BOUVET GUILLAUME

CARACTÉRISATIONS ET IMPACTS DES TRANSPOSONS À ADN CHEZ OPHIOSTOMA ULMI ET O. NOVO-ULMI, PRINCIPAUX AGENTS DE LA MALADIE HOLLANDAISE DE L'ORME

Thèse présentée

à la faculté des études supérieures de l'Université Laval dans le cadre du programme de doctorat en sciences forestières pour l'obtention du grade de Philosophiae Doctor (Ph.D.)

> DÉPARTEMENT DE FORESTERIE SCIENCES FORESTIÈRES UNIVERSITÉ LAVAL QUÉBEC

> > 2007

© Bouvet Guillaume, 2007

Remerciements...

Initialement, ma thèse devait porter sur la saturation d'une carte génétique dans le but de se rapprocher du mystique locus Pat-1, cartographié par Et-Touil en 1999. Je tiens, en tout premier lieu, à remercier le Prof. Louis Bernier, mon directeur de thèse, pour m'avoir laissé la possibilité de modifier ce projet initial et de l'avoir complèment converti, pour aboutir à cette thèse sur les éléments mobiles du génome des champignons de la maladie hollandaise de l'orme. Sans son ouverture d'esprit, qui va de pair avec la recherche scientifique, cette thèse dans son ensemble n'aurait pu avoir lieu.

Je tiens également à remercier le Dr Volker Jacobi qui a bien voulu devenir le co-directeur de cette thèse en cours de route et pour avoir fait preuve d'un très grand sens critique face mes diverses « découvertes ». Son sens critique m'a permis, au cours des années, de développer le mien.

Cette thèse a bien sur fait l'objet d'un évaluation par 3 examinateurs externes que je voudrais maintenant remercier : le Prof. François Belzile, pour son expertise sur les éléments mobiles ainsi que pour ces conseils au niveau de la partie transcriptionnelle des transposons OPHIO. Le Prof. Roger C. Levesque pour m'avoir permis de mettre un pied dans la biologie humaine et pour m'avoir aiguillé lors de certains de mes choix et enfin le Dr. Ken Dewar pour avoir été l'examinateur externe, attendu, sur ce travail.

Une thèse ne peut être effectuée sans l'aide de personnes d'encadrement, de soutien de laboratoire et de gestion, c'est pour cela que je voudrais remercier, dans le désordre Sauphie, Ginette, Josée ainsi que l'ensemble de la Bernier Team® avec qui mes rapports scientifiques et amicaux furent enrichissants : Karine, Marie-Ève, Mirella et bien d'autres.

Une thèse ne peut également voir le jour sans les mille et une rencontres hasardeuses ou non qu'elle génère. Merci donc à Claude et Annie, Frink, Sylvain, Greg, JLJ, Juan Pablo et plus récemment Julien, Seb... pardon à ceux que j'oublie.

Enfin une thèse ne peut avoir lieu sans l'énergie créatrice initiale... donc merci à mes parents de ne jamais m'avoir empêché de réaliser mes rêves et merci à l'énergie création fondamentale, Marie-Josée.

Le cheval est-il un âne complexe,

Où l'âne est-il un cheval libéré ?

Remarque préliminaire : l'abréviation « TE » (pour « *Transposable element* ») a été utilisée tout au long de cette thèse et décrit les éléments de transposition dans leur ensemble, à savoir les éléments de type 1 et de type 2.

Résumé

Des éléments mobiles de type 2 ont été mis en évidence chez Ophiostoma ulmi et O. novo*ulmi*, agents pathogènes de la maladie hollandaise de l'orme. Dans un premier temps, ces transposons à ADN, nommés OPHIO1, OPHIO2 et OPHIO3, ont été subséquemment caractérisés, tant au niveau structural qu'au niveau de leur répartition dans les espèces concernées. L'étude approfondie de leur séquence a permis de faire ressortir des mutations particulières de type RIP (Repeat induced point mutations) uniquement présentes chez OPHIO3. Ces dernières, ont par ailleurs, permis de mettre au point une nouvelle technique de visualisation de ce type de mutations (CTS visualization), applicable à l'ensemble de ces éléments mobiles. Dans un second temps, la mobilité d'OPHIO1 et OPHIO2 a fait l'objet d'une étude détaillée. Grâce à divers stress abiotiques, nous avons démontré que ces éléments sont mobiles au sein des génomes des champignons responsables de la maladie hollandaise de l'orme. En dernier lieu, des analyses bioinformatiques ont permis de mettre en évidence des zones de sélection positive au sein de domaines précis des transposases, l'enzyme requise pour la mobilité d'OPHIO1 et d'OPHIO2. L'ensemble de ces résultats a permis d'accroître la connaissance des TE ainsi que d'essayer de comprendre la dynamique complexe existant entre les transposons et leurs génomes hôtes.

Abstract

Type 2 mobile elements were detected in *Ophiostoma ulmi* and *O. novo-ulmi* sp., the causal agents of the Dutch elm disease. Firstly, the structure and the distribution in *Ophiostoma* species of these DNA transposons, named *OPHIO1*, *OPHIO2* and *OPHIO3*, were characterized. A precise analysis of their sequence demonstrated some particular RIP mutations (*Repeat induced point mutations*) in the case of *OPHIO3*. These mutations allowed us to develop a new visualization (*CTS visualization*) for these types of mutations, applicable to all mobile elements. Secondly, the mobility of *OPHIO1* and *OPHIO2* was the main investigation of a detailed study. We demonstrated that abiotic stresses have a direct impact on the induction of mobility of the transposons within *Ophiostoma* sp. To finish our investigation, bioinformatics analyses were performed on *OPHIO1* and *OPHIO2* (considered to be active transposons) and allowed the presence of regions under positive selection inside the transposase (enzyme required for their mobility). Taken together, these results lead to a better understanding of a part of the complex dynamics that links mobile elements to their host genomes.

