



Université d'Évry – Val d'Essonne  
Ecole Doctorale des Génomes Aux Organismes

## THÈSE

Présentée pour obtenir le grade de :  
**Docteur en sciences de l'université d'Évry-Val d'Essonne**  
Spécialité : Biologie Moléculaire et Cellulaire

# Génomique comparative entre *Muscadinia rotundifolia* et *Vitis vinifera* pour faciliter l'identification de gènes de résistance

Par

**Iritché Cyrille ZAH-BI**

Le 6 janvier 2014

### JURY

**Elisabeth Dirlewanger**

**Valérie Geffroy**

**Paul Schellenbaum**

**Bénédicte Sturbois**

**Patricia Faivre-Rampant**

**Anne-Françoise Adam-Blondon**

**Directrice de Recherche, INRA Bordeaux**

**Chargée de Recherche, IBP Orsay**

**Maître de conférences, Université de Haute-Alsace**

**Professeur, Université d'Évry**

**Chargée de Recherche, INRA Evry**

**Directrice de Recherche, INRA Versailles**

**Rapporteur**

**Rapporteur**

**Examineur**

**Examineur**

**Co-Directrice de thèse**

**Directeur de thèse**



## Table des matières

<b>Table des matières.....</b>	<b>3</b>
<b>LISTE DES TABLEAUX .....</b>	<b>7</b>
<b>LISTE DES FIGURES.....</b>	<b>9</b>
<b>LISTE DES ANNEXES.....</b>	<b>12</b>
<b>Remerciements.....</b>	<b>13</b>
<b>1. Introduction .....</b>	<b>14</b>
<b>1.1. LA VIGNE.....</b>	<b>16</b>
1.1.1. Classification botanique et diversité de la vigne.....	16
1.1.1.1. Les vignes du sous-genre <i>Euvtis</i> .....	16
1.1.1.2. Les vignes du sous-genre <i>Muscadinia</i> .....	19
1.1.1.3. Histoire et domestication de la vigne.....	22
1.1.1.4. La viticulture mondiale .....	24
1.1.2. Cycle de développement de la vigne .....	26
1.1.2.1. Le cycle végétatif .....	26
1.1.2.2. Le cycle reproducteur .....	26
<b>1.2. LE CONTRÔLE PHYTOSANITAIRE AU VIGNOBLE .....</b>	<b>27</b>
1.2.1. L'oïdium de la vigne.....	29
1.2.1.1. Cycle de développement .....	30
1.2.1.2. Diversité et évolution de la maladie .....	32
1.2.2. Le mildiou de la vigne .....	34
1.2.2.1. Cycle de développement .....	35
1.2.2.2. Diversité et évolution de la maladie .....	37
1.2.3. Réduction de l'usage des produits phytosanitaires.....	38
1.2.4. Amélioration variétale pour la résistance aux maladies .....	41
<b>1.3. GENETIQUE DE LA RESISTANCE AUX MALADIES CHEZ LA VIGNE.....</b>	<b>43</b>
1.3.1. Cartographie génétique des résistances aux maladies chez la vigne .....	43
1.3.1.1. Cartographie génétique des gènes de résistance de la vigne .....	43
1.3.1.1.1. Les résistances provenant du sous-genre <i>Muscadinia</i> .....	44
1.3.1.1.2. Les résistances provenant du sous-genre <i>Euvtis</i> .....	45
<b>1.4. L'INTERACTION PLANTE - AGENT PATHOGENE : MECANISME DE RESISTANCE ET EVOLUTION DES GENES DERESISTANCE .....</b>	<b>49</b>
1.4.1. L'immunité chez les plantes - Les défenses induites.....	50
1.4.1.1. Le modèle Zig-Zag.....	50
1.4.1.2. La reconnaissance spécifique des agents pathogènes .....	52
1.4.2. Les protéines de résistance végétales .....	53
1.4.2.1. Les domaines fonctionnels des protéines de résistance.....	53
1.4.2.1.1. Classification des gènes <i>R</i> .....	56
1.4.2.1.2. Activation et régulation des gènes NBS-LRR.....	59
1.4.2.2. Les voies de signalisation impliquées dans la mise en place de la défense de la plante .....	63
1.4.2.2.1. Les flux ioniques.....	64

1.4.2.2.2. Les espèces réactives d'oxygène (ROS).....	65
1.4.2.2.3. La voie de transduction des MAP Kinases .....	65
1.4.2.2.4. Les molécules "signal" .....	65
1.4.3. Organisation et évolution des gènes NBS-LRR dans le génome des plantes.....	70
1.4.3.1. Organisation génomique et évolution des gènes NBS-LRR.....	70
1.4.3.1.1. Les gènes de résistances NBS-LRR complets.....	70
1.4.3.2. Les pseudogènes.....	72
1.4.3.3. Evolution des gènes NBS-LRR chez les plantes .....	74
1.4.4. Colocalisation des gènes NBS-LRR avec les <i>loci</i> impliqués dans la résistance aux maladies chez les plantes .....	76
1.4.4.1. Cas des génomes non séquencés.....	77
1.4.4.2. Cas des génomes entièrement séquencés .....	78
<b>1.5. CONTEXTE ET OBJECTIFS DE THESE .....</b>	<b>79</b>
1.5.1. Contexte .....	79
1.5.1.1. Supercontig de clones BAC de la région du <i>locus</i> de résistance <i>Run1/Rpv1</i> .....	80
1.5.1.2. Région homologue du <i>locus</i> de résistance <i>Run1/Rpv1</i> chez <i>V. vinifera</i> .....	82
1.5.1.3. Région homologue du <i>locus</i> de résistance <i>Rpv3</i> chez <i>V. vinifera</i> .....	84
1.5.2. Les Objectifs du travail de recherche .....	84
<b>2. Résultats .....</b>	<b>87</b>
<b>2.1. Introduction.....</b>	<b>89</b>
<b>2.2. Etude des paramètres généraux du génome de <i>M. rotundifolia</i> cv Regale.....</b>	<b>89</b>
2.2.1. Taux de Guanine-Cytosine (GC) des BES de <i>M. rotundifolia</i> cv Regale .....	90
2.2.2. Contenu en éléments répétés des BES de <i>M. rotundifolia</i> cv Regale.....	90
2.2.2.2. Les éléments transposables .....	93
2.2.3. Contenu en régions codantes des BES de <i>M. rotundifolia</i> cv Regale .....	96
<b>2.3. Analyse du contenu en régions codant pour des protéines NBS-LRR dans les BES .....</b>	<b>97</b>
2.3.1. Identification à l'aide des motifs MEME spécifiques du domaine NBS.....	97
2.3.2. Identification de BES portant des séquences NBS-LRR à l'aide de domaine NB-ARC des protéines fonctionnelles .....	98
2.3.2.1. Choix d'un set de gènes de résistances fonctionnels de la famille des NBS-LRR.....	98
2.3.2.2. Identification des BES portant les domaines NB-ARC des protéines fonctionnelles .....	99
2.3.3. Analyse phylogénétique des gènes à domaine NB-ARC de <i>M. rotundifolia</i> cv Regale .....	100
2.3.3.1. Sélection des BES de <i>M. rotundifolia</i> cv Regale pour l'analyse phylogénétique .....	100
2.3.3.1.1. Sélection des domaines NB-ARC des protéines fonctionnelles .....	101
2.3.3.1.2. Classification phylogénétique des domaines NB-ARC de gènes de résistance.....	104
<b>2.4. Alignement des BES de <i>M. rotundifolia</i> cv Regale sur la séquence de <i>V. vinifera</i>.....</b>	<b>106</b>
<b>2.5. Alignement des BES de <i>M. rotundifolia</i> cv Regale sur la séquence du génome de <i>V.</i> <i>vinifera</i> cv PN40024 et sur le supercontig de la région du <i>locus</i> de résistance <i>Run1/Rpv1</i></b> .....	<b>113</b>
2.5.1. Détermination des paramètres d'alignement des BES sur le génome de <i>V. vinifera</i> .....	113
2.5.1.1. Alignement des BES de <i>M. rotundifolia</i> cv Regale sur le génome de <i>V. vinifera</i> .....	114
2.5.1.2. Identification de clones BAC de <i>M. rotundifolia</i> cv Regale dont les BES s'alignent sur le supercontig de séquence contenant les gènes <i>Run1</i> et <i>Rpv1</i> .....	116



2.5.2. Identification des clones BAC de <i>M. rotundifolia</i> cv Regale dont les BES s'alignent sur le génome de référence dans les régions de <i>Run1/Rpv1</i> et <i>Rpv2/Rpv3</i> .....	117
2.5.3. Enrichissement des <i>loci</i> de résistance <i>Run1/Rpv1</i> , <i>Rpv2/Rpv3</i> en marqueurs moléculaires .....	120
2.5.3.1. Ancrage <i>in silico</i> des marqueurs génétiques et physiques des régions génomiques .....	120
2.5.3.2. Développement des couples d'amorces uniques dans la région des <i>loci</i> de résistance .....	121
<b>2.6. CRIBLAGE PAR PCR DE LA BANQUE BAC DE <i>M. ROTUNDIFOLIA</i> CV REGALE .....</b>	<b>122</b>
2.6.1. Tests d'amplification génomiques des marqueurs sélectionnés dans les régions d'intérêt .....	122
2.6.2. Criblage par PCR de la banque BAC de <i>M. rotundifolia</i> cv Regale .....	123
2.6.3. Identification de séquences NBS dans les clones BAC candidats .....	124
2.6.4. Choix des clones BAC de <i>M. rotundifolia</i> cv Regale à séquencer .....	127
2.6.4.1. Dans la région du <i>locus</i> de résistance <i>Run1/Rpv1</i> .....	127
2.6.4.2. Dans la région du <i>locus</i> de résistance <i>Rpv3</i> (Bellin <i>et al.</i> 2009) .....	129
<b>2.7. SEQUENÇAGE ET ANNOTATION DES CLONES BAC CANDIDATS .....</b>	<b>132</b>
2.7.1. Séquençage des clones BAC candidats .....	132
2.7.2. Assemblage des contigs .....	133
2.7.3. Annotations structurale et fonctionnelle des contigs de BAC .....	140
2.7.3.1. Méthode de prédiction <i>ab initio</i> .....	141
2.7.3.2. Annotation manuelle des gènes des contigs de BAC .....	141
2.7.3.2. Annotation fonctionnelle des modèles de gène .....	142
2.7.4. Analyse de la colinéarité entre les séquences de <i>M. rotundifolia</i> et <i>V. vinifera</i> .....	146
<b>3. Discussion &amp; Conclusions .....</b>	<b>151</b>
<b>3.1. Caractéristiques du génome de <i>M. rotundifolia</i> cv Regale sur la base de l'analyse de la collection de BES .....</b>	<b>152</b>
3.1.1. Taux de Guanine-Cytosine (GC) .....	152
3.1.2. Les séquences microsatellites .....	154
3.1.3. Les éléments transposables .....	155
3.1.4. Les régions codantes .....	157
<b>3.2. Génomique comparative entre <i>M. rotundifolia</i> et <i>V. vinifera</i> .....</b>	<b>158</b>
3.2.1. Transférabilité des marqueurs entre les deux génomes .....	159
3.2.2. Recherche de zones de synténie entre <i>M. rotundifolia</i> et <i>V. vinifera</i> par alignement des paires de BES .....	160
3.2.3. La famille des gènes de résistance de type NBS-LRR .....	161
3.2.4. Etude de la synténie entre <i>M. rotundifolia</i> et <i>V. vinifera</i> dans la région de <i>clusters</i> de gènes NBS-LRR .....	165
3.2.4.1. Ebauche de carte physique dans les régions des gènes <i>Run1/Rpv1</i> et <i>Rpv3</i> .....	165
3.2.4.2. Colinéarité de séquence dans la région des <i>loci</i> de résistance <i>Run1/Rpv1</i> et <i>Rpv3</i> .....	167
<b>4. Perspectives .....</b>	<b>170</b>
<b>5. Matériels &amp; Méthodes .....</b>	<b>173</b>
<b>4.1. MATRIELS .....</b>	<b>174</b>
4.1.1. Matériel végétal et préparation de l'ADN génomique .....	174

4.1.2. Ressources génomiques.....	174
4.1.2.1. Construction de la banque BAC de <i>M. rotundifolia</i> cv Regale.....	174
4.1.2.2. Séquences d'extrémités de BAC (BES).....	174
4.1.3. Séquence du contig et des clones BAC dans la région du gène <i>Run1/Rpv1</i> .....	175
4.1.4. Séquence de référence de <i>Vitis vinifera</i> .....	175
<b>4.2. METHODES .....</b>	<b>176</b>
4.2.1. Analyse des collections de BES.....	176
4.2.1.1. Identification des séquences microsatellites.....	176
4.2.1.2. Analyse des séquences répétées connues.....	176
4.2.1.3. Analyse des régions codantes.....	176
4.2.1.4. Séquences et analyse phylogénétique.....	177
4.2.1.5. Alignement des BES sur le génome de référence.....	177
4.2.2. Criblage de la banque BAC et sélection des clones BAC candidats.....	177
4.2.2.1. Choix des marqueurs génétiques et physiques.....	177
4.2.2.2. Organisation et criblage de la banque BAC de <i>M. rotundifolia</i> cv Regale.....	178
4.2.2.3. Amplification par PCR.....	178
4.2.2.4. Identification des clones BAC positifs.....	179
4.2.3. Séquençage et annotation des contigs de BAC.....	179
4.2.3.1. Raccordement des contigs.....	179
4.2.3.2. Annotation des contigs de clones BAC.....	179
<b>Références bibliographiques.....</b>	<b>181</b>
<b>Annexes .....</b>	<b>209</b>

## LISTE DES TABLEAUX

Tableau 1. Utilisation des espèces dans les programmes de sélection au cours du dernier siècle

Tableau 2. Différences entre les sous-genres *Euvitis* et *Muscadinia*

Tableau 3. QTLs associés à la résistance contre *P. viticola* et *E. necator*

Tableau 4. Gènes et pseudogènes codant des gènes *NBS-LRR* identifiés dans différents génomes de plantes

Tableau 5. Caractéristiques de la banque BAC de *Muscadinia rotundifolia* cv Regale

Tableau 6. Caractéristiques des génomes de diverses plantes sur la base des séquences d'extrémités de BAC et du génome de *V. vinifera* cv PN40024

Tableau 7. Alignement des collections de BES contre le protéome de *V. vinifera* cv PN40024 par BLASTX. Uniquement les alignements avec 90 - 100 % d'identité sont considérés.

Tableau 8. Alignement des BES de *M. rotundifolia* cv Regale sur la séquence du génome de *V. vinifera* cv PN40024

Tableau 9. Alignement des BES de *M. rotundifolia* cv Regale dans les régions de *Run1/Rpv1* et *Rpv2/Rpv3* de la séquence du génome de *V. vinifera* cv PN40024 et du supercontig

Tableau 10. Localisation *in silico* des marqueurs génétiques et moléculaires des régions contenant les gènes de résistance *Run1*, *Rpv1*, *Rpv2* et *Rpv3* sur le génome de *V. vinifera* cv PN40024, les contigs de BAC de *M. rotundifolia* cv G52 et sur les BES de *M. rotundifolia* cv Regale

Tableau 11. Séquençage des clones BAC candidats pour les gènes *Run1*, *Rpv1*, *Rpv2* et *Rpv3* par la méthode de séquençage 454

Tableau 12. Alignement des fragments amplifiés avec les couples d'amorces définis aux extrémités du clone BAC 34D01.

Tableau 13. Résumé des séquences manquantes (*gap*) entre les différents contigs suivant leur alignement sur le génome de référence de *V. vinifera* cv PN40024 et le supercontig de *M. rotundifolia* cv G52

Tableau 14. Prédiction automatique des gènes dans les contigs de *M. rotundifolia* cv Regale par FGENESH (Solovyev *et al.* 2006)

Tableau 15. Résumé des modèles de gènes annotés dans les contigs de *M. rotundifolia* cv Regale

Tableau 16. Annotation fonctionnelle des séquences protéiques extraites des CDS de *M. rotundifolia* cv Regale

Tableau 17. Vue d'ensemble du contenu des BES de *V. vinifera* cv Cabernet Sauvignon, *M. rotundifolia* cv Regale, de la tomate (*S. lycopersicon*) et de la pomme de terre (*S. tuberosum*)

Tableau 18. Contenu des rétrotransposons les plus abondants des BES, des clones BAC et des génomes entiers de la tomate et de la pomme de terre (Zhu *et al.* 2008 ; Datema *et al.* 2008)

Tableau 19. Distribution des gènes à motifs NB-ARC du génome de *V. vinifera* cv PN40024 d'après Yang *et al.* 2008, la dernière version de l'annotation (V1) du génome de *V. vinifera* cv PN40024 et des BES de *M. rotundifolia* cv Regale.

Tableau 20. Alignement des fragments amplifiés par PCR avec les couples d'amorces définis aux extrémités des clones BAC de *M. rotundifolia* cv G52.

## LISTE DES FIGURES

Figure 1. Classification de *Muscadinia* selon l'hypothèse de Planchon, 1887

Figure 2. Origine géographique des espèces du genre *Vitis*.

Figure 3. Différences morphologiques entre les genres *Vitis* et *Muscadinia*

Figure 4. L'ivresse du patriarche Noé, par Giovanni Bellini vers 1515

Figure 5. Les 13 plus importantes productions agricoles en France pour l'année 2011.

Figure 6. Les 10 plus importantes productions vinicoles dans le monde pour l'année 2011

Figure 7. Les symptômes et les signes représentatifs de l'oïdium

Figure 8. Cycle de développement de *Erysiphe necator*.

Figure 9. Les symptômes et les signes représentatifs du mildiou

Figure 10. Cycle de développement de *Plasmopara viticola*

Figure 11. Les dépenses en pesticides dans l'agriculture française en 2006

Figure 12. Répartition des petites régions agricoles selon l'origine et le niveau de la pression pesticide (pesticides par hectare)

Figure 13. Généalogie des variétés résistantes commercialisées.

Figure 14. Modèle en Zig-Zag illustrant les mécanismes de défense mis en place contre les pathogènes (d'après Jones et Dangl, 2006)

Figure 15. Le modèle de résistance « gène pour gène »

Figure 16. Représentation schématique d'une protéine NB-LRR

Figure 17. Représentation de la localisation et de la structure des cinq classes principales de récepteurs impliqués dans l'ETI

Figure 18. Modèle d'activation des protéines R proposé par Lukasik *et al.* (2009)

Figure 19. Schéma d'élicitation des réactions de défense chez les plantes

Figure 20. Les cascades de MAPK induites par les PAMP dans la défense des plantes aux agents pathogènes bactériens et fongiques (d'après Pitzschke *et al.* 2009)

Figure 21. Le résumé du mode d'action et des mécanismes régulant l'expression du gène *N* à différents niveaux

Figure 22. Représentation schématique du contig de BAC et de la séquence correspondante dans la région des gènes *Run1/Rpv1*

Figure 23. Colocalisation du déterminisme génétique de la réponse d'hypersensibilité entre les marqueurs UDV-305 et VMC7F2, et du *locus* de résistance au mildiou, *Rpv3*, sur le chromosome 18 de l'hybride Bianca (Bellin *et al.* 2009)

Figure 24. Analyse comparative des éléments répétés dans les BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon, et dans le génome de *V. vinifera* cv PN40024.

Figure 25. Les SSR les plus abondants dans les BES de *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon.

Figure 26. Distribution des motifs SSR les plus abondants dans les BES de *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon.

Figure 27. Analyse comparative des éléments répétés majeurs dans les BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon, et dans le génome de *V. vinifera* cv PN40024.

Figure 28. Nombre de BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon contenant les motifs MEME spécifiques des TNL et des CNL.

Figure 29. Nombre de BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon identifiées par comparaison de séquences avec des domaines NB-ARC de protéines de résistance fonctionnelles (NB-ARC) et des motifs MEME spécifiques du domaine conservé NB-ARC (DOMAINE MEME).

Figure 30. Domaines PFAM détectés dans la séquence protéique extraite des BES par alignement avec des protéines connues.

Figure 31. Analyse des domaines NBS de protéines de résistance fonctionnelles avec le programme CLUSTALX

Figure 32. Analyse des domaines NB-ARC de protéines de résistance fonctionnelles avec le programme CLUSTALX

Figure 33. Proportion des BES contenant des domaines NB-ARC alignées sur le génome de référence de *V. vinifera* cv PN40024 par BLASTN

Figure 34. Organisation chromosomique des gènes de résistances à domaine NB-ARC de *V. vinifera* cv PN40024 et de *M. rotundifolia* cv Regale.

Figure 35. Alignement des BES de *M. rotundifolia* cv Regale sur la séquence de référence de *V. vinifera* cv PN40024 grâce à l'algorithme BLAT (Kent, 2002).

Figure 36. Pourcentage des chromosomes de *V. vinifera* couvert par les contigs de clone BAC de *M. rotundifolia* cv Regale.

Figure 37. Alignement des clones BAC de *M. rotundifolia* cv Regale *via* leur BES sur la séquence *M. rotundifolia* cv G52 contenant les gènes *Run1/Rpv1* et sur la séquence de référence du génome de *Vitis vinifera* dans la région correspondante du chromosome 12

Figure 38. Exemple de test d'amplification des marqueurs de la région du *locus* de résistance *Run1/Rpv1*.

Figure 39. Electrophorèse sur gel d'agarose 1% avec le couple d'amorces dégénérées RG1/RG2, spécifiques du motif NBS.

Figure 40. Résultat du criblage de la banque BAC *M. rotundifolia* cv Regale BAC par les différentes approches utilisées au cours de cette étude (alignement des BES, EPCR, PCR) avec pour cible la région contenant le pic du QTL *Rpv3*.

Figure 41. Position des BACs choisis pour être séquencés dans la région des gènes *Run1/Rpv1*.

Figure 42. Choix des BAC à séquencer dans la région du *locus Rpv3*.

Figure 43. Alignement des contigs de clones BAC *M. rotundifolia* cv Regale sur la séquence du génome de *V. vinifera*.

Figure 44. Modèle des différentes combinaisons de couples d'amorces pour l'assemblage des contigs d'un même clone BAC

Figure 45. Assemblage des contigs grâce à une EPCR sur la séquence de *V. vinifera* cv PN40024 et sur le supercontig de la région introgressée de *M. rotundifolia* cv G52

Figure 46. Assemblage des contigs grâce à une amplification par PCR du génome des clones BAC d'intérêt

Figure 47. Classification des protéines en fonction de la présence de domaines fonctionnels, répétitions, et signaux grâce aux bases de données PFAM (Finn *et al.* 2010) et InterPro (Hunter *et al.* 2011).

Figure 48. Microcolinéarité entre les clones BAC de *M. rotundifolia* cv Regale et les régions homologues chez *V. vinifera* cv PN40024.

## LISTE DES ANNEXES

Annexe 1. Classification et distribution des séquences répétées de plantes connues dans les séquences d'extrémités de BAC

Annexe 2. Motifs MEME d'acides aminés spécifiques du domaine NBS

Annexe 3. Liste des gènes de résistances fonctionnels de la famille des NBS-LRR clonés chez les végétaux en date du 13 - 02 – 2013

Annexe 4. Identification des séquences d'acides aminés déduites des BES de *M. rotundifolia* cv Regale grâce au motif PFAM spécifique du domaine NB-ARC

Annexe 5. Alignement des BES de *M. rotundifolia* cv Regale avec des modèles de gènes à motif NB-ARC

Annexe 6. Séquences des couples d'amorces définis dans les BES de *M. rotundifolia* cv Regale

Annexe 7. Les couples d'amorces définis dans la région du *locus* de résistance *Run1/Rpv1* avec une localisation unique par EPCR sur le génome de référence de *V. vinifera* cv PN40024

Annexe 8. Les couples d'amorces définis dans la région du *locus* de résistance *Rpv2/Rpv3* avec une localisation unique par EPCR sur le génome de référence de *V. vinifera* cv PN40024.

Annexe 9. Amplification des génomes de *V. vinifera* cv PN40024 et de *M. rotundifolia* cv Regale

Annexe 10. Alignement des contigs sur le génome de *V. vinifera* cv PN40024 par BLASTN

Annexe 11. La méthode de prédiction de gènes *ab initio* a été faite à l'aide du programme FGENESH (Yu *et al.* 2002 ; Goff *et al.* 2002).

Annexe 12. Protéines homologues aux séquences protéiques de *M. rotundifolia* cv Regale

Annexe 13. Les codes PFAM et les termes *GO* assignés aux séquences protéiques extraites des CDS de *M. rotundifolia* cv Regale

Annexe 14. Gènes du génome de référence de *V. vinifera* cv PN40024 couverts par les contigs de *M. rotundifolia* cv Regale

Annexe 15. PUBLICATION EN PREPARATION



## **Remerciements**

*Je tiens à exprimer mes plus vifs remerciements à Anne-Françoise et Patricia qui ont été pour moi des directrices de thèse attentives et disponibles, malgré leurs nombreuses charges. Leur compétence, leur rigueur scientifique, leur patience et leur clairvoyance m'ont beaucoup appris. Elles ont été et resteront pour moi un exemple*

*J'exprime tous mes remerciements à l'ensemble des membres de mon jury. Merci d'avoir accepté de vous plonger dans mon sujet, surtout en ces périodes de vacances...*

*J'exprime tous mes remerciements à Bénédicte Sturbois et Nathalie Boudet pour la confiance qu'elles m'ont accordée lors de mes années de monitorat. Merci surtout de m'avoir donné envie d'aller plus loin dans la Biologie du Végétal... ♣*

*J'adresse toute ma gratitude à tous ceux qui ont pu m'aider, ne serait-ce que par un petit sourire... Merci à Aurélie, Cécile, à Nathalie et Marc, à mes ami(e)s et à toutes les personnes qui m'ont aidé dans la réalisation de ce travail. Je remercie du plus profond de mon cœur Isabelle et son "fouet". Merci pour ta disponibilité, ton attention... ta bonne humeur contagieuse. Toujours dans la bonne humeur, je voudrais remercier toute l'équipe de Boulos... Harry l'inépuisable, Nathalie pour les conseils et explications, Smahane pour son aide précieuse... Merci à tous de m'avoir permis de travailler dans d'aussi bonnes conditions. Que la génomique soit avec vous.*

*J'exprime tous mes remerciements à toutes ces personnes dans l'ombre dont la contribution à mon travail est non négligeable...*

*Je remercie toutes les personnes formidables que j'ai rencontrées par le biais de l'URGV. Merci pour votre support et vos encouragements. Je pense particulièrement à tous ceux avec qui j'ai joué au foot...*

*J'exprime un profond remerciement à mes frères et sœurs de la Frat' Adoramus, particulièrement au couple Kodjo...*

*Je voudrais remercier tous mes ami(e)s pour les encouragements, les inquiétudes, les dépannages d'ordinateurs... Une pensée spéciale pour la chorale St Corbinien 🎵🎵, ainsi qu'à la communauté paroissiale de la Cathédrale d'Evry.*

*Une pensée particulière pour ma famille, particulièrement à Mr et Mme ZAH-BI. Merci pour vos prières et vos encouragements ; merci d'avoir été pour moi, le vent qui a fait avancer ma barque au large... Sans exception, je vous aime. Merci à ma chère et tendre fiancée Eva Nick Tessia ♥ Tu m'as permis de relever certains défis alors que tout semblait obscur devant moi... Merci du plus profond de mon cœur... Je t'♥*

*Je ne saurais terminer sans rendre toute la gloire à Celui en qui j'ai mis toute ma confiance  
Que toute gloire Te soit rendu Seigneur mon Dieu*

*Shalom*

# ***1. Introduction***

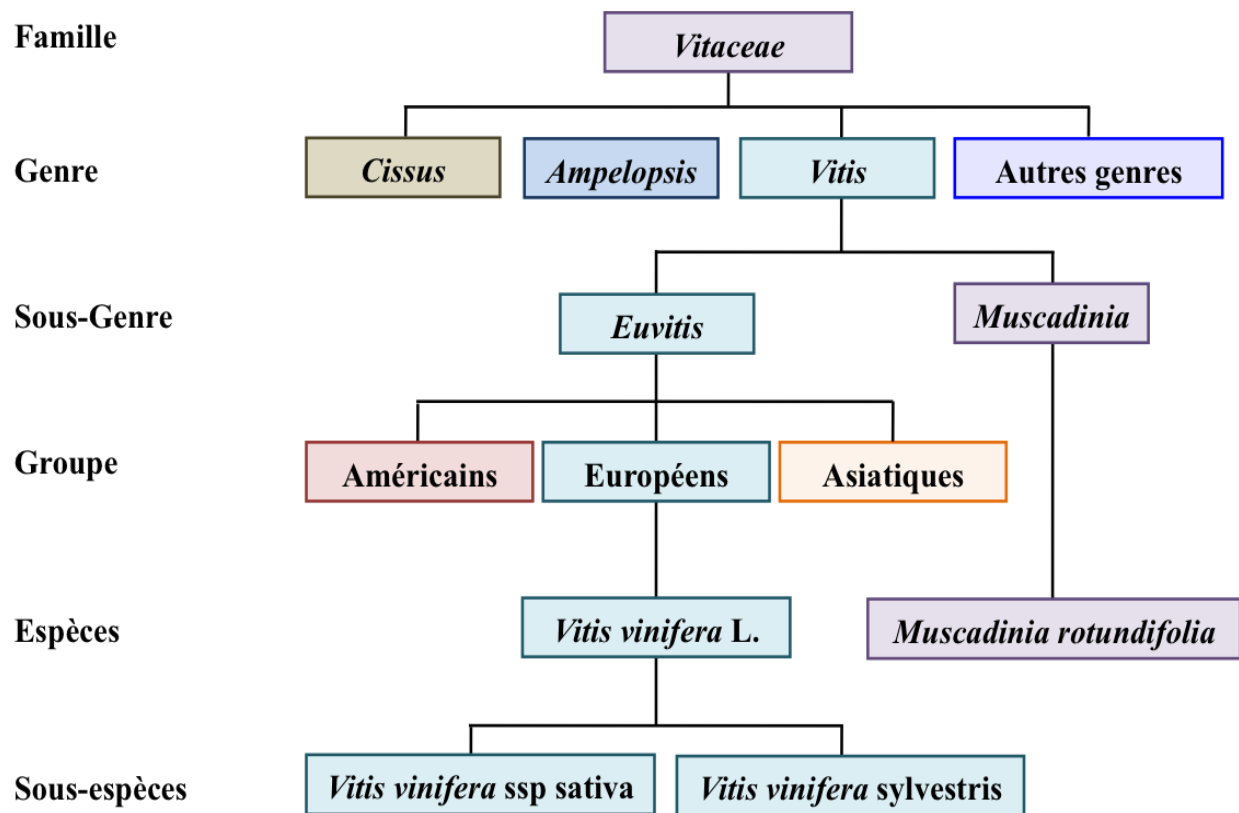


Figure 1. Classification de *Muscadinia* selon l'hypothèse de Planchon, 1887

## 1.1. LA VIGNE

### 1.1.1. Classification botanique et diversité de la vigne

La vigne est une plante angiosperme qui appartient à l'ordre des *Vitales* et à la famille des *Vitaceae*. Cette famille comprend 700 espèces réparties en 16 genres, essentiellement des lianes pérennes ou herbacées (Figure 1). Le genre *Vitis* est divisé en deux sous-genres, *Muscadinia* et *Euvitis* (Planchon, 1887). Ils se distinguent en fonction des caractéristiques morphologiques, anatomiques, cytologiques (Mullins *et al.* 1992). Seuls ces 2 sous-genres contiennent des espèces cultivées (This *et al.* 2006). D'après Péros *et al.* (2010), l'origine du genre *Vitis* est située en Eurasie et il s'est ensuite étendu vers l'Ouest sur le continent américain. La séparation des continents et les périodes de glaciations successives du pléistocène, ont alors provoqué l'isolement de populations qui a conduit à des événements de spéciations (Péros *et al.* 2010).

#### 1.1.1.1. Les vignes du sous-genre *Euvitis*

Le sous-genre *Euvitis*, pour “vraies vignes”, comprend 36 espèces diploïdes ( $2n = 38$ ) qui se trouvent dans les zones tempérées, subtropicales et tropicales de l'hémisphère nord (Figure 2). À l'origine dioïque, la plupart des cépages de *Vitis vinifera* L. est actuellement hermaphrodite ce qui en a fait une plante plus facilement cultivable (Mullins *et al.* 1992). On les classe en trois groupes, en fonction de leurs origines géographiques :

**Les vignes américaines** comptent une vingtaine d'espèces réparties en Amérique du nord et en Amérique centrale et dont 3 ont une importance sur le plan cultural. *Vitis labrusca* est une vigne du Sud du Canada et de la bordure nord-est Atlantique des États-Unis. Elle est cultivée, essentiellement pour la production de jus de fruits en Amérique du nord mais également utilisée comme parent dans l'élaboration de variétés résistantes au froid (Tableau 1). *Vitis riparia* est originaire du Nord-est des États-Unis, *Vitis berlandieri* du Texas et Nord-est du Mexique. Ces vignes, ainsi que leurs hybrides, présentent une bonne résistance aux pathogènes (Galet, 1957 ; Fischer *et al.* 2004) et sont aujourd'hui utilisées comme porte-greffe dans 85 % du vignoble mondial (Tableau 1).

**Les vignes asiatiques** représentent une dizaine d'espèces réparties en Asie orientale. Parmi elles, seule *Vitis amurensis* est cultivée pour ses baies. Certaines de ces espèces peuvent être utilisées

pour faire du jus, notamment *Vitis coignetiae* (Vigne du Japon). L'espèce *Vitis amurensis* est principalement distribuée dans les régions Nord-Est de la Chine à la Corée du Nord. Certaines accessions de *V. amurensis* montrent un haut niveau de résistance au mildiou (Tableau 1 ; Staudt et Kassemeyer, 1995 ; Cadle-Davidson, 2008). Outre la résistance au mildiou, *V. amurensis* présente de nombreuses caractéristiques favorables telles que la tolérance au froid (Ma *et al.* 2010), la résistance à l'oïdium (Wan *et al.* 2007), l'anthracnose et la pourriture blanche (Li *et al.* 2008), et la présence de composés d'intérêt médicinal (Huang *et al.* 2001 ; Ha *et al.* 2009 ; Yim *et al.* 2010). Ces caractères ont incité les sélectionneurs à inclure cette espèce dans leurs programmes de sélection.

**Les vignes européennes et du Proche Orient** ne comprennent que l'espèce *Vitis vinifera*. On distingue cependant la vigne sauvage *Vitis vinifera ssp sylvestris* et la vigne cultivée *Vitis vinifera ssp sativa* (Huglin et Schneider, 1998) pour les raisons de cuve et de table. La Vigne cultivée présente un polymorphisme remarquable, et le nombre de variétés est estimé à 5000 (la plupart sont conservées dans les collections). Ces différentes variétés de vigne sont également appelées cépages par les vignerons. En France, 249 cépages sont autorisés par la réglementation. Une quarantaine sont plus particulièrement utilisés et 12 d'entre eux couvrent 70 % des vignobles : Merlot, Grenache, Carignan, Cabernet Sauvignon, Syrah, Cabernet Franc, Gamay, Cinsault et Pinot Noir en rouge, Ugni blanc, Chardonnay et Sauvignon en blanc.

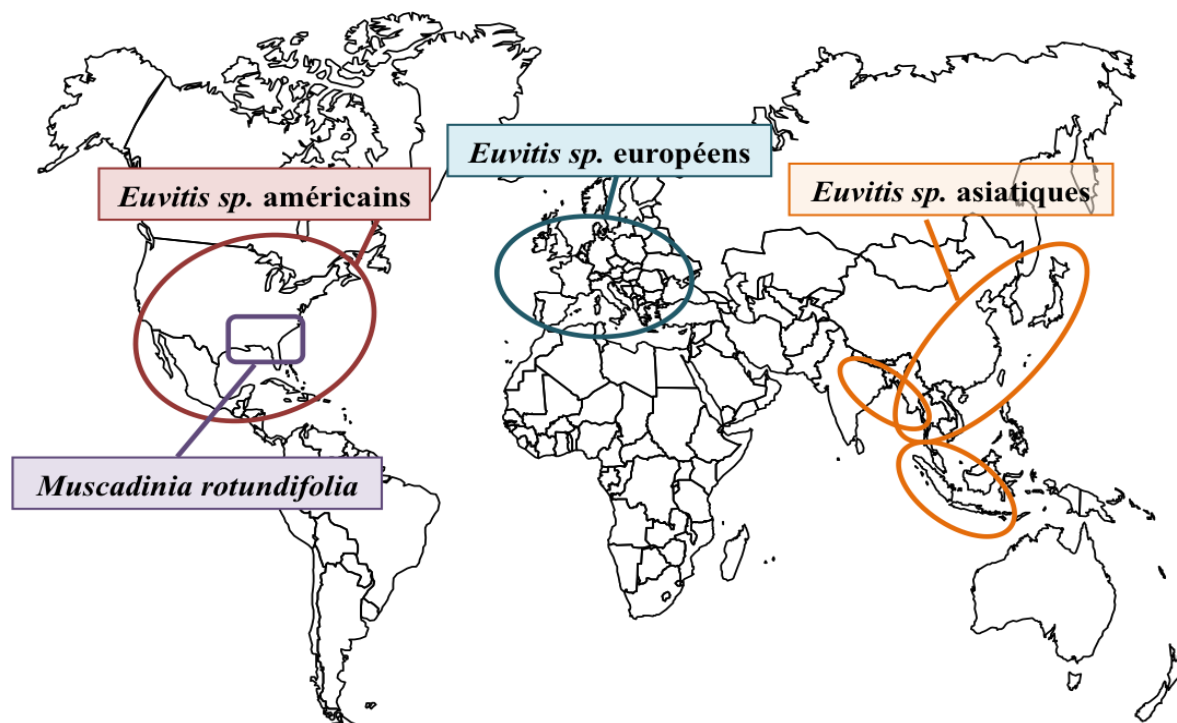


Figure 2. Origine géographique des espèces du genre *Vitis*.

Tableau 1. Utilisation des espèces dans les programmes de sélection au cours du dernier siècle

Espèces	Origine	Utilisations dans les programmes de sélection		
		Sélection de nouvelles variétés	Sélection de porte-greffes	Production de variétés hybrides interspécifiques
<i>M. rotundifolia</i>	Sud-est des États-Unis	++	+	+
<i>V. aestivalis</i>	Canada (Ontario) ; États-Unis	-	-	++
<i>V. amurensis</i>	Asie	+	-	++
<i>V. berlandieri</i>	États-Unis	+	+++	-
<i>V. candicans</i>	États-Unis	-	+	-
<i>V. caribaea</i>	Caraïbes	-	-	+
<i>V. champinii</i>	États-Unis	+	+	-
<i>V. cinerea</i>	États-Unis	-	+	++
<i>V. cordifolia</i>	États-Unis	-	+	+
<i>V. labrusca</i>	États-Unis	+++	++	+++
<i>V. longii</i>	États-Unis	+	++	-
<i>V. riparia</i>	Canada ; États-Unis	++	+++	+++
<i>V. rupestris</i>	États-Unis	++	+++	+++
<i>V. simpsonii</i>	États-Unis	-	+	-
<i>V. vinifera</i>	Europe, Asie, Bassin méditerranéen	+++++	+	++++

\* Les programmes de sélection sont faits pour la production de raisons de cuve ou de table

### 1.1.1.2. Les vignes du sous-genre *Muscadinia*

Le sous-genre *Muscadinia* ne comporte que trois espèces diploïdes *M. munsoniana*, *M. popeniae*, et *M. rotundifolia*, originaires des régions subtropicales d'Amérique du nord (Figure 1). Ces espèces ont 40 chromosomes ( $1n = 20$ ), ce qui les différencie des espèces du genre *Vitis* qui ont 38 chromosomes ( $1n = 19$ ) (Patel et Olmo 1955 ; Bouquet, 1980). Outre les différences cytogénétiques, de nombreuses caractéristiques morphologiques distinguent les espèces du genre *Muscadinia* des espèces du genre *Vitis* (Bouquet, 1980 ; Comeaux *et al.* 1987) dont quelques-unes sont résumées dans la Figure 3 et Tableau 2.

Seule l'espèce *M. rotundifolia* est cultivée pour la production de fruits, de jus de raisin et de vin dans les états du sud des Etats-Unis bordant le golfe du Mexique (Olmo 1986 ; Olien, 1990a). En effet, des efforts en vue de cultiver les espèces de vignes européennes *V. vinifera* ont échoué dans cette région, en raison de la pression des maladies fortement accentuée par le climat chaud et humide (Goldy, 1992). Or, la plupart des ravageurs et des maladies qui limitent la production mondiale des variétés de *V. vinifera* sont originaires d'Amérique du Nord. Les espèces de vignes nord-américaines ont développé une résistance aux parasites et aux maladies de la vigne, et l'espèce *M. rotundifolia* est considérée comme ayant les plus forts niveaux de résistance à la plupart d'entre eux (Olmo, 1986 ; Zecca *et al.* 2012). Les espèces du genre *Muscadinia* présentent ainsi une forte résistance aux dégâts causés par le *Phylloxera vastatrix* tant au niveau de la racine qu'au niveau des feuilles (Davidis et Olmo, 1964; Firoozabady et Olmo, 1982 ; Grzegorzczak et Walker, 1998), à la maladie de Pierce, causée par la bactérie *Xylella fastidiosa* (Olmo, 1971 ; Ruel et Walker, 2006), au nématode à galles, *Meloidogyne* spp. (Bloodworth *et al.* 1980 ; Firoozabady et Olmo 1982 ; Walker *et al.* 1994a), au nématode vecteur du virus du court-noué, *Xiphinema index* (Bouquet, 1980 ; Olmo, 1986 ; Esmenjaud *et al.* 2010); à l'antracnose (Mortensen, 1981), à la pourriture des racines, *Armillaria* (Clayton, 1975), et aux plus répandues des maladies fongiques de la vigne que sont l'oïdium, *Erysiphe necator* et le mildiou, *Plasmopara viticola* (Bouquet, 1980 ; Olmo 1986 ; Merdinoglu *et al.* 2003). Enfin, des essais ont indiqué que les hybrides F1 issus du croisement entre *V. vinifera* × *M. rotundifolia* n'ont pas les symptômes de dégénérescence du court-noué et donc semblent présenter une résistance au *Grapevine Fan Leaf Virus* (Walker *et al.* 1991, 1994b). Ces niveaux de résistance aux maladies ont stimulé plus de 100 ans d'efforts pour hybrider les sous-genres *Vitis* et *Muscadinia* (Detjen, 1919 ; Dunstan, 1962 ; Patel et Olmo, 1955 ; Olmo, 1971 ; Bouquet, 1980 ; Olmo, 1986 ; Olien,

1990b ; Goldy, 1992). La méthode de sélection classique, visant à combiner la résistance de *M. rotundifolia* avec la qualité des fruits élevée de *V. vinifera*, est difficile en raison du décalage des dates de floraison et des différences dans le nombre de chromosomes. De plus, les hybrides obtenus sont généralement stériles. La fécondité des hybrides F1 est fortement dépendante de la sélection du parent femelle *V. vinifera* (Patel et Olmo, 1955). La production de porte-greffes *Vitis* × *M. rotundifolia* a par ailleurs été limitée par leur faible capacité à l'enracinement, hérité de *M. rotundifolia* (Davidis et Olmo, 1964).



*Muscadinia rotundifolia* 1n = 20

*Vitis vinifera* 1n = 19



1. Feuilles

2. Baies ; Grappe

**Figure 3. Différences morphologiques entre les genres *Vitis* et *Muscadinia***  
<http://www.volkoomen.nl/v,w,x,y,z/vitis.htm> ; Franz Eugen Köhler, *Köhler's Medizinal-Pflanzen*, 1897

**Tableau 2. Différences entre les sous-genres *Euvitis* et *Muscadinia***

	<i>Euvitis</i>	<i>Muscadinia</i>
Fruits	Epicarpe fin Nombreux fruits par grappe	Epicarpe épais Peu de fruits par grappe
Graines	Ovoïdes, Chalaze lisse	Allongées, Chalaze plissée
Feuilles	Feuilles rondes avec des dents triangulaires	5 lobes principaux plus ou moins découpés
Lenticelles	Absentes	Présentes
Vrilles	Fourchues	Simple
Moelle	Discontinue	Continue
Phellogène	Profond	Sous-épidermique
Phloème	Tangentiel	Radial
Nombre de chromosomes	2n = 38	2n = 40
Nombre d'espèces	60	3

### 1.1.3.3. Histoire et domestication de la vigne

Bien avant sa domestication, la vigne sauvage était exploitée par l'Homme pour la consommation direct des baies de raisin (Mc Govern, 2003). L'Épopée de Gilgamesh, récit akkadien vieux de 4 000 ans parle déjà du vin, tandis que la Bible fait de Noé le premier viticulteur (Gn 9, 20-21 – Traduction œcuménique de la Bible ; Figure 4). La domestication de la vigne serait apparue au cours du Néolithique (5000 à 6000 ans av. JC.) lorsque l'Homme se sédentarisait et développait l'agriculture dans les régions du Sud du Caucase et du croissant fertile (Levadoux, 1956 ; Zohary & Hopf, 2000 ; McGovern, 2003).



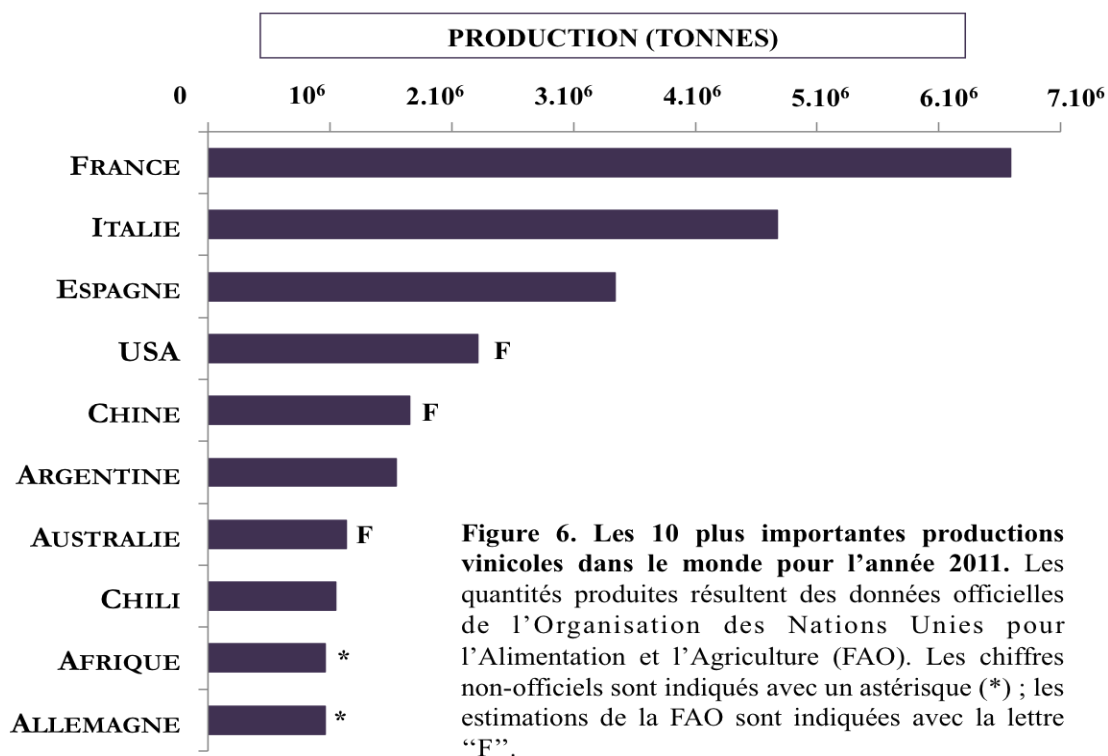
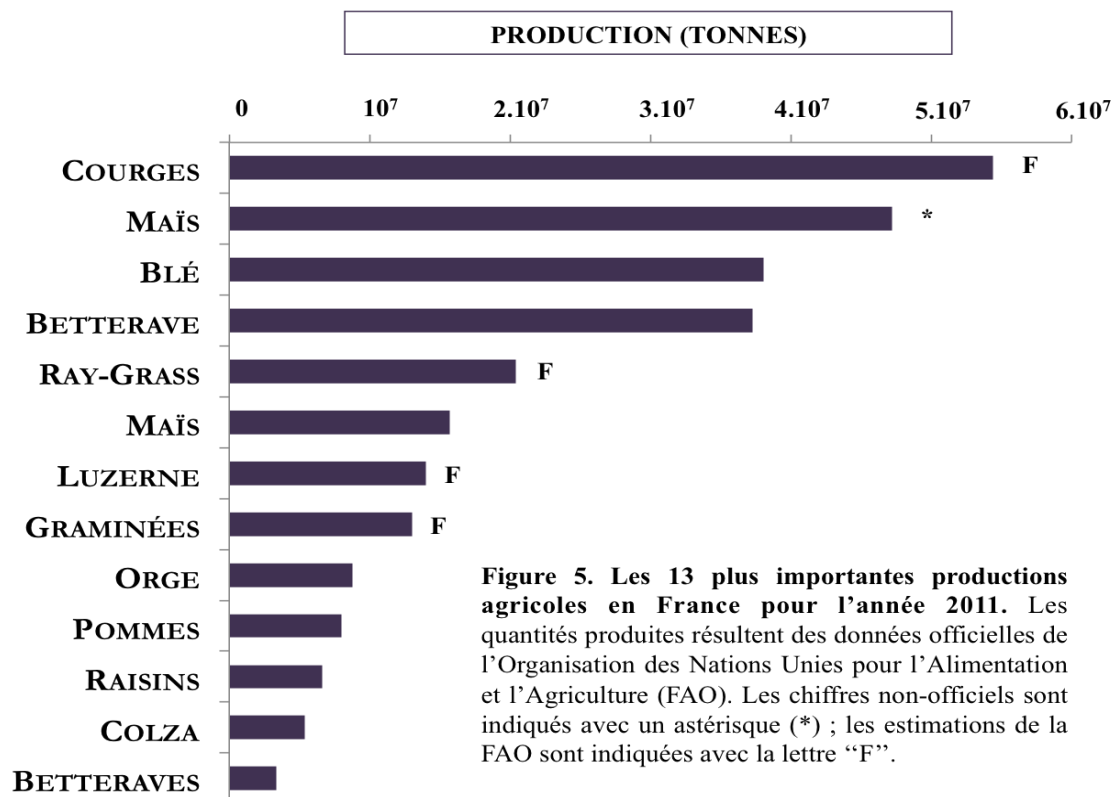
**Figure 4. L'ivresse du patriarche Noé, par Giovanni Bellini vers 1515**  
Tableau exposé au Musée des beaux-arts et d'archéologie de Besançon

A partir de ces régions, la viticulture va s'étendre au fil des civilisations autour du bassin méditerranéen et au Moyen-Orient. Vers l'Est, en Asie, ce sont les Perses et les Phéniciens qui propageront la vigne. Vers l'Ouest, la viticulture s'implante en Egypte puis en Grèce et en Italie où les Romains la diffuseront à l'ensemble de la Méditerranée. En Gaule, les premiers vignobles sont implantés par les phocéens à Massalia (Marseille, 600 av. JC ; Dion, 1982). Les Romains vont ensuite planter la viticulture dans le Languedoc et le couloir Rhodanien (125 av. JC.) avec le cépage *Allobrogica*, éventuel ancêtre du cépage Syrah ou encore le *Biturca* celte, hypothétique ancêtre du Cabernet franc (Levadoux, 1956). Ces données indiquent que les Romains pratiquaient déjà une sélection des meilleurs cépages qu'ils multipliaient de façon végétative tout en maintenant dans certains cas une multiplication sexuée. Tout au long de leur histoire, la vigne et le vin ont été associés à des considérations symboliques et religieuses. Dans l'Egypte ancienne, la vigne et le vin auraient été apportés par Osiris, chez les Grecs par Dionysos et chez les Romains par Bacchus. C'est ensuite la religion Chrétienne, en associant le vin au sang du Christ, qui sera très impliquée dans la diffusion de la viticulture. Ainsi portée par le christianisme, la viticulture va ensuite petit à petit se répandre dans le monde entier. Les vignobles du Nouveau Monde (Amérique, Australie, Afrique du Sud, Nouvelle Zélande) ont été établis dans un premier temps par les missionnaires qui ont apporté des pépins issus de croisements (Boursiquot & This, 2000 ; This *et al.* 2006) puis par les colons chrétiens durant les XVII et XVIIIème siècles avec des boutures (Royer, 1988). Aujourd'hui la vigne est cultivée sur les 5 continents et vin et viticulture restent chargés de nombreux symboles sociaux, culturels et gastronomiques. C'est tout particulièrement vrai en France avec par exemple l'inscription de Saint-Emilion au patrimoine Mondial de L'UNESCO en 1999.

Au cours de la domestication, les hommes ont sélectionné des caractères agronomiques leur permettant d'accroître le rendement et la qualité du fruit. On suppose qu'un des premiers caractères à avoir été sélectionné est l'hermaphrodisme permettant d'augmenter significativement le rendement et la régularité de production (This *et al.* 2006). D'autres caractères comme la vigueur, la tolérance, la taille ou la couleur des baies, ont été au fil des générations sélectionnés par l'Homme. Les meilleurs cépages sélectionnés ont été par la suite maintenus et conservés par reproduction végétative.

#### **1.1.1.4. La viticulture mondiale**

La vigne est cultivée à travers le monde pour ses fruits qui sont des baies à graines albuminées. Ces dernières permettent la préparation de jus de raisin, la fabrication de vins par fermentation alcoolique ou de vinaigre par acétification, la distillation du vin en liqueurs (armagnac, cognac, porto) ou peuvent être consommées comme fruit frais (raisin de table) ou séchés (raisins secs). La vigne fait partir des **30** espèces (25<sup>ème</sup> rang ; statistique de la FAO : Organisation des Nations Unies pour l'Alimentation et l'Agriculture, 2013) les plus cultivées dans le monde. Le vignoble mondial s'étend sur les cinq continents et sa surface est de **7 086 022** d'hectares. La majorité des surfaces viticoles mondiales sont situées en Europe (51,9 %), le reste étant réparti entre l'Asie (26,5 %), l'Amérique (13,7 %), l'Afrique (5,1 %) et l'Océanie (2,8 %) (FAO, 2013). La production annuelle mondiale de vin est estimée à **28 674 361** tonnes, soit 292 millions d'hl (FAO, 2013). L'Europe, avec environ 66 % de cette production occupe une place prépondérante (FAO, 2013). En France, **764 164** ha étaient des zones consacrées à la viticulture en 2011, pour une production de **6 590 810** tonnes (Figure 5). C'est la 11<sup>ème</sup> production agricole en France pour l'année 2011. On y distingue 14 grandes régions viticoles: Alsace, Beaujolais, Bordelais, Bourgogne, Champagne, Jura, Savoie, Languedoc, Roussillon, Provence, Corse, Sud-Ouest, vallée de la Loire, vallée du Rhône. La France a repris sa place de premier producteur de vin au monde en 2011 (Figure 6) une position qu'elle avait perdue en 2009 au bénéfice de l'Italie. Les exportations françaises de vins en 2012 sont en croissance par rapport à l'année 2011 ; et celles-ci ont atteint **14,98** millions d'hectolitres (soit + **5** % par rapport à 2011) pour un solde positif de **7,83** milliards d'euros en 2012 (soit + **9** % par rapport à 2011).



### 1.1.2. Cycle de développement de la vigne

Chaque année, la vigne effectue au cours de son développement deux cycles en concurrence trophique. Le cycle végétatif assure la pérennité du cep grâce au développement de l'appareil végétatif. Le cycle reproducteur permet la formation des fleurs, puis des fruits.

#### 1.1.2.1. Le cycle végétatif

De novembre à février, la vigne entre dans une période de **repos hivernal ou dormance**. La sève ne circule plus dans la plante. Au début du printemps, la reprise de l'activité végétative se manifeste par les 'pleurs', correspondant à l'écoulement de la sève brute par les plaies résultant de la taille.

En mars/avril, c'est le **débourrement** : les bourgeons commencent à se développer. C'est la période de croissance des rameaux et des feuilles. La phase de croissance se poursuit jusqu'en fin juillet à début août. Elle se caractérise par l'allongement des rameaux, l'étalement et la croissance des feuilles. L'aoûtement débute au mois d'août et se poursuit jusqu'en novembre. Il s'accompagne d'une accumulation de réserves en amidon et en lignine dans les sarments, ce qui favorise la résistance des tissus au froid et pour permettre le début d'un nouveau cycle au printemps suivant (Huglin et Schneider, 1998 ; Galet, 2000).

#### 1.1.2.2. Le cycle reproducteur

Les vignes destinées à la production de raisin sont hermaphrodites alors que celles destinées à la production de porte-greffes sont la plupart du temps dioïques. La différenciation des fleurs commence au printemps, lorsque le bourgeon reprend son activité. Peu après le débourrement, les inflorescences se développent. La **floraison** a lieu entre mai et juin, et aboutit à la libération du pollen. La pollinisation est généralement autogame.

Dès que l'ovaire commence à se développer, on dit qu'il est noué, et l'ensemble de ce phénomène pour une grappe s'appelle la **nouaison**. A partir de l'instant où il est formé, le fruit ne cessera de grossir. Cette croissance est due à des phénomènes de divisions et de grandissement cellulaires. La baie de raisin subit une très forte augmentation de taille qui est corrélée à une accumulation d'eau et d'acides organiques (notamment des acides tartriques et maliques) dans la vacuole des cellules. La concentration en sucres reste très faible (Schlosser *et al.* 2008).

En août, c'est la **véraison** ; les baies de raisins verts changent de couleur (phénomène lié à l'accumulation de pigments) et entame une phase de **maturation** qui résulte exclusivement de grandissement cellulaire. Les changements métaboliques engagés à la véraison permettent l'accumulation progressive de sucres, de composés phénoliques (pigments et tanins), de composés aromatiques et de précurseurs aromatiques et de protéines antifongiques (notamment de type PR, *Pathogenesis Related*) (Benharbit, 1995). Parallèlement, les acides organiques sont dégradés (Ollat, 1997). Tous ces composés détermineront la qualité finale du fruit et éventuellement du vin.

En septembre/octobre, c'est l'époque des vendanges. Les raisins sont récoltés soit manuellement, soit mécaniquement avec une machine à vendanger. La chute des feuilles marque le début de la période de repos végétatif de la vigne.

Les possibilités d'avoir des conditions propices au développement de maladies s'étendent du débourrement à la récolte. La gestion de certaines d'entre elles a nécessité des changements majeurs dans les pratiques viticoles au cours du temps : le greffage et l'introduction des traitements phytosanitaires.

## **1.2. LE CONTRÔLE PHYTOSANITAIRE AU VIGNOBLE**

La viticulture française était dans une phase de prospérité, lorsque l'arrivée de l'oïdium puis du phylloxéra et du mildiou dans les années 1850 - 1870, a modifié durablement la viticulture européenne (Legros, 1997). En conséquence, une réduction de la diversité est probablement survenue pour des espèces cultivées et sauvages. La diversité de l'espèce *V. vinifera*, trouvée aujourd'hui, n'est probablement que le reflet de ce qui existait avant l'invasion des maladies. En effet, depuis la fin du XVIII<sup>ème</sup> siècle, certaines vignes nord-américaines, tolérantes à l'agent pathogène de l'oïdium, *Erysiphe necator*, avaient été introduites dans de nombreuses régions de France par des botanistes et des pépiniéristes et ce, sans aucune précaution sanitaire. Au début des années 1850, cette maladie a entraîné la baisse de la production vinicole : celle-ci est passée de 39 millions d'hectolitres en 1851 à 11 millions d'hectolitres en 1854 (Legros, 1997). Des traitements à base de soufre s'avérèrent très efficaces et leur généralisation permit une gestion de l'épidémie. A la même époque (à partir de 1861), des cas de dépérissement des vignes ont également été signalés d'abord dans le Gard (Languedoc-Roussillon), puis très rapidement en Provence, puis dans le Sud-ouest, dévastant ainsi progressivement les vignobles. Le phylloxéra

(*Phylloxera vastatrix* : puceron qui attaque principalement les racines de la vigne), responsable de ce dépérissement, a été identifié en 1868. Le vignoble français a régressé de 28 % entre 1875 et 1903, perdant ainsi 660 000 ha. Cette maladie a causé la disparition définitive de nombreux cépages anciens et probablement de l'essentiel des populations de vignes sauvages (Legros, 1993 ; This *et al.* 2006). Par ailleurs, en 1872 des craintes étaient exprimées quant à la possible introduction du mildiou par ces porte-greffes américains. En 1878, le premier cas de mildiou a été détecté près de Libourne (Bordelais) jusqu'en 1892, avec des dégâts peu importants. En 1893 que le mildiou a dévasté 50 % de la récolte. Dès lors, des années "à mildiou" et des années "sans mildiou" se succédèrent selon les conditions climatiques. A l'heure actuelle, outre le *phylloxera* qui est parfaitement enrayeré grâce à l'utilisation des porte-greffes, les deux maladies cryptogamiques qui touchent sévèrement les vignobles sont le mildiou et l'oïdium. Ce sont ces trois maladies qui ont motivé les débuts de l'amélioration moderne de la vigne pour la création variétale de porte-greffes résistants au phylloxera et adaptés aux sols souvent calcaires des vignobles mais également à la création d'hybrides producteurs directs résistants au mildiou et à l'oïdium.

Plusieurs stratégies de lutte, telles que des traitements au sulfure de carbone ou des submersions totales des sols pendant au moins 40 jours, ont été tentés pour lutter contre le phylloxéra, sans beaucoup de succès. Les traitements au sulfure de carbone se sont avérés coûteux et toxiques et l'immersion n'a pu être pratiquée que dans des zones très limitées. En 1874, il a été montré que les espèces américaines *Vitis rupestris* et *Vitis riparia* étaient résistantes au phylloxéra. Deux nouvelles approches ont été développées pour lutter contre le phylloxera. La première méthode a consisté à greffer des cépages européens sur des plants résistants d'origine américaine. Cela a eu pour avantage de combiner la qualité de récolte des cépages traditionnels européens à la résistance racinaire apportée par le porte-greffe américain (Pouget, 1990). La deuxième approche a consisté à effectuer des croisements interspécifiques entre espèces américaines et européennes pour obtenir des variétés utilisées franc de pied (hybrides producteurs directs dans le jargon viticole). Ces hybrides interspécifiques ont été largement utilisés jusqu'au milieu du XX<sup>ème</sup> siècle mais ils sont maintenant assez rares. En effet, la qualité de la récolte était généralement médiocre (Pouget, 1990). Les scandales des vins frelatés dans les années 1900-1930 provenant des



hybrides ont également fortement entaché la réputation de ces derniers avec pour conséquence, en 1934, l'interdiction de six hybrides producteurs directs (Noah, Isabelle, Clinton, Herbemont, Othello et Jacquez). Par la suite, sous l'influence des politiques viticoles nationales et européennes, les surfaces cultivées d'hybrides ont fortement régressé depuis 1958. Cependant une vingtaine d'hybrides restent inscrits au catalogue officiel français (Boursiquot *et al.* 2007).

Si les résistances au *phylloxera* déployées sont stables et durables depuis plus de 100 ans maintenant, la création de variétés résistantes au mildiou et à l'oïdium reste un challenge. La sélection pour une durabilité des résistances, combinée aux critères d'excellence qualitative, demande de nombreuses générations de croisements et plusieurs années d'évaluation. En outre, sélectionner pour une durabilité des résistances nécessite de bien comprendre les populations d'agent pathogènes, leur origine, leur diversité et leurs capacités d'adaptation. Ces deux maladies cryptogamiques seront l'objet des paragraphes suivants.

### **1.2.1. L'oïdium de la vigne**

Les oïdiums sont causés par des champignons ascomycètes et parasites biotrophes obligatoires des genres *Blumeria* mais surtout *Erysiphe* (Famille des *Erysiphaceae*) et affectent de nombreuses plantes cultivées économiquement importantes : vignes, arbres fruitier, céréales à paille, houblon, nombreuses plantes ornementales. Ils sont souvent visibles en raison d'une production abondante de conidies (Figure 7). Ils infectent les feuilles, les tiges, les fleurs et les fruits de près de 10 000 espèces d'angiospermes (Braun *et al.* 2002). L'oïdium de la vigne cause par *E. necator*, anciennement nommé *Uncinula necator*, et exclusif au genre *Vitis*, a été décrit dans les vignes d'Amérique du nord en 1834, avant sa découverte en Europe en 1845 (Large, 1940). Après son introduction en Europe de l'Est de l'Amérique du Nord (Spencer, 1978), l'oïdium de la vigne a été observé dans toutes les régions viticoles du monde, dont la Californie en 1859 (Smith, 1961) et Australie, en 1866 (Emmett *et al.* 1990).



**Figure 7. Les symptômes et les signes représentatifs de l'oïdium.** (A) *Erysiphe alphitoides* chez le chêne de Garry (*Quercus garryana*). (B) *Podosphaera fuliginea* chez le kalanchoë de Blossfeld (*Kalanchoe blossfeldiana*). (C) *Erysiphe necator* chez *Vitis vinifera*

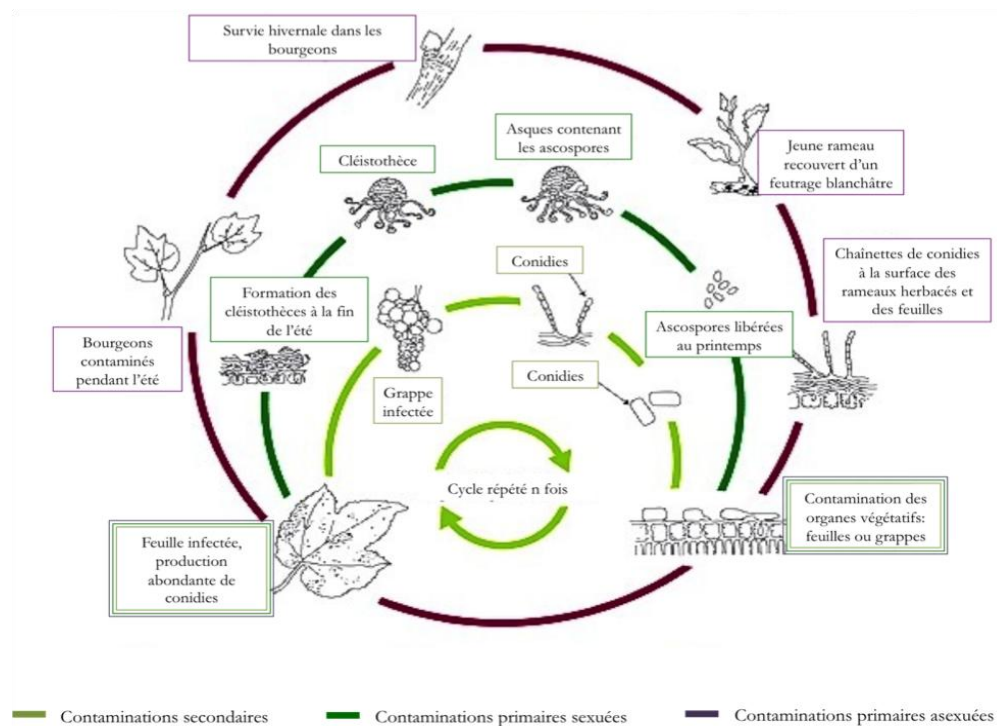
### 1.2.1.1. Cycle de développement

Le champignon *E. necator* colonise la surface de tous les organes verts de la vigne (Corio-Costet, 2007), notamment la face supérieure des feuilles, et se propage sur les baies. Durant l'hiver, le champignon se conserve sous deux formes (Figure 8) :

Forme **sexuée** ou phase sporophytique (la plus répandue) durant laquelle l'oïdium se conserve essentiellement sous la forme de cléistothèces qui contiennent des ascospores. Ces derniers se forment sur tous les organes très atteints par la maladie mais plus généralement sur les feuilles à partir de la fin de l'été et mûrissent à l'automne. La majeure partie des cléistothèces est lessivée par les pluies d'automne, seuls ceux restant sur l'écorce des cepes sont viables et formeront l'inoculum primaire. A maturité, au printemps, ils libèrent des spores (ascospores) contenues dans des asques qui sont expulsées dès que les conditions sont favorables.

Forme **asexuée** ou phase gamétophytique conduisant à la formation de conidiophores portant des conidies : Le mycélium se conserve pendant l'hiver entre les écailles des bourgeons infectés au

cours de la saison précédente. Au printemps, les filaments mycéliens se développent et contaminent très précocement les jeunes pousses en croissance qu'ils parasitent *via* des suçoirs et forment des conidies qui assurent la dissémination de la maladie. En effet, le champignon émet un pédoncule, à partir duquel se différencie par invagination de la membrane cellulaire de la plante un haustorium globuleux, grâce auquel l'agent pathogène puise ses nutriments pour se reproduire (Corio-Costet, 2007). C'est le mode de contamination primaire de la forme asexuée. Par la suite, du printemps à l'automne, de nombreux cycles de contaminations secondaires se répètent (Rumbolz *et al.* 1999). Les conidies sont libérées suite à un choc (vent, traitements, rognage, pluie violente) puis germent et forment un nouveau mycélium. La durée du cycle est variable et peut aller de 5 à 6 jours entre 23 et 30°C jusqu'à 32 jours à 7°C. La maladie est polycyclique et a un développement épidémique qui est assuré par des cycles successifs de contamination/infection. L'oïdium ne nécessite pas d'eau pour germer. Cependant, le développement des conidies nécessite une hygrométrie élevée. La température optimale pour la germination des conidies se situe entre 20 et 25°C (Corio-Costet, 2007).



**Figure 8. Cycle de développement de *Erysiphe necator*.**

### 1.2.1.2. Diversité et évolution de la maladie

L'Amérique du Nord est le centre de l'origine de nombreuses espèces sauvages de *Vitis* qui ont des niveaux relativement élevés de résistance à de nombreuses maladies et aux ravageurs de la vigne, y compris l'oïdium (Lepik, 1970 ; Lenné & Wood, 1991). À l'exception des vignes du genre *Muscadinia*, il n'y a aucune différenciation génétique entre les populations d'*E. necator* en fonction des espèces hôtes du genre *Vitis*. L'agent pathogène de l'oïdium le mieux étudié, *Blumeria graminis*, autrefois aussi appelé *Erysiphe graminis*, montre une spécificité de genres hôtes plutôt qu'une spécificité d'espèces dans un genre (Wyand & Brown, 2003 ; Inuma *et al.* 2007). Côté vigne, *E. necator* semble se comporter comme *Blumeria graminis*, avec une spécialisation incomplète au niveau des genres hôtes. Gadoury et Pearson (1991) ont en effet montré que des souches d'*E. necator* var. *ampelopsidis* prélevées chez *Parthenocissus quinquefolia* n'étaient que rarement pathogène(s) sur les espèces du genre *Vitis*, et avec une très faible agressivité. Brewer & Milgroom (2010) ont montré une absence de spécificité d'hôte parmi les populations d'agents pathogènes se développant sur des hôtes de type *Vitis* aux États-Unis. En revanche, Bien que les populations d'*E. necator* retrouvées sur *M. rotundifolia* sont distinctes de celles retrouvées sur les autres espèces de *Vitis*, par une ou deux mutations, elles présentent en outre une différence phénotypique forte avec les autres types de souches (Brewer & Milgroom, 2010). Les auteurs ont émis l'hypothèse soit d'un stade précoce de spéciation de ces populations résultant d'une spécialisation d'hôte ; soit les isolats qui infectent *M. rotundifolia* sont capables d'éviter la reconnaissance de l'agent pathogène par les défenses de l'hôte lors d'une interaction gène pour gène (Brewer & Milgroom, 2010). De même, Frenkel *et al.* (2010) ont démontré que tous les isolats sont pathogènes pour l'espèce *Vitis vinifera*, quel que soit l'hôte source. Il n'y avait aucune différence d'agressivité et de taille de lésions parmi les isolats, provenant des différents hôtes, qui ont servi à inoculer *V. vinifera*. Ce manque de spécialisation pourrait être expliqué par la colonisation récente des hôtes du genre *Vitis* par *E. necator* ou la récente diversification d'espèce *Vitis* en Amérique du Nord. Chez les espèces étroitement apparentées ou des populations qui subissent la spéciation, la divergence génétique ne peut être évidente qu'à un ou quelques *loci* impliqués dans l'adaptation (Dopman *et al.* 2005). Les populations de *E. necator* se développant sur les espèces du sous-genre *Muscadinia* ou du genre *Parthenocissus* pourraient être dans les premiers stades d'une spécialisation d'hôte et donc d'une spéciation. Alternativement, ces isolats pourraient simplement présenter des nouveaux gènes de

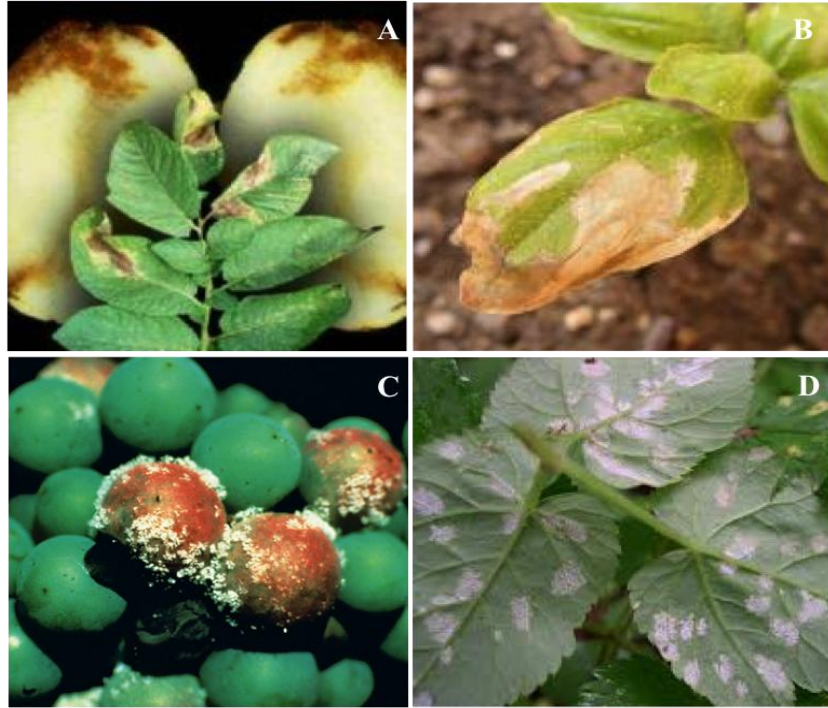
virulence qui leur permette de contourner les défenses des espèces du genre *Muscadinia* ou du genre *Parthenocissus*.

Les populations d'*E. necator* d'Europe et d'Australie sont deux groupes génétiquement distincts, A et B, identifiés avec des marqueurs génétiques de type RAPD et AFLP, ou avec des marqueurs de polymorphismes nucléotidiques de type SCAR ou SNP (Délye *et al.* 1997 ; Evans *et al.* 1997 ; Stummer *et al.* 2000 ; Miazzi *et al.* 2003 ; Péros *et al.* 2005 ; Núñez *et al.* 2006). Les deux groupes A et B peuvent être présents sur une même parcelle, sous formes de mycélium dormant ou de cléistothèces (B en France). Les populations d'oïdium du groupe A présentent peu de diversité et la plupart des isolats possèdent le même type sexuel (Corio-Costet, 2007) alors que les souches du groupe B semblent se reproduire de manière sexuée (Péros *et al.* 2005) et présentent une plus grande variabilité génétique (Corio-Costet, 2007). Toutefois, quelques études réalisées en Australie et en Italie suggèrent que les souches du groupe A peuvent se reproduire de manière sexuée (Miazzi *et al.* 2003 ; Stummer *et al.* 2006 ; Frenkel *et al.* 2010), mais les recombinants n'ont jamais été trouvés *in natura*. Par ailleurs, ces deux populations d'oïdium possèdent des exigences écologiques différentes (Délye et Corio-Costet, 1997 et Amrani et Corio-Costet, 2006) et les populations rencontrées dans la plupart des vignobles en Italie et en France appartiennent uniquement au groupe B (Délye *et al.* 1997 ; Miazzi *et al.* 2008). Lorsque les deux groupes génétiques se trouvent dans le même vignoble en début de saison, seul le groupe génétique B est retrouvé en fin de saison (Corio-Costet, 2007 ; Montarry *et al.* 2009). La présence conjointe de ces deux populations dans ce cas a conduit à des erreurs d'interprétations de la résistance aux fongicides (Corio-Costet, 2007). En effet, les populations A étant plus sensibles aux fongicides que les populations B, leur présence massive en début de saison, puis leur remplacement en cours de saison, a conduit à décrire une perte de sensibilité au fongicide, quelque soit la dose appliquée. Ce qui n'était pas le cas. Corio-Costet (2007) et Dufour *et al.* (2011) ont montré que les deux différentes populations exprimaient temporellement des niveaux de sensibilité différents aux fongicides inhibiteurs de la biosynthèse des stérols. L'étude menée par Brewer & Milgroom (2010) a montré une plus grande diversité sur le continent américain. En effet, les auteurs ont séquencés 146 isolats de *E. necator* de diverses espèces de *Vitis* sauvages et cultivées recueillies dans l'est des Etats-Unis (nord, sud et centre) et des espèces de *V. vinifera* cultivées de l'Ouest des Etats-Unis, d'Europe et d'Australie. Les isolats de *E.*

*necator* var. *Ampelopsidis* (Braun *et al.* 2000) de *Parthenocissus quinquefolia* ont été utilisés comme groupe externe lors d'une analyse phylogéographique. Les résultats de l'analyse ont montré que les haplotypes ancestraux identifiés étaient issus de l'est des Etats-Unis, dans tous les cas.

### **1.2.2. Le mildiou de la vigne**

Le mildiou de la vigne est causé par le oomycète, *Plasmopara viticola*, appartenant à la famille des *Peronosporaceae*, une petite lignée avec moins de 1 000 espèces qui inclut des parasites parmi les plus dévastateurs pour les plantes et les animaux, comme ceux appartenant aux genres *Phytophthora*, *Peronospora*, *Bremia* ou *Albugo* (Kirk *et al.* 2001 ; Kamoun, 2003). Les mildious peuvent être des biotrophes obligatoires, comme *P. viticola*, ou des hémibiotrophes, comme le mildiou de la tomate et de la pomme de terre *Phytophthora infestans*. Ces derniers sont biotrophes dans leur phase parasite initiale, puis passent ensuite en mode de vie nécrotique en tuant les tissus hôtes pour se répandre (Deacon, 2006). Les mildious constituent un groupe important qui se développe sur tous les organes vivants de la plante. Cette maladie affecte de nombreuses espèces de plantes mais prenant des proportions épidémiques dans certaines cultures qui ont un grand impact économique tels que *Plasmopara viticola* de la vigne *Vitis vinifera*, *Pseudoperonospora humuli* du houblon *Humulus lupulus*, *Peronospora tabacina* du tabac *Nicotiana* (Figure 9). Il provoque des défoliations, ainsi que le brunissement et l'assèchement des baies et des tiges. La sensibilité des plantes au mildiou dépend de l'interaction de plusieurs variables plus ou moins favorables à l'agent pathogène, telles que le climat et l'état physiologique de la plante hôte.



**Figure 9. Les symptômes et les signes représentatifs du mildiou.** (A) *Phytophthora infestans* chez la pomme de terre (*Solanum tuberosum*). (B) *Peronospora belbahrii* chez le basilic (*Ocimum basilicum*). (C) *Plasmopara viticola* chez la vigne (*Vitis vinifera*). (D) *Pseudoperonospora humuli* chez le houblon (*Humulus lupulus*).

### 1.2.2.1. Cycle de développement

*P. viticola* est un agent pathogène hétérothallique diploïde. Son cycle de développement comprend une phase sexuée qui produit des oospores permettant le passage de l'hiver et les infections primaires ; et une phase asexuée qui conduit à la production de spores nécessaires aux infections secondaires et à la dispersion de l'agent pathogène sur une courte distance (Figure 10 ; Galet, 1979) :

*P. viticola* se conserve en hiver sous forme d'oospores sur les feuilles mortes tombées à terre, au niveau des parties nécrosées. Au printemps, ces oospores germent dès que les conditions d'humidité sont favorables et que la température atteint 11°C (Lafon & Clerjeau, 1988). Cette germination donne naissance à des macroconidies qui émettent elles-mêmes de nombreuses



spores. Ces dernières, après formation de filaments mycéliens capables de pénétrer entre les cellules du tissu foliaire, provoquent la contamination primaire (Latijnhouwers *et al.* 2003). La pénétration du champignon dans le tissu foliaire reste invisible à l'œil nu. Au bout de quelques jours, apparaissent des conidies à la face inférieure des feuilles, visibles sous forme de tâches blanches correspondant à des tâches d'huile à la face supérieure de la feuille.

Ces conidies émettent de nouvelles spores qui propagent la maladie sur le même cep ou sur les ceps voisins, donnant lieu à des contaminations secondaires : c'est la phase d'invasion. Ces cycles secondaires peuvent être très courts (5-6 jours) quand les conditions sont favorables. La température optimale pour la prolifération de l'agent pathogène est de 20 à 25°C environ (Lafon & Clerjeau, 1988). Contrairement à l'oïdium, un temps pluvieux, une atmosphère humide et une température douce constituent les conditions les plus favorables au mildiou. Une forte rosée matinale est un facteur aggravant si les contaminations primaires ont eu lieu.

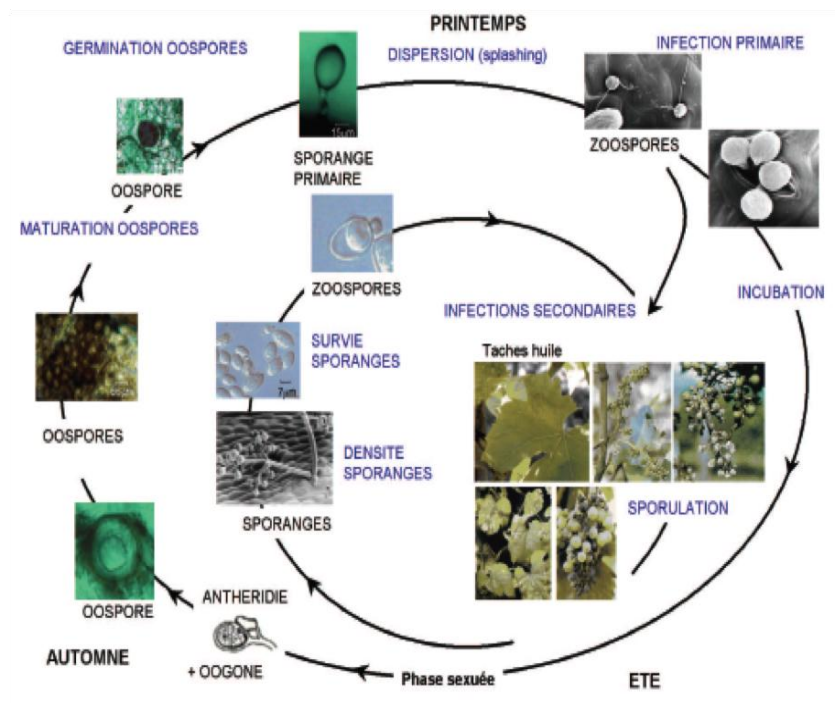


Figure 10. Cycle de développement de *Plasmopara viticola*. D'après Olivier Viret, ACW



### 1.2.2.2. Diversité et évolution de la maladie

Tous les cultivars de *V. vinifera* sont sensibles au mildiou, bien que des variations de sensibilité sont observées chez les cultivars (Boubals, 1959 ; Dubos, 2002). En revanche, de nombreuses sources de résistance au mildiou ont été décrites dans des espèces apparentées à *V. vinifera* (Boubals, 1959 ; Staudt et Kassemeyer, 1995 ; Dai *et al.* 1995 ; Brown *et al.* 1999 ; Kortekamp et Zyprian, 2003 ; Unger *et al.* 2007 ; Cadle-Davidson, 2008), dont certaines ont déjà été introduite avec succès dans la vigne pour créer des cultivars résistants (Csizmazia et Bereznoi, 1968 ; Eibach et Töpfer, 2003 ; Cadle-Davidson, 2008 ; Merdinoglu *et al.* 2009). Les espèces *Vitis* américaines et asiatiques appartenant au sous-genre *Euvinis* ou au sous-genre *Muscadinia* montrent des niveaux variables de résistance à *P. viticola*, allant de la résistance modérée, par exemple chez *V. rupestris*, à une haute résistance, par exemple en *V. rubra*, *candicans*, *V. amurensis*, *V. riparia*, *V. cinerea* ou *M. rotundifolia* (Boubals, 1959 ; Olmo, 1971 ; Staudt et Kassemeyer, 1995 ; Cadle-Davidson, 2008).

Les populations de *P. viticola* nord-américaines sont considérées comme des populations fondatrices. Des analyses phylogénétiques sur un petit nombre d'isolats ont mis en évidence l'existence de trois lignées distinctes. L'une d'entre elles contient des échantillons d'origine américaine et européenne tandis que les deux autres lignées contiennent uniquement des isolats américains (Chen *et al.* 2007 ; Schröder *et al.* 2011). Les populations européennes de *P. viticola* présentent à la fois un faible niveau de diversité au sein des gènes et une forte diversité génotypique. En général, les populations sous régime de reproduction sexuée montrent des niveaux de diversité génotypique plus importants que les populations strictement asexuées. Par exemple, Delmotte *et al.* (2002) ont montré que les populations sexuées du puceron *Rhopalosiphum padi* présentaient une plus grande diversité génotypique et génique que les populations asexuées de ce même puceron. La plupart des populations européennes peuvent se croiser entre elles, ce qui indique une reproduction sexuée fréquente. Elles présentent également un léger isolement par la distance à l'échelle du continent (Gobbin *et al.* 2006). Les populations de mildiou trouvées en France présentent une grande variabilité phénotypique avec des sensibilités différentes à plusieurs fongicides (Corio-Costet *et al.* 2012). Une étude phylogénétique réalisée à partir d'une portion du génome mitochondriale (2 281 pb) incluant le

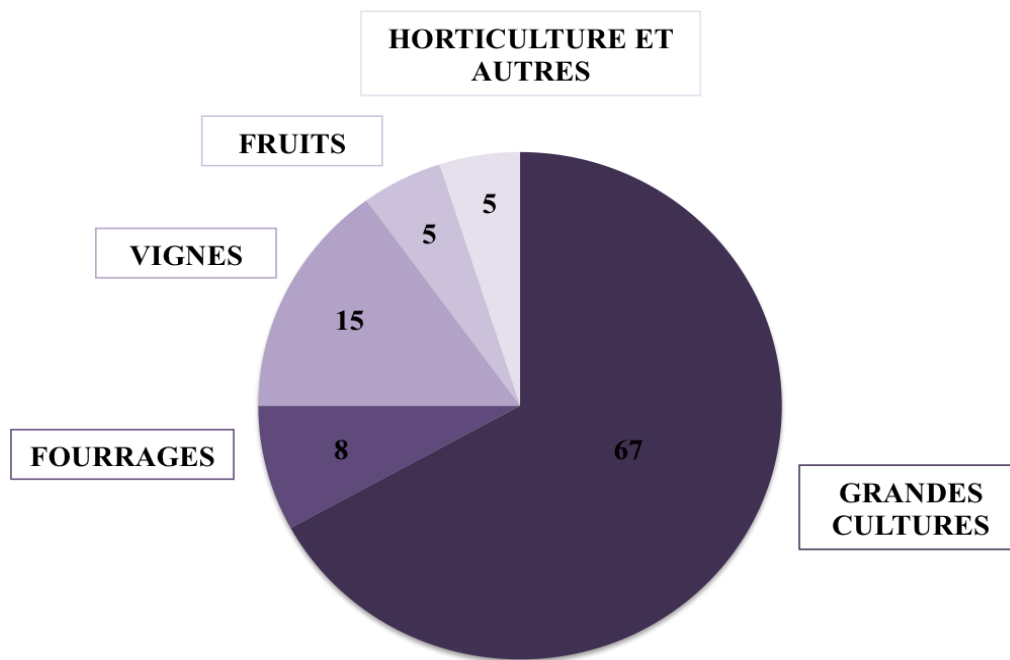
gène du cytochrome b a permis de démontrer l'existence de différents haplotypes mitochondriaux plus ou moins sensibles à certains fongicides (IS, IR, IIS, IIR) (Chen *et al.* 2007 ; Corio-Costet *et al.* 2011). En effet, Corio-Costet *et al.* (2011) ont noté une fréquence des isolats résistants de 0 % à 23,25 %, avec une moyenne de 4,64 % parmi les populations de *P. viticola* étudiées entre les années 2000 et 2004. Pour évaluer l'aptitude des isolats sensibles et résistants, une comparaison des différents paramètres biologiques, y compris la période de latence, la production de spores et la fréquence des infections a été réalisée. Les isolats résistants ont montré une plus grande fréquence d'infection que les isolats sensibles, alors qu'aucune différence significative n'a été trouvée dans la capacité de la sporulation et la période de latence entre les isolats sensibles et résistants. Pour approfondir la compétitivité entre les isolats, un essai incluant deux différentes proportions d'isolats résistants avec un isolat sensible a été mené sur huit cycles de croissance asexuée, en l'absence de fongicide QoI (*Quinone outside Inhibitor*). La compétitivité des isolats résistants variait en fonction de leurs paramètres de remise en forme, ce qui suggère qu'il n'y a pas de coût notable de la résistance QoI dans des conditions contrôlées dans *P. viticola*.

### **1.2.3. Réduction de l'usage des produits phytosanitaires**

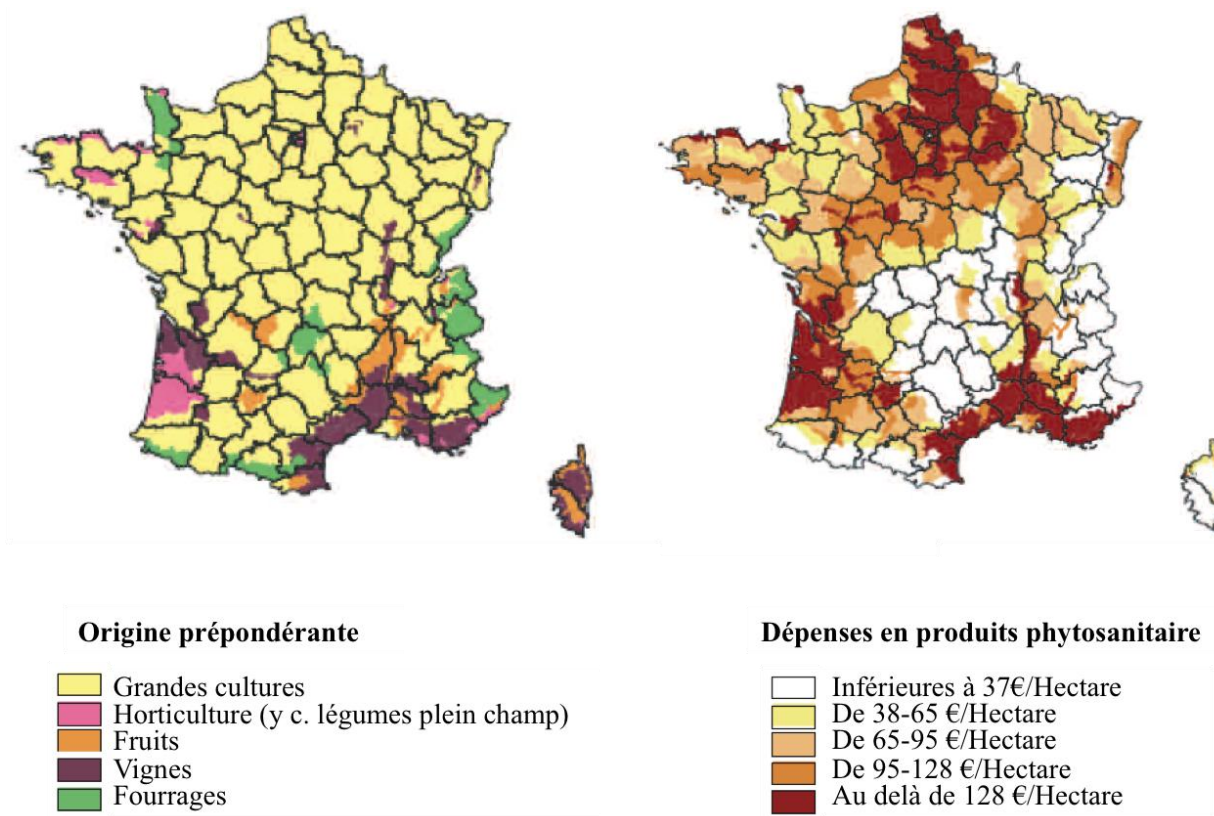
L'Union Européenne s'est engagée à réduire l'emploi des pesticides dans l'agriculture. En France, un plan interministériel de réduction des risques liés aux pesticides a été mis en place en juin 2006 ; et le Grenelle de l'environnement a confirmé les orientations. L'un des principaux objectifs de ce plan est la réduction de moitié de l'emploi de pesticides de synthèse (plan ECOPHYTO 2018). Pour atteindre cet objectif, une étude sur la faisabilité de cet objectif et des scénarios y conduisant été menées par l'INRA et le Cemagref (Aubertot *et al.* 2004 ; Butault *et al.* 2010). Un état des lieux de l'utilisation des pesticides en France a ainsi été réalisé en se basant sur des données d'enquêtes sur les pratiques culturales des grandes cultures (Guichard *et al.* 2009), de la viticulture (Gary *et al.* 2009), de la culture des légumes (Pitrat *et al.* 2009) et des fruits (Sauphanor *et al.* 2009) menées en 2006. Les dépenses en pesticides dans l'agriculture professionnelle se montaient ainsi à **2,3** milliards d'€ pour la France, en 2006 (Figure 11 ; Butault *et al.* 2010). Elles correspondaient entre autres pour **67** % (1 564 millions d'€) aux grandes cultures (hors légumes de plein champ) et pour **15** % (322 millions d'€) aux vignes.

