

**ECOLE NATIONALE SUPERIEURE D'AGRONOMIE DE MONTPELLIER
AGRO MONTPELLIER**

THESE

**Présentée à l'Ecole Nationale Supérieure Agronomique de Montpellier
pour obtenir le DIPLOME DE DOCTORAT**

FORMATION DOCTORALE: Ressources Phytogénétiques et Interactions Biologiques
ECOLE DOCTORALE: Biologie des Systèmes Intégrés, Agronomie-Environnement
LABORATOIRE D'ACCUEIL: UMR 1097 Diversité et Génomes des Plantes Cultivées, IRD Montpellier

**Origine de la diversité de plantes domestiquées par la reproduction
végétative en Amérique du Sud: reproduction sexuée résiduelle et
introgression d'espèces sauvages éloignées. Exemples de l' arracacha
(*Arracacia xanthorrhiza* Banc., Apiaceae) et du manioc (*Manihot esculenta*
Crantz, Euphorbiaceae)**

présentée et soutenue publiquement

par

Eduardo MORILLO V.

Soutenue publiquement le 4 décembre 2006 devant le jury composé de :

M. Jean-Pierre BAUDOIN , Professeur, Université de Gembloux, Belgique	Rapporteur
Mme Mathilde BRIARD , Professeur, INH, Angers	Rapporteur
M. André CHARRIER , Professeur, Agro.M	Examineur
M. Doyle McKEY , Professeur, Université Montpellier II	Examineur
M. Patrice THIS , Directeur de Recherche, INRA, Montpellier	Examineur
M. Gérard SECOND , Directeur de Recherche, IRD, Montpellier	Directeur de thèse

REMERCIEMENTS

Cette thèse est l'aboutissement de presque quatre années de travail. Je vais essayer de remercier les gens qui ont collaboré à ce travail résultant d'une coopération entre un chercheur du sud et une équipe de l'IRD (Institut de Recherche pour le Développement).

Tout d'abord je tiens à remercier l'IRD, plus concrètement l'équipe Dynamique de la Diversité des Plantes Cultivées (DYNADIV), pour m'avoir donné la chance de poursuivre en thèse mon DEA réalisé en 2001 dans les mêmes conditions. Je voudrais exprimer ma reconnaissance à Francis Khan qui a soutenu tout au début, l'idée de travailler sur une plante native de l'Equateur et de l'intégrer à la thématique de recherche de l'équipe DYNADIV.

Merci aux différents membres de l'équipe DYNADIV que j'ai pu rencontrer pendant mon stage de trois ans à l'IRD Montpellier, particulièrement à Gérard Second pour accepter la direction de cette thèse et pour son aide à l'édition de ce manuscrit. Un grand merci aussi à Jean Louis Pham, co-directeur de cette thèse et responsable de l'équipe, pour m'avoir accueilli dans l'équipe, ses conseils toujours opportuns et pour m'avoir donné la chance de participer à plusieurs congrès scientifiques en France. Je tiens à les remercier tous les deux pour avoir soutenu cette recherche par le budget de l'équipe et m'avoir donné toutes les facilités pour la développer. A propos du manuscrit, je remercie les deux rapporteurs, Mathilde Briard et Jean Pierre Baudouin ainsi que tous les membres du jury.

Je tiens ensuite à remercier Yves Vigouroux, Cédric Mariac, Jean Christophe Pintaud, Serge Tostain, Silvie Pounhet, les autres thésards et stagiaires avec qui nous avons partagé des moments au labo, au bureau ou ailleurs: Thomas Couvreur, Rommel Montufar, Clément Agbangla, Nora Scarcelli... Merci aussi à Ange Marie Risterucci pour sa collaboration lors du développement de la banque microsat au CIRAD en 2002.

Merci au petit groupe de « latinos » avec qui avons partagé des moments d'amitié loin de nos pays, Juan Carlos et Diana (Colombia), Edgardo (Costa Rica), José Bustamante (Venezuela), Leandro (Brasil), Rommel et Edwin (Equateur)...

Merci à Sandra et Sandrine, stagiaires qui ont travaillé avec moi pendant la première et la deuxième année, respectivement, et à Carol, pendant la dernière année.

Merci aux gens d'autres équipes de l'UMR avec qui j'ai pu partager souvent la paillasse, demander un avis ou conseil, ou tout simplement plaisanter. Un grand merci à Serge Hamon pour son accueil, amitié et aussi pour m'avoir donné l'opportunité de présenter une partie de mon travail au congrès de Biotechnologie au Togo en 2004 et à Quito en 2005, lors de l'anniversaire des 30 ans de présence de l'IRD en Equateur. Merci aussi à Mme. Ferard Pouget, responsable du bureau d'accueil d'étrangers à l'IRD pour sa disposition toujours positive.

Merci à la représentation de l'IRD à Quito et aux représentants que j'ai pu rencontrer, d'abord M. Francis Khan, puis M. Pierre Gondard, de m'avoir fourni toutes les facilités logistiques et administratives pour mes déplacements vers la France ou en Equateur pendant mes prospections de terrain. Je ne peux pas oublier de remercier Elicio Tapia, une personne qui mérite tout mon respect pour son énorme qualité sur le terrain. Merci à Steen Randers,

chercheur danois qui sans aucun intérêt personnel m'as fourni du matériel précieux pour mes analyses génétiques et qui m'a accueilli chez lui à Copenhague pendant ma visite en 2004.

J'exprime ma reconnaissance au DSF (Département Soutien et Formation) pour l'attribution d'une bourse de recherche. Merci aussi au FUNDACYT (*Fundación para la Ciencia y la Tecnología*) pour l'attribution d'une bourse supplémentaire qui m'as permis de prolonger mon séjour à Montpellier en 2006. Je remercie l'INIAP, en particulier le Dr. Gustavo Enriquez, Directeur général en 2002, pour m'avoir attribué une mise en disponibilité qui m'a permis de récupérer mon poste après mon séjour en France.

Enfin, je remercie très vivement ma famille, mes parents, mes frères, ma sœur, ma famille politique, pour leur appui inconditionnel pendant tout ce séjour.....

Puis j'ai ne pas de paroles pour remercier Mariana, ma femme, pas seulement pour son appui à tout moment pendant ces trois ans passés en France, sinon pour m'avoir fait le cadeau le plus précieux, mon fils Dario Leonel.....

Eduardo Morillo V.

SOMMAIRE

RESUME.....	10
PRESENTATION: Situation du sujet et structure de la thèse.....	11
CHAPITRE I: INTERACTIONS GENETIQUES AU SEIN DU COMPLEXE <i>A. xanthorrhiza</i> ET ORIGINE DE LA FORME CULTIVEE.....	14
1) SYNTHÈSE BIBLIOGRAPHIQUE.....	15
1.1. La domestication végétale dans les Andes: le cas des racines et tubercules.....	15
1.2. Présentation de l'arracacha (<i>A. xanthorrhiza</i> Banc.).....	16
1.2.1. Aspects généraux.....	16
1.2.2. Aspects biologiques.....	19
1.2.3. Agronomie et utilisation.....	21
1.2.4. Diversité et état des ressources génétiques.....	23
1.3. Taxonomie.....	25
1.3.1. La famille des Apiaceae.....	25
1.3.2. Le genre <i>Arracacia</i>	26
1.3.3. Les espèces tubéreuses d' <i>Arracacia</i>	27
1.4. Etat des connaissances sur l'origine de la domestication.....	30
2) OBJECTIFS ET DEMARCHE METHODOLOGIQUE.....	33
3) MATERIELS ET METHODES.....	34
3.1. Prospections et collectes.....	34
3.1.1. Observations de terrain.....	35
3.2. Matériel végétal.....	37
3.3. Analyses génétiques.....	38
3.3.1. Création d'une banque SSR et développement des marqueurs.....	38
3.3.2. Génotypage par microsatellites (SSR).....	39
3.3.3. Génotypage par AFLPs.....	40
3.3.4. Génotypage de l'ADN chloroplastique (ADNcp).....	41
3.4. L'analyse des données.....	41
4) RESULTATS ET DISCUSSION.....	45
I: Identification des marqueurs SSR chez <i>A. xanthorrhiza</i>	46
II: Diversité génétique et interactions au sein du complexe sauvage-cultivé de <i>A. xanthorrhiza</i> estimés par l'emploi des SSR.....	49
III: Peut-on préciser l'origine de la forme cultivée?.....	57
ARTICLE: "Cryptic genetic structure and origin of the Andean root crop arracacha (<i>Arracacia xanthorrhiza</i> Bancr. Umbellifereae) revealed by the analysis of AFLP and chloroplast DNA polymorphism analysis".....	58

IV: La diversité génétique chez l'arracacha cultivée	68
5) DISCUSSION GENERALE DU CHAPITRE I.....	72
CHAPITRE II: HYPOTHESE DE L'INTROGRESSION GENETIQUE A PARTIR DE FORMES SAUVAGES DIVERGENTES: Des espèces éloignées sont-elles représentées dans le génome de la forme cultivée?	
<i>Approche expérimentale chez le manioc et l'arracacha.....</i>	<i>77</i>
1) ANTECEDENTS ET OBJECTIFS.....	78
1.1. Le cas de l'arracacha.....	80
1.2. Le cas du manioc.....	81
2) MATERIELS ET METHODES.....	83
2.1 Matériel végétal et locus microsatellites à analyser.....	84
2.2 Approche méthodologique pour tester l'hypothèse d'introgession.....	86
2.2.1 Allongement de la séquence, marche sur le chromosome	86
2.2.2 Validation de l'allongement.....	87
2.3 Isolement et obtention de séquences d'allèles SSR allongés	87
2.4 Analyse de séquences: Reconstruction phylogénétique de locus SSR allongés.....	88
3) RESULTATS	89
3.1 Allongement des allèles SSR par marche chromosomique.....	89
3.2 Allèles isolés et séquencés pour l'analyse phylogénétique.....	90
3.3 Polymorphisme des séquences flanquantes des locus SSR.....	91
3.4 Analyse phylogénétique des locus SSR.....	91
4) DISCUSSION	96
CONCLUSIONS ET DISCUSSION GENERALE.....	105
I. Récapitulatif des principaux résultats.....	106
II. Qu'apporte notre étude chez l'arracacha?.....	109
A. Des connaissances chez une plante dite mineure originaire d'un centre majeur de domestication.....	109
B. Un scénario original de domestication: sélection d'un type sauvage avec une aptitude naturelle à la reproduction végétative.....	110
C. Flux de gènes et pratiques paysannes traditionnelles dans l'agrosystème traditionnel andin.....	111
III. L'introgession comme mécanisme de diversification des formes cultivées.....	111
IV. Implications des introgressions larges chez les formes cultivées: présence de transgènes dans les agrosystèmes traditionnels.....	114
V. Implications pour la conservation des ressources génétiques de l'arracacha.....	115
REFERENCES BIBLIOGRAPHIQUES.....	117
TABLEAUX.....	124
FIGURES.....	142
ANNEXES.....	173

LISTE DES TABLEAUX

Tableau 1. Principales plantes domestiquées dans la région andine.....	128
Tableau 2. Quelques caractéristiques des plantes andines à racines et tubercules.....	129
Tableau 3. Liste des 102 allèles observés parmi 178 formes cultivées (73) et sauvages (105) de <i>A. xanthorrhiza</i> sur 11 locus SSR.....	130
Tableau 4. Statistiques générales de la diversité SSR au sein du complexe sauvage et cultivé d' <i>A. xanthorrhiza</i>	130
Tableau 5. Fréquence de 31 allèles partagés entre les formes sauvages et cultivées d' <i>A. xanthorrhiza</i>	131
Tableau 6. Fréquences des allèles communs entre formes cultivées et sauvages d' <i>A. xanthorrhiza</i> regroupées selon leur origine géographique.....	132
Tableau 7. Valeurs observées et attendues d'hétérozygotie sur 11 locus SSR chez 178 formes tubéreuses cultivées et sauvages d'arracacha.....	133
Tableau 8. Valeurs d'appartenance (q) des individus assignés par STRUCTURE aux 2 ou 3 groupes supposés selon les données génotypiques sur 11 locus SSR.....	134
Table I-1. Plant material used in the AFLP analysis.....	135
Table I-2. Numbers of AFLP bands scored for each primer combination in 59 samples of wild and cultivated <i>A. xanthorrhiza</i> and related species.....	136
Table I-3. Pairwise matrix of Nei's genetic distance from 235 scored AFLP bands in cultivars and wild Arracacia species or varieties.....	136
Table I-4. Nei's genetic distance from 198 AFLP bands observed in the four groups of accessions distinguished among the <i>A. xanthorrhiza</i> complex by UPGMA and PCO.....	136
Tableau 9. Statistiques générales de la diversité SSR au sein de la forme cultivée d' <i>A. xanthorrhiza</i>	137
Tableau 10. Résultats préliminaires de séquençage comparatif des régions flanquantes des 4 locus SSR chez le manioc.....	138
Tableau 11. Résultats obtenus de l'allongement des locus SSR par marche chromosomique chez le manioc et l'arracacha.....	138
Tableau 12. Allèles représentés dans les analyses phylogénétiques des séquences des locus AxC87 et AxC29 chez <i>A. xanthorrhiza</i>	139
Tableau 13. Allèles enregistrés pour les accessions de manioc et formes sauvages sur les locus SSRY-3, SSRY-9 et SSRY-40 en gel et <i>in silico</i>	140
Tableau 14. Allèles représentés dans l'analyse phylogénétique des séquences du locus SSRY-3 du manioc.....	141
Tableau 15. Allèles représentés dans l'analyse phylogénétique des séquences du locus SSRY-9 du manioc.....	142
Tableau 16. Allèles représentés dans l'analyse phylogénétique des séquences du locus SSRY-40 du manioc.....	143
Tableau 17. Résultats de l'analyse phylogénétique par MP des séquences d'allèles SSR allongés chez le manioc et l'arracacha.....	144

LISTE DES FIGURES

Figure 1.	Plantes à tubercules et racines originaires des Andes.....	146
Figure 2.	Généralités sur l'arracacha.....	147
Figure 3.	Exemples de cultures de l'arracacha.....	148
Figure 4.	Floraison spontanée chez l'arracacha cultivée.....	149
Figure 5.	Formes horticoles distinguées chez l'arracacha cultivée.....	150
Figure 6.	Distribution géographique du genre <i>Arracacia</i> en Amérique du sud.....	151
Figure 7.	Populations naturelles d'espèces sauvages de <i>Arracacia</i> en Equateur.....	152
Figure 8.	Type sauvage polycarpique d' <i>A. xanthorrhiza</i> (<i>A. x. var. andina</i>).....	153
Figure 9.	Type sauvage monocarpique de <i>A. xanthorrhiza</i> (<i>A. x. var. monocarpa</i>).....	154
Figure 10.	Morphologie de l'espèce tubéreuse <i>A. incisa</i>	154
Figure 11.	Origine de la forme cultivée d'arracacha selon Blas (2005).....	155
Figure 12.	Carte des origines des formes sauvages et cultivées de <i>A. xanthorrhiza</i> analysées dans la présente étude.....	156
Figure 13.	Principe technique de la méthode de développement des marqueurs microsatellites.....	157
Figure 14.	Exemples d'amplification de 4 locus microsatellites chez <i>A. xanthorrhiza</i>	158
Figure 15.	Exemples des profils AFLPs obtenus dans l'étude de diversité parmi le complexe <i>A. xanthorrhiza</i>	159
Figure 16.	Arbre NJ basé sur la distance DAS qui montre les relations génétiques entre 178 formes tubéreuses cultivées et sauvages d'arracacha sur la base de la diversité observée sur 11 locus SSR.....	160
Figure 17.	Analyses multivariées fondées sur la diversité SSR chez les formes cultivées et sauvages d'arracacha.....	161
Figure 18.	Arbre NJ basé sur la distance DAS sur 11 locus SSR qui montre les relations génétiques entre 178 formes cultivées et sauvages d'arracacha et 13 hybrides cultivés x <i>A. x. var. andina</i>	162
Figure 19.	Analyse PCO basée sur la diversité observée à 11 locus SSR chez 178 formes tubéreuses cultivées et sauvages d'arracacha, et représentation graphique du test d'assignation réalisé avec STRUCTURE assumant K=3.....	163
Figure I-1.	UPGMA tree obtained from 235 AFLP bands scored in a set of cultivated arracacha and wild relatives.....	164
Figure I-2.	Plane defined by the first two axes of a PCO based on 198 AFLP bands.....	165
Figure I-3.	Parsimonious phylogramme of arracacha accessions obtained through the phylogenetic analysis of 24 polymorphisms observed in 2000 bp of chloroplast ADN.....	165
Figure 20.	Arbre UPGMA qui regroupe les 72 accessions cultivées selon le polymorphisme observé sur 11 locus SSR (31 allèles).....	166
Figure 21.	Analyses multivariées basées sur la diversité SSR et AFLPs parmi les formes cultivées d'arracacha.....	167
Figure 22.	Scénario hypothétique sur l'origine de la forme domestiquée de <i>A. xanthorrhiza</i> sur la base de l'information donnée par les marqueurs SSR, AFLP et l'ADN chloroplastique.....	168
Figure 23.	Principe de la méthode de marche sur le chromosome employée par le kit <i>Genome Walker Universal</i>	169
Figure 24.	Etapes pour l'isolement d'allèles SSR allongés.....	170

Figure 25. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 24 allèles du locus SSR AxC87 isolés chez <i>A. xanthorrhiza</i> et <i>A. incisa</i>	171
Figure 26. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 25 allèles du locus SSR AxC29 isolés chez <i>A. xanthorrhiza</i> et espèces sauvages apparentées.....	172
Figure 27. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 40 allèles du locus SSR Y-3 chez le manioc.....	173
Figure 28. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 27 allèles du locus SSR Y-9 chez le manioc.....	174
Figure 29. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 54 allèles du locus SSR Y-40 chez le manioc.....	175

LISTE DES ANNEXES

Annexe 1.	Liste d'échantillons utilisés pour l'étude de diversité de l'arracacha.....	177
Annexe 2.	Primer Note publiée dans <i>Molecular Ecology Notes</i>	181
Annexe 3.	Séquences des clones SSR allongés par marche chromosomique.....	185
Annexe 4.	Polymorphismes de séquences observés chez l'arracacha au niveau des régions flanquantes du locus SSR AxC87 allongé par marche chromosomique.....	188
Annexe 5.	Polymorphismes de séquences observés chez l'arracacha au niveau des régions flanquantes du locus SSR AxC29 allongé par marche chromosomique.....	189
Annexe 6.	Polymorphismes de séquences observés chez le manioc au niveau des régions flanquantes du locus SSRY-3 allongé par marche chromosomique.....	190
Annexe 7.	Polymorphismes de séquences observés chez le manioc au niveau des régions flanquantes du locus SSRY-9 allongé par marche chromosomique.....	191
Annexe 8.	Polymorphismes de séquences observés chez le manioc au niveau des régions flanquantes du locus SSRY-40 allongé par marche chromosomique.....	192
Annexe 9.	Mémoire de DEA (Morillo 2001).....	193

RESUME

Origine de la diversité de plantes domestiquées par la reproduction végétative en Amérique du Sud: reproduction sexuée résiduelle et introgression d'espèces sauvages éloignées. Exemples de l'arracacha (*Arracacia xanthorrhiza* Banc., Apiaceae) et du manioc (*Manihot esculenta* Crantz, Euphorbiaceae)

Nous nous sommes intéressés à l'étude de la domestication de l'arracacha, une carotte présentant l'originalité d'être reproduite végétativement et d'être native de la région andine tropicale. Le polymorphisme de cette espèce était réputé faible. Dans un premier temps, les marqueurs microsatellites sont donc apparus les marqueurs de choix et ont été développés. Une diversité allélique significative de la forme cultivée mais une diversité plus grande des formes sauvages con-spécifiques sont observées. Cependant, une fraction de cette diversité de la forme cultivée ne s'observe pas parmi les formes sauvages. Un criblage plus large du génome, avec des marqueurs AFLPs et un marqueur de l'ADN chloroplastique, a cependant permis d'identifier la forme sauvage la plus probablement à l'origine de la forme cultivée. En accord avec un patron de domestication fréquemment rencontré parmi les multiples espèces à racines et tubercules domestiquées de la région andine, la forme pérenne serait l'ancêtre direct. Une divergence cryptique entre deux groupes clairement définis de variétés cultivées a par ailleurs été mise en évidence. Elle est attribuable à une introgression génétique de la forme monocarpique sauvage postérieurement à la domestication. Mais une fraction de la diversité génétique de la forme cultivée reste non attribuable à ces formes ancestrales et conduit à tester l'hypothèse d'autres espèces plus éloignées représentées dans le génome de la forme cultivée par introgression génétique. L'approche méthodologique est de repérer des locus microsatellites présumés introgressés et d'en séquencer la région flanquante du motif. Une marche génomique permet d'allonger cette région et d'augmenter le polymorphisme observé. Le travail a été conduit à la fois sur l'arracacha et sur le manioc (suite aux résultats relatés dans le mémoire de DEA). La phylogénie des différents allèles observés à ces locus a permis d'en déterminer l'origine et, dans certains cas parmi les deux genres étudiés, de valider l'hypothèse d'allèles introgressés à partir d'espèces éloignées. L'introgression est aussi démontrée réciproque car des allèles de la forme cultivée sont retrouvés dans des espèces sauvages, ce qui est important des points de vue de l'impact potentiel de la biotechnologie dans les agroécosystèmes et de l'évolution de l'agrobiodiversité. Dans le cas d'un locus de l'arracacha, l'introgression aurait conduit à l'existence de deux groupes d'allèles représentés à la fois parmi les cultivars, ses formes ancestrales et chez une espèce éloignée. Des allèles recombinants sont observés parmi cette espèce sauvage, impliquant des recombinaisons répétées entre les deux groupes d'allèles. L'hypothèse d'une origine anthropique de cette espèce éloignée peut être avancée.

Mots clés: arracacha, domestication, reproduction végétative, flux de gènes, région andine, manioc, introgression

PRESENTATION

Situation du sujet et structure de la thèse

Cette thèse aborde deux aspects liés à la domestication des plantes cultivées à propagation végétative: un premier porte sur l'étude de la diversité génétique du complexe sauvage-cultivé d'une plante andine, l'arracacha (*A. xanthorrhiza*), et un deuxième aborde la possible contribution d'espèces éloignées dans la constitution génétique des formes cultivées de l'arracacha et du manioc (pour ce dernier, l'étude de la diversité génétique du complexe sauvage-cultivé en vue de la caractérisation d'introgessions éloignées a été effectuée au cours du stage de DEA –mémoire en annexe-).

Chapitre I: Interactions génétiques au sein du complexe sauvage-cultivée de *A. xanthorrhiza* et origine de la forme cultivée

La région andine en Amérique du sud représente un centre majeur d'origine des plantes cultivées et comprend des agrosystèmes traditionnels uniques au monde. On observe dans cette région une diversité des plantes cultivées natives à racines et tubercules inégalée ailleurs. A l'exception de la pomme de terre dont le succès est planétaire, ces plantes restent très peu étudiées alors qu'elles sont l'objet d'une importante érosion génétique. L'agrosystème andin traditionnel est complexe et peu d'études se sont engagées à comprendre les mécanismes opérant dans la génération et la conservation de la diversité des plantes cultivées (Ex. Quiros et al. 1992). Depuis une vingtaine d'années, les programmes de recherche agronomique de la région ont assemblé des collections de ressources génétiques de ces plantes. Des études visant à une description de la diversité ont été menées. Les résultats montrent en général une diversité morphologique importante au sein des formes cultivées (notamment au niveau des tubercules) mais l'origine de cette diversité (mutations, recombinaison, épigénétique) reste peu documentée. Par ailleurs, peu de travaux incluant des formes sauvages apparentées ont été menés, alors que leur étude peut nous renseigner sur la domestication des plantes dans cette région.

Dans ce chapitre, nous nous intéresserons à l'étude de l'une de ces plantes, l'arracacha (*Arracacia xanthorrhiza* Banc.) qui a l'avantage d'être présente à l'état sauvage et cultivé en Equateur. L'arracacha est une Ombellifère (Apiaceae) cultivée à fort potentiel agronomique mais sa diversité génétique est estimée faible. Une base génétique insuffisante est aussi invoquée pour expliquer une adaptabilité limitée responsable des échecs d'introduction sur

d'autres continents. La forme domestiquée est pérenne et sa propagation se fait exclusivement par voie asexuée. L'intérêt scientifique pour les espèces sauvages est récent : validation de la taxonomie, domestication et potentialité d'utilisation de ces ressources génétiques. Sur cette base, notre étude a les objectifs suivants:

- analyser la diversité des formes sauvages et cultivées d'*A. xanthorrhiza*
- analyser les relations génétiques entre les trois sous-espèces au sein du complexe *A. xanthorrhiza* (selon la classification de Knudsen 2003) dans l'aire probable de domestication (centre/sud de l'Equateur et nord du Pérou)
- proposer un scénario de l'origine de la forme cultivée
- déterminer les relations génétiques et phylogénétiques du complexe *A. xanthorrhiza* avec d'autres formes sauvages du genre présentes dans l'aire probable de domestication

La présentation de ce chapitre est organisée selon le plan suivant:

1) Dans une partie introductive on inclut une révision bibliographique afin de présenter l'arracacha car il s'agit d'une plante dite mineure, largement inconnue dans le milieu scientifique.

2) Dans la partie Matériel et Méthodes on présente le matériel végétal étudié et les outils d'analyse employés.

3) La partie Résultats est structurée en quatre sections. Après chaque section on inclut une discussion des résultats.

- une première section présente les résultats de l'isolement et de l'identification des marqueurs microsatellites (SSR),
- une deuxième présente les résultats de l'analyse SSR dans le complexe sauvage-cultivé,
- une troisième comprend un aperçu plus large du génome à l'aide des marqueurs AFLPs complétée par une analyse phylogénétique sur la base de l'ADN chloroplastique,
- une quatrième aborde la diversité génétique chez l'arracacha cultivée et sa structuration

4) Une discussion générale relative à la diversité et à la domestication de *A. xanthorrhiza*

Chapitre II

Chez l'arracacha et le manioc, l'observation d'allèles SSR bien distincts chez la forme cultivée qui ne sont pas présents chez ses homologues sauvages, suggère une possible participation d'autres types sauvages plus divergents. L'objectif de ce chapitre est donc de déterminer si les génomes d'autres espèces que la forme ancestrale sont représentés chez les formes cultivées. Pour cela on a développé une approche méthodologique innovante qui consiste à allonger par marche chromosomique la zone flanquante du locus en question afin d'augmenter l'information obtenue pour construire des phylogénies entre les séquences d'un groupe d'allèles.

Ce chapitre est présenté de la manière suivante:

- 1) Une partie introductive sur le rôle des introgressions dans l'origine des formes cultivées; puis le cas particulier de l'arracacha et le manioc
- 2) Une partie expérimentale où on présente l'approche méthodologique adoptée pour tester l'hypothèse d'introgressions larges
- 3) Une partie de résultats avec la phylogénie d'allèles de locus SSR observés dans des individus représentatifs d'espèces sauvages et de variétés cultivées présumées introgressées pour ces locus.
- 4) Une discussion de nos résultats

Ce document de thèse se terminera par une partie de discussion générale, conclusions et perspectives.

CHAPITRE I.

INTERACTIONS GENETIQUES AU SEIN DU COMPLEXE *A. xanthorrhiza* ET ORIGINE DE LA FORME CULTIVEE

1) SYNTHESE BIBLIOGRAPHIQUE

1.1. La domestication végétale dans les Andes: le cas des racines et tubercules

L'imposant massif des Andes sud-américaines dont une face est orientée sur l'océan Pacifique et l'autre sur le bassin amazonien présente une grande diversité de microclimats. Cette situation a permis que des environnements très différents puissent se trouver à courte distance, conduisant à une grande diversité floristique. Les débuts de l'agriculture dans cette région remontent probablement à plus de 10 000 ans. Piperno et Stochert (2003) apportent des évidences d'une émergence indépendante de l'agriculture sur les côtes de l'Equateur avec l'utilisation de formes domestiquées de Cucurbitacées dont l'origine est estimée entre 10 000 et 12 000 ans. De la même manière dans les Andes, des pratiques de domestication anciennes ont probablement eu lieu qui, compte tenu de la grande diversité disponible, ont généré un vaste groupe de plantes cultivées mentionné dans le Tableau 1. Il a été proposé que les processus développés à cette fin (domestication) appartiennent à un système cohérent, inspiré de la cosmogonie andine, qui intègre le labour dans son contexte culturel et religieux. Les nombreuses communautés indigènes possèdent leurs propres institutions, rituels, religions, langages, culture et lois ainsi que leurs propres modes d'être, de connaître et de relation au monde. Tous ces acquis ont été essentiels au développement du système d'agriculture pratiqué avec succès par les peuples andins (Gonzalez et al. 1998). Les plantes à racines et tubercules forment un groupe emblématique caractérisant l'agriculture dans cette région andine: en effet, en dehors du vaste complexe des pommes de terre, on compte au moins neuf autres espèces cultivées pour leurs organes souterrains (Arbizu et al. 1997; Figure 1). En voici une brève description (CIP 2005) :

Ullucu (*Ullucus tuberosus* C.), famille des Basellacées, plante herbacée originaire des hauts plateaux andins (altiplano bolivien-péruvien) qui produit de petits tubercules très appréciés.

Mashua (*Tropaeolum tuberosus* R. & P.), espèce herbacée de la famille de Tropaéolacées qui produit de grands tubercules de diverses formes et couleurs. Sa culture est rustique. Les tubercules ont une saveur astringente caractéristique.

Oca (*Oxalis tuberosa* Mol.) Oxalidaceae, aussi cultivée pour ses tubercules qui sont très appréciés. Cette plante peut être cultivée au-dessus de 4000 m. Des références historiques signalent sa forte présence avant l'arrivée des européens.

Yacon (*Smallanthus sonchifolius* P. & E.), membre de la famille des tournesols (Compositae). A la différence du tournesol, elle n'est pas propagée par graines. Elle est cultivée pour ses

racines tubéreuses souvent appelés « fruit de l'indien » car elles sont consommées en frais. Cette plante pousse au long des vallées andines jusqu'au niveau des 3200 m.

Mauka (*Mirabilis expansa* R. &P.), famille des Nyctaginacée. Cultivée pour ses racines tubéreuses connues comme « yuquilla ». On la trouve entre 2000 et 3000m, toujours associée à d'autres plantes cultivées. On pense que sa culture est ancienne mais elle est très restreinte à l'heure actuelle. Alors qu'elle se reproduit facilement par graines, elle est aussi propagée végétativement par bouturage.

Maca (*Lepidium meyenii* Ref): c'est la seule crucifère cultivée des Amériques. Elle est une des rares plantes cultivées au-dessus des 4000m. Il existe un intérêt récent pour cette plante à cause de propriétés sur la fertilité humaine et le système immunologique qui lui ont été attribuées.

Ahipa (*Pachyrhizus ahipa*): c'est une légumineuse cultivée entre 1000 et 2300m pour sa racine très riche en protéines. Sa culture est restreinte à la Bolivie et elle est très menacée d'érosion génétique. Elle est connue comme « *Pachyrhizus andin* » car d'autres espèces du genre sont cultivées dans des régions plus tropicales.

Achira (*Canna edulis*): c'est une Cannaceae cultivée dans les vallées andines pour ses rhizomes qui contiennent un amidon de très bonne qualité.

Arracacha (*Arracacia xanthorrhiza* B.), ombellifère à fort potentiel agronomique par sa rusticité et qualité alimentaire. Consommée traditionnellement dans les pays andins et certains de la région Caraïbe, son introduction il y a une centaine d'années au Brésil s'est avérée un succès. Elle est présentée en détail dans la suite.

On estime que la culture de ces espèces à racines et tubercules a eu un rôle très important dans la sécurité alimentaire des peuples indigènes précolombiens mais, à ce jour, elles restent sous-utilisées. On note néanmoins plusieurs cas d'une introduction réussie dans d'autres régions. On peut citer le cas de l'oca en Nouvelle Zélande, de l'achira au Vietnam, du yacon au Japon, de l'arracacha au Brésil (Hermann 1997).

1.2. Présentation générale de l'arracacha (*Arracacia xanthorrhiza* Banc.)

1.2.1. Aspects généraux

L'arracacha est une ombellifère cultivée reproduite végétativement pour ses racines tubéreuses (Figure 2). Selon Bukasov (1981) l'arracacha pourrait être parmi les plantes les plus anciennement cultivées de l'Amérique. Sa culture fut remarquée par les Européens dès leur arrivée en Amérique. Les Espagnols signalent en effet cette plante lors de leur traversée

de ce qui est aujourd'hui la Colombie, l'Equateur, le Pérou et la Bolivie (León 1964). La première description botanique de cette plante fut réalisée en 1825 mais bien plus tard en ce qui concerne ses formes sauvages. La culture de l'arracacha est encore largement restreinte à l'Amérique du sud (Andes et sud-est du Brésil) en ce qui concerne son actuelle distribution; par ailleurs on la retrouve dans certains pays centro-américains et des Caraïbes, mais la production n'est significative qu'au Costa Rica et à Puerto Rico. L'arracacha serait importante pour environ 30 millions de personnes dans les Andes et environ pour un même nombre dans les autres régions confondues. Dans les Andes, elle s'observe en monoculture de petites parcelles n'excédant pas 2 à 3ha ou encore associée à d'autres cultures (Figure 3).

Actuellement, l'arracacha est notamment produite dans quatre pays: Brésil, Colombie, Equateur et Venezuela. Sa surface cultivée serait de quelques 30 000 ha. Il est néanmoins certain que la surface cultivée était plus importante il y a peu. Hodge (1954) et Bristol (1988) signalent une culture intensive de l'arracacha en Colombie par différents groupes indigènes (Chibchas, Pijao, Sibundoy). Bukasov cité par León (1964) signale une surface de 28 000 Ha en Colombie dans les années 30. León (1964) indique que dans certaines régions, elle est même plus appréciée que la pomme de terre. Dans les pays mentionnés, la présence actuelle de l'arracacha dans les marchés urbains est courante et sa consommation est destinée à une bonne partie de la population. Dans les Andes, l'arracacha est aussi cultivée au Pérou, Venezuela et Bolivie mais sa production est généralement consommée en autarcie, seule une plus faible part étant consacrée à la vente sur les marchés locaux.

Au Brésil, la plante a été littéralement adoptée et la production dans ce seul pays égale toute la production dans les pays andins (Tableau 2, Hermann 1997, Gonçalves 2004). Des rapports signalent l'introduction de l'arracacha en Asie à la fin du 19^{ème} siècle (Inde et Sri Lanka, Hermann 1997), Afrique centrale (Burundi, Rwanda et Congo, Hodge 1954) et plus récemment en Australie (Ames 1997 cité par Blas 2005). Néanmoins, la situation actuelle de la culture de l'arracacha dans ces pays n'est pas connue (Hermann 1997). Hodge (1954) regrette qu'une plante avec un tel potentiel n'ait pas abouti à des introductions réussies en Europe. Cependant, des essais n'ont pas manqué: plusieurs essais d'introduction en Europe ainsi qu'aux Etats-Unis ont été signalés mais avec peu de résultats positifs. Hodge (1954) signale qu'aux Etats Unis les meilleurs résultats ont été observés en automne et au printemps, mais la production de racines étant très faible, la plante ne fut pas adoptée.

Plusieurs raisons conduisent à considérer l'arracacha comme l'espèce de plus fort potentiel parmi les plantes dites mineures à racines et tubercules andins: variété d'utilisations culinaires, bonnes qualités organoleptiques, bonne texture. Par ailleurs, son amidon est de digestion facile ce qui fait de l'arracacha un nutriment idéal pour les enfants (l'arracacha est utilisée au Brésil comme ingrédient dans les soupes instantanées destinées notamment à l'alimentation infantile). Le tableau 2 résume quelques propriétés des racines et tubercules andins, et montre que l'arracacha reste une plante sous-utilisée bien que son importance économique dépasse largement celle des autres espèces mentionnées.

Noms vernaculaires

Arracacia xanthorrhiza est connue majoritairement comme « arracacha » (Pérou, Bolivie, Colombie) bien que ce nom ne soit pas reconnu dans d'autres pays comme en Equateur, où on l'appelle « *zanahoria blanca* » (carotte blanche), ou au Venezuela et la région caraïbe (apio criollo) et le Brésil (*mandioquinha salsa* ou « *pastinaca* »)

Le nom *Arracacia* dériverait du mot quechua (langue des Incas) « *racacha* », tandis que le nom « *xanthorrhiza* » veut dire racine jaune en grec, une des trois couleurs avec le blanc et le violet que prend cette racine cultivée. Une diversité de noms est employée par les communautés indigènes. Par exemple, dans le nord du Pérou plus de 50 noms différents ont été répertoriés en langue quechua en ce qui concerne seulement les formes sauvages qui ont des propriétés médicinales (Valderrama et Seminario 2002). Le Pérou a sans doute aussi la plus grande diversité de noms vernaculaires pour l'arracacha. L'expansion de l'empire Inca à la fin du XV^{ème} siècle et donc du quechua, fait penser que plusieurs noms autochtones ont pu être remplacés juste avant l'arrivée des européens. D'autres noms pourraient être dérivés du nom arracacha ; comme « *aricachi* » employé par les indiens Ayoman au Venezuela ou « *arocueche* » par les indiens Muzo de Colombie, mais il n'est pas clair si ces noms datent de la conquête inca ou de leur traversée des Andes (Hermann 1997). Patiño (1964) présente des arguments selon lesquels le nom « *arocueche* » aurait une origine autochtone des indiens Muzo. Cet auteur considère comme moins probable qu'il s'agisse d'une déformation du nom *racacha* et suggère que le mot arracacha aurait son origine dans la région nord andine. Ce serait suite à son expansion dans les Andes centrales que le nom *racacha* aurait été substitué au nom Muzo original.

En général, les noms traditionnels de cultivars font référence à des variantes morphologiques (couleur des racines ou cycle de croissance). Meza (1995) indique neuf noms quechuas

différents pour les variétés de la région de Cuzco au Pérou. En Equateur, les paysannes distinguent seulement les variétés comme « *amarilla, blanca* ou *morada* » (jaune, blanche ou violette) qui correspondent aux couleurs racinaires de la forme cultivée. Hodge (1954) rapporte des équivalences de ces dernières appellations par les indiens Aymara en Bolivie.

Dans des cas plus rares, des appellations peuvent aussi faire référence à l'origine géographique (par exemple « *Salamena* » est une variété qui provient de Salamina au Pérou). D'autres appellations peuvent aussi renseigner sur le cycle de culture: « *chaucha amarilla* » ou « *chaucha blanca* » (*chaucha* est un qualificatif employé pour les variétés plus précoces). Bristol (1988) signale que les indiens Sibundoy du sud de la Colombie utilisent au moins 11 noms différents en langage *kamsa* pour différencier leurs variétés, de plus ces indiens ont leur propre appellation (*yengo*) pour l'arracacha (Hodge 1954).

D'autres dénominations contemporaines sont d'origine espagnole et portugaise. Ces appellations font référence à des ombellifères de l'ancien monde: *apio* (céleri), *zanahoria* (carotte), *pastianaca* (panais) ou *salsa* (persil). Le Brésil est le pays qui a la plus grande diversité de noms pour l'arracacha cultivée. A Rio de Janeiro on utilise *batata baroa* (la patate du baron), en référence à l'introduction de l'arracacha au Brésil par un noble qui habitait à Rio. Le nom *mandioquinha salsa*, originaire de Sao Paulo, fut adopté comme étant le nom brésilien officiel : *mandioquinha* est le diminutif portugais de manioc et *salsa* fait référence au persil (Hermann 1997).

1.2.2. Aspects biologiques

1.2.2.1. Cycle de vie

L'arracacha est une plante pérenne propagée par voie asexuée. Sans intervention humaine et en absence d'adventices compétitives, Hermann (1997) estime que des échappés de culture peuvent survivre pendant quelques années en état de croissance minimale. Cette propriété permet de repérer des échappés spontanés en floraison dans des jachères ou des champs occupés par d'autres cultures. Des pratiques culturales, comme l'élimination systématique des plantes qui fleurissent, amènent à maintenir les champs d'arracacha en état apparent purement végétatif, mais des plantes fleurissent occasionnellement. La floraison est un caractère très variable chez la forme cultivée. Cependant, il semblerait que le stress de sécheresse induit une réponse positive qui peut épuiser la plante et la conduire à périr. A partir de cette observation il a été proposé que la forme cultivée soit une plante bisannuelle (León

1964). Selon nos observations ponctuelles, les plantes que l'on observe échappées de culture ou qui survivent dans des jachères fleurissent généralement (Fig. 4, Obs. pers.). Il se peut donc que l'absence de floraison fréquemment observée au niveau des cultures corresponde à un retard dans l'induction de la floraison et que la plante soit réellement pérenne. Hermann (1997) rapporte que tous les cultivars sont aptes à la reproduction sexuée. En revanche, Knudsen (2003) rapporte ne pas avoir obtenu de floraison pour certaines accessions, même après une tentative d'induction expérimentale, ce qu'il attribue à des différences génétiques entre les clones. D'autres tentatives visant à définir les facteurs déclenchant la floraison chez l'arracacha n'ont toujours pas résolu l'énigme; néanmoins des résultats encourageants furent obtenus en combinant le stockage et la déshydratation des propagules (Knudsen 2003).

1.2.2.2. Morphologie générale

La figure 2 montre une plante cultivée dont quatre parties sont clairement distinguées: les racines tubéreuses (A), la couronne racinaire centrale (B), les tiges aériennes (dont la base est utilisée comme bouture et appelée « *colino* », équivalent au *cormel* chez les araceae tubéreuses) et les feuilles (D). Les racines tubéreuses sont utilisées pour l'alimentation, une seule plante produisant en moyenne 10 racines chacune pesant entre 100 et 300 g. Dans des bonnes conditions de culture, les racines peuvent atteindre jusque 1 kg par plante. A la différence du « *colino* », les racines ne peuvent pas être utilisées pour la propagation de la plante.

1.2.2.3. Biologie reproductive

L'arracacha est une espèce monoïque dont l'allogamie est favorisée par la protogynie (Hermann 1997, Knudsen 2003). Des descendances sexuées d'arracacha ont montré une hétérogénéité considérable ce qui suggère un taux élevé d'hétérozygotie chez la forme cultivée, cohérent aussi avec sa nature allogame (Hermann 1997). Des taux d'autogamie ont néanmoins été enregistrés. Ceci explique l'existence de graines spontanées dans un champ monoclonal au Brésil, et suggère l'auto-compatibilité de l'arracacha au moins pour certains génotypes. Knudsen (2003) signale que l'autogamie est possible grâce à la geitonogamie (pollinisation entre fleurs distinctes d'une même plante ou entre plantes du même génotype), en précisant que les taux de germination de ces graines sont moindres. Santos et al. (1990) par exemple rapportent seulement 10% de germination pour des graines récoltées en champ monoclonal au Brésil. L'arracacha est donc une plante à allogamie facultative dont les taux d'autogamie ne dépasseraient pas 4% (Knudsen, com. pers.).

L'inflorescence chez l'arracacha est une ombelle composée de 8 à 14 rayons portant chacun de 10 à 25 ombellules en fleur. Ces ombellules peuvent porter entre 3 et 5 fleurs hermaphrodites et des fleurs mâles. En général, le nombre de fleurs hermaphrodites diminue vers le centre, de manière que l'ombellule centrale est presque entièrement constituée de fleurs mâles. Le temps de floraison d'une ombelle est d'environ 15 jours (Hermann 1997).

La graine est rarement formée en champs de culture. Hodge (1954) signale que même si la graine est disponible, elle n'est pas utilisée parce qu'elle allongerait le temps de culture. Cependant, Vasquez et al. (2003) signale que des communautés indiennes en Colombie laissent fleurir l'arracacha et que l'utilisation sporadique de plantules issues de graines expliquerait la variabilité des cultivars. Au Pérou et en Equateur par contre, les paysans ne sont pas familiarisés avec l'utilisation des graines (Blas 2005, obs. pers).

1.2.2.4. Niveau de ploïdie

A. xanthorrhiza a été récemment proposée comme étant une espèce tétraploïde ($2n = 4x = 44$) sur la base des comptages de chromosomes et caryotypage (Salazar 1997). Ce même nombre de chromosomes est observé chez les types sauvages ainsi que chez d'autres espèces d'Arracacia (Blas 2005). Constance et al. (1976a, 1976b) rapporte aussi 44 chromosomes pour 15 espèces mexicaines d'Arracacia et le même nombre pour un genre proche (*Tauschia*) mais indique un nombre de base de 22. Remarquons que la reproduction uniquement sexuée de toutes les espèces du genre indique leur fertilité sexuée. Les méioses sont donc présumées régulières même si aucune étude de ségrégation allélique n'a été rapportée. Sur cette base, l'arracacha ayant le même nombre de chromosomes que le genre entier et au-delà peut donc être considérée comme une espèce diploïde issue d'un genre ne comportant que des espèces diploïdes connues.

1.2.3. Agronomie et utilisation

La forme domestiquée est propagée par bouturage de colinos. En fonction de l'état de la plante, entre 15 à 40 colinos peuvent en être obtenus. Chaque colino présente environ une douzaine de bourgeons ayant la capacité de générer une nouvelle pousse. La préparation du colino est importante pour l'obtention de bons rendements. A la différence des formes sauvages, aucune vraie dormance n'est observée chez la forme cultivée.

Les meilleures productions sont obtenues entre 1700 et 2500 m d'altitude, avec des températures fluctuant entre 15 et 20°C (Hodge 1954). En général, les plantations se font au début de la saison de pluies. Dans les Andes, la récolte a lieu au bout de 10 à 11 mois. La

culture de l'arracacha est traitée comme celle de la pomme de terre avec laquelle elle est souvent alternée (Hodge 1954). La densité des plantes varie entre 15000 et 30000 plantes/ha (Hermann 1997).

Les rendements sont variables selon les conditions de culture et les génotypes. Ils restent précaires dans l'aire andine, oscillant entre 5 à 10 t/ha, en opposition avec des rendements de 22 t/ha rapportés au Brésil avec des variétés issues de programmes de sélection. Mazon (1993) rapporte des rendements entre 9t et 27 t/ha dans une collection de 72 accessions équatoriennes et Hermann (1997) rapporte des rendements entre 35 et 52 t/ha dans une zone de production à San Jose de Minas, Equateur. Cet auteur affirme qu'en général, les rendements chez les variétés traditionnelles sont inférieurs à 20t/ha.

Concernant les maladies, l'arracacha est en général et dans les conditions traditionnelles une plante robuste avec peu de maladies graves. Un bon état phytosanitaire peut être maintenu en employant correctement des cycles de rotation.

La racine a un contenu de matière sèche variant entre 17 et 34% du poids frais. Celui-ci correspond pour 95% à des carbohydrates (95% c'est l'amidon et 5% de sucre, notamment sucrose, Hermann 1997). L'arracacha est aussi une bonne source d'acide ascorbique, vitamine A et minéraux, spécialement calcium. Par contre, elle est une faible source de protéines avec une moyenne d'environ 1% dans la matière fraîche et 4% dans la matière sèche. L'amidon est de digestion facile et moins flatulent que l'amidon de pomme de terre (Hodge 1954). Dans les Andes, où on constate une absence de produits transformés, les racines sont consommées en soupes, purée ou frites (Figure 2) alors que le Brésil est pionnier en produisant une large gamme de produits transformés.

Dans les systèmes traditionnels des Andes, l'arracacha est souvent associée avec le maïs ou d'autres plantes en fonction de l'altitude. On la retrouve par exemple avec du manioc dans les régions plus basses (1000 m) et avec d'autres tubercules dans les régions plus élevées (2000 m, obs. pers). D'autre part, l'arracacha est une plante à fort potentiel pour les régions tropicales et tempérées car c'est une plante qui reste productive sans forts investissements. Par ailleurs, il est signalé que l'arracacha peut représenter une diversification pour les petits producteurs de café arabica, avec lequel l'arracacha partage les mêmes conditions écologiques de développement (Hermann 1997).

1.2.4. Diversité et état des ressources génétiques

1.2.4.1. Variation morphologique

La racine tubéreuse est la partie la plus variable de la plante. Sur ce critère, trois formes horticoles principales sont différenciées chez la forme cultivée: jaunes, blanches et violettes (Figure 5). On trouve souvent des cultivars intermédiaires entre ces trois formes. Néanmoins la variation morphologique de l'arracacha est modeste si on la compare avec celle des autres plantes à racines et tubercules andines (Hermann 1997).

Hodge (1954) signale que les variétés jaunes sont les plus cultivées et appréciées dans certaines régions, car elles sont plus robustes et résistantes aux variations climatiques, bien que les plus tardives à la récolte. Les variétés blanches sont par contre plus appréciées dans d'autres régions, à cause de leur goût, texture et précocité. Mais elles ont souvent des rendements plus faibles. Les formes violettes, qui sont vraisemblablement des formes blanches anthocyanées, car partageant leurs autres caractéristiques, sont souvent peu appréciées par certains pour la consommation.

En ce qui concerne l'évaluation de la diversité, certaines collections ont été caractérisées à l'aide de descripteurs morpho-agronomiques. En Equateur et au Pérou, 57 et 51% de doublons ont été estimés à ce niveau (Mazon et al. 1996, Blas 2000) alors qu'au sein de collections colombiennes, une plus importante diversité morphologique a été rapportée (Rosso et al. 2002, Velez 2005). Ces résultats sont cohérents avec les observations de Bristol (1988) qui signale que les indiens Sibundoy de la Colombie cultivent un grand nombre de cultivars. Il rapporte au moins 13 variétés locales et estime que cette région est un centre de diversification de la forme cultivée (nous reviendrons sur cette observation au chapitre II). Des caractérisations agro morphologiques de collections colombiennes indiquent aussi une diversité plus élevée parmi les cultivars de ce pays (Rosso et al. 2002, Vasquez et al. 2003).

1.2.4.2. Diversité moléculaire

Il y a peu, les études de diversité chez l'arracacha étaient principalement limitées à la forme cultivée. Les premières recherches de polymorphisme ont été réalisées à l'aide des isoenzymes. Mazon (1993) sur 20 isoenzymes testés, ne repère que trois systèmes polymorphes dont seulement un parmi les formes cultivées. Erazo et al. (1996) trouvent très peu ou aucun polymorphisme en analysant un ensemble de variétés originaires de Bolivie, Brésil, Chili, Equateur et Pérou. Par contre, au niveau de l'ADN, Castillo (1997), Blas (2000, 2005) rapportent de bons pourcentages d'amorces RAPD polymorphiques (48, 85% et 50%

respectivement). Plus récemment Blas (2005) a employé les AFLPs pour déterminer le niveau de doublons au sein de la collection internationale du CIP et éclaircir la structure des formes sauvages au Pérou.

1.2.4.3. Collections de ressources génétiques

On relève actuellement des indices de perte de variabilité chez l'arracacha parce que: i) la plupart de la production destinée à son commerce se base sur l'emploi d'un nombre limité de génotypes qui ont tendance à remplacer les variétés traditionnelles et ii) la culture de l'arracacha a énormément diminué dans son aire d'origine. Pour ces raisons, plusieurs collections en champs ont été mises en place. Au Pérou et en Equateur, des collections avec 500 et 100 accessions respectivement sont conservées au champ (Hermann 1997). Cependant, pour l'arracacha, comme pour d'autres espèces à propagation asexuée, la conservation au champ est coûteuse, et les problèmes phytosanitaires entraînent des pertes. Par ailleurs, les risques de « mutations d'étiquette » sont réels. En Equateur par exemple, la collection originale de l'INIAP comptait 109 accessions en 1995 (Mazon et al. 1996) mais, selon un dernier inventaire, il n'en restait que 60 (obs. pers.). La conservation sous forme de graines a été envisagée mais il y a deux difficultés à prendre en compte: a) celui du maintien des génotypes chez une plante allogame et, b) certains cultivars ont quasiment perdu la capacité à fleurir, tout du moins dans les conditions où elles sont conservées (Knudsen 1999). La conservation *in vitro* a été essayée mais sans beaucoup de succès pour l'instant (Landázuri, 1996; obs. pers.).

1.2.4.4. Amélioration génétique

Bien que l'arracacha soit considérée comme une espèce rustique en culture traditionnelle, sa production fait face à quelques problèmes qui pourraient être résolus. La tardivité (10 à 14 mois pour certains cultivars) et la détérioration rapide des racines utiles sont les deux problèmes majeurs qui affectent les producteurs d'arracacha. Par ailleurs, il existe un certain nombre de ravageurs et de maladies qui affectent l'arracacha bien qu'à ce jour aucun parasite ne soit considéré comme un facteur limitant principal. Pour l'amélioration, Hermann (1997) propose d'exploiter la variabilité génétique de toute l'espèce (y compris des formes sauvages) et d'innover au niveau de la phytotechnie de la plante. Par ailleurs, il propose de travailler davantage sur la transformation du produit récolté. Le succès du progrès génétique se base sur la possibilité de créer de nouvelles hybridations entre génotypes complémentaires. Ceci n'est pas évident chez l'arracacha du fait que la floraison est un caractère mal maîtrisé et que

certain cultivars auraient perdu la capacité de fleurir (Knudsen 1999). En plus, les fleurs sont petites, le pollen est produit en faible quantité et peu de graines sont obtenues par pollinisation manuelle (Randers, pers.com.). Par ailleurs, la forme cultivée semblerait avoir une faible base génétique. D'où l'importance de collecter et d'étudier les ressources génétiques sauvages de cette plante qui pourraient être utilisées en croisements. A ce sujet, des travaux pionniers et encourageants ont été réalisés et des hybrides entre cultivars et des formes sauvages proches ont été ainsi obtenus (Vasconez non publié, Knudsen 2003).

Selon Hermann (1997), le seul programme d'amélioration variétale existant chez l'arracacha a débuté en 1987 au Brésil et avait pour but le développement d'un matériel adapté à différents environnements. Les clones obtenus sont tous issus d'autofécondation d'un seul clone brésilien qui avait révélé une diversité importante lorsqu'il est reproduit par graine (Santos, com. pers. à Hermann 1997). L'EMBRAPA avait sélectionné des clones prometteurs sur la base de certains caractères d'intérêt (vigueur de la plante, pigmentation jaune des racines et forme cylindrique pour faciliter l'emballage) et de nouvelles variétés ont été ainsi obtenues.

1.3. TAXONOMIE

1.3.1. La famille des Apiaceae

La famille des Apiaceae ou Ombellifères (Umbelliferae) comprend environ 450 genres et 3700 espèces (Pimenov et Leonov 1993). Les espèces ne sont pas également distribuées entre les genres: 41% des genres sont monotypiques et 26% comprennent seulement deux ou trois espèces. Un tiers des espèces sont réparties dans le Nouveau Monde (Mathias 1971, cité par Hermann 1997). La famille inclut des formes avec des cycles de vie annuels, bisannuels et pérennes, généralement caractérisées par leur inflorescence, connue comme ombelle. Les fruits contiennent souvent des huiles essentielles qui donnent un arôme particulier à de nombreuses espèces et permettent leur utilisation comme épices. Il existe une diversité d'utilisation considérable au sein de la famille qui compte de nombreuses espèces économiquement intéressantes pour l'alimentation de base ou les épices. Une liste détaillée considère quelques 80 espèces utilisées par l'homme, notamment comme légumes ou condiments (Rubatzky et al. 1999); parmi ce groupe, 37 espèces sont cultivées pour leurs racines tubéreuses. La carotte, le panais et l'arracacha sont considérés d'une importance majeure dans certaines régions. Plusieurs autres sont produites à une échelle moins importante et d'autres encore sont utilisées seulement en cueillette.

La domestication des ombellifères est intervenue principalement en Eurasie; seule l'arracacha et le *culantro* (*Eryngium foetidum*) utilisé comme condiment, sont originaires de l'Amérique. Les ombellifères domestiquées dans l'Ancien Monde sont en grande majorité biannuelles et sont propagées uniquement par voie sexuée, à la différence de l'arracacha qui fait donc exception.

1.3.2. Le genre *Arracacia*

La sous famille Apiodea dont le genre *Arracacia* fait partie comprend 404 genres avec plus de 2800 espèces (Pimenov et Leonov 1993). Le genre *Arracacia* Banc. (1825) est endémique du nouveau monde. Ce genre comprend une trentaine d'espèces pérennes ou annuelles distribuées depuis le Mexique et l'Amérique Centrale jusqu'au Pérou et la Bolivie, mais l'estimation du nombre d'espèces et les délimitations du genre par rapport aux genres voisins sont loin d'être définitifs. Constance (1949), lors de ses premiers travaux, avait identifié 24 espèces pour le genre que cet auteur signalait comme faisant partie des plus diversifiés du continent; Hiroe (1979), Pimenov et Leonov (1993 cités par Blas 2000) rapportent respectivement 28 et 55 espèces. Enfin, Downie et al. (2000) sur la base d'une phylogénie moléculaire de l'ADN chloroplastique, confirment que le genre *Arracacia* fait partie de ceux les plus diversifiés de la sous-famille Apiodea mais, sur cette base cytoplasmique, signalent qu'il n'est pas monophylétique.

Selon Constance et Affolter (1995), le genre occupe une position centrale par rapport à un certain nombre d'autres genres neotropicaux, tous ayant été confondus à certaines périodes avec le genre *Arracacia*. Les espèces Nord-Américaines du genre ont été les plus étudiées. Ce sont principalement les études de Constance entre 1949-1995 qui ont servi de référence. Cet auteur signale des difficultés pour définir les espèces du genre: herbarium inapproprié, manque d'études cytologiques et absence de fruits mûrs qui sont pourtant essentiels pour la taxonomie de la famille. Il regrette aussi que, sauf quelques exceptions, les collecteurs n'aient pas inclus de racines dans les échantillons d'herbier (Constance et Affolter 1995, Hermann 1997).

Les espèces sud-américaines d'Arracacia

Le plus grand nombre d'espèces du genre *Arracacia* se localise au Mexique, seule une dizaine étant localisée en Amérique du sud avec la plupart présentes en Colombie (Knudsen 2003, Fig. 6). Hermann (1997) fait une première compilation des collections d'herbier disponibles et

rapporte neuf espèces pour le genre *Arracacia* en Amérique du sud. Knudsen (2003) sur la base d'une révision plus approfondie, retient 10 espèces sud-américaines distribuées dans les vallées andines du Venezuela, Colombie, Equateur, Pérou et Bolivie. Il signale que cinq espèces sont endémiques en Amérique du Sud (*A. colombiana*, *A. elata*, *A. incisa*, *A. moschata* et *A. tilletii*, plus les formes sauvages de *A. xanthorrhiza*) alors que les cinq autres sont aussi présentes en Amérique centrale (*A. acuminata*, *A. bracteata*, *A. tolusensis*, *A. vaginata* et la forme cultivée de *A. xanthorrhiza*).

Les espèces sud-américaines d'*Arracacia* ont des racines pérennes qui sont soit ligneuses soit tubérisées mais pivotantes (Mathias et Constance 1976). Les racines pérennes dépassent rarement 2 cm de diamètre et les tiges 3 m de hauteur. Ces espèces à racines ligneuses se trouvent souvent dans des habitats temporairement humides ou dans les forêts de montagne à plus de 3000 m d'altitude. Elles sont aussi présentes dans des aires de végétation perturbée, par exemple le long des routes comme c'est le cas chez les espèces *A. moschata* ou *A. acuminata* montrées dans la figure 7. Ces plantes se caractérisent par une croissance végétative continue: elles peuvent fleurir et fructifier à tout moment, sauf dans les périodes d'extrême sécheresse. Ce mode de vie n'est pas celui des espèces d'*Arracacia* à racines tubéreuses. Ces dernières présentent une phénologie différente correspondant à leur adaptation à des habitats plus chauds avec une aridité saisonnière marquée, situés dans des régions bien délimitées et à des altitudes plus basses (généralement en dessous des 3000 m). Elles se caractérisent par un système racinaire épais, dépassant les 4 cm à la base, une couronne racinaire et des structures comprimées de tiges qui complètent les organes de stockage permettant à la plante de survivre à des périodes de sécheresse de plus de 8 mois.

1.3.3. Les espèces tubéreuses d'*Arracacia*

L'aire de distribution des espèces tubéreuses comprend l'Equateur, le Pérou, et la Bolivie entre 1800 et 3000 m. Elles sont pérennes et, selon leur cycle de vie, sont polycarpiques (fleurissent plusieurs fois) ou monocarpiques (meurent après la première fructification). Hermann (1997) se basant sur la nomenclature proposée par Constance (1949) signale que les formes tubéreuses du genre *Arracacia* comprennent, à part la forme cultivée *A. xanthorrhiza*, les espèces polycarpiques *A. andina*, *A. equatorialis* et *A. incisa*. Knudsen (2003), estime que *A. Andina* et *A. equatorialis* sont synonymes et les classe comme une variété sauvage d'*A. xanthorrhiza* (*A. x. var. andina*). Par ailleurs, cet auteur décrit pour la première fois le type monocarpique remarqué par Hermann (1997) et le classe comme une deuxième variété de *A. xanthorrhiza* (*A. x. var. monocarpa*). Il propose *A. xanthorrhiza* comme un complexe

intégrant trois formes à racines tubéreuses: les formes cultivées (*A. x. var. xanthorrhiza*), les formes sauvages polycarpiques (*A. x. var. andina*) et les formes sauvages monocarpiques (*A. x. var. monocarpa*). Par ailleurs, il estime qu'une deuxième espèce tubéreuse, *A. incisa*, est proche de *A. xanthorrhiza* mais la conserve dans son statut d'espèce en argumentant certaines différences botaniques au niveau floral. Dans le présent travail, nous adoptons la nomenclature proposée par Knudsen (2003).

Les formes sauvages tubéreuses se rencontrent en général dans des écosystèmes relativement différents de celui de la forme cultivée mais dans certains cas, elles se retrouvent aussi comme adventices dans les champs de culture qui peuvent occasionnellement aussi inclure l'arracacha (León 1964, Valderrama et Seminario 2002, Knudsen 2003, Blas 2005). Il a été suggéré que cette association peut conduire dans des conditions favorables à des flux de gènes mais cela n'a pas été démontré.

1.3.4.1. Les formes sauvages pérennes de *A. xanthorrhiza* (*A. x. var. andina*)

En termes de morphologie, ces formes sont les plus proches de la forme cultivée mais elles restent distinguables des plantes échappées de culture (Hermann 1997, Knudsen 2003) (Fig. 8). Ce sont les formes les plus fréquentes parmi les formes tubéreuses sauvages et leur présence est souvent associée à des aires perturbées par la présence humaine. Dans le nord du Pérou par exemple, des restes archéologiques témoignent de leur présence associée à d'anciennes populations inca (Valderrama et Seminario 2002). Ces mêmes auteurs estiment que des plantes échappées de culture ont pu donner lieu à des populations qui se sont retrouvées ultérieurement à l'état sauvage. Cependant, Hermann (1997) affirme que même en l'absence de faible compétition, des échappés de culture ont peu de chances de coloniser de nouveaux habitats et qu'ils vont seulement survivre pendant trois ou quatre années dans des conditions de croissance minimum.

Une grande variation morphologique est observée au sein de cette variété, comme reflété dans les classifications taxonomiques antérieures (Knudsen 2003, Blas 2005). Ces populations polycarpiques partagent des aires de distribution similaires tout au long des vallées andines et on ne peut pas exclure des flux génétiques entre elles (Blas 2005).

Elles montrent une période de dormance qui se déclenche après la floraison lors de l'arrivée de la saison sèche. Ce comportement est conservé en conditions de serre au Pérou, même sous humidité permanente (Randers, com. pers). Certaines observations montrent que la dormance

est provoquée par des photopériodes courtes (inférieure aux 12 heures), alors qu'au dessus de 13 heures les plantes ont une croissance continue (Hermann 1997, Knudsen 2003).

Cette variété botanique est, selon Knudsen (2003), la plus proche de la forme cultivée, les deux partageant : le mode de vie pérenne, une architecture similaire en état végétatif ou reproductif (moins de 30 ombelles) et l'aptitude à la reproduction asexuée par bouturage. Par contre, les deux types diffèrent par la taille des feuilles qui, de même que les colinos, sont généralement plus grands chez la forme cultivée, celle-ci ne présentant par ailleurs pas de dormance.

1.3.4.2. Les formes monocarpiques de *A. xanthorrhiza* (*A.x.var .monocarpa*)

Ce type ne fut que récemment documenté. Il est moins fréquent que le type polycarpique. Hermann (1997) pensait qu'il s'agissait d'un type bisannuel, puis a réalisé qu'il s'agit d'un type pérenne monocarpique dont la floraison peut parvenir après quelques années de vie (com. pers, Blas 2005). La description botanique fut réalisée par Knudsen (2003). A la différence du type polycarpique, ces plantes sont plus vigoureuses et peuvent atteindre 3.5m de haut en floraison (Fig.9). La partie souterraine est formée par de grosses racines de couleur crème qui peuvent atteindre 7 cm de diamètre et 1 m de longueur. L'inflorescence peut porter jusque 250 ombelles mais le nombre de tiges génératives dépasse rarement 2 ou 3 (Knudsen 2003, obs. pers.).

Malgré ces différences évidentes entre les deux formes, certaines populations sont difficiles à classer en type polycarpique ou monocarpique (Obs pers.). Knudsen (2003) signale cependant ne pas avoir observé de populations mixtes bien qu'il ait noté des cas où les deux types sont en proximité spatiale. Il rapporte aussi des cas peu fréquents où la forme monocarpique est sympatrique de la forme cultivée et estime que la forme monocarpique pourrait aussi être ancêtre de la forme cultivée.

1.3.4.3. L'espèce *A. incisa*

C'est la seule autre espèce du genre à former des racines tubéreuses. On ne la trouve que dans les Andes centrales du sud du Pérou où elle pousse dans des habitats similaires à ceux d'*A. x. var. andina*, avec peu de végétation (Knudsen 2003, Blas 2005). Les plantes sont pérennes polycarpiques et ne dépassent pas les 75cm de hauteur (Blas 2005, Fig. 10). Ses racines sont tubéreuses de couleur blanche ou jaune. Elle est morphologiquement proche de *A. x. var. andina* avec qui elle partage le comportement de dormance et la capacité d'être

végétativement propagée. Or malgré ces similitudes, Knudsen (2003) trouve qu'elle est facilement reconnaissable grâce à des différences notamment sur des caractères floraux. De plus, *A. incisa* présente des plantes de plus petite taille et Knudsen (2003) considère qu'elle mérite d'être maintenue comme une espèce à part du complexe *A. xanthorrhiza*. Une corroboration de ce traitement se retrouve dans une étude morphologique et moléculaire de Blas (2005).

1.3.4.4. Utilisation des formes sauvages

Les formes tubéreuses sauvages d'Arracacia sont utilisées par l'homme notamment dans la médecine traditionnelle comme souligné par plusieurs auteurs, sans que des études approfondies aient été faites. Valderrama et Seminario (2002) donnent des informations sur l'utilisation et les appellations employées dans le nord péruvien. Ils indiquent que ces formes sauvages sont souvent présentes dans les aires de culture de l'arracacha et rapportent 34 appellations différentes dont 20 dans le nord du Pérou. Ces appellations font soit référence au milieu où ces populations se trouvent (« de monte », « de cerro », « de campo », « de jalca », « orco-racacha », « sacha zanahoria »), soit à la relation avec les populations humaines (del antiguo, del abuelo, del gentil) ou encore à l'utilisation dans certaines communautés (comme « cochezanahoria » destiné à l'alimentation des porcs). D'autres appellations comme « de zorro » « de huishco » signifient dans la cosmovision de l'homme andin que ce sont des plantes sauvages produisant spontanément. Ces appellations dénotent la connaissance vis-à-vis de ces plantes et les différents rapports que l'homme andin a avec ces plantes. Il faut ajouter que ces appellations sont aussi appliquées à des espèces non tubéreuses d'Arracacia comme *A. acuminata*, *A. moschata* ou *A. elata* (Bristol 1988, Hermann 1997, Knudsen 2003, obs. pers).

Par ailleurs, plus d'une vingtaine d'applications dans la médecine traditionnelle ont été signalées pour ces plantes (Valderrama et Seminario 2002), notamment comme purgatifs, traitement du rhume, des troubles de la ménopause et d'autres affections. Une autre utilisation est l'alimentation animale; les racines sont bien consommées par les porcs et les feuilles aussi éventuellement par d'autres animaux. Finalement, ces auteurs rapportent un cas où, en l'absence de nourriture, des racines sauvages ont été utilisées pour l'alimentation humaine.

1.4. ETAT DES CONNAISSANCES SUR L'ORIGINE DE LA DOMESTICATION

La découverte au Pérou et Equateur des formes sauvages de *A. xanthorrhiza* étant relativement récente (Hermann 1997), les connaissances sur la domestication de *A.*

xanthorrhiza sont à ce jour très limitées. Nous proposons néanmoins une synthèse de l'état du sujet avant notre travail.

A quelle époque et dans quelle région la domestication de A. xanthorrhiza se situe-t-elle?

La domestication de *A. xanthorrhiza* a été suggérée comme ancienne sur une base notamment archéologique. Des restes archéologiques qui semblent appartenir à l'arracacha ont été trouvés dans des tombeaux précolombiens (León 1964); également, des céramiques ont été trouvées dans des tombeaux Incas à Nazca (Pérou) datant d'au moins 2000 ans A.J. qui présentent des dessins pouvant bien correspondre à l'arracacha (Fig. 2, Hodge 1954). Selon ce dernier, le fait que la forme cultivée produise rarement des graines viables serait également un indice suggérant une ancienne domestication par l'homme.

Concernant le site probable de domestication, les indices suivants pointent la zone Andine d'Amérique du sud:

- a) l'existence des formes sauvages tubéreuses est restreinte à l'Equateur, au Pérou et dans une moindre mesure à la Bolivie. Selon Hermann (1997) ces formes sont véritablement sauvages et non des plantes férales qui auraient développé un système racinaire moins volumineux en milieu naturel.
- b) Des données linguistiques soutiennent l'origine andine de la plante; racacha, virraca, lacache, arrecate et plusieurs autres noms en langues autochtones attestent d'une origine ancienne dans les Andes. Les évidences linguistiques amènent Hodge (1954) à suggérer que l'arracacha cultivée serait le résultat d'un seul événement de domestication et que la dispersion dans d'autres régions aurait pu se passer ultérieurement dans les temps précolombiens.

Le site précis de la domestication est aventureux à établir, cependant certains auteurs font des propositions. Hodge (1954) signale que la domestication de l'arracacha doit se situer au sein des anciennes cultures pre-Inca qui développèrent un degré élevé de spécialisation dans l'agriculture, et propose les Andes du sud de la Colombie sur la base d'une grande diversité de cultivars et des noms autochtones. Hermann (1997) penche pour la région centrale de l'Equateur suite à la découverte de populations sauvages pérennes d'une morphologie très proche de la forme cultivée. Blas (2005) de son côté propose la région sud du Pérou, ancien

centre de cultures pre-Inca, sur la base de la présence de formes sauvages et de restes archéologiques.