TABLE DES MATIÈRES

RÉSUMÉ	III
ABSTRACT	IV
TABLE DES MATIÈRES	V
GÉNÉRALITÉS	VIII
FIGURES ET SCHÉMAS	IX
TABLEAUX	X
LINTRODUCTION	1
Ι 1 Ρογαματική το	1
I 1 1 Préambule	
1.1.1.1 reambuie	1
1.1.2.1. Les régétitions estallites	
1.1.2.2. Les répétitions salemes.	3 1
1.1.2.2. Les repeutoris geantes.	4
I.1.2.4. Les pseudogenes	4 1
Les rétrotronsposons ou éléments de type 1 (mobilité de type « conjer/coller »)	4
Les transposons à ADN ou éléments de type 2 (mobilité de type « copier/coller »)	0 Q
Les transposons a ADA ou elements de type 2 (mobilité de type « couper/conet »)	
1 3 Historique de la recherche sur les álóments mobiles des gánomes	10
1.1.5. Instorique de la recherche sur les élements modres des génomes	
1.1.4. Proportions des elements mobiles dans les genomes	
1.1.5. Les elements de transpositions chez les champignons filamenteux	
1.2. LA MALADIE HOLLANDAISE DE L'ORME	
1.2.1. Généralités	24
1.2.2. Les agents causals de la maladie hollandaise de l'orme	
I.2.2.1. Ophiostoma ulmi	26
I.2.2.2. Ophiostoma novo-ulmi	27
I.2.2.3. Ophiostoma himal-ulmi	
1.2.3. Origine de la maladie hollandaise de l'orme	
I.2.4. « Homing » endonucléases ou introns mobiles : les éléments répétés actuellement prés	ents chez
O. ulmi, O. novo-ulmi sp. et O. himal-ulmi	
I.2.5. Cycle de la maladie (transmission, adhésion, colonisation et invasion)	
I.2.6. Variabilités intra et interspécifiques = impacts visibles d'éléments mobiles ?	
I.2.6.1. Polymorphisme chromosomique de la souche CESS16K	
I.2.6.2. Variabilité intraspécifique du pouvoir pathogène	35
I.2.6.3. Mating type switching de la souche VA	
I.3. OBJECTIFS ET HYPOTHÈSES DE RECHERCHE	
IL CADACTÉDICATION DES DDEMIEDS TRANSDOSONS À ADM CHEZ LES	ACENTS
II, CARACIERISATION DES PREMIERS TRANSPOSONS A ADN CHEL LES DESDONGADIES DE LA MALADIE HOLLANDAISE DE L'ODME	AGENIS
RESPONSABLES DE LA MALADIE HOLLANDAISE DE L'ORME	
II.1.AVANT PROPOS	
II.2. Résumé	
II.3. ABSTRACT	
II 4 INTRODUCTION	45
II.5. IVIATERIALS AND IVIE INOUS.	
11.5.1. Strains, mean and cumure conamons	
11.5.2. Genomic DIVA extractions, aegenerate primers and PCK conditions	
11.5.5. Recovery of OPHIO1 by genomic library ligation reaction (GLLR)-mediated PCR	
11.5.4. Recovery of OPHIO2 and additional copies of OPHIO1 by Inverse-PCR	
II.5.5. Recovery of a third transposon, OPHIO3, from a Suppression Subtractive Hybridizati	on (SSH)
library	51

II.5.6. Southern hybridizations	51
II.5.7. Sequence, phylogenetic and mapping analyses	52
II.6. RESULTS	53
II.6.1. Discovery and characterization of the first three DNA transposons in the fungal genus	
Ophiostoma	53
II.6.2. Distribution of DNA transposons in Ophiostoma ulmi and O. novo-ulmi	
II.6.3. Signature of silencing by repeat-induced point (RIP) mutations and analysis of variatio	ns in RIP-
based silencing	
II.7. DISCUSSION	
11./.1. OPHIO1, OPHIO2 and OPHIO3: three transposons from the Tc1/mariner superfamily	found in
the DED fungt	02 62
II.7.2. The potential role of interspecific hybrids in TE transmission	03 65
II.7.5. Evidence indi copies of OF HIOS have been suenced by KIF	03 70
II.0. DONNEES COMPLEMENTAIRES	·····.70 71 ASP 71
II.7. CORRELATIONS AVEC L'ETODE MENEE DANS CE CHAFTIRE SUR LE RIT CHEZ OF MOSTOM	A SF / I
III. IMPACTS DES TRANSPOSONS À ADN SUR LE GÉNOME HÔTE : ÉTUDE	DE LA
MOBILITÉ DES TRANSPOSONS <i>OPHIO1</i> ET <i>OPHIO2</i>	74
III 1 AVANT PROPOS	74
III.2. RESIME	
III.3. ABSTRACT	
III.4. INTRODUCTION	
III.5. MATERIALS AND METHODS	
III.5.1. Strains, media, culture conditions	
III.5.2. Nucleic acid extractions	
III.5.3. Primer extension and 5'RACE (Rapid Amplification of cDNA Ends) PCR	80
III.5.4. Plasmid construction and fungal transformation	
III.5.5. PCR, Inverse-PCR, RT-PCR and semi-quantitative RT-PCR	82
III.5.6. Stress factors and mutant screening	
III.5.7. Southern hybridizations and sequence analyses	85
III.6. RESULTS	85
III.6.1. Analyses of UTRs and TIRs of OPHIO transposons and amplification of OPHIO trans	cripts 85
III.6.2. RACE PCR and primer extension	89
III.6.3. Heat shock induces incomplete transposition of an exogenous OPHIO1 copy in O. nov	o-ulmi
subsp. americana W2 mutant strain	89
III.6.4. Cold shock and UV irradiation induce perfect excision of OPHIO2-int in O. novo-ulm.	i subsp.
novo-ulmi AST27	
III.6.5. Semiquantitative PCR	
III.7. DISCUSSION	
III./.1. An hsp promoter is present in OPHIO1 TIRs	
111.7.2. Expression of OPHIOI and OPHIO2 transposases	
III.7.5. Incomplete transposition of OPHIOT	90
111.7.4. Interspecific hybrids may snutte active TE between closely related species	
IV. IMPACTS DU GÉNOME HÔTE SUR LES TRANSPOSONS À ADN : MISE EN ÉVIDI	ENCE DE
RÉGIONS POSITIVEMENT SÉLECTIONNÉES AU SEIN DES TRANSPOSASES	
	102
IV 1. AVANI PROPOS	103
IV 2 A DETDACT	103 104
IV J. ABSTRACT	104
ΙΥ 5 ΜΑΤΕΡΙΔΙ S AND ΜΕΤΗΩDS	103
IV 5.1 Strains media and culture conditions	108 וווווווווווווווווווווווווווווווווווו
IV.5.2. Phylogenetic analyses	
IV.5.3. Detection of positive selection	
IV.5.4. Visualization of positive selection	109
The manufacture of positive second manufacture and the second manufacture of the second secon	

