTL (relativement «forts» pour la canne à sucre) détectés pour la floraison sur les chromosomes IV1 et IV2 (voir annexe 2) co-localisent avec des QTL (relativement «forts» pour la canne à sucre) détectés pour le nombre d'entre-nœuds (Offmann 2000).

Perspectives pour la création variétale

La mise en évidence des déterminants mendéliens responsables de caractères agronomiques complexes par l'approche QTL s'avère particulièrement délicate chez la canne à sucre puisque jusqu'à 10 ou 12 formes alléliques différentes peuvent co-exister à un même locus du fait de la polyplioïdie. Les effets individuels des différents allèles sont dilués et donc plus difficiles à détecter. De plus, dans un tel contexte génétique on s'attend à un réseau d'interactions épistatiques inextricable. La mise en pratique de la détection de QTL pour des caractères quantitatifs tels que les composantes du rendement (Hoarau et al. 2002), la richesse en sucre (Reffay 2003), la résistance à l'échaudure (Offmann, 2000) ou encore la résistance au charbon dans ce travail de thèse sont autant de tentatives qui révèlent la difficulté de l'exercice chez la canne à sucre. D'une manière générale de nombreux QTL à effets faibles caractérisés par leur instabilité entre différents environnements sont mis en évidence. Néanmoins, il est possible, même chez la canne à sucre, de mettre en évidence les gènes responsables de caractères à déterminisme génétique simple. On l'a vu pour la résistance à la

rouille (Daugrois et al., 1996 et article 2 de ce travail), la couleur des tiges (article 2 de ce travail) ou encore la résistance au yellow spot pour laquelle un QTL majeur a été mis en évidence (Al Janabi et al. 2005). Il apparaît donc nécessaire de continuer les efforts visant à construire une carte de référence d'un cultivar moderne. Cet outil permettra d'optimiser le choix des marqueurs pour tout nouveau travail de cartographie et de détection de QTL qui pourrait s'engager. Chez la canne à sucre, une stratégie QTL ne peut être efficace que si elle est basée sur l'évaluation préalable d'un grand nombre de croisements permettant de faire un choix opportuniste des croisements les plus intéressants en privilégiant les ségrégations qui autorisent des hypothèses fortes quant à un déterminisme génétique simple du caractère observé. Il est évident qu'une telle stratégie gagne à être mise en synergie avec un programme de sélection fonctionnant sur une base familiale. L'utilisation du déséquilibre de liaison s'avère plus séduisante, au moins potentiellement, parce qu'elle permet de «balayer» l'ensemble des ressources génétiques disponibles et parce que les connaissances acquises sur les ressources génétiques (le génotypage moléculaire en particulier) seront capitalisables dans le temps et pourront être mutualisées entre différents centre de recherche. La meilleure stratégie à développer est donc à la fois de réaliser un investissement de fond sur le DL et des investissements opportunistes pour les approches QTL. Mais si l'on regarde à plus long terme, c'est dans les ressources génétiques sauvages que se situent les potentiels de progrès pour l'amélioration variétale de la canne à sucre. Les ressources génétiques de *S. spontaneum* sont très vastes et ont été largement sous-exploitées dans les ressources génétiques des cultivars modernes. C'est sans doute vrai aussi dans une moindre mesure pour *S. robustum*. Toutefois, c'est à travers la compréhension du processus de domestication de *S. officinarum* à partir de *S. robustum* que l'on pourra peut-être mettre en évidence les gènes impliqués dans l'accumulation du saccharose.

Vers une exploitation optimisée du DL chez la canne à sucre:

Cette approche est particulièrement prometteuse chez la canne à sucre pour compléter les approches QTL qui s'avèrent très délicates dans le contexte polyploïde de la plante. Pour atteindre une efficacité maximum, l'ensemble des informations moléculaires, concernant les ressources génétiques en collection, produites dans les différents centres de recherche devrait pouvoir être capitalisé dans des bases de données au bénéfice de tous. Bien entendu, la mise en œuvre d'un projet de ce type butte sur les problèmes de standardisation des méthodes de

génotypage pour des travaux réalisés entre différents laboratoires. Néanmoins, les progrès technologiques devraient permettre rapidement de passer des méthodes analytiques de séparation des polymorphismes de l'ADN à des méthodes d'hybridation à haut débit sur puces à ADN. De ce point de vue, la technique dite des «Diversity arrays» (DArT) apparaît prometteuse en permettant de détecter du polymorphisme à plusieurs centaines de locus simultanément (Wenzl et al. 2004). Ce type de technologie devrait faciliter la construction de vastes bases de données contenant des informations massives sur le génotype moléculaire de nombreuses ressources génétiques. Il sera donc possible de corrélérer ces génotypes avec les phénotypes évalués dans un grand nombre d'environnements différents pour les maladies (différents pathogènes, différentes races d'un agent pathogène, différentes pressions d'un agent pathogène...), pour les conditions climatiques ou les conditions agronomiques. L'objectif sera d'identifier les haplotypes conférant une valeur agronomique supérieure.

Exploration des ressources génétiques sauvages (*S. spontaneum*):

Une grande part de la diversité génétique disponible de *S. officinarum* semble avoir été exploitée dans les ressources génétiques de cultivars modernes (Jannoo et al, 1999). En revanche, l'espèce *S. spontaneum* qui présente une très grande diversité et qui contribue d'ailleurs fortement à la diversité des cultivars modernes n'a été exploitée que très partiellement et constitue donc une source intéressante d'allèles utiles sur un plan agronomique (Lu et al. 1994). Dans notre étude sur le déséquilibre de liaison, l'absence dans les cultivars de la population d'une grande proportion des haplotypes (constituant parfois des groupes de coségrégation entiers) de 'MQ 76/53' ayant vraisemblablement une origine *S. spontaneum* en est la démonstration. Ainsi, l'identification d'un nouveau gène de résistance à la rouille chez 'MQ 76/53', localisé sur un groupe de coségrégation constitué de marqueurs absents ou très rares dans le reste de la population étudiée met en évidence l'intérêt des ressources génétiques sauvages comme réservoirs d'allèles intéressants pour la sélection. L'espèce *S. spontaneum* peut coloniser des régions extrêmement contrastées de par leurs températures, leur taux d'humidité ou leurs types de sols. Elle possède une aire de répartition très vaste (figure 1 de l'introduction). Cette espèce peut donc constituer une source de gènes de résistance ou de tolérance aux stress abiotiques (résistance à la sécheresse, au sel, à la toxicité aluminique ...) ou à des stress biotiques (maladies et ravageurs). Pour le point particulier que constitue la toxicité aluminique, il est intéressant de remarquer que des gènes de tolérance ont été cartographiés chez un certain nombre de Poacées dont le sorgho

(Magalhaes et al. 2004). *S. spontaneum* pourrait aussi constituer une source très intéressante de résistance au charbon de la canne à sucre pour les raisons discutées dans le paragraphe ci-avant «Diversité génétique du charbon». Les potentialités génétiques contenues dans les collections sauvages sont aussi réelles pour des caractères complexes tel que le rendement. La mise en évidence de ces potentialités suppose de ne pas s'intéresser uniquement aux phénotypes mais de rechercher les gènes intéressants à l'aide de la cartographie génétique (Tanksley et McCouch 1997). La première étape est de réaliser une large étude de la diversité moléculaire de l'espèce pour identifier les réservoirs d'allèles les plus riches et les plus «nouveaux».

Domestication de *S. officinarum* à partir de *S. robustum*: Quels sont les gènes qui ont été sélectionnés au cours de ce processus?

La domestication de la canne à sucre s'est déroulée en Papouasie Nouvelle Guinée, il y a plus de 2500 ans. Des formes sucrées ont été progressivement sélectionnées, à partir de *S. robustum*, pour être mâchées. Au cours de ce processus, les allèles favorables à cette caractéristique ont été accumulés. Une des conséquences de cette sélection peut être la réduction du niveau de polymorphisme dans les zones, plus ou moins grandes en fonction de l'effet «d'hitch-hiking», entourant les locus concernés par la domestication (cas de la zone de régulation du gène tb1 «teosinte branched1» chez le maïs (Wang et al. 1999)). Donc il devrait être possible de localiser les gènes contrôlant des phénotypes d'intérêt en scannant le génome pour y rechercher la signature laissée sur le polymorphisme par la sélection. Chez le maïs, par exemple, plus de 500 microsatellites issus de banques EST ont été étudiés dans ce but et 15 d'entre eux qui ont montré des traces laissées par la sélection sont des candidats pour être des gènes importants au plan agronomique (Vigouroux et al., 2002). Une telle étude est envisageable chez la canne à sucre, qui comparerait une population de clones de *S. robustum* avec une population de clones de *S. officinarum*. Il serait intéressant dans ce cadre de développer massivement des marqueurs microsatellites à partir de la banque EST de la canne à sucre (SUCEST). Plus de 2000 clusters ont été identifiés qui contiennent des séquences de type microsatellites (Pinto et al. 2004). Cela bénéficierait aussi aux autres approches (cartographie génétique et déséquilibre de liaison) pour lesquelles on dispose pour le moment de moins de 200 microsatellites. Une autre approche du problème pourrait être de mettre en place un backcross avancé de *S. officinarum* sur le parent récurrent *S. robustum* (Tanksley et

Nelson 1996) pour tenter de mettre en évidence les QTL impliqués dans les mécanismes d'accumulation du sucre.

Polyplioïdie, sélection et marquage

Les perspectives semblent nombreuses en matière d'exploitation de la diversité des ressources génétiques. Les généticiens et sélectionneurs de la canne à sucre n'ont pourtant pas ménagé leurs efforts pour tenter d'élargir la diversité des cultivars (voir Heinz 1987). La structure même des variétés modernes illustre l'un des exemples de réussite d'un brassage large les plus spectaculaires de l'amélioration variétale moderne. Cependant, vu la quantité de matériel génétique issu de l'espèce sauvage *S. spontaneum*, il s'agit plus d'«incorporation» que d'«introgression» génétique (Simmonds 1993). Les efforts de recours à des sources diverses (élargissement de la base génétique ou «base-broadening» chez les anglo-saxons) se sont ralentis pour des raisons probablement liées à la polyplioïdie. En effet, la polyplioïdie a tendance à uniformiser les comportements des combinaisons interspécifiques faisant intervenir différents parents sauvages et masque ainsi les apports spécifiques de l'un ou l'autre d'entre eux. Les marqueurs moléculaires devraient permettre de mieux comprendre la structure des espèces ancêtres, d'identifier des compartiments sous-utilisés jusqu'à présent et de reprendre des tentatives d'élargissement plus ciblées.

La polyplioïdie, grand frein à l'analyse génétique, ralentit aussi l'impact de la sélection sur la diversité génétique des populations d'amélioration. Les matériaux présents et évalués dans les programmes de sélection sont probablement suffisamment hétérozygotes pour présenter des ségrégations sur la quasi-totalité des régions du génome et permettre l'analyse génétique de la majorité des caractères agronomiques importants. Combiné au fort déséquilibre de liaison, cela permettra de valoriser les populations existantes (en cours d'évaluation ou déjà évaluées) pour marquer les gènes impliqués dans les caractères agronomiques importants. Ainsi, plus pour que la plupart des autres espèces, les progrès des technologies de génotypage auront probablement un impact positif pour l'amélioration génétique de la canne.

