



HAL
open science

Développement de marqueurs microsatellites nucléaires chez le pin maritime et utilisation dans le cadre de la traçabilité des lots de graines issus des vergers-à-graines

Christophe Plomion

► To cite this version:

Christophe Plomion. Développement de marqueurs microsatellites nucléaires chez le pin maritime et utilisation dans le cadre de la traçabilité des lots de graines issus des vergers-à-graines. [Rapport de recherche] Institut National de la Recherche Agronomique (Inra). 2005. hal-03410193

HAL Id: hal-03410193

<https://hal.inrae.fr/hal-03410193>

Submitted on 31 Oct 2021

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

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

Document rédigé par l'INRA

**DEVELOPPEMENT DE MARQUEURS MICROSATELLITES
NUCLEAIRES CHEZ LE PIN MARITIME ET UTILISATION DANS LE
CADRE DE LA TRAÇABILITE DES LOTS DE GRAINES ISSUS DES
VERGERS-A-GRAINES**

(AFOCEL : contrat C2224, convention DERF N° 61 45 80 15 02)

Projet AFOCEL BE01

**Rapport final
7 décembre 2005**

**Rédigé par
C. Plomion (INRA, Pierroton)
Avec une participation de
Luc Harvengt (AFOCEL, Nangis)**

Coordonné par C. Plomion

FICHE SIGNALÉTIQUE

1 - IDENTIFICATION

THEME : Développement de marqueurs microsatellites nucléaires chez le pin maritime et utilisation dans le cadre de la traçabilité des lots de graines issus des vergers-à-graines

CONVENTION N°:xxxxx

- date de notification :xxxxx
- durée de la convention : **36 mois**
- montant : **60 000 € TTC**

NON DU RESPONSABLE SCIENTIFIQUE : **Christophe Plomion**

Téléphone : **05 57 12 28 38**

Adresse de l'Organisme : **UMR BIOGECO, INRA, 69 route d'Arcachon , 33 610 CESTAS**

COMITE DE PILOTAGE :

Dr. Bernard Héois (CEMAGREF), Section "Arbres forestiers du CTPS"
Dr. Jean-Charles Bastien (INRA) pour le GIS VFA
Patrick Pastuzska (INRA) pour le GIS Pin maritime
Prof. Jean-Michel Favre (UMR IAM, Nancy)

2 – OBJECTIFS VISES

- Les objectifs ont-ils été atteints ? **OUI**
Sinon, pourquoi ?

- D'autres objectifs que ceux visés à l'origine ont-ils été atteints ? **OUI**

utilisation de l'outil moléculaire pour la cartographie comparée des génome des Pinacées (cf thèse de David Chagné)

- Nouvelles pistes à explorer au vu des résultats : **Développement d'outils moléculaires plus précis (microsatellites nucléaires tri- ou tétra-nucléotidiques)**

3 – RESULTATS

- Référence bibliographique du document fourni : **Rapport final (ci-joint). Développement de marqueurs microsatellites nucléaires chez le pin maritime et utilisation dans le cadre de la traçabilité des lots de graines issus des vergers-à-graines**

- Confidentiel **NON**

-Diffusion prévue : nombre d'exemplaires : **10**

cible : **membres du comité exécutif du GIS pin maritime du futur, comité de pilotage, DGFAR**

4 – VALORISATION DES RESULTATS

-PUBLICATIONS

- Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera M-T, Vendramin GG, Garcia V, Frigerio JM, Echt C, Richardson T, Plomion C (2004) Cross species transferability and mapping of genomic and cDNA SSRs in pines. Theor Appl Genet 109: 1204-1214
- Guevara MA, Chagné D, Almeida MH, Byrne M, Collada C, Favre JM, Harvengt L, Jeandroz S, Orazio C, Plomion C, Ramboer A, Rocheta M, Sebastiani F, Soto A, Vendramin GG, Cervera MT (2005) Isolation and characterization of nuclear microsatellite loci in *P. pinaster* Ait. Mol Ecol Notes 5: 57-59
- David Chagne, thèse (01/01-04/04), Université de Nancy I. Développement de marqueurs moléculaires chez le pin maritime (*Pinus pinaster* Ait.) et cartographie génétique comparée chez les conifères. 166p

-Projet d'articles :

sur l'application de l'outil moléculaire au contrôle de pollution des vergers à graines de pin maritime, INRA+AFOCEL

- Dans quelles revues scientifiques ?
TGG
- Dans quelles revues de vulgarisation ?

-PRESENTATION ORALE

1. Chagné D, Chaumeil P, Madur D, Lalanne C, Brown G, Neale D, Echt C, **Plomion C** (2002) Contribution of maritime pine for comparative mapping in conifers. 10th international conference of the Plant and Animal Genome, San Diego, CA, January 13-18. Abstr. P557 (poster)
2. Chagné D, Almeida M-H, Byrne M, Cervera M, Favre J-M, Harvengt L, Oliveira M, Orazio C, **Plomion C**, Ramboer A, Vendramin G (2003) Development of microsatellite markers in maritime pine. BIOFOR conference, Vitoria-Gasteiz, Spain, 11-14 Nov. 2002 (poster)
3. Chagné D, Almeida M-H, Byrne M, Cervera M, Favre J-M, Harvengt L, Oliveira M, Orazio C, **Plomion C**, Ramboer A, Vendramin G (2003) Development of microsatellite markers in maritime pine. 11th international conference of the Plant and Animal Genome, San Diego, CA, January 12-16. Abstr. P555 (poster)
4. Chagné D, Le Dantec L, Ramboer A, Frigerio JM, Pot D, Collada C, Cervera MT, Vendramin GG, **Plomion C** (2004) Development of molecular markers in *Pinus pinaster* based on pine ESTs. 11th international conference of the Plant and Animal Genome, San Diego, CA, January 12-16 (presentation orale W94)
5. Almeida MH, Byrne M, Cervera MT, Chagné D, Collada C, Favre JM, Frigerio JM, Garcia V, Guevara A, Harvengt L, Jeandroz S, Oliveira M, Orazio C, Pastorelli R, **Plomion C**, Ramboer R, Rocheta M, GG Vendramin (2004) Development of microsatellite markers in maritime pine (*Pinus pinaster* AIT.). 11th international conference of the Plant and Animal Genome, San Diego, CA, January 12-16 (poster P619)

- Projet de présentation scientifique (colloque...) :

- Projet de vulgarisation
(formation, journée sur le terrain...etc)

5 – RESUME

CONVENTION N° **xxxxx**

Des marqueurs moléculaires contenant des motifs microsatellites ont été identifiés chez le pin maritime dans le cadre d'un consortium international dans lequel la DGFAR et l'INRA ont joué un rôle essentiel. Une fois ces marqueurs développés, l'INRA et l'AFOCEL, aidé de la DGFAR et sur leurs fonds propres, ont utilisé cet outil moléculaire pour qualifier la pureté génétique de trois vergers-à-graines de pollinisation libre (vergers d'équivalent clone). Cette étude a clairement mis en évidence une forte hétérogénéité dans le succès reproducteur des arbres « plus » (G0) à l'origine des vergers. Ainsi, la composition génétique des vergers (arbres G1) n'est pas proportionnelle à celle des géniteurs sélectionnés (G0), certains contribuant plus que d'autres à la descendance G1. Même si elle reste minime, nous avons également pu identifier des traces de contamination pollinique lors des croisements contrôlés des G0. Par ailleurs, un flux de pollen extérieur aux vergers important (de 75 à 95%) a pu être mis en évidence par l'analyse des descendants G2. Il est très certainement sur-estimé - le système de marquage étant loin d'être parfait - mais il donne une idée de l'ampleur de la pollution pollinique venant de l'extérieur des vergers. Une baisse non négligeable (mais difficilement quantifiable) des gains génétiques espérés sur la base d'une panmixie seule à l'intérieur des vergers est donc attendu. Au pire (100% de pollen extérieur pollinisant le verger), seule la part maternelle du gain (soit 50%) pourra être garanti.

MOTS CLES : Pin maritime, marqueurs moléculaires, microsatellites, vergers à graines, pollution pollinique, test variétal, pureté variétal, gain génétique

6 – RAPPORT FINAL

C Plomion, V Léger, S Gerber pour l'INRA
L Harvengt, JF Trontin, S Quoniou, F Canlet, P Alazard pour l'AFOCEL

Objectif : Ce projet concerne d'une part la mise au point de critères d'identification variétale chez le pin maritime : les marqueurs microsatellites nucléaires, et d'autre part, l'application de cet outil moléculaire à la traçabilité du matériel forestier de reboisement. Il s'agit en particulier de quantifier le taux de pollution dans les vergers à graines de pollinisation libre (vergers d'équivalent clone).

1/ Développement de l'outil moléculaire

Nous avons suivi deux approches complémentaires afin de développer des marqueurs microsatellites (SSR) :

- la transférabilité au pin maritime, d'information d'espèces de pin phylogénétiquement proche (*P. radiata*, *P. taeda*...)
- le développement de SSR directement chez le pin maritime, en partant soit d'une banque génomique enrichie en motifs microsatellites, soit d'une banque d'ADNc classique.

Ces travaux ont été essentiellement menés par l'UMR BIOGECO (CDD de A Ramboer financé par la DERF et la thèse de D Chagné) en interaction forte avec l'INIA de Madrid et le CNR de Florence. Six autres partenaires (fédérés au sein d'un consortium international initié par l'UMR BIOGECO), ont également participé de manière plus ponctuelle. Les détails de ces études ont été publiés dans les deux articles suivants :

- Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera M-T, Vendramin GG, Garcia V, Frigerio JM, Echt C, Richardson T, Plomion C (2004) Cross species transferability and mapping of genomic and cDNA SSRs in pines. *Theor Appl Genet* 109: 1204-1214
- Guevara MA, Chagné D, Almeida MH, Byrne M, Collada C, Favre JM, Harvengt L, Jeandroz S, Orazio C, Plomion C, Ramboer A, Rocheta M, Sebastiani F, Soto A, Vendramin GG, Cervera MT (2005) Isolation and characterization of nuclear microsatellite loci in *P. pinaster* Ait. *Mol Ecol Notes* 5: 57-59

La thèse de David Chagné réalisé dans ce cadre, est également disponible en ligne à l'URL : <http://www.pierroton.inra.fr/biogeco/genetique/theses.html>. Les annexes 1 et 2 résument l'information obtenue à l'issue de ces études :

2. Choix des marqueurs pour la traçabilité

Le choix des marqueurs que nous avons utilisé pour l'application de l'outil moléculaire à la traçabilité du matériel forestier, s'est fait sur la base d'une information complémentaire (localisation chromosomique) obtenue sur le lot de marqueurs développé par Chagné et al. (2004). Cette information est en cours d'acquisition pour le lot développé par Guevara et al. (2005). Dix marqueurs indépendants (idéalement situés sur des chromosomes différents, sinon génétiquement très éloignés), portant plus de 4 allèles et de diversité génétique élevée ($H > 0.5$), ont aussi été sélectionnés pour l'application au verger à graines : RPtEST11, ssrPt_ctg275, ssrPt_ctg4363, NZPR1078, NZPR413, NZPR472, NZPR544, FRPp91, FRPp94, ITPh4516.

3. Mise au point d'une méthode de génotypage à haut débit

Devant les milliers d'analyses nécessaires pour mener à bien l'étude de traçabilité, nous avons mis au point une méthode de génotypage sur séquenceur 96 capillaires (Amersham Biosciences). Un protocole simple, fiable et reproductible a ainsi été établi au sein de l'UMR BIOGECO par V. Léger (CDD financé par un contrat FEDER). Toujours dans un souci d'amélioration de la productivité et d'une diminution des coûts, nous avons réussi à multiplexer 8 de ces 10 locus en 3 jeux indépendants (sur la base d'une homogénéité « intra jeux » des conditions d'amplification PCR). Il en a résulté une capacité de génotypage de 1500 à 2000 données / jour. Le protocole détaillé (MO-MET-005-LAB) est joint en annexe 3. Deux jeux (Jeu 1 : FRPP91, FRPP94, ITPH4516, Jeu 2 : NZPR413, NZPR1078, SSRptctg4363) ont été transféré au partenaire AFOCEL pour l'étude de traçabilité. Le jeu n°3 (NZPR472, NZPR544) n'a été utilisé qu'au sein de l'UMR BIOGECO.

4. Application à la traçabilité

Cette partie est divisée en deux grands chapitres, traitant tour à tour du verger-à-graines de Mimizan analysé par l'INRA puis de ceux de Berdillan et St Augustin analysés par l'AFOCEL.

VERGER DE MIMIZAN

Nous présentons dans un premier temps les grands types de matériel végétal qui ont été récoltés et analysés sur le plan moléculaire, à savoir :

- 1/ des populations naturelles en Aquitaine,
- 2/ les arbres élites (génération G0 de la population d'amélioration) à l'origine du verger-à-graines,
- 3/ les descendants G1 installés sur le verger,
- 4/ la semence G2 (variété améliorée récoltée sur les arbres G1)
- 5/ un échantillon d'arbres entourant le verger.

Dans un second temps, nous poserons diverses hypothèses liées au contrôle variétal et les testerons en utilisant l'information moléculaire dont nous disposons.

4.1. Matériel végétal récolté et données moléculaires

4.1.1. Diversité génétique de la population Atlantique

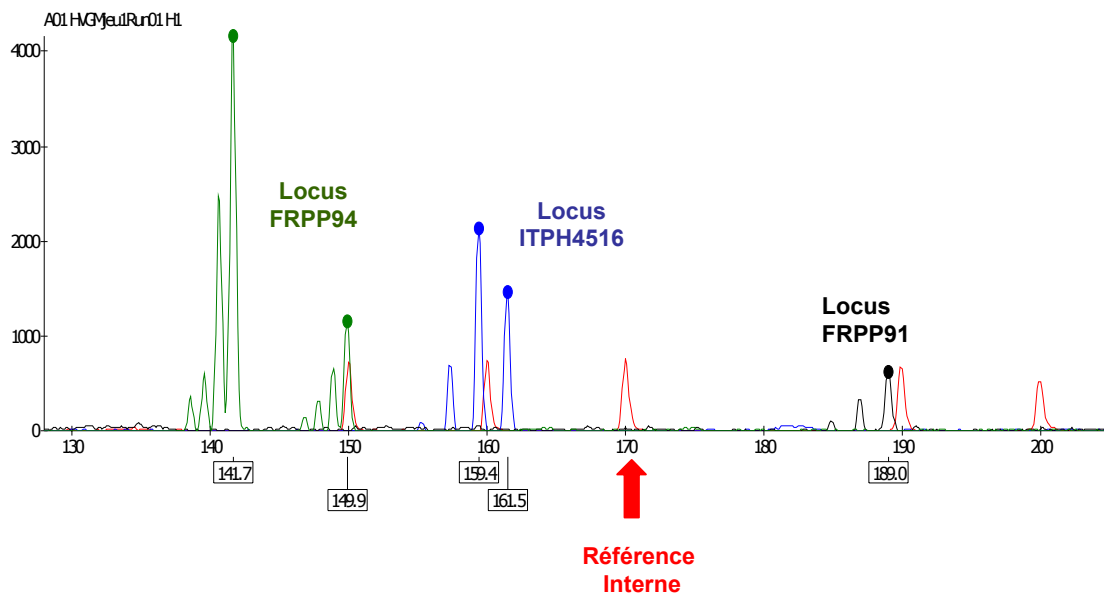
Matériel : 8 populations permettant de décrire le niveau et la structure de la diversité génétique de la population Atlantique dont sont issus les arbres élites servant de géniteurs G0 au Verger d'équivalent clone.

Populations	code	Nombre d'arbres analysés
Petrock	40	30
Mimizan	41	30
Hourtin	42	30
Le Verdon	43	30
Olonne/mer	44	30
St Jean de Monts	45	30
Pleucadec	46	30
Erdeven	47	30

Méthode :

- Récolte d'aiguille au fusil sur 30 arbres par population.
- Extraction d'ADN (protocole, cf annexe 4)
- Génotypage des 3 multiplexes.
- Analyse de la taille des fragments à l'aide du logiciel « Fragment Profiler » en utilisant une référence interne (figure 1). L'analyse des pics d'intérêt est vérifiée et modifiée le cas échéant manuellement.