Cependant, si on tient compte de la SAU concernée par ces deux types de culture en calculant

par exemple le nombre moyen de doses homologuées utilisées par hectare sur une campagne (Indice de Fréquence de Traitement ou IFT), la viticulture présente une fréquence de traitement bien plus importante (IFT = 13) que les grandes cultures (IFT = 3,8 ; (Butault *et al.* 2010). Si les dépenses les plus importantes sur le territoire français en produits phytosanitaires concernaient les grandes cultures en 2006, les vignobles, les régions spécialisées dans les cultures fruitières ou dans l'horticulture et les légumes de plein champ, étaient donc *in fine* à l'origine des pressions pesticides (matière active/ha) les plus fortes avec quelques régions de grandes cultures du nord de la France (Figure 12). Or dans le cas des vignobles, l'utilisation des produits phytosanitaires est essentiellement liée au contrôle du mildiou et de l'oïdium. Avec la nécessaire baisse des intrants phytosanitaires (plan Eco-Phyto 2018) en lien avec le développement d'une agriculture durable et les changements climatiques, la création de nouveaux cépages est donc relancée en France depuis une dizaine d'années, tout comme en Europe, où de nombreux programmes de création variétale se sont poursuivis, comme en Allemagne ou en Hongrie, (Chabin *et al.* 2008).



**Figure 11. Les dépenses en pesticides dans l'agriculture française en 2006.** L'état des lieux de l'utilisation des pesticides monte à 2,3 milliards d'€ selon les données des enquêtes des pratiques culturales du Réseau d'Information Comptable Agricole. Les chiffres, exprimés en pourcentage (%), représentent la part allouée aux cultures.



**Figure 12. Répartition des petites régions agricoles selon l'origine et le niveau de la pression pesticide (pesticides par hectare).**

Ces dernières années, des méthodes de lutte alternative sont à l'étude. La lutte biologique utilisant des champignons mycoparasites (*Ampelomyces quisqualis*) ou des traitements éliciteurs induisent des mécanismes de résistance aux maladies constituent de nouvelles méthodes de contrôle de l'oïdium. A titre d'exemple, *Ampelomyces quisqualis*, un champignon hyperparasite naturel de l'oïdium, est capable d'infecter et de former des pycnides dans les hyphes, les conidiophores, et les cléistothèces de *E. necator*. *A. quisqualis* a été l'objet de recherches pour lutter contre l'oïdium depuis plus de 50 ans (Falk *et al.* 1995). Toutefois, dans cet exemple comme dans d'autres, en dépit de la mise en évidence initiale de leur potentiel, les performances sur le terrain sont en général faibles et irrégulières. Par ailleurs, l'utilisation des fongicides fait apparaître la nécessité de développer des résistances à des maladies qui étaient contrôlées en

même temps que l'oïdium et le mildiou par les applications de fongicides comme le black rot. L'un des enjeux majeurs pour la viticulture est donc la réduction de l'utilisation des produits phytosanitaires, tout en conservant la qualité du produit. La création de variétés de vigne résistantes constitue une des voies pour répondre à cet enjeu, tout en respectant l'environnement.

#### **1.2.4. Amélioration variétale pour la résistance aux maladies**

L'amélioration de la vigne a pour objectif de combiner la résistance acquise d'espèces sauvages, avec les qualités des raisins de cuve de *V. vinifera*. Au fil du temps, cette approche a conduit au développement de variétés résistantes qui peuvent être cultivées. Les premières variétés résistantes créées à partir du début du 20<sup>ème</sup> siècle, essentiellement à partir de croisements entre des espèces américaines et européennes n'ont pas produit des variétés présentant une qualité de fruit suffisante principalement en raison de la complexité de l'hérédité de la résistance et des caractères de qualité (Eibach *et al.* 1989). Leur utilisation de ces hybrides a donc été limitée (Galet, 1988). Cependant, avec le temps, certaines variétés résistantes commencent maintenant à atteindre des critères qualitatifs suffisamment intéressants (Villard blanc ou noir, Chambourcin, Seyval blanc...) voire très proche des caractéristiques trouvées chez *V. vinifera* (Regent, Solaris,...) pour être actuellement commercialisées. La plupart de ces hybrides commerciaux ont recruté les mêmes gènes de résistance (Figure 13 ; Venuti *et al.* 2013). Cette observation soulève la question de la durabilité de ces résistances. En effet, Peressotti *et al.* (2010) ont mis en évidence, lors d'une étude visant à établir le spectre d'action des facteurs de résistance à *P. viticola* et à estimer leur potentiel de durabilité, un isolat de mildiou contournant spécifiquement *Rpv3*, un gène de résistance au mildiou de la vigne initialement identifié dans la variété Bianca (Bellin *et al.* 2009). Cette étude qui constitue la première description d'un contournement de résistance montre que, malgré la variabilité génétique réduite de *P. viticola* en Europe, le contournement des résistances monogéniques peut survenir. Il est donc nécessaire de construire des résistances durables qui combinent plusieurs mécanismes de résistance à un même pathogène et cela, pour plusieurs pathogènes différents. Cependant, la connaissance des bases génétiques, moléculaires et physiologiques des résistances naturelles chez la vigne est un pré-requis pour optimiser le processus de création variétale et offrir une résistance efficace et durable.

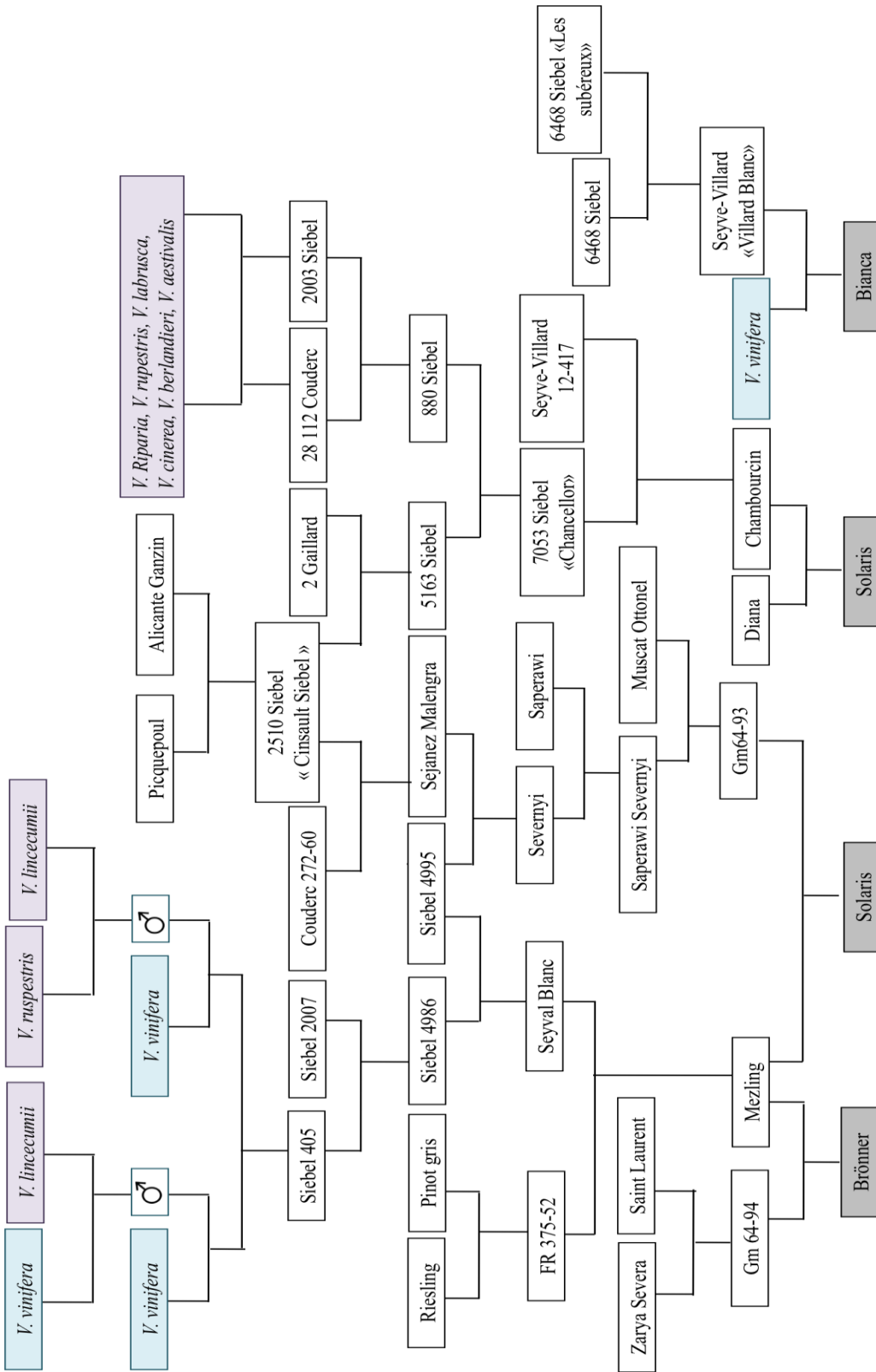


Figure 13. Généalogie des variétés résistantes commercialisées.

### **1.3. GENETIQUE DE LA RESISTANCE AUX MALADIES CHEZ LA VIGNE**

Des espèces résistantes au mildiou ayant été identifiées, l'étude de leur déterminisme génétique a été entreprise pour connaître le nombre de facteurs génétiques qui gouvernent la résistance, leur position sur le génome de la vigne, leur mode d'action. L'établissement d'une carte génétique nécessite la création d'une population en ségrégation issue d'un croisement par reproduction sexuée, ainsi que la caractérisation moléculaire des individus de la descendance. L'utilisation de méthodes de calculs mathématiques et statistiques permet de dresser la carte des *loci* liés à des caractères qualitatifs ou quantitatifs (QTL : *Quantitative Trait Loci*) et à des marqueurs moléculaires. Plus un marqueur est proche d'un gène d'intérêt, plus la probabilité de recombinaison entre le marqueur et le gène au cours de la méiose est faible. Une fois la corrélation entre un marqueur moléculaire et un phénotype spécifique établie, l'hérédité d'un caractère d'intérêt peut être suivie dans une descendance à des stades très précoces du développement des plantes (Gallais, 1993 ; Fischer *et al.* 2004). De nombreuses cartographies génétiques de QTL de résistance sont publiées ou en cours d'étude chez la vigne.

#### **1.3.1. Cartographie génétique des résistances aux maladies chez la vigne**

##### **1.3.1.1. Cartographie génétique des gènes de résistance de la vigne**

Ainsi, des QTL de résistance au mildiou et à l'oïdium ont été évalués et cartographiés dans plusieurs fonds génétiques y compris chez des hybrides obtenus à partir d'espèces *Euveitis* nord-américaines, de différentes accessions de *M. rotundifolia* (Pauquet *et al.* 2001 ; Dalbó *et al.* 2001 ; Merdinoglu *et al.* 2003 ; Fischer *et al.* 2004 ; Akkurt *et al.* 2007 ; Welter *et al.* 2007 ; Marguerit *et al.* 2009 ; Bellin *et al.* 2009 ; Moreira *et al.* 2011). Des QTL de résistance ont également été cartographiés chez les espèces originelles asiatiques (*V. amurensis*, *V. romanetii*) et les cultivars *V. vinifera* 'Kishmish vatkana' et 'Dzhandzhal kara' (Hoffmann *et al.* 2008 ; Coleman *et al.* 2009 ; Blasi *et al.* 2011 ; Schwander *et al.* 2012 ; Venuti *et al.* 2013). Les résultats publiés à ce jour sont résumés dans le Tableau 3.

#### 1.3.1.1.1. Les résistances provenant du sous-genre *Muscadinia*

En 1974, l'INRA a entrepris la sélection de variétés résistantes en introgressant dans *V. vinifera* des gènes issus de *M. rotundifolia* (Merdinoglu *et al.* 2003). Une résistance totale à l'oïdium provenant *M. rotundifolia* a été trouvée, contrôlée par un seul *locus* dominant (Bouquet, 1986) dénommé *Run1* (*Resistance to Uncinula necator 1*). Ce dernier a été cartographié sur le chromosome 12 de la vigne ; et il provoquerait une mort cellulaire programmée apparaissant 24 - 48 heures après la contamination dans les cellules épidermiques jouxtant le site d'infection.

Après évaluation du niveau de résistance d'une population issue d'un croisement entre un descendant résistant de *M. rotundifolia* et la variété Syrah, un QTL expliquant 75 % de la variation totale du niveau de résistance observé a été identifié et cartographié sur le chromosome 12 du génome de la vigne. Cette région, du fait de son effet important sur la résistance, a été considérée comme un gène majeur et appelé *Rpv1* (Merdinoglu *et al.* 2003). Le *locus* de résistance *Rpv1* confère une résistance partielle contre *P. viticola* et il est situé au même *locus* que le gène de résistance à l'oïdium *Run1* précédemment décrit par A. Bouquet (Barker *et al.* 2005 ; Merdinoglu *et al.* 2006 ; Bellin *et al.* 2009). Afin de caractériser la fonction de ces deux gènes, l'INRA a initié l'isolement par clonage positionnel et l'étude de la fonction des gènes *Run1* et *Rpv1*, en collaboration avec une équipe australienne du CSIRO (Barker *et al.* 2005 ; Dry *et al.* 2010).

A l'aide d'une seconde population de cartographie également dérivée de *M. rotundifolia*, un deuxième QTL majeur de résistance au mildiou, appelé *Rpv2*, a été identifié et localisé sur le chromosome 18 (Merdinoglu *et al.* 2006). Contrairement à *Rpv1*, le *locus* de résistance *Rpv2* est le seul QTL identifié conférant une résistance élevée dans un fond génétique sensible. L'étude des effets de ces QTL sur le développement du pathogène et les mécanismes de réponse à l'infection mis en place ont montré que ces QTL ont un mode d'action différent sur *P. viticola* ou qu'ils agissent à des stades distincts de son développement. *Rpv2* provoque un arrêt rapide du développement de *P. viticola*, souvent 24 heures post-inoculation, ce qui inhibe toute possibilité de sporulation, alors que *Rpv1* ralentit sa progression sans l'arrêter (Merdinoglu *et al.* 2003 ; Merdinoglu *et al.* 2006).

Plus récemment, plusieurs QTL de résistance à l'oïdium ont été identifiés pour les différents types de tissus (feuille, tige, rachis et baies) chez *M. rotundifolia* cv 'Magnolia' ; et ils expliquent



12,6 - 52,8 % de la variance phénotypique (Riaz *et al.* 2011). La résistance de *M. rotundifolia* cv ‘Magnolia’ a été nommée *Run2.1*. Pareillement, des QTL de résistances à l'oïdium ont également été identifiés chez *M. rotundifolia* cv ‘Trayshed’ et ils expliquent 50 % de la variance phénotypique. La résistance de *M. rotundifolia* cv ‘Trayshed’ a été nommée *Run2.2*. Contrairement au gène de résistance à l'oïdium *Run1* qui a été cartographié sur le chromosome 12, les loci *Run2.1* et *Run2.2* sont chevauchants et sur un même locus du chromosome 18. En effet, *M. rotundifolia* cv ‘Trayshed’ et ‘Magnolia’ ont en commun 68 % de marqueurs testés. Ces résultats indiquent que les deux sélections de *M. rotundifolia* sont très différentes mais partagent au moins un allèle.

#### **1.3.1.1.2. Les résistances provenant du sous-genre *Euvtis***

Dans le cas de la résistance au mildiou, un locus majeur, *Rpv3*, a été identifié chez la variété Bianca par Bellin *et al.* (2009). Le locus *Rpv3* conférerait une résistance partielle en expliquant environ 74,6 % de la variance phénotypique. Il contrôlerait la capacité de déclencher une résistance race-spécifique contre *P. viticola* (Casagrande *et al.* 2011) en limitant la densité des sporangiophores. Le locus *Rpv3* serait localisé sur la partie distale du chromosome 18, en aval du locus de résistance *Rpv2*, et entre les marqueurs UDV-305 et UDV-737, distant de 1,4 cM (Moroldo *et al.* 2008 ; Di Gaspero *et al.* 2012). Ce large intervalle de confiance a conduit Bellin *et al.* (2009) à présumer que des gènes fonctionnels, dont le locus *Rpv2*, pourraient résider dans le locus *Rpv3*. Néanmoins, une souche de *P. viticola* contournerait la résistance conférée par le locus de résistance *Rpv3* chez Bianca (Peressotti *et al.* 2010). Mais la résistance conférée par *Rpv2* reste active lors du contournement de la résistance conférée par *Rpv3* (Blasi *et al.* 2011), ce qui suggère qu’il s’agit de deux gènes distincts. En plus du locus de résistance *Rpv3*, Bellin *et al.* (2009) ont identifié un QTL mineur qui expliquerait 12,1 % de la variance phénotypique chez ‘Bianca’. Un QTL mineur a également été identifié chez *V. riparia* (Moreira *et al.* 2010) Comme le locus *Rpv3*, il limiterait la densité des sporangiophores. Ce locus pourrait augmenter la résistance au mildiou chez ces génotypes.

Fischer *et al.* (2004) ont réalisé une carte génétique des cultivars résistant ‘Regent’ et sensible ‘Lemberger’, en se basant sur les marqueurs sur les marqueurs RAPD et AFLP. Même si ces types de marqueurs permettent la construction rapide de cartes génétiques, ils ne facilitent pas le transfert d’information entre différents génotypes (Adam-Blondon *et al.* 2004). Néanmoins, en

analysant les données phénotypiques, un QTL majeur de résistance au mildiou expliquant jusqu'à 70 % de la variation phénotypique, a été localisé sur le groupe de liaison 18 ; ainsi qu'un QTL mineur qui a été localisé sur le groupe de liaison 5. En 2007, Welter *et al.* (2007) ont amélioré la carte génétique en introduisant des marqueurs microsatellites. Un QTL de résistance majeur a été localisé sur le groupe de liaison 18 mais il explique seulement 37,3 % de la variance phénotypique, avec un large intervalle de confiance. Selon les auteurs, il y aurait plusieurs gènes impliqués dans la résistance associée à cette région chromosomique. Quant au QTL mineur, il a été localisé sur le groupe de liaison 4, contrairement aux résultats publiés par Fischer *et al.* (2004). Ce dernier explique jusqu'à 22,6 % de la variance phénotypique, malgré son effet 'faible'.

En 2009, deux QTL de résistance, **Rpv5** et **Rpv6**, ont été identifiés chez l'hybride interspécifique *V. riparia* cv. Gloire de Montpellier (Marguerit *et al.* 2009). Ces QTL ont été cartographiés sur les groupes de liaison 9 et 12, respectivement ; et ils expliquent 34,4 % et 28,9 % de la variance phénotypique, respectivement. Ces QTL seraient impliqués dans la réduction du nombre de sporanges. Le locus de résistance **Rpv6** serait situé dans la même région que le locus **Rpv1** sur le groupe de liaison 12.

Schwander *et al.* (2011) ont identifié un QTL de résistance au mildiou, **Rpv10**, qui est localisé sur le groupe de liaison 9 de *V. amurensis*. Ce QTL co-localise avec le QTL de résistance **Rpv5** identifié par Marguerit *et al.* (2009) chez *V. riparia*. Le QTL de résistance **Rpv10** explique jusqu'à 50 % de la variation phénotypique. En parallèle, Blasi *et al.* (2011) et Venuti *et al.* (2013) ont identifié les QTL de résistance au mildiou, **Rpv8** et **Rpv12**, qui sont localisés sur le groupe de liaison 14 de *V. amurensis*. Ces QTL expliquent 86,3 % et 79 % de la variance phénotypique, respectivement. Le QTL de résistance **Rpv8** limiterait la sporulation de *P. viticola* et permettrait la formation rapide de nécrose au site d'infection de l'agent pathogène alors que le QTL de résistance **Rpv12** conduirait à la formation de nécrose uniquement. Ces observations phénotypiques affaiblissent l'hypothèse selon laquelle les résistances des QTL **Rpv8** et **Rpv12** seraient conférées par le même haplotypes. Néanmoins, il pourrait s'agir de différents allèles chez *V. amurensis* ou de locus paralogues (Venuti *et al.* 2013). En somme, ces études suggèrent que la résistance au mildiou chez *V. amurensis* serait due à des QTL situés sur différents chromosomes.

Dalbó *et al.* (1998) ont réalisé une carte génétique à l'aide de marqueurs RAPD et AFLP grâce à un hybride issu du croisement entre les espèces américaines *V. cinerea* et *V. rupestris* (Dalbó *et al.* 1998 ; Dalbó *et al.* 2001). Un QTL de résistance à l'oïdium, **Ren2**, et il explique 41 % de la variance phénotypique. Ces marqueurs ne permettant pas le transfert d'information entre espèces (Adam-Blondon *et al.* 2004), ils sont absents des cartes de références actuelles.

Le QTL de résistance à l'oïdium, **Ren3**, a été identifié à partir des hybrides de 'Regent', 'Villard Blanc' ou 'Gf.Ga-47-42' (Fisher *et al.* 2004 ; Welter *et al.* 2007 ; Akkurt *et al.* 2007), et il est localisé sur le groupe de liaison 15 de la vigne. Il explique jusqu'à 56,8 % et 64,5 % de la résistance au niveau des feuilles et des baies, respectivement.

**Tableau 3. QTLs associés à la résistance contre *P. viticola* et *E. necator***

	QTL	Variance expliquée (%)	Marqueurs associés	Groupe de liaison	Population (N°)	Source de résistance	Référence
<i>Plasmopara viticola</i>	<i>Rpv1</i>	75	VMC4f3.1/VMC8g9	12		<i>M. rotundifolia</i>	Merdinoglu <i>et al.</i> 2003
	<i>Rpv2</i>	100	VMC6f11/VVMD17	18	129	<i>M. rotundifolia</i>	Merdinoglu <i>et al.</i> 2006
	<i>Rpv3</i>	74,6	UDV-305/UDV-737	18	153	Bianca*	Bellin <i>et al.</i> 2009 Casagrande <i>et al.</i> 2011
	<i>Rpv4</i>	37,3	VMC7h3/VMCNg2e.1	18	153	Regent*	Welter <i>et al.</i> 2007
	<i>Rpv5</i>	34,4	VVIo52b	9	138	<i>V. riparia</i>	Marguerit <i>et al.</i> 2009
	<i>Rpv6</i>	28,9	VMC8G9	12	138	<i>V. riparia</i>	Marguerit <i>et al.</i> 2009
	<i>Rpv7</i>	12,1	UDV-097	7	116	Bianca*	Bellin <i>et al.</i> 2009
	<i>Rpv8</i>	86,3	Chr14V015	14	232	<i>V. amurensis</i>	Blasi <i>et al.</i> 2011
	<i>Rpv9</i>	5,8 - 10,7	CCoAOMT	7	174	<i>V. riparia</i>	Moreira <i>et al.</i> 2010
	<i>Rpv10</i>	50	GF09-46	9	265	<i>V. amurensis</i>	Schwander <i>et al.</i> 2012
	<i>Rpv11</i>	–	VVMD27	5	153	–	Fischer <i>et al.</i> 2004
	<i>Rpv12</i>	79	UDV014/UDV370	14	180	<i>V. amurensis</i>	Venuti <i>et al.</i> 2013
	<i>Rpv13</i>	21,2	VMC1G3.2	12	174	<i>V. riparia</i>	Moreira <i>et al.</i> 2010
<i>Erysiphe necator</i>	<i>Ren1</i>	–	UDV-020/VMC9h4-2	13	310	Kishmish vatkana	Hoffmann <i>et al.</i> 2008
	<i>Ren2</i>	41	CS25	14	58	<i>V. cinerea</i>	Dalbo <i>et al.</i> 2001
	<i>Ren3</i>	56,8 - 64,5	UDV-015b/VViv67	15	153	Regent	Welter <i>et al.</i> 2007
	<i>Ren4</i>	70	VMC7f2	18	42	<i>V. romanetii</i>	Riaz <i>et al.</i> 2011
	<i>Ren5</i>	80	VVIp05/Chr14V041	14	191	<i>M. rotundifolia</i>	Blanc <i>et al.</i> 2013
	<i>Run1</i>	100	VMC4f3.1/VMC8g9	12	161	<i>M. rotundifolia</i>	Barker <i>et al.</i> 2005
	<i>Run2.1</i>	12,6 - 52,8	UDV-108	18	97	<i>M. rotundifolia</i>	Riaz <i>et al.</i> 2011
	<i>Run2.2</i>	50	VMC7f2	18	255	<i>M. rotundifolia</i>	Riaz <i>et al.</i> 2011

\* Hybrides interspécifiques

#### 1.4. L'INTERACTION PLANTE - AGENT PATHOGENE : MECANISME DE RESISTANCE ET EVOLUTION DES GENES DERESISTANCE

Dans leur environnement, les plantes doivent faire face à l'agression de micro-organismes tels que les champignons, les bactéries, les virus ou encore les nématodes. En tant qu'êtres vivants, elles ont élaboré au cours de leur évolution des mécanismes de résistance efficaces contre les divers agents pathogènes. Ces mécanismes sont activés par la plante après reconnaissance de l'agent pathogène. Dans le cas où la mise en place de la défense n'est pas suffisante pour limiter la progression de l'agent pathogène, on parle d'**interaction compatible**. L'agent pathogène, dit **virulent** se multiplie et colonise une partie ou l'ensemble de la plante qui réagit alors plus ou moins efficacement à cette agression. La plante développe alors une maladie qui affecte son développement et elle est alors dite **sensible**. Si toutes les variétés d'une même espèce végétale sont résistantes à toutes les races d'un même agent pathogène : on parle de **résistance non-hôte**. Dans le cas d'une reconnaissance spécifique entre la plante et l'agent pathogène, les mécanismes de défense mis en place par la plante sont suffisants pour limiter la progression de l'agent pathogène, on parle alors de **résistance spécifique** de la plante. L'agent pathogène est dit avirulent et on parle alors d'**interaction incompatible**. La défense désigne l'ensemble des mécanismes mis en place par la plante afin de lutter contre une attaque pathogène. Il existe deux grands types de défenses :

- Les **défenses constitutives** : Comme leur nom l'indique, ces défenses sont présentes dans la plante même en l'absence d'un agent pathogène. La première barrière à l'infection est bien entendu une barrière physique telle que les parois ou le cytosquelette des cellules végétales (Thordal-Christensen, 2003). Il s'agit également de molécules anti-microbiennes exprimées constitutivement chez la plante. Cette protection immédiate permet surtout de laisser le temps à la plante de mettre en place ses défenses inductibles plus efficaces.

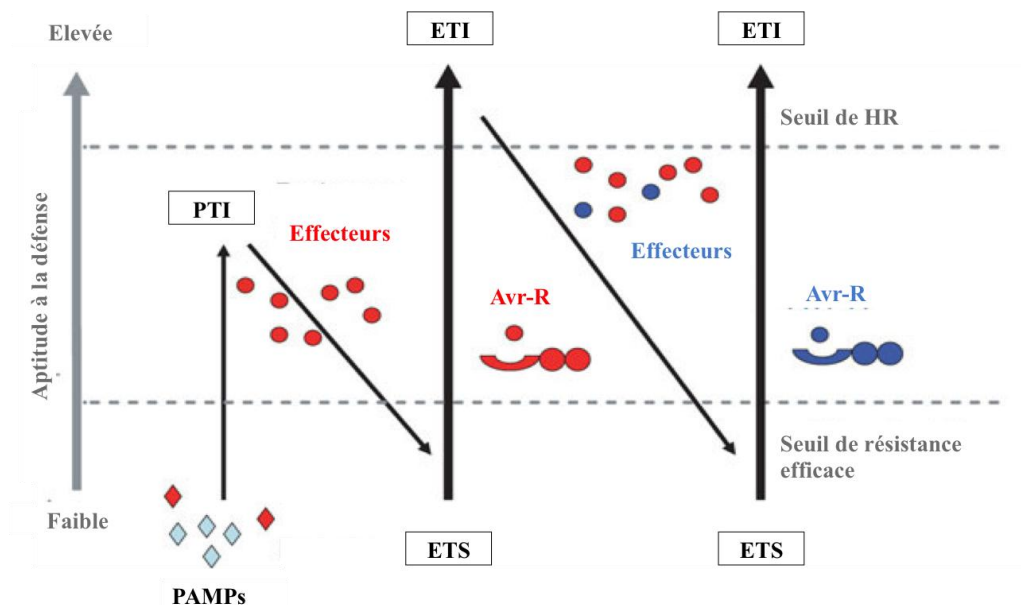
- Les **défenses induites** : Elles sont activées par la plante en réponse à une attaque par un agent pathogène (Thordal-Christensen, 2003). On parle également d'immunité chez les plantes par analogie à l'immunité observée dans les systèmes animaux.

Dans la suite de ce chapitre, nous présenterons uniquement les défenses induites. Dans un premier temps, nous verrons comment l'évolution du couple plante-agent pathogène est à l'origine de la résistance induite par certains mécanismes conservés (i), puis nous nous intéresserons particulièrement à la résistance spécifique (ii).

## 1.4.1. L'immunité chez les plantes - Les défenses induites

### 1.4.1.1. Le modèle Zig-Zag

La résistance correspond à la capacité de la plante à retarder ou supprimer l'activité et le développement du parasite. Selon la mise en place et la nature des mécanismes impliqués en réponse à une attaque d'agents phytopathogènes, différents types de résistance sont distingués en fonction (i) de l'immunité totale de l'espèce ou de la possibilité de détecter des réactions (résistance maladie) de tout ou partie des génotypes de l'espèce, (ii) du spectre d'action des résistances mises en œuvre par l'espèce végétale vis-à-vis de la diversité génétique présente au sein de l'espèce pathogènes considérée et enfin (iii) de l'intensité de la réaction de défense conduisant à une résistance complète ou partielle.



**Figure 14. Modèle en Zig-Zag illustrant les mécanismes de défense mis en place contre les pathogènes (d'après Jones et Dangl, 2006).** ETI : *Effector-Triggered-Immunity* ; ETS : *Effector-Triggered-Sensibility* ; PTI : *PAMP-Triggered-Immunity* ; PAMP : *Pathogen Associated Molecular Pattern* ; Avr-R: gène d'avirulence

Du point de vue moléculaire, la capacité des plantes à mettre en œuvre des défenses est due, en partie, à l'existence dans chaque cellule végétale de systèmes de reconnaissance. Le modèle Zig-Zag, représente de manière schématique l'évolution simultanée des protéines végétales impliquées dans la reconnaissance des agents pathogènes et des protéines produites par ces agents pathogènes, au cours du temps, ainsi que leurs conséquences pour la plante : résistance ou maladie (Jones and Dangl, 2006) (Figure 14). Les plantes détectent les agents pathogènes *via* la perception de composés très conservés d'une espèce pathogène à l'autre, nommés **PAMP** (*Pathogen-Associated Molecular Pattern*) (Zipfel and Felix, 2005). Ce premier niveau de reconnaissance déclenche la mise en place de l'immunité innée ou **PTI** (*PAMP-triggered immunity*) (Figure 14). Cette immunité regroupe à la fois la **résistance basale** et la **résistance non-hôte** :

- la résistance basale limite l'étendue de la maladie par l'agent pathogène. Elle réduit la sévérité des symptômes mais ne stoppe ni le développement de l'agent pathogène, ni la maladie.
- La résistance non-hôte qui se distingue par le fait qu'elle est efficace en réponse à toutes les souches d'un agent pathogène donné.

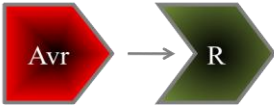
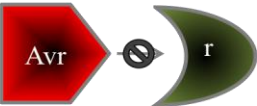
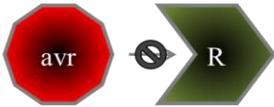

Cependant les agents pathogènes ont la capacité de supprimer la PTI soit en interférant avec l'événement de reconnaissance, soit en supprimant la signalisation en aval de la reconnaissance. Dans la plupart des cas, la suppression de la PTI fait intervenir des effecteurs de virulence sécrétés directement à l'intérieur de la cellule végétale. Il en résulte une sensibilité accrue de la plante, nommée **ETS** (*Effector-Triggered Susceptibility*) (Figure 14).

Au cours de l'évolution entre les plantes et les agents pathogènes, certaines plantes ont en outre développé un mécanisme de reconnaissance spécifique afin de détecter la présence d'agents pathogènes. La résistance spécifique qui est alors déclenchée, ou encore **ETI** (*Effector-Triggered Immunity*), est une PTI amplifiée et accélérée, qui est le plus souvent accompagnée d'une mort cellulaire limitée au niveau du site d'infection appelée **réponse hypersensible** (HR) (Figure 14). La HR est similaire au processus de mort cellulaire programmée (apoptose), qui déclenche des nécroses cellulaires et confine l'agent pathogène au niveau du site d'infection.

Enfin, les agents pathogènes ont évolué à nouveau par l'acquisition de nouveaux effecteurs afin de supprimer la résistance spécifique. Chez la plante, de nouveaux allèles R évoluent à leur tour afin de contourner les effecteurs des agents pathogènes.

### 1.4.1.2. La reconnaissance spécifique des agents pathogènes

Le concept de reconnaissance spécifique a été découvert par Flor *et al.* (1971) qui a travaillé sur la rouille (*Melampsora lini*) du lin (*Linum usitatissimum*). Il a mis en évidence que la mise en place de la résistance chez la plante nécessitait la présence d'un gène chez les deux partenaires : chez la plante un gène appelé **gène de résistance** ou **gène R**, et chez le champignon pathogène, un **gène d'avirulence** ou **gène Avr**. En l'absence d'un des acteurs, la maladie se développe (Figure 15). Sur la base de ce modèle génétique appelé **résistance gène-à-gène**, il a été proposé une interprétation du modèle selon laquelle les protéines R agiraient comme des récepteurs qui lieraient spécifiquement et de façon directe ou indirecte une protéine ligand Avr correspondante afin d'activer les mécanismes de défense de la plante (Gabriel and Rolfe, 1990). Un exemple est le gène d'avirulence de *Pseudomonas syringae*, *AvrPphB*, qui code une protéase qui peut cliver une protéine kinase de la plante hôte. Ce clivage est détecté par la protéine R apparentée (Résistance à *Pseudomonas syringae* 5, RPS5), qui devient alors active (Ade *et al.* 2007).

Génotype de la plante  Génotype de l'agent pathogène	<i>R</i>	<i>r</i>
<i>Avr</i>	 <b>RESISTANCE</b>	 <b>MALADIE</b>
<i>avr</i>	 <b>MALADIE</b>	 <b>MALADIE</b>

**Figure 15. Le modèle de résistance « gène pour gène**

La résistance n'est induite que dans le cas où la plante exprime le gène de résistance R dont le produit est impliqué dans la reconnaissance du produit du gène d'avirulence Avr exprimé par l'agent pathogène. Dans tous les autres cas, la maladie se développe. *R*, *Avr* : gènes dominants; *r*, *avr* : gènes récessifs.



### 1.4.2. Les protéines de résistance végétales

Près d'une centaine de gènes codant les protéines de résistance ont été clonées à partir des plantes mono et dicotylédones (Martin *et al.* 2003; Hammond-Kosack and Kanyuka, 2007). Malgré l'importante diversité des plantes dont ils proviennent et des parasites auxquels ils confèrent la résistance, ils appartiennent pour la grande majorité d'entre eux à trois familles de gènes impliqués dans la reconnaissance et la transduction de signaux (DeYoung and Innes, 2006).

#### 1.4.2.1. Les domaines fonctionnels des protéines de résistance

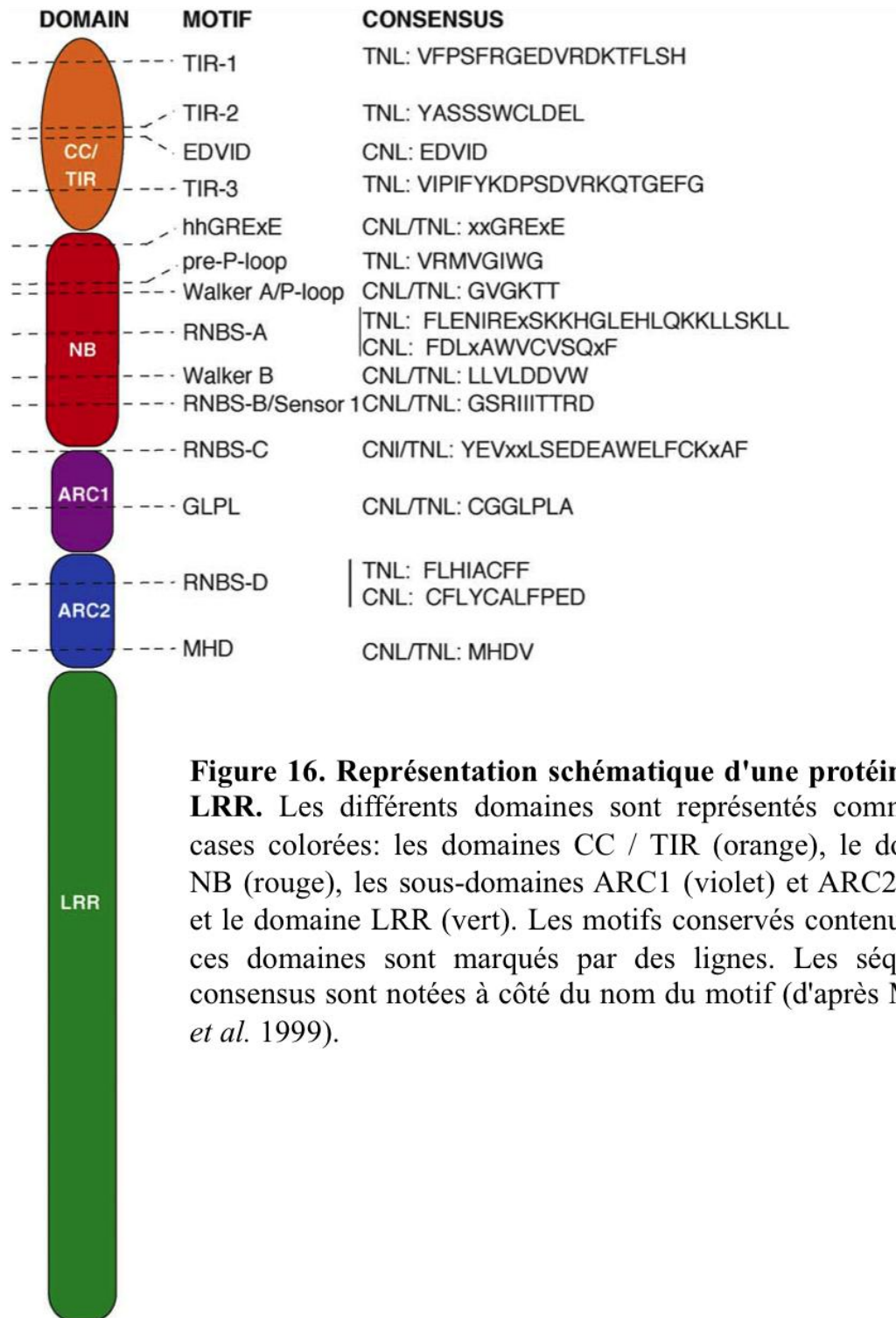
Des études comparatives de similarité de séquence entre protéines R végétales et des protéines de l'immunité innée chez les animaux ont contribué à la compréhension de la structure des protéines R, le rôle des domaines protéiques individuels, et le mécanisme par lequel les protéines R identifient et répondent aux protéines étrangères (Nürnberger *et al.* 2004 ; Takken *et al.* 2006 ; Rairdan et Moffett, 2007). La plupart des protéines codées par les gènes *R* isolés de plantes partage différents domaines protéiques intervenant dans les mécanismes de défense chez la levure, la drosophile ou encore les vertébrés (Michelmore et Meyers, 1998 ; Cohn *et al.* 2001 ; Dangl et Jones, 2001 ; Staskawicz *et al.* 2001 ; Nürnberger et Brunner, 2002). Par analogie, les protéines R combinent un domaine récepteur et un domaine effecteur assurant respectivement : la reconnaissance des molécules élicitrices et l'activation directe ou indirecte de signaux de transduction. Cinq principaux domaines structuraux conservés ont été distingués (Figure 16).

Les protéines *R* clonées présentent en position C-terminal un **domaine LRR** (*Leucine-Rich Repeats*) intra- ou extra- cytoplasmique. Les domaines LRR correspondent à la répétition d'un motif de taille variable comprenant des leucines. Les domaines LRR sont retrouvés dans des protéines de nature très diverse et sont impliqués dans les interactions protéine-protéine ou protéines-polysaccharides, ainsi que dans d'autres liaisons peptide-ligand (Kobe et Kajava, 2001). Il a été montré que la reconnaissance spécifique est déterminée par ces domaines LRR (Rairdan and Moffett, 2007 ; Van Ooijen *et al.* 2007). Ce sont des domaines très variables qui tendent à s'adapter à l'évolution constante des protéines d'agents pathogènes (Meyers *et al.* 1998b ; Michelmore et Meyers, 1998 ; Mondragon-Palomino *et al.* 2002).

Le **domaine NBS** (*Nucleotide Binding Site*), associé au domaine LRR, constitue la plus grande classe des gènes *R* clonés. Ce domaine est composé de différents motifs de type kinase, et il correspond à un site de fixation et d'hydrolyse des nucléotides triphosphates ATP et GTP. Le domaine NBS est très conservé et fait partie d'une entité plus grande appelée domaine **NB-ARC**, qui est présent dans APAF-1 (*Apoptotic protease-activating Factor1*) chez l'Homme et la protéine CED4 chez *Caenorhabditis elegans* (Van Der Biezen and Jones, 1998 ; Rairdan et Moffett 2007).

Trois autres domaines, dits TIR, CC et BED, peuvent être associés, en position N-terminal, aux protéines NBS-LRR. Le **domaine TIR** (*Toll Interleukin Receptor*) présente d'importantes homologies de séquences avec les domaines intracellulaires de récepteurs protéiques isolés chez la drosophile (récepteur *Toll*) et chez l'homme (récepteur *Interleukin-1*) (Hammond-Kosack et Jones, 1997). Sur la base de cette homologie, un rôle dans la cascade de signalisation cellulaire est attribué au domaine TIR. Le **domaine CC** (*Coiled-Coil*), qui peut varier en taille et en position, est connu pour assurer un rôle dans l'homo- ou l'hétérodimérisation des protéines. En raison de l'incertitude quant à la présence d'un motif du domaine CC, la classe de gènes CC-NB-LRR est souvent désignée comme les protéines non-TIR. Le domaine **BED** présente des homologies de séquence avec les protéines régulatrices de la transcription BEAF, DREF et Dpy-20 caractérisées chez *Drosophila melanogaster* (Zhao *et al.* 1995). Ce domaine a également présenté des homologies avec les protéines de fixation à l'ADN chez le tabac (Lam *et al.* 1990), la tomate (Coupe *et al.* 1997) et chez les transposons de plantes (Liu *et al.* 1998). La présence de la capacité de se fixer à l'ADN dans divers domaines BED a conféré une fonction à ce domaine (Lam *et al.* 1990 ; Hirose *et al.* 1996 ; Hart *et al.* 1997 ; Coupe *et al.* 1997 ; Köhler *et al.* 2008).

Enfin, le **domaine Sérine/Thréonine kinase** se présente seul comme dans le cas du gène *Pto* de la tomate contre la bactérie *Pseudomonas syringae* ; ou associé à un domaine LRR comme dans le cas du gène *Xa21* du riz contre la bactérie *Xanthomonas oryzae*. Il serait impliqué dans des réactions de phosphorylation lors de cascades de signalisation (Martin *et al.* 1993 ; Song *et al.* 1995 ; Jones and Dangl 2006 ; Mucyn *et al.* 2006).



**Figure 16. Représentation schématique d'une protéine NB-LRR.** Les différents domaines sont représentés comme des cases colorées: les domaines CC / TIR (orange), le domaine NB (rouge), les sous-domaines ARC1 (violet) et ARC2 (bleu) et le domaine LRR (vert). Les motifs conservés contenus dans ces domaines sont marqués par des lignes. Les séquences consensus sont notées à côté du nom du motif (d'après Meyers *et al.* 1999).

#### 1.4.2.1.1. Classification des gènes *R*

Les gènes *R* sont classés en fonction des différentes associations des domaines conservés (LRR, NBS, BED, CC, TIR et Ser/Thr kinase) au sein des protéines qu'ils codent. On distingue ainsi 5 classes de gènes (Figure 17) :

##### - La classe des gènes *NBS-LRR*

La majorité des gènes clonés codent des protéines cytoplasmiques de type récepteur. Ces gènes présentent à la fois un domaine LRR à l'extrémité 3' et un domaine NBS à l'extrémité 5'. C'est la classe de gènes *NBS-LRR* la plus abondante chez les plantes (McHale *et al.* 2006). Certaines protéines codées par certains gènes présentent en position amino-terminale les domaines TIR, CC ou BED (Köhler *et al.* 2008), en plus des domaines LRR et NBS. Chez *A. thaliana* la majorité des gènes *R* sont des TIR-NBS-LRR (TNL) (Meyers *et al.* 1999), contrairement aux monocotylédones n'en possèdent pas (Pan *et al.* 2000). Ceci pourrait indiquer une différence dans la co-évolution plante hôte-agent pathogène entre les monocotylédones et les dicotylédones (Van Ooijen *et al.* 2007). Pour ne citer que quelques exemples qui illustrent la diversité des rôles des gènes de cette famille, chez les solanacées trois gènes appartenant à cette classe ont été clonés : *Bs4* chez la tomate (Schornack *et al.* 2004), *Gro1-4* chez la pomme de terre (Paal *et al.* 2004) et *N* chez le tabac (Whitham *et al.* 1994) conférant la résistance à *Xanthomonas campestris*, *Globodera rostochiensis* et au virus de la mosaïque du tabac, respectivement. Un représentant de la classe des CC-NBS-LRR (CNL) est le gène *RPS2* identifié chez *A. thaliana* et qui code une protéine reconnaissant spécifiquement la protéine avr Rpt2 de *Pseudomonas syringae* pv. *maculicola* (Leister *et al.* 1996). Chez cette même plante un autre gène appartenant à la classe des CNL, *RPML1*, a été identifié conférant la résistance à des souches de *P. syringae* ayant les gènes *avrB* ou *avrRpm1* (Bent *et al.* 1994 ; Grant *et al.* 1995).

##### - La classe des gènes de résistance avec un domaine LRR extracellulaire

Cette classe est essentiellement représentée par la famille de gènes *Cf* conférant la résistance à *Cladosporium fulvum* chez la tomate (Hammond-Kosack et Jones, 1997 ; Thomas *et al.* 1997) et *Xa21* conférant la résistance à *Xanthomonas oryzae* chez le riz. Le gène *Xa21* est constitué d'un domaine LRR extracellulaire, une région transmembranaire et un domaine sérine/thréonine

kinase cytoplasmique ; contrairement au gène *Cf* qui présente un domaine LRR extracellulaire, un domaine transmembranaire et une courte région cytoplasmique. Cette dernière ne possède pas de région intracellulaire qui pourrait être impliquée dans des voies de signalisation telles que les protéines kinases.

- *Les gènes de résistance Serine-Thréonine Kinase (STK)*

Le gène *Pto* code une protéine kinase de type sérine/thréonine. La protéine *Pto* présente des similarités avec les protéines IRAK et PELLE, des kinases impliquées respectivement dans la réponse immunitaire chez les mammifères et la drosophile (Medzhitov, 2001). Ces homologies suggèrent des mécanismes de transduction communs entre les différents règnes. La protéine kinase codée par le gène *Pto* nécessite la présence de la protéine codée par le gène *Prf* (*Pseudomonas resistance and fenthion sensitivity*) pour qu'il y ait expression de la résistance après infection par *P. syringae* exprimant le gène *AvrPto* (Ellis et Jones, 1998 ; Martin *et al.* 1999). Deux autres gènes présentent un domaine de type kinase :

Le gène *Pbs1*, identifié chez *A. thaliana* confère la résistance à *P. syringae* (Swiderski *et al.* 2001). La protéine codée par ce gène présente un domaine putatif Sérine/Thréonine kinase, différent de celui identifié au sein de la protéine *Pto*, ainsi que de courtes extensions N- et C-terminales. Par ailleurs, la fonctionnalité du gène *Pbs1* est indissociable de celle du gène *RPS5* (Warren *et al.* 1999).

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

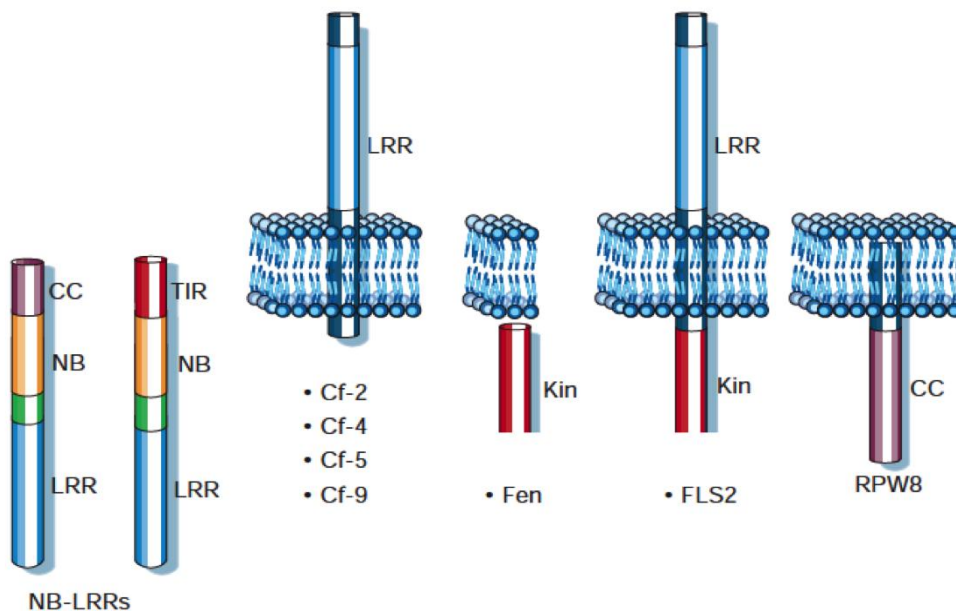
- *Autres*

**Gènes codant des toxines réductases :** Le gène *Hm1* représente la classe de gènes codant des toxines réductases. La protéine *Hm1* est capable d'inactiver, chez le maïs, la toxine du champignon agent pathogène *Cochliobolus caronum* (Johal et Briggs, 1992). Ces gènes sont

conservés chez les céréales puisque des homologues ont été retrouvés chez l'orge ou encore le blé (Han *et al.* 1997).

**Gènes codant des protéines à domaine CC et transmembranaire :** Cette classe de gènes *R* est représentée par la famille de gènes *RPW* conférant la résistance à *Erysiphe cichoracearum* chez *A. thaliana* (Xiao *et al.* 2001). Les protéines cytoplasmiques RPW8.1 et RPW8.2 sont caractérisées par la présence d'un signal d'ancrage putatif en N-terminal, suivi d'un domaine CC impliqué dans les interactions avec d'autres acteurs de voies de signalisation (Pan *et al.* 2000). Un domaine transmembranaire, situé en C-terminal, intervient dans la reconnaissance d'un facteur d'avirulence et facilite la reconnaissance d'autres molécules (Fluhr, 2001).

Un nouveau type de récepteur a été identifié chez la tomate (Kawchuk *et al.* 2001). En effet, la protéine codée par le gène *Ve* confère une résistance à *Verticillium*. Elle associe un domaine LRR, une séquence-signal potentielle d'endocytose, ainsi qu'un domaine CC ou une séquence PEST (Pro-Glu-Ser-Thr) connue pour intervenir dans l'ubiquitination, la compartimentation et la dégradation des protéines.



**Figure 17. Représentation de la localisation et de la structure des cinq classes principales de récepteurs impliqués dans l'ETI** (d'après Dangl et Jones, 2001). CC : Coiled Coil ; NB : Nucleotide Binding ; LRR: Leucine Rich Repeat; TIR: Toll Interleukine Receptor.

#### 1.4.2.1.2. Activation et régulation des gènes NBS-LRR

La structure et la fonction du domaine NBS qui contient plusieurs motifs caractéristiques des STAND (*Signal Transduction ATPases with Numerous Domains*) de la famille des ATPases, qui comprend les protéines NOD des mammifères (Leipe *et al.* 2004 ; Albrecht *et al.* 2006). Les protéines STAND fonctionnent comme des interrupteurs moléculaires dans les voies de signalisation en cas de maladie. La liaison spécifique et l'hydrolyse de l'ATP provoqueraient des changements conformationnels qui régulent la signalisation en aval (Tameling *et al.* 2002).

##### 1.4.2.1.2.1. Fonction des gènes de résistance NBS-LRR

Comme mentionné plus haut, la reconnaissance spécifique des effecteurs des agents pathogènes est due au domaine LRR qui est soumis à une diversification pour assurer la co-évolution avec les effecteurs des agents pathogènes (Qi *et al.* 2012). Bien que les domaines NBS et LRR soient connus, certaines petites régions conservées ont également été identifiées, formant le domaine “*Nucleotide-Binding adaptor shared by APAF-1, R proteins, and CED-4*” (**NB-ARC**). Ces régions conservées sont les domaines **NBS** et deux domaines **ARC** (ARC1, ARC2) (Figure 16 ; Bent *et al.* 2007). Plusieurs motifs hautement conservés composent le domaine ARC et ils ont été étudiés en détail à partir d'un point de vue fonctionnel. Le motif **P-loop** du domaine NBS est nécessaire pour la liaison des nucléotides ; et des mutations dans ce motif entraînent la perte de fonction de plusieurs protéines NBS-LRR (Williams *et al.* 2011). Par ailleurs, l'auto-activation de nombreuses protéines NBS-LRR est déterminée par des mutations dans le motif **MHD** (Méthionine-Histidine-Aspartate) qui est situé dans le domaine ARC2 et participe à des changements conformationnels (Van Ooijen *et al.* 2008).

Les connaissances actuelles sur les mécanismes d'action moléculaires des protéines NBS-LRR suggèrent un modèle fonctionnel dans lequel le domaine LRR contrôle l'état moléculaire du domaine NB-ARC (Takken *et al.* 2006 ; Van Ooijen *et al.* 2007). Les interactions entre les domaines de la protéine Rx (CNL) de la pomme de terre ont été démontrées, ainsi que les modifications de ces interactions, suite à une stimulation par une protéine Avr (Rairdan *et al.* 2008). Par conséquent, un modèle a été proposé, dans lequel les effecteurs de l'agent pathogène provoquent plus ou moins la perturbation des associations intramoléculaires, ce qui libère les domaines CC, NBS et/ou LRR (Lukasik *et al.* 2009 ; Figure 18). Ces domaines vont ainsi

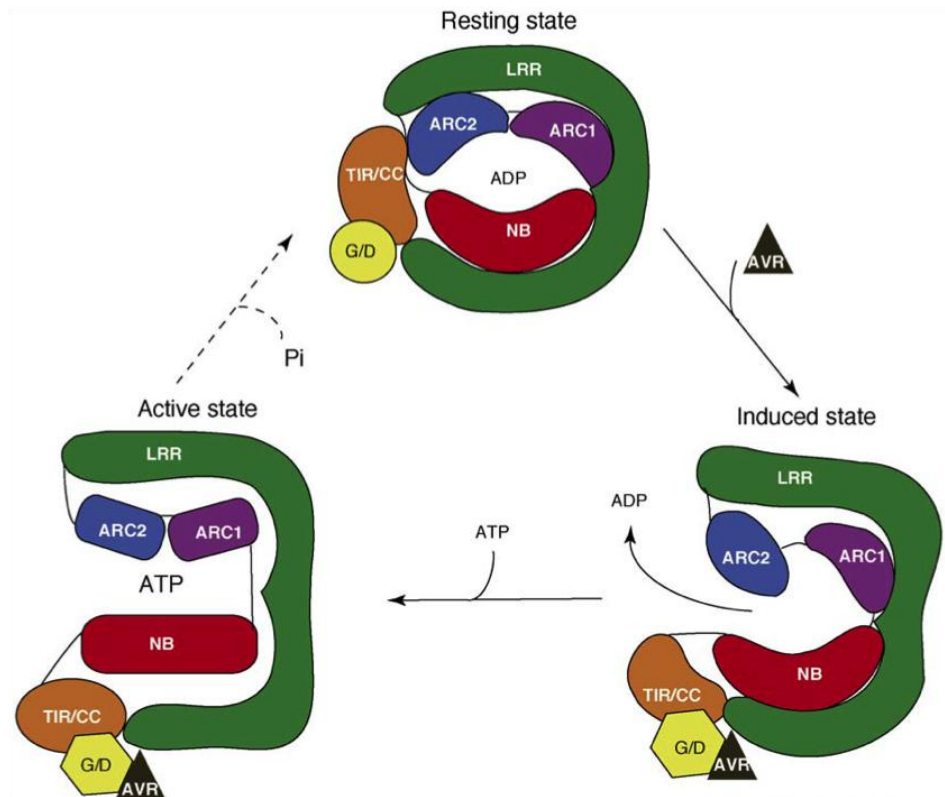
pouvoir reconnaître, permettre le changement de conformation et interagir avec d'autres protéines (Bendahmane *et al.* 2002 ; Moffett *et al.* 2002). Dodds *et al.* (2006) ont montré que la liaison d'un nucléotide est nécessaire pour permettre à la protéine NBS-LRR d'atteindre cet état actif.

Fait intéressant, la permutation des domaines LRR par des paralogues (des gènes issus de plusieurs copies ayant évolué indépendamment) se traduit souvent par une activation constitutive des protéines NBS-LRR (Hwang *et al.* 2003 ; Rairdan *et al.* 2006 ; Van Ooijen *et al.* 2008). Par conséquent, les domaines NBS et LRR doivent avoir co-évolué pour maintenir l'inhibition de l'activation automatique de la protéine NBS-LRR. Une analyse fonctionnelle détaillée a été réalisée très récemment pour la protéine de résistance aux maladies d'*Arabidopsis*, RPS5 (CNL). La protéine RPS5 est activée par le clivage la protéine AvrPphB par la protéine kinase PBS1. Qi *et al.* (2012) ont montré que la substitution du domaine CC de la protéine RPS2 par le domaine CC de la protéine RPS5 n'a pas modifié la spécificité de la protéine RPS5, mais a modérément réduit sa capacité à activer la mort cellulaire programmée ; ce qui suggère que le domaine CC n'a pas de rôle direct dans la reconnaissance du clivage par la kinase PBS1. La troncature du domaine C-terminal de la protéine RPS5 a révélé que la fonction du domaine LRR est de supprimer l'activation de la protéine RPS5 en l'absence du clivage par la kinase PBS1, et l'activation de la protéine RPS5 en sa présence.

Les changements conformationnels pourraient, à leur tour, induire l'oligomérisation ou permettre l'activation des acteurs de la signalisation en aval. L'oligomérisation dépendant d'un domaine NBS fonctionnels a été démontré pour la protéine N du tabac (Mestre *et al.* 2006).

Pour certaines protéines CNL, le domaine CC seul a été montré suffisant pour être parfois capable de déclencher la mort cellulaire. Quelques exemples ont été donnés avec les gènes *RPS2*, *RPS5*, *RPM1* et *ADR1* d'*Arabidopsis*, *NRG1* de *Nicotiana benthamiana* et *MLA10* de l'orge des gènes (Ade *et al.* 2007 ; Tao *et al.* 2000 ; Collier *et al.* 2011 ; Gao *et al.* 2011 ; Maekawa *et al.* 2011). Autrement, pour la protéine Rx qui est une protéine de résistance de type CCEDVID-NB-LRR, son domaine NBS centrale, et non son domaine CCEDVID N-terminal, est suffisant pour induire la mort cellulaire (Rairdan *et al.* 2008).





**Figure 18. Modèle d'activation des protéines R proposé par Lukasik *et al.* (2009)**

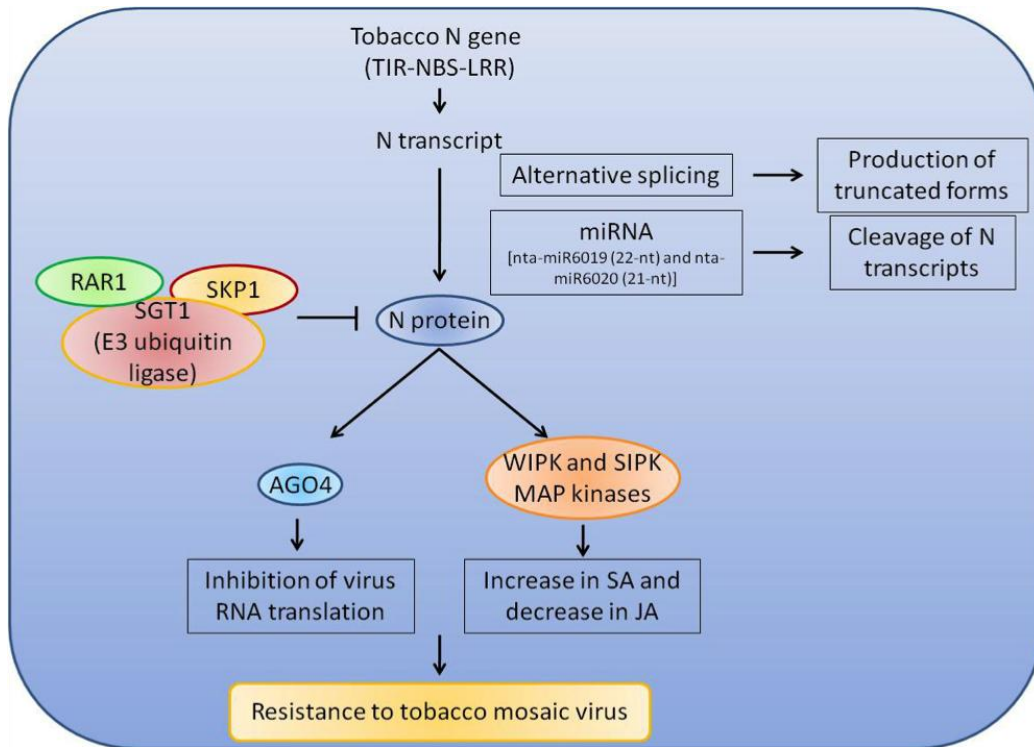
#### 1.4.2.1.2.2. La régulation des gènes NBS-LRR

Un facteur important pour la fonction des protéines NBS-LRR dans la promotion de la résistance aux agents pathogènes est leur localisation subcellulaire. Il a été montré que plusieurs protéines R sont localisées au niveau du cytoplasme et du noyau ; même si aucun signal de localisation nucléaire clair ne peut être identifié dans la plupart des séquences de protéine R. En effet, dans de nombreux cas, les protéines R s'accumulent dans le noyau en réponse à une infection pathogène (Shen *et al.* 2007 ; Meier *et al.* 2011), et cette localisation nucléaire a été démontré comme essentielle pour la résistance dans le cas des protéines MLA10 de l'orge, N du tabac, et RPS4, RRS1-R et SNC1 d'*Arabidopsis* (Shen *et al.* 2007 ; Burch-Smith *et al.* 2007 ; Deslandes *et al.* 2003 ; Wirthmueller *et al.* 2007 ; Cheng *et al.* 2009). Des données récentes ont indiqué que le découplage de la réponse immunitaire de la mort cellulaire est lié à la localisation nucléo-

cytoplasmique de protéines R. Dans une étude récente, Bai *et al.* (2012) ont montré que l'activité de la protéine MLA10, dans la signalisation de la mort cellulaire, est inhibée dans le noyau mais augmentée dans le cytoplasme ; Bien que la présence de MLA10 dans le noyau est suffisante pour permettre la résistance à l'oïdium de l'orge. En outre, ces données suggèrent que la mort cellulaire dépendante et de la résistance aux maladies est dépendante du compartiment cellulaire. Le rôle de la protéine MLA10 dans le noyau serait associé à des interactions spécifiques intramoléculaires et intermoléculaires. Une récente étude a également révélé que le domaine CC de la protéine MLA peut former un homodimère, et que ce dimère définit l'unité fonctionnelle minimale qui est requise pour déclencher la mort cellulaire chez l'orge et *N. benthamiana* (Maekawa *et al.* 2011). En outre, dans le noyau, la protéine MLA10 interagit avec les facteurs de transcription WRKY qui agissent comme des répresseurs de la défense basale MAMP. Cette interaction est donc censée supprimer l'action des facteurs de transcription WRKY sur les défenses induites par les MAMP, et ainsi favoriser une expression plus forte et plus rapide menant à une défense très efficace et une réponse hypersensibilité.

Certaines informations sont également disponibles sur les mécanismes qui contrôlent la résistance aux virus par NBS-LRR. Une réponse antivirale qui inhibe la traduction des protéines codées par le virus chez *N. benthamiana* a été démontrée pour la protéine N (Bhattacharjee *et al.* 2009). Lors de l'activation de la protéine NBS-LRR, les transcrits viraux peuvent s'accumuler mais ne pas s'associer aux ribosomes, et l'inhibition de cette interaction est médiée par des gènes AGO4. Les protéines Argonautes, déjà connues pour leur rôle dans l'ARN interférence, régulent également alternatif. Leur rôle, dans le contrôle de la traduction spécifique des transcrits viraux, a été proposé comme un facteur clé dans la résistance aux virus médiée par les protéines NBS-LRR.

L'action du gène *N* implique également une cascade de protéines MAPK. Des données récentes ont montré que la fonction des MAPK WIPK et SIPK régulent négativement la résistance locale au virus de la mosaïque du tabac, d'une part ; mais ils régulent positivement la résistance systémique, d'autre part. L'utilisation de cultivars de tabac *knockout* ou qui expriment gène *N* a suggéré que la fonction des kinases WIPK et SIPK dans la résistance des plants de tabac au virus de la mosaïque du tabac est contrôlée par le gène *N* (Kobayashi *et al.* 2010). Un résumé des mécanismes de régulation du gène *N* est présenté dans la Figure 19.



**Figure 19. Le résumé du mode d'action et des mécanismes régulant l'expression du gène N à différents niveaux. SA : Acide Salicylique ; JA : Acide Jasmonique**

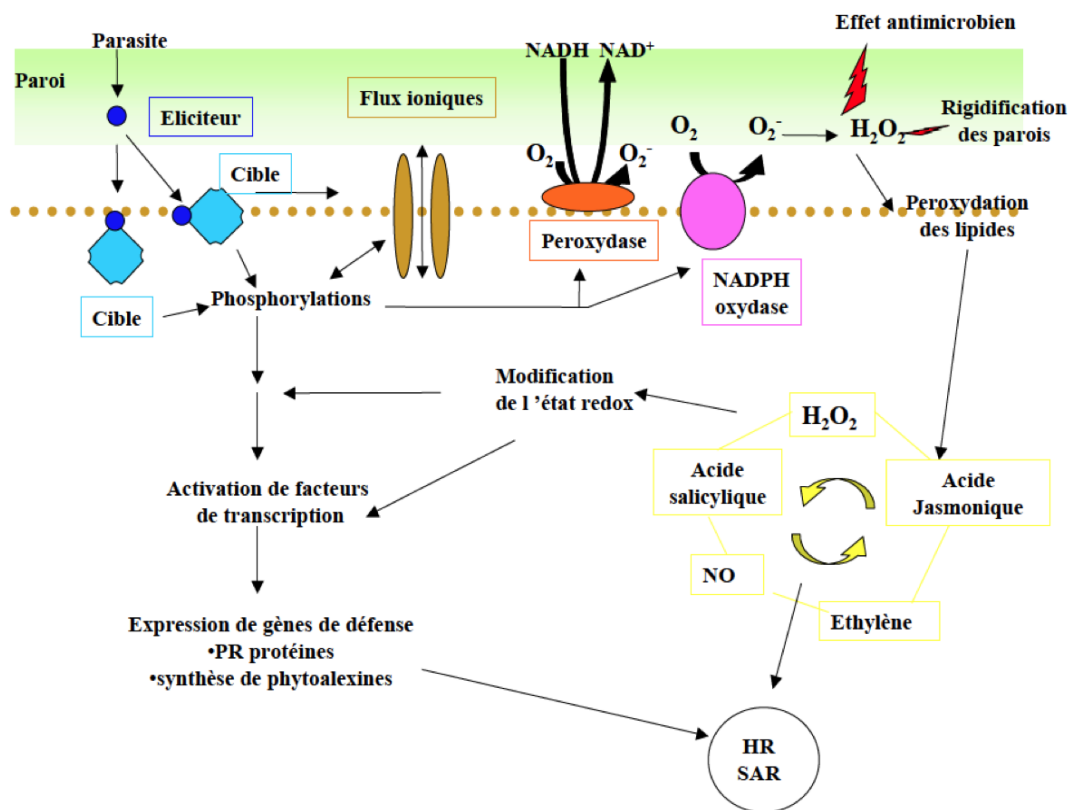
#### 1.4.2.2. Les voies de signalisation impliquées dans la mise en place de la défense de la plante

Les immunités induites (PTI/ETI) activent plusieurs mécanismes de défenses qui font intervenir les défenses passives et des mécanismes de défenses actives qui empêchent le développement du pathogène en le confinant au site de l'infection ou provoquant sa mort.

Les réponses cellulaires les plus fréquentes sont des modifications de flux ioniques (influx de  $\text{Ca}^{2+}$ , efflux de  $\text{Cl}^-$  et  $\text{K}^+$ ), de production de formes actives de l'oxygène (ROS) et des mécanismes de phosphorylation/déphosphorylation (*Mitogen-Activated Protein Kinase* ou MAPKKK, MAPKK et MAPK) (Figure 20) (Zhao *et al.* 2005 et Walters *et al.* 2007).

### 1.4.2.2.1. Les flux ioniques

Les réponses les plus rapides lors d'une attaque d'agents pathogènes se situent au niveau des membranes plasmiques des cellules végétales où des modifications de perméabilité provoquent des échanges ioniques entre les milieux intra- et extracellulaires. Ces flux d'ions peuvent impliquer des protéines G, liées aux récepteurs membranaires qui interviennent dans la transduction précoce du signal en utilisant le Guanosine Diphosphate (GDP) ou Triphosphate (GTP). Ces changements incluent des influx de  $\text{Ca}^{2+}$  et de  $\text{H}^+$  et des efflux de  $\text{K}^+$ ,  $\text{Na}^+$  et  $\text{Cl}^-$  par des canaux ioniques (Zhao *et al.* 2005). Le Calcium semble jouer le rôle de messager secondaire pour l'ouverture des canaux ioniques et l'activation des protéines Kinases calcium-dépendantes (Figure 20).



**Figure 20. Schéma d'élicitation des réactions de défense chez les plantes**

<http://vegetal.snv.jussieu.fr/documents/M2%20defense-plantes%20Dellagi.pdf>

#### **1.4.2.2.2. Les espèces réactives d'oxygène (ROS)**

La HR provient d'un mécanisme appelé “*burst oxydatif*”. Ce mécanisme correspond à une brusque augmentation de la consommation en oxygène qui a pour finalité la production de monoxyde d'azote (NO), d'anion superoxyde ( $O_2^-$ ), de perhydrol ( $H_2O_2$ ), hydroxyle ( $OH^-$ ) et l'oxydation du glucose par le *shunt* des hexoses monophosphates (Figure 20). L'activation d'une enzyme, la NADPH oxydase, est la clé de ce *burst*. Elle permet de catalyser la transformation de l' $O_2$  en  $O_2^-$  en transformant le NADH en  $NAD^+ + H^+$ . Il y a également l'activation d'enzymes antioxydantes incluant des ascorbate-peroxydases, des glutathione-peroxydases, des catalases et des antioxydants non-enzymatiques comme l'acide ascorbique, le glutathion et la riboflavine (Delledone *et al.* 1998).

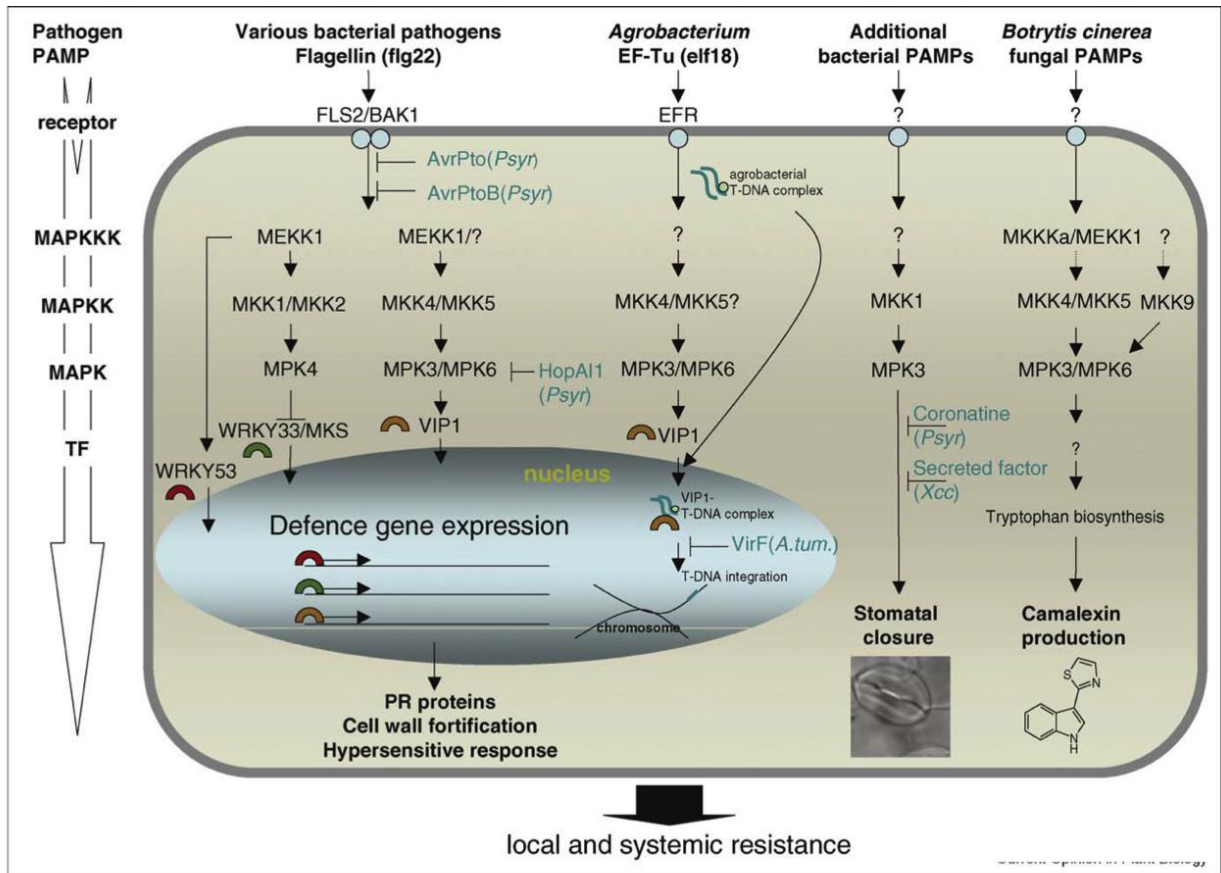
#### **1.4.2.2.3. La voie de transduction des MAP Kinases**

Des cascades de signalisation faisant intervenir des protéines kinases telles que les *Mitogen-Activated Protein Kinase* (MAPK) sont rapidement activées en réponse à des interactions plante - pathogène (Pitzschke *et al.* 2009). La cascade de MAPK est une composante importante située en aval des récepteurs et en amont des facteurs de transcription, lors de la perception des PAMP (Figure 21). La MAP kinase (MAPK) est activée par une MAPK kinase (MAPKK), qui, à son tour, est activée par une MAPKK kinase (MAPKKK). L'activation ou l'inactivation des acteurs de voies de signalisation est réalisée par phosphorylation ou déphosphorylation.

#### **1.4.2.2.4. Les molécules "signal"**

À l'issue des phénomènes précédemment décrits, certaines voies métaboliques secondaires sont activées et permettent de généraliser la réponse dans la plante entière, au moment de la mise en place de la SAR (Raynal *et al.* 1980 ; Lyon, 2007). La mise en place de la SAR nécessite des signaux à partir des tissus infectés vers les tissus sains. Des molécules comme l'Acide Salicylique (SA), l'Acide Jasmonique (JA), l'Éthylène (ET), le perhydrol ( $H_2O_2$ ), l'Éthylène et le monoxyde d'azote (NO) participent à ces différentes voies de signalisation. Ces molécules s'accumulent rapidement dans la cellule et permettent l'activation de l'expression des gènes de défense (Dong, 1998 et Pieterse ; Van Loon, 1999). Leur application de façon exogène peut

également déclencher les mécanismes de la SAR. Elles peuvent donc jouer un rôle d'éliciteur.



**Figure 21. Les cascades de MAPK induites par les PAMP dans la défense des plantes aux agents pathogènes bactériens et fongiques (d'après Pitzschke *et al.* 2009).** La perception des PAMP permet l'activation des cascades de MAPK qui conduisent à l'expression des gènes de résistance et à la fermeture des stomates. Les effecteurs des agents pathogènes interférant avec la transduction du signal sont en bleu ; les facteurs de transcription activés (TF) sont représentés par des demi-cercles ; les récepteurs et composants des cascades de MAPK inconnus sont indiqués par "?".

#### 1.4.2.2.4.1. Régulation de l'expression et de l'activité des gènes NBS-LRR

Les gènes NBS-LRR doivent être finement régulés pour garantir des réponses de résistance correctes, tout en limitant les coûts métaboliques et des effets néfastes sur la croissance des plantes. Cette régulation a lieu à différents niveaux allant de la transcription à la régulation post-traductionnelle.

- *Régulation de l'expression des gènes NBS-LRR*

La régulation au niveau transcriptionnel est destinée à contrôler le niveau d'expression et la viabilité des gènes NBS-LRR. Selon les données de la littérature, plusieurs gènes de résistance sont régulés par des processus tel que l'épissage alternatif de leurs transcrits. Plusieurs cas de régulation de gènes de résistance ont été associés aux TNL, avec quelques exemples rapportés pour les CNL chez les céréales (Mastrangelo *et al.* 2012). Il a été montré que des transcrits de gènes *R* sont synthétisés sous formes tronquées. C'est le cas du gène *RPS6* (*Resistance to P. syringae 6*) d'*Arabidopsis*, de qui trois transcrits peuvent être obtenus par épissage alternatif. Alors que la protéine complète contient les domaines fonctionnels TIR, NBS et LRR, la transcription du gène *RPS6* est caractérisée par un codon stop prématuré ; et il en résulte des protéines tronquées, avec seulement un ou deux domaines (Kim *et al.* 2009). D'autres exemples ont été passés en revue par Mastrangelo *et al.* (2012). Par ailleurs, pour plusieurs protéines tronquées, produites par épissage alternatif, des effets positifs sur la résistance ont été démontrés, comme pour les gènes *RLM3* (Staal *et al.* 2008), *N* (Dinesh-Kumar *et al.* 2000), *RPP1* (Weaver *et al.* 2006) et *RPS5* (Dangl *et al.* 2001), et pour le gène *RPS4* (Zhang *et al.* 2003). L'avantage de ces protéines tronquées peut être obtenu par la réduction de l'auto-inhibition des protéines pleine longueur (Zhang *et al.* 2003) ou par leur fonctionnement dans les événements de signalisation en aval (Duque *et al.* 2011). Xu *et al.* (2011) ont analysé la fonction du gène *MOS14* d'*Arabidopsis* qui code un récepteur nucléaire pour des protéines riches en Serine-Arginine. La perte de fonction du gène *MOS14* résulte en l'altération de l'épissage des transcrits des gènes *SNCI* et *RPS4* ; et elle compromet la résistance médiée par les protéines *snc1* et *RPS4*. Ce résultat suggère que l'importation dans le noyau cellulaire de protéines riches en Serine-Arginine par la protéine *MOS14* est nécessaire dans l'immunité des plantes. Ainsi, l'épissage alternatif pourrait participer à l'amplification et à l'évolution de la résistance des plantes face à des stress biotiques et abiotiques (Mastrangelo *et al.* 2012).

- *Régulation post-transcriptionnelle des gènes NBS-LRR*

Chez les plantes et d'autres organismes, l'ARN interférence est un mécanisme hautement conservé entre espèces qui permet d'éteindre l'expression d'un gène par la dégradation des séquences homologues aux transcrits ou par l'arrêt de la traduction de l'ARNm cible (Ryther *et al.* 2004 ; Mallory *et al.* 2008). Deux classes des petites molécules d'ARN ont été trouvées

comme étant des régulateurs séquence spécifique post-transcriptionnelle :

- les siRNA (*short interfering RNA*) représentent des petites molécules d'ARN double brin de 21-22 nucléotides générés à partir d'un long double brin d'ARN.

- les miARN (*micro-RNA*) représentent une classe d'ARN simple brin de 19-25 nucléotides, codés par le génome de la plupart des organismes multicellulaires. Ces petits ARN sont obtenus par le clivage de précurseurs d'ARN *via* l'action de la nucléase DICER (Baulcombe *et al.* 2004). Les miARN forment des complexes ribo-nucléoprotéiques avec des protéines Argonautes (AGO) qui peuvent se lier, par appariement de bases, à un transcrit cible (Bartel *et al.* 2009 ; Voinnet *et al.* 2009). Aussi, les miARN et leurs précurseurs peuvent se déplacer d'une cellule à une autre et réguler les réponses biologiques dans des cellules adjacentes (Bari *et al.* 2006 ; Pant *et al.* 2008). La régulation de l'expression des gènes aux niveaux transcriptionnel et post-transcriptionnel est impliquée dans la défense de la plante contre les agents pathogènes (Yi *et al.* 2007). Li *et al.* (2012) ont identifié des précurseurs de deux miARN [nta-miR6019 (22 nt et nta-miR6020)] qui permettent le clivage spécifique des transcrits du gène de résistance *N* du tabac. Cette résistance est atténuée lorsque le gène *N* est co-exprimé avec nta-miR6019 et nta-miR6020. En outre, en utilisant une approche bioinformatique (*in silico*), plusieurs familles de miARN ont été identifiées chez trois espèces de *Solanaceae*. Ces miARN pourraient cliver les transcrits de gènes *R* fonctionnels. Dans une étude réalisée chez la tomate, Shivaprasad *et al.* (2012) ont identifié une cascade de régulation pouvant affecter la résistance aux maladies. Les initiateurs de cette cascade appartiennent à une superfamille diversifiée, et ils ciblent spécifiquement la séquence des transcrits codant le motif P-loop dans les protéines de résistance NBS-LRR. Suite à cette interaction, le transcrit du gène NBS-LRR est morcelé en siRNA secondaires. Au moins un de ces siRNA secondaires cible les transcrits d'une protéine liée à la défense (Shivaprasad *et al.* 2012). L'ensemble de ces résultats montrent un rôle conservé des miARN et siRNA secondaires dans la régulation des gènes NBS-LRR et dans la résistance aux agents pathogènes chez les solanacées.

- *Régulation post-traductionnelle des gènes NBS-LRR*

Les réponses des plantes aux agents pathogènes doivent être bloquées en l'absence de l'agent pathogène, afin d'éviter l'auto-immunité, ce qui peut être préjudiciable à la croissance et au développement des plantes. Un exemple est celui du gène *Snc1* d'*Arabidopsis* qui code une



protéine *R* de type TNL (Zhang *et al.* 2003). Le mutant *Snc1*, qui se caractérise par une mutation gain de fonction qui se trouve dans la région entre les domaines NBS et LRR, montre des niveaux élevés d'acide salicylique et une expression constitutive de gènes liés à la pathogénèse (Li *et al.* 2001). Le phénotype associé à ces mutants est le nanisme. La fonction de certaines protéines étudiées consiste à bloquer l'action des protéines *R* en l'absence de l'agent pathogène. Il a été montré que certaines de ces protéines interagissent avec des protéines chaperons. Chez *Arabidopsis*, les protéines RAR1 (requis pour la résistance de ML-A12 résistance), SGT1 (suppresseur de l'allèle G2 de Skp1) et HSP90 (protéine de choc thermique 90) sont impliquées dans le correct repliement des protéines NBS-LRR. L'accumulation des protéines MLA de l'orge, Rx de la pomme de terre, et RPM1 et RPS5 d'*Arabidopsis* a été réduite lorsque la fonction de la protéine RAR1 est compromise (Tornero *et al.* 2001 ; Bieri *et al.* 2004 ; Holt *et al.* 2005). De même, compromettre l'activité de la protéine HSP90 a également entraîné une réduction de l'accumulation de plusieurs protéines *R*, y compris RPM1, RPS5 et Rx (Hubert *et al.* 2003 ; Lu *et al.* 2003 ; Holt *et al.* 2005). Li *et al.* (2010) ont identifié une protéine conservée, SRFR1 (suppresseur de RPS4-RLD 1), qui interagit avec la protéine SGT1 et joue un rôle dans la répression des réponses immunitaires ; celles-ci sont en effet activées de façon constitutive chez les mutants 'perte de fonction' *srfr1*.

La protéine SGT1 est une ubiquitine ligase E3, qui est un composant du système d'ubiquitination, une voie de régulation post-traductionnelle dans laquelle l'ubiquitine se lie à des résidus de lysine des protéines cibles et favorise leur dégradation par le proteasome/26S. Il a été démontré par une analyse mutationnelle que la protéine AtSGT1b est nécessaire à la résistance contre *Peronospora parasitica* chez *Arabidopsis* (Austin *et al.* 2002 ; Tör *et al.* 2002). Comme les interactions entre SGT1 et RAR1 ont été démontrées chez *Arabidopsis*, un modèle a été proposé dans lequel SGT1 participe à la dégradation de RPM1 afin de contrôler la taille de la lésion lors de la réponse d'hypersensible et à l'amplitude de la réponse au site de l'infection (Dielen *et al.* 2010). L'implication des protéines RAR1 et SGT1 dans les mécanismes de défense a également été soulignée chez plusieurs autres plantes (Dielen *et al.* 2010).

### **1.4.3. Organisation et évolution des gènes NBS-LRR dans le génome des plantes**

#### **1.4.3.1. Organisation génomique et évolution des gènes NBS-LRR**

##### **1.4.3.1.1. Les gènes de résistances NBS-LRR complets**

Les analyses génomiques comparatives ont montré que les génomes de plantes peuvent coder plusieurs centaines de gènes NBS-LRR, et qu'il y a une grande diversité dans le nombre et la répartition des différentes classes de ces gènes. À ce jour, un grand nombre de séquences codant des NBS-LRR ont été isolés à partir de diverses espèces végétales grâce à des analyses du génome entier : d'une cinquantaine de gènes chez *Carica papaya* et *Cucumis sativus* (Porter *et al.* 2009 ; Wan *et al.* 2013), à 653 chez *Oryza sativa* (Shang *et al.* 2009). Les données présentées dans le tableau 4 sont sujettes à une évolution rapide, liés aux nombreux projets de séquençage de génomes en cours, mais elles donnent d'ores et déjà une idée de la taille et de l'organisation de cette famille de gènes dans les génomes des plantes. Sauf pour *Carica papaya* dont les gènes NBS-LRR sont bien répartis entre les groupes de liaison (Porter *et al.* 2009), la distribution chromosomique des gènes NBS-LRR semble être très irrégulière dans la plupart des espèces étudiées, avec des chromosomes caractérisées par beaucoup plus de gènes NBS-LRR que d'autres. Chez la pomme de terre, par exemple, le plus grand nombre de gènes NBS-LRR se trouvent sur les chromosomes 4 et 11 (environ 15 % des gènes cartographiés), avec le plus petit nombre sur le chromosome 3 (1 %) (Jupe *et al.* 2012 ; Lozano *et al.* 2012). Chez *Brachypodium distachyon*, le chromosome 4 contient environ un tiers du total des gènes NBS-LRR qui ont été identifiés (Tan *et al.* 2012). Chez *Brassica rapa* les chromosomes 3 et 9 contiennent plus de la moitié des gènes NBS-LRR cartographiés (Mun *et al.* 2009). (Kang *et al.* (2012) ont rapporté que dans le génome du soja (*Glycine max*), le chromosome 16 a le plus grand nombre de gènes NBS-LRR. Enfin, chez *Medicago truncatula*, plus de 54 % des gènes NBS-LRR sont localisés par les chromosomes 3, 4 et 6 (Ameline-Torregrosa *et al.* 2008).

**Tableau 4. Gènes et pseudogènes codant des gènes *NBS-LRR* identifiés dans différents génomes de plantes**

Espèces	Total de gènes NBS-LRR	TNL	CNL	Nombre de pseudogènes	Références
<i>Arabidopsis lyrata</i>	185	123	38	–	Guo <i>et al.</i> 2011
<i>Arabidopsis thaliana</i>	149	94	55	10	Meyers <i>et al.</i> 2003
	159	98	50	–	Guo <i>et al.</i> 2011
<i>Brachypodium distachyon</i>	126	0	113	–	Tan <i>et al.</i> 2012
<i>Brassica rapa</i>	92	62	30	–	Mun <i>et al.</i> 2009
<i>Carica papaya</i>	54	7	6	–	Porter <i>et al.</i> 2009
<i>Cucumis sativus</i>	57	13	18	–	Wan <i>et al.</i> 2013
<i>Glycine max</i>	319	–	–	–	Kang <i>et al.</i> 2012
<i>Lotus japonicus</i>	158	32	28	62	Li <i>et al.</i> 2010
<i>Medicago truncatula</i>	333	156	177	49	Ameline-Torregrosa <i>et al.</i> 2008
<i>Oryza sativa L. spp. Indica</i>	653	–	–	184	Shang <i>et al.</i> 2009
<i>Oryza sativa L. spp. Japonica</i>	553	–	–	150	Shang <i>et al.</i> 2009
<i>Populus trichocarpa</i>	402	91	119	161	Köhler <i>et al.</i> 2008
<i>Solanum tuberosum</i>	438	77	361	–	Jupe <i>et al.</i> 2012
	435	65	370	179	Lozano <i>et al.</i> 2012
<i>Vitis vinifera</i>	459	97	203	–	Yang <i>et al.</i> 2008

‘–’ : la donnée n’est pas présente

La distribution des sous-groupes CNL et TNL est également variable entre les différentes espèces végétales. L'exemple le plus frappant est l'absence quasi-totale de gènes TNL chez les monocotylédones. Les gènes TNL sont plutôt présents dans le génome des dicotylédones, et souvent en plus grand nombre par rapport aux gènes CNL. Les génomes des espèces comme *Arabidopsis thaliana*, *Arabidopsis lyrata* et *Glycine max* contiennent deux à six fois plus de TNL que de CNL (Guo *et al.* 2011 ; Kang *et al.* 2012), avec le contraire pour le génome de la pomme de terre (Jupe *et al.* 2012 ; Lozano *et al.* 2012) et *Medicago truncatula* (Ameline-Torregrosa *et al.* 2008) qui présentent un grand nombre de CNL.

Dans le génome, les gènes NBS-LRR sont organisés soit en tant que gènes isolés, soit en groupes ou *clusters* de différentes tailles qui sont censés faciliter l'évolution rapide des gènes *R* (Hulbert *et al.* 2001). Ces *clusters* peuvent être divisés en deux types en fonction des relations phylogénétiques de leurs membres:

- (i) les *clusters* qui contiennent des gènes NBS-LRR provenant de duplications en tandem. Les gènes NBS-LRR se regroupent alors dans la même branche des arbres phylogénétiques.
- (ii) Les *clusters* qui contiennent des gènes NBS-LRR de différentes branches d'arbres phylogénétiques (McDowell *et al.* 2006), où les gènes sont issus de duplication (ectopique ou segmentaire), de transposition, avec des réarrangements locaux ultérieurs (Leister *et al.* 2004).