REFERENCES BIBLIOGRAPHIQUES

REFERENCES BIBLIOGRAPHIQUES

- Abdennadher M., Mills D. 2000. Telomere-associated RFLPs and electrophoretic karyotyping reveal lineage relationships among race-specific strains of *Ustilago hordei*. Curr. Genet. 38, 141-147.
- Aitken K.S., Jackson P.A., McIntyre C.L. 2005. A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. Theor. Appl. Genet. in press.
- Alexander K.C., Srinivasan K.V. 1966. Sexuality in *Ustilago Scitaminea* Syd. Curr. Sci. 23, 603-604.
- Alix K., Baurens F.-C., Paulet F., Glaszmann J.-C., D'Hont A. 1998. Isolation and characterization of a satellite DNA family in the *Saccharum* complex. Genome 41:854–864.
- Alix K., Paulet F., Glaszmann J.-C., D'Hont A. 1999. Inter- Alu-like species-specific sequences in the *Saccharum* complex. Theor. Appl. Genet. 99: 962–968.
- Al-Janabi S.M., McClelland M., Petersen C. 1994. Phylogenetic analysis of organellar DNA sequences in the Andropogoneae: Saccharinae. Theor. Appl. Genet. 88: 933-944.
- Al-Janabi S.M., Honeycutt R.J., McClelland M., Sobral B.W.S. 1993. A genetic linkage map of *Saccharum spontaneum* L. ‘SES 208’. Genetics 134:1249–1260.
- Al-Janabi S.M., Parmessur Y., Kross H., Dhayan S., Saumtally S., Dookun-Saumtally A. 2005. Characterization of quantitative trait loci (QTLs) contributing to field resistance to yellow spot disease in sugarcane. Proc. Int. Soc. Sugarcane Technol. 25:
- Amire O.A., Trione E.J., Schmitt R.A. 1982. Characterization of pathogenic races of the sugarcane smut fungus by neutron activation analysis. Journal of Radioanalytical Chemistry 75, 195-203.
- Antoine R., 1961. Smut. In: Martin J.P., Abbott E.V., Hughes C.G. (Eds), Sugar Cane Disease of the World. Elsevier Publishing Co., Amsterdam, pp 327-345.
- Arceneaux G. 1965. Cultivated sugarcanes of the world and their botanical derivation. Proc. Int. Soc. Sugarcane Technol. 12: 844-854.
- Asnaghi C., Paulet F., Kaye C., Grivet L., Glaszmann J.C., D'Hont A. 2000. Application of synteny across the Poaceae to determine the map location of a rust resistance gene of sugarcane. Theor. Appl. Genet. 10:962-969.
- Asnaghi C., Roques D., Ruffel S., Kaye C., Hoarau J.Y., Télismart H., Girard J.C., Raboin L.M., Risterucci A.M., Grivet L., D'Hont A. 2004. Targeted mapping of a sugarcane rust resistance gene (*Brul*) using bulked segregant analysis and AFLP markers. Theor. Appl. Genet. 108:759-764.
- Bakkeren G., Kronstad J.W. 1994. Linkage of mating type loci distinguishes bipolar from tetrapolar mating in Basidiomycetous smut fungi. Proc. Natl. Acad. Sci. USA 91, 7085-7089.
- Balance M.C., Ramos, N.M., Lopez J.M., 1995. Diallelic analysis of resistance of five sugarcane parents to *Ustilago scitaminea* Sydow. Proc. Int. Soc. Sugar cane technol: 22 : 340-345.

- Barat H. 1971. Etat Phytosanitaire des cannes à sucre de Côte d'Ivoire. Dangers résultant de la présence du charbon en Haute-Volta. Rapport IRAT.
- Bates J.F. 1975. Smut disease in Guiana. Sugarcane Pathologists' Newsletter 13/14, 8.
- Besse P., McIntyre C.L., Berding N. 1997. Characterisation of *Erianthus* sect. *Ripidium* and *Saccharum* germplasm (Andropogoneae -Saccharinae) using RFLP markers. Euphytica 93: 283-292.
- Bhat S.R., Gill B.S. 1985. The implication of 2n egg gametes in nobilisation and breeding of sugarcane. Euphytica 34:377-384.
- Billotte N., Lagoda P.J.L., Risterucci A.M., Baurens F.C. 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54 (4), 277-288..
- Bock K.R. 1964. Studies on sugar-cane smut (*Ustilago scitaminea*) in Kenya. Trans. Brit. Mycol. Soc. 47 (3): 403-417.
- Braithwaite K.S., Bakkeren G., Croft B.J., Brumbley S.M. 2004. Genetic variation in a worldwide collection of the sugarcane smut fungus *Ustilago scitaminea*. Proc. Aust. Soc. Sugar cane Technol. 26, abstract p 48.
- Brandes E.W., Sartoris G.B., Grassl C.O. 1939. Assembling and evaluating wild forms of sugarcane and closely related plants. Proc. Int. Soc. of Sugarcane Technol. 6: 128-154.
- Brandes E.W. 1958. Origin, classification and characteristics. In : Sugarcane (*Saccharum officinarum* L.), edited by E. Artschwager, and E.W. Brandes. USDA Agriculture Handbook, pp 1-35.
- Bremer G. 1930. The cytology of *Saccharum*. Proc. Int. Soc. Sugarcane Technol. 3: 408-415.
- Bremer G. 1961. Problems in breeding and cytology of sugarcane. I. A short history of sugarcane breeding - the original forms of *Saccharum*. Euphytica 10: 59-78.
- Brown J.K.M., Hovmöller M.S. 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science 297, 537-541.
- Burner D.M. 1991. Cytogenetic analyses of sugarcane relatives (Andropogoneae : Saccharinae). Euphytica 54: 125-133.
- Burner D.M., Legendre B.L. 1994. Cytogenetic and fertility characteristics of elite sugarcane clones. Sugar Cane 1: 6-10.
- Burner D.M., Legendre B.L. 1993. Chromosome transmission and meiotic stability of sugarcane (*Saccharum* spp.) hybrid derivatives. Crop Sci. 33:600-606.
- Burner D.M., 1993. Resistance of sugarcane relatives injected with *Ustilago scitaminea*. Plant disease. 77:1221-1223.
- Butterfield M.K. 2005. Molecular breeding of sugarcane using DNA markers and linkage disequilibrium mapping. Proc. Int. Soc. Sugarcane Technol , 25 :545-548.
- Butterfield M.K., D'Hont A., Berding N. 2001. The sugarcane genome : a synthesis of current understanding, and lessons for breeding and biotechnology. Proc. South African Sugar Technol. Ass. 75:1-5.
- Byther R.S., Steiner G.W. 1974. Unusual smut symptoms on sugarcane in Hawaii. Plant Dis.Rep. 58: 401-405.

- Cardon L.R., J.I. Bell. 2001. Association study designs for complex diseases. *Nature Reviews Genetics* 2: 91-99.
- Chakraborty R., Jin L. 1993. A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and populations distances. In: Pena, S.D.J., Chakraborty, R., Epplen, J.T., Jeffreys, A.J. (Eds), *DNA fingerprinting: State of the science*. Birkhäuser Verlag, Basel, pp 153-175.
- Chao C.P., Hoy J.W., Saxton A.M., Martin F.A. 1990. Heritability of resistance and repeatability of clone reactions to sugarcane smut in Louisiana. *Phytopathology* 80, 622-626.
- Chen W.H., Huang Y.J., Shen I.S., Shih S.C. 1983. Utilization of *Miscanthus* germplasm in sugarcane breeding in Taiwan. *Proc. Int. Soc. Sugarcane Technol.* 18(2) : 641-648.
- Churchill G.A., Doerge R.W. 1994. Empirical threshold values for quantitative trait-mapping. *Genetics*, 138:963-971.
- Commonwealth Mycological Institute. 1977. Distribution Maps of Plant Diseases. N°79, 4th ed.
- Comstock J. C., Heinz D.J. 1977. A new race of culmicolous smut of sugarcane in Hawaii. *Sugarcane Pathologists' Newsletter* 19, 24-25.
- Comstock J.C. 2000. Smut. In: Rott, P., Bailey, R.A., Comstock, J.C., Croft, B.J., Salem Saumtally, A. (Eds), *A guide to sugarcane diseases*. CIRAD and ISSCT, pp. 181-185.
- Cordeiro G.M., Taylor G.O., Henry R.J. 2000. Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species, *Plant Science* 155(2): 161-168.
- Cox M.C., Hogarth D. M., Smith. G. R. 2000. Cane breeding and improvement. In D. M. Hogarth, and P. G. Allsopp (eds.). *Manual of Cane Growing*. Bureau of Sugar Experiment Stations, Brisbane, Queensland, Aust., pp. 91-108.
- D'Hont A. 2005. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenet. Genome Res.* 109:27-33.
- D'Hont A., Garsmeur O., Raboin L.M., Paulet F., Begum D., Wing R., Glaszmann J.C. 2001. Chromosome walking toward a major resistance gene for common rust of sugarcane. *Proc Int Soc Sugarcane Technol* 24: 315-317.
- D'hont A., Glaszmann J.C. 2001. Sugarcane genome analysis with molecular markers, a first decade of research. *Proc. Int. Soc. Sugarcane Technol.* 24: 556-559 .
- D'Hont A., Grivet L., Feldmann P., Rao S., Berding N., Glaszmann J.C. 1996. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol. Gen. Genet.* 250: 405-413.
- D'Hont A., Ison D., Alix K., Roux C., Glaszmann J.C. 1998. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome* 41 : 221-225.
- D'Hont A., Lu Y.H., Feldmann P., Glaszmann J.C. 1993. Cytoplasmic diversity in sugarcane revealed by heterologous probes. *Sugar Cane* 1: 12-15.

- D'Hont A., Lu Y.H., Gonzales de Leon D., Grivet L., Feldmann P., Lanaud C., Glaszmann J.C. 1994. A molecular approach to unravelling the genetics of sugarcane, a complex polyploid of the Andropogoneae. *Genome* 37: 222-230.
- D'Hont A., Paulet F., Glaszmann J.C. 2002. Oligoclonal interspecific origin of 'North Indian' and 'Chinese' sugarcanes. *Chromosome Research* 10:253-262.
- D'Hont A., Rao P.S., Feldmann P., Grivet L., Islam-Faridi N., Taylor P., Glaszmann J.C. 1995. Identification and characterisation of sugarcane intergeneric hybrids, *Saccharum officinarum* x *Erianthus arundinaceus*, with molecular markers and DNA in situ hybridisation. *Theor. Appl. Genet.* 91:320-326.
- Da Silva J., Honeycutt R.J., Burnquist W., Al-Janabi S.M., Sorrells M.E., Tanksley S.D., Sobral B.W.S. 1995. *Saccharum spontaneum* L. 'SES 208' genetic linkage map combining RFLP- and PCR-based markers. *Mol. Breed.* 1:165–179.
- Da Silva J.A.G., Burnquist W.L., Tanksley S.D. 1993. RFLP linkage map and genome analysis of *Saccharum spontaneum*. *Genome* 36:782–791.
- Daniels J. and Roach B.T. 1987. Taxonomy and evolution. In : Heinz D.J. *Sugarcane improvement through breeding*. Elsevier Press, Amsterdam, 7-84.
- Daugrois J.H., Grivet L., Roques D., Hoarau J.Y., Lombard H., Glaszmann J.C., D'Hont A. 1996. A putative major gene for rust resistance linked with an RFLP marker in sugarcane cultivar R570. *Theor. Appl. Genet.* 92:1059–1064.
- de Sornay P. 1936. Manuel de la canne à sucre à l'usage des chargés de cours et des élèves des grandes écoles coloniales . Port Louis , Ile Maurice, the general Printing&stationary Cy. Ltd., 330p.
- Delassus M., 1975. Rapport sur le charbon de la canne à sucre. Situation au Niger et méthodes de lutte à envisager. Rapport de l'IRAT.
- Delmotte F., Bucheli E., Shykoff J.A. 1999. Host and parasite population structure in a natural plant-pathogen system. *Heredity* 82, 300-308.
- Dufour P., Deu M., Grivet L., D'Hont A., Paulet F., Bouet A., Lanaud C., Glaszmann J.C., Hamon P. 1997. Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid. *Theor. Appl. Genet.* 94:409-418.
- Dufour P., Grivet L., D'Hont A., Deu M., Trouche G., Glaszmann J.C., Hamon P. 1996. Comparative genetic mapping between duplicated segments on maize chromosomes 3 and 8 and homoeologous regions in sorghum and sugarcane. *Theor. Appl. Genet.* 92: 1024-1030.
- F.A.O. 2004. FAOSTAT Database Gateway. <http://apps.fao.org/>
- Farnir F., Coppieters W., Arranz J.-J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D., Georges M. 2000. Extensive Genome-wide Linkage Disequilibrium in Cattle. *Genome Research* 10: 220-227.
- Fauconnier R., Bassereau D. 1970. La canne à sucre. Maisonneuve et Larose, Paris, 441p.
- Ferreira, S.A. Comstock, J.C. 1989. Smut. In: Ricaud, C., Egan, B.T., Gillaspie, A.G., Hughes, C.G. (Ed.), Diseases of sugarcane. Elsevier, Amsterdam, pp. 211-229.
- Flint-Garcia S.A., Thornsberry J. M., Buckler IV E.S. 2003. Structure of Linkage Disequilibrium in plants. *Annu. Rev. Plant Biol.* 54: 357-374.