IV.6. RESULTS	110
IV.6.1. Visualization of positive selection using the Hermes DNA transposon	110
IV.6.2. Sequence variations and fast-evolving regions in fungal DNA transposons	115
IV.6.3. Prokaryotic insertion sequences (IS) have already undergone positive selection	124
IV.7. DISCUSSION	128
V. CONCLUSIONS ET SCHÉMA DE SYNTHÈSE	133
V.1. SYNTHESE # 1 = IMPACTS DES TRANSPOSONS OPHIO SUR LE GÉNOME D'OPHIOSTOMA SP	136
V.2. SYNTHESE # 2 = IMPACTS DU GÉNOME D'OPHIOSTOMA SP. SUR LES TRANSPOSONS OPHIO	136
V.2.1. Les mutations RIP chez les champignons de la M.H.O	137
V.2.2. Les zones de sélection positive au sein des transposases	139
V.3. SCHÉMA DE SYNTHÈSE	140
VI. PERSPECTIVES DE CES RECHERCHES – PROJETS	142
VI.1. PERSPECTIVES - PROJET #1 = LES TRANSPOSONS OPHIO, OUTILS D'ANALYSES	142
VI.2. PERSPECTIVES - PROJET #2 = LA RÉPARTITION DES TRANSPOSONS OPHIO AU SEIN DE	
POPULATIONS NATURELLES D'OPHIOSTOMA SPP	143
VI.3. PERSPECTIVES - PROJET #3 = LA MOBILITÉ DES TRANSPOSONS OPHIO	144
VI.4. Épilogue	147
BIBLIOGRAPHIE	148
ANNEXES	173

Généralités

Cette thèse est articulée autour de six chapitres qui aborderont successivement diverses questions inhérentes aux transposons à ADN (éléments mobiles de type 2). L'une d'entre elles est la simple vérification du paradigme voulant que les éléments mobiles de l'ADN soient présents au sein de tous les génomes des organismes vivants. Notre modèle d'étude a été deux ascomycètes filamenteux : Ophiostoma ulmi et O. novo-ulmi, principaux mycètes responsables de la maladie hollandaise de l'orme (M.H.O.) ou graphiose de l'orme.

L'<u>introduction</u> (*premier chapitre*) déclinera, dans un premier temps, les grandes définitions liées aux éléments de transposition et, dans un second temps, les principales caractéristiques de notre modèle d'étude pour finir sur les hypothèses de recherche de cette thèse.

Le second chapitre, décrivant la caractérisation de transposons à ADN (transposons OPHIO) au sein de génomes Ophiostomatoïdes, développera les méthodologies employées afin de mettre en évidence ces éléments mobiles des génomes ophiostomiens ainsi qu'une analyse précise, rigoureuse et comparée des divers transposons mis en évidence.

Le troisième chapitre mettra en évidence la <u>mobilité des transposons OPHIO</u> par le biais *de stress biotiques ou abiotiques.*

Le quatrième chapitre portera sur la mise en évidence de domaines de sélection positive au sein des transposases. Cette étude est née de l'analyse des transposons OPHIO et a été étoffée par la mise en évidence de zones d'évolution différentielle chez d'autres éléments mobiles présents chez les eucaryotes et les procaryotes.

Enfin, les cinquième et sixième chapitres donneront les <u>conclusions et perspectives</u> de cette thèse et feront le point sur les relations entre transposons à ADN et génomes hôtes par l'intermédiaire d'un schéma de synthèse. Le dernier chapitre tentera également de définir des perspectives plus générales à l'ensemble de ces travaux, principalement axées autour du développement de trois projets de recherche.

Figures et schémas

Figure 1. Histogramme représentant la proportion d'éléments répétés au sein des génomes séquencés
(adaptation de Kidwell et Lisch, 2002)
Figure 2. Une vision « métaphorique » du génome
Figure 3. Récapitulatif sur les éléments mobiles (adaptation de Kempken et Kück, 1998)
Figure 4. Schéma général de la rétrotransposition (adaptation de Esperenca Vieira et al., 1996)
Figure 5. Schéma général de la transposition (adaptation de Mahillon et Chandler, 1998)
Figure 6. Principales dates de l'histoire de la découverte des éléments mobiles des génomes et de la
hiologie moléculaire de 1900 à nos jours
Figure 7 Génomes et TF, mécanismes énigénétiques et naradoxe de la valeur C (adaptation de Fedoroff
1999 · Kidwell et Lisch 2002)
Figure 8 Synthèse des éléments mobiles découverts chez les championons (adaptation de Dahoussi et
Canv 2003)
Figure 9 Rénartition des rétratransposons et transposons chez les championons filamenteur
Figure 10. L'histoire d'une invasion (adaptation de Brasier, 1088 · Brasier, 1006 · Brasier et Kirk, 2000 ·
Pression et al. 2004)
Diasier et un, 2004)
Actorminána à mantin de management DADD (a dantation de Dine et al. 1005)
acterminees a partir de marqueurs KAPD (adaptation de Pipe et al. 1995)
Figure 12. Analyses comparatives ae longueurs moyennes aes peruneces (en µm) aes champignons ae la
M.H.O. (adaptation de Brasier et Kirk, 2001)
Figure 13. Cycle de la maladie hollandaise de l'orme (Adaptation d'Agrios, 2005)
Figure 14. Visualisation chromosomique et rupture chez CESS16K, souche d'Ophiostoma novo-ulmi
subsp. americana (adaptation de Dewar et Bernier, 1995)
Figure 15. Cartes génétiques des fragments introgressés de l'espèce O. ulmi vers O. novo-ulmi sp. novo-
ulmi (adaptation Et-Touil et al., 1999)
Figure 16. Commutation des gènes de reconnaissance sexuelle chez la levure par l'intermédiaire de
rupture chromosomique (inactivation du gène sexuel actif, par exemple MATa) et conversion
génique (remplacement et réactivation du nouvel allèle MAT a) (adaptation de Brown, 2006)
Figure 17. Analyses of OPHIO sequences
Figure 18. Phylogenetic relationships of DNA transposons
Figure 19. Distribution of OPHIO transposons in Ophiostoma spp. as determined by Southern
hybridization
Figure 20. Distribution of OPHIO1 and OPHIO2 transposons in Ophiostoma spp. as determined by PCR
Figure 21. Meiotic mapping of OPHIO2 DNA transposon of Ophiostoma novo-ulmi (OPHIO2-int)
Figure 22. Cumulative transition scores of RIPed forms of DNA transposons and retroelements found in
ascomycete fungi
Figure 23. Alignement des séquences d'OPHIO1 et OPHIO3 à partir d'amplifications in vivo sur O. ulmi,
O. novo-ulmi subsp. novo-ulmi et d'une séquence prédéterminées in silico (OPHIO3 de in silico).
Figure 24. Analyse par CTS de différentes copies d'OPHIO3 isolés chez diverses souches d'Ophiostoma
ulmi
Figure 25 Localisation géographique des souches d'Ophiostoma ulmi précédemment analysées par CTS7
Figure 26 Position of primers on sequences used in the study and indicated in Table 5
Figure 27 Flow diagram and nomenclature of stress treatments used to induce mobility of OPHIO
transposons
Figure 28 Alignment and analysis of terminal inverted repeats (TIRs) and untranslated regions (UTRs) of
Depine 20, magninem and analysis of terminal inverted repeats (TIKS) and antihistated regions (UTKS) of OPHIO1 and OPHIO2 transposable elements
CI HIOI and OF HIO2 transposable elements
Figure 29. KI-FOK JOF OF HIO Iransposase expression in O. uimi and O. novo-uimi
r igure 50. incomplete transposition of Or filo1 in O. novo-ulmi subsp. americana mulant strain
W2:UF FIUI
rigure 51. rerject excision of OrniO2-ini induced by UV snock and cold siress in O. novo-ulmi subsp.
novo-umi sirain AS12/