La grande majorité des gènes NBS-LRR sont retrouvés dans des *clusters*. A titre d'exemples, Li *et al.* (2010) indiquent que dans l'ensemble du génome de *L. japonicus*, 38,2 % des gènes NBS cartographiés appartiennent à huit *clusters*, 70 % des gènes NBS chez le peuplier appartiennent à 75 *clusters*, dans le génome de riz et de *B. distachyon* la moitié des gènes NBS-LRR identifiées sont en *cluster* (Yang *et al.* 2008 ; Tan *et al.* 2012). Cette proportion peut être encore plus chez d'autres espèces, comme *V. vinifera* (83,2 %, Yang *et al.* 2008) ou *M. truncatula* (80 % ; Ameline-Torregrosa *et al.* 2008). Cette proportion de gènes en *clusters* peut évoluer assez rapidement puisqu'elle peut être différente entre des espèces apparentées. Ainsi, Guo *et al.* (2011) ont identifié 159 et 185 gènes NBS-LRR chez *A. thaliana* et *A. lyrata*, respectivement, avec 71,1 % et 63,8 % de gènes organisés en *clusters* pour ces deux espèces, respectivement. Avec un plus grand nombre de *cluster* par rapport à *A. lyrata* (38 contre 35).

Bien que la majorité des *clusters* de gènes NBS-LRR soient composés de séquences ayant les mêmes domaines protéiques, un nombre non négligeable est hétérogène. En effet, environ 25 % des *clusters* chez *M. truncatula*, le riz, la pomme de terre et *Arabidopsis* comprennent à la fois les CNL et les TNL (Ameline-Torregrosa *et al.* 2008 ; Shang *et al.* 2009 ; Jupe *et al.* 2012 ; Guo *et al.* 2011).

#### **1.4.3.2. Les pseudogènes**

Les pseudogènes sont communément définis comme des séquences qui ressemblent à des gènes connus, mais qui ne peuvent pas produire des protéines fonctionnelles. Ces pseudogènes proviennent de l'évolution par duplication suivie d'évènements de mutations (substitutions, insertions ou délétions) de gènes codants des protéines (Chandrasekaran *et al.* 2008). Ces pseudogènes ne peuvent être identifiés précisément que dans le cadre de travaux d'annotation

expertisée de familles de gènes (Yang *et al.* 2006 ; Luo *et al.* 2012). Leur pourcentage dans la famille des gènes NBS-LRR est variable suivant les espèces considérées : de faibles pourcentages chez *A. thaliana* (8,05 %) (Meyers *et al.* 2003) et *M. truncatula* (14,7 %) (Ameline-Torregrosa *et al.* 2008) à des pourcentages élevés chez le coton polyploïde (24,6 %) (He *et al.* 2004) ou le riz diploïde (48-56% suivant la variété considérée ; Luo *et al.* 2012), ce qui exclue la polyploïdie comme seul vecteur d'une expansion en pseudogènes. Dans la plupart des cas, les pseudogènes sont caractérisés par des délétions importantes (Luo *et al.* 2012), qui peuvent être produites par des événements de transposition ou des sauts d'exon, liés à des changements de cadres de lecture pouvant également conduire à la synthèse d'ARNm et de protéines tronquées (Lozano *et al.* 2012). Ceci illustre la difficulté spécifique de l'annotation de cette famille de gènes qui est en général faiblement exprimée et dans des conditions assez spécifiques et est donc proportionnellement faiblement représentée dans les collections d'EST (Meyers, Morgante & Michelmore, 2002). Par ailleurs, il existe des preuves de l'expression de pseudogènes chez certaines espèces, y compris pour le riz (Monosi *et al.* 2004), le pin (Liu *et al.* 2003) et *M. truncatula* (Ameline-Torregrosa *et al.* 2008). En particulier, certains pseudogènes identifiés dans ces espèces ont des alignements presque parfaits (99 % - 100 % d'identité) avec des bases de données EST.

Enfin, des fonctions peuvent également être attribuées aux protéines partielles NBS-LRR. Chez la souris, un gène tronqué est impliqué dans la stabilité de l'ARNm de son homologue fonctionnel (Hirotsumi *et al.* 2003). Lozano *et al.* (2012) ont proposé un rôle de molécules adaptatrices pour les pseudogènes identifiés chez la pomme de terre, par lequel ils peuvent recruter ou interagir avec d'autres protéines NBS-LRR. La même hypothèse a été formulée pour les gènes *R* tronqués identifiés dans *Populus* (Köhler *et al.* 2008). Polypeptides de *R* tronqués peuvent provenir aussi par épissage alternatif, et un rôle dans la promotion de la résistance aux maladies a été démontré pour beaucoup d'entre eux, tel que révisé par Mastrangelo *et al.* (2006). Dans certains cas, les pseudogènes diffèrent des gènes NBS-LRR fonctionnels par la longueur du domaine NBS, qui peut être excessivement réduite, comme chez *L. japonicus* (Li *et al.* 2010) et *B. distachyon* (Tan *et al.* 2012).

Le nombre de pseudogènes varie également selon le sous-groupe CNL ou TNL mais en général de façon proportionnelle à l'abondance relative des gènes CNL et TNL. À titre d'exemple, sur

179 pseudogènes identifiés dans la pomme de terre, 156 (87 %) appartiennent au groupe CNL et 23 (13 %) au groupe TNL ; mais cela a été prévu que 85 % de l'ensemble des gènes NBS-LRR sont CNL et seulement 15 % sont TNL (Lozano *et al.* 2012).

Concernant la distribution des pseudogènes dans le génome, les gènes NBS-LRR tronqués sont souvent retrouvés à côté des gènes NBS-LRR complets, à l'intérieur d'un intervalle de 100 Kb, et donc distribués de la même manière que les gènes fonctionnels ; en effet, ces pseudogènes sont aussi majoritairement retrouvés dans les *clusters* de gènes NBS-LRR (Ameline-Torregrosa *et al.* 2008 ; Lozano *et al.* 2012).

Ainsi, les pseudogènes peuvent être considérés à la fois comme des gènes non fonctionnels qui seront éliminés du génome ou comme des réservoirs de diversité génétique pour des mécanismes de recombinaison ou de conversion génique (Meyers *et al.* 1999) ou bien comme des réservoirs de plasticité fonctionnelle (Michelmore and Meyers, 1998).

#### **1.4.3.3. Evolution des gènes NBS-LRR chez les plantes**

Les gènes NBS-LRR représentent l'une des plus nombreuses et des plus anciennes familles de gènes chez les plantes. Des mécanismes comme les duplications, les *crossing over* inégaux, la recombinaison ectopique, la conversion génique et la diversification de sélection ont été proposés comme contribuant aux structures des *clusters* de gènes *R* et à l'évolution des spécificités de résistance (Michelmore *et al.* 1998 ; Young *et al.* 2000 ; Sun *et al.* 2001). Cependant, la vitesse d'évolution et les effets de sélection ne sont pas homogènes au niveau des différents domaines d'un même gène, ainsi que pour les membres des différents sous-groupes de la famille de gènes. En ce qui concerne ce premier aspect, le taux de conversion génique semble avoir été plus élevé pour les domaines LRR que pour les régions NBS, qui semblent être soumis à la sélection purification (Kuang *et al.* 2004). La grande variabilité des domaines LRR est liée à leur rôle dans la reconnaissance spécifique des protéines Avr afin de promouvoir la résistance des plantes aux agents pathogènes. En effet, des taux élevés de substitutions de nucléotides ont été trouvés dans des domaines LRR, ce qui indique qu'il y aurait une diversification importante au sein de ces domaines (Mondragon-Palomino *et al.* 2002). A titre d'exemple, il y a en moyenne 14 LRR par protéine, et souvent 5 - 10 variations de séquences pour chaque répétition.

Par conséquent, dans une espèce comme *Arabidopsis thaliana*, il y a potentiellement bien plus de  $9.10^{11}$  variantes (McHale *et al.* 2006).

Par ailleurs, Kuang *et al.* 2004 proposent de classer les gènes NBS-LRR en deux classes : les gènes de type I, qui ont évolué rapidement avec des conversions géniques fréquentes entre eux et les gènes de type II qui ont évolué lentement, avec de rares événements de conversion génique entre clades. L'un des facteurs qui influencerait la conversion génique est l'organisation des gènes *R* en *cluster* puisque les plus grands *clusters* seraient des donneurs de séquences potentiels. Cependant, cette hypothèse n'est pas toujours vérifiée. Guo *et al.* (2011) ont montré des corrélations positives et significatives entre la taille du *cluster* et la fréquence de conversion génique chez *A. thaliana* mais pas chez *A. lyrata*, où la taille médiane des régions de conversion est élevée.

Les domaines NBS et les domaines LRR sont toujours trouvés dans des gènes séparés dans les génomes de bactéries, des archées, les protistes et les algues (Yue *et al.* 2012) mais des gènes NBS-LRR ont été trouvés chez deux espèces de bryophytes : *Physcomitrella patens* et *Marchantia polymorpha* (Xue *et al.* 2012). En plus des CNL et TNL, ils définissent deux nouvelles classes de gènes NBS-LRR. Le premier a été identifié dans le génome de *P. patens*, et il possédait un domaine protéine kinase à l'extrémité N-terminale (PK-NBS-LRR ; PNL). Le second a été identifié dans le génome de *M. polymorpha*, et il possédait par un domaine □/□-hydrolase à l'extrémité N-terminale (hydrolase-NBS-LRR; HNL). La classe des CNL s'est révélée être divergente des TNL, PNL et HNL chez bryophytes. En revanche, les CNL de *M. polymorpha* et certains CNL de la mousse *P. patens* forment un groupe monophylétique fortement soutenu (Xue *et al.* 2012). En outre, les CNL de monocotylédones et dicotylédones ont tendance à se regrouper, ce qui suggère une origine des CNL antérieure à la divergence des monocotylédones et dicotylédones (McHale *et al.* 2006 ; Yang *et al.* 2008). Finalement, les domaines TIR sont absents des génomes des céréales, et ils semblent avoir été perdus par ces espèces de céréales, plutôt que d'avoir apparu plus tard, après la séparation des monocotylédone et dicotylédone (Pan *et al.* 2000). Les gènes NBS-LRR semblent donc avoir émergés pour permettre aux premières lignées de plantes terrestres de faire face à des pathogènes spécifiques du milieu aérien.

Le nombre de gènes NBS-LRR unique et dupliqué reflète les petits ou les grands événements de duplication qui ont eu lieu au cours de l'évolution des génomes. Chez *Arabidopsis*, Meyers *et al.* (2003) ont rapporté que les duplications locales et distantes des TNL et CNL sont responsables de l'amplification des *clusters*, ce qui est compatible avec la duplication du génome entier au début de l'évolution de *Arabidopsis*. De même, de grandes régions de duplication ont également été identifiées chez *M. truncatula*, dont le génome a subi une duplication génomique à grande échelle (Ameline-Torregrosa *et al.* 2008). Une séparation claire entre les deux groupes a également été identifiée chez *Brassica rapa*, malgré l'événement de triplication du génome entier qui a eu lieu il ya 11 - 12 millions années, après spéciation (Mun *et al.* 2009). En outre, le nombre de gènes NBS-LRR est un peu plus grand que chez *Arabidopsis*, qui a subi une duplication du génome. Probablement, environ la moitié des gènes NBS-LRR générées après le triplement du génome de *B. rapa* a maintenant été perdue. Une perte similaire de gènes *R* a probablement eu lieu pour le génome de *Populus* chez qui des duplications segmentaires et des réarrangements chromosomiques n'ont abouti qu'à l'amplification de 5 % de gènes NBS-LRR (Köhler *et al.* 2008).

Souvent, la duplication des régions génomiques qui contiennent des gènes NBS-LRR correspond également à une redondance fonctionnelle. La présence de gènes avec des fonctions redondantes au sein de régions dupliquées a été démontrée dans le génome du soja (Kang *et al.* 2012). Cette étude a révélé que 91 gènes NBS-LRR étaient dans dix régions génomiques dupliquées, et que ces régions contenaient des *Quantitative Trait Loci* (QTL) aux maladies dupliquées.

#### **1.4.4. Colocalisation des gènes NBS-LRR avec les *loci* impliqués dans la résistance aux maladies chez les plantes**

Le clonage des gènes de résistance sur la base de cartes génétiques a démontré dans de nombreux cas l'implication des gènes NBS-LRR dans la résistance à un certain nombre d'agents pathogènes. (cf. 2.2.5.2). Cela a conduit à proposer un rôle dans la résistance aux maladies pour un nombre beaucoup plus élevé de gènes appartenant à cette famille en fonction de leur co-ségrégation avec les *loci* de résistance génétiquement cartographiés. Cette correspondance peut être effectuée entre les positions des QTL de résistance aux maladies (i) et les positions des gènes NBS-LRR identifiés dans les génomes séquencés ou (ii) pour les espèces dont le génome



n'a pas été séquencé, les positions de marqueurs moléculaires développés dans des séquences de gènes NBS-LRR.

#### **1.4.4.1. Cas des génomes non séquencés**

Dès que la famille des gènes NBS-LRR a été identifiée comme une famille importante dans la résistance aux maladies chez les plantes, de nombreuses approches de cartographies les utilisant comme gènes candidats ont été développées (Leister *et al.* 1996 ; Shen *et al.* 1998 ; Collins *et al.* 1998 ; Speulman *et al.* 1998 ; Geffroy *et al.* 2000 ; Gebhardt and Valkonen, 2001 ; Calenge *et al.* 2005 ; Yang *et al.* 2012). Cependant, dans cette première génération d'études, le nombre de marqueurs dérivés de gènes NBS-LRR cartographiés était encore un facteur limitant. Récemment, Marone *et al.* (2012) ont analysé les séquences de 2 000 marqueurs DArT (*Diversity Array Technology*) qui ont été développés chez le blé et qui avaient déjà été utilisés comme marqueurs anonymes pour la construction de plusieurs cartes génétiques. Les auteurs ont constaté qu'une forte proportion de clones DArT correspondait ou était située à proximité des séquences liées à la résistance aux maladies chez les plantes ; et en particulier proche des gènes NBS-LRR et des protéines kinases dont la plupart contenaient un domaine LRR. Ces données leur ont permis de positionner un nombre élevé de gènes NBS-LRR sur les cartes génétiques de blé. Par exemple, 40 marqueurs qui correspondent à des gènes de résistance ont été positionnés sur les génomes A et B du blé, sur une carte consensus (Marone *et al.* 2012). En particulier, le marqueur DArT qui correspond à un gène NBS-LRR (*WPT-1601*) représente le marqueur situé au niveau du pic d'un QTL majeur de résistance au virus de la mosaïque du blé qui a été cartographié sur le chromosome 2B, dans la population blé dur Meridiano × Claudio (Maccaferri *et al.* 2011). En outre, Russo *et al.* (2012) ont identifié deux marqueurs qui correspondent à des gènes NBS-LRR comme marqueurs liés à des QTL mineurs pour la même maladie sur les chromosomes 3B et 7B dans une autre population de blé dur (Cirillo × Neodur). Par ailleurs, Yu *et al.* (2011) ont montré que le marqueur DArT, wPt-8460 qui est localisé sur le chromosome 2B du blé, est significativement associé à la résistance à la rouille de la tige par une analyse de cartographie d'association. La séquence de ce marqueur n'est pas disponible, mais il a été placé à 3 cM du marqueur WPT-0189 dans une population en ségrégation de différents de blé dur (*Creso* × *Pedroso*) (Marone *et al.* 2012), qui correspond à un gène NBS-LRR.

Ces études sont d'une grande importance, car elles fournissent des gènes candidats pour plusieurs *loci* de résistance, ainsi que des marqueurs moléculaires étroitement liés qui peuvent être utilisés dans la sélection assistée par marqueurs.

#### **1.4.4.2. Cas des génomes entièrement séquencés**

L'accès aux séquences complètes de génome permet maintenant d'avoir une vision complète du rôle de cette famille de gènes dans les résistances aux maladies chez les plantes. Un exemple du premier type d'approche a été réalisé chez le soja. Kang *et al.* (2012) ont analysé les gènes NBS-LRR qui colocalisent avec les QTL de résistance aux maladies et ils ont observé qu'environ 63 % des QTL étaient localisés avec des régions de 2 Mb contenant des gènes NBS-LRR. En particulier, une région sur le chromosome 6 qui contenait des *clusters* de NBS-LRR et à laquelle a été associé à sept QTL de résistance, dont quatre confèrent une résistance à des maladies cryptogamiques, et trois confèrent une résistance aux nématodes. Un autre exemple est celui du chromosome 16, où 19 QTL de résistance (14 QTL de résistance aux maladies cryptogamiques et 5 QTL de résistance aux nématodes) colocalisent avec 40 gènes NBS-LRR dans une région de 2 Mb. Une approche similaire menée par Shang *et al.* (2009) ont identifié les relations entre les pseudogènes NBS-LRR et la sensibilité à la rouille chez le riz. Les auteurs ont analysé une population F2 issue du croisement entre la variété résistante *Oryza sativa ssp indica* et la variété sensible *Oryza sativa ssp japonica*, qui a permis de montrer la corrélation entre le pseudogène *PID3* et la sensibilité.

## 1.5. CONTEXTE ET OBJECTIFS DE THESE

### 1.5.1. Contexte

La vigne *V. vinifera* est la seule espèce cultivée en Europe pour la qualité de ces baies. Cependant, cette espèce est très sensible aux agents pathogènes majeurs de la vigne, importés en Europe du continent Américain comme l'oïdium (*E. necator*), le mildiou (*P. viticola*) et le phylloxéra (*D. vitifolia*) (Riaz *et al.* 2007). Sur le continent américain les viticulteurs doivent également faire face à la maladie de 'Pierce' (*X. fastidiosa*). Avec seulement deux exceptions connues chez *V. Vinifera* (Hoffmann *et al.* 2007 ; Coleman *et al.* 2009), les résistances génétiques à ces maladies sont trouvées dans les espèces sauvages du genre *Vitis*. Or, si la viticulture ne représente que 3 % de la surface cultivée en France, elle utilise 20 % en valeur des produits phytosanitaires, dont 80 % de fongicides (Phytowelt GmbH 2003 ; Expertise scientifique collective Inra, Cemagref, 2005). De plus, des résistances à un certain nombre de fongicides couramment utilisés ont été identifiées chez des souches fongiques (Erickson et Wilcox, 1997; Savocchia *et al.* 2004; Baudoin *et al.* 2008) et il est maintenant considéré que l'utilisation généralisée de produits phytosanitaires est à la fois préjudiciable à l'environnement et à la santé humaine. Par conséquent, l'identification des gènes qui confèrent une résistance à ces agents pathogènes et leur introgression dans *V. vinifera* seraient avantageuses tant sur plan économique qu'environnemental. Dans cette optique, une source de gènes de résistance à de nombreux agents pathogènes a été identifiée chez l'espèce *M. rotundifolia* appartenant à un des deux sous-genres du genre *Vitis*. *M. rotundifolia* est domestiquée et cultivée dans le sud-est des États-Unis (Olmo, 1986) pour la production de fruits frais, de jus de fruits et de vin. Toutefois, l'espèce *M. rotundifolia* est séparée d'un point de vue taxonomique des espèces du sous-genre *Euvitis* par quelques caractéristiques morphologiques (Planchon 1887) et par une différence dans le nombre de chromosomes (*Vitis* = 2n = 38, *Muscadinia* = 2n = 40; Branas 1932) qui explique que les tentatives pour produire des croisements interspécifiques entre *M. rotundifolia* et *V. vinifera* ont un succès limité. Pour cette raison, le transfert de gènes de résistance dans les variétés cultivées de *V. vinifera* par la sélection classique reste un défi important chez la vigne (Bisson *et al.* 2002). La cartographie comparée entre *V. vinifera* et *M. rotundifolia* permettrait d'accélérer l'identification et l'isolement des gènes de résistance de *M. rotundifolia* et mieux

comprendre la similitude entre leur génome, tant sur le plan structurel que sur le plan fonctionnel. Cependant, en dépit de son importance pour l'amélioration de la vigne cultivée, les connaissances sur le génome de *M. rotundifolia* sont encore très limitées et les ressources génomiques disponibles publiquement sont rares.

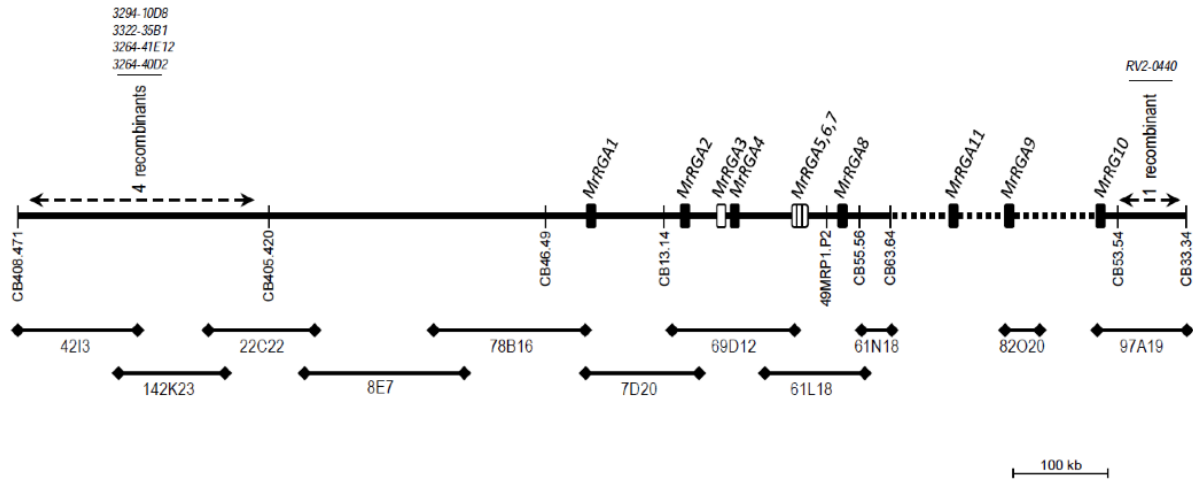
Les objectifs de cette thèse étaient d'évaluer l'intérêt d'une banque BAC du génome de *M. rotundifolia* dans des études de génomique comparée entre *V. vinifera* et *M. rotundifolia* et en particulier dans la région de *clusters* NBS-LRR en lien avec des QTL de résistance à l'oïdium et au mildiou. Ce travail était inscrit dans le cadre du projet MUSCARES (projet ANR) qui a pour but de profiter de l'existence de la séquence du génome de la vigne afin d'accélérer la caractérisation fonctionnelle de la résistance aux agents pathogènes trouvés chez plusieurs espèces de *M. rotundifolia*. Mon travail était complémentaire à celui de Sophie Blanc dont les objectifs étaient (i) de réaliser une analyse comparative des génomes de *V. vinifera* et *M. rotundifolia* par cartographie génétique et (ii) d'identifier de nouveaux facteurs de résistance chez *M. rotundifolia* utilisables à terme en sélection.

Nous nous sommes intéressés à deux régions : (i) la région introgressée sur le chromosome 12 de *V. vinifera* contenant les gènes de résistance *Run1* et *Rpv1* d'origine *M. rotundifolia* (Tableau 3) et (ii) une région du chromosome 18 dans le genre *Vitis* sur laquelle a été cartographié le QTL de résistance *Rpv3* (Tableau 3). Les ressources disponibles étaient très différentes dans les deux cas lorsque j'ai débuté ma thèse.

#### **1.5.1.1. Supercontig de clones BAC de la région du locus de résistance *Run1/Rpv1***

Le locus de résistance *Run1/Rpv1* a été introgressé de *M. rotundifolia* dans *V. vinifera* (Pauquet *et al.* 2001 ; Dry *et al.* 2011) et une banque BAC de ce génotype a été construite, à partir d'une plante résistante issue d'un *backcross* 5, contenant environ **55 000** clones, avec une taille moyenne des inserts d'environ **120** Kb, ce qui représente 5 équivalents génome haploïde (Barker *et al.* 2005). Une trentaine de marqueurs génétiques (RGA, AFLP et microsatellites), étroitement liés au locus de résistance *Run1/Rpv1* (Pauquet *et al.* 2001 ; Donald *et al.* 2002), a été utilisée pour cribler la banque BAC afin de choisir les clones BAC intéressants (Barker *et al.* 2005). Douze clones BAC ont ainsi été identifiés comme délimitant une région contigüe au locus de résistance *Run1/Rpv1*. L'analyse *in silico* de 11 d'entre eux suggère qu'une grande partie de la

séquence serait constituée d'ADN non-codant, de gènes codant des glycosyltransférases et d'éléments transposables. Un *cluster* de RGAs contenait les candidats probables pour les gènes *Run1* et *Rpv1*. Ces gènes codent des protéines de type TIR-NBS-LRR très semblables par leur structure à la protéine *Mla* qui confère la résistance à l'oïdium chez l'orge (Haltermann *et al.* 2001). Les marqueurs CB46.49, CB13.14 et 49MRP1.P2, définis dans la séquence des clones BAC candidats, encadrent le *cluster* de gènes (Donald *et al.* 2002, Barker *et al.* 2005 et Dry *et al.* 2010). Ces clones BAC étaient arrangés en 3 contigs (Figure 22): un supercontig (917 559 pb) qui couvre les RGAs 1 à 8, le clone BAC 82O20 (37 192 pb) qui porte le RGA 9 et le clone BAC 97A19 (93 632 bp) qui porte le RGA 10. Cinq RGAs (RGA-1, RGA-2, RGA-4, RGA-8 et RGA-10) sont complets et 4 autres (RGA-3, RGA-5, RGA-6 et RGA-7) sont tronqués ou non-fonctionnels. Le RGA-11 est localisé à l'intérieur d'un "trou" dans le supercontig, encadré par les marqueurs 49MRP1.P2 et CB53.54. Les tentatives en vue de combler ce trou en utilisant une banque de cosmides n'ont réussi que partiellement. Cet échec suggère que cette région serait hautement répétée. Jusqu'à présent, aucune recombinaison entre *Run1* et *Rpv1* n'avait été mise en évidence au commencement de ma thèse, ce qui suggérait que la résistance au mildiou et à l'oïdium était, soit codée par un même gène de résistance, soit par différents membres d'un même *cluster* de gènes de résistance (Dry *et al.* 2009). Une forte dépression de recombinaison dans l'haplotype résistant avait effectivement été observée (Barker *et al.* 2005). Depuis, l'ensemble des gènes candidats disponibles dans la région ont été clonés et transformés dans des fonds génétiques sensibles et cela a permis de démontrer que ces deux résistances sont codées par des gènes différents (Feechan *et al.* 2013).



**Figure 22. Représentation schématique du contig de BAC et de la séquence correspondante dans la région des gènes *Run1* et *Rpv1***

### 1.5.1.2. Région homologue du locus de résistance *Run1/Rpv1* chez *V. vinifera*

La région chromosomique délimitée par les marqueurs microsatellites VMC4F3-1 et VMC8G9 correspond à **7 311 Mb** du chromosome 12 de *V. vinifera* cv PN40024 ; et un autre marqueur microsatellite, UDV-058, a été localisé entre les deux précédents (Cipriani *et al.* (2011)). Elle est constituée de **82 % (2 292 647 pb)** du *scaffold* 55, l'ensemble du *scaffold* 34 (**372 0581 pb**) et **36 % (1 296 772 pb)** du *scaffold* 35. Ces différents *scaffolds* sont reliés par 508 'N', ce qui correspond à 0.01 % de la région d'intérêt. Cette région chromosomique compte 390 gènes selon l'annotation automatique sur le *Genome Browser* de la vigne faite par les programmes GAZE et JIGSAW (V1). Il y aurait 25 gènes qui possèdent des domaines PFAM de gènes NBS : 6 gènes avec le domaine TIR (PF01582), 15 gènes avec le domaine NB-ARC (PF00931) et 4 gènes avec les domaines TIR et NB-ARC. La plupart de ces gènes sont au niveau du *scaffold* 34, de part et d'autre du marqueur UDV-058.

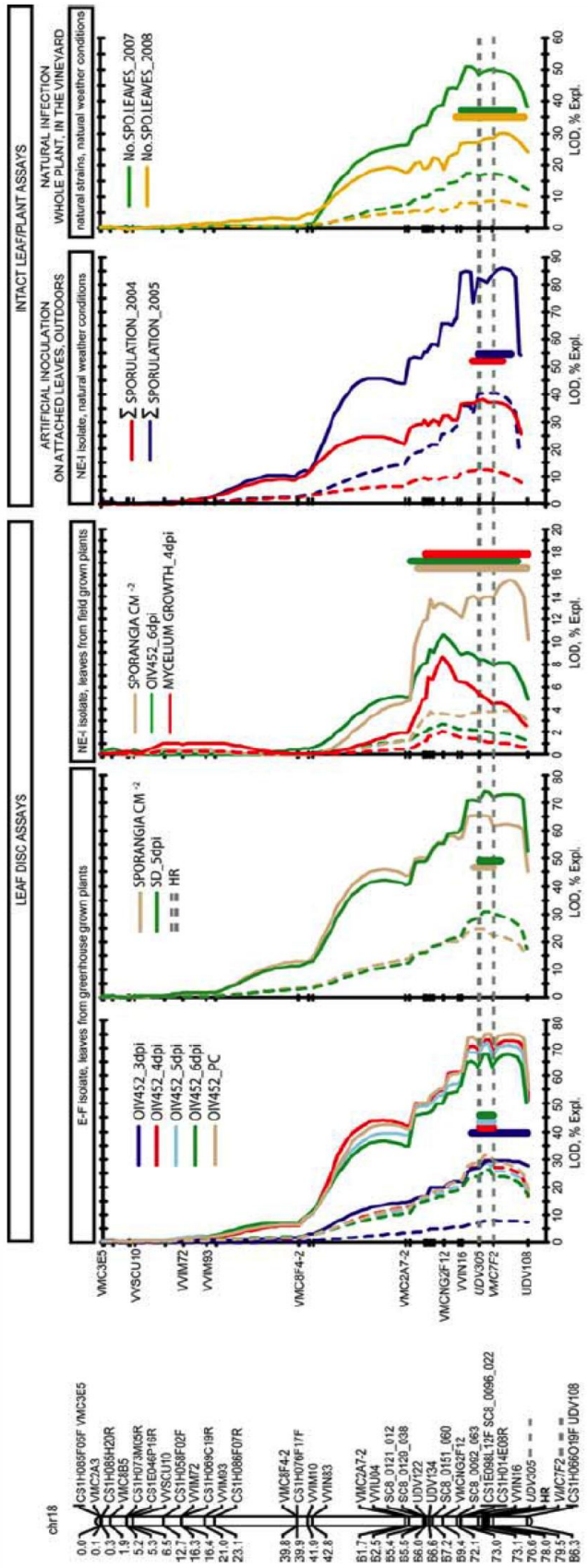


Figure 23. Colocalisation du déterminisme génétique de la réponse d'hypersensibilité entre les marqueurs UDV-305 et VMC7F2, et du locus de résistance au mildiou, *Rpv3*, sur le chromosome 18 de l'hybride Bianca. (Bellin *et al.* 2009)

### **1.5.1.3. Région homologue du *locus* de résistance *Rpv3* chez *V. vinifera***

Le *locus* de résistance à l'oïdium *Rpv3* est associé aux marqueurs microsatellites VMC7F2 et UDV-108 (Bellin *et al.* 2009 ; Riaz *et al.* 2011 ; Figure 23) et la séquence de l'haplotype résistant, provenant d'un hybride interspécifique complexe, n'est pas connue à ce jour. La région homologue chez *V. vinifera* cv PN40024 est de 2 227 Mb de la partie distale du chromosome 18 et compte 151 gènes selon la version actuelle de l'annotation automatique du génome de référence chez *V. vinifera*. Il y aurait 10 gènes qui possèdent des domaines PFAM de gènes NBS : 5 gènes avec le domaine TIR (PF01582), 1 gène avec le domaine NB-ARC (PF00931) et 4 gènes avec les domaines TIR et NB-ARC. Selon Bellin *et al.* (2009), le *locus* de résistance au mildiou *Rpv3* est confiné entre les marqueurs UDV-305 et VMC7F2. Nous avons donc choisi de considérer, comme intervalle de confiance, la région entre les marqueurs VMC2A7 et UDV-108. La région homologue chez *V. vinifera* cv PN40024 est de 15 726 Mb, pour un total de 742 gènes. Il y aurait 37 gènes qui possèdent des domaines PFAM de gènes NBS : 19 gènes avec le domaine TIR (PF01582), 9 gène avec le domaine NB-ARC (PF00931) et 9 gènes avec les domaines TIR et NB-ARC.

### **1.5.2. Les Objectifs du travail de recherche**

Au démarrage de ma thèse, une banque BAC de *Muscadinia rotundifolia* cv Regale venait d'être construite (Tableau 5) et leurs extrémités séquencées au Genoscope (ZahBi *et al.* en préparation). Les séquences des extrémités de ces BAC (BES) représentent un échantillon aléatoire de séquences du génome de *M. rotundifolia* qui peut être utilisé pour obtenir un premier aperçu de la composition de son génome ; ce qui a constitué (i) le premier objectif de ma thèse. J'ai utilisé cet échantillon de séquence pour regarder si des différences majeures étaient observées avec le génome de référence de *V. vinifera*. La majorité des gènes de résistance clonés et fonctionnels décrits dans le règne végétal étant des gènes de type NBS-LRR, je me suis également particulièrement focalisé sur cette famille de gènes dans les deux espèces. Ce travail fait l'objet de mon premier chapitre de résultats et alimente un article en préparation joint en annexe (ii) Un second objectif de cette thèse a été d'évaluer le degré de synténie entre *V. vinifera* et *M. rotundifolia* dans les deux régions présélectionnées en exploitant les ressources génomiques



disponibles, utilisables pour l'intégration des données de cartes physique et génétique entre les sous-genres *Euvitis* et *Muscadinia*. Cela fait l'objet de mon deuxième chapitre de résultats et une partie d'entre eux ont également été inclus dans l'article en préparation. Enfin, (iii) le troisième objectif, présenté dans le troisième chapitre de résultats, consistait à évaluer plus finement la synténie entre les deux genres à travers le séquençage de clones BAC de *M. rotundifolia* cv Regale, sélectionnés dans ces deux régions.

Tableau 5. Caractéristiques de la banque BAC de *Muscadinia rotundifolia* cv Regale

Enzyme	Clones (N°)	Taille moyenne des inserts	Clones vides (%)	Contaminations (%)	Couverture du génome	BES Forward (N°)	BES Reverse (N°)	Total de BES (N°)
<i>Hind</i> III	15 774	75 kb	6,27	3,8	2,1	12 800	12 677	25 477
<i>Hind</i> III	10 368	82 kb	6,81	3,9	1,5	8 683	8 600	17 283
<i>Bam</i> HI	13 440	59 kb	13,97	5,2	1,3	10 907	10 802	21 709
<i>Hind</i> III	14 592	73 kb	9	3,5	1,9	11 064	11 007	22 071
	<b>54 174</b>	<b>71,8 kb</b>	<b>9</b>	<b>4,1</b>	<b>6,8</b>	<b>43 454</b>	<b>43 086</b>	<b>86 540</b>

## ***2. Résultats***

**CHAPITRE I :**  
**CARACTERISTIQUES DU GENOME DE *Muscadinia***  
***rotundifolia* CV REGALE A TRAVERS L'ANALYSE DES BES**

## 2.1. Introduction

Pour mieux comprendre la structure et l'évolution du génome de *M. rotundifolia*, une banque de **54 174** grands fragments d'ADN du cultivar (cv) Regale a été réalisée. La banque BAC (chromosomes bactériens artificiels) était disponible au début de ma thèse. Le clonage avait été fait dans les sites de restriction *HindIII* (**40 734** clones BAC) et *BamHI* (**13 440** clones BAC). Au total, **86 810** extrémités de séquences BAC (BES) ont été obtenues, soit **43 724** séquences “*forward*” (**50,3 %**) et **43 086** séquences “*reverse*” (**49,6 %**), avec une longueur moyenne de **682** pb. Après élimination de la redondance (**270** doublons) des séquences dans les extrémités, il restait **86 540** BES uniques, représentant **59** Mb du génome de *M. rotundifolia* cv Regale. Au cours de ma thèse, je me suis attaché à exploiter cette ressource pour en tirer des enseignements sur les similarités et les différences entre les génomes de *V. vinifera* et *M. rotundifolia*. Pour cela, j'ai utilisé comme point de comparaison une collection de **77 237** BES de *V. vinifera* cv Cabernet Sauvignon (Adam *et al.* 2005 ; Lamoureux *et al.* 2006) issus de clonage fait dans des sites de restriction *EcoRI* et *HindIII*, et le génome de référence de la vigne (*V. vinifera* cv PN40024). L'analyse des différences génomiques entre *Vitis vinifera* et *Muscadinia rotundifolia* se fera sur le plan structurel par le biais des teneurs en GC, des éléments répétés, des teneurs en séquences microsatellites et du contenu en gènes des BES de *Vitis vinifera* cv Cabernet Sauvignon et *Muscadinia rotundifolia* cv Regale.

## 2.2. Etude des paramètres généraux du génome de *M. rotundifolia* cv Regale

Aucune estimation du contenu en ADN nucléaire n'est disponible pour les vignes du genre *Muscadinia*. A ce jour, il existe une seule étude qui compare la teneur en ADN nucléaire de différentes espèces de *Vitis* (Lodhi et Reisch, 1995). Cette étude comprend 21 espèces du genre *Vitis*, 21 variétés de *V. vinifera*, et une espèce représentant *Ampelopsis* et *Parthenocissus* (deux autres genres de la famille des *Vitaceae*). Ces dernières possèdent 20 chromosomes ( $2n = 40$ ) comme *Muscadinia rotundifolia* (Karkamkar *et al.* 2010). Le genre *Muscadinia* n'était pas inclus dans cette étude. Les auteurs n'ont pas observé de différence de contenu en ADN entre les genres et entre les espèces. Il n'y a donc pas de différences entre les genres à 38 chromosomes (*Vitis*) et ceux à 40 chromosomes. Par conséquent, dans la suite de ce travail, j'ai supposé que la taille du génome de *Muscadinia rotundifolia* est comparable à l'ensemble des espèces de la famille des *Vitaceae*, soit  $1C = 475$  Mb. Elle est donc égale entre *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon.

### 2.2.1. Taux de Guanine-Cytosine (GC) des BES de *M. rotundifolia* cv Regale

L'analyse des **86 540** BES de *M. rotundifolia* cv Regale, représentant **59 Mb** de son génome (12,4 %), et des **77 737** BES de *V. vinifera* cv Cabernet Sauvignon, représentant **51,8 Mb** de son génome (10,9 %), ont permis d'estimer les teneurs en GC en utilisant REPEATMASKER. La teneur moyenne en GC des BES de *M. rotundifolia* cv Regale est de **35,84 %**, ce qui est légèrement supérieur, mais comparable, à celle des BES de *V. vinifera* cv Cabernet Sauvignon (Tableau 6). Par contre, ces estimations sont légèrement supérieures à la teneur en GC du génome entier de *V. vinifera* cv PN40024 (**34,55 %** ; Giannuzzi *et al.* 2011).

**Tableau 6. Caractéristiques des génomes de diverses plantes sur la base des séquences d'extrémités de BAC et du génome de *V. vinifera* cv PN40024**

	Génome (Mb)	Ploïdie	% du génome	% de GC	Eléments répétés (%)	Eléments majoritaires	Références
<i>M. rotundifolia</i> cv Regale	475	Diploïde	11,9	35,84	27,26	Gypsy	Notre étude
<i>V. vinifera</i> cv Cab. Sauv.	475	Diploïde	10,4	35,5	32,47	Gypsy	Notre étude
<i>V. vinifera</i> cv PN40024 *	475	Diploïde	100	34,55	29,26	Gypsy	Giannuzzi <i>et al.</i> 2011
<i>Quercus robur</i>	740	Diploïde	~2	35,3	5,90	Copia	Rampant <i>et al.</i> 2011
<i>Solanum lycopersicon</i>	950	Diploïde	19	36	17,86	Gypsy	Datema <i>et al.</i> 2008
<i>Solanum tuberosum</i>	840	Diploïde	10	35,6	11,28	Gypsy	Datema <i>et al.</i> 2008
<i>Malus x domestica</i> Borkh	740	Diploïde	0,3	39	20,9	Copia	Han <i>et al.</i> 2008
<i>Cucumis melo</i>	454	Diploïde	5,7	35,2	6	Copia	Gonzalez <i>et al.</i> 2010
<i>Citrus clementina</i>	367	Diploïde	8	39	12,6	Copia	Terol <i>et al.</i> 2008
<i>Panicum virgatum</i>	1600	Tétraploïde	5,5	45,5	33,3	Gypsy	Sharma <i>et al.</i> 2012
<i>Brachypodium</i>	350	Diploïde	10,9	45,9	10,99	Gypsy	Huo <i>et al.</i> 2008
<i>Triticum aestivum</i> (Chr 3B)	995	Hexaploïde	1,1	44,5	76	Gypsy	Choulet <i>et al.</i> 2010

\* Génome entier de *V. vinifera* cv PN40024

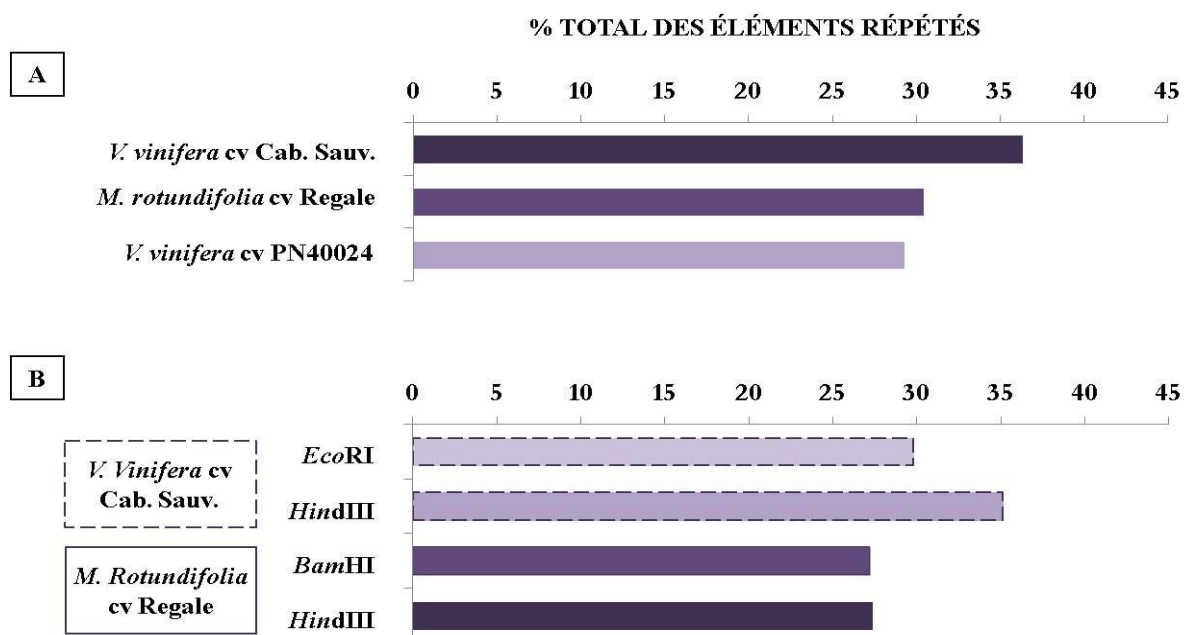
Ces résultats suggèrent donc un biais en termes d'abondance en GC dans les collections de BES par rapport aux séquences de génomes entiers qui suggère une surestimation du taux de GC dans les BES par rapport au génome de *V. vinifera* cv PN40024. Par conséquent, les taux de GC seraient similaires dans les génomes de *M. rotundifolia* cv Regale et de *V. vinifera*.

### 2.2.2. Contenu en éléments répétés des BES de *M. rotundifolia* cv Regale

Les éléments répétés ont été recherchés en parallèle dans les deux collections de BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon, avec le programme REPEATMASKER, en utilisant la section *Viridiplantae* de la banque d'éléments nucléiques répétés REPBASE. Le pourcentage d'éléments

répétés détectés dans les BES de *V. vinifera* cv Cabernet Sauvignon est estimé à **36,32 %** alors que celui dans les BES de *M. rotundifolia* cv Regale est estimé à **30,37 %**. La teneur des éléments répétés dans les BES de *V. vinifera* cv Cabernet Sauvignon est **1,2** fois supérieure à celle des BES de *M. rotundifolia* cv Regale (Figure 24A). Ces séquences répétées sont comparées à l'ensemble des répétitions du génome entier de *V. vinifera* cv PN40024 (**29,26 %** ; Giannuzzi *et al.* 2011). Un écart de **1 - 7 %** est observé entre les BES et le génome entier de *V. vinifera* cv PN40024 (Figure 24A).

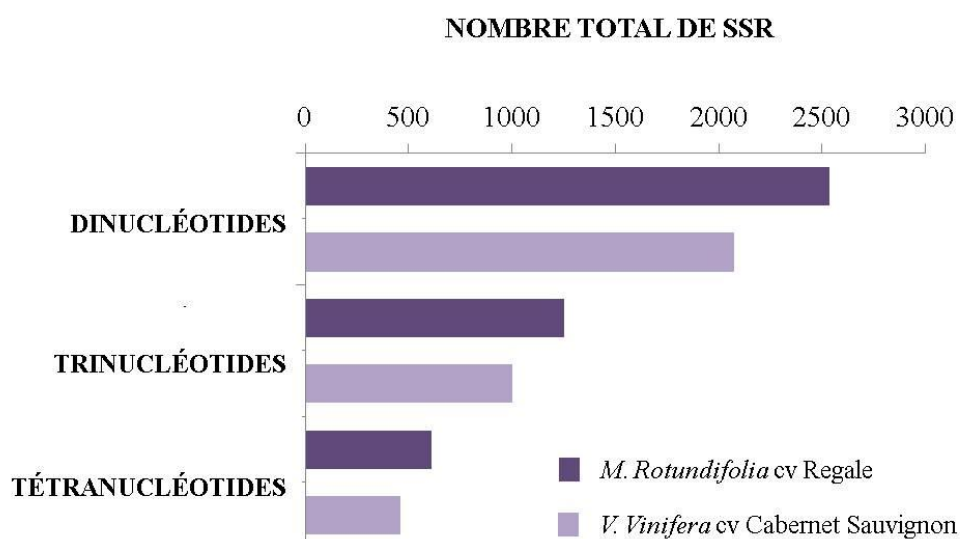
L'utilisation d'enzymes de restriction (*HindIII*, *EcoRI* et *BamHI*) lors de la construction des banques BACs pourrait favoriser des extrémités de clones dans des régions légèrement enrichies en éléments répétés. En effet, si le pourcentage de séquences répétées détectées dans les BES des sous-banques *HindIII* et *BamHI* de *M. rotundifolia* cv Regale sont équivalents (**27,17 %** et **27,34 %** respectivement; Figure 24B ; Annexe 1) une différence est observée entre les sous-banques *HindIII* et *EcoRI* de *V. vinifera* cv Cabernet Sauvignon (**30,94 %** et **24,77 %** respectivement ; Figure 24B ; Annexe 1). A ce stade, le contenu (quantitatif) en séquences répétées semble donc similaire chez *M. rotundifolia* et *V. vinifera*.



**Figure 24. Analyse comparative des éléments répétés dans les BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon, et dans le génome de *V. vinifera* cv PN40024.** Ces éléments répétés sont représentés en pourcentage de séquences occupées grâce aux estimations faites par REPEATMASKER : (A) pour la totalité des BES de *V. vinifera* cv Cabernet Sauvignon et *M. rotundifolia* cv Regale, et pour l'ensemble du génome de *V. vinifera* cv PN40024 ; (B) pour les BES issus des banques BAC *EcoRI*, *HindIII* et *BamHI*.

### 2.2.2.1. Les séquences microsatellites

Un total de **4 383** SSR, avec une longueur de motif entre **2** et **4** nucléotides, a été détectés dans les BES de *M. rotundifolia* cv Regale ; avec une taille de SSR comprise entre **18** et **196** nucléotides. Le terme “longueur de motif” est utilisé pour décrire la longueur du motif qui se répète dans le SSR. Par exemple, une répétition ATATAT a une longueur de deux motifs (avec AT étant le motif). Les SSR les plus abondants sont les dinucléotides (**2 532** SSR ; **38 %**), suivis par les trinucleotides (**1 246** SSR ; **28 %**) et enfin les tétranucléotides (**605** SSR ; **14 %** ; Figure 25). Chez *V. vinifera* cv Cabernet Sauvignon, **3 522** SSR avec une longueur de motif entre **2** et **4** nucléotides ont été trouvés, avec une taille de SSR comprise entre **18** et **648** nucléotides. Comme chez *M. rotundifolia* cv Regale, les SSR dinucléotides sont les plus abondants (**2 070** SSR ; **59 %**), suivis par les SSR trinucleotides (**996** SSR ; **28 %**) et les SSR tétranucléotides (**456** SSR ; **13 %** ; Figure 25).



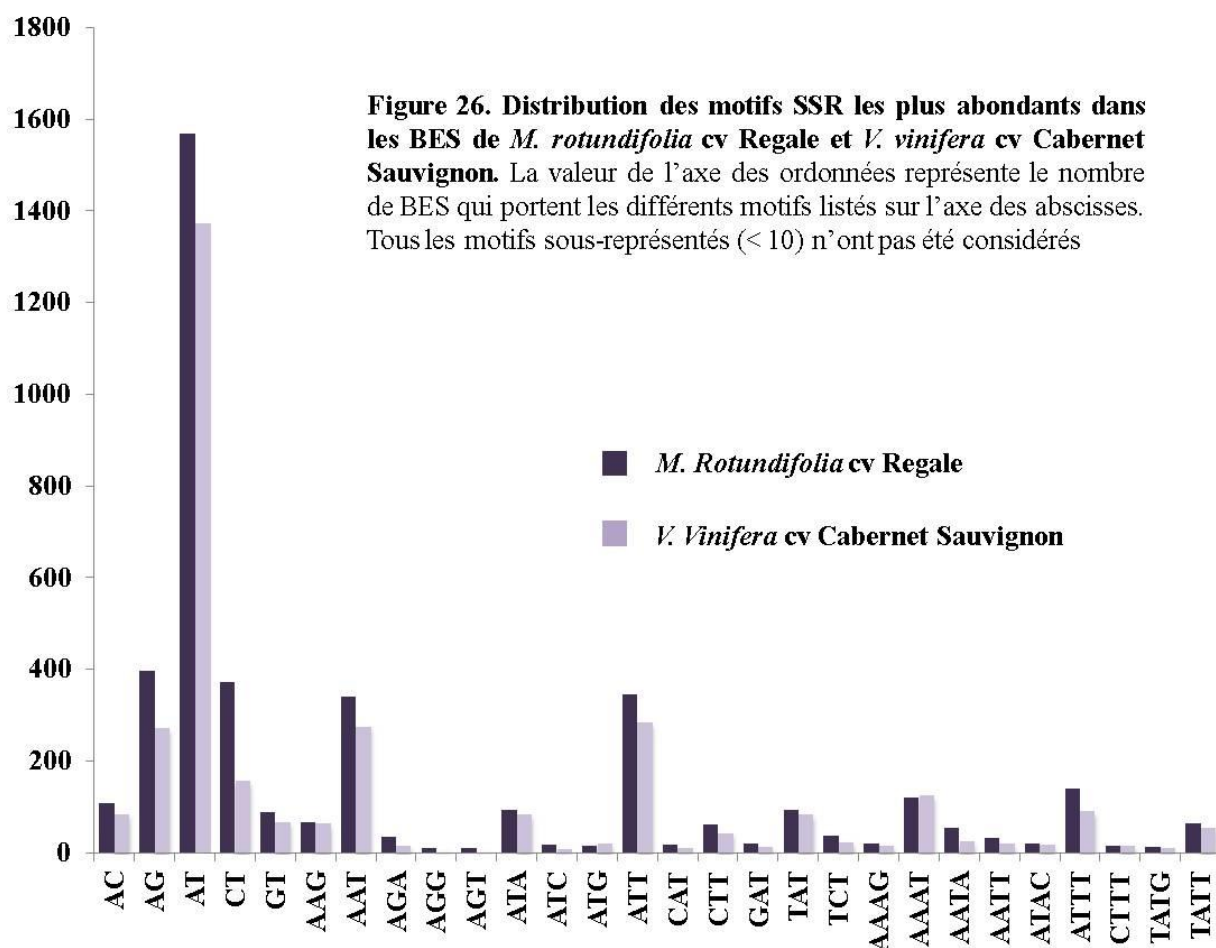
**Figure 25. Les SSR les plus abondants dans les BES de *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon.**

Parmi les dinucléotides de *M. rotundifolia* cv Regale, les motifs les plus communs étaient AT/TA (**62 %**) et AG/GA/TC/CT (**15 %**) alors que pour les motifs trinucleotides, les motifs AAT/TAA/ATT/TTA étaient les plus abondants (**28 %**). Pareillement, les motifs AT/TA (**66 %**) et AG (**13 %**) étaient les plus communs parmi les dinucléotides de *V. vinifera* cv Cabernet Sauvignon alors que pour les motifs trinucleotides, les



motifs AAT/TAA/ATT/TTA étaient les plus abondants (28 %). Aucun motif GC/CG n'a été trouvé (Figure 26).

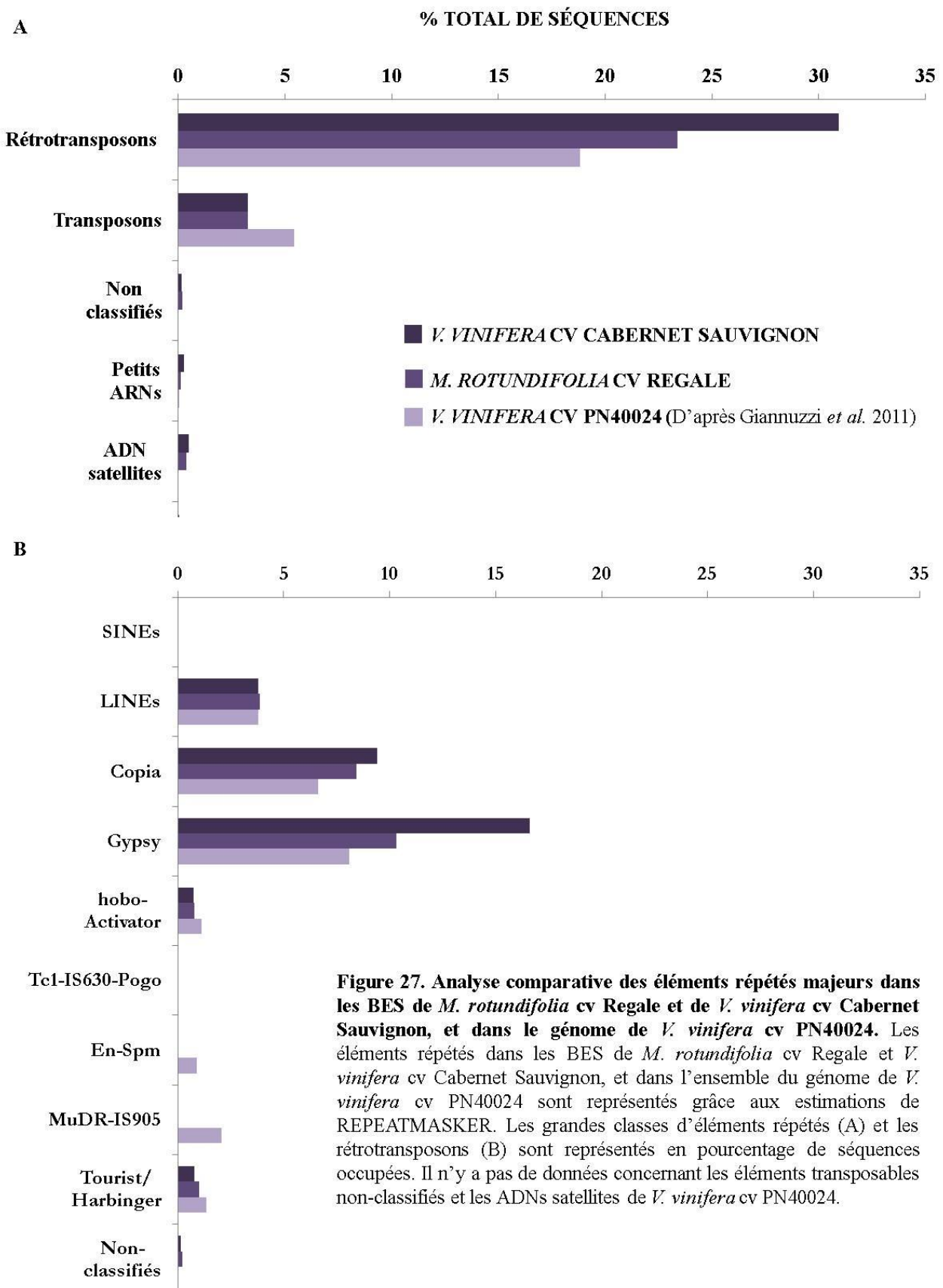
Nos résultats indiquent que les BES de *M. rotundifolia* cv Regale ont une proportion plus élevée de di-, tri- et tétranucléotides par rapport aux BES de *V. vinifera* cv Cabernet Sauvignon (Figure 26). C'est peut-être parce que certains BES de *M. rotundifolia* cv Regale sont issus de clones BAC provenant de régions centromériques du génome de *M. rotundifolia* cv Regale, car il a précédemment été montré, chez la tomate, que ces régions sont riches en séquences microsatellites (Areshchenkova *et al.* 1999). Toutefois, cette hypothèse pourrait être vérifiée si (i) les régions centromériques contenaient plusieurs sites de restriction *Hind*III et *Bam*HI que la moyenne pour le génome de *M. rotundifolia* cv Regale ; et si le contenu en régions codantes étaient moins important dans les BES de *M. rotundifolia* cv Regale.



#### 2.2.2.2. Les éléments transposables

Les familles d'éléments répétés les plus abondantes dans les BES de *M. rotundifolia* cv Regale sont les classes de rétrotransposons à LTR Gypsy et Copia (Figure 27B ; Annexe 1). Comme dans les BES de *M.*

*rotundifolia* cv Regale, la majorité des éléments répétés, dans les BES de *V. vinifera* cv Cabernet Sauvignon, sont des rétrotransposons à LTR Gypsy et Copia (Figure 27B ; Annexe 1). On note de légères différences de répartition des classes de rétrotransposons en fonction des banques *Bam*HI et *Hind*III de *M. rotundifolia* cv Regale. Puisque la détection des éléments répétés est basée sur une recherche de similarité, les différents segments d'un BES peuvent être attribués à plus d'une famille d'éléments répétés. Par conséquent, la somme des classes d'éléments répétés peut être légèrement supérieure aux teneurs totales données par REPEATMASKER. Pour éviter de comparer des ensembles de séquences avec des biais différents, les collections de BES *Hind*III de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon ont été utilisées pour quantifier les différences entre ces deux espèces. Ainsi, la teneur en éléments répétés des BES *Hind*III de *M. rotundifolia* cv Regale (**27,34** %) est inférieure à celle des BES *Hind*III de *V. vinifera* cv Cabernet Sauvignon (**35,13** %) (Figure 24B). La teneur des éléments répétés et notamment des rétrotransposons est plus importante dans les BES de *V. vinifera* cv Cabernet Sauvignon que dans le génome entier de *V. vinifera* cv PN40024, elle-même supérieure à celle des BES de *M. rotundifolia* cv Regale (Figure 27A). En revanche, le contenu en transposon et en simples répétitions du génome de référence est **2** fois et **36** fois supérieur à celui observé dans les BES, respectivement (Figure 27A). Ces observations suggèrent un enrichissement en rétrotransposons et une faible représentation des transposons dans les séquences d'extrémités de BAC. La sous-représentation des transposons pourrait correspondre à la difficulté à identifier leurs structures complètes par alignement de séquences sur des séquences courtes comme les BES. La surreprésentation des rétrotransposons de classe I dans les BES correspond à une surreprésentation des Gypsy très marquée chez *V. vinifera* cv Cabernet Sauvignon et moindre en Copia.



Les résultats obtenus dans ce travail et leur comparaison avec des travaux similaires de la littérature suggèrent donc (i) une tendance à surestimer la teneur en séquence répétées d'un génome dans les collections de BES, (ii) à y surestimer la proportion d'éléments de type Gypsy et à sous-estimer la proportion de transposon. Le taux de séquences répétées de la collection de BES de *M. rotundifolia* cv Regale étant équivalent au taux estimé dans le génome complet de *V. vinifera* cv PN40024. Il est donc possible qu'en réalité le taux de séquences répétées du génome *M. rotundifolia* soit plus faible que celui de *V. vinifera*. De même, la proportion de Gypsy, majoritaire chez *M. rotundifolia* cv Regale comme chez *V. vinifera* cv Cabernet Sauvignon, est probablement un peu surestimée. Il faut cependant noter que la méthode d'identification des éléments transposables, par homologie de séquence, a le désavantage d'être très conservative en ne recherchant que des éléments similaires à des éléments déjà connus et qu'il est possible que des éléments spécifiques de *M. rotundifolia* aient été mal détectés.

### **2.2.3. Contenu en régions codantes des BES de *M. rotundifolia* cv Regale**

Les séquences répétées identifiées dans les BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon ont été masquées grâce à REPEATMASKER. Les BES ont été alignées contre le protéome de *V. vinifera* cv PN40024 (E-VALUE =  $10^{-4}$ ) par BLASTX afin d'identifier les séquences codantes des protéines. L'analyse indique qu'on trouve des similarités entre **42 698** BES de *M. rotundifolia* cv Regale (**49 %** des BES) et **22 403** protéines, et entre **38 541** BES de *V. vinifera* cv Cabernet Sauvignon (**50 %** des BES) et **21 571** protéines. Ce qui représente **74 %** et **71 %** du protéome de *V. vinifera*. En ne considérant que les alignements avec un pourcentage d'identité compris entre **90 - 100 %**, il apparaît que le contenu en gènes est aussi élevé voire légèrement plus élevé dans les BES de *M. rotundifolia* cv Regale que dans ceux de *V. vinifera* cv Cabernet Sauvignon (Tableau 7) et que ces paramètres très stringents permettent tout aussi bien de repérer des séquences codantes chez *Vitis vinifera* que chez *M. rotundifolia*.

**Tableau 7. Alignement des collections de BES contre le protéome de *V. vinifera* cv PN40024 par BLASTX.** Uniquement les alignements avec 90 - 100 % d'identité sont considérés.

	Nombre de BES	Nombre de protéines	E-VALUE	Score
<i>M. rotundifolia</i> cv Regale Banque <i>Bam</i> HI	4 319	4 420	$1.10^{-3}$ - $4.10^{-163}$	20,4 – 570
<i>M. rotundifolia</i> cv Regale Banque <i>Hind</i> III	12 992	9 392	$1.10^{-3}$ - $4.10^{-155}$	20,4 – 597
<i>V. vinifera</i> cv Cab. Sauv Banque <i>Eco</i> RI	4 451	4 250	0 - $3.10^{-177}$	20,4 – 648
<i>V. vinifera</i> cv Cab. Sauv Banque <i>Hind</i> III	11 431	8 309	0 - $9.10^{-180}$	20,4 – 648

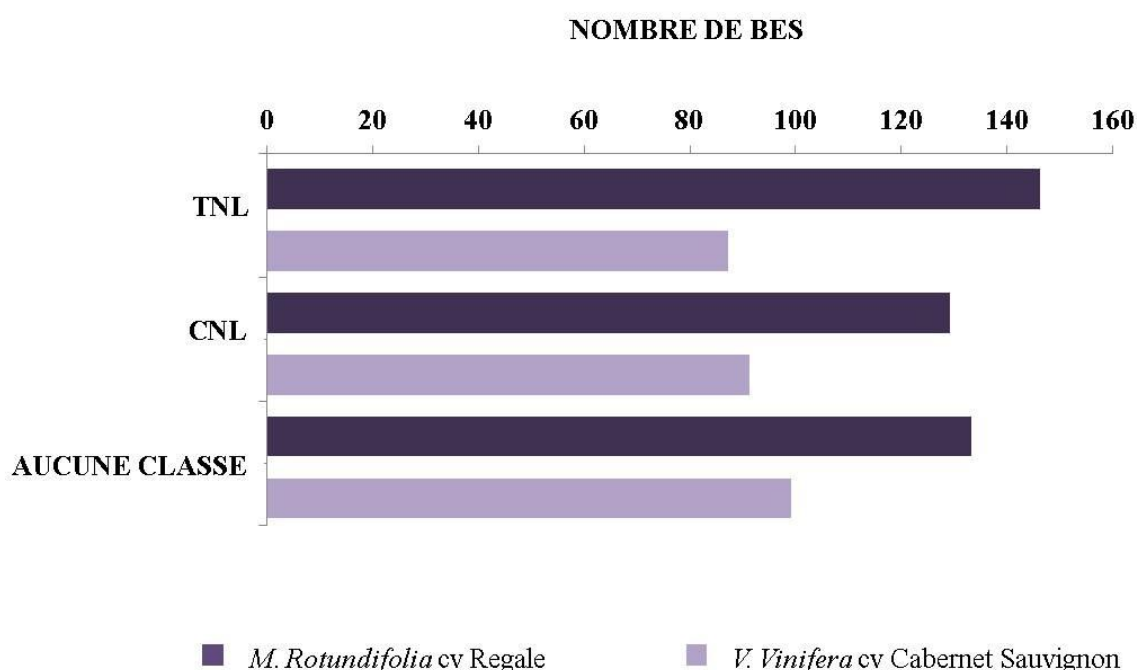
Par la suite, nous nous sommes plus particulièrement intéressés à la famille des gènes codant des protéines à motifs NBS-LRR qui constituent l'une des plus grandes familles de gènes chez les plantes et sont impliquées dans la résistance aux maladies chez les plantes. Leur identification dans les BES de *M. rotundifolia* cv Regale permettra d'avoir un aperçu significatif de l'évolution de ces gènes dans la famille des *Vitaceae* et, surtout, un modèle pour les reconnaître et les cloner plus rapidement chez les *Vitaceae*.

### 2.3. Analyse du contenu en régions codant pour des protéines NBS-LRR dans les BES

#### 2.3.1. Identification à l'aide des motifs MEME spécifiques du domaine NBS

Quarante-huit motifs (48) MEME spécifiques du domaine conservé NBS d'*A. thaliana* (Meyers *et al.* 2003) et de *P. trichocarpa* (Kölher *et al.* 2008) ont été utilisés pour caractériser les gènes NBS-LRR. Certains motifs sont associés à des domaines connus des superfamilles CNL et TNL tels que les domaines TIR, P-loop, RNBS-A, Kinase-2, RNBS-B, RNBS-C, GLPL, RNBS-D et MHD (Annexe 2). L'ensemble de ces motifs MEME a été recherché par BLASTX (E-VALUE =  $10^{-4}$ ) contre les BES de *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon. Au total, 408 BES de *M. rotundifolia* cv Regale contiennent les motifs MEME recherchés contre 277 BES de *V. vinifera* cv Cabernet Sauvignon. Les protéines de résistance de type CNL se caractérisent par la présence de motifs RNBS-A non-TIR (consensus : FDLxAWCVSxQxF), d'un résidu tryptophane (W) à la fin du motif kinase 2 et d'un motif RNBS-D non-TIR (consensus : CFLYCALFPED); pareillement, les protéines de résistance de type TNL se caractérisent par la présence de motifs RNBS-A TIR (consensus : FLENIRExSKKHGLEHLQKKLLSKLL) et d'un motif RNBS-D TIR (consensus : FLHIACFF) (Lukasik

*et al.* 2009). L'identification de ces motifs spécifiques, dans les BES identifiées avec les motifs MEME, a permis de préciser la part des deux grandes classes TNL et CNL (Figure 28). Cent vingt-neuf (**129**) BES de *M. rotundifolia* cv Regale ( $5.10^{-13} \leq E\text{-VALUE} \leq 9.10^{-4}$ ) et **91** BES de *V. vinifera* cv Sauvignon ( $7.10^{-13} \leq E\text{-VALUE} \leq 9.10^{-4}$ ) contiennent les motifs RNBS-A, RNBS-D spécifiques des CNL ; **146** BES de *M. rotundifolia* cv Regale ( $5.10^{-16} \leq E\text{-VALUE} \leq 9.10^{-4}$ ) et **87** BES de *V. vinifera* cv Sauvignon ( $5.10^{-16} \leq E\text{-VALUE} \leq 9.10^{-4}$ ) contiennent les motifs TIR et RNBS-D spécifiques des TNL.



**Figure 28. Nombre de BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon contenant les motifs MEME spécifiques des TNL et des CNL.** Les BES ont été identifiés par homologie de séquence avec les motifs MEME grâce au programme BLASTX. Les CNL possèdent les motifs RNBS-A et/ou RNBS-D non-TIR ; les TNL possèdent le motif RNBS-D TIR et/ou les motifs spécifiques du domaine TIR (Annexe h2 ; Lukasik *et al.* 2009)

### 2.3.2. Identification de BES portant des séquences NBS-LRR à l'aide de domaine NB-ARC des protéines fonctionnelles

#### 2.3.2.1. Choix d'un set de gènes de résistances fonctionnels de la famille des NBS-LRR

Une collection de gènes de résistance aux maladies des plantes est disponible en libre accès pour la communauté scientifique (Sanseverino *et al.* 2009). A ce jour, elle compte plus de **16 000** gènes de

résistance parmi lesquels **112** sont des gènes fonctionnels caractérisés pouvant permettre de prédire de façon précise de nouveaux gènes de résistance putatifs. Parmi ces **112** gènes, **70** codent des gènes de résistance à domaine NBS-LRR (**50** CNL, **15** TNL, **4** NL et **1** TN ; Annexe 3). Ils ont été choisis comme gènes de référence dans la suite du travail. Les séquences protéiques de ces gènes ont été comparées à la collection complète des domaines protéiques PFAM (Finn *et al.* 2010). Les séquences d'acides aminés correspondant au domaine NB-ARC dans ces protéines de référence ont été récupérées pour chacune des **70** protéines de résistance et la taille moyenne de ce domaine NBS est estimée à **277** acides aminés. D'une part, ces séquences vont servir pour le criblage *in silico* des BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon ; et d'autre part, pour faire l'analyse phylogénétique uniquement des BES de *M. rotundifolia* cv Regale.

### **2.3.2.2. Identification des BES portant les domaines NB-ARC des protéines fonctionnelles**

L'ensemble des BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon a été comparé aux domaines NB-ARC des protéines fonctionnelles grâce au programme BLASTX avec une E-VALUE =  $10^{-4}$ . Trois cent quinze (**315**) BES de *M. rotundifolia* cv Regale ont été identifiées contre **255** BES de *V. vinifera* cv Cabernet Sauvignon. Ces BES ont été retenues quelque soit la longueur de l'alignement et comparées au *set* de BES identifiées à l'aide des motifs MEME. Ainsi, **129** BES et **200** BES ont été identifiées à la fois par les motifs MEME et par alignement sur les domaines NB-ARC des gènes fonctionnels, chez *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon, respectivement (Figure 29). Par ailleurs, la plupart des BES identifiées appartenant à la fois aux CNL ou aux TNL, ce qui ne permet pas de les différencier.

Un nombre similaire de BES contenant des séquences codant des gènes NBS-LRR a été identifié, avec **451** (**0,5** % de la collection) pour *M. rotundifolia* et **332** pour *V. vinifera* (**0,4** % de la collection). Pareillement, un nombre similaire de CNL ou TNL a été identifié dans les deux échantillons : environ un tiers des séquences avait un domaine TIR et/ou une séquence spécifique des TNL dans le domaine NB-ARC, un tiers avait des séquences spécifiques des CNL et aucune catégorie n'a pu être proposée pour le dernier tiers.

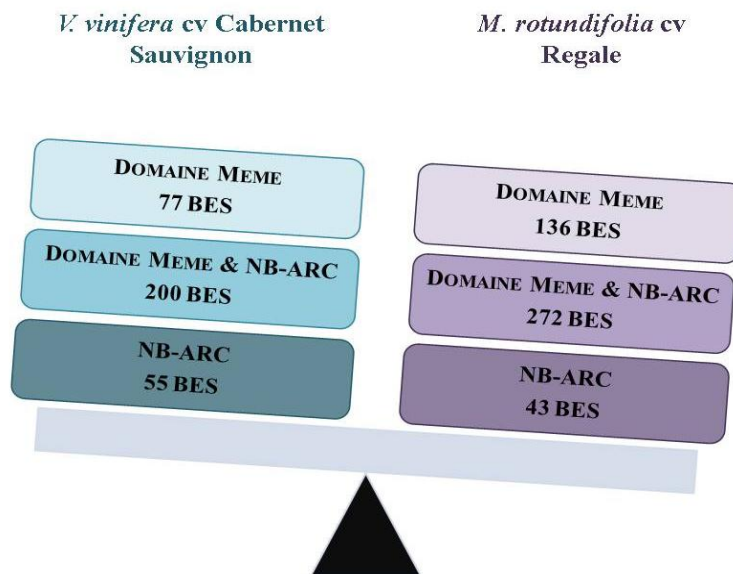


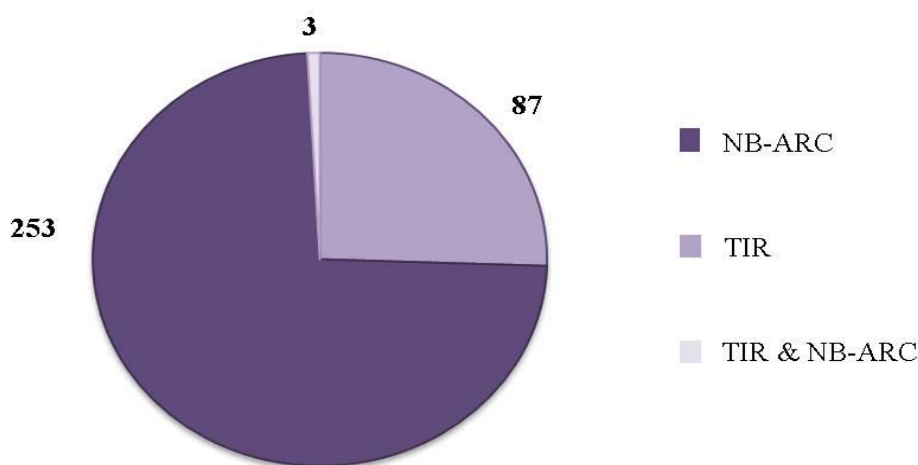
Figure 29. Nombre de BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon identifiées par comparaison de séquences avec des domaines NB-ARC de protéines de résistance fonctionnelles (NB-ARC) et des motifs MEME spécifiques du domaine conservé NB-ARC (DOMAINE MEME). Le programme BLASTX a permis d'identifier les différents BES avec une E-value =  $10^{-4}$ .

### 2.3.3. Analyse phylogénétique des gènes à domaine NB-ARC de *M. rotundifolia* cv Regale

#### 2.3.3.1. Sélection des BES de *M. rotundifolia* cv Regale pour l'analyse phylogénétique

Pour convertir les séquences nucléotidiques en séquences d'acides aminés, et en tenant compte des éventuels décalages de phase de lecture lors de la traduction, chacune des **451** BES de *M. rotundifolia* cv Regale a été comparée aux protéines expertisées SWISSPROT grâce au programme BLASTX. Seuls **343** BES (soit 75,6 %) ont une similitude avec des protéines de résistance qui permettent d'extraire une séquence protéique fiable. Les séquences d'acides aminés ont été récupérées pour ces BES de *M. rotundifolia* cv Regale puis elles ont été comparées à la banque de domaines conservés disponible au NCBI (Marchler-Bauer *et al.* 2011). Quarante-sept (**87**) BES possèdent le domaine PFAM 1582 (spécifique du domaine TIR), **251** possèdent le domaine PFAM 931 (spécifique du domaine NB-ARC) et **3** BES possèdent une partie des deux domaines (Figure 30).





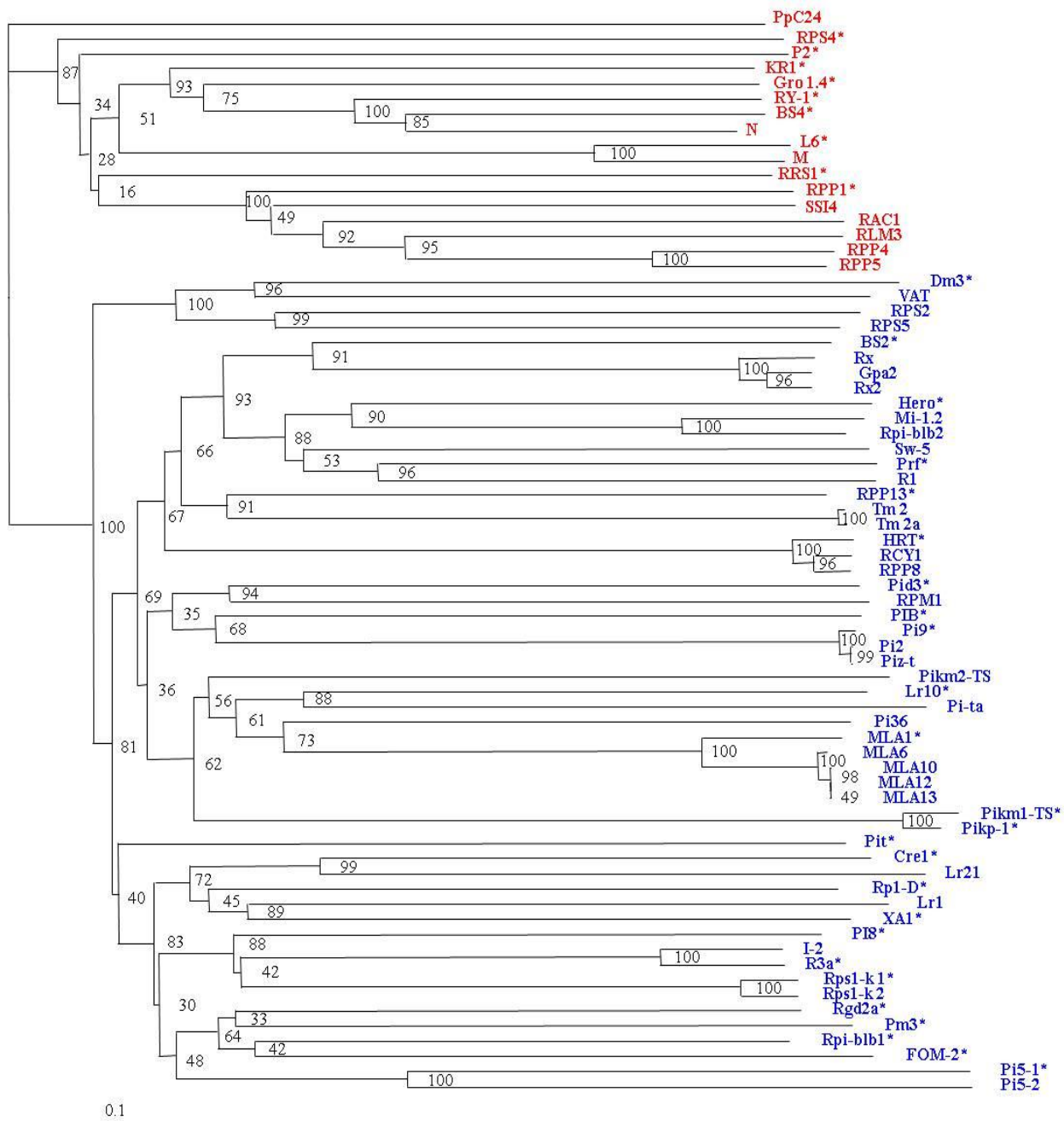
**Figure 30. Domaines PFAM détectés dans la séquence protéique extraite des BES par alignement avec des protéines connues.** Les chiffres représentent le nombre de BES de *M. rotundifolia* cv Regale

Les séquences déduites des BES de *M. rotundifolia* cv Regale ont une longueur qui varie entre **24** et **255** acides aminés. Puisque la longueur moyenne du domaine NB-ARC des protéines de résistance fonctionnelles est de **277** acides aminés, les **78** BES de *M. rotundifolia* cv Regale qui présentaient au moins **150** acides aminés similaires au domaine NB-ARC ont été sélectionnées pour l'analyse phylogénétique (Annexe 4). Ces 78 séquences contiennent donc **55** à **91** % du domaine NB-ARC des protéines fonctionnelles.

#### 2.3.3.1.1. Sélection des domaines NB-ARC des protéines fonctionnelles

Une première analyse phylogénétique a été réalisée à partir des **70** domaines NB-ARC de l'ensemble des protéines de résistance fonctionnelles (Annexe 3). Le domaine NB-ARC de la protéine de résistance PpC24 de *Physcomitrella patens* est inclus comme groupe externe. Comme attendu, l'analyse phylogénétique sépare bien les domaines NB-ARC des TNL et CNL en deux clades distincts (Figure 31) avec des valeurs de *bootstrap* de **87** et **100**, respectivement. Le clade des TNL contient **16** protéines NBS-LRR qui peuvent être divisées en **8** sous-groupes en fonction de la valeur des *bootstrap*. De même, le clade des CNL contient **54** protéines NBS-LRR qui peuvent être divisées en **25** sous-groupes en fonction de la valeur des *bootstrap*. Pour éviter une redondance d'informations et les groupes de protéines dont le

positionnement a des valeurs de *bootstrap* très basses, trente-quatre (**34**) séquences d'acides aminés ont été choisies aux extrémités des branches de l'arbre phylogénétique (Figure 31). Pour les protéines avec des séquences similaires, un membre a été choisi pour l'étude des relations phylogénétiques des domaines NB-ARC identifiés dans les BES de *M. rotundifolia* cv Regale.

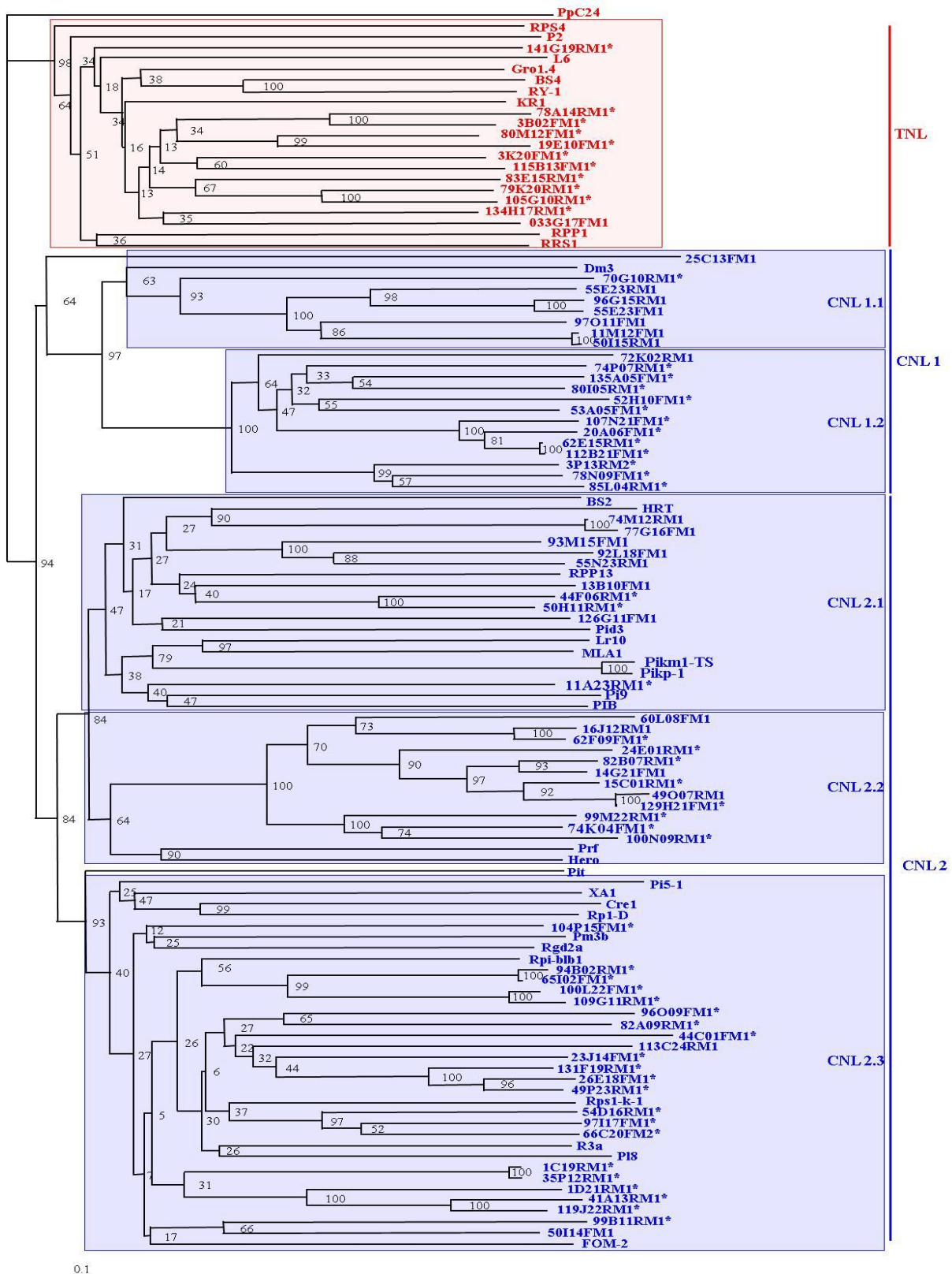


**Figure 31. Analyse des domaines NBS de protéines de résistance fonctionnelles avec le programme CLUSTALX.** Un domaine NBS de *Physcomitrella patens* (PpC24) est utilisé comme groupe externe. Les domaines NB-ARC des protéines identifiés avec un astérisque (\*) sont utilisés pour l'étude des relations phylogénétiques des domaines NBS des BES de *M. rotundifolia* cv Regale . Les chiffres indiquent la valeur des *bootstrap* de 100 répliquions.

### 2.3.3.1.2. Classification phylogénétique des domaines NB-ARC de gènes de résistance

Pour étudier les relations évolutives des gènes NBS-LRR identifiés dans les BES de *M. rotundifolia* cv Regale, un arbre phylogénétique a été construit à partir de l'alignement des domaines conservés NB-ARC grâce au programme CLUSTALX. Cette analyse a été faite sur la base d'études antérieures qui ont montré que les phylogénies construites à partir du domaine NBS étaient robustes pour distinguer les familles de TNL et non-TNL (Meyers *et al.*, 1999 ; Meyers *et al.*, 2002.). Le domaine NB-ARC de la protéine PpC24 de *P. patens* a été utilisé comme groupe externe. L'arbre phylogénétique sépare bien les TNL et les CNL en deux clades distincts et confirme les prédictions faites avec les motifs MEME pour **57** BES de *M. rotundifolia* cv Regale (Figure 32 ; Annexe 4). Pour considérer la solidité des nœuds, le seuil de valeur de *bootstrap* retenu est de **70** %. En effet, il a été montré qu'à une valeur de *bootstrap* d'au moins **70** % correspondait une probabilité très élevée que le clade correspondant soit exact (Hillis et Bull, 1993). Ainsi, le clade des TNL compte **12** BES de *M. rotundifolia* cv Regale qui sont dans le même groupe phylogénétique que les protéines de résistance fonctionnelles L6, P2, Gro1.4, BS4, RY-1, KR1 identifiées chez le lin, la pomme de terre et le soja (Figure 32 ; Annexe 3).

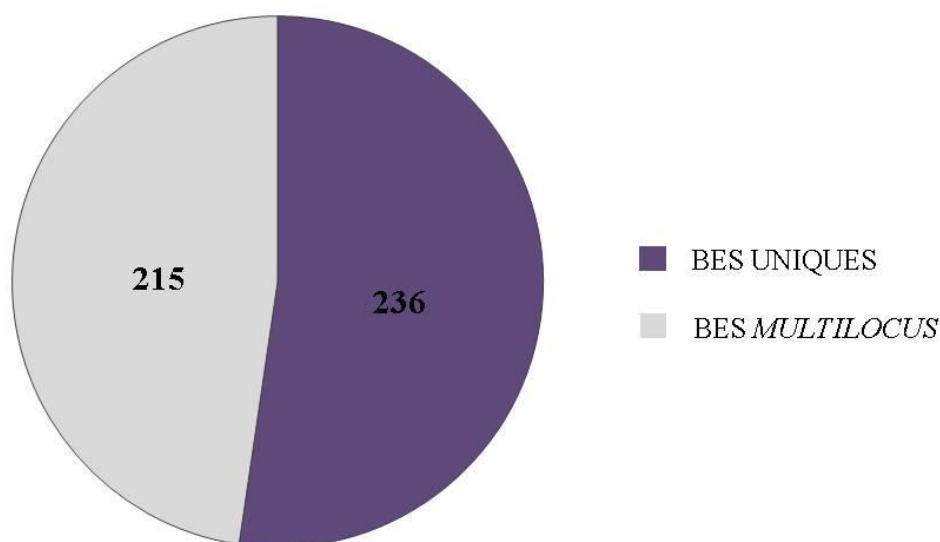
La branche des CNL compte **66** BES de *M. rotundifolia* cv Regale qui se divisent en 2 clades distincts CNL 1 et CNL 2. Le clade CNL 1 contient **21** BES qui se subdivisent en 2 sous-groupes, CNL 1.1 et CNL 1.2. Le sous-groupe CNL 1.1 compte **8** BES qui sont proches de la protéine de résistance au mildiou de la laitue (Dm3) alors que le sous-groupe CNL 1.2 compte **13** BES de *M. rotundifolia* cv Regale. Le clade CNL 2 compte **45** BES de *M. rotundifolia* cv Regale qui se subdivisent en 3 sous-groupes, CNL 2.1, CNL 2.2 et CNL 2.3 (Figure 32). Le sous-groupe CNL 2.1 compte **10** BES de *M. rotundifolia* cv Regale proches des protéines de résistance fonctionnelles BS2, HRT, RPP13, Pid3, Pikm1-TS, Pikip-1, Pi9, PIB, Lr10, MLA1 identifiées chez le piment, *A. thaliana*, le riz, le blé et l'orge. Le sous-groupe CNL 2.2 compte **12** BES de *M. rotundifolia* cv Regale proches des protéines de résistance Prf et Hero de la tomate et de la pomme de terre. Le sous-groupe CNL 2.3 compte **23** BES de *M. rotundifolia* cv Regale proches des protéines de résistance Pi5-1, Xa1, Cre1, Rp1-D, Pm3b, Rgd2a, Rpi-blb1, R3a, Rps1-k-1, P18 et FOM-2. Riz, *Aegilops tauschii*, du maïs, du blé, l'orge, de la tomate, du soja, du tournesol et du melon. A l'exception du sous-groupe CNL 1.2, qui ne présente aucune similitude de séquence avec une protéine de résistance fonctionnelle utilisée pour notre analyse, tous ces sous-groupes ont des valeurs de *bootstrap* élevées, ce qui suggère un ancêtre commun.