- Flores S.C. 1978. Smut in Florida and Belize - A serious threat to the mexican sugar industry. Sugarcane Pathologists' Newsletter 21, 1-2.
- Flores S.C. 1980. Sugarcane smut in Mexico. Sugarcane Pathologists' Newsletter 24: 8-10.
- Gallacher D.J. 1997. Evaluation of sugarcane morphological descriptors using variance components analysis. Aust. J. Agric. Res. 48:769-773.
- Gallais A. 1990. Théorie de la sélection en amélioration des plantes. Masson editions, Paris. 588p.
- Garris A.J., McCouch S.R., Kresovich S. 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus in rice (*Oryza sativa* L.). Genetics 165: 759-769.
- Genetix version 4.0. Logiciel sous Windows TM pour la génétique des populations . Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France.
- Girard J.C., Payet J. 1997. La maîtrise des maladies de la canne à sucre à la Réunion : un effort continu de lutte depuis plus de 60 ans. Proceedings du 4^{ème} congrès international de l'ARTAS et des 2^{ème} rencontres en langue française de l'AFCAS, pp 249-259.
- Glaszmann J.-C., Dufour P., Grivet L., D'Hont A., Deu M., Paulet F., Hamon P. 1997. Comparative genome analysis between several tropical grasses. Euphytica 96: 13-21.
- Glaszmann J.-C., Jannoo N., Grivet L., D'Hont A. 2003. Sugarcane. In Genetic diversity of cultivated tropical plants. Hamon, P., Séguin, M., Perrier, X. and Glaszmann, J.-C. (Eds) CIRAD, Montpellier. pp 337-359.
- Grassl C.O. 1972. Taxonomy of *Saccharum* relatives: *Sclerostachya*, *Narenga* and *Erianthus*. Proc. Int. Soc. of Sugarcane Technol. 14: 240-248.
- Griffiths, S., Dunford R.P., Coupland G., Laurie. D.A. 2003. The evolution of constans-like gene families in barley, rice and *Arabidopsis*. Plant Physiology 131: 1855-1867.
- Grisham, M.P. 2001. An international project on sugarcane variability within sugarcane smut. Proc. Int. Soc. Sugar Cane Technol. 24 (2), 459-461.
- Grivet L., Arruda P. 2001. Sugarcane genomics: depicting the complex genome of an important tropical crop. Current Opinion in Plant Biology 5:122-127.
- Grivet L., D'Hont A., Roques D., Feldmann P., Lanaud C., Glaszmann J.C. 1996. RFLP mapping in cultivated sugarcane (*Saccharum spp.*) : genome organization in a highly polyploid and aneuploid interspecific hybrid. Genetics 142: 987-1000.
- Grivet L., Daniels C., Glaszmann J.C., D'Hont A. 2004. A review of recent molecular genetics evidence for sugarcane evolution and domestication. Ethnobotany Research and Applications 2:9-17.
- Grivet L., Glaszmann J.C., Vincentz M.A., da Silva F., Arruda P. 2003. ESTs as a source for sequence polymorphism discovery in sugarcane: example of the Adh genes. Theor. Appl. Genet. 106: 190-197.
- Guimaraes C.T., Sills G.R., Sobral B.W.S. 1997. Comparative mapping of Andropogoneae: *Saccharum* L. (sugarcane) and its relation to sorghum and maize. Proc. Natl. Acad. Sci. 94:14261-14266.

- Gupta P.K., Rustgi S., Kulwal P.L. 2005 Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Molecular Biology* 57: 461-485.
- Ha S., Moore P.H., Heinz D., Kato S., Ohmido N. and Fukui K. 1999. Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence hybridization and imaging methods. *Mol. Breed.* 39:1165-1173.
- Hansen M., Kraft T., Ganestam S., Säll T., Nilsson N.-O. 2001. Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. *Genet. Res. Camb.* 77: 61-66.
- Hayama R., Coupland G. 2004. The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice. *Plant Physiology* 135: 677-684.
- Heinz D.J. (editor). 1987. Sugarcane improvement through breeding. Elsevier, Amsterdam. 603 p.
- Heinze B.S., Thokoane L.N., Williams N.J., Barnes J.M., Rutherford R.S. 2001. The smut-sugarcane interaction as a model system for the integration of marker discovery and gene isolation. *Proc. South African Sugar Technol. Ass.* 75: 88-93.
- Henderson C. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31 : 423-447.
- Hirshhorn E. 1949. Un nuevo metodo de infección artificial con el carbon de la cana de azucar. *Revista de Investigaciones Agricolas.* III (4), 335-344.
- Hoarau J.Y., Payet J., Hellmann M. 2002. La lutte contre la maladie de la mosaïque à la Réunion. L'action agricole de la Station de Génétique de la Bretagne. In : Actes des Rencontres internationales pluridisciplinaires : «Perspectives de développement de la canne à sucre en milieu insulaire:approches technico-économiques, sociales, culturelles », du 2 au 5 octobre 2002, Muséum Stella Matutina, St-Leu, Réunion. (sous presse) .
- Hoarau J.Y., Offmann B., D'Hont A., Risterucci A.M., Roques D., Glaszmann J.C., Grivet L. 2001. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp). I. Genome mapping with AFLP markers. *Theor. Appl. Genet.* 103:84–97.
- Hoarau J.Y., Grivet L., Offmann B., Raboin L.M., Diorflar J.P., Payet J., Hellmann M., D'Hont A., Glaszmann J.-C. 2002. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp). II. Detection of QTLs for yield components. *Theor. Appl. Genet.* 105:1027-1037.
- Hoisington D: Laboratory protocols 1992. CIMMYT applied molecular genetics laboratory, CIMMYT, Mexico, D.F..
- Hood M.E., Antonovics J. 2000. Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity* 85, 231-241.
- Hughes C.G. 1964. Gumming disease. In : Martin J.P., Abbott E.V. Hughes C.G. (eds). *Sugar cane diseases of the world*. Elsevier, Amsterdam vol.I, 55-76.
- Ivandic V., Thomas W.T.B., Nevo E., Zhang Z., Forster B.P. 2003. Associations of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. *Plant Breeding* 122: 300-304.

- James G.L. 1976. Sugarcane smut infection in Guyana and Martinique. The Sugar Journal May, 17.
- James N.I. 1978. Sugarcane smut - An eminent threat to the U.S. mainland. Proc. Amer. Soc. Sugar Cane Technol. 7, 63-66.
- Jannoo N., Grivet L., Seguin M., Paulet F., Domaingue R., Rao P.S., Dookun A., D'Hont A., Glaszmann J.C. 1999a. Molecular investigation of the genetic base of sugarcane cultivars. *Theor. Appl. Genet.* 99: 171-184.
- Jannoo N., Grivet L., Dookun A., D'Hont A., Glaszmann J.C. 1999b. Linkage disequilibrium among modern sugarcane cultivars. *Theor. Appl. Genet.* 99: 1053-1060.
- Jannoo N., Grivet L., David J., D'Hont A., Glaszmann J.C. 2004. Differential chromosome pairing affinities at meiosis in polyploid sugarcane revealed by molecular markers. *Heredity* 93:460-467.
- Jordan D.R., Casu R.E., Besse P., Carroll B.C., Berding N., McIntyre C.L. 2004. Markers associated with stalk number and suckering in sugarcane collocate with tillering and rhizomatousness QTLs in sorghum. *Genome* 47:988-993.
- Kaltz O., Shykoff J.A. 1999. Selfing versus outcrossing propensity of the fungal pathogen *Microbotryum violaceum* across *Silene latifolia* host plants. *J. Evol. Biol.* 12, 340-349.
- Kojima S., Takahashi Y., Kobayashi Y., Monna L., Sasaki T., Araki T., Yano M. 2002. Hd3a, a rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. *Plant Cell Physiol.* 43(10):1096-1105.
- Kraakman A.T.W., Niks R.E., Van den Berg P.M.M.M., Stam P., Van Eeuwijk F.A. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168: 435-446.
- Kruglyak L. 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics* 22: 139-144.
- Lambat A.K., Chenulu V.V., Chona B.L., 1968. Morphological variation in the sugarcane smut *Ustilago scitaminea* Syd. *Mycopathologia and Mycologia Applicata* 36, 301-304.
- Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S.E., Newburg L. 1987. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174-181 .
- Lander E.S., Schork N.J. 1994. Genetic dissection of complex traits. *Science* 265: 2037-2048.
- Lecunff L., Garsmeur O., Raboin L.M., Grivet L., Glaszmann J.C., D'Hont A. 2004. Construction of a sugarcane BAC contig covering the target resistance gene Brul through exploitation of synteny with sorghum and rice. *Plant and Animal Genomes XII conference San Diego*.
- Lee C.S., Yuan C.H., Liang Y.G. 1999. Occurrence of a new pathogenic race of culmicolous smut of sugarcane in Taiwan. *Proc. Int. Soc. Sugar Cane Technol.* 23(2) : 406-407 .
- Lee-Lovick G. 1978. Smut of sugarcane - *Ustilago scitaminea*. Review of Plant Pathology of the Commonwealth Mycological Institute 57 (5), 181-187. .

- Leonards-Schippers C., Gieffers W., Schafer-Pregl R., Ritter E., Knapp S.J., Salamini F., Gebhardt C. 1994. Quantitative resistance to *Phytophthora infestans* in potato: a case study for mapping in an allogamous plant species. *Genetics* 137:67-77.
- Leu L.S. 1972. Culmicolous smut of sugar cane in Taiwan III. Germination and storage of teliospores, and compatibility of *Ustilago scitaminea* Sydow. Report of Tai^{wan} sugarcane experimental station 56, 37-45.
- Leu L.S. 1978. Culmicolous smut of sugar cane in Taiwan VI. New pathogenic strain obtained by artificial hybridization and further studies on compatibility of *Ustilago scitaminea* Sydow. *Ann. Phytopath. Soc. Japan* 44, 321-324.
- Leu L.S., Teng W.S. 1972. Pathogenic Strains of *Ustilago scitaminea* Sydow. Sugarcane Pathologists' Newsletter 8, 12-13.
- Leung H., Nelson R.J., Leach J.E. 1993. Population structure of plant pathogenic fungi and bacteria. *Adv. Plant Pathol.* 10, 157-205.
- Li H.W., Price S. 1967. Chromosome numbers of noble sugarcane clones. *Proc. Int. Soc. of Sugarcane Technol.* 12 : 884-886.
- Lima M.L.A., Garcia A.A.F., Oliveira K.M., Matsuoka S., Arizono H., de Souza C.L., de Souza A.P. 2002. Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*Saccharum* spp.). *Theor. Appl. Genet.* 104:30-38.
- Linning R., Lin D., Lee N., Abdennadher M., Gaudet D., Thomas P., Mills D., Kronstad J.W., Bakkeren G. 2004. Marker-based cloning of the region containing the UhAvr1 avirulence gene from the basidiomycete barley pathogen *Ustilago hordei*. *Genetics* 166, 99-111.
- Liu L.-J. 1981. Sugarcane smut in Puerto Rico. Sugarcane Pathologists' Newsletter 26, 52.
- Lloyd H.L. and Naidoo G. 1983. Chemical assay potentially suitable for determination of smut resistance of sugarcane cultivars. *Plant disease*. 67:1103-1105.
- Lu Y.H., D'Hont A., Walker D.I.T., Rao P.S., Feldamnn P. & Glaszmann J.C. 1994a. Relationships among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. *Euphytica* 78: 7-18.
- Lu Y.H., D'Hont A., Paulet F., Grivet L., Arnaud M. & Glaszmann J.C. 1994b. Molecular diversity and genome structure in modern sugarcane varieties. *Euphytica* 78:217-226.
- Machado G.R. 2001. Sugarcane variety notes – An international directory 7th revision, Piracicaba, Int. Soc. Sugar Cane Techn. Februray 2001, 132 p.
- Mackay T.F.C. 2001. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* 35: 303-339.
- Magalhaes J.V., Garvin D.F., Wang Y.-H., Sorrells M.E., klein P.E., Schaffert R.E., Li L., Kochian L.V. 2004. Comparative mapping of a major aluminium tolerance gene in sorghum and other species in the poaceae. *Genetics* 167: 1905-1914.
- Martinez-Espinoza A.D., Garcia-Pedrajas M.D., Gold S.E. 2002. The Ustilaginales as plant pests and model system. *Fungal Genet. Biol.* 35, 1-20.
- Mather K. 1957. The Measurement of Linkage in Heredity. Methuen & Co. LTD, London .
- McDonald B.A., Linde C. 2002. Pathogen population genetics, evolutionary potential , and durable resistance. *Annu. Rev. Phytopathol.* 40, 349-379.