Quelle forme sauvage est-elle ancêtre de la forme cultivée?

L'identification de la forme sauvage qui a donné lieu à l'arracacha cultivée est controversée. Les espèces *A. andina* et *A. equatorialis*, ont été en premier suggérées comme possibles ancêtres de la forme cultivée (Constance 1949). Ultérieurement, ce même auteur suggère la synonymie de *A. andina* et *A. xanthorrhiza* et signale que *A. equatorialis* n'est probablement pas à l'origine de la forme cultivée (In Arbizu et al. 1997). Hermann est le premier à s'intéresser de près à la domestication de l'arracacha depuis la découverte des populations sauvages pérennes dans la région centrale de l'Equateur (com. pers.). Dans sa monographie, il décrit ces populations (retrouvées à Huigra et Sibambe, province de Chimborazo et Bolivar dans le centre de l'Equateur) qu'il considère comme les plus probables formes ancestrales de la forme cultivée, et écarte les populations du nord du Pérou à cause de leurs racines moins volumineuses et de leur forte astringence. Knudsen (2003) à son tour conclut que les deux types sauvages de *A. xanthorrhiza* (polycarpique et monocarpique) sont tous deux proches du type cultivé et représentent de possibles formes ancestrales, en soulignant cependant que le type polycarpique est plus affine au niveau botanique alors que le type monocarpique n'a probablement pas participé à la domestication à cause de son inaptitude à la reproduction asexuée (com. pers). Blas (2005) étudie des populations pérennes et monocarpiques du Pérou à l'aide de marqueurs AFLPs et conclut qu'à cause des différents processus d'hybridation entre les formes sauvages au cours de la domestication, il n'est pas possible de préciser laquelle des deux formes sauvages est à l'origine de la forme cultivée (Fig. 11).

Il est évident qu'en l'état des connaissances actuelles, la domestication de l'arracacha reste un sujet à élucider. Notre étude vise à proposer un scénario de la domestication et à déterminer s'il comprend une forme ancestrale majeure ou si un scénario alternatif doit être envisagé. Chez le manioc par exemple (dont la domestication est aussi abordée dans le chapitre II de cette thèse), une forme ancestrale majeure a été identifiée mais l'hypothèse d'une participation probable d'espèces éloignées par introgression génétique justifie une documentation plus précise.

2. OBJECTIFS ET DEMARCHE METHODOLOGIQUE

Comme indiqué dans la partie introductive, l'arracacha est une plante entièrement domestiquée. En effet, bien qu'une démarcation claire des caractéristiques morpho anatomiques entre formes sauvages et domestiquées soit difficile à cerner, la forme cultivée ne semble pas pouvoir survivre sans l'intervention de l'homme (Hodge 1954, Hermann 1997) et il apparaît un syndrome net de domestication: taille des propagules de reproduction végétative, racines tubéreuses, mode de vie, etc. A ce propos, bien que l'intérêt scientifique pour cette plante ait augmenté durant ces dernières années, l'état des connaissances sur l'histoire de la domestication et de la diversité des ressources génétiques de l'arracacha est presque inexistant. L'existence de formes sauvages en Equateur offre la possibilité d'aborder le sujet de l'origine de l'arracacha et d'étendre nos résultats d'une façon plus large à la question de la domestication des plantes dans cette région.

Les objectifs de cette thèse sont de répondre à des questions d'ordre général lorsqu'on aborde le sujet de domestication d'une plante cultivée:

- Quel impact la domestication a-t-elle eu sur la diversité de la forme cultivée?
- Existe-il encore actuellement une dynamique de la diversité dans l'aire de domestication; flux géniques entre cultivars, ou entre cultivars et ancêtres sauvages directs, ou entre cultivars et d'autres espèces sauvages?
- Les formes considérées comme sauvages sont-elles de vraies espèces sauvages ou plutôt des formes dérivées de la forme domestiquée ?
- Quel est ou quels sont le ou les ancêtre(s) sauvage(s) ?
- Y a-t-il eu un seul ou plusieurs événements de domestication? Un centre ou un non-centre de domestication (Harlan 1971)?

Afin de répondre à ces questions, nous avons procédé par étapes. Une première étape fut de constituer une collection de travail pour les analyses génétiques. Nous avons effectué un travail de prospection des formes sauvages et cultivées en Equateur afin de compléter une collection existante des formes cultivées en Equateur. A cette occasion, nous avons aussi cherché à identifier des pratiques traditionnelles de domestication qui pourraient être entretenues ou des pratiques paysannes spécifiques à la culture de l'arracacha. Cette collection fut complétée avant d'entreprendre des analyses génétiques avec des outils moléculaires. Le

choix des marqueurs a été fait en fonction des objectifs de l'étude mais il s'agit de marqueurs devenus habituels dans les études de diversité des plantes cultivées. Dans un premier temps, nous avons développé des marqueurs microsatellites (SSR) afin de disposer d'un marqueur très polymorphe et de surcroît codominant, car un des objectifs de la thèse est de déterminer si des hybridations spontanées et/ou des introgressions ont lieu dans l'aire de domestication. De plus, on s'attendait à peu de diversité compte tenu des analyses précédentes (connues à l'époque) qui suggéraient une faible base génétique de la forme cultivée. Ce sont donc les raisons pour lesquelles nous avons opté a priori pour les SSR comme marqueurs de choix. Par la suite, un ensemble d'échantillons représentatifs de la diversité observée avec les SSR a été caractérisé sur un nombre plus important de locus à l'aide des marqueurs AFLPS. A la différence des SSR dont le polymorphisme analysé se base sur un petit nombre de locus, mais un nombre relativement important d'allèles, les AFLPs sont des marqueurs multi-locus qui fournissent une couverture plus large du génome en permettant une évaluation plus globale de la diversité des génotypes étudiés. Enfin, un génotypage de l'ADN chloroplastique sur la base de deux séquences intergéniques a permis d'établir une première phylogénie d'haplotypes représentant les lignées femelles (si l'on admet que *A. xanthorrhiza* hérite son ADN chloroplastique maternellement comme dans le cas général des angiospermes).

Avec l'information procurée, nous tenterons de donner quelques réponses aux questions posées. Nous espérons approfondir les connaissances sur la diversité génétique de l'arracacha afin d'en retirer des informations directes pour l'exploitation et la conservation des ressources génétiques; également apporter des éléments de réponse quant à l'origine de la forme cultivée et contribuer aux questions plus générales sur la domestication et la diversification des racines et tubercules andins qui restent à ces jours peu étudiés, à l'exception de la pomme de terre.

3. MATERIEL ET METHODES

3.1. Prospections et collectes

Des missions de prospection dans les Andes équatoriennes ont été organisées sur la base des références d'herbier et des collectes précédentes. Nous avons notamment visé la collecte des populations des formes sauvages tubéreuses bien que d'autres formes sauvages moins apparentées, sans racines tubéreuses, et quelques cultivars aient aussi été échantillonnés. En fonction de la population trouvée, 5 à 10 plantes furent échantillonnées végétativement. Notre

échantillonnage est donc constitué de populations regroupées par région, selon les quatre tournées ayant eu lieu en 2002 et 2003 et dont le détail est le suivant:

1ère mission: 02-14/Août/2002. Destination: Sud de l'Equateur (Provinces de Loja, Azuay, Cañar, et Zamora). Nous avons récolté 27 individus sur 8 populations de *A. x.var. andina* et échantillonné quatre accessions des espèces non tubéreuses.

2ème mission: 18-22/Novembre/2002. Destination: Nord de l'Equateur (Provinces de Carchi, Imbabura et Pichincha). Aucune forme sauvage tubéreuse n'a été observée dans cette région et seules des espèces non tubéreuses (6) ont été échantillonnées.

3ème mission: 9-12/Décembre/2002. Destination: Centre de l'Equateur (Provinces de Cotopaxi, Tungurahua et Chimborazo). Nous avons échantillonné 2 populations d'*A. x. var. andina* et deux accessions des espèces non tubéreuses.

4ème mission: 08-18/Décembre/2003. Destination: Sud de l'Equateur (province de Loja). Nous avons récolté 49 individus sauvages sur 16 populations d'*A. x. var. andina* et *A. x. var. monocarpa* et échantillonné quatre accessions des espèces non tubéreuses.

Une liste détaillée des données de passeport de ces collections est présentée dans l'annexe 1.

3.1.1. Observations de terrain

Bien que notre travail de terrain n'ait pas été exhaustif, nous avons effectué quelques observations intéressantes lors de nos prospections des populations sauvages et visites à San José de Minas, petite localité qui produit intensivement l'arracacha au nord de l'Equateur. En voici quelques unes:

Concernant les formes sauvages

En Equateur, les formes sauvages tubéreuses sont présentes dans les régions centrales et sud des Andes. Elles deviennent plus fréquentes au sud, notamment dans la province de Loja, limite avec le Pérou. La taille des populations est variable. Celle de Huigra (centre de l'Equateur, Province du Chimborazo) par exemple s'étend sur plusieurs kilomètres de long en surplomb d'une route entre 1450 et 2500 m d'altitude (Hermann 1997). A Loja il s'agit souvent de petites populations formées d'une dizaine d'individus séparées de quelques kilomètres. Assez fréquemment, une enquête est nécessaire pour les découvrir. Elles sont souvent bien isolées entre elles et préservées de l'activité humaine (si ce n'est le pâturage) car sur les crêtes ventées des piedmonts andins.

Il est intéressant pour les prospecteurs que les habitants locaux sachent où poussent ces plantes. En effet, bien que des pratiques traditionnelles de domestication n'aient pas été observées, ils cueillent les formes sauvages souvent pour des applications médicinales. En aucun cas nous n'avons rencontré des paysans déclarant s'être intéressés à mettre en culture des plantes sauvages. Même dans les cas où les organes souterrains sont vendus au marché, les gens préfèrent aller les chercher eux-mêmes, devant dans certains cas employer quelques heures de marche pour les cueillir. Les gens se justifient en expliquant qu'il n'est pas nécessaire de les planter car elles (les formes sauvages) « poussent toutes seules » et de toute façon ne tiendraient pas en culture.

Concernant l'interaction possible entre formes sauvages et cultivées

Il arrive, rarement, qu'une plante sauvage ou un ensemble de plantes sauvages repoussent au milieu d'un champ cultivé. Le cas observé chez un paysan était intéressant. Sans le savoir disait-il, il avait dû défricher pour y cultiver l'arracacha un lot comprenant des plantes sauvages en dormance. En effet, comme observé après l'arrivée des pluies, des formes sauvages et cultivées ont poussé dans le même champ. Comme le lot était en mélange il a préféré abandonner la culture. Dans d'autres cas, des lots cultivés avec l'arracacha peuvent se trouver adjacents à des populations sauvages (Fig. 8). Cependant les paysans distinguent bien la forme sauvage « qui ne se mange pas » de la forme cultivée, que l'on appelle « *del noble* » (ennoblie).

Concernant la forme cultivée:

Une pratique traditionnelle de la culture d'arracacha est d'arracher systématiquement les plantes qui fleurissent. Il arrive cependant, que les plantes d'un même lot fleurissent massivement comme c'était le cas lors de notre visite à San José de Minas en décembre 2003. D'après les paysans, cette situation est plutôt rare et dépend des conditions environnementales. Si une période de sécheresse est suivie d'une saison des pluies courte, les plantes se mettent à fleurir. Dans de telles conditions, le lot n'est plus intéressant pour le producteur qui peut abandonner la culture en laissant les plantes former des graines en relative abondance. Malgré cela, les producteurs ne sont pas familiarisés avec la graine et pour certains c'était même la première fois qu'ils l'observaient. Par ailleurs ils affirment n'avoir jamais vu de plantules issues de ces graines. Nous avons repéré seulement le cas d'un paysan

qui s'était intéressé à récupérer les graines et à les faire pousser « pour voir ce que ça donnerait ».

Par ailleurs, on a souvent pu constater des plantes échappées de culture dans des jachères en état de floraison et porteuses de graines (Fig.4). L'origine de ces plantes spontanées est certainement à partir des plantes ou propagules qui restent après les récoltes lorsque les champs sont laissés, bien qu'on ne puisse écarter la possibilité que des graines dans de bonnes conditions arrivent à y germer.

3.2. Matériel végétal

Notre collection biologique fut assemblée à partir des trois sources suivantes:

- 1) Une représentation de la diversité morphologique de la collection d'arracacha de l'INIAP (43 échantillons),
- 2) les prospections INIAP-IRD (44);
- 3) la collection de S. R. Knudsen (ex-KVL University, Copenhague, Danemark) (60, sous forme de feuilles séchées)

Ainsi notre collection de travail représente:

- La forme cultivée avec 77 accessions (61 équatoriennes et 16 péruviennes) collectées dans 11 provinces de l'Equateur (y compris Zamora dans l'Amazonie) et à Cajamarca, département nord du Pérou limitrophe avec l'Equateur. La plupart des accessions (26) proviennent des trois provinces du sud (Loja Azuay, et Zamora), suivies de 16 accessions des provinces centrales (Tungurahua, Bolivar, Cotopaxi, Chimborazo et Cañar) et 14 accessions du nord (Pichincha, Imbabura et Carchi). Pour deux accessions l'origine est inconnue et 14 accessions sont du nord du Pérou. Une accession représente un clone collecté dans un champ cultivé. Nous avons échantillonné une plante par accession choisie au hasard.
- Des formes sauvages tubéreuses: 48 populations représentées par 1 à 10 individus; 19 ont été collectées en Equateur et 29 en Pérou (en tout, 115 plantes). Parmi ces collections, 35 populations ont été préliminairement identifiées comme *A. x. var. andina* (P), 12 comme *A. x. var. monocarpa* (M) et une correspondent à l'espèce *A.*

incisa, avec 3 plantes. Ces populations ont été récoltées dans les provinces du sud de l'Equateur (14), la région centrale (7) et le nord du Pérou (départements de Cajamarca et Amazonas (27).

- Des hybrides artificiels (13) obtenus de deux croisements entre le type cultivé et *A. x. var. andina*. Les parents utilisés dans les croisements sont aussi représentés.
- Des formes sauvages non tubéreuses; 9 collections appartenant à trois des espèces présentes en Equateur (*A. elata*, *A. moschata*, *A. acuminata*). Ces espèces ont été collectées tant au nord qu'au sud du pays.

Une liste détaillée des données de passeport est présentée dans l'annexe 1. Le numéro de référence indiqué sur cette liste est utilisé dans la présentation des résultats: les formes cultivées sont numérotées C1, C2, etc. et les formes sauvages sont numérotées S1, S2, etc. (pour les individus d'une même population une numérotation additionnelle est signalée, Ex. S1-1, S1-2, etc.). Treize hybrides artificiels sont numérotés H1 à H13. Une carte montrant les sites de collection est présentée dans la figure 12.

3.3. Analyses génétiques

Pour l'étude de diversité génétique, des marqueurs SSR et AFLPs ont été utilisés en génotypage. Par ailleurs la variation moléculaire de l'ADN chloroplastique (ADNcp) a été abordée en séquençant deux fragments de séquences intergéniques à l'aide d'amorces universelles. Le travail de laboratoire s'est entièrement déroulé au laboratoire Genetrop de l'IRD à Montpellier, à l'exception de ce qui concerne la création de la banque microsatellite qui s'est déroulée au laboratoire BIOTROP du CIRAD.

Extraction d'ADN génomique

L'extraction d'ADN fut faite à partir de feuilles séchées dans du silica gel et broyées à l'aide d'un broyeur mécanique. Des tests d'extraction d'ADN en essayant différents protocoles et agents (MTAB, metalxanthates -PEX- et sorbitol) ont été conduits. Une modification de la méthode décrite par Jhingan (1992) qui se base sur l'utilisation de PEX a été finalement sélectionnée en fonction du rendement et de la qualité d'ADN obtenue.

3.3.1. Création d'une banque microsatellite et identification des marqueurs

L'élaboration d'une banque SSR s'est basée sur la méthode publiée par Billotte et al. (1999) (Fig. 13). L'ADN d'un cultivar (code d'accèsion MST-047) a été utilisé pour le développement de la banque. Cette démarche comprend plusieurs étapes qui sont:

1. Digestion de l'ADN génomique
2. Ligation des adaptateurs
3. Pré-amplification
4. Sélection des fragments contenant les motifs SSR
5. Amplification des fragments sélectionnés
6. Ligation et clonage
7. Transformation
8. PCR des inserts clonés
9. Contrôle sur gel et transfert sur membrane
10. Identification des clones contenant les motifs SSR par hybridation

Les clones positifs du test d'hybridation sont sélectionnés pour séquençage. Pour les séquences estimées exploitables (motifs peu complexes et longueur suffisante de séquences flanquantes pour le dessin d'amorces), des amorces ciblant le motif SSR sont ensuite dessinées. Nous avons pour cela utilisé le logiciel Primer Select sur DNASTAR (2003). Un test d'amplification afin de vérifier la taille du produit PCR a été réalisé avec un gradient de température et des conditions standard du mélange d'amplification. Dans certains cas, une mise au point additionnelle a été nécessaire afin d'améliorer la qualité du produit d'amplification en testant des gradients de concentration d'ADN, d'amorces et de $MgCl_2$ dans la réaction PCR. Avec les amorces amplifiant un produit PCR dont la taille correspondait à l'attendu, un criblage de polymorphisme fut ensuite réalisé à l'aide de marquage radioactif au P33. L'ADN de dix accessions (sauvages et cultivés) a été utilisé pour la sélection de locus polymorphes.

3.3.2. Génotypage par microsatellites (SSR)

Matériel végétal: un ensemble de 178 échantillons de plantes individuelles comprenant les formes tubéreuses sauvages et cultivés a été utilisé pour l'étude de diversité SSR. La forme cultivée (FC) fut ainsi représentée par 72 individus représentant 72 cultivars, la forme sauvage polycarpique (P) par 76 individus (35 populations) et la forme sauvage monocarpique (M) par

30 individus (9 populations). L'espèce *A. incisa* fut représentée par trois individus d'une seule population. Treize hybrides F1 FC x SP ont aussi été analysés.

SSRs étudiés: 14 SSR polymorphiques ont été utilisées. Les amplifications ont été réalisées sur 5-10 ng d'ADN en utilisant la méthode du « *M13 tailing* ». Cette méthode consiste à appliquer une révélation fluorescente grâce à l'ajout en extrémité 5' d'une queue M13 fluorescente (5'CACGACGTTGTAAAACGAC3') à l'une des amorces (dans notre cas l'amorce *Forward*). Le segment M13 fluorescent est ajouté au mélange PCR durant l'amplification. La migration et l'acquisition de l'image du gel se fait grâce au *LI-COR IR² automated DNA sequencer* (LI-COR, Lincoln, NE, USA, Biosciences). Le principe consiste à détecter les produits PCR fluorescents par deux lasers qui transmettent l'image des bandes au serveur. Le dépôt des produits PCR se fait sur un gel dénaturant d'acrylamide 6.5% (KB^{plus} 6.5%). L'image brute est exportée sur un logiciel d'analyse (SAGATM) afin d'obtenir l'assignation de taille des allèles sur la base d'une comparaison avec un marqueur de taille. Un exemple d'amplification est montré sur la figure 14.

Une phase préliminaire de mise au point fut nécessaire pour définir les SSR combinables en duplex pour le génotypage. Les combinaisons suivantes de SSR utilisées sont: AxD82-AxC87, AxC27-AxD43, AxC64-AxD34, AxC38-AxD13, AxC85-AxD72. Les SSR suivants ont été utilisés individuellement: AxD55 et AxD85.

3.3.3 Génotypage par AFLPs

L'ADN génomique fut digéré avec les enzymes EcoRI et TruI puis les réactions de ligation et préamplification (PCR+1) ont suivi le protocole de Vos et al. (1995). Un criblage avec 16 ADN (arracacha cultivée et formes sauvages tubéreuses) fut réalisé afin d'identifier les SSR les plus polymorphes sur 24 testés. Un exemple d'amplification est montré sur la figure 15. Les amorces E furent marquées par fluorescence en 700 ou 800 nm pour être détectées par le séquenceur LI-COR. Les produits d'amplification sélective (PCR+3) furent déposés sur un gel d'acrylamide à 6.5% pour une migration de 3 heures à 1500 V. L'acquisition des images des gels a été effectuée par le séquenceur pour sa lecture sur le logiciel AFLP-QuantarTM Pro 1.0 (KeyGene)

Les sept couples d'amorces retenus (AAT-CGA, AAT-CGG, AAT-CGT, ACT-CAC, ACT-CGA, AGC-CCA et AGC-CGT) ont été utilisés sur 59 échantillons choisis sur la base de la

diversité observée avec les marqueurs SSR. Ainsi, les formes cultivées, sauvages tubéreuses et sauvages non tubéreuses furent représentées par 27, 31 et 4 plantes, respectivement.

Les bandes AFLP polymorphes ont été codées comme présentes (1), absentes (0) ou manquantes (9) en cas d'ambiguïté. Seules les bandes claires et présentes chez au moins deux accessions différentes ont été considérées.

3.3.4. Génotypage de l'ADN chloroplastique: Relations phylogénétiques entre les membres de *A. xanthorrhiza* et d'autres formes sauvages du genre *Arracacia* estimées par séquençage de des espaceurs chloroplastiques (ADNcp)

Cette étude est basée sur le séquençage de deux espaceurs intergéniques chloroplastiques (*trnQ-rps16* et *trnS-trnfM*) à l'aide d'amorces universelles *trnQ-rps16* et *trnS-trnfM* (Demesure et al. 1995; Hahn 2002). L'intérêt principal est le positionnement du type cultivé vis-à-vis de leurs types sauvages et d'autres espèces du genre aussi présentes dans l'aire de domestication. Pour ce faire, nous avons analysé deux représentants du type cultivé (des groupes G1 et G2 révélés par les AFLPs, voir partie III), deux représentants du type polycarpique et monocarpique et un des espèces *A. incisa*, *A. elata* et *A. acuminata*. Les produits PCR ont été obtenus en suivant les conditions d'amplification détaillées dans Demesure et al. (1995) et Hahn (2002), puis séquencés directement par l'emploi des deux amorces utilisées pour la PCR.

3.4. L'analyse des données

3.4.1. Données microsatellites

a) Diversité allélique:

Comme expliqué précédemment, les calculs de diversité génétique ont été effectués en se basant sur l'hypothèse de diploïdie de *A. xanthorrhiza*. Cette hypothèse semble se confirmer par l'observation d'un maximum de deux bandes pour les 11 locus retenus dans ce calcul. La diversité génétique a été estimée à l'aide des paramètres de diversité allélique (comme le PIC, *Polymorphism Information Content*) et d'hétérozygotie. Le calcul de l'hétérozygotie pour chaque locus (observée H_o et attendue H_e à l'équilibre de Hardy-Weinberg, HW) permet une estimation de la valeur F_{is} . Le F_{is} est un index de déviation par rapport aux proportions attendues à l'équilibre dans les sous-populations considérées, avec un déficit (valeur positive) ou un excès (valeur négative) d'hétérozygotes, l'hypothèse nulle étant l'association aléatoire

des allèles. Les valeurs de Fis ont été obtenues sur GENETIX (Belkhir et al. 2000) et leur signification statistique fut testée sur GENEPOP ver. 3.1 (Raymond et Rousset 1995).

b) Structure génétique:

Pour décrire les relations génétiques au sein du complexe sauvage et cultivé d'*A. xanthorrhiza*, des arbres ont été construits à partir des fréquences alléliques selon les méthodes suivantes: Neighbor Joining (NJ) (Saitou et Nei 1987) à partir des matrices de distances entre individus calculés avec les distances de CORD (Dc), (Cavalli-Sforza et Edwards 1967) et DAS (Chakraborty et Jin 1993). Ces distances se basent sur la proportion d'allèles partagés qui semble appropriée pour des populations ayant récemment divergé. Les arbres ont été générés avec le logiciel POWERMARKER 3.23 (Liu et Spencer 2001) et la visualisation a été faite avec le logiciel TREEVIEW 1.6.6 (Page 1996).

La structure génétique du complexe *A. xanthorrhiza* a été aussi visualisée par des analyses multivariées. Des Analyses en Coordonnées Principales (PCO) ont été obtenues par NTSYS ver.2.1 (Rohlf 2002). La PCO est une technique multivariée qui permet de trouver des axes majeurs de variation (et non plus un seul comme dans UPGMA ou NJ). Le calcul mathématique permet de définir des axes orthogonaux (extrayant chacun une fraction indépendante de la variance globale), sur lesquels se distribuent les individus observés. Elle s'apparente à l'Analyse en Composantes Principales mais s'applique, non pas aux données d'observation, mais à une matrice de similarité de ces observations entre les individus analysés. Typiquement les 2 ou 3 premiers axes orthogonaux sont considérés. Dans des analyses successives, on se concentre sur différents sous-ensembles d'individus entre lesquels on veut vérifier s'il existe différents axes majeurs de variation.

Afin d'estimer la différenciation génétique entre compartiment sauvage et cultivé, l'estimateur non-biaisé du Fst (Weir et Cockerman 1984) a été calculé avec FSTAT ver.2.9.3.2 (Goudet 2001). Pour des marqueurs présumés neutres tels que les SSRs, le Fst est un indicateur de la cohésion de l'ensemble des sous-populations considérées; dans ce cas le compartiment sauvage ou cultivé. La signification statistique du Fst fut testée par le test G après les corrections de Bonferroni sur le même logiciel. Ce test est suggéré comme plus approprié pour des groupes à effectifs déséquilibrés et/ou peu de locus (Raymond et Rousset 1995).

c) Interactions sauvage-cultivé: Identification d'individus hybrides

Un test d'assignation a été appliqué afin de confirmer le statu présumé hybride de certains échantillons. Ce test est une méthode qui consiste à assigner des individus à des populations à partir de données moléculaires en utilisant des approches statistiques comme les distances génétiques, les fréquences alléliques ou les méthodes bayésiennes. Dans notre cas, nous avons employé le logiciel STRUCTURE ver. 2.0 (Pritchard et al. 2000) qui utilise la méthode bayésienne pour inférer des populations (K) à partir des fréquences alléliques. Cette méthode fournit une probabilité d'origine de chaque individu dans chaque population (valeur q). Elle considère que toutes les populations sont à l'équilibre de HW et que les locus sont indépendants à l'intérieur des populations. Les individus sont assignés à chaque population si la probabilité d'origine calculée est supérieure à un seuil choisi ($T = 0.95$ à 0.999). Le modèle utilisé pour l'assignation des individus a été le « *Correlated Allele Frequencies Model* » qui se base sur la notion selon laquelle la fréquence d'un allèle dans deux populations différentes est en relation, soit avec la probabilité qu'elles partagent un ancêtre commun, soit à l'existence d'individus migrants entre ces deux populations.

Nous avons fait des assignations avec et sans information *a priori* sur l'appartenance des individus (cultivés ou sauvages) et avec un nombre donné de populations (K) de 2 et 3. Les paramètres du calcul ont été choisis selon les recommandations données par Pritchard et Wen (2004).

3.4.2. Données AFLPs

La structure génétique a été analysée avec un Cluster et une PCO à l'aide des logiciels PAUP (Swofford 2001) et GenALEx ver. 6 (Peakall et Smousse 2001), respectivement. Le Cluster Analysis a été obtenu avec la méthode UPGMA à partir d'une matrice de distances calculée avec l'option « *Mean Character Difference* » sur PAUP. Une analyse bootstrap a été calculée afin de déterminer la robustesse des groupes après 100 répétitions. Pour la PCO sur GenALEx, l'algorithme d'Orlocus (1978) a été utilisé. Sur ce même logiciel, une décomposition de la variance moléculaire (AMOVA) est proposée. L'AMOVA est une méthode statistique qui permet l'estimation de statistiques-F et/ou de ses analogues, par la décomposition de la variation entre groupes (populations, régions...). Le calcul de l'AMOVA sur GenALEx est adapté à différents types de marqueurs génétiques et un test statistique par permutation aléatoire est aussi proposé. Ce logiciel permet ainsi une estimation de la différenciation entre populations sur des données binaires comme les AFLPs. Le paramètre Φ_{PT} , analogue au traditionnel F_{st} , détermine des valeurs positives oscillant entre 0 et 1 selon

la différenciation génétique entre populations (Hartl et Clark 1997). Dans notre analyse, le support statistique de Φ_{PT} fut évalué avec 1000 permutations aléatoires. Une AMOVA a été calculée entre formes sauvages et cultivées, puis entre les groupes déterminés par l'analyse UPGMA.

Finalement, la distance génétique de Nei (Nei 1978) a été calculée sur GenAlEx entre espèces et les groupes révélés par l'analyse Cluster. Ce paramètre est fréquemment utilisé pour estimer la distance génétique entre populations.

3.4.3. Séquences chloroplastiques

L'alignement des séquences a été réalisé sur DNASTAR et l'analyse phylogénétique sur PAUP ver. 4b8. Les insertions/délétions (indels) ont été codées en mode binaire et ajoutées à la matrice des nucléotides (Pintaud, com. pers). Le critère de maximum de parcimonie a été appliqué, puis les options « *random step wise addition* » et « *tree bisection reconstruction* » ont été suivies. Un arbre consensus a été obtenu par la méthode « consensus strict » et une analyse bootstrap a été appliquée avec 100 permutations.

4. RESULTATS ET DISCUSSION

La présentation des résultats est structurée en quatre sections de la manière suivante.

- i) une première partie avec les résultats d'isolation et d'identification des SSR utilisés pour l'étude de la diversité au sein du complexe *A. xanthorrhiza*. Ces résultats sont présentés sous la forme d'une Primer Note parue dans *Molecular Ecology*; puis une réflexion concernant le niveau de ploïdie chez *A. xanthorrhiza* sur la base des profils microsatellites observés.
- ii) une deuxième partie avec les résultats de la diversité SSR du complexe sauvage-cultivé de *A. xanthorrhiza* permet, en comparant la diversité génétique des deux compartiments, d'évaluer l'impact de la domestication et de possibles interactions entre formes sauvages et cultivées dans l'aire présumée de domestication.
- iii) une troisième partie avec les résultats des AFLPs et du séquençage/génotypage chloroplastique. Cette information permet d'identifier la forme sauvage ancestrale probablement à l'origine du type cultivé. Par ailleurs, les AFLPs révèlent une information additionnelle intéressante; il s'agit d'une divergence au sein de la forme cultivée que l'on explique par l'introgession du type sauvage monocarpique. On présente ces résultats sous forme d'un article.
- iv) une discussion générale et les conclusions concernant ce chapitre.

I partie: Identification des marqueurs SSR chez *A. xanthorrhiza*

OBJECTIF

Pour disposer de marqueurs SSR utiles pour l'étude de diversité chez *A. xanthorrhiza*, nous avons développé une banque microsatellite à partir de l'ADN génomique d'un cultivar traditionnel.

RESULTATS

Les détails concernant l'isolement et l'identification des marqueurs SSR chez *A. xanthorrhiza* sont précisés dans une Primer Note parue en 2004 (Annexe 2). Lors de l'élaboration d'une banque de clones SSR, 14 SSR ont révélé du polymorphisme au sein du complexe *A. xanthorrhiza*. Pour 11 d'entre eux, un maximum de deux bandes est observé par génotype ce qui permet de présumer un profil diallélique à 11 locus différents. Néanmoins, pour les trois autres, un profil de 2, 3 ou 4 bandes est observé. Dans tous les cas, les bandes sont d'intensité sensiblement égale pour un même génotype. Nous présentons ci-dessous une réflexion concernant le niveau de ploïdie d'*A. xanthorrhiza* sur la base des profils d'amplification SSR observés.

DISCUSSION

L'observation de profils d'amplification à 3 et 4 bandes pour certains SSR peut être expliquée par plusieurs hypothèses: tétraploïdie, séquences flanquantes dupliquées, chimères somatiques.

Bien que nous ayons conclu de l'analyse bibliographique que le genre *Arracacia* était diploïde, nous ne pouvons exclure que ce soit un ancien tétraploïde diploïdisé. La polyploïdisation est fréquente chez les plantes. Le séquençage global de génomes comme *Arabidopsis* et *Oryza*, ou partiel comme dans le cas du blé, montre que l'histoire de ces génomes inclue de nombreuses duplications de segments chromosomiques. Des amorces de PCR définies sur la base d'une sonde SSR peuvent, dans ces conditions, amplifier plus d'un locus. Dans notre cas, sur 14 SSR retenus pour révéler du polymorphisme, 11 ne révèlent au maximum que deux bandes par génotype alors que le total de bandes observées atteint 5 à 14 sur l'ensemble des génotypes. Sur cette base et compte tenu de la reproduction sexuée

allogame de notre plante, l'interprétation parcimonieuse de ces résultats est que chacun de ces 11 SSR ne révèle la diversité qu'à un seul locus et que pour les 3 restants où un nombre de bandes allant jusqu'à trois ou quatre est observé, il est possible que nos amorces révèlent des segments chromosomiques dupliqués.

Un facteur par ailleurs intéressant chez *A. xanthorrhiza* est que le profil à 3 et 4 bandes chez deux locus SSR distinguent les formes sauvages des cultivés. En effet pour le locus AxC87, les formes cultivées présentent de 1 à 4 bandes alors que les formes sauvages montrent une ou deux bandes. De même, pour le locus AxC64, les formes cultivées présentent de 1 et 4 bandes et les sauvages généralement une ou deux bandes (sauf chez certains individus qui révèlent aussi 3 et 4 bandes, mais qui se montrent intermédiaires, voir dans les résultats de la détection d'individus hybrides). Le séquençage des fragments amplifiés montre que dans le cas du locus AxC87, les séquences flanquantes sont homologues et suggèrent l'allélisme des fragments amplifiés. Une hypothèse pour expliquer le patron observé serait donc que, vue l'occurrence d'allèles nuls chez les formes sauvages, il pourrait s'agir d'un locus dupliqué et que nos amorces n'amplifient qu'un seul locus chez les formes sauvages alors que deux sont amplifiés chez les formes cultivés. En revanche pour le locus AxC64 notre interprétation reste limitée car les séquences flanquantes des fragments amplifiés révèlent un manque d'homologie bien que l'on retrouve les sites d'hybridation d'amorces.

En revanche, pour un troisième locus AxD55, deux zones alléliques sont bien distinguées pour tous les individus de *A. xanthorrhiza* analysés, cultivés ou sauvages, voir chez d'autres espèces divergentes du genre. Ces zones sont décalées de 150pb l'une de l'autre, et sur chacune un ou deux allèles par individu sont révélés. Dans le cas de ce locus, l'hypothèse d'une duplication ancienne semble appropriée car par ailleurs l'indépendance entre ces deux zones est aussi suggérée statistiquement (par une analyse de déséquilibre de liaison dans la comparaison deux à deux entre les différents loci). On a donc considéré chacune de ces zones alléliques comme un locus indépendant (Voir Primer Note en annexe 2).

Dans l'état actuel, le niveau de ploïdie d'*A. xanthorrhiza* reste à confirmer. Pour l'instant, il nous semble plus adéquat d'assumer l'hypothèse de diploïdie chez *A. xanthorrhiza* (compatible avec celle d'un ancien tétraploïde diploïdisé) en se basant sur les 11 locus à 1 ou 2 allèles malgré la détection d'amorces qui amplifient deux locus simultanément (car la plupart des logiciels utilisés sont adaptés à des données de type diploïde). Cette observation

peut être attribuée à des duplications ancestrales bien qu'on ne puisse pas écarter l'occurrence possible de chimères parmi les formes cultivées comme démontré chez la vigne (Franks et al. 2002).

Par la suite, il conviendra d'analyser la ségrégation de ces marqueurs lorsqu'une descendance sera disponible. Cette information permettrait d'éclaircir les patrons d'hérédité de ces marqueurs.

II partie: Diversité génétique et interactions au sein du complexe sauvage-cultivé de *A. xanthorrhiza* estimées par l'emploi des marqueurs SSR

OBJECTIFS

Les objectifs de cette partie sont les suivants:

1) Etudier la structure génétique du complexe sauvage-cultivé de *A. xanthorrhiza* à l'aide des marqueurs SSR. Comme déjà souligné, les marqueurs SSR sont appropriés car la diversité chez les formes sauvages et cultivées était presque inconnue avant notre travail, mais présumée faible. Ces marqueurs permettent aussi une quantification de la diversité allélique entre ces deux compartiments allogames (bien que le pourcentage d'allèles nuls soit largement sous estimé comme on le verra dans le chapitre II). Cette information nous permettra d'estimer la réduction de diversité liée au processus de domestication chez *A. xanthorrhiza*.

2) Déterminer si un flux génique entre formes sauvages et cultivées a eu lieu dans l'aire de domestication. En effet, les formes sauvages sont généralement isolées mais occasionnellement, des situations sympatriques avec la forme cultivée ont été rapportées et nous en avons observé un cas. La possibilité de flux génétique a été proposée par certains mais aucune étude n'a été menée jusqu'à présent.

RESULTATS

1) Diversité SSR et structure génétique dans le complexe *A. xanthorrhiza*

a) Diversité génétique

L'analyse de diversité a pu être menée exclusivement entre les formes tubéreuses (Complexe *A. xanthorrhiza* et l'espèce *A. incisa*), car sur d'autres espèces sauvages les taux de transférabilité d'amorces furent très faibles (voir Primer Note en Annexe 2). Comme précisé précédemment nous nous limitons ici à l'analyse de diversité aux 11 locus ne révélant pas plus de deux bandes chez un même individu. Chez les 178 individus du complexe *A. xanthorrhiza*, le nombre total de bandes et donc d'allèles présumés est de 102. L'espèce *A. incisa*, représentée par trois plantes qui se révèlent du même génotype, partage avec *A. xanthorrhiza* 14 allèles sur les 11 locus mais montre des allèles nuls sur deux locus (AxD55A et AxD43) et un allèle privé (sur le locus AxD72); cette espèce a donc été considérée comme un groupe externe.

Le nombre d'allèles distingués par locus oscille entre 5 et 14 avec une moyenne de 9.9. Le tableau 3 indique la répartition du nombre d'allèles entre les différents locus et entre formes sauvages et cultivées d'*A. xanthorrhiza*. Le tableau 4 indique pour chaque locus et en moyenne, la fréquence de l'allèle le plus commun, le nombre de génotypes ainsi que les valeurs d'hétérozygotie et de PIC (*Polymorphism Information Content*) observées. La valeur PIC dépasse dans tous les cas 0.50 en variant jusque 0.86. Le nombre de génotypes distingués par locus varie de 10 à 49. Au niveau multilocus, le même génotype n'a pas été observé plus d'une fois parmi les individus sauvages, par contre parmi les cultivars, cinq génotypes sont représentés par plus d'une accession et correspondent donc à des clones présumés.

La fréquence de l'allèle le plus commun aux différents locus dépasse dans un seul cas 0.77 (locus AxD13). Aux autres locus, cette fréquence varie de 0.21 à 0.60 (Ex. locus AxD82 qui a la valeur la plus élevée de PIC). La plus grande différence de taille entre allèles d'un même locus a été observée pour le locus AxC85 avec 48 pb. Les locus AxC27 et AxD13 présentent la plus petite différence entre allèles observés avec 10 pb. Les autres locus ont des allèles distingués au maximum par 16 et 40 pb. Le nombre d'allèles par locus augmente en fonction de cette différence de taille; ainsi le locus AxC85 est un de ceux qui révèlent le plus d'allèles, par opposition aux locus AxC27 et AxD13 où seulement 5 allèles ont été observés.

Par ailleurs, les tests de déséquilibre de liaison entre les différents locus ont montré parmi toutes les combinaisons possibles, une déviation significative entre les locus AxD34 et AxD85 ($p < 0.001$). Comme on ne dispose d'aucune information quant au positionnement chromosomique de ces locus, on ne peut pas conclure sur l'origine du déséquilibre.

Polymorphisme des SSR chez les formes sauvages et cultivées

Tous les allèles observés dans le complexe *A. xanthorrhiza* le sont dans les formes sauvages, mais seuls 31 des 102 allèles le sont parmi les formes cultivées. Par ailleurs, des allèles nuls (absence d'amplification) ont été observés parmi les individus sauvages avec une fréquence significative: 10.4% des locus/individu, alors que cette estimation néglige les hétérozygotes comprenant un allèle nul et un allèle exprimé. La forme cultivée présente une moyenne de 2.1 allèles exprimés par locus alors que cette valeur augmente à 10 chez les formes sauvages. Ceci est traduit par les valeurs de polymorphisme obtenus (PIC): 0.35 pour les cultivés et 0.60 pour les sauvages (Tableau 4). Le tableau 5 compare les fréquences de 31 allèles partagés entre les formes sauvages et cultivées d'*A. xanthorrhiza*, et le tableau 6 montre les fréquences de ces allèles parmi les formes sauvages classées en trois groupes selon l'origine géographique : Centre Equateur, Sud Equateur et Nord Pérou. On note des différences très importantes de fréquences alléliques entre ces groupes et la forme cultivée. Dix allèles sur les 31 ne sont pas représentés dans le groupe de sauvage Centre Equateur alors qu'ensemble, les deux groupes Sud Equateur et Nord Pérou présentent tous les 31 allèles observés dans la forme cultivée.

Le tableau 7 compare l'hétérozygotie et les valeurs *Fis* observées entre les formes sauvages et cultivées. L'hétérozygotie observée est en moyenne de 0.46 parmi les formes cultivées et de 0.43 chez les formes sauvages alors que chez ces dernières une valeur de 0.69 est attendue en panmixie. Les valeurs *Fis* (qui reflètent un déficit ou excès d'hétérozygotes parmi les individus analysés) sont en effet significativement positives chez les formes sauvages. Ce résultat suggère un large déficit en génotypes hétérozygotes au sein de ce compartiment. Chez la forme cultivée en revanche, ces valeurs sont le plus souvent négatives et seulement significatives dans deux cas (AxD82 et AxD55B, $p < 0.01$), ce qui suggère au contraire des formes sauvages, un excès d'hétérozygotes dans ces deux locus.

b) *Structure génétique du complexe A. xanthorrhiza*

La structure du complexe *A. xanthorrhiza* est représentée par un arbre NJ et le premier plan d'une analyse PCO. L'arbre NJ montré (Fig. 16) a été obtenu à partir de la matrice de distances DAS mais un regroupement similaire est observé avec la distance Dc. Les formes sauvages sont généralement portées par des branches longues liées à leur plus grande diversité. Les formes cultivées se regroupent indépendamment du groupe sauvage et sont portées par des branches courtes. Il existe néanmoins un petit groupe de transition entre ces deux groupes, avec un cultivar (C60) et cinq individus sauvages S3, S18, S23, S26 et S37. Ces échantillons ont été collectés dans la région de Cajamarca au Pérou, à l'exception de l'échantillon S3 originaire de la province de Loja (qui est en Equateur la Province la plus proche de Cajamarca) et prélevé près d'un petit champ où la forme cultivée était présente, alors qu'une population sauvage abondante (S4) se trouvait à une centaine de mètres de ce champ. Sur l'arbre NJ, les accessions S4 (1 à 4) sont aussi proches des formes cultivées. Il est remarquable que le reste des plantes sauvages tend à se regrouper selon les populations d'appartenance et les populations selon leur origine géographique alors qu'une structuration selon leur distinction botanique (en *A. x. var. andina* ou *A. x. var. monocarpa*) n'est pas clairement observée. On constate cependant que les formes monocarpiques sont regroupées au centre de l'ensemble des formes sauvages et moins dispersées. Les trois échantillons de l'espèce tubéreuse *A. incisa*, représentés par un seul génotype, sont situés comme attendu en groupe externe.

La distribution en PCO des échantillons de *A. xanthorrhiza* est présentée sur la figure 17a sur le plan des deux premiers axes qui représentent près d'un tiers de la variation totale observée (27.6%). L'axe 1, responsable de 19.6% de la variation, sépare les formes sauvages et cultivées. Les individus sauvages trouvés ci-dessus affins du groupe des cultivés sont aussi localisés par la PCO en position intermédiaire. Les formes sauvages, étalées tout au long du deuxième axe, révèlent une diversité importante, alors que les formes cultivées, à l'exception de deux cultivars qui semblent plus affins des formes intermédiaires (C60 et C16), forment un groupe plus dense.

Au contraire de l'arbre NJ où un regroupement des formes sauvages par origine géographique est davantage visible qu'une distinction botanique et où les formes monocarpiques sont donc dispersées, on constate dans la PCO un regroupement des formes monocarpiques originaires des différentes localités au niveau de l'axe 2. On retrouve aussi en position centrale, des

formes pérennes de Loja et Cajamarca associées à certains individus identifiés comme monocarpiques, et dans la partie négative de l'axe on retrouve uniquement les formes pérennes du centre de l'Equateur.

Une deuxième analyse PCO comprenant uniquement les formes sauvages tubéreuses est montrée sur la figure 17b. Les deux premiers axes extraient chacun environ 10% de la variation génétique. On observe un groupe pur de formes monocarpiques distingué dans la figure 17a et globalement la même structure que précédemment. Cependant les zones géographiques Nord Pérou, Sud Equateur et Nord Equateur sont mieux distinguées (indiqués sur la figure). Le reste des formes monocarpiques n'est pas mieux distingué que dans la fig. 17A mais reste proche uniquement des formes polycarpiques du Nord Pérou et Sud Equateur.

Estimation de la différenciation génétique

La valeur F_{st} calculée entre les types sauvages et cultivés suggère une forte différenciation entre les deux compartiments ($F_{st} = 0.23$; avec un intervalle de confiance de 0.15 - 0.31 selon les paramètres F de Wei&Cockerham 1984). Cependant, en distinguant dans cette comparaison les formes polycarpiques et monocarpiques « purs », on obtient une valeur de différenciation génétique inférieure avec le type polycarpique ($F_{st} = 0.240$, $p < 0.01$) qu'avec le type monocarpique ($F_{st} = 0.388$, $p < 0.01$). En revanche, ces deux types sauvages se montrent moins différenciés entre eux ($F_{st} = 0.125$, $p < 0.01$).

2) Interactions entre formes sauvages et cultivées: détection d'hybrides spontanés

Les analyses de structure génétique (NJ et PCO) montrent une différenciation entre formes sauvages et cultivées mais suggèrent l'existence d'individus intermédiaires entre les deux compartiments. La figure 18 montre l'arbre NJ incluant, en plus du matériel analysé ci-dessus, 13 hybrides expérimentaux entre *A. x. var. andina* et l'arracacha cultivée. La structure de l'arbre NJ est similaire à celle du précédent (figure 16), avec les hybrides artificiels formant un groupe intermédiaire où sont aussi compris d'autres individus (les individus sauvages S36 et S37 et les cultivars C60 et C66). L'analyse PCO (Fig. 19a) montre de la même façon les hybrides artificiels dans une position centrale entre les types sauvages et cultivés, de même que des individus intermédiaires (S37, S26, C60).

Test d'assignation des génotypes:

Les analyses de structure génétique, avec ou sans hybrides expérimentaux entre *A. xanthorrhiza* sauvage (var. *andina*) et l'arracacha cultivée, suggèrent l'existence de certains individus intermédiaires spontanés. Le tableau 8 montre les valeurs d'appartenance (valeurs Q) données à chaque individu par le test d'assignation du programme STRUCTURE avec un nombre prédéfini de populations (K) de 2 et 3. On peut constater sur ce tableau que lorsque K = 2, les individus sont classés en types sauvage et cultivé sauf pour huit individus qui présentent des valeurs intermédiaires, dont six plantes sauvages et deux hybrides artificiels. Une représentation graphique de ces résultats est présentée dans la figure 19b. Lorsque K=3, les individus sauvages sont par contre différenciés en pérennes ou monocarpiques avec certains intermédiaires entre les deux.

Des admixtures sont donc observées comme suit:

1) entre le type cultivé et A.x.var.andina

C'est le cas des hybrides artificiels, de quatre échantillons attribués à *A. x. var. andina* (S18, S26, S36 et S37) et d'une accession cultivée (C60). Les huit hybrides du premier croisement (C68 x S37) présentent des valeurs d'appartenance entre 31 et 67 % pour le type cultivée et *A. x. var. andina* respectivement, alors que chez les cinq hybrides du deuxième croisement (C60xS32) ces valeurs sont de 21 et 77%. Cette différence peut s'expliquer par la nature du parent cultivé utilisé dans le deuxième croisement (C60, AM) car ce cultivar est lui-même le seul qui apparaisse en admixture avec 80% d'appartenance au type cultivée et 20% d'appartenance à *A. x. var. andina*. Parmi les formes sauvages intermédiaires, trois (S18, S36 et S37) sont assignés avoir un quart d'appartenance au type cultivée (21-29%), alors que chez un quatrième (S26) cette valeur est le double (51%).

2) Entre les deux types sauvages:

La plupart des accessions sauvages sont classées selon leur identification botanique, forme pérenne ou monocarpique. Cependant, quatre accessions (S4, S16, S43 et S44) sont assignées mixtes entre les deux types sauvages. Les deux premières ont été collectées en Equateur et leur identification *A. x. var. andina* était douteuse, alors que les deux dernières, collectées vers Cajamarca au Pérou, avaient été identifiées comme *A. x. var. monocarpa* (Knudsen, com. pers) bien que toutes les quatre soient majoritairement assignées à *A. x. var. andina*.

DISCUSSION

Diversité SSR des formes sauvages et cultivées de *A. xanthorrhiza*

La diversité chez la forme cultivée de *A. xanthorrhiza* est significativement moins importante que chez les formes sauvages. Seuls 30% de la diversité allélique microsatellite observée dans le complexe *A. xanthorrhiza* sont présents chez la forme domestiquée. Ce résultat conforte l'hypothèse que la forme cultivée représente seulement une portion de la diversité du compartiment sauvage de l'espèce. Cependant, une forte divergence génotypique entre formes sauvage et cultivée est suggérée par la valeur de différenciation génétique ($F_{st} = 0.23$). Ce résultat peut s'expliquer à partir de plusieurs facteurs:

- i) une grosse portion de la diversité SSR évaluée est absente de la forme cultivée,
- ii) les fréquences des marqueurs partagés entre formes sauvages et cultivés sont différentes au sein de chaque compartiment,
- iii) la présence d'allèles nuls observés chez certaines formes sauvages,
- iv) la recombinaison génétique entre origines géographiques différentes des formes cultivées, entre elles et avec les sauvages locaux.

La présence d'allèles nuls chez certaines formes sauvages (et plus rarement chez certains cultivés) peut être expliquée par des mutations ponctuelles dans la séquence flanquante du motif SSR responsable de l'absence d'amplification ou, dans certains cas, par une plus faible amplification (Pemberton et al. 1995). Cela peut conduire à une sous-estimation des individus hétérozygotes (un hétérozygote avec un allèle nul sera compté comme un homozygote et un homozygote nul sera visualisé comme une absence d'amplification). Chez les formes sauvages de *A. xanthorrhiza*, des déficits en hétérozygotes sont suggérés, et la présence des allèles nuls (qui ne s'expriment qu'à l'état homozygote) peut être liée à ce résultat. Cependant, d'autres hypothèses peuvent aussi l'expliquer: autofécondation ou faible taille effective des populations (effet Wahlund). Au contraire des formes sauvages, l'arracacha cultivée ne présente pas de déficit en hétérozygotes. Ceci peut être lié à la sélection de plantes fortement hétérozygotes soumises à la reproduction asexuée et suggère d'autre part un brassage génétique au sein du type cultivé comme observé par exemple, chez le manioc (Second et Iglesias 2001). Des allèles nuls rares auraient alors peu de chance d'être observés.

En conclusion, la base génétique de la forme cultivée de *A. xanthorrhiza* est plus importante que présumée. Elle est néanmoins modérée si on la compare avec la diversité globale du complexe *A. xanthorrhiza*. Cette observation peut s'expliquer par deux hypothèses:

- a) la domestication a pu impliquer un faible nombre d'individus sauvages et,
- b) il existe une faible incorporation de diversité chez la forme cultivée à partir du compartiment sauvage

Ces hypothèses sont discutées dans la discussion globale du chapitre.

Le compartiment sauvage de *A. xanthorrhiza* explique-t-il toute la diversité chez la forme cultivée? Les cas particuliers de 3 locus SSR

Dans le cas d'une domestication directe à partir d'un type sauvage, on s'attend à que la diversité présente chez le type cultivée soit expliquée entièrement par son homologue sauvage. Or, chez l'arracacha, ce postulat ne s'applique pas sur 3 locus SSR (AxC87, AxC64 et AxC29) dont le profil d'amplification observée révèle une particularité intéressante: des allèles spécifiques à la forme cultivée montrent un décalage en taille avec ceux observés chez les homologues sauvages de l'espèce. A cause de ces caractéristiques, ces SSR permettent une discrimination entre les types cultivés et sauvages, cependant l'origine de ces différences est inconnue. Nous étudierons ce point dans le chapitre II de la thèse.

III Partie: Peut-on préciser l'origine de la forme cultivée?

L'origine de la forme cultivée de *A. xanthorrhiza* reste à confirmer. Les SSR ne sont pas assez résolutifs pour déterminer quel type sauvage est à l'origine de la forme cultivée. En effet, les deux types sauvages sont partiellement distingués par les SSR mais cette distinction reste faible. Nous avons donc utilisé en complément des marqueurs AFLPs qui nous permettent d'élargir notre évaluation de la diversité du génome. Un ensemble d'échantillons représentatifs de la diversité observée au niveau SSR est utilisé dans cette analyse. D'autres espèces sauvages présentes dans la même aire de distribution mais qui n'ont pu être étudiés avec les SSR (compte tenu d'une faible transférabilité d'amorces) sont aussi analysées. Une analyse phylogénétique à partir de l'ADN chloroplastique est également abordée.

Les résultats sont présentés sous forme d'un article Accepté par GENOME sous réserve de modifications.

Article: "Cryptic genetic structure and origin of the Andean root crop arracacha (*Arracacia xanthorrhiza* Bancr. Umbellifereae) revealed by the analysis of AFLP and chloroplast DNA polymorphism analysis"

Cryptic genetic structure and origin of the Andean root crop arracacha (*Arracacia xanthorrhiza* Bancr. Umbellifereae) revealed by AFLP and chloroplast DNA polymorphism analysis

Authors:

E. MORILLO ^{1,2}, M. HERMANN ³, J.L. PHAM ² and G. SECOND ²

¹ Departamento Nacional de Recursos Fitogenéticos y Biotecnología (DENAREF), Instituto Nacional Autónomo de Investigaciones Agropecuarias (INIAP), Casilla postal 17-01-340. Quito, Ecuador

² IRD (Institut de Recherche pour le Développement), UMR DGPC/DYNADIV, 911 Av. Agropolis BP 64501, 34394 Montpellier Cedex 5, France

³ IPGRI, Regional Office for the Americas (AMS), CIAT, Apartado Aéreo 6713, Cali (Colombia)

Corresponding author: E.Morillo. E-mail: morillo@mpl.ird.fr, denaref@ecnet.ec

Abstract

Andean arracacha (*Arracacia xanthorrhiza* Bancr.) is a valuable but poorly known vegetatively reproduced root crop whose origin is still unresolved. AFLP was monitored in a representative sample of this crop and its wild relatives from the presumed areas of domestication, i.e. Ecuador and Peru. Wild species with tuberous and non-tuberous roots were included, as well as perennial and monocarpic life history forms of the presumed wild ancestor. While the two life history varieties of wild *A. xanthorrhiza* were closest to the cultivars, they were well distinguished by AFLPs. Unexpectedly, two clearly distinct groups among the cultivars were also detected, one of which was closer to the monocarpic wild ancestor. However, the chloroplast DNA sequences showed higher similarity between all cultivars and the wild perennial *A. xanthorrhiza*. These results, combined with the morphological and life history features, support the hypothesis that arracacha domestication started from the perennial form, followed by introgression from the monocarpic form, with differentiation of an unsuspected group of cultivars, well distinguished at molecular level.

Key words: plant domestication, crop introgression, asexual propagation, arracacha, perennial, Umbellifereae, AFLPs, chloroplast DNA

INTRODUCTION

A large proportion of crops domesticated by ancient people in South America, as compared to other centers of origin, are propagated vegetatively. Besides potatoes (tuberous *Solanum* spp.), at least another seven species (each belonging to a different botanical family) have been cultivated since pre-Columbian times through vegetative means, especially in the Andes (Hawkes 1989). They include arracacha (*A. xanthorrhiza* Bancr.), the only Umbellifereae (syn. Apiaceae) species domesticated in the New World, while numerous Old World domesticates of the same family are generally propagated by sexual seeds. Arracacha is now grown for its edible and starchy storage roots in many parts of the highlands of Central America, the Caribbean region and the high plateau of southern and southeastern Brazil. A peculiar arracacha organ derived from the fleshy stems, and locally called “colino”, is used for propagation. Colinos emerge from a central main root (rootstock), and are homologous and functionally similar to cormels in araceous root crops (Hermann 1997).

Arracacha belongs to the New World genus *Arracacia*, whose range extends from Mexico to Bolivia. Out of a total of about 30 species, 10 are recognized as being from the South American Andes region, and five of these are endemic (Knudsen 2003). Of these, the wild tuberous forms are considered to be the most closely related to cultivated arracacha. Wild tuberous forms that grow in seasonally dry areas form large central roots, which together with the tuberous storage roots attached to them, enable the plants to survive through dry periods generally lasting several months. According to Valderrama and Seminario (2002) and our field observations in Ecuador and Peru, storage roots of wild tuberous forms are consistently reported by local people to have specific medicinal uses. In ancient times, humans may have gathered wild tuberous forms as a food source, which may eventually have led to the domestication of arracacha. Wild tuberous forms are reported from Peru, Ecuador and Bolivia and, according to their life history, are perennial or monocarpic. The perennials flower and bear fruit yearly (they are also called polycarpic) once their persistent storage roots have grown to a sufficiently large size to support the formation of vigorous generative shoots, which may subsequently bear several hundreds of seeds per plant. By contrast, the monocarpic forms only flower once and an abundant seed set exhausts the plant and concludes the life cycle. In greenhouse cultivation, monocarpic plants behave as biannuals, but in nature more than 2 years of vegetative growth may be needed before the plants flower. Wild perennial and monocarpic plants occur in large distinct and geographically isolated

populations, but mixed populations of perennial and monocarpic plants have not been observed (Hermann 1997; Knudsen 2003; Blas 2005).

Wild perennial forms have been described under several species names (i.e. *A. andina* Britton, *A. equatorialis* Constance, *A. peruviana* Wolff, *A. incisa*, Hermann 1997). However, more recently, Knudsen (2003) claimed that these species are synonymous (with the exception of *A. incisa*) and proposes the species name *A. xanthorrhiza* to encompass a complex including the cultivars (*A. x. var. xanthorrhiza*) and two wild varieties, the perennial (*A. x. var. andina*) and monocarpic (*A. x. var. monocarpa*) life forms. The life history form of wild arracacha from which these cultivars originated is unknown. Both *A. x. var. andina* and *A. x. var. monocarpa* grow throughout a broad-ranging area in the Andes and are a priori possible ancestors of cultivated *A. xanthorrhiza* (Knudsen 2003, Blas 2005). However, there is higher morphological and phenological similarity between the cultivars and the wild perennial (Hermann 1997; Knudsen 2003; Blas 2005). We attempted to unravel the origin of cultivated *A. xanthorrhiza* through an amplified fragment length polymorphism (AFLP) analysis and chloroplast DNA (cpDNA) sequence survey of a representative set of accessions of cultivars and wild species from the presumed domestication area. Our results highlighted an unsuspected strong but cryptic genetic structure among the cultivars, and provided clues to their origin over the course of their domestication.

MATERIAL AND METHODS

Plant material. A total of 58 plant samples were used in the AFLP analysis (Table I-1) to represent the following species: *A. xanthorrhiza*, *A. incisa*, *A. elata*, *A. moschata* and *A. acuminata*. From Ecuador, these included morphologically representative cultivars from a collection that is vegetatively maintained at INIAP (Mazón et al. 1996), and wild species as well as additional cultivars gathered during collection trips in 2002 and 2003. From Peru, leaf samples were provided by S. R. Knudsen (KVL University, Denmark) for both wild species and cultivars. Overall, the accessions assembled for the tuberous species were as follows: cultivated *A. xanthorrhiza* (27), wild perennial *A. xanthorrhiza* (14), wild monocarpic *A. xanthorrhiza* (11) and *A. incisa* (1). Live material of the Ecuadorian accessions is conserved at INIAP (Quito, Ecuador) and Peruvian materials are conserved at the Universidad Nacional de Cajamarca (CPUN).

AFLP analysis: Young leaves dried over silica gel were used for DNA extraction following a metalxanthate method (Jhingan 1992). The original AFLP protocol (Vos et al. 1995) was used. Seven primer pair combinations out of 24 pairs tested were selected according to the polymorphism observed: E-AAT/M-CGA, E-AAT/M-CGG, E-AAT/M-CGT, E-ACT/M-CAC, E-ACT/M-CGA, E-AGC/M-CCA and E-AGC/M-CGT. AFLP products were separated according to the protocol of Roy et al. (1996) using automated infrared fluorescence technology with a LI-COR sequencer (IR2; LI-COR Biosciences). Gel analyses were carried out using AFLP-Quantar™ Pro 1.0 (KeyGene).

Chloroplast DNA sequencing (CpDNA): PCR was performed using two pairs of universal primers amplifying non-coding chloroplast regions: trnQ-rps16 and trnS-trnfM (Demesure et al. 1995; Hahn 2002). The amplification products were purified and sequenced in both directions using forward and reverse primers. In order to screen for the presence of a specific Q16 insertion, we designed the following primer combination: RPS16-R (5'CAAGTCCGACGTTGCTTTCTACCACATCGTTT3') and Q16-ExF1 (5'TTTTTGTCAATCCATTTATCTGCT3').

Data analysis: AFLP bands were scored as 1 if present, 0 if absent and 9 in case of ambiguity. Only bands of similar intensity across samples were recorded in order to minimize missing data in the data matrix. Bands observed in a single accession were not recorded. Genetic structure was analyzed using cluster and a principal coordinate analysis (PCoA) based on a distance pairwise comparison. For the cluster analysis, we used the mean character difference and the UPGMA method in the PAUP ver. 4b8 software package (Swofford 2001); the bootstrap values were derived from 100 replications. PCoA was carried out with GenAIEx ver. 6 software, which uses Orloci's algorithm (Peakall and Smouse 2001). Nei genetic distances (Nei 1972) and an analysis of molecular variance (AMOVA) between the examined species and intraspecific groups determined were also performed. AMOVA provides an estimate of Φ_{PT} that is an analogue of F_{ST} when binary data are analysed (Peakall and Smouse 2001). A significance test of Φ_{PT} values was performed using 1000 random permutations.

For the CpDNA survey, DNA sequences were aligned with DNASTAR ver.5.07 software (DNASTAR 2003). Insertions or deletions (indels) were jointed as binary data to the nucleotide

matrix for the analysis. The phylogenetic analysis was carried out using PAUP with the maximum parsimony method.

RESULTS

AFLP polymorphism

The seven AFLP primer combinations produced 235 scored polymorphic bands from the 58 samples examined. The polymorphism rate ranged from 42 to 69% for the different primer combinations, as detailed in Table I-2. Out of the 235 scored AFLPs, 198 fragments were polymorphic among the tuberous forms (including *A. incisa* and *A. xanthorrhiza*) and 104 among the cultivars themselves. Cultivated *A. xanthorrhiza* showed less polymorphism than the wild species: 42% of the scored bands were polymorphic among the cultivars, 59% among the wild perennials and 64% among the wild monocarpic forms. No AFLP fragment was diagnostic (i.e. one fragment displayed by all accessions from one group and none from the other) of the cultivars or the wild *A. xanthorrhiza*. However 10 and 65 bands were observed only in the cultivars and wild forms, respectively.

Genetic relationships revealed by AFLPs

A UPGMA tree obtained via the distance matrix in a pairwise comparison of accessions is shown in Figure I-1. The tuberous forms were clearly clustered with a bootstrap value of 100, and four groups of accessions were distinguished as follows: G1 and G2 included only cultivars and were supported by bootstraps of 73 and 77, respectively. G1 consisted of 16 Ecuadorian accessions and G2 pooled 11 accessions, including the six Peruvian cultivars analyzed. G3 included all the wild perennial *A. x. var. xanthorrhiza* samples, along with three collections from Cajamarca (S42, S43 and S44) identified as *A. x. var. monocarpa* (Knudsen pers. com). G4 pooled the remaining monocarpic samples and the *A. incisa* sample. The non-tuberous species were located in a more distant cluster (G5). This divergence between the cultivars and the various wild species was also revealed by the Nei's pairwise genetic distance results (Table I-3). However, among the wild tuberous forms, *A. incisa* appeared to be more distant from the cultivars whereas the perennial *A. xanthorrhiza* was closer. The Nei's genetic distance was 0.123 when comparing only the cultivated and wild *A. xanthorrhiza* samples. AMOVA determined a Φ_{PT} value of 0.26 ($p = 0.001$), which means that 26% of the overall variation could be explained by differences between the two compartments and the remaining 74% was within the cultivated and wild *A. xanthorrhiza*.

Within the *A. xanthorrhiza* complex, the plane defined by the first two axes of the PCoA (Figure I-2) more clearly highlighted the relationship between the same four groups distinguished in the UPGMA tree. The first axis (33% of the total variance) separated the G1 cultivar group and the wild monocarpic accessions (G4), leaving the wild perennial accessions (G3) and the G2 cultivar group in an intermediate position. These latter two groups were separated along the second axis (30% of the total variance). The Nei's genetic distance matrix shown in Table I-4 confirmed that G1 was closer to the wild perennial whereas G2 was closer to the wild monocarpic life forms.

The genetic divergence of the two groups of wild *A. xanthorrhiza* (G3 and G4) was supported by a Φ_{PT} value of 0.24 ($p = 0.001$), while the two cultivar groups were also supported by a Φ_{PT} value of 0.47 ($p = 0.001$), so 47% of the variation among cultivars was between these groups. 18.2% of the AFLP bands polymorphic among the cultivars were specific to the G1 or G2 cultivars groups, respectively.

Phylogenetic relationships revealed by cpDNA

We sequenced two cultivated accessions (C1 and C62, representing the two cultivar groups distinguished by AFLP), two wild perennial accessions (P2 and P6), two wild monocarpic accessions (M4 and M7) and one sample of each of the following species: *A. incisa* (RBS-001), *A. elata* (MST-036) and *A. acuminata* (MST-015). The Q16 and S-FM primer pairs amplified a 1250 bp and 1110 bp PCR product, respectively. This represented a total of 2000 aligned base pairs, 1969 of which were constant, 7 were not parsimoniously informative and 24 were informative, including five observed indels. The Q16 locus was most informative with 16 informative data. No polymorphism was observed between the cultivated and wild tuberous arracacha, except for a 13 bp insertion located in the Q16 region. Interestingly, the sequences showed that this insertion was present in the cultivated and the wild perennial *A. xanthorrhiza*, *A. incisa* and *A. acuminata*, but absent in the wild monocarpic *A. xanthorrhiza* and *A. elata* accessions.

The parsimonious phylogram based on the 24 informative SNPs and indels is shown in Figure I-3. The species *A. incisa* was closest to the *A. xanthorrhiza* cluster whereas *A. elata* was most divergent. A single branch, with a bootstrap value of 80, pools the cultivated and the wild perennial *A. xanthorrhiza*, separated from the branch bearing the wild monocarpic samples only by the described insertion. Screening of this insertion in the whole set of samples by PCR (in which an amplified fragment of approximately 260 bp or 247 bp was obtained)

showed that: 1) this insertion was not observed in the monocarpic accessions, except in the three accessions included in the wild perennial group (group G3); and, 2) this insertion was present in all the cultivated and wild perennial samples (groups G1, G2 and G3, Fig. 1).

DISCUSSION

Since plant domestication is known to be very recent in relation to biological evolution, domesticated forms would be expected to be closely related genetically to their wild ancestors. However, mutations and selection pressure or genetic drift in the course of domestication, as well as possible introgression from other wild relatives, could account for the differentiation between a cultigen and its wild ancestors. Although no wild *Aracaccia* species include all of the observed molecular variation of the cultivars, the present results confirmed the most direct wild ancestors of arracacha, as described below.

Both AFLP and cpDNA sequence polymorphism showed that the non-tuberous species *A. acuminata*, *A. elata* and *A. moschata* were the wild forms the least related to cultivated arracacha. Among the wild tuberous forms, both AFLP and cpDNA distinguished *A. incisa* from the *A. xanthorrhiza* samples, showing that this is the most distant species from the cultivars, so it could not be considered as a direct ancestor of cultivated arracacha. Our results, in agreement with those of Blas (2005) who analyzed a higher number of *A. incisa* samples, reinforced the position of Knudsen (2003) that this type should be considered as a separate species close to the *A. xanthorrhiza* complex. The wild forms of *A. xanthorrhiza* thus appeared to be the most directly related to the cultivars, although clearly distinguished at AFLP level. This could be explained by higher polymorphism of the wild compartment with many specific markers but also specific markers of the cultivars, as well as substantial differences in the common marker frequencies.

Among wild *A. xanthorrhiza*, AFLP and cpDNA divergence also appeared between the monocarpic and the perennial life forms, except in three monocarpic plants from the northern department of Cajamarca in Peru. This latter result could be explained by a more recent independent derivation of monocarpic populations from perennial ancestral populations (Knudsen, pers. com.), spontaneous gene flow between both wild forms, or an error in the labeling of the material. Overall, however, the genetic results agree with the morphological differences already reported to distinguish the two life forms (Knudsen 2003, Blas 2005); the perennials differ by the presence of compressed basal stems (cormels) with several shoots and

fewer umbels, whereas the monocarpic forms are characterized by an absence of thickened starchy stems and they have only one (rarely two) reproductive shoot per plant. Moreover, monocarpic plants are taller (to 3.5 m) and have a vigorously branching shoot carrying up to 250 umbels (Knudsen 2003).

Since arracacha was reported to have a narrow genetic base, it was unsuspected that AFLPs would structure the cultivars in two well separated groups. One of these groups was found to be closer to the wild perennial and the other one closer to the wild monocarpic form. This suggests the possibility of direct domestication of the wild monocarpic forms or cultivar diversification through introgression by the monocarpic life form. However, since both groups of cultivars share the same cytotype with the wild perennial, we favour the second hypothesis, which is also more parsimonious given the likely long original domestication process. Furthermore, the wild perennial is morphologically closer to the cultivated form and both can be vegetatively propagated, whereas the monocarpic form presents no evidence of vegetative reproduction (Hermann 1997; Knudsen 2003; Blas 2005). Combined evidence suggests that arracacha was domesticated from the wild perennial type. Hermann (1997) speculated that the ability of vegetative propagation (replanting the cormel) of the wild perennial forms could have been accidentally discovered, for instance during food preparation. Based on this scenario, we propose that the wild monocarpic populations are not ancestral to the domesticated forms but have likely contributed to the genetic diversity of this crop through introgression. This is an important feature from a genetic resource conservation and breeding standpoint. Cultivated *A. xanthorrhiza* has been suggested to have a weak genetic base, while our results showed that there are at least two main diversity groups. The detected diversity groups in the *A. xanthorrhiza* complex are thus also promising for arracacha breeding.