Figure 1 : Exemple d'un typage en multiplex sur le jeu n°1



Résultats : Fréquences alléliques de la population Atlantique (cf annexe 5)

4.1.2. Empreinte génétique des parents G0 du verger d'équivalent clones

Matériel : récolte manuelle d'aiguilles des 34 arbres élités sur de jeunes clones. Ces individus ont été utilisés comme père et/ou comme mère lors de croisements contrôlés (cf tableau ci-dessous). Ils sont à l'origine de la population G1 du verger de Mimizan.

Code des clones	Clone G0	utilisé comme mère	utilisé comme père
1	0003	x	x
2	0007	x	x
3	0008	non	x
4	0014	x	x
5	0022	x	non
6	0024	x	x
7	0029	x	x

8	0054	x	x
9	0063	x	x
10	0123	x	x
11	0137	x	x
12	0142	x	x
13	0157	x	x
14	0601	x	x
15	1309	x	x
16	1317	x	non
17	1319	x	x
18	1328	x	x
19	1901	x	x
20	3102	x	x
21	3110	x	x
22	3803	x	x
23	3836	x	x
24	3841	x	x
25	3847	x	x
26	4301	x	x
27	4701	x	x
28	NAU3 16T5	x	x
29	NAU3 12T4	x	x
30	LBR1 2T2	x	x
31	CRE1 7T2	x	x
32	CAPT 20T4	x	x
33	CREON 8T2	x	x
34	LEG2 13T1	x	non

Méthode :

- extraction d'ADN des aiguilles
- génotypage des trois multiplex
- Analyse automatique de la taille des fragments
- Vérification manuelle des résultats de l'analyse

Résultat : Fréquences alléliques de la population G0 (Annexe 6)

4.1.3. Récolte et analyse moléculaire des arbres G1 dans les enclos 1 et 2 du verger à graines de Mimizan

Matériel et méthodes :

- Récolte (6 et 13 Oct. 2004) au fusil d'aiguilles pour 200 arbres par enclos (un arbre sur quatre prélevé dans une rangée sur 5) correspondant à une intensité de sélection de 5% environ.
- Extraction de l'ADN des aiguilles
- Génotypage de 96 arbres par enclos sur les trois multiplex
- Analyse automatique de la taille des fragments
- Vérification manuelle des résultats de l'analyse

Résultat : Fréquences alléliques de la population G1 (cf Annexe 7)

Les paramètres de diversité aux 8 locus microsatellites pour la génération G1 (arbres du verger) sont donnés dans le tableau suivant (H0 : hétérozygotie observée, He : hétérozygotie attendue, A : nombre d'allèles observés, Ae : nombre d'allèle efficace, Fis : indice de fixation, Fst : indice de différenciation)

Ces données ont été calculées à l'aide du logiciel GENETIX v.4.03 (Laboratoire Génome, Populations, Interactions -CNRS UMR5000-Montpellier).

Locus	H0	He	A	Ae	Fis
FRPP94	0,4737	0,5319	6	2,1428	+0.1121ns
ITPH4516	0,6684	0,8325	12	6,0487	+0.1996 ***
FRPP91	0,7421	0.8013	11	5,0865	+0.0765 ***
NZPR413	0,3737	0,3684	4	1,5857	-0.0117 ns
NZPR 1078	0,6684	0,6482	3	2,8561	-0.0286 ns
Ssr Pt ctg 4363	0,6947	0,6687	4	3,0343	-0.0363 ns
NZPR472	0,6296	0,6607	6	2,9323	+0.0497 ns
NZPR544	0,2211	0,3874	3	1,6350	+0.4315 ***
Moyenne multi locus	0.5590	0.6124	6.1250	2.5799	0.0871

4.1.4. Récolte et analyse moléculaire de la variété améliorée (semences G2)

Matériel et méthodes :

- Les cônes de 100 arbres G1 répartis sur les enclos 1 et 2 ont été récoltés le 27 janvier 2005. Trois graines par arbre ont été mises à germer en chambre de culture et 75 ont donné au moins un descendant G2 dont 37 provenaient d'une mère G1 génotypée, les autres étant pris au hasard dans des G1 non génotypés (c'est-à-dire sans information du génotype maternel). Cependant nous n'avons pas tenu compte de cette information dans nos tests d'affectation des G2.
- Extraction d'ADN à partir des plantules G2
- Génotypage de 75 plantules G2 pour les 3 multiplex
- Analyse automatique de la taille des fragments
- Vérification manuelle des résultats de l'analyse

Résultat : Fréquences alléliques de la population G2 (cf Annexe 8)

4.1.5. Récolte et analyse moléculaires d'un échantillon d'arbres extérieurs au verger

Cette population sera dénommée population HV pour Hors Verger.

Matériel et méthodes

- Prélèvement (17 Nov. 2004) d'aiguilles (tir au fusil) de 49 arbres autour des enclos 1 et 2 selon le plan de prélèvement fourni en annexe 9.
- Extraction d'ADN des aiguilles
- Génotypage des 49 arbres pour les 3 multiplex

- Analyse automatique de la taille des fragments
- Vérification manuelle des résultats de l'analyse

Résultat : Fréquences alléliques de la population HV (cf Annexe 10)

L'annexe 11 reprend l'ensemble des allèles trouvés aux 8 locus pour l'ensemble des populations.

4.2. Questions de recherche

Ce matériel nous a permis de répondre à trois questions principales :

4.2.1. Les populations G0 et HV sont-elles représentatives de la population Atlantique ?

Pour répondre à cette question, nous avons étudié la différenciation génétique de l'ensemble des 8 populations naturelles, de la population de 34 arbres G0 et de la population de 49 arbres HV. Les indices de différenciation entre populations sont extrêmement faibles (annexe 12). Sur la base de 8 marqueurs SSR on peut donc considérer ces populations comme faisant partie de la même métapopulation (Fst d'environ 0.001). Pour donner plus de poids à cette analyse, nous avons aussi génotypé 4 populations Corses avec 5 locus microsatellites (FRPP91, ITPH4516, FRPP94, NZPR413, NZPR1078). L'annexe 13 récapitule les résultats locus par locus et pour l'ensemble de ces 5 locus. Globalement, le groupe de populations Corses est homogène (Fst d'environ 0.02) et très différencié du groupe Atlantique (Fst d'environ 0.23) (annexe 14).

Décision : sur la base de cette analyse il a été décidé de calculer les fréquences alléliques (nécessaires aux analyses de parenté) sur l'ensemble de ces 8 populations Atlantiques ainsi que de la population HV. Nous avons cependant exclu la population G0 pour ne pas entraîner de biais dans les assignements (cf 4.2.2).

Résultat : Paramètre de diversité aux 8 locus microsatellites pour la métapopulation ainsi constituée (H0 : hétérozygotie observée, He : hétérozygotie attendue, A : nombre d'allèles observés, Ae : nombre d'allèle efficace, Fis : indice de fixation, Fst : indice de différenciation) calculés à l'aide du logiciel GENETIX v.4.03 (Laboratoire Génome, Populations, Interactions -CNRS UMR5000-Montpellier).

locus	H0	He	A	Ae	Fis
FRPP94	0.6115	0.6176	11	2.6151	+0.012
ITPH4516	0.6799	0.8420	13	6.3291	+0.194
FRPP91	0.6245	0.8236	22	5.6689	+0.243
NZPR413	0.3561	0.3507	5	1.5401	-0.014
NZPR 1078	0.6201	0.6629	4	2.9665	+0.066
Ssr Pt ctg 4363	0.6353	0.6178	5	2.6164	-0.026
NZPR472	0.6563	0.6914	7	3.2404	+0.053
NZPR544	0.2795	0.3396	5	1.5142	+0.179
Moyenne	0.5579	0.6182	9.6250	2.6192	+0.0975

multi locus					
-------------	--	--	--	--	--

4.2.2. Succès reproducteur en G1

- Dans la G1 nous n'avons pas détecté de nouvel allèle qui n'existait pas dans la G0, sur la base de notre échantillonnage. Cela ne prouve pas qu'il n'y ai pas eu de pollution contaminant les G1 (étant donné la proximité de la composition allélique de la G1 à la population Aquitaine). D'un autre côté, l'absence d'allèles autres que celle des G0 reste rassurante !
- Par ailleurs, nous ne retrouvons pas l'ensemble des allèles G0 dans la G1 : c'est la cas des allèles 155 au locus ITPH4516 (1 allèle G0), et de l'allèle 160 au locus FRPP94 (1 allèle G0). Notons que ces deux allèles n'existaient que dans un seul G0 à l'état hétérozygote (i.e. à une fréquence de 1/68). S'il y avait eu une transmission équi-proportionnelle des allèles G0 dans la G1, avec un effectif de 200 G1 analysés la probabilité de les détecter restait très faible, 1.5%.
- Maintenant se pose la question plus générale du succès reproducteur en G1 :

Une comparaison des fréquences alléliques entre la population G0 et la population G1 permet de tester l'écart éventuel de la fréquence observée des allèles dans la G1 par rapport à l'attendu G0. Pour cela nous avons utilisé un test du chi-2 :

$$\chi^2 = \sum (N_i - N_{i\text{ref}})^2 / N_{i\text{ref}}$$

avec N_i le nombre d'allèles i à un locus dans l'échantillon G1 et $N_{i\text{ref}}$ le nombre d'allèles i à un locus dans la population de référence G0.

Les résultats de cette analyse sont synthétisés au tableau suivant. Ils montrent que la représentation équi-proportionnelle des allèles n'est pas vérifiée dans la G1 pour l'ensemble des locus, la moitié d'entre eux montrant un fort écart par rapport à l'attendu panmictique. Il s'agit notamment des locus FRPP91 et ITPH4516 pour lesquels un Fis très significatif avait été observé dans la G1. Il sera donc important de prendre en compte la fréquence des allèles dans le verger, plutôt que dans la G0, pour effectuer un contrôle variétal. Cependant il convient de tempérer cette conclusion basée sur un simple calcul statistique ; en effet la comparaison des histogrammes de fréquences alléliques entre G0 et G1 (annexe 15) ne fait pas état d'un écart trop important entre ces deux générations : les allèles les plus fréquents dans la G0 restant majoritaires dans la G1, les allèles mineurs dans la G0 restant également peu représentés dans la G1.

locus	chi-2	ddl	signification
FRPP91	33.77	10	***
ITPH4516	37.48	12	***
FRPP94	30.46	6	***
NZPR1078	2.505	2	ns
NZPR413	5.936	3	ns
ctg4363	21.54	3	***
NZPR472	4.979	5	ns
NZPR544	1.49	2	ns

Fort de ce premier résultat, il nous a semblé important de voir si l'outil SSR permettait de retrouver les couples de parents des G1, partant de l'hypothèse que les ascendants appartiennent forcément à la G0. Pour répondre à cette question nous avons utilisé le

logiciel FAMOZ mis en place par Sophie Gerber au sein de notre unité (Gerber et al. 2003). Cet outil informatique permet de mesurer la capacité d'un système de marquage moléculaire à reconstituer les filiations. Il présente entre autre l'avantage de permettre d'introduire dans les calculs une probabilité d'erreur de génotypage et ainsi d'accepter des relations d'apparentement même si les génotypes ne sont pas 100% compatibles, ce qui est le cas avec le type de marqueurs utilisé (erreur de lecture possible due aux différences de taille parfois mineures entre les allèles).

Avant toute chose, pour avoir une idée de la puissance du test de parenté (donc de la qualité du marquage), nous avons calculé la probabilité d'exclusion (capacité moyenne de n'importe quel système de marquage génétique à exclure une relation donnée entre individus) ainsi que la probabilité d'identité des parents G0 (probabilité que deux individus aient le même génotype multilocus). Le tableau suivant récapitule les probabilités d'exclusion pour chacun des 8 locus. Elles sont de 91.9% pour un seul parent (donc pas très bonne) et de 99.9% pour un couple de parents (donc très bonne, 1 chance sur mille d'attribuer le mauvais couple de G0 à un descendant G1). Pour ce qui est de la probabilité d'identité, il est à noter que les 3 locus (ITPH4516, FRPP91, NZPR472) permettent une identification unique des 34 G0.

Probabilités d'exclusion cumulées (locus classés par exclusion croissante)

Locus n°:	Un seul parent	couples parents
ITPH4516	0.546751	0.876039
FRPP91	0.756496	0.977171
NZPR472	0.814543	0.990496
NZPR1078	0.854675	0.995363
Ctg4363	0.884904	0.997741
FRPP94	0.904153	0.998849
NZPR413	0.912872	0.999264
NZPR544	0.919453	0.999476

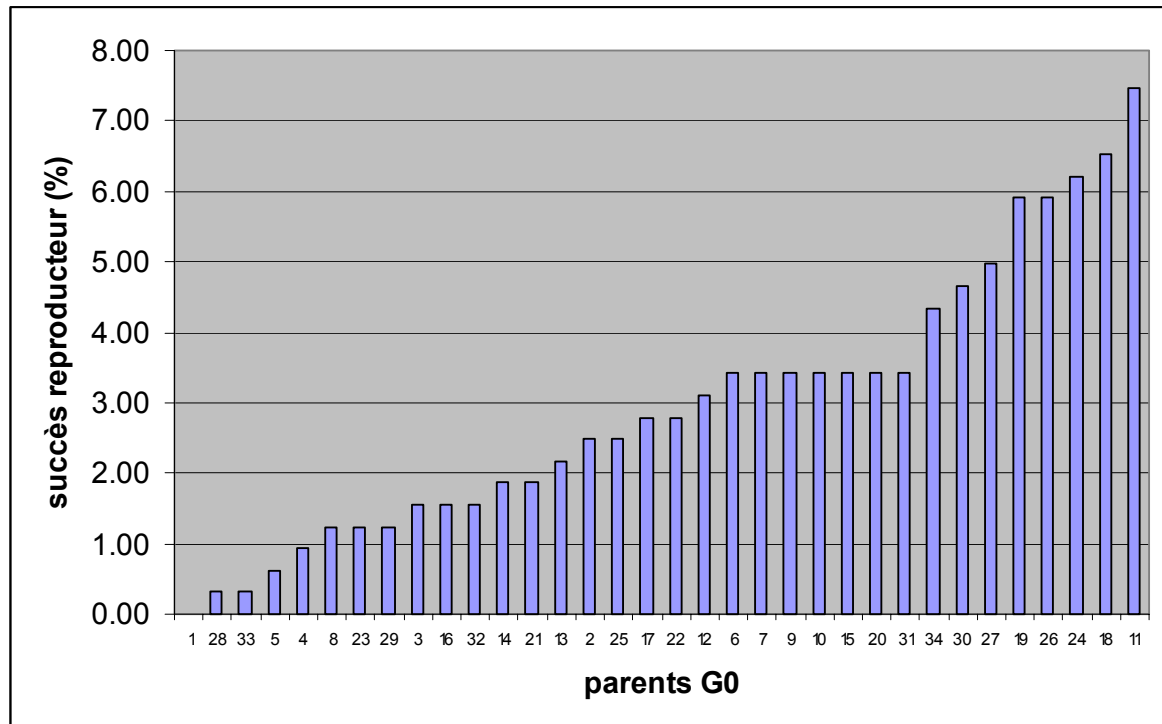
En se basant sur un calcul de LOD score, Nous avons recherché le couple de parent de plus probable des G1. Le LOD score est défini comme le ratio de deux probabilités : la probabilité d'un G0 particulier d'être parent d'un descendant G1, divisée par la probabilité de n'importe quel autre G0 d'être parent de ce même G1. Cette seconde probabilité est calculée en se basant sur les fréquences alléliques de la population G0, et non de la population Atlantique, car nous partons ici de l'hypothèse nulle que les G1 sont forcément issus des G0.

Sur les 190 G1, nous avons recherché le couple de parents le plus probable. Dans 143 cas nous avons pu assigner un couple, dans 3 cas deux couples étaient ex-aequo, les 44 G1 restant n'ayant pas de couple attribué, ce qui représente 23% des cas pour lesquels nous n'avons pas retrouvé les parents, alors qu'ils étaient forcément dans la population G0 ! Ce résultat suggère que le système de marquage n'est pas assez discriminant. Pour quantifier le succès reproducteur des G0 en G1, nous avons comptabilisé les parents G0 au travers des 143 couples précédemment trouvés. Un test du chi-2 à 33 ddl a permis de tester l'équité représentativité des 34 parents dans la G1. Il ressort de cette analyse hautement significative ($\chi^2=122$) que certains parents G0 ont contribué plus que d'autres à la

génération G1 (voir figure 2), certains n'ayant pas du tout contribué. Pour calculer le succès reproducteur (en %) de chaque G0 dans la G1, nous avons tenu compte de la moindre contribution des 4 arbres G0 (qui ne sont intervenus que comme mère ou comme père seulement). La contribution cumulée des G0 montre que 50% des gamètes sont issus de 9 reproducteurs! Une telle distribution en « L inversé » est classiquement reportée dans la littérature sur le succès reproducteur des arbres.