- Mehta C.R., Patel N.R. 1983. A network algorithm for performing Fisher's exact test in rxc contingency tables. *J. Am. Statist. Assoc.* 78:427-434.
- Meyer J. 1989. Histoire du sucre. Les éditions Desjonquières, Paris, 335p.
- Michelmore R. 1995. Molecular approaches to manipulation of disease resistance genes. *Annu. Rev. Phytopathol.* 15: 393-427.
- Ming R., Del Monte T.A., Hernandez E., Moore P.H., Irvine J.E., Paterson A.H. 2002. Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genomes. *Genome* 45:794-803.
- Ming R., Liu S.C., Bowers J.E., Moore P.H., Irvine J.E., Paterson A.H. 2002. Construction of a *Saccharum* consensus genetic map from two interspecific crosses. *Crop Sci.* 42:570-583.
- Ming R., Liu S.-C., Lin Y.-R., Da Silva J., Wilson W., Braga D., Van Deynze A., Wenslaff T.F., Wu K.K., Moore P.H., Burnquist W., Sorrells M.E., Irvine J.E., Paterson A.H. 1998. Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. *Genetics* 150:1663–1682.
- Ming R., Liu S.C., Moore P.H., Irvine J.E., Paterson A.H. 2001 QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. *Genome Res.* 11:2075–2084.
- Morgante M., Salamini L.V. 2003. From plant genomics to breeding practice. *Curr. Opin. Biotech.* 14: 214-219.
- Mudge J., Andersen W.R., Kehrer R.L., Fairbanks D.J. 1996. A RAPD genetic map of *Saccharum officinarum*. *Crop Sci.* 36: 1362-1363 .
- Muhammad S.; Kausar A.G. 1962. Preliminary studies on the genetics of sugarcane smut, *Ustilago scitaminea* Sydow. *Biologia* 8, 65-74.
- Neale D.B., Savolainen O. 2004. Association genetics of complex traits in conifers. *Trends Plant Sci.* 9: 325-330.
- Nei M., Li W. 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.* 76 : 427-434.
- Neuveglise C., Brygoo Y, Vercambre B., Riba G. 1994. Comparative analysis of molecular and biological characteristics of strains of *Beauveria brongnartii* isolated from insects. *Mycol. Res.* 98:322-328.
- Nordborg M., Tavaré S. 2002. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 18: 83-90.
- Nordborg M., Borevitz J.O., Bergelson J., Berry C.C., Chory J., Hagenblad J., Kreitman M., Maloof J.N., Noyes T., Oefner P.J., Stahl E.A., Weigel D. 2002. The extent of linkage disequilibrium in *Arabidopsis Thaliana*. *Nature Genetics* 30: 190-193.
- Offmann B. 2000. Caractérisation et analyse génétique de la résistance de la canne à sucre *Xanthomonas albilineans*. Thèse de doctorat en sciences, Université de la Réunion, 116 p.
- Olsen K.M., Halldorsdottir S.S., Stinchcombe J.R., Weinig C., Schmitt J., Purugganan M.D. 2004. Linkage disequilibrium mapping of *Arabidopsis* CRY2 flowering time alleles. *Genetics* 167: 1361-1369.

- Panje R.R., Babu C.N. 1960. Studies in *Saccharum spontaneum*. Distribution and Geographical association of chromosome numbers. *Cytologia* 25:152-172.
- Paterson A.H., Lin Y.-R., Li Z., Schertz K.F., Doebley J.F., Pinson S.R.M., Liu S.-C., Stansel J.W., Irvine J.E. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269: 1714-1718 .
- Patterson HD, Williams ER (1976) A new class of resolvable incomplete block designs. *Biometrika* 63:83-92.
- Péros J.P. 1981. Contribution à l'étude du charbon de la canne à sucre (*Saccharum sp.*) causé par *Ustilago scitaminea* (Syd.). Variabilité du parasite et physiologie de l'infection. Thèse de doctorat en sciences, Université de Paris-sud, 93 p.
- Péros J.P., Baudin P. 1983. Etude de la variabilité d'*Ustilago scitaminea* Syd., agent du charbon de la canne à sucre. *L'Agronomie Tropicale* 38 (3), 234-239.
- Perrier X., Flori A., Bonnot F. 2003. Data analysis methods. In : Hamon, P., Seguin, M., Perrier, X. ,Glaszmann, J. C. Ed., Genetic diversity of cultivated tropical plants. Enfield, Science Publishers. Montpellier. pp 43 - 76.
- Pfleiger S., Lefebvre V., Causse M. 2001. The candidate gene approach in plant genetics: a review. *Mol. Breed.* 7:275-291.
- Pinto L.R., Oliveira K.M., Ulian E.C., Garcia A.A., de Souza A.P. 2004. Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome* 47(5):795-804.
- Piperidis G., Christophe M.J., Carroll B.J., Berging N. and D'Hont A. 2000. Molecular contribution to selection of intergeneric hybrids between sugarcane and the wild species *Erianthus arundinaceus*. *Genome*, 43:1033-1037.
- Price S. 1963. Cytogenetics of modern sugar canes. *Economic Botany* 17: 97–105.
- Price S. 1957. Cytological studies in *Saccharum* and allied genera. II. Geographic distribution and chromosome numbers in *S. robustum*. *Cytologia*, 22: 40-52.
- Price S. 1965. Cytology of *Saccharum robustum* and related sympatric species and natural hybrids. U.S. Dep. Agric. Res. Serv., Tech. Bull. 133, 47p.
- Price S., Daniels J. 1968. Cytology of South Pacific sugarcane and related grasses, with special reference to Fiji. *J. Hered.* 59: 141-145.
- Prospero J.M., Bonatti E., Schubert C., Carlson T.N., 1970. Dust in the Caribbean atmosphere traced to an african dust storm. *Earth and Planetary Science letters* 9, 287-293.
- Rafalski A., Morgante M. 2004. Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* 20: 103-111.
- Rands R.D., Abbott E.V. 1964. Sereh. In : Hughes c.G., Abbott E.V. and Wismer C.A. (eds). *Sugar cane diseases of the world*. Elsevier, Amsterdam, 183-189.
- Reffay N. 2003. Etude des facteurs génétiques contrôlant le rendement en sucre et la teneur en fibre dans deux populations de canne à sucre (*Saccharum spp.*) de l'île de la Réunion et d'Australie. Thèse de doctorat en Sciences, Université de la Réunion, 146 p.
- Ricaud C., Egan B.T., Gillaspie-jr A.G., Hugues C.G. (Eds). 1989. Diseases of sugarcane. Elsevier Science Publishers B.V, Amsterdam. 400 p.

- Riley I.T., Jubb T.F., Egan B.T., Croft, B.J. 1999. First outbreak of sugarcane smut in Australia. Proc. Int. Soc. Sugar Cane Technol. 23 (2), 333-336.
- Roach B.T. 1972. Nobilization of sugarcane. Proc. Int. Soc. of Sugarcane Technol. 14: 206-216.
- Roach B.T. 1986. Evaluation and breeding use of sugarcane germplasm. Proc. Int. Soc. Sugarcane Technol. 19: 492-501.
- Rochoux J.Y. 2002. La canne à sucre à la Réunion, une production opportune? In : Actes des Rencontres internationales pluridisciplinaires : «Perspectives de développement de la canne à sucre en milieu insulaire:approches technico-économiques, sociales, culturelles », du 2 au 5 octobre 2002, Muséum Stella Matutina, St-Leu, Réunion. (sous presse).
- Rogers S., Bendich A. 1988. Extraction of DNA from plant tissues. In: Gelvin, S.B., Schilperoort, R.A. (Eds), Plant Molecular Biology Manual A6. Kluwer Academic Publishers, Dordrecht, pp 1-10.
- Rossi M., Araujo P.G., Paulet F., Garsmeur O., Dias V.M., Chen H., Van Sluys M.A., D'hont A. 2003. Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. Mol. Gen. Genet. 269:406-419.
- Rott P., Bailey R.A., Comstock J.C., Croft B.J., Salem Saumtally A. (Eds). 2000. A guide to sugarcane diseases. CIRAD and ISSCT. 340p.
- Ruiz M., Rouard M., Raboin L.M., Lartaud M., Lagoda P., Courtois B. 2004. Tropgene-DB, a multitropical crop information system. Nucleic Acids Res. 32:D364-D367.
- SAS Institute Inc. 1990. SAS procedures guide, version 6, third edition. SAS Intitute Inc, Cary, NC, USA, 705 pp.
- SAS Institute Inc. 1996. SAS/STAT® Software: Changes and enhancements through release 6.11, Cary , NC:SAS Institute Inc. 1104 pp.
- Saxena S.K., Khan A.M. 1963. Studies on sugarcane smut caused by *Ustilago scitaminea* Syd. I. Effect of temperature on spore germination. J. Ind. Bot. Soc. 42:195-203.
- Simko I. 2004. "One potato, two potato: haplotype association mapping in autotetraploids." Trends Plant Sci. 9(9): 441-448.
- Simko I., Costanzo S., Haynes K., Christ B., Jones R. 2004. Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor. Appl. Genet. 108: 217-224.
- Simmonds N.W. 1993. Introgression and incorporation strategies for the use of crop genetic ressources. Biol. Rev. 68: 539-562.
- Skinner J.C., Hogarth D.M., Wu K.K. 1987. Selection methods, criteria and indices. In : Heinz D.J. Sugarcane improvement through breeding. Elsevier Press, Amsterdam, 409-453.
- Sobral B.W.S., Braga D.P.V., LaHood E.S., Keim P. 1994. Phylogenetic analysis of chloroplast restriction enzyme site mutations in the *Saccharinae* Griseb. subtribe of the *Andropogoneae* Dumort. tribe. Theor. Appl. Genet. 87: 843-853.
- Soula B., Goebel R., Caplong P., Karimjee H., Tibere R., Tabone E. 2003. *Trichogramma chilonis* (Hymenoptera: trichogrammatidae) as a biological control agent of *Chilo*