Figure 32. Semi-quantitative RT-PCR analysis of Ophiostoma novo-ulmi subsp. novo-ulmi AST27	
survivors using primers amplifying the actin gene and the OPHIO2-int transposase gene	95
Figure 33. Three-D visualization of the positively selected site analyses on Hermes in insects	.114
Figure 34. Comparison of Fot1-like TEs and OPHIO transposon sequences	. 116
Figure 35. Phylogenetic trees relating to actin genes and Tc1/mariner DNA transposase family proteins ascomycete fungi	s in .120
Figure 36. Split analyses of Tc1/mariner TEs divided in two distinct groups: the Fot1-like and Aft1-like groups	, .122
Figure 36 (suite). Split analyses of Tc1/mariner TEs divided in two distinct groups: the Fot1-like and A	ft1-
like groups	. 123
Figure 37. Positively selected site analysis on IS5 in proteobacteria	. 126
Figure 37 (suite). Positively selected site analysis on IS5 in proteobacteria	. 127
Figure 38. Comparaison des transposons OPHIO1 et OPHIO2 par visualisation CTS	. 135
Figure 39. Courbe CTS d'OPHIO1 par rapport à OPHIO2 (en rouge) et inversement (en vert)	. 138
Figure 40. Schéma de synthèse.	. 141
Figure 41. Principaux mécanismes de silencing chez les champignons	.174

Tableaux

Tableau 1. Récapitulatif des stresses environnementaux ayant pour conséquence la mobilité des TE	22
Table 2. Ophiostoma ulmi and O. novo-ulmi strains used	48
Table 3. Primers used for PCR cloning of DNA transposons in Ophiostoma sp	49
Table 4. RIP indices of various DNA transposons and retroelements in ascomycete fungi	61
Table 5. Primers used in this study	79
Table 6. Various mobile elements activated under stress condition in both prokaryotic and eukaryotic	
organisms	. 102
Table 7. LTR and Bayesian tests calculated from the transposable element datasets	. 117

Chapitre 1 : Introduction

I.1. Préambule et caractéristiques générales

I.1.1. Préambule

Les transposons à ADN étudiés dans cette thèse font appel à diverses notions qui vont être développées dans ce préambule. La première notion clé est celle de « répétition ». On considérera comme répétition toute séquence d'ADN présente en au moins deux copies similaires au sein d'un génome. La similarité entre deux séquences se caractérise par un pourcentage d'identité entre les copies. Nous considérerons comme copie d'une même répétition, toute séquence d'ADN présentant un minimum de 95 % d'identité entre les deux séquences comparées. Il faut toutefois différencier « copie identique » et « répétition fortuite » de l'ADN. En 1995 a été séquencé le premier génome, Haemophilus influenza, composé de 1,83 millions de bases. Il est statistiquement très peu probable que deux séquences de plus de 20 bases soient présentes en deux copies identiques au sein d'un génome. Or, l'analyse subséquente du génome d'H. influenza a démontré clairement la présence de nombreuses séquences en copies répétées de tailles de loin supérieures à 20 bases. Il est donc nécessaire de postuler qu'il existe, au sein des génomes, des mécanismes duplicatifs non aléatoires capables de générer des copies d'une même séquence d'ADN avec un pourcentage d'identité supérieur à 95 %. Depuis une vingtaine d'années, les génomes d'une multitude d'organismes, procaryotes ou eucaryotes (modèles génétiques pour la plupart), ont été séquencés et subséquemment annotés. Leur analyse a permis de vérifier que tous ces génomes possédaient des séquences répétées : Saccharomyces cerevisiae (12 Mb - 1996), Escherichia coli (5 Mb - 1997), Caenorhabditis elegans (100 Mb - 1998), Drosophila melanogaster (180 Mb - 1999), Arabidopsis thaliana (125 Mb - 2000), Populus trichocarpa (480 Mb - 2004), une grande majorité des séquences consensuelles d'Homo sapiens (3,29 Gb - 2004) ainsi que les séquences consensuelles de Pan troglodytes (3,02 Gb - 2005). Les analyses préliminaires des séquences de ces génomes ont permis d'extraire une quantité considérable d'informations : elles constitueraient notamment entre 5 et 95 % des génomes. La Figure 1 illustre, pour partie, cette observation.



(Kidwell et Lisch, 2002)

Chaque barre verticale représente un génome séquencé dans sa totalité (en millions de paires de bases). Par ordre de taille du plus grand au plus petit : orge, maïs, humain, riz, drosophile, vesce, ver et levure. La zone grisée représente la proportion d'éléments répétés au sein de chacun des génomes. Dans génomes, les éléments chacun des transposables (TE *Transposable* pour *Elements*) représentent une partie significative de ces derniers, allant de 3 % pour le génome le plus petit à plus de 70 %. Le « record » est atteint par le genre Lilium henryi, qui totaliserait plus de 95 % de son génome en éléments répétés, avec plus de 13 000 copies du transposon del (Sentry et al., 1985).