**Figure 32. Analyse des domaines NB-ARC de protéines de résistance fonctionnelles avec le programme CLUSTALX.** Un domaine NB-ARC de *Physcomitrella patens* (PpC24) est utilisé comme groupe externe. Les chiffres indiquent la valeur des *bootstrap* de 100 répliques. Les BES identifiées avec les motifs MEME sont indiquées avec une étoile (\*). Les TNL sont en rouge et les CNL sont en bleu

#### 2.4. Alignement des BES de *M. rotundifolia* cv Regale sur la séquence de *V. vinifera*

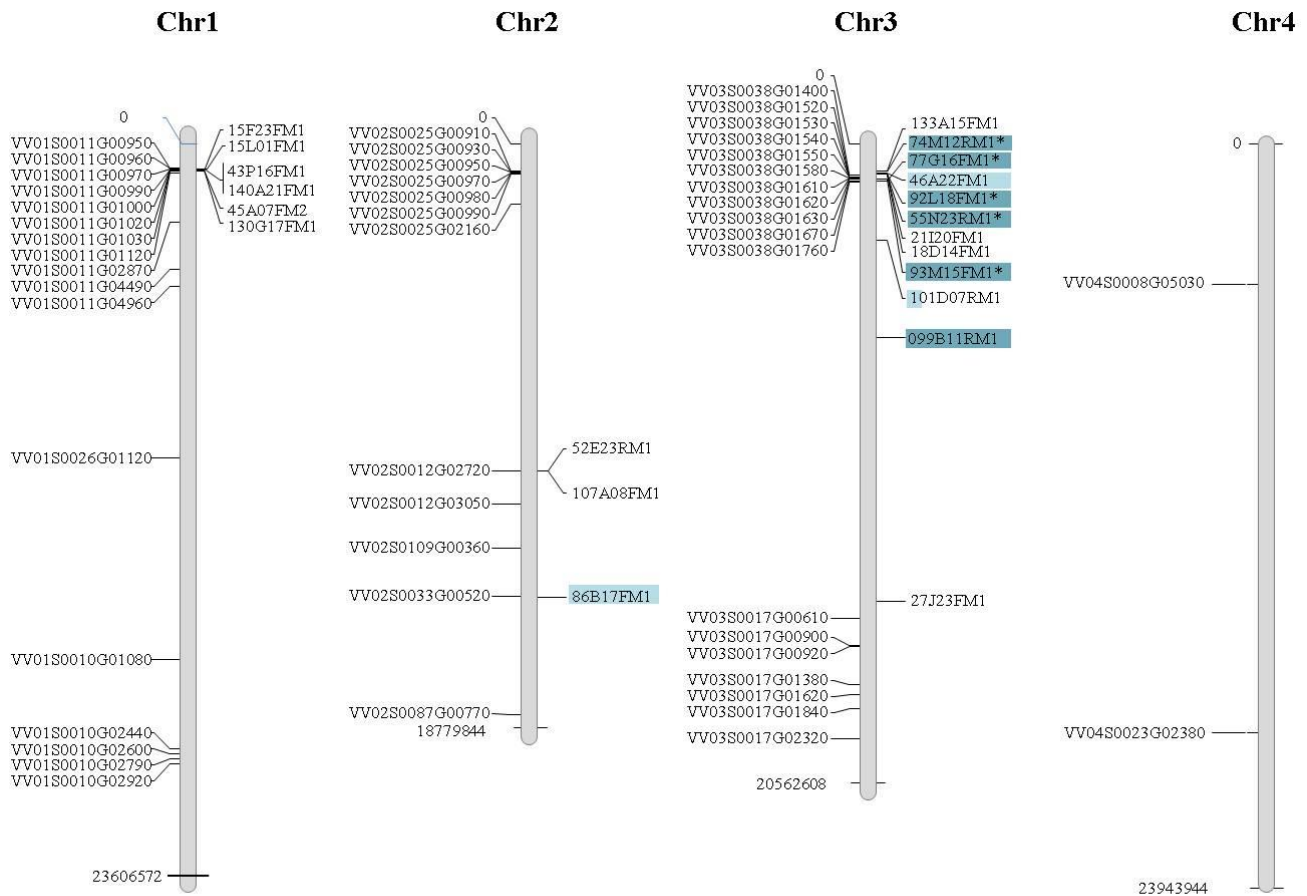
L'ensemble des **451** BES de *M. rotundifolia* cv Regale montrant une homologie de séquence avec le domaine NB-ARC ont été alignées sur la séquence de référence de *V. vinifera* cv PN40024 par BLASTN (E-VALUE =  $10^{-4}$ ) ; et la position de ces séquences sur le génome de référence a été visualisée en utilisant MAPCHART (Voorrips *et al.* 2001). Les BES avec des alignements *multiloci* ou ambigus ont été éliminées (Figure 33). Ainsi, les positions sur le génome de référence de *V. vinifera* de **236** BES uniques et **4** BES dupliquées ont été analysées (Figure 33). Il s'agit de **30** TNL et **142** CNL prédits par les motifs MEME et l'arbre phylogénétique, et de **72** gènes NBS qui n'ont pu être attribués aux classes des CNL ou TNL par l'utilisation des motifs MEME.



**Figure 33. Proportion des BES contenant des domaines NB-ARC alignés sur le génome de référence de *V. vinifera* cv PN40024 par BLASTN.**

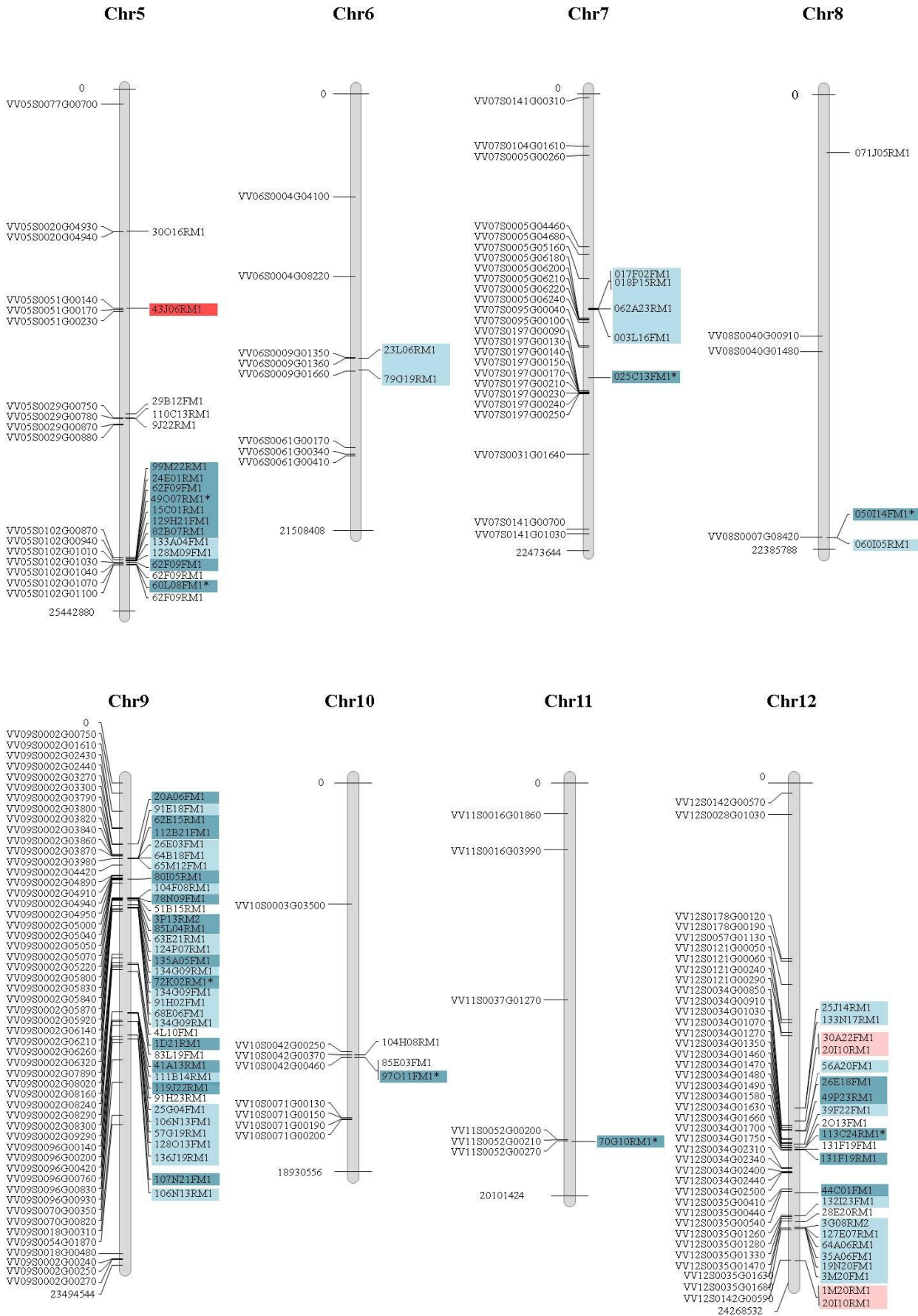
La plupart des *clusters* de gènes à motif NB-ARC de *V. vinifera* cv PN40024 sont recouverts par les BES de *M. rotundifolia* cv Regale contenant des séquences à motif NB-ARC. Les BES contenant les TNL (Figures 32 et 34), s'alignent majoritairement sur le chromosome 18 de *V. vinifera*, à l'exception de quelques membres qui sont sur les chromosomes 13, 15 et sur les séquences non assignées à un chromosome. Quant aux CNL, ils sont majoritairement localisés sur les chromosomes 3, 5, 9, 12, 13 et 19. Les sous-groupes CNL 1.1, CNL 1.2, CNL 2.1, CNL 2.2 et CNL 2.3 sont principalement localisés sur les chromosomes 14, 9, 3, 5 et 13, respectivement (Figures 32 et 34). En ne considérant que les gènes NBS utilisés pour l'analyse phylogénétique, il apparaît que les BES appartenant à un même groupe de

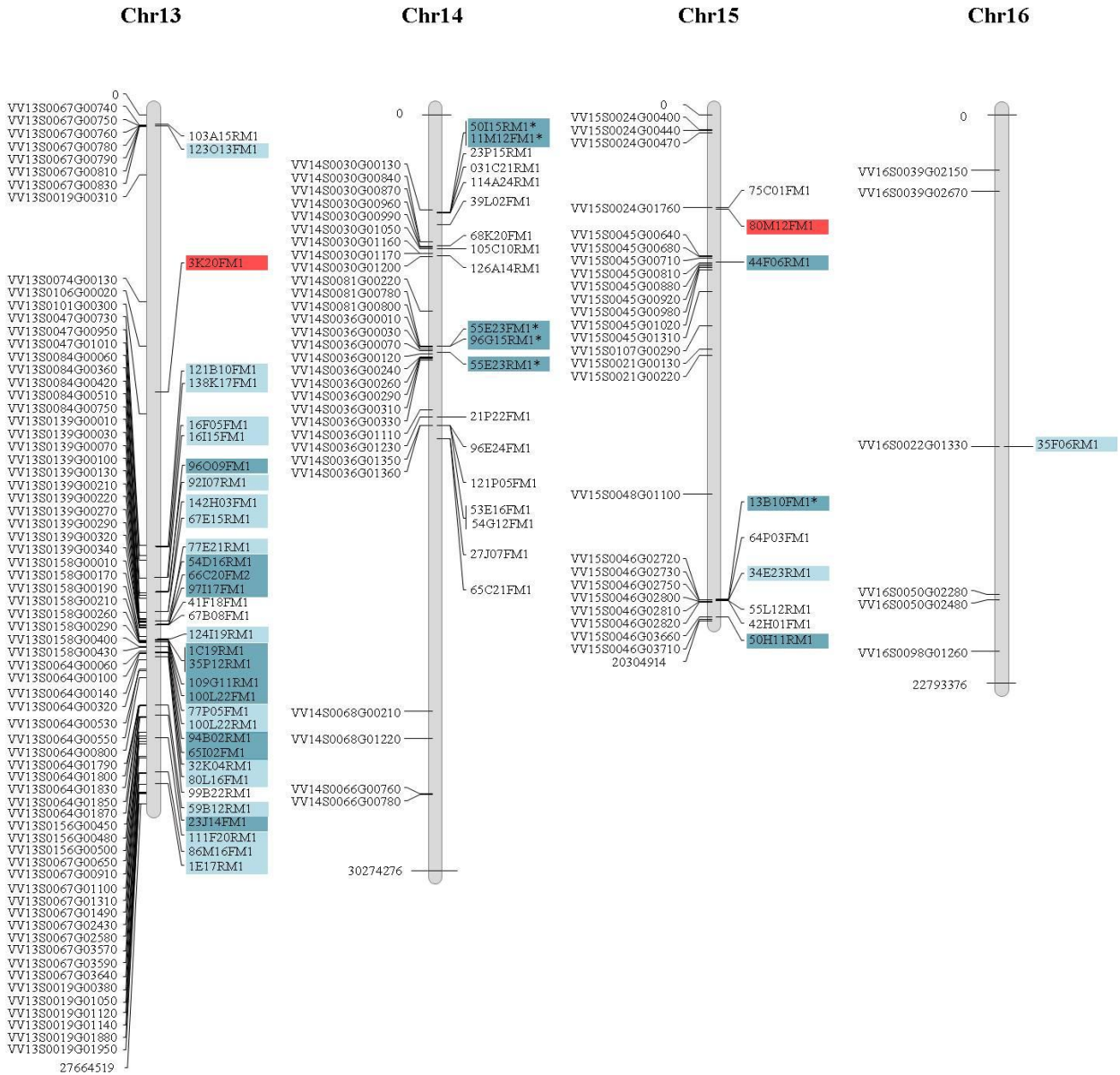
cette analyse sont en général sur le même chromosome, voire le même *cluster* de gènes. Par ailleurs, les coordonnées d'alignement des BES de *M. rotundifolia* cv Regale chevauchent dans certains cas les coordonnées de modèles de gènes proposés par l'annotation automatique du génome de référence de *V. vinifera* cv PN40024 (Annexe 5). Parmi les **30** TNL prédits, **3** chevauchent des domaines TIR et TIR-NB-ARC, **3** des gènes contenant un domaine NBS, **3** des protéines kinases et **21** ne chevauchent aucun modèle de gène ; parmi les **142** CNL prédits, **52** chevauchent des gènes NBS, **1** une phosphoprotéine et **87** ne chevauchent aucun modèle de gène ; parmi les **72** gènes NBS non-classifiés, **13** correspondent à des gènes NBS, **1** à un TIR-NB-ARC et **58** n'ont aucune correspondance. Plusieurs groupes de gènes NBS de *M. rotundifolia* cv Regale sont localisés dans des régions chromosomiques où la résistance aux maladies fongiques, comme le mildiou et l'oïdium, ont été précédemment détectées. C'est le cas des chromosomes 12 et 18 (Fischer *et al.* 2004 ; Barker *et al.* 2005 ; Coleman *et al.* 2009), les chromosomes 14 et 15 (Akkurt *et al.* 2007 ; Fischer *et al.* 2004 ; Dalbó *et al.* 2001). Ceci confirme qu'à cette échelle une synténie existe entre *V. vinifera* et *M. rotundifolia* qui pourra être exploitée dans les étapes de cartographie génétique de résistance aux maladies qui sont en général réalisées dans un contexte de complexe d'espèces (Di Gaspero *et al.* 2007). Dans la suite de mon travail, j'ai voulu vérifier comment cette synténie se maintient à une échelle plus fine.

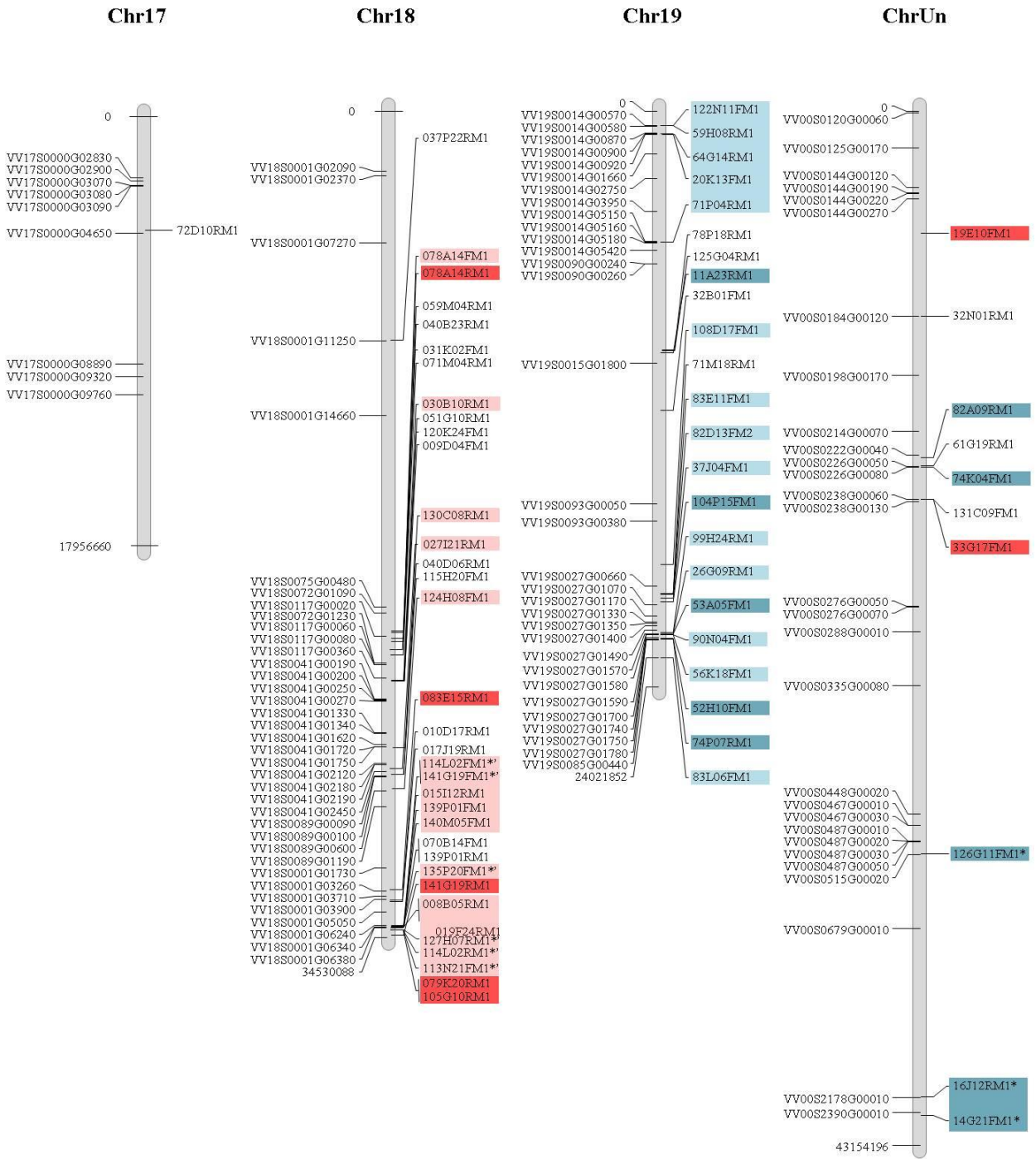


**Figure 34. Organisation chromosomique des gènes de résistances à domaine NB-ARC de *V. vinifera* cv PN40024 et de *M. rotundifolia* cv Regale.** Les séquences nucléotidiques des BES de *M. rotundifolia* cv Regale, qui présentent une similarité avec les domaines NB-ARC de protéines de résistance fonctionnelles, sont localisées sur les groupes de liaison de *V. vinifera* cv PN40024 après un alignement par BLASTN. Les BES assignées aux CNL par l'analyse phylogénétique sont en turquoise sombre et celles assignées aux CNL par les motifs MEME RNBS-A et RNBS-D non-TIR sont en turquoise claire. Pareillement, les BES assignées aux TNL par l'analyse phylogénétique sont en rouge sombre et celles assignées aux TNL par les motifs MEME RNBS-A et RNBS-D TIR, sont en rouge clair. L'ensemble des BES utilisées dans l'analyse phylogénétique ont les motifs RNBS-A et RNBS-D, à l'exception de celles désignées par une étoile (\*). Les différents groupes de liaison de *V. vinifera* cv PN40024 sont nommés chr1-Un. Les gènes de résistances à motif NB-ARC de *V. vinifera* cv PN40024 sont à gauche des différents chromosomes et ceux de *M. rotundifolia* cv Regale sont à droite.









**CHAPITRE II :**  
**ETUDE DE LA MICROSYPNTENIE ENTRE *V. VINIFERA* ET**  
***M. ROTUNDIFOLIA* AU NIVEAU DES *LOCI* MAJEURS DE**  
**RESISTANCE AU MILDIOU ET À L'OÏDIUM DE LA VIGNE**

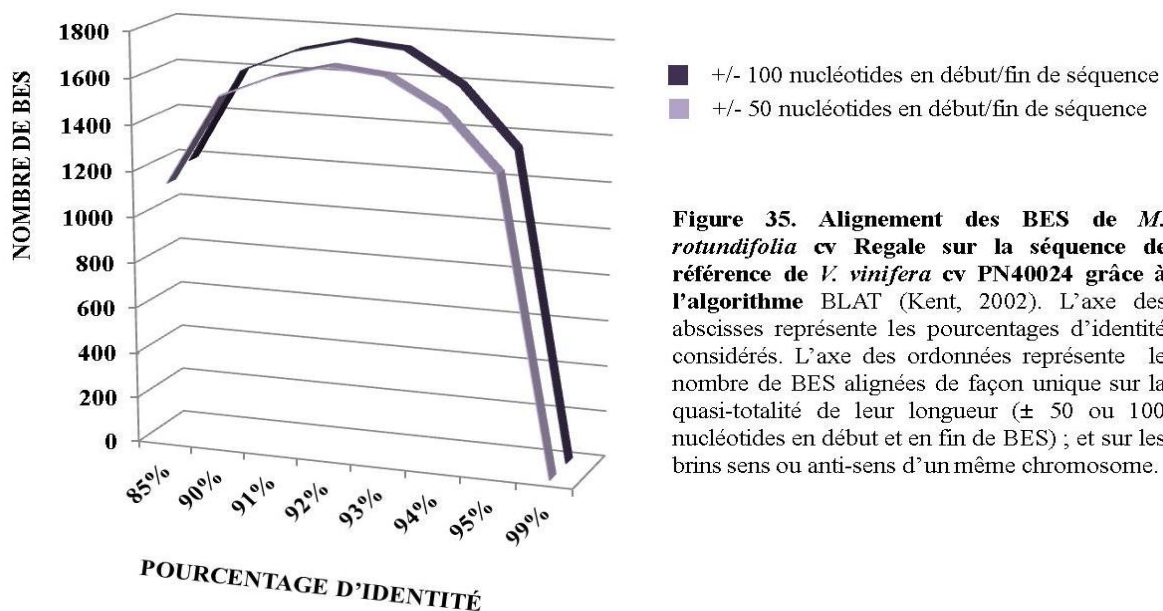
Le travail en cours se concentre sur deux régions des chromosomes **12** et **18** où des gènes/QTLs de résistance à l'oïdium ou au mildiou ont été détectés, co-localisant avec des *clusters* de gènes contenant des motifs NBS-LRR sur la séquence de référence du génome de *V. vinifera* (Moroldo *et al.* 2008). La région du chromosome 18 à laquelle nous nous sommes intéressés est située entre les marqueurs génétiques VMC2A7 et UDV-108 et contient les *loci* de résistance au mildiou *Rpv2* détecté chez *M. rotundifolia* cv 'Trayshed' et *Rpv3* détectés chez l'hybride interspécifique Bianca (Merdinoglu *et al.* 2003 ; Fischer *et al.* 2004 ; Welter *et al.* 2007 ; Bellin *et al.* 2009). Celle du chromosome 12 contient le *locus* de résistance au mildiou, *Rpv1* génétiquement co-localisé avec le *locus* de résistance à l'oïdium, *Run1* (Barker *et al.* 2005 ; Dry *et al.* 2009). Ce *locus* de résistance *Run1/Rpv1* a été introgressé de *M. rotundifolia* cv G52 dans *V. vinifera* (Pauquet *et al.* 2001 ; Dry *et al.* 2011). Une carte physique locale a été construite dans cette région (Barker *et al.* 2005) et en partie séquencée : nous disposons notamment d'un supercontig (917 559 pb) et de deux clones BAC (**82O20** : 37 192 pb ; **97A19** : 93 632 pb) qui couvrent les deux *loci* (Feechan *et al.* 2013). Les marqueurs CB46.49, CB13.14 et 49MRP1.P2, définis dans la séquence de clones BAC de cette région candidate, encadrent un *cluster* de gènes de type TNL (Donald *et al.* 2002, Barker *et al.* 2005 et Dry *et al.* 2010) et la région séquencée contient 11 gènes de résistance type NBS-LRR. L'objectif de cette partie du travail était de déterminer si la séquence du génome de *V. vinifera* peut être utilisée pour (i) accélérer l'identification de gènes de type NBS-LRR chez *M. rotundifolia* cv Regale, (ii) accélérer la constitution de contigs de clones BAC *M. rotundifolia* cv Regale et (iii) étudier la microsyténie entre *V. vinifera* et *M. rotundifolia* à l'aide de clones BAC séquencés.

## **2.5. Alignement des BES de *M. rotundifolia* cv Regale sur la séquence du génome de *V. vinifera* cv PN40024 et sur le supercontig de la région du *locus* de résistance *Run1/Rpv1***

### **2.5.1. Détermination des paramètres d'alignement des BES sur le génome de *V. vinifera***

Dans une première approche nous avons simplement analysé l'information qu'apporte l'alignement des BES de *M. rotundifolia* cv Regale sur le génome de référence de *V. vinifera* cv PN40024. L'ensemble des BES de *M. rotundifolia* cv Regale a été aligné sur la totalité de la séquence de référence de *V. vinifera* cv PN40024 en utilisant l'algorithme BLAT (Kent, 2002). Nous avons recherché les conditions d'alignement optimales de courtes séquences génomiques de *M. rotundifolia* cv Regale avec le génome de référence de *V. vinifera* en faisant varier le pourcentage d'identité de **85 %** à **99 %** (Figure 35). Les alignements des BES devaient être uniques et : (1) les deux membres d'une même paire devaient être cartographiés sur

un même chromosome, (2) sur la quasi-totalité de leur longueur (+/- 50 ou 100 nucléotides en début et en fin) et (3) correctement orientées l'une par rapport à l'autre. Un seuil en terme de pourcentage d'identité trop faible a pour conséquence de diminuer le nombre de paires dont l'alignement est conservé car le nombre de séquences présentant des alignements multiples augmente, tandis qu'un pourcentage d'identité trop fort réduit le nombre de paires conservées par manque de sensibilité. Nous cherchions donc le pourcentage qui permettait de récolter le maximum de paires. En rapport avec ces critères, le plus grand nombre de BES de *M. rotundifolia* cv Regale alignées sur la séquence de référence de *V. vinifera* cv PN40024, et dans une bonne orientation, est obtenu avec un minimum de **92 %** d'identité. Ce seuil permettrait donc de révéler au mieux les similitudes existantes entre ces deux espèces.



**Figure 35. Alignement des BES de *M. rotundifolia* cv Regale sur la séquence de référence de *V. vinifera* cv PN40024 grâce à l'algorithme BLAT (Kent, 2002).** L'axe des abscisses représente les pourcentages d'identité considérés. L'axe des ordonnées représente le nombre de BES alignées de façon unique sur la quasi-totalité de leur longueur ( $\pm$  50 ou 100 nucléotides en début et en fin de BES) ; et sur les brins sens ou anti-sens d'un même chromosome.

### 2.5.1.1. Alignement des BES de *M. rotundifolia* cv Regale sur le génome de *V. vinifera*

Les BES de *M. rotundifolia* cv Regale ont donc été alignées sur la séquence du génome de *V. vinifera* cv PN40024 avec un minimum de **92 %** d'identité à l'URGI. Trois filtres supplémentaires du moins stringent (Filtre 1) vers le plus stringent (Filtre 3) ont été appliqués (cf. Matériels et Méthodes). Les résultats peuvent être consultés sur le navigateur du génome de la vigne (Figure 36 ;



[http://urgi.versailles.inra.fr/gb2/gbrowse/vitis\\_12x\\_pub/](http://urgi.versailles.inra.fr/gb2/gbrowse/vitis_12x_pub/)). Sur toute la collection de BES, **14 406** BES (correspondant à 7 203 clones BAC soit 14,6 % de la banque BAC), **11 754** (5 877 clones BAC soit 12,9 % de la banque BAC) et **8 488** (4 244 clones BAC soit 8,6 % de la banque BAC) (Tableau 8) ont ainsi été alignés en utilisant les filtres 1, filtre 2 et filtre 3, formant **1 111** (Filtre 3) à **1 553** (Filtre 1) contigs, s'étendant sur des régions du génome de *V. vinifera* de **143** Kb en moyenne (Tableau 8).

**Tableau 8. Alignement des BES de *M. rotundifolia* cv Regale sur la séquence du génome de *V. vinifera* cv PN40024 selon 3 critères différents d'alignement, du moins stringent (filtre 1) vers le plus stringent (filtre 3 ; cf. Matériels et Méthodes)**

	<b>Filtre 1</b>	<b>Filtre 2</b>	<b>Filtre 3</b>
<b>Couple de BES alignées de façon unique</b>	7 203	5 877	4 244
<b>Nombre de contigs de BAC (&gt; 2 BAC)</b>	1 553	1 340	1 111
<b>Taille moyenne des contigs (Kb)</b>	147,1	137	145,7
<b>Taille maximale (Kb)</b>	5 269,20	917,5	931,8
<b>Taille minimale (Kb)</b>	0,6	21,6	46,5

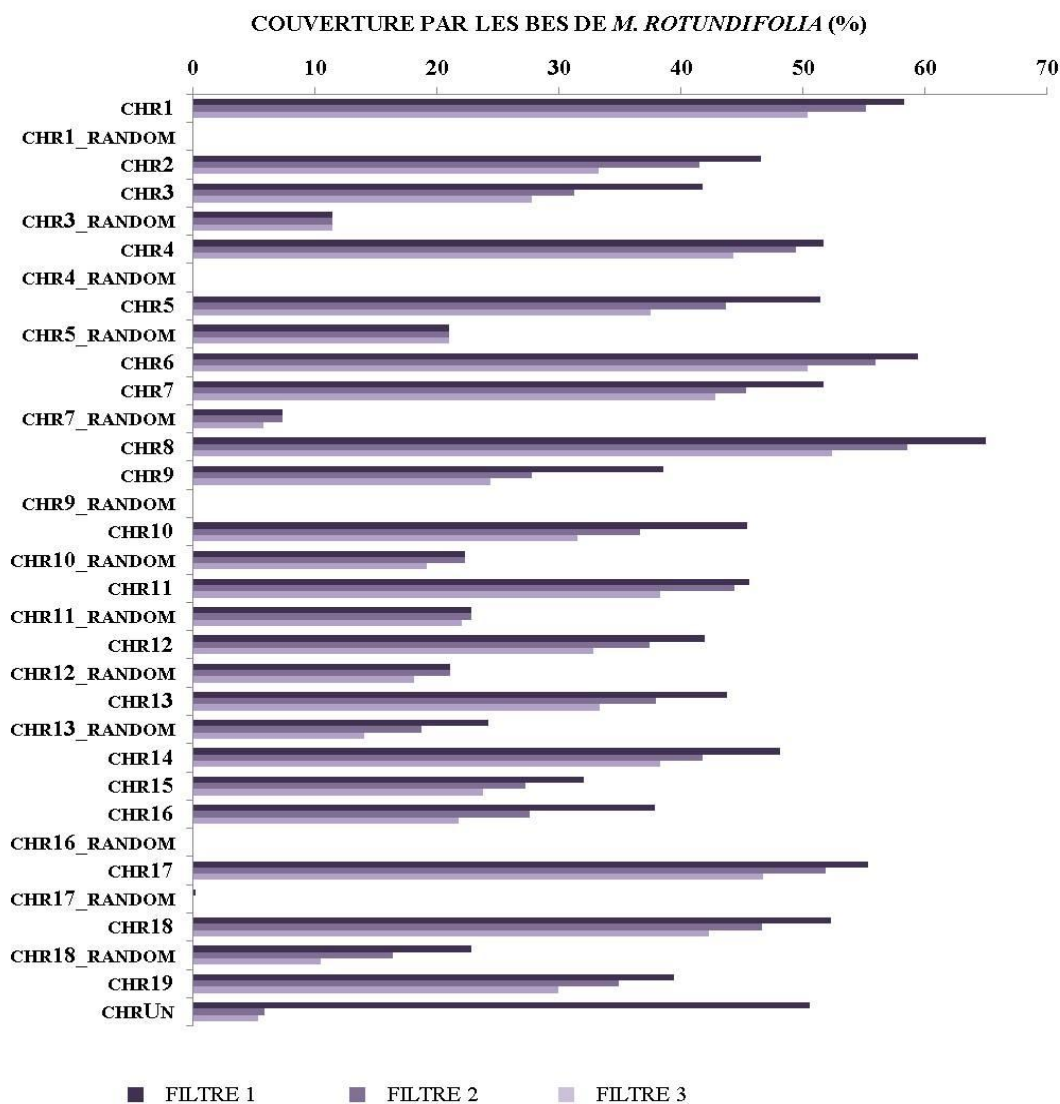
La couverture des chromosomes de *V. vinifera* cv PN40024 est variable (Figure 36). Les *random* chromosome, qui correspondent aux séquences non orientées mais assignées à un chromosome, et le chrUn, qui correspond aux séquences non assignées à un chromosome (Adam-Blondon *et al.* 2011), étaient beaucoup moins couverts (**10** % en moyenne pour les trois filtres) que le reste des chromosomes (**42** % en moyenne pour les trois filtres). Cela peut être dû à une plus grande proportion de séquences répétées dans les séquences non orientées mais également à des ruptures de contigs ou de couplage entre paires de BES sur les chromosomes *random* et ChrUn, sur lesquels l'ordre des *scaffolds* est aléatoire. Selon les critères d'alignement, les régions du génome de *V. vinifera*, sur lesquelles les contigs de BAC s'alignent, contiennent entre **19,5** % et **44,2** % de modèles de gènes prédits sur au moins **50** % (tableau 8).

Ces résultats suggèrent cependant qu'il y aurait des régions très similaires entre ces deux génomes. Il pourrait s'agir de régions fonctionnelles de gènes qui sont relativement préservées par l'évolution. La recherche de similarité de séquences, au niveau des *loci* de résistance déjà identifiés, permettrait de connaître l'homologie qui existe entre les gènes de résistance dans ces régions. Pour notre étude, uniquement les alignements des filtres 2 et 3 seront considérés.

### **2.5.1.2. Identification de clones BAC de *M. rotundifolia* cv Regale dont les BES s'alignent sur le supercontig de séquence contenant les gènes *Run1* et *Rpv1***

Sur toute la collection de BES, 7 clones BAC et 6 clones BAC (Figure 37) ont montré un alignement unique sur le supercontig contenant les gènes *Rpv1* et *Run1* en utilisant respectivement les filtres 2 et 3. Les BES du clone BAC MRRE2H068K18 (Filtre 2) étaient distants de **668** Kb ; ce qui est deux fois supérieur à la taille maximale des fragments d'ADN qui peuvent être clonés dans un BAC et au moins dix fois supérieur à la taille moyenne des clones de la banque. Ce clone BAC n'a pas été considéré pour la suite de notre étude. Ainsi, six clones BAC étaient alignés parmi lesquels les clones BAC MRRE2H062H06, MRRE3H117N19 et MRRE2H066F16 couvrent partiellement le *cluster* de gènes à motif NBS-LRR identifié par Barker *et al.* (2005). Ces résultats d'alignement des BES de *M. rotundifolia* cv Regale contre la séquence de l'espèce de *M. rotundifolia* cv G52 utilisée pour l'introgession montrent une faible couverture de la région chromosomique considérée. On s'attendrait à une plus grande similarité entre ces deux variétés du genre *M. rotundifolia*.





**Figure 36. Pourcentage des chromosomes de *V. vinifera* couvert par les contigs de clone BAC de *M. rotundifolia* cv Regale.**

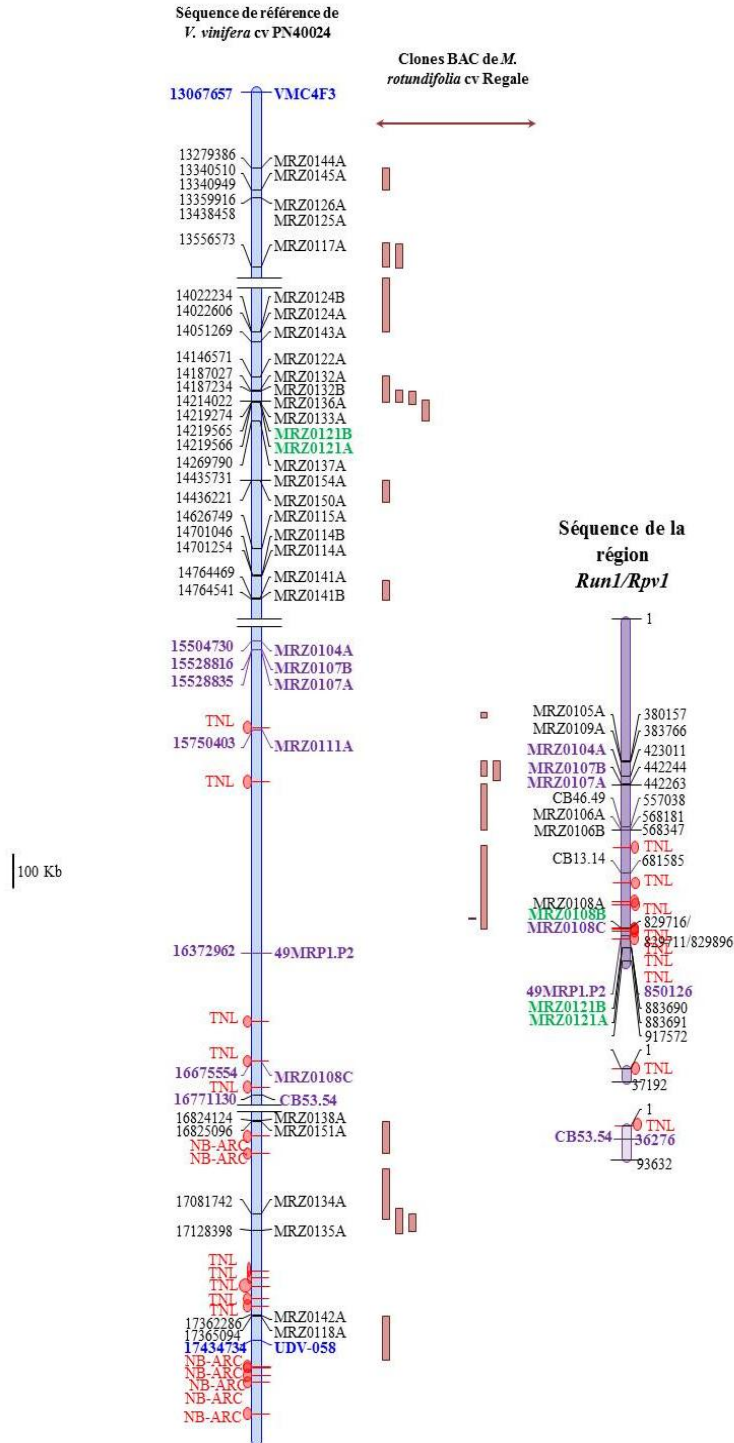
### 2.5.2. Identification des clones BAC de *M. rotundifolia* cv Regale dont les BES s'alignent sur le génome de référence dans les régions de *Run1/Rpv1* et *Rpv2/Rpv3*

En considérant les critères d'alignement les plus stricts, **30** clones BAC de *M. rotundifolia* cv Regale ont été alignés de façon unique sur la séquence de *V. vinifera* cv PN40024 entre les marqueurs VMC4F3-1 et VMC8G9 (région de **7,3 Mb** ; Tableau 9, Figure 37) et **30** clones BAC entre les marqueurs VMC2A7 et UDV-108 (région de **15,8 Mb** ; tableau 9) qui encadrent respectivement les gènes *Run1/Rpv1* et *Rpv2/Rpv3*. Ces régions contiennent respectivement **25** et **37** modèles de gènes contenant des motifs

NBS-LRR. Aucun des clones BAC alignés sur le supercontig ne s'aligne aussi sur la séquence de *V. vinifera* cv PN40024, et *vice versa*. Cette observation suggère une divergence de séquences dans cette région du chromosome 12 entre *V. vinifera* et *M. rotundifolia*. Ce résultat confirme la rupture de synténie observée lors de l'alignement du supercontig sur le génome de *V. vinifera* cv PN40024 (Feechan *et al.* 2013).

**Tableau 9. Alignement des BES de *M. rotundifolia* cv Regale dans les régions de *Run1/Rpv1* et *Rpv2/Rpv3* de la séquence du génome de *V. vinifera* cv PN40024. et du supercontig**

	Chromosome 12		Chromosome 18
	VMC4F3-1/VMC8G9	Supercontig	VMC2A7/UDV-108
Couple de BES alignées de façon unique	30	6	30
Nombre de contigs de BAC ( $\geq 2$ BAC)	6	2	4
Taille moyenne des contigs (Kb)	134	139	151
Taille maximale (Kb)	183	221	180
Taille minimale (Kb)	56	1,5	22



**Figure 37.** Alignement des clones BAC de *M. rotundifolia* cv Regale via leur BES sur la séquence *M. rotundifolia* cv G52 contenant les gènes *Run1/Rpv1* (à droite) et sur la séquence de référence du génome de *Vitis vinifera* dans la région correspondante du chromosome 12 (à gauche). Les BAC sont représentés par un rectangle marron dont la taille correspond à la distance entre l'alignement des deux BES (cela ne correspond pas à la taille réelle des clones BAC). En violet, les marqueurs communs en position conservée entre la séquence de la région *Run1/Rpv1* et la région correspondante du chromosome 12 chez *V. vinifera*. En vert, des marqueurs communs mais en position non conservée (le marqueur MRZ108B dans une région non représentée sous le marqueur UDV-058).

### 2.5.3. Enrichissement des *loci* de résistance *Run1/Rpv1, Rpv2/Rpv3* en marqueurs moléculaires

#### 2.5.3.1. Ancrage *in silico* des marqueurs génétiques et physiques des régions génomiques

L'ensemble des marqueurs génétiques de la région VMC8G9, UDV-058, VMC4F3-1, CB46.49, CB13.14, 49MRP1.P2 et CB53.54 (Barker *et al.* 2005 ; Dry *et al.* 2010 ; Cipriani *et al.* 2011) ont été alignés par EPCR sur le génome de *V. vinifera* cv PN40024 et sur la séquence du supercontig de BAC d'origine *M. rotundifolia* cv G52 (Tableau 10 ; Figure 37). L'ensemble de ces marqueurs a ensuite été utilisé pour cribler par EPCR la collection de BES de *M. rotundifolia* cv Regale et seul le clone BAC **101L21** contient la séquence de VMC8G9 à son extrémité *Forward*.

**Tableau 10. Localisation *in silico* des marqueurs génétiques et moléculaires des régions contenant les gènes de résistance *Run1, Rpv1, Rpv2 et Rpv3* sur le génome de *V. vinifera* cv PN40024, les contigs de BAC de *M. rotundifolia* cv G52 et sur les BES de *M. rotundifolia* cv Regale**

	Marqueurs	Coordonnées	Taille de l'amplicon (pb)	Coordonnées sur les BES de <i>M. rotundifolia</i> cv Regale
<i>V. vinifera</i> cv PN40024	VMC2A7	Chr18 : 13456636..13456778	143	
	VMCNG2F12	Chr18 : 19902915..19903113	199	
	VVIN16	Chr18 : 23389686..23389846	161	
	UDV-305	Chr18 : 24868065..24868384	361	
	VMC7F2	Chr18 : 26896790..26896989	200	
	UDV-108	Chr18 : 29123113..29123353	241	
	VMC4F3-1	Chr12 : 13067657..13067837	181	
	49MRP1.P2	Chr12 : 16372962..16373169	208	
	CB53.54	Chr12: 16771129..16771506	378	
	UDV-058	Chr12 :17434734..17434887	154	
	VMC8G9	Chr12 : 20378485..20378656	172	101L21FM1: 415..551
<i>M. rotundifolia</i> cv G52	CB46.49	Supercontig : 557038..557584	547	
	CB13.14	Supercontig : 681585..681813	229	
	49MRP1.P2	Supercontig : 850126..850332	207	
	CB53.54	97A19:33253..33612	360	

Pareillement, les marqueurs génétiques qui délimitent la région des *loci* de résistance *Rpv2* et *Rpv3*, localisés sur le chromosome 18 du génome de *V. vinifera* cv PN40024, ont été utilisés pour cribler la collection de BES sans résultat positif.

Ces résultats suggèrent que le génome *M. rotundifolia* a divergé du génome de *V. vinifera* dans la région du *locus* de résistance *Run1/Rpv1*. Ces résultats sont très cohérents avec le faible taux de recombinaison observé dans cette région (Pauquet *et al* 2001 ; Adam-Blondon *et al.* 2011). La densité de marqueurs moléculaires n'est pas assez élevée dans la région d'intérêt pour faciliter l'analyse comparative entre *V. vinifera* et *M. rotundifolia*. Il est donc nécessaire pour améliorer la couverture physique de ces régions, (i) de développer de nouveaux marqueurs et (ii) de cribler par PCR la banque BAC avec l'ensemble des marqueurs disponibles dans ces régions.

#### **2.5.3.2. Développement des couples d'amorces uniques dans la région des *loci* de résistance**

Les BES des zones cibles ont été utilisées pour définir des couples d'amorces (Annexe 6). Ces couples ont ensuite été alignés sur le génome de référence de *V. vinifera* cv PN40024 grâce à une EPCR ou par BLASTN. Sur **76** couples d'amorces définis dans la région du *locus* de résistance *Run1/Rpv1*, **75** couples ont finalement été sélectionnés pour leur alignement unique sur le génome de référence, et entre les marqueurs VMC4F3-1 et VMC8G9 (Annexe 7). Ce sont précisément **21** sur **22** marqueurs définis dans les BES alignées sur le supercontig de *M. rotundifolia* cv G52 et les **54** marqueurs définis dans les BES alignées sur le génome de référence de *V. vinifera* cv PN40024. Le couple d'amorce MRZ0108B, défini dans une BES alignée sur le supercontig de *M. rotundifolia* cv G52, s'aligne de façon unique sur le génome de référence mais en dehors de la région délimitée par VMC4F3-1 et VMC8G9. Nous l'avons donc laissé de côté. Enfin, **2** marqueurs ont également été définis en utilisant la séquence du supercontig. Sur l'ensemble des **23** marqueurs définis dans les BES s'alignant sur le supercontig ou directement sur la séquence du supercontig, **15** s'alignent sur une zone d'environ 1,4 Mb (1 387 969 pb) du génome de *V. vinifera* qui délimite la région homologue entre les deux génomes sur laquelle nous avons ensuite travaillé plus particulièrement.

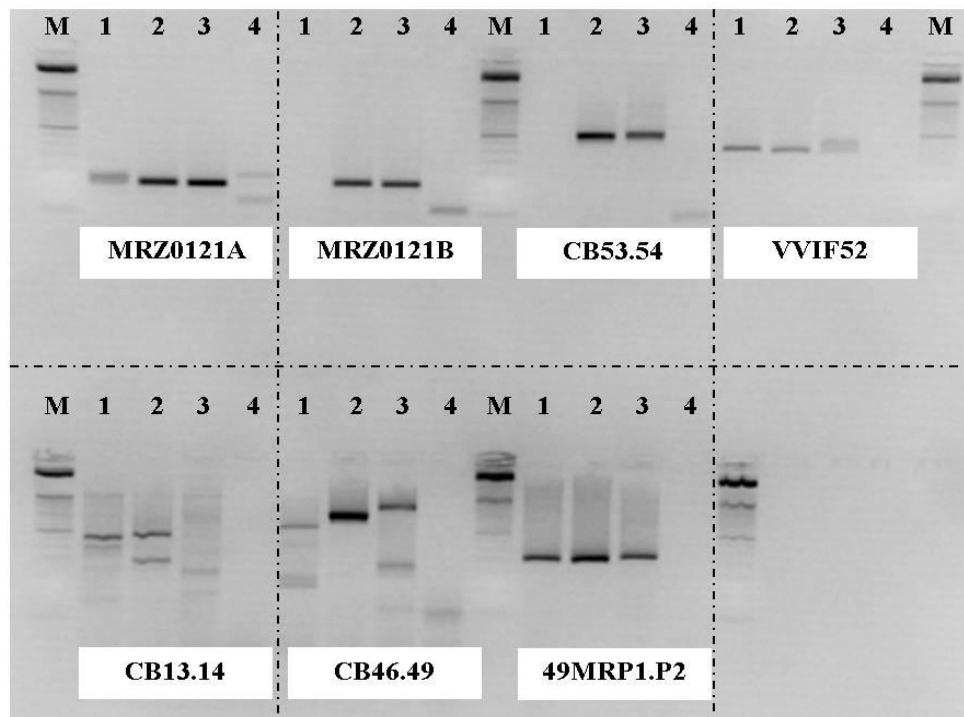
De la même façon, sur **102** couples d'amorces définis dans la région des *loci* de résistance *Rpv2/Rpv3*, **93** couples d'amorces ont été sélectionnés (Annexe 8) entre les marqueurs VMC7F2 et UDV-108. Les **9** autres soient n'amplifiaient pas, soient présentaient des amplifications multiples sur le génome de référence. Nous les avons donc laissés de côté.

## 2.6. CRIBLAGE PAR PCR DE LA BANQUE BAC DE *M. ROTUNDIFOLIA* CV REGALE

Les marqueurs nouvellement définis sur le *locus* de résistance *Run1/Rpv1* de *M. rotundifolia* cv G52 et dans la région des *locus* de résistance *Rpv2/Rpv3* ont été utilisés pour initier la construction d'une carte physique locale. L'objectif sera d'identifier les marqueurs de *M. rotundifolia* cv Regale qui sont uniques *in vitro* sur l'ensemble du génome de *V. vinifera* cv PN40024 pour positionner des clones BAC de *M. rotundifolia* cv Regale dans les régions d'intérêt, en criblant la banque BAC de *M. rotundifolia* cv Regale.

### 2.6.1. Tests d'amplification génomiques des marqueurs sélectionnés dans les régions d'intérêt

Pour valider l'approche *in silico*, les tests d'amplification ont été effectués avec un sous-ensemble de marqueurs disponibles dans les deux régions d'intérêt ; correspondant pour la région du *locus* de résistance *Run1/Rpv1*, aux marqueurs VMC4F3-1, UDV-058, VMC8G9, CB46.49, CB13.14, 49MRP1.P2 et CB53.54 (Barker *et al.* 2005), aux **22** marqueurs définis dans les BES alignées sur le supercontig de *M. rotundifolia* cv G52, et aux **2** marqueurs définis dans le supercontig ; pour la région du *locus* de résistance *Rpv2/Rpv3*, aux marqueurs VMC2A7, VMC7F2, VMCNG2F12, VVIN16, UDV-108 et UDV-305 (Bellin *et al.* 2009) ainsi qu'aux **102** marqueurs définis à partir des BES de *M. rotundifolia* cv Regale alignées dans cette région (Figure 38). Sur l'ensemble des **139** marqueurs testés, **119** ont amplifié des fragments uniques chez *V. vinifera* cv PN40024 contre **124** chez *M. rotundifolia* cv Regale ; **9** marqueurs sont dupliqués chez *V. vinifera* cv PN40024 contre **8** marqueurs dupliqués chez *M. rotundifolia* cv Regale (Annexe 9). Parmi l'ensemble de ces marqueurs, **110** marqueurs ont amplifié des fragments de tailles identiques à la fois chez *V. vinifera* cv PN40024 et *M. rotundifolia* cv Regale. Les marqueurs CB53.54, MRZ0121B, MRZ0069A, MRZ0094A et MRZ0161A n'ont amplifié que le génotype résistant *M. rotundifolia* cv Regale ; malgré les répétitions de l'expérience, les marqueurs MRZ0104A, MRZ0030A et MRZ0054A qui sont uniques chez *M. rotundifolia* cv Regale semblent être dupliqués ou sont hétérozygotes chez *V. vinifera* cv PN40024.



**Figure 38. Exemple de test d'amplification des marqueurs de la région du locus de résistance *Run1/Rpv1*.** Les ADN génomiques de *V. vinifera* cv PN40024 (1), *M. rotundifolia* cv Regale (2) et *V. vinifera* cv Bianca (3) ont été amplifiés par PCR et les fragments ont été séparés par électrophorèse sur un gel d'agarose 3%. Les noms des différents marqueurs sont écrits en gras. Le marqueur VVIF52 est utilisé comme contrôle positif et un contrôle négatif a été fait avec de l'eau (4). La taille des fragments est estimée grâce à un marqueur de taille de 50 pb (M).

Ces résultats montrent globalement une bonne fiabilité des marqueurs dans l'identification de marqueurs uniques sur les deux génomes, valorisables pour construire un contig de clones BAC de *M. rotundifolia* cv Regale dans la région des loci de résistance *Run1/Rpv1*, *Rpv2* et *Rpv3*. Au final, la totalité des marqueurs testés a été utilisée pour le criblage de la banque BAC de *M. rotundifolia* cv Regale.

### 2.6.2. Criblage par PCR de la banque BAC de *M. rotundifolia* cv Regale

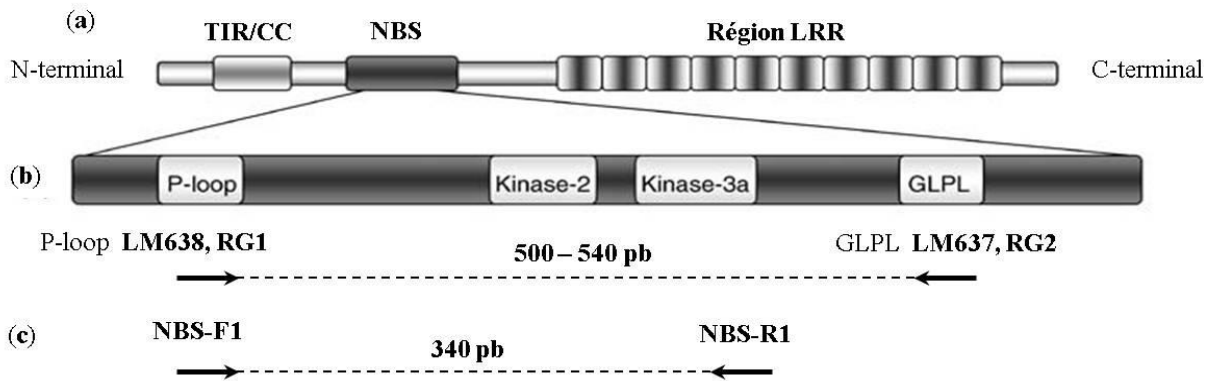
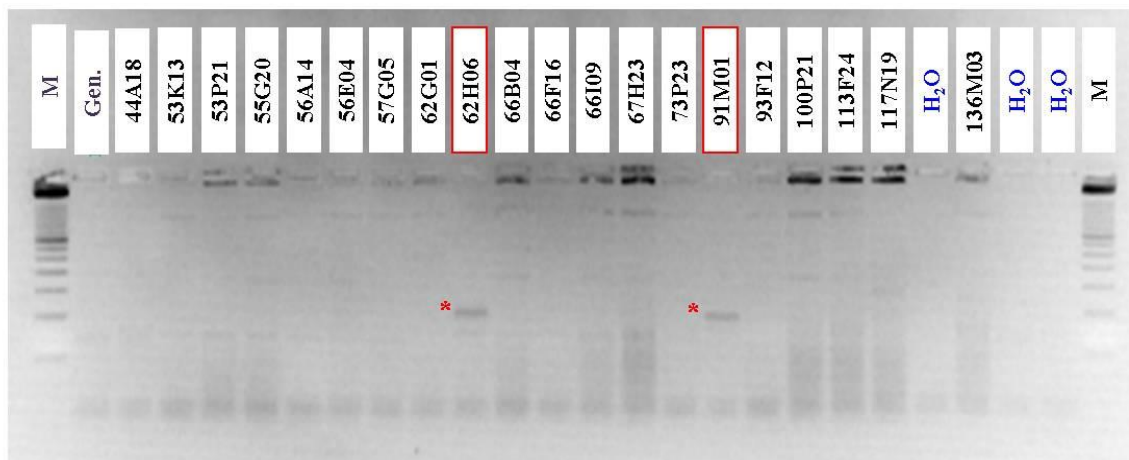
Les marqueurs physiques et génétiques liés aux loci de résistance *Run1/Rpv1*, *Rpv2* et *Rpv3* ont été utilisés pour cribler la totalité de la banque BAC de *M. rotundifolia* cv Regale par PCR (Annexe 9). L'amplification avec les marqueurs CB46.49 et CB13.14 résulte en un "smear" malgré les efforts d'adaptation des conditions PCR. Aucun clone BAC n'a été isolé avec les marqueurs 49MRP1.P2, VMC8G9 et UDV-305 alors que **15**, **15** et **5** clones BAC positifs ont été identifiés avec les marqueurs

VMC4F3-1, UDV-058 et CB53-54 du chromosome 12 ; et **5, 3, 4, 11 et 10** clones BAC positifs ont été identifiés avec les marqueurs VMC7F2, VMCNG2F12, VMC2A7, VVIN16 et UDV-108 du chromosome 18, respectivement. Pareillement, l'ensemble des marqueurs uniques sur le supercontig de *M. rotundifolia* cv G52 et dans la région du *locus* de résistance *Rpv2/Rpv3* a été utilisé pour cribler *in vitro* la banque BAC de *M. rotundifolia* cv Regale. Ainsi, **632** clones BAC ont été identifiés par PCR couvrant potentiellement les gènes NBS-LRR. Pour augmenter la couverture des *loci* de résistance *Run1/Rpv1* et *Rpv3*, une EPCR a été effectuée sur la collection de BES de *M. rotundifolia* cv Regale avec la totalité des marqueurs.

### **2.6.3. Identification de séquences NBS dans les clones BAC candidats**

Trois (**3**) couples d'amorces dégénérées NBS-F1 (P-loop)/NBS-R1 (Kinase 3), RG1 (P-loop)/RG2 (GLPL) et LM638 (P-loop)/LM637 (GLPL) ciblant les domaines P-loop, Kinase 2, Kinase 3a et le domaine hydrophobe GLPL de gènes NBS-LRR (Yu *et al.* 1996 ; Aarts *et al.* 1998 ; Kanazin *et al.* 1998 ; Figure 39A) ont été utilisés sur **43** clones BAC identifiés dans les deux régions. Un exemple de l'électrophorèse est présenté dans la figure 39B. Le seul couple qui a permis d'obtenir des résultats positifs est le couple RG1/RG2 Ces marqueurs ont permis d'identifier **4** clones BAC (**62H06, 25N15 et 32K01** sur le chromosome 12 et le clone BAC **91M01** sur le chromosome 18) qui ont présenté une faible amplification. La taille des fragments amplifiés est unique et d'environ **410 - 450** pb chez *M. rotundifolia* cv Regale, une taille légèrement inférieure à celle observée chez *V. amurensis*, *V. riparia*, le tabac, le lin et *A. thaliana* (**530** pb ; Di Gaspero *et al.* 2002). L'absence d'amplification observée chez le témoin *M. rotundifolia* cv Regale ainsi qu'avec les deux autres couples d'amorces dégénérées pose cependant problème. Ceci pourrait être expliqué par une compétition accrue des séquences amplifiables dans l'ADN génomique ou bien être le témoignage d'une amplification artefactuelle sur nos ADN de clone BACs. En effet, un niveau élevé de dégénérescence dans les oligonucléotides pourrait diminuer l'efficacité de l'amplification des brins spécifiques par compétition avec des amplifications moins spécifiques.

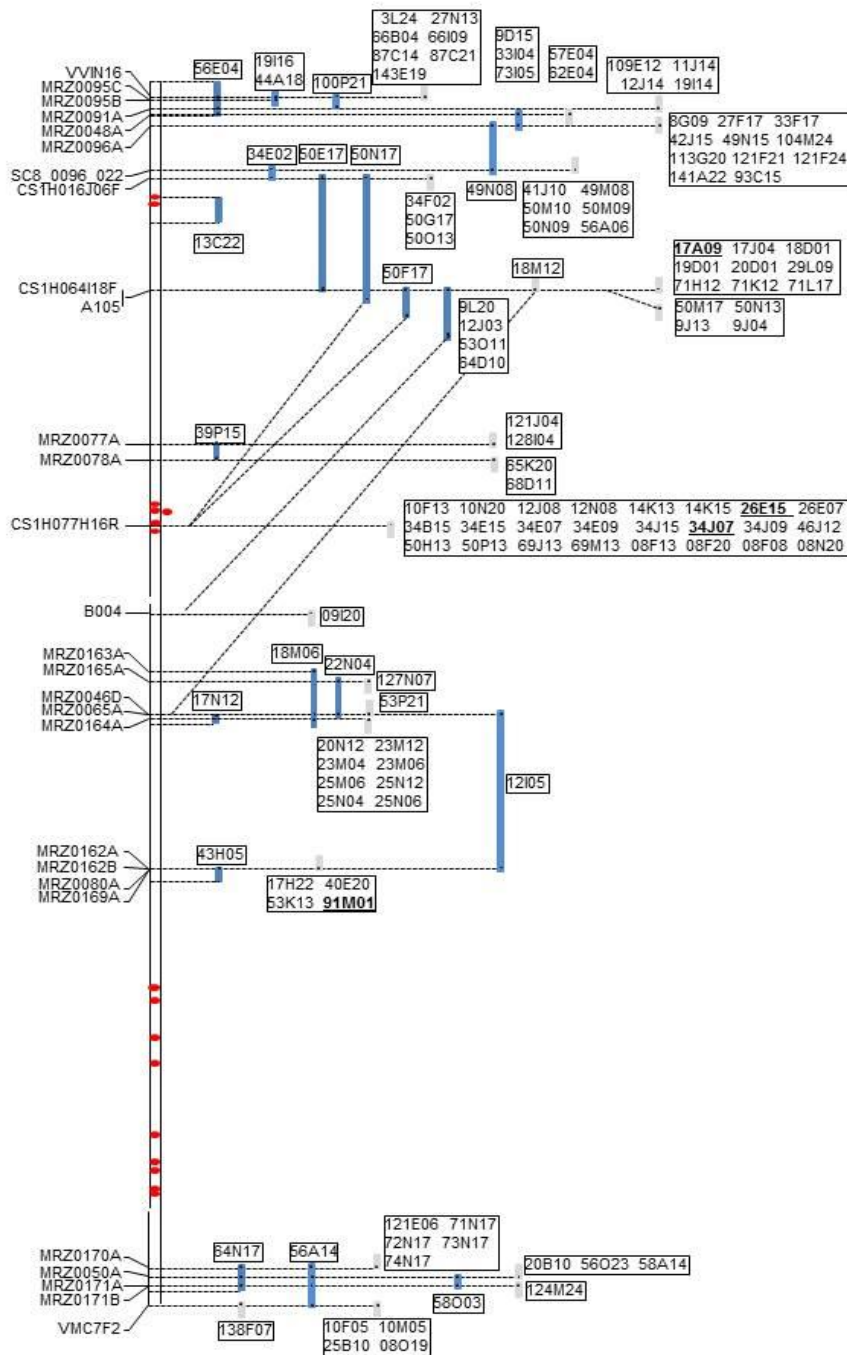


**A****B**

**Figure 39. Electrophorèse sur gel d'agarose 1% avec le couple d'amorces dégénérées RG1/RG2, spécifiques du motif NBS.** (A) Modèle de gènes de résistance de type NBS-LRR avec un domaine TIR ou CC à l'extrémité N-terminal (Di Gaspero *et al.* 2002). Les positions relatives des (a) caractéristiques fonctionnelles des protéines de résistance NBS-LRR, (b) les motifs les plus conservés au sein du domaine NBS, et (c) des amorces dégénérées obtenues à partir des études antérieures et utilisées pour l'amplification des clones BAC par PCR (Yu *et al.* 1996 ; Aarts *et al.* 1998 ; Kanazin *et al.* 1998). (B) Un exemple d'amplifications des clones BAC. Les clones BAC sélectionnés sont marquées par un astérisque.

Malgré l'évolution importante de la séquence entre les deux sous-espèces dans la région des gènes *Rpv1/Run1*, il a été possible de développer des marqueurs amplifiant les deux séquences. Nous avons pu le démontrer rapidement grâce à la connaissance d'une séquence complète de la région chez la Muscadine qui a permis de développer des marqueurs dans la région mais il semble possible d'utiliser des va-et-vient entre les deux séquences pour avancer dans la construction d'une carte physique. Il faut cependant rester prudent : des translocations de séquences portant des marqueurs (MRZ121A, MRZ121B, MRZ108C ; figure 37) ont été mises en évidence. Dans la région du chromosome 18, un remaniement plus important est mis en évidence entre la région délimitée par les marqueurs CS1H077H16R et MRZ0046D et la région autour des marqueurs CS1H064I18F et A105 puisque des marqueurs développés dans ces deux régions ancrent les

mêmes séries de BAC (Figure 40).



**Figure 40.** Résultat du criblage de la banque BAC *M. rotundifolia* cv Regale BAC par les différentes approches utilisées au cours de cette étude (alignement des BES, EPCR, PCR) avec pour cible la région contenant le pic du QTL *Rpv3*. Sur la gauche, la portion du chromosome 18 qui porte le *locus* chez *V. vinifera*. La position des modèles de gènes NBS-LRR est indiquée avec des ronds rouges. Les BAC ou groupes de BACs sont indiqués par des barres bleues (orientés) ou grises (non orientés). Les noms des BACs en gras et soulignés correspondent à ceux qui contiennent des homologues de NBS-LRR dans leurs BES ou par amplification avec des amorces dégénérées.

#### **2.6.4. Choix des clones BAC de *M. rotundifolia* cv Regale à séquencer**

L'ensemble des précédentes analyses a permis de (i) choisir des clones BAC à séquencer selon leur proximité aux gènes NBS-LRR cartographiés sur la séquence de référence de *V. vinifera* et le supercontig de *M. rotundifolia* cv G52 ; et d'étudier plus finement la synténie entre *V. vinifera* et *Muscadinia* dans la région des *clusters* de gènes de résistance. Les critères de choix sont détaillés dans les deux paragraphes suivants.

##### **2.6.4.1. Dans la région du *locus* de résistance *Run1/Rpv1***

Neuf (9) clones BAC de *M. rotundifolia* cv Regale ont été sélectionnés pour être séquencés dans la région du *locus* de résistance *Run1/Rpv1* (Figure 41). En se référant à l'annotation automatique du génome de *V. vinifera* cv PN40024 et de la région des gènes *Run1/Rpv1* (Feechan *et al.* 2013), ces clones BAC de *M. rotundifolia* cv Regale seraient des porteurs potentiels de séquences NBS-LRR.

Le clone BAC **62H06** est ancré *in silico* sur la séquence du supercontig *via* ses BES qui sont distants de **224** Kb. Les marqueurs MRZ0106A et MRZ0107A ont été définis dans ses extrémités *Forward* et *Reverse*, respectivement. Le clone BAC 62H06 a été identifié lors du criblage de la banque BAC avec le marqueur MRZ0106A et il possède éventuellement un gène à motifs NBS (Figure 39B).

Le clone BAC **66F16** est ancré *in silico* sur la séquence du supercontig *via* ses BES qui sont distants de **227** Kb. Il couvre la région contenant les RGA1 - RGA6 du *cluster* de TNL identifiés sur le supercontig du *locus* de résistance *Run1/Rpv1*. Enfin, il a été identifié lors du criblage de la banque BAC avec le marqueur MRZ0108A qui a été défini dans son extrémité *Reverse*.

Le clone BAC **20I10** a été positif lors du criblage de la banque BAC et de la collection de BES avec le marqueur MRZ0108A. En effet, ce dernier, qui a initialement été développé dans l'extrémité *Reverse* du clone BAC 66F16, est retrouvé *in silico* dans l'extrémité *reverse* du clone BAC 20I10.

Les clones BAC **34D01** et **62G01** ont été identifiés lors du criblage de la banque BAC avec les marqueurs MRZ0121A et MRZ0121B qui ont été développés en fin de séquence du supercontig. Ces clones couvriraient donc les RGA7 et RGA8 du *cluster* de TNL identifiés sur le supercontig du *locus* de résistance *Run1/Rpv1*.

Les clones BAC 25N15, 32K01, 11K09 et 27L20 sont porteurs du marqueur CB53.54 qui est lié au *locus* de résistance *Run1/Rpv1* de *M. rotundifolia* cv G52 (Barker *et al.* 2005). Aussi, les clones BAC 25N15 et 32K01 possèdent éventuellement un gène à motifs NBS (Figure 39B).

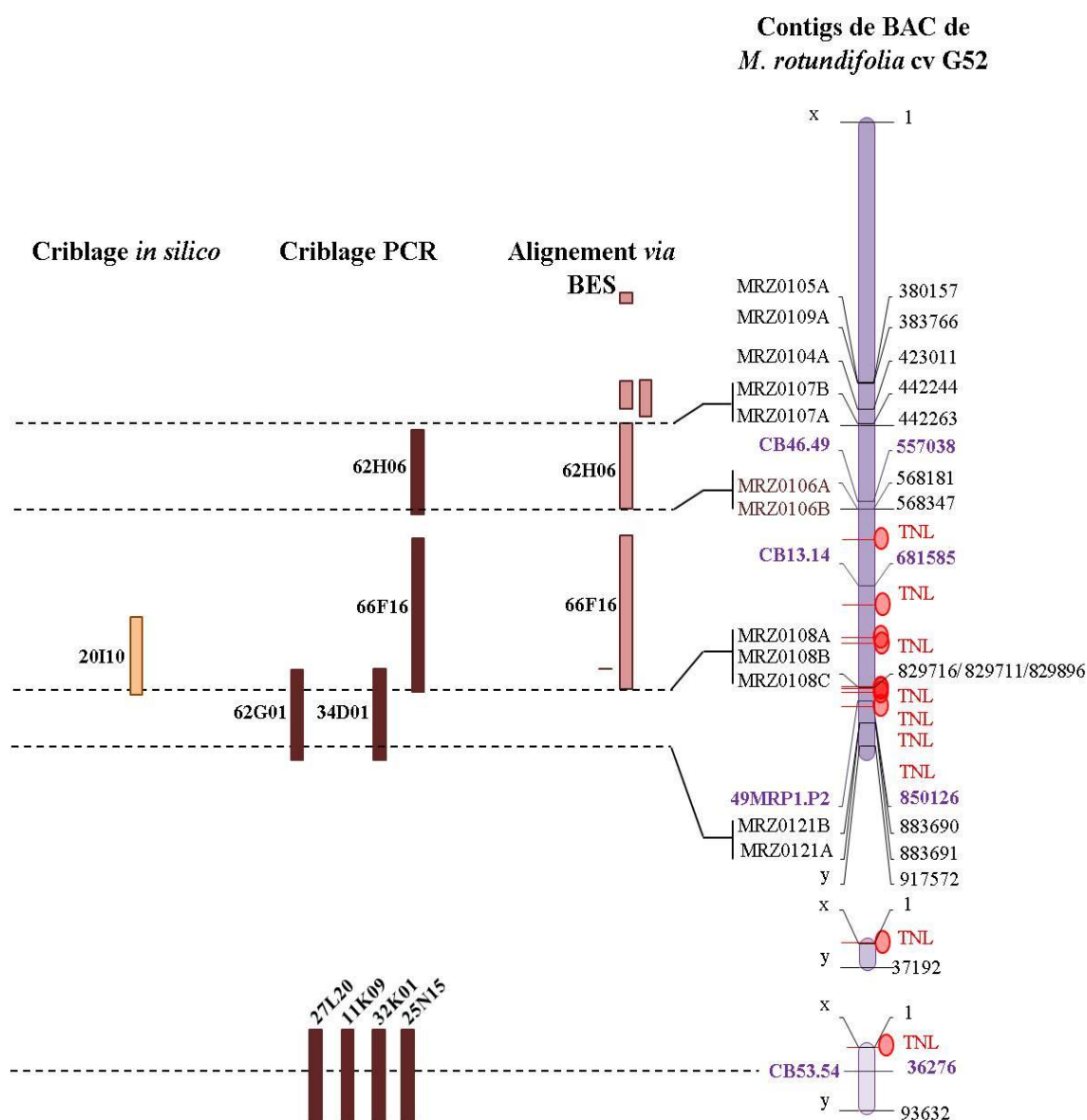


Figure 41. Position des BACs choisis pour être séquencés dans la région des gènes *Run1/Rpv1*.

#### 2.6.4.2. Dans la région du *locus* de résistance *Rpv3* (Bellin *et al.* 2009)

La région du chromosome 18 qui entoure le *locus* de résistance *Rpv2/Rpv3* dans l'hybride interspécifique complexe *V. vinifera* cv Bianca, est bornée par les marqueurs VMC2A7 et UDV-108, avec le pic de la résistance entre les marqueurs UDV-305 et VMC7F2 (Bellin *et al.* 2009). Cette région est relativement grande avec très peu de marqueurs et de BES alignées dans les régions où des *clusters* de gènes NBS-LRR ont été identifiés sur la séquence de référence de *V. vinifera*. Ainsi, les choix de **9** clones BAC ont été fait entre les marqueurs VVIN16 et UDV-108 et à proximité des *clusters* de gènes de résistance :

Les clones BAC **13C22** et **58B07** sont ancrés sur la séquence de *V. vinifera* cv PN40024 *via* leurs BES. Selon les coordonnées d'alignement, le clone 13C22 couvre des gènes à domaine NBS-LRR et le clone BAC 58B07 est dans une région chromosomique où sont localisés des gènes à domaine NBS-LRR (Figure 42). Ce dernier pourrait permettre d'agrandir la couverture de cette région chromosomique.

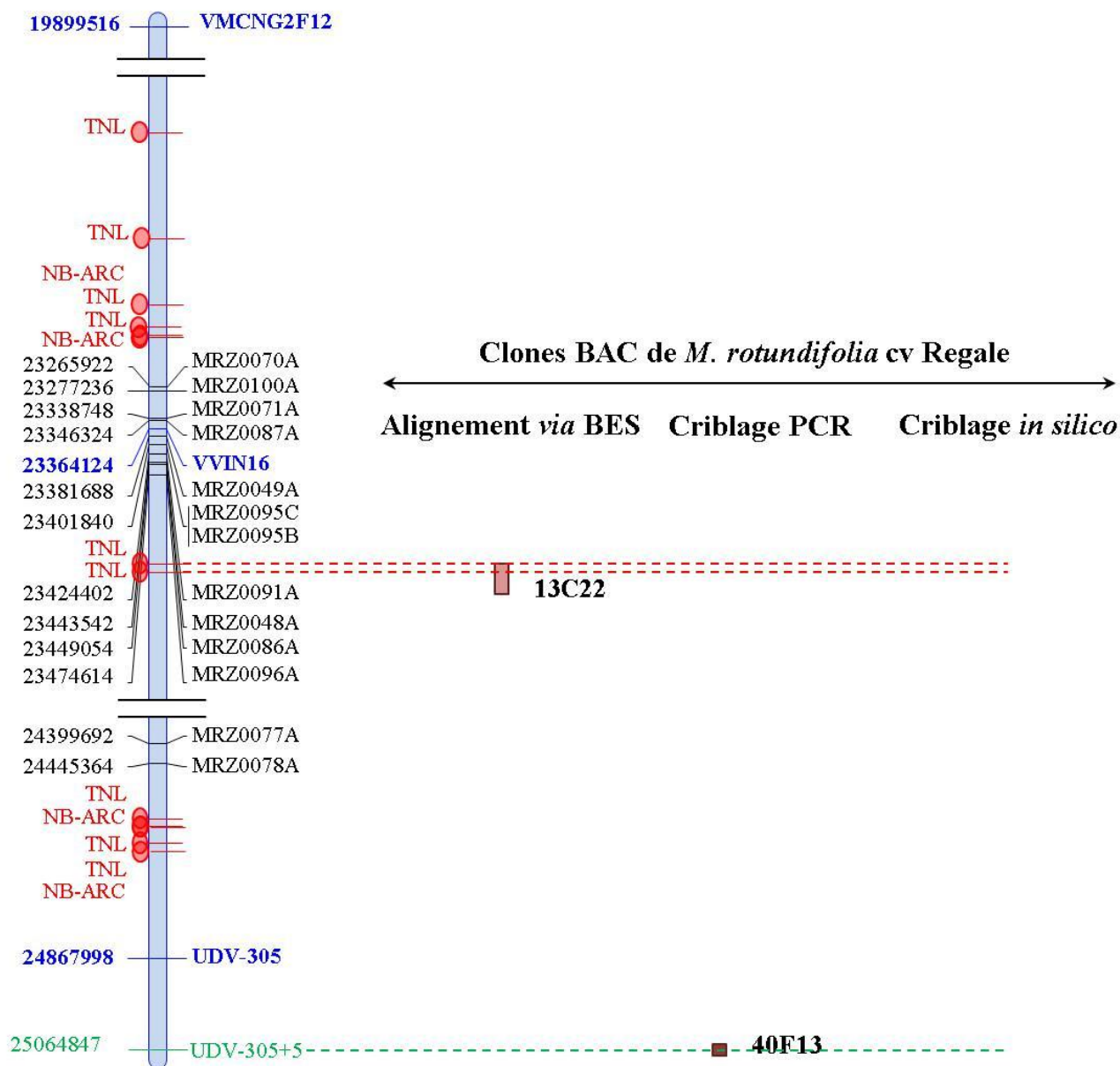
Le clone BAC **40F13** porte le marqueur UDV-305+5, marqueur lié au *locus* de résistance *Rpv3* chez *V. vinifera* cv Bianca (Thèse de Dario Copetti, Università Degli Studi di Udine, Italia ; Figure 42).

Les clones BAC **17N12**, **18M06** et **22N04** sont chevauchants et portent les marqueurs MRZ0065A et MRZ0046D (Figure 42). Ces clones pourraient permettre d'agrandir la couverture de la région entre les marqueurs UDV-305 et VMC7F2.

Le clone BAC **43H05** porte le marqueur MRZ0080A qui a permis d'identifier *in silico* la BES *Reverse* du clone BAC 91M01 (Figure 39B).

Les clones BAC **64N17** et **56A14** ont été identifiés lors du criblage de la banque BAC avec le marqueur MRZ0050A et le clone BAC 56A14 a également été identifié *in silico* avec ce même marqueur. Ces deux clones BAC encadrent un *cluster* de gènes NBS-LRR avec le clone BAC **58B07** (Figure 42).

**Séquence de référence de *V. vinifera* cv PN40024**



**Figure 42. Choix des BAC à séquencer dans la région du locus *Rpv3*.**



## 2.7. SÉQUENÇAGE ET ANNOTATION DES CLONES BAC CANDIDATS

### 2.7.1. Séquençage des clones BAC candidats

Pour ce faire, la présence des marqueurs spécifiques de chaque clone a d'abord été vérifiée par PCR sur les clones isolés à partir de colonies et des résultats positifs ont été obtenus pour seulement 8 clones BAC sur les 18 sélectionnés au départ. En parallèle, la taille des inserts a été estimée après digestion des ADN de BAC par l'enzyme *NotI*. La taille des inserts est comprise entre **40** et **105** Kb (Tableau 11). Les tailles observées pour ces clones BAC étaient différentes de celles estimées par alignement des BES sur la séquence de référence de *V. vinifera* cv PN40024, illustrant le polymorphisme de séquence entre ces deux espèces. Les 9 clones BAC ont été envoyés à séquencer avec la méthode **454** Roche (<http://www.454.com/>) au CNRGV de Toulouse, avec une couverture de 30 fois la séquence de chaque clone BAC analysé. Tous les *reads* ont été assemblés avec l'outil NEWBLER et plusieurs contigs ont été obtenus pour chaque clone (tableau 11).

**Tableau 11. Séquençage des clones BAC candidats pour les gènes *Run1*, *Rpv1*, *Rpv2* et *Rpv3* par la méthode de séquençage 454.**

Clone BAC	Chr	Insert (Kb)	Contigs (N°)	Somme des contigs (Kb)	BES disponibles (N°)	Distance entre BES (Kb)
<b>62H06</b>	12	120	5	88	2	41
<b>34D01</b>	12	70	2	47	2	-
<b>20I10</b>	12	100	5	85	2	-
<b>11K09</b>	12	80	3	66	2	-
<b>13C22</b>	18	80	2	66	2	71
<b>40F13</b>	18	78	5	46	2	-
<b>43H05</b>	18	45	2	25	2	30
<b>56A14</b>	18	105	3	83	2	-
<b>58B07</b>	18	40	2	18	2	22

- : information non disponible

Les sommes des contigs obtenus pour chaque clone BAC sont inférieures à la taille des inserts clonés. Cette différence de taille entre l'estimation réalisée après électrophorèse de digestions *NotI* et les



séquences assemblées varie entre **14** Kb (11K09 et 13C22) et **32** Kb (62H06 et 40F13). Elle pourrait être due, soit à une mauvaise estimation de la taille des inserts sur les électrophorèses ou à des problèmes d'assemblage, en particulier en relation avec la présence de séquences répétées. L'assemblage de ces contigs en une séquence complète de BAC permettrait donc de vérifier la taille des fragments obtenus. Nous allons donc tenter de combler ces *gaps* par des méthodes *in silico* et moléculaire.

### **2.7.2. Assemblage des contigs**

Nous avons donc tenté d'ordonner et orienter ces contigs en utilisant différentes approches.

Dans un premier temps, les séquences d'extrémités de BAC connues ont été recherchées dans les contigs de séquences pour repérer les extrémités. Puis les programmes d'assemblage PHRAP (Gordon *et al.* 2004) et MATCHER (Longden *et al.* 1999) ont été utilisés pour identifier les contigs chevauchants. Selon les résultats d'alignement, il n'y a pas de chevauchement entre les contigs d'un même clone BAC. Nous avons alors aligné, par BLASTN sur la séquence de référence de *V. vinifera* cv PN40024, tous les contigs de *M. rotundifolia* cv Regale (Figure 43 ; Annexe 10). Ces deux opérations ont permis de proposer un ordre et une orientation des contigs d'emblée pour les clones BAC 62H06, 34D01, 11K09, 13C22, 40F13, 43H05, 56A14 et 58B07. Par ailleurs, certaines coordonnées d'alignements sont chevauchantes. Les contigs 2 et 3 des clones BAC 62H06 et 56A14 auraient en commun **15** Kb et **19** paires de bases, respectivement. De même, les contigs 4 et 3 du clone BAC 40F13, ainsi que les contigs 1 et 2 du clone BAC 58B07 ont **4** Kb et **3** bases en commun, respectivement. En revanche, il y a des réarrangements manifestes entre la séquence de *V. vinifera* et les séquences du clone BAC 20I10 de *M. rotundifolia* cv Regale. Dans le cas du clone BAC 40F13, les meilleurs scores d'alignement sont obtenus sur le chromosome 8, au lieu du chromosome 18, comme attendu. Des expériences supplémentaires seraient nécessaires pour ordonner et orienter les contigs entre eux.

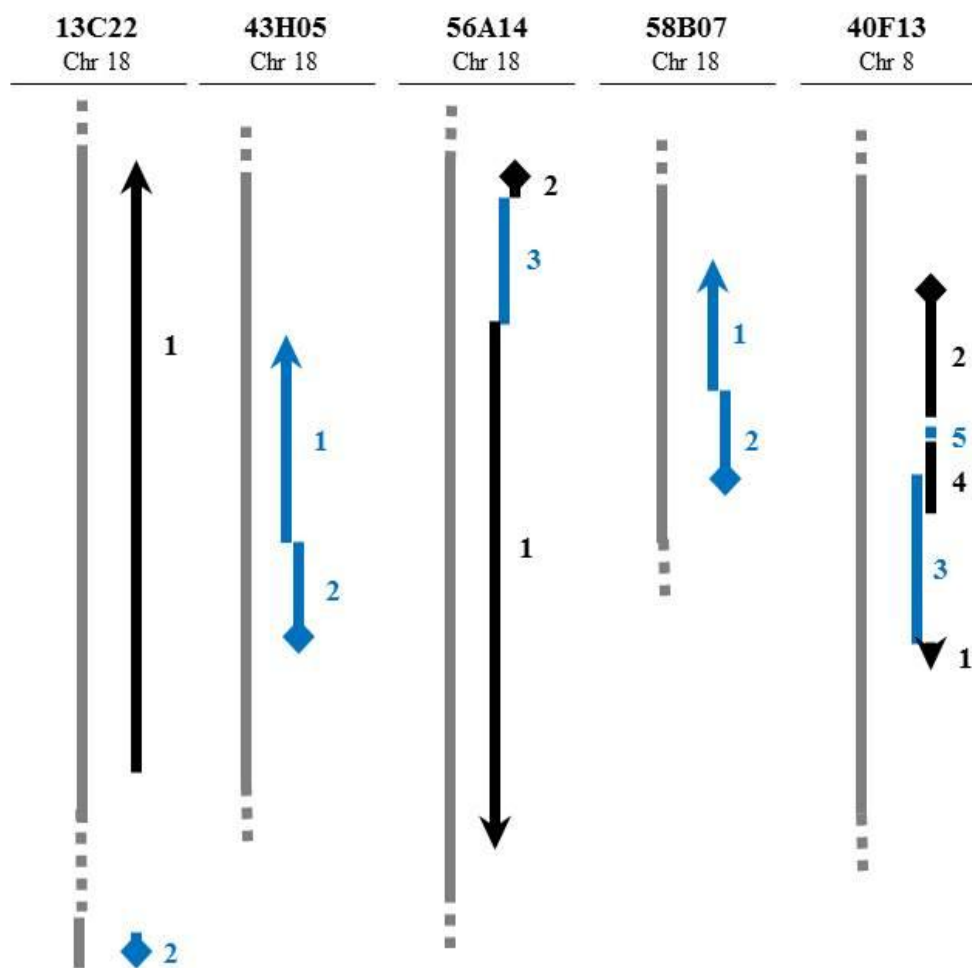


Figure 43. Alignement des contigs de clones BAC *M. rotundifolia* cv Regale sur la séquence du génome de *V. vinifera*. Les chiffres représentent les numéros de contig de chaque clone BAC; et les lignes bleues et noires les

contigs, avec les extrémités *Forward* et *Reverse* des clones BAC représentées par un losange et une flèche, respectivement. Les contigs en noir sont sur le brin sens et les contigs en bleu sont sur le brin anti-sens.

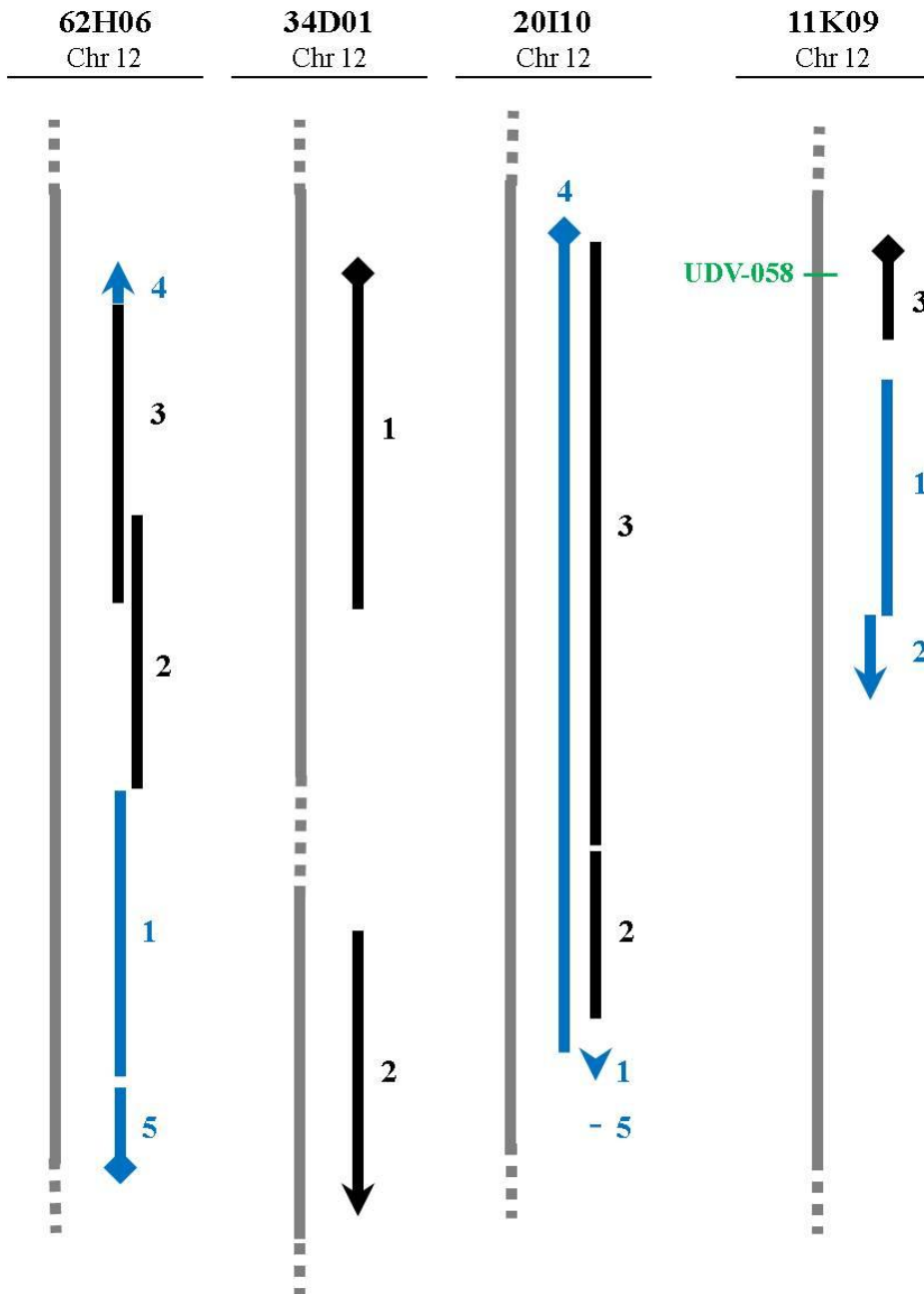


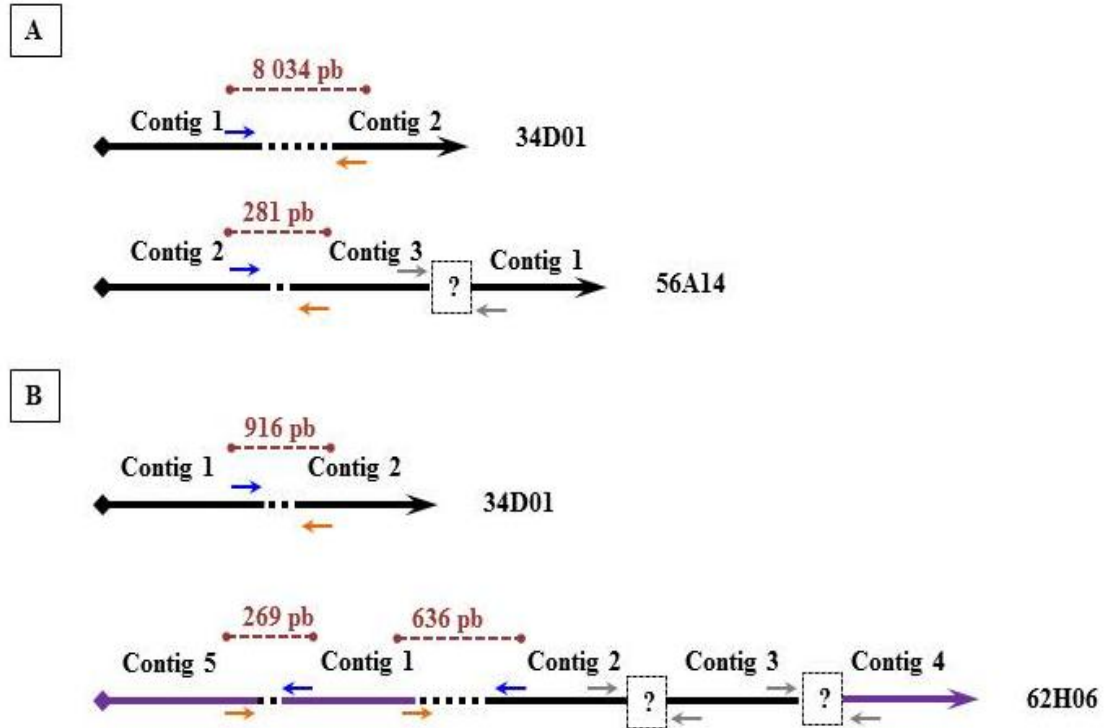
Figure 43. Suite.

Dans un deuxième temps, j'ai donc tenté à la fois de combler les régions manquantes par amplification PCR de l'ADN génomique et des clones BAC individuels et d'obtenir des indications sur l'ordre des contigs entre eux par des expériences PCR. Pour ce faire, les couples d'amorces ont été développés à l'extrémité de chaque contig, et de part et d'autre de la région manquante (Figure 44) ; et toutes les combinaisons de paires de ces oligonucléotides ont été testées en ne tenant pas compte de l'ordre proposé par l'alignement des *contigs* sur la séquence de *V. vinifera* cv PN40024. L'amplification PCR ne donnera de résultat positif que si les deux amorces correspondent à des séquences prises sur deux brins complémentaires et séparées de moins de quelques milliers de nucléotides.



**Figure 44. Modèle des différentes combinaisons de couples d'amorces pour l'assemblage des contigs d'un même clone BAC.** Les couples d'amorces sont développés à l'extrémité de chaque contig en chevauchant la région manquante. Les couples d'amorces sont représentés par des flèches désignant les amorces *Forward* (Bleu) et *Reverse* (Orange). Les extrémités *Forward* et *Reverse* des clones BAC sont représentées par un losange et une flèche, respectivement.

Sur **192** couples d'amorces définis (**66** avec la première méthode et **56** avec la deuxième), **55** couples ont finalement été sélectionnés pour leur alignement unique sur l'ensemble des contigs par BLASTN. Ces couples d'amorces ont tout d'abord été testés par EPCR sur le génome de *V. vinifera* cv PN40024 et sur le supercontig de la région introgressée de *M. rotundifolia* cv G52, et des résultats positifs ont été obtenus pour certains des couples d'amorces définis sur les séquences des clones BAC 34D01, 56A14 et 62H06 (Figure 45).



**Figure 45.** Assemblage des contigs grâce à une ePCR sur la séquence de *V. vinifera* cv PN40024 (A) et sur le supercontig de la région introgressée de *M. rotundifolia* cv G52 (B). Les couples d'amorces sont représentés par des flèches désignant les amorces *Forward* (Bleu) et *Reverse* (Orange). La taille des régions manquantes est donnée en paire de bases (pb) et elles sont représentées par des pointillés. Les extrémités *Forward* et *Reverse* des clones BAC sont représentées par un losange et une flèche, respectivement. Les contigs reverses/complémentés sont en violet.