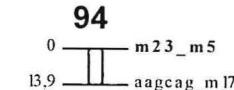
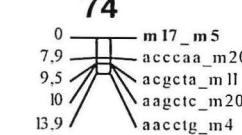
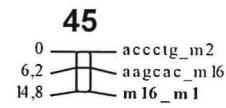
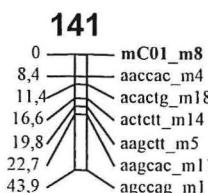
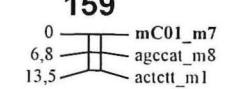
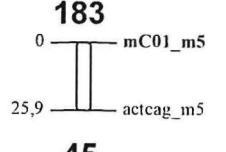
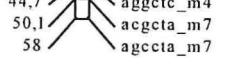
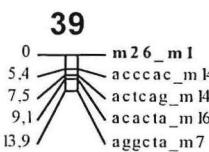
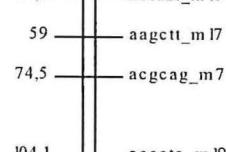
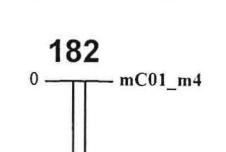
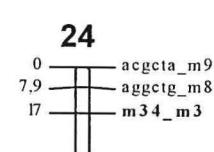
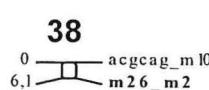
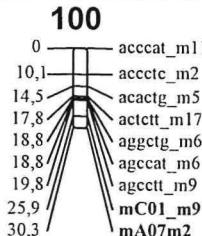
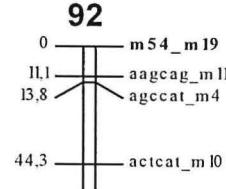
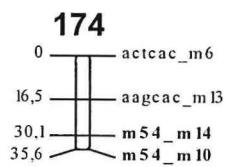
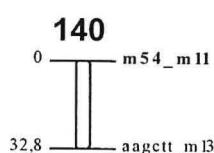
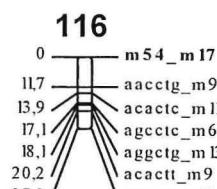
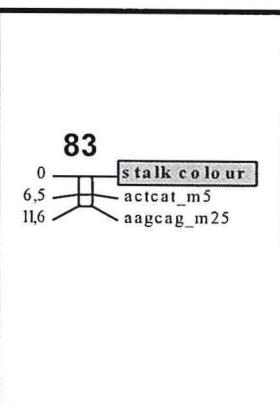
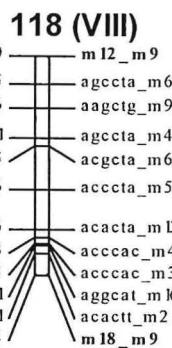
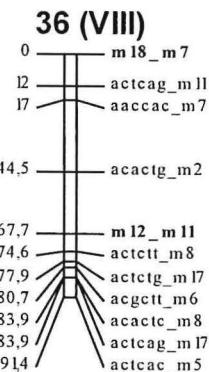
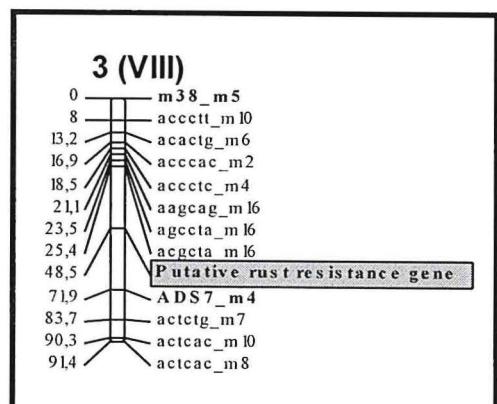
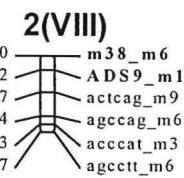
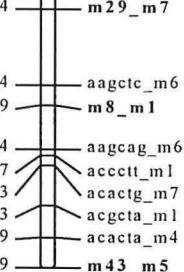
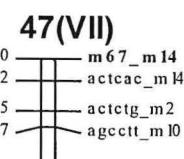
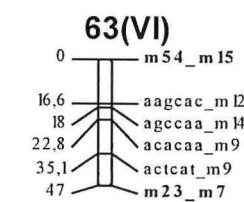
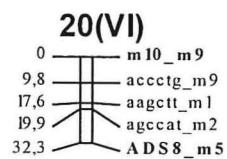
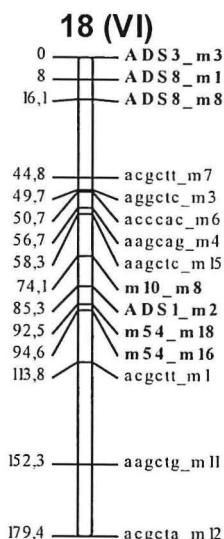
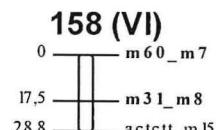
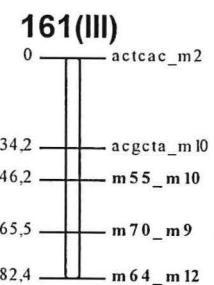
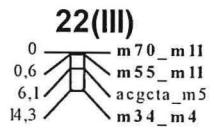
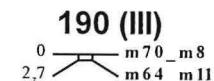
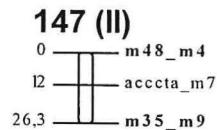
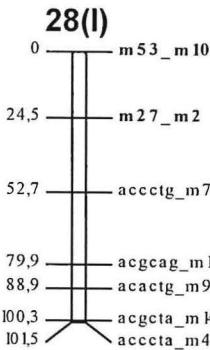
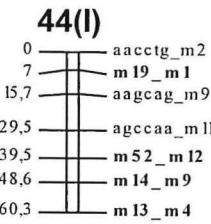
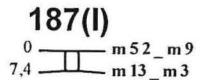
- sacchariphagus* (Lepidoptera:crambidae) in Reunion Island. Initial field trials. Proc. South African Sugar Technol. Ass. 77: 278-283.
- Sreenivasan T.V. Ahloowalia B.S., Heinz D.J. 1987. Cytogenetics. In : Heinz D.J. Sugarcane improvement through breeding. Elsevier, Amsterdam, 211-253.
- Srinivasan K.V., Alexander K.C. 1971. Sources of resistance to red rot and smut in the species of *Saccharum*. Sugarcane pathologist's newsletter 6, 6-7.
- Stevenson G.C. 1965 Genetics and Breeding of Sugar cane. Longmans, London 284 p.
- Tai P.Y.P., Miller J.D., Legendre B.L. 1995. Evaluation of the world collection of *Saccharum spontaneum* L. Proc. Int. Soc. of Sugarcane Technol. 22(2): 250-259.
- Tai P.Y.P., Miller J.D., Dean J.L. 1981. Inheritance of resistance to rust in sugarcane. Fields Crops Res. 4:261-268.
- Takahashi Y., Shomura A., Sasaki T. et Yano M. 2001. Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the subunit of protein kinase CK2. Proc. Natl. Acad. Sci. 98(14): 7922-7927.
- Tanksley S.D., McCouch S.R. 1997. Seed banks and molecular maps : unlocking genetic potential from the wild. Science 277: 1063-1066.
- Tanksley S.D., Nelson J.C. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. 92: 191-203.
- Tenaillon M.I., Sawkins M.C., Long A.D., Gaut R.L., Doebley J.F., Gaut.S. 2001. Patterns of DNA sequence polymorphism along chromosome of maize (*Zea mays* ssp. *mays* L.). Proc. Natl. Acad. Sci. 98: 9161-9166.
- Toffano W.B. 1976. Estudos sobre polaridade sexual e possíveis raças fisiológicas de *Ustilago scitaminea* Syd., no estado de São Paulo. Arq. Inst. Biol., São Paulo 43, 65-79.
- Tomkins J.P., Yu Y., Miller-Smith H., Frisch D.A., Woo S.S., Wing R. 1999. A bacterial artificial chromosome library for sugarcane. Theor. Appl. Genet. 99:419-424.
- Trione E.J. 1980 Teliospore formation by *Ustilago scitaminea* in Sugarcane. Phytopathology 70:513-516.
- Trione E.J. 1990. Growth and sporulation of *Ustilago scitaminea*, in vivo and in vitro. Mycol. Res. 94 (4), 489-493.
- Veiga F.M. 1972. Smut in Brazil. Sugarcane Pathologists' Newsletter 9 (25), 17.
- Vettore A.L., da Silva F.R., Kemper E.L., Souza G.M., da Silva A.M., Ferro M.I., Henrique-Silva F., Giglioti E.A., Lemos M.V., Coutinho L.L., Nobrega M.P., Carrer H., Franca S.C., Bacci Junior M., Goldman M.H., Gomes S.L., Nunes L.R., Camargo L.E., Siqueira W.J., Van Sluys M.A., Thiemann O.H., Kuramae E.E., Santelli R.V., Marino C.L., Targon M.L., Ferro J.A., Silveira H.C., Marini D.C., Lemos E.G., Monteiro-Vitorello C.B., Tambor J.H., Carraro D.M., Roberto P.G., Martins V.G., Goldman G.H., de Oliveira R.C., Truffi D., Colombo C.A., Rossi M., de Araujo P.G., Sculaccio S.A., Angella A., Lima M.M., de Rosa Junior V.E., Siviero F., Coscrito V.E., Machado M.A., Grivet L., Di Mauro S.M., Nobrega F.G., Menck C.F., Braga M.D., Telles G.P., Cara F.A., Pedrosa G., Meidanis J., Arruda P. 2003. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Res. 13(12):2725-2735.

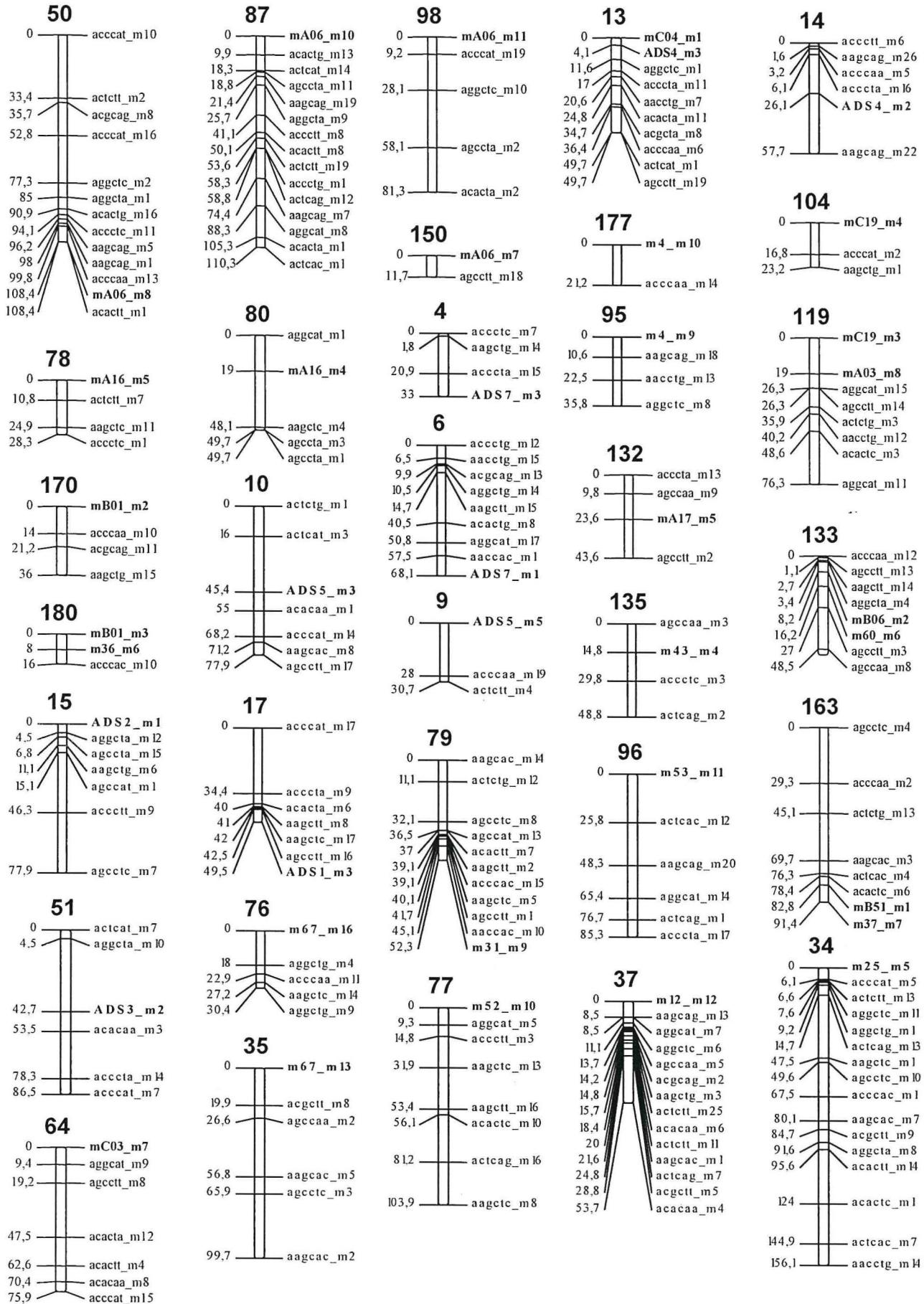
- Vigouroux Y., McMullen M.D., Hittinger C.T., Houchins K., Schulz L., Kresovich S., Matsuoka Y., Doebley J.F.. 2002. Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc. Natl. Acad. Sci. 99: 9650-9655.
- Vincentz M., Cara F., Okura V., da Silva F.R., Pedrosa G.L., Hemerly A.S., Capella A.N., Marins M., Ferreira P.C., Franca S.C., Grivet L., Vettore A.L., Kemper E.L., Burnquist W.L., Targon M.L., Siqueira W.J., Kuramae E.E., Marino C.L., Camargo L.E., Carrer H., Coutinho L.L., Furlan L.R., Lemos M.V., Nunes L.R., Gomes S.L., Santelli R.V., Goldman M.H., Bacci M. Jr, Giglioti E.A., Thiemann O.H., Silva F.H., Van Sluys M.A., Nobrega F.G., Arruda P., Menck C.F. 2004. Evaluation of Monocot and Eudicot Divergence Using the Sugarcane Transcriptome. Plant Physiology, Vol. 134: 951-959.
- Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414 .
- Waller J.M. 1969. Sugarcane smut (*Ustilago scitaminea*) in Kenya (I) Epidemiology. Trans. Brit. Mycol. Soc. 52: 139-151.
- Waller J.M. 1970. Sugarcane smut (*Ustilago scitaminea*) in Kenya II. Infection and resistance. Trans. Brit. Mycol. Soc. 54: 405-414.
- Walker D.T., 1987. Breeding for disease resistance . In Heinz D.J(editor), Sugarcane improvement through Breeding. Elsevier Amsterdam, Netherlands. 455-502.
- Wang R.-L., Stec A., Hey J., Lukens L., Doebley J.F. 1999. The limits of selection during maize domestication. Nature 398: 236-239.
- Wenzl P., Carling J., Kudrna D., Jaccoud D., Huttner E., Kleinhofs A., Kilian A. 2004. Diversity arrays technology (DArT) for whole-genome profiling of barley. Proc. Natl. Acad. Sci. 101: 9915-9920.
- Whittle A.M., Walker D.I.T. 1982. Interpretation of sugarcane smut susceptibility trials. Trop. Pest. Manag. 28: 228-237.
- Williams J.R., Metcalfe J.R., Mungomery R.W., Mathes R. 1969. Pests of sugarcane. Amsterdam: Elsevier. 568p.
- Wu K.K., Burnquist W., Sorrels M.E., Tew T.L., Moore P.H., Tanksley S.D. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. 83:294–300.
- Wu K.K., Heinz D.J., Hogarth D.M. 1988. Association and heritability of sugarcane smut resistance to races A and B in Hawaii. Theor. Appl. Genet. 75, 754-760.
- Wu K.K., Heinz D.J., Meyer H.K. 1983. Heritability of sugarcane smut resistance and correlation between smut grade and yield components. Crop Science 23: 54-56.
- Yamauch S. 1978. Epidemic of smut (*Ustilago scitaminea* Syd.) In Okinawa and its control. Sugarcane Pathologists' Newsletter 21, 4-5.
- Yano M., Katayose M., Yamanouchi U., Monna L., Fuse T., Baba T., Yamamoto K., Umehara Y., Nagamura Y., Sasaki T.. 2000. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *constans*. The Plant Cell 12: 2473-2483.