Hence, the domestication of arracacha from a perennial ancestor could be representative of a more general domestication pattern in the Andean region concerning various root and tuber crops and involving vegetative reproduction. This suggests a preference for a perennial type with asexual propagation potential, while a monocarpic life form was also available for domestication.

Acknowledgements

The first author thanks the IRD research team “*Anthropisation et Dynamique de la Diversité Génétique des Plantes (DYNADIV)*” for supporting this research; Steen Randers Knudsen (ex

KVL University, Denmark) for providing Peruvian samples for DNA analysis and his cooperation; Elicio Tapia for his invaluable help in the field collections, J. C. Pintaud for his advice on the cpDNA approach, and Y. Vigouroux for his comments on the manuscript.

References

- Blas, R. 2005. Diversity of Arracacia species in Peru. Ph.D. thesis. Gembloux Agricultural University, Belgium. 154p.
- Demesure, B . Sodzi N. and R.J. Petit. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*. 4, 129-131.
- DNASTAR software ver. 5.07. 2003. Inc., Madison, WI. <http://www.dnastar.com>
- Hahn, W. 2002. A phylogenetic analysis of the arecoid line of palms based on plastid DNA sequence data. *Molecular Phylogenetics and Evolution*. 23:189-204.
- Hawkes, J.G. 1989. The domestication of roots and tubers in the America tropics. Pp. 481-503. In. *Foraging and Farming*, Harris and Hillman eds., Unwin Hyman, London.
- Hermann, M. 1997. Arracacha. In: M. Hermann and J. Heller eds., *Andean roots and tubers: Ahipa, arracacha, maca and yacon. Promoting the conservation and use of underutilized and neglected crops*. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome, Italy. p. 75–172.
- Jhingan, A. 1992. A novel technology for DNA isolation. *Methods Molecular Cellular Biology*. 3:15-22.
- Knudsen, S.R. 2003. A revision of the South American species of the genus *Arracacia* Bancroft Apiaceae. Ph.D. thesis, Chapter I. Department of Ecology. KVL University, Denmark. Available in <http://www.ecol.kvl.dk/~srk/arracacia/side2.html>
- Mazón, N., R. Castillo, M. Hermann y Espinosa. 1996. La arracacha o zanahoria blanca *Arracacia xanthorrhiza* Bancroft en Ecuador. Publicación miscelánea N° 67, DENAREF. INIAP. Quito, Ecuador. 41p.
- Nei, M. 1972. Genetic distance between populations. *American Naturalist*, 106, 283-392.
- Peakall, R. and Smouse P.E. 2001. GenAlEx V5: Genetic Analysis in Excel. Australian National University, Canberra, Australia. <http://www.anu.edu.au/BoZo/GenAlEx/>.
- Roy R, Steffens D.L, Gratside B, Jang G.Y, Brumbaugh J.A. 1996. Producing STR locus patterns from bloodstains and other forensic samples using an infrared fluorescent automated DNA sequencer. *Journal of Forensic Sciences* 41:418-424.

- Swofford, D. 2001. PAUP software. Phylogenetic analysis using parsimony. Version 4b8. Sinaur, Sunderland, MA.
- Valderrama, M. and J. Seminario. 2002. Los parientes silvestres de la arracacha (*Arracacia xanthorrhiza* Bancroft) y su uso en medicina tradicional, en el norte peruano. *Arnaldoa*, Trujillo, Peru. 9(1):67-91.
- Vos P., Hogers R., Bleeker M., Reijans M., T van de Lee, Hornes M, Frijters A., Pot J., Peleman J., and Kuiper M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 11; 2321: 4407–4414

IV Partie: Diversité génétique chez l'arracacha cultivée

La connaissance de la variabilité génétique de l'arracacha cultivée était avant notre travail essentiellement restreinte à des analyses agro-morphologiques et à l'analyse du polymorphisme d'isoenzymes. Peu d'études sur la diversité moléculaire de *A. xanthorrhiza* ont été publiées. Elles se basent sur l'emploi des marqueurs RAPDs bien qu'un travail avec des AFLP ait été récemment réalisé (Blas 2005).

Les objectifs de cette partie sont d'évaluer la diversité génétique de l'arracacha cultivée par les marqueurs SSR et AFLP et de comprendre comment elle peut être structurée. Cette diversité moléculaire, *a priori* sélectivement neutre, a été par ailleurs comparée aux données morphologiques et géographiques dont on dispose.

RESULTATS

Nous avons donc étudié la diversité génétique de 72 accessions cultivées de *A. xanthorrhiza* collectées dans tout l'Equateur et au nord du Pérou avec 11 marqueurs SSR, puis un ensemble de 27 cultivars choisis sur la base de la diversité observée avec les SSR a été caractérisé plus largement par le polymorphisme révélée avec 7 couples d'amorces AFLPs.

Nos résultats avec les SSR montrent une variation génétique non négligeable au sein de la forme cultivée d'*A. xanthorrhiza* (Figure 20). Au niveau du polymorphisme, le locus AxD82 qui se montre comme le plus informatif dans le complexe *A. xanthorrhiza* est également le plus polymorphe chez la forme cultivée: parmi les 72 accessions analysées, 9 géotypes différents sont ainsi mis en évidence pour ce locus (Tableau 9). Pour le reste des locus, entre 3 et 6 géotypes différents sont distingués, à l'exception du locus AxD13 qui est monomorphe dans la forme cultivée bien que polymorphe dans le complexe *A. xanthorrhiza*.

L'arbre UPGMA (Figure 20) montre le regroupement des cultivars selon le polymorphisme SSR. Sur cette figure, l'origine géographique et leur classification morphologique selon Mazon (1993) parmi les 17 morphotypes définis au sein de la collection originale de l'INIAP, (Ex. M1 pour morphotype 1) sont aussi indiquées pour les accessions pour lesquelles on dispose de l'information. Sur les 72 accessions caractérisés, 33 représentent 13 des 17 morphotypes définis. Ces 33 accessions représentent la forme horticole blanche avec 7 morphotypes (20 accessions assignées à M1, M2, M3, M4, M5, M6 et M7), la forme horticole jaune avec 2

morphotypes (6 accessions assignées à M10 et M12) et la forme horticole violette avec 4 morphotypes (7 accessions assignées à M13, M15, M16 et M17).

On pourrait s'attendre à que les accessions classées dans un même morphotype présentent le même génotype multilocus. Cependant, les SSR montrent au contraire que des accessions appartenant à un même morphotype peuvent présenter des génotypes distincts. Prenons par exemple les 4 morphotypes les plus représentés (M6, M7, M10 et M15); les six accessions classées M6 présentent toutes des génotypes multilocus distincts, les cinq accessions M7 présentent 4 génotypes. Par contre, les 4 accessions M15 sont identiques et les cinq accessions M10 ne présentent que deux génotypes. On peut donc présumer que les accessions partageant les mêmes morphotype et génotype SSR multilocus correspondent probablement à des clones (seulement 8 cas sont observés, voir Fig. 20). Concernant la distribution géographique de ces clones, on constate qu'en général ils ont été collectés dans une même région géographique (par exemple dans les provinces du sud, Loja et Azuay, ou du Nord, Imbabura et Carchi); cependant d'autres montrent une distribution plus large. Ainsi par exemple, les clones du morphotype 10 ont été collectés au centre et au Nord de l'Equateur. Il est à ce sujet intéressant de noter le positionnement au niveau SSR des 13 cultivars péruviens analysés pour lesquels aucune donnée morphologique n'est disponible. Six d'entre eux se montrent affines entre eux et sont localisés à la base de l'arbre, alors que les 7 autres se regroupent avec des cultivars équatoriens. L'accession péruvienne AM (C60), que nous avons observée intermédiaire avec *A. x. var. andina*, se localise à une distance importante vis-à-vis du reste des cultivars analysés.

Par ailleurs la diversité SSR révélée chez la forme cultivée d'*A. xanthorrhiza* ne semble pas particulièrement structurée comme observée avec les AFLPs, bien que *a posteriori* on puisse remarquer une certaine corrélation entre les deux marqueurs (Fig. 21). Les AFLPs structurent la diversité des cultivars en deux groupes clairement différenciés (Fig. 21b) (cette différenciation est faible au niveau SSR, Fig. 21a), pour lesquels les hypothèses sur leur origine sont présentées dans la partie III.

DISCUSSION

Les deux types de marqueurs, SSR et AFLPs, ont été utiles pour la détection de polymorphisme au sein du type cultivé. Les résultats nous montrent que la base génétique de la forme cultivée est plus importante que ce qui avait été présumé (Mazon et al. 1996, Erazo et al. 1996; Castillo 1997). Ce résultat est en accord avec celui de Blas (2001) qui à l'aide de marqueurs RAPDs suggérait qu'une diversité cachée pourrait être présente chez la forme cultivée. Par ailleurs, aucune association entre cultivars et leur provenance ou morphologie n'a pu être établie, à l'exception des cultivars du nord du Pérou qui constituent un groupe plus différencié.

Chacun des marqueurs a fourni une information importante. Les SSRs ont révélé une diversité peu structurée mais recombinaison. Ainsi les 34 allèles observés sur les 11 locus analysés se trouvent recombinaison en 56 génotypes multilocus, ce qui indique 80% de génotypes originaux parmi les 72 accessions caractérisés. Ce résultat est significativement plus élevé que d'autres estimations basées sur des marqueurs RAPD ou isoenzymatiques (57 et 51%, Blas 2000, Mazon et al. 1996). Il peut être expliqué par i) la nature plus polymorphe du marqueur utilisé au contraire des isoenzymes par exemple qui se sont montrés très peu polymorphes (Mazon et al. 1996, Erazo et al. 1996), ii) une bonne représentativité de notre échantillonnage, bien que la majorité de accessions aient été collectée en Equateur et seulement une partie représente la zone nord du Pérou, et iii), l'existence d'un brassage génétique vraisemblablement lié à la reproduction sexuée.

On peut supposer que la dispersion ancienne du type cultivé dans la région andine, a été associée à la recombinaison de génotypes et que ce brassage a homogénéisé la diversité génétique, plus structurée géographiquement au niveau de l'ancêtre sauvage. La forme cultivée est donc constituée de génotypes recombinaison à laquelle correspond aussi une diversité morphologique rapportée au sein des différentes collections de germplasm. Dans la collection équatorienne par exemple, 17 morphotypes ont été définis parmi 102 accessions caractérisées (Mazon 1993). Notre analyse par SSR indique que l'homogénéité génétique de ces morphotypes est variable. De 5 morphotypes représentés par 3 à 6 accessions dans notre analyse, deux sont constitués de clones alors que les trois autres présentent plusieurs génotypes. Ce résultat est important en matière de conservation des ressources génétiques, car il montre que la caractérisation morphologique sous-estime la diversité et qu'il est important de combiner différentes approches pour la détermination des véritables doublons au sein des collections.

De leur côté, les AFLPs, avec un plus grand nombre de marqueurs, ont été utiles pour la structuration de la diversité en révélant deux groupes de cultivars comme représentés par la PCO de la Fig. 21b. Dans l'état actuel, nous sommes cependant incapable de déterminer si des différences au niveau morphologique ou phénologique entre ces deux groupes existent; la forte plasticité phénotypique observée chez la forme domestiquée (Hermann 1997) pourrait dissimuler cette structure génétique.

La reproduction sexuée joue t-elle un rôle actuel dans l'aire d'origine ?

L'hypothèse d'un brassage génétique chez la forme cultivée de *A. xanthorrhiza* implique l'utilisation de la reproduction sexuée dans une plante à propagation végétative. L'arracacha, espèce préférentiellement allogame, comme un grand nombre (sinon toutes) les espèces cultivées à propagation végétative, a conservé la capacité de se reproduire par voie sexuée bien que sa culture soit basée sur la propagation asexuée. Concernant l'utilisation de la reproduction sexuée, il y a peu d'observations chez l'arracacha; cependant les échappés de culture en jachère fleurissent bien et produisent des graines en bonne quantité (environ entre 300 à 500 graines par plante, Knudsen 2003, Obs. pers.). La question qui s'impose est donc de savoir si l'utilisation de la reproduction sexuée dans les systèmes traditionnels est restée active grâce à la sexualité résiduelle de certains génotypes ou s'il s'agit de pratiques anciennes ayant pu perdre leur efficacité actuellement.

L'incidence actuelle de la sexualité dans la région andine reste à préciser. Comme déjà mentionné, des observations chez des indiens Sibundoy du sud de la Colombie font référence à l'utilisation sporadique des plantules issues de graines (Vasquez et al. 2003, Lobo M. CORPOICA, com. pers.). Au Brésil par ailleurs, région où l'arracacha est introduite depuis un peu plus d'un siècle, l'utilisation de la sexualité de la part des améliorateurs a permis l'obtention de plusieurs variétés à partir d'un faible nombre de clones.

5. DISCUSSION GENERALE DU CHAPITRE I

Choix des marqueurs moléculaires

Le choix des marqueurs génétiques pour les études de diversité est souvent conduit en fonction des questions posées et des possibilités d'utilisation que compte le manipulateur. Du point de vue technique, les SSR sont plus coûteux à mettre en oeuvre si la démarche est, comme dans notre étude, développée entièrement, c'est-à-dire avec la création d'une banque de clones SSR et l'identification de marqueurs. Nos résultats montrent l'importance d'employer différentes approches méthodologiques complémentaires, SSR et AFLPs dans notre cas. En effet, ces deux types de marqueurs sont puissants pour révéler une diversité génétique pas négligeable au sein du complexe *A. xanthorrhiza* et particulièrement chez l'arracacha cultivée. En revanche, les marqueurs SSR sont moins informatifs pour révéler plus précisément la structuration de cette diversité. Ceci est certainement dû au nombre de marqueurs considérés car pour les AFLPs près de 200 marqueurs ont été considérés alors que pour les SSR, le polymorphisme à 11 locus a été analysé. En accord avec la théorie, la résolution de la diversité devient plus fine lorsqu'on augmente le nombre de paramètres pour comparer génétiquement deux individus. Par contre la nature génétique codominante des SSR a permis une comparaison quantitative de la diversité sauvage et cultivée à l'aide des estimations de richesse allélique et d'hétérozygotie qui n'est pas possible avec les AFLPs.

Structure génétique du compartiment sauvage de *A. xanthorrhiza*

Chez les formes sauvages de *A. xanthorrhiza*, la diversité révélée par les marqueurs génétiques apparaît distribuée en un continuum où les deux types sauvages polycarpique et monocarpique sont reconnus aux extrêmes de cette distribution. Les AFLP se sont montrés cependant plus résolutifs que les SSR pour la distinction entre ces deux types sauvages (ce qui est reflété par les valeurs de différenciation obtenues avec chaque type de marqueur ($F_{st} = 0.240$ et 0.125 pour les AFLPs et SSR respectivement)).

Du point de vue botanique, les deux variétés sauvages de *A. xanthorrhiza* se ressemblent mais restent différenciables notamment à la floraison (Knudsen 2003, Obs. pers). Selon notre expérience cependant, certaines populations sont d'assignation difficile en utilisant les caractères botaniques diagnostiques proposés par Knudsen (2003), alors qu'Hermann (com. pers.) mentionne que le mode de vie polycarpique ou monocarpique n'a pas d'ambiguïté possible. Selon nos résultats, pour certaines collections, l'identification botanique n'est pas

toujours cohérente avec la distinction génétique. On peut expliquer ce résultat par deux hypothèses:

i) une plasticité morphologique des populations sauvages comme observée chez la forme cultivée (Hermann 1997).

ii) l'existence de flux génétique entre formes sauvages qui aurait donné lieu à une diversité des formes intermédiaires spontanées. Concernant cette hypothèse, des croisements entre les types sauvages monocarpique et polycarpique ont été obtenus facilement au niveau expérimental, et les graines hybrides semblent avoir une bonne viabilité (Knudsen, com. pers.). Le même auteur (2003) signale cependant n'avoir pas retrouvé de populations mixtes dans le Nord du Pérou bien que dans certains cas les deux types sauvages soient en proximité spatial ce qui est aussi le cas dans le Sud de l'Equateur (obs. pers).

Les marqueurs SSR indiquent par ailleurs une diversité plus importante au sein du type polycarpique, cette diversité semblant associée à l'origine géographique des populations (Fig. 17b). On distingue ainsi les types pérennes du Centre de l'Equateur, du Sud de l'Equateur (Loja), et du Nord du Pérou (Cajamarca). Chacun de ces trois types présentent certaines particularités au niveau morphologique (Hermann 1997), raison pour laquelle les taxonomistes les ont autrefois catalogué comme des espèces différentes (*A. andina*, *A. equatorialis* et *A. peruviana*). Knudsen (2003) considère que ces trois types représentent la variation morphologique du type polycarpique de *A. xanthorrhiza* ce qui semble être confirmé par nos résultats génétiques.

Impact de la domestication chez *A. xanthorrhiza*

Chez les plantes domestiquées, il est courant de retrouver une diversité génétique réduite par rapport à celle des formes sauvages ancestrales dans l'aire d'origine, en relation avec le goulot d'étranglement provoqué par la domestication. Nous observons qu'environ un tiers de la diversité SSR révélée dans le complexe *A. xanthorrhiza* est présente chez la forme domestiquée bien que l'on détecte une portion de diversité au niveau de 3 locus SSR qui est particulière au type cultivé. Ce résultat est cohérent avec ce qui a été rapporté chez d'autres plantes à propagation végétative, comme le manioc par exemple où 22% des allèles SSR de l'espèce ancestrale postulée ont été observés dans la forme cultivée (Olsen et Schall 1999).

Par ailleurs, la diversité SSR suggère un scénario de domestication à partir des populations de Loja et Cajamarca. Notre échantillonnage a été cependant biaisé en faveur des accessions équatoriennes puisqu'un faible nombre de cultivars péruviens ont été analysés. L'analyse des cultivars péruviens d'autres régions géographiques ainsi que d'accessions d'autres pays andins (Bolivie et Colombie) permettrait de déterminer si une nouvelle diversité peut être mise en évidence ou si ces cultivars proviennent du centre de domestication que nous avons identifié. Comme déjà mentionné, on peut supposer qu'une dispersion ancienne par la dissémination des autochtones américains au cours de leurs migrations soit associée au fait que la forme cultivée se retrouve dans des aires plus éloignées du centre de domestication comme la côte Caraïbe du Venezuela.

Interactions entre formes sauvages et cultivées de *A. xanthorrhiza*: les SSR suggèrent un flux génétique faible dans l'aire de domestication présumée

La divergence génétique entre formes sauvages et cultivées suggère une faible interaction entre les deux compartiments dans l'aire de domestication. Nous avons cependant détecté un petit nombre de formes intermédiaires parmi les échantillons analysés. Le test d'assignation ainsi que la comparaison avec des hybrides artificiels entre *A. x. var. andina* et le type cultivé, ont permis de confirmer qu'il s'agit d'individus hybrides spontanés entre ces deux types *A. xanthorrhiza*. Nous attribuons ce résultat à un flux génétique spontané. Au niveau expérimental, des croisements entre les différents types de *A. xanthorrhiza* ont été réussis avec une certaine facilité (Vasconez com. pers, Knudsen 2003). Cependant, compte tenu du faible nombre d'individus intermédiaires spontanés détectés, nous estimons que l'incidence de ce flux en conditions naturelles est limitée.

Nous argumentons aussi ceci sur la base des observations suivantes:

- un faible contact spatial entre types sauvages et cultivés. En effet, les deux types sont retrouvés dans des écosystèmes relativement différents, bien que des associations notamment dans le nord du Pérou aient été rapportées (Valderrama et Seminario 2002, Knudsen 2003, Blas 2005). C'est en effet dans cette région où nous détectons des formes intermédiaires.
- la forme cultivée fleurit mal, de plus ce caractère est contre sélectionné par les paysans au profit du rendement (Herman 1997, Knudsen 2003, Obs. pers). D'autre part la synchronisation de la floraison entre formes sauvages et cultivées n'est pas assurée.

- une faible présence actuelle de pollinisateurs pourrait selon Knudsen (2003) être aussi un facteur limitant dans les zones andines.

Par ailleurs, les individus analysés avérés intermédiaires sauvage-cultivé par les SSR, ont été tous à l'exception du cultivar AM, collectés à l'état sauvage. Le cultivar AM a été collecté à Cajamarca au nord du Pérou, sans autre information. Par contre, on sait qu'il a été utilisé comme parent d'un des croisements analysés pour sa meilleure aptitude au croisement avec *A. x. var. andina* (Knudsen, com.pers).

D'autre part, on peut présumer qu'en fonction de la forte ressemblance entre les types pérennes cultivés et sauvages (Fig. 8), des formes intermédiaires spontanées sont lors des collections cataloguées comme arracacha cultivée si elles se trouvent à l'état cultivé (le cas de l'accession AM), ou comme *A. x. var. andina* si elles se trouvent en milieu sauvage, bien qu'il puisse s'agir de plantes échappées de culture. Cependant, parmi tous les individus tubéreux collectés à l'état sauvage et analysés, aucune plante ne présente un génotype cultivé. Ce résultat peut contribuer à soutenir l'hypothèse d'Hermann (1997) selon laquelle des échappés de culture auraient du mal à coloniser des populations en milieu sauvage car elles ne seraient pas compétitives dans ce milieu, et de plus elles présentent une mauvaise floraison. Par contre, si des échappés de culture se retrouvent en contact avec des populations sauvages, des graines hybrides auront plus de chance à survivre. En effet, les résultats expérimentaux montrent que la viabilité de ces graines est plutôt favorisée ce qui a été attribué à la vigueur hybride (Knudsen 2003). De plus, des observations en serre des hybrides expérimentaux montrent que ces hybrides fleurissent bien et semblent aussi avoir une bonne fertilité (Knudsen, com. pers). On peut donc présumer que les hybrides spontanés pourront introgresser les populations sauvages comme cela est suggéré par nos résultats.

Scénario de domestication de *A. xanthorrhiza*

Avec l'information fournie par les SSR, AFLPs et ADN chloroplastique, un scénario illustré de la domestication chez *A. xanthorrhiza* peut être présenté (Figure 22). Les formes sauvages tubéreuses (*A. x. var. andina*, *A. x. var. monocarpa* et *A. incisa*) ont été observées proches au niveau génétique de l'arracacha cultivée (tant au niveau AFLP que SSR). Cependant les AFLPs et le génotypage de l'ADN chloroplastique permettent d'identifier *A. x. var. andina* comme la forme sauvage la plus proche du type cultivé. Avec l'information fournie par les SSR, la domestication est suggérée avoir eu lieu à partir des populations pérennes de la zone

localisée entre Loja et/ou Cajamarca où le type monocarpique est aussi retrouvé. En effet, le type sauvage pérenne présente une distribution géographique plus importante que les autres types sauvages tubéreux. Des populations sauvages pérennes de trois régions géographiques principales ont été représentées: Centre Equateur (provinces de Chimborazo, Cañar et Bolivar), Sud Equateur (province de Loja) et Nord Pérou (Département de Cajamarca). Les résultats SSR suggèrent que la forme domestiquée a été domestiquée à partir des populations de la région Sud Equateur et Nord Pérou (Cajamarca) et non à partir des populations du Centre Equateur comme proposé par Mathias et Constance (1976) et Hermann (1997) sur la base d'observations morphologiques. En effet, les 31 allèles observés sur les 11 locus analysés chez les formes cultivées sont retrouvés parmi les populations de Loja et Cajamarca, alors que dans les populations du centre Equateur, 10 des 31 allèles cultivés n'ont pas été observés. Chez les populations du centre Equateur, on retrouve par ailleurs une diversité moins importante par rapport à celle du Sud (Loja et Cajamarca en Pérou), ainsi qu'une moins bonne amplification SSR. Au niveau morphologique, les populations de Loja et Cajamarca ont été par contre décrites comme étant moins affines du type cultivé; par exemple leurs racines sont moins volumineuses et possèdent une astringence majeure, bien que leur écorce soit fine comme chez la forme cultivée (Hermann 1997). Par contre selon le même auteur, les populations du Centre Equateur sont les plus ressemblantes à la forme cultivée, ce qui l'avait amené à les proposer comme étant à l'origine de la forme cultivée.

L'accessibilité à la propagation asexuée du type sauvage pérenne a pu accélérer le processus de domestication, dans une région où la reproduction végétative des plantes domestiquées est un patron dominant pour les espèces à racines et tubercules. Les AFLPs révèlent l'existence de deux groupes de cultivars, l'un plus proche du type pérenne et l'autre plus proche du type monocarpique. L'hypothèse favorisée par les résultats au niveau de l'ADN chloroplastique est que ce dernier résulte de l'introgession du type monocarpique. Cet événement qui a pu avoir lieu au cours de la domestication a diversifié le groupe de cultivars en deux types bien différenciés au niveau moléculaire mais restés néanmoins cryptiques jusqu'à notre analyse.

Par ailleurs, tant au niveau AFLP que SSR, la forme cultivée montre de plus une fraction de diversité qui n'est pas retrouvée dans le compartiment sauvage du complexe *A. xanthorrhiza*, et dont l'origine est inconnue. L'hypothèse que cette diversité résulterait de la participation d'une autre espèce sauvage à la domestication est considérée dans le deuxième chapitre de ce travail.

CHAPITRE II.

HYPOTHESE DE L'INTROGRESSION GENETIQUE A PARTIR DE FORMES SAUVAGES ELOIGNEES:

Des espèces éloignées sont-elles représentées dans le génome de la forme cultivée?

**APPROCHE EXPERIMENTALE CHEZ LE
MANIOC ET L'ARRACACHA**

1. ANTECEDENTS ET OBJECTIFS

L'objectif de ce deuxième chapitre est de déterminer si une portion de la diversité génétique observée qui est propre à la forme cultivée chez l'arracacha et le manioc, par rapport à leurs ancêtres sauvages directs présumés, trouve sa source dans l'introggression à partir d'espèces relativement très divergentes (même genre ou genres voisins).

Il s'agit de valider l'hypothèse émise sur la base de la diversité AFLP et SSR :

- i) Chez l'arracacha car nos analyses moléculaires montrent une diversité originale de la forme cultivée dont l'origine est inconnue,
- ii) Chez le manioc car c'est chez cette plante que l'hypothèse a d'abord été postulée en fonction de plusieurs évidences. Par ailleurs, on dispose d'une collection biologique de travail utilisée lors de notre étude de DEA alors que les résultats acquis méritent d'être amplifiés pour valider cette hypothèse.

D'autre part, ces deux plantes (arracacha et manioc) ont quelques caractéristiques en commun:

- elles appartiennent au même centre de domestication Sud américain bien que le manioc soit une plante tropicale et l'arracacha une plante sub-tropicale
- elles sont des plantes à racines propagées végétativement,
- la délimitation d'espèces au sein du genre botanique et les genres eux-mêmes ont été sujets à controverses. Dans le cas du manioc, le genre botanique a été proposé de radiation récente
- une espèce ancestrale majeure a été identifiée mais la forme cultivée montre une fraction de diversité non expliquée par celle-ci.
- la forme cultivée partage une large zone de distribution avec plusieurs espèces sauvages éloignées et des conditions de sympatrie se rencontrent.

On note aussi des différences entre les deux plantes:

- La culture du manioc est beaucoup plus répandue que celle de l'arracacha et mieux étudiée expérimentalement.
- Le manioc domestiqué présente une grande diversité génétique dans laquelle on distingue deux pôles, « variétés amères » et « variétés douces », mais un continuum de variation. Chez l'arracacha cultivée, la diversité semble plus modeste mais on retrouve cependant une structure génétique au niveau AFLP avec deux groupes bien définis.
- Le manioc a conservé une reproduction sexuée dans son aire d'origine plus importante que l'arracacha (voir néanmoins plus loin); l'utilisation de plantules issues de la reproduction sexuée entretenant un processus de gestion dynamique a été démontrée (Empéaire et al. 1998, Elias 2000).
- Chez le manioc, aucune barrière génétique forte à l'hybridation n'a été mentionnée (Nassar 2003).
- Chez le manioc, des hybridations spontanées avec des espèces éloignées sont documentées et les racines de certains de ces hybrides sont même consommées (Second 2000).

INTRODUCTION

Selon Anderson (1961), l'hybridation introgressive a joué un rôle important dans l'origine des plantes cultivées. En particulier, l'introgession dans les aires perturbées par l'homme serait fréquemment un mécanisme à l'origine de la génération de génotypes recombinants d'intérêt représentant une diversité nouvelle soumise à sélection. Cet auteur montre la connexion entre introgression et perturbation des habitats naturels, et affirme que la variabilité existante chez les formes cultivées ressemble à la variation observée chez les populations naturelles introgressées. Selon lui, dans les centres d'origine des plantes cultivées, la possibilité d'introgession entre formes domestiquées ou en cours de domestication avec des formes sauvages est énorme et ne doit pas être négligée.

Sur cette base, l'objectif principal de ce chapitre est de confirmer si le schéma proposé par Anderson s'applique aux cas de l'arracacha et du manioc.

1.1 Le cas de l'arracacha: la forme ancestrale présumée est-elle la seule représentée dans la forme cultivée ?

Nos résultats confirment que la forme sauvage polycarpique (*A. x. var. andina*), qui ressemble le plus morphologiquement à la forme cultivée (Hermann 1997, Knudsen 2003), est aussi au niveau AFLP et microsatellites la plus proche de la forme cultivée et donc sa forme ancestrale présumée.

Cependant, les mêmes marqueurs moléculaires montrent que la forme cultivée présente une fraction de diversité qui lui est privée. Un criblage large du génome avec des AFLPs met en évidence une divergence au sein de la forme cultivée; ce résultat est attribué à la participation du type sauvage monocarpique à la constitution du génome d'un groupe de variétés. Le type monocarpique est aussi une forme proche de la forme cultivée mais qui n'aurait pas été directement domestiquée, ce résultat restant néanmoins à confirmer sur un plus grand nombre de marqueurs. Cette introgression n'explique cependant pas une fraction de diversité qui reste spécifique au type cultivé dans nos analyses. Au niveau des marqueurs microsatellites, sur 14 locus, 11 présentent une diversité compatible avec l'hypothèse que les formes sauvages et cultivées sont con-spécifiques, mais trois montrent des profils qui différencient bien les deux compartiments. Cette observation et celle d'une moins bonne amplification des mêmes locus

SSR chez la forme sauvage convergent pour indiquer que les séquences flanquantes pourraient être différentes au niveau de ces locus entre la forme cultivée et son homologue sauvage.

L'hypothèse proposée pour expliquer ces observations est que d'autres formes (espèces, genres?) sauvages plus divergentes pourraient être représentées dans le génome de la forme cultivée. A ce sujet, plusieurs espèces du genre *Arracacia* partagent l'aire de distribution de la forme cultivée, et il se peut que des situations de sympatrie se mettent en place. Bristol (1988) notamment, rapporte que l'arracacha cultivée est souvent en contact avec l'espèce *A. acuminata* dans le sud-est de la Colombie. Il mentionne aussi que les indiens Muzo habitant cette zone, ont la tradition d'utiliser la graine pour la reproduction des cultivars ce qui expliquerait selon lui, l'énorme diversité de cultivars natifs de cette même région qui serait un centre de diversification de la forme cultivée. La question suivante émerge: dans de telles situations, des introgressions entre autres espèces sauvages éloignées et cultivars ont elles eu lieu?

1.2 Le cas du manioc: un ancêtre sauvage direct est confirmé mais d'autres espèces sont sans doute représentées dans le génome de la forme cultivée

Le genre *Manihot* comprend près d'une centaine d'espèces toutes monoïques et allogames. Alors que les limites du genre semblent définies, la délimitation de ses espèces est sujette à controverse. Sa distribution est limitée au continent américain, du Sud des Etats-Unis au Nord de l'Argentine. L'espèce cultivée est fréquemment rencontrée en sympatrie avec des espèces sauvages (Rogers et Appan 1973). Cette association se retrouve dans le cas d'un champ cultivé dans l'aire de répartition de la forme sauvage ou, comme il arrive souvent, dans les cas où ces espèces sont «réfugiées» le long des routes et chemins, eux-mêmes souvent en bordure de champs cultivés en manioc. De plus, parmi le manioc cultivé (appelé *manioc* par la suite), l'existence de formes hybrides interspécifiques spontanées est avérée, en particulier par l'existence de formes «arbres manioc» récoltées après plusieurs années. Les barrières reproductives interspécifiques semblent peu prononcées dans le genre entier, et aucune barrière génétique forte à l'hybridation n'a été mentionnée (Nassar 2003).

Jusqu'à récemment, l'origine du manioc était inconnue. Alors que Rogers et Appan (1973) concluaient à une origine hybride du manioc, Allem (1994) a proposé le concept d'une espèce ancestrale directe qu'il a nommée *M. esculenta* ssp. *flabellifolia*, syn. *M. esculenta* ssp.

peruviana (appelée *M. flabellifolia* par la suite). Cette forme avait été considérée rudérale par Rogers et Appan (1973). D'autres données moléculaires au niveau de l'ADN chloroplastique et des gènes ribosomiques nucléaires (Fregene et al. 1994) confirment aussi l'hypothèse de Allem (1994) sans toutefois écarter la possibilité d'hybridations avec d'autres espèces sauvages. Ainsi, d'autres études se basant sur de nombreux marqueurs arbitraires et ayant inclus une large représentation tant de l'espèce cultivée que du genre *Manihot* (bien que principalement dans ses espèces sud-américaines), tout en confirmant l'hypothèse de Allem, obtiennent des indices de contribution à la diversité génétique du manioc de certaines autres espèces sauvages, en particulier *M. glaziovii* (Colombo 1997, Second et al. 1997, Second et Iglesias 2001). Cette espèce, originaire de l'Est du Brésil est une des espèces du genre *Manihot* les plus éloignées du manioc; cependant, ses hybrides avec le manioc se rencontrent fréquemment dans le Nord-Est brésilien où ils sont connus sous le nom de « manioc de sept ans » ou « arbres manioc » et certains sont récoltés pour leurs racines (Second et al. 1997). *M. glaziovii* apparaît génétiquement très proche de *M. carthaginensis* (Second et al. 1997) qui est aussi connu pour former des hybrides avec le manioc (Rogers et Appan 1973). En accord avec l'analyse morphologique, l'ensemble de ces formes (*M. carthaginensis*, *M. glaziovii* et les arbres manioc) représentent le complexe d'espèces *M. carthaginensis* s. l. (Second 2000). *M. glaziovii* a été aussi introduit en Afrique il y a une centaine d'années et l'existence d'hybrides spontanés fertiles en Afrique a été confirmée (Lefèvre 1989). Une situation similaire à cette situation en Afrique de l'Ouest se retrouve donc dans l'aire d'origine du manioc au Nord-Est brésilien. Il est alors possible que *M. glaziovii* comme d'autres espèces du genre *Manihot* ait pu contribuer à la diversité génétique du manioc par introgressions génétiques (rétrocroisements successifs de l'hybride inter-spécifique sur l'espèce cultivée). L'incidence de telles introgressions sur la diversité du manioc n'a pas encore été éclaircie. Olsen et Schaal (1999) rejettent l'hypothèse d'introgression tandis que Second et al. (1997, 1999) penchent pour l'hypothèse d'une introgression non négligeable d'espèces sauvages du genre *Manihot* parmi les formes cultivées.

De nouvelles évidences d'introgression, basées sur la présence de bandes AFLPs ou allèles SSR partagés entre le manioc et des espèces sauvages mais absentes chez l'ancêtre présumé, sont rapportées dans notre étude de DEA (Morillo 2001). Nous avons génotypé sur la base de 27 locus SSR, 60 échantillons incluant le manioc et diverses espèces sauvages. L'analyse multivariée de 269 allèles observés confirme l'existence d'un groupe d'accessions de manioc affines de *M. glaziovii*, ce qui est aussi observé en utilisant des marqueurs AFLP. Nos

résultats suggèrent en particulier l'introgession parmi des variétés traditionnelles sud américaines à partir de *M. glaziovii* d'un fragment chromosomique du groupe de liaison D. Cette zone chromosomique a aussi été indépendamment introgressée de manière contrôlée en Afrique pour la résistance à la mosaïque virale africaine qu'il apporte (la résistance croisée à la bactériose, répandue en Amérique, et à la virose africaine peut expliquer ce résultat). L'un des parents de la population utilisée pour déterminer la carte génétique du manioc présente ce fragment artificiellement introgressé (Jorge et al. 2000). Des allèles SSR suspectés d'introgession observés au niveau de ce fragment chromosomique justifient son intérêt dans notre contexte.

La présence d'allèles SSR sauvages dans des variétés traditionnelles de manioc peut être expliquée par deux hypothèses:

- i) mutations donnant des allèles SSR différents mais homoplasiques de ceux observés dans les populations sauvages;
- ii) allèles spécifiques de formes ou espèces sauvages introgressés dans des variétés cultivées suite à des hybridations.

Nous détaillons dans la suite l'approche méthodologique utilisée pour trancher entre ces deux hypothèses.

2. MATERIELS ET METHODES

2.1 Matériel végétal et locus microsatellites à analyser

A) Chez l'arracacha

Matériel végétal: Nous avons utilisé neuf accessions représentant les deux groupes de cultivars révélés par les AFLPs, puis cinq représentant *A. x. var. andina*, trois *A. x. var. monocarpa* et cinq représentant d'autres espèces du genre (*A. incisa*, *A. acuminata*, *A. elata* et *A. moschata*).

Locus SSR: Nous avons analysé trois locus présentant un profil d'amplification qui distingue les formes cultivés et sauvages (Fig. 14). Ces locus sont les suivants:

AxC87: L'amplification de ce locus résulte en un profil à trois ou quatre bandes chez les formes cultivées alors que chez les formes sauvages on note une absence d'amplification (chez les espèces éloignées) ou une ou deux bandes (chez l'ancêtre *A. xanthorrhiza* sauvage, à l'exception des individus « sauvages » ayant un génotype intermédiaire sauvage-cultivé, voir chapitre I, partie II). Pour ce type de locus, on parlera d'allèles putatifs car on ne sait pas si les bandes obtenues sont homologues du même locus. Au sein du complexe *A. xanthorrhiza*, dix allèles putatifs variant entre 106 et 140 pb sont observés. Les formes sauvages présentent seulement des bandes à 110, 112 et 114pb (présentes aussi chez certaines accessions cultivées) alors que chez les formes cultivées des bandes de taille plus élevée sont en général observées (122, 136, 140 et 142pb).

AxC64: une situation similaire à celle du locus précédent est observée. Alors que jusqu'à 4 allèles putatifs sont observés chez les formes cultivées, une absence d'amplification ou au maximum deux bandes sont observées chez les formes sauvages. Par ailleurs, les bandes privées des formes cultivées sont décalées de 20pb des bandes observées chez les formes sauvages.

AxC29: la désignation d'allèles en gel fut difficile car des bandes multiples apparaissent, similaires aux cas interprétés comme résultant d'un glissement (*slippage*) de la Taq polymérase, ce qui pourrait être dû à la nature du motif concerné (haut nombre de répétitions comme confirmé par le séquençage de ce clone montré dans l'annexe 3). Un décalage est cependant observé entre les formes cultivées et sauvages.

B) Chez le manioc

Matériel végétal: La forme cultivée a été représentée par 18 accessions dont six présumées introgressées. Cette liste comprend des cultivars traditionnels de la collection du CIAT (Centro International de Agronomia Tropical) et de l'IAC (Institut Agronomique de Campinas, Brésil) et aussi des « cultivars à grosses tiges » observés au Venezuela et échantillonnés par feuilles. De plus, sont également inclus deux cultivars artificiellement introgressés de *M. glaziovii* (qu'on appellera AI) ce qui leur a conféré la résistance à la Mosaïque virale africaine (accessions 5.59 et 5.70). La forme ancestrale *M. flabellifolia* est représentée par 5 accessions. Le complexe *M. carthagenensis* est représenté par des arbres manioc (5), *M. glaziovii* (4 accessions, deux de chaque groupe révélé par les AFLPs, Morillo 2001), et *M. carthagenensis* (2). Deux autres espèces sont aussi représentées: *M.*

quinquepartita (2) et *M. brachyloba* (2). Des détails sur l'origine de ces accessions sont disponibles dans Morillo (2001).

Locus SSR: Parmi différents locus présumés présenter des allèles introgressés dans des variétés traditionnelles de manioc cultivé (Morillo 2001), trois locus ont été étudiés. Ils ont été choisis sur la base de leur localisation sur la carte génétique du manioc et appartiennent au groupe de liaison D qui présente, chez un des parents utilisés pour établir la carte génétique, un fragment chromosomique présumé introgressé artificiellement de *M. glaziovii* (Jorge et al. 2000).

Locus SSRY-3: 11 allèles ont été observés en génotypage, dont quatre sont présents chez la forme cultivée: 204, 216, 226 et 233 pb (Morillo 2001, voir tableau 13 pour une comparaison entre taille d'allèles SSR en gel et *in silico* après séquençage). Les deux premiers (204 et 216) sont suspectés introgressés à cause des observations suivantes: i) ils sont présents à faible fréquence chez la forme cultivée, le premier étant uniquement présent chez les variétés introgressées par voie artificielle (5.59, 5.70) et le deuxième chez une seule variété traditionnelle (5.46); ii) ils sont absents chez *M. flabellifolia* (dans notre échantillonnage); iii) ils sont observés parmi *M. carthaginensis* et, iv) leur taille allélique est nettement différente par rapport à celle des allèles observés chez *M. flabellifolia* (226 et 228 pb).

Locus SSRY-9: De huit allèles observés à ce locus en génotypage, six sont présents chez la forme cultivée : 244, 256, 257, 259, 262 et 270. Les quatre derniers sont aussi présents chez *M. flabellifolia*. L'allèle 256, proche en taille de ceux observés chez *M. flabellifolia*, n'est pas présumé introgressé. L'allèle 244 est suspecté introgressé à cause de son absence chez *M. flabellifolia* et de sa taille allélique nettement distincte. Par ailleurs cet allèle est présent chez des accessions cultivées comme SRT1316 (5.64) et Col22 (5.62) présumées introgressées selon des indices suggérés par d'autres marqueurs: AFLPs (Second et al. 1999) et RAPDs (Colombo et al. 1998). De plus, la variété Col22 est une variété réputée souvent distincte au niveau moléculaire dans l'expérience du CIAT (J. Tohme, comm. verbale à G. Second). La présence de cet allèle parmi le complexe *M. carthaginensis* est restreinte à un arbre manioc (5.31).

Locus SSRY-40: Parmi dix allèles observés en génotypage pour ce locus, six sont présents chez la forme cultivée: 218, 221, 226, 229, 231, et 236pb. Ces allèles sont communs à *M. flabellifolia* et *M. glaziovii* et aucun allèle n'a été particulièrement suspecté d'introgression.

2.2 Approche méthodologique pour tester l'hypothèse d'introgression

L'approche choisie est le séquençage comparé des régions flanquantes du motif SSR entre les différentes espèces. Selon le degré de divergence du locus en question, le positionnement phylogénétique des séquences flanquantes entre les différents allèles peut apporter la preuve de l'introgression. Il faut néanmoins pouvoir écarter la possibilité de polymorphisme ancestral de l'ancêtre direct assumé, *M. flabellifolia*. La concordance de divergence importante à plusieurs locus (en comparaison avec les données de nombreux marqueurs tels que RAPD ou AFLP) ou à plusieurs locus génétiquement liés (le cas des locus SSR utilisés) n'est pas en accord avec l'hypothèse du polymorphisme ancestral.

Un premier essai de comparaison de séquences flanquantes fut effectué en dessinant des amorces aux extrémités distales de la séquence des clones SSR, alors que les amorces de génotypage sont choisies pour amplifier un fragment de moindre longueur. La séquence presque complète des clones SSR fut ainsi comparée sur 4 locus et des résultats encourageants furent obtenus (Morillo et al. 2001). Cependant cette approche présente une limitation liée au faible polymorphisme observé (Tableau 10). Nous avons donc entrepris la comparaison de séquences plus informatives en allongeant au moins une des zones flanquante des clones SSR.

2.2.1 Allongement de la séquence des clones SSR

Plusieurs méthodes basées sur la PCR sont publiées pour l'allongement de séquences génomiques flanquant une séquence d'ADN connue: PCR inverse, TAIL PCR et *PCR-walking* (Balzergue et al. 2001). Dans un premier temps nous avons essayé la PCR inverse (sans bons résultats), puis la marche chromosomique (*PCR-Walking*) a été adoptée.

Marche sur le chromosome

Le principe en est expliqué dans la figure 23. Cette technique fut appliquée selon le protocole du kit *Universal GenomeWalker*™ (CLONTECH, No. 638904). La méthode consiste en trois étapes principales: i) une restriction de l'ADN génomique par des digestions séparées

avec au moins 4 enzymes (dans notre cas, six enzymes ont été utilisées), ii) la ligation d'un adaptateur asymétrique pour les produits de digestion, et iii) une PCR emboîtée (« *Nested PCR* ») qui a pour objectif d'améliorer la spécificité du produit PCR obtenu en réamplifiant le fragment par des amorces internes aux premières employées. Les fragments candidats obtenus avec chaque enzyme (entre 1 et 3 Kb, dans notre pratique) sont clonés et séquencés afin de valider l'allongement du locus en question.

2.2.2 Validation de l'allongement

Une première vérification de l'allongement se fait par alignement de la séquence du fragment candidat avec la séquence connue; une centaine de bases permettent de confirmer une homologie entre les deux séquences. Sur la base de la séquence complète on peut alors procéder au dessin de nouvelles amorces. La validation définitive se fait par l'emploi de ces nouvelles amorces pour amplifier l'ADN génomique et le séquençage de l'ensemble du locus SSR, y compris la partie allongée des différents individus ; une absence d'amplification peut être observée notamment chez les espèces éloignées (Fig. 24). Elle est interprétée par un défaut d'hybridation de l'amorce due à la divergence de la séquence et d'autres combinaisons d'amorces sont alors essayées.

2.3 Isolement et obtention des séquences des allèles SSR allongés

Séquençage direct

Une première tentative fut de séquencer directement les bandes amplifiées en les séparant (y compris chez les individus hétérozygotes) par prélèvement après électrophorèse sur un gel d'agarose de haute résolution (Methaphor). Les chromatogrammes de séquence obtenus n'étaient pas de bonne qualité probablement à cause: i) de la faible quantité d'ADN amplifié (qui diminue en fonction de la taille du produit PCR), et/ou, ii) de la présence résiduelle de dimères dans le produit d'amplification (malgré une purification tentée). Nous avons donc opté pour le clonage des produits d'amplification.

Clonage, criblage et séquençage des allèles SSR allongés

Le clonage fut effectué à l'aide d'un kit commercial (pGEM-T, Promega A-1380). Le nombre d'allèles par individu dépend du génotype (homozygote ou hétérozygote) et du locus concerné.

Pour deux locus analysés chez l'arracacha (AxC87 et AxC64), jusqu'à 4 bandes ou allèles putatifs ont été amplifiés. Chez le manioc, un ou deux allèles ont été obtenus. Pour limiter le nombre de réactions de clonage, une amplification multiplexe (combinant au maximum les trois locus du manioc) a été faite pour chaque individu. Les allèles ont été ensuite différenciés après amplification spécifique du locus par leur taille connue (Fig. 24). Dans un premier temps, six allèles clonés par individu ont été séquencés puis dans une deuxième étape, des allèles particulièrement intéressants ont été séquencés ou reséquencés parmi un nombre accru d'individus. Dans le cas du locus AxC29 chez l'arracacha par exemple, deux allèles se sont révélés recombinants entre les deux groupes principaux d'allèles lors de la première étape de séquençage et des analyses phylogénétiques préliminaires (un individu de chaque espèce *A. acuminata*, *A. x. var. andina*). Pour confirmer ces résultats et écarter la possibilité d'artefacts, ces allèles ont été reséquencés et 10 accessions supplémentaires, cultivées, *A. x. var. andina* et *A. acuminata* ont été analysés. Les résultats ont confirmé la nature recombinante de ces allèles, et d'autres allèles recombinants ont été observés chez *A. acuminata* comme détaillés ultérieurement dans les résultats.

Une partie des chromatogrammes de séquences a été obtenue sur le séquenceur LICOR du laboratoire et le reste par un laboratoire commercial (SEQWRIGTH, US).

2.4 Analyse de séquences : Reconstruction phylogénétique de locus SSR allongés

Les séquences brutes (chromatogrammes) des allèles allongés ont été assemblées, nettoyées et éditées sur DNASTAR (2003). Une séquence consensus pour chaque allèle est ainsi obtenue et les allèles redondants (d'un même individu) éliminés. L'alignement des séquences a été fait sur MEGALIGN; les indels (insertions ou délétions, considérés à partir de 5pb) ont été codés en caractères binaires (absence ou présence) à la fin de la séquence nucléotidique et les motifs SSR ont été écartés de l'analyse en supprimant les éléments répétitifs.

Pour la reconstruction phylogénétique de chaque locus, le logiciel PAUP a été utilisé en employant deux méthodes d'analyse:

- i) la méthode du Maximum de Parcimonie (MP) qui se base sur l'hypothèse selon laquelle les relations phylogénétiques entre les taxa sont le plus probablement représentées par l'arbre comportant le minimum de mutations (nucléotide unique ou indel). Les options « *random step wise addition* » et « *tree bisection reconstruction* »

(TBR) » ont été utilisées pour la recherche heuristique des arbres parcimonieux. Cette méthode détecte quels caractères de la séquence sont informatifs pour une phylogénie. Un arbre consensus est obtenu par la méthode « consensus strict ». La robustesse des branches de l'arbre a été calculée par une analyse bootstrap avec 100 répétitions

- ii) le critère de Maximum de Vraisemblance (*Likelihood* -ML): cette méthode fournit l'arbre ayant la plus forte probabilité d'aboutir aux données, qui sera l'arbre le plus vraisemblable. Selon Whelan et al. 2001 (cité par Leotard 2003), cette méthode qui est plus longue en temps de calcul, permet une modélisation de l'évolution des séquences plus réaliste que les méthodes des MP. Les options précisées pour la méthode MP ont été aussi employées pour la recherche de l'arbre ML.

3. RESULTATS

Afin de faciliter la présentation des résultats, nous précisons, pour chaque locus étudié, les données de polymorphisme puis l'arbre obtenu par la méthode de ML car celle-ci fournit un seul arbre qui, dans tous les cas, a présenté une bonne similarité avec l'arbre consensus obtenu avec la méthode MP. Cet arbre est présenté comme un phylogramme non raciné dont les longueurs de branches sont proportionnelles au pourcentage de divergence estimé par le programme. Les pourcentages de bootstrap après 100 réplifications sont indiqués pour les groupes pour lesquelles ces valeurs sont supérieures à 50 (ces valeurs ont été obtenues par la méthode MP car la méthode ML demande un temps de calcul énorme avec nos moyens de calcul, de l'ordre de 200h par locus). Chaque allèle est présenté par le code de l'accension dont il a été isolé, puis sa taille *in silico* obtenue avec les amorces utilisés en génotypage.

3.1 Allongement des allèles SSR par marche chromosomique

Le tableau 11 montre les résultats obtenus pour les six locus SSR soumis à allongement. Pour quatre locus (dont trois chez le manioc) l'extension dépasse les 1000 pb alors que pour les deux locus restants (AxC87 et AxC64) l'allongement est moindre, malgré plusieurs tentatives pour obtenir un allongement plus important. Les taux d'allongement obtenus sont de l'ordre de 0.5 à 3.6 de la taille originale de la séquence flanquante du clone SSR. Les couples d'amorces sélectionnés pour l'amplification d'allèles allongés à partir de l'ADN génomique sont indiqués sur le tableau 11; les séquences des locus SSR allongés dans l'annexe 3 et les polymorphismes observés pour chaque locus analysé dans l'annexe 4, 5, 6, 7 et 8.

3.2 Allèles isolés et séquencés pour l'analyse phylogénétique

3.2.1. Chez l'arracacha:

Le tableau 12 montre les individus utilisés et leur génotype pour le locus analysé. Les séquences allongées obtenues avec les nouvelles amorces dessinées pour le locus AxC64 n'ont pas été exploitables à cause d'un manque d'homologie entre elles qui dénotait un problème probable avec les amorces. Ce locus a été donc écarté de notre analyse.

Au locus AxC87, 13 allèles sont représentés dans notre analyse (Tableau 12). Pour le locus AxC29, alors que le nombre d'allèles distingués sur les gels n'a pu être précisé à cause des bandes multiples observées, les séquences distinguent 8 allèles selon la longueur d'un fragment homologue *in silico* parmi les individus séquencés.

3.2.2. Chez le manioc:

Le tableau 13 montre les individus utilisés et leur génotype codé aux trois locus analysés (selon Morillo 2001). Il faut préciser que la taille des allèles dans Morillo (2001) a été estimée en gel, et qu'elle ne correspond pas nécessairement à la taille *in silico* utilisée dans la présentation de nos résultats (Voir ce tableau pour la correspondance entre allèles codés en génotypage et *in silico* sur les données des séquences). Nous avons réussi à représenter dans l'analyse phylogénétique, 75%, 49% et 89% des allèles distingués sur les gels pour les locus SSRY-3, SSRY-9 et SSRY-40, respectivement. De plus, des allèles « nuls » en génotypage sur gel (absence de bande, probablement à cause d'un polymorphisme dans le site des amorces qui empêche l'amplification) ont été amplifiés avec les nouvelles amorces et sont donc présentés (Tableau 13).

Chez le locus SSRY-3, 20 allèles sont représentés dans notre phylogénie. Les allèles 224, 238 et 244 sont les plus représentés. En tout, 40 allèles isolés de 28 accessions ont été séquencés (Tableau 14).

Chez le locus SSRY-9, 11 allèles sont représentés dans notre phylogénie. L'allèle 262, isolé chez 7 accessions est le plus représenté. En tout, 27 allèles isolés à partir de 22 accessions ont été séquencés (Tableau 15).

Au locus SSRY-40, 14 allèles sont représentés dans notre phylogénie. Les allèles 229, 225 et 223, isolés à partir de 7 accessions chacun, sont les plus représentés. En tout, 54 allèles isolés à partir de 36 accessions ont été séquencés (Tableau 16).

3.3 Polymorphisme des séquences flanquantes des locus SSR

En fonction du nombre de caractères informatifs, le taux de polymorphisme des régions flanquantes des 5 locus allongés est en moyenne de 5.6% chez le manioc. Le locus SSRY-3 est le plus polymorphe avec 8.2% de polymorphisme alors que le locus SSRY-9 montre seulement 3.2%. Chez l'arracacha, le locus AxC29 est le plus polymorphe avec 7.9% mais seulement 2.3% pour le locus AxC87; ce dernier résultat est peut-être lié au faible allongement obtenu au niveau de ce locus (Tableau 11). Pour chaque locus, la séquence des bases nucléotidiques des régions flanquantes retenues est montrée dans l'annexe 4 à 8. Une synthèse des résultats du polymorphisme observé au deux locus SSR analysés chez l'arracacha et au trois locus chez le manioc est montrée dans le tableau 17.

3.4 Analyse phylogénétique des locus SSR

3.4.1. Chez l'arracacha

Locus AxC87:

De l'analyse de 633 bases de séquences flanquantes alignées sur 21 allèles séquencés, seulement 14 caractères polymorphes ont été détectés (dont deux indels de 40 et 47 pb chacun), soit 2.2% de polymorphisme informatif (Tableau 17). L'arbre obtenu par ML (figure 25) montre que les allèles rencontrés chez les individus cultivés sont classés en deux groupes (A et B) mais distingués par des valeurs faibles de bootstrap (58 pour le groupe B). Ces groupes correspondent néanmoins à des longueurs de motifs bien différentes. Les allèles 91, 101 et 109 des cultivars se regroupent aux allèles rencontrés chez les individus sauvages. Au contraire, le groupe A avec les allèles 113, 115, 117 et 119 (taille déterminée *in silico*) est spécifique aux cultivars. La taille de l'allèle SSR et la phylogénie sur la base des séquences flanquantes convergent donc pour faire de A un groupe distinct d'allèles qui ne sont pas observés parmi les formes sauvages analysées.

Locus AxC29 :

Il s'agit d'un locus donnant des bandes multiples en gel, probablement à cause d'un glissement de la Taq polymérase, ce qui ne permet pas une bonne détermination en gel de la taille des allèles. Celle-là a par contre été déterminée *in silico* à partir de la séquence du locus allongé. Sur les 18 individus séquencés, 25 allèles ont été isolés dont huit se distinguent selon la longueur des motifs: trois sont présents chez les individus cultivés, cinq chez les formes sauvages de *A. xanthorrhiza*, et quatre chez d'autres formes sauvages, *A. incisa* et *A. acuminata* (Tableau 12). L'amplification du locus allongé à 1405 pb chez ces individus résulte en 7.9% de polymorphisme sur la séquence alignée (Tableau 17).

L'arbre ML de la figure 26 montre deux groupes de séquences soutenus par des valeurs de bootstrap proches de 100 (A et B) mais plusieurs allèles intermédiaires sont aussi mis en évidence. Le groupe A comprend les allèles les plus grands (134, 135, 136, 138 et 140) isolés de i) 4 cultivars, dont deux sont hétérozygotes avec un des allèles se retrouvant dans B, ii) 5 accessions *A. x. var. andina*, une *A. incisa* et deux *A. acuminata* (ces deux dernières sont hétérozygotes, avec un des allèles se retrouvant en position intermédiaire entre A et B).

Les allèles plus petits (117 et 125) se retrouvent dans le groupe B qui comprend une accession *A. acuminata*, une *A. x. var. andina*, une *A. x. var. monocarpa* et six cultivars, dont les deux hétérozygotes appartenant aux deux groupes (C1 et C2) distingués par les AFLP. A part ces deux cultivars qui sont hétérozygotes entre les groupes A et B, on note que le groupe A présente 2 cultivars C1 et le groupe B, 4 cultivars C2. Il se pourrait donc que les groupes A et B correspondent généralement aux groupes de cultivars C1 et C2, respectivement mais ceci reste à préciser. Cinq allèles sont intermédiaires entre ces deux groupes (de taille 124, 134 et 135). Ils sont tous isolés d'espèces sauvages, quatre *A. acuminata* et un *A. x. var. andina* (population de Huigra, code MST-041).

3.4.2. Chez le manioc

Locus SSRY-3

Sur 11 allèles observés à ce locus, 9 sont représentés dans l'analyse phylogénétique en plus de trois allèles nuls en génotypage mais révélés par séquençage. L'analyse par MP de 1110pb alignées retient 92 caractères informatifs (dont cinq indels) faisant de ce locus le plus polymorphe parmi les 3 locus analysés chez le manioc. La figure 27 montre le phylogramme

obtenu par ML. Tous les allèles, à l'exception de deux allèles de *M. brachyloba* (232 et 238), sont regroupés en deux clades principaux A et B distingués avec de fortes valeurs de bootstrap (94 et 90 respectivement). Le groupe A comprend la plupart des allèles observés chez des cultivars, tous ceux isolés chez *M. flabellifolia* et l'allèle 253 d'un arbre manioc (5.19). Le groupe B comprend tous les allèles observés dans le complexe *M. carthaginensis* (sauf l'arbre manioc 5.19 hétérozygote), la plupart des arbres manioc, le seul individu de l'espèce *M. quinquepartita* analysé, mais aussi deux cultivars traditionnels (Col22 et Pan 51) ainsi que deux cultivars africains introgressés artificiellement de *M. glaziovii* (5.59 et 5.70). Ces deux derniers se regroupent avec un individu *M. glaziovii* (5.42) du groupe G2 (le plus divergent du manioc parmi les deux distingués dans *M. glaziovii*, Morillo 2001) et un arbre manioc du Brésil (5.33, Mandioca de 7 anos) représenté par ses deux allèles (228 et 230). L'allèle 226 de *M. quinquepartita* de Guyane est regroupé avec trois allèles de *M. glaziovii*. La présence de deux cultivars traditionnels américains dans ce groupe indique une introgression spontanée, mais à partir d'allèles différents par la séquence flanquante et la taille du motif, donc indépendante, bien qu'homologue, de celle obtenue artificiellement en Afrique. Les arbres manioc, qui représentent des cultivars particuliers intermédiaires avec des formes sauvages indiquent clairement une voie possible de cette introgression.

Locus SSRY-9:

Pour ce locus, 27 allèles ont été isolés à partir de 22 individus. Tous les allèles observés en génotypage (8) sont représentés dans l'analyse phylogénétique avec de plus deux allèles nuls en génotypage qui ont été amplifiés par l'emploi des nouvelles amorces. Sur 1143pb alignées, l'analyse détermine 37 polymorphismes faisant de ce locus le moins polymorphe parmi les trois analysés chez le manioc (Annexe 2).

Le phylogramme obtenu par ML est montré dans la figure 28. On note que les allèles des cultivars sont distribués dans trois clades A, B et C à l'exception d'un allèle isolé du cultivar 28 (Manicoba, un triploïde confirmé être un hybride avec *M. carthaginensis*, Morillo 2001). Le clade A localisé en haut de l'arbre comprend l'allèle 250 (nommé allèle 244 en génotypage et présumé introgressé dans certains cultivars dont 5.62 (Col22) et trois autres, 5.49, 5.65 et 5.72 eux-mêmes suspectés introgressés, Morillo 2001) associé à l'allèle 266 du cultivar africain artificiellement introgressé (5.59), aussi retrouvé dans un arbre manioc (5.31) et associés à un cultivar à grosses tiges (5.47) qui présente l'allèle 264. Les séquences

flanquantes des allèles 250 et 266 diffèrent seulement pour deux SNP (positions 180 et 253) comme montré dans l'annexe 7.

Le clade A n'est pas soutenu par les valeurs de bootstrap mais est le plus proche de la majorité des allèles observés uniquement chez des formes sauvages éloignées. Le reste des allèles de cultivars (y compris arbres manioc) se localise dans les clades B et C, soutenues par des valeurs de bootstrap de 62 et 82, avec les 3 accessions de *M. flabellifolia* représentés et aussi un *M. glaziovii*.

Locus SSRY-40

Pour ce locus, 49 allèles ont été isolés à partir de 34 individus. Tous les allèles observés en génotypage ont été représentés ainsi que trois allèles nuls révélés en séquençage par l'emploi des nouvelles amorces. La phylogénie (Fig. 29) montre quatre groupes d'allèles principaux (A, B, C et D) avec des groupes internes montrant des valeurs de bootstrap de 69% ou plus. Les allèles de cultivars sont présents dans tous ces groupes et associés avec *M. flabellifolia* dans deux de ces groupes (B et D). Le groupe A, soutenu à 94% (à l'exception de la variété triploïde 5.28), présente aussi de bonnes valeurs bootstrap de ses branches internes. Il inclut des allèles d'espèces éloignées (*M. brachyloba*, *M. quinquepartita* et *M. carthaginensis*), deux arbres manioc (5.32 et 5.33) et trois cultivars : un introgressé artificiellement (5.59) et deux sont traditionnels, 5.52 (Pan51) déjà rencontré dans un clade n'incluant pas *M. flabellifolia* pour le locus SSRY-3 et 5.54 rencontré pour la première fois dans cette situation.

Le groupe B, soutenu pour sa plus grande part à 73% inclut aussi des espèces éloignées (*M. glaziovii*, *M. carthaginensis*) et deux variétés cultivées traditionnelles associées à ces mêmes espèces pour l'un (5.49, Cub51) ou les deux (5.62, Col 22) loci précédents pour lesquels elles ont été analysées. On serait ainsi tenté de conclure à leur introgression à partir du complexe *M. carthaginensis*. Néanmoins, les deux allèles d'une accession *M. flabellifolia* (5.91) sont également trouvées dans ce groupe B, ce qui contredirait l'hypothèse précédente. Il semble donc difficile de trancher entre les trois possibilités suivantes :

- i) ce groupe B représente des allèles ancestraux communs à *M. flabellifolia* et au complexe *M. carthaginensis* qui n'apporterait donc pas d'information phylogénétique ;
- ii) ce groupe représente des allèles du complexe *M. carthaginensis* et l'accession *M. flabellifolia* (5.91) est elle-même introgressée. Son origine géographique (Etat du Maranhon) est compatible avec cette hypothèse.

iii) ce groupe représente des allèles du complexe *M. flabellifolia* et les deux accessions *M. glaziovii* (5.7 et 5.42) sont introgressées du manioc ou de son ancêtre direct (*M. flabellifolia*).

Le fait que, pour les deux autres loci analysés du même groupe de liaison, les accessions en question ont été caractérisées comme introgressées (5.62, Col22 pour les locus SSRY-3 et SSRY-9, et 5.49, Pan51, pour le locus SSRY_9) ou pas introgressées (*M. glaziovii* et *M. flabellifolia*) pointe pour la première possibilité. La position relative des loci sur le même groupe de liaison rendrait néanmoins plausible l'hypothèse d'un fragment d'introgression n'incluant pas ce locus dans certain cas (les locus SSR-3 et 9 sont les plus contigus dans le groupe de liaison D (Mba 2001) et les plus proches des QTL de résistance partielle à la bactériose (Jorge 2000), alors que le locus SSR-40 est le plus éloigné de ces mêmes repères. Cependant, cette position relative pourrait être remise en cause (M. Fregene, pers. com.) et n'apparaît donc pas très fiable.

4. DISCUSSION

Qu'apportent de plus les séquences des régions flanquantes d'allèles SSR?

A l'heure actuelle, peu d'études se sont intéressées au polymorphisme des régions flanquantes des allèles SSR chez les plantes cultivés. Butuler et al. (1999) par exemple s'intéressent au séquençage d'allèles SSR afin d'identifier des mécanismes de mutations SSR chez les polypléides. Notre étude est à notre connaissance la première à appliquer le séquençage des régions flanquantes d'allèles SSR pour essayer de démontrer des introgressions. Grâce à la technique de marche chromosomique, notre étude est aussi la première à examiner ces régions avec une taille de séquences permettant de construire des phylogénies des locus SSR sur la base du polymorphisme observé.

Par ailleurs, l'analyse de séquences d'allèles nous a permis de détecter d'autres éléments de polymorphisme qui ne sont pas distinguables par génotypage SSR conventionnel.

- i) des substitutions de bases (SNP),
- ii) des variations dans la taille des allèles qui ne résultent pas nécessairement du polymorphisme des motifs mais des indels (insertions et délétions) dans les régions flanquantes (Ex. l'allèle 431 chez l'accèsion 5.90 de *M. flabellifolia* pour le locus SSRY-40),
- iii) des changements dans la composition de l'élément répétitif.

De plus, la reconstruction de la phylogénie des locus SSR permet de détecter plusieurs cas d'homoplasie probable d'allèles, c'est-à-dire des allèles homologues en taille mais avec une histoire évolutive différente. Sur tous les locus analysés, on observe aussi que des polymorphismes dans les régions flanquantes sont responsables de la non amplification de certains allèles (définis comme allèles nuls en génotypage). Nos résultats s'accordent ainsi avec Butuler et al. (1999) dans le sens qu'une assignation d'allèles basée seulement sur la taille allélique peut biaiser significativement les relations génétiques entre différentes espèces en l'absence des données de séquences. Ces auteurs concluent à la nécessité d'interpréter les résultats avec précaution lorsque l'on compare la diversité entre espèces différentes à cause justement de l'homoplasie des SSR.

Des allèles particuliers chez l'arracacha cultivée proviennent-ils d'espèces sauvages?

Comme déjà discuté, les analyses de diversité révèlent une portion de diversité propre à la forme cultivée *A. xanthorrhiza* qui est absente des formes sauvages de l'espèce ancestrale pérenne et monocarpique. L'étude des régions flanquantes menée en parallèle avec le manioc, en analysant deux locus pour lesquels des différences significatives en taille des allèles entre formes sauvages et cultivés sont observées, nous apporte certains éléments pour interpréter ces observations.

Parmi les deux locus analysés, le locus AxC87 était le plus intrigant car il distingue bien les formes cultivées et sauvages (Fig. 14). En effet, les allèles présents chez les formes cultivées dépassent d'une vingtaine de pb ceux amplifiés chez les formes sauvages, qu'elles soient pérennes ou monocarpiques. Le séquençage de ces allèles, effectué malgré l'allongement court obtenu pour ce locus, nous indique que:

- a. la différence de taille entre les allèles sauvages et cultivés est due au polymorphisme du motif SSR;
- b. bien que le polymorphisme des régions flanquantes soit faible (compte tenu de la courte extension obtenue pour ce locus), l'analyse phylogénétique montre que certains allèles cultivés (tous ceux de plus grande taille) ne sont pas regroupés avec ceux des formes sauvages de l'espèce. Toutefois, les valeurs bootstrap obtenues ne donnent pas une robustesse statistique à cette structuration. On peut néanmoins reconnaître que le groupe d'allèles A présente sans exception les motifs SSR longs, avec un total de 6 accessions cultivées contre 6 autres cultivées dans le groupe B avec des motifs plus courts.

Bien qu'il y ait une concordance du polymorphisme du motif SSR et de celui d'une de ses séquences flanquantes pour reconnaître une origine divergente entre les deux groupes d'allèles A et B, il n'est pas possible dans l'état actuel de trancher entre les hypothèses qui pourraient expliquer cette différence entre formes sauvages et cultivées. D'une part, les séquences montrent que la différence en taille est due au nombre de répétitions du motif SSR et on peut penser à une évolution depuis la domestication qui pourrait être ancienne comme suggérée par Bukasov (1981). Chez l'igname cultivée par exemple, Scarcelli (2005) détecte des mutations des allèles SSR dans l'évolution des clones, et des mutants décalés de quelques

répétitions du motif original ont été observés. Par ailleurs, il est possible aussi que dans un seul événement de mutation, des locus SSR incorporent plusieurs répétitions, bien que ce type de mutation ait une occurrence rare (Vigouroux, com. pers). D'autre part, le fait que l'analyse phylogénétique regroupe les allèles des cultivars selon leur taille, conduit à l'hypothèse que ces allèles pourraient avoir leur origine chez d'autres formes sauvages du genre. Il pourrait s'agir de formes géographiques d'*A. xanthorrhiza* non représentées dans notre échantillonnage, ou d'espèces sauvages plus éloignées. Ayant eu l'occasion de nous procurer des échantillons d'herbier d'Amérique centrale (*California Academy of Sciences*) nous y avons amplifié ce locus SSR. Les résultats montrent une bonne transférabilité des amorces du locus AxC87 sur 8 des 11 espèces testées (*A. bracteata*, *A. ebracteata*, *A. bandegeei*, *A. nelsonii*, *A. rigida*, *A. tolescencis*, *A. vaginata* et *A. dounell*), mais chez aucune d'entre elles, les allèles recherchés n'ont été observés. La question posée trouverait probablement une réponse dans une séquence flanquante plus longue et un échantillonnage plus important des formes sauvages.

Concernant le locus AxC29, le séquençage d'allèles fournit des résultats inattendus et très intéressants. La phylogénie de ce locus montre deux groupes d'allèles clairement différenciés au niveau génétique et bien soutenus statistiquement. Des allèles cultivés et sauvages de *A. xanthorrhiza* et de l'espèce *A. acuminata* sont présents dans chacun des groupes et des allèles de forme sauvage, principalement *A. acuminata*, forment une série d'intermédiaires en continuum. L'examen des séquences montre que cette position intermédiaire correspond à des séquences recombinantes. La répétition de l'amplification PCR du clonage et du séquençage conduit à la confirmation de ces séquences.