Ces résultats montrent une différence importante de taille de la région manquante pour le clone BAC 34D01, en utilisant les séquences de *V. vinifera* cv PN40024 et de *M. rotundifolia* cv G52 comme matrice pour l'ePCR (Figure 45). En considérant l'éloignement des contigs du clone BAC 34D01 sur le génome de référence (Figure 43), il n'est pas surprenant d'obtenir une telle différence en ePCR. Ces séquences amplifiées *in silico* ont été extraites puis elles ont été alignées l'une contre l'autre, et sur le génome de référence de *V. vinifera* cv PN40024 par BLAT (Tableau 12).

**Tableau 12. Alignement des fragments amplifiés avec les couples d'amorces définis aux extrémités du clone BAC 34D01.** Les fragments de 8034 pb et de 916 pb ont été obtenus par EPCR en utilisant les séquences de *V. vinifera* cv PN40024 et de *M. rotundifolia* cv G52 par EPCR comme matrice, respectivement.

	Origine	Taille (pb)	Score	Identité (%)	Chr	Coordonnées sur le génome de <i>V. vinifera</i> cv PN40024
<b>Amplification EPCR</b>	<i>M. rotundifolia</i> cv G52	916	394	93,3	16	Chr16:6672003..6673178
<b>Amplification EPCR</b>	<i>V. vinifera</i> cv PN40024	8034	8034	100	12	Chr12:15742168..15750402
<b>34D01 Contig 2</b>	<i>M. rotundifolia</i> cv Regale	21692	11964	94,3	12	Chr12:15897583..16770278
<b>34D01 Contig 1</b>	<i>M. rotundifolia</i> cv Regale	25648	1191	87,1	12	Chr12:6975650..6978577

**Tableau 13. Résumé des séquences manquantes (gap) entre les différents contigs suivant leur alignement sur le génome de référence de *V. vinifera* cv PN40024 et le supercontig de *M. rotundifolia* cv G52.** Les alignements des contigs ont été faits par EPCR ou par BLAT.

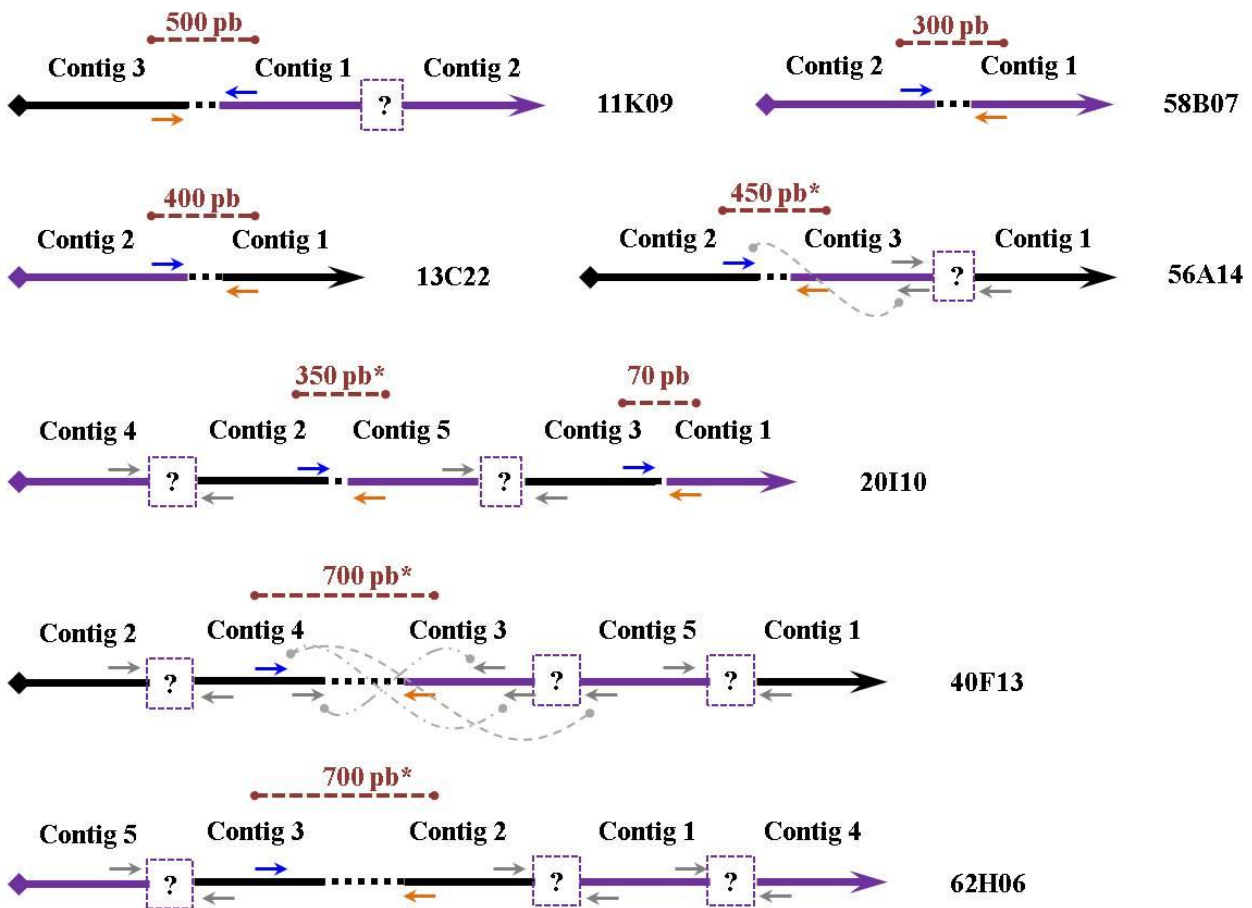
	Distance entre contigs estimée par BLAT (pb)		Distance entre contigs estimée par EPCR (pb)	
	Contigs conjoints*	<i>V. vinifera</i> cv PN40024	<i>V. vinifera</i> cv PN40024	<i>M. rotundifolia</i> cv G52
	Contig4 - Contig3	602	–	–
<b>62H06</b>	Contig2 - Contig1	455	–	636
	Contig1 - Contig5	2 271	–	269
<b>34D01</b>	Contig2 - Contig1	9 794 629	8 034	916
<b>11K09</b>	Contig3 - Contig1	4 358	–	–
	Contig1 - Contig2	1	–	–
<b>13C22</b>	Contig1 - Contig2	1 088 037	–	–
<b>43H05</b>	Contig1 - Contig2	0	–	–
<b>56A14</b>	Contig2 - Contig3	–	281	–
	Contig3 - Contig1	70	–	–

\* : Ordre des contigs selon le génome de référence

Pour valider l'approche *in silico*, les marqueurs uniques par BLASTN ont été utilisés pour un test d'amplification *in vitro* du génome de *M. rotundifolia* cv Regale et des clones BAC d'intérêt (Figure 46). Sur l'ensemble des 55 marqueurs testés, 24 amplifient uniquement le génome de *M. rotundifolia* cv

Regale, **6** les clones BAC d'intérêt et **1** couple d'amorces amplifie à la fois le génome de *M. rotundifolia* cv Regale et le clone BAC 11K09. Les résultats d'amplification confirment des ordres et orientation proposées sur la base des analyses *in silico* pour les BAC 11K09, 13C22, 56A14 et 58B07.

Pour rallonger les contigs obtenus, les fragments amplifiés avec les **7** marqueurs positifs sur les clones BAC ont été envoyés à séquencer par la méthode de SANGER (Sanger *et al.* 1975). Une fois les séquences obtenues, les programmes d'assemblage PHRAP (Gordon *et al.* 2004) et MATCHER (Longden *et al.* 1999) ont été utilisés pour identifier les éventuels chevauchements. Selon les résultats, les séquences manquantes n'arrivent toujours pas à être comblées. Les fragments de séquences amplifiés ont ensuite été alignés sur la séquence de référence de *V. vinifera*. Il en résulte que les séquences obtenues sont toutes localisées sur plusieurs chromosomes, et aucune ne s'aligne significativement avec le contig de séquence de la région des gènes *Run1/Rpv1*. Comme observé avec les séquences obtenues entre les deux contigs du clone BAC 34D01 (Tableau 12), les fragments amplifiés correspondent aussi à des éléments répétés lorsqu'ils sont alignés sur le génome de référence de *V. vinifera*. Ces observations soulèvent des interrogations sur le contenu en éléments répétés dans cette région chromosomique.



**Figure 46. Assemblage des contigs grâce à une amplification par PCR du génome des clones BAC d'intérêt.** Les combinaisons d'amorces, qui ont permis d'obtenir une amplification unique, sont représentés par des flèches désignant les amorces *Forward* (Bleu) et *Reverse* (Orange). Les couples d'amorces qui ont amplifiés plusieurs fragments sont en gris. La taille des régions manquantes est donnée en paire de bases (pb) et elles sont représentées par des pointillés; et les extrémités *Forward* et *Reverse* des clones BAC sont représentées par un losange et une flèche, respectivement. Les contigs reverse/complémentés sont en violet. Les produits d'amplifications qui ont été séquencés sont marqués par un astérisque.

### 2.7.3. Annotations structurale et fonctionnelle des contigs de BAC

Les contigs d'un même clone BAC ont été collés dans une seule et un nombre aléatoire de 'N' a été ajouté comme séparateur entre deux contigs, avec les BES en début et en fin de contig. L'annotation des gènes par comparaison de séquences, été faite en utilisant ARTEMIS (Rutherford *et al.* 2000 ; Carver *et al.* 2008) comme outil de visualisation et d'intégration des résultats.



### 2.7.3.1. Méthode de prédiction *ab initio*

La prédiction des gènes a été faite avec le programme FGENESH (Solovyev *et al.* 2006) entraîné pour *V. vinifera* pour identifier automatiquement les modèles de gènes dans les contigs de *M. rotundifolia* cv Regale. Ainsi, selon l'annotateur automatique, les longueurs minimale et maximale d'une ORF sont de **27** nucléotides (sur les contigs 13C22 et 20I10) et de **3 756** (sur le contig 11K09), respectivement ; soit des protéines prédites de **9** acides aminés et **1 252** acides aminés, respectivement (Annexe 11). Pour l'ensemble des contigs, FGENESH a prédit **18** modèles de gènes à **1** seul exon et **35** modèles de gènes contenant **2** à **16** exons ; avec une moyenne de **6** exons par gène, ce qui est comparable à la moyenne de **5** exons par gène rapporté par Jaillon *et al.* (2007 ; Tableau 14). Les modèles de gènes proposés ont ensuite été vérifiés manuellement en intégrant d'autres informations.

**Tableau 14. Prédiction automatique des gènes dans les contigs de *M. rotundifolia* cv Regale par FGENESH (Solovyev *et al.* 2006)**

	CHROMOSOME 12				CHROMOSOME 18				
	20I10	34D01	11K09	62H06	13C22	40F13	43H05	56A14	58B07
<b>Longueur de séquence (pb)</b>	85 362	47 498	65 898	88 284	65 587	46 555	25 191	83 582	17 800
<b>Nombre de gènes</b>	9	5	5	13	7	3	2	6	2
<b>Nombre d'exons</b>	51	21	12	41	36	10	6	28	11
<b>Taille moyenne des exons (pb)</b>	336	483	561	449	504	218	454	196	328

### 2.7.3.2. Annotation manuelle des gènes des contigs de BAC

Les régions des contigs de *M. rotundifolia* cv Regale similaires aux protéines connues ont été recherchées par alignement BLASTX de la banque de données SWISSPROT (alignements supérieurs à **90** % d'identité). Les structures intron-exon ont ensuite été modélisées (i) en recherchant des séquences associées aux EST (*Expressed Sequenced Tags*) dans les collections d'EST disponibles au NCBI avec le programme SIM4 (Florea *et al.* 1998) ; et (ii) en prédisant les sites d'épissage potentiels avec le programme de prédiction SPLICE MACHINE (Degroeve *et al.* 2004). L'ensemble de ces analyses a permis de valider **16** modèles de gène codant des protéines fonctionnelles, **1** gène incomplet et **6** pseudogènes (Tableau 15). Cette dernière catégorie se compose principalement de protéines avec des codons stop prématurés. Ces gènes et pseudogènes prédits ne représentent que **37** % des gènes prédits par l'annotateur automatique. Cette différence pourrait s'expliquer par le fait que je n'ai gardé que les modèles de gènes pour lesquels j'avais des confirmations par des séquences homologues/EST. Le reste correspondait à des éléments

transposables, la présence de séquences sans homologue dans les banques utilisées, de séquences ayant une composition éloignée par rapport à leurs homologues, et l'absence d'ARNm et d'EST connus pour la plupart de gènes. Puisqu'une homologie entre une protéine prédite et une protéine connue sous-entend en général une homologie probable de fonction (Dardel et Képès, 2002), les séquences protéiques directement déduites des CDS issus de l'annotation structurale ont été extraites comme point de départ à l'annotation fonctionnelle. On pourra ainsi attribuer à chacune de ces protéines une fonction putative.

### **2.7.3.2. Annotation fonctionnelle des modèles de gène**

Les séquences protéiques extraites des CDS ont été alignées sur les séquences présentes dans les bases de données SWISSPROT et INTERPRO. (Hunter *et al.* 2011) afin de les classifier suivant l'identification d'éventuelles protéines homologues, de signatures de domaine et de signaux protéiques. Chacune des **23** séquences protéiques de *M. rotundifolia* cv Regale s'aligne de façon significative ( $0 \leq E\text{-VALUE} \leq 2.10^{-6}$ ) avec au moins une protéine dans SWISSPROT (Annexe 12). Parmi ces **23** séquences, **22** ont été associées à au moins un terme *GO* (Tableau 16 ; Figure 47, Annexe 13). Une plus grande proportion de termes *GO* de la catégorie des fonctions moléculaires représente les protéines impliquées dans les interactions protéine-protéine (23 %) et dans la fixation de l'ADP (18 % ; Figure 47C). De même, les protéines impliquées dans les mécanismes de défense représentent une grande proportion des processus biologiques (31 % ; Figure 47B).

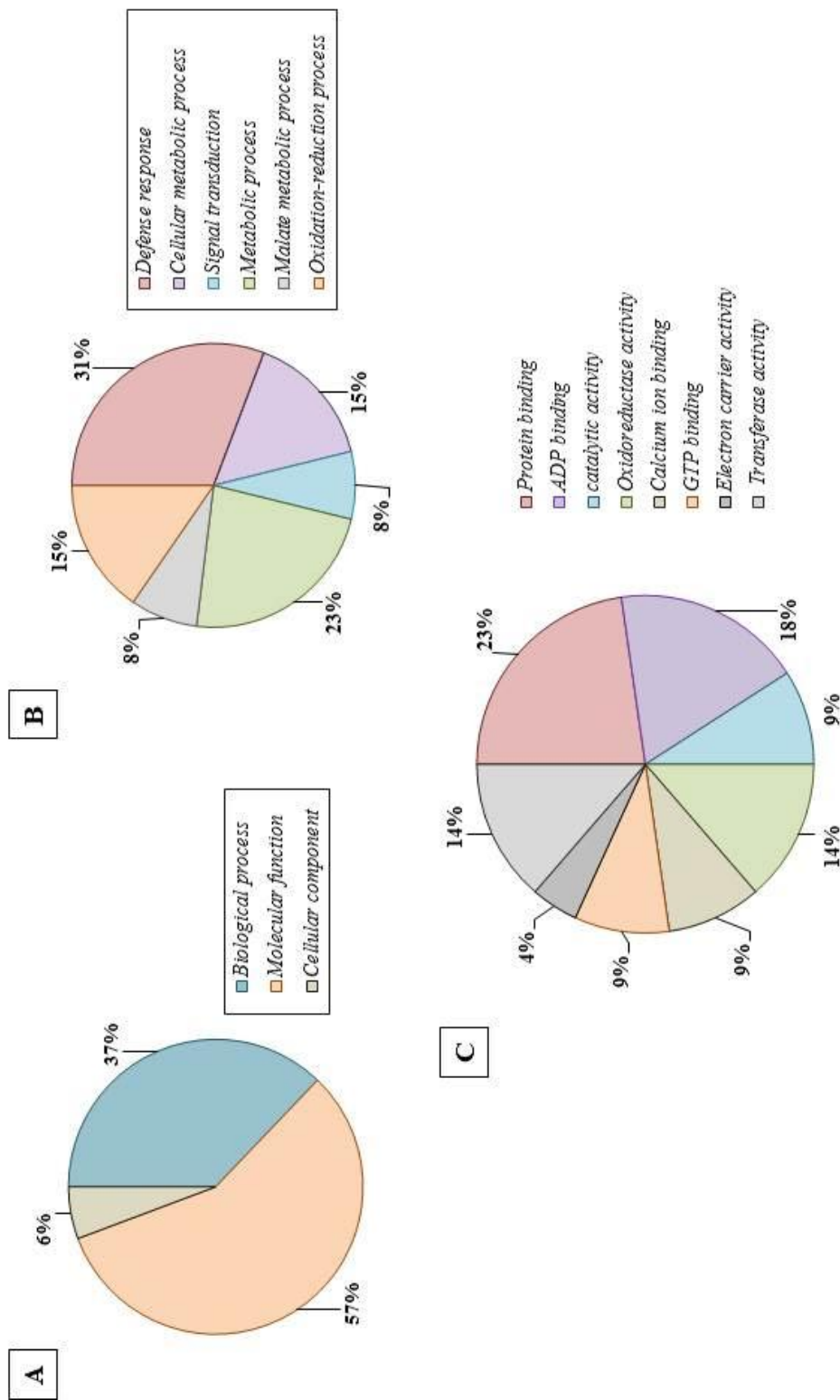
**Tableau 15. Résumé des modèles de gènes annotés dans les contigs de *M. rotundifolia* cv Regale**

Clones BAC	20I10	34D01	11K09	62H06	13C22	40F13	43H05	56A14	58B07
<b>Contigs (pb)</b>	85362	47498	65898	88284	65587	46555	25191	83582	17800
<b>FGENESH</b>									
<b>Nombre de gènes</b>	9	5	5	13	7	3	2	6	2
<b>Nombre d'exons</b>	51	21	12	41	36	10	6	28	11
<b>Moyenne des exons (pb)</b>	336	483	561	449	504	218	454	196	328
<b>SWISSPROT</b>									
<b>Nombre de hit</b>	12	10	6	8	15	7	2	11	10
<b>Moyenne des hits (pb)</b>	222	158	878	687	1059	262	1088	70	249
<b>E-VALUE maximale</b>	$5.10^{-5}$	$3.10^{-15}$	$4.10^{-5}$	$6.10^{-9}$	$5.10^{-5}$	$2.10^{-7}$	$9.10^{-40}$	$6.10^{-6}$	$7.610^{-5}$
<b>E-VALUE minimale</b>	$5.10^{-64}$	$8.10^{-37}$	$5.10^{-163}$	$1.10^{-169}$	$5.10^{-159}$	$2.10^{-56}$	$4.10^{-48}$	$2.10^{-43}$	$1.310^{-149}$
<b>EST/ADNc</b>									
<b>Nombre de hit</b>	4	4	10		12	9	2		
<b>Moyenne des hits (pb)</b>	698	338	235		420	352	424		
<b>Couverture* (%)</b>	0	0	13		0	0	52		
<b>Modèles de gènes</b>									
<b>Gènes complets</b>	0	1	2	4	1	3	2	2	1
<b>Gènes tronqués</b>	1	1	0	0	0	0	0	0	0
<b>Pseudogènes</b>	0	0	0	0	3	0	0	1	1
<b>Couverture** (%)</b>	33	40	40	31	43	67	100	33	100

\* : Pourcentage de régions homologues à des protéines connues (SWISSPROT) couvertes par des EST

\*\* : Pourcentage de gènes prédits par FGENESH et couverts par des régions homologues à des protéines connues (SWISSPROT)

En somme, les séquences protéiques identifiées chez *M. rotundifolia* cv Regale correspondent à **4** protéines de résistance de type NBS-LRR, **1** domaine TIR isolé, **2** protéines AIG1 qui sont impliquées dans la résistance aux bactéries chez *A. thaliana*, **3** protéines *F-box* qui sont généralement impliqués dans des interactions entre protéines lors de l'inhibition de l'expression des gènes *via* les miARN (Jones-Rhoades *et al.* 2006), **2** protéines de la famille des *Calcium Binding Protein*, **10** protéines des différentes classes d'enzymes et **1** protéine liée à la membrane du réticulum endoplasmique (Tableau 16). Les régions codantes identifiées en début de contig chez les clones BAC 20I10 et 34D01 sont homologues à la protéine N (TNL) de *Nicotiana glutinosa*, et elles seraient tronquées en leur extrémité N-terminale. Les séquences des clones BAC vont permettre d'examiner en détail la microsynténie entre ces contigs de *M. rotundifolia* et les régions des chromosomes 12 et 18 de *V. vinifera*.



**Figure 47. Classification des protéines en fonction de la présence de domaines fonctionnels, répétitifs, et signaux grâce aux bases de données Pfam (Finn *et al.* 2010) et InterPro (Hunter *et al.* 2011). Les protéines prédites ont été classées en fonction de processus biologiques, des fonctions moléculaires et des composants cellulaires (A). Les différents processus biologiques et fonctions moléculaires sont représentés par les camemberts B et C, respectivement.**

**Tableau 16. Annotation fonctionnelle des séquences protéiques extraites des CDS de *M. rotundifolia* cv Regale.** Les séquences ont été comparées aux bases de données PFAM (Finn *et al.* 2010) et INTERPRO (Hunter *et al.* 2011)

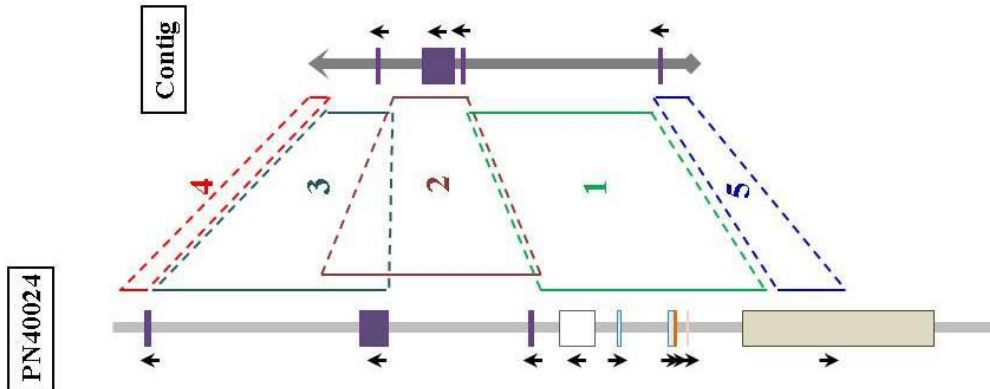
<b>PFAM FAMILY</b>		<b>GENE ONTOLOGY (GO)</b>		
<b>PFAM code</b>		<b>Biological process</b>	<b>Molecular function</b>	<b>Cellular component</b>
<b>62H06</b>	<i>Glucuronosyltransferase</i>	<i>Metabolic process</i>	<i>Transferase activity</i>	–
	<i>Glucuronosyltransferase</i>	<i>Metabolic process</i>	<i>Transferase activity</i>	–
	<i>Glucuronosyltransferase</i>	<i>Metabolic process</i>	<i>Transferase activity</i>	–
	<i>Glucuronosyltransferase</i>	<i>Metabolic process</i>	<i>Transferase activity</i>	–
<b>34D01</b>	<i>NB-ARC/LRR</i>	<i>Defense response</i>	<i>ADP binding</i>	–
	<i>Malic enzyme</i>	<i>Oxidation-reduction process</i>	<i>Oxidoreductase activity</i>	–
<b>20I10</b>	<i>NB-ARC/LRR</i>	<i>Defense response</i>	<i>ADP binding</i>	–
<b>11K09</b>	<i>F-box-like</i>	–	<i>Protein binding</i>	–
	<i>NB-ARC</i>	<i>Defense response</i>	<i>ADP binding</i>	–
<b>13C22</b>	<i>TIR</i>	<i>Signal transduction</i>	<i>Protein binding</i>	<i>Intracellular</i>
	<i>Epimerase</i>	<i>Cellular metabolic process</i>	<i>Catalytic activity</i>	–
	<i>Epimerase</i>	<i>Cellular metabolic process</i>	<i>Catalytic activity</i>	–
	<i>TIR/NB-ARC/LRR</i>	<i>Defense response</i>	<i>Protein binding/ADP binding</i>	–
<b>40F13</b>	<i>DIOX_N/2OG-FeII_Oxy</i>	<i>Oxidation-reduction process</i>	<i>Oxidoreductase activity</i>	–
	<i>EF-hand_7</i>	–	<i>Calcium ion binding</i>	–
	<i>EF-hand_1</i>	–	<i>Calcium ion binding</i>	–
<b>43H05</b>	<i>F-box/F-box associated</i>	–	<i>Protein binding</i>	–
	<i>F-box</i>	–	<i>Protein binding</i>	–
<b>56A14</b>	<i><math>\alpha/\beta</math>hydrolase_6</i>	–	–	–
	<i>AIG1</i>	–	<i>GTP binding</i>	–
	<i>AIG1</i>	–	<i>GTP binding</i>	–
<b>58B07</b>	<i>Reticulon</i>	–	–	<i>Endoplasmic reticulum</i>
	<i>Fer2/Ald_Xan_dh_C2</i>	<i>Oxidation-reduction process</i>	<i>Oxidoreductase activity</i>	–

#### 2.7.4. Analyse de la colinéarité entre les séquences de *M. rotundifolia* et *V. vinifera*

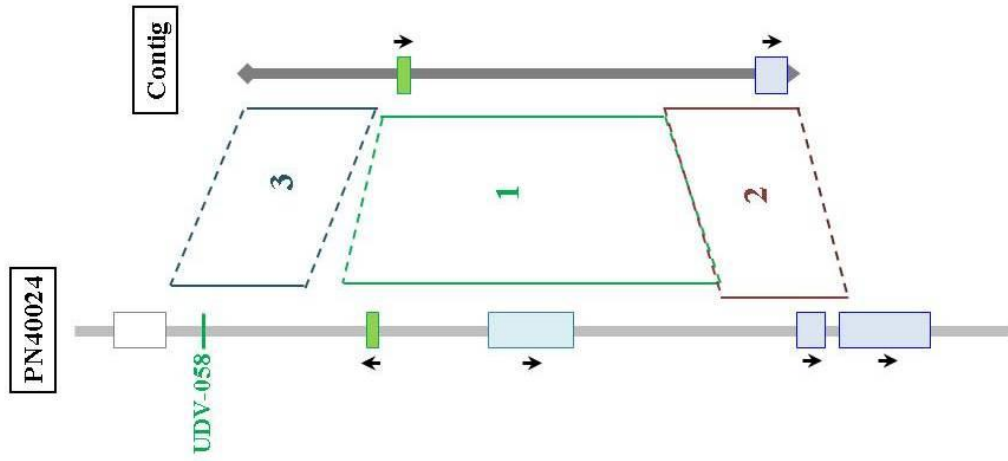
Les gènes codant des protéines, dans les régions microsynténiques entre les contigs de *M. rotundifolia* cv Regale et *V. vinifera* cv PN40024, ont été listés dans l'annexe 14. En parallèle, les séquences protéiques des modèles de gènes ont été recherchées dans le protéome de *V. vinifera* par BLASTP (Annexe 12). Les homologues correspondant de **14 (48 %)** protéines de *M. rotundifolia* cv Regale ont été identifiées sur le génome de *V. vinifera* cv PN40024, et dans les régions d'alignement des contigs. Les autres homologues ont été identifiés soit sur le même chromosome mais en dehors des régions couvertes par les contigs (**4 ; 17 %**), soit sur des chromosomes différents (**5 ; 22 %**). L'ordre des gènes de *M. rotundifolia* cv Regale et leur orientation transcriptionnelle ont été comparés avec leurs régions homologues chez de *V. vinifera* cv

PN40024. La figure 48 montre une représentation, à l'échelle, de la microcolinéarité entre 7 clones BAC de *M. rotundifolia* cv Regale et les régions correspondantes sur la séquence de référence. Aucun gène prédit chez *V. vinifera* cv PN40024 a été identifié dans la région synténique avec les clones BAC 34D01 et 20I10. En général, les régions semblent plus longues chez *V. vinifera* cv PN40024, même en tenant compte des régions manquantes dans les séquences des BACs. Dans l'ensemble, l'ordre, les orientations transcriptionnelles et les structures des gènes de *M. rotundifolia* cv Regale semblent être conservées. Par contre, les familles de gènes *AIG1* et Tyrosine kinase de *V. vinifera* cv PN40024 semble être tronqués chez les clones BAC 56A14 et 40F13 de *M. rotundifolia* cv Regale, respectivement. Malgré divers réarrangements locaux dans ces régions, y compris des inversions, des translocations, des délétions et des insertions, un bon niveau de microcolinéarité a été observé en termes de contenu en gènes. Des réarrangements manifestes ont été observés majoritairement dans les régions synténiques des clones BAC 13C22 et 20I10 qui correspondent à des régions où des gènes NBS-LRR ont été identifiées chez *M. rotundifolia* cv G52. Cette colinéarité réduite dans ces régions semble être due à la faible représentation des séquences codantes.

62H06  
Chr 12



11K09  
Chr 12



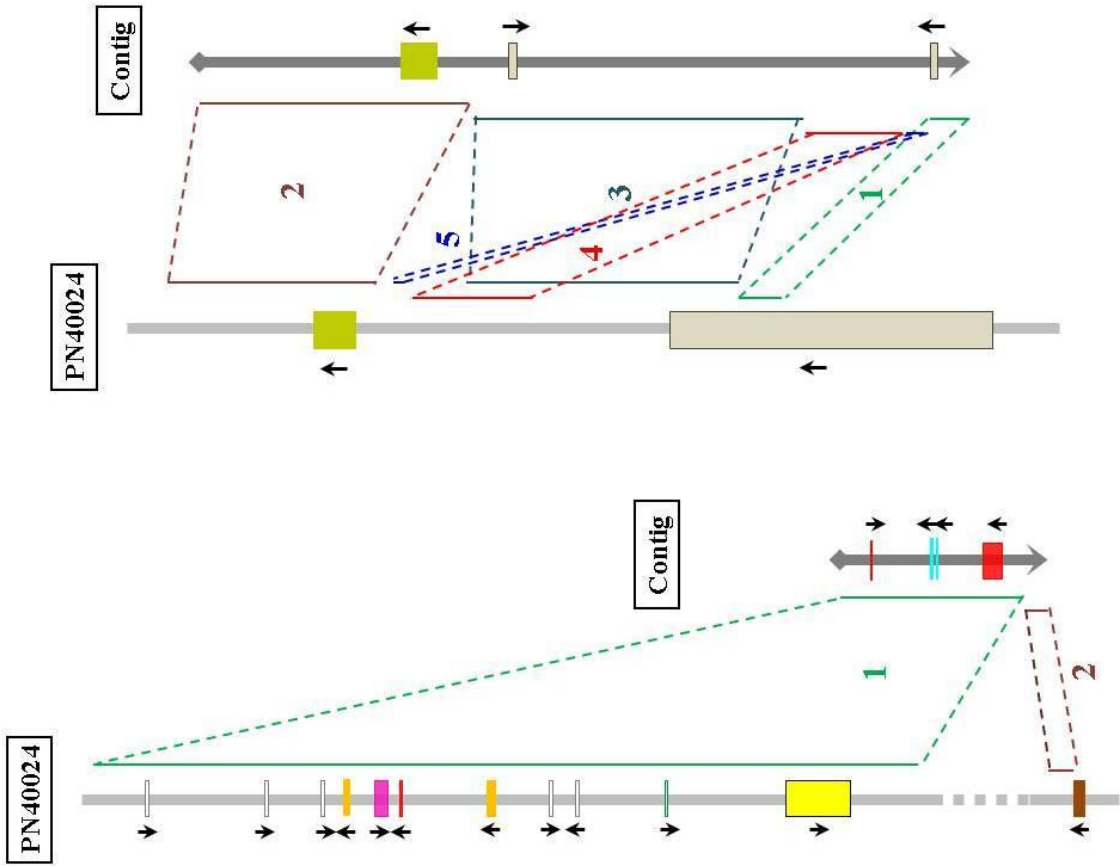
Nom des domaines PFAM identifiés dans les gènes

- *UDP-glucosyl transferase / Transferase activity*
- *FA\_desaturase\_2*
- *Methyltransf\_7*
- *DUF4589*
- *Kinase/Tyrosine kinase*
- *F-box/FBA\_1/Kelch\_1*
- *NB-ARC domain*
- *TNL / TIR domain*
- *Epimerase / Catalytic activity*
- *Cu-oxidase/Cu-oxidase\_2*
- *Smg4\_UPF3 family*
- *14-3-3 protein/Hydrolase*
- *PolyA\_pol*
- *FAR1 DNA-binding domain*
- *2OG-Fe(II) oxygenase superfamily*
- *α/β-hydrolase\_1*
- *AIG1 family*
- *Ald\_Xan\_dh\_C/Ald\_Xan\_dh\_C2*
- *Ferredoxin/[2Fe-2S] binding domain*
- *Reticulon*
- *Famille inconnue*



13C22  
Chr 18

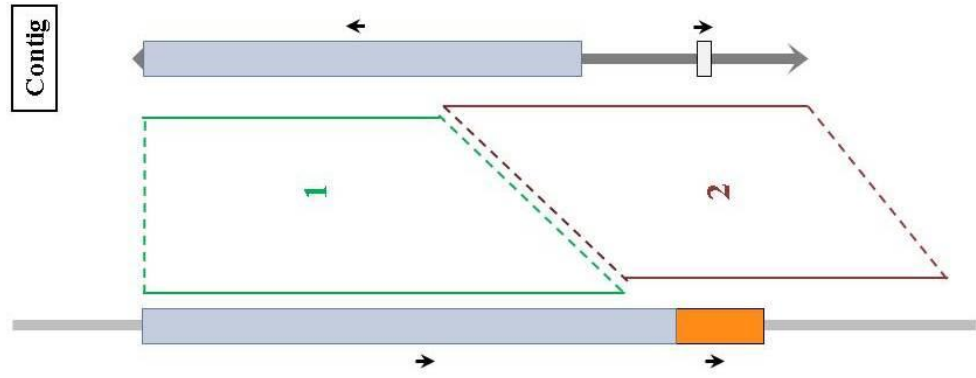
40F13  
Chr 18



**Figure 48. Microcolinéarité entre les clones BAC de *M. rotundifolia* cv Regale et les régions homologues chez *V. vinifera* cv PN40024.** Cases colorées le long de la localisation physique représentent les gènes de chaque espèce et des flèches au-dessus des cases indiquent l'orientation de la transcription. Les gènes homologues ont la même couleur. Les traits représentent les régions couvertes par chaque contig, et les chiffres désignent les numéros de contig. Les extrémités *Forward* et *Reverse* des clones BAC sont représentées par un losange et une flèche, respectivement. Les lignes en pointillées représentent des ruptures dans la contiguïté pour permettre une vue des larges régions chromosomiques. Des informations détaillées sur les régions synténiques sont données dans l'annexe 14.

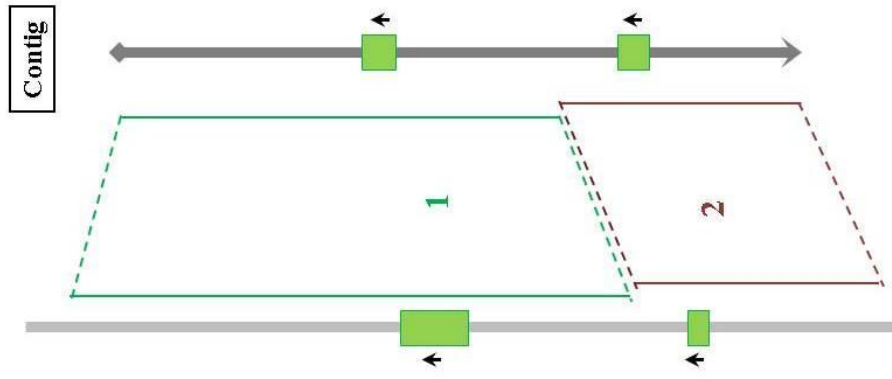
58B07  
Chr 18

PN40024



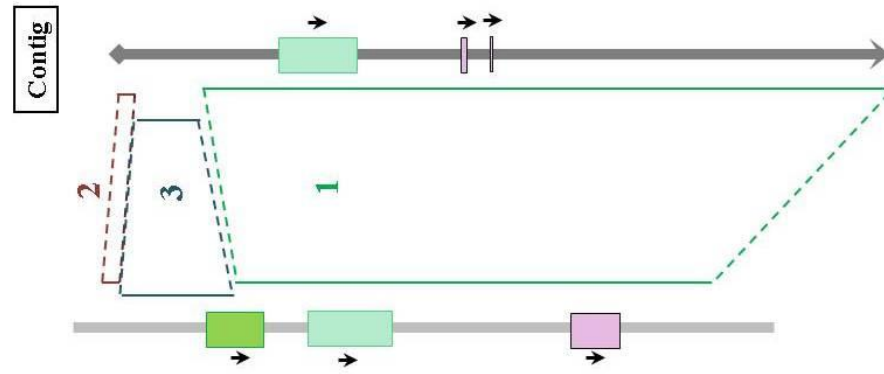
43H05  
Chr 18

PN40024



56A14  
Chr 18

PN40024



# ***3. Discussion & Conclusions***

### 3.1. Caractéristiques du génome de *M. rotundifolia* cv Regale sur la base de l'analyse de la collection de BES

Une banque BAC de *M. rotundifolia* cv Regale, représentant 7 équivalents génome et présentant un faible taux de contamination par des ADN chloroplastique et mitochondriaux, était disponible au début de ma thèse. Cette banque était de plus associée à une collection de **86 540** séquences d'extrémités de BAC (BES), d'une longueur moyenne de **682** pb (Zah-Bi *et al.* in prep). Je me suis attaché durant ma thèse à utiliser ces deux ressources pour une comparaison entre les génomes des deux espèces du genre *Vitis*, *V. vinifera* et *M. rotundifolia*. Tout d'abord, la collection de BES de *M. rotundifolia* cv Regale a permis de faire une analyse comparative des caractéristiques globales du génome en termes de pourcentages de GC, de séquences répétées et codantes, en comparaison avec le génome de *Vitis vinifera*. Pour cela, les analyses ont été effectuées simultanément avec les BES de *M. rotundifolia* cv Regale et *V. Vinifera* cv Cabernet sauvignon puis comparées au génome de référence. Les résultats obtenus dans la collection de clones BAC *M. rotundifolia* cv Regale ont toujours été comparés avec ceux des mêmes analyses menées sur une collection de BES de taille équivalente de *V. vinifera* cv Cabernet Sauvignon et sur le génome de référence de *V. vinifera*.

#### 3.1.1. Taux de Guanine-Cytosine (GC)

La teneur moyenne en GC des BES de *M. rotundifolia* cv Regale qui est de **35,84** %, est comparable à celle des BES de *V. vinifera* cv Cabernet Sauvignon (**35,5** %) au taux de GC du génome de référence (**34,5** % ; Tableau 6). Ce taux de GC des BES de *M. rotundifolia* cv Regale est bien dans la norme des taux observés dans des collections de BES d'autres dicotylédones (**35,3** à **39** %) et inférieur à ceux observés chez les monocotylédones (**44,5** à **45,9** %) (Tableau 6). Par ailleurs, la légère différence observée entre le taux de GC observé dans les deux collections de BES avec celui du génome entier de *V. vinifera* semble pouvoir être lié à un biais d'échantillonnage dans les collections de séquences d'extrémités en relation avec les différentes étapes de leur construction (digestion par les enzymes de restriction, sélection de taille, ligation...). La nature des enzymes de restriction *HindIII*, *EcoRI* et *BamHI* ne semblent pas avoir d'effet sur le taux de GC dans les BES de vigne (Tableau 17). Des résultats similaires sont d'ailleurs observés chez la tomate et la pomme de terre (Datema *et al.* 2008), et chez le ver à soie (Mita *et al.* 2004). Les collections de BES *HindIII* (**144 307** BES pour la tomate et **76 930** BES

pour la pomme de terre) et *EcoRI* (**77 141** BES pour la tomate et **51 889** chez la pomme de terre) ont des taux de GC compris entre **35 %** et **36 %**, sans différence notable d'effet entre les deux enzymes (Tableau 17). Or, comme chez la vigne dans cette étude, les taux de GC des génomes de la tomate et de la pomme de terre estimés à partir de séquences complètes de BAC, sont respectivement de **33,6 %** et **34,2 %**, soit légèrement plus bas que ceux obtenus avec les BES. Cette surestimation du taux de GC dans des BES apparait également dans une étude menée chez le ver à soie (*Bombyx mori* ; Suetsugu *et al.* 2007). En effet, deux groupes de BES issus de clonage fait dans des sites de restriction *EcoRI* (**61 696** BES) et *BamHI* (**33 208** BES) ont des taux de GC de **37,45 %** et **40,30 %**, respectivement. Par ailleurs, chez cet organisme, le taux de GC dans les régions codantes est estimé à **43,2 %** tandis que celui du génome entier du ver à soie est de **32,54 %** (Mita *et al.* 2004). Le taux de GC des BES issus de la banque BAC *BamHI* serait donc plus proche du taux observé dans les régions codantes que sur l'ensemble du génome du ver à soie. Il n'est donc pas exclu que l'étape de digestion avec les enzymes de restriction ait un rôle dans cet enrichissement en GC ; les enzymes utilisées pour la construction de ces banques ayant un pouvoir filtrant équivalent.

Dans la suite de cette analyse, le contenu en séquences répétées du génome de *M. rotundifolia* a été caractérisé de façon comparative avec *V. vinifera*, et avec la même méthode. L'intérêt porté aux éléments transposables se focalise sur deux points principaux : l'évolution des éléments transposables eux-mêmes et leur impact sur les génomes. En effet, les séquences répétées évoluent souvent beaucoup plus vite que les autres séquences et sont donc relativement spécifiques de chaque génome (Bennetzen, 2000 ; Feschotte *et al.* 2002).

**Tableau 17. Vue d'ensemble du contenu des BES de *V. vinifera* cv Cabernet Sauvignon, *M. rotundifolia* cv Regale, de la tomate (*S. lycopersicon*) et de la pomme de terre (*S. tuberosum*)**

	N° de séquences	Longueur totale (pb)	Taux de GC (%)
<b><i>Muscadinia rotundifolia</i> cv Regale</b>	<b>86 540</b>	<b>59 124 757</b>	<b>35,84</b>
Regale ( <i>BamHI</i> )	21 709	15 308 404	35,94
Regale ( <i>HindIII</i> )	64 831	43 816 353	35,74
<b><i>Vitis vinifera</i> cv Cab Sauvignon</b>	<b>77 237</b>	<b>51 832 602</b>	<b>36</b>
Cab Sauv. ( <i>EcoRI</i> )	19 736	14 660 109	35,29
Cab Sauv. ( <i>HindIII</i> )	57 501	37 172 493	36
<b><i>Solanum lycopersicon</i></b>	<b>22 1448</b>	<b>136 047 970</b>	<b>36,1</b>
Eco ( <i>EcoRI</i> )	77 141	46 398 406	35,2
HBa ( <i>HindIII</i> )	144 307	89 649 564	35,5
<b><i>Solanum tuberosum</i></b>	<b>12 8819</b>	<b>86 972 687</b>	<b>35,6</b>
PPT ( <i>EcoRI</i> )	51 889	34 276 989	35
POT ( <i>HindIII</i> )	76 930	52 695 698	36

### 3.1.2. Les séquences microsatellites

Les séquences microsatellites (SSR) sont une classe de marqueurs moléculaires qui sont souvent polymorphes et sont largement utilisés pour produire des cartes génétiques. En l'absence de la séquence du génome, les BES s'avèrent être très utiles pour concevoir des marqueurs SSR chez les plantes (Frelichowski *et al.* 2006 ; Shultz *et al.* 2007 ; Cheng *et al.* 2009). Les dinucléotides représentent la classe la plus abondante de SSR dans les BES des deux collections que j'ai analysées, les trinucleotides étant la seconde classe la plus abondante. Ces observations sont cohérentes avec la distribution des dinucléotides observée chez d'autres espèces végétales, comme *A. thaliana* (Katti *et al.* 2001), *Brassica rapa* (Hong *et al.* 2006), *Carica papaya* (Lai *et al.* 2006), *Glycine max* (Shult *et al.* 2007), *Musa acuminata* (Cheung *et al.* 2007), *Quercus robur* (Faivre Rampant *et al.* 2011) et *Coffee canephora* (Dereeper *et al.* 2013). Il existe quand même des exceptions à cette règle puisque, les dinucléotides sont moins fréquents dans le génome de *Medicago truncatula*, chez qui les hexa- et heptanucléotides sont les plus abondants (Mun *et al.* 2006).

Une proportion similaire de dinucléotides est observée dans les BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon. Parmi les motifs identifiés, les dinucléotides AT/TA ont été les plus abondants avec **62** % et **66** % des BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon, respectivement. Quant au dinucléotides AG/GA/TC/CT, ils ont été les plus abondants avec **15** % des BES de *M. rotundifolia* cv Regale et **13** % des BES de *V. vinifera* cv Cabernet Sauvignon. Les régions codantes présentent une plus grande fréquence de dinucléotides AG/CT par rapport aux AT/TA alors que ce rapport est inversé dans des régions non codantes comme par exemple les centromères, riches en séquences de microsatellites en général (Areshchenkova *et al.* 1999). Contrairement à la distribution de dinucléotides, une même proportion de motifs trinucleotides est observée dans les deux collections de BES (**28** %). Par ailleurs, ce classement peut aussi changer en fonction du compartiment du génome analysé : fraction codante ou non codante. Ainsi les trinucleotides sont classiquement plus abondants dans les séquences codantes que dans les séquences non codantes et c'est le cas chez la vigne avec **45,4** % de trinucleotides dans les EST de *V. vinifera* disponible au NCBI, contre **26,6** % de dinucléotides (Huang *et al.* 2011). Cette différence de motifs observée entre région codante (motifs AG/GA) et non codante (motifs AT/TA) a été observée chez d'autres espèces. Fait intéressant, les motifs GC étaient rares dans tous les génomes analysés (Mun *et al.* 2006 ; Faivre Rampant *et al.* 2011 ; Dereeper *et al.* 2013). La rareté de ces motifs GC a aussi été observée dans un large éventail d'espèces, de la levure aux vertébrés et chez les plantes (Toth *et al.* 2000). Les auteurs ont émis l'hypothèse que cette faible fréquence de GC dans les différents génomes pourrait être attribuée à la méthylation des cytosines, ce qui peut augmenter les taux

de mutation de la thymine. Mais cela n'explique pas la rareté des motifs GC chez *C. elegans*, la drosophile ou la levure chez qui la méthylation de la cytosine est rare (Katti *et al.* 2001). Une autre explication est que les motifs GC, dans les régions codantes, seraient sélectionnés en raison de leur stabilité dans les structures *hairpin* (Mun *et al.* 2006).

### 3.1.3. Les éléments transposables

Une proportion importante de la séquence des collections de BES correspond à des séquences d'éléments transposables : **26,82** % chez *M. rotundifolia* cv Regale et **34,36** % chez *V. vinifera* cv Cabernet Sauvignon. Ces séquences représentent **24,24** % du génome entier de *V. vinifera* cv PN40024. Les rétrotransposons à LTR (Gypsy et Copia) sont les éléments répétés de loin les plus nombreux dans le genre *Vitis* (Figure 27B ; Annexe 1), même si leur proportion par rapport aux transposons peut être un peu surestimée. En effet, selon Malik and Eickbush (1998), il est souvent difficile d'identifier correctement les rétrotransposons non-LTR sur la base de quelques copies seules. Contrairement aux rétrotransposons à LTR, les copies de rétrotransposons non-LTR sont souvent tronquées à leurs extrémités 5', et beaucoup génèrent des sites de duplication variables, voire des suppressions. Toutes les copies tronquées vont, avec le temps, accumuler des mutations éliminant leurs ORF, ce qui explique pourquoi les rétrotransposons non-LTR passent parfois inaperçues ou peuvent être confondu aux insertions de type SINE (Malik and Eickbush, 1998). De fait, une annotation expertisée des transposons a récemment permis d'avoir estimation plus juste de leur proportion dans le génome de la vigne (Benjak *et al.* 2008). Il n'en reste pas moins que cette prédominance des rétrotransposons dans les génomes des plantes est générale (Tableau 6), seule la famille de rétrotransposons dominants change entre Gypsy, comme chez *V. vinifera*, *S. lycopersicon*, *S. tuberosum*, *P. virgatum*, *Brachypodium* et *T. aestivum* ; et copia, comme chez *Quercus robur*, *Malus x domestica* Borkh, *Cucumis melo* et *Citrus clementina* (Tableau 6). Le ratio Copia/Gypsy dans les BES de *M. rotundifolia* cv Regale (**4:5**) est semblable à celui du génome de *V. vinifera* cv PN40024 (**4:5**) mais différent de *V. vinifera* cv Cabernet Sauvignon (**4:7**). L'analyse du contenu en éléments transposables des collections de BES chez la tomate et la pomme de terre a montré une tendance à la surreprésentation des rétrotransposons de type Gypsy dans les BES issus des banques BAC *HindIII* en comparaison aux banques BAC *EcoRI* (Tableau 18 ; Datema *et al.* 2008). De notre côté, nous avons observé un biais dans la représentation en rétrotransposons à LTR entre les banques BAC *HindIII*, *BamHI* et *EcoRI* (Figure 24B ; Annexe 1). Une étude similaire a été effectuée chez les mêmes espèces sur des collections de BES de pomme de terre (**87,14** Mb) et de tomate (**2,2** Mb), des séquences complètes de 18 clones BAC de pomme de terre choisis au hasard (**2,2** Mb) et de 16 clones BAC choisis dans des régions euchromatiques chez la tomate (**1,69** Mb) par Zhu *et al.* (2008). Les auteurs observent que les

rétrotransposons de type Gypsy sont plus abondants dans les génomes de la tomate (950 Mb) et de la pomme de terre (840 Mb) que dans les BES, où ils sont plus abondants que dans les séquences complètes de clones BAC (Tableau 18). En revanche, si le ratio Copia/Gypsy est comparable aussi bien dans les collections de BES, les séquences de BAC que dans la séquence du génome de pomme de terre, il est inversé dans les clones BAC de tomate par rapport aux deux autres types de séquences (Tableau 18), ce qui peut être expliqué par l'échantillonnage non aléatoire de ces BAC (dans des régions euchromatiques), contrairement à l'étude chez la pomme de terre. L'ensemble de ces résultats souligne la nécessité de construire des banques BAC basées sur différents sites de restriction pour surmonter les artéfacts dus au clonage avec l'utilisation d'une seule enzyme, tout en montrant que cela ne suffit pas à éliminer tous les biais. Le taux de séquences répétées de la collection de BES de *M. rotundifolia* cv Regale étant équivalent au taux estimé dans le génome complet de *V. vinifera* cv PN40024. Il est donc possible qu'en réalité le taux de séquences répétées du génome *M. rotundifolia* soit plus faible que celui de *V. vinifera*. De même, la proportion de Gypsy, majoritaire chez *M. rotundifolia* cv Regale comme chez *V. vinifera* cv Cabernet Sauvignon, est probablement un peu surestimée. Il faut cependant noter que la méthode d'identification des éléments transposables, par homologie de séquence, a le désavantage d'être très conservative en ne recherchant que des éléments similaires à des éléments déjà connus et qu'il est possible que des éléments spécifiques de *M. rotundifolia* aient été mal détectés.

**Tableau 18. Contenu des rétrotransposons les plus abondants des BES, des clones BAC et des génomes entiers de la tomate et de la pomme de terre (Zhu *et al.* 2008 ; Datema *et al.* 2008)**

		Copia	Gypsy	Ratio Copia/Gypsy
<b>TOMATE</b> <i>Solanum lycopersicon</i>	BES (%)	3,75	4,61	0,81
	Clones BAC (%)	3,48	0,99	3,52
	Génome (%)	6,3	19,7	0,32
<b>POMME DE TERRE</b> <i>Solanum tuberosum</i>	BES (%)	1,82	4,6	0,40
	Clones BAC (%)	1,92	3,34	0,57
	Génome (%)	3,8	15,2	0,25
<i>V. vinifera</i> cv Cab. Sauv.	BES (%)	9	15,59	0,58
<i>M. rotundifolia</i> cv Regale	BES (%)	8,38	10,01	0,84
<i>V. vinifera</i> cv PN40024	Génome (%)	6,64	8,09	0,82



### 3.1.4. Les régions codantes

Comme le contenu en séquences répétées et le taux de GC ne sont pas fondamentalement différents entre les deux sous-espèces, et en faisant l'hypothèse que la taille de leur génome est comparable (Lodhi et Reisch, 1995 ; Karkamkar *et al.* 2010), on suppose que le contenu en gènes entre *M. rotundifolia* et *V. vinifera* n'est pas différent. Au total, **49** % des BES de *M. rotundifolia* cv Regale et **50** % des BES de *V. vinifera* cv Cabernet Sauvignon ont une homologie significative avec des protéines du génome de la vigne ; et **19,5** à **44,2** % des BES de *M. rotundifolia* cv Regale sont localisées dans des régions contenant des gènes chez *V. vinifera*. Contrairement aux observations de Dereeper (2013) les enzymes de restrictions *HindIII*, *BamHI* et *EcoRI* ne semblent pas avoir d'effet sur le contenu en régions codantes dans les BES de *M. rotundifolia* et *V. vinifera*. Outre l'effet des enzymes de restriction, une autre limite serait le fait que la base de données protéique utilisée ne représenterait pas la totalité du protéome de *V. vinifera*. L'analyse comparative des gènes prédits entre les différentes versions du génome de *V. vinifera* a effectivement révélé des écarts importants entre deux versions d'annotation automatique du génome (Grimplet *et al.* 2012) et des problèmes de sensibilité de ces annotations dans les familles multigéniques organisées en tandem (Martin *et al.* 2010).

Les régions codant des NBS-LRR ont été recherchées pour voir s'il y a une différence quantitative de représentation de cette famille de gènes dans les génomes de *M. rotundifolia* et *V. vinifera*, à travers l'étude des collections de BES. Nous avons utilisé différentes approches *in silico* pour identifier **451** BES pour *M. rotundifolia* cv Regale et **332** BES pour *V. vinifera* cv Cabernet Sauvignon avec des séquences associées aux protéines TNL ou CNL. La proportion de l'ensemble des BES de *M. rotundifolia* cv Regale et *V. vinifera* cv Cabernet Sauvignon avec des séquences codant des NBS-LRR est de **0,5** % et **0,4** % de la collection de BES, ce qui est dans la norme des estimations pour d'autres espèces de plantes qui se situent entre 0,6-1,8 % (Mun *et al.* 2009). Il n'y a donc pas d'expansion ou rétraction importante de cette famille de gène chez *Muscadinia*. Bien que notre analyse utilise des motifs MEME spécifiques d'*A. thaliana* et de *P. trichocarpa* (Annexe 2), elle a permis de distinguer les sous-classes de NBS-LRR. L'analyse phylogénétique, qui a été effectuée sur le domaine conservé NB-ARC, a soutenu cette distinction (Annexe 4) et est compatible avec les observations antérieures pour d'autres espèces de plantes (Meyers *et al.* 1999 ; Meyers *et al.* 2003 ; McHale *et al.* 2006 ; Rairdan *et al.* 2008 ; Yang *et al.* 2006 et 2008 ; Mun *et al.* 2009 ; Tarr DE *et al.* 2009, Guo *et al.* 2011 ; Collier *et al.* 2011). Ainsi, les BES de *M. rotundifolia* cv Regale possèdent **4,7** fois plus de CNL que de TNL. Une proportion similaire a été trouvée chez *V. vinifera* (**3,8** fois plus de CNL que de TNL) et il n'y a donc pas non plus de différence qualitative pour cette famille entre les deux sous espèces. Ce ratio est plus faible chez le peuplier (**1,7** fois

plus de CNL que de TNL ; Yang *et al.* 2008 ; Malcarne *et al.* 2012), chez les *Brassicaceae* *A. thaliana*, *A. lyrata* et *B. rapa* (0,5 ; Meyers *et al.* 2003 ; Mun *et al.* 2009 ; Guo *et al.* 2011). Leister *et al.* (2006) a suggéré que ces différences de proportion entre TNL et CNL pourraient être le reflet de l'adaptation des espèces face à leurs principaux agents pathogènes. En effet, la surreprésentation des CNL chez la pomme de terre peut être assimilée à une réponse à certains des agents pathogènes les plus néfastes tels que *P. infestans*, qui est généralement contrôlé par les CNL (Ballvora *et al.* 2002 ; Huang *et al.* 2005 ; Van *et al.* 2005 ; Wang *et al.* 2008).

Nos résultats ont révélé la présence des sept motifs précédemment identifiés (P-loop, Kinase-2, RNBS-B, RNBS-C, GLPL, MHDV et RNBS-D) chez *A. thaliana* (Meyers *et al.* 2003) et *P. trichocarpa* (Kolher *et al.* 2008) dans la plupart des gènes. Toutefois, certains motifs MEME n'ont pas toujours été détectés, suggérant soit un biais d'échantillonnage dans les BES, soit que ces motifs ont évolué chez les *Vitaceae*. La deuxième hypothèse est la plus probable car ces motifs concernent essentiellement la famille des CNL et il a été montré que cette famille est très diversifiée et datent d'avant la scission entre les gymnospermes et angiospermes (Cannon *et al.* 2002 ; Bai *et al.* 2002 ; Tarr *et al.* 2009). En revanche, il en découle que la famille des TNL est plus homogène et chez les dicotylédones, ce qui suggère qu'elle a surgi après la divergence des monocotylédones et trouvée uniquement chez les dicotylédones. Par ailleurs, les analyses phylogénétiques que nous avons réalisées avec les séquences identifiées dans les BES de muscadine, regroupent les TNL en deux clades composés des TNL canoniques et du seul gène RPS4 d'*A. thaliana*. La branche des CNL forme également deux clades CNL1 et CNL2 qui se subdivisent en 2 et 3 sous-groupes, respectivement. Le sous-groupe CNL2.2 est soutenu par une valeur de *bootstrap* élevé et de courtes longueurs de branches, ce qui suggère un ancêtre commun récent entre les séquences qui le composent. Par ailleurs, aucun des gènes CNL de fonction connue inclus dans l'arbre n'est trouvé dans le sous-groupe CNL1.2, ce qui suggère un clade peut-être plus spécifique ou à fonction originale par rapport aux gènes de référence utilisés.

### **3.2. Génomique comparative entre *M. rotundifolia* et *V. vinifera***

Mon travail de thèse apporte une vision complémentaire sur la divergence entre les génomes de *V. vinifera* et *M. rotundifolia* et son impact sur des études de génomique comparative à différents niveaux.

### 3.2.1. Transférabilité des marqueurs entre les deux génomes

Ainsi, les résultats obtenus en termes de développement de marqueurs à partir des séquences d'extrémité de BAC donnent des indications sur le degré d'évolution entre les séquences des génomes des deux sous-genres mais également sur les limites techniques auxquelles nous sommes confrontés dans ce type d'étude. Ainsi, dans la région du *locus* de résistance *Run1/Rpv1*, les marqueurs physiques CB46.49 et CB13.14 cartographiés chez *M. rotundifolia* cv G52 (Barker *et al.* 2005) n'ont pas été localisés *in silico* sur la séquence de *V. vinifera* mais ils ont amplifié *in vitro* le génome de *V. vinifera*. Cette observation met en lumière des zones manquantes dans l'assemblage du génome de la vigne.

Sur l'ensemble des **132** marqueurs définis dans les BES de *M. rotundifolia* cv Regale et testés pour l'amplification du génome de *V. vinifera* cv PN40024, **117** (88 %) et **110** (83 %) ont montré des amplifications uniques *in silico* et *in vitro*, respectivement (Annexe 9). Dans la plupart des cas, les tailles des fragments amplifiés chez *V. vinifera* et *M. rotundifolia* étaient identiques. Nous avons cependant observé plusieurs cas d'amplification d'un nombre de bandes différentes entre *V. vinifera* cv PN40024 et *M. rotundifolia* cv Regale, faisant penser à des variations de nombre de séquences. Ces variations ont également été observées lors de l'établissement de la carte génétique de *M. rotundifolia* cv Regale (Blanc *et al.* 2013). En effet, **451** marqueurs de type microsatellite, initialement conçus sur *V. vinifera*, ont été utilisés pour amplifier l'ADN génomique de *M. rotundifolia* cv Regale. Sur l'ensemble de ces marqueurs, **77** (17 %) étaient absents du génome de *M. rotundifolia* cv Regale, et **197** (43,6 %) marqueurs polymorphes chez *V. vinifera* ne l'étaient pas chez *M. rotundifolia* cv Regale. La proportion des marqueurs absents comparable à celle que nous avons observée dans notre étude. Dans une étude menée chez les *Eucalyptus*, le transfert de l'amplification et du polymorphisme des quelques *loci* microsatellites testés était de 100 % en restant dans la même section, et seulement de 50 % ou de 0 % en passant à des espèces d'autres sections (Brondani *et al.* 2002), résultats confirmés dans une étude dans le genre *Medicago* (Eujayl *et al.* 2004). Dans d'autres études, Rossetto (2001) a résumé des informations concernant la transférabilité des microsatellites de plantes et rapporté un succès d'en moyenne 35 %. Les comparaisons de séquences à différents *loci* ont montré que des réarrangements locaux, y compris des délétions, insertions, duplications et des transferts ont eu lieu au niveau des *loci* de génomes proches et qui semblaient colinéaires à la cartographie génétique (Feuillet *et al.* 2002).

Dans notre étude, outre la présence/absence de marqueurs, deux inversions ont été observées dans l'ordre des marqueurs entre la carte génétique de référence de *V. vinifera* et celle de *M. rotundifolia*. L'une des régions concernée par cette inversion est celle contenant le *cluster* de gènes de résistance *Run1/Rpv1* du chromosome 12 de *M. rotundifolia* cv G52. Ces résultats suggèrent des réarrangements locaux ou une

évolution de cette région. Malgré ces variations de séquence, l'ancrage des marqueurs définis dans les extrémités des BAC alignés dans les deux régions a permis d'étoffer un peu la construction de contigs de BAC. Dans l'ensemble, une bonne conservation globale des marqueurs et de leur ordre a été observée dans la région des gènes *Run1/Rpv1* avec toutefois la translocation de petites séquences (Figure 37 ; Annexe 15). Dans la région du gène *Rpv3*, un réarrangement plus important est déjà observé dans notre carte physique grossière (Figure 40 ; annexe 15).

### 3.2.2. Recherche de zones de synténie entre *M. rotundifolia* et *V. vinifera* par alignement des paires de BES

Une autre façon de regarder l'impact des variations de séquences entre les génomes des deux sous-genres sur des études de génomique comparative est de regarder la couverture du génome de *V. vinifera* par des contigs de BAC construits *in silico* par alignement des paires de BES sur le génome. Avec les critères d'alignement que nous avons posés, cette couverture va être impactée :

- Par des variations locales de séquence qui feront passer l'alignement de séquence en deçà des seuils d'identité (92% d'identité) et de longueur d'alignement
- Par l'existence de séquences répétées conduisant à des alignements multiples des BES
- Par des réarrangements structuraux plus importants qui éloigneront ou réorienteront les deux membres d'une paire de séquence

Il faut noter que la qualité variable de l'assemblage du génome peut également affecter la couverture globale de ces contigs.

Les BES de *M. rotundifolia* cv Regale ont donc été alignées sur la séquence du génome de *V. vinifera* cv PN40024 et suivant la stringence des critères d'alignement (Cf. Matériels et Méthodes), **14,6 %**, **12,9 %** et **8,6 %** de paires BES ont pu être placée de façon unique sur la séquence du génome de *V. vinifera*. Ces valeurs indiquent une situation favorable aux études de génomique comparative si on les compare aux résultats d'autres études similaires. Par exemple, dans la famille des *Poaceae*, les BES de *Saccharum spontaneum* (Kim *et al.* 2012) et de *Panicum virgatum* (Sharma *et al.* 2012) ont été alignées sur le génome de *Sorghum bicolor* (Paterson *et al.* 2009) qui a été utilisé comme génome de référence. Un seuil **70 %** d'identité sur toute la longueur du BES et un score d'alignement supérieur à 100 ont été considérés. Cette cartographie comparée a donné un pourcentage plus élevé (**8 %**) avec l'alignement des BES de *P. virgatum* qu'avec ceux de *Saccharum* ssp. (**6,4 %**). L'étude faite avec les BES de *P. virgatum* ne considérait pas nécessairement uniquement les paires de BES contrairement à celle faite avec les BES de

*Saccharum spontaneum*. Les résultats obtenus entre *M. rotundifolia* cv Regale et *V. vinifera* cv PN40024 (8,6 à 14,6 % de la collection de BES) sont comparables, voire un peu meilleurs si on tient compte des seuils choisis pour rejeter un alignement.

Ces résultats peuvent être influencés, soit par l'évolution locale des séquences telles que des mutations, des insertions/délétions d'éléments transposables, des duplications de gènes ; soit par des remaniements structuraux plus importants dans les génomes telles que des duplications segmentales, des translocations, des fusions/fissions de chromosomes, des inversions. Ces variations sont *a priori* d'autant plus importantes entre espèces qu'elles sont éloignées phylogénétiquement. Ainsi, une autre étude, visant à analyser la synténie entre espèces apparentées et éloignées, a analysé les alignements de 3 744 BES de *Malus × domestica* Borkh sur les séquences les génomes de *P. trichocarpa*, *Arabidopsis thaliana* et *O. sativa* (Han, Korban, 2008). Il en résulte que 154, 107 et 106 paires de BES ont été identifiées comme étant homologues à *P. trichocarpa*, *A. thaliana*, et *O. sativa*, respectivement, dans une région de 10 - 300 Kb. Parmi ces paires de BES, 8 (5,2 %), 3 (2,8 %) et 0 paires ont été cartographiées sur les génomes de *Populus*, *Arabidopsis* et *O. sativa*, respectivement, en cohérence totale avec leur degré d'éloignement phylogénétique avec le pommier. Cependant, Dereeper *et al.* (2013) ont montré que les BES de *C. canephora* et le génome de *V. vinifera* partageaient plus régions microsyténiques (331) qu'avec les génomes de la tomate (270) et d'*Arabidopsis* (88). Or, les génomes de *C. canephora* et de la tomate sont plus proches phylogénétiquement. Ceci souligne que le taux de divergence entre les séquences des génomes de deux espèces n'est pas linéairement corrélée au temps de séparation entre ces deux espèces mais dépend également de la biologie de l'espèce et notamment de son temps de génération et de l'histoire des génomes (Cenci *et al.* 2013). Ainsi *Canephora* et *Vitis vinifera* sont toutes deux des espèces pérennes et *A. thaliana* cumule un temps de génération court et une longue histoire de duplication segmentaire et de remaniement de son génome (Blanc *et al.* 2000 ; Mahé *et al.* 2007). Enfin, elles peuvent être fortement affectée par la qualité des assemblages des génomes de référence et dans notre étude, nous observons effectivement une mauvaise couverture en BAC positionnés *in silico*, dans les régions morcelées et dont l'assemblage chromosomique est aléatoire, représenté par les chromosomes dit random dans le génome de référence de la vigne (Figure 36 ; Annexe 15).

### 3.2.3. La famille des gènes de résistance de type NBS-LRR

La famille des gènes NBS-LRR est connue pour être organisée en *clusters* de gènes qui évoluent très vite, aussi bien au niveau de la séquence des gènes que de la structure de la région les portant, rendant difficile des approches de génomique comparative (Yue *et al.* 2011). Nous avons voulu avoir une première

évaluation de ce problème avec deux approches : (i) en regardant la couverture des *clusters* de gènes NBS-LRR de *Vitis vinifera* par les contigs de BAC définis par alignement de leurs BES sur le génome ; (ii) en regardant de façon plus fine où se trouvent les orthologues des gènes présents dans les séquences de BES. La première approche montre que cette famille est moins bien couverte par rapport à l'ensemble des gènes (quasiment deux fois moins dans les alignements les plus stringents ; Zah-BI *et al.* Tableau 8). Pour la deuxième approche, un ensemble de **236** BES de *M. rotundifolia* cv Regale montrant une homologie de séquence avec le domaine NB-ARC et uniques sur la séquence de référence de *V. vinifera* cv PN40024 ont été analysées. Dans une étude antérieure, Yang *et al.* (2008) ont rapporté que chez *V. vinifera*, la majorité (**66,7** %) des gènes de résistance NBS-LRR cartographiés sont regroupés sur les chromosomes 9, 12, 13, et 18 avec 49, 48, 58 et 80 gènes, respectivement (Tableau 19). Les TNL (**80,4** %) et les TN (**71,4** %) étaient localisés sur les chromosomes 5, 12 et 18. Les résultats d'alignement des séquences de *M. rotundifolia* cv Regale ont montré que les BES contenant les TNL s'alignent en majorité sur le chromosome 18 (Figure 34) qui contient bien un *cluster* important de TNL chez *Vitis vinifera* (Yang *et al.* 2008).

Globalement, les nombres et qualités des BES de *M. rotundifolia* cv Regale à motif NB-ARC s'alignant sur le génome de *V. vinifera* cv PN40024 sont cohérents avec la distribution de cette famille sur le génome de référence de la vigne (Figure 34). Par ailleurs, les coordonnées d'alignement des BES de *M. rotundifolia* cv Regale ne chevauchent pas toujours les coordonnées de modèles de gènes proposés par l'annotation automatique du génome de référence de *V. vinifera* cv PN40024 (Annexe 5). A quelques exceptions près, les répartitions chromosomiques des gènes à domaines NB-ARC proposées par Yang *et al.* (2008) sont similaires à celle de la dernière version de l'annotation du génome de *V. vinifera*. En effet, des différences d'annotations sont observées sur les chromosomes 2, 7, 14, 15 et 19 pour lesquels 11 à 25 gènes à domaine NB-ARC ont été détectés en plus sur la dernière version de l'annotation, et sur le chromosome 18 pour lequel 43 de ces gènes n'ont pu être identifiés par la dernière version de l'annotation (Tableau 19). Ces résultats illustrent tout d'abord le besoin d'une annotation expertisée de cette famille de gènes chez *V. vinifera*. En effet, Martin *et al.* (2010) ont montré que les programmes utilisés pour l'annotation automatique des gènes codant des Terpènes Synthases chez *V. vinifera* cv PN40024 n'ont pas détecté une grande partie de ces gènes (Howe *et al.* 2002 ; Jaillon *et al.* 2007). En parallèle, l'annotation expertisée des gènes de résistance à motif NB-ARC en cours à l'URGV a montré que plus de la moitié des gènes de cette famille n'ont pu être détectés par l'annotation automatique (Cécile Guichard et Nadia Bentahar, communication personnelle).

**Tableau 19. Distribution des gènes à motifs NB-ARC du génome de *V. vinifera* cv PN40024 d'après Yang *et al.* 2008, la dernière version de l'annotation (V1) du génome de *V. vinifera* cv PN40024 et des BES de *M. rotundifolia* cv Regale.**

Chr	Génome de <i>V. vinifera</i> (Yang <i>et al.</i> 2008)				BES de <i>M. rotundifolia</i> cv Regale				Génome de <i>V. vinifera</i> (V1)
	TNL/TN	CNL/CN	XNL/XN	Total/Chr	TNL	CNL	XNL/XN	Total/Chr	PF00931
1	2	6	2	10	0	0	6	6	17
2	0	0	1	1	0	0	3	3	12
3	0	6	8	14	0	8	4	12	18
4	0	1	0	1	0	0	0	0	2
5	12	2	11	25	1	11	6	18	17
6	2	4	7	13	0	2	0	2	8
7	0	0	0	0	0	5	0	5	25
8	0	1	1	2	0	2	1	3	3
9	0	35	14	49	0	32	4	36	53
10	0	0	0	0	0	1	2	3	8
11	0	4	2	6	0	1	0	1	6
12	14	14	20	48	4	16	3	23	39
13	1	30	27	58	1	29	4	34	69
14	0	5	2	7	0	5	14	19	29
15	0	8	5	13	1	4	4	9	25
16	0	0	2	2	0	1	0	1	6
17	0	3	2	5	0	0	1	1	9
18	62	4	14	80	21	0	14	35	37
19	0	12	7	19	0	19	4	23	32
ChrUn	18	93	70	181	5	2	3	10	29
<b>Total</b>	<b>111</b>	<b>228</b>	<b>195</b>	<b>534</b>	<b>33</b>	<b>138</b>	<b>73</b>	<b>244</b>	<b>444</b>

Outre les limites de l'annotation automatique, les différences de nombre de gènes observées pourrait être le fait de l'évolution structurale des *clusters* de gènes NBS-LRR. Cela a été démontré pour la première fois chez la vigne récemment (Feechan et 2013) et depuis longtemps chez un grand nombre d'espèces comme par exemple au *locus* de résistance  $S_{H3}$  dans les génomes de *C. eugenoides*, *C. Arabica* et *C. canephora* avec la présence d'un nombre variable de gènes NBS-LRR appartenant à la sous-classe des CNL (Ribas *et al.* 2011) comme dans l'étude de Feechan et al (2013). En parallèle, les auteurs ont comparé le *locus* de résistance  $S_{H3}$  à la région orthologue dans le génome de *S. lycopersicum* qui est, à ce jour, l'espèce la plus proche de *C. Arabica* pour laquelle la séquence du génome entier est disponible. Une microsyténie a été trouvée entre le *locus* de résistance  $S_{H3}$  et deux régions du génome de *S. lycopersicum* qui partageaient **53,2** et **23,4** % des gènes de *C. Arabica*, respectivement, mais aucun gène CNL n'a été trouvé dans ces régions du génome chez *S. lycopersicum*. Par conséquent, les auteurs ont supposé que les CNL du *locus* de résistance  $S_{H3}$  y ont été insérés après la divergence entre *S. lycopersicum* et *C. Arabica*. Dans une étude similaire menée par David *et al.* (2009), les auteurs ont comparé le *locus* de résistance  $B4$  de *P. vulgaris* (qui a 26 gènes CNL) avec trois génomes séquencés de légumineuses, *M. truncatula*, *L. japonicus* et *Glycine max*. Leur analyse a révélé la conservation d'une microsyténie entre les espèces de légumineuses, à l'exception des CNL, qui semblait être complètement absents dans les régions correspondantes de *M. truncatula* et *L. japonicus*, et seulement un gène CNL tronqué dans les régions synténiques de *Glycine max* (David *et al.* 2009). Ils ont suggéré que les CNL ont été insérés dans l'ancêtre du *locus* de résistance  $B4$  après la divergence de *M. truncatula*, *L. japonicus* et *P. vulgaris*, mais avant la divergence de *P. vulgaris* et *Glycine max* à travers un événement de recombinaison ectopique entre chromosomes non homologues. L'analyse phylogénétique incluant les séquences des CNL de légumineuses et toutes les CNL du *locus* de résistance  $B4$  connus ont soutenu cette hypothèse (David *et al.* 2009).

Plusieurs études ont comparé des *clusters* de gènes de résistance dans différents haplotypes ou espèces, et il en résulte que l'évolution de ces gènes est un processus dynamique impliquant principalement des duplications, des délétions, des échanges de séquences, des mutations ponctuelles, des recombinaisons, des conversions géniques diversifiés et des insertions d'éléments transposables (Noel, 1999 ; Parniske *et al.* 1999 ; Kruijt *et al.* 2004 ; Xiao *et al.* 2004 ; Kuang *et al.* 2005 ; Seah *et al.* 2007 ; Innes *et al.* 2008 ; Kuang *et al.* 2008 ; David *et al.* 2009 ; Dai *et al.* 2010). Ainsi, les *clusters* représentent un important réservoir de diversité et une source de variation génétique permettant la génération de nouvelles spécificités de résistance (Michelmore *et al.* 1998 ; Hulbert *et al.* 2001). Dans cette première partie de notre étude, l'évolution structurale des *clusters* de gènes NBS-LRR n'est pas détectable, car à ce stade nous n'avons réalisé aucune cartographie des séquences analysées. Nous n'avons aucune indication sur la



localisation réelle sur le génome de *M. rotundifolia* des homologues des gènes de *V. vinifera* identifiés par alignement de séquences. La seule indication structurale est que le pourcentage de paires de BES dont les alignements sont conservés dans les régions portant des NBS-LRR est plus faible que la moyenne (Annexe 15). Or nos critères d'alignement favorisent les régions non-arrangées entre les deux génomes. Pour aller plus loin, nous avons commencé une démarche de cartographie physique comparée dans deux régions portant des *clusters* de NBS LRR, sur les chromosomes 12 et 18.

### 3.2.4. Etude de la synténie entre *M. rotundifolia* et *V. vinifera* dans la région de *clusters* de gènes NBS-LRR

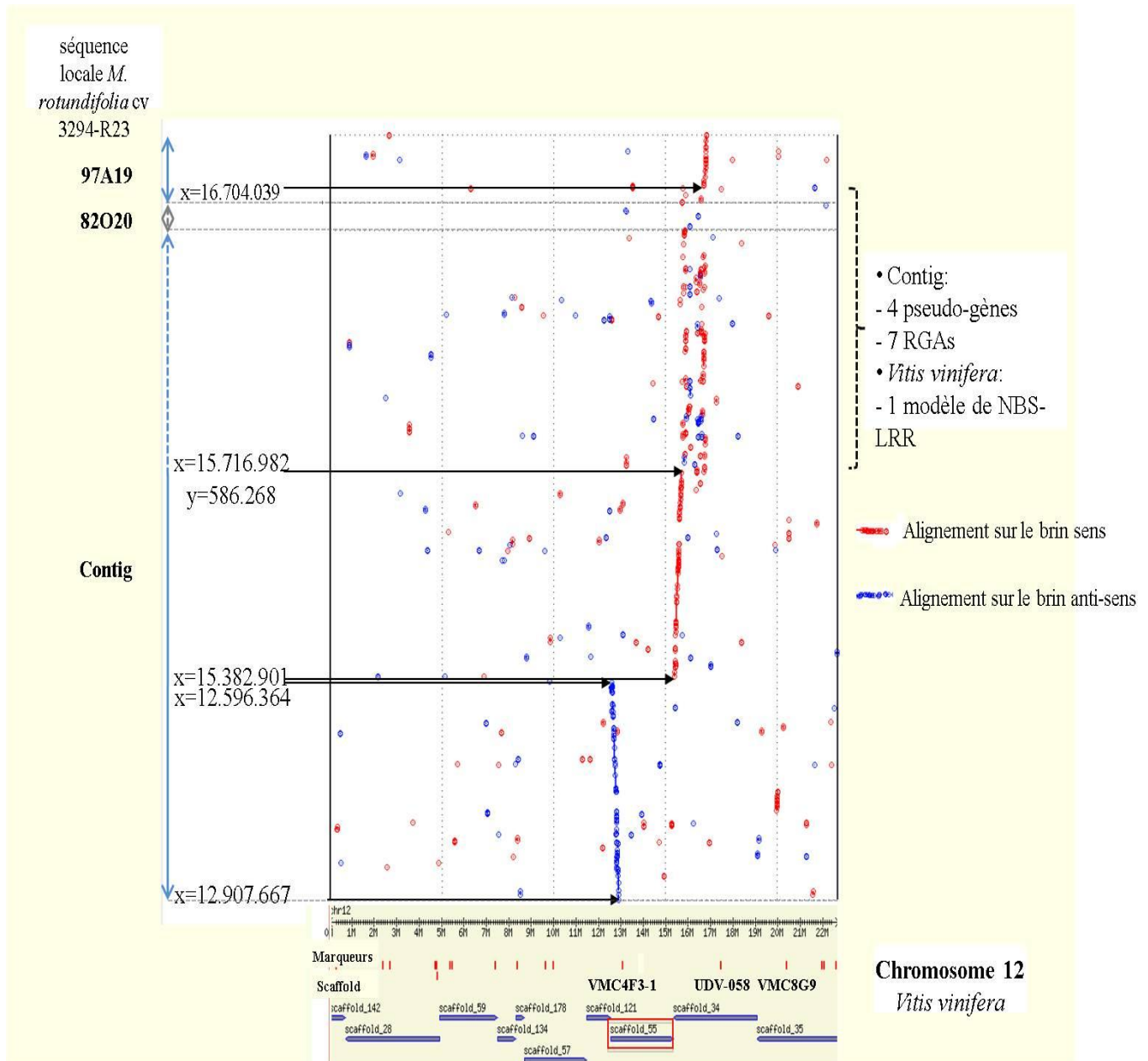
Les informations disponibles dans les deux régions choisies étaient de niveau assez différent :

- Disponibilité d'une séquence d'origine *M. rotundifolia* cv G52 portant les deux gènes de résistance *Run1* et *Rpv1* qui ont été fonctionnellement caractérisés dans le courant de ma thèse (Feechan et al 2013) sur le chromosome 12. Une rupture de synténie dans la région du *cluster* de gènes NBS-LRR était donc connue dès le départ (Figure 49) grâce à un alignement de la séquence portant les gènes de résistance sur la séquence du chromosome 12 de *V. vinifera*. Une inversion du scaffold 55 du génome du PN40024 par rapport à la séquence de *M. rotundifolia* cv G52 est ainsi observée et peut être liée à une erreur d'orientation de ce contig dans l'assemblage chromosomique de la séquence de référence de la vigne. Une autre petite inversion au centre de cette zone correspond en revanche sans doute à une différence réelle entre les deux espèces. La comparaison de ces séquences révèle aussi une rupture importante de l'alignement correspondant à la région du *locus Run1/Rpv1* où le *cluster* de gène NBS-LRR a été identifié.
- Un QTL de résistance *Rpv3* bien cartographié sur le chromosome 18 (Bellin et al 2009) sans information sur la séquence d'un haplotype portant la résistance.

#### 3.2.4.1. Ebauche de carte physique dans les régions des gènes *Run1/Rpv1* et *Rpv3*.

Sur l'ensemble des BES de *M. rotundifolia* cv Regale qui s'alignent de façon unique sur le génome de *V. vinifera* cv PN40024, **164** et **465** paires ont été cartographiées sur les chromosomes 12 et 18, respectivement. Parmi ces paires, **30** paires (**18** % des BES alignées sur le chromosome 12) et **30** paires (**6** % des BES alignées sur le chromosome 18) colocalisent dans un intervalle de **7,3** Mb et **15,8** Mb de la région des *loci* de résistance *Run1/Rpv1* et *Rpv3*, respectivement. Par ailleurs, **6** paires de BES de *M. rotundifolia* cv Regale étaient alignées sur le supercontig de *M. rotundifolia* cv G52 (**918** Kb). Dans le cas

du locus de résistance *Run1/Rpv1*, les paires de BES qui s'alignent sur le génome de *V. vinifera* sont différentes de celles qui s'alignent sur la région correspondante du supercontig. Par ailleurs, sur le chromosome 18 comme sur le 12, ces paires sont généralement en bordure des régions contenant des gènes NBS-LRR, illustrant la rupture de synténie observée dans ces *clusters* et illustrée par la figure 49.



**Figure 49.** Alignement des séquences des clones BAC identifiés (axe des ordonnées) sur le locus de résistance à l'oïdium (Barker *et al.*, 2005) sur le groupe de liaison 12 (axe des abscisse) de la séquence de référence de *Vitis vinifera*, à l'aide du système d'alignement MUMmer. Les marqueurs génétiques sont représentés par des traits rouges et les scaffolds par des barres bleues.

Finalement, par cette approche *in silico*, seuls trois BAC couvrant potentiellement une région comprenant des NBS-LRR ont été identifiés 13C22 (chr18), 11K09 et 66F06 (chr12). Un quatrième a été rajouté grâce à des expériences d'ancrage PCR de marqueurs (34D01, chr12). Deux d'entre eux ont été séquencés (13C22 et le 11K09) par la suite et présentent effectivement des séquences NBS-LRR.

#### 3.2.4.2. Colinéarité de séquence dans la région des *loci* de résistance *Run1/Rpv1* et *Rpv3*

Un set de **9** clones BAC de *M. rotundifolia* cv Regale a en effet été choisi sur différents critères et séquencé. Les **29** contigs de séquences de BAC représentent au total **286** Kb et **238** Kb non continus des régions des *loci* de résistance *Run1/Rpv1* et *Rpv3*, respectivement. Les alignements de ces séquences de *M. rotundifolia* cv Regale sur la séquence du génome de *V. vinifera* et le supercontig de *M. rotundifolia* cv G52 ont permis de mettre en évidence des réarrangements plus ou moins importants dans ces régions des chromosomes 12 et 18 (Figure 48 ; Tableau 13 ; Annexe 10). Les contigs de séquences de l'un des BAC (40F13), ancré par un marqueur du chromosome 18 (UDV-305+5) s'alignent en réalité sur le chromosome 8, démontrant un réarrangement de ce marqueur. Les tentatives d'assemblage des différents contigs d'un même clone BAC ont révélé que l'existence de séquences répétées ou de moindre complexité serait à l'origine des *gaps* observés. En effet, les séquences manquantes entre les contigs ont été amplifiées par PCR puis séquencées pour certains clones BAC (Figure 46). Sur la totalité des fragments amplifiés, représentant environ **2** Kb de séquence, **1,2** Kb (54 %) correspondent essentiellement à des séquences répétées annotées par RepeatMasker sur le génome de référence de *V. vinifera* (Tableau 20). Cette proportion de séquences répétées est comparable à celle observée par Carrier *et al.* (2012) dans l'analyse du polymorphisme au sein du sous-genre *Euvitis*. En effet, pour identifier les polymorphismes impliqués dans les mutations somatiques chez *V. vinifera*, quatre accessions de Pinot Noir ont été sélectionnées pour leurs phénotypes distincts (Boursiquot *et al.* 2007), puis séquencées. La comparaison des séquences obtenues avec la séquence de référence de *V. vinifera* a montré que **25** % des séquences étaient *multiloci*. Les résultats de cette étude ont montré que les polymorphismes somatiques étaient principalement dus à des insertions d'éléments transposables (**35,2** par Mb) (Carrier *et al.* 2012).

**Tableau 20. Alignement des fragments amplifiés par PCR avec les couples d'amorces définis aux extrémités des clones BAC de *M. rotundifolia* cv G52.** Les produits d'amplifications ont été séquencés par la méthode de Sanger puis alignés par BLAT sur la séquence de *V. vinifera* cv PN40024

Contig conjoints	Taille estimée sur gel (pb)	Séquences obtenues (pb)	Score	Identité (%)	Chromosome	Coordonnées sur le génome de <i>V. vinifera</i> cv PN40024
<b>56A14</b> Ct 2 - Ct 3	450	463	328	93,5	18	Chr18:26806612..26807218
<b>20I10</b> Ct 2 - Ct 5	350	308	282	96,2	12	Chr12:16691853..16692394
<b>40F13</b> Ct 3 - Ct 4	700	197	52	82,3	8	Chr8:3892591..3892903
<b>62H06</b> Ct 1 - Ct 4	700	216	205	97,7	12	Chr12:15624308..15624723

De fait, Feechan *et al.* (2013) ont annoté 1 Mb du supercontig de *M. rotundifolia* cv G52, et ont révélé que **49,8 %** de la séquence correspondait à des transposons et des rétrotransposons. Ces séquences restent malheureusement à annoter de façon expertisée sur nos contigs. En effet, l'alignement des contigs sur la séquence de référence de *V. vinifera* est sans doute un indicateur de présence en éléments répétés dans ces régions du génome (Figure 48). Par exemple, les régions du chromosome 12 comprises entre les contigs de séquence du BAC 11K09 3 et 1 et 1 et 2 (figure 43) sont respectivement riches en élément de type Copia et en éléments inconnus (à ce stade de l'annotation automatique du génome de la vigne). Ce type d'observation est également quasiment toujours vrai pour les autres BACs. Une annotation expertisée des éléments répétés sur le génome de référence est également nécessaire et en cours. Enfin, les contigs de séquence de Muscadine semblent quasi systématiquement plus courts que la région homologue sur le génome de *V. vinifera* (Figure 48). Nous n'avons pas eu le temps de vérifier si cela correspondait à une situation réelle ou à un problème d'assemblage des séquences de BAC. Nous avons en effet vu que la somme des tailles des séquences de contigs était souvent inférieure à la taille attendue du BAC, estimée sur gel après une digestion *NotI* (Tableau 11). Une autre explication pourrait être également la présence différentielle d'éléments répétés. Ces régions sont effectivement riches en tels éléments sur le génome de référence mais une analyse approfondie reste à faire.

Mon annotation expertisée a révélé la présence de **16** gènes codant potentiellement des protéines fonctionnelles, **6** pseudogènes et **1** fragment de gène. Nous avons alors regardé si nous retrouvions des

gènes homologues chez *V. vinifera*. Même en tenant compte des trous dans les séquences des BAC et des possibilités d'erreur d'orientation ou d'ordre des contigs, nous avons plusieurs exemple de gènes ou de groupes de gènes présents chez *V. vinifera* et absents chez *M. rotundifolia* : les gènes sur les contigs 1 des BAC 62H06, 11K09 et 13C22 et la région homologue correspondante (Figure 48) :

- Région 62H06 : pas de gène chez *M. rotundifolia* et 5 gènes et le fragment d'un autre sur le génome de *V. vinifera*.
- Région de 11K09 : une F-Box chez *M. rotundifolia* et 1 F-Box et une FA-desaturase sur le génome de *V. vinifera*.
- Région 13C22 : 2 NBS-LRR et 2 épimérasés chez *M. rotundifolia* versus 11 gènes chez *V. vinifera*.

Il est possible que mon expertise ait été trop stringente dans ces deux régions. De fait, **4** et **2** ORF sont proposées par les logiciels d'annotation *de novo* sur les contig 1 des clones BAC 62H06 et 11K09, respectivement (Annexe 11). Les recherches d'homologies dans la base de données protéiques SWISS-PROT révèlent que ces séquences sont généralement similaires à des éléments transposables ou ne sont pas homologues à des protéines connues. Rien de supplémentaire n'est détecté sur la séquence du 13C22 dans cette région. Il faudrait donc expertiser la séquence de *V. vinifera* également dans ces régions. Enfin des gènes semblent dupliqués chez *M. rotundifolia* par rapport à *V. vinifera*: une protéine kinase sur le BAC 40F13 et une protéine de type AIG1 sur le BAC 56A14. Dans les deux cas, le gène chez *V. vinifera* a de très longs introns contenant des éléments transposables dans l'état actuel de l'annotation.

Globalement les remaniements entre les séquences des deux génomes sont donc assez nombreux et soulignent l'intérêt de disposer de ressources génomiques spécifiques.

## ***4. Perspectives***

La banque BAC de *M. rotundifolia* cv Regale développée est une ressource génomique qui convient à une large gamme d'applications dans la recherche génétique et génomique dans la famille des *Vitaceae*. En outre, les analyses et les données obtenues dans cette étude fournissent un premier aperçu de la constitution du génome de *M. rotundifolia*. Par rapport au génome de référence de *V. vinifera*, un niveau élevé de microsynténie a été observé entre les deux génomes suggérant une conservation des microstructures du génome. En ce qui concerne l'initiative actuelle de faciliter l'identification des gènes de résistance du génome de *M. rotundifolia*, la présente étude semble utile. En effet, les paires de BES générées à partir des clones BAC de *M. rotundifolia* devraient considérablement faciliter la construction de contigs, et les marqueurs définis à partir de BES pourraient être utilisés pour l'ancrage des cartes physiques et génétiques dans le genre *Vitis*. En outre, la construction d'une base de données d'éléments transposables résultant de l'analyse des BES de *M. rotundifolia* et de *V. vinifera* améliorerait grandement les procédures de masquage dans l'assemblage des séquences génomiques des deux sous-genres. Néanmoins, deux exigences majeures sont à satisfaire, surtout dans les régions des *clusters* de gènes NBS-LRR : (i) Tout d'abord, l'augmentation de la couverture des régions où des QTL/gènes de résistance ont été identifiés doit être suffisante pour une prédiction fiable de la séquence consensus. (ii) Deuxièmement, la couverture des clones BAC doit être suffisamment élevée pour être en mesure de reconnaître les allèles alternatifs et de les regrouper en contigs spécifiques pour chacune des régions homologues. Dans ce processus, lorsque des séquences répétées sont présentes, la taille des contigs assemblés devrait dépasser la longueur du segment répété pour permettre la jonction des régions uniques. Cependant, des lectures obtenues avec la technique HiSeq (Illumina) du génome de *Muscadinia rotundifolia* cv Regale ont été obtenues et pourraient être exploitées pour avoir une meilleure couverture des régions des gènes de résistance, et ainsi permettre une comparaison plus fine de ces régions d'intérêt. Elles pourraient également être exploitées pour améliorer l'assemblage des contigs de séquence des clones BAC. Ces lectures, les BES et des contigs de clones BAC de *M. rotundifolia*, pourraient permettre de produire des marqueurs de type SNP qui peuvent être utilisés efficacement dans diverses expériences génétiques et des études d'association. Le séquençage de plus de clones BAC dans les régions des *loci* de résistance pourraient générer des données pertinentes sur les insertions et délétions pour mieux comprendre la dynamique évolutive des éléments répétées dans le génome des espèces du genre *Vitis*. Une autre ressource qui aurait pu être exploitée, dans ce sens, est la séquence de la vigne hétérozygote, *V. vinifera* cv Pinot Noir (Velasco *et al.* 2006). En parallèle de cette approche *in silico*, l'identification des gènes de résistance chez *M. rotundifolia* pourrait être possible en définissant des marqueurs spécifiques du domaine conservé NB-ARC de *M. rotundifolia* pour amplification du génome de *M. rotundifolia* par

PCR. Cette approche est soutenue par le fait que l'utilisation des amorces dégénérées spécifiques du domaine NB-ARC n'a pas donné de résultats satisfaisants dans notre étude.