Figure 1. Histogramme représentant la proportion d'éléments répétés au sein des génomes séquencés



Imaginons le génome d'un organisme tel un iceberg, la partie émergée représentant la partie exprimée de ce dernier, à savoir les zones codantes transcrites : les ARN messagers totaux (± 20 % des génomes). Le sommet de l'iceberg serait plus précisément la fraction des ARN apte à subir les mécanismes post-transcriptionels et à être traduite en protéines (environ 5%). Toutes proportions gardées. immergée de l'iceberg la partie correspondrait à l'ensemble des ADN non codants communément appelé « Junk DNA » ou « ADN poubelle » et correspondant approximativement à 80 % des génomes (Iceberg of Newfoundland, St Johns, www.defiant.corban.edu/gtipton/netfun/iceberg.html).

Figure 2. Une vision « métaphorique » du génome

D'une manière générale, les génomes sont constitués de deux sortes d'ADN: l'euchromatine (ou la forme hypocondensée de la chromatine, transcriptionnellemen active) et l'hétérochromatine (correspondant à la chromatine à l'état hypercondensé). La différence entre ces deux types de chromatine réside essentiellement au niveau de leur activité transcriptionnelle : l'hétérochromatine, souvent hypoacétylée ou hyperméthylée, n'est transcriptionnellement active, contrairement à l'euchromatine. pas L'hétérochromatine peut également contenir de l'ADN poubelle ou « junk DNA », mais quelle est la composition de l'ADN poubelle (Fig. 2) ? Une définition simple nous indique qu'il est « composé de l'ensemble des séquences d'ADN des chromosomes n'ayant pas de fonctions ou de structures actuellement identifiées ». À ce jour, on inclut dans cette catégorie l'ensemble des séquences d'ADN qui ne possèdent pas de structures codantes apparentes, les pseudogènes, ainsi que l'ensemble des séquences d'ADN répétées.

I.1.2. Caractéristiques générales des différents types de séquences répétées

I.1.2.1. Les répétitions satellites

Ce sont des séquences d'ADN formées d'une répétition de motifs composés de *n* nucléotides. Ces structures ont été mises en évidence dans l'hétérochromatine, principalement près des centromères. Elles sont essentiellement divisées en trois grandes classes, basées sur la longueur (*n*) de l'unité de répétition : les <u>microsatellites</u> (ou SSR pour *Simple Sequence Repeats*), les <u>mini-satellites</u> (ou VNTR pour *Variable Number of Tandem Repeats*) et les <u>satellites</u>, également appelés répétitions centromériques. Leurs rôles et impacts sur le génome sont très différents et chacune des catégories semble être caractérisée par un mécanisme duplicatif particulier. Ces séquences ont été appelées *satellites* car, lors de la séparation de l'ADN total par gradients de densité, elles sédimentent en bandes distinctes qui se séparent de la majeure partie de l'ADN génomique en raison de leur composition particulière en bases nucléiques.

I.1.2.2. Les répétitions géantes

Les génomes des cellules eucaryotes possèdent un certain nombre de séquences répétées de très grande taille dites « géantes », pouvant couvrir plusieurs centaines de kilobases (Kb), telles que des duplications de segments d'ADN, de chromosomes (hyperploïdisation) ou encore de génomes entiers (polyploïdisation). La première répétition géante (isolée chez *Saccharomyces cerevisiae*) a été localisée au niveau des centromères des chromosomes III et XIV (Friedman et Hughes, 2001). Elle est décrite comme une région synténique entre ces deux chromosomes. La synténie n'est cependant pas parfaite puisque s'insèrent, dans ces régions, des gènes uniques.

I.1.2.3. Les pseudogènes

Les pseudogènes sont des séquences d'ADN de taille variable ayant un certain pourcentage d'identité avec des gènes connus. Toutefois, ces derniers ont perdu la capacité de coder leur(s) protéine(s) (insertion de codons stop, décalage du cadre de lecture, absence de promoteurs...) ou ne sont pas exprimés dans la cellule (non fonctionnalité). Ils peuvent provenir de divers mécanismes génétiques, divisés en trois catégories : (1) la rétrotransposition, (2) la duplication ou (3) le dysfonctionnement de gènes initialement codants en copies multiples. Ces dysfonctionnements peuvent être dus à des insertions, des mutations non synonymes, au décalage du cadre de lecture, etc. Ils peuvent toutefois avoir un rôle dans les réarrangements génomiques (via la recombinaison homologue), comme régulateur de séquences codantes (Hirotsune et al., 2003) ou dans le phénomène d'ARN interférence (Mitrovich et Anderson, 2005).

I.1.2.4. Les éléments de transposition

Ces éléments sont des séquences d'ADN de taille variable, allant de 200 à 7 000 nucléotides. Ils sont actuellement divisés en deux grandes familles : les <u>rétrotransposons</u> et les <u>transposons à ADN</u>. C'est la compréhension de leur mécanisme de transposition qui a permis cette division (Fig. 3).

retrotransposon of Ty1-Copia family R1 pol retrotransposon of Gypsy family Int RNa: 989 PBS pol retroposon / LINE-like RNaseH 81 939 (TAA) SINE-like (TAA) " Tc1/Mariner and Fot1/Pogo families -AT AT TSD TIR hAT family transposase TSD TIR

mini transposon Guest

(Kempken et Kück, 1998)

Figure 3. Récapitulatif sur les éléments mobiles (adaptation de Kempken et Kück, 1998)

Les diverses familles d'éléments mobiles sont représentées dans cette figure et divisées en deux grandes catégories : les <u>rétrotransposons</u> (éléments mobiles de type I), constitués des principales familles Ty1-Copia, gypsy, LINE-like et SINE-like, et les <u>transposons à ADN</u> (éléments mobiles de type II) constitués des familles Tc1/mariner, Fot1/pogo, hAT et MITE ou mini-transposons (Feschotte et al., 2002). La nomenclature est actuellement la même utilisée pour ces deux familles d'éléments mobiles. Toutefois, une troisième famille a fait son apparition avec la découverte de nouveaux éléments tels que les cryptons ou encore polintrons (C.f. page 13).