La détection d'allèles partagés entre *A. acuminata* et *A. xanthorrhiza* (sauvage et cultivé), peut être expliquée par deux hypothèses : introgression ou polymorphisme ancestral. A propos de la première hypothèse, *A. acuminata* est une espèce divergente par rapport au complexe *A. xanthorrhiza*. Elle a été récemment replacée au sein du genre *Arracacia* après avoir été classée dans un genre proche (*Neonelsonia*); nos données de diversité AFLP et chloroplastique confirment sa divergence génétique. C'est une espèce fréquente distribuée en Amérique centrale et dans les Andes de la Colombie, Equateur et Pérou (Knudsen 2003). Bristol (1988) signale l'avoir observée rarement en association avec l'arracacha cultivée dans les vallées Sibundoy du sud de la Colombie. Dans notre expérience, bien que souvent on puisse trouver des populations proches des champs de production comme à San José de Minas au nord de

l'Equateur, cette association au champ n'a pas été observée. D'un autre côté, Vasconez et Hermann (1997) n'ont pas obtenu de résultats après 20 tentatives de croisements expérimentaux entre une accession de *A. acuminata* et un cultivar.

L'hypothèse de polymorphisme ancestral apparaît peu applicable car d'autres espèces aussi divergentes comme *A. elata* ou *A. mochata* n'ont pas amplifié ce locus. L'espèce *A. acuminata* est la seule forme sauvage non tubéreuse pour laquelle une amplification du locus allongé fut obtenue, et cela chez toutes les accessions essayées (les autres espèces mentionnées n'ont conduit à aucune amplification malgré les tentatives avec ce couple d'amorces).

Il est d'autant plus intéressant de noter que l'observation de deux groupes identifiés par l'analyse de ce locus semble congruente avec la divergence des cultivars révélée par les AFLPs (chapitre I, partie III). Les allèles de *A. xanthorrhiza* cultivée et sauvage sont localisés dans les deux groupes et on note un recoupement des groupes alléliques avec les groupes de cultivars G1 et G2 définis sur la base des AFLPs. En effet, les accessions du groupe 1 et 2 des AFLPs sont bien différenciées au niveau des groupes observés sur ce locus (appelés A et B) avec deux accessions hétérozygotes pour ce locus qui présentent un allèle reparté dans chacun des groupes (ECU-1201 et ECU-1229). Les groupes alléliques A et B comprennent respectivement trois cultivars du groupe G1 et quatre cultivars du groupe G2 des AFLPs. Ceci reste néanmoins à vérifier sur un nombre plus important de cultivars.

On retrouve donc au niveau phylogénie de ce locus, des allèles recombinants entre les deux groupes plus divergents d'allèles (A et B), ces allèles sont présents quasi exclusivement chez l'espèce *A. acuminata* (5 accessions) avec l'exception d'un allèle recombinant isolé d'une accession de *A. x. var. andina* (MST-041, population de Huigra). L'observation du polymorphisme de séquence de ces allèles montre leur structure en mosaïque entre les deux groupes A et B (Annexe 4) comme si un grand nombre de recombinaisons alléliques avait eu lieu. Comment expliquer cette grande variabilité allélique au niveau de ce locus, particulièrement chez l'espèce *A. acuminata* ? Une hypothèse possible est qu'une espèce sauvage ou un *A. acuminata* "ancestral" se soit croisée avec *A. xanthorrhiza*. La forme *A. acuminata* étudiée serait dérivée de ce croisement. Un des allèles recombinants aurait été sélectionné dans les cultivars mais plusieurs auraient survécu dans *A. acuminata* dont on peut estimer qu'elle se reproduit par allofécondation

Phylogénie des locus SSR analysés chez le manioc et positionnement des allèles suspectés d'introgession du complexe *M. carthagenensis*

La phylogénie du locus SSRY-3 montre que mis à part deux allèles isolés chez l'espèce *M. brachyloba*, le reste des allèles est classé en deux groupes bien soutenus par les valeurs de bootstrap. Un premier groupe (A) comprend les allèles longs (238 à 257), et un deuxième (B) les allèles courts (214 à 238). Les allèles longs sont associés avec les cultivars de manioc et sa forme sauvage (*M. flabellifolia*), et dans un seul cas un allèle long a été isolé à partir d'un membre du complexe *M. carthagenensis* (l'accession 5.19 qui est un arbre manioc certainement d'origine hybride cultivé x *M. glaziovii* en accord avec nos résultats de SSR, Morillo 2001). De leur côté, les allèles courts sont associés avec le complexe *M. carthagenensis* (*M. glaziovii* des groupes 1 et 2 et *M. carthagenensis*) et l'espèce *M. quinquepartita*. On remarque aussi la présence dans ce groupe de deux variétés artificiellement introgressés (5.59 et 5.70) ainsi que deux cultivars traditionnels (5.52 et 5.62). Dans ce cas l'hypothèse d'introgession est bien soutenue. Le fait qu'on retrouve l'introgession chez deux variétés artificiellement introgressés de *M. glaziovii* est cohérente avec leur résistance à la mosaïque virale africaine, suite à une introgression dirigée à partir de *M. glaziovii*. Cette résistance introgressée est en effet connue pour être portée par le groupe de liaison D auquel appartient le locus SSRY-3. On peut donc raisonnablement rejeter l'hypothèse d'un polymorphisme ancestral expliquant le regroupement observé dans cet arbre. Les différences de longueurs du motif indiquent d'autre part que les deux couples de cultivars, traditionnels et améliorés, représentent deux événements d'introgession différents.

Concernant le locus SSRY-9, de sa phylogénie il ressort que l'allèle 250 (préssumé introgressé) se montre divergent par rapport aux autres observés chez le manioc (260, 262 et 264 qui sont regroupés avec des allèles de *M. flabellifolia*). Les allèles 250 des cultivés sont regroupés avec le même allèle isolé d'un arbre manioc et l'allèle 266 de l'accession artificiellement introgressé 5.59. En l'état actuel, deux limitations ne permettent pas de confirmer complètement l'hypothèse d'introgession:

- i) la robustesse de ce groupe n'est pas validée par une bonne valeur bootstrap à cause du faible polymorphisme observé; et;

- ii) Cet allèle n'est pas regroupé avec le reste des allèles du groupe *M. carthaginensis* bien qu'il intègre un arbre manioc; pour l'instant on est incapable de déterminer son origine.

Cependant, le fait que l'allèle 244 soit partagé par des variétés présumées introgressées par ailleurs (5.62, 5.68 et 5.65, Morillo 2001) et qu'il soit présent chez la variété artificiellement introgressée (5.59) sous un allèle différent (262) soutient l'hypothèse d'introgression par rapport à celle d'un polymorphisme ancestral.

Une même conclusion est atteinte dans le cas du locus SSRY-40: sa phylogénie révèle une diversité intéressante des cultivars car ils sont représentés dans les quatre groupes principaux définis dans le phylogramme. Particulièrement intéressant est le groupe A, le plus divergent vis-à-vis du groupe D (qui inclue la plupart des allèles cultivés et de *M. flabellifolia*). Le clade A regroupe les allèles des espèces les plus divergentes par rapport au manioc rencontrées dans cette étude, en l'occurrence *M. brachyloba*, *M. quinquepartita* et *M. carthaginensis*, et aussi remarquablement la variété artificiellement introgressée (5.59) ainsi que deux cultivars traditionnels, Pan51 dont l'affinité avec le complexe *M. carthaginensis* et *M. quinquepartita* a été aussi mise en évidence au niveau du locus SSRY-3 (Fig. 27), et 5.54 rencontré pour la première fois dans cette situation.

L'introgression d'allèles du complexe *M. carthaginensis* est-elle donc confirmée chez le manioc cultivé?

Différentes études de diversité moléculaire menées chez le manioc apportent des éléments qui confortent l'hypothèse d'une contribution d'espèces sauvages, outre la forme ancestrale postulée (*M. flabellifolia*) au génome de la forme cultivée. Notre étude de DEA (Morillo 2001) contribue dans ce sens en apportant de nouveaux indices à partir du complexe d'espèces *M. carthaginensis* signalé comme un des plus distants par rapport au manioc cultivé (Second et Iglesias 2001). Comme le gène le plus séquencé chez le manioc (G6pdh) le montre, la phylogénie des espèces du genre *Manihot* n'est pas toujours soutenue par de fortes valeurs de bootstrap (Olsen et Schaal 1999, Leotard 2003). Dans plusieurs cas, nos phylogénies montrent des valeurs de bootstrap comparables ou supérieures à celles observées au locus G6pdh, tout en indiquant des clades communs au complexe *M. carthaginensis* et à certains cultivars qui n'incluent pas l'ancêtre direct présumé du manioc, *M. flabellifolia*. Cependant, nos données concernent un nombre trop limité d'accessions de *M. flabellifolia*.

Mais, en choisissant trois locus dans le même groupe de liaison D, on permet de présumer des introgressions du complexe *M. carthagenensis* dans des cultivars traditionnels à plus d'un locus du même segment chromosomique et donc d'atténuer le risque de confondre introgression et polymorphisme ancestral. De plus, ce groupe de liaison D porte le locus de résistance à la mosaïque virale introgressé artificiellement de *M. glaziovii* et donc permet de vérifier que nous sommes à même de mettre en évidence cette introgression, tout en s'intéressant à un segment chromosomique fortement soumis à sélection.

Les arbres manioc, qui représentent des cultivars particuliers hybrides avec des formes sauvages indiquent clairement une voie possible de cette introgression.

Parmi les trois allèles qui étaient notamment suspectés introgressés dans des cultivars (nommés 204 et 216 pour le locus SSRY-3 et 244 pour le locus SSRY-9 dans Morillo 2001, et qui correspondent *in silico* aux allèles 216, 241 et 250 respectivement), l'allèle 216 du locus SSRY-3 supporte bien l'hypothèse d'introgression à partir de *M. glaziovii* chez les variétés artificiellement introgressés bien que cet allèle ne soit pas présent chez les cultivars traditionnels analysés. L'allèle 241 par contre rejette l'hypothèse (Fig. 27). Cependant, d'autres allèles qui n'étaient pas suspectés introgressés par leur taille (238 et 240 chez les cultivars 5.52 et 5.62 respectivement) sont regroupés par le polymorphisme de la région flanquante dans le clade d'allèles du complexe CTH, soutenu par une bonne valeur bootstrap (90), qui inclue également l'allèle 216 au niveau de ce locus.

Pour l'allèle 244 suspecté introgressé au locus SSRY-9 (taille 250 au niveau de la séquence, Fig. 28), les résultats obtenus au niveau séquence seul ne permettent pas d'accepter complètement l'hypothèse à cause du faible polymorphisme mis en évidence. Néanmoins, le fait que cet allèle est partagé par des variétés traditionnelles présumées introgressées comme Col22, la variété artificiellement introgressée et un arbre manioc, mais pas par *M. flabellifolia* va dans le sens d'une introgression. Sur la base d'un regroupement similaire mais au niveau marqueurs AFLP et RAPD, Second (1998) suggère la possibilité que *M. glaziovii* ait contribué spontanément à la diversité du manioc dans l'aire d'origine. Leotard (2003) corrobore une hypothèse d'introgression réciproque en repérant le cas inverse, un haplotype du manioc dans *M. glaziovii*.

Par ailleurs, la reconstruction phylogénétique de locus SSR chez le manioc cultivé a permis la détection d'autres allèles regroupés avec de fortes valeurs bootstrap aux espèces éloignées ; sur le locus SSRY-3, les cultivars 5.62 et 5.52 sont regroupés avec le complexe *M. carthaginensis*; et plus étonnant encore, des allèles des accessions 5.52 et 5.54 regroupés avec l'espèce *M. quinquepartita* sur le locus SSRY-40.

Des résultats similaires ont été rapportés par d'autres auteurs. Olsen (2002) retrouve ainsi des haplotypes partagés entre l'espèce *M. pruinosa* et le manioc cultivé en attribuant ce résultat à des introgressions récentes ou au polymorphisme ancestral entre ces deux espèces; dans ce cas, la divergence entre *M. pruinosa* et *M. flabellifolia* est cependant beaucoup moins importante que celle avec le complexe *M. carthaginensis*. Leotard (2003) en examinant les relations phylogénétiques entre divers taxa par le séquençage d'un gène nucléaire (G3PDH), retrouve certains allèles cultivés associés à des allèles d'espèces plus distantes. Cet auteur retrouve une manque de congruence entre la généalogie de haplotypes et les taxa analysés et mentionne la persistance de polymorphisme ancestral comme un facteur dont il faut tenir grand compte.

Le polymorphisme ancestral est en effet un facteur à considérer, d'autant plus que le genre *Manihot* est proposé s'être diversifié récemment (Rogers et Appan 1973), ce qui explique l'absence de barrières reproductives entre les différentes espèces. Leotard (2003) conclut cependant que les deux hypothèses (polymorphisme ancestral et flux de gènes) s'appliquent pour expliquer ses résultats en soulignant néanmoins comme très probable l'existence d'introgressions vue l'occurrence d'hybrides spontanées entre le manioc domestiqué et des taxons sauvages.

Comme suggéré par Leotard (2003), il sera par la suite intéressant d'analyser nos résultats sur des modèles de coalescence disponibles en génétique des populations. Ces modèles intègrent les informations issues de fréquences alléliques et les analyses phylogénétiques afin de séparer la structure de populations de leur histoire. Ces outils devraient permettre d'évaluer l'hypothèse des flux de gènes entre les compartiments sauvage et domestiqué et de trancher entre un scénario avec ou sans introgression post-domestication.

L'approche méthodologique utilisée est-elle appropriée pour repérer les introgressions?

Olsen (2002) signale que la capacité à repérer les introgressions est critique pour la compréhension de l'histoire évolutive des populations de plantes. Dans l'étude présente, nous avons utilisé une approche méthodologique innovante pour tester l'hypothèse d'introgression, qui consiste à comparer les régions flanquantes allongées des allèles suspectés d'introgression entre les différentes espèces. Ces allèles ont été repérés sur la base de leur taille allélique et de leur absence chez l'ancêtre postulé. Chez le manioc leur présence dans des variétés présumées introgressées pour d'autres raisons et l'analyse multivariée de données de polymorphisme de marqueurs moléculaires notamment, ont été aussi prises en compte.

L'efficacité de l'approche utilisée peut être évaluée en analysant le cas des allèles suspectés d'introgression. Chez le manioc, des trois cas suspectés, le locus SSRY-3 supporte fortement l'hypothèse d'introgression et les résultats des locus SSRY-9 et 40, situés sur le même segment chromosomique que SSRY-3, recourent bien ce résultat. On peut donc considérer l'hypothèse d'introgression spontanée comme prouvée chez le manioc. Chez l'arracacha, pour le locus AxC87 l'hypothèse d'introgression n'est pas entièrement résolue, alors que pour le locus AxC29 l'hypothèse peut être acceptée.

En conclusion, l'approche utilisée permet de soutenir l'hypothèse d'introgression dans certains cas, mais a ses limites dans les taux de polymorphisme des régions flanquantes qui restent faibles malgré l'allongement. Cependant, ceci peut varier selon les espèces analysées. Buteler et al. (1999) par exemple rapportent des taux élevés de mutation dans les régions flanquantes des allèles SSR de la patate douce cultivée (*Ipomoea batatas*) et de ses homologues sauvages. Le niveau de polymorphisme varie aussi selon le locus étudié. Chez le manioc par exemple le locus SSRY-9 fut le moins polymorphe parmi les trois examinés. Cette différence peut être liée à sa localisation dans le génome car des locus moins polymorphes sont probablement proches de gènes soumis à sélection (Vigouroux, com. pers).

En conclusion, la capacité de démontrer une introgression par notre approche dépend du degré de divergence entre les espèces étudiées et de la possibilité d'étudier des locus liés.

CONCLUSIONS ET DISCUSSION GENERALE

Cette dernière partie commence par un récapitulatif des principaux résultats que nous discuterons en relation leur avec la connaissance sur la domestication des plantes, notamment dans la région andine. Nous discuterons ensuite plus généralement de l'introgession génétique d'espèces éloignées et de ses implications en matière d'agrobiodiversité.

I. Récapitulatif des principaux résultats

Identification des marqueurs microsatellites (SSR) pour l'étude de la diversité chez *A. xanthorrhiza*

Quatorze locus ont été identifiés comme polymorphes au sein du complexe *A. xanthorrhiza*. Seule l'information donnée par onze marqueurs à profil diploïde a été utilisée car les analyses de diversité ont été faites sous l'hypothèse de diploïdie chez cette espèce.

Structure génétique et analyse d'interactions génétiques parmi le complexe sauvage-cultivé de *A. xanthorrhiza*, par l'emploi des SSR

Les SSR montrent une structuration génétique forte entre formes sauvages et cultivées alors que parmi les formes sauvages une faible divergence entre les deux sous-espèces pérenne et monocarpique est observée. Le compartiment sauvage reflète une grande diversité suggérant une forte sélection au cours du processus de domestication. Un critère de sélection appliqué à l'arracacha a pu être la réponse à la reproduction végétative, technique de propagation courante dans la domestication des plantes dans les Andes. Sur la base de l'origine des allèles de la forme cultivée, une zone probable de domestication est identifiée, en écartant des populations qui avaient été proposées auparavant. La forme cultivée présente, par rapport aux formes sauvages, une faible diversité qui est néanmoins plus importante que ce que l'on pensait préalablement. Au niveau de 11 locus SSR, cette diversité n'apparaît pas fortement structurée ni associée à la morphologie des cultivars, même si quelques cultivars du Pérou semblent un peu différenciés. La diversité allélique observée est recombinaisonnée, notamment en relation avec la distribution géographique de la diversité génétique du parent sauvage, ce qui suggère les effets d'un brassage génétique dans le compartiment cultivé au cours d'événements postérieurs à la domestication.

Nous montrons l'existence de flux génétique spontané entre formes sauvages et cultivées dans l'aire de domestication. Des individus intermédiaires issus de ce processus sont confirmés (en

milieu sauvage notamment) et comprennent un cultivar introgressé du type sauvage pérenne. Le faible nombre de génotypes détectés suggère cependant une faible occurrence de ce processus.

Origine de la domestication et structuration insoupçonnée des formes cultivées

La forme sauvage pérenne de l'espèce *A. xanthorrhiza* est bien la forme la plus apparentée à l'arracacha cultivée et donc sa forme ancestrale présumée avec une diversité supérieure, mais celle-ci n'expliquant pas à elle seule toute celle du type domestiqué. Le choix de la forme pérenne a probablement été favorisé au cours de la domestication pour son aptitude naturelle à la reproduction végétative.

Deux groupes génétiques, qui n'avaient pas été notés préalablement et que nous considérons donc cryptiques, sont clairement distingués parmi les formes cultivées au niveau moléculaire. Cette structuration, significative par rapport aux perspectives d'amélioration génétique de l'espèce, est révélée par les AFLPs et retrouvée au niveau du polymorphisme des séquences flanquantes d'allèles d'un locus SSR (AxC29). Un des groupes s'avère plus proche de la forme sauvage ancestrale alors que le deuxième se montre plus proche du type monocarpique. Selon les données de l'ADN chloroplastique, on explique cette divergence par la possible participation du type monocarpique après la domestication. Ce résultat nous suggère une dynamique génétique d'origine sexuée dans l'aire d'origine d'une espèce cultivée par propagation végétative et nous fournit une information précieuse en matière de conservation et d'utilisation future des ressources génétiques de l'arracacha.

Hypothèse de l'introgession génétique réciproque chez l'arracacha et le manioc cultivé à partir de formes sauvages éloignées

Chez l'arracacha, nos résultats de diversité suggèrent un scénario de domestication avec une forme ancestrale majeure mais qui n'explique pas toute la diversité du compartiment cultivé. Ce scénario qui rappelle celui du manioc, nous a mené à tester l'hypothèse de la participation d'autres formes sauvages par introgression génétique chez les deux plantes. L'étude a consisté à mettre au point une méthodologie visant à déterminer l'origine de certains allèles SSR observés dans les formes domestiqués des deux plantes. Sur la base principale de leur absence chez les formes sauvages ancestrales, mais aussi de différences parfois importantes de la taille du locus SSR, l'hypothèse de travail était que ces allèles pouvaient avoir leur origine chez d'autres espèces sauvages phylogénétiquement éloignées bien que faisant partie du même genre botanique. Méthodologiquement, il s'agissait d'allonger par marche chromosomique,

des clones SSR de manière à accéder à une séquence “flanquante” plus longue. Ces séquences flanquantes ont été alors séquencées dans certaines espèces sauvages disponibles du même genre pour étudier leur phylogénie en relation avec celle des formes cultivées. Le séquençage de ces fragments allongés a été mis en place sur 268 allèles clonés à partir de 30 accessions de manioc et 18 d’arracacha sélectionnés sur la base des analyses de diversité. La phylogénie basée sur le polymorphisme des séquences flanquantes d’un groupe d’allèles pour chacun des locus SSR analysés, montre en conclusion que:

a) chez l’arracacha, l’hypothèse d’introgession n’est pas entièrement résolue pour l’un des locus SSR analysés (AxC87), alors que pour un deuxième (AxC29) l’hypothèse est soutenue;

b) chez le manioc, un locus fournit la preuve définitive d’introgession chez deux cultivars traditionnels de manioc à partir de *M. glaziovii* (SSRY-3) et un deuxième supporte cette hypothèse (SSRY-9) en associant des observations acquises durant notre travail de DEA (Annexe 9). Sur l’ensemble des trois locus, qui appartiennent au même groupe de liaison, l’hypothèse de polymorphisme ancestral est rejetée et l’introgession est irréfutable. Ce résultat a une application possible immédiate dans la résistance présumée de ces cultivars à la mosaïque virale africaine (CMV). En effet, une résistance à la CMV introgressée artificiellement en Afrique concerne le même segment chromosomique et la résistance croisée à la CMV et à la bactériose vasculaire originaire d’Amérique est connue (Jorge 2000). Or l’introduction de la CMV en Amérique est redoutée et ces cultivars traditionnels pourraient donc trouver leur importance pour s’en prémunir.

II. Qu'apporte notre étude chez l'arracacha?

A. Des connaissances chez une plante dite mineure originaire d'un centre majeur de domestication

La domestication des plantes dans la région andine et périandine est à l'origine de plusieurs plantes cultivées d'importance dans le monde et d'autres d'importance régionale comme les racines et tubercules. L'étude de l'arracacha nous apporte des éléments sur l'évolution des plantes domestiquées dans cet agrosystème traditionnel, la plupart étant marginales depuis la découverte du Nouveau Monde (NRC 1989). Pour l'arracacha, les connaissances historiques suggèrent son importance avant l'arrivée des Européens, notamment dans la partie septentrionale des Andes où des ethnies indigènes ayant conservé une forte tradition de sa culture ont été signalées (Hodge 1954, Bristol 1988). Comme résumé dans « 1491 » (Mann 2005), on peut estimer que les populations du Nouveau Monde étaient très nombreuses et culturellement avancées; l'arrivée des Européens s'est traduite par leur effondrement, en grande partie suite aux épidémies dont elles avaient été protégées par leur isolement géographique et à la coercition du colonisateur qui a suivi. Chez le manioc par exemple, des traditions comme celles représentant une gestion dynamique de la diversité génétique peuvent être retrouvées parmi des populations de l'Amazonie brésilienne et guyanaise (Empeiraire 1998, Elias 2000) mais aussi du côté Pacifique comme en Equateur. Il n'en reste pas moins que de nombreuses autres traditions touchant à la domestication des plantes ont pu disparaître au cours de cette catastrophe, en particulier celles qui avaient cours principalement sur les côtes océaniques et la zone andine où se sont d'abord installés les Européens. Il est probable que ceux-ci sont parvenus dans un hémisphère déjà profondément anthropisé, en particulier dans la zone de domestication andine, mais que de nombreuses populations humaines n'ont malheureusement survécu que suite à un délabrement important de leur tissu social.

Il est donc important de considérer que des traditions paysannes ont pu se perdre suite à la désintégration ou extinction des groupes amérindiens andins. A titre d'exemple on peut citer deux groupes indigènes réputés pour être de grands agriculteurs d'arracacha, les indiens Chibcha et Pijao de la région sud de la Colombie (Hodge 1954, Bristol 1988). Ces groupes ont été dévastés ou déplacés pendant la conquête des Espagnols et avec eux, des pratiques culturelles se sont probablement perdues en laissant l'héritage de l'arracacha aux métis et colons qui habitent à ce jour ces régions (Hodge 1954). Dans la région de Tunja en Colombie par exemple, zone où s'est développé la culture indigène Chibcha mentionnée, la population

indigène a été réduite de huit fois en deux siècles, selon des recensements démographiques de l'époque coloniale (Morales 2005).

B. Un scénario original de domestication: sélection d'un type sauvage avec une aptitude naturelle à la reproduction végétative

L'étude de l'arracacha nous suggère un scénario original de domestication avec la sélection (inconsciente?) d'un type sauvage pérenne pour son aptitude à la reproduction végétative. Ainsi l'arracacha, qui est la seule ombellifère cultivée à être propagée principalement par voie asexuée, caractérise avec les autres plantes à racines et tubercules andins et aussi le manioc un « patron de domestication » dans cette région. Parmi les autres plantes andines qui partagent ce syndrome de domestication, citons : le *yacon*, propagé asexuellement alors que les autres espèces de la famille Compositae domestiquées dans d'autres régions du nouveau monde sont cultivées par graine (tournesol et topinambour, ce dernier cultivé aussi pour ses racines tubéreuses); la *mauka* et l'*achira*, alors qu'elles se propagent bien par graine, sont aussi cultivées avec des propagules végétatives obtenues de la partie basale des tiges. On peut donc conclure à un patron de domestication appliqué par l'homme andin comme suggéré par Hawkes (1989). Hermann (1997) suggère que la capacité à la reproduction végétative chez l'arracacha aurait été découverte par accident (et c'est probablement le cas d'autres plantes), par exemple après la cueillette de racines lorsque des couronnes étaient éliminées et donnaient lieu à de nouvelles plantes. Il est possible que dans les premiers stades de domestication il y ait eu participation de différentes formes sauvages, y compris des formes monocarpiques de l'espèce, mais le critère de sélection de la reproduction végétative appliqué a conduit à ce qu'une seule forme d'individus aboutisse à la constitution du stock cultivé. Chez les formes monocarpiques, elles ne sont pas par nature aptes au bouturage et la seule façon de propager ce type de plantes serait l'utilisation des graines. Leur production de graines est abondante et de bonne viabilité (Knudsen 2003) et des plantules issues des graines peuvent être retrouvées en abondance dans les alentours des plantes adultes manifestant une croissance racinaire précoce; cependant le temps pour aboutir à la floraison n'a pas été caractérisé mais semble pouvoir s'étaler sur quelques années dans les conditions naturelles (Blas 2005). Le cycle biologique limite ainsi la possibilité de domestication de cette variété sauvage, même par l'utilisation des graines.

C. Flux de gènes et pratiques paysannes dans l'agrosystème traditionnel andin

Selon Groot et al. (2003), le flux de gènes est un processus commun dans la nature et il n'est pas surprenant qu'il se retrouve entre plantes cultivées et formes sauvages apparentées (formes ancestrales comprises). Ellstrand et al. (1999) concluent que de tels processus sont courants chez les plantes cultivées car ceci a été documenté chez 12 des 13 plantes parmi les plus importantes du monde. Ces auteurs attribuent ceci à la récente divergence évolutive entre formes cultivées et leurs ancêtres sauvages. Chez l'arracacha, malgré les limitations déjà explicitées qui inhibent le flux de gènes, on montre que ce processus a bien eu lieu, bien que nos résultats suggèrent d'une faible fréquence actuelle dans l'aire d'origine. Chez d'autres plantes à propagation végétative comme le manioc ou l'igname, ceci a été aussi mis en évidence grâce à leur régime reproductif mixte qui combine asexualité et sexualité (Elias 2000, Leotard 2003, Scarcelli 2005). De même, chez l'arracacha, malgré ses faibles taux de floraison dans l'aire d'origine, nos résultats suggèrent un rôle de la reproduction sexuée dans la diversification de la forme cultivée. On repère ainsi un groupe cryptique de cultivars clairement différenciés au niveau moléculaire, attribué à une divergence postérieure à la domestication. L'existence de ce groupe de cultivars, des cas d'introgession repérés et le brassage de la diversité génétique de l'ancêtre sauvage impliquent une participation de la reproduction sexuée et suggèrent l'existence ancienne de pratiques paysannes traditionnelles qui lui correspondent dans l'agrosystème andin. Comme déjà mentionné, il est possible que ces pratiques ont disparu avec l'effondrement des cultures indigènes suite à la colonisation européenne. Cependant, certaines pratiques ont été observées de nos jours ; par exemple chez la pomme de terre, l'utilisation sporadique de la reproduction sexuée expliquerait la grande diversité morphologique des tubercules (Quiros et al. 1992). Ceci pourrait être aussi le cas d'autres tubercules andins mineurs (*oca*, *melloco*, *mashua*) qui montrent une énorme diversité morphologique, probablement d'origine ancienne car de telles pratiques ne sont pas observées. Chez l'arracacha, l'utilisation des plantules issues de graines a été exceptionnellement observée chez les indiens Sibundoy du sud de la Colombie (Vasquez et al. 2003, Lobo com. pers) et c'est d'ailleurs dans cette région qu'une importante diversité des cultivars a été signalée (Hodge 1954, Bristol 1988).

III. L'introgession comme mécanisme de diversification des formes cultivées

L'hybridation et l'introgession interspécifique ont été observées dans de nombreux groupes d'espèces végétales et ce sont des processus naturels essentiels pour leur spéciation (Groot et al. 2003). L'introgession est définie comme l'assimilation des gènes d'une population ou

espèce dans le pool génétique d'une autre espèce. Pour que l'introgression soit atteinte, les hybrides de première génération doivent être suffisamment adaptés pour se recroiser avec un des parents jusqu'à retrouver leur fertilité. Ce scénario est applicable au manioc où des hybrides spontanés avec une espèce divergente (*M. glaziovii*) ont été documentés (Second 2000). Sur cette base et des antécédents déjà détaillés, nous avons testé l'hypothèse d'introgression éloignée chez le manioc, en intégrant aussi le cas de l'arracacha chez qui une portion de diversité propre au type domestiqué est d'origine inconnue. L'approche méthodologique utilisée a été longue à mettre en œuvre mais elle s'est révélée efficace car elle favorise pour un cas dans chaque plante l'hypothèse d'introgression éloignée.

Chez l'arracacha dans le cas d'un locus (AxC29), l'introgression aurait conduit à l'existence de deux groupes d'allèles représentés à la fois parmi les cultivars, ses formes ancestrales et chez une espèce éloignée (*A. acuminata*). Des allèles recombinants sont observés parmi cette espèce sauvage, impliquant des recombinaisons répétées entre les deux groupes d'allèles. Pris dans leur ensemble, ces résultats suggèrent l'hypothèse d'une origine anthropique de cette espèce sauvage divergente, à tester lors de futures observations.

Chez le manioc, l'introgression à partir d'espèces éloignées du genre *Manihot* (tel que *M. glaziovii*) était suggérée avoir une incidence dans la diversification de la forme cultivée. Nos résultats fournissent au moins pour un locus la preuve irréfutable d'introgression (deux cultivars traditionnels de manioc, 5.62 (Col22) et 5.52 (Pan51) sont introgressés de l'espèce *M. glaziovii* au niveau du locus SSR Y-3). Chez le manioc on dispose de plus d'éléments pouvant être associés à nos résultats. De nombreux hybrides interspécifiques sont connus dans ce genre divers et largement répandu qui ne semble pas présenter de barrières reproductives strictes. Par exemple, le fait que des hybrides spontanés soient rencontrés entre des espèces très divergentes comme le manioc cultivé et *M. glaziovii*, permettait d'envisager l'hypothèse d'introgressions larges. Notre étude de DEA (Morillo 2001) a contribué dans ce sens en apportant de nouvelles observations qui suggéraient: a) l'introgression entre ces deux taxa à l'origine de certains polymorphismes au sein des cultivars et, b) que l'introgression est réciproque: des gènes du taxon domestiqué sont passés au taxon sauvage (certaines accessions de *M. glaziovii* sont introgressées du manioc cultivé et forment un groupe plus proche des arbres manioc).

L'introggression d'allèles provenant de formes sauvages éloignées dans des variétés traditionnelles de manioc serait un processus rendu possible par:

- 1) le régime mixte de reproduction du manioc, normalement reproduit végétativement mais aussi incorporant la reproduction sexuée dans la gestion traditionnelle de sa diversité par les amérindiens (Elias 2000),
- 2) l'hypothèse que le genre *Manihot*, bien que vaste et très variable, représente en fait un seul complexe d'espèces, au sein duquel la diffusion de gènes ne sera pas limitée par des barrières génétiques strictes,
- 3) le fait que la plupart des espèces sauvages du genre peuvent, dans leur aire d'origine, se trouver fréquemment au contact du manioc cultivé, rendant possible les flux de gènes et l'introggression

D'autant plus intéressant est le fait de retrouver, bien que pour des allèles différents, cette introggression chez une variété artificiellement introgressée en Afrique. Cette observation suggère que des introggressions spontanées à partir de croisements différents peuvent aussi conduire à des patrons d'introggression non aléatoire et concerner, de manière déterministe certaines zones du génome, comme décrit chez le genre *Helianthus* (Rieseberg et al. 1998). L'hypothèse de la stabilité des patrons d'introggression chez le manioc reste à vérifier par la suite.

Notre travail permet de confirmer que d'autres espèces, parmi les plus divergentes du vaste genre *Manihot*, ont contribué à la constitution génétique du manioc cultivé par introggression et ont elles-mêmes introgressé de l'ADN du manioc cultivé.

D'un autre côté, nos résultats suggèrent que des hybridations larges, combinées avec l'utilisation de la reproduction végétative, pourraient être deux mécanismes en interaction dans l'évolution de certaines plantes cultivées. Il est possible que cette association puisse se retrouver chez d'autres plantes à reproduction asexuée dont la plupart sont allogames et très hétérozygotes (Zohary 2004). Rabinowitz et al. (1990) ont montré chez la pomme de terre qu'une allogamie élevée favorise l'hybridation interspécifique, et que ce processus est suivi de la sélection humaine de phénotypes désirables par reproduction végétative. L'adoption de ce régime permet ainsi de sélectionner des « génotypes exceptionnels » et de les multiplier

massivement par clonage (Zohary 2004). L'hypothèse de cette association entre introgressions larges et reproduction végétative reste néanmoins à tester sur d'autres plantes.

En conclusion, notre travail fournit des résultats pionniers concernant l'introgression à partir des formes éloignées comme un mécanisme possible dans la diversification de certaines formes cultivées et/ou sauvages. On montre chez les deux plantes étudiées, que des introgressions entre formes cultivées et espèces divergentes ont eu lieu. Ce résultat est important pour la connaissance de l'évolution de l'agrobiodiversité et peut s'appliquer notamment dans le cadre de la gestion de l'impact des OGM dans les agrosystèmes traditionnels. Nous allons aborder brièvement ceci dans la partie finale de cette discussion.

IV. Implications des introgressions larges chez les formes cultivées: présence de transgènes dans les agrosystèmes traditionnels

Les contacts entre plantes cultivées et formes sauvages apparentées sont des événements favorisant le flux de gènes depuis le début de l'agriculture. Groot et al. (2003), après une compilation d'études de cas, concluent que ce genre de processus serait plus la règle que l'exception chez les plantes cultivées. En effet, même parmi les espèces réputées autogames comme le haricot, dont le niveau d'allogamie ne dépasserait pas 5%, ce processus a été démontré (Papa et Gepts 2004).

L'introgression, qui fait suite à l'hybridation initiale est cependant un processus plus difficile à déceler, notamment dans les générations plus avancées car le nombre de marqueurs diagnostiques doit être plus élevé (Santoni et al. 2000). Malgré cela, ce phénomène a été démontré chez certaines plantes (Ex. la carotte, le tournesol, la fraise et le navet), bien qu'impliquant plus ou moins des formes génétiquement proches.

Sur la base de ces évidences, le flux de gènes et l'introgression entre plantes génétiquement modifiés (OGM) et formes apparentées (ou variétés cultivées conventionnelles) est un risque réel. Ce sujet est aussi d'intérêt pour le grand public. Groot et al. (2003) concluent que l'évaluation de risques possibles d'introgression de transgènes dans des populations naturelles est un sujet qui reste encore à étudier. Ces auteurs mentionnent parmi les effets négatifs ou risques:

- i) l'établissement de la plante transgénique en milieu agricole et naturel,

- ii) Le risque de fuite du transgène dans le milieu naturel par hybridation et introgression spontanée,
- iii) Une perte de la diversité originale des formes sauvages (érosion génétique).

Par ailleurs, ils concluent que les niveaux de flux génétique et d'introgression peuvent varier selon plusieurs facteurs :

- a. selon la compatibilité génétique entre la forme cultivée et son homologue sauvage ou espèces apparentées du même genre ou genre voisins,
- b. la compatibilité génétique entre espèces pouvant varier selon les génotypes concernés,
- c. l'incompatibilité qui peut comprendre barrières prezygotiques et postzygotiques,
- d. l'existence de barrières temporelles (Ex. floraison asynchronisée) et spatiales,
- e. les conditions environnementales (vent, insectes pollinisateurs)

Nos résultats apportent des connaissances concernant le premier de ces points, et suggèrent que la propagation de transgènes peut comprendre un genre entier, comme chez le genre *Manihot* qui représenterait en fait un seul complexe d'espèces. Il conviendrait de considérer le risque d'introgression de transgènes même chez les formes sauvages les moins apparentées. Par ailleurs le cas du manioc, où l'introgression implique un fragment chromosomique lié à un caractère conférant une résistance, confirme que des portions spécifiques du génome sont plus aptes à se conserver par introgression car elles confèrent une valeur additionnelle à l'adaptation des individus introgressés, ce qui est aussi le cas de certains transgènes.

V. Implications pour la conservation des ressources génétiques de l'arracacha

Les plantes à racines et tubercules andines, dont l'arracacha fait partie, est un groupe de plantes cultivées d'une énorme importance dans la sécurité alimentaire de cette région depuis des siècles. Cependant de nos jours, plusieurs facteurs ont contribué à l'érosion génétique de ces plantes; citons parmi d'autres : une dévalorisation et remplacement par d'autres plantes (introduites notamment) et la perte des traditions de consommation et des pratiques paysannes. Dans ce contexte, le rôle des banques génétiques par la voie de la conservation *ex situ*, et plus récemment *in situ*, a été important afin de contribuer à éviter la perte de la diversité de ces cultivars traditionnels conservée au cours des siècles et qui pourrait disparaître à jamais.

Les expériences de conservation *ex situ* de germplasm de ces plantes andines, montre que les collections vivantes en champs ou par technologies *in vitro* sont des méthodes coûteuses et qui par ailleurs induisent des pertes (pression des facteurs biotiques ou abiotiques, erreurs de manipulation menant à des « mutations d'étiquettes »). L'arracacha cultivée ayant un mode de vie pérenne, sa conservation en collections vivantes en champ s'avère néanmoins moins problématique que celle de tubercules andins (*oca*, *mellico* et *mashua*) qui ont besoin d'être plantés et récoltés lors de chaque cycle de culture. Cependant des pertes considérables en champs ont été observées après une vingtaine d'années de conservation.

Pour que la conservation de germplasm puisse être durable dans le temps, la conservation par graines chez l'arracacha est une méthode à prendre en compte car la graine est orthodoxe, et se peut se conserver en chambre froide (Hermann 1997, Knudsen 2003). Toutefois certains génotypes sont récalcitrants à la floraison même en conditions expérimentales d'induction et une méthode efficace d'induction florale a été difficile à mettre en place (Knudsen 1999, Knudsen 2003). Par la suite, il est nécessaire d'étudier la ségrégation de graines pour déterminer si cette méthode est vraiment appropriée et n'induit pas de sélection sur certains génotypes.

Hermann (1997) invoque le besoin de connaître la structure génétique de la diversité de l'arracacha afin d'évaluer les effets de l'érosion génétique et affiner la gestion de ressources génétiques. Notre travail fournit des connaissances nouvelles à ce sujet en montrant que la diversité est plus importante à ce qui était présumé et que cette-ci se trouve recombiniée et structurée en deux groupes insoupçonnés de cultivars. Ces connaissances ont des applications directes en matière de conservation et utilisation des ressources génétiques de l'arracacha. Par exemple, par la suite il sera intéressant d'évaluer ces deux groupes génétiques, de les utiliser en croisements et d'évaluer des descendance pour leur potentiel agronomique. De même, les formes sauvages de *A. xanthorrhiza*, monocarpiques et pérennes sont une source d'une grande diversité qui pourra être utilisée en croisement. Il devient donc important de privilégier la conservation de ces formes sauvages tant en conditions *ex situ* (par graines) comme *in situ* avec la participation des populations locales.

De façon plus générale, notre démonstration de l'introgession des formes cultivées de manioc et arracacha par des espèces éloignées montre la possibilité que ces dernières soient des ressources génétiques utilisables en amélioration variétale. Cela renforce donc la nécessité d'intégrer ces espèces sauvages dans les stratégies de gestion des ressources génétiques de l'arracacha et du manioc.

REFERENCES BIBLIOGRAPHIQUES

- Allem A.C. 1994. The origin of *Manihot esculenta* Crantz (Euphorbiaceae). Genetic Resources and Crop Evolution 41: 133-150
- Anderson E. 1961. The analysis of variation in cultivated plants with special reference to introgression. Euphytica 10:79-86
- Arbizu C, Huaman Z and Golmirziae A. 1997. Other Andean roots and tubers. In: Fucillo, D., L. Sears, and P. Stapleton (eds.). Biodiversity in trust: Conservation and use of plant genetic resources in CGIAR centres. Cambridge University Press, Cambridge, UK. P. 39-56
- Balzegue S, Dubreucq B, Chauvin S, Le-Clainche I, Le Boulaire F, de Rose R, Samson F, Biaudet V, Lecharny A, Cruaud C, Weissenbach J, Caboche M, Lepiniec L. 2001. Improved PCR-walking for large-scale isolation of plant T-DNA borders. Biotechniques 30 (3):496-8
- Belkhir K., Goudet J., Chikhi L. and Bonhomme F. 2000. Genetix Ver. 4.01, logiciel sous windows™ pour la génétique des populations. Montpellier, France
- Billote N, Lagoda P., Risterucci A., Baurens F. 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits p. 277-288
- Blas R. 2000. Diversité génétique et éléments de physiologie chez l'arracacha (*Arracacia xanthorrhiza* Bancroft). Mémoire de DEA. Faculté de Sciences Agronomiques de Gembloux. Belgique. 79p.
- Blas R. 2005. Diversity of *Arracacia* species in Peru. Ph.D. Thesis. Gembloux Agricultural University, Belgium. 154p
- Bristol ML. 1988. Edible arracachas of the Sibundoy. Rev. Academia Colombiana de Ciencias Exactas, Físicas y Naturales 16 (63):107-110
- Bukasov SM. 1981. The cultivated plants of Mexico, Guatemala and Colombia. Bull. Appl. Bot., Genet. Plant-Breed. (Leningrad), 47th; Supplement, pp. 470-532
- Butuler L, Jarret R and LaBonte D. 1999. Sequence characterization of microsatellites in diploid and polyploid Ipomoeae. TAG 99:123-132
- Castillo R. 1997. Caracterización molecular de 29 morfotipos de arracacha (*Arracacia xanthorrhiza* Bancroft) de la colección ecuatoriana. IX Congreso Internacional de Cultivos Andinos, Univ. Nacional de San Antonio Abad del Cusco. Libro de resúmenes p.42
- Cavalli-Sforza L. and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. American Journal of Human Genetics 19:233-257
- CIP 2005. About ARTC Homepage: <http://www.cipotato.org/artc/artc.htm>
- Colombo C. 1997. Etude de la diversité génétique de manioacs américains (*Manihot esculenta* Crantz) par les marqueurs moléculaires (RAPD et AFLP). Thèse de doctorat, ENSAM, Montpellier, France, 144 p.
- Colombo C, Second G, Losada Valle T. et Charrier A. 1998. Genetic diversity characterization of cassava cultivars (*Manihot esculenta* Crantz). I) RAPD markers. Genetics and Molecular Biology 21, 11 105-113
- Constance L. 1949. The South American species of *Arracacia* (Umbelliferae) and some related genera. Bull. of the Torrey Botanical Club 76 (1):39-52
- Constance L, Tsan-Iang Cg, and R. Bell. 1976a. Chromosome numbers in umbelliferae. V. Amer. J. Bot. 63 (5): 608-625
- Constance L, T.L. Chuang and R.A. Bye. 1976b. Chromosome numbers in Chihuahuan Umbelliferae. Botanical Museum Leaflets (Harvard University) 24(8):241-247
- Constance L. and J.M. Affolter. 1995. Three new species and a new combination in *Arracacia* Bancroft (Umbelliferae/Apiaceae). Brittonia 47(3):320-327

- Chakraborty R. and Jin L. 1993. A unified approach to study hypervariable polymorphisms: statistical consideration of determining relatedness and populations distances. In: Chakraborty R. et al. (eds), DNA Fingerprinting: State of Science. Basel, Switzerland, pg. 153-175
- Demesure B., Sodji N. and R.J. Petit. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4, 129-131
- DNASTAR software ver. 5.07. 2003. Inc., Madison, WI. <http://www.dnastar.com>
- Downie S., D. Katz-Downie and M. Watson. 2000. A phylogeny of the flowering plant family Apiaceae based on chloroplast DNA rpl16 and rpo1 intron sequences: towards a suprageneric classification of subfamily Apioidea. *American Journal of Botany* 87(2): 273-292
- Elias M. 2000. Sélection naturelle, sélection humaine, et diversité chez une plante domestiquée propagée de façon végétative : le cas de la culture du manioc par les indiens Makushi du Guyana. Synthèse. Thèse Doctorat. Université Montpellier II. pp.22-25.
- Ellstrand N. C., H. C. Prentice, J. F. Hancock. 1999. Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* 30: 539-563
- Empereire E., Pinton F., Second G. 1998. Gestion dynamique de la diversité variétale du manioc en Amazonie du nord-ouest. *Nature, science et société* 6 (2) :27-42
- Erazo MC, F. Dorregaray and M. Hermann. 1996. Electroforesis de proteínas e isoenzimas en *Arracacia xanthorrhiza*, *Canna edulis* y *Oxalis tuberosa*. In. Memorias del II Congreso Ecuatoriano de Botánica, PUCE, Quito, Ecuador
- Franks T., Botta R. and Thomas M. R. 2002. Chimerism in grapevines: implications for cultivar identity, ancestry and genetic improvement. *Theoretical and Applied Genetics* 104: 192-199
- Fregene M.A., J. Vargas, F. Angel, J. Tohme, R. A. Asiedu, M. O. Akorada, and W.M. Roca. 1994. Chloroplast DNA and nuclear ribosomal DNA variability in cassava (*Manihot esculenta* Crantz) and its wild relatives. *TAG* 89 : 719-727
- Gonçalves S. 2004. Brasileira por adoção. *Revista Globo rural*. Edição 229. Editora Globo S. A. 4p. (<http://www.globorural.globo.com>)
- Gonzalez T., Chambi N. and Machaca M. 1998. Nurturing the seed in the Peruvian Andes. *Revista Seedling, Grain* Vol. 15, N°2
- Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>
- Groot M., Van de Wiel C., Tienderen P., and Nijs H. 2003. Hybridisation and introgression between crops and wild relatives. Current knowledge and research priorities in lieu of impending introductions of GM crops. University of Amsterdam and Plant Research International. Wageningen, Netherlands. 54p.
- Hahn W. 2002. A phylogenetic analysis of the arecoid line of palms based on plastid DNA sequence data. *Molecular Phylogenetics and Evolution* 23: 189-204
- Harlan J.R. 1971. Agricultural origins: Centers and noncenters. *Science* 174: 468-474
- Hartl DL, Clark AG. 1997. Principles of Population Genetics 3rd Ed, Sunderland, Massachusetts: Sinauer Associates, Inc.
- Hawkes J.G. 1989. The domestication of roots and tubers in the Andean tropics. Pp. 481-503. In Foraging and farming (Harris and Hillman eds.). Unwin Hyman, London
- Hermann M. 1997. Arracacha. In: M. Hermann and J. Heller eds., Andean roots and tubers: Ahipa, arracacha, maca and yacon. Promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome, Italy. Pg. 75-172

- Hodge W.H. 1954. The edible arracacha. A little known root crop of the Andes. *Econ. Bot.* 8:195-221
- Jhingan A. 1992. A novel technology for DNA isolation. *Methods Mol. Cell. Biol.* 3:15-22
- Jorge V., Fregene M.A., Duque M.C, Bonierbale M.W, Tohme J. and Verdier V. 2000. Genetic mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz). *Theor Appl Genet* 101:865-872
- Knudsen S.R. 1999. Flower induction in the Andean root crop arracacha (*Arracacia xanthorrhiza* Bancroft). A description and evaluation of the morphological changes following dehydration. M. Sc. Botany Thesis. KVL University, Copenhagen, Denmark. pg. 43
- Knudsen S.R. 2003. Reproduction biology of the Andean root crop arracacha (*Arracacia xanthorrhiza* Bancroft var. *xanthorrhiza*) and the taxonomic status of the South American *Arracacia* Bancroft species with special emphasis on the position of the cultivated arracacha and related wild species. Ph.D. thesis. Department of Ecology. KVL University, Denmark. Available from <http://www.ecol.kvl.dk/~srk/arracacia/side2.html>
- Landázuri P.E. 1996. Introducción in vitro, micropropagación, conservación y aclimatización de zanahoria blanca (*Arracacia xanthorrhiza*) y achira (*Canna edulis*). Tesis Lic. Ciencias Biológicas, Pontificia Universidad Católica del Ecuador, Quito, 116 p.
- Lefèvre F. 1989. Ressources génétiques et amélioration du manioc, *Manihot esculenta* Crantz, en Afrique. Paris, France, Orstom, Travaux et documents microédités No. 57, 175p.
- León J. 1964. Plantas alimenticias andinas. Instituto Interamericano de Ciencias Agrícolas Zona Andina, Lima, pp. 50-56
- Leotard G. 2003. Phylogéographie et origine de la domestication du manioc (*M. esculenta* Crantz, Euphorbiaceae): les apports d'un échantillon élargi à l'écotone nord de l'Amazonie. DEA, U.Montpellier II, France.
- Liu K. and Spencer Muse. 2001. PowerMarker: new genetic data analysis software. Version 3.23. <http://www.powermarker.net>
- Mann C. 2005. 1491: New revelations of the Americas before Columbus
- Mathias M.E. and L. Constance. 1976. *Arracacia* Bancr. Pp. 42-47 in *Flora of Ecuador*, No.5, 145. Umbelliferae (G. Harling and B. Sparre, eds). Lund, Denmark
- Mazon N.1993. Análisis de la variación morfológica e isoenzimática de la colección ecuatoriana de zanahoria blanca (*Arracacia xanthorrhiza* B.). Tesis Ing. Agr. Escuela Superior Politécnica del Chimborazo, Riobamba, Ecuador. 135 p.
- Mazon N., R. Castillo, M. Hermann and P. Espinosa. 1996. La zanahoria blanca o arracacha (*Arracacia xanthorrhiza* Bancroft) en Ecuador. Instituto Nacional Autónomo de Investigaciones Agropecuarias (INIAP), Publicación Miscelánea No. 67, Quito, Ecuador, 41 p.
- Meza G. 1995. Variedades nativas de virraca (*Arracacia xanthorrhiza* Bancroft) en Cusco. Centro de Investigaciones en Cultivos Andinos (CICA), Universidad Nacional de San Antonio Abad del Cusco, Perú, 14 p.
- Morales F. 2005. Antropología et historia de Tunja. Universidad pedagógica y tecnológica de Colombia. <http://www.monografias.com>
- Morillo E. 2001. Caractérisation des introgressions du complexe d'espèces *M. carthaginensis* S.l. incluant divers « arbres manioc » dans certaines variétés traditionnelles américaines et africaines du manioc (*M. esculenta* Crantz). Memoire DEA RPI, ENSAM, Montpellier, France
- Morillo E., Fuenmayor F., De Carvalho P. et Second G. 2001. AFLP and SSR polymorphisms evidence levels of introgression from *Manihot glaziovii* and *M. carthaginensis* into traditional varieties of cassava in their area of origin. In. Fauquet CM and Taylor NJ, eds.

- Proceedings of the 5th International Meeting of the Cassava Biotechnology Network, St Louis Mo. USA. [Compact Disc]
- Nassar N. 2003. Gene flow between cassava, *Manihot esculenta* Crantz, and wild relatives. *Genet. Mol. Res.* 2(4):334-347
- Nei M. 1972. Genetic distance between populations. *American Naturalist* 106: 283-392
- Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590
- NRC (National Research Council). 1989. Lost Crops of the Incas: Little known Plants of the Andes with Promise for Worldwide Cultivation. National Academy Press, Washington, D.C. pg. 1, pg. 47
- Olsen K. and Schaal B. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proc Natl. Acad. Sci. USA* 96: 5586-5591
- Olsen K. 2002. Population history of *Manihot esculenta* (Euphorbiaceae) inferred from nuclear DNA sequences. *Molecular Ecology* 11: 901-911
- Orloci L. 1978. Multivariate analysis in vegetation research. The Hague: Dr W. Junk B. V.
- Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12: 357-358
- Papa R. and P. Gepts. 2004. Asymmetric gene flow and introgression between wild and domesticated populations. In: Nijs, J. C. M., Bartsch, D. and Sweet, J. editors. *Introgression from Genetically Modified Plants into wild relatives and its consequences*, CABI Publishing, Oxon, UK
- Patiño V. M. 1964. Umbelíferas. Pp. 57-62 In: *Plantas cultivadas y animales domésticos en América Equinoccial*, tomo II (primera edición): Plantas alimenticias. Imprenta Departamental, Cali, Colombia
- Peakall R. and Smousse P.E. 2001. GenAlEx V5: Genetic Analysis in Excel. <http://www.anu.edu.au/BoZo/GenAlEx/>
- Pemberton J., Slate, D., Bancroft, D.R and Barret, J.A. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and populations studies. *Molecular Ecology* 4: 249-252
- Pimenov M.G. and Leonov M.V. 1993. The genera of Umbelliferae. Royal Botanical Gardens, Kew. Richmond, Surrey, UK
- Piperno D. R. and Stothert, K. 2003. Phytolith evidence for early Cucurbita domestication in southwest Ecuador. *Science* 299: 1054
- Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of populations structure using multilocus genotype data. *Genetics* 155: 945-959
- Pritchard J. and Wen William. 2004. Documentation for Structure software: version 2. Department of Human Genetics, Chicago, USA. 33p. <http://pritch.bsd.uchicago.edu>
- Quiros C. F. R., Ortega L., Van Raamsdonk M., E. Herrera, P. Cisneros, E. Schmidt and S. B. Brush. 1992. Increase of potato genetic resources in their center of diversity: the role of natural outcrossing and selection by the Andean farmer. *Genetic Resources and Crop Evolution* 39: 107-113
- Rabinowitz D., Linder CR. and Ortega R. 1990. High levels of interspecific hybridization between *Solanum sparsipilum* and *S. stenotomum* in experimental plots in the Andes. *American Potato Journal* 67: 73 81
- Raymond M. and Rousset F. 1995. GENEPOP Ver. 0.2. Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248-249
- Rieseberg L.H., Sinervo B., Linder C.R., Ungerer M.C. and Arias D.M. 1998. Role of gene interaction in hybrid speciation: evidence from ancient and experimental hybrids. *Science* 272:741-745

- Rohlf F. J. 2002. NTSYSpc: Numerical Taxonomy System, ver. 2.1. Exeter Publishing, Ltd.: Setauket, NY. Sokal, R.R and P.H.A. Sneath. 1963. Principles of Numerical Taxonomy. Freeman San Francisco. 359 p.
- Rogers D.J., Appan M., 1973. Manihot, Manihotoides (Euphorbiaceae). New York, Etats-Unis, Hafner Press, Flora Neotropica Monograph n°13, 274 p.
- Rosso C.A., Medina C. and Lobo M. 2002. Morphological characterization and agronomic evaluation of a Colombian collection of arracacha (*Arracacia xanthorrhiza* Bancroft). Plant Genetic Resources Newsletter. No. 132: 22-29
- Rubatzky V.E., C.F. Quiros and P.W. Simon. 1999. Carrots and related vegetable Umbelliferae. CABI Publishing. New York, NY. 294 p.
- Salazar C. 1997. Evaluacion y caracterizacion citogenetica de veinte entradas ecuatorianas de zanahoria blanca (*Arracacia xanthorrhiza* Banc.). Tesis Ing. Agr. Universidad Central, Quito, Ecuador. p. 56
- Santoni S., Faivre-Rampant, Prado E. et Prat D. 2000. Marqueurs moléculaires pour l'analyse des ressources génétiques et l'amélioration des plantes. Cahiers Agriculture 9: 311-327
- Scarcelli N. 2005. Structure et dynamique de la diversité d'une plante cultivée a multiplication végétative: le cas des ignames au Bénin (*Dioscorea* sp.) Thèse de doctorat, ENSAM, Montpellier, France, 175 p.
- Saitou N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425
- Santos FF, Vieira J.V, Reis N.V.B. 1990. Ocorrencia de germinação de sementes de mandioquinha-salsa (*Arracacia xanthorrhiza* Bancroft) a nivel de campo, em Brasilia. Horticultura Brasileira, Brasilia, v8, n1, pg 65,. Resumo
- Second G., Allem A., Mendes R., Carvalho L., Emperaire L., Ingram C. and Colombo C. 1997. Molecular Marker (AFLP)-based Manihot and cassava numerical taxonomy and genetic structure analysis in progress: Implications for their dynamic conservation and genetic mapping. Genome. Vol. 2: 140-147
- Second G. 1998. *Manihot glaziovii* contributed to the genetic make-up of cassava and represents an example of dinamic conservation and on-farm breeding of genetic resources. In Proceedings of IV International Scientific Meeting of Cassava Biotechnology Network
- Second G., Raffailac J. P. and C. Colombo. 1999. Manioc. In: Diversité génétique des plantes cultivées tropicales. Eds: Hamon, P. M. Seguin; X. Perrier; J. C. Glaszmann. Collection Repères. CIRAD. 385p.
- Second G. 2000. *Manihot glaziovii* contributed to the genetic make-up of cassava and represents an example of dinamic conservation and on-farm breeding of genetic resources. In Proceedings of IV International Scientific Meeting of Cassava Biotechnology Network
- Second G. and Iglesias C. 2001. The state of use of cassava genetic diversity and a proposal to enhance it. In Broadening the Genetic Basis of Crop Production (eds. Cooper HD, Spillane D, Hodgkin T., pg. 201-221
- Swofford D. 2001. PAUP software. Phylogenetic analysis using parsimony. Version 4b8. Sinaur, Sunderland, MA
- Valderrama M. and J. Seminario. 2002. Los parientes silvestres de la arracacha (*Arracacia xanthorrhiza* Bancroft) y su uso en medicina tradicional, en el norte peruano. Arnaldoa, Trujillo, Peru. 9(1):67-91
- Vasconez J.J y Hermann M. 1997. Experimento de Polinización Interespecífica en Arracacia. CIP Quito. Datos no publicados.
- Vasquez N. Medina C. y Lobo M. 2003. Caracterización morfológica de la colección colombiana (Tolima, Huila, Boyaca, Cauca) de arracacha (*Arracacia xanthorrhiza*). In. J. Seminario (ed.). Raíces Andinas: Contribuciones al conocimiento y a la capacitación: Una

- década de investigación para el desarrollo (1993-2003) No. 6. UNC, CIP, COSUDE. Lima, Perú. pg. 166-178
- Vos P., Hogers R., Bleeker M., Reijans M., T van de Lee, Hornes M, Frijters A., Pot J., Peleman J., and Kuiper M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 11; 2321: 4407–4414
- Weir B.S., Cockerham C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370
- Zohary D. 2004. Unconscious selection and the evolution of domesticated plants. *Economic Botany* 58 (1) pg. 5-10

TABLEAUX

Tableau 1. Principales plantes domestiquées dans la région andine (Source: NRC 1989)

TYPE DE PLANTE	NOM SCIENTIFIQUE	ANGLAIS	ESPAGNOL	QUECHUA (LANGUE DES INCAS)
Pseudo-céréales	<i>Chenopodium quinoa</i>	quinoa	quinua	kiuña
	<i>C. pallidicaule</i>	-	cañahua	qanawi
	<i>Amaranthus caudatus</i>	amaranth	achita	achita, achis, coimi, okiru, quihuicha
Légumineuses	<i>Phaseolus vulgaris</i>	common bean	frijol, poroto	purutu
	<i>Canavalia ensiformis</i>	jack bean		
	<i>Arachis hypogea</i>	peanut, groundnut	maní	inchis
	<i>Lupinus mutabilis</i>	lupin	chocho	tarwi
Plantes tubéreuses	<i>Xantoma</i> spp.		unchuca	uncucha
	<i>Canna edulis</i>		achira	achira
	<i>Pachyrrhizus</i> spp.		jíquama	ajipa
	<i>Oxalis tuberosa</i>		oca	oqa
	<i>Tropaeolum tuberosum</i>		mashua	allausu, añu
	<i>Arracacia xanthorrhizus</i>		arracacha	raccacha
	<i>Manihot esculenta</i>	cassava	yuca	ruma
	<i>Solanum</i> spp.	potato	papa	apichu, kumara
	<i>Ullucus tuberosus</i>		ullucu, papalisa	ulluku
	<i>Polymnia sonchifolia</i>		yacón	yakon, yakuma
Stimulant	<i>Erythroxylon coca</i>	coca	coca	coca, cuca

Tableau 2. Quelques caractéristiques des plantes andines à racines et tubercules (Source: Andean roots and tubers at the crossroads, M. Herman and J. Heller, Institute of Plant Genetics and Crop Plant Research, IPK).

Nom commun	Production (T/ha)	Temps de culture (mois)	Utilisateurs dans les Andes (millions)	Utilisateurs hors les Andes (millions)
<i>Climat tempéré (2500-4000m attitude)</i>				
Oca	10-40	6-9	15	<2
Ulluco	5-20	6-9	30	-
Mashua	30-60	6-9	<10	-
<i>Climat subtropical (1000-2500m attitude)</i>				
Arracacha	12-16	10-14	30	>30
Yacon	40-60	10-12	<1	No data
Achira	30-80	10-12	<1	>50
Mauka	20-50	12-18	<1	-
Ahipa	30-50	5-10	<1	-
<i>Puna (4000-4500m altitude)</i>				
Maca	10-15	8-9	<1	-

Tableau 3. Liste des 102 allèles observés parmi 178 formes cultivées (73) et sauvages (105) de *A. xanthorrhiza* sur 11 locus SSR ne présentant pas plus de deux allèles chez un même individu. Les allèles présents chez la forme cultivée (31) sont surlignés en gris.

Locus	D85	D34	C85	D55A	D55B	D72	D82	C27	D43	C38	D13
No. allèles	8	9	14	10	8	14	11	5	12	6	5
1	180 ¹	206 ⁹	266 ¹⁸	400 ³²	252 ⁴²	166 ⁵⁰	170 ⁶⁴	144 ⁷⁵	184 ⁸⁰	106 ⁹²	182 ⁹⁸
2	184 ²	208 ¹⁰	268 ¹⁹	414 ³³	258 ⁴³	168 ⁵¹	172 ⁶⁵	148 ⁷⁶	186 ⁸¹	108 ⁹³	184 ⁹⁹
3	186 ³	210 ¹¹	280 ²⁰	416 ³⁴	262 ⁴⁴	170 ⁵²	174 ⁶⁶	150 ⁷⁷	188 ⁸²	110 ⁹⁴	186 ¹⁰⁰
4	188 ⁴	212 ¹²	286 ²¹	418 ³⁵	272 ⁴⁵	172 ⁵³	176 ⁶⁷	152 ⁷⁸	190 ⁸³	112 ⁹⁵	188 ¹⁰¹
5	190 ⁵	214 ¹³	284 ²²	420 ³⁶	274 ⁴⁶	174 ⁵⁴	178 ⁶⁸	154 ⁷⁹	192 ⁸⁴	114 ⁹⁶	192 ¹⁰²
6	192 ⁶	216 ¹⁴	286 ²³	422 ³⁷	276 ⁴⁷	176 ⁵⁵	180 ⁶⁹		194 ⁸⁵	116 ⁹⁷	
7	196 ⁷	218 ¹⁵	288 ²⁴	424 ³⁸	278 ⁴⁸	178 ⁵⁶	182 ⁷⁰		196 ⁸⁶		
8	198 ⁸	220 ¹⁶	290 ²⁵	426 ³⁹	280 ⁴⁹	180 ⁵⁷	184 ⁷¹		198 ⁸⁷		
9		222 ¹⁷	292 ²⁶	428 ⁴⁰		182 ⁵⁸	186 ⁷²		200 ⁸⁸		
10			294 ²⁷	434 ⁴¹		184 ⁵⁹	188 ⁷³		202 ⁸⁹		
11			296 ²⁸			186 ⁶⁰	190 ⁷⁴		206 ⁹⁰		
12			298 ²⁹			190 ⁶¹			208 ⁹¹		
13			302 ³⁰			200 ⁶²					
14			304 ³¹			206 ⁶³					

Tableau 4. Statistiques générales de la diversité SSR au sein du complexe sauvage et cultivé d'*A. xanthorrhiza* (N = 178) PIC : *Polymorphism Index Content*

Locus	Fréquence Maximale d'un allèle	Nombre de génotypes	Hétérozygotie Observé	PIC
AxD85	0.48	19	0.57	0.66
AxD34	0.36	27	0.58	0.75
AxC85	0.43	35	0.48	0.72
AxD55A	0.47	18	0.31	0.68
AxD55B	0.37	15	0.48	0.67
AxD72	0.60	29	0.43	0.59
AxD82	0.21	49	0.80	0.86
AxC27	0.46	9	0.35	0.59
AxD43	0.39	29	0.36	0.74
AxC38	0.38	15	0.22	0.72
AxD13	0.77	10	0.08	0.36
Moyenne	0.45	23.18	0.42	0.67

Tableau 5. Fréquence de 31 allèles partagés par les formes sauvages et cultivées d'*A. xanthorrhiza*. On précise la fréquence relative de ces allèles au sein de chaque compartiment (cultivée et sauvage) ainsi que dans le complexe *A. xanthorrhiza* (Ax).

LOCUS	ALLELE	Cultivée	Sauvage	Complexe Ax
AxD85	180	0.321	0.046	0.179
	190	0.400	0.613	0.510
	192	0.278	0.166	0.220
AxD34	206	0.006	0.088	0.054
	208	0.349	0.088	0.197
	214	0.630	0.181	0.368
AxC85	266	0.464	0.051	0.237
	292	0.007	0.040	0.025
	294	0.528	0.431	0.474
AxD55A	422	0.291	0.194	0.237
	426	0.708	0.355	0.512
AxD55B	258	0.558	0.260	0.385
	276	0.441	0.361	0.395
AxD72	168	0.013	0.014	0.014
	170	0.006	0.053	0.034
	172	0.790	0.470	0.602
	180	0.171	0.029	0.088
	184	0.010	0.014	0.017
AxD82	170	0.267	0.053	0.142
	174	0.280	0.165	0.213
	182	0.404	0.043	0.193
	184	0.040	0.087	0.065
AxC27	148	0.464	0.053	0.222
	150	0.091	0.730	0.468
	152	0.443	0.196	0.297
AxD43	184	0.006	0.005	0.006
	196	0.076	0.312	0.211
	206	0.910	0.036	0.473
AxC38	106	0.826	0.115	0.456
	108	0.173	0.275	0.226
AxD13	186	1.000	0.786	0.886

Tableau 6. Fréquences des allèles communs entre formes cultivées (FC) et sauvages (FS) d'*A. xanthorrhiza* regroupées selon leur origine géographique: Centre Equateur (Provinces de Chimborazo, Cañar et Bolivar), Sud de l'Equateur (province de Loja) et Nord du Pérou (Département de Cajamarca). L'absence des certains allèles est surligné en gris.

LOCUS	ALLELE	FC (N = 73)	FS		
			Centre Equateur (N = 23)	Sud Equateur (N = 61)	Nord Pérou (N = 23)
AxD85	180	0.321	0.333	0.030	0.050
	190	0.400	0.000	0.642	0.625
	192	0.278	0.000	0.173	0.150
AxD34	206	0.006	0.027	0.086	0.142
	208	0.349	0.111	0.051	0.142
	214	0.630	0.194	0.155	0.166
AxC85	266	0.464	0.166	0.022	0.025
	292	0.007	0.000	0.000	0.125
	294	0.528	0.305	0.488	0.400
AxD55-A	422	0.291	0.352	0.090	0.357
	426	0.708	0.058	0.454	0.321
AxD55-B	258	0.558	0.421	0.060	0.590
	276	0.441	0.105	0.550	0.227
AxD72	168	0.013	0.000	0.017	0.022
	170	0.006	0.055	0.017	0.136
	172	0.787	0.805	0.465	0.295
	180	0.171	0.000	0.008	0.113
	184	0.020	0.000	0.025	0.000
AxD82	170	0.267	0.103	0.025	0.023
	174	0.280	0.210	0.112	0.190
	182	0.404	0.000	0.025	0.142
	184	0.034	0.000	0.112	0.095
AxC27	144	0.464	0.000	0.025	0.000
	150	0.091	0.210	0.879	0.761
	152	0.443	0.657	0.094	0.119
AxD43	184	0.006	0.000	0.009	0.000
	196	0.076	0.375	0.305	0.261
	206	0.916	0.031	0.000	0.095
AxC38	106	0.826	0.055	0.052	0.277
	108	0.173	0.111	0.250	0.472
AxD13	186	1.000	1.000	0.802	0.738

Tableau 7. Valeurs observées (Ho) et attendues (He) d'hétérozygotie sur 11 loci SSR chez 178 formes cultivées (FC) et sauvages (FS) d'arracacha. La signification statistique est indiquée par *** (p< 0.001), ** (p< 0.01), * (p< 0.05)

LOCUS	FC (N=73)			FS (N=105)		
	He	Ho	Fis	He	Ho	Fis
AxC27	0.57	0.56	0.024	0.42	0.21	0.470**
AxC38	0.28	0.29	-0.015	0.71	0.23	0.691***
AxC85	0.50	0.53	-0.057	0.77	0.51	0.360***
AxD13	0.00	0.00	--	0.36	0.17	0.578**
AxD34	0.48	0.54	-0.132	0.80	0.62	0.218**
AxD43	0.15	0.16	-0.061	0.78	0.53	0.324**
AxD72	0.34	0.39	-0.122	0.71	0.45	0.357***
AxD82	0.68	0.82	-0.215 **	0.88	0.78	0.117*
AxD85	0.65	0.77	-0.154	0.58	0.45	0.216**
AxD55A	0.41	0.38	0.060	0.78	0.30	0.620***
AxD55B	0.49	0.67	-0.363**	0.73	0.40	0.442***
Moyenne	0.41	0.46	-0.09	0.69	0.43	0.40

Tableau 8. Valeurs d'appartenance (q) des individus assignés par STRUCTURE aux 2 (K=2) ou 3 (K=3) groupes supposés, selon les données génotypiques sur 11 loci SSR (sans assignation a priori des individus identifiés comme à l'annexe 1). Les valeurs surlignées en gris déterminent leur assignation en cultivé (C), *A. x. var. andina* (A) ou *A. x. var. monocarpa* (M).