Ce résultat reste vrai si nous incluons dans l'analyse le parent le plus probable pour chacun des 44 G1 pour lesquels aucun couple n'avait été identifié. En effet, sur ces 44 individus, nous avons pu assigner un parent le plus probable pour 36 d'entre eux...ce qui au final nous donne un pourcentage d'assignation de $(144*2+36)/380 = 85\%$.

Figure 2 : Succès reproducteur en G1



Il serait maintenant intéressant d'obtenir cette information sur d'autres tranches du verger pour valider ces différences notables de succès reproducteur.

L'impact de ce résultat sur le gain génétique attendu a été calculé sur la base de cette distribution du succès reproducteur.

1^{ère} conclusion :

Le génotypage de deux cents arbres G1 uniformément répartis sur le VEC de Mimizan a permis d'accéder à la composition génotypique du verger et de tester l'équiproportionnalité de la contribution des parents G0 dans la descendance G1. Sur la base de cet échantillon, nous concluons à une hétérogénéité dans le succès reproducteur des G0, ce qui se traduit par un verger qui ne représente pas exactement la composition génétique des parents G0. Cette étude a été réalisée au stade « arbre G1 » et non « graine G1 », elle rend donc compte non seulement du succès reproducteur mais aussi de la sélection gamétique et sporophytique.

Pour quantifier la qualité de cette analyse, nous avons simulé 10 000 G1 sur la base des génotypes des G0. Dans 88.5% des cas, le logiciel retrouve le bon couple de G0. Ce qui signifie que le bon couple n'est pas le plus vraisemblable dans 11.5% des cas. Nous pouvons donc attribuer une relative confiance (dans 11.5% des cas nous avons attribué un couple qui n'est pas le bon) à nos conclusions sur le succès reproducteur des G0.

Pour améliorer le pouvoir discriminant de l'outil moléculaire, nous proposons de développer dans un futur projet des marqueurs microsatellites qui contiennent des motifs tri- ou tétra-nucléotidiques qui seront moins sujets aux erreurs de génotypage.

4.2.3. Quel est le taux de pollution gamétique en G2 ?

Pour tenter d'estimer le flux de gènes extérieur au verger, nous avons utilisé un test d'assignation des 75 individus G2 à l'une des trois populations de référence suivantes : G1, Atlantique ou HV « hors verger ». Le logiciel GeneClass2 (Piry et al. 2004) qui intègre une adaptation de l'approche Bayésienne de Rannala et Montain (1997) a été utilisé avec les 8 locus précédemment décrits. Deux analyses ont été réalisées: soit le calcul d'un score d'appartenance à telle ou telle population de référence, soit le calcul d'une probabilité (par simulation) qu'un individu appartienne à chacune des populations de référence (Paetkau et al. 2004). Nous présentons ici le résultat de la première approche, qui permet de visualiser le degré d'appartenance soit à la G1 soit aux deux autres populations de référence (figure 3). Sur la base des 3 populations de référence 41 des 75 G2 ont un score d'appartenance plus fort pour la G1.

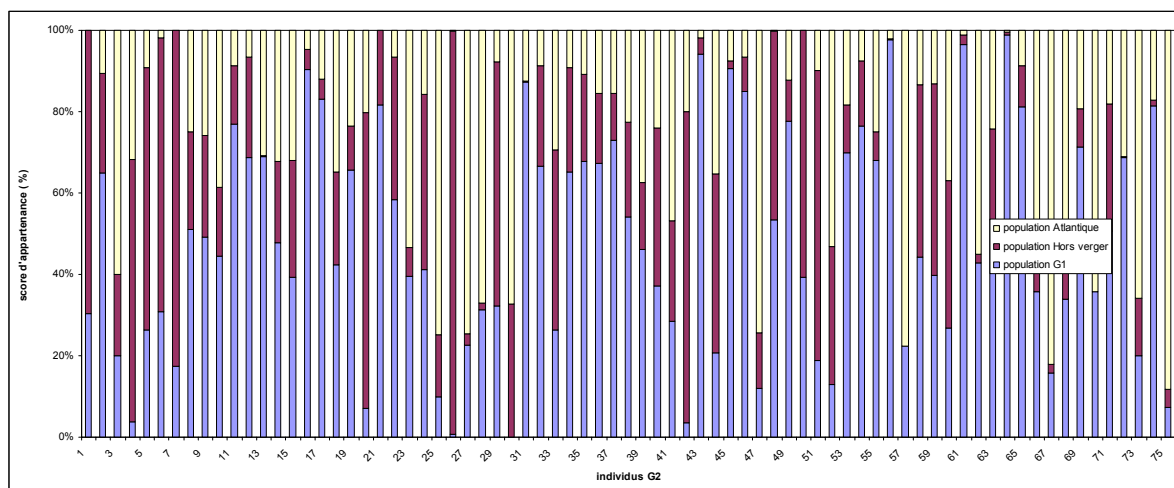


Figure 3 : Taux d'appartenance (score) des individus G2 à l'une des 3 populations de référence.

La distinction entre Atlantique versus HV pouvant être remise en question étant donnée la proximité génétique de ces deux populations, nous avons également utilisé cette approche pour assigner les 75 individus G2 soit à la population G1 soit à la population « Atlantique+HV » (Figure 4). Sur la base des 2 populations de référence 51 des 75 G2 ont un score d'appartenance plus fort pour la G1.

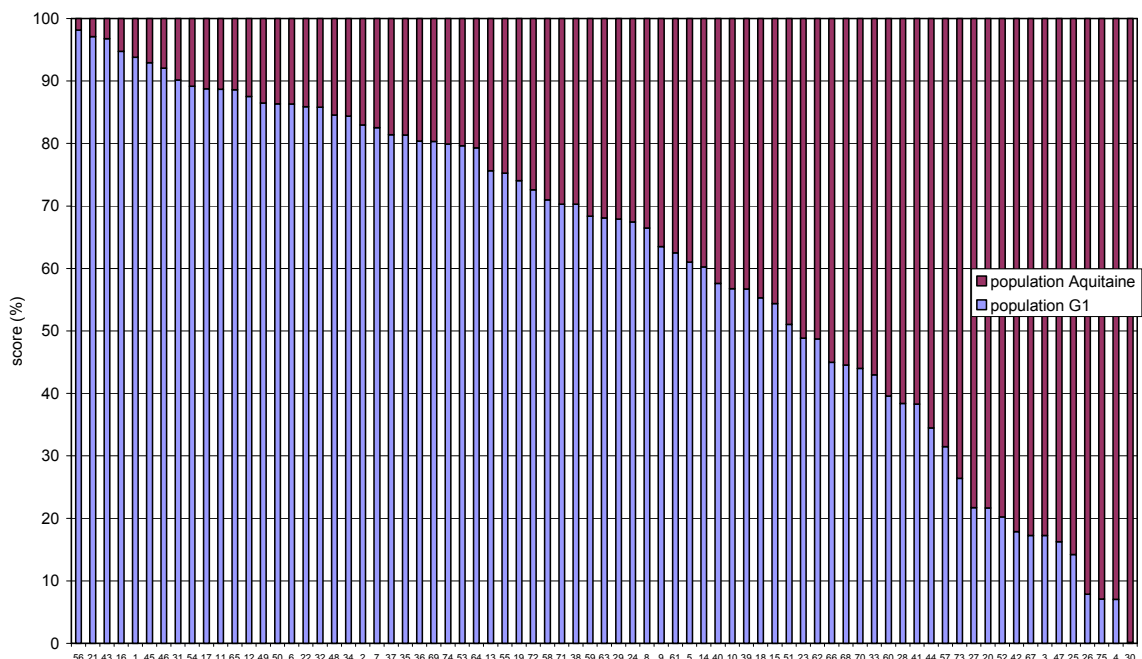


Figure 4 : Taux d'appartenance (score) des individus G2 à l'une des 2 populations de référence.

2^{ème} conclusion :

Le génotypage de 75 arbres G2 uniformément répartis sur 2 tranches du VEC de Mimizan combiné à l'utilisation d'un test d'assignation a permis d'estimer indirectement le flux de gènes extérieur au verger. Sur la base de cet échantillon, nous concluons à un flux génique important puisque 32% des G2 sont plus probablement assignés à un parent « Atlantique » qu'à un parent G1.

VERGERS DE Berdillan et St Augustin

4.1. Matériel végétal récolté et données moléculaires

L'AFOCEL a réalisé l'analyse des variations à 6 locus microsatellites nucléaires de 743 génotypes de Pin Maritime. Ils correspondent aux 34 individus G0 (population initiale) à l'origine des vergers VG2 de Berdillan et St Augustin, 40 individus G1 (population parentale) et 290 individus G2 (population de descendants) échantillonnés dans chacun d'eux ainsi qu'à 49 individus G2 issus d'un lot de graines commercial originaire du verger de Berdillan.

Matériel et méthodes :

- Récolte à l'aide d'une nacelle des aiguilles et des cônes (5-10 par arbre) de 40 arbres par verger en gardant l'identité de l'arbre mère pour chaque cône
- Localisation par GPS des semenciers récoltés (cf cartographie en annexe 16)
- Forçage à l'étuve de l'ouverture des cônes
- Mise en germination de 200 graines par cône.
- Extraction des ADN à partir des aiguilles des semenciers et de 30 semis x 2 cônes x 5 arbres pour chaque verger (l'excédent de matériel végétal a été gardé pour des analyses ultérieures)
- Génotypage à l'aide des deux multiplex recommandés par l'INRA sur séquenceur monocapillaire ABI 310
- Analyse automatique de la taille des fragments et vérification manuelle
- Vérification et analyse des données AFOCEL seules, correction des données en fonction des données de ségrégation
- Second tour d'analyse en incorporant les fréquences alléliques de la population naturelle déterminées par l'INRA (après transposition pour tenir compte des différences de taille attribuées aux mêmes ADN par nos deux systèmes d'analyse)
- Logiciels utilisés : Genescan (Applied Biosystems), GenAEx version 5.1 (Peakall et Smouse, 2001), GenePop version 3.4 (Raymond et Rousset, 1995), FaMoz (Gerber *et al.*, 2003),

Résultat : Fréquences alléliques des 6 locus dans les populations G1 et G2 des deux vergers, détection d'allèles absents chez les G0 et G1 (cf Annexe 17 et 18)

Le détail de ces résultats a été exposé dans le mémoire de S Deбилle (2005) (Quantification du taux de pollution pollinique des vergers à graines de pin maritime par du pollen extérieur par des microsattellites nucléaires - étude préliminaire. Rapport de DESS, 40 pages, Univ Montpellier II- ENSAM).

Contribution des individus G0 à la structure génétique des vergers VG2

L'analyse des 80 arbres G1 génotypés montre deux allèles supplémentaires par rapport aux G0, l'un au locus FRPP91 (allèle « nul », 6 arbres) et l'autre au locus ITPH4516 (1 arbre). Concernant le locus ITPH4516, l'allèle à 163 pb n'est observable que chez un seul individu. Nous pouvons soupçonner un problème de contamination par du pollen extérieur pendant le croisement contrôlé des G0.

La comparaison des distributions alléliques entre les populations G0 et G1 (voir annexe 17) de Berdillan et St Augustin montre que 37 des 41 allèles G0 sont présents chez les G1. Les 4 allèles non représentés, 156 pb au locus FRPP94 et 145 pb, 151 pb et 161 pb au locus ITPH4516 sont très peu fréquents chez les G0 (1 arbre par allèle). Par ailleurs, les allèles les plus fréquents chez les G0 sont généralement les plus fréquents chez les G1. La différence de fréquence observée entre allèles est également assez faible et relativement stable. Enfin, la diversité observée chez les G0 ($H_0 = 0,598$) est très similaire à celle observées chez les G1 ($H_0 = 0,608$ pour Berdillan et $H_0 = 0,679$ pour St Augustin, tableaux en annexe 18).

Polymorphisme des locus microsatellites étudiés

Population parentale G1 (tableaux III et VI)

Le nombre de locus observés (A_0) apparaît plus important pour les locus du jeu 1 (6 à 12) comparé aux locus du jeu 2 (3 à 4). Le nombre d'allèles efficaces (A_e) reflète le polymorphisme observé ; les locus les plus variables étant ITPH4516 ($A_e = 6,9$ à $7,5$) et FRPP91 ($A_e = 5,4$ à $5,6$), trois autres apparaissent intermédiaires (locus FRPP94, NZPR1078, *ssrPtctg4363*, $A_e = 2,3$ à $2,9$) et le dernier faiblement variable (locus NZPR413, $A_e = 1,4$). L'hétérozygotie observée (H_0) révèle un classement similaire des locus avec des valeurs variant de $0,875-0,925$ (ITPH4516) à $0,250-0,325$ (NZPR413). La variation observée dans la population de St Augustin ($H_0 = 0,679$) apparaît légèrement supérieure à celle de Berdillan ($H_0 = 0,608$). Excepté au locus ITPH4516 pour la population de St Augustin, les déviations observées à chaque locus par rapport à l'équilibre de Hardy-Weinberg (F) ne sont pas significatives. Globalement, l'excès d'homozygotes observé dans la population de Berdillan ($F = 0,048$) ou le déficit détecté dans celle de St Augustin ($F = -0,042$) ne sont pas significatifs.

La forte probabilité d'exclusion ($EP = 0,98$) et la faible probabilité d'identité ($IP = 0,000003$) cumulés sur l'ensemble des locus suggéraient que le système de marqueurs utilisé est suffisamment polymorphe pour être exploité dans une recherche de paternité et/ou de parenté. Il est en particulier suffisant pour distinguer tous les individus des populations initiales G0 et G1.

Populations de descendants G2 (tableaux IV et V)

Comme précédemment, les locus ITPH4516 et FRPP91 apparaissent les plus variables ($A_0 = 13$ à 15 , $A_e = 4,8$ à $7,4$ et $H_0 = 0,75$ à $0,90$) devant les locus FRPP94, NZPR1078, *ssrPtctg4363* ($A_0 = 3$ à 8 , $A_e = 2,1$ à $2,9$ et $H_0 = 0,48$ à $0,67$) et NZPR413 ($A_0 = 4$, $A_e = 1,5$ à $1,6$ et $H_0 = 0,32$ à $0,44$). Le léger excès d'homozygotes observé dans les populations de Berdillan ($F = 0,003$) et de St Augustin ($F = 0,007$) est significatif ($p < 0,01$). Il est probablement lié à la présence d'allèles nuls non détectés (confirmé par les données INRA).

4.2. Question de recherche

Evaluation de la pollution pollinique (flux de gènes extérieurs aux vergers)

Parmi les 49 allèles repérés dans les populations G2 analysées de Berdillan et St Augustin (2 x 290 individus), 6 sont absents des populations parentales G1 et G0 (175, 192 et 196 pb au locus FRPP91, 142 pb au locus FRPP94, 172 pb au locus ITPH4516, 95 pb au locus *ssrptctg4363*). Ces allèles étrangers aux vergers démontrent donc l'existence d'un flux de gènes extérieurs. Ils sont portés à l'état hétérozygote par 10 individus G2 de Berdillan et 8 individus G2 de St Augustin. Ce qui correspond respectivement à 1,7 et 1,4 % des individus. L'analyse par simulation présentée ci-après tente d'évaluer le flux total de gènes

extérieurs au verger prenant en compte non seulement les gamètes impliquant des allèles observés dans les vergers mais également les gamètes qui n'impliquent que des allèles présents dans les vergers.

a. Flux apparent minimal

Les résultats des recherches de paternité (effectuées à l'aide du logiciel Famoz) portant sur 10 descendances G2 de chaque verger sont donnés dans les tableaux VII (Berdillan) et VIII (St Augustin). Les 290 arbres G2 échantillonnés dans chaque cas correspondent à des prélèvements de 29 graines pour chacun des 10 arbres maternels G1 sélectionnés parmi les 40 qui ont été génotypés.

En choisissant de prendre les fréquences alléliques observées dans les populations naturelles comme fréquences de référence extérieures aux vergers, le pourcentage de descendants G2 pour lesquels un père a pu être attribué est de 19,0 % (Berdillan ; LOD score seuil de 4,1 ; 79,1 % de classifications exactes des pères les plus probables) ou de 39,6 % (St Augustin ; seuil = 4,1 ; 80,9 % de bonnes classifications).