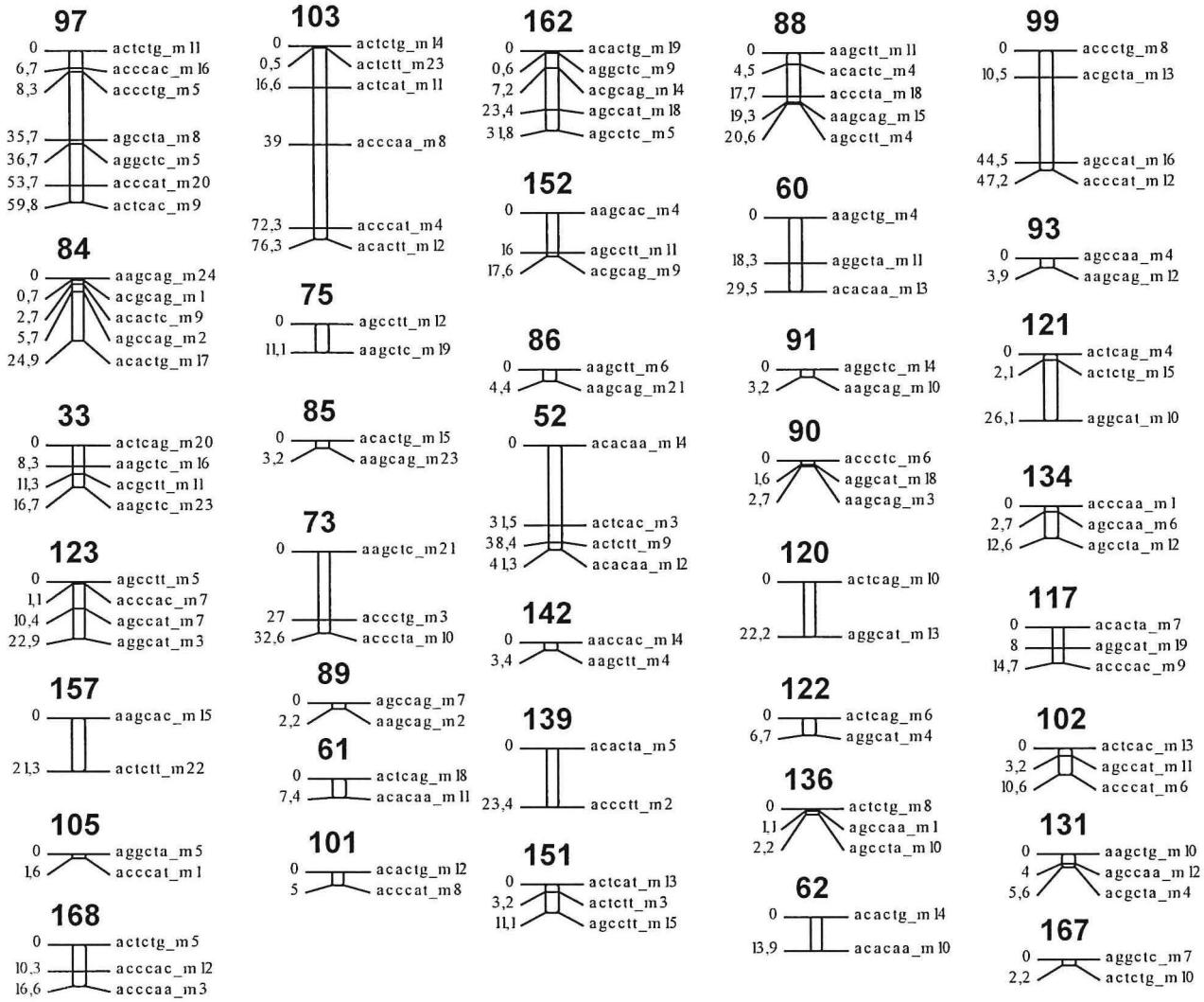
Young N.D. 1996. QTL mapping and quantitative disease resistance in plants. *Annu. Rev. Phytopathol.* 34: 479-501.

ANNEXES

ANNEXE 1: carte génétique du clone MQ 76/53







ANNEXE 2: Détection de QTLs pour les autres caractères

Détection de QTLs pour le nombre de tiges usinables

marker ¹	CG ²	effect ³	cM	HG	Rossi ²	JYH ²	n(+) ⁴	n(-) ⁴	LP 2001				LP 2002				LP 2003				LP 2004				Gol 2004						
									avg(+)	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P			
R570 and MQ76/53																															
aaggac_m23	DDgroup10	-		VI	U51	54	148	46	31.5	34.0	0.063 ***		37.3	42.3	0.043 ***						42.9	51.8	0.042 ***		33.2	38.0	0.043 ***				
actcag_rm7	DDgroup10	-		VI	U51	54	146	46	31.5	33.7	0.045 ***										42.6	51.7	0.046 ***		33.2	37.6	0.037 **				
maccat_rm8	DDgroup40	-		VII	VII ⁹	138	58										38.3	44.3	0.043 ***												
aaccac_rm14	DDgroup2	+		VI	VI1a	26	146	46									39.6	34.4	0.048 ***	41.5	35.2	0.041 ***		46.8	37.9	0.043 ***					
aaggac_rm25	-			U-54	31	188	6										38.1	51.1	0.050 ***												
aggcat_rm23	+			VI	VI-2	28	144	49	32.7	30.3	0.060 ***										47.4	38.8	0.042 **								
aggcta_rm9	+			U-5		112	143	50												41.7	35.6	0.042 ***		47.0	38.8	0.039 **					
aggdctc_rm2	-			U-41		119	147	48									37.3	42.1	0.041 ***	38.8	44.5	0.035 **		32.9	38.5	0.062 ***					
m67_rm1	-			VII	VII-1b	130	60	31.5	33.3	0.038 **		37.0	41.4	0.039 **		38.1	43.8	0.040 **		42.3	50.1	0.040 **		32.7	37.4	0.048 ***					
R570																															
aggdctc_r14	group171	+	0	U-30	111	99	96										40.6	36.4	0.042 ***					48.9	41.0	0.047 ***		36.7	31.8	0.064 ***	
aaggac_r23	group171	+	4.9	U-30	111	94	101										41.1	36.1	0.060 ***	42.9	37.7	0.038 **		49.5	40.8	0.056 ***		36.7	32.0	0.061 ***	
ADS4_r6	group11	-	0	VI		86	58	31.4	33.7	0.065 ***																					
accctl_r9	group11	-	26.4	VI	VI-c	57	108	80	31.3	33.1	0.043 ***																				
aggccat_r4	group108	-	20.1	I	I-9	2	115	74	31.5	33.2	0.039 ***																				
accccc_r5	group54	+	47.1	I			107	88									40.5	36.1	0.045 ***												
acccta_r2	group49	+	10.4	VII			100	92																							
m841m2r	group129	+	0	VII			95	94																							
aggcat_r33	-			U-48	93	98	98													42.9	37.3	0.047 ***		48.7	41.4	0.040 **		36.3	32.2	0.045 ***	
aggdctc_r6	-					131	64																42.3	50.5	0.044 ***						
m31_r1	-					105	68																								
mC01m1r	+					89	105																								
MQ76/53																															
aaggac_m7	group87	+	74.4				95	99																			36.5	32.3	0.048 ***		
aggcat_m8	group87	+	88.3				84	106																			36.6	32.4	0.046 ***		
agccaa_m3	group135	-	0				95	98																			32.1	36.2	0.043 ***		
m43_m4	group135	-	14.8				107	88																			32.3	36.6	0.048 ***		
agccaa_m1	group134	+	0				90	102									43.0	37.4	0.046 ***	48.9	41.0	0.048 ***		36.4	32.2	0.047 ***					
agccaa_m6	group134	+	2.7				91	100								40.5	36.5	0.039 **	43.4	37.3	0.054 ***		49.3	40.9	0.054 ***		36.6	32.1	0.054 ***		
aggcta_m12	group134	+	12.6				94	97									42.6	37.4	0.040 ***	48.9	40.4	0.055 ***		36.2	32.1	0.045 ***					
acccaa_m12	group133	-	0				92	100	31.0	32.9	0.050 ***																				
accctl_m13	group133	-	1.1				91	102	31.1	32.9	0.046 ***																				
accctl_m14	group133	-	2.7				90	102	31.2	32.9	0.041 ***																				
mB06m2q	group133	-	8.2				94	102	31.1	32.9	0.047 ***																				
m60_m6	group133	-	16.2				96	98	30.9	33.1	0.067 ***																				
aggctt_m3	group133	-	27				96	97	31.0	33.2	0.068 ***																32.4	36.3	0.041 ***		
agccaa_m8	group133	-	48.5				96	88	30.9	33.3	0.083 ***																32.4	36.6	0.047 ***		
m18_m8	group133	-	71				98	91	31.0	33.1	0.059 ***																				
m54_m17	group116	-	0				93	103									37.2	42.7	0.042 ***		40.4	49.0	0.055 ***		32.3	36.0	0.036 **				
aggcata_m13	group99	+	10.5				94	97								40.7	36.5	0.043 ***									36.6	32.2	0.051 ***		
acccaa_m14	group52	0					81	76	31.4	33.3	0.051 ***																				
acccg_m14	group162	+	7.2				99	96									42.8	37.5	0.041 **		48.6	41.3	0.041 ***								
actctl_m22	group157	+	21.3				95	98	32.9	31.1	0.046 ***																				
m48_m4	group147	-	0	II			90	105								36.3	40.8	0.050 ***		37.7	43.1	0.043 **		41.3	49.1	0.047 ***		32.2	36.3	0.049 ***	
aaggctc_m10	-						85	93									36.3	40.8	0.050 ***		37.7	43.1	0.043 **		41.3	49.1	0.047 ***		32.2	36.2	0.049 ***
actct_m5	-						92	104																					32.1	36.2	0.043 ***