# ***5. Matériels & Méthodes***

## 4.1. MATERIELS

### 4.1.1. Matériel végétal et préparation de l'ADN génomique

Les boutures de *M. rotundifolia* cv Regale ont été cultivées dans une serre à la station expérimentale de l'Institut National de la Recherche Agronomique (INRA) de Colmar, et mis à l'obscurité 24 heures avant la récolte. 10 g de jeunes feuilles ont été récoltées, immédiatement congelé dans l'azote liquide et stocké à -80 °C. L'ADN génomique a été extrait des feuilles congelées conformément au protocole décrit par Peterson *et al.* 2000.

### 4.1.2. Ressources génomiques

#### 4.1.2.1. Construction de la banque BAC de *M. rotundifolia* cv Regale

La banque BAC de *M. rotundifolia* cv Regale a été construite suivant le protocole modifié de Peterson *et al.* (2000), à partir de l'ADN génomique de *M. rotundifolia* cv Regale. Deux enzymes de restriction, *Hind*III et *Bam*HI, ont été utilisés pour la digestion partielle de l'ADN, cloné dans le vecteur pIndigoBAC536, et transformé dans la souche *E. coli*. Au total, **4** sous-banques contenant **54 174** clones ont été obtenus dont **40 734** ont été clonés dans des sites de restriction *Hind*III et **13 440** dans des sites de restriction *Bam*HI (Tableau 5). Les clones recombinants ont été prélevés et stockés dans des plaques **384** puits contenant 90 µl de milieu 2YT + glycérol (6 %) avec 12 µg/ml de chloramphenicol et 0,4 mM de MgSO<sub>4</sub>. Les plaques sont incubées à 37 °C pendant 42 heures et stockées à -80 °C en double exemplaire.

La taille moyenne des inserts a été estimée à partir de **128** clones BAC (**32** clones par sous-banque) suivant le protocole de Chalhoub *et al.* (2004). Les profils génétiques ont été générés pour une quarantaine de clones par digestion de l'ADN du plasmide avec l'enzyme de restriction *Hind*III, pendant 3 heures à 37 °C. L'ADN digéré a ensuite été séparé sur un gel d'agarose 1 % SeaKem (GTG). L'électrophorèse a été réalisée à 75 V pendant 16 heures dans une solution tampon TAE 0,5 X.

#### 4.1.2.2. Séquences d'extrémités de BAC (BES)

Le séquençage de **54 912** clones BAC a été effectué comme décrit dans Lamoureux *et al.* (2006). Les séquences de **86810** BES ont été récupérées et sont disponibles sur le site web : <http://urgi.versailles.inra.fr/Species/Vitis/cDNA-and-BACs-libraries>. Les données de séquences ont été filtrées *in silico* pour éliminer les séquences de *E. coli* et des génomes des organelles. Lorsque la totalité des BES d'un clone BAC est similaire aux séquences des mitochondries (Groremykin *et al.* 2009) et des

chloroplastes (Jansen *et al.* 2006) de vigne, l'insert est considéré comme issu d'un ADN chloroplastique/mitochondrial. Les clones BAC, pour lesquels une BES est similaire aux séquences des mitochondries et des chloroplastes, sont aussi considérés comme issus d'ADN chloroplastique/mitochondrial, si aucune séquence n'est disponible pour la seconde BES de la paire.

#### **4.1.3. Séquence du contig et des clones BAC dans la région du gène *Run1/Rpv1***

Le gène de résistance à l'oïdium, *Run1*, a été introgressé de *M. rotundifolia* cv G52 dans *V. vinifera* (Pauquet *et al.* 2001 ; Dry *et al.* 2011). Un contig de BAC, provenant d'un génotype issu d'un pseudobackcross, a été construit dans la région entre les marqueurs microsatellites VMC4F3-1 et VMC8G9 qui encadrent le gène *Run1/Rpv1* sur le chromosome 12 (Figure 22 ; Barker *et al.* 2005), dans le cadre d'un projet de clonage positionnel mené en collaboration le CSIRO. Ce contig de BAC contient encore des séquences manquantes. Les marqueurs CB46.49, CB13.14 et 49MRP1.P2, spécifiques de *M. rotundifolia*, encadrent un *cluster* de gènes NBS-LRR (Donald *et al.* 2002, Barker *et al.* 2005 et Dry *et al.* 2010). Une série de BACs s'alignant sur ce *cluster* de RGA a été séquencée correspondant à des séquences (Figure 22). Il est actuellement en 3 contigs : (i) une séquence de 917 559 bp couvrant les RGAs 1-8 ; (ii) La séquence du BAC 82O20 : 37192 bp couvrant le RGA 9 ; (iii) La séquence du BAC 97A19: 93632 bp couvrant le RGA 10.

Pour notre stratégie de recherche de région synténique sur le couple *V. vinifera* (séquence de référence) et *M. rotundifolia*, les 3 séquences ont été collés dans une seule. 500 N ont été ajoutés comme séparateur entre deux séquences.

#### **4.1.4. Séquence de référence de *Vitis vinifera***

La séquence complète du génome de la vigne est disponible depuis 2007 (Jaillon *et al.* 2007). La séquence de référence (Jaillon *et al.* 2007) a été obtenue à partir d'un génotype quasi-homozygote (97 %) PN40024 créé par l'INRA de Colmar (Bronner et Olivera, 1990). La stratégie choisie était un séquençage global (*Whole genome shotgun*) en Sanger avec une couverture de 12 équivalents génomiques. La version 12X V0 de la séquence et la version V1 de l'annotation sont disponibles (<http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>).

## 4.2. METHODES

### 4.2.1. Analyse des collections de BES

#### 4.2.1.1. Identification des séquences microsatellites

L'outil précédemment développé pour la détection de séquences microsatellites SSRSEARCH (Pai *et al.* 2009) a été utilisé pour l'identification des séquences microsatellites contenues dans l'ensemble de la collection de **86 810** BES de *M. rotundifolia* cv Regale. Un total de **77 237** BES de *V. vinifera* cv Cabernet Sauvignon (Adam *et al.* 2005 ; Lamoureux *et al.* 2006) a été utilisé comme point de comparaison. Une analyse similaire d'identification de séquences microsatellites a été réalisée sur cette collection de BES de *V. vinifera* cv Cabernet Sauvignon.

#### 4.2.1.2. Analyse des séquences répétées connues

Pour tenter d'identifier les séquences répétées, les collections de BES de *M. rotundifolia* cv Regale et de *V. vinifera* cv Cabernet Sauvignon ont été analysées avec la version open-3.3.0 de REPEATMASKER, en utilisant la base de données REPBASE Update 20120418; <http://repeatmasker.org>). L'analyse a été effectuée en utilisant les paramètres par défaut.

#### 4.2.1.3. Analyse des régions codantes

Le contenu en gènes a été estimé dans les BES avec les séquences répétées masquées.

Les BES de *V. vinifera* cv Cabernet Sauvignon et de *M. rotundifolia* cv Regale ont été comparées au protéome de *Vitis vinifera* (<http://genomes.cribi.unipd.it/grape/>) avec le programme BLASTX, suivant une E-VALUE limite de  $10^{-6}$ , un score d'alignement  $\geq 50$  et une longueur d'alignement  $\geq 20$  acide aminés. La part de ces gènes correspondant aux séquences NBS -LRR a été estimé avec le programme BLASTX et suivant une E-VALUE limite de  $10^{-4}$ , les (i) 48 motifs MEME spécifiques de cette famille de gènes (Meyers *et al.* 2003 ; Kohler *et al.* 2008) et (ii) 70 domaines NB-ARC de protéines correspondant aux gènes de résistance fonctionnellement caractérisés (Sanseverino *et al.* 2010 ; <http://prgdb.crg.eu> ; Annexe 2).

#### 4.2.1.4. Séquences et analyse phylogénétique

Les séquences d'acides aminés des domaines NB-ARC identifiés dans les BES de *M. rotundifolia* cv Regale et des gènes de résistances fonctionnels clonés chez les végétaux, en date du 13/02/2013 ([http://prgdb.crg.eu/wiki/Main\\_Page](http://prgdb.crg.eu/wiki/Main_Page)) ont été extraites puis alignées à l'aide de CLUSTALX version 2.0.8 (Thompson *et al.* 1997). L'analyse a été faite en utilisant la méthode du *Neighbour Joining* (Saitou *et al.* 1987). Un *bootstrap* de 100 réplicats a été utilisé pour évaluer l'appartenance à clade dans l'arbre phylogénétique. La valeur seuil de *bootstrap* retenue est de **70 %** (Hillis et Bull, 1993).

#### 4.2.1.5. Alignement des BES sur le génome de référence

Plusieurs paramètres ont été utilisés pour l'alignement des BES de *M. rotundifolia* cv Regale sur le génome de *V. vinifera*, en utilisant l'algorithme BLAT (Kent, 2002). Dans tous les cas, le pourcentage d'identité minimum était de 92 %. Les clones BAC ont été alignés sur le génome de *V. vinifera* selon les critères suivants : (i) Aucune stringence (Filtre 1) : Les deux extrémités d'un même clone doivent s'aligner de façon unique sur un même chromosome, en respectant les orientations 'sens' et 'anti-sens' sur l'ADN matrice ; (ii) Stringence intermédiaire (Filtre 2) : Les deux extrémités d'un même clone doivent s'aligner de façon unique sur un même chromosome, en respectant les orientations 'sens' et 'anti-sens' sur l'ADN matrice. La distance entre deux BES d'un même clone varie entre 20 kb et 150 kb, et l'alignement d'un BES se fait sur toute sa longueur +/- 100 paires de base en début et/ou à la fin du BES ; Forte stringence (Filtre 3) : Les deux extrémités d'un même clone doivent s'aligner de façon unique sur un même chromosome, en respectant les orientations 'sens' et 'anti-sens' sur l'ADN matrice. La taille entre deux BES d'un même clone varie entre 45 kb et 120 kb, et l'alignement d'un BES se fait sur toute sa longueur +/- 50 paires de base en début et/ou à la fin du BES. Les résultats des alignements peuvent être visualisés sur le site web <http://urgi.versailles.inra.fr/Projects/Muscares>.

#### 4.2.2. Criblage de la banque BAC et sélection des clones BAC candidats

##### 4.2.2.1. Choix des marqueurs génétiques et physiques

Un set de **15** marqueurs a été choisi dans les régions où des QTL de résistance ont été détectés dans des études impliquant différents croisements (Barker *et al.* 2005 ; Welter *et al.* 2007 ; Akkurt *et al.* 2007 ; Hoffmann *et al.* 2008 ; Marguerit *et al.* 2009 ; Bellin *et al.* 2009 ; Cipriani *et al.* 2011). Un autre set de trois paires d'amorces dégénérées a été choisi sur la base des résultats d'études antérieures afin d'amplifier des séquences d'ADN couvrant les domaines les plus conservés du domaine : (i) NBS NBS-F1/NBS-R1 (Yu *et al.* 1996), (ii) RG1/RG2 (Aarts *et al.* 1998) et (iii) LM637/LM638 (Kanazin *et al.* 1996).

En parallèle, les paires de BES *M. rotundifolia* qui s'alignaient sur le génome de référence de *V. vinifera*

de façon unique et dans les régions d'intérêt ont été utilisés pour développer des marqueurs afin de cribler la banque BAC. La conception d'amorces a été faite en plusieurs étapes : (i) Recherche de zones non répétées dans les BES *M. rotundifolia* cv. Regale par BLASTN sur le génome de référence de *V. vinifera* et sur la séquence du contig *Run1/Rpv1* ; (ii) Définition d'un couple d'amorces sur les zones uniques et dans un même BES ; (iii) Vérification de l'unicité des amorces grâce à une amplification *in silico* via l'outil EPCR (Schuler *et al.* 1997). Une validation supplémentaire a été effectuée par BLASTN sur le génome de *V. vinifera* et les séquences des BES *M. rotundifolia* cv. Regale.

La taille des amorces varie entre 18-26 nucléotides, avec une température d'hybridation optimale de 55°C pour du haut débit et un pourcentage de GC optimal de 50%. Tous les marqueurs nommés MRZ~ ont été développés à partir de BES *M. rotundifolia* cv. Regale. La liste de ces marqueurs est donnée dans l'annexe 6.

#### **4.2.2.2. Organisation et criblage de la banque BAC de *M. rotundifolia* cv Regale**

Les clones individuels sont conservés dans 141 plaques 384 puits. Le nombre de clones étant conséquent, il convient donc de faire des pools de clones afin d'optimiser le nombre de PCR nécessaire pour cribler la banque BAC *M. rotundifolia* cv. Regale.

La banque a été organisée en 3 dimensions comme décrit dans Lamoureux *et al.* 2006, à l'exception que les superpools ont consisté en 11 superpools de 12 plaques chacune et un autre de 9 plaques. Le regroupement des clones a été automatisé en utilisant les robots APRICOT et TECAN.

Le criblage de la banque BAC a été effectué avec les six premiers superpools ; et lorsqu'aucun clone positif n'était trouvé, les six autres superpools étaient criblés.

#### **4.2.2.3. Amplification par PCR**

La PCR est effectuée avec 4 µL de suspension bactérienne diluée 2000 fois. En plus de la matrice d'ADN, le milieu réactionnel comprend : 0.1 µL de TAQ Polymérase à 0.5 U, 1 µL de Tampon 10X, 0.64 µL de MgCl<sub>2</sub> à 25 mM, 0.4 µL de dNTP à 10 mM, 1 µL de chaque amorce dont la concentration initiale est à 12 µM, 2 µL de rouge de crésol et de l'eau q.s.p. 15 µL. Le programme d'amplification comprend un cycle de 5 min à 94°C puis 10 cycles d'amplification : 20 sec de dénaturation à 94°C, 20 sec d'hybridation de 60°C à 51°C avec 1°C de diminution à chaque cycle, 40 sec d'élongation à 72°C ; suivies de 40 cycles d'amplification : 20 sec de dénaturation à 94°C, 20 sec d'hybridation à 55°C, 40 sec d'élongation à 72°C, et enfin un dernier cycle d'élongation à 72°C pendant 7 min.

#### **4.2.2.4. Identification des clones BAC positifs**

Les produits amplifiés par PCR sont analysés par électrophorèse en conditions non dénaturantes sur gel d'agarose (3%) selon la taille des fragments à analyser, dans du tampon TAE 0,5 X (Tris-HCl 20 mM, EDTA 0,5 mM, acétate de sodium 2,5 mM). L'électrophorèse a été faite à 300 mA pendant 45 min. Les gels contenant du bromure d'éthidium sont analysés sous lumière UV dans l'appareil «GEL DOC apparatus» permettant la visualisation des acides nucléiques.

L'amplification par PCR permet l'identification directe des clones positifs mais les coordonnées des clones candidats sont vérifiées par une PCR supplémentaire sur BAC individuel. De nouveaux marqueurs uniques sont définis dans les BES des clones candidats pour cribler à nouveau la banque BAC *M. rotundifolia* cv. Regale. Ce nouveau crible permettra de construire un contig de BAC qui couvrirait les *loci* de résistance.

#### **4.2.3. Séquençage et annotation des contigs de BAC**

##### **4.2.3.1. Raccordement des contigs**

Le séquençage des clones BAC a été réalisé au CNRGV de Toulouse grâce à la technologie 454 conçue par J. Rothberg (Margulies, Egholm *et al.* 2005). Les séquences ont été obtenues en plusieurs contigs. Les programmes d'assemblage PHRAP (Gordon *et al.* 2004) et MATCHER (Longden *et al.* 1999) ont donc été utilisés pour identifier les contigs chevauchants.

Les séquences manquantes entre les contigs d'un même clone BAC ont été recherchées par une amplification PCR. L'ADN est alors amplifié en utilisant toutes les combinaisons de paires d'oligonucléotides uniques comme amorces. L'amplification PCR ne donne de résultat positif que si les deux amorces correspondent à des séquences prises sur deux brins complémentaires. Dans le cas d'un résultat positif, la méthode permet donc non seulement de positionner deux contigs voisins, mais elle permet de déterminer la séquence manquante par séquençage du fragment d'ADN obtenu par PCR.

##### **4.2.3.2. Annotation des contigs de clones BAC**

Une prédiction automatique des gènes contenus dans les contigs de clones BAC a d'abord été annotés avec le programme FGENESH (Solovyev *et al.* 2006) entraîné pour *V. vinifera*. Ensuite, les prédictions ont été vérifiées manuellement (i) en recherchant des similarités avec des protéines connues dans la banque de données SWISSPROT par BLASTX (avec un seuil d'identité de **90** %), (ii) avec des EST (*Expressed*

*Sequenced Tags*) dans les collections d'EST disponibles au NCBI avec le programme SIM4 (Florea *et al.* 1998), et (iii) en prédisant les sites d'épissage potentiels avec le programme de prédiction SPLICEMACHINE (Degroeve *et al.* 2004). Les contigs de *M. rotundifolia* cv Regale ont été annotés fonctionnellement par comparaison avec les bases de données de famille de protéines et de domaines protéiques PFAM (Finn *et al.* 2010) et INTERPRO. (Hunter *et al.* 2011).



# *Références bibliographiques*

- Aarts, M. G., Lintel Hekkert, B. T., Holub, E. B., Beynon, J. L., Stiekema, W. J., & Pereira, A. (1998). Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, *11*(4), 251-258.
- Adam-Blondon A-F, Bernole A, Faes G, Lamoureux D, Pateyron S, Grando MS, Caboche M, Velasco R, Chalhoub B (2005) Construction and characterization of BAC libraries from major grapevine cultivars. *Theor Appl Genet*, *110* : 1363-1371
- Adam-Blondon A-F, Jaillon O, Vezzulli S, Zharkikh A, Troggio M, Velasco R (2011) Genome Sequence Initiatives. *in*: A-F Adam-Blondon, JM Martinez-Zapater, Chittaranjan Kole (eds) *Genetics, Genomics and Breeding of Grapes*. Science Publishers and CRC Press. Pp 211-234
- Adam-Blondon, A. F., Roux, C., Claux, D., Butterlin, G., Merdinoglu, D., & This, P. (2004). Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. *Theoretical and Applied Genetics*, *109*(5), 1017-1027.
- Ade, J., DeYoung, B. J., Golstein, C., & Innes, R. W. (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences*, *104*(7), 2531-2536.
- Albrecht, D. R., Underhill, G. H., Wassermann, T. B., Sah, R. L., & Bhatia, S. N. (2006). Probing the role of multicellular organization in three-dimensional microenvironments. *Nature methods*, *3*(5), 369-375.
- Ammiraju, J. S., Luo, M., Goicoechea, J. L., Wang, W., Kudrna, D., Mueller, C., ... & Wing, R. A. (2006). The *Oryza* bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus *Oryza*. *Genome Research*, *16*(1), 140-147.
- Akkurt M, Welter L, Maul E, Töpfer R, Eva Zyprian E (2007) Development of SCAR markers linked to powdery mildew (*Uncinula necator*) resistance in grapevine (*Vitis vinifera* L. and *Vitis* sp.). *Mol Breeding*, *19*:103–111
- Ameline-Torregrosa, C.; Wang, B.B.; O'Bleness, M.S. (2008) Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. *Plant Physiol*, *146*, 5–21.
- Amrani L, Corio-Costet MF (2006). A single nucleotide polymorphism in the beta-tubulin gene distinguishing two genotypes of *Erysiphe necator* expressing different symptoms on grapevine. *PlantPathol*, *55*:505-512.
- Areshchenkova, T., & Ganal, M. W. (1999). Long tomato microsatellites are predominantly associated with centromeric regions. *Genome*, *42*(3), 536-544.

- Aubertot J.N., Pinochet X., Doré T. (2004). Analysis of the effects of sowing date and nitrogen availability during vegetative stages on phoma stem canker (*Leptosphaeria maculans*) development on two winter oilseed rape cultivars. *Crop Protection*, 23, 635-645
- Austin, M.J.; Muskett, P.; Kahn, K.; Feys, B.J.; Jones, J.D.; Parker, J.E. (2002) Regulatory role of SGT1 in early R gene-mediated plant defenses. *Science*, 295, 2077–2080.
- Bai, S.; Liu, J.; Chang, C.; Zhang, L.; Maekawa, T.; Wang, Q.; Xiao, W.; Liu, Y.; Chai, J.; Takken, F.L.W.; *et al.* (2012) Structure-function analysis of barley NLR immune receptor MLA10 reveals its cell compartment specific activity in cell death and disease resistance. *PLoS Pathog.*, 8, e1002752.
- Bari, R.; Datt Pant, B.; Stitt, M.; Scheible, W.-R. (2006) PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol*, 141, 988–999.
- Barker CL, Donald T, Pauquet J, Ratnaparkhe A, Bouquet A, Adam-Blondon A-F, Thomas MR, Dry I (2005) Genetic and physical mapping of the grapevine powdery mildew resistance gene, *Run1*, using a bacterial artificial chromosome library. *Theor Appl Genet*, 111 : 370-377
- Bartel, D.P. (2009) MicroRNAs: Target recognition and regulatory functions. *Cell*, 136, 215–233.
- Baudoin, A., Olaya, G., Delmotte, F., Colcol, J. F., & Sierotzki, H. (2008). QoI resistance of *Plasmopara viticola* and *Erysiphe necator* in the mid-Atlantic United States. *Plant Health Progress*, (February), 0211-02.
- Baulcombe, D. (2004) RNA silencing in plants. *Nature*, 431, 356–363.
- Bellin D, Peressotti E, Merdinoglu D, Wiedemann-Merdinoglu S, Adam-Blondon AF, Cipriani G, Morgante M, Testolin R, Di Gaspero G.(2009) Resistance to *Plasmopara viticola* in grapevine 'Bianca' is controlled by a major dominant gene causing localised necrosis at the infection site. *Theor Appl Genet*, 120 : 163-176
- Bendahmane, A.; Querci, M.; Kanyuka, K.; Baulcombe, D.C. (2000) Agrobacterium transient expression system as a tool for the isolation of disease resistance genes: Application to the *Rx2* locus in potato. *Plant J*, 21, 73–81.
- Benharbit El Alami N (1995). Recherches sur le développement normal et anormal du gynécée de *Vitis vinifera* L. du bouton floral à la fleur fécondée. Importance des tanins vacuolaires et relation avec la coulure. *Thèse de l'Université de Bordeaux II*, France, 126p
- Bennetzen, J. L. (2000). Transposable element contributions to plant gene and genome evolution. *Plant molecular biology*, 42(1), 251-269.
- Bennetzen J. L. (2000) Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. *The Plant cell* 12: 1021–1029.

- Bent, A.F.; Mackey, D. (2007) Elicitors, effectors, and *R* genes: The new paradigm and a lifetime supply of questions. *Annu. Rev. Phytopathol.*, 45, 399–436.
- Bhattacharjee, S., Zamora, A., Azhar, M. T., Sacco, M. A., Lambert, L. H., & Moffett, P. (2009). Virus resistance induced by NB–LRR proteins involves Argonaute4-dependent translational control. *The Plant Journal*, 58(6), 940-951.
- Bieri, S.; Mauch, S.; Shen, Q.H.; Peart, J.; Devoto, A.; Casais, C.; Ceron, F.; Schulze, S.; Steinbiß, H.-H.; Shirasu, K.; Schulze-Lefert, P. (2004) RAR1 positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance. *Plant Cell*, 16, 3480–3495.
- Bisson, L. F., Waterhouse, A. L., Ebeler, S. E., Walker, M. A., & Lapsley, J. T. (2002). The present and future of the international wine industry. *Nature*, 418(6898), 696-699.
- Blanc G, Barakat A, Guyot R, Cooke R, Delseny M (2000) Extensive duplication and reshuffling in the Arabidopsis genome. *Plant Cell* 12:1093–1101
- Blanc S, Wiedemann-Merdinoglu S, Dumas V, Mestre P, Merdinoglu D (2012) A reference genetic map of *Muscadinia rotundifolia* and identification of Ren5, a new major locus for resistance to grapevine powdery mildew. *Theor Appl Genet*, 125:1663–1675
- Blasi, P., Blanc, S., Wiedemann-Merdinoglu, S., Prado, E., Rühl, E. H., Mestre, P., & Merdinoglu, D. (2011). Construction of a reference linkage map of *Vitis amurensis* and genetic mapping of Rpv8, a locus conferring resistance to grapevine downy mildew. *Theoretical and applied genetics*, 123(1), 43-53.
- Bloodworth PJ, Nesbitt WB, Barker KR (1980) Resistance to rootknot nematodes in *Euvinis* × *Muscadinia* hybrids. In: Proceedings of third international symposium on grape breeding, University of California, Davis, 15–18 June 1980, pp 275–292
- Boubals D (1959). Amélioration de la résistance de la vigne au mildiou (*Plasmopara viticola* (Berk et Curt.) Berlese et de Toni). Recherche de géniteurs de résistance. *Annales de l'amélioration des plantes* 6:481–525
- Bouquet A (1980). *Vitis* × *Muscadinia* hybridization: a new way in grape breeding for disease resistance in France (Acta Hort). In: Proceedings of third international symposium on grape breeding, University of California, Davis, 15–18 June 1980, pp 42-61
- Boursiquot JM, Audeguin L, Charmont S, Desperrier JM, Dufour MC, Jacquet O, Lacombe T, Leguay M, Moulliet C, Ollat N, Schneider C, Serreno C. (2007). *Catalogue des variétés et clones de vignes cultivés en France*. Institut Français de la Vigne et du Vin 2
- Boursiquot JM, This P. (2000). Essai de définition du cépage. *PAV* (94):5-7.

- Braun U, Cook RTA, Inman AJ, Shin HD (2002). The Taxonomy of the Powdery Mildew Fungi. In BÉLANGER RR et al. (eds) *The Powdery Mildews. A Comprehensive Treatise*, pp. 13–55. APS Press, St. Paul.
- Braun, U. & Takamatsu, S. (2000). Phylogeny of *Erysiphe*, *Microsphaera*, *Uncinula* (*Erysipheae*) and *Cystotheca*, *Podosphaera*, *Sphaerotheca* (*Cystothecaceae*) inferred from rDNA ITS sequences – some taxonomic consequences. *Schlechtendalia* 4: 1-33.
- Brewer, M. T., Cadle-Davidson, L., Cortesi, P., Spanu, P. D., & Milgroom, M. G. (2011). Identification and structure of the mating-type locus and development of PCR-based markers for mating type in powdery mildew fungi. *Fungal Genetics and Biology*, 48(7), 704-713.
- Brewer, M. T., & Milgroom, M. G. (2010). Phylogeography and population structure of the grape powdery mildew fungus, *Erysiphe necator*, from diverse *Vitis* species. *BMC evolutionary biology*, 10(1), 268.
- Brondani, R., Brondani, C., & Grattapaglia, D. (2002). Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Molecular Genetics and Genomics*, 267(3), 338-347.
- Brown M, Moore JN, Fenn P, McNew RW (1999) Comparison of leaf disk, greenhouse, and Weld screening procedures for evaluation of grape seedlings for downy mildew resistance. *HortScience* 34:331–333
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., ... & Kleinhofs, A. (2002). The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences*, 99(14), 9328-9333.
- Burch-Smith, T. M., Schiff, M., Caplan, J. L., Tsao, J., Czymmek, K., & Dinesh-Kumar, S. P. (2007). A novel role for the TIR domain in association with pathogen-derived elicitors. *PLoS biology*, 5(3), e68.
- Butault J.P., Dedryver C.A., Gary C., Guichard L., Jacquet F., Meynard J.M., Nicot P., Pitrat M., Reau R., Sauphanor B., Savini I., Volay T., (2010), ECOPHYTO R&D : quelles voies pour réduire l’usage des pesticides. Rapport de synthèse. 91 p.
- Cadle-Davidson, L. (2008). Variation within and between *Vitis* spp. for foliar resistance to the downy mildew pathogen *Plasmopara viticola*. *Plant Disease* 92: 1577–1584.
- Calenge, F., Van der Linden, C. G., Van de Weg, E., Schouten, H. J., Van Arkel, G., Denance, C., & Durel, C. E. (2005). Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. *Theoretical and Applied Genetics*, 110(4), 660-668.

- Carver, T., Berriman, M., Tivey, A., Patel, C., Böhme, U., Barrell, B. G., ... & Rajandream, M. A. (2008). Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics*, 24(23), 2672-2676.
- Casagrande, K., Falginella, L., Castellarin, S. D., Testolin, R., & Di Gaspero, G. (2011). Defence responses in Rpv3-dependent resistance to grapevine downy mildew. *Planta*, 234(6), 1097-1109.
- Chandrasekaran, C.; Betrán, E. (2008) Origins of new genes and pseudogenes. *Nat. Educ*, 1, 1.
- Chen, W. J.; Delmotte, F.; Richard-Cervera, S.; Douence, L.; Greif, C.; Corio-Costet, M. F. (2007). At least two origins of fungicide resistance in grapevine downy mildew populations. *Applied and Environmental Microbiology*, 73: 5162-5172
- Cheng, X., Xu, J., Xia, S., Gu, J., Yang, Y., Fu, J., ... & Liu, K. (2009). Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. *Theoretical and applied genetics*, 118(6), 1121-1131.
- Cheng, Y.T.; Germain, H.; Wiermer, M.; Bi, D.; Xu, F.; García, A.V.; Wirthmuellere, L.; Després, C.; Parkere, J.E.; Zhang, Y.; *et al.* (2009) Nuclear pore complex component MOS7/Nup88 is required for innate immunity and nuclear accumulation of defense regulators in *Arabidopsis*. *Plant Cell*, 21, 2503–2516.
- Cheung F, Town CD (2007) A BAC-end view of the *Musa acuminata* genome. *BMC Plant Biol* 7:29.
- Choulet, F., Wicker, T., Rustenholz, C., Paux, E., Salse, J., Leroy, P., ... & Feuillet, C. (2010). Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *The Plant Cell Online*, 22(6), 1686-1701.
- Cipriani G, Di Gaspero G, Canaguier A, Jusseaume J, Tassin J, Lemainque A, Vincent Thareau V, Adam-Blondon A-F, Testolin R (2011) Molecular Linkage Maps: Strategies, Resources and Achievements. *in: A-F Adam-Blondon, JM Martinez-Zapater, Chittaranjan Kole (eds) Genetics, Genomics and Breeding of Grapes*. Science Publishers and CRC Press. pp 111-136
- Clayton CN (1975) Diseases of muscadine and bunch grapes in North Carolina and their control. *NC Agr Expt Stn Bull* 451:31
- Coleman, C., Copetti, D., Cipriani, G., Hoffmann, S., Kozma, P., Kovács, L., ... & Di Gaspero, G. (2009). The powdery mildew resistance gene REN1 co-segregates with an NBS-LRR gene cluster in two Central Asian grapevines. *BMC genetics*, 10(1), 89.
- Collier, S.M.; Hamel, L.P.; Moffett, P. (2011) Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Mol. Plant Microbe Interact*, 24, 918–931.
- Collins, N. C., Webb, C. A., Seah, S., Ellis, J. G., Hulbert, S. H., & Pryor, A. (1998). The isolation and mapping of disease resistance gene analogs in maize. *Molecular plant-microbe interactions*, 11(10), 968-978.

- Comeaux BL, Nesbitt WB, Fantz PR (1987) Taxonomy of the native grapes of North Carolina. *Castanea* 52:197–215
- Cooke, R., Piègu, B., Panaud, O., Guyot, R., Salse, J., Feuillet, C., & Delseny, M. (2007). From rice to other cereals: comparative genomics. In *Rice Functional Genomics* (pp. 429-479). Springer New York.
- Corio-Costet, M. F. ; Delmotte, F. ; Martinez, F. ; Giresse, X. ; Raynal, M. ; Richart-Cervera, S. ; Douence, L. ; Panon, M.L. ; Chen, W.J. (2006). Resistance aux Qois du mildiou de la vigne (*Plasmopara viticola*) : origine et diversité. 8th Int. Conf on Pest and Diseases, AFPP, Tours, 5-6 décembre, France). pp 612-620 CD-Rom.
- Corio-Costet, M. F. (2007). (ed) *Erysiphe necator*. Tec/Doc Lavoisier, Paris
- Corio-Costet, M. F. ; Martinez, F. ; Delmotte, F. ; Douence, L. ; Richart Cervera, S. ; Chen, W-J. (2008). Resistance of *Plasmopara viticola* to QoI fungicides: Origin and Diversity. In: Dehne, H.W, Gisi, U., Kuck, K.H., Russell, P.E. and Lyr H (ed.), *Modern fungicides and Antifungal compounds V* , DPG Selbstverlag, Braunschweig, pp. 107-112.
- Corio-Costet M, Dufour M, Cigna J, Abadie P, Chen W (2011) Resistance of *Plasmopara viticola*: diversity and fitness of resistant isolates to QoI fungicides. In: Lebeda A, Holmes GJ, Mauch-Mani B and Jeger MJ, eds. *The downy Mildews- Biology, Mechanisms of Resistance and Population Ecology*. Springer, 183-197.
- Corio-Costet MF Fungicide resistance in *Plasmopara viticola* in France and anti resistance measures (2012). CAB international 2012. In: Fungicide Resistance in crop protection : Risk and management (ed. TS Thind), 157-171. Sous presse.
- Csizmazia J, Bereznai L (1968) A szőlő *Plasmopara viticola* és a *Viteus vitifolii* elleni rezisztencia nemesítés eredményei. Orsz Szől Bor Kut Int Évkönyve, Budapest, pp 191–200
- Dai GH, Andary C, Mondolot-Cosson L, Boubals D (1995). Histochemical studies on the interaction between three species of grapevine, *Vitis vinifera*, *V. rupestris*, *V. rotundifolia* and the downy mildew fungus, *Plasmopara viticola*. *Physiol Mol Plant Pathol* 177–188
- Dai, L., Wu, J., Li, X., Wang, X., Liu, X., Jantasuriyarat, C., ... & Wang, G. L. (2010). Genomic structure and evolution of the Pi2/9 locus in wild rice species. *Theoretical and Applied Genetics*, 121(2), 295-309.
- Dalbó, M. A., Ye, G. N., Weeden, N. F., Wilcox, W. F., & Reisch, B. I. (2001). Marker-assisted selection for powdery mildew resistance in grapes. *Journal of the American Society for Horticultural Science*, 126(1), 83-89.
- Dangl, J.L.; Jones, J.D.G. (2001) Plant pathogens and integrated defence responses to infection. *Nature*, 411, 826–833.

- Datema, E., Mueller, L., Buels, R., Giovannoni, J., Visser, R., Stiekema, W., & van Ham, R. (2008). Comparative BAC end sequence analysis of tomato and potato reveals overrepresentation of specific gene families in potato. *BMC plant biology*, 8(1), 34.
- David P, Chen NWG, Pedrosa-Harand A, Thareau V, Sevignac M, Cannon SB, Debouck D, Langin T, Geffroy V. (2009). A Nomadic Subtelomeric Disease Resistance Gene *Cluster* in Common Bean. *Plant Physiology*, 151(3):1048-1065.
- Davidis UX, Olmo HP (1964) The *Vitis vinifera* × *V. rotundifolia* hybrids as *phylloxera* resistant rootstocks. *Vitis* 4:129–143
- Deacon, J.W., (2006). Fungal Biology. Blackwell, Malden, MA
- Degroeve, S. (2004). *Design and Evaluation of a Linear Classification Strategy for Gene Structural Element Recognition* (Doctoral dissertation, PhD thesis, Universiteit Gent, Faculty of Sciences, Gent, Belgium).
- Delledonne M, Xia Y, Dixon R and Lamb C (1988) Nitric oxide functions as a signal in plant disease resistance. *Nature* 394:585–588
- Delmotte, F., Leterme, N., Gauthier, J. P., Rispé, C., & Simon, J. C. (2002). Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Molecular Ecology*, 11(4), 711-723.
- Délye C, Laigret F, Corio-Costet MF (1997). RAPD analysis provides insight into the biology and epidemiology of *Uncinula necator*. *Phytopathology*, 87:670-677.
- Dereeper, A., Guyot, R., Tranchant-Dubreuil, C., Anthony, F., Argout, X., de Bellis, F., Combes, M., Gavory, F., Kochko, A., Kudrna, D., Leroy, T., Poulain, J., Rondeau, M., Song, X., Wing, R., & Lashermes, P. (2013). BAC-end sequences analysis provides first insights into coffee (*Coffea canephora* P.) genome composition and evolution. *Plant molecular biology*, 1-13.
- Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S., Marco, Y. (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences*, 100(13), 8024-8029.
- Detjen LR (1919) The limits in hybridization of *Vitis rotundifolia* with related species and genera. *NC Agric Expt Sta Bull* 17:42
- DeYoung, B. J., & Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature immunology*, 7(12), 1243-1249.
- Di Gaspero, G., Cipriani, G., Adam-Blondon, A. F., & Testolin, R. (2007). Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. *Theoretical and Applied Genetics*, 114(7), 1249-1263.



- Di Gaspero G, Copetti D, Coleman C, Castellarin SD, Eibach R, Kozma P, Lacombe T, Gambetta G, Zvyagin A, Cindri P, Kovács L, Morgante M, Testolin R (2012) Selective sweep at the *Rpv3* locus during grapevine breeding for downy mildew resistance. *Theor Appl Genet*, 124:277–286
- Dielen, A.S.; Badaoui, S.; Candresse, T.; German-Retana, S. (2010) The ubiquitin/26S proteasome system in plant-pathogen interactions: A never-ending hide-and-see game. *Mol. Plant Pathol*, 11, 293–308.
- Dinesh-Kumar, S.P.; Baker, B. (2000) Alternatively spliced N resistance gene transcripts: Their possible role in tobacco mosaic virus resistance. *Proc. Natl. Acad. Sci. USA*, 97, 1908–1913.
- Dion R. (1982). Histoire de la vigne et du vin en France des origines au 19<sup>ème</sup> siècle. Ed. Flammarion: Paris XII
- Dodds, P.N.; Lawrence, G.J.; Catanzariti, A.M.; The, T.; Wang, C.I.; Ayliffe, M.A.; Kobe, B.; Ellis, J.G. (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc. Natl. Acad. Sci. USA*, 103, 8888–8893.
- Donald, T. M., Pellerone, F., Adam-Blondon, A. F., Bouquet, A., Thomas, M. R., & Dry, I. B. (2002). Identification of resistance gene analogs linked to a powdery mildew resistance locus in grapevine. *Theoretical and Applied Genetics*, 104(4), 610-618.
- Dong, X. (1998). SA, JA, ethylene, and disease resistance in plants. *Current opinion in plant biology*, 1(4), 316-323.
- Dopman EB, Pérez L, Bogdanowicz SM, Harrison RG (2005). Consequences of reproductive barriers for genealogical discordance in the European corn borer. *Proc Natl Acad Sci USA* 2005, 102:14706-14711.
- Dry, I. B., Feechan, A., Anderson, C., Jermakow, A. M., Bouquet, A., ADAM-BLONDON, A. F., & Thomas, M. R. (2010). Molecular strategies to enhance the genetic resistance of grapevines to powdery mildew. *Australian Journal of Grape and Wine Research*, 16(s1), 94-105.
- Dubos B (2002). Maladies cryptogamiques de la vigne. Féret edn, Bordeaux
- Dunstan RT (1962) Some fertile hybrids of bunch and muscadine grapes. *J Hered* 53:299-303
- Duque, P. (2011) A role for SR proteins in plant stress responses. *Plant Signal. Behav*, 6, 49–54.
- Eibach R, Diehl H, Alleweldt G (1989) Untersuchungen zur Vererbung von Resistenzeigenschaften bei Reben gegen *Oidium tuckeri*, *Plasmopara viticola* und *Botrytis cinerea*. *Vitis* 28:209-228
- Eibach R, Töpfer R (2003) Success in resistance breeding: ‘Regent’ and its steps into the market. *Acta Hort* 687–691

- Ellis, J., and Jones, D. (1998). Structure and function of proteins controlling strain-specific pathogen resistance in plants. *Curr. Opin. Plant Biol.* 1, 288–293.
- Emmett RW, Wicks TJ, Magarey PA, Madge DG (1990). Recent advances in grapevine powdery mildew management. Australian and New Zealand. *Wine Industry Journal*, 5:213-217
- Erickson, E. O., & Wilcox, W. F. (1997). Distributions of sensitivities to three sterol demethylation inhibitor fungicides among populations of *Uncinula necator* sensitive and resistant to triadimefon. *Phytopathology*, 87(8), 784-791.
- Esmenjaud D, Van Ghelder C, Roger V, Bordenave L, Decroocq S, Bouquet A, Ollat N (2010) Host suitability of *Vitis* and *Vitis*–*Muscadinia* material to the nematode *Xiphinema index* over one to four years. *Am J Enol Vitic* 61:96–101
- Eujayl I, Sledge MK, Wang L, Chekhoskiy K, Zwonitzer JC, Mian MAR (2004) *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. *Theor Appl Genet* 108:414–422
- Evans KJ, Whisson DL, Stummer BE, Scott ES (1997). DNA markers identify variation in Australian populations of *Uncinula necator*. *Mycol Res*, 101:923-932.
- Faivre-Rampant P.F, Lesur I, Boussardon C, Bitton F, Martin-Magniette M-L, Bodénès C, Le Provost G, Bergès H, Fluch S, Kremer A, Plomion C (2011) Analysis of BAC end sequences in oak, a keystone forest tree species, providing insight into the composition of its genome. *BMC Genomics*, 12:292
- Falk, S., Gadoury, D., Pearson, R., and Seem, R., (1995). Partial control of grape powdery mildew by the mycoparasite *Ampelomyces quisqualis*. *The American Phytopathological Society* 79:483-490.
- Feschotte, C., Jiang, N., & Wessler, S. R. (2002). Plant transposable elements: where genetics meets genomics. *Nature Reviews Genetics*, 3(5), 329-341.
- Feuillet C, Keller B (2002) Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. *Annals of botany* 89: 3–10.
- Firoozabady E, Olmo HP (1982) Resistance to grape *Phylloxera* in *Vitis vinifera* × *Vitis rotundifolia* grape hybrids. *Vitis* 21:1–4
- Fischer, B. M., Salakhutdinov, I., Akkurt, M., Eibach, R., Edwards, K. J., Töpfer, R., & Zyprian, E. M. (2004). Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. *Theoretical and Applied Genetics*, 108(3), 501-515.
- Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual review of phytopathology*, 9(1), 275-296.
- Florea, L., Hartzell, G., Zhang, Z., Rubin, G. M., & Miller, W. (1998). A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome research*, 8(9), 967-974.

- Fluhr, J. W., Kao, J., Jain, M., Ahn, S. K., Feingold, K. R., & Elias, P. M. (2001). Generation of free fatty acids from phospholipids regulates stratum corneum acidification and integrity. *Journal of investigative dermatology*, 117(1), 44-51.
- Frelichowski JE Jr, Palmer MB, Main D, Tomkins JP, Cantrell RG, Stelly DM, Yu J, Kohel RJ, Ulloa M (2006) Cotton genome mapping with new microsatellites from *Acala* 'Maxxa' BAC-ends. *Mol Genet Genomics* 275:479–491.
- Frenkel, O., Brewer, M. T., & Milgroom, M. G. (2010). Variation in pathogenicity and aggressiveness of *Erysiphe necator* from different *Vitis* spp. and geographic origins in the eastern United States. *Phytopathology*, 100(11), 1185-1193.
- Gabriel, D. W., & Rolfe, B. G. (1990). Working models of specific recognition in plant-microbe interactions. *Annual review of phytopathology*, 28(1), 365-391.
- Gadoury DM, Pearson RC (1991). Heterothallism and pathogenic specialization in *Uncinula necator*. *Phytopathology*, 81:1287-1293.
- Galet, P. (1988). Les hybrides producteurs directs. In Galet P (ed) *Cépages et Vignobles de France*. Imprimerie Charles Dehan, Parc Euromédecine. Montpellier, France, vol 1, pp. 369-523.
- Galet P (2000). *A Practical Ampelography: Grapevine Identification* Ithaca, New York: Cornell University Press.
- Galet, P. (2000). *Précis de viticulture* (7 édition).
- Gallais, A. (1993). La sélection assistée par marqueurs. *Le sélectionneur français* 43 : 43-62
- Gao, Z.; Chung, E.-H.; Eitas, T.K.; Dangl, J.L. (2011) Plant intracellular innate immune receptor Resistance to *Pseudomonas syringae* pv. *maculicola* 1 (RPM1) is activated at, and functions on, the plasma membrane. *Proc. Natl. Acad. Sci. USA*, 108, 7619–7624.
- Gary C., Mézière D., et al., 2009 « Analyse comparative de différents systèmes en viticulture ». *In/ ECOPHYTO R&D : vers des systèmes de culture économes en produits phytosanitaires*. Rapport d'expertise. Inra ed., Tome III, 57 p + annexes.
- Gebhardt, C., & Valkonen, J. P. (2001). Organization of genes controlling disease resistance in the potato genome. *Annual review of Phytopathology*, 39(1), 79-102.
- Geffroy, V., Sévignac, M., De Oliveira, J. C., Fouilloux, G., Skroch, P., Thoquet, P., ... & Dron, M. (2000). Inheritance of partial resistance against *Colletotrichum lindemuthianum* in *Phaseolus vulgaris* and co-localization of quantitative trait loci with genes involved in specific resistance. *Molecular Plant-Microbe Interactions*, 13(3), 287-296.

- Gobbin, D.; Rumbou, A.; Linde, C. C.; Gessler, C. (2006). Population genetic structure of *Plasmopara viticola* after 125 years of colonization in European vineyards. *Molecular Plant Pathology*, 7: 519–531.
- Goldy RG (1992) Breeding muscadine grapes. *Hort Rev* 14:357–405
- Gordon, D. (2003). Viewing and editing assembled sequences using Consed. *Current protocols in bioinformatics*, 11-2.
- Grzegorzczak W, Walker MA (1998) Evaluating resistance to grape *phylloxera* in *Vitis* species with an in vitro dual culture assay. *Am J Enol Vitic* 49:17–22
- Giannuzzi G, D'Addabbo P, Gasparro M, Martinelli M, Carelli FN, Antonacci D, Ventura M (2011) Analysis of high-identity segmental duplications in the grapevine genome. *BMC Genomics*, 12:436
- Guichard L., Brunet N., et al., 2009, « Analyse comparative de différents systèmes en grandes cultures » In/ *ECOPHYTO R&D : vers des systèmes de culture économes en produits phytosanitaires*. Rapport d'expertise. Inra ed., Tome II, 133 p.
- Guo, W., Wang, W., Zhou, B., & Zhang, T. (2006). Cross-species transferability of G. arboreum-derived EST-SSRs in the diploid species of *Gossypium*. *Theoretical and Applied Genetics*, 112(8), 1573-1581.
- Guo, Y.L.; Fitz, J.; Schneeberger, K.; Ossowski, S.; Cao, J.; Weigel, D. (2011) Genome-wide comparison of nucleotide-binding site-leucine-rich-repeat-encoding genes in *Arabidopsis*. *Plant Physiol*, 157, 757–769.
- Ha, D. T., Kim, H., Thuong, P. T., Ngoc, T. M., Lee, I., Hung, N. D., & Bae, K. (2009). Antioxidant and lipoxygenase inhibitory activity of oligostilbenes from the leaf and stem of *Vitis amurensis*. *Journal of ethnopharmacology*, 125(2), 304-309.
- Hammond-Kosack, K. E., & Kanyuka, K. (2007). Resistance genes (R genes) in plants. *ELS*.
- Han, Y., & Korban, S. S. (2008). An overview of the apple genome through BAC end sequence analysis. *Plant molecular biology*, 67(6), 581-588.
- Han, Y. M., Yang, F., Zhou, A. G., Miller, E. R., Ku, P. K., Hogberg, M. G., & Lei, X. G. (1997). Supplemental phytases of microbial and cereal sources improve dietary phytate phosphorus utilization by pigs from weaning through finishing. *Journal of Animal Science*, 75(4), 1017-1025.
- He, L.; Du, C.; Covalada, L.; Xu, Z.; Robinson, A.F.; Yu, J.Z.; Kohel, R.J.; Zhang, H.B. (2004) Cloning, characterization, and evolution of the NBS-LRR-encoding resistance gene analogue family in polyploid cotton (*Gossypium hirsutum* L.). *Mol. Plant Microbe Interact*, 17, 1234–1241.
- Hillis, D. M., & Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic biology*, 42(2), 182-192.

- Hirotsune, S.; Yoshida, N.; Chen, A.; Garrett, L.; Sugiyama, F.; Takahashi, S.; Yagami, K.; Wynshaw-Boris, A.; Yoshiki, A. (2003) An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature*, 423, 91–96.
- Hofmann S, Di Gaspero G, Kovács L, Howard S, Kiss E, Galbács Z, Testolin T, Kozma P (2008) Resistance to *Erysiphe necator* in the grapevine ‘Kishmish vatkana’ is controlled by a single locus through restriction of hyphal growth. *Theor Appl Genet*, 116:427–438
- Holt, B.F.; Belkhadir, Y.; Dangl, J.L. (2005) Antagonistic control of disease resistance protein stability in the plant immune system. *Science*, 309, 929–932.
- Hong, C. P., Piao, Z. Y., Kang, T. W., Batley, J., Yang, T., Hur, Y., ... & Lim, Y. P. (2007). Genomic distribution of simple sequence repeats in *Brassica rapa*. *Molecules and cells*, 23(3), 349.
- Huang KS, Lin M and Cheng GF (2001). Anti-inflammatory tetramers of resveratrol from the roots of *Vitis amurensis* And the conformations of the seven-membered ring in some oligostilbenes. *Phytochemistry* 58:357-362, 2001
- Huglin, P., & Schneider, C. (1998). Biology and ecology of the grapevine (No. Ed. 2). *Technique et Documentation Lavoisier*.
- Hubert, D.A.; Tornero, P.; Belkhadir, Y.; Krishna, P.; Takahashi, A.; Shirasu, K.; Dangl, J.L. (2003) Cytosolic HSP90 associates with and modulates the *Arabidopsis* RPM1 disease resistance protein. *EMBO J*, 22, 5679–5689.
- Hulbert SH, Webb CA, Smith SM, Sun Q. (2001). Resistance gene complexes: Evolution and Utilization. *Annual Review of Phytopathology*, 39(1):285-312.
- Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K., Bateman, A., ... & Yong, S. Y. (2012). InterPro in 2011: new developments in the family and domain prediction database. *Nucleic acids research*, 40(D1), D306-D312.
- Hwang, C.F.; Williamson, V.M. (2003) Leucine-rich repeat-mediated intramolecular interactions in nematode recognition and cell death signaling by the tomato resistance protein Mi. *Plant J*, 34, 585–593.
- Innes, R. W., Ameline-Torregrosa, C., Ashfield, T., Cannon, E., Cannon, S. B., Chacko, B., ... & Young, N. D. (2008). Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. *Plant Physiology*, 148(4), 1740-1759.
- Inuma T, Khodaparast SA, Takamatsu S. (2007). Multilocus phylogenetic analyses within *Blumeria graminis*, a powdery mildew fungus of cereals. *Mol Phylogenet Evol*, 44(2):741-751.
- Jaillon O, Aury J-M, Noel B, Policriti A, et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449 : 463-468

- Johal, G. S., & Briggs, S. P. (1992). Reductase activity encoded by the HM1 disease resistance gene in maize. *Science*, 258(5084), 985-987.
- Jones, J. D., & Dangl, J. L. (2006). The plant immune system. *Nature*, 444(7117), 323-329.
- Jupe, F.; Pritchard, L.; Etherington, G.J.; MacKenzie, K.; Cock, P.J.A.; Wright, F.; Sharma, S.K.; Bolser, D.; Bryan, G.J.; Jones, J.D.G.; *et al.* (2012) Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics*, 13, 75.
- Kamoun, S. (2003). Molecular genetics of pathogenic oomycetes. *Eukaryot. Cell* 2:191–199.
- Kang, Y.J.; Kim, K.H.; Shim, S.; Yoon, M.Y.; Sun, S.; Kim, M.Y.; Van, K.; Lee, S.H. (2012) Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. *BMC Plant Biol*, 12, 139.
- Karkamkar, S. P., Patil, S. G., & Misra, S. C. (2010). Cyto-morphological studies and their significance in evolution of family Vitaceae. *The Nucleus*, 53(1-2), 37-43.
- Katti, M. V., Ranjekar, P. K., & Gupta, V. S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution*, 18(7), 1161-1167.
- Kawchuk, L. M., Hachey, J., Lynch, D. R., Kulcsar, F., Van Rooijen, G., Waterer, D. R., ... & Prüfer, D. (2001). Tomato Ve disease resistance genes encode cell surface-like receptors. *Proceedings of the National Academy of Sciences*, 98(11), 6511-6515.
- Kim, S.H.; Kwon, S.I.; Saha, D.; Anyanwu, N.C.; Gassmann, W. (2009) Resistance to the *Pseudomonas syringae* effector HopA1 is governed by the TIR-NBS-LRR protein RPS6 and is enhanced by mutations in SRFR11. *Plant Physiol*, 150, 1723–1732.
- Kirk, P. M., Cannon, P. F., David, J. C., & Stalpers, J. A. (2001). *Ainsworth & Bisby's dictionary of the fungi* (9th ed.). Wallingford: CAB International.
- Kobayashi, M.; Seo, S.; Hirai, K.; Yamamoto-Katou, A.; Katou, S.; Seto, H.; Meshi, T.; Mitsuhashi, I.; Ohashi, Y. (2010) Silencing of WIPK and SIPK mitogen-activated protein kinases reduces tobacco mosaic virus accumulation but permits systemic viral movement in tobacco possessing the N resistance gene. *Mol. Plant Microbe Interact*, 23, 1032–1041.
- Kobe, B., & Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Current opinion in structural biology*, 11(6), 725-732.
- Kohler, A.; Rinaldi, C.; Duplessis, S.; Baucher, M.; Geelen, D.; Duchaussoy, F.; Meyers, B.C.; Boerjan, W.; Martin, F. (2008) Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Mol. Biol*, 66, 619–636.
- Kortekamp A, Zyprian E (2003). Characterization of *Plasmopara* resistance in grapevine using *in vitro* plants. *J Plant Physiol* 160:1393–1400

- Kruijt, M., Brandwagt, B. F., & de Wit, P. J. (2004). Rearrangements in the Cf-9 disease resistance gene cluster of wild tomato have resulted in three genes that mediate Avr9 responsiveness. *Genetics*, 168(3), 1655-1663.
- Kuang HH, Wei FS, Marano MR, Wirtz U, Wang XX, Liu J, Shum WP, Zaborsky J, Tallon LJ, Rensink W, *et al* (2005) The R1 resistance gene cluster contains three groups of independently evolving, type I R1 homologues and shows substantial structural variation among haplotypes of *Solanum demissum*. *Plant Journal*, 44(1):37-51.
- Kuang, H.; Woo, S.S.; Meyers, B.C.; Nevo, E.; Michelmore, R.W. (2004) Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. *Plant Cell*, 16, 2870–2894.
- Kuang H, Caldwell KS, Meyers BC, Michelmore RW (2008) Frequent sequence exchanges between homologs of RPP8 in Arabidopsis are not necessarily associated with genomic proximity. *Plant Journal*, 54(1):69-80.
- Lafon R, Clerjeau M (1988) Downy mildew. In RC Pearson, AC Goheen, eds, Compendium of Grape Diseases. American Phytopathological Society, St. Paul, MN, pp 11–13
- Lai CWJ, Yu Q, Hou S, Skelton RL, Jones MR, Lewis KLT, Murray J, Eustice M, Guan P, Agbayani R, Moore PH, Ming R, Presting GG. (2006). Analysis of papaya BAC-end sequences reveals first insights into the organization of a fruit tree genome. *Mol Genet Genomics* 276:1–12.
- Lam, E., & Chua, N. H. (1990). GT-1 binding site confers light responsive expression in transgenic tobacco. *Science*, 248(4954), 471-474.
- Lamoureux D, Bernole A, Le Clainche I, Tual, S, Thareau V, Paillard S, Legeai F, Dossat C, Wincker P, Oswald M, Merdinoglu D, Vignault C, Delrot S, Caboche M, Chalhoub B, Adam-Blondon A-F (2006) Anchoring a large set of markers onto a BAC library for the development of a draft physical map of the grapevine genome. *Theor Appl Genet*, 113 : 344-356
- Large EC (1940). The Advance of the Fungi New York: *Henry Holt and Co.*
- Latijnhouwers, M., De Wit, P. J., & Govers, F. (2003). Oomycetes and fungi: similar weaponry to attack plants. *Trends in microbiology*, 11(10), 462-469.
- Legros JP. (1993). L'invasion du vignoble par le phylloxéra. *Academie des sciences et lettres de Montpellier Bull.* n°24(205-22).
- Legros JP. (1997). Le phylloxéra, une histoire sans fin. *Association française pour l'avancement des sciences* 97(1): 32-39.
- Leipe, D. D., Koonin, E. V., & Aravind, L. (2004). STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual

- phyletic patterns, and evolution by horizontal gene transfer. *Journal of molecular biology*, 343(1), 1-28.
- Leister, D. (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet*, 20, 116–122.
- Leister, D., Kurth, J., Laurie, D. A., Yano, M., Sasaki, T., Devos, K., ... & Schulze-Lefert, P. (1998). Rapid reorganization of resistance gene homologues in cereal genomes. *Proceedings of the National Academy of Sciences*, 95(1), 370-375.
- Lenné JM, Wood D (1991). Plant diseases and the use of wild germplasm. *Phytopathol*, 29:35-63
- Lepik EE (1970). Gene centers of plants as sources of resistance. *Annu Rev Phytopathol* 1970, 8:323-344
- Levadoux, L. (1956). *Les populations sauvages et cultivées des Vitis vinifera L.* Institut national de la recherche agronomique.
- Li, D., Wan, Y., Wang, Y., & He, P. (2008). Relatedness of resistance to anthracnose and to white rot in Chinese wild grapes. *Vitis*, 47(4), 213-215.
- Li, F.; Pignatta, D.; Bendix, C.; Brunkard, J.O.; Cohn, M.M.; Tung, J.; Sun, H.; Kumar, P.; Baker, B. (2012) MicroRNA regulation of plant innate immune receptors. *Proc. Natl. Acad. Sci. USA*, 109, 1790–1795.
- Li, X.; Clarke, J.D.; Zhang, Y.; Dong, X. (2001) Activation of an EDS1-mediated *R*-gene pathway in the *sncl* mutant leads to constitutive, NPR1-independent pathogen resistance. *Mol. Plant Microbe Interact*, 14, 1131–1139.
- Li, Y.; Li, S.; Bi, D.; Cheng, Y.T.; Li, X.; Zhang, Y. (2010) SRRFR1 Negatively regulates plant NB-LRR resistance protein accumulation to prevent autoimmunity. *PLoS Pathog*, 6, e1001111.
- Liu, J.J.; Ekramoddoullah, A.K. (2003) Isolation, genetic variation and expression of TIR-NBS-LRR resistance gene analogs from western white pine (*Pinus monticola* Dougl. ex. D. Don.). *Mol. Genet. Genomics*, 270, 432–441.
- Lodhi, M. A., & Reisch, B. I. (1995). Nuclear DNA content of *Vitis* species, cultivars, and other genera of the *Vitaceae*. *Theoretical and Applied Genetics*, 90(1), 11-16.
- Lozano, R.; Ponce, O.; Ramirez, M.; Mostajo, N.; Orjeda, G. (2012) Genome-wide identification and mapping of NBS-encoding resistance genes in *Solanum tuberosum* group Phureja. *PLoS One*, 7, e34775.
- Lu, R.; Malcuit, I.; Moffett, P.; Ruiz, M.T.; Peart, J.; Wu, A.-J.; Rathjen, J.P.; Bendahmane, A.; Day, L.; Baulcombe, D.C. (2003) High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *EMBO J*, 22, 5690–5699.



- Lukasik, E., & Takken, F. L. (2009). STANDING strong, resistance proteins instigators of plant defence. *Current opinion in plant biology*, 12(4), 427-436.
- Luo, S.; Zhang, Y.; Hu, Q.; Chen, J.; Li, K.; Lu, C.; Liu, H.; Wang, W.; Kuang, H. (2012) Dynamic nucleotide-binding-site and leucine-rich-repeat-encoding genes in the grass family. *Plant Physiol*, 159, 197–210.
- Lyon, D. (2007). *Surveillance studies: An overview*. Polity.
- Ma, H., Kong, Z., Fu, B., Li, N., Zhang, L., Jia, H., and Ma, Z.(2011). Identification and mapping of a new powdery mildew resistance gene on chromosome 6D of common wheat. *Theoretical and Applied Genetics* 123: 1099–106.
- Maccaferri, M.; Ratti, C.; Rubies-Autonell, C.; Vallega, V.; Demontis, A.; Stefanelli, S.; Tuberosa, R.; Sanguineti, M.C. (2011) Resistance to soil-borne cereal mosaic virus in durum wheat is controlled by a major QTL on chromosome arm 2BS and minor loci. *Theor. Appl. Genet*, 123, 527–544.
- Maekawa, T., Cheng, W., Spiridon, L. N., Töller, A., Lukasik, E., Saijo, Y., ... & Schulze-Lefert, P. (2011). Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell host & microbe*, 9(3), 187-199.
- Mallory, A.C.; Bouché, N. (2008) MicroRNA-directed regulation: To cleave or not to cleave. *Trends Plant Sci*, 13, 359–367.
- Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., ... & Bryant, S. H. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research*,39(suppl 1), D225-D229.
- Marguerit E, Boury C, Manicki A, Donnart M, Butterlin G, Némorin A, Wiedemann-Merdinoglu S, Merdinoglu D, Ollat N, Decroocq S (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. *Theor Appl Genet*, 118:1261–1278
- Marone, D.; Panio, G.; Ficco, D.B.M.; Russo, M.A.; de Vita, P.; Papa, R.; Rubiales, D.; Cattivelli, L.; Mastrangelo, A. (2012) Characterization of wheat DArT markers: Genetic and functional features. *Mol. Genet. Genomics*, 287, 741–753.
- Mastrangelo, A.M.; Marone, D.; Laidò, G.; De Leonardis, A.M.; de Vita, P. (2012) Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. *Plant Sci*, 185–186, 40–49.
- McDowell, J.M.; Simon, S.A. (2006) Recent insights into *R* gene evolution. *Mol. Plant Pathol*, 7, 437–448.
- McGovern PE (2003). Ancient wine: the search for the origins of viticulture. *Ed Princeton Univ Princeton, New Jersey, USA*, 400p

- McHale, L.; Tan, X.; Koehl, P.; Michelmore, R.W. (2006) Plant NBS-LRR proteins: Adaptable guards. *Genome Biol*, 7, 212.
- Meier, I., & Somers, D. E. (2011). Regulation of nucleocytoplasmic trafficking in plants. *Current opinion in plant biology*, 14(5), 538-546.
- Merdinoglu D, Wiedeman-Merdinoglu S., Coste P., Dumas V., Haetty S., Butterlin G., Greif C. (2003). Genetic analysis of downy mildew resistance derived from *Muscadinia rotundifolia*. *Acta Horticulturae* 603: 451-456
- Merdinoglu D, Merdinoglu-Wiedemann S, Mestre P, Prado E, Schneider C (2009). Apport de l'innovation variétale dans la réduction des intrants phytosanitaires au vignoble: exemple de la résistance au mildiou et à l'oïdium. *Progrès Agric Vitic* 12:290–293
- Mestre, P.; Baulcombe, D.C. (2006) Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell*, 18, 491–501.
- Meyers, B.C.; Dickerman, A.W.; Michelmore, R.W.; Sivaramakrishnan, S.; Sobral, B.W.; Young, N.D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J*, 20, 317–332.
- Meyers, B. C., Kozik, A., Griego, A., Kuang, H., & Michelmore, R. W. (2003). Genome-wide analysis of NBS-LRR–encoding genes in Arabidopsis. *The Plant Cell Online*, 15(4), 809-834.
- Meyers, B. C., Morgante, M., & Michelmore, R. W. (2002). TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in Arabidopsis and other plant genomes. *The Plant Journal*, 32(1), 77-92.
- Miazzi M, Hajjeh H, Faretra F (2003). Observations on the population biology of the grape powdery mildew fungus *Uncinula necator*. *J Plant Pathol*, 85:123-129.
- Michelmore, R. W., & Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, 8(11), 1113-1130.
- Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., ... & Sasaki, T. (2004). The genome sequence of silkworm, *Bombyx mori*. *Dna Research*, 11(1), 27-35.
- Moffett, P.; Farnham, G.; Peart, J.; Baulcombe, D.C. (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *EMBO J*, 21, 4511–4519.
- Mondragón-Palomino, M., Meyers, B. C., Michelmore, R. W., & Gaut, B. S. (2002). Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. *Genome Research*, 12(9), 1305-1315.
- Monosi, B.; Wissner, R.J.; Pennill, L.; Hulbert, S.H. (2004) Full-genome analysis of resistance gene homologues in rice. *Theor. Appl. Genet*, 109, 1434–1447.

- Montarry, J., Cartolaro, P., Richard-Cervera, S., & Delmotte, F. (2009). Spatio-temporal distribution of *Erysiphe necator* genetic groups and their relationship with disease levels in vineyards. *European Journal of Plant Pathology*, *123*(1), 61-70.
- Moreira, F. M., Madini, A., Marino, R., Zulini, L., Stefanini, M., Velasco, R., ... & Grando, M. S. (2011). Genetic linkage maps of two interspecific grape crosses (*Vitis* spp.) used to localize quantitative trait loci for downy mildew resistance. *Tree Genetics & Genomes*, *7*(1), 153-167.
- Moroldo M, Paillard S, Marconi R, Legeai F, Canaguier A, Cruaud C, De Berardinis V, Guichard C, Brunaud V, Le Clainche I, Scalabrin S, Testolin R, Di Gaspero G, Morgante M., Adam-Blondon A.F (2008) A physical map of the heterozygous grapevine ‘Cabernet Sauvignon’ allows mapping candidate genes for disease resistance. *BMC Plant Biol*, *8*:66
- Mortensen JA (1981). Sources and inheritance of resistance to anthracnose in *Vitis*. *J Hered* *72*:423–426
- Mullins MG, Bouquet A, Williams LE (1992). Biology of the grapevine. *Cambridge Univ. Press, New York*
- Mun, J. H., Kim, D. J., Choi, H. K., Gish, J., Debellé, F., Mudge, J., Denny R., Endré, G., Saurat, O., Dubez, A. M., Kiss, G. B., Roe, R., Young, N. D., & Cook, D. R. (2006). Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. *Genetics*, *172*(4), 2541-2555.
- Mun, J.H.; Yu, H.J.; Park, S.; Park, B.S. (2009) Genome-wide identification of NBS-encoding resistance genes in *Brassica rapa*. *Mol. Genet. Genomics*, *282*, 617–631.
- Noël, L., Moores, T. L., van der Biezen, E. A., Parniske, M., Daniels, M. J., Parker, J. E., & Jones, J. D. (1999). Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of *Arabidopsis*. *The Plant Cell Online*, *11*(11), 2099-2111.
- Núñez Y, Gallego J, Ponz F, Raposo R (2006). Analysis of population structure of *Erysiphe necator* using AFLP markers. *Plant Pathol*, *55*:650-656.
- Nürnbergger, T., & Brunner, F. (2002). Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. *Current opinion in plant biology*, *5*(4), 318-324.
- Nurnberger Jr, J. I., Wiegand, R., Bucholz, K., O'Connor, S., Meyer, E. T., Reich, T., ... & Porjesz, B. (2004). A family study of alcohol dependence: coaggregation of multiple disorders in relatives of alcohol-dependent probands. *Archives of general psychiatry*, *61*(12), 1246.
- Olien WC (1990a) Muscadine-a classic southeastern fruit. *Hort-Science* *25*:726–831
- Olien WC (1990b) The muscadine grape: botany, viticulture, history, and current industry. *HortScience* *25*:732–739

- Ollat, N., & Gaudillère, J. (1997, May). Carbon balance in developing grapevine berries. In *V International Symposium on Grapevine Physiology 526* (pp. 345-350).
- Olmo HP (1971) *Vinifera* × *Rotundifolia* hybrids as wine grapes. *Am J Enol Vitic* 22:87–91
- Olmo HP (1986) The potential role of (*vinifera* × *rotundifolia*) hybrids in grape variety improvement. *Experientia* 42:921–926
- Pastrana-Bonilla E, Akoh C, Sellappan S, Krewer G (2003) Phenolic Content and Antioxidant Capacity of Muscadine Grapes. *J. Agric. Food Chem*, 5: 5497-5503
- Pai, T. W., Chen, C. M., Hsiao, M. C., Cheng, R., Tzou, W. S., & Hu, C. H. (2009). An online conserved SSR discovery through cross-species comparison. *Advances and applications in bioinformatics and chemistry: AABC*, 2, 23.
- Pan, Q., Wendel, J. and Fluhr, R. (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *J. Mol. Evol.* 50, 203–213.
- Pant, B.D.; Buhtz, A.; Kehr, J.; Scheible, W.R. (2008) Micro-RNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J*, 53, 731–738.
- Parniske, M., & Jones, J. D. (1999). Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. *Proceedings of the National Academy of Sciences*, 96(10), 5850-5855.
- Patel GI, Olmo HP (1955) Cytogenetics of Vitis: I. The hybrid of *V. vinifera* × *V. rotundifolia*. *Am J Bot* 42:141–159
- Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, et al. (2000) Comparative genomics of plant chromosomes. *Plant Cell* 12: 1523–1540.
- Pauquet, J., Bouquet, A., This, P., & Adam-Blondon, A. F. (2001). Establishment of a local map of AFLP markers around the powdery mildew resistance gene Run1 in grapevine and assessment of their usefulness for marker assisted selection. *Theoretical and Applied Genetics*, 103(8), 1201-1210.
- Peressotti, E., Wiedemann-Merdinoglu, S., Delmotte, F., Bellin, D., Di Gaspero, G., Testolin, R., ... & Mestre, P. (2010). Breakdown of resistance to grapevine downy mildew upon limited deployment of a resistant variety. *BMC plant biology*, 10(1), 147.
- Péros JP, Berger G, Portemont A, Boursiquot J-M, Lacombe T (2011) Genetic variation and biogeography of the disjunct *Vitis* subg. *Vitis* (*Vitaceae*). *Journal of Biogeography* (J. Biogeogr.) 38, 471–486
- Péros JP, Troulet C, Guerriero M, Michel-Romiti C, Notteghem JL (2005). Genetic variation and population structure of the grape powdery mildew fungus, *Erysiphe necator*, in southern France. *Eur J Plant Pathol*, 113:407-416.

- Phytowelt GmbH (2003) Study on the use of the varieties of interspecific vines. Germany
- Pieterse, C. M., Van Wees, S. C., Van Pelt, J. A., Knoester, M., Laan, R., Gerrits, H., ... & Van Loon, L. C. (1998). A novel signaling pathway controlling induced systemic resistance in *Arabidopsis*. *The Plant Cell Online*, 10(9), 1571-1580.
- Pitrat M., Nicot P., Brismontier E. et al., 2009, « Analyse comparative de différents systèmes en cultures légumières. » In/ *ECOPHYTO R&D : vers des systèmes de culture économes en produits phytosanitaires*. Rapport d'expertise. Inra ed., Tome V, 63 p.
- Pitzschke, A., Schikora, A., & Hirt, H. (2009). MAPK cascade signalling networks in plant defence. *Current opinion in plant biology*, 12(4), 421-426.
- Porter, B.W.; Paidi, M.; Ming, R.; Alam, M.; Nishijima, W.T.; Zhu, Y.J. (2009) Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. *Mol. Genet. Genomics*, 281, 609–626.
- Pouget R. (1990). Histoire de la lutte contre le phylloxéra de la vigne en France : 1868-1895. Institut national de la recherche agronomique.
- Qi, D.; de Young, B.J.; Innes, R.W. (2012) Structure-function analysis of the coiled-coil and leucine-rich repeat domains of the RPS5 disease resistance protein. *Plant Physiol.*, 158, 1819–1832.
- Rairdan, G.J.; Collier, S.M.; Sacco, M.A.; Baldwin, T.T.; Boettrich, T.; Moffett, P. (2008) The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. *Plant Cell*, 20, 739–751.
- Rairdan, G., & Moffett, P. (2007). Brothers in arms? Common and contrasting themes in pathogen perception by plant NB-LRR and animal NACHT-LRR proteins. *Microbes and Infection*, 9(5), 677-686.
- Raynal, G., & Ravise, A. (1980). Action du tris-O-éthylphosphonate d'aluminium (phoséthyl d'aluminium) sur la pathogénie de *Plusnropuru viticola* et sur la stimulation des re 'actions de défense de la vigne. *Ann. Phyfopofhol*, 12(3), 163-175.
- Riaz, S., Tenscher, A. C., Ramming, D. W., & Walker, M. A. (2011). Using a limited mapping strategy to identify major QTLs for resistance to grapevine powdery mildew (*Erysiphe necator*) and their use in marker-assisted breeding. *Theoretical and applied genetics*, 122(6), 1059-1073.
- Riaz, S., Vezzulli, S., Harbertson, E. S., & Walker, M. A. (2007). Use of molecular markers to correct grape breeding errors and determine the identity of novel sources of resistance to *Xiphinema index* and Pierce's disease. *American Journal of Enology and Viticulture*, 58(4), 494-498.
- Rossetto (2001) Sourcing of SSR markers from related plant species. In: Henry RJ (ed) *Plant genotyping: the DNA fingerprinting of plants*. CABI, Wallingford, pp 211–224

- Royer C. (1988). Mouvement historiques de la vigne dans le monde. In la vigne et le vin. *La Manufacture et la Cité des sciences et de l'industrie*: 15–25.
- Ruel JJ, Walker MA (2006) Resistance to Pierce's disease in *Muscadinia rotundifolia* and other native grape species. *Am J Enol Vitic* 57:158–165
- Rumbolz, J., Kassemeyer, H.-H., Steinmetz, V., Deising, H.B., Mendgen, K., Mathys, D., Wirtz, S., and Guggenheim, R. (1999). Differentiation of infection structures of the powdery mildew fungus *Uncinula necator* and adhesion to the host cuticle. *Can. J. Bot.* 78, 409–421
- Russo, M.A.; Ficco, D.B.M.; Marone, D.; de Vita, P.; Vallega, V.; Rubies-Autonell, C.; Ratti, C.; Ferragonio, P.; Giovanniello, V.; Pecchioni, N.; *et al.* (2012) A major QTL for resistance to soil-borne cereal mosaic virus derived from an old Italian durum wheat cultivar. *J. Plant Int*, 7, 290–300.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., & Barrell, B. (2000). Artemis: sequence visualization and annotation. *Bioinformatics*, 16(10), 944-945.
- Ryther, R. C. C., Flynt, A. S., Phillips, J. A., & Patton, J. G. (2004). siRNA therapeutics: big potential from small RNAs. *Gene Therapy*, 12(1), 5-11.
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, 4:406–425.
- Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology*, 94(3), 441-448.
- Sanseverino W, Roma G, De Simone M, Faino L, Melito S, Stupka E, Frusciante L, Ercolano MR (2010) PRGdb: a bioinformatics platform for plant resistance gene analysis. *Nucleic Acids Res*, 38(Database issue): D814-21
- Sauphanor B., Dirwimmer C., *et al.*, 2009, « Analyse comparative de différents systèmes en arboriculture fruitière ». In/ *ECOPHYTO R&D: vers des systèmes de culture économes en produits phytosanitaires*. Rapport d'expertise. Inra ed., Tome IV, 49 p.
- Savocchia, S., Stummer, B. E., Wicks, T. J., Van Heeswijck, R., & Scott, E. S. (2004). Reduced sensitivity of *Uncinula necator* to sterol demethylation inhibiting fungicides in southern Australian vineyards. *Australasian Plant Pathology*, 33(4), 465-473.
- Schlosser, J., Olsson, N., Weis, M., Reid, K., Peng, F., Lund, S., & Bowen, P. (2008). Cellular expansion and gene expression in the developing grape (*Vitis vinifera* L.). *Protoplasma*, 232(3-4), 255-265.
- Schröder, S.; Telle, S.; Nick, P.; Thines, M. (2011). Cryptic diversity of *Plasmopara viticola* (*Oomycota*, *Peronosporaceae*) in North America. *Organisms diversity and Evolution*, 11: 3-7

- Schwander, F., Eibach, R., Fechter, I., Hausmann, L., Zyprian, E., & Töpfer, R. (2012). Rpv10: a new locus from the Asian *Vitis* gene pool for pyramiding downy mildew resistance loci in grapevine. *Theoretical and Applied Genetics*, *124*(1), 163-176.
- Seah, S., Telleen, A. C., & Williamson, V. M. (2007). Introgressed and endogenous Mi-1 gene clusters in tomato differ by complex rearrangements in flanking sequences and show sequence exchange and diversifying selection among homologues. *Theoretical and Applied Genetics*, *114*(7), 1289-1302.
- Shang, J.; Tao, Y.; Chen, X.; Zou, Y.; Lei, C.; Wang, J.; Li, X.; Zhao, X.; Zhang, M.; Lu, Z.; *et al.* (2009) Identification of a new rice blast resistance gene, *Pid3*, by genome wide comparison of paired nucleotide-binding site leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes. *Genetics*, *182*, 1303–1311.
- Sharma, M. K., Sharma, R., Cao, P., Jenkins, J., Bartley, L. E., Qualls, M., ... & Ronald, P. C. (2012). A genome-wide survey of switchgrass genome structure and organization. *PLoS One*, *7*(4), e33892.
- Shen, K. A., Meyers, B. C., Islam-Faridi, M. N., Chin, D. B., Stelly, D. M., & Michelmore, R. W. (1998). Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. *Molecular Plant-Microbe Interactions*, *11*(8), 815-823.
- Shen, Q. H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., ... & Schulze-Lefert, P. (2007). Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science*, *315*(5815), 1098-1103.
- Shivaprasad, P.V.; Chen, H.-M.; Patel, K.; Bond, D.M.; Santos, B.A.C.M.; Baulcombe, D.C. (2012) A microRNA superfamily regulates nucleotide binding site leucine-rich repeats and other mRNAs. *Plant Cell*, *24*, 859–874.
- Shultz, J. L., Kazi, S., Bashir, R., Afzal, J. A., & Lightfoot, D. A. (2007). The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. *Theoretical and Applied Genetics*, *114*(6), 1081-1090.
- Smith RE (1961). Grape mildew as viewed in the early agricultural press of California. *Plant Dis Rep*, *45*:700-702
- Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol*, *7*(Suppl 1), S10.
- Spencer, D. M. (1978). *The powdery mildews*. Academic press, New York
- Speulman, E., Bouchez, D., Holub, E. B., & Beynon, J. L. (1998). Disease resistance gene homologs correlate with disease resistance loci of *Arabidopsis thaliana*. *The Plant Journal*, *14*(4), 467-474.
- Staal, J.; Dixelius, C. (2008) RLM3, a potential adaptor between specific TIR-NB-LRR receptors and DZC proteins. *Commun. Integr. Biol*, *1*, 59–61.

- Staskawicz, B. J., Mudgett, M. B., Dangl, J. L., & Galan, J. E. (2001). Common and contrasting themes of plant and animal diseases. *Science*, 292(5525), 2285-2289.
- Staudt, G. and Kassemeyer, H.H. (1995). Evaluation of downy mildew resistance in various accessions of wild *Vitis* species. *Vitis* 34: 225–228.
- Stummer BE, Zanker T, Scott ES, Whisson DL (2000). Genetic diversity in populations of *Uncinula necator*: Comparison of RFLP- and PCR-based approaches. *Mycol Res*, 104:44-52.
- Suetsugu, Y., Minami, H., Shimomura, M., Sasanuma, S. I., Narukawa, J., Mita, K., & Yamamoto, K. (2007). End-sequencing and characterization of silkworm (*Bombyx mori*) bacterial artificial chromosome libraries. *BMC genomics*, 8(1), 314.
- Sun, Q.; Collins, N.C.; Ayliffe, M.; Smith, S.M.; Drake, J.; Pryor, T.; Hulbert, S.H. (2001) Recombination between paralogues at the Rp1 rust resistance locus in maize. *Genetics*, 158, 423–438.
- Swiderski, M. R., & Innes, R. W. (2001). The Arabidopsis PBS1 resistance gene encodes a member of a novel protein kinase subfamily. *The Plant Journal*, 26(1), 101-112.
- Takken, F. L., Albrecht, M., & Tameling, W. I. (2006). Resistance proteins: molecular switches of plant defence. *Current opinion in plant biology*, 9(4), 383-390.
- Tameling, W. I., Elzinga, S. D., Darmin, P. S., Vossen, J. H., Takken, F. L., Haring, M. A., & Cornelissen, B. J. (2002). The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *The Plant Cell Online*, 14(11), 2929-2939.
- Tan, S., & Wu, S. (2012). Genome wide analysis of nucleotide-binding site disease resistance genes in *Brachypodium distachyon*. *Comparative and Functional Genomics*, 2012.
- Tao, Y.; Yuan, F.; Leister, RT.; Ausubel, F.M.; Katagiri, F. (2000) Mutational analysis of the *Arabidopsis* nucleotide binding site-leucine-rich repeat resistance gene RPS2. *Plant Cell*, 12, 2541–2554.
- This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes. *Trends in Genet*, 22: 511-519
- Thomas, C. M., Jones, D. A., Parniske, M., Harrison, K., Balint-Kurti, P. J., Hatzixanthis, K., & Jones, J. D. (1997). Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. *The Plant Cell Online*, 9(12), 2209-2224.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25:4876–4882.



- Thordal-Christensen, H. (2003). Fresh insight into processes of nonhost resistance. *Curr. Opin. Plant Biol.* 6, 351–357.
- Tör, M.; Gordon, P.; Cuzick, A.; Eulgem, T.; Sinapidou, E.; Mert, F.; Can, C.; Dangl, J.L.; Holub, E.B. (2002) *Arabidopsis* SGT1b is required for defense signaling conferred by several downy mildew (*Peronospora parasitica*) resistance genes. *Plant Cell*, 14, 993–1003.
- Tornero, P.; Merritt, P.; Sadanandom, A.; Shirasu, K.; Innes, R.W.; Dangl, J.L. (2002) RAR1 and NDR1 contribute quantitatively to disease resistance in *Arabidopsis*, and their relative contributions are dependent on the *R* gene assayed. *Plant Cell*, 14, 1005–1015.
- Unger S, Büche C, Boso S, Kassemeyer HH (2007). The course of colonization of two diVerent *Vitis* genotypes by *Plasmopara viticola* indicates compatible and incompatible host-pathogen interactions. *Phytopathology* 97:780–786
- van Loon, H., & Shea, D. J. (1999). A probable signal of the 11-year solar cycle in the troposphere of the northern hemisphere. *Geophysical Research Letters*, 26(18), 2893-2896.
- Van Ooijen, G.; Mayr, G.; Kasiem, M.M.A.; Albrecht, M.; Cornelissen, B.J.C.; Takken, F.L.W. (2008) Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.* 59, 1383–1397.
- Van Ooijen, G.; van den Burg, H.A.; Cornelissen, B.J.; Takken, F.L. (2007). Structure and function of resistance proteins in solanaceous plants. *Annu. Rev. Phytopathol.*, 45, 43–72
- Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, Fitzgerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Demattè L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando SM, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2:e1326
- Venuti, S., Copetti, D., Foria, S., Falginella, L., Hoffmann, S., Bellin, D., ... & Di Gaspero, G. (2013). Historical Introgression of the Downy Mildew Resistance Gene Rpv12 from the Asian Species *Vitis amurensis* into Grapevine Varieties. *PloS one*, 8(4), e61228.
- Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136, 669–687.
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity*, 93(1), 77-78.
- Walker MA, Lider LA, Goheen AC, Olmo HP (1991) VR 039–16. *HortScience* 26:1224–1225

- Walker MA, Ferris H, Eyre M (1994a) Resistance in *Vitis* and *Muscadinia* species to *Meloidogyne incognita*. *Plant Dis* 78:1055–1058
- Walker MA, Wolpert JA, Weber E (1994b) Viticultural characteristics of VR hybrid rootstocks in a vineyard site infected with grapevine fanleaf virus. *Vitis* 33:19–23
- Walter, S., Letiembre, M., Liu, Y., Heine, H., Penke, B., Hao, W., ... & Fassbender, K. (2007). Role of the toll-like receptor 4 in neuroinflammation in Alzheimer's disease. *Cellular Physiology and Biochemistry*, 20(6), 947-956.
- Wan, H.; Yuan, W.; Bo, K.; Shen, J.; Pang, X.; Chen, J. (2013) Genome-wide analysis of NBS-encoding disease resistance in *Cucumis sativus* and phylogenetic study of NBS-encoding genes in Cucurbitaceae crops. *BMC Genomics*, 14, 109.
- Wan, Y., Schwaninger, H., He, P., & Wang, Y. (2007). Comparison of resistance to powdery mildew and downy mildew in Chinese wild grapes. *VITIS-GEILWEILERHOF*, 46(3), 132.
- Warren, R. F., Merritt, P. M., Holub, E., & Innes, R. W. (1999). Identification of three putative signal transduction genes involved in R gene-specified disease resistance in *Arabidopsis*. *Genetics*, 152(1), 401-412.
- Weaver, L.M.; Swiderski, M.R.; Li, Y.; Jones, J.D.G. (2006) The *Arabidopsis thaliana* TIR-NB-LRR R-protein, RPP1A; protein localization and constitutive activation of defence by truncated alleles in tobacco and *Arabidopsis*. *Plant J*, 47, 829–840.
- Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Toepfer R, Zyprian E (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). *Mol Breed* 20: 359-374.
- Williams, S.J.; Sornaraj, P.; deCourcy-Ireland, E.; Menz, R.I.; Kobe, B.; Ellis, J.G.; Dodds, P.N.; Anderson, P.A. (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Mol. Plant Microbe Interact*, 24, 897–906.
- Wirthmueller, L.; Zhang, Y.; Jones, J.D.G.; Parker, J.E. (2007) Nuclear accumulation of the *Arabidopsis* immune receptor RPS4 is necessary for triggering EDS1-dependent defense. *Curr. Biol*, 17, 2023–2029.
- Wyand RA, Brown JKM (2003). Genetic and forma specialis diversity in *Blumeria graminis* of cereals and its implications for host-pathogen co-evolution. *Mol Plant Pathol*, 4:187-198.
- Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M., & Turner, J. G. (2001). Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science*, 291(5501), 118-120.

- Xiao, S., Emerson, B., Ratanasut, K., Patrick, E., O'Neill, C., Bancroft, I., & Turner, J. G. (2004). Origin and maintenance of a broad-spectrum disease resistance locus in *Arabidopsis*. *Molecular biology and evolution*, *21*(9), 1661-1672.
- Xu, S.; Zhang, Z.; Jing, B.; Gannon, P.; Ding, J.; Xu, F.; Li, X.; Zhang, Y. (2011) Transportin-SR is required for proper splicing of resistance genes and plant immunity. *PLoS Genet.*, *7*, e1002159.
- Xue, J.-Y.; Wang, Y.; Wu, P.; Wang, Q.; Yang, L.-T.; Pan, X.-H.; Wang, B.; Chen, J.-Q. (2012) A primary survey on bryophyte species reveals two novel classes of nucleotide-binding site (NBS) genes. *PLoS One*, *7*, e36700.
- Yang, S.; Feng, Z.; Zhang, X.; Jiang, K.; Jin, X.; Hang, Y.; Chen, J.Q.; Tian, D. (2006) Genome-wide investigation on the genetic variations of rice disease resistance genes. *Plant Mol. Biol*, *62*, 181–193.
- Yang, S.; Zhang, X.; Yue, J.X.; Tian, D.; Chen, J.Q. (2008) Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Genet. Genomics*, *280*, 187–198.
- Yi, H.; Richards, E.J. (2007) A cluster of disease resistance genes in *Arabidopsis* is coordinately regulated by transcriptional activation and RNA silencing. *Plant Cell*, *19*, 2929–2939.
- Yim N, Ha do T, Trung TN, Kim JP, Lee S, Na M et al (2010) The antimicrobial activity of compounds from the leaf and stem of *Vitis amurensis* against two oral pathogens. *Bioorg Med Chem Lett* *20*:1165–1168
- Young, N.D. (2000) The genetic architecture of resistance. *Curr. Opin. Plant Biol*, *3*, 285–290.
- Yu, L.X.; Lorenz, A.; Rutkoski, J.; Singh, R.P.; Bhavani, S.; Huerta-Espino, J.; Sorrells, M.E. (2011) Association mapping and gene–gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theor. Appl. Genet*, *123*, 1257–1268.
- Yu, Y. G., Buss, G. R., & Maroof, M. A. (1996). Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. *Proceedings of the National Academy of Sciences*, *93*(21), 11751-11756.
- Yue, J.X.; Meyers, B.C.; Chen, J.Q.; Tian, D.; Yang, S. (2012) Tracing the origin and evolutionary history of plant NBS-LRR genes. *New Phytol*, *193*, 1049–1063.
- Zhang, X.C.; Gassmann, W.(2003) RPS4-mediated disease resistance requires the combined presence of RPS4 transcripts with full-length and truncated open reading frames. *Plant Cell*, *15*, 2333–2342.
- Zhang, Y.; Goritschnig, S.; Dong, X.; Li, X. (2003) A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of npr1-1, constitutive 1. *Plant Cell*, *15*, 2636–2646.

- Zhao H., Shimohata T., Wang J. Q., Sun G., Schaal D. W., Sapolsky R. M. and Steinberg G. K. (2005) Akt contributes to neuroprotection by hypothermia against cerebral ischemia in rats. *J. Neurosci.* 25, 9794–9806.
- Zecca G, Richard Abbott J, Sun W-B, Spada A, Sala F, Grassi F (2012) The timing and the mode of evolution of wild grapes (*Vitis*). *Mol Phylogenet Evol* 62:736–747
- Zipfel, C., & Felix, G. (2005). Plants and animals: a different taste for microbes?. *Current opinion in plant biology*, 8(4), 353-360.
- Zohary D, Hopf M. (2000). Domestication of plants in the Old World, 3rd edn. New York: Oxford University Press, 151 – 159.
- Zohary, D., Hopf, M., & Weiss, E. (2012). Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. *Oxford University Press*.