Le flux de gènes apparent extérieur aux vergers apparaît donc très élevé et au minimum de 81 % dans le cas de Berdillan et de 60,4 % dans le cas de St Augustin.

L'introduction d'un taux d'erreur de typage de 0,1 % dans l'analyse selon les recommandations de Gerber *et al.* (2000) réduit légèrement le résultat (77,2 % pour Berdillan ; 53,8 % pour St Augustin).

Si nous considérons les résultats pour lesquels la recherche de paternité est la plus concluante (fréquences dans les populations naturelles, 0,1 % d'erreur de typage), les pourcentages d'individus G2 pour lesquels un père a effectivement été trouvé parmi les individus G1 génotypés sont très variables d'un arbre mère à l'autre aussi bien dans le cas de Berdillan (3,5 à 58,6 %) que dans le cas de St Augustin (10,3 à 82,8 %).

Les taux observés dans le verger de **Berdillan** sont similaires entre les deux arbres testés de chaque zone sauf dans le cas de la zone 4 (51,7 % de différence). Si l'on exclut ce dernier cas, **il semble que le flux de gènes extérieurs « suive » le sens du vent** (cline de prélèvement, figure 10 du matériel et méthode), la zone 5 étant la plus « polluée » (94,8 %) suivie de la zone 2 et 3 (85,5 % en moyenne) et de la zone 1 (50,0 %).

Des différences « intrazone » importantes sont notées dans les zones 3, 4 et 5 (37,9 à 65,5 % de différence) du **verger de St Augustin. Aucune relation ne peut être mise en évidence** en fonction de l'axe de prélèvement, les zones 4 et 5 situées aux extrémités de l'axe de prélèvement étant les plus « polluées » (63,8 et 56,9 %, respectivement) tandis que les zones 1 et 3 situées dans la partie « centrale » du cline semblent moins touchées (48,3 et 36,2 %, respectivement).

b. Flux caché et déduction du flux total (apparent + caché)

Les flux de gènes apparents extérieurs aux vergers qui viennent d'être estimés sont une évaluation minimale. L'éventuel flux de gènes caché (« cryptique ») correspondant à la proportion de paternité faussement accordée à un arbre des populations G1 parentales alors qu'elle appartient à un arbre des populations extérieures au verger, doit être évaluée afin d'affiner l'estimation.

Le flux de gènes caché a été évalué pour chaque verger (tableaux IX et X) par simulation de 10 000 descendants G2 à partir d'une population théorique de géniteurs (N = 500 ou 1000) comportant les 40 arbres G1 génotypés. Les seuils de décision quant à la signification des LOD scores calculés sont ceux qui ont été déterminés par simulation lors de l'étape précédente de construction du test. Le flux de gènes apparent obtenu lors de ces simulations sous-estime le flux de gènes réel (proche du flux de gènes attendu) d'environ 20 % dans le cas du verger de Berdillan et de 23 à 43 % dans le cas du verger de St Augustin.

Dans les meilleures conditions d'estimation du flux de gènes réel par le flux de gènes apparent, le flux de gènes caché minimal est alors estimé à 19,9 % (Berdillan) ou 27,9 % (St Augustin) du flux de gènes apparent.

Le calcul des flux totaux de gènes extérieurs aux vergers doit tenir compte du flux caché estimé par simulation à l'aide de Famos. Si nous considérons les estimations de flux apparents pour lesquels la recherche de paternité est la plus concluante dans le cas de Berdillan (77,2 %) et de St Augustin (53,8 %), les flux totaux sont alors estimés au minimum à 94,8 % et 74,8 %, respectivement.

Ces chiffres apparaissent 17,8 à 41,2 % plus élevés que ceux rapportés pour un autre verger VG2 étudié à l'aide de 6 microsatellites chloroplastiques (Plomion *et al.*, 2001 ; flux de gènes minimum de 36 %). Si des chiffres aussi élevés ont parfois été obtenus (Pakkanen *et al.*, 2000 : 14 à 74 % chez *Picea abies*), ils apparaissent exagérés.

c. Le niveau élevé des valeurs de pollution pollinique obtenues peut être partiellement expliqué par les faiblesses des marqueurs utilisés

Faible performance du test de paternité

Les simulations montrent que dans nos conditions, s'il n'y avait aucune erreur ni difficulté de typage, les marqueurs utilisés ne permettent de retrouver le père véritable que dans 77,7 % à 81,1 % des cas. Ce taux est relativement faible par rapport aux données de la littérature (par exemple, 95,5 % chez le chêne par Gerber *et al.*, 2000).

Faible niveau de diversité de certains locus

La faiblesse du test de paternité développé peut être expliquée par une diversité génétique faible à certains locus étudiés, en particulier pour le jeu 2. Ainsi, sur les 34 individus G0 et les 80 G1 analysés, tous sauf 2 peuvent être différenciés par le seul jeu 1. Les 2 individus restants sont discriminés par n'importe quel marqueur du jeu 2. Cette faible diversité génétique induit une difficulté à distinguer les individus du verger par rapport aux individus extérieurs.

Présence d'allèles nuls

L'analyse de ségrégation des 10 G1 de chaque verger a montré la présence d'allèles nuls. D'autre part, un déficit d'hétérozygotes a été observé pour les locus du jeu 1. La présence d'allèles nuls non détectés lors de la lecture des profils pourrait être à l'origine de ce déséquilibre. Ces allèles agissent sur l'analyse de parenté, parce qu'ils peuvent

conduire à de fausses exclusions quand l'hétérozygote nul est analysé comme un homozygote non nul (Jones *et al.*, 2003). L'existence d'allèles nuls pour les locus FRPP94 et ITPH4516 a été détectée par Derory *et al.* (2002) et au moins un allèle nul présent au locus FRPP91 a été décelé dans notre travail (confirmé par V. Léger, communication personnelle). Leur présence semble être fréquente chez les conifères (Fisher *et al.*, 1998 ; Varshney *et al.*, 2005).

Erreurs de typage

Les problèmes de typage existent puisque la prise en compte d'un taux d'erreur de 0,1 % entraîne une amélioration des résultats. De plus, lors des comparaisons des données AFOCEL avec les données INRA pour les génotypes des G0, des différences ont été observées pour certains individus (23,5 %). Notre analyse a été effectuée manuellement avec toutes les précautions d'usage mais des problèmes ont été rencontrés notamment pour évaluer la taille des allèles à 1 pb près. Cependant, la correspondance avec les données INRA pour les populations naturelles n'a pas posé de problème (cohérence avec nos données pour les populations G0+G1+G2).

d. Impact sur le gain génétique réel

Le flux de gènes extérieurs élevé entraîne une perte de gain génétique de 50 % au maximum. Le taux de contamination minimum par du pollen extérieur observé pour Berdillan et St Augustin correspond donc à une diminution minimale du gain génétique de 38,6 % et 26,9 % respectivement.

Puisque le gain génétique espéré pour les vergers à graines VG2 est de 30 % (Alazard, 2001), le calcul décrit par Plomion *et al.* (2001) nous permet de situer le gain génétique effectif maximum espéré à 18,4 % pour Berdillan et 21,9 % pour St Augustin ($\Delta g = \Delta G_0(1-C/2)$ où C est le taux de contamination minimum par du pollen extérieur et ΔG_0 est le gain génétique espéré en absence de contaminations).

Notre estimation du flux de gènes extérieurs semble être très élevée par rapport à certaines données de la littérature. Les problèmes répertoriés lors de cette analyse (locus peu polymorphes, erreurs de typage, présence d'allèles nuls, effectif étudié trop faible, etc.) peuvent expliquer ce résultat. Pour améliorer notre étude, nous pouvons essayer de diminuer les erreurs de typage, modifier nos paramètres de simulations et typer des individus supplémentaires. La résolution de ces faiblesses semble un élément crucial à considérer. Cependant, elle risque d'impliquer la mise au point d'un nouveau jeu de locus polymorphes avec un coût de mise développement très important.

Postérieurement à ce travail, nous tenterons de parfaire notre système de marquage en affinant nos conditions d'analyse. Il s'agira notamment d'une étude minimale de la ségrégation des allèles observés afin de préciser leur taille réelle et de caractériser les allèles nuls afin de neutraliser leur impact sur la signification des données de typage. Nous creuserons aussi l'avantage qu'il pourrait y avoir à utiliser les autres microsatellites développés par Guévara et al (2005). Une nouvelle analyse des individus déjà typés ainsi

que l'étoffement du nombre de G2 (matériel végétal disponible) devrait alors permettre d'acquérir rapidement des données plus solides.

3^{ème} conclusion :

Le génotypage de 580 arbres G2 uniformément répartis sur 1 tranche des VEC de Berdillan et Saint-Augustin combiné à l'utilisation d'un test d'assignation a permis d'estimer indirectement le flux de gènes extérieur au verger. Sur la base de cet échantillon, nous concluons à un flux génique apparent extrêmement important (81 % pour Berdillan et 60,4 % pour St Augustin) encore augmenté si on tente de prendre en compte le flux caché.

Ce niveau élevé semble indiquer que le système de marquage moléculaire actuellement disponible est insuffisamment performant pour servir à l'analyse précise du taux de pollution au niveau de lots de graines commerciaux (parents « en mélange »).

Ce manque de performance est reflété par les problèmes répertoriés lors de cette analyse (locus peu polymorphes, erreurs de typage, présence d'allèles nuls, effectif étudié trop faible, etc.).

4.3 Perspective de service en routine aux améliorateurs et exploitants de vergers à graines

Le système de marqueurs utilisé est insuffisamment polymorphe pour être exploité de manière satisfaisante dans une recherche de paternité et/ou de parenté à partir de graines de mère inconnue. Une amélioration est possible par l'abandon des marqueurs moins polymorphes et leur remplacement par d'autres plus performants. Des tests en ce sens sont en cours actuellement à l'AFOCEL. Ils visent par ailleurs également à augmenter l'automatisation des analyses (multiplexage de toutes les PCR en un seul tube et recours à des logiciels améliorés). Le système de marqueurs utilisé pour cette étude est par contre suffisant pour distinguer tous les individus des populations étudiées. Il peut donc d'ors et déjà être utilisé pour vérifier l'identité des boutures et greffes utilisées dans les plantations expérimentales (tests clonaux) et les vergers de clones.

Voici quelques estimations de coûts réalisées sur base du système non optimisé utilisable actuellement en semi-routine. CES COÛTS SONT ASSEZ ELEVES MAIS SONT SUSCEPTIBLES D'ETRE RAPIDEMENT REDUITS D'UN FACTEUR 2 A 5 SUITE AUX OPTIMISATION EN COURS.

Dans le cas où la mère est connue, un effectif de 100 graines serait suffisant. L'analyse suppose alors de faire germer les graines pendant quelques jours avant d'en récupérer l'ADN. Le délai total d'analyse serait d'environ un mois pour un coût d'environ 3000 euros HT.

On peut imaginer utiliser cette dernière stratégie pour **approcher la qualification d'un lot de graines. Il suffirait au moment de la récolte du lot de graines commerciales, de récupérer des échantillons d'aiguilles d'arbres mères assortis de quelques unes de leurs graines. On aurait ainsi des couples semencier-graines permettant de vérifier l'ampleur de la pollution pollinique.** En tablant sur 3 graines par arbre pour 50 semenciers (50 arbres + 150 graines à analyser), on arriverait à un coût de l'ordre de 6500 euros HT pour un délai de réalisation d'environ deux semaines.

Dans le cas d'une demande d'analyse à façon, le délai de réponse serait de l'ordre d'une semaine pour un nombre réduit d'analyses, avec un coût d'environ 100 euros HT par échantillon (résultat brut et interprétation).

ANNEXES

Annexe 1 : Type de marqueurs, localisation sur la carte génétique du pin maritime et niveau de diversité génétique de 19 locus microsatellites développés par Chagné et al. (2004).

Marker type	Locus ID	Mapping pedigree				Genetic diversity	
		INRA-G2	INRA-F2	AFOCEL-F1	INIA-F1	H	A
cDNA-SSR	ssrPp_cn524	6	1	P	M	0.81	5
	ssrPt_ctg275	P/UL	P/UL	P	P	0.74	8
	ssrPt_ctg4363	M	12	P	M	0.68	4
	ssrPt_ctg7824	10	M	M	M	0.35	2
	ssrPt_ctg988	11	M	P	M	0.55	3
	RPtEST11	5	2	P	M	0.74	4
	RPtEST13	10	M	M	M	0.66	3
	ssrPt_ctg1525	M	11	M	M	0.16	2
	ssrPt_ctg64	3	3	M	P	0.68	4
	<i>P. radiata</i> genomic SSR	NZPR1078	2	7	P	M	0.68
NZPR114		M	5	M	P	0.68	5
NZPR1702_b		11	6	P	M	0.38*	2*
NZPR413		4	8	P	P	0.58	4
NZPR472		1	M	P	P	0.67	4
NZPR544		M	3	M	P	0.41	4
<i>P. pinaster</i> and <i>P. halepensis</i> genomic SSR	NZPR823_a	5	M	P	P	0.67	3
	FRPp91	1	9	P	P	0.85	9
	FRPp94	10	5	P	P	0.80	8
	ITPh4516	3	3	P	P	0.84	8

Annexe 2 : 20 locus microsatellites développés par le consortium « microsatellites pin maritime » et niveau de diversité associé (Guevara et al. 2005).

Locus name (Accession no.)	Primer sequences 5' → 3'	PCR protocol	Repeated motif	Clone size (bp) (allele size range)*	No. of alleles	H_O	H_E
A3G05r	F-CATAACCCACCGTAGTAGATGC	b	AC ₁₇	288	6	0.667	0.720
CR377747	R-GGAGATTAAACACGTTAACTGATGC			(282–302)			
A4A09	F-GCCTAACTGTTTATGCAGAATGG	b	AC ₁₁	286	1		
CR377658	R-GGACCAGTTTCCCAATCTAAGG						
A5B07	F-CGTGCAGATCAGAGATTGG	a	AC ₁₉	247	6	0.455	0.773
CR354492	R-CAATTACAACCATCAGAGTGACC			(231–243)			
A5B10	F-TCTTGGCTCTCAGCTCTTGG	a	AC ₁₃	261	6	0.528	0.587
CR354484	R-GAACTTGGGTAGAGCAACTCG			(250–272)			
A5G07	F-CACCGGATAGACTTGTAGG	a	AC ₁₄	205	1		
CR377443	R-CATCATCGCATGTAAAAGATCC						
A6D04	F-ACCATATAGGCTTAGTGGATG	a	AC ₁₄	157	6	0.375	0.797
CR377497	R-TCTTGCCTTGCATCAATGTCC			(139–163)			
A6E08	F-TTGTTTTACAGCAGGTGACTAGG	a	AC ₁₂	159	9	0.722	0.745
CR377513	R-CCAGGTGGATTGTTCTAGGG			(124–152)			
A6F03	F-CCTGAAAATCGACGGATCG	b	AC ₁₇	256	6	0.222	0.537
CR377521	R-ATGGTATTTTGGCGGGTTGC			(242–254)			
A7E07	F-CCTTGGCCCTTAGGAATAGC	b	AC ₁₀	213	1		
CR377571	R-CCGGTTAGTGTAGATTAGATCC						
A5A11	F-ATCACAATTTGGCTTGGTTCC	a	AC ₂₂	299	9	0.750	0.740
CR354475	R-TGCCACAAGTAAGGACAGG			(266–330)			
A5A12	F-TAAGATCACACGCCAGATGC	a	AC ₁₄	239	3	0.333	0.512
CR354476	R-CCCAACCCAACTATCTAAGG			(245–249)			
A5G11	F-CTCAGGCCGTATGAGATTCC	a	AC ₂₂	172	7	0.875	0.807
CR354527	R-TCATCATCACATGAAGAGATCC			(150–178)			
A5B01	F-ACAACGCACACAATAGTCC	a	AC ₁₄	222	8	0.813	0.826
CR354487	R-GTGTGGAACAGTCCGATAGC			(214–240)			
A7E03	F-TAATGTACCCAACCCCTTCG	a	AC ₁₄	117	12	0.739	0.840
CR377565	R-PTGCACCTCATTAGTCAACG			(104–142)			
B3A01r	F-CATGTCCACGAAGAAACAGC	a	AG ₁₈	218	6	0.667	0.556
CR377868	R-TTTTGGCCCTAGAAATGCAACC			(200–218)			
B3D02	F-TTGTAAAGTCGGCTTGTGG	a	AG ₁₄	262	5	0.652	0.664
CR377890	R-CAAAATTAAGGGAGGGAAAAAGG			(266–286)			
A1C03r	F-TGGTAAAGGAAGGAATACAAAAGTG	a	AC ₁₅	172	5	0.647	0.699
CR377594	R-AAGTGGTGCCCTACGATGAG			(180–188)			
B4C06	F-GCAAGAGGACACAAGATGTGG	a	AG ₁₈	295	3	0.706	0.593
CR377922	R-GCAACATGATGACGAGATGC			(272–298)			
B4D05	F-TTTCGGCATCACAAACAGC	a	AG ₂₇	212	4	0.765	0.718
CR377930	R-GTTTGGAAAGCTGGAAGTTGG			(218–234)			
B4E05	F-GGAGGTCTTCTCCTCAITTTGG	a	AG ₁₉	159	4	0.647	0.621
CR377935	R-GCCCTAAAATCTCGCTCTTCC			(160–174)			
B4F04	F-CGAACAAGCATGGAGATCG	a	AG ₂₀	174	4	0.391	0.339
CR377939	R-CCTCTGCTTGGCTTTACTGC			(178–196)			
B4F08	F-GCACTTTGATTTGTTCATCG	a	AG ₁₃	189	7	0.723	0.740
CR377943	R-GTGGCTGATGTCCAAATGC			(179–193)			
B5A02	F-TCTCAGTGGGTATGGGTATGC	a	AG ₁₃	183	1		
CR377945	R-CTTCCACAGGCACTAAAAGG						
B5A03	F-CACTTTCCAGTATGAGAAGGATATG	a	(AG) ₁₁ ATAA (AG) ₂₀	264	14	0.813	0.850
CR377947	R-CCCACATTGTTTCTTTTAAATC			(266–304)			

Annexe 3 : Protocole d'amplification de microsatellites en multiplexe chez le pin maritime et dépôt sur séquenceur MEGABACE.