Abbréviations : TSD = target site duplication ou séquences cible du TE dupliquée après l'insertion de la copie del'élément ; LTR = long terminal repeat ou séquences de 200 à 400-nt chez les éléments de type 1 possédant les promoteurs permettant l'expression des gènes requis pour la mobilité des retroéléments. Ces séquences servent également de zone d'intégration d'une nouvelle copie ; gag, Pr, int, RT and RNAseH = différents domaines du rétroélément : gag : domaine structurant, Pr : protéase, int : intégrase, RT : reverse transcriptase et RNAseH : ARN synthase ; (TAA)n = signifiant la présence de μ -satellites terminaux TAA et TIR = terminal inverted repeat, région palimdromique encadrant la transposase et nécessaire pour la fixation de cette dernière et la mobilité de ces éléments

Toutefois, les recherches actuelles ont démontré qu'il existait d'autres types d'éléments de transposition (éléments *rolling-circle*, *Foldback*, *TU*, *crypton*, *Polintron*, *Helitron etc...*) qui n'entrent pas dans les deux premières catégories. Ces derniers sont maintenant classés dans une nouvelle classe à part entière que nous appelerons ici Type 3.

Les caractéristiques majeures de ces éléments sont les suivantes : la *mobilité*, l'*ubiquité* et la capacité de *duplication*. Ces trois notions seront reprises au **chapitre IV**. L'étude de ces éléments a permis de les relier à des maladies génétiques (par exemple l'insertion d'une copie du rétrotransposon *LINE1* dans le gène codant pour le facteur *VIII* induisant l'hémophilie A (Singer, 1994). Il a été également démontré que l'insertion d'éléments ALU génère plus de 15 maladies génétiques chez l'homme (Deininger et Batzer, 1999).

• Les rétrotransposons ou éléments de type 1 (mobilité de type « copier/coller »).

Les rétrotransposons sont les plus abondants des éléments mobiles au sein des eucaryotes. Ce sont des séquences d'ADN allant de 2,4 Kb pour *marsu* chez *Fusarium oxysporum* à 8 Kb pour *grasshopper* chez *Magnaporthe oryzae*. Ils se distinguent par leur mécanisme de transposition qui met en jeu un ARN intermédiaire qui sera, par la suite, rétrotranscrit en ADN complémentaire, par l'enzyme *RNA-dependent DNA polymerase* (découvert simultanément par Baltimore et par Temin et Mizutani en 1970). Diverses études évolutives intègrent les rétrovirus en tant que sous-classe de ces éléments. Les rétrotransposons sont divisés en deux sous-groupes, fonction ou non de la présence de longues séquences répétées (les *LTR* ou *Long Terminal Repeat*) encadrant les diverses zones codantes (domaines *gag* et *pol*) : (1) Les rétrotransposons à *LTR* synthétisent un ADN complémentaire (ADNc) à partir de leur ARNm (via la transcriptase inverse comprise dans le domaine *pol*) dans des particules pseudovirales (codée par le domaine *gag*), puis intègrent cet ADNc à un nouveau locus chromosomique. (2) Les rétrotransposons sans *LTR* synthétisent leur ADNc directement au site cible d'intégration, provoquant parfois la coupure du double brin d'ADN nécessaire à cette intégration.



Figure 4. Schéma général de la rétrotransposition (adaptation de Esperença Vieira et al., 1996)

Imaginons une copie d'un rétroélément (par exemple le rétrotransposon L1) insérée dans le génome. Cette luite. Le brin d'ARN messager (ARNm) est initié à l'extrémité 5' du copie, si Nouveau rétrotransposon LTR 5' (r ntinue vers le domaine interne, et se termine au niveau 3' du LTR 3', où il est polyadénylé. L'ARNm ainsi obtenu sera utilisé, d'une part, pour la synthèse des enzymes nécessaires à la transposition (réverse transcription, RNaseH, intégrase), d'autre part pour la synthèse d'une nouvelle copie de l'élément (Copier). Ces produits permettront également de former une particule pseudo virale (notamment grâce au domaine gag pour group associated antigen) qui protègera la séquence d'ARN de possible dégradations. Cet ARN mobile sera alors réverse-transcrit en ADN complémentaire et pourra être circularisé par la ligation de ses extrémités. Selon Brown et col. (Brown et al., 1989), la forme circulaire du rétrotransposon semble être plus appropriée pour l'étape d'intégration. Cette intégration est associée à la duplication de 4 à 6 paires de bases de la séquence d'insertion du génome hôte. La taille de la duplication du site cible semble être spécifique pour chaque famille de rétrotransposons. La copie circularisée formera alors une nouvelle copie autonome localisée aléatoirement dans le génome (Coller). Les grandes familles de rétrotransposons sont les suivantes : éléments avec LTR = Ty3/gypsy, Ty1/copia ; éléments sans LTR = LINEet LINE-like (également appelés rétroposons) et SINE et SINE-like (Eickbush, 1994).

Chez les mammifères, ces éléments sont nommés LINE (*Long Interspersed Nuclear Elements*) ou SINE (*Short Interspersed Nuclear Elements*), en fonction de leur taille et de leur structure (Wichman *et al.*, 1992). Le nombre de copies des éléments *L1* (principale famille des LINE) chez les mammifères peut aller jusqu'à 100 000 (Evans et Palmiter, 1991). Il est important de remarquer qu'aucun rétrotransposon n'a été découvert à ce jour chez les procaryotes, ce qui situe certainement leur apparition après l'émergence des eucaryotes. Le mécanisme de rétrotransposition est décrit à la Fig. 4.

• Les transposons à ADN ou éléments de type 2 (mobilité de type « couper/coller »)

Les transposons à ADN représentent un très large ensemble d'éléments mobiles se retrouvant chez la totalité des génomes eucaryotes. Leur taille varie de 1,5 à 3 Kb. Chez les eucaryotes, les transposons à ADN se regroupent en quelques familles dont la plus grande est *Tc1/mariner*, présente dans tous les phylums (Plasterk et al., 1999). L'élément *Tc1*, découvert en 1983, chez C. elegans (Emmons et al., 1983), fut ensuite rapproché de l'élément mariner de Drosophila mauritiana (Jacobson et al., 1986), puis d'autres familles comme les éléments pogo de D. melanogaster (Tudor et al., 1992). Une seconde famille bien connue des transposons à ADN eucaryotes est celle qui englobe l'élément P (Bingham et al., 1982). Cet élément a envahi les génomes de D. melanogaster il y a quelques décennies (Engels, 1996). Il a été, par la suite, décrit chez de nombreuses espèces de drosophiles, et plus récemment chez d'autres espèces, y compris H. sapiens (Hagemann et Pinsker, 2001). Les transposons diffèrent des éléments d'insertion bactériens (IS) par le fait même qu'ils sont des éléments eucaryotes. Ils diffèrent également des rétrotransposons par l'absence de la rétrotranscription dans leur mécanisme de transposition. Ils possèdent des structures particulières appelées séquences répétées inversées (TIR pour Terminal Inverted Repeats allant de 10 à 500nt) situées à leurs extrémités et des sites dupliqués encadrant les TIR (TSD pour Target Site Duplications allant de 2 à 20nt). Les transposons à ADN bougent via la transcription et traduction de leur séquence codant pour une enzyme responsable de cette mobilité : la transposase. Cette dernière est composée de trois domaines fondamentaux : les domaines HTH (pour Helix-turn-Helix) et CENPB