Ind.	K=2		K=3			Ind.	K=2		K=3			Ind.	K=2		K=3		
	C	S	C	A	M		C	S	C	A	M		C	S	C	A	M
C1	0.996	0.004	0.994	0.004	0.002	C71	0.997	0.003	0.995	0.002	0.002	S28	0.008	0.992	0.004	0.989	0.007
C2	0.996	0.004	0.994	0.004	0.003	C73	0.996	0.004	0.993	0.004	0.003	S29	0.033	0.967	0.015	0.967	0.018
C3	0.996	0.004	0.994	0.003	0.003	S1-1	0.006	0.994	0.004	0.022	0.974	S30	0.012	0.988	0.006	0.961	0.033
C4	0.977	0.023	0.966	0.019	0.015	S1-2	0.005	0.995	0.004	0.007	0.989	S34	0.007	0.993	0.004	0.946	0.049
C5	0.996	0.004	0.992	0.005	0.003	S1-3	0.005	0.995	0.004	0.008	0.988	S35	0.033	0.967	0.008	0.968	0.024
C6	0.996	0.004	0.994	0.003	0.002	S1-4	0.005	0.995	0.004	0.024	0.973	S36	0.243	0.757	0.220	0.773	0.007
C7	0.995	0.005	0.992	0.004	0.004	S1-5	0.003	0.997	0.003	0.005	0.993	S37	0.292	0.708	0.215	0.782	0.003
C8	0.974	0.026	0.983	0.009	0.008	S1-6	0.006	0.994	0.004	0.004	0.992	S39-1	0.006	0.994	0.003	0.005	0.991
C9	0.997	0.003	0.995	0.003	0.002	S1-7	0.009	0.991	0.019	0.008	0.973	S39-2	0.009	0.991	0.005	0.021	0.974
C10	0.995	0.005	0.992	0.004	0.003	S2	0.019	0.981	0.006	0.950	0.043	S39-3	0.005	0.995	0.003	0.007	0.989
C11	0.997	0.003	0.995	0.003	0.002	S3	0.138	0.862	0.017	0.934	0.049	S39-4	0.007	0.993	0.005	0.008	0.987
C12	0.996	0.004	0.994	0.003	0.002	S4-1	0.056	0.944	0.074	0.703	0.223	S39-5	0.009	0.991	0.005	0.013	0.982
C13	0.995	0.005	0.994	0.003	0.003	S4-2	0.025	0.975	0.012	0.680	0.308	S39-6	0.009	0.991	0.005	0.019	0.975
C14	0.996	0.004	0.994	0.003	0.003	S4-3	0.022	0.978	0.014	0.843	0.143	S39-7	0.007	0.993	0.005	0.029	0.966
C15	0.995	0.005	0.991	0.005	0.004	S5-1	0.016	0.984	0.008	0.984	0.008	S40-1	0.006	0.994	0.004	0.004	0.992
C16	0.987	0.013	0.959	0.023	0.017	S5-2	0.015	0.985	0.005	0.990	0.005	S40-2	0.004	0.996	0.003	0.004	0.994
C17	0.997	0.003	0.995	0.003	0.002	S5-3	0.041	0.959	0.008	0.984	0.008	S40-3	0.008	0.992	0.005	0.009	0.986
C18	0.987	0.013	0.974	0.012	0.014	S6-1	0.019	0.981	0.008	0.983	0.008	S40-4	0.005	0.995	0.003	0.005	0.992
C19	0.988	0.012	0.975	0.011	0.014	S6-2	0.012	0.988	0.005	0.985	0.009	S40-5	0.004	0.996	0.003	0.004	0.994
C20	0.996	0.004	0.994	0.004	0.003	S6-3	0.057	0.943	0.011	0.980	0.009	S40-6	0.004	0.996	0.003	0.004	0.993
C21	0.996	0.004	0.994	0.003	0.003	S6-4	0.009	0.991	0.004	0.991	0.005	S40-7	0.005	0.995	0.003	0.005	0.992
C22	0.996	0.004	0.994	0.003	0.002	S6-5	0.028	0.972	0.01	0.983	0.008	S40-8	0.004	0.996	0.003	0.007	0.990
C23	0.997	0.003	0.995	0.003	0.002	S7	0.077	0.923	0.012	0.983	0.005	S41	0.007	0.993	0.005	0.024	0.971
C24	0.990	0.010	0.977	0.01	0.012	S8	0.012	0.988	0.004	0.990	0.006	S42	0.100	0.900	0.072	0.061	0.867
C25	0.997	0.003	0.995	0.003	0.002	S9-1	0.006	0.994	0.003	0.992	0.004	S43	0.015	0.985	0.006	0.659	0.335
C26	0.997	0.003	0.995	0.003	0.002	S9-2	0.006	0.994	0.003	0.991	0.005	S44	0.035	0.965	0.009	0.846	0.146
C27	0.996	0.004	0.994	0.003	0.003	S10-1	0.107	0.893	0.054	0.921	0.025	S45	0.008	0.992	0.01	0.011	0.979
C28	0.983	0.017	0.964	0.011	0.025	S10-2	0.137	0.863	0.05	0.937	0.013	S47	0.012	0.988	0.007	0.053	0.940
C29	0.996	0.004	0.994	0.003	0.002	S11-1	0.004	0.996	0.006	0.071	0.924	S48-1	0.005	0.995	0.005	0.057	0.938
C30	0.995	0.005	0.993	0.004	0.003	S11-2	0.014	0.986	0.014	0.018	0.968	S48-2	0.005	0.995	0.005	0.057	0.939
C31	0.996	0.004	0.994	0.003	0.002	S12-1	0.004	0.996	0.003	0.017	0.980	H1-1	0.806	0.194	0.393	0.603	0.004
C32	0.996	0.004	0.993	0.003	0.003	S12-2	0.004	0.996	0.003	0.006	0.991	H1-2	0.809	0.191	0.348	0.652	0.007
C33	0.996	0.004	0.994	0.004	0.002	S12-3	0.007	0.993	0.005	0.008	0.987	H1-3	0.804	0.196	0.223	0.717	0.005
C34	0.996	0.004	0.994	0.003	0.003	S12-4	0.006	0.994	0.004	0.01	0.986	H1-4	0.793	0.207	0.336	0.659	0.005
C35	0.996	0.004	0.993	0.004	0.002	S12-5	0.004	0.996	0.003	0.005	0.992	H1-5	0.810	0.190	0.295	0.705	0.007
C36	0.996	0.004	0.994	0.003	0.003	S12-6	0.005	0.995	0.006	0.066	0.927	H1-6	0.826	0.174	0.305	0.695	0.009
C37	0.996	0.004	0.994	0.004	0.003	S12-7	0.022	0.978	0.001	0.075	0.909	H1-7	0.994	0.006	0.289	0.701	0.010
C38	0.994	0.006	0.992	0.004	0.005	S12-8	0.007	0.993	0.004	0.006	0.990	H1-8	0.784	0.216	0.346	0.647	0.006
C39	0.996	0.004	0.994	0.003	0.003	S13-1	0.011	0.989	0.014	0.015	0.971	H2-1	0.938	0.062	0.217	0.783	0.009
C40	0.997	0.003	0.995	0.003	0.002	S13-2	0.006	0.994	0.004	0.01	0.986	H2-2	0.964	0.036	0.235	0.765	0.005
C41	0.997	0.003	0.995	0.003	0.002	S13-3	0.006	0.994	0.010	0.032	0.958	H2-3	0.144	0.856	0.221	0.769	0.010
C42	0.997	0.003	0.995	0.003	0.002	S13-4	0.006	0.994	0.004	0.007	0.989	H2-4	0.837	0.163	0.211	0.789	0.012
C43	0.997	0.003	0.995	0.003	0.002	S13-5	0.009	0.991	0.006	0.113	0.881	H2-5	0.833	0.167	0.206	0.779	0.015
C44	0.997	0.003	0.995	0.003	0.002	S13-6	0.013	0.987	0.015	0.016	0.970						
C45	0.997	0.003	0.995	0.003	0.002	S14-1	0.005	0.995	0.003	0.005	0.991						
C46	0.997	0.003	0.995	0.003	0.002	S14-2	0.006	0.994	0.003	0.008	0.988						
C47	0.997	0.003	0.995	0.003	0.002	S14-3	0.004	0.996	0.003	0.004	0.993						
C48	0.996	0.004	0.994	0.003	0.003	S14-4	0.005	0.995	0.003	0.005	0.993						
C49	0.996	0.004	0.994	0.003	0.003	S14-5	0.006	0.994	0.009	0.008	0.982						
C50	0.997	0.003	0.995	0.003	0.002	S14-6	0.004	0.996	0.003	0.005	0.993						
C51	0.995	0.005	0.992	0.004	0.003	S15-1	0.005	0.995	0.003	0.006	0.990						
C52	0.982	0.018	0.972	0.018	0.018	S15-2	0.004	0.996	0.003	0.007	0.990						
C53	0.996	0.004	0.994	0.003	0.003	S15-3	0.006	0.994	0.004	0.051	0.946						
C54	0.997	0.003	0.994	0.003	0.003	S15-4	0.006	0.994	0.004	0.006	0.990						
C55	0.996	0.004	0.992	0.005	0.003	S15-5	0.006	0.994	0.004	0.008	0.988						
C57	0.995	0.005	0.991	0.005	0.004	S15-6	0.004	0.996	0.003	0.004	0.992						
C58	0.996	0.004	0.990	0.007	0.003	S16-1	0.022	0.978	0.015	0.528	0.457						
C59	0.996	0.004	0.992	0.004	0.003	S16-2	0.057	0.943	0.029	0.621	0.379						
C60	0.893	0.107	0.792	0.198	0.012	S17	0.013	0.987	0.008	0.010	0.982						
C61	0.995	0.005	0.993	0.004	0.003	S18	0.419	0.581	0.296	0.697	0.008						
C62	0.995	0.005	0.993	0.003	0.003	S19	0.016	0.984	0.005	0.978	0.017						
C63	0.994	0.006	0.99	0.006	0.005	S20	0.165	0.835	0.04	0.944	0.017						
C64	0.996	0.004	0.993	0.004	0.003	S21	0.015	0.985	0.004	0.988	0.008						
C65	0.996	0.004	0.992	0.004	0.003	S22	0.017	0.983	0.008	0.981	0.011						
C66	0.995	0.005	0.984	0.016	0.006	S23	0.135	0.865	0.012	0.982	0.006						
C67	0.995	0.005	0.993	0.004	0.003	S24	0.008	0.992	0.004	0.99	0.006						
C68	0.996	0.004	0.99	0.005	0.005	S25	0.019	0.981	0.011	0.945	0.045						
C69	0.995	0.005	0.992	0.004	0.004	S26	0.749	0.251	0.512	0.476	0.013						
C70	0.996	0.004	0.995	0.002	0.003	S27	0.009	0.991	0.004	0.988	0.008						

Table I-1. Plant material used in the AFLP analysis

ID No.	Acc. code	Taxa	Country	Locality	Latitude	Longitude	Alt.(m)
C1	ECU-1148	<i>A. x. var. xanth</i>	ECU	Pichincha, Palmito	N 00.11	W 78.40	1500
C2	ECU-1154	<i>A. x. var. xanth</i>	ECU	Bolívar, Negroyaco	S 01.34	W 79.00	2760
C4	ECU-1157	<i>A. x. var. xanth</i>	ECU	Cañar, Biblian	S 02.42	W 78.52	2700
C5	ECU-1158	<i>A. x. var. xanth</i>	ECU	Cañar, Biblian	S 02.42	W 78.56	2800
C6	ECU-1161	<i>A. x. var. xanth</i>	ECU	Imbabura, Sta Clara	N 00.20	W 78.14	2400
C7	ECU-1162	<i>A. x. var. xanth</i>	ECU	Cañar, Biblian	S 02.42	W 78.56	2750
C9	ECU-1173	<i>A. x. var. xanth</i>	ECU	Azuay, Cofrada	S 03.08	W 79.08	2300
C12	ECU-1179	<i>A. x. var. xanth</i>	ECU	Azuay, Giron	S 03.11	W 79.10	2110
C13	ECU-1180	<i>A. x. var. xanth</i>	ECU	Pichincha, Palogillo	S 00.13	W 78.20	2575
C16	ECU-1188	<i>A. x. var. xanth</i>	ECU	Chimborazo, Sasapui	S 01.43	W 78.35	2990
C17	ECU-1193	<i>A. x. var. xanth</i>	ECU	Pichincha, Cayambe	S 03.00	W 78.16	3400
C27	ECU-1226	<i>A. x. var. xanth</i>	ECU	Azuay, Sigsig	S 03.03	W 78.48	2600
C28	ECU-1229	<i>A. x. var. xanth</i>	ECU	Cotopaxi, Fliguilo	S 00.42	W 78.52	2700
C32	ECU-6660	<i>A. x. var. xanth</i>	ECU	Loja, Shalshi	N.A.	N.A.	2400
C35	ECU-8758	<i>A. x. var. xanth</i>	ECU	Bolívar	S 01.35	W 79.01	2900
C37	JEE-041	<i>A. x. var. xanth</i>	ECU	N.A.	N.A.	N.A.	N.A.
C46	MST-020	<i>A. x. var. xanth</i>	ECU	Zamora, Zumba	S 04.51.599	W 79.08.104	950
C52	MST-030	<i>A. x. var. xanth</i>	ECU	Zamora, Timbara	S 04.02.618	W 78.53.656	2000
C55	MST-047	<i>A. x. var. xanth</i>	ECU	Pichincha, SJ. Minas	N 00.05.582	W 72.24.156	1960
C62	A	<i>A. x. var. xanth</i>	PER	Cajamarca, Minas	S 06.17.10	W 78.41.05	2200
C63	NA	<i>A. x. var. xanth</i>	PER	Cajamarca, Minas	S 06.16.53	W 78.41.25	2470
C67	UNC-120	<i>A. x. var. xanth</i>	PER	Cajamarca	N.A.	N.A.	N.A.
C68	UNC-171	<i>A. x. var. xanth</i>	PER	Cajamarca	N.A.	N.A.	N.A.
C69	UNC-123	<i>A. x. var. xanth</i>	PER	Caj., Huañambra	S 06.53.05	W 78 09 54	2780
C74	MTC 7-8	<i>A. x. var. xanth</i>	PER	Cajamarca, Minas	S 06.16.53	W 78 41 25	2470
C75	ECU-1196	<i>A. x. var. xanth</i>	ECU	Loja, Celica	S 04.06	W 79 57	1990
C76	ECU-3295	<i>A. x. var. xanth</i>	ECU	Loja	S 03.59	W 79 11	2060
S1	MST-006	<i>A. x. var. andina</i>	ECU	Loja, Saraguro	S 03.36.908	W 79 14 155	2400
S2	MST-011	<i>A. x. var. andina</i>	ECU	Loja, Gualel	S 03.46.240	W 79 22 913	2000
S3	MST-025	<i>A. x. var. andina</i>	ECU	Loja, Loja	S 04.20.741	W 79 10 927	2600
S4	MST-039	<i>A. x. var. andina?</i>	ECU	Cañar, Tambo	S 02.28.676	W 79 00 156	2950
S5	MST-041	<i>A. x. var. andina</i>	ECU	Cañar, Alausí	S 02.11.740	W 78 50 692	2400
S6*	MST-042	<i>A. x. var. andina</i>	ECU	Cañar, Huigra	S 02.14.205	W 78 54 315	2330
S7	MST-044	<i>A. x. var. andina</i>	ECU	Tung., Ambato	N.A.	N.A.	N.A.
S8	MST-045	<i>A. x. var. andina</i>	ECU	Cotopaxi, Pujilí	N.A.	N.A.	N.A.
S9	MST-057	<i>A. x. var. andina</i>	ECU	Bol., Jatunpamba	S 01.32.372	W 78 57 900	3050
S11	MST-070	<i>A. x. var. andina</i>	ECU	Loja, Oña	N.A.	N.A.	2900
S12	MST-072	<i>A. x. var. andina</i>	ECU	Loja, Saraguro	N.A.	N.A.	2500
S13	MST-074	<i>A. x. var. monoc</i>	ECU	Loja, Manu	N.A.	N.A.	2350
S14	MST-075	<i>A. x. var. monoc</i>	ECU	Loja, Loja	N.A.	N.A.	2600
S15	MST-076	<i>A. x. var. monoc</i>	ECU	Loja, Payanchi	S 04.00.532	W 79.15.459	2600
S16	MST-083	<i>A. x. var. andina</i>	ECU	Loja, Calvas	S 04.18.739	W 79.31.554	2000
S17	MST-085	<i>A. x. var. andina</i>	ECU	Loja Calvas	S 04.20.179	W 79.36.723	2050
S39*	MST-081	<i>A. x. var. monoc</i>	ECU	Loja, Gimbura	S 04.35.566	W 79.26.566	1600
S40	MST-082	<i>A. x. var. monoc</i>	ECU	Loja, Gimbura	S 04.37.483	W 79.27.611	2000
S42	SRK-183	<i>A. x. var. monoc</i>	PER	Cajamarca	S 06.17.680	W 78.51.300	N.A.
S43	SRK-500	<i>A. x. var. monoc</i>	PER	Cajamarca	S 06.18.940	W 78.41.030	N.A.
S44	SRK-510	<i>A. x. var. monoc</i>	PER	Cajamarca	S 06.17.060	W 78.41.430	N.A.
S45	75-1	<i>A. x. var. monoc</i>	PER	Cuzco	S 13.24.490	W 71.50.650	N.A.
S47	31-2	<i>A. x. var. monoc</i>	PER	Amazonas	S 06.43.160	W 77.51.900	N.A.
S48*	RBS-001	<i>A. incisa</i>	PER		N.A.	N.A.	N.A.
S49	MST-015	<i>A. acuminata</i>	ECU	Loja, Yangana	S 04.28.490	W 79.08.897	2520
S51	MST-036	<i>A. elata</i>	ECU	Azuay, Cuenca	S 03.41.801	W 78.53.685	3050
S53	MST-050	<i>A. acuminata</i>	ECU	Imbabura, Otavalo	N 00.12.314	W 78.21.377	3000
S54	MST-051	<i>A. moschata</i>	ECU	Carchi, El Ángel	N 00.38.998	W 77.53.404	3400

(ECU: Ecuador, PER: Peru, NA: Data not available)

* Populations represented by two plants in the analysis

Table I-2. Numbers of AFLP bands scored for each primer combination in 59 samples of wild and cultivated *A. xanthorrhiza* and related species

Primer combination	No. of scored bands*	Amplified bands **	Polymorphic bands **	Polymorphism rate (%)**
E-AAT/M-CGA	39	49	34	69.4
E-AAT/M-CGG	43	73	36	49.3
E-AAT/M-CGT	24	35	21	60.0
E-ACT/M-CAC	36	48	29	60.4
E-ACT/M-CGA	20	33	14	42.4
E-AGC/M-CCA	35	62	32	51.6
E-AGC/M-CGT	38	45	32	71.1
Total	235	345	198	57.4

* Including all tuberous and non-tuberous samples

** Including only the tuberous samples

Table I-3. Pairwise matrix of Nei's genetic distance from 235 scored AFLP bands in cultivars (C) and wild *Arracacia* species or varieties.

Taxa	C	P	M	A.i	A.e	A.a
Wild perennial <i>A. xanthorrhiza</i> (P)	0.115	0.000				
Wild monocarpic <i>A. xanthorrhiza</i> (M)	0.146	0.119	0.000			
<i>A. incisa</i> (A.i)	0.241	0.207	0.135	0.000		
<i>A. elata</i> (A.e)	0.499	0.370	0.423	0.502	0.000	
<i>A. acuminata</i> (A.a)	0.467	0.369	0.410	0.481	0.383	0.000
<i>A. moschata</i> (A.m)	0.511	0.371	0.356	0.421	0.308	0.432

Table I-4. Nei's genetic distance from 198 AFLP bands observed in the four groups of accessions distinguished among the *A. xanthorrhiza* complex by UPGMA and PCoA. C: cultivar groups G1 and G2, W: wild groups G3 (perennial) and G4 (monocarpic).

	C (G1)	C (G2)	W (G3)
C (G1)	0.000	0.119	
W (G3)	0.117	0.180	0.000
W (G4)	0.221	0.109	0.119

Tableau 9. Statistiques générales de la diversité SSR au sein de la forme cultivée d'*A. xanthorrhiza* (N = 72)

Marqueur	Fréquence maximale d'un allèle	Nb. de génotypes observés	PIC
AxD85	0.40	6	0.59
AxD34	0.64	4	0.36
AxC85	0.53	4	0.38
AxD55A	0.71	3	0.33
AxD55B	0.53	4	0.45
AxD72	0.79	6	0.31
AxD82	0.40	9	0.62
AxC27	0.48	6	0.48
AxD43	0.92	3	0.14
AxC38	0.83	3	0.24
AxD13	1	1	0
Moyenne	0.66	4.45	0.35

Tableau 10. Résultats préliminaires de séquençage comparatif des régions flanquantes des 4 loci SSR chez le manioc

LOCUS (Clon)	TAILLE (pb)	TAILLE. ANALYSEE	C. C		C. NI		C. I	
SSRY-3 (23-120)*	520 pb	331 pb	301	90.9%	13	3.9%	17	5.6%
SSRY-40	600 pb	395 pb	347	87.8%	35	8.8%	13	3.2%
SSRY-9	350 pb	287 pb	264	91.9%	17	5.9%	6	2%

* Séquence consensus de trois loci (SSRY3, 23 et 120) qui se sont révélées être redondantes

CC : Nombre de nucléotides non polymorphes

C. NI : Nombre de polymorphismes non informatifs

C. I : Nombre de polymorphismes informatifs

Tableau 11. Résultats obtenus pour l'allongement des loci SSR par marche chromosomique chez le manioc et l'arracacha. Les amorces sélectionnées pour l'amplification d'allèles par PCR sont indiquées

LOCUS	TAILLE ORIGINAL (pb)	APRES EXTENSION (pb)	TA	AMORCE F (SEQUENCE 5'-3')	AMORCE R (SEQUENCE 5'-3')	Temp. Annealing	PRODUIT PCR (pb)
AxC87	473	704	0.5	<i>GSP1-A87</i> CTCTCAGATCATTTCATAAGGTA	<i>C87-R2</i> TCACTAGGGTCTTTACTAAGTCCAAAA	60	673
AxC64	385	789	1.0	<i>C64-Ex1F</i> AAGAAGAGCGGAGAGAGAGGAT	<i>C64-Ex1R</i> CAACTGCAGATCAAACCTATACGG	60	713
AxC29	409	1430	2.5	<i>C29-Ex1-F</i> TACTATGCCAATCCCACCTGTT	<i>C29-R</i> GAAATATCTGCTGGAAAGAGC	60	1405
SSRY-3	520	1501	2.6	<i>3-5'</i> CATACTCATGTATGACCCATTG	<i>C3-Ex2R</i> GGCTTGAAGGCTTAAGAAAACT	60	1155
SSRY-9	350	1641	3.6	<i>SSRY-9F</i> ACAATTCATCATGAGTCATCAACT	<i>C9-Ex2R</i> ATGTGAATGGTGGCTTTGTG	60	1157
SSRY-40	600	1836	2.0	<i>SSRY-40F</i> TGCATCATGGTCCACTCACT	<i>C40-Ex7R</i> ATCAGTGTGAGCAGCAACCAT	60	1078

TA = Taux d'allongement

Tableau 12. Allèles représentés dans les analyses phylogénétiques des séquences des loci AxC87 et AxC29 chez *A. xanthorrhiza*. La taille des allèles a été établie sur les données des séquences (*in silico*). Les formes cultivées sont différenciées en C1 ou C2 selon les deux groupes révélés par les AFLPs (Voir chapitre I, partie III).

Type	Code acc.	No. alleles*	ALLELES SEQUENCES DU LOCUS AxC87 (24)												
			119	117	115	113	109	107	105	101	99	96	91	89	85
Cultivé (C2)	1157	3				X									
Cultivé (C2)	1193	2			X					X					
Cultivé (C1)	1226	3	X	X						X					
Cultivé (C1)	1229	4								X					
Cultivé (C2)	UNC-43	2			X		X								
Cultivé (C2)	A	2		X						X					
Cultivé (C2)	UNC171	3								X	X				
A.x.var.andina	S9	3						X	X		X				
A.x.var.andina	S133	2								X			X		
A.x.var.monoc	081	1								X					
A.x.var.monoc	S500	2								X			X		
A.x.var.monoc	31-2	1												X	
A.x.var.monoc	S105	0									X				
A. incisa	RBS	1													X
TOTAL			1	2	2	1	1	1	1	8	2	1	2	1	1

* Nombre d'allèles observés en gel

Type	Code acc.	No. allèles	ALLELES SEQUENCES DU LOCUS AxC29										
			117	124	125	134	135	136	138	140			
Cultivée (C2)	1193	1			X								
Cultivée (C2)	1201	2			X				X				
Cultivée (C2)	A	1			X								
Cultivée (C2)	UNC171*	2*			X								
Cultivée (C1)	1229	2			X				X				
Cultivée (C1)	1226	1									X		
Cultivée (C1)	1161	1									X		
A.x.var. andina	041	1						X					
A.x.var. andina	011	2			X				X				
A.x.var. andina	S105	1											X
A.x.var. andina	S507	1											X
A.x.var. andina	S133	1											X
A.x.var.monoca	081	1	X										
A. incisa	RBS001	1						X					
A.acuminata	015	2		X			X						
A.acuminata	029	1		X									
A.acuminata	043	1					X						
A.acuminata	050	1			X								
A.acuminata	053	2		X			X						
TOTAL		25	1	3	7	3	2	3	2	3	2	3	

* Cette accession révèle deux allèles qui présentent le même motif mais certaines différences au niveau des régions flanquantes (Voir annexe 3).

Tableau 13. Allèles enregistrés pour les accessions de manioc et formes sauvages sur les loci SSRY-3, SSRY-9 et SSRY-40 en gel¹ (Morillo 2001) et *in silico*² (sur les données des séquences). Les allèles représentés dans l'analyse phylogénétique sont représentés dans les arbres et tableaux par leur taille *in silico*.

	Sp.	Code	SSRY-3 ¹		SSRY-3 ²		SSRY-9 ¹		SSRY-9 ²		SSRY-40 ¹		SSRY-40 ²		
1	B	5-2	221		238		253		262		221		223		
2	B	5-4	209	214	232		253		262		221				
3	Q	5-24	209				253		260		221		223	227	
4	Q	5-25	209		226		253				236		237		
5	Cth	5-78	213				254		264		216		225		
6	Cth	5-85	208	217	236		254		264		216		217		
7	G	5-6	204	216	216	224	254		260	263	217		225		
8	G	5-7	216	226	214	224	254				250		229		
9	G	5-9	213				254				218		217		
10	G	5-42	204		228		254				218	221	227		
11	T	5-19	216	233	253		259	262			218		219		
12	T	5-30	209	233	222	224	253	259	260		221	236	225	229	
13	T	5-31	216		214	224	244	259	250		218		217	219	
14	T	5-32	204	233			256	259	256	257	218		225	231	
15	T	5-33	216	233	228	230	256				218	231	216	231	
16	E	5-28	228				259		260	264	229	236	227	231	237
17	E	5-46	216		241		259				236		231	237	
18	E	5-47	228				259	262	262	264	221		225		
19	E	5-49	228	233			244	259	250	262	226	229	229		
20	E	5-51	228		245		270				226		227		
21	E	5-52	228	233	238	242	259	262			226		219	223	
22	E	5-53	228		244		262	270	278		221		223		
23	E	5-54	228		239	244	262	270			221		223		
24	E	5-56	226		238		257	262			221	236	237		
25	E	5-58	228	233	243	257					221	229	223	229	
26	E	5-59-AI	204	226	216	239	262		266		231	236	231	237	
27	Ei ?	5-62	226	228	240		244	262	250		221	229	223	229	
28	Ei ?	5-63	228		241	243	259	262			218	236	225	235	
29	Ei ?	5-64	228				256	262			218	236			
30	Ei ?	5-65	228	233			244	262	250		226	236	235	249	
31	Ei ?	5-68	228		244	246	259	262	262		218	236	219	237	
32	E	5-70-AI	204		216	238	259		262		231	236	233		
33	Ei ?	5-72	233		242		244	262	250		226				
34	F	5-88	228		245		270				226	238	239		
35	F	5-89	226		240		270		274		221	231	225	229	
36	F	5-90	228		244		259	262	268		226	236	431		
37	F	5-91	228		246		257				236		229	233	
38	F	5-94	226				257		262		221		219		

B = *M. brachyloba*, Q = *M. quinquepartita*, Cth = *M. carthaginensis*, T = arbres manioc, E = *M. esculenta* (? pour les suspects introgressés, AI pour les artificiellement introgressés), F = *M. flabellifolia*

Tableau 14. Allèles représentés dans l'analyse phylogénétique des séquences du locus SSRY-3 du manioc. Les allèles suspectés d'introgession sont en gras. La taille des allèles a été établie *in silico*.

Sp.	Code	No. allèles	214	216	222	224	226	228	230	232	236	238	239	240	241	242	243	244	245	246	253	257
B	5-2	1										X										
B	5-4	1								X												
Q	5-25	1					X															
Cth	5-85	1									X											
G1	5-6	2		X		X																
G1	5-7	2	X			X																
G2	5-42	1						X														
T	5-19	1																			X	
T	5-30	2			X	X																
T	5-32	2	X			X																
T	5-33	2						X	X													
C	5-46	1													X							
C	5-51	1																		X		
C	5-52	2										X				X						
C	5-53	1																X				
C	5-54	2											X					X				
C	5-56	1										X										
C	5-58	2															X					X
Ci	5-59 (AI)	2		X									X									
Ci ?	5-62	1												X								
Ci ?	5-63	2													X		X					
Ci ?	5-68	2																X			X	
Ci	5-70 (AI)*	2		X								X										
Ci ?	5-72	1														X						
F	5-88	1																		X		
F	5-89	1												X								
F	5-90	1																X				
F	5-91	1																			X	
TOTAL		40	2	3	1	4	1	2	1	1	1	4	2	2	2	2	2	4	2	2	1	1

Tableau 15. Allèles représentés dans l'analyse phylogénétique des séquences du locus SSRY-9 du manioc. Les allèles suspectés d'introgessions sont en gras. La taille des allèles a été établie *in silico*.

Sp.	Code	No. allèles	278	274	268	266	264	263	262	260	257	256	250
B	5-2	1							X				
B	5-4	1							X				
Q	5-24	1								X			
Cth	5-78	1					X						
Cth	5-85	1					X						
G1	5-6	2						X		X			
T	5-30	1								X			
T	5-31	1											X
T	5-32	2									X	X	
C	5-28	2					X			X			
C	5-47	2					X		X				
C	5-49	2							X				X
C	5-53	1	X										
Ci	5-59-AI	1				X							
Ci ?	5-62	1											X
Ci ?	5-65	1											X
Ci ?	5-68	1							X				
Ci	5-70-AI	1							X				
Ci ?	5-72	1											X
F	5-89	1		X									
F	5-90	1			X								
F	5-94	1							X				
TOTAL		27	1	1	1	1	4	1	7	3	1	1	5

Tableau 16. Allèles représentés dans l'analyse phylogénétique des séquences du locus SSRY-40 du manioc. La taille des allèles a été établie *in silico*.

Sp.	Code	No. allèles	431	249	239	237	235	233	231	229	227	225	223	219	217	216
B	5-2	1											X			
Q	5-24	2									X		X			
Q	5-25	1				X										
Cth	5-78	1										X				
Cth	5-85	1													X	
G1	5-6	1										X				
G1	5-7	1							X							
G2	5-9	1													X	
G2	5-42	1									X					
T	5-19	1												X		
T	5-30	3			X					X		X				
T	5-31	2												X	X	
T	5-32	2							X			X				
T	5-33	2							X							X
C	5-28	3				X			X		X					
C	5-46	2				X			X							
C	5-47	1										X				
C	5-49	1								X						
C	5-51	1									X					
C	5-52	2											X	X		
C	5-53	1											X			
C	5-54	1											X			
C	5-56	1				X										
C	5-58	2								X			X			
Ci	5-59-AI	2				X			X							
Ci?	5-61	1									X					
Ci?	5-62	2								X			X			
Ci?	5-63	2					X					X				
Ci?	5-65	2		X			X									
Ci?	5-68	2				X								X		
Ci	5-70-AI	1						X								
F	5-88	1			X											
F	5-89	2								X		X				
F	5-90	1	X													
F	5-91	2						X		X						
F	5-94	1												X		
TOTAL		54	1	1	2	6	2	2	5	7	5	7	7	5	3	1

Tableau 17. Résultats de l'analyse phylogénétique par MP des séquences d'allèles SSR allongés chez le manioc et l'arracacha

Locus	Allèles Séquencés	Caractères alignés	Caractères Constants	C. NI	C. I.	Polymorphisme (%)
AxC87	24	633	581	38	14	2.2
AxC29	25	1413	1093	169	152	10.7
SSRY-3	40	1110	891	130	89	8.2
SSRY-9	27	1143	1027	79	37	3.2
SSRY-40	54	856	705	94	57	6.6

C. NI : Nombre de polymorphismes non informatifs (Comprend des sites où un polymorphisme est présent chez un seul individu)

C. I : Nombre de polymorphismes informatifs (Comprend des sites polymorphes chez au moins deux individus et des indels codés comme présents ou absents chez au moins deux individus)

FIGURES



Figure 1. Plantes à tubercules et racines originaires des Andes. Toutes ces espèces, sauf la maca (5) et l'ahipa (6) sont propagées végétativement. 1: pomme de terre (Solanaceae), 2: mashua (Tropaeolaceae), 3: oca (Oxalidaceae), 4: ulluco (Basellaceae), 5: maca (Cruciferae), 6: ahipa (Legumineuse), 7: yacon (Compositae), 8: mauka (Nyctagynaceae), 9: achira (Cannaceae), 10: arracacha (Apiaceae) (Photos INIAP, CIP).



Figure 2. Généralités sur l'arracacha: 1) Parties botaniques distinctives chez une plante adulte d'arracacha, 2) Organes utilisés pour la propagation asexuelle de l'arracacha. Les propagules de la rangée inférieure sur la photo bas donneront lieu à des plantes plus productives que ceux de la rangée supérieure; 3) Exemples de préparations culinaires avec l'arracacha (purée, frites, «tortillas») (Photos: INIAP 1996), 4) Illustrations relevées dans des restes de céramique retrouvés à Nazca, Pérou, qui suggèrent une représentation de l'arracacha (Source: Hermann 1997).

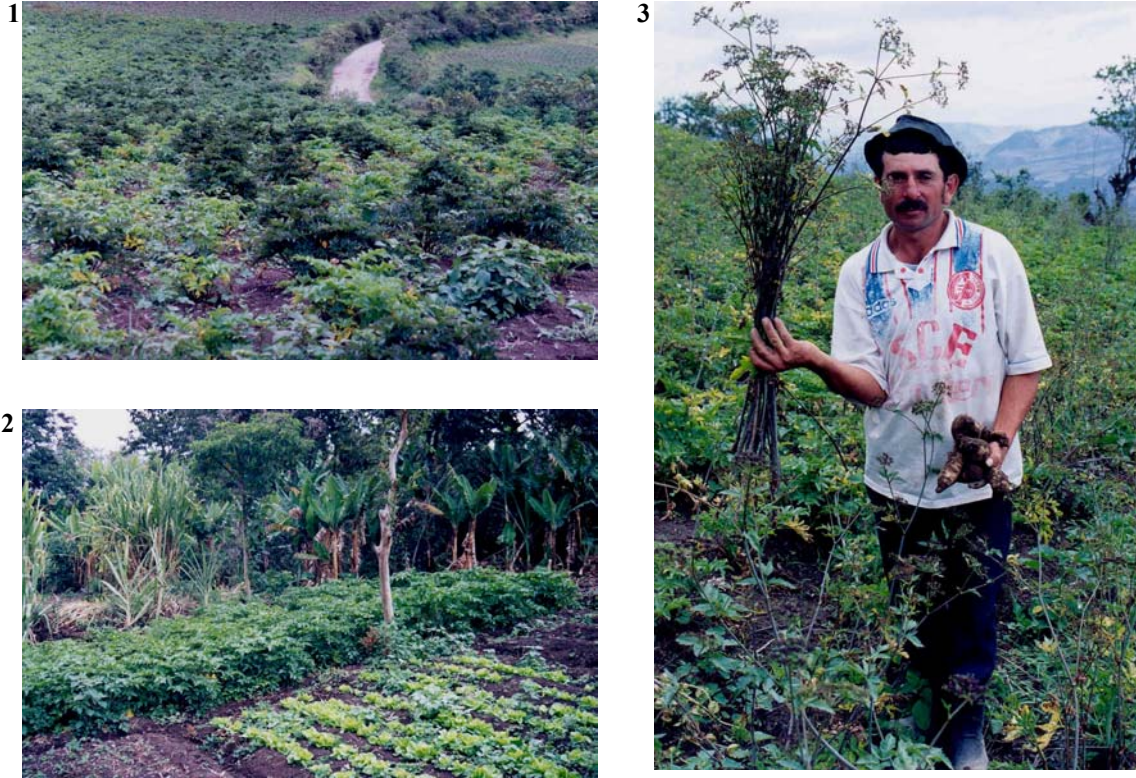


Figure 3. Exemples de cultures de l'arracacha: 1) Champs de production d'arracacha (1-2 Ha); au moins deux cultivars sont distinguables. 2) Parcelle d'arracacha dans une ferme; 3) Parcelle d'arracacha abandonnée à cause de la floraison massive des plantes (Photos: E. Morillo, San José de Minas, Loja, Equateur 2002, 2003).



Figure 4. Floraison spontanée chez l'arracacha cultivée. A gauche: échappée de culture en pleine fructification dans une jachère. Certains fruits sont vides et ne portent pas d'embryons. A droite: Paysan montrant la différence de production de racines consommables entre une plante en floraison et une en état végétatif (Photos: E. Morillo, Loja et San José de Minas, Equateur, 2002, 2003).



Figure 5. Formes horticoles distinguées chez l'arracacha cultivée (blanches, jaunes et violettes) (Photo: INIAP, 1996).

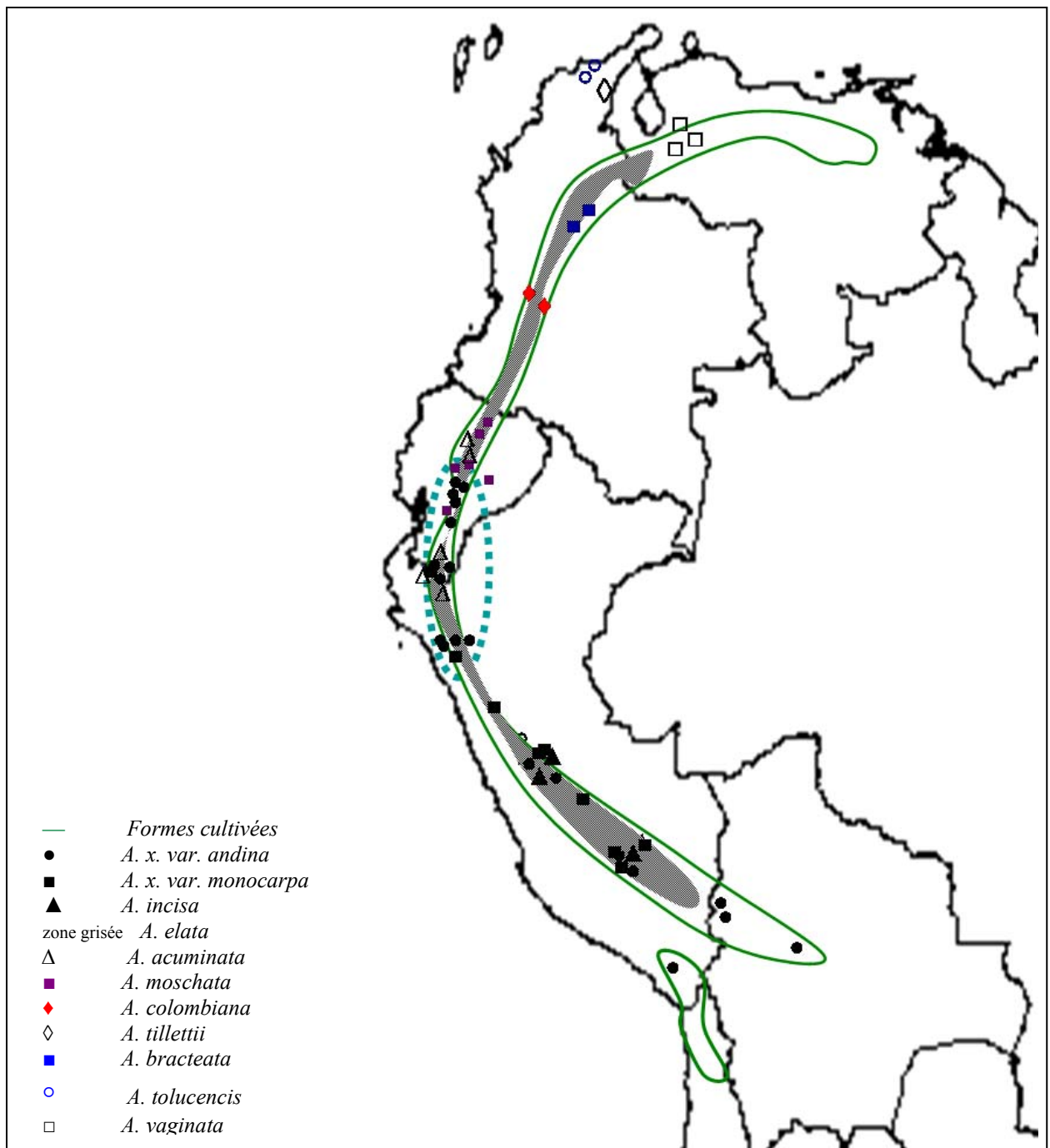


Figure 6. Distribution géographique du genre *Arracacia* en Amérique du sud (Adapté d'Hermann, 1997 et Knudsen, 2003). L'aire d'étude est encerclée en pointillés.



Figure 7. Populations naturelles d'espèces sauvages d'Arracacia en Equateur (dans les forêts de montagne, à côté de routes). En haut: *A. acuminata* (Route San Jose de Minas-Otavalo, Pichincha, accession MST-052). En bas: *A. moschata* « urco zanahoria » (Olmedo, Imbabura, accession MST-054). Photos E. Morillo 2002-2003.



Figure 8. Type sauvage polycarpique d'*A. xanthorrhiza* (*A. x. var. andina*). 1) Système racinaire d'une population retrouvée à Loja (accession MST-011); 2) Population trouvée proche du type cultivé; 3) Plantules issues de graines (deux tiges sont distinguables; 4) Paysage typique où poussent ces populations. Photos: E. Morillo, G. Second (2002-2003).



Figure 9. Type sauvage monocarpique de *A. xanthorrhiza* (*A. x. var. monocarpa*). 1) Plantes en floraison (MST-081); 2) Population en état végétatif (MST-082); 3) Système racinaire, 4) Plantules issues de graines (Une seule tige est distinguable) (Photos E. Morillo, Loja, 2003).



Figure 10. Morphologie de l'espèce tubéreuse *A. incisa*. Gauche: Plantes à l'état végétatif. Droite: Plante en floraison (Photos extraites de Blas 2005).

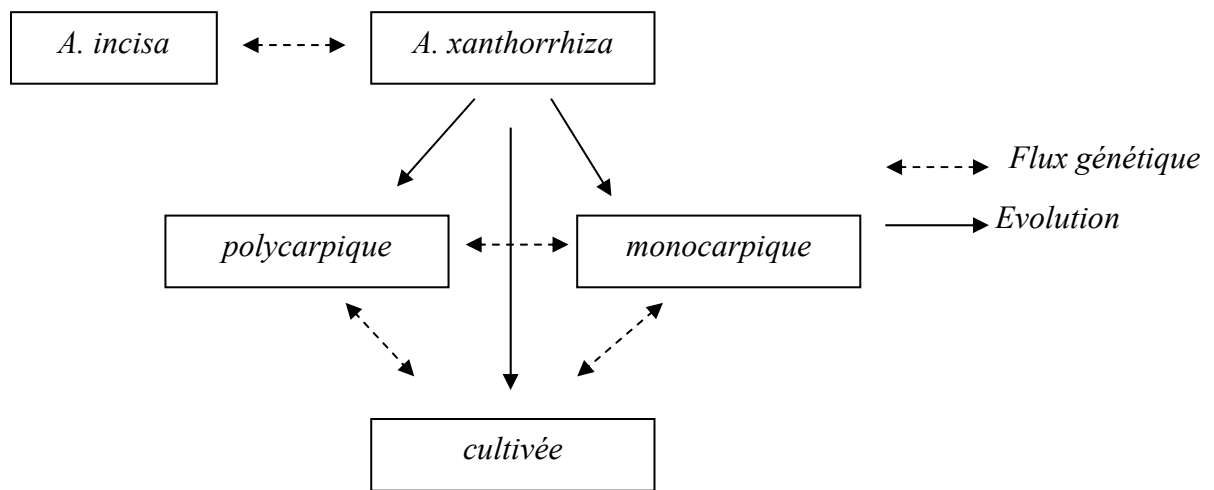


Figure 11. Origine de la forme cultivée d'arracacha selon Blas (2005).

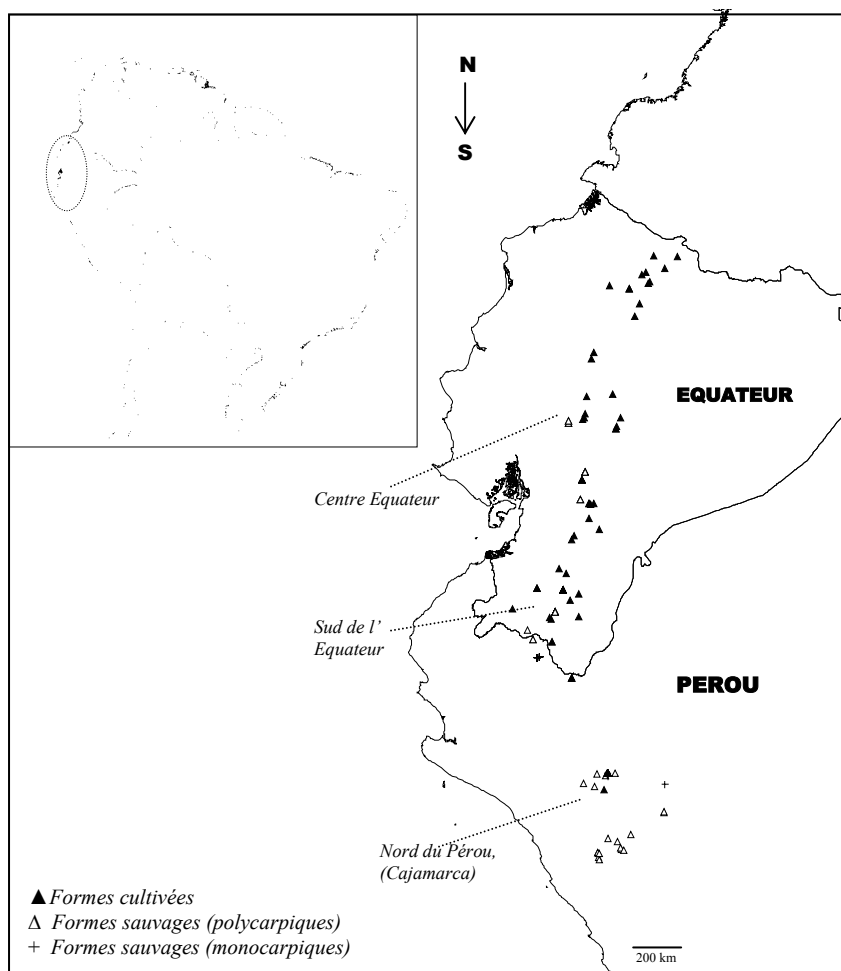


Figure 12. Carte des origines des formes sauvages et cultivées de *A. xanthorrhiza* analysées dans la présente étude.

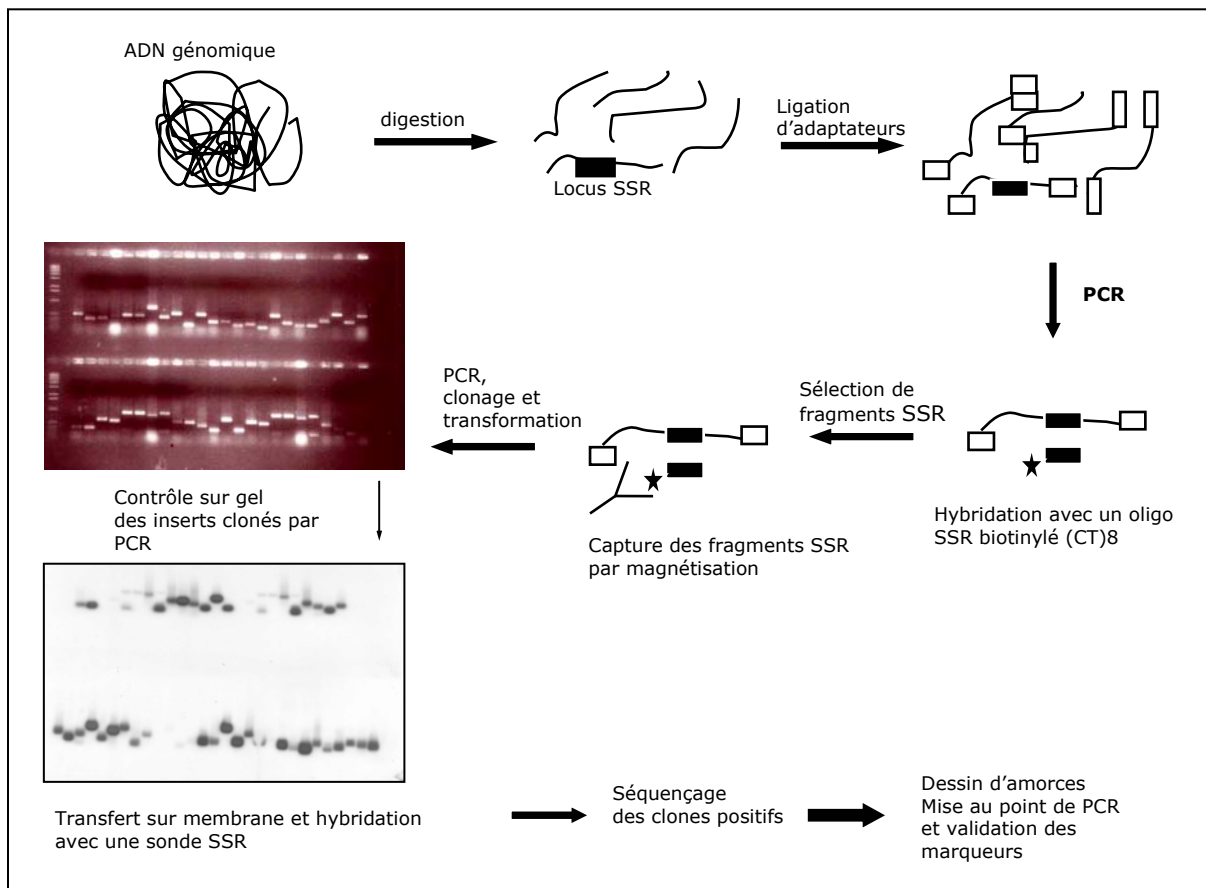


Figure 13. Principe technique de la méthode de développement des marqueurs microsatellites selon Billotte et al. (1999).

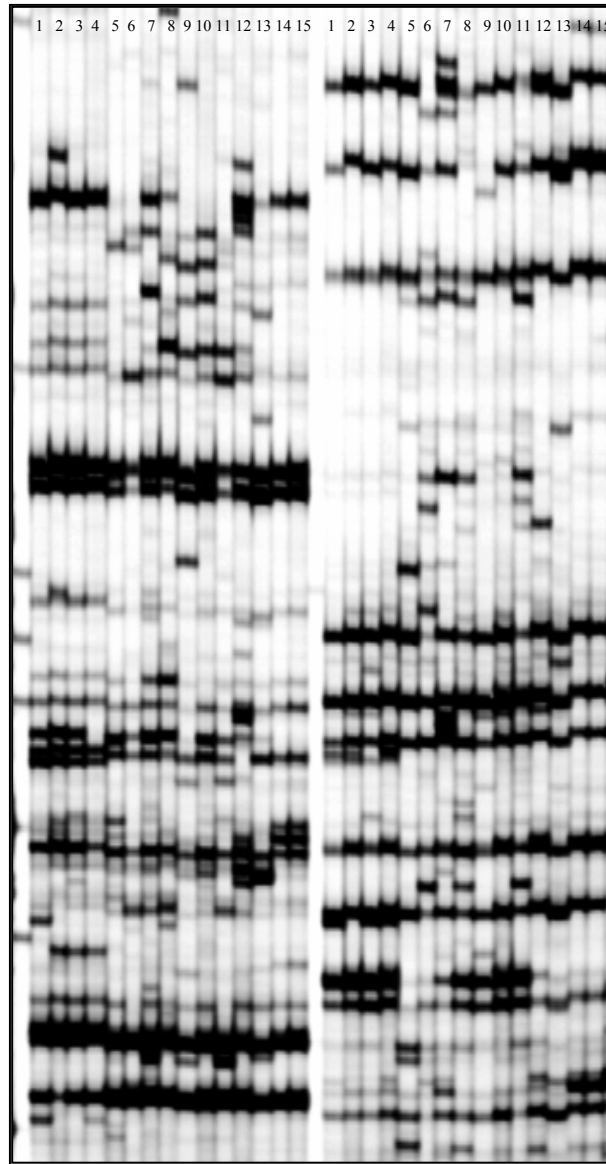


Figure 15. Exemples des profils AFLPs obtenus dans l'étude de diversité parmi le complexe *A. xanthorrhiza* (à gauche couple d'amorces E-AAT/M-CCC et à droite E-AAT/M-CTG). Echantillons 1-4: Formes cultivées, 5-9: *A. x. var. andina*, 10-13: *A. x. var. monocarpa*, 14-15: *A. incisa*.

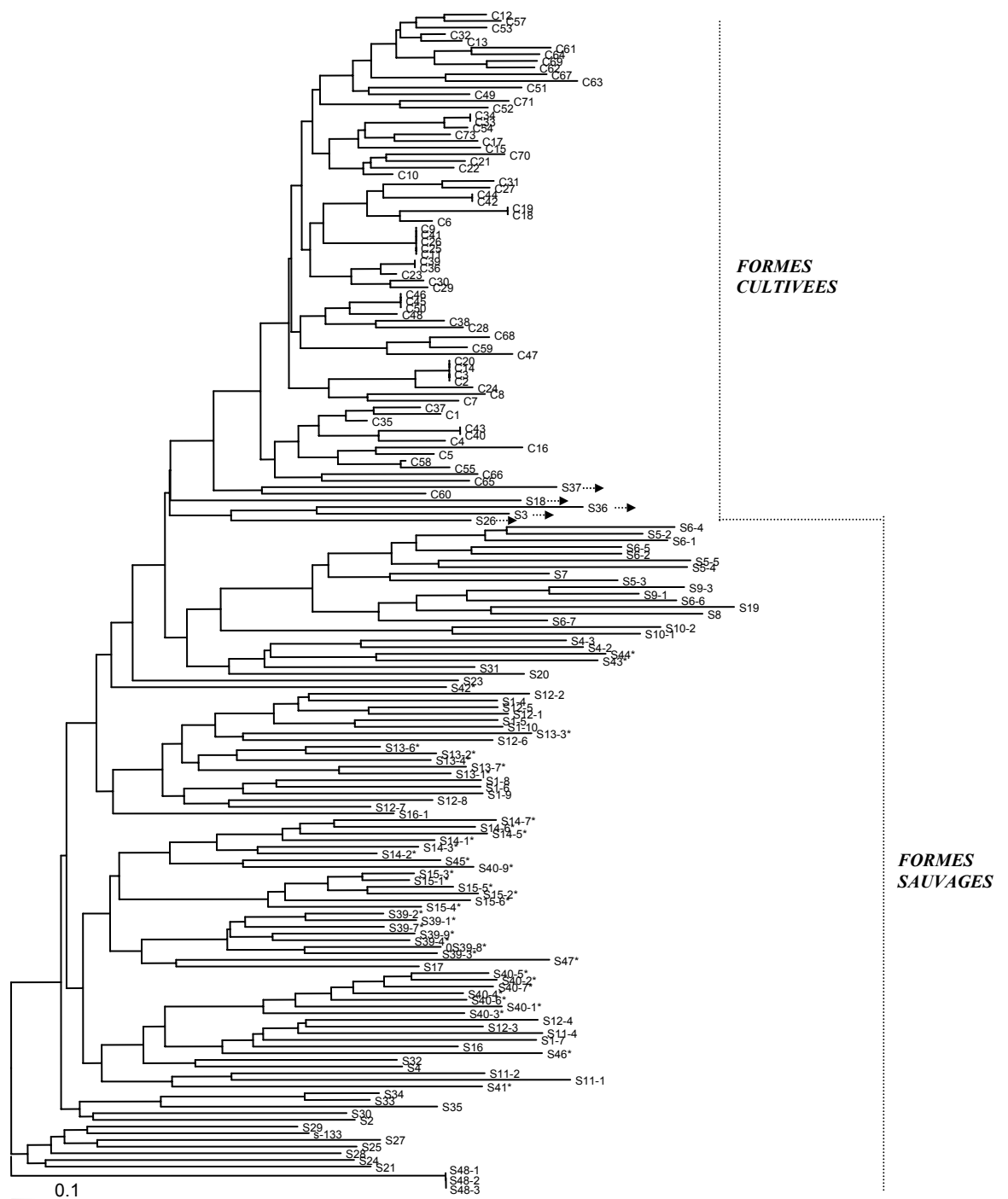


Figure 16. Arbre NJ basé sur la distance DAS qui montre les relations génétiques entre 178 formes tubéreuses cultivées et sauvages d'arracacha sur la diversité observée sur 11 loci SSR. L'étiquette correspond au numéro (No.) donnée dans l'annexe 1. Les formes sauvages monocarpiques sont distinguées par un astérisque (*). L'arbre est raciné en utilisant l'espèce *A. incisa* (S48) comme groupe externe. Les flèches indiquent les formes sauvages les plus affines des formes des cultivés.

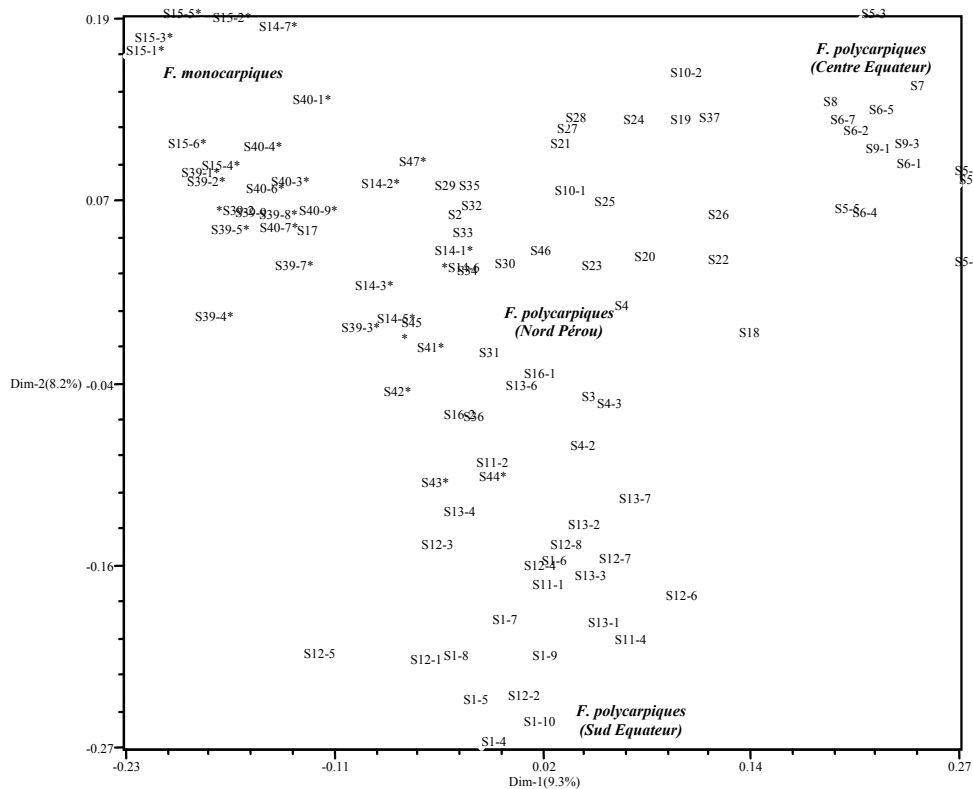
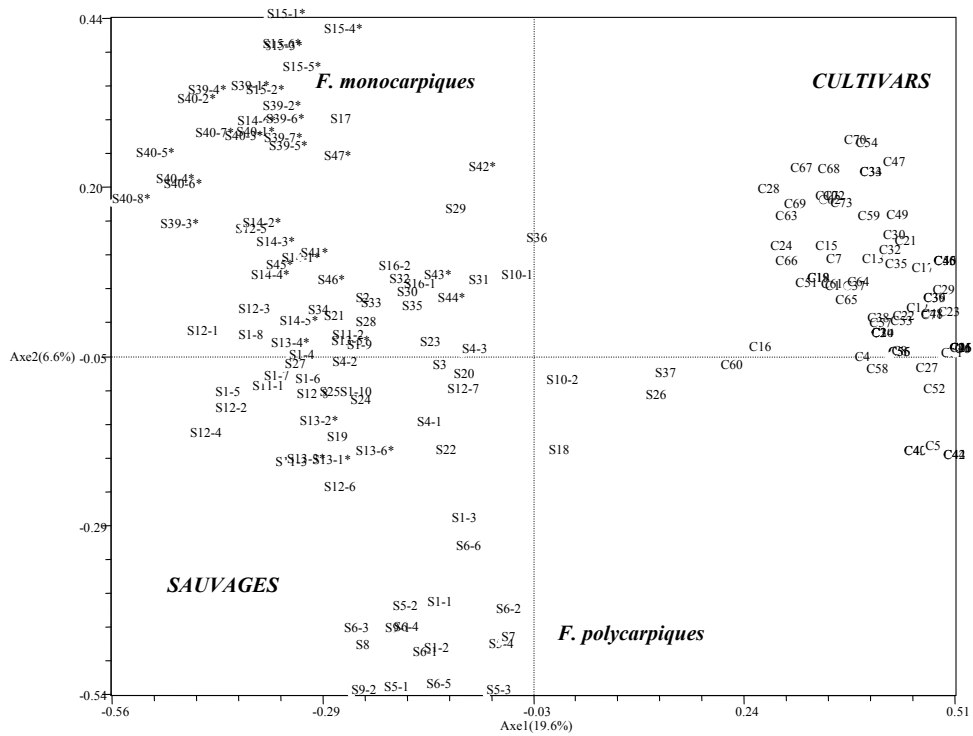


Figure 17. Analyses multivariées fondées sur la diversité SSR chez les formes cultivées et sauvages d'arracacha. En haut: Analyse PCO (coefficient DICE) fondée sur la diversité observée sur 11 loci SSR chez 178 formes tubéreuses cultivées (C) et sauvages (S) d'arracacha. En bas : PCO avec les 106 arracachas sauvages (102 allèles). Les types identifiés comme *A. x. var. monocarpa* portent un astérisque et les collections d'assignation douteuse sont suivies de (?). Chaque individu est identifié par son code d'accésion.



Figure 18. Arbre NJ basé sur la distance DAS qui montre les relations génétiques entre 178 formes tubéreuses cultivées et sauvages d'arracacha sur la base de la diversité observée sur 11 locus SSR. L'étiquette correspond au numéro (No.) donné sur l'annexe 1. Les formes sauvages monocarpiques sont distingués par un astérisque (*). L'arbre est raciné en utilisant l'espèce *A. incisa* (S48) comme groupe externe. Les flèches indiquent des formes sauvages ou cultivées regroupées avec les hybrides artificiels.

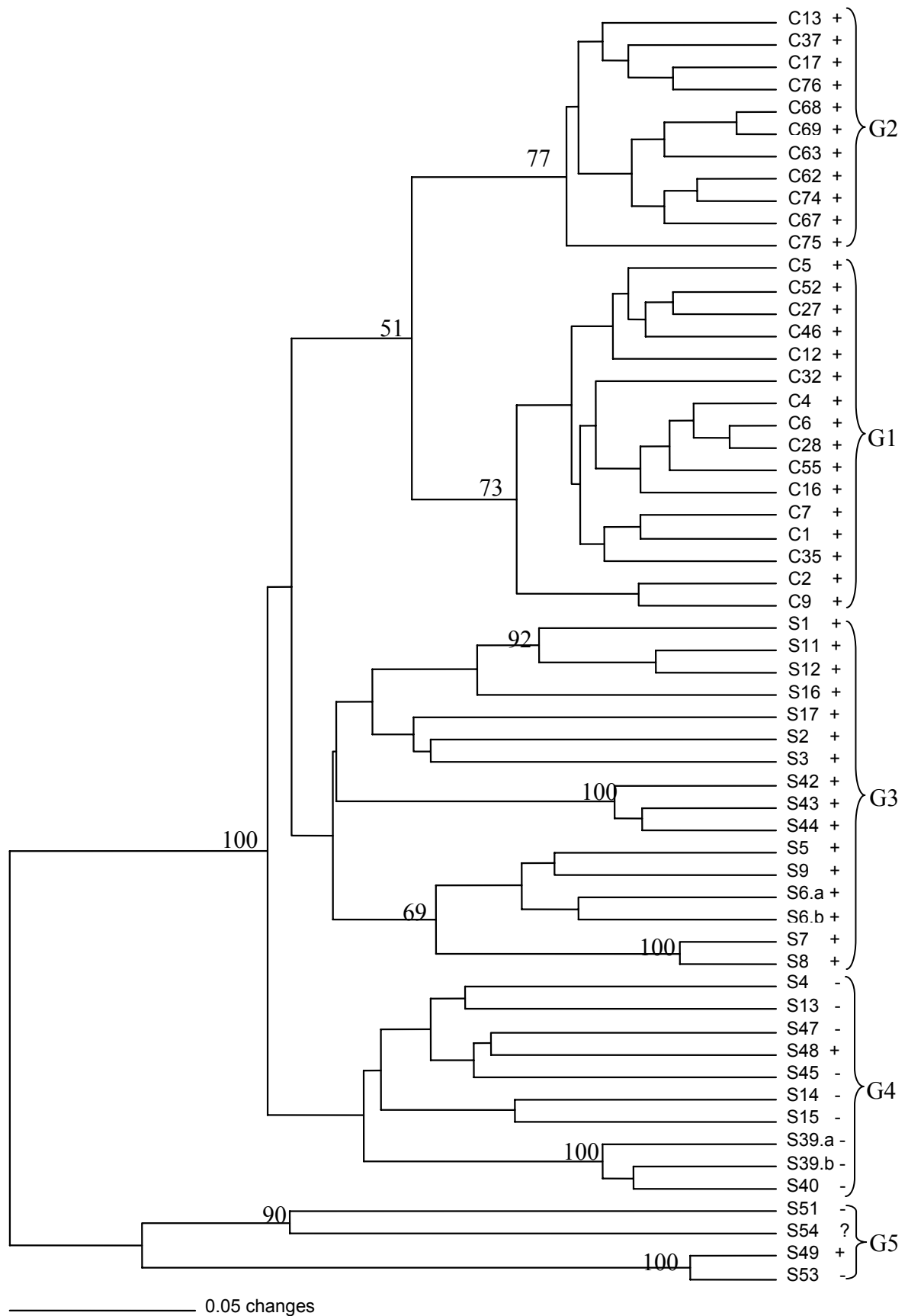


Figure I-1. UPGMA tree obtained from 235 AFLP bands scored in a set of cultivated arracacha and wild relatives. The groups are described in the text. Bootstrap values were obtained from 100 replicates (only values over 50 are shown). Presence or absence of the Q16cp indel is indicated by + or - for each accession.

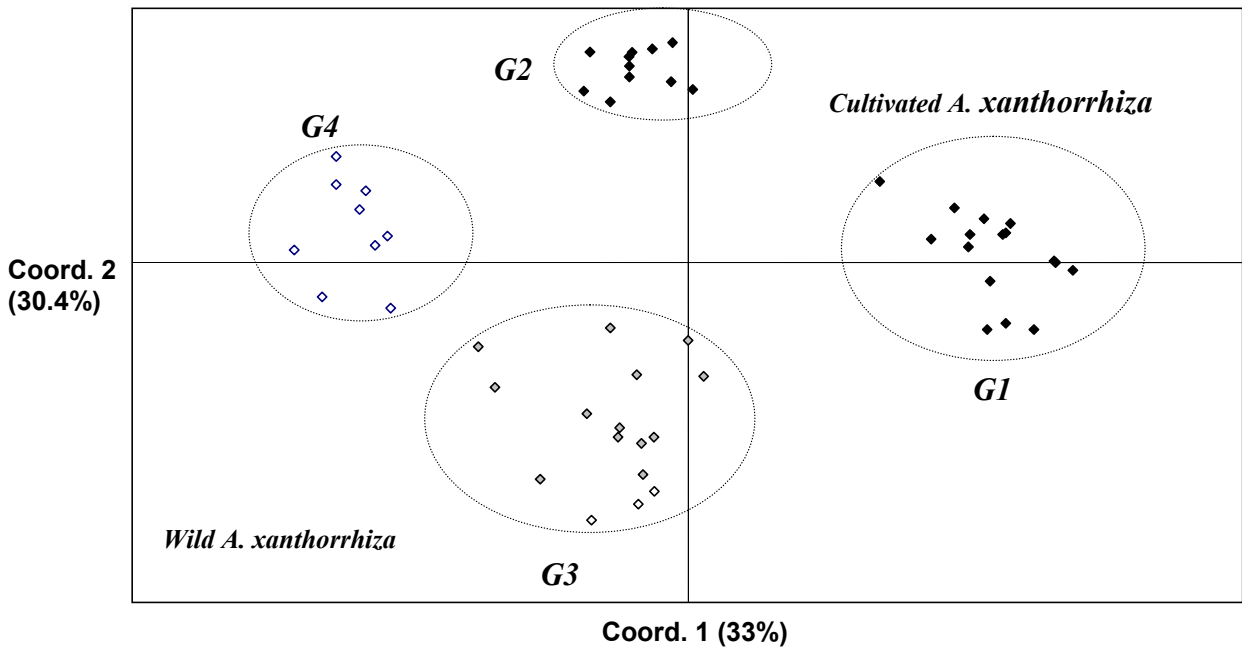


Figure I-2. Plane defined by the first two axes of a PCO based on 198 AFLP bands. Cultivated accessions are represented by black symbols and wild perennial and monocarpic *A. xanthorrhiza* by gray and white symbols, respectively. The dotted circles (G1-G4) refer to the clusters numbered as in Fig. 1.