# *Annexes*

**Annexe 1. Classification et distribution des séquences répétées de plantes connues dans les séquences d'extrémités de BAC**

Enzymes de restrictions	<i>V. VINIFERA</i> cv CAB. SAUVIGNON		<i>M. ROTUNDIFOLIA</i> cv REGALE		<i>V. VINIFERA</i> cv PN40024*
	<i>EcoRI</i> (%)	<i>HindIII</i> (%)	<i>BamHI</i> (%)	<i>HindIII</i> (%)	Genome entier (%)
<b>RETROTRANSPOSONS DE CLASSE I</b>	<b>24,77</b>	<b>30,94</b>	<b>21,73</b>	<b>23,37</b>	<b>18,81</b>
<b>Rétrotransposons non-LTR</b>	<b>2,95</b>	<b>3,82</b>	<b>4,1</b>	<b>3,88</b>	<b>3,8</b>
SINEs	0,01	0,01	0,01	0,01	
LINEs	2,94	3,81	4,09	3,87	3,8
RTE/Bov-B	0,01	0,02	0,02	0,03	
L1/CIN4	2,93	3,79	4,07	3,84	
<b>Rétrotransposons à LTR</b>	<b>21,83</b>	<b>27,12</b>	<b>17,63</b>	<b>19,49</b>	<b>15,03</b>
Ty1/Copia	7,81	<b>9,4</b>	8,11	<b>8,42</b>	6,64
Gypsy/DIRS1	13,09	<b>16,61</b>	9,07	<b>10,33</b>	8,09
<b>TRANSPOSONS A ADN DE CLASSE II</b>	<b>3,91</b>	<b>3,26</b>	<b>3,54</b>	<b>3,25</b>	<b>5,42</b>
hobo-Activator	0,98	0,75	0,78	0,78	1,11
Tc1-IS630-Pogo	0,02	0,01	0,01	0,01	
En-Spm	0	0,01	0	0	0,89
MuDR-IS905	0,02	0,01	0,02	0,01	2,07
Tourist/Harbinger	0,97	<b>0,79</b>	0,97	<b>1,01</b>	1,34
<b>NON-CLASSIFIES</b>	<b>0,17</b>	<b>0,16</b>	<b>0,14</b>	<b>0,2</b>	
<b>ELEMENTS TRANSPOSABLES TOTAUX</b>	<b>28,85</b>	<b>34,36</b>	<b>25,42</b>	<b>26,82</b>	<b>24,24</b>
<b>PETITS ARNs</b>	<b>0,79</b>	<b>0,26</b>	<b>1,21</b>	<b>0,12</b>	<b>0,01</b>
<b>ADN SATELLITES</b>	<b>0,18</b>	<b>0,49</b>	<b>0,54</b>	<b>0,4</b>	
<b>SIMPLES REPETITIONS</b>	<b>0,02</b>	<b>0,03</b>	<b>0,02</b>	<b>0,02</b>	<b>0,71</b>
<b>TOTAUX PAR SOUS-BANQUES</b>	<b>29,81</b>	<b>35,13</b>	<b>27,17</b>	<b>27,34</b>	
<b>TOTAUX</b>	<b>36,32</b>		<b>30,37</b>		<b>29,26</b>

\* Données sur le génome entier de *Vitis vinifera* d'après Giannuzzi *et al.* 2011

**Annexe 2. Motifs MEME d'acides aminés spécifiques du domaine NBS**

Motif	Séquences	Domaine	Groupe	Similaire	Référence
1	YDVFLSFRGEDTRNNFTDHL Y TALCQAGIHTFRDD	TIR	TNL	TIR-1	Köhler <i>et al.</i> 2008
2	ELPRGEEISPHLWKAIQESRISIIVFSKDYASPWCLDELVKI	TIR	TNL	TIR-2	Köhler <i>et al.</i> 2008
3	CKK <sub>x</sub> TGQIVLPVFDVDPDVRKQTGSFA	TIR	TNL	TIR-3	Köhler <i>et al.</i> 2008
4	VQRWRDALTEAANLSGWD	TIR	TNL	TIR-4	Köhler <i>et al.</i> 2008
5	VGIYGMGGIGKTTIAKVVYNQIR <sub>x</sub> Q	NB-ARC	TNL	P-loop	Köhler <i>et al.</i> 2008
6	IPIVGMGGVGKTTLAQLVYND	NB-ARC	CNL	P-loop	Köhler <i>et al.</i> 2008
7	STIGIYGMGGVGKTTLLQHIHNELLQRPD	NB-ARC	BNL	P-loop	Köhler <i>et al.</i> 2008
8	VRLQEQLLSDILKER	NB-ARC	TNL	RNBS-A	Köhler <i>et al.</i> 2008
9	KIWVCVSQDFDV <sub>x</sub> RIQKAILE	NB-ARC	CNL	RNBS-A	Köhler <i>et al.</i> 2008
10	HVYWVTVSQDFSNRLQNLIAKRLGLDLS	NB-ARC	BNL	RNBS-A	Köhler <i>et al.</i> 2008
11	IKDRLCRKKVLIVLDDV	NB-ARC	TNL	Kinase-2	Köhler <i>et al.</i> 2008
12	LKGKKFLLVLDDVWN	NB-ARC	CNL	Kinase-2	Köhler <i>et al.</i> 2008
13	RKKQKWILILDDLWNN	NB-ARC	BNL	Kinase-2	Köhler <i>et al.</i> 2008
14	WFGPGSRIITTRDK	NB-ARC	TNL	RNBS-B	Köhler <i>et al.</i> 2008
15	DGARGSKILVTRNE <sub>x</sub> VA	NB-ARC	CNL	RNBS-B	Köhler <i>et al.</i> 2008
16	VPLKGCKLIMTTRSERVC	NB-ARC	BNL	RNBS-B	Köhler <i>et al.</i> 2008
17	YEVEELNDDDALQLFSWHAFKNDQP	NB-ARC	TNL	RNBS-C	Köhler <i>et al.</i> 2008
18	VH <sub>x</sub> LG <sub>x</sub> LSEEDCWSLFFK <sub>x</sub> AF	NB-ARC	CNL	RNBS-C	Köhler <i>et al.</i> 2008
19	PLSEGEAWTLFMEKLGHDIAL	NB-ARC	BNL	RNBS-C	Köhler <i>et al.</i> 2008
20	VVDYANGLPLALEVLGSSLYGR	NB-ARC	TNL	GLPL	Köhler <i>et al.</i> 2008
21	EIAKKCKGLPLAAKTLGGLLR	NB-ARC	CNL	GLPL	Köhler <i>et al.</i> 2008
22	RIAKDVARECAGLPLGIITMAGSLRGV	NB-ARC	BNL	GLPL	Köhler <i>et al.</i> 2008
23	EEKKIFLDIACFFIGMKKDYV	NB-ARC	TNL	RNBS-D	Köhler <i>et al.</i> 2008
24	RCFAYCAIFPKDYEI <sub>x</sub> KEELVRLWMAEGF	NB-ARC	CNL	RNBS-D	Köhler <i>et al.</i> 2008
25	QQCLLYCALFPEDHRIEREELIGYLIDEG	NB-ARC	BNL	RNBS-D	Köhler <i>et al.</i> 2008
26	IVSDDKLWMHDLLRDMGREIV	Linker	TNL	MHDV	Köhler <i>et al.</i> 2008
27	FKMHDLVHDLAQSIA	Linker	CNL	MHDV	Köhler <i>et al.</i> 2008
28	KMHDLIRDMAIQILQENSQGM	Linker	BNL	MHDV	Köhler <i>et al.</i> 2008
29	DVFPSFRGEDVRKTFLSHLLKEF	TIR	TNL	TIR-1	Meyers <i>et al.</i> 2003
30	IGPELIQAIRESRIAIVLSKNYASSWCLDELVEIMKC	TIR	TNL	TIR-2	Meyers <i>et al.</i> 2003
31	ELGQIVMPIFYGVDPDVRKQ	TIR	TNL	TIR-3	Meyers <i>et al.</i> 2003
32	WRKALTDVANIAGEHS	TIR	TNL	TIR-4	Meyers <i>et al.</i> 2003
33	VGIWGPPGIGKTTIARALF	NB-ARC	CNL/TNL	P-loop	Meyers <i>et al.</i> 2003
34	VGIYGMGGVGKTTLARQIF	NB-ARC	CNL/TNL	P-loop	Meyers <i>et al.</i> 2003
35	DYGMKLHLQEQLSEILNQDKIKI <sub>x</sub> HLGV	NB-ARC	TNL	RNBS-A	Meyers <i>et al.</i> 2003
36	VK <sub>x</sub> GFDIVIWVVVSQEFTLKKIQDILEK	NB-ARC	CNL	RNBS-A	Meyers <i>et al.</i> 2003
37	RLKDKKVLIVLDDVD	NB-ARC	CNL/TNL	Kinase-2	Meyers <i>et al.</i> 2003
38	KRFLVLDDIW	NB-ARC	CNL/TNL	Kinase-2	Meyers <i>et al.</i> 2003
39	QLDALAGET <sub>x</sub> WFGPGSRIIVTTEDK	NB-ARC	CNL/TNL	RNBS-B	Meyers <i>et al.</i> 2003
40	NGCKVLFTTRSEEVC	NB-ARC	CNL/TNL	RNBS-B	Meyers <i>et al.</i> 2003
41	NHIYEV <sub>x</sub> FPS <sub>x</sub> EEALQIFCQYAFGQNSPP	NB-ARC	CNL/TNL	RNBS-C	Meyers <i>et al.</i> 2003
42	KVECLTPEEAWELFQRKV	NB-ARC	CNL/TNL	RNBS-C	Meyers <i>et al.</i> 2003
43	EVA <sub>x</sub> LAGGLPLGLKVL	NB-ARC	CNL/TNL	GLPL	Meyers <i>et al.</i> 2003
44	EVAKKCGGLPLALKVI	NB-ARC	CNL/TNL	GLPL	Meyers <i>et al.</i> 2003
45	EDKDLFLHIACFFNG	NB-ARC	TNL	RNBS-D	Meyers <i>et al.</i> 2003
46	CFLYCALFPEDYEI <sub>x</sub> KEKLIDYWIAEGFI	NB-ARC	CNL	RNBS-D	Meyers <i>et al.</i> 2003
47	MHNLLQQLGREIV	Linker	CNL/TNL	MHDV	Meyers <i>et al.</i> 2003
48	VKMHDVVREMAWIA	Linker	CNL/TNL	MHDV	Meyers <i>et al.</i> 2003

**Annexe 3. Liste des gènes de résistances fonctionnels de la famille des NBS-LRR clonés chez les végétaux en date du 13 - 02 - 2013**

<b>Protéines</b>	<b>Classe</b>	<b>Espèces</b>	<b>Agents pathogènes</b>	<b>Références</b>
Bs2	CNL	<i>Capsicum chacoense</i>	<i>Xanthomonas campestris</i>	Tai <i>et al.</i> 2000
Bs4	TNL	<i>Solanum lycopersicum</i>	<i>Xanthomonas campestris</i>	Schornack <i>et al.</i> 2004
Cre1	CNL	<i>Aegilops tauschii</i>	<i>Heterodera avenae</i>	Majnik <i>et al.</i> 2003
Dm3	CNL	<i>Lactuca sativa</i>	<i>Bremia lactucae</i>	Shen <i>et al.</i> 2002
FOM-2	CNL	<i>Cucumis melo</i>	<i>Fusarium oxysporum</i>	Joobeur <i>et al.</i> 2004
Gpa2	CNL	<i>Solanum tuberosum</i>	<i>Globodera pallida</i>	van der Vossen <i>et al.</i> 2000
Gro1.4	TNL	<i>Solanum tuberosum</i>	<i>Globodera rostochiensis</i>	Paal <i>et al.</i> 2004
Hero	CNL	<i>Solanum lycopersicum</i>	<i>Globodera</i>	Ernst <i>et al.</i> 2002
HRT	CNL	<i>Arabidopsis thaliana</i>	'Turnip crinkle virus'	Cooley <i>et al.</i> 2000
I-2	NL	<i>Solanum lycopersicum</i>	<i>Fusarium oxysporum</i>	Simons <i>et al.</i> 1998
KR1	TNL	Glycine max	Soybean mosaic virus	He <i>et al.</i> 2003
L6	TNL	<i>Linum usitatissimum</i>	<i>Melampsora lini</i>	Lawrence <i>et al.</i> 1995
Lr1	CNL	<i>Triticum aestivum</i>	<i>Puccinia triticina</i>	Cloutier <i>et al.</i> 2007
Lr10	CNL	<i>Triticum aestivum</i>	<i>Puccinia triticina</i>	Feuillet <i>et al.</i> 2003
Lr21	CNL	<i>Triticum aestivum</i>	<i>Puccinia triticina</i>	Huang <i>et al.</i> 2009
M	TNL	<i>Linum usitatissimum</i>	<i>Melampsora lini</i>	Anderson <i>et al.</i> 1997
Mi-1.2	CNL	<i>Solanum lycopersicum</i>	<i>Meloidogyne</i>	Milligan <i>et al.</i> 1998
MLA1	CNL	<i>Hordeum vulgare</i>	<i>Blumeria graminis</i>	Jorgensen <i>et al.</i> 1994
MLA10	CNL	<i>Hordeum vulgare</i>	<i>Blumeria graminis</i>	Shen <i>et al.</i> 2007
MLA12	CNL	<i>Hordeum vulgare</i>	<i>Blumeria graminis</i>	Shen <i>et al.</i> 2003
MLA13	CNL	<i>Hordeum vulgare</i>	<i>Blumeria graminis</i>	Halterman <i>et al.</i> 2003
MLA6	CNL	<i>Hordeum vulgare</i>	<i>Blumeria graminis</i>	Halterman <i>et al.</i> 2001
N	TNL	<i>Nicotiana glutinosa</i>	'Tobacco Mosaic Virus'	Whitham <i>et al.</i> 1994
P2	TNL	<i>Linum usitatissimum</i>	<i>Melampsora lini</i>	Doods <i>et al.</i> 2001
Pi-ta	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Bryan <i>et al.</i> 2000
Pi2	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Zhou <i>et al.</i> 2006
Pi36	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Liu <i>et al.</i> 2007
Pi5-1	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Lee <i>et al.</i> 2009
Pi5-2	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Lee <i>et al.</i> 2009
PI8	CNL	<i>Helianthus annuus</i>	<i>Plasmopara halstedii</i>	Radwan <i>et al.</i> 2004
Pi9	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Qu <i>et al.</i> 2006
PIB	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Wang <i>et al.</i> 1999
Pid3	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Shang <i>et al.</i> 2009
Pikm1-TS	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Ashikawa <i>et al.</i> 2008
Pikm2-TS	CNL	<i>Oryza sativa</i>	<i>Magnaporthe grisea</i>	Ashikawa <i>et al.</i> 2009
Pikp-2	CNL	<i>Oryza sativa japonica</i>	<i>Magnaporthe oryzae</i>	Yuan <i>et al.</i> 2011
Pit	CNL	<i>Oryza sativa japonica</i>	<i>Magnaporthe oryzae</i>	Hayashi <i>et al.</i> 2009
Piz-t	CNL	<i>Oryza sativa</i>	<i>Magnaporthe oryzae</i>	Zhou <i>et al.</i> 2006
Pm3	CNL	<i>Triticum aestivum</i>	<i>Blumeria graminis</i>	Yahiaoui <i>et al.</i> 2004
Prf	CNL	<i>Solanum pimpinellifolium</i>	<i>Pseudomonas syringae</i>	Chang <i>et al.</i> 2002
R1	CNL	<i>Solanum demissum</i>	<i>Phytophthora infestans</i>	Ballvora <i>et al.</i> 2002
R3a	NL	<i>Solanum tuberosum</i>	<i>Phytophthora infestans</i>	Huang <i>et al.</i> 2005
RAC1	TNL	<i>Arabidopsis thaliana</i>	<i>Albugo candida</i>	Borhan <i>et al.</i> 2004
RCY1	CNL	<i>Arabidopsis thaliana</i>	'Cucumber Mosaic Virus'	Takahashi <i>et al.</i> 2002
Rdg2a	CNL	<i>Hordeum vulgare</i>	<i>Pyrenophora graminea</i>	Bulgarelli <i>et al.</i> 2010
RLM3	TN	<i>Arabidopsis thaliana</i>	<i>Leptosphaeria maculans</i>	Staal <i>et al.</i> 2008
Rp1-D	CNL	<i>Zea mays</i>	<i>Puccinia sorghi</i>	Collins <i>et al.</i> 1999



Rpi-blb1	CNL	<i>Solanum bulbocastanum</i>	<i>Phytophthora infestans</i>	Wang <i>et al.</i> 2008
Rpi-blb2	CNL	<i>Solanum bulbocastanum</i>	<i>Phytophthora infestans</i>	Van <i>et al.</i> 2005
RPM1	CNL	<i>Arabidopsis thaliana</i>	<i>Pseudomonas syringae</i>	Grant <i>et al.</i> 1995
RPP1	TNL	<i>Arabidopsis thaliana</i>	<i>Peronospora parasitica</i>	Botella <i>et al.</i> 1998
RPP13	CNL	<i>Arabidopsis thaliana</i>	<i>Peronospora parasitica</i>	Bittner Eddy <i>et al.</i> 2000
RPP4	TNL	<i>Arabidopsis thaliana</i>	<i>Peronospora parasitica</i>	Van der Biezen <i>et al.</i> 2002
RPP5	TNL	<i>Arabidopsis thaliana</i>	<i>Peronospora parasitica</i>	Parker <i>et al.</i> 1997
RPP8	CNL	<i>Arabidopsis thaliana</i>	<i>Peronospora parasitica</i>	McDowell <i>et al.</i> 1998
Rps1-k-1	NL	<i>Glycine max</i>	<i>Phytophthora sojae</i>	Gao <i>et al.</i> 2008
Rps1-k-2	NL	<i>Glycine max</i>	<i>Phytophthora sojae</i>	Gao <i>et al.</i> 2005
RPS2	CNL	<i>Arabidopsis thaliana</i>	<i>Pseudomonas syringae</i>	Bent <i>et al.</i> 1994
RPS4	TNL	<i>Arabidopsis thaliana</i>	<i>Pseudomonas syringae</i>	Gassmann <i>et al.</i> 1999
RPS5	CNL	<i>Arabidopsis thaliana</i>	<i>Pseudomonas syringae</i>	Warren <i>et al.</i> 1998
RRS1	TNL	<i>Arabidopsis thaliana</i>	<i>Ralstonia solanacearum</i>	Deslandes <i>et al.</i> 2002
Rx	CNL	<i>Solanum tuberosum</i>	'Potato virus X'	Bendahmane <i>et al.</i> 1999
Rx2	CNL	<i>Solanum acaule</i>	'Potato virus X'	Bendahmane <i>et al.</i> 2000
RY-1	TNL	<i>Solanum tuberosum</i>	'Potato virus Y'	Vidal <i>et al.</i> 2002
SSI4	TNL	<i>Arabidopsis thaliana</i>	<i>Pseudomonas syringae</i>	Shirano <i>et al.</i> 2002
Sw-5	CNL	<i>Solanum lycopersicum</i>	Tospovirus	Prins <i>et al.</i> 2000
Tm-2	CNL	<i>Solanum lycopersicum</i>	'Tomato Mosaic Virus'	Lanfermeijer <i>et al.</i> 2003
Tm-2a	CNL	<i>Solanum lycopersicum</i>	'Tomato Mosaic Virus'	Young <i>et al.</i> 1988
VAT	CNL	<i>Cucumis melo</i>	<i>Aphis gossypii</i>	Carzo <i>et al.</i> 2001
XA1	CNL	<i>Oryza sativa</i>	<i>Xanthomonas oryzae</i>	Yoshimura <i>et al.</i> 1998

**Annexe 4. Identification des séquences d'acides aminés déduites des BES de *M. rotundifolia* cv Regale grâce au motif PFAM spécifique du domaine NB-ARC**

Requête	De	A	E-Value	Score	Alignement (acides aminés)	NB-ARC*	Motifs MEME**	CLUSTALX
MRRE1B072K02RM1	1	155	7,5.10 <sup>-31</sup>	113,196	155	CNL	+	CNL
MRRE1B074K04FM1	1	173	2,15.10 <sup>-35</sup>	125,907	173	+	CNL	CNL
MRRE1B074M12RM1	1	152	1,21.10 <sup>-41</sup>	141,7	152	+	+	CNL
MRRE1B074P07RM1	1	237	3,78.10 <sup>-64</sup>	203,332	237	+	CNL	CNL
MRRE1B077G16FM1	1	180	1,39.10 <sup>-31</sup>	116,277	180	+	+	CNL
MRRE1B078A14RM1	2	181	3,15.10 <sup>-24</sup>	95,8616	180	+	TNL	TNL
MRRE1B078N09FM1	1	211	1,4.10 <sup>-63</sup>	201,021	211	+	CNL	CNL
MRRE1B079K20RM1	1	233	1,19.10 <sup>-32</sup>	120,9	233	+	+	TNL
MRRE1B080I05RM1	9	203	1,55.10 <sup>-58</sup>	187,539	195	+	CNL	CNL
MRRE1B080M12FM1	1	162	1,04.10 <sup>-26</sup>	102,41	162	+	TNL	TNL
MRRE1B082A09RM1	1	214	6,05.10 <sup>-44</sup>	150,175	214	+	CNL	CNL
MRRE1B082B07RM1	1	159	1,95.10 <sup>-28</sup>	107,032	159	+	CNL	CNL
MRRE1B083E15RM1	1	177	1,29.10 <sup>-25</sup>	100,099	177	+	TNL	TNL
MRRE1B085L04RM1	1	164	1,58.10 <sup>-42</sup>	144,782	164	+	CNL	CNL
MRRE1B092L18FM1	1	189	3,32.10 <sup>-36</sup>	128,604	189	+	+	CNL
MRRE1B093M15FM1	1	204	9,38.10 <sup>-55</sup>	178,294	204	+	+	CNL
MRRE1B094B02RM1	1	170	2,52.10 <sup>-48</sup>	159,805	170	+	CNL	CNL
MRRE1B096G15RM1	16	177	8,5.10 <sup>-22</sup>	88,928	162	+	+	CNL
MRRE1B096O09FM1	1	218	1,82.10 <sup>-45</sup>	154,412	218	+	CNL	CNL
MRRE1B097I17FM1	1	244	5,08.10 <sup>-67</sup>	211,036	244	+	CNL	CNL
MRRE1B097O11FM1	1	186	1,04.10 <sup>-34</sup>	124,752	186	+	+	CNL
MRRE1B099B11RM1	1	158	1,5.10 <sup>-31</sup>	115,507	158	+	CNL	CNL
MRRE1B099M22RM1	5	178	4,93.10 <sup>-27</sup>	103,566	174	+	CNL	CNL
MRRE1B100L22FM1	1	190	1,2.10 <sup>-52</sup>	171,746	190	+	CNL	CNL
MRRE1B100N09RM1	1	224	7,93.10 <sup>-36</sup>	128,989	224	+	CNL	CNL
MRRE1B104P15FM1	1	169	2,3.10 <sup>-38</sup>	133,996	169	+	CNL	CNL
MRRE1B105G10RM1	1	214	6,53.10 <sup>-30</sup>	112,81	214	+	+	TNL
MRRE1H001C19RM1	2	167	8,39.10 <sup>-40</sup>	137,463	166	+	CNL	CNL
MRRE1H001D21RM1	1	234	6,29.10 <sup>-63</sup>	200,251	234	+	CNL	CNL
MRRE1H003B02FM1	1	159	1,83.10 <sup>-20</sup>	84,6908	159	+	TNL	TNL
MRRE1H003K20FM1	1	205	2,65.10 <sup>-34</sup>	124,366	205	+	+	TNL
MRRE1H003P13RM2	1	158	1,19.10 <sup>-43</sup>	147,093	158	+	CNL	CNL
MRRE1H011A23RM1	1	175	1,03.10 <sup>-33</sup>	121,67	175	+	CNL	CNL
MRRE1H011M12FM1	10	167	3,45.10 <sup>-27</sup>	104,336	158	+	+	CNL
MRRE1H013B10FM1	1	165	3,61.10 <sup>-37</sup>	130,53	165	+	+	CNL
MRRE1H014G21FM1	1	196	1,96.10 <sup>-23</sup>	94,3208	196	CNL	-	CNL
MRRE1H015C01RM1	1	152	4,16.10 <sup>-23</sup>	92,0096	152	CNL	CNL	CNL
MRRE1H016J12RM1	1	187	1,63.10 <sup>-32</sup>	118,974	187	+	+	CNL
MRRE1H019E10FM1	1	155	1,2.10 <sup>-21</sup>	87,7724	155	+	+	TNL
MRRE1H020A06FM1	1	216	7,51.10 <sup>-67</sup>	209,495	216	+	CNL	CNL
MRRE1H023J14FM1	1	227	1,33.10 <sup>-59</sup>	191,391	227	+	CNL	CNL
MRRE1H024E01RM1	1	192	1,37.10 <sup>-28</sup>	108,573	192	CNL	CNL	CNL
MRRE1H025C13FM1	17	197	7,46.10 <sup>-08</sup>	49,6377	181	+	+	CNL
MRRE1H026E18FM1	1	164	5,93.10 <sup>-48</sup>	158,649	164	+	CNL	CNL
MRRE1H033G17FM1	5	161	1,37.10 <sup>-18</sup>	79,298	157	+	+	TNL
MRRE1H035P12RM1	2	167	2,64.10 <sup>-39</sup>	136,308	166	+	CNL	CNL

MRRE1H041A13RM1	1	186	$2,71.10^{-39}$	136,693	186	+	CNL	CNL
MRRE2H044C01FM1	1	166	$1,91.10^{-23}$	93,1652	166	+	CNL	CNL
MRRE2H044F06RM1	1	168	$1,83.10^{-52}$	170,59	168	+	CNL	CNL
MRRE2H049O07RM1	1	218	$2,64.10^{-24}$	97,4024	218	+	-	CNL
MRRE2H049P23RM1	2	169	$2,5.10^{-37}$	130,915	168	+	CNL	CNL
MRRE2H050H11RM1	1	192	$6,59.10^{-53}$	172,516	192	+	CNL	CNL
MRRE2H050I14FM1	1	155	$2,42.10^{-29}$	109,344	155	+	+	CNL
MRRE2H050I15RM1	1	253	$2,45.10^{-38}$	136,693	253	+	+	CNL
MRRE2H052H10FM1	1	169	$1,15.10^{-40}$	139,774	169	+	CNL	CNL
MRRE2H053A05FM1	1	190	$5,29.10^{-57}$	182,917	190	+	CNL	CNL
MRRE2H054D16RM1	1	168	$1,2.10^{-46}$	155,568	168	+	CNL	CNL
MRRE2H055E23FM1	1	210	$1,74.10^{-27}$	105,877	210	+	+	CNL
MRRE2H055E23RM1	5	225	$9,25.10^{-39}$	136,693	221	+	+	CNL
MRRE2H055N23RM1	1	237	$3,88.10^{-66}$	208,725	237	+	+	CNL
MRRE2H060L08FM1	1	153	$4,38.10^{-20}$	83,15	153	+	+	CNL
MRRE2H062E15RM1	1	176	$1,41.10^{-58}$	186,769	176	+	CNL	CNL
MRRE2H062F09FM1	1	178	$2,13.10^{-31}$	115,507	178	+	CNL	CNL
MRRE2H065I02FM1	1	189	$9,66.10^{-57}$	182,532	189	+	CNL	CNL
MRRE2H066C20FM2	1	168	$1,68.10^{-44}$	149,79	168	+	CNL	CNL
MRRE2H070G10RM1	17	244	$8,9.10^{-42}$	145,552	228	+	CNL	CNL
MRRE3H107N21FM1	2	169	$1,6.10^{-48}$	160,19	168	+	CNL	CNL
MRRE3H109G11RM1	1	228	$7,47.10^{-59}$	189,465	228	+	CNL	CNL
MRRE3H112B21FM1	1	176	$3,6.10^{-57}$	182,917	176	+	CNL	CNL
MRRE3H113C24RM1	1	201	$2,48.10^{-37}$	132,07	201	+	+	CNL
MRRE3H115B13FM1	1	199	$1,21.10^{-24}$	97,7876	199	+	+	TNL
MRRE3H119J22RM1	1	199	$3,5.10^{-36}$	128,989	199	+	CNL	CNL
MRRE3H126G11FM1	1	159	$1,46.10^{-40}$	139,004	159	+	+	CNL
MRRE3H129H21FM1	11	167	$2,23.10^{-23}$	93,1652	157	+	CNL	CNL
MRRE3H131F19RM1	1	169	$1,68.10^{-46}$	155,182	169	+	CNL	CNL
MRRE3H134H17RM1	1	169	$3,09.10^{-27}$	103,951	169	+	TNL	TNL
MRRE3H135A05FM1	1	171	$6,3.10^{-46}$	153,642	171	+	CNL	CNL
MRRE3H141G19RM1	1	220	$1,44.10^{-26}$	103,566	220	+	TNL	TNL

\* BES identifiées (+) ou non (-) avec les domaines NB-ARC des 78 protéines fonctionnelles

\*\* BES identifiées (+) ou non (-) avec les motifs MEME spécifiques des TNL ou CNL

**Annexe 5. Alignement des BES de *M. rotundifolia* cv Regale avec des modèles de gènes à motif NB-ARC**

<b>BES de <i>M. rotundifolia</i> cv Regale</b>	<b>Coordonnées de l'alignement sur la séquence de <i>V. vinifera</i> cv PN40024</b>	<b>Motifs présents sur l'intervalle de séquence de <i>V. vinifera</i> cv PN40024</b>
MRRE1H015L01FM1	chr1:812812..812547	NB-ARC: 812008..816090
MRRE2H045A07FM2	chr1:865242..865860	NB-ARC:865318..867702
MRRE3H130G17FM1	chr1:872788..873174	NB-ARC: 872585..888273
MRRE2H046A22FM1	chr3:1046148..1046006	NB-ARC: 1044307..1047456
MRRE2H055N23RM1	chr3:1077915..1078512	NB-ARC: 1077896..1078534
MRRE1H027J23FM1	chr3:16159668..16160112	NB-ARC:16159523..16162175
MRRE1B074M12RM1	chr3:979296..979705	NB-ARC: 978958..982622
MRRE1B077G16FM1	chr3:980085..979702	NB-ARC: 978958..982622
MRRE2H043J06RM1	chr5:10575993..10576629	NB-ARC:10572519..10577127
MRRE1H023L06RM1	chr6:13014491..13015130	NB-ARC:13009732..13016430
MRRE1H017F02FM1	chr7:11039096..11039708	NB-ARC:11038146..11040905
MRRE1H018P15RM1	chr7:11039096..11039841	NB-ARC:1038146..11040905
MRRE2H062A23RM1	chr7:11095806..11096494	NB-ARC:11093802..11097518
MRRE1H003L16FM1	chr7:11130636..11130065	NB-ARC: 11128916..11131720
MRRE1H025C13FM1	chr7:14610839..14610078	NB-ARC: 14608192..14612028
MRRE3H106N13RM1	chr9:12490165..12490823	NB-ARC:12490597..12491130
MRRE1B091E18FM1	chr9:3689092..3689634	NB-ARC:3688933..3691944
MRRE2H062E15RM1	chr9:3690255..3689697	NB-ARC:3688933..3691944
MRRE3H112B21FM1	chr9:3690286..3689697	NB-ARC: 3688933..3691944
MRRE1H026E03FM1	chr9:3690765..3690060	NB-ARC: 3688933..3691944
MRRE2H064B18FM1	chr9:3690765..3690085	NB-ARC: 3688933..3691944
MRRE2H065M12FM1	chr9:3690765..3690114	NB-ARC: 3688933..3691944
MRRE1B080I05RM1	chr9:4723833..4723091	NB-ARC:4722218..4726125
MRRE1H003P13RM2	chr9:5643861..5643099	NB-ARC:5642255..5645259
MRRE2H063E21RM1	chr9:5692015..5691586	NB-ARC: 5688927..5697812
MRRE1B085E03FM1	chr10:13368550..13369062	NB-ARC:13357268..13369931
MRRE1B097O11FM1	chr10:13368550..13369247	NB-ARC:13357268..13369931
MRRE2H070G10RM1	chr11:17453431..17454234	NB-ARC:17451588..17454789
MRRE1H030A22FM1	chr12:16411310..16411992	Kinase:16375557..16417330
MRRE2H049P23RM1	chr12:17549297..17550050	NB-ARC:17549019..17552648
MRRE1H039F22FM1	chr12:17549309..17549996	NB-ARC:17549019..17552648
MRRE3H131F19RM1	chr12:17811541..17812287	<i>Nucleotide Binding</i> (GO):17812003..17813997
MRRE2H044C01FM1	chr12:19890042..19890645	ATP binding:19889093..19890495
MRRE3H132I23FM1	chr12:21051602..21052068	<i>Nucleotide Binding</i> (GO): 21046696..21054369
MRRE1H003M20FM1	chr12:21649813..21649446	NB-ARC:21647419..21655095
MRRE1H035A06FM1	chr12:21649875..21649268	NB-ARC: 21647419..21655095
MRRE1H019N20FM1	chr12:21649886..21649399	NB-ARC:21647419..21655095
MRRE1H003K20FM1	chr13:11134956..11150436	TIR:11134956..11150436
MRRE1H016F05FM1	chr13:18566240..18566928	NB-ARC:18564206..18567244
MRRE2H067E15RM1	chr13:19948495..19947849	NB-ARC:19944625..19949505
MRRE2H066C20FM2	chr13:20456314..20455525	NB-ARC:20452421..20460206
MRRE1B097I17FM1	chr13:20465769..20466639	NB-ARC:20462915..20467293
MRRE3H124I19RM1	chr13:21068330..21067512	NB-ARC:21065902..21069368
MRRE1H001C19RM1	chr13:21068332..21069014	NB-ARC:21065902..21069368
MRRE1H035P12RM1	chr13:21068332..21069028	NB-ARC:21065902..21069368
MRRE1B100L22FM1	chr13:21081066..21081729	NB-ARC:21080351..21081799

MRRE1B100L22RM1	chr13:21158133..21157420	NB-ARC: 21154357..21160200
MRRE1H032K04RM1	chr13:21573542..21572887	NB-ARC: 21571875..21576340
MRRE1B080L16FM1	chr13:21604684..21603993	NB-ARC: 21601517..21610881
MRRE3H123O13FM1	chr13:466573..465718	NB-ARC:463533..468515
MRRE1H021P22FM1	chr14:12106982..12106681	NB-ARC:2105751..12118092
MRRE2H054G12FM1	chr14:12464203..12464852	NB-ARC:12463695..12466138
MRRE2H053E16FM1	chr14:12464203..12464891	NB-ARC:12463695..12466138
MRRE1H039L02FM1	chr14:5257961..5258380	NB-ARC:5254285..5275694
MRRE3H126A14RM1	chr14:5672751..5672074	NB-ARC:5671014..5674017
MRRE2H055E23FM1	chr14:9279512..9280309	NB-ARC:9273608..9280327
MRRE2H064P03FM1	chr15:19528194..19528837	NB-ARC:19528135..19530195
MRRE1B075C01FM1	chr15:3728979..3732547	NB-ARC: 3728979..3732547
MRRE2H044F06RM1	chr15:5930122..5930771	<i>Nucleotide binding (GO):</i> 5929307..5931886
MRRE1H035F06RM1	chr16:13303539..13304285	NB-ARC:13303988..13305889
MRRE3H130C08RM1	chr18:26581908..26581235	NB-ARC: 26576660..26585729
MRRE3H124H08FM1	chr18:28316070..28316482	TIR:28305735..28320509
MRRE1H026G09RM1	chr19:21812354..21813102	<i>Nucleotide binding (GO):</i> 21811437..21812969
MRRE1B090N04FM1	chr19:21841212..21841643	NB-ARC:21839338..21841479
MRRE1B083L06FM1	chr19:22805589..22806362	NB-ARC:22805628..22809970
MRRE2H064G14RM1	chr19:908286..908987	NB-ARC: 897695..916258
MRRE1H020K13FM1	chr19:908338..908944	NB-ARC:897695..916258
MRRE1B082A09RM1	chrUn:14469306..14468642	<i>Nucleotide binding (GO):</i> 14436373..14476776
MRRE2H061G19RM1	chrUn:14813993..14813418	NB-ARC: 14811601..14815280
MRRE1B074K04FM1	chrUn:14859674..14858903	NB-ARC:14856735..14860814
MRRE1H033G17FM1	chrUn:16216919..16217511	TIR/NB-ARC:16213588..16218226
MRRE3H126G11FM1	chrUn:31019675..31020398	NB-ARC:31016731..31020673
MRRE1H019E10FM1	chrUn:5113539..5113029	<i>Protein binding (GO):</i> 5110267..5115965
MRRE1H032N01RM1	chrUn:8551638..8550889	TIR/NB-ARC:8548903..8552625

**Annexe 6. Séquences des couples d'amorces définis dans les BES de *M. rotundifolia* cv Regale**

<b>Marqueurs</b>	<b>Amorce <i>Forward</i></b>	<b>Amorce <i>Reverse</i></b>
MRZ0030A	GTGCGAAACCATAGGAAG	CTGGAAGAAACGTGAGGA
MRZ0031A	TGGGAGCAGTAGTGATGAAG	GAAAACAGGTCTCTCTCAGG
MRZ0032C	TGTCCATGCCAGCTTAGT	TCTGGTGAGCACATACCC
MRZ0033A	TGCATGACTACCTCTTTCT	ACCCTACTCCTAACCTCTGG
MRZ0034A	GCTCATTCGCAGACTACAC	GACAAACCGGGAAGAAAG
MRZ0036D	TGAGAGCATCTTGGTTTAGG	AAAAAGTTGAAAAGTCAGAAGG
MRZ0040A	CAGTTGAGAACCTTGACCTG	GGTTTGAAGGGGTTTGAC
MRZ0041A	CCTTGCTGAGGTAATTGG	GTACGGAGGTGAATTGTCTG
MRZ0042D	ATCCACATGCTTGCTCAC	CAGTATCAATCACGCCAGAG
MRZ0046D	ACCACTTCAAACCTTTTCTC	TGGTATTTGGGATTTGTCTG
MRZ0047A	CTCCACAAAACCTCTCCCTCT	CCATTAGGTCTGGATGCT
MRZ0048A	AGCACACTTCCAATATGACC	CATCTCCAGCCTAATTCTTG
MRZ0049A	CCATTTTCTTGCGTATGTTT	TTCCTAACTCCTCAACTCCA
MRZ0050A	CTTTTGATGTGTTTGCTGAA	GCAATAAGTGGAAGCATTGT
MRZ0051A	TGACAGCATGATAAATGGAG	GTCTGGTTTGGTGGAGTATG
MRZ0052A	CGTTTTGGACTAACTTCCAG	CTGTCATGTGGTGAATGTGT
MRZ0052B	AGTTGAGAACCTTGACCTGA	GTTAGTCCAAAACGAGCTGA
MRZ0053A	ATACTGAACCAATGCCAAAG	CCGCAACAACCATCTATTAC
MRZ0053B	AGGTTATGGAGACAAGTGG	CGTTTGCCTATGTTGCTATT
MRZ0053C	ATGCTATGTTTCCTTTGTCC	TTTGCCTATGTTGCTATTCA
MRZ0053D	TGAATAGCAACATAGGCAAA	GTCACCTCAGCACTCATCTC
MRZ0054A	TAAGAGCCTTCATCTCCTTG	AATTTCCAATCATGTGTCCT
MRZ0055A	GGTAGCTTTACCATCTTCCA	ACGGTCTGGATTTTCTAGTTT
MRZ0056A	AACTCAAACACCAAGAAAGG	CTTCACGCTTTAGGAACATC
MRZ0057A	GGAAACATCTGAAACCAAGA	ATGTCAAAGGGAATCACAAG
MRZ0057B	GAAACCAAGAACAAGAATGG	ATGTCAAAGGGAATCACAAG
MRZ0058A	CTTGGGCTAAATCCCTAAAA	AATAAATTGCACCATCCAAC
MRZ0058B	ACACTGGATCAAAGAGATGTG	TCTTCCTATTGCTGCTTTTT
MRZ0059A	ATAGCATTCCATGAGACCAC	AGATGCCACTGAGAAATGAG
MRZ0059B	TCACATCACCAATCCTTACA	GGATATGCAACAAATGGAGA
MRZ0060A	CTATCTTTGGCCTCTTTTGA	ATTGGTCTTTTATGCCTTTT
MRZ0061A	TGGGAGTTGAGATAGACTTGA	CTTTGAAGGTGGAGAAAATG
MRZ0062A	CGGGGTAACCTCTGACAATAA	TTCACTGCTATCACCTGTTC
MRZ0063A	TGTGCATTTCTTACATTTGG	AACCAGAGGAGCCTAAAATC
MRZ0064A	AATTAGCATCCACATTGACC	CTTGCTTCATTTTCGCTATT
MRZ0064B	AATAGCGAAAATGAAGCAAG	AAGAAAACAAGGACAGCAAA
MRZ0065A	ACAAAAGCACAAAATCTCC	TGGTTGGATCTGTGTATGAA
MRZ0066A	GAGCTTTTTCCCTTTTACC	TGGTTTCTTCTTCTCGTTCT
MRZ0067A	CCTCACTGGTAATTCGTCTC	ACACAGCTTGAACAGAAAAAG
MRZ0068A	TTGTTGGCTAGGAAGAACAC	TCCACCTGCATACATACAAA
MRZ0068B	AAGAACACGGACATCATTTT	CGAGGCTGAGGAGTTAATAG



MRZ0068C	ATTGTTGGCTAGGAAGAACA	AATCAAAACATAACAGGCAAA
MRZ0069A	GAGTTGAGAGCATCTTGTT	AAAAGTTGAAAAGTCAGAAGGA
MRZ0070A	CCACTCCATGACAATATCC	GATATGAAGGGAGGAGAGATG
MRZ0071A	CCATAAGTGGGTAAATCTGG	TGAATAAGCCTCCAATTCTT
MRZ0072A	TAAATACTAGCAGGTTTGAGTGG	GTGGTAGTTTTAATTGTTGTGG
MRZ0073A	TCACATAACAACCCGACAAG	GGATCTCCTGAAAAATGACT
MRZ0074A	TTGCATAAGGCTAGAAATCC	AAACTAGATCATGGCCCTTT
MRZ0075A	AAGCAATATACCCAAACCAA	ATTTACAAGGGATGAAAGCA
MRZ0076A	GAATCAACCATACCCATGTC	CTTGGGGTTTCTCTCTCTTT
MRZ0076B	ATGAGGATGTTGATGATGGT	TGTGAACTAGGGTTTCAAGG
MRZ0076C	CTTGGTTCATCCTCATCACT	ATTCACCATCATCAACATCC
MRZ0077A	GAAAACCTCTTCTCTTGGT	GTTTAAGCTCCCAATCTATCC
MRZ0078A	CCCAAGCAAAAATAAAGACA	ACTCAAAGAAAGGGAGGTTC
MRZ0079A	TGCGTATATGATGAGGTCTTT	CTCAAGTTACCCAAGGAGTG
MRZ0079B	TGCGTATATGATGAGGTCTTT	TGAATTTGGTCTCCTTCAAC
MRZ0080A	TGTGTTATCAAAGGCCATAA	TTTCTCCATATTCCAAGTCAG
MRZ0081A	TCACTTTCATATCAGCCAAA	GAATAATGGGAAGAAAGCAG
MRZ0082A	GGGTGCTCTACAAAAGAAAA	CCTGAAAAACTGACCTTTGA
MRZ0083A	CCCTTGTTGGTATGTTATTGG	CATAGGTGGTTTCTTTTGCT
MRZ0084A	CAATGAAACCTCACTCCACT	TTGCTCCTAGCTCTATGGTT
MRZ0085A	AGAAAAACCACCAAACCTGAA	ACCCAGCTTTATGACAACAC
MRZ0086A	CTTTTGGACATCCTTTTTGT	GTTCCATTTTTGCACTGATT
MRZ0087A	ATACTTCTTGCTCGCATAA	TACCTGACTTCTTGGGTTC
MRZ0088A	GAAACAAAGAGGAACCAGAA	ATCTCTCCACCAAATTCACA
MRZ0088B	ATGCAACAGCCTACCTAAAC	TTTAATGGTGGTCAAATGCT
MRZ0089A	GAACTAAGGCCAATCCTACC	ATTAGCAAGTGGCCTACAAA
MRZ0090A	CAATCCCTCATCTTCCTAA	TTCCATCCTCAACTGAACT
MRZ0090B	CACCTTCATAACCCTTACCTT	GAAGCAGAAGAAAGCACACT
MRZ0090C	CAATCATCTCATCCTTCTGG	ATTTCTCTGGTCATGGTGT
MRZ0091A	GTTCCATCCCCATACTTAGA	AAAGAGAAGAGAGTTGGTTGG
MRZ0092A	AAATGCAAATTAAGCCAAG	TGGATCAAAACATCAACAAG
MRZ0093A	ATGCAACCATGAAGGATTT	TTTGGTTCCTTAGTTTCTCT
MRZ0094A	CACCACTACATCTTCAACCA	TGACTTGGAATTTGGAAAAC
MRZ0095A	AGCTCGCTTTATACCTTTCC	TCTAGCAAATGGTGGTTACA
MRZ0095B	TACGCCATAGTGATTCCAAG	GGACTGTTTGTGTGTTACGA
MRZ0095C	TTTAAGAGCCAGATGCAAGA	GGACTGTTTGTGTGTTACGA
MRZ0096A	GAAAGAAATGAATCCAAAAGA	ATGGAAGAAGTTTTGTAAGTGA
MRZ0097A	TAGCTCCACAAGCCATAAGT	AGAAAATTCAAAACCCCTGT
MRZ0097B	ATCATTCTGTGATTTGCTT	CAACTGTCTTCTTTGCTTC
MRZ0098A	TCAAGTGGGAATAAAAATGAA	TTTGAGAAGCAATGTGAAAA
MRZ0099A	GTGTACTTTGCTGTGCCTCT	GTTCCCTCCTGAAATGTTG
MRZ0100A	TCTGAATCTATCTCGCATTTT	GCATCTCTAGCCTCAATTTTT
MRZ0101A	CACAGAAACAAATCAGCAAA	GTTGGACAGATCAGAAATCC
MRZ0102A	GAAACCTTTAACCCATCCTT	TCATTCTCTCCCTGTTTCATC
MRZ0103A	CCTCTTCTCTTTCACCTTT	ATAACCCTCTTCCAACCATC

MRZ0104A	GTAGTAGTGGGTGTGGAGGA	TGCGACCAAGGTTGGGACCA
MRZ0105A	TGTCGGTAGTGGAGTTTTTC	AGATAAATTAGGGCAATGTA
MRZ0106A	GGCATCAACTGCTAAAGAGA	CAAATAGGTTTTCCCATGTA
MRZ0106B	TTTCCTTGTTCCCAAGGT	ACCATCATTAGAAAACATCTCT
MRZ0107A	AGACGAGTTTTATTGTTTTCA	TGCCCTTGATTATGTAGCC
MRZ0107B	AATTTAATGCTAAATAATGAGAC	TTCATTTTTTCATTTTTACCC
MRZ0108A	ACTCCTTGATGTTGCATGTT	ACAATCTGCTCCATTTGTCT
MRZ0108B	TTGATGTTGCATGTTTTCTT	ACAATCTGCTCCATTTGTCT
MRZ0108C	GTGAAAAATTTCCCTAATGAGC	GTTTATATCTTCAACTAGAGCTT
MRZ0109A	TGCCTTATTTGGGGGGGAC	GTGTTGGTAATAGTTGAAACT
MRZ0110A	GGCAGGTATGAGTCCAATAA	ATCCCAACAATACGAACATC
MRZ0111A	GTGCTTAAAAGTAGCTATGATGGA	TCAAATCATGCATACCTATACAGT
MRZ0156A	TGATATGGTGCCTATACGA	ACTCCAGGCTTCTGGTTACCCAA
MRZ0156B	GGAGAAATTCTGTTACCTGAAACTA	AGAAACTTGATGTTGGTATTTCAG
MRZ0157A	ACACCAAATGATGGTGCACCTGAGA	GGAAGATTCAGTACAGCAGTGCC
MRZ0157B	ACGCGAAGACAAAGATGGAGAGGCTCA	GCATCTGCAGAATGAGGATGAGG
MRZ0158A	ACCATGAACTAGGAGTTCAAG	TGTTAGATTACAGACATCAATTTAGG
MRZ0159A	ATCTAGTGTTCAGTAGCTGCCTT	GAGTATACCTCTCTTGCCTCTCTTGC
MRZ0160A	TCTCCACCACCACCACCAATAGC	GGCTCAATGTGGTAGCATCAGGT
MRZ0160B	AGTCTGCAATATTTCAAAGTA	TCCATTCTAACGGGGAAAAATGTT
MRZ0160C	TTGCTCAGGATTCTAAGTAAGT	GACGTGTATGGTGTAGGGATAGT
MRZ0161A	AGCTTTCTCTCCTCGAGAACTACAT	GTTATTAATGAAGTAGGATGAT
MRZ0162A	AGAAATCCTAATTCTAAAGCA	GAACAAAATTTGGATTTTCCATTCTC
MRZ0162B	TGACTGATCTTGATACCCATTCAGG	AGCCCTAAGTCTTCAAATTTGTTCTC
MRZ0163A	ACTGGAAGCAAGGAACTCGGTTTAGC	CAAAGTAGGGCCACGAAAGAGTTGG
MRZ0164A	AGGTTCAGAAGCAGATCCAACCTATC	TTGCCTGCATGATGCTCATTACCAA
MRZ0165A	TTGATTAACCAACTATTGAGAAG	AGCAATTAGCTCCCTTCACATAT
MRZ0166A	GCATAATGTAAAGGGCCATGATCT	GCATATCTCCATTTGTATCCGCCAC
MRZ0166B	GCGTATGGGGACCATATTGAGAGTCC	CCCATACATGTTTTGGTGATAGAGA
MRZ0167A	CAAGAAGTACCTCAACCTCCCTAAATC	TGCCATGCTGAAGATGAAGTACCTAC
MRZ0168A	TCATCCCCTCCCTAGATTCTA	GAAGGCGAAGTAGGCGGCAT
MRZ0168B	AGGGTTTTATACTGTTGTCAGTC	GAGGAGAGCTTCTTGGACTGCTTG
MRZ0169A	GCTGTTGCTCATTTTACTTCTTGA	TACTGACTTGGAATATGGAGAAAG
MRZ0170A	CAATCCTACCCACAGGAGATAAC	CTTAGAGCCTTCAGGATGATCTTTGG
MRZ0171A	GGTCCTCACACCGCTAACAC	GGATCATCAGTCCATCTGGACTTTG
MRZ0171B	GATGGTTCAGCTAAGGGCAGTGAA	GTCATGATTGGTGGGATCAAACCT
MRZ0172A	CATCTTTCCCAATGTAGGTTTCTGG	TGGATGTGGGTGATTTTAATATGTC
MRZ0173A	GTCAACTCAGCTATGAGGGGAACCAT	CCTCTTCTCCCAAAGGCTTCTTTGGT



**Annexe 7. Les couples d'amorces définis dans la région du *locus* de résistance *Run1/Rpv1* avec une localisation unique par EPCR sur le génome de référence de *V. vinifera* cv PN40024. Les marqueurs définis dans les BES alignées sur le supercontig de *M. rotundifolia* cv G52 sont en violet.**

Marqueurs	Coordonnées sur le génome de référence	Taille de l'amplicon (pb)
<b>VMC4F3-1</b>	Chr12 : 13067657..13067837	181
MRZ0144A	Chr12 : 13279386..13279578	193
MRZ0145A	Chr12 : 13340510..13340732	223
MRZ0126A	Chr12 : 13359916..13360106	191
MRZ0125A	Chr12 : 13438457..13438605	149
MRZ0117A	Chr12 : 13556573..13556806	234
MRZ0124B	Chr12 : 14022233..14022411	179
MRZ0124A	Chr12 : 14022606..14022793	188
MRZ0143A	Chr12 : 14051269..14051367	99
MRZ0122A	Chr12 : 14146571..14146645	75
MRZ0132A	Chr12 : 14187027..14187216	190
MRZ0132B	Chr12 : 14187233..14187447	215
MRZ0136A	Chr12 : 14214022..14214178	157
MRZ0133A	Chr12 : 14219273..14219485	213
MRZ0121B*	Chr12 : 14219565..14219731	167
MRZ0121A*	Chr12 : 14219566..14219731	166
MRZ0137A	Chr12 : 14269790..14269992	203
MRZ0154A	Chr12 : 14435732..14435925	194
MRZ0150A	Chr12 : 14436221..14436372	152
MRZ0115A*	Chr12 : 14626749..14626883	135
MRZ0114B	Chr12 : 14701047..14701256	210
MRZ0114A	Chr12 : 14701255..14701443	189
MRZ0141A	Chr12 : 14764469..14764626	158
MRZ0141B	Chr12 : 14764541..14764757	217
MRZ0105A*	Chr12 : 15433396..15433558	163
MRZ0109A*	Chr12 : 15457550..15457763	214
MRZ0104A	Chr12 : 15504729..15504854	126
MRZ0107B	Chr12 : 15528816..15529039	224
MRZ0107A	Chr12 : 15528835..15529012	178
MRZ0106A*	Chr12 : 15696209..15696413	205
MRZ0106B*	Chr12 : 15696375..15696565	191
MRZ0111A	Chr12 : 15750403..15750598	196
MRZ0108A*	Chr12 : 15750451..15750649	199
<b>49MRP1.P2</b>	Chr12 : 16372962..16373169	208
MRZ0110A*	Chr12 : 16674532..16674679	148
MRZ0108C	Chr12 : 16675553..16675686	134
MRZ0138A	Chr12 : 16824123..16824327	205
MRZ0151A	Chr12 : 16825097..16825330	234

MRZ0134A	Chr12 : 17081742..17081934	193
MRZ0135A	Chr12 : 17128398..17128649	252
MRZ0142A	Chr12 : 17362286..17362509	224
MRZ0118A	Chr12 : 17365094..17365264	171
<b>UDV-058</b>	Chr12 : 17434734..17434887	154
MRZ0153A	Chr12 : 18038458..18038619	162
MRZ0152B	Chr12 : 18055204..18055357	154
MRZ0152A	Chr12 : 18055726..18055886	161
MRZ0140A	Chr12 : 18278168..18278327	160
MRZ0139A	Chr12 : 18411257..18411464	208
MRZ0112A	Chr12 : 18502167..18502542	376
MRZ0131A*	Chr12 : 18570990..18571075	86
MRZ0113A*	Chr12 : 18605721..18605894	174
MRZ0123A	Chr12 : 18668490..18668649	160
MRZ0123B	Chr12 : 18668722..18668885	164
MRZ0128A	Chr12 : 19449490..19449649	160
MRZ0149A	Chr12 : 19468888..19469038	151
MRZ0129A	Chr12 : 19498515..19498750	236
MRZ0119A	Chr12 : 19528237..19528405	169
MRZ0120A	Chr12 : 19604178..19604348	171
MRZ0120B	Chr12 : 19604238..19604417	180
MRZ0148A	Chr12 : 19711880..19712047	168
MRZ0148B	Chr12 : 19712003..19712153	151
MRZ0147A	Chr12 : 19728780..19728946	167
MRZ0130A	Chr12 : 20281747..20281920	174
MRZ0146A	Chr12 : 20302569..20302741	173
MRZ0127B	Chr12 : 20307660..20307888	229
MRZ0127A	Chr12 : 20307858..20308050	193
MRZ0116B	Chr12 : 20330174..20330342	169
MRZ0116A	Chr12 : 20330249..20330494	246
MRZ0116C	Chr12 : 20330322..20330494	173
MRZ0155A	Chr12 : 20330733..20330890	158
<b>VMC8G9</b>	Chr12 : 20378485..20378656	172
MRZ0108B	Chr12 : 23204018..23204234	217

\* Marqueurs alignés par BLASTN

**Annexe 8. Les couples d'amorces définis dans la région du *locus* de résistance *Rpv2/Rpv3* avec une localisation unique par EPCR sur le génome de référence de *V. vinifera* cv PN40024.**

Marqueurs	Coordonnées sur le génome de référence	Taille de l'amplicon (pb)
<b>VMC2A7</b>	Chr18 : 13397376..13397518	142
MRZ0081A	Chr18 : 13486282..13486369	88
MRZ0059A	Chr18 : 13489961..13490183	223
MRZ0072A	Chr18 : 13850407..13850565	159
MRZ0031A	Chr18 : 13858229..13858456	228
MRZ0062A	Chr18 : 13858280..13858528	249
MRZ0094A	Chr18 : 14145661..14145832	172
MRZ0054A*	Chr18 : 14220508..14220679	172
MRZ0055A	Chr18 : 14295700..14296006	307
MRZ0064A	Chr18 : 15868203..15868402	200
MRZ0064B	Chr18 : 15868383..15868553	171
MRZ0063A	Chr18 : 15970263..15970427	165
MRZ0082A	Chr18 : 15985322..15985530	209
MRZ0058B	Chr18 : 18296291..18296540	250
MRZ0058A	Chr18 : 18296648..18296861	214
MRZ0099A	Chr18 : 18297917..18298117	201
MRZ0034A	Chr18 : 18834685..18834856	172
MRZ0036D	Chr18 : 19264021..19264245	225
MRZ0068B	Chr18 : 19479009..19479256	248
MRZ0068A*	Chr18 : 19479035..19479268	234
MRZ0068C	Chr18 : 19479090..19479269	180
MRZ0085A*	Chr18 : 19479752..19479973	222
MRZ0042D	Chr18 : 19484243..19484569	327
MRZ0066A	Chr18 : 19490513..19490689	177
MRZ0090B	Chr18 : 19527599..19527771	173
MRZ0090C	Chr18 : 19527638..19527815	178
MRZ0090A	Chr18 : 19527771..19527993	223
MRZ0041A	Chr18 : 19559694..19559892	199
MRZ0057A	Chr18 : 19559794..19560030	237
MRZ0057B	Chr18 : 19559794..19560020	227
MRZ0067A	Chr18 : 19587260..19587425	166
MRZ0060A	Chr18 : 19888868..19889102	235
<b>VMCNG2F12</b>	Chr18 : 19902915..19903113	198
MRZ0076C	Chr18 : 19920406..19920611	206
MRZ0076B	Chr18 : 19920588..19920749	162
MRZ0076A	Chr18 : 19920807..19920995	189
MRZ0061A	Chr18 : 20007593..20007756	164
MRZ0075A	Chr18 : 20042037..20042231	195
MRZ0053D	Chr18 : 20332731..20332931	201

MRZ0053B	Chr18 : 20332910..20333126	217
MRZ0053C	Chr18 : 20332912..20333104	193
MRZ0053A	Chr18 : 20333029..20333249	221
MRZ0097A	Chr18 : 20521410..20521623	214
MRZ0097B	Chr18 : 20521566..20521726	161
MRZ0098A	Chr18 : 20611688..20611952	265
MRZ0051A	Chr18 : 20763772..20763957	186
MRZ0052A	Chr18 : 20806305..20806586	282
MRZ0040A	Chr18 : 20806522..20806758	237
MRZ0052B	Chr18 : 20806573..20806757	185
MRZ0047A	Chr18 : 21257986..21258187	202
MRZ0103A	Chr18 : 21358520..21358698	179
MRZ0084A	Chr18 : 21363379..21363539	161
MRZ0083A	Chr18 : 21381938..21382152	215
MRZ0079A	Chr18 : 21552120..21552313	194
MRZ0079B	Chr18 : 21552138..21552313	176
MRZ0070A	Chr18 : 23265922..23266153	232
MRZ0100A	Chr18 : 23277235..23277494	260
MRZ0071A	Chr18 : 23338747..23339036	290
MRZ0087A	Chr18 : 23346324..23346516	193
MRZ0049A	Chr18 : 23381687..23381865	179
<b>VVIN16</b>	Chr18 : 23389686..23389846	160
MRZ0095C	Chr18 : 23401840..23402021	182
MRZ0095B	Chr18 : 23401865..23402021	157
MRZ0091A	Chr18 : 23424402..23424577	176
MRZ0048A	Chr18 : 23443542..23443743	202
MRZ0086A	Chr18 : 23449054..23449296	243
MRZ0096A	Chr18 : 23474614..23474915	302
MRZ0077A	Chr18 : 24399691..24399884	194
MRZ0078A	Chr18 : 24445364..24445523	160
<b>UDV-305</b>	Chr18 : 24868065..24868384	361
MRZ0046D	Chr18 : 25182566..25182734	169
MRZ0065A	Chr18 : 25182739..25182919	181
MRZ0080A	Chr18 : 25633486..25633648	163
MRZ0089A	Chr18 : 26792863..26792958	96
MRZ0050A	Chr18 : 26817620..26817839	220
MRZ0088A	Chr18 : 26859017..26859173	157
MRZ0088B*	Chr18 : 26859176..26859420	245
<b>VMC7F2</b>	Chr18 : 26896790..26896989	199
MRZ0093A	Chr18 : 28744584..28744739	156
MRZ0074A	Chr18 : 28882041..28882213	173
MRZ0101A	Chr18 : 29069393..29069552	160
MRZ0102A	Chr18 : 29100762..29100940	179
<b>UDV-108</b>	Chr18 : 29123113..29123353	240

\* Marqueurs alignés par BLASTN

**Annexe 9. Amplification des génomes de *V. vinifera* cv PN40024 et de *M. rotundifolia* cv Regale.** L'ensemble des marqueurs génétiques et physiques s'alignant sur le supercontig de *M. rotundifolia* cv G52 et sur le *locus* de résistance *Rpv2/Rpv3* a été utilisé pour amplifier la séquence de référence de *V. vinifera* cv PN40024 (EPCR), les génomes de *V. vinifera* cv PN40024 et *M. rotundifolia* cv Regale (PCR), et pour cribler la banque BAC de *M. rotundifolia* cv Regale.

Marqueurs	Amplification <i>In silico</i>	Amplification <i>In vitro</i>		Criblage de la banque BAC
	EPCR (pb) Séquence de référence	ADNg (pb) PN40024	ADNg (pb) Regale	Clone positif (N°) Test individuel
<b>49MRP1-P2</b>	208	220	220	–
<b>CB13.14</b>	–	300/250	310/200	–
<b>CB46.49</b>	–	400	540	–
<b>CB53.54</b>	–	–	350	5**
<b>VMC4F3-1</b>	181	200	220	15**
<b>UDV-058</b>	154	150	150	15**
<b>VMC8G9</b>	172	180	160	–
MRZ0104A	126	120/110	120	3
MRZ0105A*	163	180	180	–
MRZ0106A*	205	220	220	15
MRZ0106B*	191	220/150	220/150	1
MRZ0107A	178	180	180	1
MRZ0107B	224	220	220	2
MRZ0108A*	199	220	220	5
MRZ0108B	217	220	220	–
MRZ0108C	134	120	120	–
MRZ0109A*	214	230	230	–
MRZ0110A*	148	160	160	–
MRZ0111A	196	190	190	–
MRZ0121A	166	130	130	5
MRZ0121B	167	–	130	5
<b>MRZ0156A'</b>	280	280	280	4
<b>MRZ0156B'</b>	193	190	190	3
<b>MRZ0157A'</b>	135	130	130	6
<b>MRZ0157B'</b>	137	140	140	5
<b>MRZ0158A'</b>	218	220	230	5
<b>MRZ0159A'</b>	168	170	170	–
<b>MRZ0160A'</b>	261	260	260	6
<b>MRZ0160B'</b>	215	220	220	9
<b>MRZ0160C'</b>	141	140	140	5
<b>MRZ0161A'</b>	129	–	130	–

Région du *locus* de résistance *Ruv1/Rpv1*

<b>UDV-108</b>	240	250	220/230	10**
<b>UDV-305</b>	361	<i>Multilocus</i>	<i>Multilocus</i>	–
<b>VMC2A7</b>	142	–	–	4**
<b>VMC7F2</b>	199	200	200	5
<b>VMCNG2F12</b>	198	200	180	3
<b>VVIN16</b>	160	160	160	11**
MRZ0030A	–	200/270	270	3
MRZ0031A	228	220	220	15
MRZ0032C	–	130	130	2
MRZ0034A	172	150	150	11
MRZ0036D	225	200	200	7
MRZ0040A	237	230	230	10
MRZ0041A	199	200	200	3
MRZ0042A	–	240	240	2
MRZ0042D	327	250	250	2
MRZ0046D	169	160	160	4
MRZ0047A	202	200	200	4
MRZ0048A	202	200	200	6
MRZ0049A	179	180	180	8
MRZ0050A	220	220	220	6
MRZ0051A	186	190	190	3
MRZ0052A	282	280	280	10
MRZ0052B	185	280	280	11
MRZ0053A	221	220	220	2
MRZ0053B	217	220	220/230	3
MRZ0053C	193	190	190	–
MRZ0053D	201	200	200	3
MRZ0054A*	172	190/200	190	18
MRZ0055A	307	310	310	3
MRZ0056A	–	–	–	–
MRZ0057A	237	240	240	8
MRZ0057B	227	230	230	4
MRZ0058A	214	210	210	3
MRZ0058B	250	250	250	3
MRZ0059A	223	210	210	1**
MRZ0059B	–	–	–	–
MRZ0060A	235	230	230	–
MRZ0061A	164	160	160	4
MRZ0062A	249	250	250	5
MRZ0063A	165	160	160	2
MRZ0064A	200	200	190	6
MRZ0064B	171	170	170	7
MRZ0065A	181	180	180	3
MRZ0066A	177	140	140	–
MRZ0067A	166	170	170	3

Région du *locus* de résistance *Rpv2/Rpv3*

MRZ0068A*	234	230	230	1
MRZ0068B	248	250	250	–
MRZ0068C	180	180	180	2
MRZ0069A	–	–	230	3
MRZ0070A	232	230	230	3
MRZ0071A	290	290	290	5**
MRZ0072A	159	160	160	3
MRZ0073A	–	–	–	–
MRZ0074A	173	170	170	6
MRZ0075A	195	190	190	6
MRZ0076A	189	190	190	7
MRZ0076B	162	160	160	4
MRZ0076C	206	210	210	5
MRZ0077A	194	190	190	1
MRZ0078A	160	160/200	160/200	3
MRZ0079A	194	190	190	5
MRZ0079B	176	180	–	27
MRZ0080A	163	160	160	5
MRZ0081A	88	90	90	8
MRZ0082A	209	210	210	5
MRZ0083A	215	210	210	9
MRZ0084A	161	160	160	6
MRZ0085A*	222	240	240	2
MRZ0086A	243	240	240	4
MRZ0087A	193	160/190	160/190	8**
MRZ0088A	157	160/200	140/160	5
MRZ0088B*	245	250	250	13
MRZ0089A	96	100	100	1
MRZ0090A	223	220	220	4
MRZ0090B	173	170	170	6
MRZ0090C	178	180	180	8
MRZ0091A	176	180	180	4
MRZ0092A	–	–	–	–
MRZ0093A	156	160	160	2**
MRZ0094A	172	–	170	5
MRZ0095A	–	130	130	–
MRZ0095B	157	160	160	3
MRZ0095C	182	180	180	3
MRZ0096A	302	300	300	20**
MRZ0097A	214	210	210	8
MRZ0097B	161	160	160	8
MRZ0098A	265	230/260	230/260	1
MRZ0099A	201	200	200	3
MRZ0100A	260	260	260	5
MRZ0101A	160	160	160	4

MRZ0102A	179	180	180	5
MRZ0103A	179	180	180	35
MRZ0162A'	176	180	180	4
MRZ0162B'	151	150	150	4
MRZ0163A'	208	210	210	1
MRZ0164A'	112	110	110	10
MRZ0165A'	208	210	210	–
MRZ0166A'	196	200	200	2
MRZ0166B'	158	160	160	5
MRZ0167A'	197	200	200	22
MRZ0168A'	337	340	340	–
MRZ0168B'	235	240	240	17
MRZ0169A'	236	240	240	1
MRZ0170A'	137	140	140	6
MRZ0171A'	243	240	240	2
MRZ0171B'	179	180	180	1
MRZ0172A'	183	180	180	21
MRZ0173A'	144	140	140	–

\*: Marqueurs alignés par BLASTN; \*\*: Criblage sur la banque entière; \_: Absence d'amplification



**Annexe 10. Alignement des contigs sur le génome de *V. vinifera* cv PN40024 par BLASTN**

Contig n°	Contig n°	Orientation	% identité	Coordonnées d'alignement Génome de référence	Score	Taille du contig (pb)	Présence de BES
<b>62H06</b>	Contig4	-	97	Chr12:15528606..15536260	7136	8009	<i>Reverse</i>
	Contig3	+	94,1	Chr12:15536863..15591272	8451	15177	-
	Contig2	+	95,7	Chr12:15575939..15625876	5590	17646	-
	Contig1	-	95,8	Chr12:15626332..15678609	16937	42659	-
	Contig5	-	96,5	Chr12:15680881..15696590	4150	4553	<i>Forward</i>
<b>34D01</b>	Contig2	+	93,7	Chr12:15897583..16770278	11881	21692	<i>Reverse</i>
	Contig1	+	87,3	Chr12:6975650..6978577	1125	25648	<i>Forward</i>
<b>20I10</b>	Contig4	-	93,6	Chr12:15750201..16682671	6249	7527	<i>Forward</i>
	Contig3	+	94,7	Chr12:15761487..16448185	14200	18960	-
	Contig2	+	89,7	Chr12:16454499..16642836	6524	26744	-
	Contig1	-	91,7	Chr12:16701441..16709423	3405	30413	<i>Reverse</i>
	Contig5	-	93,7	Chr12:16766615..16767921	1097	1467	-
<b>11K09</b>	Contig3	+	94,8	Chr12:17430841..17446931	12991	15682	<i>Forward</i>
	Contig1	-	95,6	Chr12:17451290..17496452	26212	33863	-
	Contig2	-	94,2	Chr12:17496452..17511794	12644	16122	<i>Reverse</i>
<b>13C22</b>	Contig1	+	90,7	Chr18:22400355..22660414	5823	57957	<i>Reverse</i>
	Contig2	-	95,7	Chr18:23748452..23756784	6420	7591	<i>Forward</i>
<b>40F13</b>	Contig2	+	93,5	Chr8:7592020..7604591	9491	16388	<i>Forward</i>
	Contig5	-	91,7	Chr8:7605536..7606338	674	1156	-
	Contig4	+	92,8	Chr8:7606822..7613989	2691	5993	-
	Contig3	-	94,2	Chr8:7610082..7626561	13536	20054	-
	Contig1	+	96,9	Chr8:7626541..7629324	2343	2609	<i>Reverse</i>
<b>43H05</b>	Contig1	-	95,6	Chr18:25633057..25653749	14275	16189	<i>Reverse</i>
	Contig2	-	95,1	Chr18:25653750..25662922	7960	8928	<i>Forward</i>
<b>56A14</b>	Contig2	+	97,7	Chr18:26792531..26794688	1954	2088	<i>Forward</i>
	Contig3	-	96,4	Chr18:26794670..26807017	5448	6804	-
	Contig1	+	95,8	Chr18:26806948..26858638	34151	74546	<i>Reverse</i>
<b>58B07</b>	Contig1	-	96,8	Chr18:27491500..27504403	7269	7948	<i>Reverse</i>
	Contig2	-	94,8	Chr18:27504400..27513058	5717	9766	<i>Forward</i>

**Annexe 11. La méthode de prédiction de gènes *ab initio* a été faite à l'aide du programme FGGENESH (Yu *et al.* 2002 ; Goff *et al.* 2002).**

11K09							
N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1	-	1 CDS <sub>o</sub>	4326	4622	23,59	4326-4622	297
2*	+	1 CDS <sub>f</sub>	18102	18807	35,98	18102-18806	705
2*	+	2 CDS <sub>i</sub>	18945	19225	18,06	18947-19225	279
2*	+	3 CDS <sub>i</sub>	19324	19560	11,94	19324-19560	237
3	-	1 CDS <sub>i</sub>	30944	31088	0,43	30944-31087	144
3	-	2 CDS <sub>i</sub>	31200	31276	6,31	31202-31276	75
3	-	3 CDS <sub>i</sub>	31952	32164	16,73	31952-32164	213
3	-	4 CDS <sub>i</sub>	33181	33251	2,53	33181-33249	69
3	-	5 CDS <sub>i</sub>	37527	37662	3,39	37528-37662	135
3	-	6 CDS <sub>f</sub>	40636	40851	8,76	40636-40851	216
4	-	1 CDS <sub>o</sub>	57962	58567	17,13	57962-58567	606
5*	+	1 CDS <sub>o</sub>	60731	64486	201,07	60731-64486	<b>3756</b>
13C22							
N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1	+	1 CDS <sub>o</sub>	46	825	51,14	46-825	780
2	-	1 CDS <sub>i</sub>	1261	1817	12,58	1261-1815	555
2	-	2 CDS <sub>i</sub>	2915	2999	0,35	2916-2999	84
2	-	3 CDS <sub>i</sub>	3366	3782	5,04	3366-3782	417
2	-	4 CDS <sub>i</sub>	4250	4407	4,78	4250-4405	156
2*	-	5 CDS <sub>i</sub>	15219	17256	55,8	15220-17256	2037
2*	-	6 CDS <sub>i</sub>	17355	17672	14,56	17355-17672	318
2*	-	7 CDS <sub>i</sub>	17823	18238	14,19	17823-18236	414
2*	-	8 CDS <sub>i</sub>	18436	18917	25,36	18437-18916	480
2*	-	9 CDS <sub>f</sub>	19591	20051	31,82	19593-20051	459
3	-	1 CDS <sub>i</sub>	37793	38127	10,26	37793-38125	333
3	-	2 CDS <sub>i</sub>	38224	38838	39,82	38225-38836	612
3	-	3 CDS <sub>i</sub>	38947	39186	9,93	38948-39184	237
3	-	4 CDS <sub>i</sub>	39223	39408	5,03	39224-39406	183
3	-	5 CDS <sub>i</sub>	40986	41364	23,99	40987-41364	378
3	-	6 CDS <sub>f</sub>	41458	42336	73,16	41458-42336	879
4*	-	1 CDS <sub>i</sub>	42933	43455	16,85	42933-43454	522
4*	-	2 CDS <sub>i</sub>	43533	44362	30,58	43535-44362	828
4*	-	3 CDS <sub>i</sub>	44516	46021	41,37	44516-46021	1506
4*	-	4 CDS <sub>i</sub>	46618	47662	48,8	46618-47661	1044
4*	-	5 CDS <sub>i</sub>	48947	49187	12,6	48949-49185	237
4*	-	6 CDS <sub>i</sub>	49677	49877	4,96	49678-49875	198
4*	-	7 CDS <sub>i</sub>	49956	50284	7,97	49957-50283	327
4*	-	8 CDS <sub>i</sub>	50517	50547	0,5	50519-50545	<b>27</b>
4*	-	9 CDS <sub>i</sub>	50770	51244	15,27	50771-51244	474

4*	-	10 CDSi	52337	52551	18,28	52337-52549	213
4*	-	11 CDSf	52937	52991	1,11	52938-52991	54
5	+	1 CDSo	54059	54394	12,9	54059-54394	336
6	+	1 CDSf	56294	56734	13,37	56294-56734	441
6	+	2 CDSi	56834	57155	11,39	56834-57154	321
6	+	3 CDSi	57562	57830	4,23	57564-57830	267
6	+	4 CDSi	58338	58703	2,37	58338-58703	366
6	+	5 CDSi	58755	59411	33,27	58755-59411	657
6	+	6 CDSi	59481	60083	30,44	59481-60083	603
6	+	7 CDSi	60120	60521	11,41	60120-60521	402
7*	-	1 CDSo	64563	65564	21,61	64563-65564	1002

### 20110

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1	-	1 CDSi	9578	9837	14	9578-9835	258
1	-	2 CDSf	11848	12028	5,86	11849-12028	180
2*	+	1 CDSf	13027	13365	20,8	13027-13365	339
2*	+	2 CDSi	13813	13843	2,08	13813-13842	30
2*	+	3 CDSi	13943	14901	50,48	13945-14901	957
2*	+	4 CDSi	15088	15837	55,9	15088-15837	750
2*	+	5 CDSi	15964	16240	4,98	15964-16239	276
2*	+	6 CDSi	16388	16437	0,79	16390-16437	48
2*	+	7 CDSi	16565	16900	14,76	16565-16900	336
3	-	1 CDSi	18045	18398	11,92	18045-18398	354
3	-	2 CDSi	18668	18844	7,08	18668-18844	177
3	-	3 CDSi	18941	19045	0,62	18941-19045	105
3	-	4 CDSi	19306	19362	4,82	19306-19362	57
3	-	5 CDSf	19484	19699	5,85	19484-19699	216
4	+	1 CDSf	20948	21376	7,37	20948-21376	429
4	+	2 CDSi	21556	21669	6,15	21556-21669	114
4	+	3 CDSi	21811	22303	1,7	21811-22302	492
4	+	4 CDSi	22475	22705	5,12	22477-22704	228
4	+	5 CDSi	22991	23240	12,29	22993-23238	246
4	+	6 CDSi	23409	23599	2,58	23410-23598	189
4	+	7 CDSi	23933	24282	16,64	23935-24282	348
5	+	1 CDSf	43691	43717	0,59	43691-43717	27
5	+	2 CDSi	43811	43867	4,85	43811-43867	57
5	+	3 CDSi	43956	44015	5,08	43956-44015	60
5	+	4 CDSi	45144	45263	14,01	45144-45263	120
5	+	5 CDSi	45740	45808	-4,17	45740-45808	69
6*	+	1 CDSf	55568	55713	6,11	55568-55711	144
6*	+	2 CDSi	57252	58044	36,82	57253-58044	792
6*	+	3 CDSi	58364	58637	3,82	58364-58636	273
6*	+	4 CDSi	58803	59119	5,6	58805-59119	315
6*	+	5 CDSi	59275	59592	4,2	59275-59592	318
6*	+	6 CDSi	61277	61510	21,53	61277-61510	234

6*	+	7 CDSi	61685	61869	4,45	61685-61867	183
6*	+	8 CDSi	62544	62777	4,8	62545-62775	231
6*	+	9 CDSi	63835	63950	-1,12	63836-63949	114
6*	+	10 CDSi	64241	64453	3,19	64243-64452	210
6*	+	11 CDSi	64706	65217	30,27	64708-65217	510
6*	+	12 CDSi	65659	66690	60,17	65659-66690	1032
6*	+	13 CDSi	66748	66955	16,6	66748-66954	207
6*	+	14 CDSi	66997	67548	32,03	66999-67547	549
6*	+	15 CDSi	67609	67974	16,4	67611-67973	363
6*	+	16 CDSi	68243	69681	98,41	68245-69681	1437
7	+	1 CDS <sub>o</sub>	70726	71655	54,54	70726-71655	930
8	+	1 CDS <sub>f</sub>	74194	74324	4,56	74194-74322	129
8	+	2 CDSi	76590	76836	24	76591-76836	246
8	+	3 CDSi	76870	77112	-5,2	76870-77112	243
9*	-	1 CDSi	82318	82569	7,49	82318-82569	252
9*	-	2 CDSi	82746	83045	9,01	82746-83045	300
9*	-	3 CDSi	83124	84218	68,42	83124-84218	1095
9*	-	4 CDSi	84533	84703	10,13	84533-84703	171
9*	-	5 CDS <sub>f</sub>	84849	85223	13,85	84849-85223	375

#### 34D01

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1*	+	1 CDSi	48	517	11,33	50-517	468
1*	+	2 CDSi	660	1019	10,68	660-1019	360
1*	+	3 CDSi	1114	1911	46,88	1114-1911	798
1*	+	4 CDSi	1993	2748	39,27	1993-2748	756
1*	-	1 CDSi	7648	7878	5,55	7648-7878	231
1*	-	2 CDSi	8095	8241	9,29	8095-8241	147
1*	-	3 CDSi	9279	10682	62,93	9279-10682	1404
2	-	4 CDSi	10752	11687	44,34	10752-11687	936
2	-	5 CDSi	11746	11795	5,01	11746-11793	48
2	-	6 CDSi	11855	12227	10,14	11856-12227	372
2	-	7 CDSi	12670	13417	44,57	12670-13416	747
2	-	8 CDSi	13463	14313	31,12	13465-14313	849
2	-	9 CDS <sub>f</sub>	14404	14631	17,01	14404-14631	228
3	-	1 CDSi	17229	17471	-0,64	17229-17471	243
3	-	2 CDSi	17505	17754	22,75	17505-17753	249
3	-	3 CDSi	18218	18263	4,26	18220-18261	42
3	-	4 CDSi	19107	19338	12	19108-19338	231
3	-	5 CDS <sub>f</sub>	19399	20145	42,71	19399-20145	747
4	-	1 CDSi	36406	36708	4,19	36406-36708	303
4	-	2 CDS <sub>f</sub>	36804	36965	13,64	36804-36965	162
5*	-	1 CDS <sub>o</sub>	39407	40237	20,31	39407-40237	831

#### 40F13

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
----	------	-----	-------	-----	-------	-----	-------------

1*	-	1 CDSi	2656	2822	10,88	2656-2820	165
1*	-	2 CDSi	12083	12293	11,19	12084-12293	210
1*	-	3 CDSi	12377	12707	35,71	12377-12706	330
1*	-	4 CDSf	13871	14370	33,16	13873-14370	498
1*	+	1 CDSf	17667	17869	-1,25	17667-17867	201
1*	+	2 CDSi	17976	18033	2,5	17977-18033	57
1*	+	3 CDSi	18739	18969	32,97	18739-18969	231
1*	+	4 CDSi	19049	19198	5,17	19049-19198	150
3	-	1 CDSi	44092	44313	15,29	44092-44313	222
3	-	2 CDSi	44405	44520	5,17	44405-44518	114

#### 43H05

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1*	+	1 CDS <sub>o</sub>	5757	6833	17,62	5757-6833	1077
2*	+	1 CDSf	15047	16229	68,08	15047-16228	1182
2*	+	2 CDSi	16668	16800	-4,11	16670-16798	129
2*	+	3 CDSi	17301	17394	6,13	17302-17394	93
2	+	4 CDSi	18944	19040	-0,91	18944-19039	96
2	+	5 CDSi	19109	19254	0,52	19111-19254	144

#### 56A14

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1	+	1 CDS <sub>o</sub>	10901	12154	52,95	10901-12154	1254
2*	+	1 CDSf	17555	17800	24,08	17555-17800	246
2*	+	2 CDSi	19683	19922	17,32	19683-19922	240
2*	+	3 CDSi	20024	20167	8,64	20024-20167	144
2*	+	4 CDSi	20770	20904	7,67	20770-20904	135
2*	+	5 CDSi	23421	23498	2,72	23421-23498	78
2*	+	6 CDSi	25762	26034	15,84	25762-26034	273
3*	+	1 CDSf	37212	37443	13,5	37212-37442	231
3*	+	2 CDSi	37653	37930	23,66	37655-37930	276
3*	+	3 CDSi	38417	38614	3,4	38417-38614	198
3*	+	4 CDSi	39090	39122	1,88	39090-39122	33
3*	+	5 CDSi	40466	40630	9,71	40466-40630	165
3*	+	6 CDSi	40716	40850	11,18	40716-40850	135
3*	+	7 CDSi	41121	41333	22,03	41121-41333	213
4	+	1 CDSf	49898	50017	7,52	49898-50017	120
4	+	2 CDSi	50121	50192	6,33	50121-50192	72
4	+	3 CDSi	51822	52121	9,82	51822-52121	300
5*	+	1 CDSf	55879	56229	15,96	55879-56229	351
5*	+	2 CDSi	56718	56777	7,49	56718-56777	60
5*	+	3 CDSi	57329	57423	8,65	57329-57421	93
5*	+	4 CDSi	57896	57965	3,71	57897-57965	69
5*	+	5 CDSi	58094	58210	0,39	58094-58210	117
6	-	1 CDSi	66563	66784	2,88	66563-66784	222
6	-	2 CDSi	66988	67029	1,52	66988-67029	42

6	-	3 CDSi	67468	67567	13,79	67468-67566	99
6	-	4 CDSi	67792	67853	4,1	67794-67853	60
6	-	5 CDSi	68115	68196	5,57	68115-68195	81
6	-	6 CDSf	72397	72578	13,08	72399-72578	180

#### 58B07

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1*	-	1 CDS <sub>o</sub>	2627	2989	27,01	2627-2989	363
2*	+	1 CDSf	6095	6248	8,03	6095-6247	153
2*	+	2 CDSi	6924	6994	6,1	6926-6994	69
2*	+	3 CDSi	7056	8077	53,7	7056-8075	1020
2*	+	4 CDSi	8183	8603	20,74	8184-8603	420
2*	+	5 CDSi	9960	10085	3,22	9960-10085	126
2*	+	6 CDSi	10272	10490	12,24	10272-10490	219
2*	+	7 CDSi	10706	11101	12,29	10706-11101	396
2*	+	8 CDSi	11154	11672	21,23	11154-11672	519
2*	+	9 CDSi	13183	13281	9,93	13183-13281	99
2*	+	10 CDSi	13655	13882	5,72	13655-13882	228

#### 62H06

N°	Brin	CDS	Start	End	Score	ORF	Taille (pb)
1	+	1 CDSf	57	922	44,53	57-920	864
1	+	2 CDSi	1156	1273	5,19	1157-1273	117
1	+	3 CDSi	1716	2189	13,39	1716-2189	474
1	+	4 CDSi	2235	2625	17,47	2235-2624	390
1	+	5 CDSi	3006	3254	10,96	3008-3253	246
1	+	6 CDSi	3300	3832	17,99	3302-3832	531
2*	-	1 CDS <sub>o</sub>	15340	16767	98,22	15340-16767	1428
3	+	1 CDS <sub>o</sub>	20040	20603	24,35	20040-20603	564
4	-	1 CDSi	22756	23080	11,32	22756-23079	324
4	-	2 CDSi	23330	23481	8,42	23332-23481	150
4	-	3 CDSi	23535	23597	3,31	23535-23597	63
4	-	4 CDSi	23815	24107	8,15	23815-24105	291
4	-	5 CDSf	24192	24627	23,47	24193-24627	435
5*	-	1 CDSi	26797	27084	2,97	26797-27084	288
5*	-	2 CDSi	27593	28149	28,99	27593-28147	555
5*	-	3 CDSi	28553	29989	63,01	28554-29987	1434
5*	-	4 CDSi	30050	30119	4,27	30051-30119	69
5*	-	5 CDSi	30732	31678	1,33	30732-31676	945
5*	-	6 CDSi	32015	32765	12,37	32016-32765	750
5*	-	7 CDSf	33176	33460	23,21	33176-33460	285
6*	-	1 CDS <sub>o</sub>	34785	36203	112,41	34785-36203	1419
7	-	1 CDSi	40441	40566	-2,35	40441-40566	126
7	-	2 CDSi	40814	40876	5,85	40814-40876	63
7	-	3 CDSi	41091	41175	7,52	41091-41174	84
7	-	4 CDSi	43551	43846	0,31	43553-43846	294

7	-	5 CDSi	45856	46218	19,11	45856-46218	363
7	-	6 CDSf	47534	47593	0,16	47534-47593	60
8	+	1 CDSf	49083	49176	-0,07	49083-49175	93
8	+	2 CDSi	49814	49888	2,67	49816-49887	72
8	+	3 CDSi	50559	50801	3,9	50561-50800	240
8	+	4 CDSi	51140	51279	5,59	51142-51279	138
8	+	5 CDSi	52401	52634	14,88	52401-52634	234
9	-	1 CDS <sub>o</sub>	56047	57048	88,6	56047-57048	1002
10	-	1 CDSi	59453	59904	12,57	59453-59902	450
10	-	2 CDSi	59992	60639	33,01	59993-60637	645
10	-	3 CDSi	60709	60851	4,55	60710-60850	141
10	-	4 CDSi	60932	61646	17,66	60934-61644	711
10	-	5 CDSf	62029	62149	-0,32	62030-62149	120
11	+	1 CDS <sub>o</sub>	77660	77968	16,58	77660-77968	309
12*	+	1 CDS <sub>o</sub>	80497	81927	99,63	80497-81927	1431
13	+	1 CDS <sub>o</sub>	86827	87045	12,65	86827-87045	219

\* : Gènes/Exons avec des séquences homologues dans les banques de données SwissProt  
CDS<sub>o</sub> : Gène à un seul exon ; CDS<sub>i</sub> : Premier exon ; CDS<sub>i</sub> : Exon interne ; CDS<sub>f</sub> : Dernier exon

**Annexe 12. Protéines homologues aux séquences protéiques de *M. rotundifolia* cv Regale.** Les séquences extraites des CDS ont été recherchées dans la base de données SWISSPROT par BLASTP Chaque ligne représente une protéine

<b>BLASTP</b>					
	<b>Gène*</b>	<b>Accession</b>	<b>Espèces</b>	<b>E-value</b>	<b>Protéines homologues</b>
<b>62H06</b>	CP	Q2V6K0	<i>Fragaria x ananassa</i>	$2.10^{-179}$	UDP-glucose flavonoid 3-O-glucosyltransferase 6
	CP	Q66PF3	<i>Fragaria x ananassa</i>	$5.10^{-102}$	UDP-glucose flavonoid 3-O-glucosyltransferase 3
	CP	Q2V6K0	<i>Fragaria x ananassa</i>	0	UDP-glucose flavonoid 3-O-glucosyltransferase 6
	CP	Q2V6K0	<i>Fragaria x ananassa</i>	0	UDP-glucose flavonoid 3-O-glucosyltransferase 6
<b>34D01</b>	TR	Q40392	<i>Nicotiana glutinosa</i>	$6.10^{-80}$	TMV resistance protein N
	CP	P51615	<i>Vitis vinifera</i>	$7.10^{-21}$	NADP-dependent malic enzyme
<b>20I10</b>	TR	Q40392	<i>Nicotiana glutinosa</i>	$1.10^{-77}$	TMV resistance protein N
<b>11K09</b>	CP	Q9SAC4	<i>Arabidopsis thaliana</i>	0	F-box protein
	CP	Q9LRR4	<i>Arabidopsis thaliana</i>	0	Putative disease resistance RPP13-like protein 1
<b>13C22</b>	PS	Q40392	<i>Nicotiana glutinosa</i>	$7.10^{-170}$	TMV resistance protein N
	PS	Q9LPG6	<i>Arabidopsis thaliana</i>	$1.10^{-24}$	UDP-L-rhamnose synthase
	PS	Q9SYM5	<i>Arabidopsis thaliana</i>	$9.10^{-25}$	Probable rhamnose biosynthetic enzyme 1
	CP	Q40392	<i>Nicotiana glutinosa</i>	$5.10^{-8}$	TMV resistance protein N
<b>40F13</b>	CP	D4N501	<i>Papaver somniferum</i>	$3.10^{-71}$	2-oxoglutarate/Fe(II)-dependent dioxygenase
	CP	Q8W4I7	<i>Arabidopsis thaliana</i>	$4.10^{-78}$	Calcium-dependent protein kinase 13
	CP	Q8W4I7	<i>Arabidopsis thaliana</i>	$5.10^{-80}$	Calcium-dependent protein kinase 13
<b>43H05</b>	CP	Q8GXC7	<i>Arabidopsis thaliana</i>	$2.10^{-47}$	F-box/kelch-repeat protein
	CP	Q9LIR8	<i>Arabidopsis thaliana</i>	$2.10^{-37}$	F-box/kelch-repeat protein
<b>56A14</b>	CP	Q940H7	<i>Arabidopsis thaliana</i>	0	Putative methylesterase 12
	CP	P54120	<i>Arabidopsis thaliana</i>	$2.10^{-63}$	Protein AIG1
	PS	P54120	<i>Arabidopsis thaliana</i>	$2.10^{-6}$	Protein AIG1
<b>58B07</b>	CP	O82352	<i>Arabidopsis thaliana</i>	$5.10^{-21}$	Reticulon-like protein B5
	PS	Q7G9P4	<i>Arabidopsis thaliana</i>	0	Aldehyde oxidase 3

\* CP = Gènes complets ; TR = Gènes tronqués ; PS = Pseudogènes



**Annexe 13. Les codes PFAM et les termes GO assignés aux séquences protéiques extraites des CDS de *M. rotundifolia* cv Regale. Chaque ligne représente une protéine**

	<i>PFAM FAMILY</i>	<i>GENE ONTOLOGY (GO)</i>		
	<i>PFAM code</i>	<i>Biological process</i>	<i>Molecular function</i>	<i>Cellular component</i>
<b>62H06</b>	PF00201	0008152	0016758	
	PF00201	0008152	0016758	
	PF00201	0008152	0016758	
<b>34D01</b>	PF00931/PF13855/PF13855	0006952	0043531	
	PF00390	0006108/0055114	0004470/0016616/0016619	
<b>20I10</b>	PF00931/PF07725/PF12799	0006952	0043531	
<b>11K09</b>	PF12937		0005515	
	PF00931	0006952	0043531	
<b>13C22</b>	PF01582/PF00931/PF13855	0007165/0006952	0005515/0043531	
	PF01370	0044237	0003824/0050662	
	PF01370	0044237	0003824/0050662	
	PF01582	0007165	0005515	0005622
<b>40F13</b>	PF14226/PF03171	0055114	0005506/0016491/0016706	
	PF13499		0005509	
	PF00036		0005509	
<b>43H05</b>	PF12937/PF07734		0005515	
	PF00646		0005515	
<b>56A14</b>	PF12697			
	PF04548		0005525	
	PF04548		0005525	
<b>58B07</b>	PF02453			0005783
	PF00111/PF02738	0055114	0009055/0051536/0016491	

**Annexe 14. Gènes du génome de référence de *V. vinifera* cv PN40024 couverts par les contigs de *M. rotundifolia* cv Regale**

Clone BAC Regale	Contig (n°)	Gènes annotés PN40024	Code PFAM /Terme GO	Famille de protéines
<b>62H06</b>	4	Vv12s0034g00130	PF00201	<i>Glucuronosyltransferase</i>
	2	Vv12s0034g00140	PF00201	<i>Glucuronosyltransferase</i>
	3	Vv12s0034g00140	PF00201	<i>Glucuronosyltransferase</i>
		Vv12s0034g00160	PF00201	<i>Glucuronosyltransferase</i>
	1	Vv12s0034g00160	PF00201	<i>Glucuronosyltransferase</i>
		Vv12s0034g00170	–	–
		Vv12s0034g00180	PF03405	<i>FA_desaturase_2</i>
		Vv12s0034g00190	PF03405	<i>FA_desaturase_2</i>
		Vv12s0034g00200	PF03492	<i>Methyltransf_7</i>
		Vv12s0034g00210	PF15252*	<i>DUF4589</i>
		Vv12s0034g00240	PF00069/PF07714	<i>Pkinase/Tyrosine kinase</i>
	5	Vv12s0034g00240	PF00069/PF07714	<i>Pkinase/Tyrosine kinase</i>
	<b>34D01</b>	2	–	–
1		–	–	–
<b>20I10</b>	4	–	–	–
	3	–	–	–
	2	–	–	–
	1	–	–	–
	5	Vv12s0034g00950	PF03801/PF10174*	<i>Ndc80_HEC/Cast</i>
<b>11K09</b>	3	Vv12s0034g01420	–	–
	1	Vv12s0034g01430	PF12937*	<i>F-box-like</i>
		Vv12s0034g01440	PF01105/GO:0006810;0016020	<i>EMP24_GP25L</i>
	2	Vv12s0034g01460	PF00931/GO:0000166;0004721	<i>NB-ARC domain</i>
	Vv12s0034g01470	PF00931/GO:0000166;0005515	<i>NB-ARC domain</i>	
<b>13C22</b>	1	Vv18s0075g00740	–	–
		Vv18s0075g00750	–	–
		Vv18s0075g00760	–	–
		Vv18s0075g00770	–	–
		Vv18s0075g00780	PF00394/PF07731	<i>Cu-oxidase/Cu-oxidase_2</i>
		Vv18s0075g00790	PF03467	<i>Smg4_UPF3</i>
		Vv18s0075g00800	PF01582	<i>TIR</i>
		Vv18s0075g00810	PF00394/PF07731	<i>Cu-oxidase/Cu-oxidase_2</i>
		Vv18s0075g00820	–	–
		Vv18s0075g00830	–	–
		Vv18s0075g00840	PF00244*/PF13246*	<i>14-3-3 protein/Hydrolase_like2</i>
		Vv18s0075g00880	PF01743	<i>PolyA_pol</i>
		Vv18s0117g00390	PF03101	<i>FAR1 DNA-binding domain</i>
	2	Vv18s0117g00390	PF03101	<i>FAR1 DNA-binding domain</i>
<b>40F13</b>	2	Vv8s0105g00380	PF03171	<i>2OG-Fe(II) oxygenase superfamily</i>

	1	Vv8s0105g00390	PF00036/PF00069	<i>EF-hand_1/Protein kinase domain</i>
	5	–	–	–
	4	–	–	–
	3	Vv8s0105g00390	PF00036, PF00069	<i>EF-hand_1/Protein kinase domain</i>
<b>43H05</b>	1	Vv18s0041g01090	PF00646	<i>F-box</i>
	2	Vv18s0041g01100	PF07734/PF01344*	<i>FBA_1/Kelch_1</i>
<b>56A14</b>	2	–	–	–
	3	Vv18s0041g01830	PF02826/GO:0003824;0004617	<i>2-Hacid_dh_C</i>
		Vv18s0041g01840	PF00646	<i>F-box</i>
	1	Vv18s0041g01840	PF00646	<i>F-box</i>
		Vv18s0041g01850	PF00561	<i>α/β-hydrolase_1</i>
		Vv18s0041g01860	PF04548	<i>AIG1</i>
<b>58B07</b>	1	Vv18s0041g02390	PF01315/PF02738	<i>Ald_Xan_dh_C/Ald_Xan_dh_C2</i>
	2	Vv18s0041g02400	PF00111/PF01799	<i>Ferredoxin/[2Fe-2S] binding domain</i>

\* : Les alignements non significatifs

## ANNEXE 15. PUBLICATION EN PREPARATION

### **Development of genomic resources for *Muscadinia rotundifolia* in order to assist comparative mapping with *Vitis vinifera***

Zah-Bi CI<sup>1</sup>, Canaguier A<sup>1</sup>, Choisne N<sup>2</sup>, Le Clainche I<sup>1</sup>, Bras M<sup>2</sup>, Couturat L<sup>3</sup>, Dry I<sup>4</sup>, Gouyvenoux M<sup>3</sup>, Poulain J<sup>3</sup>, Adam-Blondon A-F<sup>1,2\*</sup>

<sup>1</sup> UMR-INRA-UEVE-CNRS Génomique Végétale, 2 rue Gaston Crémieux, BP5708, 91 057 Evry cedex, France

<sup>2</sup> Unité de Recherches Génomique-Info, route de Saint-Cyr 78026 Versailles, France

<sup>3</sup> Genoscope, CEA-IG, 2 rue Gaston Crémieux, BP 5708, 91057 Evry cedex, France

<sup>4</sup> CSIRO Plant Industry, PO Box 350, Glen Osmond SA 5064, Australia

\*corresponding author

Grapes are cultivated in Mediterranean and subtropical areas throughout the world for fruit, juice, raisin and wine production. All the cultivated grapes belong to the *Vitis* genus that is divided in two sub-genera, *Euvtis* and *Muscadinia* (Patel and Olmo 1955). Their phylogenetic and paleogeographic origins are better understood now and confirm their status of sister genus (Peros et al 2011). The most widespread cultivated species is *Vitis vinifera*, the only species originating from Europe among the 36 species of the *Euvtis* subgenus. It is adapted to Mediterranean climates and has been introduced in all continents over the ages (Mullins et al 1992; This et al 2006; Bouquet 2011). During this period of historical expansion, the cultivation of *Vitis vinifera* has been threaten by pathogens with which it had not co-evolved, such as fungi (e.g. *Erishyphne necator*, *Plasmopara viticola*, ...), bacteria (e.g. *Xyllella fastidiosa*) or insects (e.g. *Phylloxera vastatrix*; Mullins et al 1992). A few other species of the *Euvtis* subgenus are cultivated for their berries either as a species or as parents of interspecific hybrids (e.g. *V. labrusca*, *V. amurensis*) (Bouquet 2011; Töpfer et al 2011; Di Gaspero et al 2012) while others (e.g. *V. riparia*, *V. berlandieri*) are widely used for the breeding of rootstocks (Mullins et al 1992; Bouquet 2011; Töpfer et al 2011; Di Gaspero et al 2012). *Muscadinia rotundifolia* is the only species from the *Muscadinia* subgenus which is cultivated, in subtropical areas of South-East USA.