(*CENtromeric Protein B*) qui sont deux domaines de fixation à l'ADN et le domaine catalytique DDE servant à cliver l'ADN. Les différentes étapes de la mobilité des transposons à ADN seront décrites plus précisément dans le chapitre 3 de cette thèse, consacré à la mobilité (Fig. 5).



Figure 5. Schéma général de la transposition : mécanisme de type Couper/Coller (adaptation de Mahillon et Chandler, 1998)

Imaginons une copie d'un transposon à ADN (par exemple le transposon *pogo*) insérée dans le génome. Cette copie, si elle est exprimée, est alors transcrite et traduite. Un ou plusieurs produits d'expression (transposases) sont alors exprimés. Les transposases proviendraient initialement de l'évolution d'endonucléases ancestrales ayant développée la spécificité de cibler des séquences d'ADN situées dans les séquences répétées inversées (TIR pour *Terminal Inverted Repeats* ou séquences palindromiques) encadrant la zone codant la transposase. La reconnaissance de ces séquences semblerait s'effectuer par l'intermédiaire d'un domaine *Helix-Turn-Helix (HTH domain* en vert dans le schéma) qui se fixerait sur une ou deux zones précises des TIR (*box1* et *box2*, Feschotte et al., 2005). Une fois la transposase fixée aux deux extrémités de l'élément, le domaine catalytique (*DDE domain*, en bleu) clive l'ADN à l'extrémité des séquences répétées inversées. Un complexe « transposase-transposon » se forme alors (prenant parfois le nom de *transpososome*). Ce complexe reconnaîtrait alors un nouveau site d'intégration (par exemple, le dinucléotides TA) qui lui permettrait de s'insérer dans l'ADN (Plasterk et Luenen, 2003).

Le mécanisme général (décrit à la Fig. 5) de la transposition peut toutefois être divisé en deux mécanismes distincts.

<u>Mécanisme non réplicatif</u> : ce dernier est initié par l'hydrolyse simple brin des extrémités de l'élément. L'extrémité libre 3'OH se lie à un site d'intégration aléatoire ou spécifique (selon les familles). Mizuuchi propose que le transfert de brin se fasse en une seule étape de transesterification (Mizuuchi, 1992). La transposition laisse une cassure double brin au site d'excision qui est réparée par les mécanismes de l'hôte.

<u>Mécanisme réplicatif</u> : il existe dans les génomes eucaryotes des mécanismes pouvant expliquer la duplication de ces éléments. La coupure double-brins peut être réparée par copie de la chromatide soeur (85% des cas pour l'élément *P* [Engels, 1996]) ou du chromosome homologue ce qui conduit à la duplication du transposon. Cependant, si la réparation par copie est avortée, seules les extrémités du transposon à ADN sont copiées et raboutées. Les MITE (*Miniature Inverted repeat Transposable Elements*), très répandus dans les génomes eucaryotes (Surzycki et Belknap, 2000), présentent ce type de structure. Il a d'ailleurs été montré chez *A. thaliana* que certains MITE sont des transposons à ADN tronqués mobilisables en *trans* par des transposases d'éléments complets (Feschotte et Mouches, 2000) (adapté de la thèse d'Achaz, 2002).

0

Les éléments de type 3

La troisième catégorie d'éléments mobile englobe l'ensemble des éléments ne pouvant être inclus dans les deux précédentes catégories. Ce paragraphe rapportera 2 types d'éléments mobiles de cette classe. Premièrement, les <u>cryptons</u>. Ce sont des éléments mobiles qui possèdent une séquence codante unique, coupée par plusieurs introns. Ces éléments ne possèdent pas de séquences palindromiques, ce qui les différentient des éléments de type 2 mais uniquement des répétitions directes encadrant la région codante (Goodwin et al., 2003). Deuxièmement, les <u>polintons</u> qui peuplent les génomes des protistes, mycètes, et une grande partie du règne animal. Les Polintons sont caractérisés par un ensemble unique de protéines nécessaires pour leur transposition, comprenant une AND polymérase, une intégrase rétroviral, protéase et ATPase. Les polintons sont également caractérisés par la

10

présence de séquences cibles dupliquées de 6-bp, de régions répétées inverses de plusieurs centaines de nucléotides de longueur et possèdent les dinucléotides 5'-AG et TC-3' à leur extrémité et peuvent être autonomes ou non. Les publications actuelles suggèrent que ces éléments pourraient avoir évolué d'un plasmide linéaire après acquisition d'une intégrase rétrovirale (Kapitonov et Jurka, 2006).

Pour terminer cette partie sur les éléments de transposition, il est important de mentionner que ces derniers sont des éléments *cryptiques* des génomes : (1) leur transposition est génétiquement invisible car ils bougent à des fréquences extrêmement faibles et (2) ils induisent généralement des recombinaisons illégitimes à des fréquences plus faibles que celles induites par les séquences homologues.

I.1.3. Historique de la recherche sur les éléments mobiles des génomes

Les études de mutations instables (ou réversibles) remontent aux découvertes de **DeVries** (école mutationniste), au début des années 1900 et portaient essentiellement sur le maïs (*Zea mays* L.), le modèle principal de l'époque (Fig. 6). Ces travaux lui ont permis de décrire une série de mutations ne respectant pas les lois mendéliennes et ont abouti au développement de plusieurs variétés anormales dont les variétés *ever-sporting* (Shull, 1911). Il faut attendre **Emerson** qui, de 1914 à 1925, donna les premières interprétations biologiques de ces transmissions non mendéliennes en étudiant le locus P du maïs. Sa première publication sur le sujet introduisit la définition de *variégation* ou phénotype mosaïque : « ...ce phénotype se distingue des autres patrons de couleurs par son incorrigible irrégularité ». Emerson fut capable d'intégrer ces mutations instables au sein du paradigme mendélien en postulant que la variégation se met en place lors de l'association temporaire d'un inhibiteur et d'un locus particulier requis pour la pigmentation.