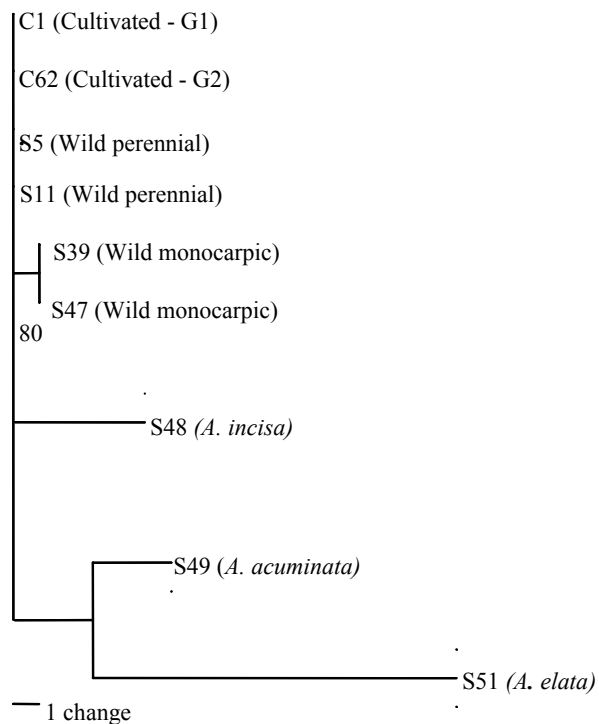


Figure I-3. Parsimonious phylogramme of arracacha accessions obtained through the phylogenetic analysis of 24 polymorphisms observed in 2000 bp of chloroplast ADN

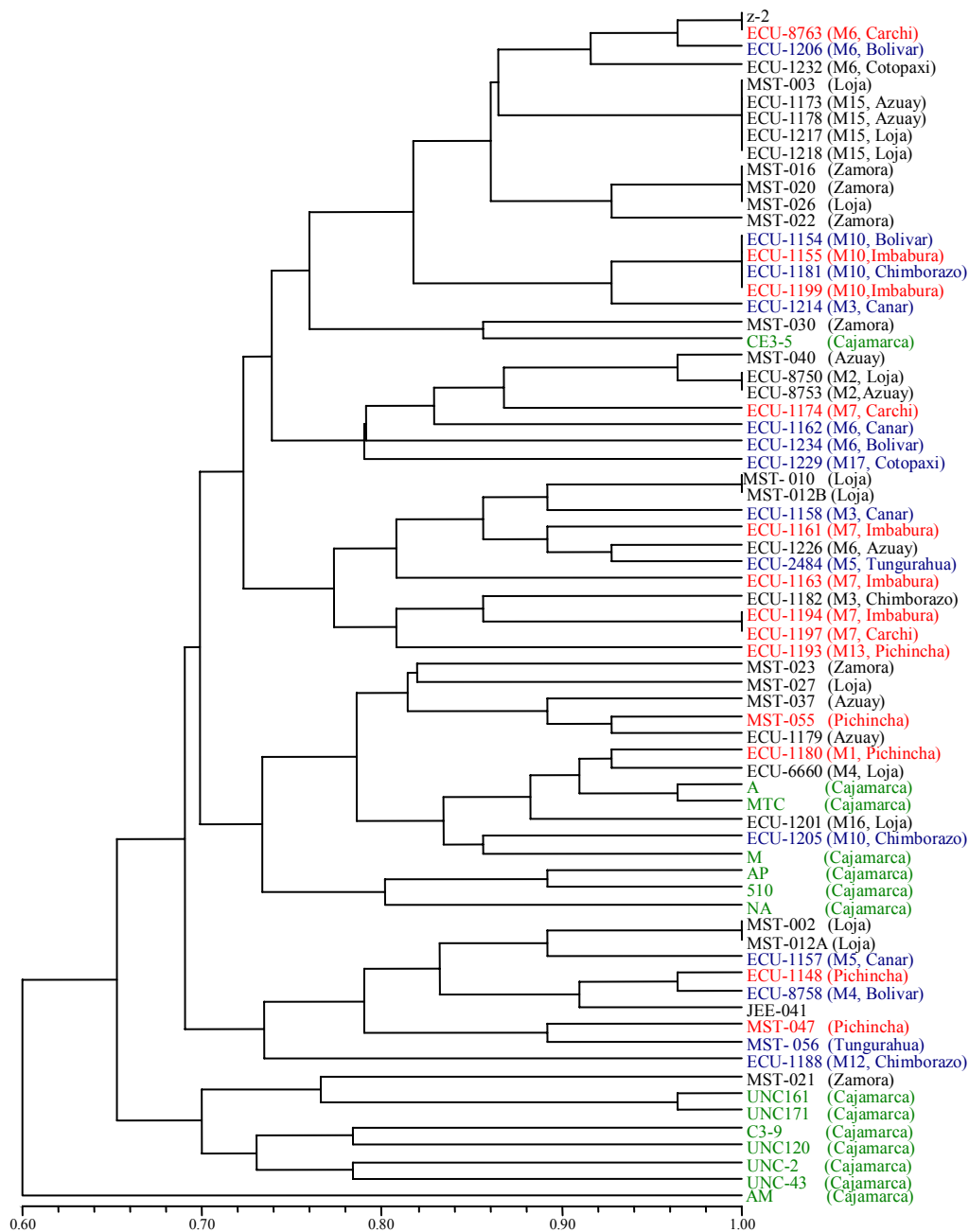


Figure 20. Arbre UPGMA qui regroupe les 72 accessions cultivées selon le polymorphisme observé sur 11 loci SSR (31 allèles). Pour chaque accession, son code est suivi lorsque cette information était disponible par le numéro de morphotype accordé par Mazon (1996) suivie de la province ou département où elles ont été collectées. Les accessions collectées au sud de l'Equateur sont en couleur noire, les accessions collectées dans les provinces du centre sont en bleu et celles provenant du nord du pays sont en rouge. Les accessions collectées dans le nord du Pérou sont en vert.

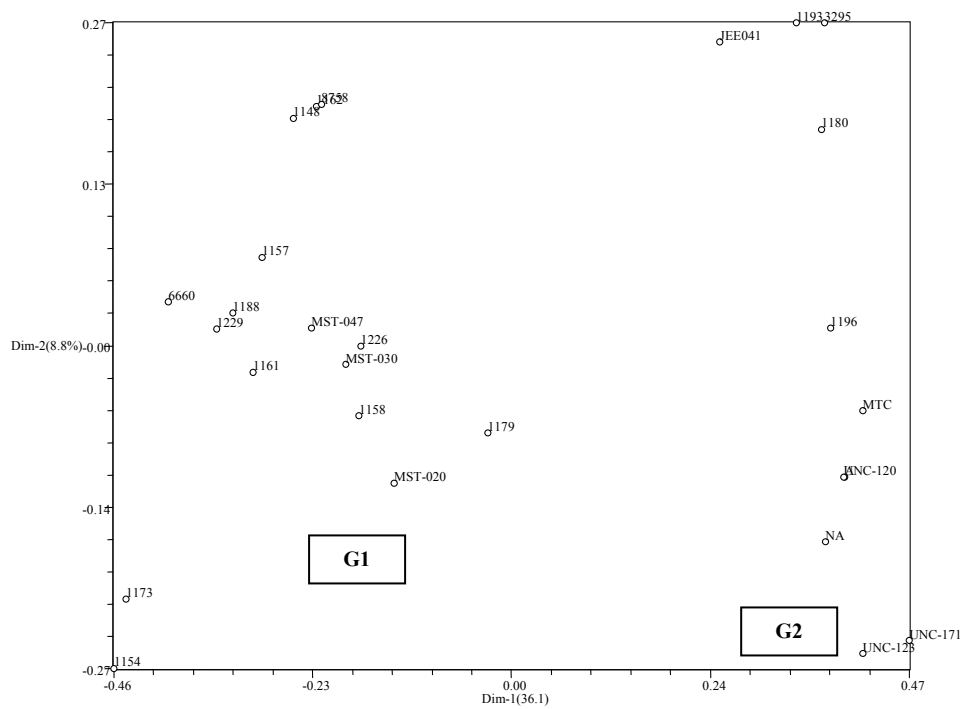
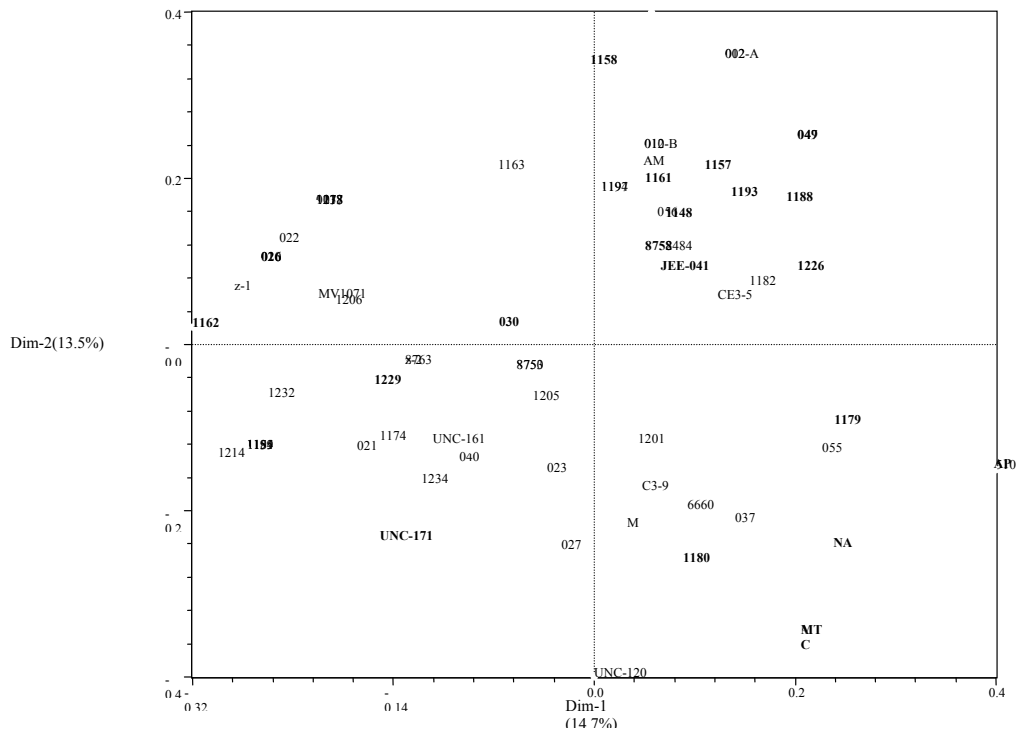


Figure 21. Analyses multivariées basées sur la diversité SSR et AFLPs parmi les formes cultivées d'arracacha. a) En haut : Analyse PCO fondée sur la diversité observée sur 11 loci SSR chez 72 accessions de la forme cultivée d'arracacha, b) En bas: PCO fondé sur la diversité observée sur 7 couples AFLPs sur un ensemble de 27 cultivars sélectionnés (indiqués en gras sur la figure a). Le code d'accésion pour chaque cultivar est donné.

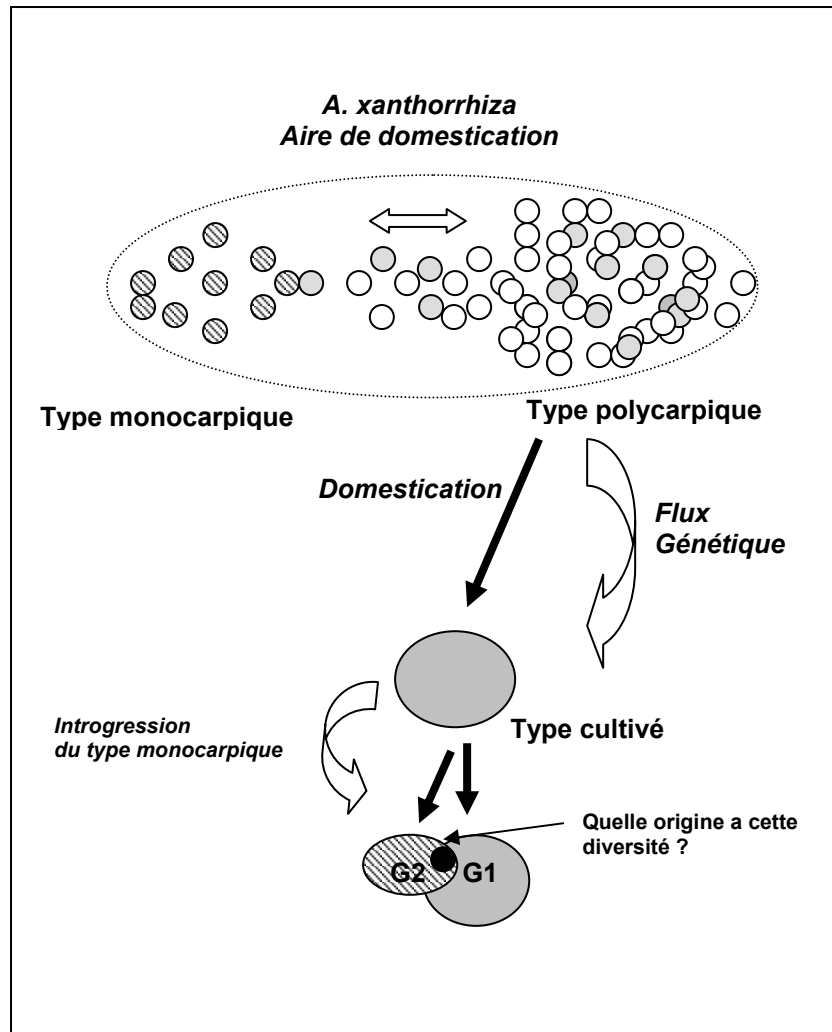


Figure 22. Scénario hypothétique sur l'origine de la forme domestiquée de *A. xanthorrhiza* sur la base de l'information donnée par les marqueurs SSR, AFLP et l'ADN chloroplastique : Parmi les deux formes sauvages de *A. xanthorrhiza* présentes dans le centre de domestication (Sud de l'Equateur et Nord du Pérou), le type polycarpique ou pérenne aurait donné lieu à la forme cultivée. Par ailleurs la participation post-domestication du type sauvage monocarpique à la diversité de la forme cultivée serait à l'origine d'une divergence cryptique parmi les cultivars. Cependant une portion de la diversité du type domestiqué n'est pas expliquée par une contribution de ses homologues sauvages (mono ou polycarpique).

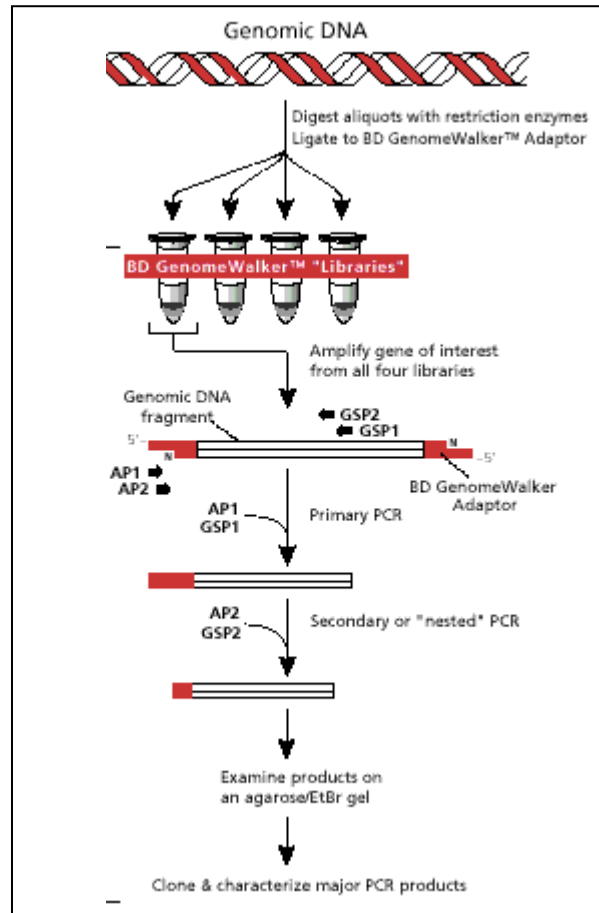


Figure 23. Principe de la méthode de marche sur le chromosome employée par le kit *Genome Walker Universal* (Source: Manuel Technique). Voir le texte pour une explication de la technique.

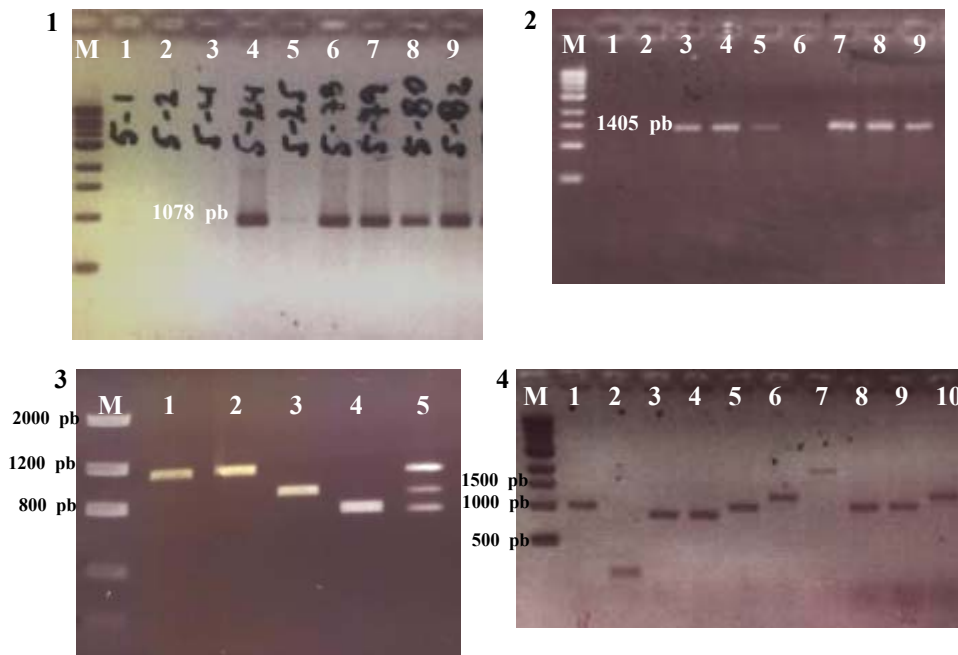


Figure 24. Etapes pour l'isolement d'allèles SSR allongés:

1 et 2) Exemples de transférabilité des nouvelles amorces dessinés sur la séquence du locus SSR allongé par marche chromosomique: 1) Chez le manioc pour le locus SSRY-40 (1078pb), on remarque l'absence d'amplification chez *M. brachyloba* (piste 1 et 2) et *M. quinquepartita* (piste 3), 2) Chez l'arracacha pour le locus AxC29 (1405pb), on remarque l'absence d'amplification chez deux accessions de *A. x. var. monocarpa* (piste 1 et 2) et *A. elata* (piste 6) et une bonne amplification chez l'arracacha cultivée (piste 3, 4), *A. acuminata* (piste 5, 7 et 8) et *A. incisa* (piste 9).

3) Amplification de différents loci SSR allongés chez une même accession (M: marqueur de taille, clone SSRY-3:piste1, clone SSRY-9: piste 2, clone SSRY-40: piste 3, SSRY-108 piste 4,) et multiplexage des produits PCR pour le clonage d'allèles (piste 5),

4) Criblage d'allèles allongés et clonés pour le séquençage: allèles du locus SSRY40 (piste 1, 5, 8 et 9), locus SSRY-3 (piste 6 et 10), locus SSRY-9 (piste 3 et 4).

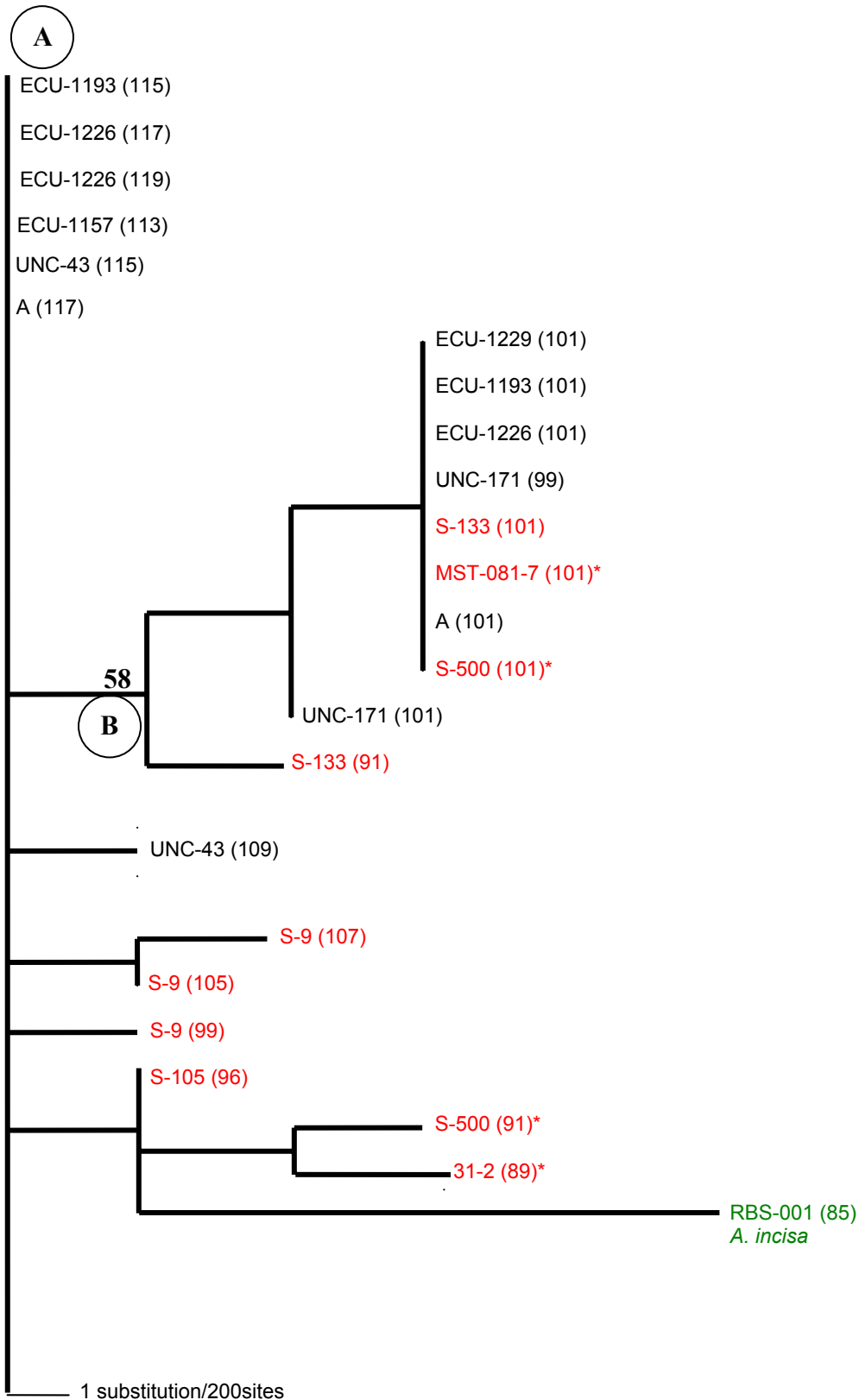


Figure 25. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 24 allèles du locus SSR AxC87 isolés chez *A. xanthorrhiza* et *A. incisa* (en vert). Les allèles sont indiqués par l'accèsion dont ils ont été isolés, puis la taille obtenue *in silico* avec les amorces utilisées en génotypage sur les données de séquence. Les formes cultivées sont en noir et les formes sauvages sont en rouge (les formes monocarpiques sont distingués par*).

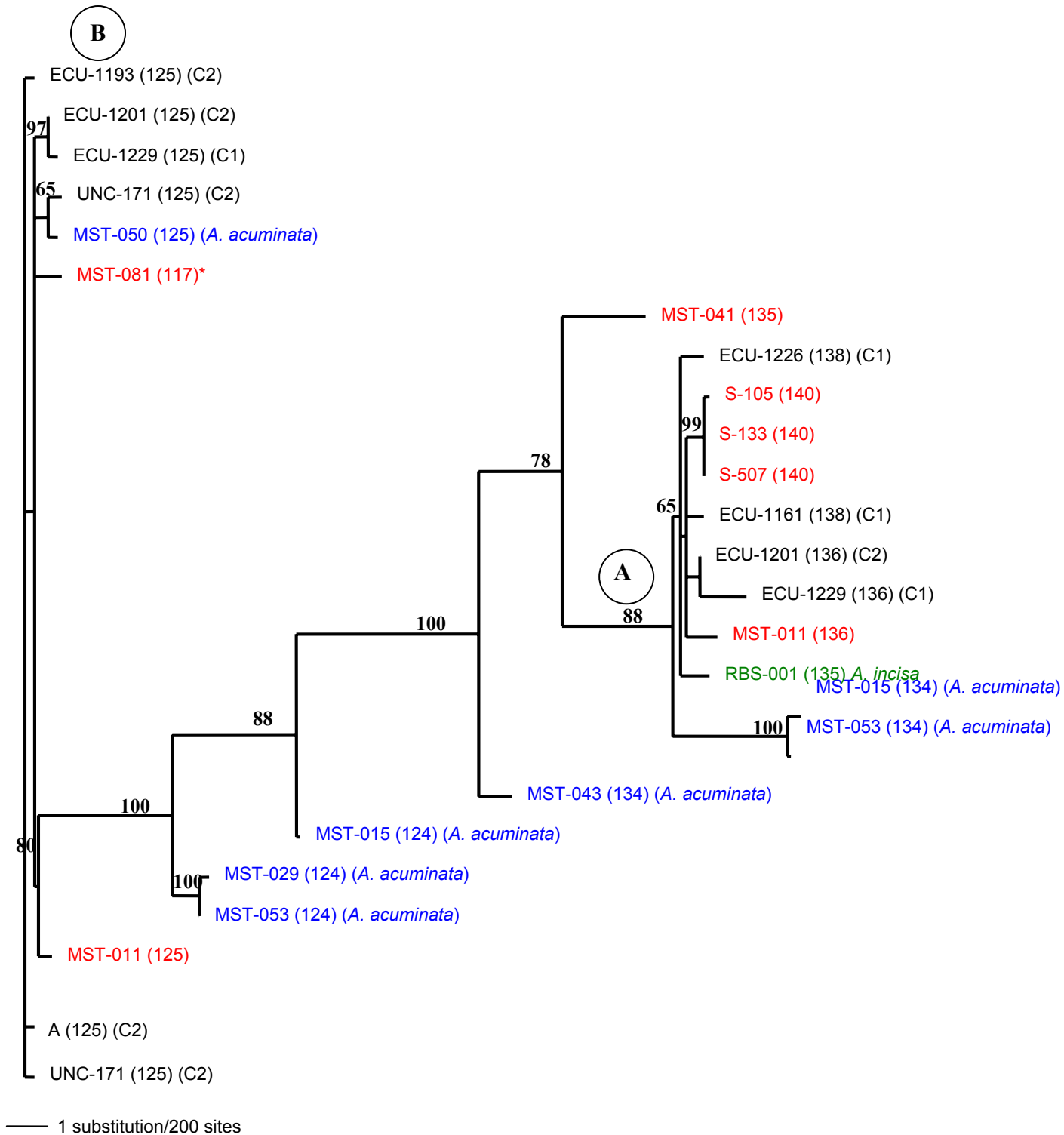


Figure 26. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 25 allèles du locus SSR AxC29 isolés chez *A. xanthorrhiza* et espèces sauvages apparentées (*A. incisa* et *A. acuminata*, les autres espèces essayées n'ayant pas amplifié l'allèle). Les allèles sont indiqués par l'accèsion à partir de laquelle ils ont été isolés, puis la taille obtenue *in silico* avec les amorces utilisées en génotypage sur les données de séquence. Les formes cultivées sont en noir et classées entre parenthèse en C1 ou C2 selon les deux groupes révélés par les AFLPs. Les formes sauvages de *A. xanthorrhiza* sont en rouge dont les monogamiques distingués par *. Les pourcentages de *bootstrap* ont été calculés après 100 répliquions.

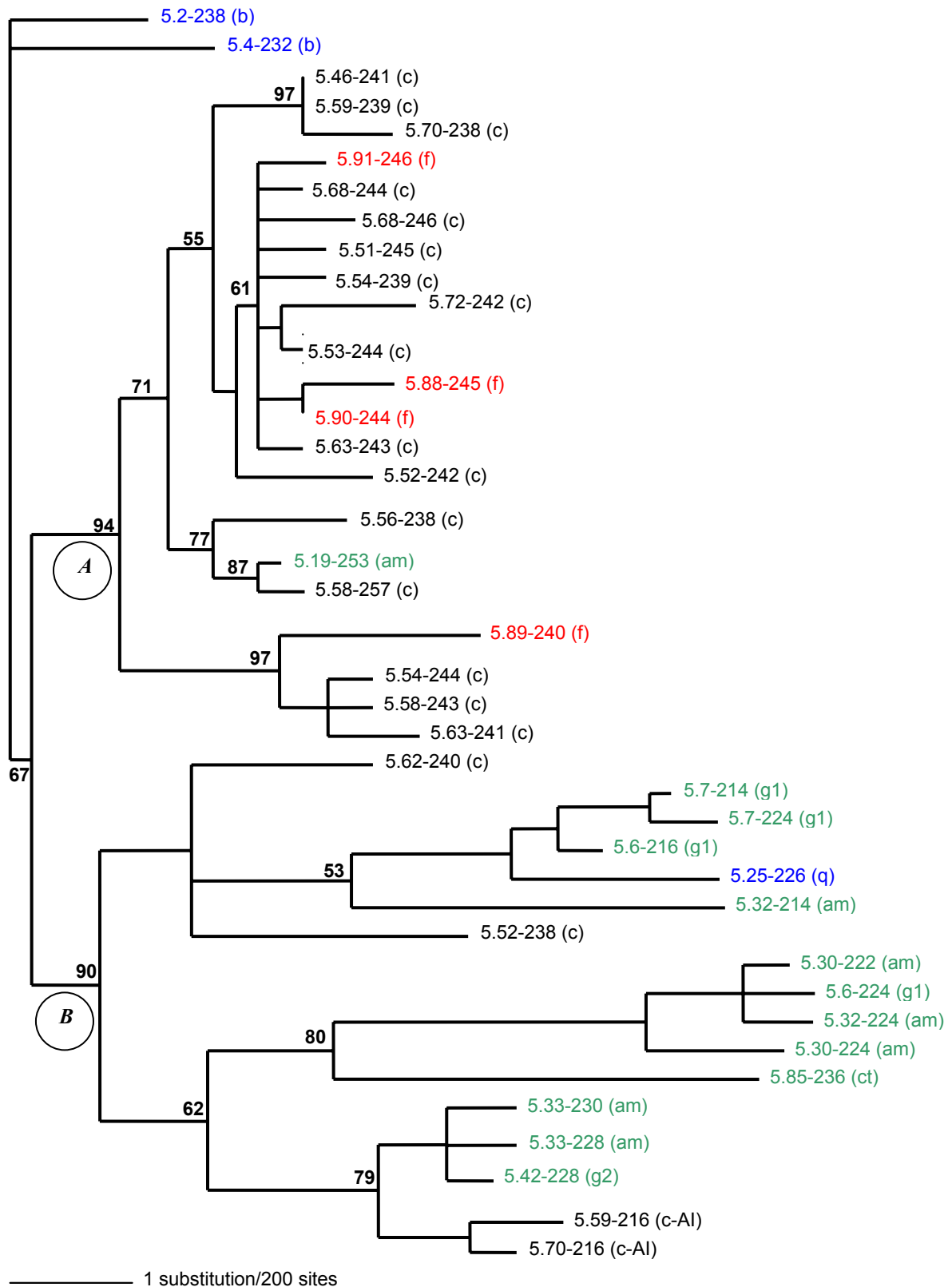


Figure 27. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 40 allèles du locus SSRY-3 chez le manioc. Les allèles sont indiqués par l'accèsion dont ils ont été isolés, puis la taille obtenue *in silico* avec les amorces utilisées en génotypage sur les données de séquence. Les formes cultivées sont indiquées par C et de plus AI pour celles artificiellement introgressées et am pour les « arbres manioc »; les formes sauvages par f pour *M. flabellifolia*, g pour *M. glaziovii*, ct pour *M. carthagenensis*, b pour *M. brachyloba* et q pour *M. quinquepartita*. Les pourcentages de *bootstrap* ont été calculés après 100 réplifications.

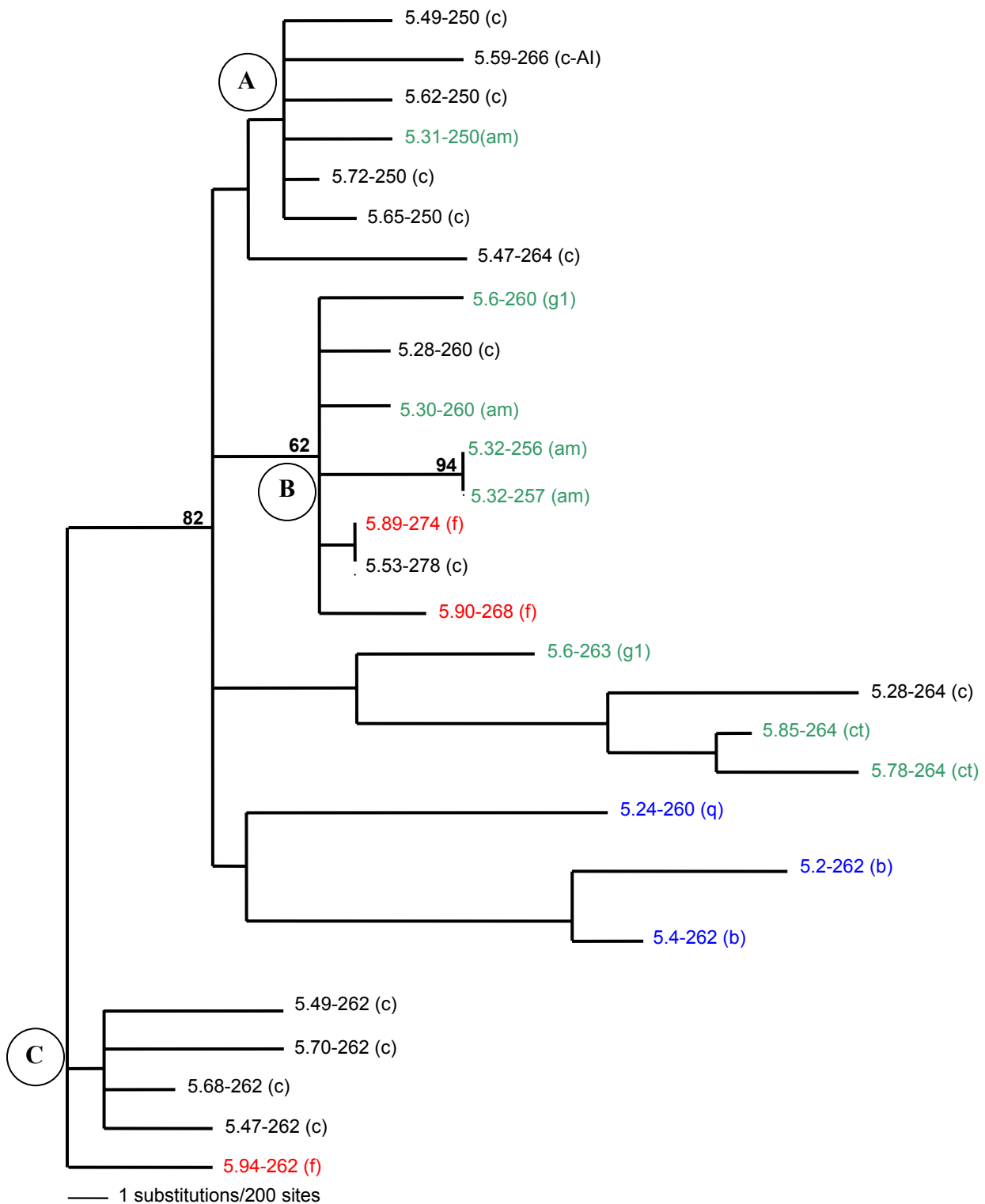


Figure 28. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 27 allèles du locus SSRY-9 chez le manioc. Les allèles sont indiqués par l'accèsion dont ils ont été isolés, puis la taille obtenue *in silico* avec les amorces utilisées en génotypage sur les données de séquence. Les formes cultivées sont indiquées par C et de plus AI pour celles artificiellement introgressées et am pour les « arbres manioc »; les formes sauvages par f pour *M. flabellifolia*, g pour *M. glaziovii*, ct pour *M. carthagensis*, b pour *M. brachyloba* et q pour *M. quinquepartita*. Les pourcentages de *bootstrap* ont été calculés après 100 réplifications

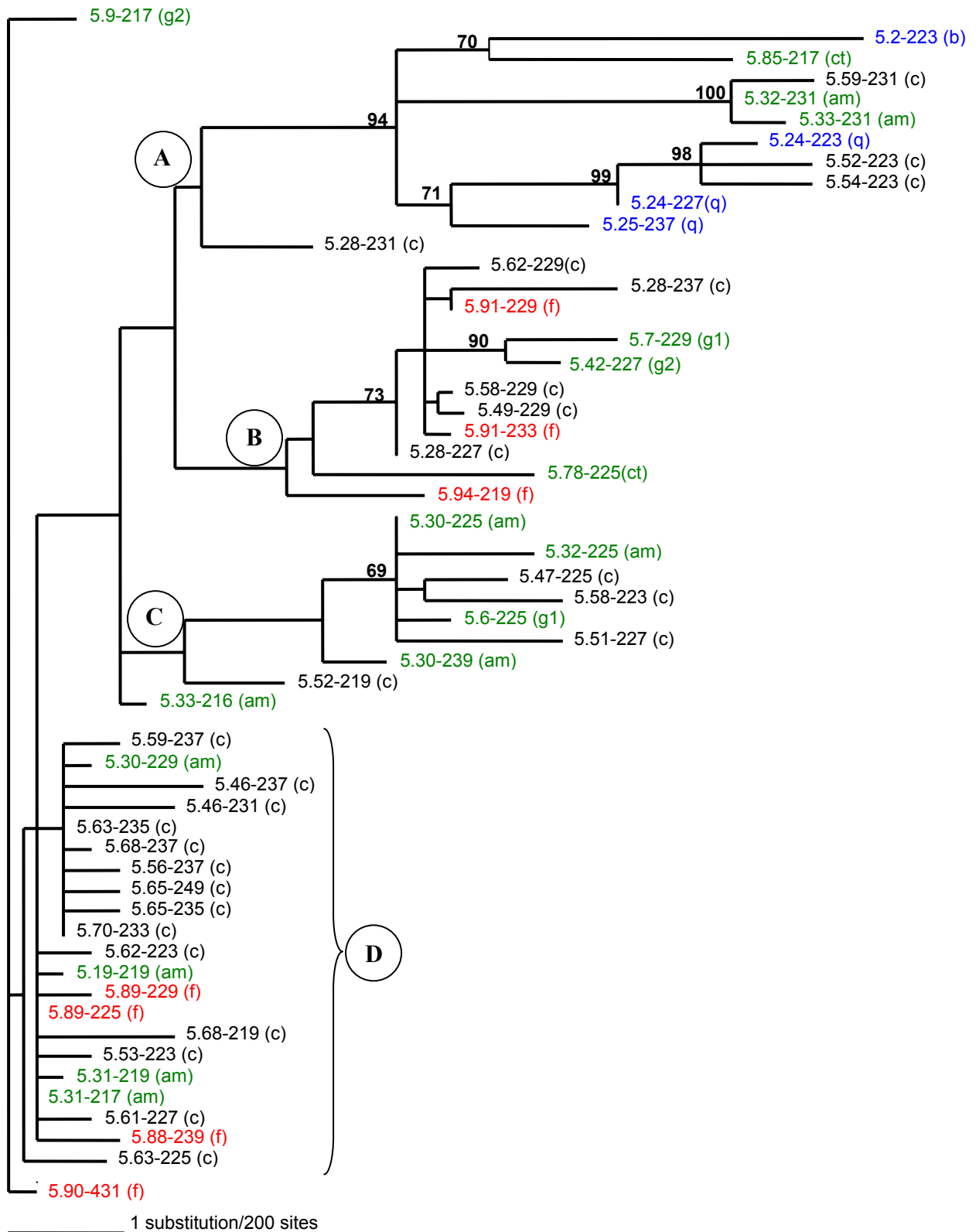


Figure 29. Phylogramme obtenu par Maximum de Vraisemblance à partir des régions flanquantes des 54 allèles du locus SSRY-40 chez le manioc. Les allèles sont indiqués par l'accésion dont ils ont été isolés, puis la taille obtenue *in silico* avec les amorces utilisées en génotypage sur les données de séquence. Les formes cultivées sont indiquées par C et de plus AI pour celles artificiellement introgressées et am pour les « arbres manioc »; les formes sauvages par f pour *M. flabellifolia*, g pour *M. glaziovii*, ct pour *M. carthaginensis*, b pour *M. brachyloba* et q pour *M. quinquepartita*. Les pourcentages de *bootstrap* ont été calculés après 100 répliques.

ANNEXES

Annexe 1. Liste d'échantillons utilisés pour l'étude de diversité de l'arracacha. Collections source: INIAP, INIAP-IRD, SRK

No. ref.	Code Accesion	Status	Variété	N	Collection	Pays	Province	Canton	Localité	ALT (m)	LAT	LONG	Date	Collecteurs	SSR	AFLP
C1	ECU- 1148	Cultivée	C. blanca	1	INIAP	ECU	Pichincha	Quito	Palmito Pamba	1500	00.11N	78.40W	1983.10	R. Castillo, C. Cañizares	x	x
C2	ECU- 1154	Cultivée	C. amarilla	1	INIAP	ECU	Bolívar	Guaranda	Negroyaco	2760	01.34S	79.00W	1984.02	E. Peralta, J. Urbano, N. Diaz	x	x
C3	ECU- 1155	Cultivée	C. amarilla	1	INIAP	ECU	Imbabura	Otavaló	Angla	2920	00.13N	78.09W	1983.06	R. Castillo, G. Garcia, E. Cazar	x	
C4	ECU- 1157	Cultivée	C. blanca	1	INIAP	ECU	Cañar	Biblian	Aguilan	2700	02.42S	78.52W	1982.10	R. Castillo, J. Rea, R. Padrún	x	x
C5	ECU- 1158	Cultivée	C. blanca	1	INIAP	ECU	Cañar	Biblian	Mangon	2800	02.42S	78.56W	1982.10	R. Castillo, J. Rea, R. Padrún	x	x
C6	ECU- 1161	Cultivée	C. blanca	1	INIAP	ECU	Imbabura	Antonio Ante	Santa Clara	2400	00.20N	78.14W	1984.03	R. Castillo	x	
C7	ECU- 1162	Cultivée	C. blanca	1	INIAP	ECU	Cañar	Biblian	Mangon	2750	02.42S	78.56W	1982.11	C. Nieto	x	x
C8	ECU- 1163	Cultivée	C. blanca	1	INIAP	ECU	Imbabura	Ibarra	El Carmen	2340	00.22N	78.11W	1984.03	R. Castillo	x	
C9	ECU- 1173	Cultivée	C. morada	1	INIAP	ECU	Azuay	Girón	Cofrada	2300	03.08S	79.08W	1984.06	E. Peralta, E. Cazar	x	x
C10	ECU- 1174	Cultivée	C. blanca	1	INIAP	ECU	Carchi	Mira	Guaquer	2150	00.35N	78.05W	1984.04	R. Castillo, J. Sanchez	x	
C11	ECU- 1178	Cultivée	C. morada	1	INIAP	ECU	Azuay	Cuenca	Ochangata	2500	02.54S	78.56W	1984.06	E. Peralta, C. Cazco	x	
C12	ECU- 1179	Cultivée	C. morada	1	INIAP	ECU	Azuay	Giron	Leocapac Chico	2110	03.11S	79.10W	1986.04	E. Peralta, E. Cazar	x	x
C13	ECU- 1180	Cultivée	C. blanca	1	INIAP	ECU	Pichincha	Quito	Palogillo	2575	00.13S	78.20W	1983.06	C. Nieto, G. Garcia, N. Diaz	x	x
C14	ECU- 1181	Cultivée	C. amarilla	1	INIAP	ECU	Chimborazo	Riobamba	Balcash	3075	01.41S	78.34W	1984.08	R. Castillo, G. Garcia	x	
C15	ECU- 1182	Cultivée	C. blanca	1	INIAP	ECU	Chimborazo	Penipe	Matus Alto	2800	01.34S	78.31W	1984.08	R. Castillo, G. Garcia	x	
C16	ECU- 1188	Cultivée	C. amarilla	1	INIAP	ECU	Chimborazo	Riobamba	Sasapui	2990	01.43S	78.35W	1983.07	C. Nieto, G. Garcia, M. Sola	x	x
C17	ECU- 1193	Cultivée	C. morada	1	INIAP	ECU	Pichincha	Cayambe	Monteserrín Alto	3400	00.03S	78.16W	1987.01	C. Nieto, N. Diaz, G. Garcia	x	x
C18	ECU- 1194	Cultivée	C. blanca	1	INIAP	ECU	Imbabura	Pimampiro	Mataqui	1800	00.25N	77.56W	1983.07	A. Carrera	x	
C19	ECU- 1197	Cultivée	C. blanca	1	INIAP	ECU	Carchi	Montufar	Piartal	2840	00.34N	77.46W	1983.07	R. Castillo, A. Chulde	x	
C20	ECU- 1199	Cultivée	C. amarilla	1	INIAP	ECU	Imbabura	Otavaló	Cubilche	2870	00.14N	78.08W	1983.06	C. Nieto, N. Diaz, G. Garcia	x	
C21	ECU- 1201	Cultivée	C. morada	1	INIAP	ECU	Loja	Gonzanami	La Avenida	2040	04.14S	79.26W	1984.07	A. Ortega, R. Castillo, J. Sanchez	x	
C22	ECU- 1205	Cultivée	C. amarilla	1	INIAP	ECU	Chimborazo		Curilla	2990	01.43S	78.35W	1985.01	R. Castillo, J. Sanchez, F. Paredes	x	
C23	ECU- 1206	Cultivée	C. blanca	1	INIAP	ECU	Bolívar	Guaranda	Pungoloma	3040	01.31S	78.59W	1985.01	R. Castillo, J. Sanchez, F. Paredes	x	
C24	ECU- 1214	Cultivée	C. blanca	1	INIAP	ECU	Cañar	Biblian	Zhalao	2700	02.43S	78.55W	1984.09	E. Peralta, C. Nieto	x	
C25	ECU- 1217	Cultivée	C. morada	1	INIAP	ECU	Loja	Saraguro	Buenaventura	2750	03.34S	79.20W	1986.07	E. Peralta, J. Tola	x	
C26	ECU- 1218	Cultivée	C. morada	1	INIAP	ECU	Loja	Saraguro	Las Lagunas	2580	03.38S	79.14W	1986.07	E. Peralta, J. Tola	x	
C27	ECU- 1226	Cultivée	C. amarilla	1	INIAP	ECU	Azuay	Sigsig	Las Acacias	2600	03.03S	78.48W	1987.02	F. Leon	x	x
C28	ECU- 1229	Cultivée	C. morada	1	INIAP	ECU	Cotopaxi	Latacunga	Fligüilo	2700	00.42S	78.52W	1987.02	C. Nieto, M. Huaman, F. Leon	x	x
C29	ECU- 1232	Cultivée	C. blanca	1	INIAP	ECU	Cotopaxi	Pujili	Itoalú	2500	00.47S	78.54W	1987.04	M. Rivera	x	
C30	ECU- 1234	Cultivée	C. blanca	1	INIAP	ECU	Bolívar	Guaranda	Chuquisungo	3600	01.17S	78.58W	1987.06		x	
C31	ECU- 2484	Cultivée	C. blanca	1	INIAP	ECU	Tungurahua	Ambato	Mercado Ambato	2550	01.15S	78.37W	1990.02		x	
C32	ECU- 6660	Cultivée	C. blanca	1	INIAP	ECU	Loja	Loja	Shalshi	2400			1992.02	C. Tapia, J. Velasquez	x	x
C33	ECU- 8750	Cultivée	C. blanca	1	INIAP	ECU	Loja	Loja	Barrio La Guangara	2400			1993.09	C. Tapia, E. Cazar, J. Velasquez	x	
C34	ECU- 8753	Cultivée	C. blanca	1	INIAP	ECU	Azuay	Sigsig	S.J. de Paranga	2900			1993.09	C. Tapia, E. Cazar, J. Velasquez	x	
C35	ECU- 8758	Cultivée	C. blanca	1	INIAP	ECU	Bolívar	Guaranda	Julio Moreno	2900	01.35S	79.01W	1993.09	C. Tapia, E. Cazar, J. Velasquez	x	x

C36	ECU-	8763	Cultivée	C. blanca	1	INIAP	ECU	Carchi	Bolivar	Alor	2300				1993.09	C. Tapia, N. Mazon	x	
C37	JEE-041		Cultivée	C. amarilla	1	INIAP	ECU										x	x
C38	Z-1		Cultivée	C. morada	1	INIAP	ECU										x	
C39	Z-2		Cultivée	C. blanca	1	INIAP	ECU										x	
C40	MST-002		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Loja	Saraguro	Apuguin	2400	3.36.908 S	79.14.155 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C41	MST-003		Cultivée	C. morada	1	INIAP-IRD	ECU	Loja	Saraguro	Apuguin	2400	3.36.908 S	79.14.155 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C42	MST-010		Cultivée	C. morada	1	INIAP-IRD	ECU	Loja	Loja	Gualel	2650	3.46.240 S	79.22.913 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C43	MST-012-A		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Loja	Loja	Gualel	2650	3.46.240 S	79.22.913 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C44	MST-012-B		Cultivée	C. blanca	1	INIAP-IRD	ECU	Loja	Loja	Gualel	2650	3.46.240 S	79.22.913 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C45	MST-016		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Zamora Ch.	Palanda		1300	2.11.740 S	78.50.692 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C46	MST-020		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Zamora Ch.	Zumba	Las Cidras	950	4.51.599 S	79.08.104 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x	
C47	MST-021		Cultivée	C. blanca	1	INIAP-IRD	ECU	Zamora Ch.	Zumba	Las Cidras	950	4.51.599 S	79.08.104 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C48	MST-022		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Zamora Ch.	Zumba	Las Cidras	950	4.51.599 S	79.08.104 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C49	MST-023		Cultivée	C. blanca	1	INIAP-IRD	ECU	Zamora Ch.	Zumba	Las Cidras	2000	4.51.599 S	79.08.104 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C50	MST-026		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Loja	Loja	Quebrada de Suro	1940	4.20.741 S	79.10.927 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C51	MST-027		Cultivée	C. blanca	1	INIAP-IRD	ECU	Loja	Loja	San Fr.Alto	950	4.11.136 S	79.14.814 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C52	MST-030		Cultivée	C. blanca	1	INIAP-IRD	ECU	Zamora Ch.	Zamora	Timbara	2000	4.02.618 S	78.53.656 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x	
C53	MST-037		Cultivée	C. morada	1	INIAP-IRD	ECU	Azuay	Cuenca	Vía Soldados	2900	3.41.801 S	78.53.685 W	2002.08	E.Morillo, G.Second, E.Tapia	x		
C54	MST-040		Cultivée	C. amarilla	1	INIAP-IRD	ECU	Azuay	Gualaceo		2350			2002.08	E.Morillo, G.Second, E.Tapia	x		
C55	MST-047		Cultivée	C. blanca	1	INIAP-IRD	ECU	Pichincha	SJ Minas	Lote La Calera	1960	0.08.58 N	72.24.15 W	2002.11	E.Morillo, G.Second	x	x	
C56	MST-049		Cultivée	C. blanca	1	INIAP-IRD	ECU	Pichincha	SJ Minas	Lote La Calera	1960	0.08.58 N	72.24.15 W	2002.11	E.Morillo, G.Second	x		
C57	MST-055		Cultivée	C. morada	1	INIAP-IRD	ECU	Pichincha	SJ Minas	Lote La Calera	1960	0.08.58 N	72.24.15 W	2002.11	E.Morillo, G.Second	x		
C58	MST-056		Cultivée	C. blanca	1	INIAP-IRD	ECU	Tungurahua	Baños		2300			2002.12	E.Morillo, G.Second	x		
C59	UNC161		Cultivée		1	SRK	PER	Cajamarca	Cutervo	Minas	2200	6.17.10S	78.41.05W		Juan Seminario	x		
C60	AM		Cultivée	C. amarilla	1	SRK	PER	Cajamarca	Cutervo	Minas	2470	6.16.53S	78.41.28W	2001.09	S. Knudsen	x		
C61	AP		Cultivée		1	SRK	Per	Cajamarca		Minas	2470	6.16.53S	78.41.25W	2001.09	S.Knudsen	x		
C62	A		Cultivée	C. amarilla	1	SRK	Per	Cajamarca		Minas	2200	6.17.10S	78.41.05W		J.Seminario	x	x	
C63	NA		Cultivée		1	SRK	Per	Cajamarca		Minas	2470	6.16.53S	78.41.25W	2001.09	S.Knudsen	x	x	
C64	510		Cultivée		1	SRK	Per									x		
C65	UNC-2		Cultivée		1	SRK	Per									x		
C66	UNC-43		Cultivée		1	SRK	Per	Cajamarca	chota	Churucancha	2050	6.30.20S	78.43.59w	35065	J.Seminario	x		
C67	UNC-120		Cultivée		1	SRK	Per									x	x	
C68	UNC-171		Cultivée		1	SRK	Per									x	x	
C69	MTC 7-8		Cultivée		1	SRK	Per	Cajamarca		Minas	2470	6.16.53S	78.41.25W	2001.09	S.Knudsen	x	x	
C70	M		Cultivée	C. morada	1	SRK	Per	Cajamarca		Minas	2470	6.16.53S	78.41.25W	2001.09	S.Knudsen	x		
C71	CE3-5		Cultivée		1	SRK	Per	Cajamarca		Minas	2200	6.17.10S	78.41.05W		J.Seminario	x		
C72	CM		Cultivée		1	SRK	Per	Cajamarca		Minas	2200	6.17.10S	78.41.05W	2001.09	J.Seminario	x		
C73	C3-9		Cultivée		1	SRK	Per	Cajamarca		Minas	2470	6.16.53S	78.41.125W		S.Knudsen	x		

C74	UNC-123	Cultivée		1	SRK	Per	Cajamarca	Celendin	Huañambra	2780	6.53.05S	78.09.54W		M.Valderrama		x
C75	ECU-1196	Cultivée	C. blanca	1	INIAP	ECU	Loja	Celica	Barrio San Vicente	1996	04.06S	79.57W	1984.07	R. Castillo, A. Ortega, J. Sanchez		x
C76	ECU-3295	Cultivée	C. blanca	1	INIAP	ECU	Loja		Loja	2064	03.59S	79.11W	1990	R. Castillo		x
S1	MST-004/006	Sauvage	Polycarpique ?	10	INIAP-IRD	ECU	Loja	Saraguro	Apuguin	7900	3.36.908 S	79.14.155 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S2	MST-011	Sauvage	Polycarpique	1	INIAP-IRD	ECU	Loja	Loja	Gualel	2650	3.46.240 S	79.22.913 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S3	MST-025	Sauvage	Polycarpique	1	INIAP-IRD	ECU	Loja	Loja	Quebrada de Suro	2000	4.20.741 S	79.10.927 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S4	MST-039	Sauvage	Polycarpique ?	3	INIAP-IRD	ECU	Cañar	Tambo		2950	2.28.676 S	79.00.156 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S5	MST-041	Sauvage	Polycarpique	4	INIAP-IRD	ECU	Cañar	Alausí		2405	2.11.740 S	78.50.692 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S6	MST-042	Sauvage	Polycarpique	6	INIAP-IRD	ECU	Cañar	Huigra		2330	2.14.205 S	78.54.315 W	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S7	MST-044	Sauvage	Polycarpique	1	INIAP-IRD	ECU	Tungurahua	Ambato (marché)		-	-	-	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S8	MST-045	Sauvage	Polycarpique	1	INIAP-IRD	ECU	Cotopaxi	Pujilí (marché)		-	-	-	2002.08	E.Morillo, G.Second, E.Tapia	x	x
S9	MST-057	Sauvage	Polycarpique	4	INIAP-IRD	ECU	Bolivar	Guaranda	Com.Jatunpamba	3050	1.32.372 S	78.57.900 W	2002.12	E.Morillo, G.Second	x	x
S10	MST-058	Sauvage	Polycarpique	4	INIAP-IRD	ECU	Bolivar	Guaranda		3100	1.29.439 S	79.01.691 W	2002.12	E.Morillo, G.Second	x	
S11	MST-070	Sauvage	Polycarpique?	3	INIAP-IRD	ECU	Loja	Ona		2900			2003.12	E.Morillo, E.Tapia	x	x
S12	MST-072	Sauvage	Polycarpique?	8	INIAP-IRD	ECU	Loja	Saraguro		2500			2003.12	E.Morillo, E.Tapia	x	x
S13	MST-074	Sauvage	Monocarpique ?	6	INIAP-IRD	ECU	Loja	Manu		2350			2003.12	E.Morillo, E.Tapia	x	x
S14	MST-075	Sauvage	Monocarpique	7	INIAP-IRD	ECU	Loja	Loja		2600			2003.12	E.Morillo, E.Tapia	x	x
S15	MST-076	Sauvage	Monocarpique	6	INIAP-IRD	ECU	Loja	Payanchi		2600			2003.12	E.Morillo, E.Tapia	x	x
S16	MST-083	Sauvage	Polycarpique?	2	INIAP-IRD	ECU	Loja	Calvas		2000			2003.12	E.Morillo, E.Tapia	x	x
S17	MST-085	Sauvage	Polycarpique ?	1	INIAP-IRD	ECU	Loja	Calvas		2050			2003.12	E.Morillo, E.Tapia	x	x
S18	S-9	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Cutervo	Socota	2230	06.17.42S	78.41.59W	2000.10	S.Knudsen	x	x
S19	S-35	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Cajamarca	Cumbe Mayo	3550	07.11.36S	78.33.18W	2000.06	S.Knudsen	x	
S20	S-105	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Contumaza	El tunel, along road	2700	07.23.67S	78.47.09W	2000.07	S.Knudsen	x	
S21	S-115	Sauvage	Polycarpique	1	SRK	PER	Cajamarca			2675	06.48.02S	77.56.50W	2000.07	S.Knudsen	x	
S22	S-133	Sauvage	Polycarpique	1	SRK	PER	Cajamarca			3200-3250	07.05.30S	78.20.39W	2000.06	S.Knudsen	x	
S23	S-141	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Contumaza	Ollada verde	2950	7.20.58S	78.48.49W	2000.07	S.Knudsen	x	
S24	S-152	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Contumaza	Vuelta de los tajos	2720	7.19.70S	78.48.62W	2000.07	S.Knudsen	x	
S25	S-161	Sauvage	Polycarpique	1	SRK	PER	Cajamarca		Ingañan	2420	7.17.27S	78.29.82W	2000.07	S.Knudsen	x	
S26	S-182	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Cutervo	Socota	2080	06.17.99S	78.42.42W	2000.10	S.Knudsen	x	
S27	S-185	Sauvage	Polycarpique	1	SRK	PER	Amazonas		along road	2800	06.47.29s	77.56.69w	2000.07	S.Knudsen	x	
S28	S-199	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Cajamarca	Chetilla	2700	07.08.60S	78.40.62		S.Knudsen	x	
S29	S-507	Sauvage	Polycarpique	1	SRK	PER	Cajamarca			2600	07.16.35S	78.31.08W	2000.07	S.Knudsen	x	
S30	S-512	Sauvage	Polycarpique	1	SRK	PER								S.Knudsen	x	
S31	S-513	Sauvage	Polycarpique	1	SRK	PER				2385	06.17.3S	78.34.90X	2000.10	S.Knudsen	x	
S32	S-514	Sauvage	Polycarpique	1	SRK	PER								S.Knudsen	x	
S33	S-515	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Cajamarca	San Juan	2580	07.18.26S	78.27.70w	2000.07	S.Knudsen	x	
S34	S-516	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Contumaza		1840	7.26.02S	78.47.54W	2000.07	S.Knudsen	x	
S35	S-517	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Contumaza	Vuelta de los tajos	2720	7.19.70S	78.48.62W	2000.07	S.Knudsen	x	

S36	S-518	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Contumaza		2400	7.20.32S	78.47.34W	2000.07	S.Knudsen	x	
S37	MV-1071	Sauvage	Polycarpique	1	SRK	PER	Cajamarca	Chota	Cusilguan	2300	6.25.00S	79.00.32W		J.Seminario, M. Valderram.	x	
S38	MV-1075	Sauvage	Polycarpique	1	SRK	Per	Cajamarca	Celendin	Huañambra	2780	6.53.05S	78.09.54W	2001.03	M.Valderrama	x	
S39	MST-081	Sauvage	Monocarpique	8	SRK	ECU	Loja	Gimbura		1600	4.35.566S	79.26.566W	2003.12	E.Morillo, E.Tapia	x	x
S40	MST-082	Sauvage	Monocarpique	8	SRK	ECU	Loja	Gimbura		2000	4.37.483S	79.27.611W	2003.12	E.Morillo, E.Tapia	x	x
S41	S-168	Sauvage	Monocarpique	1	SRK	PER	Amazonas		Ruinas de Kuelap	3070	06.24.95S	77.55.41W	2000.07	S.Knudsen	x	
S42	S-183	Sauvage	Monocarpique	1	SRK	PER	Cajamarca	Cutervo	Minas Altas	2480	06.17.68S	78.51.30W	2000.01	S.Knudsen	x	x
S43	S-500	Sauvage	Monocarpique	1	SRK	PER	Cajamarca	Cutervo	Socota	1910	06.18.94S	78.41.03W	2000.10	S.Knudsen	x	x
S44	S-510	Sauvage	Monocarpique	1	SRK	PER	Cajamarca	Cutervo		2300	06.17.06S	78.41.43W	2000.06	S.Knudsen	x	x
S45	75-1	Sauvage	Monocarpique	1	SRK	PER	Cuzco								x	x
S46	75-2	Sauvage	Monocarpique	1	SRK	PER									x	
S47	31-2	Sauvage	Monocarpique	1	SRK	PER	Amazonas								x	x
S48	RBS-001	Sauvage	A.incisa	3	SRK	PER									x	x
S49	MST-015	Sauvage	A.acuminata	1	INIAP-IRD	ECU	Loja	Loja		2250	4.28.490 S	79.08.897 W	2002.08	E.Morillo, G.Second, E.Tapia		x
S50	MST-029	Sauvage	A.acuminata	1	INIAP-IRD	ECU	Loja	Loja	Podocarpus	2100	3.58.522 S	79.08.486 W	2002.08	E.Morillo, G.Second, E.Tapia		
S51	MST-036	Sauvage	A.elata	1	INIAP-IRD	ECU	Azuay	Cuenca	Laguna Yabiuco	3050			2002.08	E.Morillo, G.Second, E.Tapia		x
S52	MST-043	Sauvage	A.acuminata?	1	INIAP-IRD	ECU	Chimb.	Colta	Talalag	2100	1.46.658 S	78.59.381 W	2002.08	E.Morillo, G.Second, E.Tapia		
S53	MST-050	Sauvage	A.acuminata	1	INIAP-IRD	ECU	Imbabura	Otavallo		3000	0.12.314 S	78.21.377 W	2002.12	E.Morillo, G.Second		x
S54	MST-051	Sauvage	A.moschata	2	INIAP-IRD	ECU	Carchi	El Angel	Paramo El Angel	3400	0.38.998 S	77.53.404 W	2002.12	E.Morillo, G.Second		x
S55	MST-052	Sauvage	A.acuminata	1	INIAP-IRD	ECU	Imbabura	Otavallo	Lago Cuicocha	3050			2002.12	E.Morillo, G.Second		
S56	MST-053	Sauvage	A.acuminata	1	INIAP-IRD	ECU	Imbabura	Otavallo	San Juan Inguincho	3200			2002.12	E.Morillo, G.Second		
S57	MST-054	Sauvage	A.moschata	1	INIAP-IRD	ECU	Pichincha	Cayambe	Com. Pesilla	3200			2002.12	E.Morillo, G.Second		

No. ref.	Plant code	Status	Parents	Source	SSR
H1	208	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H2	209	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H3	211	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H4	214	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H5	219	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H6	221	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H7	225	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H8	226	Presumed F1 hybrid	UNC123 (C) x MV1071(SP)	SRK	x
H9	229	Presumed F1 hybrid	AM (C) x S-514(SP)	SRK	x
H10	230	Presumed F1 hybrid	AM (C) x S-514(SP)	SRK	x
H11	813	Presumed F1 hybrid	AM (C) x S-514(SP)	SRK	x
H12	820	Presumed F1 hybrid	AM (C) x S-514(SP)	SRK	x
H13	822	Presumed F1 hybrid	AM (C) x S-514(SP)	SRK	x

Annexe 2. Primer Note publiée dans Molecular Ecology Notes

**“Development of DNA microsatellite markers in the Andean root crop
arracacha: *Arracacia xanthorrhiza* Banc. (Apiaceae)”**

Eduardo MORILLO, Gérard SECOND, Jean Louis PHAM et A. Marie
RISTERUCCI

Molecular Ecology Notes, 4: 680-682 (2004)

PRIMER NOTE

Development of DNA microsatellite markers in the Andean root crop arracacha: *Arracacia xanthorrhiza* Banc. (Apiaceae)

E. MORILLO,*† G. SECOND,† J. L. PHAM† and A. M. RISTERUCCI‡

*Departamento Nacional de Recursos Fitogenéticos y Biotecnología (DENAREF), Instituto Nacional Autónomo de Investigaciones Agropecuarias (INIAP), Quito, Ecuador, †IRD (Institut de Recherche pour le Développement), UMR 1097 DGPC/DYNADIV, BP 64501, 34394 Montpellier Cedex 5, France, ‡CIRAD-AMIS/BIOTROP UMR 1096, 34398 Montpellier, France

Abstract

A microsatellite-enriched library was constructed in the Andean root crop arracacha (*Arracacia xanthorrhiza* B). Of 18 loci tested, 14 were found to be polymorphic after screening for diversity in different cultivars and related wild forms. Allelic diversity in the crop was low but the transferability of the primers to closely related wild forms was good. The loci reported here are the first genetic markers to be published for this species and will be useful for future germplasm characterization and studies of genetic diversity.

Keywords: Andean crop, arracacha, *Arracacia xanthorrhiza*, microsatellites, zanahoria blanca

Received 14 July 2004; revision accepted 3 August 2004

Arracacha (*Arracacia xanthorrhiza* B., tetraploid species: $2n = 4x = 44$) is a vegetatively propagated carrot and the only Apiaceae species domesticated in the New World. Among nine minor Andean root and tuber species, arracacha is considered the most promising crop (Hermann 1997). To date, its genetic variability has been estimated using morphological characteristics, isozymes and random amplified polymorphic DNA (RAPD) markers, leading to contrasting conclusions. In order to study the genetic structure of the cultivated pool and its relation to wild relatives (*A. xanthorrhiza* species complex; Randers *et al.* 2003), we have isolated and characterized microsatellite markers, which we present here.

The microsatellite-enriched library was built following the procedure of Billote *et al.* (1999). Total genomic DNA was extracted from a white root cultivar using the PEX method (Jhingan 1992). Prior to amplification, DNA was digested with *RsaI* and ligated with adapters (5'-CTCTT-GCTTACGCGTGGACTA-3' and 5'-TAGTCCACGCG-TAAGCAAGAGCACA-3'). Selection of simple sequence repeat (SSR) fragments was carried out with a biotinylated SSR primer. Selected fragments were cloned into a pGEM-

T plasmid (Promega A3600) and transformed in *Escherichia coli* XL1-Blue MRF' supercompetent cells (Stratagene). In total, 192 white transformant clones were transferred to Hybon-N⁺ nylon membranes (Amersham), and hybridized using labelled P32 microsatellite oligoprobe (GA)₁₅ and (GT)₁₅. For the 74 clones that gave a satisfactory positive signal, the insert was sequenced in one direction. DNA sequence alignments and primer design were performed using DNASTAR 1.2. Primer pairs were synthesized for 26 microsatellite sequences of which 18 amplified the expected size fragment. SSR amplification was performed on 58 DNA samples of arracacha from Ecuador. SSR products were analysed following the protocol of Roy *et al.* (1996) applied with the automated infrared fluorescence technology of a sequencer (IR2; LI-COR Biosciences). For a given SSR locus, the forward primer was 5'-end labelled with a M13 extension (5'-CACGACGTTGTAACACGAC-3'). Amplifications were performed in 10 µL of 20 ng genomic DNA, 0.02 µM of the M13-labelled primer, 0.3 µM of the reverse primer, 0.06 µM of M13 primer-fluorescent dye IR700 or IR800, 2.5 mM MgCl₂, 0.2 mM dNTP and 0.5 U *Taq* DNA polymerase (Promega). A Biometra thermocycler was used with the following cycling conditions: 94 °C (1 min), 30 cycles at 94 °C for 60 s, 45–60 °C for 60 s, 72 °C for 60 s and a final elongation step of 72 °C for 7 min. IR700 or IR800-labelled polymerase chain reaction (PCR) products were diluted 10-fold in its loading buffer, subject to electrophoresis in a 6.5% polyacrylamide gel and then sized by the

Correspondence: E. Morillo, IRD (Institut de Recherche pour le Développement), UMR DGPC/DYNADIV, 911 Av. Agropolis BP 64501, 34394 Montpellier, Cedex 5, France.
E-mail: morillo@mpl.ird.fr or denaref@cnet.ec

2 PRIMER NOTE

Table 1 Characteristics of 14 microsatellite loci isolated from *Arracacia xanthorrhiza* B

Locus	GenBank Accession no.	Clone	Primer sequences (5'-3')	Motif	Annealing temp (°C)	PCR band size (bp)	H_O	H_E
AxC27	AY530811	338	F: AAGTTCGTATTGTGCTGCTGTAT R: ATTGTGGCGTGGATGTGAAAAAG	(GT) ₈	58	152, 150, 148	0.53	0.50
AxC38	AY530812	492	F: TAAAGCCTATAAACATCAAAA R: ACCCTCTCCCATCATAAC	(GT) ₈	48	186, 184	0.28	0.24
AxC85	AY530813	405	F: TTGGGACTGGAACCTTGT R: GTGTGGTGTATTAATAAT	(GT) ₈ N ₁₂ (GT) ₁₀	58	294, 292, 266	0.55	0.50
AxD34	AY530814	435	F: AAGCTACGGATATTTACTACAT R: AGCGGCTCTGATTTGAG	(GA) ₁₃	50	209, 215	0.50	0.48
AxD43	AY530815	279	F: AATGGTGGTGTAGGTTTGAAG R: AATTGTTTATCTGAGTCCGTTGGTA	(CA) ₁₅	58	206, 196, 184	0.13	0.15
AxD72	AY530816	571	F: GATACCAATAGCCGAAAGGAG R: AGGGGTGGAGTAGCAATGTT	(TA) ₈ (GT) ₈	58	184, 180, 172, 168	0.40	0.35
AxD82**	AY530817	503	F: TGGAGAGGCTAATGCAAAATACT R: ATAAGCAGAACGCAAAACGACAT	(GT) ₁₂ (GA) ₁₁	58	182, 174, 170	0.84	0.63
AxD85*	AY530818	570	F: TGCACGCATTGTGAACT R: CAAATGGACCTGGTATGT	(GA) ₈	48	192, 190, 180	0.77	0.65
AxD13	AY675515	392	F: AACTTTGAACATGGTCTATTACTT R: TGATGCCACGACAAAGATA	(CA) ₈	48	186	0.00	0.00
AxD55†	AY675514	570	F: AACCCGACTGAAATCCCAAAAT R: GCAAAAAGACCCGACCAATCAA	(GT) ₈	54	426, 422, 404, 276, 258	—	—
AxC64†	AY530819	385	F: GAAATGATCTGCAACTGGAT R: TCCATCCTTCTTACATAT	(CT) ₂₀	48	196, 194, 192, 175, 169, 163	—	—
AxC87†	AY530820	473	F: CTCGCAGATCATCTATAAAGT R: TTAACCTGCAAAAGGAGCAC	(CT) ₁₃	58	142, 140, 136, 128	—	—
AxC8	AY675512	531	F: GATCATTTCGCAAGGTAAGTCTC R: AGTCACCTTATGAAATGTTCTGTA	(CT) ₁₂	58	135 (?)	—	—
AxC29	AY675513	409	F: GCCCAATAGCCACAAG R: GAAATATCTGCTGGAAAGAGC	(CT) ₁₃ (TA) ₇	58	186 (?)	—	—

Data of PCR band size and observed (H_O) and expected (H_E) heterozygosity values were obtained in 58 cultivated accessions.