A cost-effective, environment friendly alternative to the use of chemicals to cope with crop diseases is the use of varieties resistant to pathogens. It is sometimes the only solution when there is no possibility of chemical control (e.g. control of *Phylloxera*, nematode, viruses, phytoplasms,...). For grapes, the resistance needs to be introduced from adapted species of the *Vitaceae* family through breeding programs ensuring wine or fruit quality regarding the scion or adaptation to environment regarding the rootstock (Bouquet et al 2011; Töpfer et al 2011; Di Gaspero et al 2012). The species of the *Vitaceae* family have been and are still widely used to create rootstocks resistant to *Phylloxera vastatrix* (Boubals 1966) and to nematodes either detrimental by themselves or because they transmit viruses (Laimer et al 2009; Esmenjaud et al 2010). They are also sources of resistance for the breeding of new grapevine varieties resistant to the mildews (Eibach et al 1989, Eibach 1994) and the Pierce disease (Riaz et al 2007). Among the *Vitaceae* germplasm, *M. rotundifolia* is a source of high level resistances to several major pathogens of grapevine (Bouquet 2011). It has also been shown to be an interesting source of nutraceutical compounds (Pastrana-Bonilla et al 2003). Although the sequencing of the grapevine genome has shown a slow evolution since the formation of its hexaploid ancestor (Jaillon et al 2007), the number of chromosomes varies within the family (Branas 1932). The different number of chromosomes between *V. vinifera* ( $2n=38$ ) and *M. rotundifolia* ( $2n=40$ ) indeed makes the first generations of crosses between these two species poorly efficient (Detjen 1919, Patel and Olmo 1955). Comparative mapping between *V. vinifera* and *M. rotundifolia* would speed up the identification and isolation the different resistance genes from *M. rotundifolia*, would allow a better understanding of the mechanisms associated to the introgression of genome segments from *M. rotundifolia* into *V. vinifera* and would give clues on genome evolution among the *Vitaceae*. However, despite its importance for

grapevine breeding, knowledge about the *M. rotundifolia* genome is yet very limited. The first genetic map was only very recently published (Blanc et al 2012) and the genomic resources publically available are scarce. On the other hand, the genome of *Vitis vinifera* is available (Jaillon et al 2007) and many genetic markers and genomic resources have been developed during the last two decades (Adam-Blondon et al 2011). Here we present the development of a Bacterial Artificial Chromosome (BAC) library and a collection of BAC end sequences (BES) to facilitate physical mapping in *M. rotundifolia*, comparative analysis with *V. vinifera* and evolutionary analysis in the *Vitis* genus.

## **Material and methods**

### ***Plant material and DNA preparation***

Cuttings of *M. rotundifolia* cv. Regale were grown in a greenhouse at the experimental station of INRA Colmar and put into the dark 24 hours before harvest. 10g of young leaves were harvested, immediately frozen in liquid nitrogen and stored at 80°C. High molecular weight genomic DNA enriched was extracted from frozen leaves according to Peterson et al 2000.

### ***Construction of the BAC library and BAC end sequencing***

A BAC library of *M. rotundifolia* cv Regale was constructed according to a protocol modified from Peterson et al 2000 using *Hind*III and *Bam*HI digested nuclear DNA. Prior to enzymatic digestion, the plugs were washed according to Peterson et al 2000 and incubated for one hour on ice in the incubation buffer: 1X enzyme buffer, 1mM DTT, 100µM BSA, 4mM Spermidine. The plugs were then cut into pieces and dispatched in several tubes containing the same buffer and respectively 1u, 5u, 10u, 15u, 20u, 30u and 60u of either *Hind*III or *Bam*HI. White recombinant colonies were picked and arrayed in 384-well plates containing of 90 µl of 2YT+glycerol (6%) medium supplemented with 12 µg/ml chloramphenicol and 0.4 mM MgSO<sub>4</sub>. Plates were incubated for 42 hours at 37°C and stored at - 80°C before replication.

The average size of the inserts was estimated from 32 clones per sub-library according to Chalhoub et al 2004. Fingerprints were generated for some clones by digesting the plasmid DNA with *Hind*III for 3h at 37°C. The digested DNA was then loaded on an agarose gel 1% SeaKem (GTG). Electrophoresis was performed at 75 V for 16 h in 0.5X TAE.

The BAC-end sequences (BES) were obtained as described in Lamoureux et al 2006 and are available at the <http://urgi.versailles.inra.fr/Species/Vitis/cDNA-and-BACs-libraries> web site. The percentage of inserts corresponding to chloroplastic DNA was estimated through *in silico* analysis of the BES: when the two BES of a clone were aligned on the grapevine chloroplast (Jansen et al 2006) and mitochondrial (Groreymykin et al 2009) sequences on their whole length, the insert was considered as derived from chloroplastic DNA. The clones for which an alignment on their whole length on the

grapevine chloroplastic sequence was obtained only for one BES were also considered as derived from chloroplastic DNA, if no sequence for the other end was available.

### ***Alignment of the BES on the *V. vinifera* reference genome sequence and analysis of their content in coding sequences and in repeats***

Different sets of parameters were used for the alignment of the BES on the *V. vinifera* genome using the BLAT algorithm (Kent 2002). In all cases, the minimum % of identity was 92%, the two extremities of the same clone were aligned in a forward and reverse orientation of the same chromosome and the match was unique. In Filter 1, no additional filter was added. In Filter 2 and Filter 3, the minimum distance between the two paired sequences were respectively 20kb and 45kb, the maximum distance 150kb and 120kb and each sequence was to be aligned on its whole length respectively +/- 100bp and 50bp at each extremities of the sequence. The results can be visualized at <http://urgi.versailles.inra.fr/Projects/Muscares>. The gene content of the BES was estimated by BlastX of masked BES sequences (RepeatMasker version open-3.3.0, viridiplantae RepBase Update 20120418; <http://repeatmasker.org>) against the *Vitis vinifera* proteome (<http://genomes.cribi.unipd.it/grape/>) with a E-Value threshold at 10<sup>-6</sup>, a score >50, an alignment length >20aa. The content in NBS-LRR sequences was estimated by alignment on the BES by BlastX (Evalue<10<sup>-4</sup>) of (i) 48 MEME motifs specific of this gene family (Meyers et al 2003; Kohler et al 2008) and (ii) 70 NBS LRR proteins corresponding to functionally characterized resistance genes (Sanseverino et al 2010; <http://prgdb.crg.eu>; Supplementary table 2). The module “Operate on genomic intervals” of the Galaxy web-based platform ([galaxyproject.org](http://galaxyproject.org)) was used to compare the position of features (gene models, blast/blat hits, etc...) on the reference genome sequence.

### ***Construction of local physical maps***

Three dimensional pools were prepared according to Lamoureux et al 2006 except that the super-pools consisted into 11 super-pools of 12 plates each and one super pool of 9 plates. The pooling of the clones was made using Apricot and TECAN robots. Library screening for specific clones was conducted in a hierarchical fashion beginning with the first 6 super-pools and when no positive clone was found, the 6 other super-pool were screened. 94 markers were chosen on the reference grapevine genetic map (Cipriani et al 2011) or in regions where QTLs for resistance have been detected in studies involving different crosses (Marguerit et al 2009, Welter et al 2007, Akkurt et al 2007, Hoffmann et al 2008, Bellin et al 2009, Barker et al 2005) and are listed in supplementary table 1. 55 new markers were developed from *Muscadinia rotundifolia* BES aligning in specific regions using the following procedure (listed in Supplementary tables 2 and 3). Primer pairs were defined using the Primer 3 software (<http://primer3.sourceforge.net/>) with the following parameters: expected size of the amplified fragment between 100 and 600 bp, primer size between 18 and 25 bases and T<sub>m</sub> between 55°C and

60°C. The uniqueness of the primer pairs was checked by *in silico* amplification with the e-PCR software (Schuler *et al.* 1997) and by BLASTn on the grapevine reference genome sequence.

The pooled clones were diluted 2000 times in TE 10-1 and 4 µl of the dilution was added to a 10-µl reaction mix containing 1.6 mM MgCl<sub>2</sub>, 0.5 U AmpliTaq DNA polymerase, 400 µM dNTP, 1× Taq polymerase buffer, 1.85 ng/µl of each primer and 20% (v/v) of loading buffer [60% (p/v) sucrose, 5 mM Cresol Red in water. Amplifications were performed using a GeneAmp PCR System 9700 (PE Applied Biosystems) programmed as follows: one cycle of 5 min at 94°C; then 10 cycles of 20 s at 94°C, 20 s touchdown from 60°C to 51°C with a decrease of 1°C at each cycle, 20 s at 72°C; a final 40 cycles of 20 s at 94°C, 20 s at 50°C and 40 s at 72°C. The program ended with a 7 min step at 72°C. PCR products were loaded in 0.5× TAE buffer onto 3% MetaPhor (primer testing) or agarose gels (library screening). Electrophoresis were performed at 300 mA for 45 min in 0.5× TAE buffer. The gels were stained with ethidium bromide and pictures were taken using a Gel Doc apparatus (Bio-Rad, Hercules, Calif.).

## Results

### *Construction of a BAC library of M. rotundifolia cv. Regale*

Four sub-libraries were obtained, three using the *Hind*III digested DNA, one using the *Bam*HI digested DNA and stored in one hundred and twelve 384 plates. The *M. rotundifolia* BAC library thus consists of 54,174 clones. The characteristics of each sub-library are given in table 1. The average size of inserts, ranging from 59 to 82kb was lower to the one observed for other grapevine libraries by Adam-Blondon *et al* 2005. The percentage of empty clones, ranging between 6% to 14%, was higher than in this former work whereas the chloroplastic and mitochondrial contaminations, estimated through the alignment of BAC end sequences to the *V. vinifera* chloroplastic and mitochondrial genomes (Jansen *et al* 2006; Groremykin *et al* 2009), was comparable (table 1). Taking into account all these parameters, the present BAC library should represent 6.8X the *M. rotundifolia* genome, giving a 91.5% probability of identifying a clone corresponding to any *M. rotundifolia* DNA sequence.

The coverage was checked through the PCR screening of the library with 91 markers chosen along the 19 chromosomes of the *V. vinifera* reference genetic map (Cipriani *et al* 2011) and/or marking QTL regions for disease resistance in various studies (Marguerit *et al* 2009; Welter *et al* 2007; Akkurt *et al* 2007; Hoffmann *et al* 2008; Bellin *et al* 2009; Barker *et al* 2005). The markers were screened on the 6 first super pools that should correspond to a 3.9X coverage. Eighteen markers did not give any hit in this part of the library (Supplementary table 1). The 73 remaining primer pairs allowed potentially anchoring 609 BAC clones, estimation based on the 3D coordinates (Supplementary table 1). Ten of the primer pairs amplified two distinct bands that were scored independently. Depending on whether these two bands were considered as alleles or duplicated loci, these results respectively corresponded to



an average of 6.6 to 6.1 BAC clones per marker. To check if there was any bias in such an estimation based on pool screening, the individual BACs carrying the loci were identified for 28 of them, yielding a total of 141 individual BAC (Supplementary table 1), which corresponded to an average of 5 BAC clone per locus.

**Table 1.** Characteristics of the *Muscadinia rotundifolia* cv Regale **BAC library**.

Name of the sub-library	Enzyme	Number of clones	Average insert size	% of empty clones	% of chloroplastic and mitochondrial clones	Genome coverage	Number of forward end sequences	Number of reverse end sequences	Total number of BES
MRRE1H	<i>Hind</i> III	15 774	75kb	6.27%	3.8%	2.1	12800	12677	25477
MRRE2H	<i>Hind</i> III	10 368	82kb	6.81%	3.9%	1.5	8683	8600	17283
MRRE1B	<i>Bam</i> HI	13 440	59kb	13.97%	5.2%	1.3	10907	10802	21709
MRRE3H	<i>Hind</i> III	14 592	73kb	9%	3.5%	1.9	11064	11007	22071
<b><i>Whole library</i></b>		<b>54 174</b>	<b>71.8kb</b>	<b>9%</b>	<b>4.1%</b>	<b>6.8</b>	<b>43454</b>	<b>43086</b>	<b>86540</b>

### ***Development and analysis of a collection of M. rotundifolia BAC end sequences (BES)***

After sequencing the ends of all BACs, a total of 86,540 unique BES, 682 bp long in average after trimming, were obtained, corresponding to 44,800 individual BAC clones. Taking into account the percentage of empty clones, this collection represent 88% of the total expected number of BES, with at least one sequence for 91% of the BAC clones of the library. The library was used to make a first assessment of the main features of the Muscadine genome. For this purpose, all the analysis were carried out in parallel in the *M. rotundifolia* BES collection (59 Mb of total sequence length representing 12.4 % of the genome) and in a similar set of 77,237 BES obtained from *V. vinifera* cv Cabernet Sauvignon (Lamoureux et al 2006; 51.8 Mb of total length, representing 10.9 % of the genome sequence). The two collections presented a similar average G+C percent (GC%; table 2).

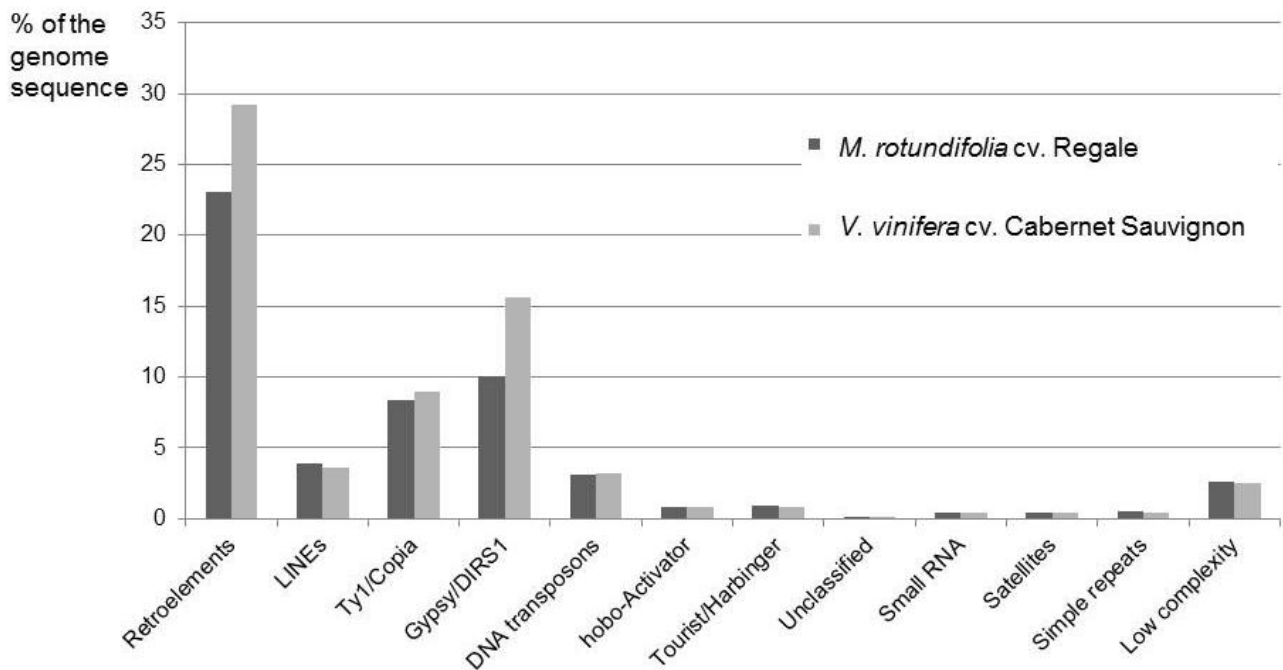
A search for repeated elements was performed and the results are detailed in table 2. The proportion of the BES sequences that corresponded to a repeat was slightly higher in the *V. vinifera* BES collection with 36.32% of the bases found in a repeated sequence for 30.4% in the BES collection from *M. rotundifolia*. Most of the difference could be explained by a higher number of Gypsy elements in the *V. vinifera* collection of sequences, with 5.6% more bases involved in such elements as compared to *M. rotundifolia* (Table 2). Except for this difference, the profiles of repeats of the two sets of BES were similar (Figure 1).

The collection of 33,856 predicted proteins from the *V. vinifera* genome was then used to search for coding regions in the *V. vinifera* and *M. rotundifolia* masked BES. With our criteria of analysis, 38,541 *V. vinifera* and 42,698 *M. rotundifolia* BES carried a sequence potentially coding for respectively 21,571 and 22,403 predicted proteins. Most of these regions corresponded to incomplete genes. MEME motifs specific from the NBS domains (Meyers et al 2003; Köhler et al 2008) and a collection of 70 NBS domains from a list of functionally characterized resistance genes of the NBS-LRR category were used to search for BES containing NBS-LRR sequences in both collections (Supplementary table 2. Similar numbers of BES containing NBS-LRR coding sequences were identified, with 451 (0.5% of the collection) *M. rotundifolia* and 332 *V. vinifera* (0.4% of the collection). A similar proportion of sequences with CNL or TNL signatures was found in both samples: roughly one third of the sequences presented a TIR domain and/or TNL signatures in the NB-ARC domain, one third CNL signatures and no category could be proposed for the other third (data not shown).

**Table 2.** Repeat content assessed with RepeatMasker version open-3.3.0 (sensitive mode) in two sets of BAC end sequences: *M. rotundifolia* cv Regale (86,810 sequences; 35.8% of GC; 59,273,295 bp total length without N among which 17,998,569 masked bases) and *V. vinifera* cv Cabernet Sauvignon (77,237 sequences; 36.2% of GC; 51,832,602 bp total length without N among which 18,827,146 masked bases).

	Regale			Cabernet Sauvignon		
	Number of elements*	Length occupied	% of sequence	Number of elements*	Length occupied	% of sequence
<b>Retroelements</b>	31729	13647891	23,03	34049	15150568	29,23
<b>SINEs:</b>	54	6344	0,01	23	2902	0,01
<b>Penelope</b>	1	49	0	0	0	0
<b>LINEs:</b>	6147	2338850	3,95	4735	1851309	3,57
<b>R2/R4/NeSL</b>	1	42	0	0	0	0
<b>RTE/Bov-B</b>	75	11872	0,02	62	11019	0,02
<b>L1/CIN4</b>	6065	2326597	3,93	4673	1840290	3,55
<b>LTR elements:</b>	25528	11302697	19,07	29291	13296357	25,65
<b>Ty1/Copia</b>	11336	4969091	8,38	10577	4665992	9
<b>Gypsy/DIRS1</b>	13274	5935875	10,01	17584	8078784	15,59
<b>DNA transposons</b>	7738	1862473	3,14	6356	1674926	3,23
<b>hobo-Activator</b>	1713	467249	0,79	1359	413834	0,8
<b>Tc1-IS630-Pogo</b>	34	4516	0,01	25	3150	0,01
<b>En-Spm</b>	3	990	0	10	5029	0,01
<b>MuDR-IS905</b>	14	6837	0,01	12	5351	0,01
<b>Tourist/Harbinger</b>	2711	555790	0,94	1961	403536	0,78
<b>Other (Mirage, P-element, Transib)</b>	2	122	0	0	0	0
<b>Unclassified:</b>	417	109603	0,18	272	81404	0,16
<b>Total TE</b>	<i>39884</i>	<i>15619967</i>	<i>26,35</i>	<i>40677</i>	<i>16906898</i>	<i>32,62</i>
<b>Small RNA:</b>	575	239241	0,4	615	216264	0,42
<b>Satellites:</b>	505	252730	0,43	408	203643	0,39
<b>Simple repeats:</b>	8677	316512	0,53	5832	224418	0,43

<b>Low complexity:</b>	27410	1579471	2,66	22545	1280600	2.47
------------------------	-------	---------	------	-------	---------	------

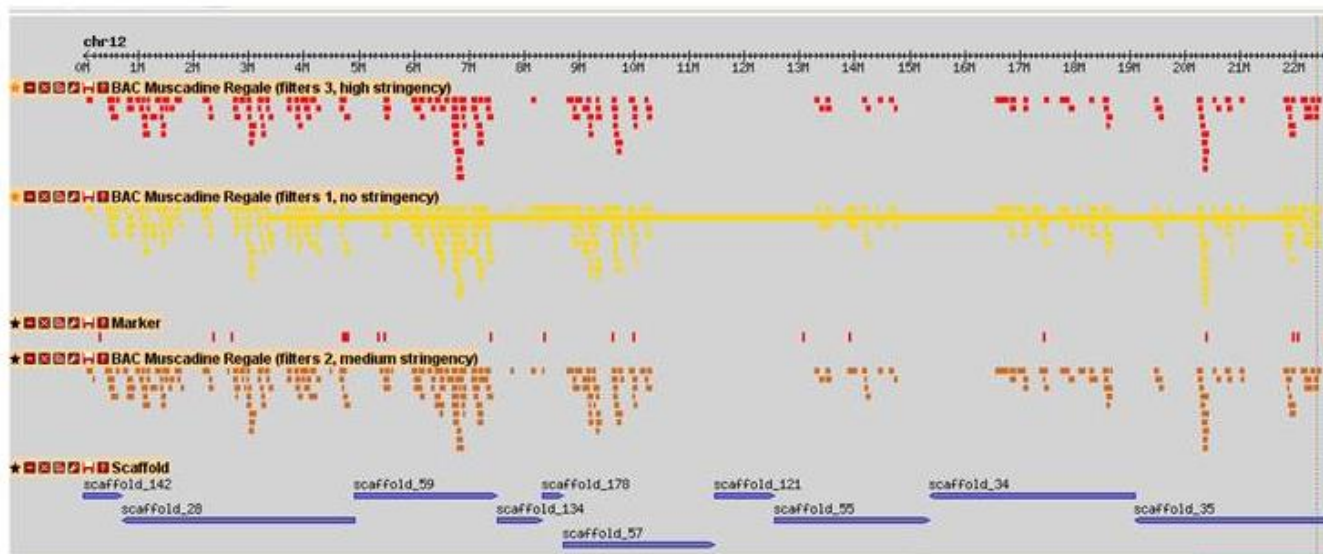


**Figure 1.** Percentage of the BES sequence (without N) involved in different categories of repeats in the *V. vinifera* cv Cabernet Sauvignon and *M. rotundifolia* cv Regale collections of BES.

#### ***Alignment of the M. rotundifolia BES on the V. vinifera reference genome sequence***

In order to better characterize the potential of the BES collection to isolate *M. rotundifolia* genes in syntenic regions, the BES were aligned on the *V. vinifera* reference genome sequence (Figure 2). Out of the whole BES collection, 14,406 BES (7,203 BAC clones, 14.6% of the whole library), 11,754 BES (5877 BAC clones; 12.9% of the whole library) and 8,488 (4,244 BAC clones; 8.6% of the whole library) showed a unique match on a single chromosome using respectively Filter 1, Filter 2 and Filter 3 criteria of alignment on the reference genome sequence (Figure 2). The non-alignment of the rest of the library was either due to one of the expected paired sequences missing (4787 clones, or 9% of the library), to matches of the two paired sequences on more than one chromosome, to rearranged/duplicated areas (as illustrated on Figure 2 with one of the BAC clones which extremities are found at 18.8Mb of distance on chromosome 12 with the Filter 1 parameters). The BACs that were successfully aligned on the *Vitis vinifera* genome formed a total of 1111 (Filter 3) to 1553 (Filter1) contigs spanning *V. vinifera* genome areas of 143 kb in average, half of them being smaller than 100 kb (Table 3). The coverage of the *V. vinifera* chromosomes with the *M. rotundifolia* BAC contigs was variable (Figure 3). The “random chromosomes”, which corresponds to the portion of the genome assembly which is assigned to a chromosome but not oriented and or poorly ordered, and the chUkn, which corresponds to all the scaffolds in the genome assembly that could not be assigned to a chromosome (Adam-Blondon et al 2011), were much less covered (10% in average for the three

criteria of alignment) than the rest of the chromosomes (42% in average for the three criteria of alignment). This is most probably due to both incorrect scaffolds order in the “random” portion of the assembly and to their higher



**Figure 2.** GBrowse view of the alignment of the *M. Rotundifolia* BAC clones using their BAC end sequences on the *V. vinifera* reference genome sequence with three criteria of alignment (see material and methods). A plain bar is drawn between the two BES to underline the region of possible synteny between *M. rotundifolia* and *V. vinifera*. It does not represent the real size of the BAC clone.

proportion of repetitive sequences, which also explains the difficulty to design unique genetic markers for their proper ordering. Depending on the criteria of alignment, the regions of the *V. vinifera* genome on which the BAC contigs aligned contained 19.5% to 44.2% of the predicted gene model on at least 50% of their length while 53.7% to 76.3% are not covered at all (Table 3). The gene models containing a NB-ARC domain are however less well covered in average (Table 3). The 451 *M. rotundifolia* BES containing NBS-LRR sequences were again aligned by blast on the *V. vinifera* genome sequence and the position of the alignments compared to the coordinates of 444 predicted NB-ARC domain containing gene models of the reference genome annotation. 374 out of the 451 BES, of which 229 also contain a partial or total NB-ARC domain, were both aligning on the genome sequence and overlapping with 249 gene models from the reference genome sequence with a NB-ARC domain (Supplementary figure 1). The present resource can therefore help to quickly identify a BAC clone carrying a given gene in *M. rotundifolia* for a quarter to a half of its genes and may also help starting physical mapping studies aiming at the identification of resistance gene using comparative approaches between *V. vinifera* and *M. rotundifolia*.

**Table 3.** *M. rotundifolia* BAC build of contigs by alignment of their BAC ends on the *V. vinifera* reference genome sequence using three different criteria of alignment from the less stringent (Filter1) to the most stringent (Filter3; see material and methods).

	<b>Filter1*</b>	<b>Filter2</b>	<b>Filter3</b>
<b><i>Regions of the V. vinifera reference genome sequence covered by M. rotundifolia BAC Contigs</i></b>			
Number	1553	1340	1111
Average size (kb)	147.1	137.0	145.7
Median size (kb)	100.6	103.2	109.3
Maximum size (kb)	5,269.2	917.5	931.8
Minimum size (kb)	0.6	21.6	46.5
Average % of the region corresponding to a predicted gene	37.7	24.9	25.9
% of <i>V. vinifera</i> predicted genes not covered by a BAC contig	53.7	74.2	76.3
% of <i>V. vinifera</i> predicted genes covered by a BAC contig at least on 50% of their length length (% for the NBARC domain containing gene models)	44.2 (37.6)	23.5 (14.2)	21.6 (13.9)
% of <i>V. vinifera</i> predicted genes covered by a BAC contig on 100% of their length (% for the NBARC domain containing gene models)	42.0 (36.3)	21.0 (12.8)	19.5 (11.9)
<b><i>Gaps</i></b>			
Number	1586	1373	1144
average size (kb)	162.5	220.4	283.5
median size (kb)	51.3	66.4	90.5
maximum size (kb)	7,676.2	17,497.2	17,497.2

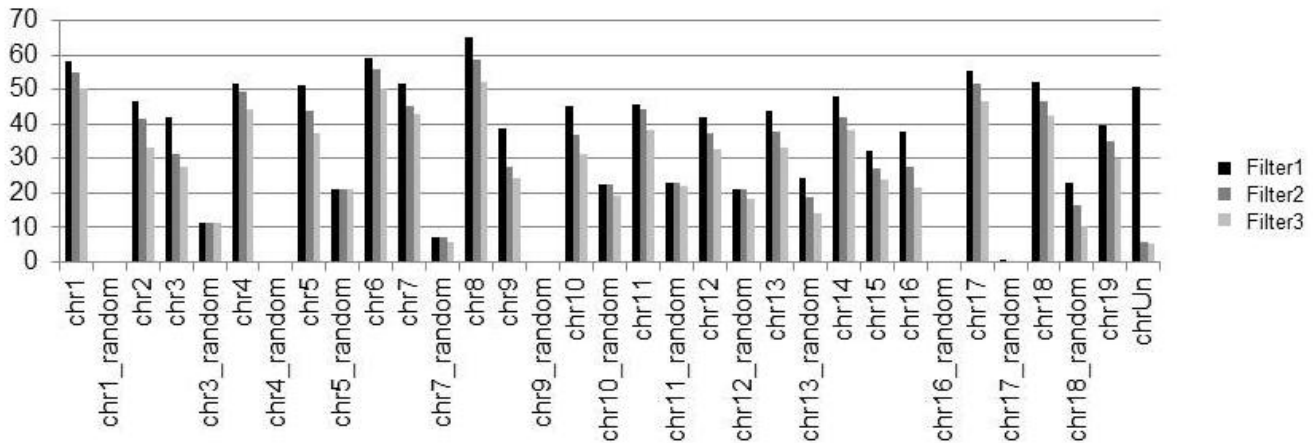
---

minimum size (kp)	0.001	0.001	0.001
-------------------	-------	-------	-------

---

\* The calculations are done after elimination of one BAC supposed to span 18Mb of the *V. vinifera* genome.





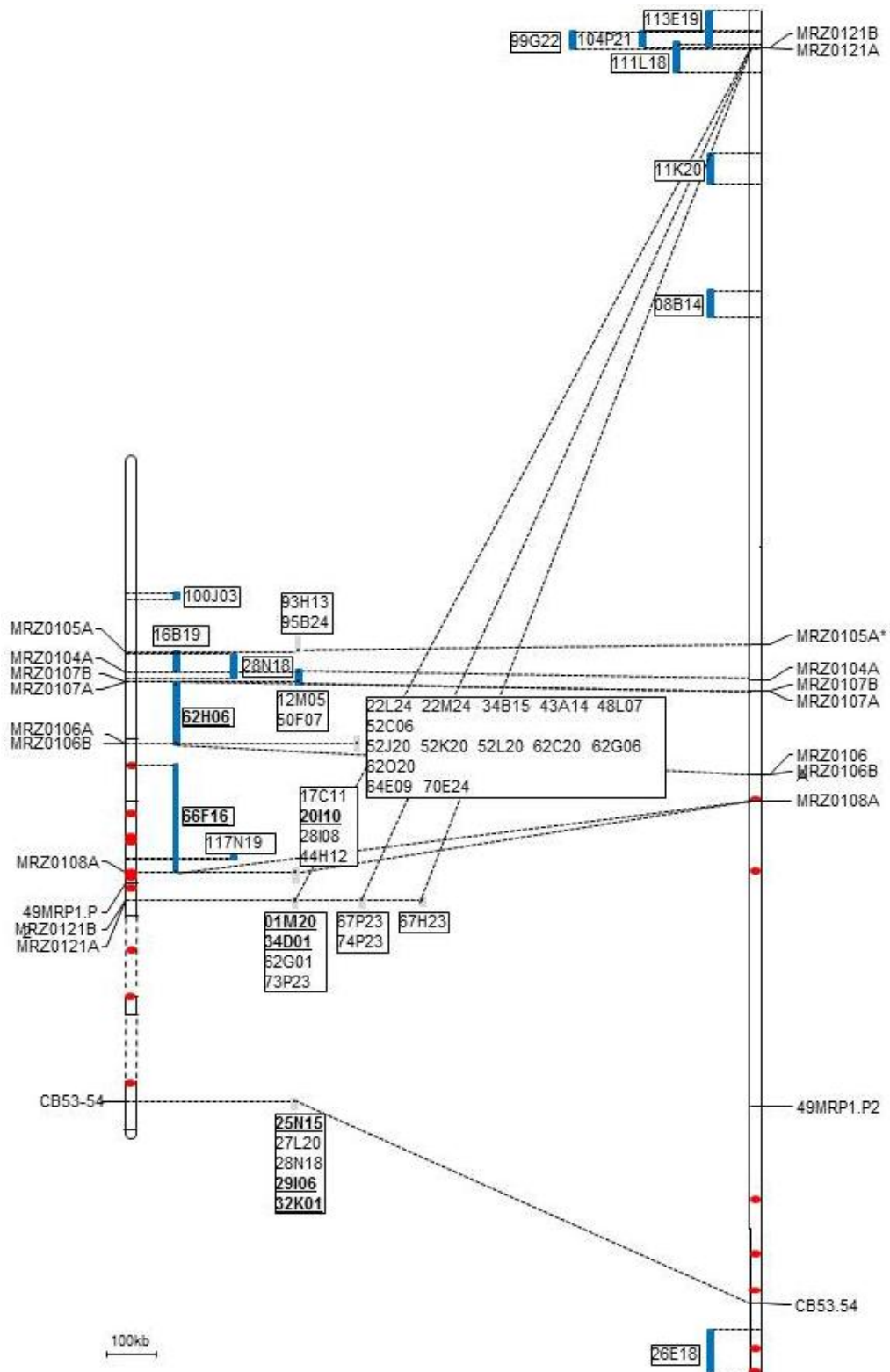
**Figure 3.** Percentage of the *V. vinifera* chromosome sequence covered by the *M. rotundifolia* contigs of BACs depending on the criteria of BES alignment.

### *Comparative physical mapping between M. rotundifolia and V. vinifera in two regions containing clusters of NBS-LRR genes in V. vinifera*

The level of synteny between *M. rotundifolia* and *V. vinifera* was further studied in two regions containing NBS-LRR clusters in *V. vinifera*, one on chromosome 12 and one on chromosome 18.

*Run1*, a single dominant gene for resistance to powdery mildew and *Rpv1*, a partial resistance to downy mildew present at the same locus in *M. rotundifolia*, have been introgressed into *V. vinifera* chromosome 12 (Bouquet 1986; Merdinoglu et al 2003, Barker et al 2005). Genetic and physical mapping allowed Barker et al (2005) to construct a BAC contig from an introgressed individual and spanning a region between the SSR markers VMC4f3.1 and VMC8g9. This BAC contig still contains a gap and corresponds to a region with a cluster of 11 NBS-LRR encoding genes in the introgressed region (Barker et al 2005, Feechan et al 2013). Twenty-five gene models corresponding to a cluster of NBS LRR are found in the current version of the annotation of the corresponding 7,310,990 bp long *V. vinifera* region (Supplementary table 3) and around 50 genes and pseudo genes after manual curation (Feechan et al 2013). First, the SSR markers VMC4F3-1, VMC8G9 and UDV-058 (which was mapped between the two other markers, Cipriani et al 2011) and the markers developed by Barker et al (2005), CB53.54, CB46.49, CB13.14 and 49MRP1.P2, were aligned by e-PCR on the *V. vinifera* genome sequence and on the *M. rotundifolia* BAC-end sequences (Supplementary table 1). CB46.49 and CB13.14 were not found on the *V. vinifera* genome sequence. One BAC (MRRE1B101L21) contained the sequence of VMC8G9 in its extremity. In parallel, 30 BACs were aligned through their BES on the *V. vinifera* reference genome sequence between the markers VMC4F3-1 and VMC8G9 using the most stringent criteria (Filter 3) and 6 on the sequence spanning the *Run1/Rpv1* region. VMC4F3-1, VMC8G9, UDV-058, CB46.49, CB13.14, CB53.54, and 49MRP1.P2 were then used to screen by PCR the whole *M. rotundifolia* BAC library. The amplification of CB46.49, CB13.14 on Regale genomic DNA resulted in smears despite our efforts to adapt the PCR conditions. No positive BAC clone was obtained for 49MRP1.P2 while

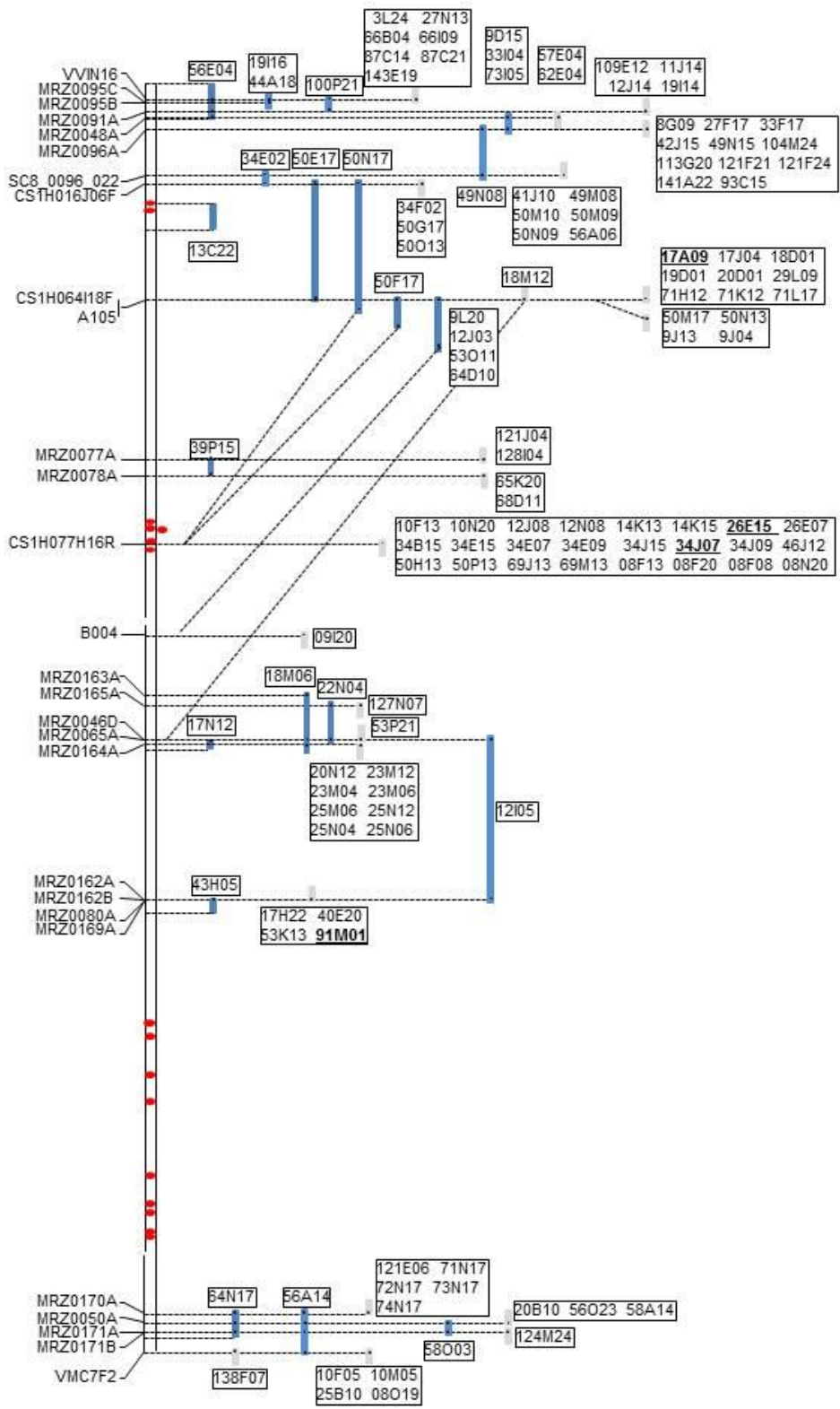
respectively 1, 15, 15 and 5 positive BAC clones were yielded for VMC8G9, VMC4F3-1, UDV-058 and CB53-54 (Supplementary table 1; Figure 4). Fourteen new markers were developed from the BES available from the clones identified and from the sequence of the BAC contig (Supplementary table 4) spanning the *Run1/Rpv1* region (Feechan et al 2013) and used to screen the BAC library and the BES collection by ePCR and by PCR (Supplementary table 4). Two new clones were identified by ePCR and 73 by PCR (Supplementary table 4; Figure 4). Finally the same 14 markers were aligned by ePCR or BLAST on the sequence of the BAC contig from the *Run1/Rpv1* region and on the *V. vinifera* genome reference sequence (Supplementary table 4; Figure 4): 12 (all but the two designed on the *Run1/Rpv1* contig sequence) aligned on the *V. vinifera* genome in the expected region while only 8 aligned on the sequence of the *Run1/Rpv1* region, the two designed using this sequence and 6 of the markers designed from the BES sequences (Supplementary table 4; Figure 4). This showed that the microsynteny in regions containing NBS-LRR is limited between *M. rotundifolia* and *V. vinifera* and underlines the importance of developing BAC libraries from resistant individuals such as *M. rotundifolia*. Differences were also observed between two different *M. rotundifolia* genome sequences as 6 out of the 12 markers developed from the Regale sequences did not match to the corresponding region originating from *M. rotundifolia* G52 (Supplementary table 4; Figure 4). On the whole 116 non redundant BACs were identified in this area and a focus on the *Run1/Rpv1* sequenced region is presented in Figure 4.



**Figure 4.** Result of the screening of the *M. rotundifolia* cv Regale BAC library using several means: BES alignment, ePCR screening and PCR in the region of the *Run1/Rpv1* genes. The left vertical bar corresponds to the contig of sequence of the *Run1/Rpv1* region (Feechan et al 2013) and the right one to the corresponding portion of the *V. vinifera* genome sequence of chromosome 12. The position of the NBS LRR gene models (*V. vinifera* genome sequence) and genes (*Run1/Rpv1* gene region) are figured by plain red circles and the position of markers by black bars. The BAC or groups of BAC clones that were identified are figured by blue (oriented) and grey (non

oriented) bars in the middle. The bold and underlined BAC clone names correspond to BAC clone, which contain sequences homologous to NBS-LRR sequences in their BES.

The same approach was carried out on chromosome 18 along a region where a QTL of resistance to downy mildew was detected in a cross between a complex interspecific hybrid, Bianca, and *V. vinifera* cv. Chardonnay between the SSR markers VVIN16 and VMC7F2 (Bellin et al 2009). First, the SSR markers VVIN16, UDV-305, SC8-00296-022, CS1H016J06F, CS1H064I18F, CS1H077H16R, B004, A105 and VMC7F2 which were developed and eventually mapped in the interval (Bellin et al 2009; Cipriani et al 2011; Canaguier A. unpublished results) were aligned by e-PCR on the *V. vinifera* genome sequence and delimited an interval of 5,733,667 bp between VVIN16 and VMC7F2 (Supplementary table 1) containing 17 NBS-LRR gene models in the current genome annotation (Supplementary table 3) and of 2,028,924 bp between UDV-305 and VMC7F2, which corresponds to the peak of the QTL (Bellin et al 2009). The test of the markers on *M. rotundifolia* cv Regale genomic DNA showed that UDV-305 was multilocus and it was not further used. The ten other markers were used to screen by e-PCR the BES sequences with only one BAC containing the CS1H077H16R1 marker in its extremity (supplementary table 1). In parallel, 5 BAC were aligned through their BES on the *V. vinifera* reference genome sequence between the markers VVIN16 and VMC2F12 (figure 5). Finally, VVIN16, SC8-00296-022, CS1H016J06F, CS1H064I18F, CS1H077H16R, B004, A105 and VMC7F2 were used to screen the *M. rotundifolia* BAC library, yielding 83 positive BAC clones (Supplementary table 1). Twenty new markers were developed from the BES from clones identified in the interval between VVIN16 and VMC7F2 and used to screen the BAC library by PCR and the BES collection by ePCR (Supplementary table 5). Eight clones were identified by ePCR and 92 by PCR (Supplementary table 5; Figure 5). On the whole, 135 non redundant BAC clones were placed in the area in seven contigs (Figure 5). All the markers developed from the *M. rotundifolia* BES could be localized on the *V. vinifera* genome in this case, which was probably due to the fact that they all were developed from BAC which BES aligned on the genome and BAC anchored by markers existing in the *V. vinifera* genome. A rearrangement is suggested between the region delimited by the markers CS1H077H16R and MRZ0046D and the region around the markers CS1H064I18F and A105 (figure 5).



**Figure 5.** Result of the screening of the *M. rotundifolia* cv Regale BAC library using several means: BES alignment, ePCR screening and PCR screening in the region of the *Rpv3* gene. The left vertical bar corresponds to the portion of the *V. vinifera* genome sequence of chromosome 18 containing this locus. The position of the NBS LRR gene models (*V. vinifera* genome sequence) are figured by plain red circles and the position of markers by black bars. The BAC or groups of BAC clones that were identified are figured by blue (oriented) and grey (non-oriented) bars in the middle. The bold and underlined BAC clone names correspond to BAC clone, which contain sequences homologous to NBS-LRR sequences in their BES.

## Discussion

A unique BAC library of *M. rotundifolia* cv. Regale is presented here, which covers, depending of the method assessment: 5 times the genome when the estimation was based on PCR screening experiments with SSR markers to close to 7 times the genome, based on the estimation of the total length of the collection of DNA fragments. The latter calculation was based on the assumption that the size of the *M. rotundifolia* genome is similar to the size of *V. vinifera*. Indeed, flux cytometry-based measurement of the genome size yielded similar size for a large diversity of species and varieties within the *Euveitiss* sub-genus ( $2n=38$ ) but also for an individual from the *Parthenocissus* genus ( $2n=40$ ; Lodhi and Reisch 1995). An individual from the *Ampelopsis* ( $2n=40$ ) genus showed the only significant difference with the rest of the sampling and its genome was 1.2 to 1,6 time bigger (Lodhi and Reisch 1995). The genome size variation in the Vitaceae therefore does not seem to be directly linked to the number of chromosomes. Finally, a first genetic map of the *M. rotundifolia* genome was recently published (Blanc et al 2012) and presented a total size in the range of the size of equivalent genetic maps in *V. vinifera* (Cipriani et al 2011). The alignment of this map on the *V. vinifera* genome showed a very good level of macrosyntheny between the two species (Blanc et al 2012).

A collection of 86,540 non redundant BES was generated from this BAC library, in order to facilitate the use of the BAC library but also to allow a first assessment of the *M. rotundifolia* genome organization by comparison with the well-known genome of *V. vinifera*. In order to take into account possible biases of the sequences sampling in BES collections, the analysis were done in parallel with the *M. rotundifolia* BES collection and with a collection of similar size of *V. vinifera* BES (Lamoureux et al 2006). No striking differences were observed in the two collections in term of GC% or repeats and coding sequences contents. The *M. rotundifolia* GC% (35.8%) was slightly higher to the *V. vinifera* reference genome one (34.5%; Giannuzzi et al 2011) but is still in the range of what is observed in the dicotyledons (e.g., 36% in tomato and 35.6% in potato, Datema et al 2008; 39% in clementine, Terol et al 2008; 35.2% in melon, Gonzalez et al 2010 and 35.3% in Oak, Rampant et al 2011). Consistently with other studies in the *Solanum* genus (Datema et al 2008) and in *Bombyx mori* (Mita et al 2004; Suetsugu *et al.* 2007), we observe a slight over estimation of the GC% in the *V. vinifera* BES collection (35.5%) compared to the same parameter calculated in the entire genome. The most abundant repeats in both BES collections are retro-transposons, as in the entire *V. vinifera* genome (Jaillon et al 2007; Velasco et al 2007). Again, the BES collections seem to offer a slightly biased representation with a possible over-representation of Gypsy elements in the *V. vinifera* BES collection, the *HindIII* sub-collection accounting for most of it (Supplementary figure 2). This would be consistent with another study showing similar biases (Datema et al 2008), although the direction of the biase would be plant family dependent (under-estimation in Datema et al 2008 and over-estimation in the present study) and perhaps, in the present case variety dependent. On the whole, all our observations suggest that the genome organization of *M. rotundifolia* is very close to the *V. vinifera* genome organization. Potentially, the *M. rotundifolia* BES collection and the BAC library will facilitate the access to the coding sequences and their regulatory elements in this species, either because they are partially present in the BES (66% of the *V. vinifera*

predicted gene models) or because they are in a syntenic position with *V. vinifera* and spanned by a BAC contig (19.5% to 44.2% of the gene models, depending of the stringency of the definition of the potential syntenic areas). The extent of microsynteny was further explored using the BAC library and the BES collection, in particular in relation with NBS-LRR gene clusters. There is no indication either of an expansion of the NBS-LRR gene family in the Muscadine genome based on our sampling of NBS LRR in the two BES collections (0.4% to 0.5% of the BES containing NBS LRR related sequences) or of a different balance between the different families of NBS-LRR genes. The use of markers developed from the *M. rotundifolia* BES for ePCR experiments on the *V. vinifera* genome sequence and on a sequence from another *M. rotundifolia* genotype yielded failures in both cases. This is an indication that, even in regions known for a particularly quick evolution with at least one example now documented in grapevine (Feechan et al 2013, the genomes of the two sub-genera are perhaps not so substantially more different than two diverse genotypes of the same species.

### Aknowledgements

This work was funded by the CEA - Institut de Génomique, the Institut National de la Recherche Agronomique (INRA) and the ANR project MUSCARES. The authors thank the Colmar experimental station for its assistance in producing leaves of *M. rotundifolia* cv. Regale.

### References

- Adam-Blondon A-F, Bernole A, Faes G, Lamoureux D, Pateyron S, Grando MS, Caboche M, Velasco R, Chalhoub B (2005) Construction and characterization of BAC libraries from major grapevine cultivars. *Theor Appl Genet*, 110 : 1363-1371
- Adam-Blondon A-F, Jaillon O, Vezzulli S, Zharkikh A, Troggio M, Velasco R (2011) Genome Sequence Initiatives. *in*: A-F Adam-Blondon, JM Martinez-Zapater, Chittaranjan Kole (eds) *Genetics, Genomics and Breeding of Grapes*. Science Publishers and CRC Press. pp 211-234
- Akkurt M, Welter L, Maul E, Töpfer R, Eva Zyprian E (2007) Development of SCAR markers linked to powdery mildew (*Uncinula necator*) resistance in grapevine (*Vitis vinifera* L. and *Vitis* sp.). *Mol Breeding*, 19:103–111
- Barker CL, Donald T, Pauquet J, Ratnaparkhe A, Bouquet A, Adam-Blondon A-F, Thomas MR, Dry I (2005) Genetic and physical mapping of the grapevine powdery mildew resistance gene, *Run1*, using a bacterial artificial chromosome library. *Theor Appl Genet*, 111 : 370-377

- Bellin D, Peressotti E, Merdinoglu D, Wiedemann-Merdinoglu S, Adam-Blondon AF, Cipriani G, Morgante M, Testolin R, Di Gaspero G.(2009) Resistance to *Plasmopara viticola* in grapevine 'Bianca' is controlled by a major dominant gene causing localised necrosis at the infection site. *Theor Appl Genet*, 120 : 163-176
- Blanc S, Wiedemann-Merdinoglu S, Dumas V, Mestre P, Merdinoglu D (2012) A reference genetic map of *Muscadinia rotundifolia* and identification of Ren5, a new major locus for resistance to grapevine powdery mildew. *Theor Appl Genet*, 125:1663–1675
- Boubals D (1966) Etude de la distribution et des causes de la resistance au phylloxera radicicole chez les Vitacees. *Ann. Amelior. Plantes*
- Bouquet A (1986) Introduction dans l'espèce *Vitis vinifera* L. d'un caractère de résistance à l'o (*Uncinula necator* Schw. Burr.) issu de l'espèce *Muscadinia rotundifolia* (Michx.) Small. *Vignevine*, 12 (suppl):141–146
- Bouquet A (2011) Grapevines and viticulture. *in*: A-F Adam-Blondon, JM Martinez-Zapater, Chittaranjan Kole (eds) *Genetics, Genomics and Breeding of Grapes*. Science Publishers and CRC Press. pp 1-29
- Branas M (1932) Sur la caryologie des Ampélidées. *Compte Rendus de l'Académie des Sciences de Paris*, 194 : 121-123
- Cipriani G, Di Gaspero G, Canaguier A, Jusseaume J, Tassin J, Lemainque A, Vincent Thareau V, Adam-Blondon A-F, Testolin R (2011) Molecular Linkage Maps: Strategies, Resources and Achievements. *in*: A-F Adam-Blondon, JM Martinez-Zapater, Chittaranjan Kole (eds) *Genetics, Genomics and Breeding of Grapes*. Science Publishers and CRC Press. pp 111-136
- Datema, E ; Mueller, LA; Buels, R; Giovannoni, JJ; Visser, RGF ; Stiekema, WJ ; van Ham, RCHJ (2008) Comparative BAC end sequence analysis of tomato and potato reveals overrepresentation of specific gene families in potato. *BMC Plant Biol*, 8: 34
- Detjen LR (1919) The limits in hybridization of *Vitis rotundifolia* with related species and genera. *North Carolina Agric. Exp. Sta. Bull.* 17: 1-25
- Di Gaspero G, Copetti D, Coleman C, Castellarin SD, Eibach R, Kozma P, Lacombe T, Gambetta G, Zvyagin A, Cindri P, Kovács L, Morgante M, Testolin R (2012) Selective sweep at the *Rpv3* locus during grapevine breeding for downy mildew resistance. *Theor Appl Genet*, 124:277–286
- Doligez A; Adam-Blondon A-F; Cipriani G; Di Gaspero G; Laucou V; Merdinoglu D; Meredith C; Riaz S; Roux C; This P (2006) An integrated SSR map of grapevine based on five mapping populations. *Theor Appl Genet*, 113 : 369-382
- Eibach R (1994) Investigations about the genetic resources of grapes with regard to resistance characteristics to powdery mildew (*oidium tuckeri*). *Vitis*, 33: 143-150
- Eibach R, Diehl H, Alleweldt G (1989) Untersuchungen zur Vererbung von Resistenzeigenschaften bei Reben gegen *Oidium tuckeri*, *Plasmopara viticola* und *Botrytis cinerea*. *Vitis* 28:209-228

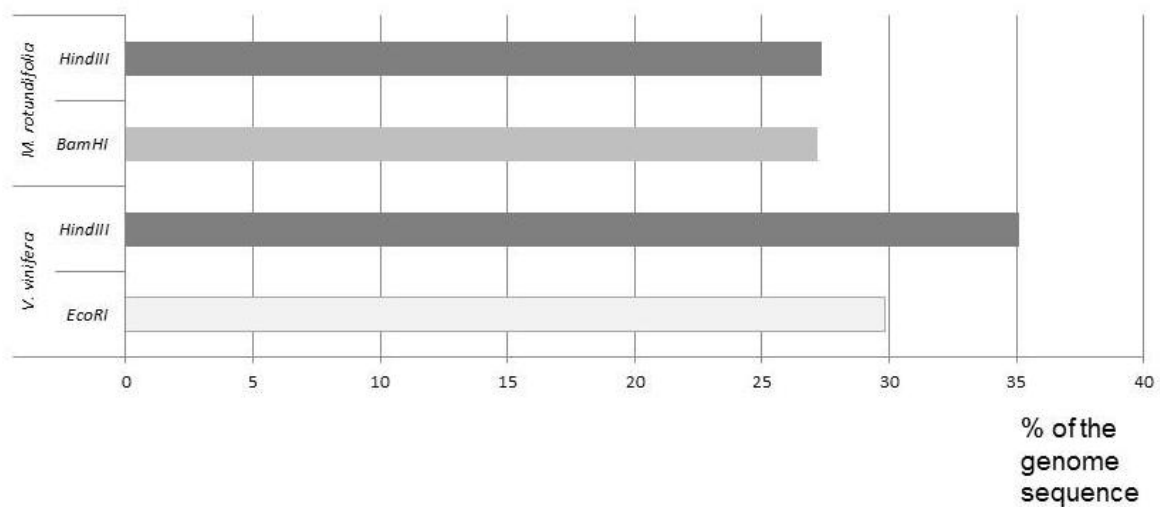


- Esmenjaud D, Bouquet A, Demangeat G, Van Helden M (2011) Nematode-Resistant Rootstocks as a Major Component of the Management Alternative for Grapevine FanLeaf Virus Control in Grape. *Acta hort*, 904: 111-115
- Feechan A, Anderson C, Torregrosa L, Jermakow A, Mestre P, Wiedemann-Merdinoglu S, Merdinoglu D, Walker A, Cadle-Davidson L, Reisch B, Aubourg S, Bentahar N, Shrestha B, Bouquet A, Adam-Blondon A-F, Thomas MR, Dry I (2013) Genetic dissection of a TIR-NB-LRR locus from the wild North American grapevine species *Muscadinia rotundifolia* identifies paralogous genes conferring resistance to major fungal and oomycete pathogens in cultivated grapevine. *Plant J*, 76 : 661–674
- Giannuzzi G, D’Addabbo P, Gasparro M, Martinelli M, Carelli FN, Antonacci D, Ventura M (2011) Analysis of high-identity segmental duplications in the grapevine genome. *BMC Genomics*, 12:436
- Gonzalez VM, Benjak A, Henaff EM, Mir G, Casacuberta JM, Garcia-Mas J, Puigdomenech P (2010) Sequencing of 6.7 Mb of the melon genome using a BAC pooling strategy. *BMC Plant Biol*, 10: 246
- Goremykin VV, Salamini F, Velasco R, Viola R (2009) Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. *Mol Biol Evol*, 26:99-110
- Hofmann S, Di Gaspero G, Kovács L, Howard S, Kiss E, Galbács Z, Testolin T, Kozma P (2008) Resistance to *Erysiphe necator* in the grapevine ‘Kishmish vatkana’ is controlled by a single locus through restriction of hyphal growth. *Theor Appl Genet*, 116:427–438
- Jaillon O, Aury J-M, Noel B, Policriti A, et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449 : 463-468
- Jansen RK, Kaittanis C, Saski C, Lee SB, Tomkins J, Alverson AJ, Daniell H. (2006) Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. *BMC Evol Biol*, 6:32.
- Kent WJ (2002) BLAT : The BLAST-Like Alignment Tool. *Genome Res*, 12: 656-664
- Kohler A, Rinaldi, C, Duplessis S, Baucher M, Geelen D, Duchaussoy F, Meyers BC, Boerjan W Martin F (2008) Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Mol Biol*, 66: 619-636
- Laimer M, Lemaire O, Herrbach E, Goldschmidt V, Minafra A, Bianco P, Wetzel T (2009) Resistance to viruses, phytoplasmas and their vectors in the grapevine in Europe: a review. *J Plant Pathol*, 91: 7-23
- Lamoureux D, Bernole A, Le Clainche I, Tual, S, Thareau V, Paillard S, Legeai F, Dossat C, Wincker P, Oswald M, Merdinoglu D, Vignault C, Delrot S, Caboche M, Chalhoub B, Adam-Blondon A-F (2006) Anchoring a large set of markers onto a BAC library for the development of a draft physical map of the grapevine genome. *Theor Appl Genet*, 113 : 344-356
- Lodhi MA, Reisch BI (1995) Nuclear DNA-content of *Vitis* species, cultivars and other genera of the Vitaceae. *Theor Appl Genet*, 90:11-16

- Marguerit E, Boury C, Manicki A, Donnart M, Butterlin G, Némorin A, Wiedemann-Merdinoglu S, Merdinoglu D, Ollat N, Decroocq S (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. *Theor Appl Genet*, 118:1261–1278
- Meyers BC, Kozik A, Griego A, Kuang HH, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell*, 15:809-834
- Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikumar G, Shimomura M, Nagamura Y, Shin-I T, Abe H, Shimada T, Morishita S, Sasaki T (2004) The genome sequence of silkworm, *Bombyx mori*. *DNA Res*, 11: 27-35
- Moroldo M, Paillard S, Marconi R, Legeai F, Canaguier A, Cruaud C, De Berardinis V, Guichard C, Brunaud V, Le Clainche I, Scalabrin S, Testolin R, Di Gaspero G, Morgante M., Adam-Blondon A.F (2008) A physical map of the heterozygous grapevine ‘Cabernet Sauvignon’ allows mapping candidate genes for disease resistance. *BMC Plant Biol*, 8:66
- Mullins MG, Bouquet A, Williams LE (1992) *Biology of the grapevine*. Cambridge Univ. Press, New York
- Pastrana-Bonilla E, Akoh C, Sellappan S, Krewer G (2003) Phenolic Content and Antioxidant Capacity of Muscadine Grapes. *J. Agric. Food Chem*, 5: 5497-5503
- Patel GI, Olmo HP (1955) Cytogenetics of *Vitis*: I. The hybrid of *V. vinifera* 9 *V. rotundifolia*. *Am J Bot* 42:141–159
- Pauquet, J, Bouquet, A, This, P, Adam-Blondon, A-F (2001) Establishment of a local map of AFLP markers around the powdery mildew resistance gene *Run1* in grapevine and assessment of their usefulness for Marker Assisted Selection, *Theor. Appl. Genet*, 103 : 1201-1210
- Péros J-P, Berger G, Portemont A, Boursiquot J-M, Lacombe T (2011) Genetic variation and biogeography of the disjunct *Vitis* subg. *Vitis* (Vitaceae). *Journal of Biogeography (J. Biogeogr.)* (2011) 38, 471–486
- Peterson, D.G., Tomkins, J.P., Frisch, D.A., Wing, R.A. and Paterson, A.H. (2000) Construction of Plant Bacterial Artificial Chromosome (BAC) Libraries: An Illustrated Guide. *J. Agric. Genomics*, 5. [www.ncgr.org/research/jag](http://www.ncgr.org/research/jag).
- Faivre Rampant P, Lesur I, Boussardon C, Bitton F, Martin-Magniette M-L, Bodénès C, Le Provost G, Bergès H, Fluch S, Kremer A, Plomion C (2011) Analysis of BAC end sequences in oak, a keystone forest tree species, providing insight into the composition of its genome. *BMC Genomics*, 12:292
- Riaz S, Vezzulli S, Harbertson ES, Walker MA (2007) Use of molecular markers to correct grape breeding errors and determine the identity of novel sources of resistance to *Xiphinema index* and Pierce’s disease. *Am. J. Enol. Vitic.* 58:4
- Sanseverino W, Roma G, De Simone M, Faino L, Melito S, Stupka E, Frusciante L, Ercolano MR (2010) PRGdb: a bioinformatics platform for plant resistance gene analysis. *Nucleic Acids Res*, 38(Database issue): D814-21

- Terol J, Naranjo MA, Ollitrault P, Talon M (2008) Development of genomic resources for *Citrus clementina*: Characterization of three deep-coverage BAC libraries and analysis of 46,000 BAC end sequences. *BMC Genomics*, 9: 423
- This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes. *Trends in Genet*, 22: 511-519
- Töpfer R, Hausmann L, Eibach R (2011) Molecular Breeding. *in*: A-F Adam-Blondon, JM Martinez-Zapater, Chittaranjan Kole (eds) *Genetics, Genomics and Breeding of Grapes*. Science Publishers and CRC Press. pp 160-185
- Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, Fitzgerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Demattè L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando SM, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R **2007**. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2:e1326
- Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Toepfer R, Zyprian E (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). *Mol Breed* 20: 359-374.
- Suetsugu Y, Minami H, Shimomura M, Sasanuma S, Narukawa J, Mita K, Yamamoto K (2007) End-sequencing and characterization of silkworm (*Bombyx mori*) bacterial artificial chromosome libraries. *BMC Genomics*, 8:314

**Supplementary material**



**Supplementary figure 1.** Percentage of the BES sequence (without N) involved in retro-elements in the *V. vinifera* cv Cabernet Sauvignon and *M. rotundifolia* cv Regale sub-collections of BES.

**Supplementary table 1:** Markers chosen from grapevine genetic maps used for the screening of the *M. rotundifolia* cv Regale BAC library and results of the screening.

Marker name	Chromosome	Type*	nb 3D\$	nb 1D\$	ePCR size on <i>V. vinifera</i> (bp)	ePCR localization on the <i>V. vinifera</i> reference genome (12X V0)	ePCR localization on <i>M. rotundifolia</i> BES	Marker origin
VVCS1H030L17F	1	chr	4	.	100	chr1:19533484..19533583		SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
VMC4F8	1	chr	7	.	117	chr1:632258..632374		Vitis Microsatellite Consortium
VMC8A7	1	chr	4	.	159	chr1:3423317..3423475		Vitis Microsatellite Consortium
VVIF52	1	chr	A: 17-B:22	.	259	chr1:22273039..22273297		Merdinoglu <i>et al.</i> 2005
UDV-027	2	chr	6	.	177	chr2:3610264..3610440		Di Gaspero <i>et al.</i> 2005
UDV-076	2	chr	0	.	161	chr2:2136446..2136606		Di Gaspero <i>et al.</i> 2005
UDV-061	3	chr	0	.	203	chr3:1985912..1986114		Di Gaspero <i>et al.</i> 2005
VMC8F10	3	chr	0	.	210	chr3:3324131..3324340		Vitis Microsatellite Consortium
VVMD28	3	chr	0	.	237	chr3:11617504..11617740		Bowers <i>et al.</i> 1996, 1999
VMC4D4	4	chr	7	.	161	chr4:3753977..3754137		Vitis Microsatellite Consortium
VMC6G10	4	chr	8	.	178	chr4:23105311..23105488		Vitis Microsatellite Consortium
VMC7H3	4	R	0	.	135	chr4:4719044..4719178		Vitis Microsatellite Consortium
VMCNG2E1	4	R	1	.	109	chr4:5232859..5232967		Vitis Microsatellite Consortium
VVIP77	4	chr	6	.	209	chr4:19317923..19318131		Merdinoglu <i>et al.</i> 2005
VVIR46	4	chr	37	.	377	chr4:1079759..1080135		Merdinoglu <i>et al.</i> 2005
VVCS1E104J11F	5	R	0	.	179	chr5:6769019..6769197	MRRE1H039C16 FM1:199..366	SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
UDV-041	5	R	9	.	186	chr5:6510484..6510669		Di Gaspero <i>et al.</i> 2005
UDV-053	5	R	2	.	120	chr5:8541098..8541217*		Di Gaspero <i>et al.</i> 2005
UDV-106	5	chr	10	.	196	chr5:4373195..4373390		Di Gaspero <i>et al.</i> 2005
VMC4C6	5	chr	3	.	163	chr5:24864769..24864931		Vitis Microsatellite Consortium
VRZAG79	5	R	3	.	234	chr5:5692460..5692693*		Vitis Microsatellite Consortium
VVC6	5	chr	6	.	112	chr5:441633..441744		Decroocq <i>et al.</i> 2003
VVC71	5	R	A: 5- B:3	.	87	chr5:14531568..14531654*		Decroocq <i>et al.</i> 2003
VVIN40	5	chr	6	.	232	chr5:23800844..23801075		Merdinoglu <i>et al.</i> 2005
VVIT68	5	R	3	.	263	chr5:7099947..7100209		Merdinoglu <i>et al.</i> 2005
VVMD27	5	R	9	.	180	chr5:4472022..4472201		Bowers <i>et al.</i> 1996, 1999

VMCNG4B9	6	chr	2	3	178	chr6:18403250..18403427		Vitis Microsatellite Consortium
SC8_0087_070	6	chr	7	3	246	chr6:20263406..20263651		SSR ( <i>V. vinifera</i> PN40024)
VVCS1H087J20R	7	R	9	.	100	chr7:14483946..14484045		SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
SC8_0114_030	7	R	5	.	240	chr7:9430216..9430455		SSR ( <i>V. vinifera</i> PN40024)
VMC8D11	7	R	5	.	135	chrUKN: 30710310..30710444		Vitis Microsatellite Consortium
VRZAG62	7	chr	6	.	235	chr7:1766878..1767112*		Vitis Microsatellite Consortium
VVIN56	7	chr	1	.	175	chr7:19408271..19408445		Merdinoglu <i>et al.</i> 2005
VVMD7	7	chr	0	.	246	chr7:1178552..1178797		Bowers <i>et al.</i> 1996, 1999
VMC2H10	8	chr	3	1	108	chr8:20816914..20817021		Vitis Microsatellite Consortium
UDV-075	8	chr	10	3	147	chr8:2400696,,2400842		Di Gaspero <i>et al.</i> 2005
UDV-132	9	R	1	.	170	chr9:4044565..4044734		Di Gaspero <i>et al.</i> 2005
VVCS1H004K08R	10	chr	6	.	252	chr10:17425418..17425669		SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
SC8_0282_007	10	chr	3	.	150	chr10:1058613..1058762		SSR ( <i>V. vinifera</i> PN40024)
VVIV37	10	chr	0	.	169	chr10:11032334..11032502		Merdinoglu <i>et al.</i> 2005
VVIP02	11	chr	30	.	273	chr11:13539541..13539813		Merdinoglu <i>et al.</i> 2005
VVMD8	11	chr	4	.	154	chr11:19674992..19675145		Bowers <i>et al.</i> 1996, 1999
49MRP1-P2	12	R	0 <sup>E</sup>	0 <sup>E</sup>	208	no e-PCR product		Baker <i>et al.</i> 2005
CB53.54	12	R	3+2 <sup>E</sup>	3+2 <sup>E</sup>	378	chr12:16771129..16771506		Baker <i>et al.</i> 2005
VVCS1E047J11R	12	R	0	.		no e-PCR product		SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
VVCS1E098E15R	12	R	A: 9 -B: 1	6	291	chr12:22026504..22026794	MRRE1H021M04 FM1: 202..504	SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
SC8_0301_003	12	chr / R	7	2	221	chr12: 22579191..22579411		SSR ( <i>V. vinifera</i> PN40024)
UDV-058	12	R	10+12 <sup>E</sup>	6+9 <sup>E</sup>	154	chr12:17434734..17434887		Di Gaspero <i>et al.</i> 2005
<b>VMC4F3-1</b>	12	R	6+4 <sup>E</sup>	2+13 <sup>E</sup>	181	chr12:13067657..13067837		Vitis Microsatellite Consortium
VMC5C6	12	chr	A: 1 -B: 6	4	79	chr12:2682377..2682455		Vitis Microsatellite Consortium
VMC8G6	12	chr	3	2	150	chr12:2354142..2354291		Vitis Microsatellite Consortium
<b>VMC8G9</b>	12	R	0+1 <sup>E</sup>	1 <sup>E</sup>	172	chr12:20378485..20378656	MRRE1B101L21F M1: 415..551	Vitis Microsatellite Consortium
SC8_0149_025	13	chr	8	.	330	chr13:24031290..24031619		SSR ( <i>V. vinifera</i> PN40024)
VMC8E6	13	chr	5	.	254	chr13:22501796..22502049		Vitis Microsatellite Consortium

VVIH54	13	chr	0	.	166	chr13:3333487..3333652	Merdinoglu <i>et al.</i> 2005
VMC9H4-2	13	R	20	.	no e-PCR product	no e-PCR product	Vitis Microsatellite Consortium
VMCNG4E10-1	13	R	5	.	246	chr13:18413046..18413291	Vitis Microsatellite Consortium
UDV-095	14	R	3	.	140	chr14:26102129..26102268	Di Gaspero <i>et al.</i> 2005
VMC2A5	14	R	7	.	186	chr14:24080928..24081113	Vitis Microsatellite Consortium
VVC62	14	chr	14	.	206	chr14:1426793..1426998	SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
VVII51	14	chr	6	.	262	chr14:28772163..28772424	Merdinoglu <i>et al.</i> 2005
VVIN94	14	chr	6	.	290	chr14:28772134..28772423	Merdinoglu <i>et al.</i> 2005
VVIP05	14	chr	0	.	250	chr14:3275203..3275452	Merdinoglu <i>et al.</i> 2005
VVIV69	14	R	7	.	288	chr14:24080814..24081101	Merdinoglu <i>et al.</i> 2005
SC8_0040_088	15	chr	14	.	330	chr15:18924868..18925197	SSR ( <i>V. vinifera</i> PN40024)
ScORA7-32	15	R	0	.	no e-PCR product	no e-PCR product	Akkurt <i>et al.</i> 2007
ScORN3-R	15	R	0	.	no e-PCR product	no e-PCR product	Akkurt <i>et al.</i> 2007
VMC8G3-2	15	chr	1	.	292	chr15:17666601..17666892	Vitis Microsatellite Consortium
VVIV67	15	R	3	.	367	chr15:10897392..10897758	Merdinoglu <i>et al.</i> 2005
VVIN52	16	chr	8	.	104	chr16:619678..619781	Merdinoglu <i>et al.</i> 2005
VVMD5	16	chr	5	.	227	chr16:20785556..20785782	Bowers <i>et al.</i> 1996, 1999
VVIB09	17	chr	0	.	280	chr17:9295347..9295626	Merdinoglu <i>et al.</i> 2005
VVIP16	17	chr	14	.	106	chr17:11011994..11012099	Merdinoglu <i>et al.</i> 2005
VVSCU06	17	chr	0	.	171	chr17:3158735..3158905	Scott <i>et al.</i> 2000
VVMD17	18	R	8	7	211	chr18:19873209..19873419	Bowers <i>et al.</i> 1996, 1999
VMCNG2F12	18	R	11	3	199	chr18:19902915..19903113	Vitis Microsatellite Consortium
SC8_0092_063	18	R	25	.	122	chr18:22627154..22627275	SSR ( <i>V. vinifera</i> PN40024)
VVIN16	18	R	13+4 <sup>E</sup>	7+4 <sup>E</sup>	161	chr18:23389686..23389846	Merdinoglu <i>et al.</i> 2005
SC8_0096_022	18	R	8	8	216	chr18:23605106..23605321	SSR ( <i>V. vinifera</i> PN40024)
CS1H016J06F	18	R	9	6	222	chr18:23626880..23627101	SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
CS1H064I18F	18	R	A:5-B:14	9	293	chr18:23952419..23952711	SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
CS1H077H16R	18	R	A:3-B:13	28	100	chr18:24636885..24636984	MRRE1H026E15 RM1: 624..723; MRRE2H061M16 FM1: 384..486 SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
UDV-305	18	R	-	.	361	chr18:24868065..24868384	Di Gaspero <i>et al.</i> 2005
B004	18	R	9	5	393	chr18:24889316..24889708	SSR marker ( <i>V. vinifera</i> )
A105	18	R	9	10	273	chr18:24889436..24889708	SSR marker ( <i>V. vinifera</i> )
VMC7F2	18	chr / R	6+4 <sup>E</sup>	5+1 <sup>E</sup>	200	chr18:26896790..26896989	Vitis Microsatellite Consortium
UDV-108	18	chr / R	5+4 <sup>E</sup>	6+4 <sup>E</sup>	241	chr18:29123113..29123353	Di Gaspero <i>et al.</i> 2005
CS1H085H20R	18	chr	2	3	367	chr18:877731..878097	SSR from <i>V. vinifera</i> cv Cabernet Sauvignon BES
SC8_0096_097	18	R	9	5	332	chr18_random:3098265..3098596	SSR ( <i>V. vinifera</i> PN40024)
VMC2A7	18	R	3+12	1+4 <sup>E</sup>	143	chr18:13397376..13397518	Vitis Microsatellite Consortium
VVIN74	19	chr	5	3	290	chr19:213243..213532	Merdinoglu <i>et al.</i> 2005

\*chr: markers chosen to represent the chromosome; R: markers bordering a gene/QTL for disease resistance

<sup>s</sup>nb 3D: estimated number of positive BAC clones based on 3D pools screening; nb 1D: number of positive clones found from the individual test of all the 3D candidates

<sup>E</sup>the whole library tested

A-B : when two bands of different sizes are observed

**Supplementary table 2.** List of 70 functionally characterized NBS LRR resistance genes used for searching BES containing NBS LRR sequences

Voir Annexe 3 de la thèse



**Supplementary table 3.** NBS LRR gene models of the current annotation of the *V. vinifera* genome annotation in two regions of interest on chromosome 12 and chromosome 18. The position of the SSR markers that delimit the region is also indicated.

Gene id/Marker name	Chromosome	Gene/Marker	Start	Stop	Pfam domain id	Pfam domain
<b>VMC4F3-1</b>	<b>chr12</b>	<b>Marker</b>	<b>13067657</b>	<b>13067837</b>		
Vv12s0034g00270.t01	chr12	Gene	15747458	15748106	PF01582	TIR domain
Vv12s0034g00370.t01	chr12	Gene	15894835	15895537	PF01582	TIR domain
Vv12s0034g00760.t01	chr12	Gene	16561902	16567192	PF01582	TIR domain
Vv12s0034g00850.t01	chr12	Gene	16671797	16677604	PF00931, PF01582	NB-ARC domain, TIR domain
Vv12s0034g00910.t01	chr12	Gene	16746048	16753289	PF00931, PF01582	NB-ARC domain, TIR domain
<b>CB53.54</b>	<b>chr12</b>	<b>Marker</b>	<b>16771129</b>	<b>16771506</b>		
Vv12s0034g01030.t01	chr12	Gene	16865807	16881373	PF00931	NB-ARC domain
Vv12s0034g01070.t01	chr12	Gene	16912313	16915246	PF00931	NB-ARC domain
Vv12s0034g01250.t01	chr12	Gene	17238554	17239125	PF01582	TIR domain
Vv12s0034g01260.t01	chr12	Gene	17258174	17266288	PF01582	TIR domain
Vv12s0034g01270.t01	chr12	Gene	17283683	17289426	PF00931, PF01582	NB-ARC domain, TIR domain
Vv12s0034g01310.t01	chr12	Gene	17317387	17318217	PF01582	TIR domain
Vv12s0034g01350.t01	chr12	Gene	17338041	17344555	PF00931, PF01582	NB-ARC domain, TIR domain
<b>UDV-058</b>	<b>chr12</b>	<b>Marker</b>	<b>17434734</b>	<b>17434887</b>		
Vv12s0034g01460.t01	chr12	Gene	17505429	17508679	PF00931	NB-ARC domain
Vv12s0034g01470.t01	chr12	Gene	17510489	17521321	PF00931	NB-ARC domain
Vv12s0034g01480.t01	chr12	Gene	17531014	17535018	PF00931	NB-ARC domain
Vv12s0034g01490.t01	chr12	Gene	17549019	17552648	PF00931	NB-ARC domain
Vv12s0034g01580.t01	chr12	Gene	17637984	17640074	PF00931	NB-ARC domain
Vv12s0034g01660.t01	chr12	Gene	17777713	17780786	PF00931	NB-ARC domain
Vv12s0034g01750.t01	chr12	Gene	17879196	17887865	PF00931	NB-ARC domain
Vv12s0034g02310.t01	chr12	Gene	18720504	18724583	PF00931	NB-ARC domain
Vv12s0034g02340.t01	chr12	Gene	18771329	18773036	PF00931	NB-ARC domain
Vv12s0034g02400.t01	chr12	Gene	18903515	18910168	PF00931	NB-ARC domain
Vv12s0034g02440.t01	chr12	Gene	18937893	18944121	PF00931	NB-ARC domain
Vv12s0034g02500.t01	chr12	Gene	18978545	18982936	PF00931	NB-ARC domain
Vv12s0035g00440.t01	chr12	Gene	19766685	19771021	PF00931	NB-ARC domain
<b>VMC8G9</b>	<b>chr12</b>	<b>Marker</b>	<b>20378485</b>	<b>20378656</b>		
<b>VVIN16</b>	<b>chr18</b>	<b>Marker</b>	<b>23389686</b>	<b>23389846</b>		
<b>SC8_0096_022</b>	<b>chr18</b>	<b>Marker</b>	<b>23605106</b>	<b>23605321</b>		
<b>VVCS1H016J06F1-1</b>	<b>chr18</b>	<b>Marker</b>	<b>23626880</b>	<b>23627101</b>		
Vv18s0117g00360.t01	chr18	Gene	23682046	23686889	PF00931, PF01582	NB-ARC domain, TIR domain
Vv18s0117g00380.t01	chr18	Gene	23701169	23701660	PF01582	TIR domain
<b>VVCS1H014E08R1-1</b>	<b>chr18</b>	<b>Marker</b>	<b>23949454</b>	<b>23949810</b>		

<b>VVCS1H064I18F1-1</b>	<b>chr18</b>	<b>Marker</b>	<b>23952419</b>	<b>23952711</b>		
Vv18s0041g00190.t01	chr18	Gene	24574674	24580098	PF00931, PF01582	NB-ARC domain, TIR domain
Vv18s0041g00200.t01	chr18	Gene	24591355	24594186	PF00931, PF07725	NB-ARC domain, Leucine Rich Repeat
Vv18s0041g00210.t01	chr18	Gene	24595027	24596218	PF01582	TIR domain
Vv18s0041g00250.t01	chr18	Gene	24633183	24636942	PF00931, PF01582	NB-ARC domain, TIR domain
<b>VVCS1H077H16R1-1</b>	<b>chr18</b>	<b>Marker</b>	<b>24636885</b>	<b>24636984</b>		
Vv18s0041g00270.t01	chr18	Gene	24653492	24657134	PF00931	NB-ARC domain
<b>B004</b>	<b>chr18</b>	<b>Marker</b>	<b>24889316</b>	<b>24889708</b>		
<b>A105</b>	<b>chr18</b>	<b>Marker</b>	<b>24889436</b>	<b>24889708</b>		
Vv18s0041g01330.t01	chr18	Gene	25979007	26000298	PF00931, PF01582	NB-ARC domain, TIR domain
Vv18s0041g01340.t01	chr18	Gene	26016813	26021251	PF00931, PF01582	NB-ARC domain, TIR domain
Vv18s0041g01380.t01	chr18	Gene	26121589	26124946	PF01582	TIR domain
Vv18s0041g01440.t01	chr18	Gene	26197203	26223822	PF01582	TIR domain
Vv18s0041g01570.t01	chr18	Gene	26402820	26405256	PF01582	TIR domain
Vv18s0041g01620.t01	chr18	Gene	26481427	26487339	PF00931, PF07725	NB-ARC domain, Leucine Rich Repeat
Vv18s0041g01650.t01	chr18	Gene	26507474	26507989	PF01582	TIR domain
Vv18s0041g01710.t01	chr18	Gene	26562393	26562908	PF01582	TIR domain
Vv18s0041g01750.t01	chr18	Gene	26576660	26585729	PF00931, PF07725	NB-ARC domain, Leucine Rich Repeat
<b>VMC7F2</b>	<b>chr18</b>	<b>Marker</b>	<b>26896790</b>	<b>26896989</b>		

Supplementary table 4. STS markers developed from *M. rotundifolia* sequences aligning on *V. vinifera* chromosome 12

Marker name	nb 1D <sup>s</sup>	ePCR size on <i>V. vinifera</i> (bp)	ePCR localization on the <i>V. vinifera</i> reference genome (12X V0)	ePCR size on the Run1/Rpv1 sequenced region (bp)	ePCR localization on the introgressed <i>Run1</i> region (size= 886114 bp)	ePCR localization on <i>M. rotundifolia</i> BES (bp)	Marker origin
MRZ0121A	6	no e-PCR product	no e-PCR product	111	Contig_Run1: 883691..883801	no e-PCR product	Contig_Run1 (Feechan et al, in prep)
MRZ0121B	5	no e-PCR product	no e-PCR product	112	Contig_Run1: 883690..883801	no e-PCR product	Contig_Run1 (Feechan et al, in prep)
MRZ0104A	3	126	chr12:15504729..15504854	126	Contig_Run1:423011..423136	MRRE1H016B19FM1:259..384	MRRE1H016B19FM1
MRZ0105A	2*	182	chr12:15433396..15433577	182	Contig_Run1:380157..380338	MRRE1H016B19RM1:317..498	MRRE1H016B19RM1
MRZ0106A	15	225	chr12:15696209..15696432	224	Contig_Run1:568181..568404	MRRE2H062H06FM1:166..390	MRRE2H062H06FM1
MRZ0106B	1	211	chr12:15696375..15696583	211	Contig_Run1:568347..568557	MRRE2H062H06FM1:13..223	MRRE2H062H06FM1
MRZ0107A	3	178	chr12:15528835..15529012	178	Contig_Run1:442263..442440	MRRE2H062H06RM1:237..414	MRRE2H062H06RM1
MRZ0108A	5	227	chr12R:470125..470351	227	Contig_Run1:829711..829937	MRRE1H020I10RM1:256..473 MRRE2H066F16RM1:256..482	MRRE2H066F16RM1
MRZ0156A	4	282	chr12:15865423..15865702	no e-PCR product	no e-PCR product	MRRE2H059I15RM1:251..532 MRRE1B073P23FM1:430..711	MRRE1B073P23FM1
MRZ0156B	4*	195	chr12:15865452..15865644	no e-PCR product	no e-PCR product	MRRE2H059I15RM1:309..503 MRRE1B073P23FM1:488..682	MRRE1B073P23FM1
MRZ0157A	6	138	chr12:17430927..17431061	no e-PCR product	no e-PCR product	MRRE1H011K09FM1:93..230	MRRE1H011K09FM1
MRZ0158A	5	206	chr12:17494983..17495200	no e-PCR product	no e-PCR product	MRRE1H020O16RM1:263..468	MRRE1H020O16RM1
MRZ0160B	9	216	chr12:17365172..17365386	no e-PCR product	no e-PCR product	MRRE2H055G20FM1:360..575	MRRE2H055G20FM1
MRZ0160C	5	141	chr12:17365202..17365342	no e-PCR product	no e-PCR product	MRRE2H055G20FM1:391..531	MRRE2H055G20FM1

<sup>s</sup>nb 1D: number of positive clones found from the individual test of all the 3D candidates; \*the whole library tested

**Supplementary table 5.** STS markers developed from *M. rotundifolia* BAC end sequences aligning on *V. vinifera* chromosome 18

Marker name	nb 1D <sup>§</sup>	Marker origin	ePCR size on <i>V. vinifera</i> (bp)	ePCR localization on the <i>V. vinifera</i> reference genome	ePCR localization on <i>M. rotundifolia</i> BES
MRZ0095C	3	MRRE3H120J20RM1	182	chr18:23401840..23402021	MRRE3H121F24FM1:285..466
MRZ0095B	3	MRRE3H120J20RM1	157	chr18:23401865..23402021	MRRE3H121F24FM1:310..466
MRZ0091A	4	MRRE3H109E12FM1	176	chr18:23424402..23424577	MRRE1B100P21FM1:125..300 MRRE3H109E12FM1:504..679
MRZ0048A	6	MRRE1B076J20FM1	202	chr18:23443542..23443743	MRRE1B076J20FM1:228..429
MRZ0096A	20*	MRRE3H121F24RM1	302	chr18:23474614..23474915	MRRE3H121F24RM1:316..619
MRZ0077A	3	MRRE1H039P15FM1	194	chr18:24399691..24399884**	MRRE1H039P15FM1:153..346
MRZ0078A	3*	MRRE1H039P15RM1	160	chr18:24445364..24445523**	MRRE1H039P15RM1:200..359
MRZ0163A	1	MRRE1H018M06FM1	208	chr18:25058276..25058483	MRRE1H018M06FM1:272..480
MRZ0165A	0*	MRRE1H022N04RM1	208	chr18:25086114..25086321	MRRE1H022N04RM1:162..369 MRRE3H127N07RM1:162..369
MRZ0046D	5	MRRE1H017N12RM1	152	chr18:25182566..25182734	MRRE1H017N12RM1:430..581
MRZ0065A	4	MRRE1H017N12RM1	180	chr18:25182739..25182919	MRRE1H017N12RM1:586..765
MRZ0164A	10	MRRE1H018M06RM1	112	chr18:25195784..25195895	MRRE1H018M06RM1:566..677
MRZ0162A	4	MRRE1H017H22FM1	176	chr18:25631226..25631401	MRRE1H017H22FM1:421..596
MRZ0162B	4	MRRE1H017H22FM1	151	chr18:25631321..25631471	MRRE1H017H22FM1:516..666
MRZ0080A	5	MRRE2H043H05RM1	163	chr18:25633486..25633648	MRRE2H043H05RM1:436..598 MRRE1B091M01RM1:220..382 MRRE2H043H05RM1:576..811
MRZ0169A	1	MRRE2H053K13FM1	236	chr18:25633626..25633861	MRRE2H053K13FM1:218..453 MRRE1B091M01RM1:360..595 MRRE2H056A14FM1:348..484 MRRE2H064N17RM1:348..484
MRZ0170A	7*	MRRE2H056A14FM1	137	chr18:26792873..26793009	MRRE2H056A14FM1:348..484 MRRE2H064N17RM1:348..484
MRZ0050A	6	MRRE1B081L13FM1	220	chr18:26817620..26817839	MRRE1B081L13FM1:53..272
MRZ0171A	2	MRRE2H058O03FM1	244	chr18:26845670..26845912	MRRE2H058O03FM1:264..507 MRRE3H124M24FM1:160..403
MRZ0171B	1	MRRE2H058O03FM1	179	chr18:26845695..26845873	MRRE2H058O03FM1:303..482 MRRE3H124M24FM1:199..378

<sup>§</sup>nb 1D: number of positive clones found from the individual test of all the 3D candidates

\*all the library tested

\*\* Blast alignment