Principe de la méthode

Amplification simultanée de plusieurs locus microsatellites à l'aide d'amorces fluorescentes permettant un marquage différentiel spécifique de chaque locus .La détection des fragments est obtenue grâce à un séquenceur capillaire. Leur taille peut-être évaluée par comparaison à une référence interne que l'on fait co-migrer.

REALISATION DE LA PCR

Volume réactionnel: 15 µl

- Dans chaque puits de la plaque PCR , déposer un volume d'ADN équivalent à 20ng (ne pas excéder un volume de 5 µl)
- Ajuster à 5µl avec de l'eau milliQ si nécessaire
- Conserver dans la glace le temps nécessaire à la préparation des autres réactifs

Préparation du Mix PCR :

Ce mix peut être préparé à l'avance (hors amorces et Taq polymérase) pour un plus grand nombre d'échantillons et congelé.

- Dans un tube Eppendorff, préparer le mélange de réactifs donné dans le tableau suivant pour le nombre d'échantillons + 10%

Les volumes suivants sont donnés pour 1 échantillon :

Réactifs	Volume (en µl)	concentration
Tampon 10 X	1.5	1X
Mg Cl ₂ (50 mM)	0.6	2 mM
dNTPs (5 mM)	0.6	0.2 mM
Taq polymérase	0.075	5U
Eau milliQ	qsp 10µl	

- Ajouter dans le même tube les volumes d'amorces suivants (les volumes sont identiques pour chaque amorce du couple)

Jeu 1		Jeu 2	
nom	Volume (µl)	nom	Volume (µl)
ITPH4516	0.05	NZPR413	0.10
FRPP91	0.08	NZPR1078	0.05
FRPP94	0.08	Ssr Ptctg4363	0.03
Jeu 3			
NZPR472	0.06		
NZPR544	0.09		

ITPH4516 : marquage FAM de l'amorce F

FRPP94 : marquage HEX de l'amorce F

FRPP91 : marquage NED de l'amorce F

- Distribuer 15µl de ce pool dans chaque puits de la plaque PCR
- Fermer hermétiquement avec un film spécial PCR
- Placer dans l'appareil PCR
- Fermer le couvercle
- Vérifier le programme et le lancer

94°C 4min

94°C 30 sec

Tm°C 1 min

72°C 1 min

30 cycles

(Tm=60°C jeu 1 et 55°C autres jeux)

72°C 15 min

4°C stockage

- Lorsque le programme est achevé, retirer la plaque
- Congeler si nécessaire (si le dépôt ne peut pas être fait le même jour)

PREPARATION DU DEPOT

A faire le jour du dépôt

- les plaques PCR utilisables pour le dépôt des échantillons sont les **plaques PCR AB800** et **uniquement** celles-ci. Le séquenceur ne peut pas accueillir d'autres types de plaques .
- Dans une plaque de type ELISA, diluer les échantillons au 1/5 pour le jeu 1 ou au 1/4 pour les autres jeux.
- Reprendre 1µl de cette dilution dans une plaque AB800
- Préparer un mix avec de l'eau milliQ et le standard de taille (0.20 µl de standard de taille et 10µl d'eau pour un échantillon)
- Répartir 10 µl dans chaque puits
- Couvrir la plaque
- Dénaturer les échantillons à 94°C pendant 5 minutes et conserver obligatoirement dans la glace pour éviter la renaturation. Cette opération peut se faire pendant le pré-run.

Annexe 4 : Protocole d'extraction d'ADN génomique à partir d'aiguilles

Principe de la méthode

Effectuer une lyse des cellules végétales afin de libérer les composants cellulaires. A chaque étape, les protéines sont éliminées par dissolution dans le chloroforme ou par précipitation. L'ADN est purifié par précipitation dans l'isopropanol et par lavages successifs.

Préparation du tampon d'extraction CTAB

Préparer seulement la quantité de tampon nécessaire à l'extraction

CTAB pH 5.5 : (100ml de solution)

1% PVP360 soluble	1g
2% CTAB	2 g
EDTA 20mM	4 ml d'EDTA 0.5M
NaCl 1.4M	28 ml de NaCl 5M
TRIS-HCl 100 mM	10 ml de TRIS-HCl 1M pH8
Eau milliQ	qsp 100 ml
1% β mercapto ethanol	2μl/ml de solution

- Peser le PVP.360 et le CTAB
- Mettre les poudres dans un flacon, puis ajouter un peu d'eau milliQ
- Mélanger en agitant le flacon afin d'obtenir une solution homogène sans agrégats. Essayer d'éviter le plus possible la formation de mousse qui rend la manipulation du produit plus difficile.
- Ajouter les autres liquides et remettre un peu d'eau milliQ
- Placer un barreau aimanté
- Mettre à dissoudre en chauffant sur un agitateur
- Quand les produits sont bien dissous, retirer le barreau magnétique et ajuster le volume avec le restant d'eau milliQ
- Autoclaver la solution si elle n'est pas utilisée immédiatement
- **Ajouter le β mercapto éthanol avant de réaliser l'extraction**

Extraction

En raison de l'utilisation de nombreux produits volatiles et toxiques, il est vivement recommandé de manipuler sous Sorbonne en portant une blouse et des gants.

Avant de commencer l'extraction :

- Préparer une série de tubes numérotés de 2ml et de 1.5ml qui serviront à recueillir les différentes fractions.
- Faire chauffer à l'avance le bain-marie ou le Brunswick à 65°C et y placer le flacon de solution d'extraction.
- Lorsque la solution est bien mélangée, prélever uniquement la quantité de tampon nécessaire à l'extraction du jour (cas où les extractions seraient échelonnées sur plusieurs jours)

Attention : il est impératif d'utiliser des tubes Eppendorff de 2 ml pour le broyage sinon les billes métalliques se coincent au fond du tube.

1^{ère} étape :

- Dans un tube Eppendorff de 2 ml, préalablement identifié au numéro de l'échantillon, placer 2 billes métalliques de taille moyenne.
- Prélever une demi-aiguille de pin (4 cm maxi) et la fractionner à l'aide d'un scalpel ou d'une paire de ciseaux. Placer les morceaux dans le tube Eppendorff et ajouter 300µl de tampon d'extraction.
- Renouveler l'opération pour tous les échantillons en nettoyant les instruments à l'alcool entre chaque nouveau prélèvement
- Placer les tubes dans les portoirs du broyeur en les équilibrant l'un par rapport à l'autre.
- Broyer pendant 5minutes à la vitesse de 30hz
- Sortir les tubes et regarder si les échantillons sont suffisamment broyés, on doit obtenir une poudre fine. Sinon prolonger le broyage.
- Centrifuger brièvement pour faire retomber le matériel végétal au fond du tube : On limitera ainsi les risques de dispersion et de contamination à l'ouverture
- Ajouter 600 µl de tampon d'extraction maxi
- Mélanger
- Placer au bain-marie ou au Brunswick à 65°C pendant au moins une heure
- Sortir les tubes et les laisser refroidir
- Centrifuger brièvement pour faire redescendre le matériel dans le tube.
- Ajouter dans chaque tube, 900µl (1 volume) du mélange 24:1 chloroforme-isoamylalcool
- Mélanger lentement les tubes pendant 10 minutes manuellement ou en utilisant le broyeur (3hz)
- Centrifuger 10 minutes à vitesse maxi (13000t/min)
- Récolter environ 800µl de surnageant dans un nouveau tube Eppendorff de 1.5 ml
- Ajouter 2/3 de volume (600µl) d'isopropanol très froid (conserver le flacon à -20°C) et mélanger manuellement en inversant le tube.
- Placer à -20°C pour 20 minutes minimum, voire pour la nuit.

Ne pas oublier de récupérer les billes métalliques restées dans le culot avant de jeter les tubes usagés de 2ml : Vider le contenu des tubes 2 ml (culot) dans un bécher. A l'aide de la tige aimantée, récupérer les billes. Evacuer les produits dans le bidon des déchets chimiques prévu à cet effet. Nettoyer les billes sous l'eau et les ranger.

2^{ième} étape :

- Centrifuger les tubes à vitesse maximum pendant 5 minutes et vider le surnageant
- Laver le culot d'ADN à l'éthanol 76% : ajouter de l'éthanol dans le tube et laver le culot en retournant le tube plusieurs fois. Eliminer l'éthanol. Aspirer les derniers millilitres à l'aide d'une micro pipette.

Attention : cette étape est délicate car le culot se décolle parfois du tube. Prendre garde à ne pas le perdre.

- Sécher le culot au speed-vac pendant 5 minutes
- Ajouter 600µl de Nucleic Lysis Solution (kit Millipore)
- Incuber 20 minutes au bain-marie à 65°C. Mélanger de temps en temps jusqu'à ce que le culot soit totalement dissout.
- Sortir le tube et le laisser refroidir à température ambiante ou mieux au frigo
- Quand le tube est froid, ajouter 200µl de Protein Precipitation Solution (kit Millipore)
- Mélanger immédiatement et vortexer pendant 15 secondes.
- Centrifuger 5 minutes à vitesse maxi si possible à 4°C
- Prélever 750µl de surnageant dans un nouveau tube
- Ajouter ¾ volume (560µl) d'isopropanol
- Homogénéiser et placer à -20°C pendant 20 minutes minimum (la nuit si possible)
- Centrifuger à nouveau et ôter le surnageant
- Laver le culot à l'éthanol 76%
- Sécher le culot au speed-vac ou le laisser une nuit à température de la pièce
- Suspendre le culot dans 100µl d'eau.
- On peut également traiter le culot à la RNase A (10 µg/ml) en incubant à 37°C pendant 1 heure minimum.
- Quand l'ADN est bien dissout, on peut vérifier sa qualité sur gel d'agarose et le doser au spectrophotomètre.
- Stocker les échantillons au frigo pour une utilisation immédiate et au congélateur (-20°C) pour un stockage long.

Annexe 5 : Fréquences alléliques de la population Atlantique (cf annexe 5)

Jeu n°1

Locus	allèles	Fréquences alléliques
FRPP94	134	0.051
	135	0.004
	142	0.379
	146	0.002
	148	0.004
	150	0.489
	152	0.013
	154	0.007
	156	0.002
	158	0.042
ITPH4516	160	0.007
	153	0.217
	155	0.022
	157	0.002
	160	0.250
	162	0.115
	164	0.057
	166	0.030
	168	0.017
	170	0.050
172	0.148	
174	0.074	
176	0.013	
178	0.004	

FRPP91	169	0.011
	172	0.004
	174	0.002
	175	0.004
	176	0.009
	177	0.235
	178	0.007
	179	0.015
	181	0.009
	186	0.009
	188	0.013
	189	0.209
	190	0.011
	191	0.026
	192	0.002
	193	0.009
	194	0.178
	196	0.037
	198	0.202
	200	0.007
202	0.002	

Jeu n°2

Locus	allèles	Fréquences alléliques
NZPR413	163	0.127
	165	0.011
	171	0.818
	173	0.002
	195	0.042
NZPR 1078	331	0.316
	333	0.284
	335	0.002
	341	0.398
Ssr Pt ctg 4363	94	0.128
	96	0.447
	98	0.007
	100	0.406
	102	0.011

Jeu n°3

Locus	allèles	Fréquences alléliques
NZPR472	308	0,070
	310	0,111
	312	0,002
	316	0,435
	318	0,029
	320	0,339
	322	0,014
NZPR544	244	0.002
	246	0.002
	248	0.774
	252	0.214
	254	0.007

Annexe 6 : Fréquences alléliques des G0

<i>Locus</i>	<i>allèles</i>	<i>Fréquences alléliques</i>
FRPP94	136	0,029
	142	0,250
	150	0,618
	152	0,029
	154	0,029
	158	0,029
	160	0,015
ITPH4516	147	0,015
	149	0,015
	153	0,147
	155	0,015
	157	0,015
	160	0,235
	162	0,118
164	0,044	

	166	0,015
	170	0,059
	172	0,176
	174	0,132
	178	0,015
FRPP91	177	0,265
	181	0,015
	186	0,015
	188	0,088
	189	0,191
	190	0,015
	191	0,015
	193	0,015
	194	0,191
	198	0,176
	202	0,015
NZPR413	162	0,176
	164	0,015
	170	0,735
	194	0,074
NZPR 1078	331	0,338
	333	0,397
	341	0,265
Ssr Pt ctg 4363	94	0,147
	96	0,471
	100	0,353
	102	0,029
NZPR472	308	0,106
	310	0,106
	316	0,500
	320	0,273
	322	0,015
NZPR544	244	0,032
	248	0,774
	252	0,226

Annexe 7 : Fréquences alléliques des 8 locus dans le Verger (pop G1)

<i>locus</i>	<i>allèles</i>	<i>Fréquences alléliques</i>
FRPP94	136	0,042
	142	0,205
	150	0,647
	152	0,008
	154	0,037
	158	0,061
ITPH4516	147	0,003
	149	0,005
	153	0,213
	157	0,005