A polymorphic locus for which the number of alleles is not known is indicated by (?).

*Significant deviation from Hardy–Weinberg equilibrium ($P < 0.01$).

** Significant deviation from Hardy–Weinberg equilibrium ($P < 0.001$).

†Locus with 3 or 4 bands displayed per sample (H_E and H_O values not calculated).

IR fluorescence scanning system of the LI-COR sequencer. Scoring of allele size was performed using the software AFLP-QUANTAR™ PRO 1.0 (Keygen) by comparison with a 50–350 bp ladder (Cat. No. LI-COR 829-05343).

Of the 18 primer pairs that amplified successfully, 14 were polymorphic in the *A. xanthorrhiza* species complex. Three loci showed more than two alleles per sample and are currently being investigated (*AxD55*, *AxC64* and *AxD55*). Two loci (*AxC8* and *AxC29*) displayed slippage profiles and the number of alleles could not be determined exactly. Table 1 shows allelic variation in a survey of the diversity in 58 accessions from the INIAP germplasm bank. The number of alleles detected ranged from two to six. Observed and expected heterozygosities (H_O , H_E) under Hardy–Weinberg equilibrium were calculated for the loci displaying a diploid profile using GENEPOP 1.2 (Raymond & Rousset 1995). Only two loci (*AxD82* and *AxD85*) deviated significantly owing to an excess of heterozygotes. The

average H_E of 0.43 indicates significant levels of polymorphism given that no isozyme polymorphism has been observed in this crop (Mazon 1993) and that a higher polymorphism can be observed in the wild pool. Only two loci (*AxD34* and *AxD85*) showed significant linkage disequilibrium out of all possible pairs of loci. This linkage disequilibrium, as heterozygote excess mentioned above may be due, like observed in some other cultivated species, to the clonal propagation of the crop. Finally, the primers were tested on other species of the genus *Arracacia* collected in Ecuador, including the recently established wild forms of the *A. xanthorrhiza* complex (Randers *et al.* 2003): *A. x. var. andina* and *A. x. var. monocarpa*. Successful amplification in the closely related wild forms was observed in comparison with little or no amplification for other *Arracacia* species (Table 2). These SSR markers will be useful for studies of diversity of the cultivated pool in comparison to wild tuberous populations.

Table 2 Primer transferability of 14 SSR loci developed in arracacha (*Arracacia xanthorrhiza* Banc.) to other South American *Arracacia* species

	AxC27	AxC38	AxC85	AxD34	AxD43	AxD72	AxD82	AxD85	AxD13	AxD55	AxC64	AxC87	AxC8	AxC29
<i>A. x. var. xanthorrhiza</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>A. x. var. andina</i>	+	+	+	+	+	+	+	+	+	+	+–	+–	+	+
<i>A. x. var. monocarpa</i>	+	+	+	+	+	+	+	+	+	+	+–	+	+	+
<i>A. incisa</i>	+	+	+	+	–	+	+	+	+	–	+	+	–	+
<i>A. elata</i>	+	+	–	+	–	–	+	+	+	+	–	–	–	–
<i>A. acuminata</i>	+	+	+	+	+–	–	+	+–	+–	+	–	–	–	–
<i>A. moschata</i>	–	+–	–	+–	–	–	–	–	–	+	–	–	–	–

+ indicates amplification of expected product; – indicates no amplification.

Acknowledgements

The first author thanks the research team 'Anthropisation et Dynamique de la Diversité Génétique des Plantes (DYNADIV)' of the IRD for supporting this research and Steen Randers Knudsen for providing *A. incisa* samples.

References

- Billote N, Lagoda P, Risterucci A, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*, **54**, 277–288.
- Hermann M (1997) Arracacha (*Arracacia xanthorrhiza* Bancroft). In: *Andean Roots and Tubers: Ahipa, Arracacha, Maca and Yacon* (eds Hermann M, Heller J), pp. 75–172. International Plant Genetic Resources Institute, Rome.
- Jhingan A (1992) A novel technology for DNA isolation methods. *Molecular Cell Biology*, **3**, 15–22.
- Mazon N (1993) Análisis de la variación morfológica e isoenzimática de la colección ecuatoriana de zanahoria blanca (*Arracacia xanthorrhiza* B.). Thesis, Ing. Agr. ESPOCH, Riobamba, Ecuador.
- Randers SR, Sørensen M, Hermann M (2003) A revision of the South American species of the genus *Arracacia* Bancroft (Apiaceae). PhD thesis, Department of Ecology, KVL University, Denmark.
- Raymond M, Rousset F (1995) GENEPOP Ver. 0.2. Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Roy R, Steffens DL, Gratside B, Jang GY, Brumbaugh JA (1996) Producing STR locus patterns from bloodstains and other forensic samples using an infrared fluorescent automated DNA sequencer. *Journal of Forensic Sciences*, **41**, 418–424.

CLON SSRY-3 (de 520 à 1545 pb)

ATTAACTCATACTCATGTATGACCCATTGAGTGGCGTCGGGATTATGGAAAACCAAAGTTCTCTTTATTCCAATCTTCTCATCGCTATCTTGAGCCGTGACCA
AGCGGAGGGTACCTGTGCATTTCCAGTATCCACCTTTTGCCTTCTCTTAGCCAGGCCACTGTTCTTGTGATACAGGCGAAGACAAAAGAAATACCACACCT
TATCCCTAGATTTATTCTTGAAAGATCTACAAAAGCACATATCACCGTTAATTGTAGTCTGCGTAACAACAAAAAAAAAAAAAAAAAACACACACACACACAC
ACACACACACACACACACAGAGGAAGTGTATATAGGAATAGTTGTTACCTGGAAGATCCTTGGGATCATAATCGCATATTTGAACCTCGCCGATGTCCAAGC
CATCTATAGGATCACCACGCCTCTTCCGCTTCAAATAATGATCAACCAGTTCTTCATCGGACGGGTAGAATCTCCATCCAGGCGGCAAGTTTGGCGGGACAA
ACTGCGGTGAGCCTGTCGTACAGTTTAAAAAAAAACAAGAGCCACATTAGTTTCAGACATGAATTTAATACTAAATATAAGAATTTGACAACTTCTCTTCT
GAGCTAGCTTATACAGTACAGTAACGTGCGATCCAACTCTCTCCAGAAATTAAGTGGGGGAAAGAGAAAACAGCAATCGAATAGAATAAGCAAAAATGCACTCACCT
GAGGAGGAGTAGGAGGAGTAGGAGTTGTTGTTCACCCTCTTCTTAAAAACAGCAAGCCTCGGAATTGAAAAAGCAATAGAGTGAAGATCGTAATGTAAGTACTC
TTCCAGAAACAGCAAGTGCAGGAACTGAAATGTTACGAGGGAGAGGAAATTGCGATCGAGCGGAAGGGCATTAAAAAGGAAGCAGAAGCACTAGTGATCACT
TCTTTGAGGGTAGTGATGAGTGAGTCACTGTTCACTTACGTCACCCCTTATTGAAGAAGGGGGAGTGAAAAATATTTCGACGCGGGAATAGAACTTCCACTG
AGAGTGAGACCCATTAGTTCCACTTTTTTATTCTTGGTGCCTACTCGCAACTTCAACTCGCTATCATATGGAACAATCTCGACTTTTCGAGCATGTGCTTTGT
CAGTTTTCTTAAGCCTTCAAGCCTTAGAGGAAAATAGCCAATATAACATTTTAAATAATCTTCTTAAAATTACTTTATTTAAAATTTAAAAGAACTATATCATTTT
AATTTTTTAAAATAATTTTTTTAAATAAAAGTTAAATACTTTAATTTCAAAAATAAAATTAATAAAAAGTAAATAATGAAACCAGCCCGGGCCGTCG

CLON SSRY-9 (de 350 à 1641 pb)

AGAGATAGTGCATGCACAGGTTGAGCCTTGGTTTCAGAAAAGAGTTTTATTTTCACAGAAAATGTATGATCATGTATGAGGTTTACAGGTATACAGAGAGTAT
AGCAGGCTTGCTACGGTTCTGACGGCCTTAAGCCGACCTGAATCCTAGCGCCGGTGACGGTCCATTTTTGGGTCGTTACATATATTTACTCCTGTTACAGCC
TCCAGTATTGTTGTAGATTATTAGGATGATCCACATGAGACCAAGACATAGTGAATGGAAAGGACCACCCGCTAGTGTGCAACATAAGAATCAAACCTTTGGA
ATACACATGATGCTATTGTCTTTACTACAATTCATCATGAGTCACTCAACT*AATAAGGATTAACAAAAATCCCATTATTCTACATGTGCATGCATGCTGAGAG
AAGAGGGACATAAGTTGGCATAACAAAGAGGAGCATGAAACAGAATCAGAAAACAAAAATGAGAGTGAGTGATAAGGTTACGGCAACTGCCAAAACAGAC
ACACACACACACACACACACACACACACACTTAGAGGAGCAATCCTCATTTCACTTCAAAACTACTAGCACTACTGCAGTAGGACCAGGAACAATAACGGTA
ACACTGGATTATGATGAGGCATCATTGCTAGACGCAGACATAAGATAGTAATGGCTCCAAATACTATTACCTTGGCGATTGGCAACATCCCAGTACTTTCAG
CAGACACAGGTAGAATATTTAGAAGACCCTCATCAACAAAATTTCTGCAACTGCACCCTTTGCAATGAGCCAAGCTTTGGGACAGAGGATTCAGCCAATCTCG
CAAATGCAGTTGCTTGGCTGTGCACAATACATGAATCTGAGGCTTTGAACCAGAACCACAAAAAGTAAAAACAAAATCAGAGGAAGGAAAGAGAGAAACAAAAC
ATTAATAACTTAAATACCACTACTTCCATAGAGAACTTTGAAGAGATATAAACTTTCACAATAAAAACCTTGGAAATTGTAGCTGATGATATACTTCTCCATTT
GTTTGTTCAGTATAAGGCTTGAAGCAACAGCTTTCGGTGTATCCTGTGCATCTTGGTTTTCAGGAACACTGCTGGCTAAATCTTCAACAAGAAAAGCTAGAAAC
AGTTAGAAACCAAACTGCAGACTTGGAAAATATACAAGAAGATTGAAAATACCATTAGAACCAGATGAAGGATCAACGTTATCTACAAGCCTCCTAGTTAATG
GACCAGGATTCAAGATGGATATGCTTCAATTTCATGACGTATGGCTTCAAGTTGAGCCATTGTATCTTGAAGTTCTTGTGGACCTAAAACAAAAACTGGTATA
CATCTCAAGAAACATAGTCTTTGCAATAATATCATAAATCCCACATTCAAACGCAACAACATAACCATATCCGACACCATTACGCACTCAATATTTATAATGCA
CAAAGCCACCATTCAATTTGACTCCGTAATTTGCACTTAAAGAGAATTACTTTAAAAAAAAAAAAAAAAAGATTGAAAAAGGTAACCTGGCGAGCTTGAGATTGCT
GCATAACATTCTCAAATTTGGCCACGAGCCAACTGGACATAGCCAATGGCTCGACCAGCTAATCTCCAGACCAGCCCGGGCCGTCG

CLON SSRY-40 (de 600 à 1836 pb)

AACGAAACGAACAAAGGGATTCATTTATTTTTACCCTTTAACTTTTCATATGGGCAAACCTTCTTATTCAATTTGCTTCTTGATCCATTTTCAAATGAGCAACAC
TGCGTGCATCATGGTCCACTCACTC*ATTGGCGGTGACGGAACCCCTCACAAGCTCAAACCGGTGATCTGCCAGTGCACCCGCTCACAAAATCTCAGACAGG
AGAGAGAGAGAGAGAGAGAGAGAGACAGAGAAGTGTGTTGCATGTATGTGACGAAGTGACAGTAAGGGGAAACGAAAAAGAAGAAAAAATATGTTTTATAAT
GCATATCTCTTATTTGATGGAATGCCGAAAAGAATGTTATGCAATACATGCTTATTATGTGTAGCTTAAGTTTTGGAAAGTATTTTATCAAGAATCAAGATG
CATAAACTTGACTTCTCTTGCATTGCTAGCTGAGAAATCAGTTCACAGCTTTATGATAAAGAATCTAGTTGCGACCATTGTAGATTTAGTGGGTTTAGAGAAA
GAGAGAAAAAGAATGTATGGAGAAGGGGGAGGAGAAGAATAAGGAGATTGTTCTAGTTATTTTCCTTTCACTAGCGGTATTTTTATGGAGAGACCTCCAAT
TTGAATCATTATAAAACTTGGCTAATCTTGAGTTTTTAGAAGTAGAGGGACATACCTATTAATCATGAGATAAGCGAGTGAATAAAAAATTTTTCCCTAATA
TTTTTATGCATTTTGCTTACTGCTTAAATCTGCCTGCTCTGATATAATTGGCCTTTATCACTCAAATGAATGTACAACCTGCACTTTTGCTGCTGTAGACTGTGAAG
TTCAAGTCATATACTTTTTGGTGCTTTGTGGTATTCATTCTTGGAGTTAAAAATAGATAAAAATATTTATTAAGTGATTAATTTGCCCAATGAGGTTCTATATAT
TTGTTAATCTTGTGAATCTTGTGCTTTCTATGKCTTGKAGTTTCTTACTATGCCTTATTTTCGTGATTGTAATTAATACTGTGTTGCACACTTGAATTTGTGTT
GGTTAATCTTTTTGAACCTCTAATTATGATTGGTTGCATAATCATGCAGGGCATTTTCTGCCGGAGGAGGATGACAAGTTTCTTGAGAGAGTTCGGGCTGCAG
TTGAGGAAGAGGAGCGCAAGCAATGGTTGCTGCTGACACTGATGCTGCCAAGGATGCCATTGCAACTGCTGAGGGATCGCGGAAAAACAACCCAGACTCACGG
GCCTAAGGATTCAAGTGCTGATAATGTTGGAACCTAAGGAAAGTACAGGCCAAGTAACTGATCGCAAAGACAGTGCAGGCTCTGGTGCGGTCACTGATCCATCTG
AAGAAAGAGGAACTGAAGGTCAAAGTTATAGCGGAGCATATGATTCTGTGGCAAACCTACCATTGGAATCTACCACTATTATTATGGCAGCAATAATGATATGG
GCACACTTATTGAGGTGATTTTAAACTGCTTATTCCACTGATTCTCTGTATGGTGCAGAATACATCATCTCACAAAGCACAAAGTATTCTTAATATTAGTGGC
AGTCTTTCTTACTATTAACACAAAACCTGCAGGTTGACTCCATAGTGAGTTGATAATTATAGCGAGAGCTAATCATTCTTACAAGTTCTATCATCTTGCCTGCA
GGTTAGAAGAACCTGGGATGCATATATAAGACCGGGGGGAAGGTAAGCCTGTAAATCCTTTTAGCCGGAATTATGCAAACCGTTATGTCAATCCCCACACTAA
CTTTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTGACCATATGGGAGAGCTCCCAACGCGTG

Annexe 4. Polymorphismes de séquences observés chez l'arracacha au niveau des régions flanquantes du locus SSR AxC87 allongé par marche chromosomique. Les séquences suivent l'ordre de l'arbre phylogénétique (Figure 25). La taille a été obtenue à partir de la séquence de l'allèle avec les amorces utilisées en génotypage du locus en question.

	Accession	Ref. seq	Taille	1*	2	3	4	5	6	7	8	9	10	11	12	13	14
				28**	39	46	56	85	95	118	141	166	197	264	265	IN-1	IN-2
1	ECU-1193	1193c5	115	T	T	T	C	C	C	G	T	G	C	A	A	0	0
2	ECU-1226	1226c1	117	T	T	T	C	C	C	G	T	G	C	A	A	0	0
3	ECU-1226	1226c5	119	T	T	T	C	C	C	G	T	G	C	A	A	0	0
4	ECU-1157	1157c12	113	C	T	C	C	C	C	G	T	G	C	A	A	0	0
5	UNC43	UNC43C7	115	T	T	T	C	C	C	G	T	G	C	A	A	0	0
6	A	AC8	117	T	T	T	C	C	C	G	T	G	C	A	A	0	0
7	ECU-1229	1229C1	101	T	C	T	T	C	A	G	T	A	T	A	A	0	0
8	ECU-1193	1193c4	101	T	C	T	T	C	A	G	T	A	T	C	A	1	0
9	ECU-1226	1226c9	101	T	C	T	T	C	A	G	T	A	T	C	A	1	0
10	UNC171	UNC171C3	99	T	C	T	T	C	A	G	T	A	T	C	A	1	0
11	S133	S133c5	101	T	C	T	T	C	A	G	T	A	T	C	A	1	0
12	MST-081-7	081-7-C4	101	T	C	T	T	C	A	G	T	A	T	C	A	1	0
13	A	AC15	101	T	C	T	T	C	A	G	T	A	T	C	A	1	0
14	S-500	S500c6	101	T	C	T	-	-	A	G	T	A	T	C	C	1	1
15	UNC171	UNC171C13	101	T	C	T	C	C	A	G	T	A	T	C	A	1	0
16	S133	S133c6	91	C	C	T	C	A	-	G	T	A	C	A	A	0	0
17	UNC43	UNC43C4	109	T	T	T	C	C	C	G	T	G	C	A	A	0	0
18	S-9	S9C5	107	T	C	T	C	C	-	A	T	G	C	A	A	0	0
19	S-9	S9C20	105	T	C	T	T	C	-	A	T	G	C	A	A	0	0
20	S-9	S9C20	99	T	C	T	T	C	-	G	T	G	C	A	A	0	0
21	S-105	S105c10	96	T	C	T	T	C	C	G	T	G	C	A	A	0	0
22	S-500	S500c2	91	T	C	T	T	A	C	G	A	G	T	A	A	0	0
23	31-2	31-2c5	89	T	C	T	T	C	-	G	A	G	C	A	A	0	0
24	RBS-001	RBSC7	85	T	C	A	T	C	-	G	T	G	C	C	C	1	1

** signale le positionnement du polymorphisme dans la séquence alignée

* signale le numéro en série du polymorphisme

Annexe 7. Polymorphismes de séquences observés chez le manioc au niveau des régions flanquantes du locus SSRY-9 allongé par marche chromosomique. Les séquences suivent l'ordre de l'arbre phylogénétique (Figure 28). La taille a été obtenue à partir de la séquence de l'allèle avec les amorces utilisées en génotypage du locus en question.

	Accession	Ref. seq	taille	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	
				33**	47	111	134	142	161	173	180	253	312	470	529	559	571	587	611	634	645	691	696	720	733	739	748	751	887	903	941	978	1011	1030	1034	1046	1094	1103	1117	IN1	
1	5.49	5.49c3	250	A	G	A	G	G	T	G	-	G	C	G	T	T	T	G	A	C	C	A	C	G	A	G	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
2	5.59	5.59C2b	266	A	G	A	G	G	T	G	A	A	C	G	T	T	T	G	A	C	C	A	C	G	A	G	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
3	5.62	5.62C10	250	A	G	A	G	G	T	G	-	G	C	G	T	T	T	G	A	C	C	A	C	G	A	G	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
4	5.31	5.31c3	250	A	G	A	G	G	T	G	-	G	C	G	T	T	T	G	A	C	C	A	C	G	A	G	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
5	5.72	5.72c16	250	A	G	A	G	G	T	G	-	G	C	G	T	T	T	G	A	C	C	A	C	G	A	G	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
6	5.65	5.65c9a	250	A	G	A	G	G	T	G	-	G	C	G	T	T	T	G	A	C	C	A	C	G	A	G	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
7	5.47	5.47C18	264	A	G	A	G	G	T	G	A	A	C	G	T	T	C	A	A	C	T	A	C	G	A	C	C	T	C	C	G	A	T	A	C	A	C	G	A	0	
8	5.6	5.6c1	260	A	A	A	G	G	T	C	A	G	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
9	5.28	5.28C12	260	A	A	A	G	G	T	C	A	G	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
10	5.3	5.30c3	260	A	A	A	G	G	T	C	A	G	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
11	5.32	5.32c1	256	A	A	G	G	G	C	C	A	G	C	G	T	T	T	G	G	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
12	5.32	5.32C5	257	A	A	G	G	G	C	C	-	G	C	G	T	T	T	G	G	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
13	5.89	5.89C10	274	A	A	A	G	G	T	C	A	G	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
14	5.53	5.53C2	278	A	A	A	G	G	T	C	A	G	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
15	5.9	5.90C9	268	A	A	A	G	G	T	C	A	G	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
16	5.6	5.6C11	263	C	G	A	C	G	T	G	T	A	C	G	T	T	T	G	A	C	C	A	C	G	G	C	C	T	C	C	G	A	C	A	C	A	C	G	A	0	
17	5.28	5.28c2	264	C	G	A	C	G	T	G	T	A	T	T	T	T	T	G	A	A	C	A	T	G	A	A	C	G	T	C	G	G	C	A	C	G	G	G	C	1	
18	5.85	5.85c10	264	C	G	A	C	G	T	G	T	A	T	T	T	T	T	G	A	A	C	A	T	G	A	C	C	G	T	C	G	G	C	A	C	G	G	G	C	1	
19	5.78	5.78c3	264	G	G	A	C	G	T	G	T	A	T	T	T	T	T	G	A	A	C	A	T	G	A	C	A	G	T	C	G	G	C	A	C	G	G	G	C	1	
20	5.24	5.24c13	260	A	G	A	G	G	T	G	A	G	T	G	T	T	T	G	A	C	C	A	C	G	A	C	A	T	C	C	G	A	C	A	C	A	C	A	C	G	A
21	5.2	5.2c6	262	A	G	A	G	G	T	G	A	G	T	G	C	C	T	G	A	C	C	A	G	C	T	A	C	C	T	C	T	G	G	C	G	C	G	C	A	0	
22	5.4	5.4c1	262	A	G	A	G	G	T	G	A	G	T	G	C	C	T	G	A	C	C	A	C	G	C	T	A	C	C	T	C	T	G	G	C	G	C	G	C	A	0
23	5.49	5.49c3	262	A	G	A	G	C	T	G	A	G	C	G	T	T	C	G	A	C	T	A	C	G	A	C	C	T	C	C	T	A	C	A	A	A	A	C	G	A	0
24	5.7	5.70c5	262	A	G	G	G	C	T	G	A	G	C	G	T	T	C	G	A	C	T	A	C	G	A	C	C	T	C	C	T	A	C	A	A	A	A	C	G	A	0
25	5.68	5.68c30	262	A	G	A	G	C	T	G	A	G	C	G	T	T	C	G	A	C	T	A	C	G	A	C	C	T	C	C	T	A	C	A	A	A	A	C	G	A	0
26	5.47	5.47C19	262	A	G	A	G	C	T	G	A	G	C	G	T	T	C	A	A	C	T	A	C	G	A	C	C	T	C	C	T	A	C	A	A	A	A	C	G	A	0
27	5.94	5.94c3b	262	A	G	A	G	C	T	G	A	G	C	G	T	T	T	G	A	C	T	A	C	G	A	C	C	T	C	C	T	A	C	A	A	A	A	C	G	A	0

** signale le positionnement du polymorphisme dans la séquence alignée

* signale le numéro en série du polymorphisme

Annexe 8. Polymorphismes de séquences observés chez le manioc au niveau des régions flanquantes du locus SSRY-40 allongé par marche chromosomique. Les séquences suivent l'ordre de l'arbre phylogénétique (Figure 29). La taille a été obtenue à partir de la séquence de l'allèle avec les amorces utilisées en génotypage du locus en question.

Accession	Ref. seq.	Taille	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	
1	5.9	5.9	217	N	N	N	N	N	N	N	N	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	A	T	A	A	T	T	T	T	C	A	1
2	5.2	5.2F	223	N	N	N	N	N	N	N	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	A	T	A	A	T	T	T	T	C	A	1	
3	5.85	5.85c1	217	G	A	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	A	T	A	A	T	T	T	T	C	A	1		
4	5.59	5.59c5	231	G	G	G	C	C	A	A	A	T	A	G	C	T	T	A	C	C	T	T	G	G	T	A	G	G	T	T	C	C	G	A	G	T	A	A	C	C	G	A	A	T	A	C	G	A	A	T	A	A	T	T	T	C	A	1		
5	5.32	5.32c2	231	G	G	G	C	C	A	A	A	T	A	G	C	T	T	A	C	C	T	T	G	G	T	A	G	G	T	T	C	C	G	A	G	T	A	A	C	C	G	A	A	T	A	C	G	A	A	T	A	A	T	T	T	C	A	1		
6	5.33	5.33c3	231	G	G	G	C	C	A	A	A	T	A	G	C	T	T	A	C	C	T	T	G	G	T	A	G	G	T	T	C	C	G	A	G	T	A	A	C	C	G	A	A	T	A	C	G	A	A	T	A	A	T	T	T	C	A	1		
7	5.24	5.24c14	223	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	C	T	A	G	T	A	A	T	A	C	G	T	A	A	T	T	T	T	C	A	0		
8	5.52	5.52c3	223	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	C	T	A	G	T	A	A	T	A	C	G	T	A	A	T	T	T	C	A	0			
9	5.54	5.54c16	223	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	C	T	A	G	T	A	A	T	A	C	G	T	A	A	T	T	T	C	A	1			
10	5.24	5.24c15	227	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	C	T	A	G	T	A	A	T	A	C	G	T	A	A	T	T	T	C	A	1			
11	5.25	5.25c1	237	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	C	T	A	G	T	A	A	T	A	C	G	T	A	A	T	T	T	C	A	1			
12	5.28	5.28c11	231	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	C	T	A	G	T	A	A	T	A	C	G	T	A	A	T	T	T	C	A	1			
13	5.62	5.62c11	229	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
14	5.28	5.28c9	237	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
15	5.91	5.91c16	229	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
16	5.7	5.7c16	229	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
17	5.42	5.42c14	227	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
18	5.58	5.58c10	229	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
19	5.49	5.49c8b	229	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
20	5.91	5.91c3	233	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
21	5.28	5.28c12	227	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	T	G	G	G	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
22	5.78	5.78c6	225	G	T	G	T	G	A	A	A	G	G	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	C	G	A	A	T	A	C	A	G	G	A	T	A	T	T	T	C	A	1		
23	5.94	5.94c9b	219	G	T	G	T	G	A	A	A	G	A	T	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	G	T	A	A	A	C	G	A	A	T	A	C	A	G	A	G	T	A	T	T	T	C	A	1		
24	5.3	5.30c11	225	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	G	A	G	T	A	C	A	G	G	A	T	A	A	T	T	T	T	C	A	1			
25	5.32	5.32c13	225	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	G	A	G	T	A	C	A	G	G	G	A	A	A	T	T	T	T	T	C	A	1		
26	5.47	5.47c1	225	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	G	A	G	T	A	C	A	G	G	G	A	A	A	T	T	T	T	T	C	A	1		
27	5.58	5.58c7	223	A	A	A	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	T	T	G	G	C	T	A	A	A	A	C	G	A	A	T	A	C	A	G	G	A	T	A	A	T	T	T	T	C	A	1		
28	5.6	5.6c2	225	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	G	A	A	T	A	C	A	G	G	A	T	A	A	T	T	T	T	C	A	1			
29	5.51	5.51c6	227	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	D	G	A	G	T	A	C	A	G	G	A	T	A	A	T	T	T	T	C	A	1		
30	5.3	5.30c12	239	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	G	A	A	T	A	C	A	G	G	A	T	A	A	T	T	T	T	C	A	1			
31	5.52	5.52c6	219	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	D	G	A	A	T	A	C	A	G	G	A	T	A	A	T	T	T	C	A	1			
32	5.33	5.33c10	216	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	T	T	A	G	T	A	A	A	C	G	A	A	T	A	C	A	G	G	A	T	A	A	T	A	T	T	T	C	A	1		
33	5.59	5.59c6	237	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	G	T	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	G	A	T	A	A	T	T	T	T	C	A	1	
34	5.3	5.30c13	229	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	G	A	T	A	A	T	T	T	C	A	1		
35	5.46	5.46c1	237	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	G	A	T	A	A	T	T	T	C	A	1		
36	5.46	5.46c12	231	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	G	A	T	A	A	T	T	T	C	A	1		
37	5.63	5.63c6	235	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C	T	C	A	A	C	A	G	G	G	G	C	C	A	G	T	A	A	A	C	G	A	A	T	A	C	G	G	G	A	T	A	A	T	T	T	C	A	1		
38	5.68	5.68c27	237	G	T	G	T	G	A	A	A	T	A	G	T	G	C	A	C	C																																								

Diplôme d'Etudes Approfondies
de Ressources Phytogénétiques et Interactions
biologiques

Ecole Doctorale Biologie Intégrative

***Caractérisation des introgressions
du complexe d'espèces *Manihot carthaginensis* s.l,
incluant divers «arbres manioc», dans certaines
variétés traditionnelles américaines et africaines
de manioc (*M. esculenta* Crantz)***

Eduardo MORILLO

Laboratoire Gene Trop
IRD
Montpellier

Responsable de stage : Gérard SECOND

Soutenu à Montpellier
Le 31 Août 2001

Caractérisation des introgressions du complexe d'espèces *Manihot carthaginensis* s.l., incluant divers «arbres manioc», dans certaines variétés traditionnelles américaines et africaines de manioc (*M. esculenta* Crantz)

Résumé

Plusieurs types d'observations suggèrent que certaines variétés de *manioc* seraient introgressées du complexe d'espèces *M.carthaginensis* s.l.. Cette étude s'est basée sur l'utilisation des marqueurs moléculaires de type SSR et AFLP pour préciser la structure génétique de ce complexe d'espèces et ses introgressions dans le manioc cultivé. De nouvelles évidences d'introgression basées sur l'absence de quelques bandes chez l'ancêtre du manioc postulé *M.esculenta* ssp. *flabellifolia* sont reportées. En particulier, l'utilisation des SSR cartographiés montrent qu'un fragment du chromosome D, introgressé artificiellement de *M.glaziovii* dans le parent femelle de la carte génétique, serait aussi présent dans certaines variétés traditionnelles originaires d'Amérique du Sud. Le fait que ce segment chromosomique soit considéré comme portant des facteurs de résistance non seulement à la bactériose vasculaire du manioc (CBB) mais aussi à la virose africaine (CMD) permet d'interpréter cette observation. L'analyse d'autres régions chromosomiques suggère que ces introgressions correspondent à plusieurs fragments dans le génome du manioc cultivé et que d'autres espèces ont pu aussi participer à la diversification du manioc dans son aire d'origine. La capacité de chaque technique pour repérer ces introgressions est discutée et les perspectives présentées. Par ailleurs, le séquençage d'un AFLP cloné a fourni la preuve d'introgression de *M.glaziovii* chez quelques cultivars de *M.esculenta*.

Mots clés: Manioc, *M. esculenta*, introgression, diversification, domestication, AFLP, SSR

Abstract

Many observations suggest that some varieties of cassava are introgressed from *M. carthaginensis* s.l complex species. This study is based on the utilization of SSR and AFLP markers to study the genetic structure of these species and their introgressions in some cultivars of *M.esculenta*. New evidences based on the absence of some molecular markers in the presumed ancestor of cassava, *M.esculenta* ssp. *flabellifolia*, are shown. Interestingly, the mapped SSRs show that a chromosome segment located on the D linkage group was introgressed from *M.glaziovii*, not only in one of the parent of the cassava genetic map in which it resulted from artificial introgression but also in various traditional cultivars originated in South America. The fact that this chromosome segment is believed to bear resistance genes to both cassava bacterial blight (CBB) and African cassava mosaic disease (ACMD) allow interpreting this observation. The analysis of mapped SSRs suggests also introgressions in some other chromosomic regions and show that other species could also have participated to cassava diversification. The capacity of each molecular marker to evidence introgressions is also discussed and the perspectives presented. In addition, an AFLP clone sequence confirms the introgression from *M.glaziovii* in some cassava cultivars.

Key words: Cassava, *M. esculenta*, introgression, diversity, domestication, AFLP, SSR

INTRODUCTION

Le manioc (*Manihot esculenta* Crantz, Euphorbiaceae) est un arbuste pérenne originaire de l'Amérique et cultivé pour ses racines tubérisées. Il constitue la base de l'alimentation pour plus de 600 millions d'habitants des régions tropicales et se classe au cinquième rang mondial des productions végétales alimentaires derrière le maïs, le riz, le blé et la pomme de terre (FAO, 2000).

Le genre *Manihot* comprend 98 espèces toutes monoïques et allogames, dont la délimitation est sujette à controverses. Sa distribution est limitée au continent américain, du sud des États-Unis au nord de l'Argentine. L'existence de formes hybrides interspécifiques spontanées est avérée et les barrières reproductives interspécifiques semblent peu prononcées bien que mal connues. L'espèce cultivée est fréquemment rencontrée en sympatrie avec des espèces sauvages (Rogers et Appan, 1973) mais, encore récemment, l'ancêtre sauvage et le rôle possible des hybridations interspécifiques dans le processus de la domestication du manioc n'avaient pas été caractérisés. Si aucune barrière génétique forte à l'hybridation n'a été mentionnée, il existe cependant des barrières géographiques en particulier entre les centres de diversité des espèces sauvages centrés sur le Brésil et le Mexique, respectivement, qui limitent les hybridations.

Jusqu'à ces dernières années, l'origine du manioc était inconnue. Plusieurs espèces sauvages avaient été proposées comme ancêtre du manioc: *M. aesculifolia* distribué en Amérique centrale (Rogers et Appan, 1973), *M. carthaginensis*, originaire de la Colombie et du Venezuela (Bertram, 1993) enfin *M. esculenta ssp. flabellifolia* et *M. esculenta ssp. peruviana* de la région amazonienne (Allem, 1994). Le développement de l'utilisation des marqueurs moléculaires (Colombo *et al.*, 1998, Roa *et al.*, 1997, Roa *et al.*, 2000 Second *et al.*, 1997), et le séquençage d'un gène nucléaire (Olsen et Schaal, 1999) ont permis de préciser la structure et l'origine génétiques du manioc. Ces études montrent que la diversité moléculaire du manioc est peu structurée et confirment la forme sauvage du manioc (*M. esculenta ssp. flabellifolia* syn. *peruviana*) comme l'ancêtre le plus direct du manioc. Cette hypothèse est en désaccord avec celle de Rogers et Appan (1973) selon laquelle *M. esculenta* serait issu d'hybridations introgressives entre de nombreuses espèces sauvages sans un correspondant sauvage direct. D'autres données moléculaires au niveau de l'ADN chloroplastique et des gènes ribosomiques nucléaires (Fregene *et al.*, 1994) confirment aussi l'hypothèse d'Allem (1994) sans toutefois écarter la possibilité d'hybridations avec d'autres espèces sauvages. Ainsi, Colombo (1997), Second *et al.*, (1997) et Second et Iglesias (2000) se basant sur de nombreux marqueurs arbitraires et ayant inclus une large représentation tant de l'espèce cultivée que du genre *Manihot* (bien que principalement dans ses espèces Sud-américaines) obtiennent des indices de contributions à la diversité génétique du manioc de certaines espèces sauvages, en particulier *M. glaziovii*. Cette espèce, appelée « maniçoba », originaire de l'Est du Brésil est une des espèces du genre *Manihot* la plus éloignée du manioc. Cependant leurs hybrides se rencontrent fréquemment dans le nord-est brésilien où ils sont connus sous le nom de « manioc de sept ans » (*mandioca de sete anos*) ou « arbres manioc » et certains sont récoltés pour leurs racines (Second *et al.*, 1997). *M. glaziovii* apparaît génétiquement très proche de *M. carthaginensis* (Second *et al.*, 1997) qui est aussi connu pour former des hybrides avec le manioc (Rogers et Appan 1973). En accord avec l'analyse morphologique (A.C. Allem pers. com.), l'ensemble de ces formes (*M. carthaginensis*, *M. glaziovii* et *M. epruinosa* –syn. *glaziovii* et les arbres manioc) représentent le complexe

d'espèces *M. carthagenensis s. l.* (Second, 2000). *M. glaziovii* a été aussi introduit en Afrique il y a une centaine d'années où elle a fait l'objet d'une tentative de production de latex à la fin de XIX siècle (Serier, 1989). Cette espèce est aussi la source de résistance la plus utilisée pour l'amélioration de la résistance tant à la mosaïque virale africaine (CMV) qu'à la bactériose vasculaire (CBB) qui sont les deux contraintes phytosanitaires majeures de cette culture dans le monde (Hahn, 1984). L'existence d'hybrides spontanés fertiles en Afrique a été confirmée par un travail pionnier à l'aide de marqueurs isoenzymatiques sur l'origine d'introgessions spontanées du manioc sur le continent africain (Lefèvre, 1989). Une situation similaire à celle qui a été décrite en Afrique de l'Ouest par Lefèvre (1989) se retrouve donc dans le Nord-Est brésilien. *M. glaziovii* et probablement d'autres espèces ont pu contribuer à la diversité génétique du manioc par introgessions résultants de rétrocroisements successifs de l'hybride inter-spécifique sur l'espèce cultivée (Second, 1999). L'incidence de telles introgessions sur la diversité du manioc n'a pas encore été éclaircie. Les travaux de Roa et al., (1997, 2000) et Olsen et Schaal (1999, 2000) conduisent ces auteurs à considérer que son effet est plutôt mineur tandis que Second et al. (1997, 1999) penchent pour l'hypothèse d'une introgession non négligeable d'espèces sauvages du genre *Manihot* parmi les formes cultivées.

L'objectif de cette étude est donc de préciser la structure du complexe d'espèces *Manihot carthagenensis s.l.* et de démontrer ses introgessions dans le manioc en utilisant comme référence une variété artificiellement introgessée. Deux approches ont été essayées; la première est basée sur l'utilisation des marqueurs SSR cartographiés (Mba et al., 2001) localisés dans des régions chromosomiques qui portent des QTLs de résistance à la CBB (Jorge et al., 2000). Malgré l'homoplasie fréquente des marqueurs SSR, la taille ainsi que les fréquences observées sont souvent très différentes entre les espèces *M. carthagenensis s.l.* et *M. esculenta* (Roa et al., 2000) et permettent donc de présumer des introgessions. Une seconde approche a consisté à démontrer l'introgession postulée en séquençant certaines bandes AFLPs et SSR observées parmi nos échantillons chez le manioc cultivé et dans le complexe *M. carthagenensis s.l.* mais pas chez l'ancêtre direct postulé (*M. esculenta ssp.flabellifolia*).

MATERIELS ET METHODES

Matériel végétal

Un ensemble de 76 échantillons détaillés dans le tableau 1 a été utilisé. Soixante de ceux là ont été retenus pour l'analyse par SSRs. Ils correspondent au manioc (35) y compris une variété triploïde traditionnelle (Carvalho et al., 1999), *M. e. ssp. flabellifolia* (6), des hybrides interspécifiques spontanés « arbres manioc » (12), *M. glaziovii* du nord-est du Brésil et Afrique (10), *M. carthagenensis* (6), *M. brachyloba* (3) et *M. quinquepartita* (2).

Extraction d'ADN

Une extraction d'ADN total a été faite à partir de feuilles séchées conservées au sec en gel de silice (Tableau 1). Le protocole détaillé par Colombo (1998) a été utilisé en remplaçant le CTAB (Cétyl-triméthyl ammonium bromide) par du MTAB (Mixed Alkyltrimethylammonium bromide) à la même concentration. La concentration et la qualité de l'ADN ont été déterminés par fluorescence.

Analyse par AFLPs (Amplified Fragment Length Polymorphism)

Les analyses ont été contractées auprès du laboratoire AGROGENE (<http://www.agrogene.com>) y compris la lecture des bandes observées. Les protocoles de

digestion, ligation et pre-amplification employés ont été réalisés selon Vos *et al.* (1995). Pour la séquence arbitraire des amorces, les combinaisons E-AGA/M-CAG, E-AGT/M-CTC et E-AAT/M-CCC ont été utilisées. Les deux premières ont montré un important polymorphisme chez le genre *Manihot* (Second, 1997). La troisième combinaison a été choisie d'un *criblage* réalisé antérieurement. Les produits AFLPs ont été révélés par marquage radioactif et les polymorphismes codés à l'aide du logiciel *Keygene*. Les mêmes préamplifications ont été utilisées pour utiliser, après complétion du même protocole dans nos laboratoires, la révélation du gel au nitrate d'argent et le repérage des bandes présumés introgressés afin de les réamplifier. La coloration a été effectuée selon les recommandations du fournisseur (Kit Promega). Les bandes d'ADN ont été découpées du gel et transférées dans 50 μ l.

Clonage d'un AFLP

Le clonage d'un fragment de 230 pb révélé avec le couple E-AGA/M-CAG a été appliqué. Il est présent chez *M. glaziovii*, les arbres manioc et 13 variétés de *M. esculenta* mais absent de notre échantillonnage chez *M. e. flabellifolia*. Cette bande a été découpée dans le gel et amplifiée par PCR à partir du couple d'amorces initialement utilisé lors de l'amplification sélective en AFLP. Les conditions d'amplification sont les suivantes : après dénaturation de l'ADN à 94°C pendant 5 min, celui-ci subit 12 cycles de 30s à 94°C, 30s à 64°C (-0.7°C par cycle), 1min à 72°C suivi de 32 cycles de 30s à 94°C, 30s à 56°C et 1min à 72°C. Les ADN ainsi amplifiés ont été purifiés avec le kit Amersham (Ref.No. 27-9602-01) et clonés avec le kit Promega (Ref.A1380).

Analyse par SSR (Single Sequence Repeats)

L'ADN de chaque accession a été amplifié pour chacun des 27 locus microsatellites détaillés dans le tableau 2. Dix correspondent aux paires d'amorces de Chavarriaga *et al.* (1998) et 17 ont été choisies parmi les 36 nouveaux marqueurs cartographiés par Mba *et al.* (2001) en fonction de leur localisation dans les régions des QTLs de résistance à la CBB (Chromosomes D, F, G et L, Jorge *et al.*, 2000).

Pour la révélation de microsatellites, les amorces et un marqueur de taille (30-330 pb Promega) ont été marqués au [γ^{33} P]-ATP (AH9968 Amersham). La concentration finale de la réaction de PCR en volumes de 20 μ l a été de 25 ng d'ADN total, 0.4 μ M de chaque amorce marqué, 200 μ M dNTPs, 2 mM MgCl₂ et 0.04 U/ μ l de Taq DNA polymérase (Promega). L'amplification correspond aux cycles suivants: 2 min à 94°C, 1 min à 94°C, 1 min à 55°C ou 45°C en fonction des amorces (tableau 2) et 2 min à 72 °C pendant 30 cycles et 5 min à 72°C. Après dénaturation des produits d'amplification, ceux-ci sont séparés sur gel de polyacrylamide 6% dénaturant. Les gels sont séchés sur un papier Wattman de dimension 33x40 cm et exposés pendant 24H à un film Kodak Bio Max X-Ray. Pour la lecture des bandes, seulement les bandes les plus intenses ont été considérées. La taille des allèles a été estimée par rapport à la distance de migration du marqueur de référence et calculée en utilisant le logiciel Lenght III (Templeton, 1988).

Analyses statistiques

La présence ou l'absence des bandes AFLP et SSR a été codée par 1 et 0 respectivement. Les matrices ainsi élaborées ont été utilisées pour calculer une matrice de similarité en utilisant les coefficients de Jaccard ou *Simple Matching* (SM) dans les comparaisons de géotypes deux à deux. Les matrices de similarités sont ensuite utilisées pour construire des dendrogrammes avec la méthode UPGMA ou pour effectuer des analyses en coordonnées principales (PCO). Le calcul d'une corrélation (r) par le test de Mantel (Sokal et Rolph, 1995) permet d'estimer la validité de l'analyse Cluster (« indice cophénétique ») et la corrélation entre l'information

donnée par chaque type de marqueur. Les valeurs au dessus de 0.8 signifient une bonne concordance de la corrélation. Le programme NTSYS (Exeter ver. 2.0, Rolph, 1998) a été utilisé.

Le programme Genetix ver. 4.02 (Belkhir *et al.*, 2001) a été utilisé avec les SSR pour effectuer une Analyse Factorielle des Correspondances (AFC) afin de mettre en correspondance les allèles et les individus sur les mêmes coordonnées et identifier les allèles responsables de la structure observée. Tous les individus ont été considérés diploïdes. Les fréquences alléliques par locus pour chaque espèce et les valeurs de *F_{is}* pour chaque locus et chaque espèce ont été calculées à l'aide des estimateurs de Robertson&Hill et des valeurs des paramètres de polymorphisme comme l'hétérozygotie estimée sous l'hypothèse d'équilibre de Hardy-Weinberg (*H_{exp}*) et l'hétérozygotie observée (*H_{obs}*), la proportion de locus polymorphes (*P*) (calculés à 0.95 et 0.99) et le nombre moyen d'allèles (tous les allèles détectés dans l'échantillon) par locus.

Séquençage des bandes présumés introgressées

Après clonage des bandes AFLP isolées au sein du vecteur plasmidique, une mini-préparation d'ADN plasmidique a été effectuée à partir du kit QIAprep (QIAGEN). Le séquençage du clone AFLP a été effectué par le laboratoire GENOME EXPRESS (CEA Grenoble). Les séquences de *M. esculenta* et *M. glaziovii* ont été comparées et analysées avec le logiciel MAGLIGN (*DNASTAR package ver. 1.2*, <http://www.dnastar.com>). Afin de vérifier la séquence de ce fragment chez *M. e. flabellifolia* un nouveau couple d'amorces a été dessiné et synthétisé à partir de la séquence de l'espèce cultivée. À partir de la séquence obtenue avec la variété de manioc Col22 (29), les séquences 5'CACAAATATCGAATATTCATAG3' et 5'CACTATTTCTTCACATTTCCC3' ont permis l'amplification de ce fragment par PCR. L'amplification correspond aux cycles suivants: 5 min à 94°C, 34 cycles de 1 min à °C, 1 min 30s à 56°C et 5 min à 72°C. Après vérification des produits d'amplification par fluorescence sous UV sur gel d'agarose, les produits ont été purifiés et séquençés. Les séquences ont été alignées et analysées avec la méthode du logiciel PAUP ver. 4.0. Un test de *bootstrap* a été aussi appliqué.

Pour le séquençage des allèles SSR, des amplifications non radioactives ont été effectuées. On a isolé 23 allèles correspondant à 5 locus microsatellites. Les bandes ont été découpées d'un gel d'agarose, purifiées et a séquençées. Les séquences ont été alignées et le pourcentage de similarité et de divergence calculés avec MAGLIGN.

RESULTATS

ANALYSE DES AFLPS

Un total de 224 bandes AFLPs ont été codées pour les trois couples d'amorces utilisés. La combinaison AGA/CAG (1) avec 89 bandes (39.7%) permet de révéler le plus grand nombre de polymorphismes suivi du couple AGT/CTC (2) avec 78 bandes (34.8%) et le couple AAT/CTC (3) avec 57 (25.4%). La figure 1 représente le dendrogramme obtenu avec les AFLPs ($r = 0.89$). Quatre groupes principaux sont reconnus ; le manioc cultivé et la *ssp. flabellifolia* (C1), *M. glaziovii* et les arbres manioc (C2); *M. brachyloba* et *M. quinquepartita* (C3) et *M. carthaginensis* (C4). Des analyses de similarité conduites indépendamment pour chaque couple d'amorces AFLP utilisé ont permis d'observer une cohérence entre les couple 1 et 3 ce qui est ratifié avec la corrélation obtenue entre ces deux matrices de similarités ($r = 0.83$). Les valeurs de corrélation de ces deux couples avec le couple 2 sont plus faibles ($r = 0.63$ et 0.67 respectivement).

La figure 2 montre la distribution des 76 échantillons en fonction des deux premiers axes de l'analyse de la même matrice en PCO qui expriment 30% de la variance totale observée. L'axe 1 sépare le manioc et son ancêtre sauvage direct du complexe d'espèces *M. carthaginensis* s.l. dont la divergence est ainsi confirmée. L'existence de deux groupes identifiables dans *M. glaziovii*, déjà apparente sur la figure 3 est aussi confirmée. Les deux échantillons d'Afrique (17 et 42) se retrouvent dans le même sous-groupe *M. glaziovii*, le plus éloigné de *M. esculenta* et n'apparaissent donc pas être hybrides avec le manioc. Par contre l'échantillon originaire d'Indonésie (11) classifié comme *M. glaziovii* se regroupe avec les arbres manioc, intermédiaires entre *M. esculenta* et *M. glaziovii*. Parmi ceux-là il est remarquable que l'arbre manioc de Guyane, réputé avoir été introduit récemment du Surinam, se regroupe étroitement avec celui du Venezuela (37 et 41). Alors que le « maniçoba » (28) apparaît intermédiaire entre le manioc et les arbres manioc, la variété cultivée triploïde (29) se regroupe étroitement avec une autre variété « manipeba de corda » (30) du Nord-Est brésilien et apparaît avoir plus d'affinité avec *M. carthaginensis* que les « arbres manioc » comme cela est confirmé par l'examen des distances génétiques effectivement observées. Les espèces *M. brachyloba* et *M. quinquepartita* se regroupent et apparaissent moins divergentes que *M. carthaginensis* par rapport au manioc.

La figure 3 représente la distribution des échantillons de *M. esculenta* en fonction du premier et troisième axes de variation qui expriment les valeurs de variance attendues. La dispersion des échantillons montre une faible structure génétique de ces échantillons. L'axe 1 qui extrait 11.4 de la variance totale observée sépare les échantillons en deux groupes d'accessions. Sur les valeurs négatives de l'axe 1, on retrouve des variétés regroupés avec NGA2: Col22, SRT1316, Res2, Res1, Bra 110, Ven45 qui correspondent à celles identifiées comme candidates introgressées dans un travail antérieur (Second 1999). On remarque aussi qu'une variété mexicaine (99, 66, 67) est isolée dans un pôle opposé à celles-la. Notons aussi que la variété Sansan, seule dans notre échantillonnage à représenter les variétés « amères » est mal classée dans ce plan.

POLYMORPHISME DES MICROSATELLITES

Un total de 269 allèles ont été révélés chez les 60 individus étudiés. Le plus grand nombre d'allèles a été observés chez *M. esculenta* (166) suivi des arbres manioc (156), *M. glaziovii* (154), *M. e. ssp. flabellifolia* (137), *M. carthaginensis* (82), *M. quinquepartita* (57) et *M. brachyloba* (53). Vingt-deux locus présentent entre 10 et 18 allèles et seulement quatre locus

présentent moins de dix allèles. Du total d'allèles observés, 78 sont partagés entre le manioc et les espèces sauvages, tandis que 115 (43%) sont uniques à une espèce déterminée. *M. glaziovii* avec 27 allèles est l'espèce qui présente le plus grand nombre d'allèles uniques (8.5%) suivi de *M. esculenta* avec 23 (Fig.4). Dix allèles uniques a été le nombre maximum détecté pour un seul locus (SSRY- 147).

Le niveau de polymorphisme pour chacune des espèces varie entre 0.59 pour *M. brachyloba* et 0.96 pour le manioc avec une moyenne de polymorphisme observée de 0.87. Les locus SSRY- 31 et SSRY- 145 avec 23 et 20 allèles respectivement sont les plus polymorphes (Fig.5) tandis que le locus SSRY- 146 est le moins polymorphe avec 6 allèles. Le locus SSRY- 68 est le plus polymorphe pour le manioc cultivé puisque 10 allèles ont été révélés. Au niveau du nombre d'allèles par locus, le manioc est le plus polymorphe avec une moyenne de 6.1 allèles par locus suivi de *M. glaziovii* et les arbres manioc avec 5.7.(Table 3).

Taille, fréquences et distribution des allèles SSR

La plus grande différence de taille entre allèles d'un même locus a été observée pour le locus SSRY- 100 avec 62 pb tandis que le locus SSRY- 146 avec 11 pb représente la moindre différence entre allèles observés. Les autres locus ont des allèles distribués entre 16 et 58 pb. La figure 6 montre comme exemple les différences en taille des allèles entre *M. glaziovii*, *M. e. ssp. flabellifolia* et *M. esculenta* observés pour le locus SSRY-23 et SSRY-40. On peut voir que la taille des allèles pour chaque espèce diffère dans certains cas en accord avec les observations de Roa et al. (2000).

La plus part des allèles ont une fréquence inférieure à 0.3 y compris au niveau intra-spécifique, sauf pour l'allèle SSR38-114pb dont la fréquence est élevée chez *M. esculenta*, *M. e.ssp.flabellifolia*, les arbres manioc et *M. glaziovii*. Trois cas d'allèles nuls (pas d'amplification) ont été observés; le locus SSRY- 9 chez *M. glaziovii* et SSRY- 88 chez *M. brachyloba* et *M. carthaginensis*.

Niveau d'hétérozygotie

Seuls *M. esculenta* et les arbres manioc présentent des valeurs similaires d'hétérozygotie observée et attendue. Contrairement à ce que l'on observe chez le manioc cultivé, *M. glaziovii*, *M. carthaginensis* et *M. e.ssp. flabellifolia* présentent un déficit en génotypes hétérozygotes ce qui reflète une structuration de la diversité dans ces espèces. Chez les arbres manioc, le génotype des individus est largement hétérozygote (hétérozygotie moyenne observée: 0.81) et révèle notamment une structure *M. esculenta* x *M. glaziovii* bien que l'on distingue aussi des locus suggérant une hybridation *M. esculenta* x *M. carthaginensis*.

En ce qui concerne les valeurs Fis, on ne peut pas les estimer lorsque une espèce ou une population est monomorphe pour un locus donné (marqué comme PA sur le tableau 3). Ainsi, un 57.1 % de valeurs Fis sont supérieures que zéro ce qui signifie qu'il n'y a pas d'association aléatoire des allèles chez les espèces testées

Polymorphisme comparé du manioc cultivé et de M. carthaginensis s.l.

Comme le montre le tableau 3, le manioc se caractérise par une forte diversité allélique avec le plus grand nombre d'allèles par locus polymorphe et une hétérozygotie moyenne observée plus élevée que chez *M. glaziovii* (Tableau 3). Du total d'allèles observés, il existe 14 allèles communs entre *M. esculenta* et *M. glaziovii* et six avec *M. carthaginensis* qui sont absents chez *M. e. ssp. flabellifolia*. Parmi ces allèles, cinq sont exclusifs de *M. glaziovii* et deux de *M. carthaginensis* car les 9 restants sont aussi présents chez une autre espèce sauvage. En regardant les fréquences de ces allèles par espèce on distingue trois situations; soit l'allèle en

commun est fréquent (>0.1) chez les manioc et *M. glaziovii*, soit l'allèle en commun est fréquent chez *M. glaziovii* et rare chez *M. esculenta* et soit l'allèle en commun est plutôt rare chez les deux espèces (Tableau 4). Le tableau 4 présente aussi les moyennes des fréquences de ces allèles calculées pour chaque espèce en fonction du nombre d'allèles partagés avec le manioc. Dans la plus part des cas ces allèles sont rares parmi les échantillons de manioc analysés et il pourrait bien s'agir d'allèles introgressés (Tableau 4).

Le tableau 4 présente le total d'allèles partagés entre *M. esculenta*, et les espèces sauvages autres que *M. e. ssp flabellifolia* utilisées dans notre étude (15.6% du total observés). On note que 30 de ces allèles sont présents chez les arbres manioc, 15 chez *M. glaziovii*, six chez *M. carthaginensis*, et quatre chez *M. quinquepartita* et *M. brachyloba*. En ce qui concerne exclusivement les allèles candidats d'introgession de *M. glaziovii* chez le manioc, différents niveaux d'introgession peuvent être attendus puisqu'un hybride *M. esculenta* x *M. glaziovii* requière trois rétrocroisements pour retrouver un manioc cultivé (Jennings 1957 in Lefèvre, 1989). Le détail des clones qui présentent ces allèles se trouvent dans le tableau 4. Les variétés brésiliennes SRT1316 (64), SRT 1344 (65), BR12 (50) MBRA110 (68) et la variété introgressée NGA2 (75) présentent entre 25 et 27% de ces allèles.

Analyse multivariée

La figure 7 représente un dendrogramme obtenu à partir de la matrice de similarité pour les 26 locus SSR. Quatre grands groupes sont observés. Le groupe 1 correspond à *M. brachyloba*, le groupe 2 à *M. quinquepartita*, le groupe 3 comprend le manioc, les arbres manioc et *M. e. ssp flabellifolia*, et le groupe 4 correspond au complexe *M. carthaginensis* s.l. avec les mêmes sous groupes de *M. glaziovii* reconnus précédemment avec les AFLP. La valeur de corrélation cophénétique ($r= 0.94$) valide le regroupement des accessions. Les niveaux de diversité intra-spécifique observés sont importants puisque la similarité entre accessions d'une même espèce oscille entre 0.9 pour deux clones de manioc (MBRA110 et BRA383) et 0.15 pour les deux accessions de *M. quinquepartita*.

La figure 8 montre l'analyse en coordonnées principales (Axes 1 et 3) obtenue avec l'ensemble de SSR employés sur tous les échantillons. On reconnaît une structure de l'ensemble des espèces assez homologues à celle obtenue avec les AFLP (Figure 2). L'axe 1 qui extrait 10% de la variance observée sépare le complexe *M. carthaginensis* s.l. du manioc cultivé. Les arbres manioc, et les espèces *M. flabellifolia* et *M. quinquepartita* apparaissent comme intermédiaires et ces deux dernières se regroupent et apparaissent dans le graphique moins divergentes que *M. carthaginensis* s.l. par rapport au manioc cultivé. Comme avec les AFLP, le « manicoba » (28) présente une affinité étroite avec la variété « manipeba de corda » (30) et apparaît intermédiaire entre le manioc et les arbres manioc. On observe parmi le groupe du manioc des variétés présumées introgressées qui s'apparentent avec la variété Manicoba (28) et à la variété artificiellement introgressée NGA2 (75).

Des analyses ne retenant que les locus SSR localisés sur un même groupe de liaison ont été menées pour observer les relations entre espèces par rapport à un seul fragment chromosomique (Tableau 2). La figure 9 représente la distribution obtenue avec les 7 locus SSR du groupe de liaison D chez *M. esculenta* et *M. glaziovii*. Le premier axe qui exprime 21% de la variance sépare le manioc de *M. glaziovii* mais au long de l'axe 2 trois groupes de *M. esculenta* sont évidents. Un premier groupe (C1) qui est le plus proche de *M. glaziovii* est formé par des variétés présumées introgressées comme montré sur les figures et d'autres échantillons africains (70, 59) mais aussi originaires d'Argentine et Venezuela (46, 56, 73). Ces échantillons partagent un important nombre d'allèles avec *M. glaziovii*. Le groupe 2 (C2)

comprend un ensemble de variétés comme SRT1316 (64), Col22 (62), BRA383 (63), MBRA110 (68) et le Manicoba (28).

De la même façon, la figure 10 montre la PCO obtenu avec 5 locus du groupe de liaison F chez le manioc et *M. glaziovii*. Au long de l'axe 1 on observe plusieurs échantillons intermédiaires entre le manioc et *M. glaziovii* avec trois groupes d'accessions intermédiaires formés au long de l'axe 2. Un premier groupe (C1) qui est le plus proche de *M. glaziovii* comprend le manicoba (28), Res2 (71) et Mex69 (74) ; un deuxième groupe (C2) formé par une variété colombienne (53), SRT1316 (64) et MCUB51 (49) et le groupe C3 qui est le moins proche de *M. glaziovii* avec quatre variétés introgressées du groupe 1 de la figure 8 et quelques autres comme la variété colombienne Col22 (62).

Des associations différentes des échantillons de manioc sont obtenues lorsqu'on analyse les échantillons par rapport aux autres fragments chromosomiques (Groupes G, L, K et SSR non cartographiés). En général aucune structuration forte n'est observé (Analyses non montrées). Par ailleurs, les résultats de l'analyse factorielle des correspondances (graphiques non montrés) a permis de déterminer les allèles partagés (présumés introgressés) entre le manioc et le complexe d'espèces *M. carthaginensis s.l.*

Comparaison entre microsatellites et AFLPs

Pour comparer l'information donnée par chaque type de marqueur, les 60 accessions pour lesquelles on dispose des données de microsatellites et des marqueurs AFLPs ont été comparées en utilisant le test de Mantel qui permet d'obtenir un indice de corrélation entre les matrices de similarité générées par chaque technique. Les coefficients de corrélation obtenus entre les matrices de similarité ainsi que les matrices cophénétiques calculées pour chaque type de marqueur sont statistiquement significatifs, $r = 0.85$ et 0.80 respectivement.

Séquençage de bandes présumées introgressées

NB : Ces résultats ne peuvent être présentés qu'à titre provisoire car cette partie du travail n'est pas terminée.

La PCR avec un nouveau couple d'amorces a révélé la présence de cette séquence chez neuf accessions qui ne présentaient pas le fragment AFLP original (quatre accessions de *M. esculenta*, une accession *M. brachyloba*, une *M. quinquepartita*, deux accessions de *M. glaziovii* et deux accessions de *M. flabellifolia*). La séquence de cette bande chez certaines accessions ainsi qu'un phylogramme par rapport à cette séquence se présente sur la figure 11. Les variétés SRT1344 (65), Col22 (62) et SRT1366 (64) montrent une affinité avec une accession de *M. glaziovii* de Brésil sous une valeur de *bootstrap* (63) qui fournit une bonne robustesse à ce groupe. Le reste des accessions se localisent sur la partie basale du phylogramme avec *M. flabellifolia* qui a été utilisé comme *outgroup*.

Les coefficients de similarité et de divergence des régions flanquantes de SSR analysés sont présentés dans la table 5. Pour les allèles de trois locus (SSRY40, SSRY55 et SSRY135) le pourcentage de similarité de séquences de *M. esculenta* est majeur avec *M. flabellifolia* et leur pourcentage de divergence moindre par rapport à *M. glaziovii*. Pour deux locus (SSRY100 et SSRY31) la similarité des séquences de *M. esculenta* est similaire avec celles des espèces sauvages mais le pourcentage de divergence est moindre par rapport la séquence de *M. flabellifolia*. Finalement, seulement le locus 151 montre une introgression puisque la majeure similarité est donné entre *M. esculenta* et *M. glaziovii* ainsi que le moindre pourcentage de divergence.

DISCUSSION

Polymorphisme comparé de M. esculenta par rapport à M. carthaginensis s.l.

L'utilisation des AFLPs et SSRs a mis en évidence la relation génétique du manioc par rapport à un des complexes d'espèces le plus divergent parmi le genre *Manihot*. Nos résultats confirment cette divergence entre *M. esculenta* et le complexe d'espèces *M. carthaginensis* ainsi que la proximité génétique de *M. e. ssp. flabellifolia* considérée comme l'ancêtre du manioc cultivé en concordance avec Roa et al. (1997, 2000) et Second (1999). Les analyses phénétiques en dendrogrammes avec les deux types de marqueurs regroupent bien les accessions par espèces ou complexe d'espèces. Toutefois, l'emploi de l'analyse en coordonnées principales permet de mieux visualiser les relations entre les accessions d'un même groupe et d'observer les relations multiples entre les groupes par l'examen de certains axes de variance (Second, 1999).

L'utilisation des marqueurs co-dominants comme les SSRs permet le calcul d'autres paramètres de la diversité qui ne sont pas disponibles avec les marqueurs AFLPs (Santoni, 1999). Notre étude confirme l'amplification au niveau inter-spécifique des marqueurs microsatellites développés chez le manioc par Mba et al. (2001). Les résultats montrent que *M. esculenta* se caractérise par une forte diversité allélique et un niveau d'hétérozygotie plus élevé que les espèces du complexe *M. carthaginensis s.l.* Cette observation doit cependant être tempérée par le fait que notre échantillonnage n'est pas exhaustif. En particulier, *M. e. ssp. flabellifolia* a été représenté par un faible nombre d'accessions même si elle sont d'origine géographique très variées. Il s'agissait de connaître la position relative de l'espèce pour les marqueurs considérés et non sa richesse allélique. Les taux d'hétérozygotie observés chez le manioc qui sont équivalents aux valeurs attendues suggèrent un brassage génétique entre les populations en accord avec la tradition amérindienne qui rassemble de la diversité dans un même champ et utilise les germinations de graines sexuées pour créer de nouvelles variétés (Empéaire et al., 1998 ; Elias, 2000). En outre, le grand nombre d'allèles spécifiques observé pour *M. glaziovii* et *M. carthaginensis* suggère que la séparation de ces espèces du manioc cultivé à l'échelle de l'évolution n'est pas récente et que la présence d'allèles communs peuvent être issus d'événements d'hybridations, mêmes si les risques d'homoplasie entre allèles ne peuvent être écartés. L'existence des formes intermédiaires ou hybrides entre *M. esculenta* et *M. carthaginensis s.l.* révèle la possibilité de flux génique entre ces espèces. Les génotypes des hybrides sont généralement hétérozygotes et montrent notamment une composition *M. esculenta* x *M. glaziovii* en concordance avec les génotypes hybrides révélés par les isoenzymes (Lefèvre, 1989). Néanmoins, quelques hybrides ont révélé un génotype *M. esculenta* x *M. carthaginensis* à certains locus. Ces allèles ont été observés surtout chez quelques variétés originaires de Colombie où le flux génique est favorisé par la sympatrie des espèces. La diversité observée des hybrides peut être attribuée à plusieurs raisons : d'abord l'existence de plusieurs niveaux d'inter croisement ou de retro-croisements sur l'espèce cultivée mais aussi la diversité des génotypes *M. esculenta* et *M. glaziovii* impliqués dans les processus de flux géniques entre les deux espèces (Lefèvre, 1989). La présence de certaines accessions de *M. glaziovii* plus proches du manioc cultivé que d'autres, à la fois au niveau AFLP et SSR, suggère que la variabilité de cette espèce peut aussi être due à des introgressions à partir du manioc cultivé.

Les combinaisons génotypiques des hybrides incluent aussi des allèles uniques qui n'ont pas été observés parmi notre matériel. Ceci peut s'expliquer par trois hypothèses, la plus probable étant que ces allèles sont présents chez *M. glaziovii* ou *M. esculenta* mais n'ont pas été

échantillonnés, ou que ces allèles proviennent d'autres espèces du genre, ou enfin qu'il s'agit de nouveaux allèles produits suite à l'hybridation.

Variabilité intra-spécifique du manioc cultivé

Le niveau de polymorphisme observé chez *M. esculenta* reflète une bonne représentativité génétique du matériel, même si dans cette étude la représentation de l'espèce a été largement biaisée en faveur des manioc doux, dans lesquels avaient été faites les premières observations d'introgession. On sait que la divergence principale entre variétés de manioc se trouve d'abord entre les variétés amérindiennes douces et amères (Second et al 2000). Il conviendra donc d'étendre la représentation de l'espèce cultivée aux manioc amers dans la poursuite de cette étude, d'autant que l'origine de la divergence de ceux-là n'est pour l'instant pas éclaircie. L'analyse multivariée basée sur les AFLPs (Figure 12) rappelle le regroupement précédemment obtenu avec des marqueurs RAPDs et AFLP (Second, 1999) et suggère d'autre part que d'autres divergences comme celle de la variété Mex59 pourrait avoir leur origine dans des introgressions autres que celles qui ont été abordées ici. Les regroupements de variétés sud-américaines avec la variété introgressée artificiellement en Afrique suggère cependant que des introgressions spontanées à partir de croisements différents peuvent aussi conduire à des patrons d'introgession non aléatoire comme documenté dans le genre *Helianthus* (Reiseberg et al., 1998). Par contre dans le cas des marqueurs SSR, l'analyse multivariée basée sur l'ensemble des locus montre une dispersion des variétés sans structure apparente. Des groupes de variétés qui s'expliquent par des introgressions apparaissent seulement au niveau de segments chromosomiques individualisés. Il apparaît donc que la diversité infra spécifique du manioc est influencée par des introgressions qui sont mieux repérées globalement dans notre cas par les AFLP car le nombre de locus révélés est plus élevé. Par contre les SSR cartographiées permettent une caractérisation plus fine mais qui nécessite l'utilisation d'un ensemble de marqueurs couvrant le génome entier.