¹ name of markers. The suffix of the marker name indicates if the marker is present in R 570 (_r), in MQ 76/53 (_m), in both of them (_rm)

² cosegregation groups in Raboin et al (submitted) map, Horau et al. (2001) map and Rossi et al. (2003) map

³ + effect increasing measured trait, - effect diminishing measured trait

⁴ number of individuals presenting (+) or not (-) a given marker

⁵ average smut resistance parameter among clones presenting the marker (+) or not (-)

R² multiple regression

	0.28	(159)	0.32	(163)	0.36	(154)	0.39	(150)	0.45	(146)
--	------	-------	------	-------	------	-------	------	-------	------	-------

Détection de QTLs pour le diamètre des tiges

(en gras marqueurs de R 570 déjà détectés par Hoarau et al. (2002))

marker ¹	CG ²	effect ³	cM	HG	Rossi ²	JYH ²	n(+) ⁴	n(-) ⁴	LP 2002				LP 2004			
									avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P
R570 and MQ76/53																
aaccat_rm19	DDgroup30	-	I	I-8	68	132	47		2.357	2.430	0.055 ***		2.262	2.344	0.055 ***	
agccat_rm26	-	-	VII	U-12	53	151	41		2.381	2.297	0.038					
agccct_rm25	-	-	IV	IV-2	9	173	19		2.365	2.470	0.043 ***					
agccct_rm1	-	-				176	13									
aacact_rm11	+	-		U-14	11	144	48		2.360	2.434	0.041 ***		2.304	2.232	0.041 ***	
aacact_rm16	-	-				166	28									
m37_rm1	+	-	VII	VI-8		136	56		2.391	2.323	0.059 ***		2.304	2.242	0.035 **	
R570																
agccat_r4	group108	+	20.1	I	I-9	2	115	74	2.392	2.337	0.045 ***		2.309	2.247	0.040 **	
actctg_r3	group108	+	28.9	I	I-9	2	109	83					2.314	2.250	0.043 ***	
m34_r4	group25	-	15.8				102	91					2.253	2.321	0.049 ***	
m34_r3	group23	+	20.5	II	II-8		97	80					2.317	2.249	0.049 ***	
m25_r2	group21	-	33.5	II	II-6		97	99	2.344	2.396	0.042 ***					
agccit_r23	DDgroupR18	-				125	68		2.351	2.410	0.049 ***					
m54_r3	+	-				151	45						2.305	2.220	0.054 ***	
MQ76/53																
aagcac_m5	group35	+	56.8				99	92	2.400	2.340	0.055 ***		2.327	2.239	0.081 ***	
agccct_m3	group35	+	65.9				98	91	2.405	2.337	0.071 ***		2.327	2.242	0.078 ***	
aagcac_m2	group35	+	99.7				95	97	2.401	2.344	0.050 ***		2.323	2.250	0.056 ***	
acccac_m1	group34	-	67.5				77	96	2.333	2.408	0.085 ***		2.249	2.316	0.046 ***	
acgcct_m9	group34	-	84.7				53	90	2.338	2.399	0.052 ***					
agccct_m3	group18	+	49.7	VII			96	99					2.324	2.248	0.060 ***	
acccac_m6	group18	+	50.7	VII			95	97					2.321	2.255	0.048 ***	
aagcac_m4	group18	+	56.7	VII			102	92					2.324	2.241	0.074 ***	
aaagctc_m15	group18	+	58.3	VII			100	93					2.324	2.242	0.073 ***	
m10_m8	group18	+	74.1	VII			101	94					2.318	2.249	0.050 ***	
ADS1_m2	group18	+	85.3	VII			103	87					2.317	2.243	0.057 ***	
m54_m18	group18	+	92.5	VII			107	89					2.314	2.251	0.042 ***	
acccaa_m12	group133	+	0				92	100	2.402	2.345	0.049 ***					
agccct_m13	group133	+	1.1				91	102	2.400	2.347	0.044 ***					
acgcct_m14	group133	+	2.7				90	102	2.396	2.345	0.041 ***					
acacaa_m1	group10	+	55				83	102	2.403	2.351	0.043 ***					
acacaa_m3	group51	+	53.5				94	99					2.316	2.254	0.042 ***	
acgcag_m12	group28	+	79.9	I			87	108	2.401	2.347	0.044 ***					
acgcct_m8	group24	-	0				91	101	2.344	2.399	0.046 ***					
acacca_m14	+	-					146	46					2.305	2.226	0.048 ***	
m14_m8	+	-					104	91					2.316	2.247	0.051 ***	
m52_m11	+	-					142	52					2.306	2.231	0.048 ***	
R ² multiple regression										0.370				0.560		
										(143)				(132)		

¹ name of markers. The suffix of the marker name indicates if the marker is present in R 570 (_r), in MQ 76/53 (_m), in both of them (_rm)

² cosegregation groups in Raboin et al (submitted) map, Horau et al. (2001) map and Rossi et al. (2003) map

³ + effect increasing measured trait, - effect diminishing measured trait

⁴ number of individuals presenting (+) or not (-) a given marker

⁵ average smut resistance parameter among clones presenting the marker (+) or not (-)

Détection de QTLs pour le brix

(en gras marqueur de R 570 déjà détecté par Hoarau et al. (2002) mais son effet en simple dose était favorable au brix)

marker ¹	CG ²	effect ³	cM	HG	Rossi ²	JYH ²	n(+) ⁴	n(-) ⁴	LP 2001			LP 2002			LP 2004				
									avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²
R570 and MQ76/53																			
acactg_rm26	DDgroup21	-			U-44		150	45	17.6	18.1	0.040 **		17.9	18.3	0.053 ***				
acgcta_rm17	-			II	II-9	4	140	51	17.6	18.1	0.045 ***		17.8	18.3	0.043 ***		18.1	18.5	0.039 **
aggctc_rm3							184	11	17.7	18.5	0.042 ***		17.9	19.0	0.081 ***				
actcac_rm21	+			VI	VI-d	61	149	42									18.3	17.7	0.073 ***
m62_rm6				I	I-6		135	55					17.8	18.3	0.047 ***				
m67_rm1	-			VII	VII-1b		130	60	17.6	18.0	0.045 ***					18.1	18.4	0.038 **	
R570																			
acgcta_ra	group67	+	88.6	VIII	VIII-1	14	103	89	17.9	17.5	0.051 ***								
aggctg_r13	group67	+	90.7	VIII	VIII-1	14	105	88	17.9	17.5	0.055 ***								
acacgt_r3	group67	+	98.3	VIII	VIII-1	14	99	94	17.9	17.5	0.043 ***								
m12_r4	group67	+	111.3	VIII	VIII-1		104	92	17.9	17.5	0.044 ***								
acactc_r19	group53		72.9	III	III-3	23	94	98	17.5	17.9	0.049 ***								
m64_r10	group82	+	15.5	III			114	82	17.9	17.5	0.048 ***		18.1	17.8	0.037 **				
accatc_rx2	+						182	12	17.8	16.9	0.047 ***								
acccta_r9	+						138	46	17.8	17.4	0.044 ***		18.1	17.6	0.041 **				
actctt_r3	+						140	54	17.8	17.4	0.037 **		18.1	17.7	0.045 ***				
MQ76/53																			
acacaa_m3	group51	-	53.5				94	99	17.5	17.9	0.056 ***								
acccac_m7	group51	+	86.5				92	98				18.2	17.8	0.043 ***		18.4	18.0	0.053 ***	
acccac_m12	group168	-	10.3				98	94	17.5	17.9	0.044 ***					18.0	18.4	0.035 **	
acccaa_m3	group168	-	16.6				95	97	17.5	18.0	0.055 ***		17.8	18.2	0.041 ***				
aagcac_m2	group35	+	99.7				95	97								18.4	17.9	0.070 ***	
actcac_m14	group47	+	8.2	VII			98	94								18.4	18.0	0.051 ***	
actcac_m3	group52	-	31.5				102	86	17.5	18.0	0.052 ***								
actctt_r22	group157	+	21.3				96	98	17.9	17.5	0.046 ***								
aggccaa_m9	group132	+	9.8				91	104								18.4	18.0	0.039 ***	
actcac_m5	group183	*	25.9				94	97	17.9	17.5	0.045 ***								
actcat_m5	group83	+	6.5				85	106				18.2	17.8	0.048 ***					
aggccaa_r12	-						139	55	17.8	17.4	0.039 ***		18.1	17.6	0.051 ***		18.4	17.8	0.080 ***
aggctg_m11	-						144	50				17.9	18.3	0.039 ***					
accctt_m7	-						102	88				17.7	18.2	0.077 ***					

R² multiple regression

0.48 (151) 0.36 (161) 0.25 (169)

¹ name of markers. The suffix of the marker name indicates if the marker is present in R 570 (_r), in MQ 76/53 (_m), in both of them (_rm)

² cosegregation groups in Raboin et al (submitted) map, Horau et al. (2001) map and Rossi et al. (2003) map

³ + effect increasing measured trait, - effect diminishing measured trait

⁴ number of individuals presenting (+) or not (-) a given marker

⁵ average smul resistance parameter among clones presenting the marker (+) or not (-)

Détection de QTLs pour le taux de floraison

(en gras marqueurs de R 570 déjà détectés par Offmann (2000) pour le nombre d'entre-nœuds. Il est à noter que les CG IV1 et IV2 s'apparent préférentiellement à la meiose)