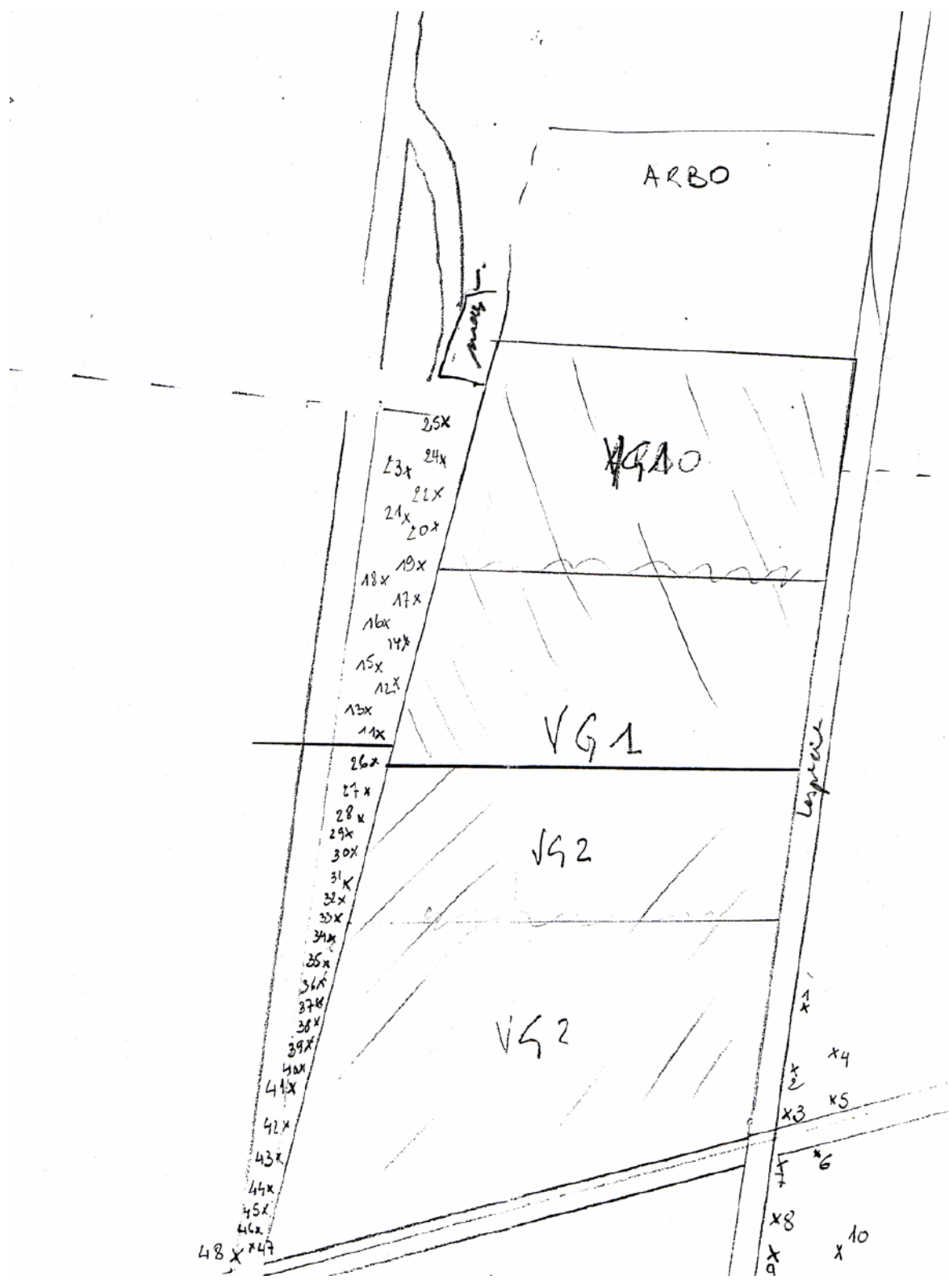
	160	0,258
	162	0,089
	164	0,037
	166	0,016
	170	0,084
	172	0,168
	174	0,100
	178	0,021
FRPP91	177	0,237
	181	0,011
	186	0,008
	188	0,051
	189	0,282
	190	0,024
	191	0,024
	193	0,003
	194	0,186
	198	0,165
	202	0,011
NZPR413	162	0,158
	164	0,003
	170	0,776
	194	0,063
NZPR 1078	331	0,353
	333	0,418
	341	0,229
Ssr Pt ctg 4363	94	0,230
	96	0,405
	100	0,336
	102	0,029
NZPR472	308	0,082
	310	0,127
	316	0,505
	318	0,024
	320	0,246
	322	0,016
NZPR544	244	0,024
	248	0,742
	252	0,234

Annexe 8 : Fréquences alléliques des 8 locus (pop G2)

<i>locus</i>	<i>allèles</i>	<i>Fréquences alléliques</i>
FRPP94	136	0,039
	140	0,007
	142	0,250
	146	0,007
	148	0,007
	150	0,579
	152	0,020
	154	0,039
	158	0,046
	160	0,007
ITPH4516	153	0,218
	157	0,014
	160	0,232
	162	0,155
	164	0,049
	166	0,070
	170	0,007
	172	0,183
	174	0,063
178	0,007	
FRPP91	177	0,296
	180	0,007
	181	0,020
	186	0,007
	188	0,059
	189	0,224
	190	0,013
	191	0,026
	193	0,033
	194	0,178
	196	0,007
	198	0,118
	202	0,013
NZPR413	162	0,151
	170	0,808
	194	0,041
NZPR 1078	331	0,268
	333	0,428
	341	0,304
Ssr Pt ctg 4363	92	0,000
	94	0,230
	96	0,459
	100	0,291
	102	0,020
NZPR472	308	0,075
	310	0,110
	316	0,575
	318	0,014
	320	0,219
	322	0,007
NZPR544	244	0,007
	248	0,807

	252	0,187
--	-----	-------

Annexe 9 : Plan de prélèvement autour du VEC de Mimizan



ARBO

V9A0

VG1

VG2

VG2

Loprow

- 25x
- 23x 24x
- 21x
- 24x 20x
- 18x 19x
- 17x
- 16x 14x
- 15x 12x
- 13x 11x
- 26x
- 27x
- 28x
- 29x
- 30x
- 31x
- 32x
- 33x
- 34x
- 35x
- 36x
- 37x
- 38x
- 39x
- 40x
- 41x
- 42x
- 43x
- 44x
- 45x
- 46x
- 48x x 47

- 1x
- x4
- x3 x5
- x6
- x8
- x9 x10

Annexe 10 : Fréquences alléliques des 8 locus (pop HV)

<i>locus</i>	<i>allèles</i>	<i>Fréquences alléliques</i>
FRPP94	136	0,092
	142	0,367
	150	0,459
	152	0,010
	154	0,010
	158	0,061
ITPH4516	153	0,214
	155	0,020
	157	0,020
	160	0,276
	162	0,122
	166	0,041
	168	0,041
	170	0,082
	172	0,112
	174	0,061
178	0,010	
FRPP91	177	0,276
	180	0,010
	186	0,010
	188	0,051
	189	0,306
	190	0,020
	193	0,020
	194	0,173
	196	0,041
	198	0,082
	200	0,010
NZPR413	162	0,214
	164	0,020
	170	0,704
	194	0,061
NZPR 1078	331	0,306
	333	0,337
	341	0,357
Ssr Pt ctg 4363	92	0,010
	94	0,133
	96	0,490
	100	0,367
NZPR472	308	0,106
	310	0,181
	316	0,394
	318	0,043
	320	0,255
	322	0,021
NZPR544	244	0,010
	248	0,844
	252	0,146

Annexe 11 : Fréquences alléliques pour chaque locus et pour l'ensemble des populations.

<i>Fréquences alléliques</i>						
<i>Locus</i>	<i>Allèles</i>	<i>Aquitaine</i>	<i>G0</i>	<i>G1</i>	<i>G2</i>	<i>HV</i>
FRPP94	134	0,051	0,000	0,000	0,000	0,000
	136	0,004	0,029	0,042	0,039	0,092
	140	0,000	0,000	0,000	0,007	0,000
	142	0,379	0,250	0,205	0,250	0,367
	146	0,002	0,000	0,000	0,007	0,000
	148	0,004	0,000	0,000	0,007	0,000
	150	0,489	0,618	0,647	0,579	0,459
	152	0,013	0,029	0,008	0,020	0,010
	154	0,007	0,029	0,037	0,039	0,010
	156	0,002	0,000	0,000	0,000	0,000
	158	0,042	0,029	0,061	0,046	0,061
	160	0,007	0,015	0,015	0,000	0,007
ITPH4516	147	0,000	0,015	0,003	0,000	0,000
	149	0,000	0,015	0,005	0,000	0,000
	153	0,217	0,147	0,213	0,218	0,214
	155	0,022	0,015	0,000	0,000	0,020
	157	0,002	0,015	0,005	0,014	0,020
	160	0,250	0,235	0,258	0,232	0,276
	162	0,115	0,118	0,089	0,155	0,122
	164	0,057	0,044	0,037	0,049	0,000
	166	0,030	0,015	0,016	0,070	0,041
	168	0,017	0,000	0,000	0,000	0,041
	170	0,050	0,059	0,084	0,007	0,082
	172	0,148	0,176	0,168	0,183	0,112
	174	0,074	0,132	0,100	0,063	0,061
	176	0,013	0,000	0,000	0,000	0,000
178	0,004	0,015	0,021	0,007	0,010	
FRPP91	169	0,011	0,000	0,000	0,000	0,000
	172	0,004	0,000	0,000	0,000	0,000
	174	0,002	0,000	0,000	0,000	0,000
	175	0,004	0,000	0,000	0,000	0,000
	176	0,009	0,000	0,000	0,000	0,000
	177	0,235	0,265	0,237	0,296	0,276
	178	0,007	0,000	0,000	0,000	0,000
	179	0,015	0,000	0,000	0,000	0,000
	180	0,000	0,000	0,000	0,007	0,010
	181	0,009	0,015	0,011	0,020	0,000
	186	0,009	0,015	0,008	0,007	0,010
	188	0,013	0,088	0,051	0,059	0,051
	189	0,209	0,191	0,282	0,224	0,306
	190	0,011	0,015	0,024	0,013	0,020
	191	0,026	0,015	0,024	0,026	0,000
	192	0,002	0,000	0,000	0,000	0,000
	193	0,009	0,015	0,003	0,033	0,020
	194	0,178	0,191	0,186	0,178	0,173
196	0,037	0,000	0,000	0,007	0,041	
198	0,202	0,176	0,165	0,118	0,082	
200	0,007	0,000	0,000	0,000	0,010	
202	0,002	0,015	0,011	0,013	0,000	
NZPR413	162	0,127	0,176	0,158	0,151	0,214
	164	0,011	0,015	0,003	0,000	0,020

	170	0,818	0,735	0,776	0,808	0,704
	172	0,002	0,000	0,000	0,000	0,000
	194	0,042	0,074	0,063	0,041	0,061
NZPR1078	331	0,316	0,338	0,353	0,268	0,306
	333	0,284	0,397	0,418	0,428	0,337
	335	0,002	0,000	0,000	0,000	0,000
	341	0,398	0,265	0,229	0,304	0,357
ssr Ptcg4363	92	0,000	0,000	0,000	0,000	0,010
	94	0,128	0,147	0,230	0,230	0,133
	96	0,447	0,471	0,405	0,459	0,490
	98	0,007	0,000	0,000	0,000	0,000
	100	0,406	0,353	0,336	0,291	0,367
	102	0,011	0,029	0,029	0,020	0,000
NZPR472	308	0,069	0,103	0,082	0,075	0,106
	310	0,110	0,103	0,127	0,110	0,181
	312	0,002	0,000	0,000	0,000	0,000
	316	0,435	0,485	0,505	0,575	0,394
	318	0,029	0,000	0,024	0,014	0,043
	320	0,340	0,265	0,246	0,219	0,255
	322	0,014	0,015	0,016	0,007	0,021
NZPR544	244	0,002	0,032	0,024	0,007	0,010
	246	0,002	0,000	0,000	0,000	0,000
	248	0,774	0,774	0,742	0,807	0,844
	252	0,214	0,226	0,234	0,187	0,146
	254	0,007	0,000	0,000	0,000	0,000

Annexe 12 : Fst (estimation par la méthode Weir et Cockerham 1984) entre les 8 populations Aquitaines, la population G0 et la population HV, calculés sur la base de 8 marqueurs SSR.

Number of samples detected: 10
 Number of loci detected: 8

Code for pop names:

```

-----
 1  40
 2  41
 3  42
 4  43
 5  44
 6  45
 7  46
 8  47
 9   1
10   2
-----
  
```

Estimates for each locus:

```

-----
NZPR413:
-----
pop  1  2  3  4  5  6  7  8  9
2 -0.0033
3  0.0055 -0.0169
4 -0.0142 -0.0089 -0.0003
5 -0.0153 -0.0054  0.0004 -0.0121
6 -0.0112 -0.0087 -0.0027 -0.0099 -0.0054
7 -0.0158  0.0086  0.0199 -0.0086 -0.0108 -0.0028
8 -0.0037 -0.0061 -0.0016 -0.0083 -0.0108  0.0058  0.0026
9 -0.0128  0.0047  0.0137 -0.0091 -0.0142  0.0027 -0.0118 -0.0091
10 -0.0006  0.0312  0.0442  0.0066  0.0006  0.0220 -0.0053  0.0096 -0.0073
  
```

```

-----
NZPR1078:
-----
pop  1  2  3  4  5  6  7  8  9
2  0.0052
3  0.0026  0.0465
4 -0.0128 -0.0097  0.0130
5  0.0236 -0.0095  0.0444 -0.0019
6 -0.0058 -0.0038  0.0022 -0.0142 -0.0081
7  0.0148  0.0126  0.0063  0.0007 -0.0039 -0.0130
8 -0.0045  0.0244 -0.0160 -0.0009  0.0214 -0.0098 -0.0047
9 -0.0156  0.0229  0.0036 -0.0008  0.0426  0.0058  0.0255  0.0003
10 -0.0101  0.0124 -0.0053 -0.0073  0.0180 -0.0095  0.0002 -0.0113 -0.0049
  
```

```

-----
FRPP91:
-----
pop  1  2  3  4  5  6  7  8  9
2  0.0129
3  0.0049  0.0126
4  0.0085  0.0094 -0.0039
  
```

5 0.0029 0.0028 0.0241 0.0127
6 0.0012 0.0084 0.0069 0.0006 0.0043
7 0.0056 0.0163 0.0103 -0.0046 0.0090 -0.0008
8 0.0006 -0.0007 -0.0010 -0.0064 0.0019 -0.0048 -0.0053
9 0.0344 0.0327 0.0357 0.0350 0.0381 0.0260 0.0302 0.0246
10 0.0546 0.0534 0.0495 0.0431 0.0557 0.0252 0.0388 0.0391 0.0014

FRPP94:

pop 1 2 3 4 5 6 7 8 9
2 -0.0052
3 -0.0130 0.0027
4 -0.0091 -0.0121 -0.0043
5 -0.0078 -0.0002 -0.0146 -0.0024
6 -0.0034 0.0008 0.0033 -0.0121 0.0094
7 -0.0120 0.0022 -0.0139 -0.0094 -0.0066 -0.0073
8 -0.0075 0.0102 -0.0157 -0.0001 -0.0095 0.0050 -0.0120
9 0.0196 0.0297 0.0326 0.0092 0.0457 -0.0050 0.0113 0.0314
10 -0.0090 -0.0014 -0.0089 -0.0052 -0.0086 0.0003 -0.0058 -0.0051 0.0267

ITPH4516:

pop 1 2 3 4 5 6 7 8 9
2 -0.0021
3 0.0036 -0.0193
4 -0.0021 -0.0091 -0.0041
5 0.0302 0.0139 0.0128 0.0218
6 0.0127 -0.0035 -0.0071 0.0056 0.0004
7 0.0158 -0.0079 -0.0112 0.0071 -0.0012 -0.0083
8 -0.0055 -0.0016 0.0048 -0.0015 0.0159 -0.0026 0.0130
9 -0.0029 -0.0116 -0.0073 -0.0124 0.0202 -0.0036 0.0057 -0.0111
10 -0.0029 -0.0089 -0.0080 -0.0025 0.0031 -0.0006 -0.0025 -0.0010 -0.0035

NZPR472:

pop 1 2 3 4 5 6 7 8 9
2 -0.0150
3 -0.0093 -0.0132
4 0.0071 0.0033 -0.0089
5 -0.0173 -0.0129 -0.0092 0.0055
6 -0.0115 -0.0069 -0.0090 -0.0128 -0.0112
7 -0.0116 -0.0070 -0.0096 -0.0089 -0.0140 -0.0163
8 0.0030 -0.0010 -0.0131 -0.0138 -0.0043 -0.0075 -0.0124
9 -0.0139 -0.0099 -0.0076 0.0154 -0.0133 -0.0025 -0.0049 0.0051
10 0.0017 -0.0023 -0.0081 0.0133 -0.0036 0.0102 -0.0004 -0.0047 -0.0017

NZPR544:

pop 1 2 3 4 5 6 7 8 9
2 -0.0165
3 -0.0211 -0.0194
4 -0.0239 -0.0215 -0.0244
5 -0.0068 -0.0173 -0.0116 -0.0122
6 0.0306 -0.0009 0.0206 0.0209 0.0012
7 0.0364 0.0033 0.0260 0.0266 0.0049 -0.0247
8 -0.0012 -0.0193 -0.0070 -0.0086 -0.0142 -0.0176 -0.0149

9 -0.0149 -0.0212 -0.0193 -0.0195 -0.0161 0.0016 0.0057 -0.0166
10 0.0359 0.0059 0.0263 0.0277 0.0069 -0.0163 -0.0155 -0.0102 0.0073

Ptctg4363:

pop 1 2 3 4 5 6 7 8 9
2 -0.0139
3 0.0445 0.0147
4 0.0003 -0.0129 0.0125
5 -0.0086 -0.0226 0.0033 -0.0183
6 -0.0097 -0.0164 0.0331 0.0044 -0.0087
7 -0.0119 -0.0162 0.0384 0.0004 -0.0096 -0.0168
8 0.0005 -0.0147 0.0012 -0.0144 -0.0210 0.0033 0.0019
9 -0.0131 -0.0196 0.0272 -0.0060 -0.0152 -0.0163 -0.0162 -0.0054
10 -0.0108 -0.0138 0.0329 -0.0027 -0.0100 -0.0090 -0.0104 -0.0031 -0.0114

Estimates for all loci:

pop 1 2 3 4 5 6 7 8 9
2 -0.0031
3 0.0031 0.0035
4 -0.0034 -0.0059 -0.0018
5 0.0030 -0.0046 0.0090 0.0022
6 -0.0000 -0.0030 0.0052 -0.0026 -0.0019
7 0.0027 0.0015 0.0061 -0.0006 -0.0033 -0.0101
8 -0.0022 0.0005 -0.0056 -0.0063 -0.0001 -0.0033 -0.0028
9 0.0004 0.0055 0.0112 0.0043 0.0153 0.0021 0.0079 0.0043
10 0.0078 0.0101 0.0135 0.0093 0.0103 0.0042 0.0031 0.0037 0.0005

Annexe 13 : Fst (estimation par la méthode Weir et Cockerham 1984) entre les 14 populations sur la base de 5 marqueurs SSR.