Indices d'introgession

Cette étude apporte de nouvelles observations confirmant, bien que dans l'ensemble encore indirectement, l'existence d'introgessions à l'origine de certains polymorphismes du manioc. La présence d'allèles communs entre le manioc cultivé et *M. carthagenensis s.l* concerne généralement des allèles fréquents chez le pool sauvage et rare chez le manioc cultivé. L'utilisation d'une variété artificiellement introgressé –la variété NGA2 (75) qui a été utilisé comme parent de la carte génétique du manioc (Fregene et al., 1997)- a permis d'établir des marqueurs SSRs d'introgession. Basé sur le nombre de marqueurs d'introgessions présents, une variété africaine (59) et la variété RES1 (70) montre un niveau d'introgession équivalent au parent de la carte. Ces résultats sont aussi corroborés au niveau AFLPs. D'autres variétés qui seraient aussi introgressées présenteraient un niveau moindre d'introgession. Ceci est en accord avec Lefèvre (1989) qui a trouvé différents niveaux d'introgession de *M. glaziovii* chez *M. esculenta*. Ces indices confirment que *M. carthagenensis* bien qu'étant l'espèce la plus divergente du manioc en Amérique du Sud a participé à sa diversité génétique. L'existence d'allèles d'autres espèces comme *M. quinquepartita* dans certains génotypes hybrides spontanés et du manioc cultivé suggèrent aussi la contribution d'autres espèces du pool secondaire à la diversité actuelle de *M. esculenta*. La contribution possible d'autres espèces comme *M. fructiculosa* (Second, 1999) pourrait être étudiée, bien que leur proximité plus grande du manioc complique encore leur mise en évidence certaine.

L'introgession d'allèles « sauvages » dans le génome cultivé serait possible grâce à l'importance de la reproduction sexuelle dans l'évolution du manioc (Elias, 2000). Il reste néanmoins à préciser si les hybridations interspécifiques ont put être à l'origine de la

domestication du manioc en stimulant les potentialités d'évolution ou si elles sont plutôt à l'origine d'adaptations secondaires comme pour la résistance à la bactériose par exemple.

Corrélations observées entre locus d'un même fragment chromosomique

Selon Lefèvre (1989), une carte génétique saturée avec plusieurs groupes de liaisons par chromosome offre une meilleure possibilité de suivre l'introgression de gènes sauvages et le marquage de caractères d'intérêt agronomique. Notre étude a appliqué cette notion et des analyses par apport aux locus d'un même fragment chromosomique ont été menées. Les analyses du groupe de liaison D confirment l'hypothèse de Jorge (2000) selon laquelle ce fragment correspond à une introgression de *M. glaziovii* chez un des parents de la carte génétique, mais montrent aussi que ce fragment introgressé est présent chez quelques autres variétés qui n'ont a priori pas de relation généalogique avec elle. En utilisant le parent de la carte comme indicateur, nos résultats montrent que le groupe de liaison F pourrait aussi correspondre à une région d'introgression de *M. glaziovii*. Le fait qu'un ensemble de locus anonymes comme les AFLPs reflètent une homologie avec ces résultats permet de postuler l'hypothèse suivante: le regroupement de variétés issues d'événements d'introgression indépendants s'explique par le fait qu'elles partagent une fraction significative du génome total, correspondant à plusieurs fragments chromosomiques conservés et qui pourraient correspondre à des patrons stables d'introgression comme montré dans le genre *Helianthus* (Reiseberg et al., 1998).

Méthodologie de démonstration des introgressions

La présence d'allèles « sauvages » dans les variétés étudiées du manioc peut être expliquée par deux hypothèses: mutations donnant des allèles différents mais homoplasiques de ceux observés dans les populations sauvages, où bien il peut s'agir d'allèles spécifiques de formes ou espèces sauvages qui ont introgressé des variétés cultivées par hybridation. Pour tester ces hypothèses, le séquençage comparé entre les différentes espèces qui présentent l'allèle (régions flanquantes des SSR ou bandes AFLP) a été tenté. Les résultats obtenus au niveau de régions flanquantes du locus microsatellite ont confirmé l'introgression d'un allèle de *M. glaziovii* chez la variété brésilienne SRT1316 et dans les autres cas on ne peut pas conclure définitivement car la taille de la séquence conduit à un faible nombre de polymorphismes observés. La capacité de transférer les amorces d'amplification des SSR du manioc cultivé aux espèces génétiquement éloignées met d'ailleurs en évidence le niveau de conservation des régions flanquantes. Comme les amorces pour l'amplification par PCR d'un locus microsatellite sont en général dessinées le plus proche possible de la région du microsatellite, pour la suite de cette étude on envisage de dessiner de nouvelles amorces à partir de la séquence des clones à l'origine du développement des SSR (Mba, com. pers.). Ceci permettra l'amplification de régions flanquantes plus grandes et informatives.

La conversion d'un fragment AFLP en marqueur d'introgression, par le dessin de nouvelles amorces et l'amplification d'une portion du fragment dans l'espèce *M. e. ssp. flabellifolia* a permis le séquençage comparé et de confirmer l'introgression de *M. glaziovii* (Brésil) chez trois variétés sud-américaines. Cette séquence pourra être cartographiée et le segment chromosomique correspondant mieux exploré. La méthodologie de capture de tels fragments apparaît cependant complexe et d'un faible rendement, notamment par le fait qu'elle implique la coloration à l'argent moins efficace que la radioactivité pour la révélation des AFLP et aussi par la limitation liée à la taille des fragments en AFLP. Une poursuite de ce travail pourrait inclure la modification de ce protocole mais aussi l'utilisation d'autres marqueurs.

Les RAPDs par exemple offrent la possibilité d'obtenir des fragments de taille plus grands et ont aussi montré leur capacité à repérer les introgressions (Colombo et al., 1997).

On peut aussi envisager l'utilisation d'autres marqueurs qui ont été déjà cartographiés. Fregene et al. (2000) ont identifié un fragment AFLP associé à la résistance à la mosaïque virale africaine (CMV). Cette résistance a aussi été obtenue d'une introgression de *M. glaziovii*. La confirmation que ce fragment correspond à une introgression pourrait être faite en appliquant la méthodologie que nous proposons. L'hybridation *in situ* à travers la technique de *chromosome painting* pourrait aussi être appliquée pour comparer la localisation des ces marqueurs chez différents variétés introgressés, bien que la petite taille du génome du manioc représente une limitation certaine à l'utilisation de cette méthode. Le développement de sondes à partir de ces fragments permettra d'incorporer un élément supplémentaire et très utile dans l'étude de la domestication et des flux géniques parmi le genre *Manihot*. L'analyse de séquences répétées par la technique des *microarray* (un test de cette approche est en cours au laboratoire du Pr. Heslop-Harisson et inclue le présent matériel *Manihot* (JC Pintaud et al., en prep.) et l'utilisation des « *diversity arrays* » (Jaccoud et al., 2000) déjà adoptée au CIAT (Cali, Colombie) sont aussi des outils potentiels d'un grand intérêt pour analyser les introgressions. Une caractérisation plus complète du manioc à ce niveau fournirait des renseignements précieux d'ordre phylogénétique et permettrait de mieux comprendre le phénomène de spéciation chez le genre *Manihot*. N'oublions pas enfin que, dans la mesure ou nombre d'introgressions s'expliquent par l'avantage adaptatif qu'elles confèrent, leur caractérisation représente aussi une stratégie pour identifier des gènes d'importance directe dans le génome.

CONCLUSIONS

Notre travail a précisé la structure génétique des espèces du complexe *M. carthaginensis* s.l. par rapport au manioc cultivé et son espèce ancestrale postulée, ainsi que la structure génétique du groupe des hybrides spontanées. De nouveaux indices d'introgression chez le manioc ont été apportés et on confirme l'introgression de *M. glaziovii* chez un des parents de la carte génétique ainsi que la présence spontanée de ce fragment dans d'autres variétés introgressés en Amérique Tropicale. Nos observations indiquent que ces introgressions correspondent à plusieurs fragments dans le génome du manioc et qu'ils peuvent se trouver à différents niveaux dans les variétés introgressés spontanées ce qui contribue à la diversité génétique du manioc. On a aussi montré que les introgressions sont mieux repérées, à travail égal, par les marqueurs AFLPs car ils révèlent un plus grand nombre de locus. A partir de la variété artificiellement introgressée, on propose des marqueurs microsatellites d'introgression qui peuvent être utilisés pour le marquage et le suivi de introgressions ou de caractères d'intérêt agronomique. La participation d'autres espèces du pool secondaire du manioc à la diversité du manioc ne doit pas être écartée. Il se pourrait en particulier que l'espèce *M. guaranitica* qui se trouve dans le région du Chaco, entre la Bolivie le Paraguay et l'Argentine, fasse partie du complexe *M. carthaginensis* s.l. et cette région est riche en hybrides interspécifiques de manioc (G. Second Com. pers.). D'autre part, il est remarquable que les variétés de manioc originaires du Mexique et Guatemala soient particulièrement riches en allèles SSR par ailleurs rares dans l'espèce (Chavariaga et al., 1999) et que nos données AFLP vont aussi dans le sens que cette région pourrait être à l'origine d'une diversité originale qui pourrait trouver sa source dans l'introgression.

L'existence des introgressions spontanées parmi certaines variétés traditionnelles de manioc cultivé est partiellement démontrée. Le séquençage de bandes présumés introgressés semble

bien pouvoir fournir la preuve d'introgression de *M. glaziovii* chez le manioc si bien que le séquençage et la recherche d'autres marqueurs d'introgression doit être poursuivie. La cartographie de ces marqueurs permettra tester l'hypothèse de la stabilité des patrons d'introgression chez le manioc. L'inclusion de matériel d'autres origines comme l'Amérique Centrale et le Chaco sud américain devrait être envisagée. Sans aucun doute, l'impact des introgressions sur la diversification du manioc en Amérique du Sud est un thème d'intérêt à plusieurs titres.

Références bibliographiques

- Allem A.C., 1994. The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resources and Crop Evolution*, 41 : 133-150.
- Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. 2001 GENETIX 4.02, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
- Bertram R.B., 1993. Application of molecular techniques to genetic resources of cassava (*Manihot esculenta* Crantz, Euphorbiaceae)/ interspecific evolutionary relationships and intraspecific characterization. Thèse PhD, University of Maryland, Etats-Unis, 465 p.
- Carvalho R., Guerra M. and Carvalho P.C.L. 1999. Occurrence of spontaneous triploidy in *Manihot esculenta* Crantz. *Cytologia* 64 : 000 -000.
- Chavarriga-Aguirre P., Maya M.M., Bonierbale M.W., Kresovich S., Fregene M.A., Tohme J., Kockert G., 1998. Microsatellites in cassava (*Manihot esculenta* Crantz) : discovery, inheritance and variability. *TAG*, 97 (3): 493-501.
- Chavarriga-Aguirre P., Maya M.M., Tohme J., Duque M. C., Iglesias I., Bonierbale M.W., Kresovich S., Kockert G. 1999. Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. *Molecular breeding* 5 : 263-273.
- Colombo C., 1997. Etude de la diversité génétique de maniocs américains (*Manihot esculenta* Crantz) par les marqueurs moléculaires (RAPD et AFLP). Thèse de doctorat, ENSAM, Montpellier, France, 144 p.
- Colombo C., Second G., Losada Valle T. et Charrier A., 1998. Genetic diversity characterization of cassava cultivars (*Manihot esculenta* Crantz). I) RAPD markers. *Genetics and Molecular Biology* 21, 11 105-113.
- Colombo C., Second G., Losada Valle T. et Charrier A., 1998. Genetic diversity characterization of cassava cultivars (*Manihot esculenta* Crantz). I) RAPD markers. *Genetics and Molecular Biology* 21, 11 105-113.
- Elias M. 2000. Sélection naturelle, sélection humaine, et diversité chez une plante domestiquée propagée de façon végétative : le cas de la culture du manioc par les indiens Makushi du Guyana. Synthèse. Thèse Doctorat. Université Montpellier II. pp.22-25.
- Empereire E., Pinton F., Second G. 1998. Gestion dynamique de la diversité variétale du manioc en Amazonie du nord-ouest. *Nature, science et société*, 6(2) :27-42.
- FAO, 2000. [HTTP://WWW.FAO.ORG/AG/MAGAZINE/SP1.HTM](http://www.fao.org/ag/magazine/sp1.htm)
- Fregene, M.A., J. Vargas, F. Angel, J. Tohme, R. A. Asiedu, M. O. Akorada, and W.M. Roca. 1994. Chloroplast DNA and nuclear ribosomal DNA variability in cassava (*Manihot esculenta* Crantz) and its wild relatives. *TAG* 89 : 719-727.

- Fregene M., Angel F., Gomez.R., Rodriguez F., Chavarriaga P., Roca W., Tohme J. and Bonierbale M. 1997. A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor Appl Genet* 95:431-441
- Fregene M., Bernal A., Duque M., Dixon A., Tohme J. 2000. AFLP analysis of African cassava (*Manihot esculenta* Crantz) germplasm resistant to the cassava mosaic disease (CMD^o. TAG 100, 678-685.
- Hahn S. K., 1984. Les plantes a racines et tubercules tropicales : amelioration et utilisation. Ibadan, Nigeria, IITA, Rapport de conférence n^o2, 32p.
- Hanlon P.&Peakall R. 2000. A simple method for the detection of size homoplasy among amplified fragment length polymorphism fragments. *Molecular Ecology* 9, 815-816.
- Jaccoud D., Peng K., Feinstein D. and Kilian A. 2001. Diversity arrays : a solid state technology for sequence information independant genotyping. *Nucleic acids research* 29, 4 e25.
- Jorge, V., Fregene, M.A., Duque M.C, Bonierbale M.W, Tohme J. and Verdier V. 2000. Genetic mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz). *Theor Appl Genet* 101 :865-872.
- Lefevre F., 1989. Ressources génétiques et amelioration du manioc, *Manihot esculenta* Crantz, en Afrique. Paris, France, Orstom, Travaux et documents microédités n^o 57, 175p.
- Mba R.E.C, Stephenson P., Edwards K., Melzer S., Nkumbira J., Gullberg U., Appel K., Gale M., Tohme J. and Fregene M. 2001. Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome : towards an SSR-based molecular genetic map of cassava. *Theor Appl Genet* 102 :21-31.
- Olsen K. et Schaal B. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proc Natl. Acad. Sci. USA*. Vol. 96. Pp. 5586-5591.
- Reiseberg, L.H., Sinervo, B., Linder, C.R., Ungerer, M.C. and Arias, D.M. Role of gene interaction in hybrid speciation: evidence from ancient and experimental hybrids. *Science* 272:741-745, 1998.
- Roa, A.C., M. M. Maya, M.C.Duque, J. Tohme, A. C ., Allem, and M. W. Bonierbale, 1997. AFLP analysis of relationships among cassava and other *Manihot* species. *Theoretical and Applied Genetics* 95 : 741-750.
- Roa A.C., P. Chavarriaga-Aguire, M. C. Duque, M.M. Maya, M.W. Bonierbale, C.I. Iglesias and J. Tohme. 2000. Cross-species amplification of cassava (*Manihot esculenta*) microsatellites : allelic polymorphism and degree of relationship. *Amer. J. Bot.* 87(11) : 1647-1655
- Rogers D.J., Appan M., 1973. *Manihot, Manihotoides* (Euphorbiaceae). New York, Etats-Unis, Hafner Press, *Flora Neotropica Monograph* n^o13, 274 p.
- Rolph, F. J. 1994. NTSYS-pc: numerical taxonomy and multivariate system, version 1.80. Exeter Software, Setauket, New York, USA.
- Santoni S., Faivre-Rampant P., Prado E. et Prat D. 2000. Marqueurs moléculaires pour l'analyse des ressources génétiques et l'amelioration des plantes. *Cahiers Agricultures*. 9:311-327.
- Second G., Allem A., Mendes R., Carvalho L., Emperaire L., Ingram C. and Colombo C. 1997. Molecular Marker (AFLP)-based *Manihot* and cassava numerical taxonomy and genetic structure analysis in progress : Implications for their dynamic conservation and genetic mapping. *Genome*. Vol. 2: Pages 140-147.
- Second G., Raffailac J. P. and C. Colombo. 1999. Manioc. In: *Diversité génétique des plantes cultivées tropicales*. Eds: Hamon, P.; M. Seguin; X. Perrier; J. C. Glaszmann. Collection Repères. CIRAD. 385p.

- Second, 2000. *Manihot glaziovii* contributed to the genetic make-up of cassava and represents an example of dynamic conservation and on-farm breeding of genetic resources. In Proceedings of IV International Scientific Meeting of Cassava Biotechnology Network.
- Second G, Iglesias C. 2000. The state of use of cassava genetic diversity and a proposal to enhance it. In Broadening the Genetic Basis of Crop Production (eds. Cooper HD, Spillane D, Hodgkin T^o, pp. 201-221.
- Serier J.B., 1989. Historico da disseminação da maniçoba fora do Brasil. In : Primeiro encontro nordestino da maniçoba. Recife, Brésil, IPAA, p.89-95.
- Sokal, R. R.& J. Rolph. 1995. Biometry : the principles and practice of statistic in biological research, 3rd. Ed. W.H. Freeman, New York, New York, USA.
- Templeton, A.; Lawrance, J. 1988. FRAG-LENGTH III: algorithm for the calculation of molecular weights. Software distributed by the authors.
- Vos P. Hogers, R. Bleeker, M. Reijnders, M. van de Lee, T. Hornes, M. Fritjers, A. Pot, J. Peleman, J. Kuiper, M. & Zabeau M. 1995. AFLP : a new technique for DNA fingerprinting. Nucleic Acid Research 21 : 4407-4414.

Tableau 1. Liste d'accessions de *Manihot* sélectionnées pour l'étude d'introgessions.

Code IRD (Man 5-)	N° collection d'origine	ESPECE	Région d'origine	Coll.d'origine	Retenus pour SSRs
0	Sansan	<i>M. esculenta</i>	Guyane	GS	-
1	BLO-001	<i>M. brachyloba</i>	-	CIAT	x
2	BLO-402	<i>M. brachyloba</i>	-	CIAT	x
4	-	<i>M. brachyloba</i>	Guyane	DL	x
5	5/00-6	<i>M. glaziovii</i>	Brésil	PCC	-
6	5/00-7	<i>M. glaziovii</i>	Brésil	PCC	x
7	5/00-8	<i>M. glaziovii</i>	Brésil	PCC	x
8	5/00-9	<i>M. glaziovii</i>	Brésil	PCC	x
9	5/00-10	<i>M. glaziovii</i>	Brésil	PCC	x
10	GLA-IAC	<i>M. glaziovii</i>	Brésil	IAC	x
11	IND4	<i>M. glaziovii</i>	Indonésie	IRD	-
17	127G	<i>M. glaziovii</i>	Afrique	IRD	x
18	167	<i>M. glaziovii</i>	Afrique	IRD	x
19	5/00-17	<i>M. pseudoglaziovii</i>	PSE005	CIAT	x
20	5/00-18	<i>M. pseudoglaziovii</i>	PSE011	CIAT	-
22	A3150	<i>M. epruinosa</i>	Brésil	CENARGEN	-
24	5/00-24	<i>M. quinquepartita</i>	-	CIAT	x
25	GUY18	<i>M. quinquepartita?</i>	Guyane	GS	x
28	5/00-5	<i>Manicoba</i>	Brésil	CENARGEN	x
29	5/00-16	<i>Manipeba triploid</i>	Brésil	PCC	-
30	W579	<i>Manipeba de corda</i>	Brésil	CENARGEN	x
31	W846	<i>Arbre manioc de 7 anos</i>	Brésil	CENARGEN	x
32	5/00-11	<i>Arbre manioc de 7 anos</i>	Brésil	PCC	x
33	5/00-12	<i>Arbre manioc de 7 anos</i>	Brésil	PCC	x
34	5/00-13	<i>Arbre manioc de 7 anos</i>	Brésil	PCC	-
35	5/00-14	<i>Arbre manioc de 7 anos</i>	Brésil	PCC	x
37	5/00-1	<i>Arbre manioc</i>	Guyane	GS	x
41	5/00-35	<i>Arbre manioc?</i>	Venezuela	FF/GS	-
42	UGA1	<i>M. glaziovii</i>	Uganda	GS	-
44*	5/00-30	<i>M. esculenta</i>	Venezuela	FF/GS	-
45*	5/00-31	<i>M. esculenta</i>	Venezuela	FF/GS	x
46*	5/00-32	<i>M. esculenta</i>	Venezuela	FF/GS	x
47*	5/00-33	<i>M. esculenta</i>	Venezuela	FF/GS	x
48	76 F Ouhoud	<i>M. esculenta</i>	Guyane	FOR	-
49	MCUB51	<i>M. esculenta</i>	-	CIAT	x
50	BR12	<i>M. esculenta</i>	-	CIAT	x
51	ECU82	<i>M. esculenta</i>	-	CIAT	x
52	PAN51	<i>M. esculenta</i>	-	CIAT	x
53	COL1438	<i>M. esculenta</i>	-	CIAT	x
54	COL2066	<i>M. esculenta</i>	-	CIAT	x
55	TAI1	<i>M. esculenta</i>	-	CIAT	x
56	ARG7	<i>M. esculenta</i>	-	CIAT	x
57	CM2177-2	<i>M. esculenta</i>	-	CIAT	x
58	Abyssinica	<i>M. esculenta</i>	Afrique	IRD	x
59	30395	<i>M. esculenta</i>	Afrique	IRD	x
60	Viro3	<i>M. esculenta</i>	Afrique	IRD	-
61	Col22 (E49)	<i>M. esculenta</i>	-	IRD	-
62	MCol22(E53)	<i>M. esculenta</i>	-	IRD	x
63	Bra383(E55)	<i>M. esculenta</i>	-	IRD	x
64	SRT1316	<i>M. esculenta</i>	Brésil	CENARGEN	x
65	SRT1344	<i>M. esculenta</i>	Brésil	CENARGEN	x
66	MMEX59	<i>M. esculenta</i>	-	CIAT	-
67	Mex59(E54)	<i>M. esculenta</i>	-	IRD	-
68	MBRA110	<i>M. esculenta</i>	-	CIAT	x
69	MCOL22	<i>M. esculenta</i>	-	CIAT	-
70	Res1	<i>M. esculenta</i>	Uganda	GS	x
71	Res2	<i>M. esculenta</i>	Uganda	GS	x
72	BRA110	<i>M. esculenta</i>	-	CIAT	x
73	VEN45	<i>M. esculenta</i>	-	CIAT	x
74	MEX69	<i>M. esculenta</i>	-	CIAT	x
75	NGA2	<i>M. esculenta</i>	-	CIAT	x
76	TME5	<i>M. esculenta</i>	-	CIAT	x
78	5/00-21	<i>M. carthaginensis</i>	CTH017-A	CIAT	x
79	5/00-23	<i>M. carthaginensis</i>	CTH190	CIAT	x
80	5/00-25	<i>M. carthaginensis</i>	Venezuela	FF-GS	x
82	5/00-27	<i>M. carthaginensis</i>	Venezuela	FF-GS	x
83	5/00-28	<i>M. carthaginensis</i>	Venezuela	FF-GS	x
85	5/00-34	<i>M. carthaginensis</i>	CTH13	CIAT	x
88	A4004	<i>M. flabellifolia</i>	Brésil, Rondonia	CENARGEN	x
89	A4001	<i>M. flabellifolia</i>	Brésil, Rondonia	CENARGEN	x
90	A4169	<i>M. flabellifolia</i>	Brésil, Goias	CENARGEN	x
91	A4245	<i>M. flabellifolia</i>	Brésil, Maranhon	CENARGEN	x
93	82 F Ouhoud	<i>M. flabellifolia?</i>	Brésil, Amapa	FOR	-
94	GUY3	<i>M. flabellifolia?</i>	Guyane	ME	x
97	5/00-19	<i>M. tristis</i>	TST026	CIAT	x
99	MMEX59	<i>M. esculenta</i>	-	CIAT	-

Collections d'origine : Références : CIAT, CENARGEN, IAC

Collections de feuilles : IRD, DL (Denis Loubry), FF-GS (Francia Fuenmayor-Gérard Second), PCC (Paulo César de Carvalho), FOR (François Ouhoud-Renoux), ME (Marianne Elias), GS (Gérard Second)

*: Variétés à grosses tiges

Tableau 2. Liste de loci microsatellites sélectionnés pour cette étude.

No.	Code Life Tech.*	Taille Produit PCR (pb)	Temp. Annealing.	Séquence du motif	Chromosome	Reference
1	SSRY-3	247	55	(CA)17	D	Mba <i>et al.</i> , 2001
2	SSRY-9	278	55	(GT)15	D	Mba <i>et al.</i> , 2001
3	SSRY-23	247	45	(CA)17	D	Mba <i>et al.</i> , 2001
4	SSRY-40	231	55	(GA)16	D	Mba <i>et al.</i> , 2001
5	SSRY-108	203	55	(CT)24CCT	D	Mba <i>et al.</i> , 2001
6	SSRY-120	139	55	(CA)7AA(GA)8	D	Mba <i>et al.</i> , 2001
7	SSRY-169	100	55	(GA)19A3(GAA)2	D	Mba <i>et al.</i> , 2001
8	SSRY-31	188	55	(GA)21	F	Mba <i>et al.</i> , 2001
9	SSRY-68	287	55	(CT)12CC(CT)17	F	Mba <i>et al.</i> , 2001
10	SSRY-179	226	55	(GA)28	F	Mba <i>et al.</i> , 2001
11	SSRY-30	220	55	(CT)22	G	Mba <i>et al.</i> , 2001
12	SSRY-38	122	55	(CA)17	G	Mba <i>et al.</i> , 2001
13	SSRY-103	272	55	(GA)22	G	Mba <i>et al.</i> , 2001
14	SSRY-135	253	45	(CT)16	G	Mba <i>et al.</i> , 2001
15	SSRY153	117	45	GA-131	G	Chavarriaga <i>et al.</i> , 1998
16	SSRY-55	145	55	(GA)16	L	Mba <i>et al.</i> , 2001
17	SSRY-88	243	55	(CA)7(N)6(GA)22	L	Mba <i>et al.</i> , 2001
18	SSRY-100	210	55	(CT)17TT(CT)7	L	Mba <i>et al.</i> , 2001
19	SSRY-144	117	45	GAGG-5	Q	Chavarriaga <i>et al.</i> , 1998
20	SSRY-145	143	45	GA-12	pd	Chavarriaga <i>et al.</i> , 1998
21	SSRY-146	139	45	GA-13	pd	Chavarriaga <i>et al.</i> , 1998
22	SSRY-147	113	45	GA-16	pd	Chavarriaga <i>et al.</i> , 1998
23	SSRY-148	114	45	GA-21	pd	Chavarriaga <i>et al.</i> , 1998
24	SSRY-151	182	45	GA-126	K	Chavarriaga <i>et al.</i> , 1998
25	SSRY-152	233	45	GA-127	K	Chavarriaga <i>et al.</i> , 1998
26	SSRY-154	318	45	GA-134	pd	Chavarriaga <i>et al.</i> , 1998
27	SSRY-155	158	45	GA-136	pd	Chavarriaga <i>et al.</i> , 1998

* = Ref.: *Cassava Map Pairs, Life Technologies*

Tableau 3. Valeurs d'heterozygotie (Hexp et Hobs), proportion de locus polymorphes, nombre moyen d'allèles par locus et valeurs Fis obtenus.

LOCUS	Espèces (Nombre d'accessions)													
	BRACH (3)		QUINQ (2)		CARTH (6)		GLAZ (9)		HYB (7)		ESC (27)		FLA (6)	
	Hexp	Hobs	Hexp	Hobs	Hexp	Hobs	Hexp	Hobs	Hexp	Hobs	Hexp	Hobs	Hexp	Hobs
SSRY-3 Fis	0.61	0.66	0	0	0.51	0.16	0.73	0.33	0.70	0.57	0.64	0.40	0.61	0
	0.04		PA		0.46		0.57		0.30		0.54		1.20	
SSRY-9 Fis	0	0	0	0	0	0	0	0	0.77	0.71	0.71	0.72	0.62	0.16
	PA		PA		PA		PA		0.22		0.04		0.46	
SSRY-23 Fis	0.44	0	0.5	0	0.51	0.16	0.82	0.55	0.66	0.57	0.65	0.38	0.56	0.33
	1.5		2.00		0.46		0.38		0.34		0.24		0.35	
SSRY-40 Fis	0	0	0.5	0	0	0	0.77	0.44	0.45	0.42	0.78	0.62	0.72	0.50
	PA		2		PA		0.51		0.01		0.18		0.22	
SSR-108 Fis	0.50	0.33	0.50	0	0.65	0.16	0.88	0.55	0.74	0.85	0.66	0.59	0.77	0.66
	0.66		2		0.71		0.37		-0.03		0.025		0.24	
SSRY-120 Fis	0.61	0.33	0.50	0	0.62	0.50	0.79	0.55	0.71	0.85	0.81	0.44	0.73	0.16
	0.66		2		0.16		0.26		-0.06		0.43		0.65	
SSRY-169 Fis	0.61	0.33	0.50	0	0.62	0.16	0.82	0.88	0.44	0.57	0.29	0.33	0.73	0.16
	0.66		2.0		0.46		0.06		-0.01		-0.03		0.65	
SSRY-31 Fis	0.61	0.33	0.62	0.50	0.77	0.83	0.83	0.55	0.84	0.85	0.78	0.85	0.83	0.66
	0.66		0.50		0		0.33		0.08		-0.05		0.22	
SSRY-68 Fis	0.66	0	0.37	0.50	0.68	0.16	0.47	0.22	0.71	0.85	0.84	0.85	0.77	0.33
	1.5		0		0.65		0.39		-0.07		0.03		0.65	
SSRY-179 Fis	0.44	0	0.62	0.5	0	0	0.42	0.44	0.54	0.42	0.75	0.88	0.80	0.50
	1.5		0.50		PA		-0.06		0.32		-0.08		0.43	
SSRY-30 Fis	0	0	0.37	0.50	0.61	0	0.76	0.22	0.75	1	0.55	0.55	0.79	0.50
	PA		0.0		1.2		0.58		-0.13		-0.06		0.43	
SSRY-38 Fis	0.72	0.66	0.50	0	0.56	1	0.72	1	0.62	0.71	0.41	0.48	0.51	0.33
	0.25		2.00		-0.45		-0.11		-0.05		-0.13		0.10	
SSRY-103 Fis	0.61	0.33	0.62	0.50	0.15	0.16	0.45	0.11	0.83	1	0.80	0.70	0.61	0.33
	0.66		0.50		-0.0		0.74		-0.09		0.07		0.33	
SSRY-135 Fis	0	0	0.50	0	0	0	0.77	0.33	0.55	0.57	0.78	0.62	0.81	0.16
	PA		2.00		PA		0.45		-0.00		0.248		0.90	
SSRY-55 Fis	0.66	0.66	0	0	0.48	0.33	0.65	0.77	0.79	1	0.79	0.92	0.80	0.33
	0.25		PA		0.32		-0.44		-0.07		-0.11		0.613	
SSRY-88 Fis	0	0	0.62	0.50	0	0	0.50	0.22	0.73	1	0.76	0.62	0.75	0.33
	PA		0.50		PA		0.68		-0.13		0.05		0.49	
SSRY-100 Fis	0.66	0	0.50	0	0.66	0.33	0.67	0.44	0.79	1	0.81	0.88	0.81	0.33
	1.5		2.0		0.28		0.29		-0.09		-0.05		0.68	
SSRY144 Fis	0.44	0	0.62	0.50	0	0	0.69	0	0.66	1	0.40	0.55	0.15	0.16
	1.5		0.50		PA		1.12		-0.27		-0.37		-0.00	
SSRY145 Fis	0	0	0.5	0	0.86	0.83	0.87	0.77	0.78	0.71	0.67	0.62	0.80	0.33
	PA		2		0.10		0.11		0.095		0.05		0.46	
SSRY146 Fis	0	0	0.5	0	0.27	0	0.49	0	0.68	0.85	0	0	0.44	0
	PA		2.0		1.20		1.12		-0.03		PA		1.2	
SSRY147 Fis	0	0	0.50	1	0.76	0.50	0.83	0.22	0.56	0.71	0.53	0.74	0.41	0.50
	PA		-1.0		0.42		0.61		-0.03		-0.09		-0.02	
SSRY148 Fis	0.27	0.33	0.50	0	0.37	0.50	0.74	0.44	0.72	1	0.47	0.44	0.76	0.33
	0.0		2		-0.26		0.535		-0.18		0.07		0.55	
SSRY151 Fis	0	0	0.62	0.50	0	0	0	0	0.64	1	0.74	0.70	0.81	0.66
	PA		0.50		PA		PA		-0.27		0.011		0.17	
SSRY152 Fis	0	0	0.37	0.50	0.56	0.16	0.80	0.55	0.87	1	0.66	0.77	0.72	0
	PA		0.0		0.54		0.29		-0.03		-0.05		1.2	
SSRY153 Fis	0.5	1	0.62	0.50	0.56	1	0.62	0.44	0.82	1	0.75	0.66	0.83	0.83
	-1.0		0.50		-0.45		0.22		-0.07		0.27		0.03	
SSRY154 Fis	0	0	0.50	0	0.51	0.16	0.80	0.44	0.64	0.85	0.47	0.48	0.69	0.50
	PA		2.0		0.46		0.352		-0.17		0.08		0.1821.5	
SSRY155 Fis	0.44	0	0.50	0	0.15	0.16	0.63	0.66	0.81	0.85	0.67	0.70	0.68	0.66
	1.5		2.0		-0.0		-0.04		0.08		0.00		0.00	
<i>Moy.H exp</i>	0.32		0.46		0.40		0.65		0.70		0.63		0.69	
<i>Moy.H obs.</i>	0.18		0.22		0.27		0.41		0.81		0.61		0.36	
<i>P (0.95)</i>	0.59		0.88		0.74		0.92		1.0		0.96		1.0	
<i>P (0.99)</i>	0.59		0.88		0.74		0.92		1.0		0.96		1.0	
<i>Moy.al./locus</i>	1.9		2.1		3.0		5.7		5.7		6.1		5.0	

Tableau 4. Fréquences des allèles partagés entre *M.esculenta*, *M.glaziovii*, *M.carthaginensis* et des hybrides interspécifiques spontanées (absents chez *M.flabellifolia*).

Chrom.	Locus	Allèle	M.esc	Hybrides	M.glaziovii	<i>M.cart h.</i>	<i>M.qui n</i>	<i>M.bra ch</i>	Clones (Code IRD 5-, voir tableau 1)
D	3	204	0.07	0.07	0.44	-	-	-	75,50,59
		216	0.03	0.28	0.11	-	-	-	46
	9	244	0.1	0.07	-	-	-	-	62,49,74,65,72
		256	0.04	0.21	-	-	-	-	74
	23	214	0.05	-	0.16	-	-	-	75,50,59
		228	0.03	-	0.16	0.08	-	0.33	75,70
		235	0.5	0.14	-	-	0.5	-	62,64,74,71,28,47,65,45,50,68,53,54,51,57,52,58,63
	40	218	0.11	0.71	0.38	-	-	-	64,74,50,68,63
		229	0.11	-	0.16	-	-	-	62,49,28,76,58
	120	115	0.05	-	0.16	-	-	-	59
		143	0.11	-	0.22	-	-	0.5	54,51,57,52
		154	0.14	0.07	-	-	-	-	65
	169	91	0.11	-	-	0.08	-	-	64,65,72,51,57,56
		92	0.01	-	-	-	-	0.5	50
F	31	173	0.03	0.07	-	-	-	-	62
	68	239	0.01	-	0.66	0.16	-	0.33	71
		258	0.14	0.35	-	-	-	-	64,47,65,72,63
		262	0.05	0.14	-	-	-	-	54
		281	0.25	0.07	-	-	-	-	47,65,46,55,45,50,68,53,57,52,56,63
179	233	0.01	-	-	-	0.02	-	53	
G	38	126	0.14	0.14	-	-	-	-	64,47,46,68,57,63
	103	271	0.18	0.14	-	-	-	-	64,74,65,46,55,68,72,52,63
	135	263	0.07	-	0.05	-	-	-	62,73
	153	102	0.20	0.07	0.55	0.08	-	-	75,70,47,65,55,53,54,51,57,59
		124	0.03	0.07	-	-	-	-	58
L	55	134	0.09	0.07	-	-	-	-	50,68,72,56
	88	239	0.01	0.07	-	-	0.25	-	50
	100	206	0.07	-	0.05	-	-	-	75,59
		215	0.09	0.28	-	-	-	-	64,74,65,76,58
		235	0.01	0.07	-	-	-	-	47
K	151	176	0.03	0.07	-	-	-	-	75,59
		181	0.42	0.28	-	-	-	-	62.64.74.28.73.47.65.55.50.68.53.54.57.56.58.63
		188	0.01	0.14	-	-	-	-	64
	152	215	0.07	0.07	0.16	-	-	-	62,70,51,58
		224	0.11	0.21	-	-	-	-	59,63,75,70,68,64
		229	0.05	0.07	-	-	-	-	47,46,45
I N C O N N U	144	126	0.27	0.21	-	-	-	-	75,49,64,28,65,55,45,50,72,54,51,58,59
	147	114	0.33	0.07	0.16	-	-	-	62.75.49.71.47.65.46.45.50.68.76.72.51.57.52.56.59.63
		126	0.01	-	-	0.16	-	-	54
	154	308	0.24	0.07	-	-	0.5	-	49,71,46,68,76,53,54,51,57,52,59
		331	0.01	0.27	-	-	-	-	73
155	153	0.05	0.14	0.16	0.08	-	-	75,64,28	
Moyennes			0.09	0.15	0.24	0.1	0.31	0.4	

Tableau 5. Tableaux de pourcentage de similarité et de divergence (en italique) calculés pour les séquences de régions flanquants de 6 loci SSR chez différentes accessions. Le code de l'accession (IRD, Table1) est en gras.

SSRY-40	76	5	88
76 (NGA2)	-	79.8	84
5 (M.glaz)	7.3	-	89.4
88 (M.fla)	4.5	6.8	-

SSRY-31	75	17	88
75 (NGA2)	-	97.6	97.6
17 (M.glaz)	19.8	-	100
88 (M.fla)	8.1	24	-

SSRY-151	25	17	64	88
25 (M.quin)	-	23.2	35.4	20.7
17 (M.glaz)	32.9	-	40.5	36.7
64 (SRT1316)	24	9	-	34
88 (M.fla)	39.3	21.6	25.6	--

SSRY-100	62	25	17	88
62 (Mcol22)	-	94.4	96.4	96.3
25 (M.quin)	6.6	-	92.7	98.1
17 (M.glaz)	5.3	8.5	-	92.7
88 (M.fla)	0	2.7	1.6	

SSRY-135	28	17	88	79
28 (Manicoba)	-	35.4	68.2	38.5
17 (M.glaz)	1.1	-	34.7	31.1
88 (M.fla)	2.6	1.1	-	46.8
79 (M.cth)	1.1	0	1.6	-

SSRY-55	79	10	50	91
79 (M.cth)	-	79	82	89.2
10 (M.glaz)	2.8	-	66.7	71.6
50 (BR12)	5.7	4.4	-	91.0
91 (M.fla)	4.3	4.2	3.4	-

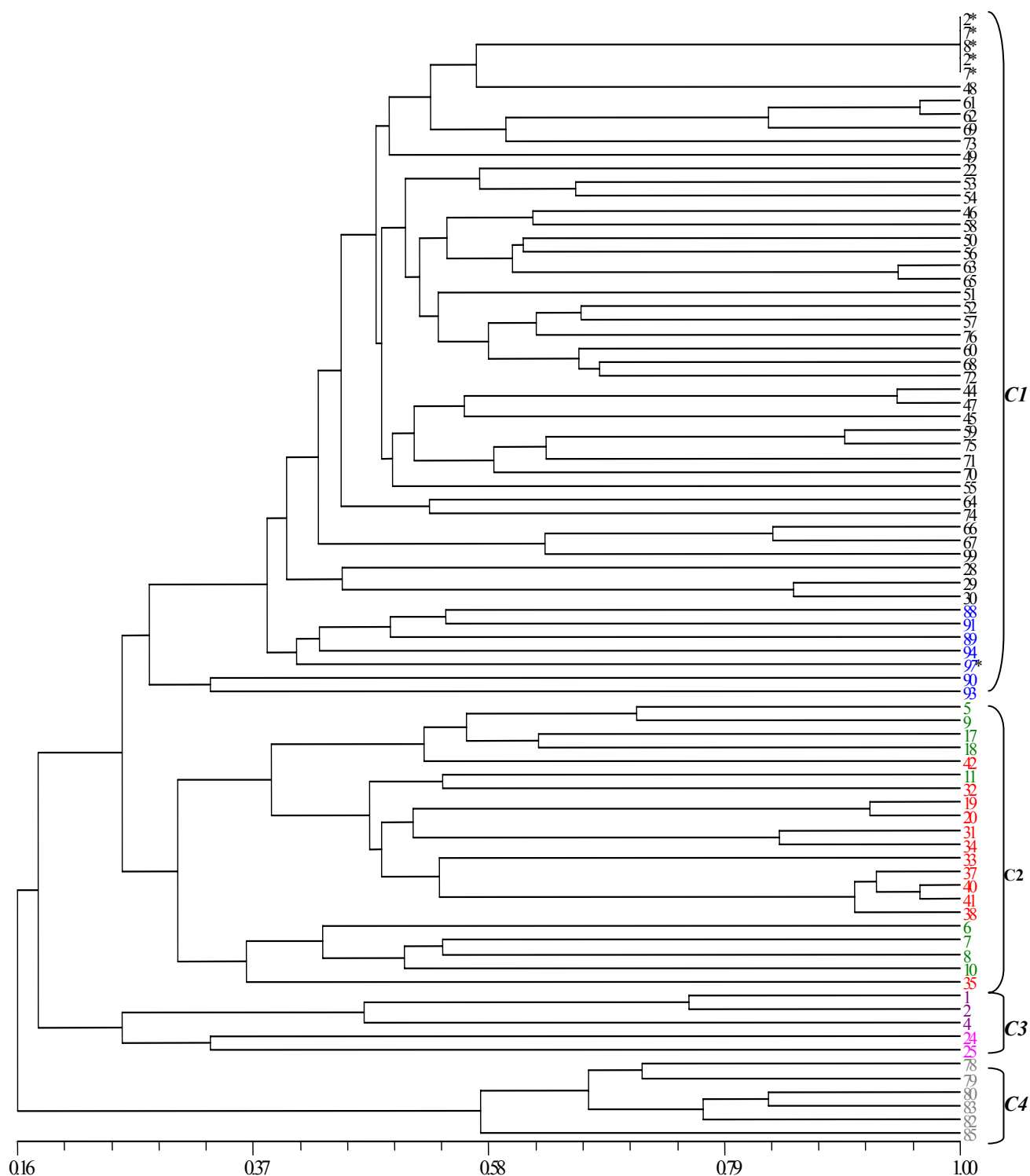


Figure 1. Dendrogramme basé sur 224 AFLPs qui montre la structure génétique de 82 accessions de manioc, hybrides et quelques espèces sauvages. En noir : *M.esculenta* ; blue : *M.flabellifolia* (**M.tristis*), rouge : arbres manioc, vert : *M.glaziovii* ; gris : *M.carthaginensis*; rose : *M.quinquepartita*, violet : *M.brachyloba* et marron clair : *M.epruinosa*..

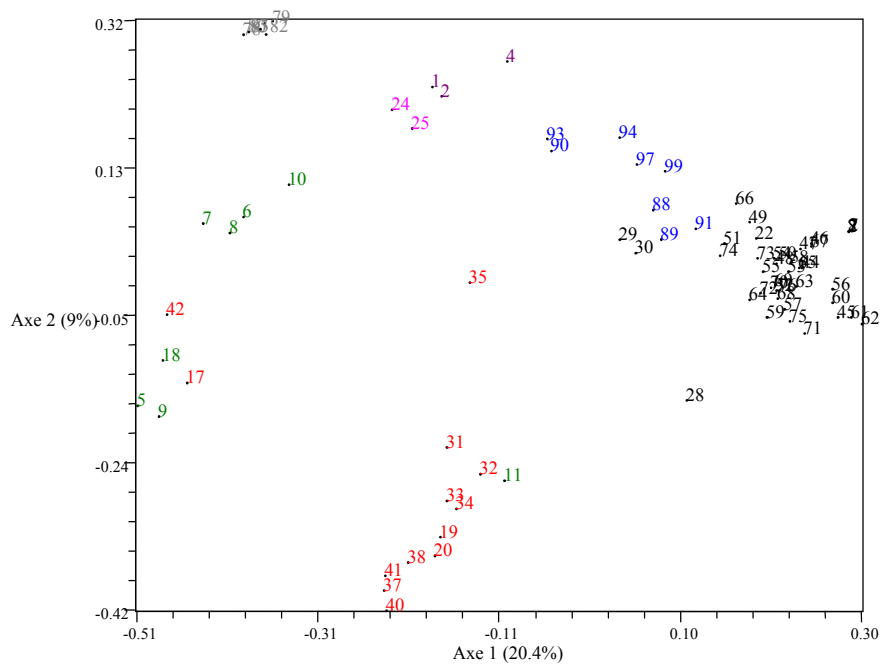


Figure 2. Distribution dans le plan défini par les deux premiers vecteurs *eigen* (le pourcentage de la variance exprimé est indiqué sous l'axe correspondant) du PCO basé sur le coefficient de similarité SMC dans les comparaisons deux a deux de 76 génotypes de manioc avec 224 AFLPs, hybrides et quelques espèces sauvages. En noir : *M.esculenta* ; bleu : *M.flabellifolia*, rouge : arbres manioc, vert : *M.glaziovii* ; gris : *M.carthagenensis*; rose : *M.quinquepartita*, et violet : *M.brachyloba*.

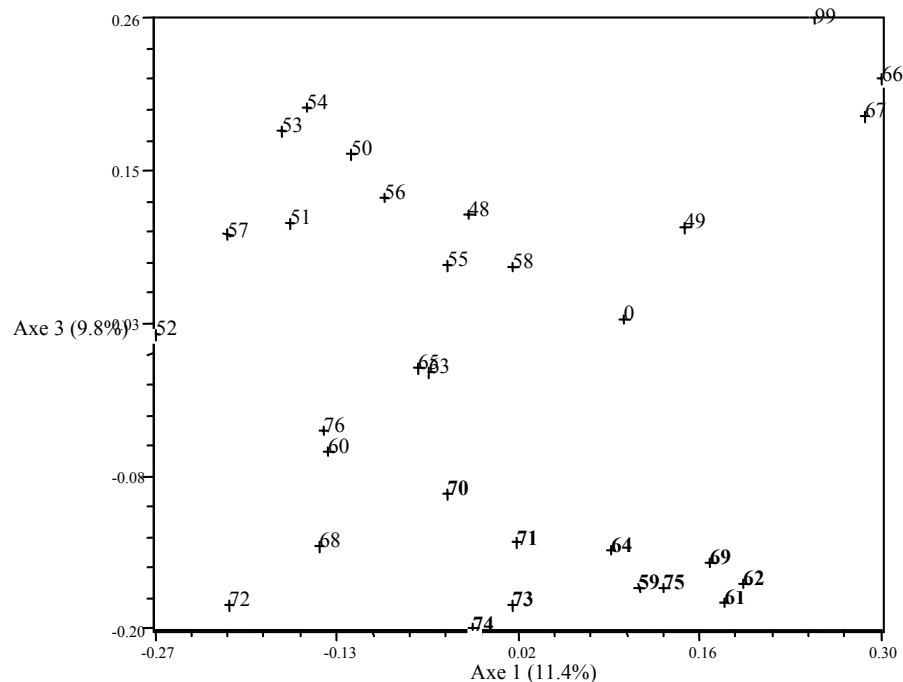


Figure 3. Distribution dans le plan défini par le premier et le troisième axe du PCO basé sur le coefficient de similarité Jaccard dans les comparaisons deux a deux de 35 génotypes de manioc avec 224 AFLPs. En gras les variétés présumés introgressés.

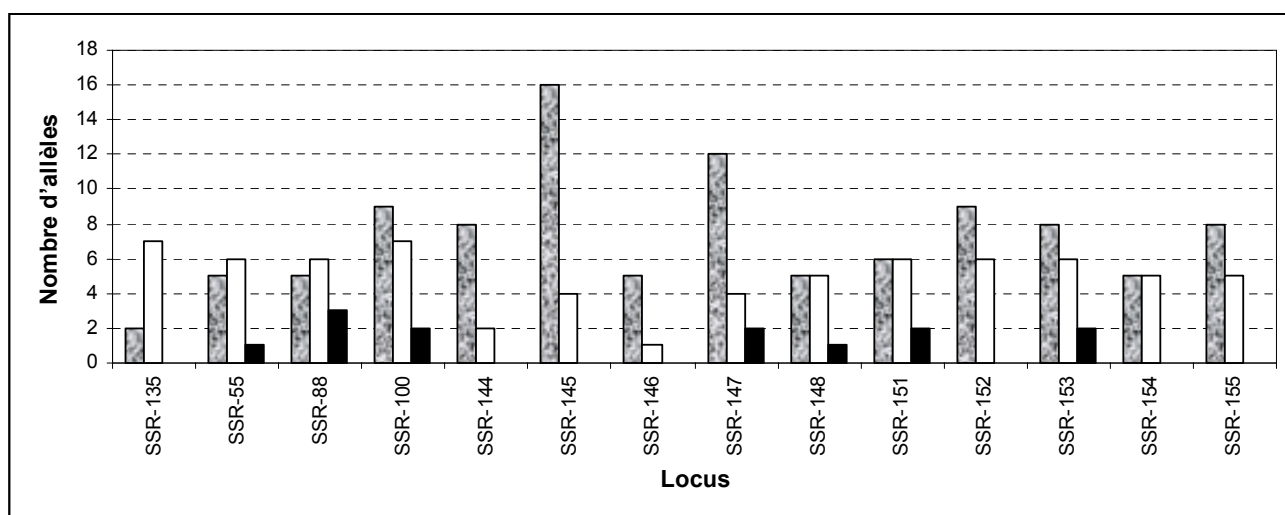
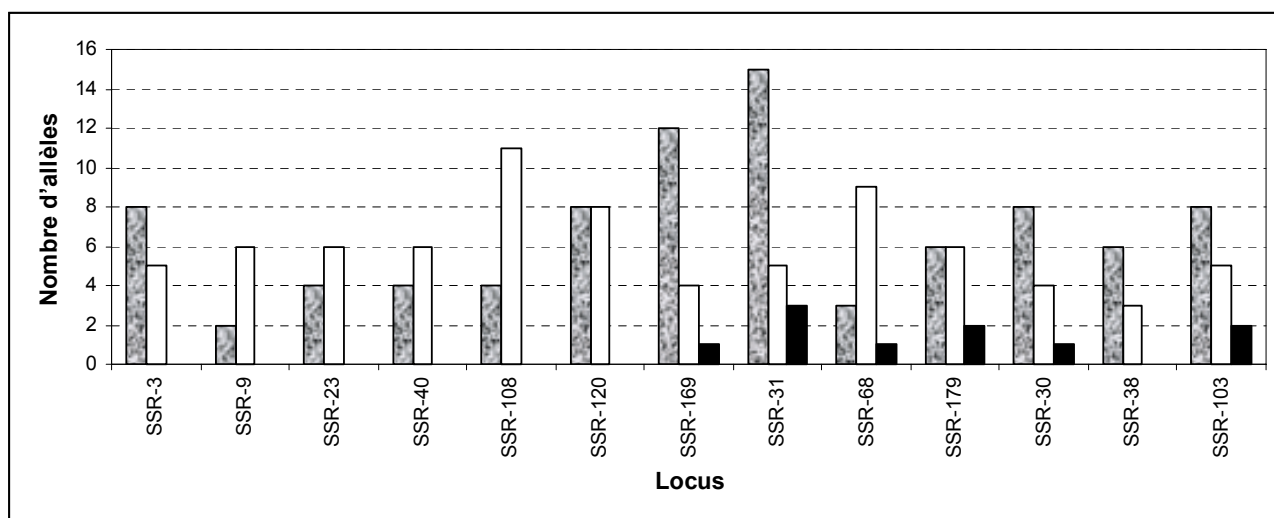


Figure 4. Nombre et distribution d'allèles observés chez 7 espèces *Manihot* dans 27 loci microsatellites étudiés. Barres noires: allèles exclusives du manioc, barres grises: allèles présents que chez les espèces sauvages, barres blanches, allèles partagés entre le manioc et les espèces sauvages.

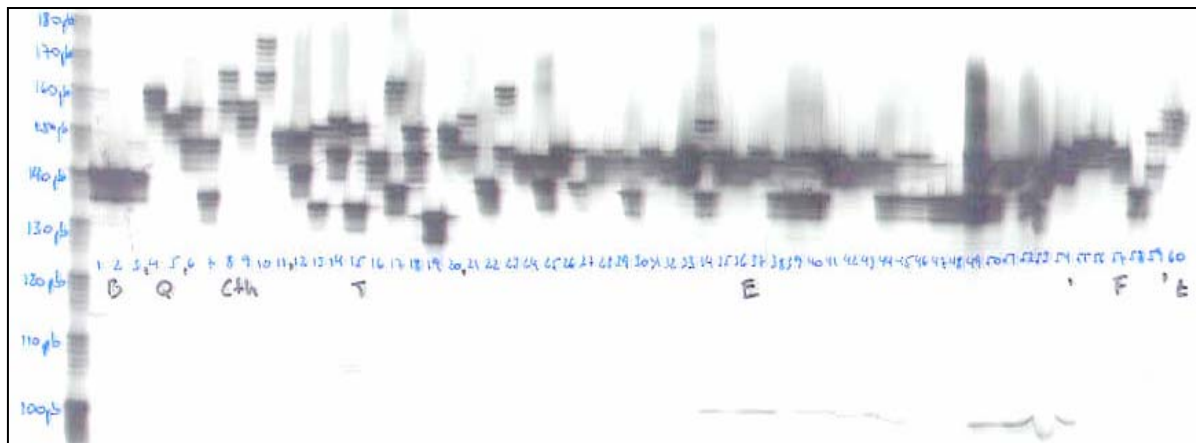
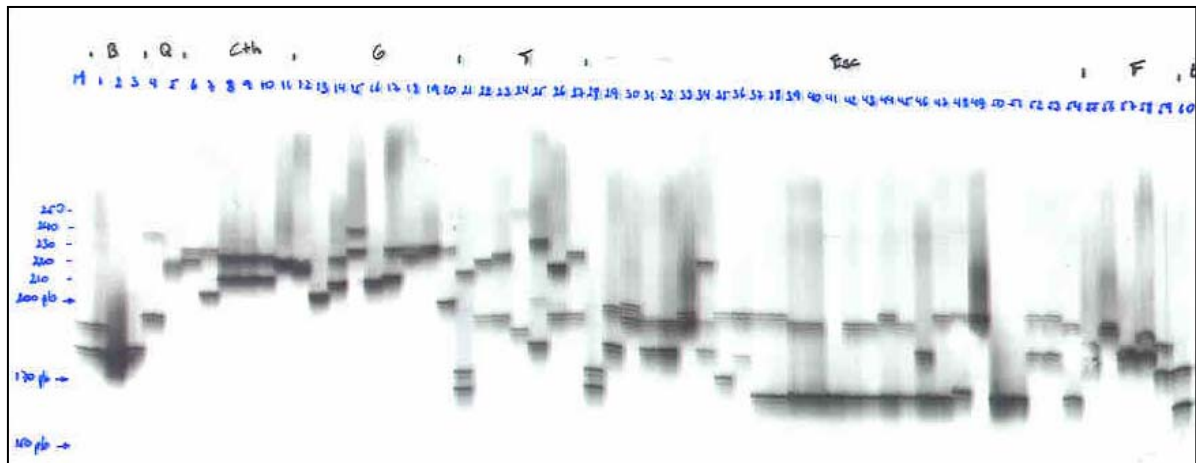


Figure 5. Produits SSR radiomarqués obtenus a partir d'ADN génomiques de manioc cultivé et quelques espèces sauvages. En haut les profils avec le couple d'amorces SSRY-31 et en bas avec le couple SSRY-145.

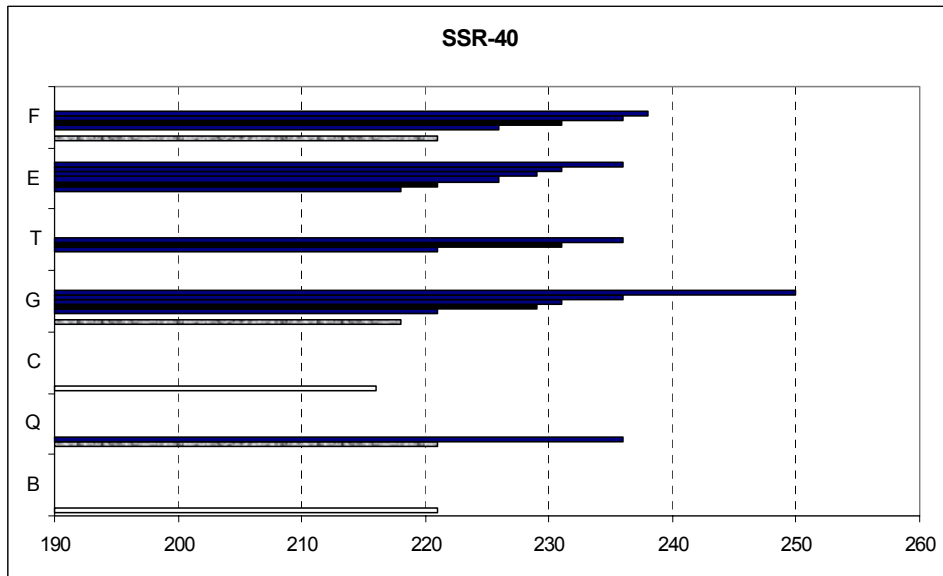
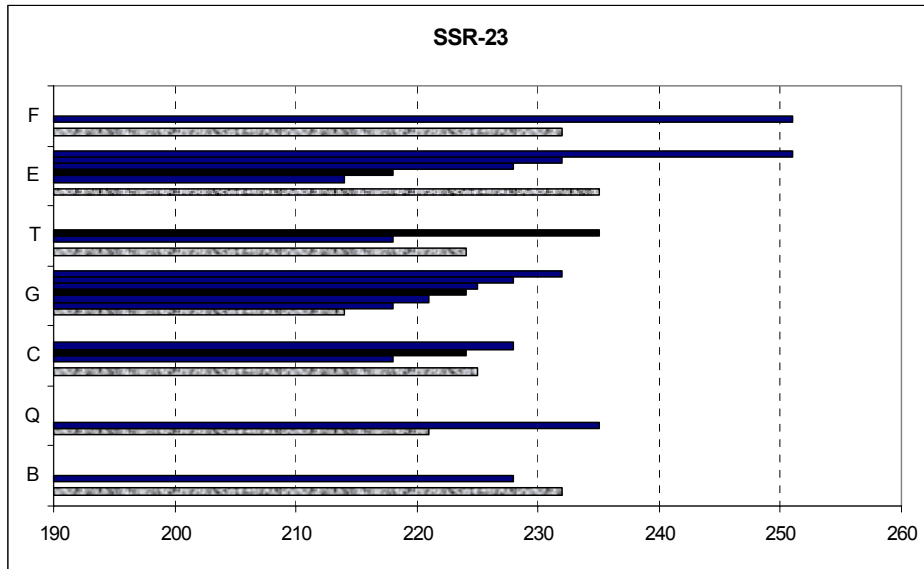


Figure 6. Distribution allélique et fréquences observés pour les locus SSRY-23 et SSR-40. En noir: fréquences 0-0.33, gris: freq. 0.3-0.6 et blanc: fraq. 0.6-1.

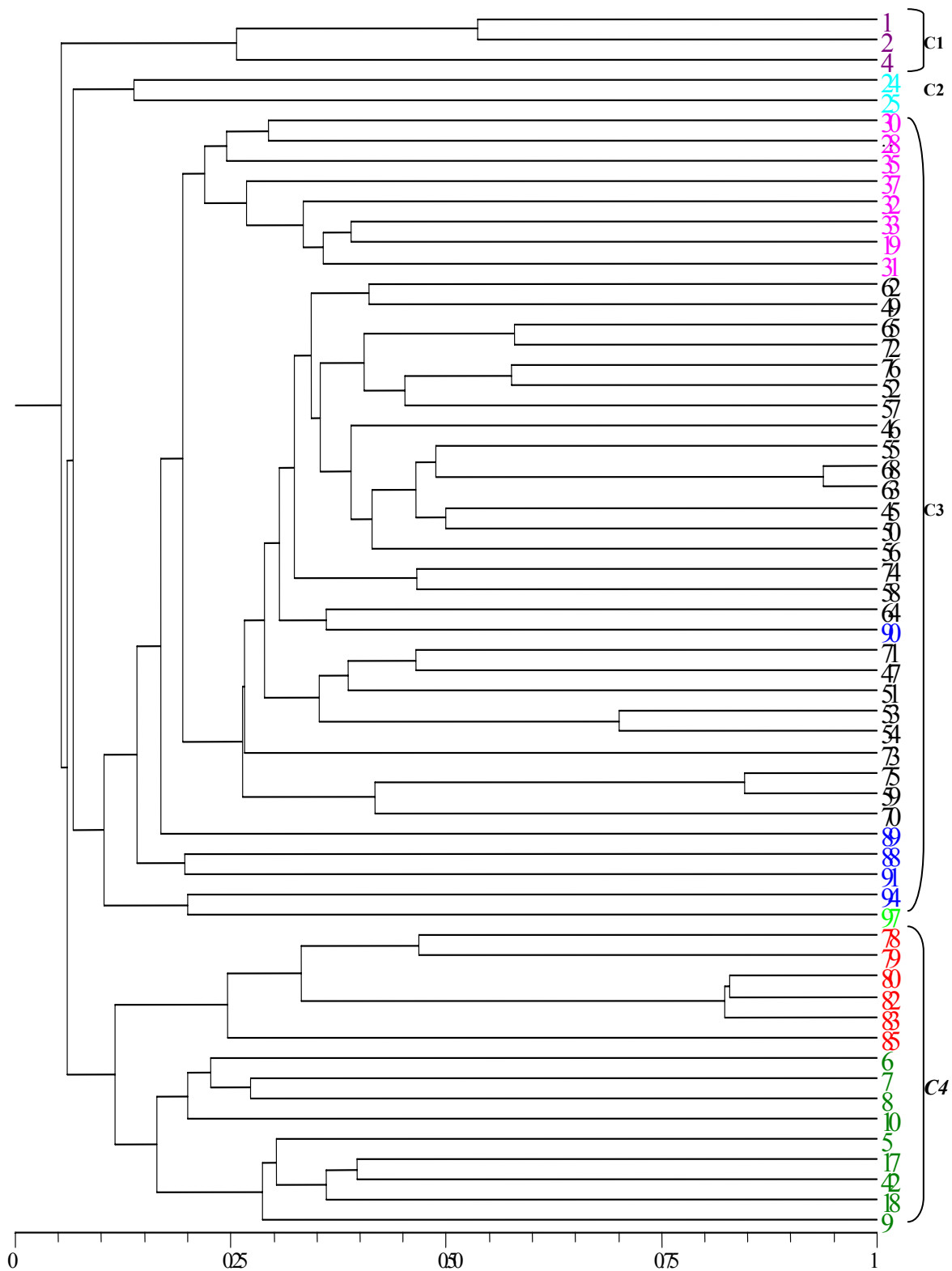


Figure 7. Dendrogramme UPGMA qui montre la diversité moléculaire des 60 accessions à partir de 27 loci SSR. Noir : *M.esculenta*, bleu: *M.flabellifolia*, vert foncé: *M.glaziovii*, rouge: *M. carthaginensis*, rose: hybrides, bleu ciel: *M.quinquepartita*, violet: *M.barchyloba* et vert clair: *M.tristis*.

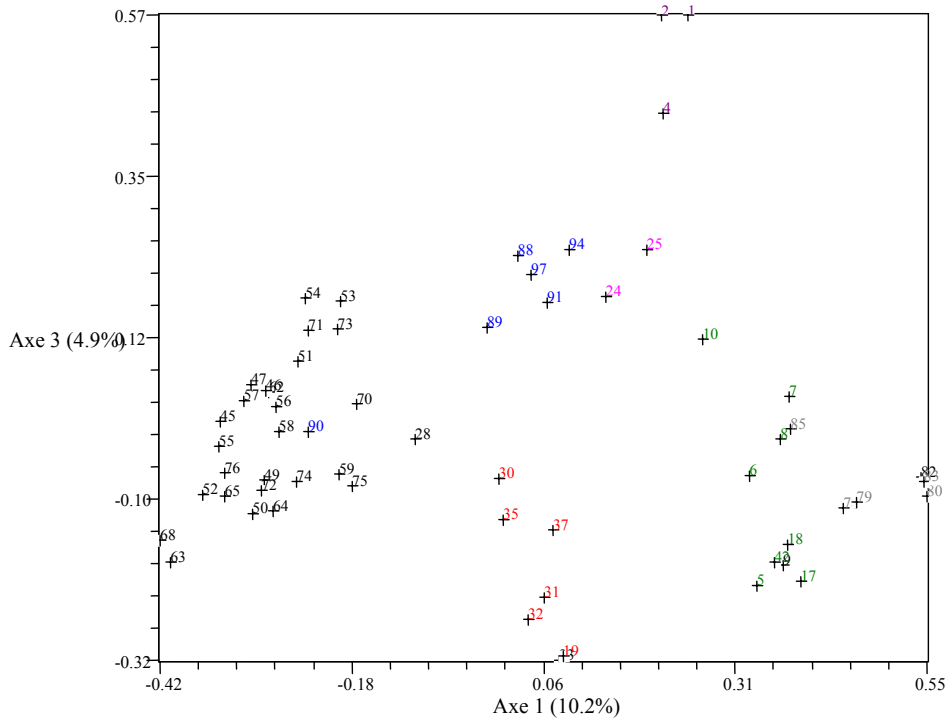


Figure 8. PCO fondée sur 27 loci SSR chez les 60 accessions étudiés en utilisant le coefficient de *Jaccard*.

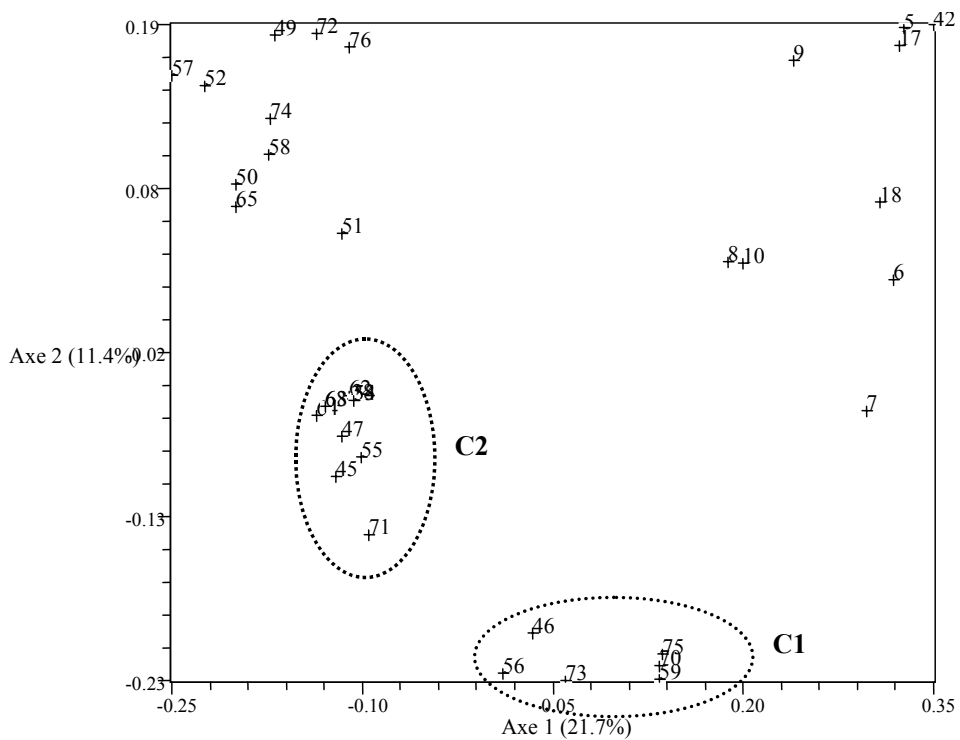


Figure 9. PCO basé sur le SMC qui graphique la distribution de *M. esculenta* et *M. glaziovii* en fonction des 7 loci SSR du groupe de liaison D.

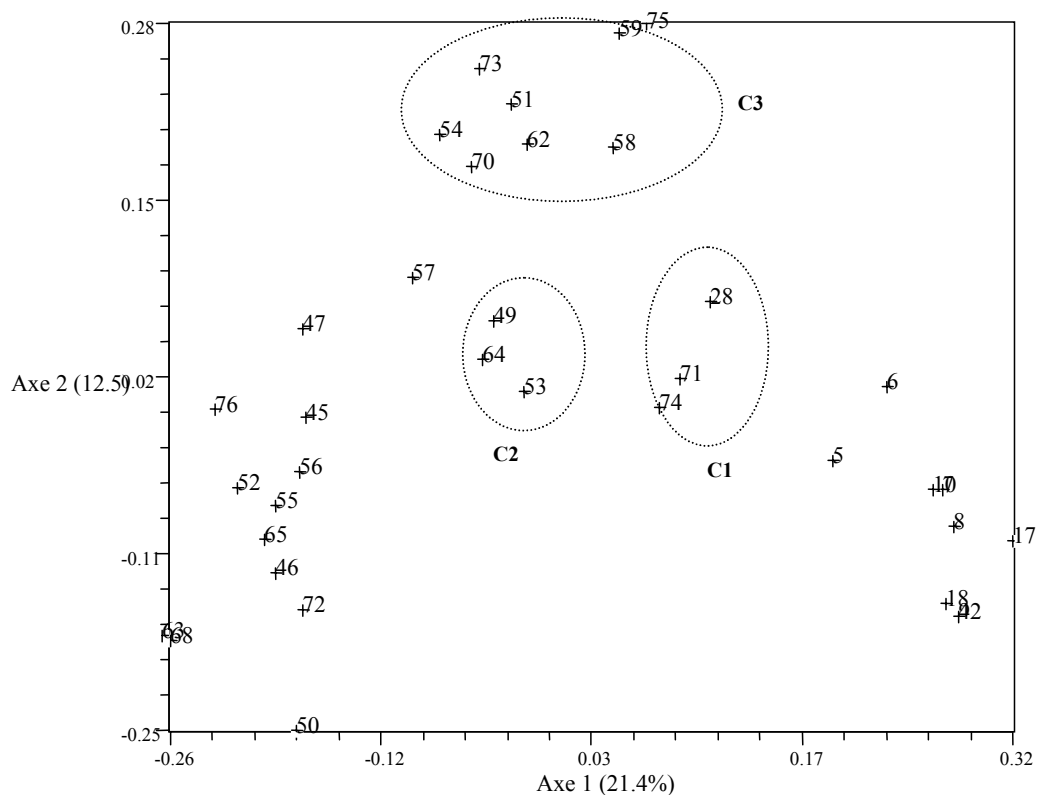


Figure 10. PCO basé sur le SMC qui graphique la distribution de *M.esculenta* et *M.glaziovii* en fonction du 5 loci SSR du groupe de liaison F.