marker ¹	CG ²	effect ³	cM	HG	Rossi ²	JYH ²	n(+) ⁴	n(-) ⁴	LP 2001 (vierge)				LP 2002				LP 2003				LP 2004				
									avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	avg(+) ⁵	avg(-) ⁵	R ²	P	
R570 and MQ76/53																									
aagcag_rm5	DDgroup8	+	II	II-7			138	56	3.508	1.586	0.046 ***		8.310	3.716	0.036 **										
acactt_rm14	DDgroup8	+	II	II-7			139	55	3.470	1.649	0.041 ***														
m67_rm8	DDgroup36	-	VII	VII-8			122	68	2.276	4.259	0.054 ***		5.471	9.860	0.036 **										
actctg_rm1	DDgroup23	-					163	28	2.507	5.461	0.066 ***		5.849	12.811	0.050 ***										
accat_rm41	-						135	26					6.386	13.286	0.049 **		5.332	11.412	0.057 ***						
accacag_rm17	-		VII	VII-8	63	178	17									5.337	12.352	0.047 ***							
aggcgtc_rm19	-		U-18		49	168	27		2.488	5.779	0.079 ***		6.040	12.614	0.043 ***								7.347	11.522	0.035 **
acactc_rm17	-		U4		69	167	27						6.034	12.827	0.046 ***		5.262	10.395	0.038 **						
accctg_rm2	-		VI	VI-1b	33	173	21		2.540	6.346	0.085 ***		6.214	13.289	0.040 **		5.313	11.435	0.045 ***						
m4m_rm4	-		VII	VII-a			146	50					5.612	10.745	0.042 ***										
R570																									
mA06m4r	group81	-	0	IV			97	98					4.238	9.569	0.059 ***		3.539	8.294	0.068 ***		6.134	9.623	0.051 ***		
aagcag_r28	group81	-	6.2	IV	IV-1	8	95	99					4.910	8.975	0.034 **		3.997	7.876	0.045 ***		6.389	9.460	0.039 **		
agcctc_r24	group81	-	11.1	IV	IV-1	8	91	99					4.697	9.266	0.042 ***		3.538	8.431	0.071 ***		6.061	9.475	0.050 ***		
aggcgt_r20	group81	-	12.6	IV	IV-1	8	93	102					4.611	9.083	0.041 ***		3.477	8.203	0.067 ***		5.972	9.705	0.058 ***		
aggcta_r21	group106	+	0	IV	IV-2	9	64	68					10.321	4.470	0.068 ***		9.247	3.290	0.106 ***		10.346	5.638	0.095 ***		
mA06m2r	group106	+	7.4	IV			98	97					9.577	4.229	0.059 ***		8.306	3.527	0.069 ***		9.686	6.071	0.055 ***		
actcat_r11	group106	+	17.1	IV	IV-2	9	99	92					9.596	4.311	0.057 ***		8.680	3.078	0.093 ***		9.918	5.769	0.073 ***		
accctg_r7	group106	+	19.8	IV	IV-2	9	100	93					9.289	4.551	0.046 ***		8.428	3.363	0.076 ***		9.839	5.978	0.062 ***		
accacat_r8	group106	+	21.9	IV	IV-2	9	100	93					9.222	4.605	0.044 ***		8.322	3.338	0.074 ***		9.845	5.838	0.067 ***		
accacag_r5	group173	+	0	VI	VI-2	28	100	94					9.129	4.688	0.041 ***										
actctg_r9	group125	+	0	VI			86	91								3.923	8.174	0.052 ***		6.170	9.738	0.053 ***			
aggcag_r5	DDgroupR28	-					143	48											7.142	10.692	0.039 ***				
MQ76/53																									
aacctg_m13	group95	+	22.5				82	105					9.758	4.489	0.059 ***		8.168	4.108	0.050 ***		9.615	6.504	0.040 ***		
aggcgc_m8	group95	+	35.8				78	116					10.024	4.835	0.053 ***		8.820	4.000	0.067 ***		10.239	6.364	0.060 ***		
m29_m7	group47	-	52.4	VII			99	91								3.941	8.075	0.051 ***							
aggcgc_m6	group47	-	81.4	VII			95	99								3.927	7.942	0.048 ***		6.345	9.502	0.042 ***			
m8_m1	group47	-	88.9	VII			84	94								3.870	8.156	0.053 ***		6.201	9.906	0.055 ***			
aaagcag_m6	group47	-	105.4	VII			88	106	2.094	3.667	0.037 **					3.679	7.883	0.052 ***		5.804	9.742	0.064 ***			
accctt_m1	group47	-	107	VII			86	108	1.964	3.742	0.047 ***		4.102	9.272	0.054 ***		3.097	8.268	0.079 ***		5.390	9.999	0.088 ***		
acacclg_m7	group47	-	111.3	VII			84	110	2.104	3.602	0.033 **		4.313	9.023	0.045 ***		3.357	7.976	0.063 ***		5.894	9.530	0.054 ***		
ADS4_m2	group14	+	26.1				65	77	4.159	2.166	0.052 ***														
aaagcag_m22	group14	+	57.7				94	100	3.820	2.139	0.043 ***														
acactg_m5	group100	+	14.5				97	97	3.884	2.024	0.053 ***		9.218	4.730	0.041 ***		8.077	3.876	0.053 ***		9.707	6.204	0.051 ***		
actctt_m17	group100	+	17.8				99	95	3.814	2.045	0.047 ***		9.095	4.531	0.043 ***		7.622	4.025	0.039 **						
aggcctg_m6	group100	+	18.8				100	93	3.767	2.093	0.042 ***		9.156	4.717	0.040 **		7.788	4.017	0.042 ***		9.419	6.447	0.037 **		
aggcat_m6	group100	+	18.8				99	93	3.797	2.039	0.047 ***		9.243	4.566	0.045 ***		7.860	4.014	0.044 ***		9.465	6.360	0.040 **		
mC01m9q	group100	+	25.9				100	95	3.684	2.166	0.035 **		9.371	4.405	0.051 ***		7.944	3.877	0.050 ***		9.543	6.257	0.045 ***		
mA07m2q	group100	+	30.3				97	99	3.755	2.129	0.040 ***		9.402	4.491	0.050 ***		7.958	3.955	0.048 ***		9.388	6.467	0.036 **		
aacctg_m14	group34	+	156.1				87	103								8.159	4.274	0.044 ***							
ADS3_m2	group51	+	42.7				74	79					5.118	10.063	0.044 **		4.218	8.777	0.055 ***						
aacctg_m9	group116	-	11.7				93	95					4.572	9.145	0.045 ***										
acactt_m7	-						158	36	2.490	5.003	0.056 **					4.950	10.476	0.055 ***		7.082	11.102	0.042 ***			
R² multiple regression																						0.310			
									0.380				(180)			0.460				(114)			(165)		
																				(118)					

¹ name of markers. The suffix of the marker name indicates if the marker is present in R 570 (_r), in MQ 76/53 (_m), in both of them (_rm)

² cosegregation groups in Raboin et al (submitted) map, Horau et al. (2001) map and Rossi et al. (2003) map

³ + effect increasing measured trait, - effect diminishing measured trait

⁴ number of individuals presenting (+) or not (-) a given marker

⁵ average smut resistance parameter among clones presenting the marker (+) or not (-)

Summary:

Sugarcane smut, caused by *Ustilago scitaminea* Syd., is present in all sugarcane growing areas with the exception of Papua New Guinea, Fiji and the eastern side of Australia. Breeding for smut resistance is efficient because this trait is fairly heritable but it requires complicated screenings. Moreover, the genetic control of smut resistance is still unknown. With the objective to identify the mendelian factors involved in sugarcane resistance to smut, two strategies have been implemented (1) QTL mapping in a bi-parental progeny derived from a cross between a resistant cultivar 'R 570' and a highly susceptible clone 'MQ 76/53', evaluated in Reunion island for resistance to smut (2) Association study in a population of cultivars, evaluated in Burkina Faso for resistance to smut.

First, the worldwide genetic structure of *U. scitaminea* was investigated. A total of 142 single-teliospore isolates from 15 countries worldwide were analysed using 17 polymorphic microsatellite loci. All isolates but one were homozygous for all loci, indicating that selfing is the highly preferential reproductive mode of *U. scitaminea*. In America and Africa, genetic diversity was found to be extremely low and all isolates belonged to a single inbred lineage. This inbred lineage was also found in some parts of the Asian continent where most *U. scitaminea* genetic diversity was detected. These observations support the hypothesis that the fungus originated in Asia. Thus, the bi-parental population and the population of cultivars used to study the genetic determinism of smut resistance have been evaluated towards isolates from the worldwide lineage.

The genetic maps of the two parents of the bi-parental progeny, 'R 570' and 'MQ 76/53', were constructed using a population of 198 progeny. A total of 1666 polymorphic markers were produced using 37 AFLP primer pairs combinations, 46 SSRs and 9 RFLP probes. Linkage analysis allowed the construction of 86 cosegregation groups for 'R 570' and 105 cosegregation groups for 'MQ 76/53' encompassing 424 and 536 single dose markers respectively. The cumulative length of 'R 570' map was 3144 cM. The cumulative length of 'MQ 76/53' map was 4329 cM. On 'MQ 76/53' map, a gene controlling the red colour of the stalks was identified at 6.5 cM of an AFLP marker and a new brown rust resistance gene was identified at 23.1 cM of an AFLP marker.

The structure of linkage disequilibrium in the population of cultivars was investigated using 1626 AFLP markers, among which 408 have known positions on 'R 570' genetic map. Thus, it was possible to study the relationship between genetic distances and statistical associations between markers in the population of cultivars, calculated using the Fisher exact test. This analysis confirmed that linkage disequilibrium in sugarcane extends over distances of tenth of centiMorgans but drops sharply for distances over 5 cM. This order of magnitude indicates that genome-wide association studies are achievable in sugarcane.

Field trials and greenhouse trials using different inoculation methods were conducted in order to characterize the resistance of the 198 progeny clones from the bi-parental population. The distribution of disease scores observed in all those trials were highly unbalanced toward the resistant parent indicating the segregation of multiple dominant resistance factors. A QTL detection was performed using the 1666 available markers allowing the identification of a few genomic zones with small effects. Besides, the association study performed in the population of cultivars (constituted of two sub-population, one highly susceptible to smut and the other highly resistant to smut) revealed interesting haplotypes associated with resistance. Two QTLs have been detected through both approaches. The progress obtained toward a better understanding of the genetic determinism of sugarcane resistance to smut are modest. The potential of association studies in sugarcane appeared interesting although much more markers would have to be used to make the most of it.

Key words : *Saccharum*, *Ustilago scitaminea*, smut disease, Polyploidy, genetic mapping, QTL, Linkage disequilibrium, association study, AFLP

Résumé:

Le charbon de la canne à sucre, causé par *Ustilago scitaminea* Syd., est présent dans toutes les zones de culture de la canne à l'exception de la Papouasie Nouvelle Guinée, Fidji et de la zone est de l'Australie. La sélection pour la résistance au charbon, caractère assez héritable, est efficace mais nécessite des essais lourds et le contrôle génétique de la résistance n'est pas connu. Pour identifier les zones du génome impliquées dans la résistance au charbon, deux stratégies ont été mises en oeuvre (1) la cartographie de QTL dans un croisement bi-parental entre un cultivar résistant 'R 570' et un clone extrêmement sensible 'MQ 76/53', évalué à la Réunion pour la résistance au charbon (2) l'étude d'associations dans une population de cultivars évaluée au Burkina Faso pour la résistance au charbon.

En préalable, la structure génétique mondiale de *U. scitaminea* a été étudiée. Pour cela, 142 isolats, issus de téliospores collectées dans 15 pays à travers le monde, ont été analysés à l'aide de 17 marqueurs microsatellites. Tous ces isolats sauf un étaient homozygotes à tous les loci étudiés ce qui indique que le mode de reproduction très préférentiel de *U. scitaminea* est l'autofécondation. En Amérique et en Afrique, la diversité génétique est extrêmement faible et tous les isolats se sont révélés appartenir à une unique lignée mondiale consanguine. Cette lignée consanguine a aussi été trouvée en Asie, continent sur lequel a été détecté l'essentiel de la diversité génétique du champignon. Ces observations indiquent que *U. scitaminea* est originaire d'Asie. La population bi-parentale et la population de cultivars utilisées pour étudier le déterminisme génétique de la résistance ont donc été évaluées vis-à-vis d'isolats de la lignée mondiale.

Les cartes génétiques des parents du croisement bi-parental, 'R 570' et 'MQ 76/53', ont été construites à partir d'une population de 198 descendants. Mille six cent soixante six marqueurs polymorphes ont été produits à l'aide de 37 couples d'amorces AFLP, 46 marqueurs SSRs et 9 sondes RFLP. L'analyse des liaisons en couplage a permis de construire 86 groupes de co-ségrégation pour 'R 570' (taille cumulée = 3144 cM) et 105 groupes de co-ségrégation chez 'MQ 76/53' (taille cumulée = 4329 cM) qui rassemblent respectivement 424 et 536 marqueurs. Sur la carte génétique de 'MQ 76/53', un gène associé à la couleur rouge des tiges a été identifié à 6.5 cM d'un marqueur AFLP et un nouveau gène de résistance à la rouille a été identifié à 23.1 cM d'un marqueur AFLP.

Dans la population de cultivars, la structure du déséquilibre de liaison a été étudiée à l'aide de 1626 marqueurs AFLP dont 408 sont cartographiés chez 'R 570'. Pour ces derniers, il a donc été possible d'établir la relation entre distances génétiques et associations statistiques entre marqueurs dans la population de cultivars, mesurées à l'aide du test exact de Fisher. Cette analyse confirme que le déséquilibre de liaison chez la canne est maintenu sur des distances pouvant atteindre plusieurs dizaines de cM mais l'intensité du déséquilibre de liaison baisse fortement lorsque les distances génétiques dépassent 5 cM. La structure du déséquilibre de liaison chez la canne à sucre apparaît donc bien adaptée pour les études d'associations à l'échelle de la totalité du génome.

Des essais au champ et en serre utilisant différentes méthodes d'inoculation ont permis de caractériser la résistance des 198 clones de la descendance bi-parentale. Les ségrégations observées dans tous ces essais sont très fortement décalées vers le parent résistant indiquant la présence de plusieurs facteurs de résistance dominants. La détection de QTL, à l'aide des 1666 marqueurs polymorphes disponibles, a permis d'identifier un certain nombre de QTLs à effets faibles impliqués dans la résistance. Par ailleurs, dans la population de cultivars (constituée de deux sous-populations, l'une très résistante et l'autre très sensible au charbon) l'étude des associations avec la résistance au charbon a permis de révéler un certain nombre d'haplotypes intéressants. Deux zones du génome impliquées dans la résistance sont communes entre ces deux approches. Bien que les progrès obtenus dans la compréhension du déterminisme génétique de la résistance au charbon restent modestes, le potentiel des études d'associations chez la canne à sucre est apparu intéressant. Cependant, son exploitation optimale nécessitera un changement d'échelle en ce qui concerne le nombre de marqueurs à analyser.

Mots clés : *Saccharum*, *Ustilago scitaminea*, maladie du charbon, Polyploïdie, cartographie génétique, QTL, déséquilibre de liaison, étude d'associations, AFLP