Number of samples detected: 14
 Number of loci detected: 5

Code for pop names:

```

-----
1  2
2  10
3  11
4  15
5  40
6  41
7  42
8  43
9  44
10 45
11 46
12 47
13  3
14  4
-----
    
```

Estimates for each locus:

NZPR413:

```

-----
pop 1  2  3  4  5  6  7  8  9  10  11  12  13
2  0.0522
3  0.0556 0.0192
4  0.0160 0.1022 0.0366
5  0.2907 0.4033 0.2703 0.1485
6  0.3147 0.4328 0.2965 0.1704 -0.0033
7  0.3325 0.4460 0.3061 0.1865 0.0055 -0.0169
8  0.2806 0.4010 0.2702 0.1397 -0.0142 -0.0089 -0.0003
9  0.2699 0.3774 0.2411 0.1285 -0.0153 -0.0054 0.0004 -0.0121
10 0.3398 0.4571 0.3248 0.1948 -0.0112 -0.0087 -0.0027 -0.0099 -0.0054
11 0.2840 0.3947 0.2641 0.1433 -0.0158 0.0086 0.0199 -0.0086 -0.0108 -0.0028
12 0.2280 0.3397 0.2059 0.0939 -0.0037 -0.0061 -0.0016 -0.0083 -0.0108 0.0058 0.0026
13 0.2440 0.3503 0.2181 0.1051 -0.0128 0.0047 0.0137 -0.0091 -0.0142 0.0027 -0.0118 -0.0091
14 0.2338 0.3362 0.2123 0.0989 -0.0006 0.0312 0.0442 0.0066 0.0006 0.0220 -0.0053 0.0096 -0.0073
    
```

NZPR1078:

```

-----
pop 1  2  3  4  5  6  7  8  9  10  11  12  13
2  0.0307
3 -0.0184 0.0059
4 -0.0118 0.0631 0.0032
5  0.0845 0.1111 0.0682 0.1581
6  0.0685 0.1497 0.0702 0.1279 0.0052
7  0.0462 0.0312 0.0228 0.1103 0.0026 0.0465
8  0.0567 0.1083 0.0497 0.1204 -0.0128 -0.0097 0.0130
9  0.0226 0.1134 0.0309 0.0659 0.0236 -0.0095 0.0444 -0.0019
10 0.0255 0.0727 0.0195 0.0815 -0.0058 -0.0038 0.0022 -0.0142 -0.0081
    
```

11 0.0063 0.0519 0.0024 0.0527 0.0148 0.0126 0.0063 0.0007 -0.0039 -0.0130
12 0.0348 0.0430 0.0170 0.0947 -0.0045 0.0244 -0.0160 -0.0009 0.0214 -0.0098 -0.0047
13 0.0972 0.1101 0.0774 0.1718 -0.0156 0.0229 0.0036 -0.0008 0.0426 0.0058 0.0255 0.0003
14 0.0503 0.0715 0.0361 0.1132 -0.0101 0.0124 -0.0053 -0.0073 0.0180 -0.0095 0.0002 -0.0113 -0.0049

FRPP91:

pop 1 2 3 4 5 6 7 8 9 10 11 12 13
2 -0.0156
3 0.0136 -0.0020
4 -0.0142 -0.0086 0.0173
5 0.3391 0.3711 0.4278 0.3110
6 0.3702 0.3984 0.4548 0.3377 0.0380
7 0.3601 0.3898 0.4454 0.3319 -0.0138 0.0323
8 0.3271 0.3556 0.4162 0.3002 -0.0016 0.0329 -0.0052
9 0.3243 0.3535 0.4119 0.2946 -0.0007 0.0179 0.0168 0.0070
10 0.3479 0.3746 0.4328 0.3188 0.0094 0.0366 0.0168 0.0012 0.0038
11 0.3309 0.3570 0.4150 0.3031 0.0232 0.0626 0.0224 -0.0030 0.0247 -0.0057
12 0.3261 0.3536 0.4131 0.2963 0.0055 0.0091 0.0017 -0.0121 0.0001 -0.0093 -0.0061
13 0.3276 0.3534 0.4078 0.2986 0.0024 0.0062 0.0030 -0.0035 0.0004 -0.0038 0.0051 -0.0137
14 0.3383 0.3627 0.4125 0.3119 0.0165 0.0443 0.0346 0.0136 0.0087 -0.0054 0.0088 0.0080 0.0014

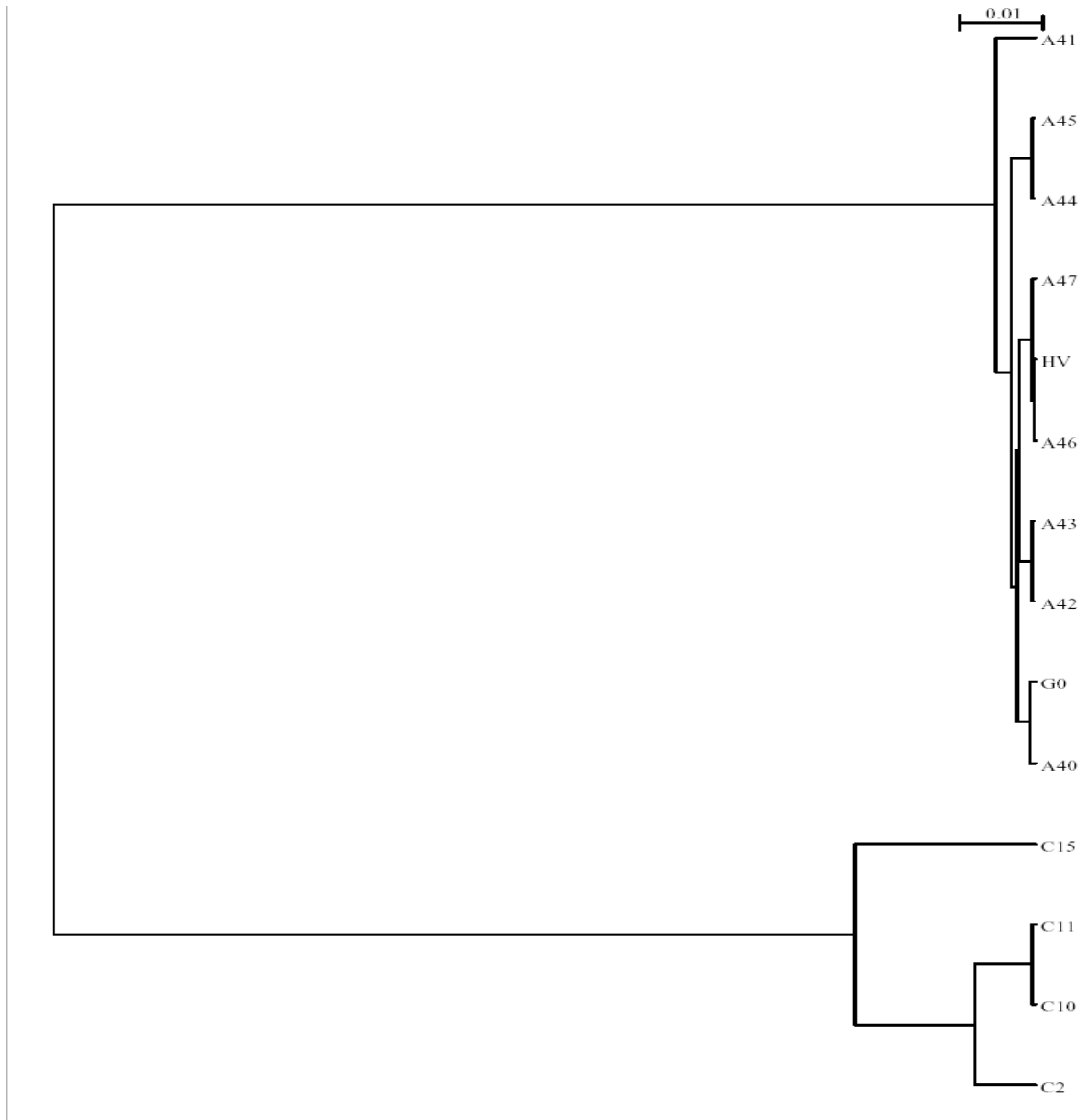
FRPP94:

pop 1 2 3 4 5 6 7 8 9 10 11 12 13
2 -0.0144
3 -0.0131 -0.0143
4 0.0699 0.0355 0.0467
5 0.1433 0.1093 0.0712 0.0967
6 0.1458 0.1101 0.0787 0.1012 -0.0052
7 0.1827 0.1420 0.1036 0.1130 -0.0130 0.0027
8 0.1537 0.1162 0.0831 0.1172 -0.0091 -0.0121 -0.0043
9 0.1862 0.1433 0.1097 0.1044 -0.0078 -0.0002 -0.0146 -0.0024
10 0.1436 0.1082 0.0763 0.1231 -0.0034 0.0008 0.0033 -0.0121 0.0094
11 0.1749 0.1372 0.0969 0.1309 -0.0120 0.0022 -0.0139 -0.0094 -0.0066 -0.0073
12 0.1965 0.1550 0.1161 0.1288 -0.0075 0.0102 -0.0157 -0.0001 -0.0095 0.0050 -0.0120
13 0.1600 0.1318 0.0915 0.1780 0.0196 0.0297 0.0326 0.0092 0.0457 -0.0050 0.0113 0.0314
14 0.1509 0.1120 0.0822 0.0886 -0.0090 -0.0014 -0.0089 -0.0052 -0.0086 0.0003 -0.0058 -0.0051 0.0267

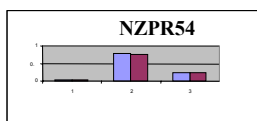
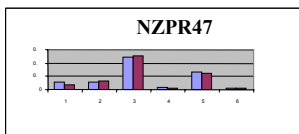
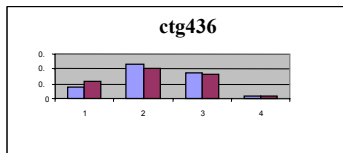
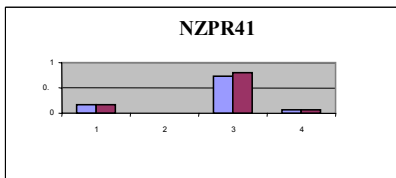
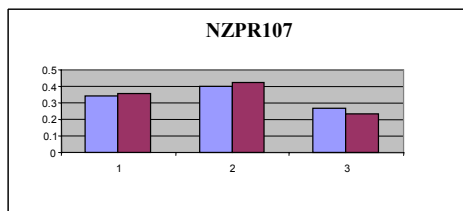
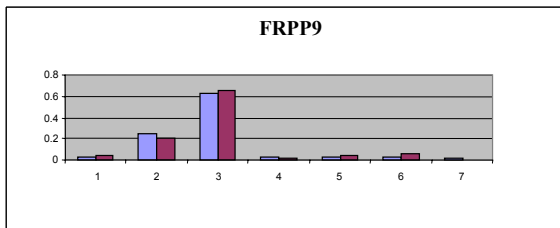
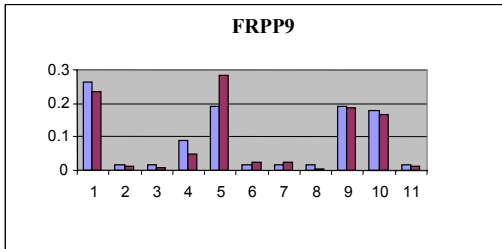
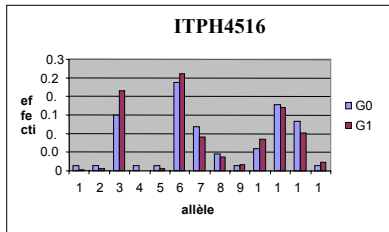
ITPH4516:

pop 1 2 3 4 5 6 7 8 9 10 11 12 13
2 0.0127
3 0.0386 -0.0048
4 0.0323 0.0949 0.1187
5 0.2280 0.3205 0.3546 0.1603
6 0.2761 0.3569 0.3859 0.1967 -0.0021
7 0.2815 0.3648 0.3937 0.2019 0.0036 -0.0193
8 0.2346 0.3166 0.3437 0.1582 -0.0021 -0.0091 -0.0041
9 0.2687 0.3398 0.3701 0.1922 0.0302 0.0139 0.0128 0.0218
10 0.2534 0.3344 0.3637 0.1704 0.0127 -0.0035 -0.0071 0.0056 0.0004
11 0.2694 0.3462 0.3755 0.2030 0.0158 -0.0079 -0.0112 0.0071 -0.0012 -0.0083
12 0.2471 0.3323 0.3617 0.1651 -0.0055 -0.0016 0.0048 -0.0015 0.0159 -0.0026 0.0130
13 0.2467 0.3284 0.3536 0.1635 -0.0029 -0.0116 -0.0073 -0.0124 0.0202 -0.0036 0.0057 -0.0111
14 0.2339 0.3065 0.3304 0.1619 -0.0029 -0.0089 -0.0080 -0.0025 0.0031 -0.0006 -0.0025 -0.0010 -
0.0035

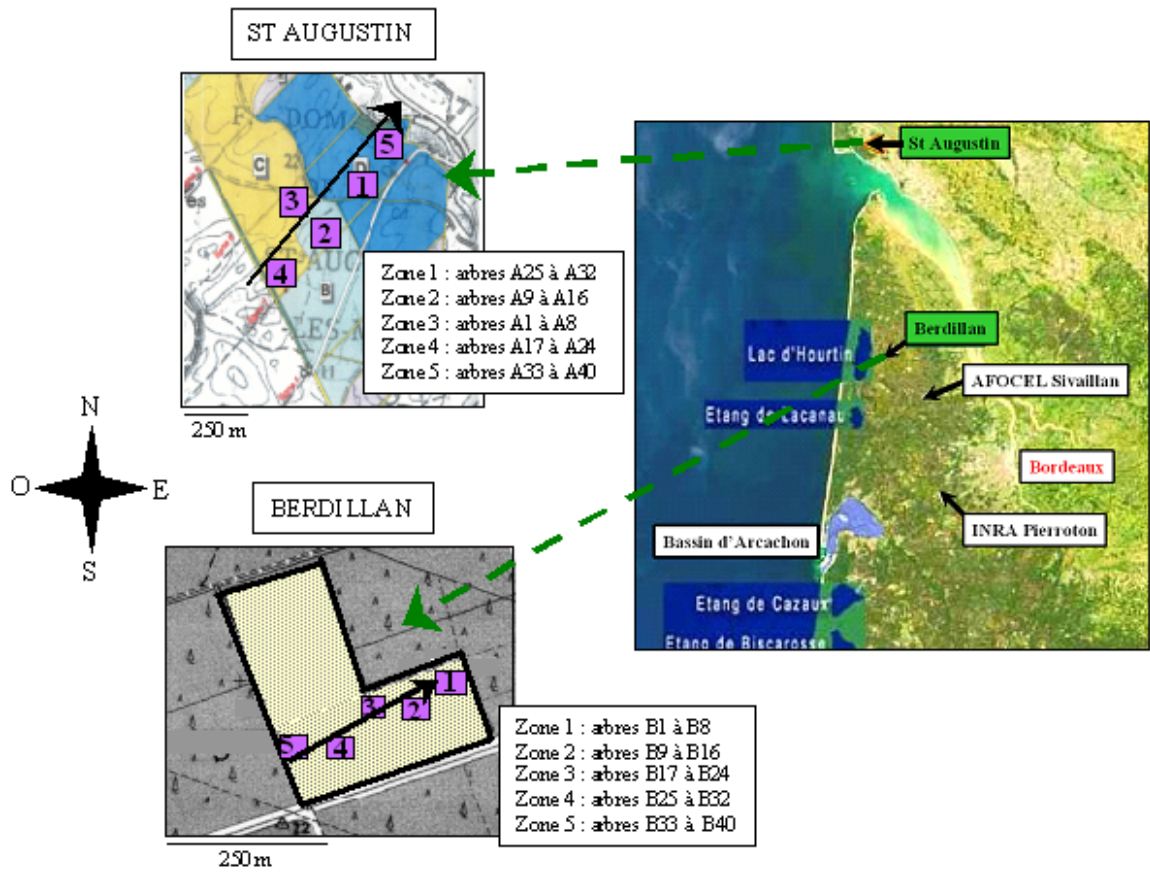
Annexe 14 : Dendrogramme des populations Aquitaines (A), Corses (C), G0 et HV.



Annexe 15 : Histogrammes des fréquences alléliques dans la G0 et la G1



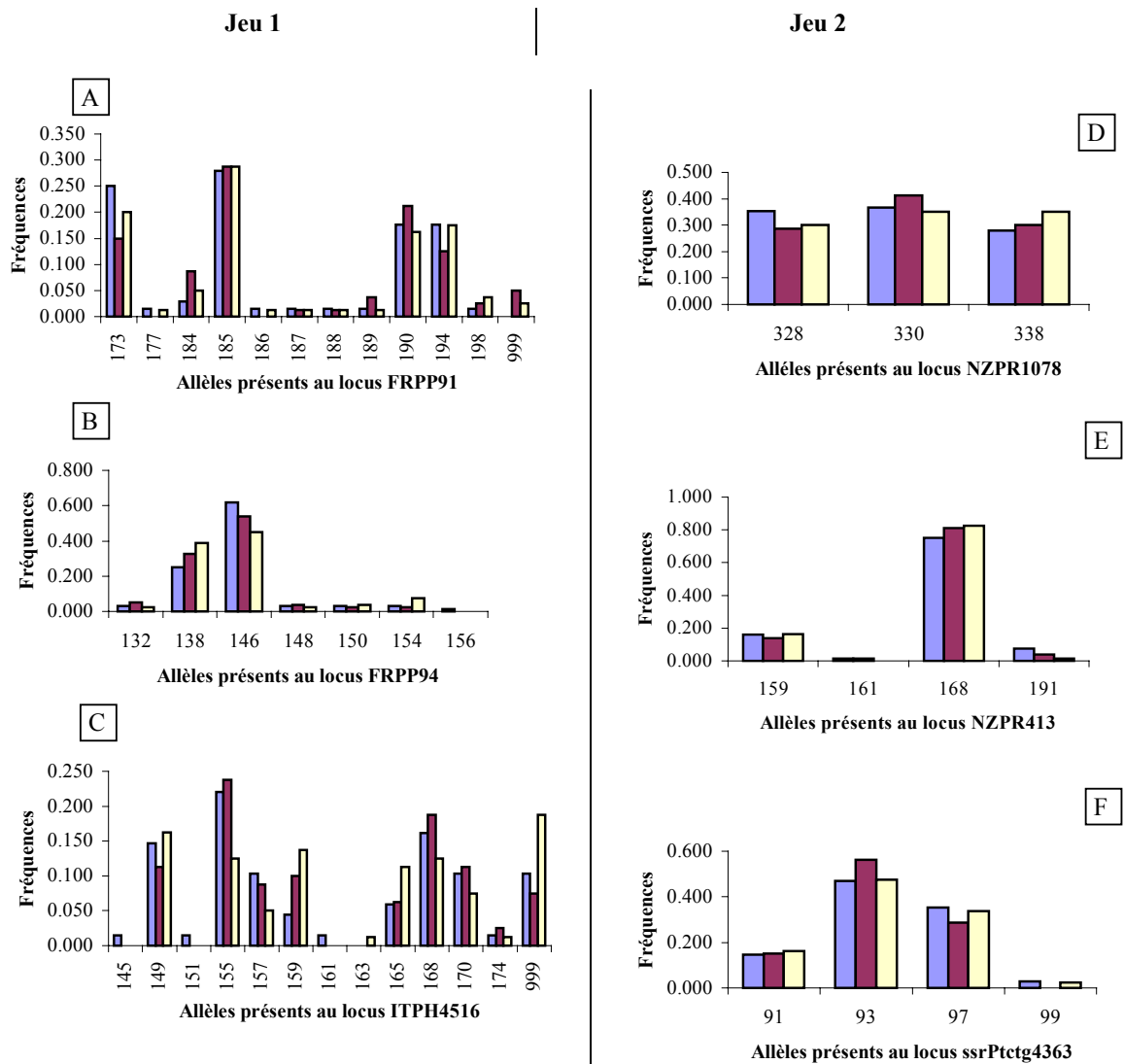
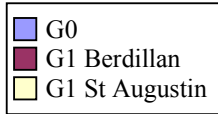
Annexe 16



Localisation et plan d'échantillonnage des vergers de Berdillan et de Saint Augustin. Les flèches pleines indiquent l'axe de prélèvement (cline).

Annexe 17

Comparaison de la distribution des allèles aux 6 loci étudiés dans les populations G0, G1 Berdillan et G1 St Augustin.



Annexe 18

Tableau III : Paramètres de diversité aux 6 loci microsatellites pour la population de 40 G1 échantillonnés dans le verger de Berdillan.

Locus	A ₀	A _e	H ₀	H _e	F
FRPP91	10	5.604	0.775	0.822	0.057 ^{ns}
FRPP94	6	2.502	0.550	0.600	0.084 ^{ns}
ITPH4516	9	6.911	0.925	0.855	-0.081 ^{ns}
NZPR1078	3	2.917	0.575	0.657	0.125 ^{ns}
NZPR413	4	1.469	0.325	0.319	-0.018 ^{ns}
ssrPtctg4363	3	2.372	0.500	0.578	0.136 ^{ns}
Ensemble des loci	5.8	3.629	0.608	0.639	0.048 ^{ns}

^{ns} = non significatif, * = p < 0.05, ** = p < 0.01

[A₀ = nombre d'allèles observés, A_e = nombre d'allèles efficaces, H₀ = hétérozygotie observée, H_e = hétérozygotie attendue, F = indice de fixation]

Tableau IV : Paramètres de diversité aux 6 loci microsatellites pour la population de 40 G1 échantillonnés dans le verger de St Augustin.

Locus	A ₀	A _e	H ₀	H _e	F
FRPP91	12	5.405	0.775	0.815	0.049 ^{ns}
FRPP94	6	2.771	0.675	0.639	-0.056 ^{ns}
ITPH4516	10	7.529	0.875	0.867	-0.009 ^{ns}
NZPR1078	3	2.985	0.750	0.665	-0.128 ^{ns}
NZPR413	3	1.414	0.250	0.293	0.146 ^{ns}
ssrPtctg4363	4	2.728	0.750	0.633	-0.184 ^{ns}
Ensemble des loci	6.3	3.805	0.679	0.652	-0.042 ^{ns}

^{ns} = non significatif, * = p < 0.05, ** = p < 0.01

[A₀ = nombre d'allèles observés, A_e = nombre d'allèles efficaces, H₀ = hétérozygotie observée, H_e = hétérozygotie attendue, F = indice de fixation]

Tableau V : Paramètres de diversité aux 6 loci microsatellites pour la population de 290 G2 issus du verger de Berdillan.

Locus	A ₀	A _e	H ₀	H _e	F
FRPP91	15	4.881	0.759	0.795	0.046 ^{ns}
FRPP94	8	2.171	0.486	0.539	0.099 ^{ns}
ITPH4516	13	7.402	0.903	0.865	-0.044 ^{ns}
NZPR1078	3	2.824	0.672	0.646	-0.041 ^{ns}
NZPR413	4	1.525	0.328	0.344	0.049 ^{ns}
ssrPtctg4363	5	2.363	0.607	0.577	-0.052 ^{ns}
Ensemble des loci	8	3.528	0.626	0.628	0.003 ^{ns}

^{ns} = non significatif, * = p < 0.05, ** = p < 0.01

[A₀ = nombre d'allèles observés, A_e = nombre d'allèles efficaces, H₀ = hétérozygotie observée, H_e = hétérozygotie attendue, F = indice de fixation]

Tableau VI : Paramètres de diversité aux 6 loci microsatellites pour la population de 290 G2 issus du verger de St Augustin.

Locus	A ₀	A _e	H ₀	H _e	F
FRPP91	13	6.945	0.797	0.856	0.069 ^{ns}
FRPP94	8	2.811	0.590	0.644	0.085 ^{ns}
ITPH4516	13	6.410	0.834	0.844	0.011 ^{ns}
NZPR1078	3	2.993	0.672	0.666	-0.010 ^{ns}
NZPR413	4	1.645	0.441	0.392	-0.126 ^{ns}
ssrPtctg4363	5	2.230	0.593	0.552	-0.075 ^{ns}
Ensemble des loci	7.7	3.839	0.655	0.659	0.007 ^{ns}

^{ns} = non significatif, * = p < 0.05, ** = p < 0.01

[A₀ = nombre d'allèles observés, A_e = nombre d'allèles efficaces, H₀ = hétérozygotie observée, H_e = hétérozygotie attendue, F = indice de fixation]

Tableau VII : Résultats de la recherche de paternité parmi les G1 génotypés pour 290 descendants du verger de Berdillan.

Population de référence*	Zone**	Arbre maternel	N	Seuil	Classification correcte (%)	N1	N1/N (%)	Seuil	Classification correcte (%)	N1'	N1'/N (%)
Population naturelle	5	B36	29	4,1	79,13	5	17,24	4,35	77,68	1	3,45
	5	B40	29			2	6,90			2	6,90
	4	B27	29			5	17,24			1	3,45
	4	B30	29			12	41,38			16	55,17
	3	B18	29			2	6,90			3	10,34
	3	B20	29			2	6,90			4	13,79
	2	B9	29			5	17,24			5	17,24
	2	B15	29			2	6,90			5	17,24
	1	B1	29			7	24,14			12	41,38
	1	B5	29			13	44,83			17	58,62
	Total	-	290			55	18,97			66	22,76
G0+G1+G2	5	B36	29	4,5	79,20	2	6,90	4,5	78,46	2	6,90
	5	B40	29			1	3,45			0	0
	4	B27	29			5	17,24			4	13,79
	4	B30	29			7	24,14			7	24,14
	3	B18	29			1	3,45			1	3,45
	3	B20	29			0	0			0	0
	2	B9	29			3	10,34			3	10,34
	2	B15	29			1	3,45			0	0
	1	B1	29			5	17,24			6	20,69
	1	B5	29			8	27,59			8	27,59
	Total	-	290			33	11,38			31	10,69

N : nombre de plantules analysées, N1 : nombre de plantules ayant 1 père potentiel dans le verger, N1' : nombre de plantules ayant 1 père potentiel dans le verger en tenant compte d'un taux d'erreur de typage de 0,1%.

* = population de référence considérée pour calculer les fréquences alléliques extérieures au verger

** = zone indiquée dans l'ordre du cline de prélèvement

Tableau VIII : Résultats de la recherche de paternité parmi les G1 génotypés pour 290 descendants du verger de St Augustin.

Population de référence *	Zone **	Arbre maternel	N	Seuil	Classification correcte (%)	N1	N1/N (%)	Seuil	Classification correcte (%)	N1'	N1'/N (%)
Population naturelle	4	A17	29	4,1	80,95	13	44,83	3,9	80,66	17	58,62
	4	A20	29			4	13,79			4	13,79
	2	A10	29			8	27,59			9	31,03
	2	A15	29			7	24,14			12	41,38
	3	A2	29			16	55,17			24	82,76
	3	A4	29			16	55,17			13	44,83
	1	A28	29			7	24,14			13	44,83
	1	A30	29			16	55,17			17	58,62
	5	A37	29			7	24,14			3	10,34
	5	A40	29			21	72,41			22	75,86
	Total	-	290			115	39,66			134	46,21
G0+G1+G2	4	A17	29	4,35	81,08	11	37,93	4,5	80,52	8	27,59
	4	A20	29			2	6,90			2	6,90
	2	A10	29			1	3,45			1	3,45
	2	A15	29			7	24,14			6	20,69
	3	A2	29			16	55,17			17	58,62
	3	A4	29			8	27,59			6	20,69
	1	A28	29			5	17,24			4	13,79
	1	A30	29			9	31,03			10	34,48
	5	A37	29			3	10,34			1	3,45
	5	A40	29			20	68,97			15	51,72
	Total	-	290			82	28,28			70	24,14

N : nombre de plantules analysées, N1 : nombre de plantules ayant 1 père potentiel dans le verger, N1' : nombre de plantules ayant 1 père potentiel dans le verger en tenant compte d'un taux d'erreur de typage de 0,1%.

* = population de référence considérée pour calculer les fréquences alléliques extérieures au verger

** = zone indiquée dans l'ordre du cline de prélèvement

Tableau IX: Estimation du flux de gènes caché et déduction du flux de gènes total pour 10 000 descendants simulés pour la population de Berdillan

Population de référence	Erreur de typage (%)	Seuil	Taille de la population (N)	Flux de gènes extérieurs				Décisions correctes (%)	%	% c/a	Flux de gènes total
				Attendu	Réel (r)	Apparent (a)	Caché (c)				
Population naturelle	0	4,1	500	9200	9163	7687	1751	78,58	-19,20	22,78	99,49
			1000	9600	9604	8003	1730	80,82	-20,00	21,62	98,55
	0,1	4,35	500	9200	9261	7723	1760	79,10	-19,91	22,79	94,84
			1000	9600	9595	7838	1898	79,09	-22,42	24,22	95,95
G0+G1+G2	0	4,5	500	9200	9203	7714	1751	78,79	-19,30	22,70	>100
			1000	9600	9614	7871	1869	79,41	-22,14	23,75	>100
	0,1	4,5	500	9200	9221	7789	1718	79,01	-18,38	22,06	>100
			1000	9600	9589	8122	1618	81,61	-18,06	19,92	>100

Tableau X: Estimation du flux de gènes caché et déduction du flux de gène total pour 10 000 descendants simulés pour la population de St Augustin

Population de référence	Erreur de typage (%)	Seuil	Taille de la population (N)	Flux de gènes extérieurs				Décisions correctes (%)	%	% c/a	Flux de gènes tot
				Attendu	Réel (r)	Apparent (a)	Caché (c)				
Population naturelle	0	4,1	500	9200	9210	6537	2839	68,64	-40,89	43,43	86,55
			1000	9600	9616	6920	2773	70,98	-38,96	40,07	84,53
	0,1	3,9	500	9200	9240	6445	2930	68,13	-43,37	45,46	78,25
			1000	9600	9596	6965	2716	71,35	-37,77	38,99	74,77
G0+G1+G2	0	4,35	500	9200	9203	7051	2404	72,39	-30,52	34,09	96,17
			1000	9600	9616	7208	2534	72,86	-33,41	35,16	96,94
	0,1	4,5	500	9200	9206	7440	2074	75,29	-23,74	27,88	97,01
			1000	9600	9625	7233	2501	73,40	-33,07	34,58	>100

Tableau XI: Caractéristiques de 3 loci microsatellites développés par Guevara *et al.*(2005) proposés pour la poursuite de l'étude.

Nom du locus	Motif répété	A ₀	H ₀	H _e
A5G11	AC ₂₂	7	0,875	0,807
A5B01	AC ₁₄	8	0,813	0,826
A7E03	AC ₁₄	12	0,739	0,840

A₀ = nombre d'allèles observés, H₀ = hétérozygotie observée, H_e = hétérozygotie attendue

Tableau XII: Pourcentage de flux externe apparent obtenu par simulation pour les 2 vergers à partir de 3 populations de référence différentes.

Population de référence	Berdillan	St Augustin
40*	77,24	53,79
1080**	51,38	36,90
5080**	51,38	31,03

* individus G1 génotypés spécifiques de chaque verger

** 80 G1 génotypés et 1000 ou 5000 individus simulés à partir des populations G0 et G1