Figure 6. Principales dates de l'histoire de la découverte des éléments mobiles des génomes et de la biologie moléculaire de 1900 à nos jours



Barbara McClintock (1902-1992) : ses travaux, des années 40 et 50 sur la cytogénétique du maïs, menèrent B. McClintock à développer une théorie sur la « transposabilité » des gènes, c'est-à-dire leur capacité à se déplacer de façon autonome au sein et entre des chromosomes. McClintock a construit cette inférence en observant les variations de coloration des grains de maïs. Initialement, l'idée de mobilité des gènes n'a pas fait l'unanimité mais sa persévérance et l'affinement des techniques moléculaires à la fin des années 70 et au début des années 80 lui ont permis de confirmer sa découverte. L'ensemble de ses recherches, tant sur les transposons que l'architecture des génomes lui ont valu l'attribution, en 1983, du prix Nobel de physiologie et de médecine. McClintock a été la première femme à remporter un prix Nobel aux États-Unis (Shapiro, 2001).

Il explique également la réversion des caractères par la perte ou la non-association de l'inhibiteur. Malgré l'avancée de ces découvertes, d'éminents scientifiques, tels que Correns et Goldschmidt, observèrent ces mutations qu'ils qualifièrent de « gènes malades », estimant toutefois que l'étude de ces gènes n'aurait pas de grands impacts sur le fonctionnement des gènes conventionnels. A l'inverse, d'autres scientifiques tel que Rhoades supportaient les idées d'Emerson en s'appuyant sur le fait qu'il n'y a aucune différence entre des mutations stables/instables, et que l'instabilité est conditionnelle (i.e. dépendante d'une partie annexe). Rhoades démontra, en analysant le locus récessif al chez le maïs (codant pour la voie de biosynthèse des anthocyanines), que ce dernier se révélait instable en fonction du matériel génétique utilisé. La clé de cette déstabilisation fut démontrée par la présence du locus Dt, qui induisait une réversion de l'allèle al vers le phénotype sauvage. Toujours vers la fin des années 30 et le début des années 40, McClintock commença ses études sur les ruptures chromosomiques. Elle développa une méthode de mutagenèse qu'elle appliqua sur une lignée. Chacune des mutations induisait la cassure du segment terminal et du télomère du chromosome 9 du maïs, générant une modification de la pigmentation. En cherchant des mutants dans diverses descendances, elle observa de fortes fréquences de mutants « mosaïques » (pigmentation variable) et ce, quel que soit le croisement. Elle remarque que, malgré le peu de données concernant des allèles mutants, elle pouvait isoler plus de 14 cas d'instabilité génétique au sein de ces lignées. Elle choisit alors de suivre le devenir d'un locus particulier, appelé Ds pour Dissociation, ayant pour impact d'induire la cassure péri-centromérique du petit bras du chromosome 9. Elle remarqua, par la suite, que cette cassure nécessitait la présence d'un autre locus, non lié, qu'elle désigna sous le nom de locus Activator (Ac). En 1948, McClintock avait assez d'assurance pour démonter que le locus Ds bouge : « Il est maintenant connu que le locus Ds pourrait changer de position au sein d'un chromosome ». La relation entre la cassure chromosomique et le phénotype mosaïque émerge lorsqu'elle analyse la descendance d'un nouveau phénotype mosaïque, provenant d'une mutation du locus C intervenant dans la pigmentation kernel des grains (coloration jaune ou orange).



Elle regroupa de nombreuses études cytologiques et génétiques sur cette mutation, nommée c-ml, et démontra qu'elle était conditionnée et dépendante du locus Ac précédemment décrit. Elle mit en évidence que l'origine de la mutation instable coïncidait avec la transposition du locus Ds de sa position initiale, située près du centromère, vers un nouveau site au sein du locus C, et que la réversion serait due au mécanisme inverse. Après cette démonstration, elle inféra que le phénotype mosaïque somatique reflétait les fréquentes transpositions du locus Ds durant le développement. Ces transpositions permettaient aussi d'expliquer les observations de Demerc et Rhoades (Fedoroff, 1999 ; Fedoroff, 2002). Il est toutefois intéressant de remarquer que ces idées novatrices ne furent pas admises par l'ensemble de la communauté scientifique. En effet, la première tentative de publication par McClintock, de sa théorie dans le journal Proceedings of the National Academy of Sciences of the U.S.A. en 1950 fut rejetée, avec comme explication, que « ... la présente thèse et les données qu'elle contient ne peuvent être acceptées par la majorité des généticiens et autres biologistes... ». Ce n'est que quelques d'années plus tard que McClintock publia sa théorie sur les éléments mobiles du génome (McClintock, 1956) dans les même années que la théorie de Jacob & Monod sur les opérons bactériens (Jacob et Monod, 1959). À partir des années 60, de grandes avancées dans la compréhension des mécanismes des éléments mobiles allaient voir le jour : (1) ils sont assimilables et maniables comme des entités génétiques propres, malgré leur capacité de mouvement, (2) les TE sont des éléments régulateurs des génomes mais sont silencieux. En 1969, Shapiro et Adhya publient la découverte du premier élément mobile ou IS (insertion sequence) chez la bactérie E. coli. Dans les années 1970, l'affinement des techniques de biologie moléculaire va bon train et permet à Fedoroff, en 1983, d'isoler par clonage, les premiers transposons : les éléments Ac et Ds décrits précédemment par McClintock chez le maïs (Fedoroff et al., 1983). La multiplicité des études sur de nombreux organismes change de beaucoup la vision de ces éléments, maintenant isolables. On met alors en lumière que les TE sont des séquences en multi-copies dont une partie ne sont toutefois que des formes tronquées. Diverses études publiées au cours des années 1984 à 1989 montrent qu'il existe des variations au sein des patrons de méthylation entre les TE actifs et TE non actifs : chez l'élément Ac en 1984 ou l'élément Spm. Enfin, à partir du milieu des années 90, arrivent les séquences des génomes entiers, et avec ces derniers, l'évidence que les TE y sont présents en proportion non négligeable.

Les spéculations vont alors bon train quant à leurs rôles ou leurs origines. Et s'ils sont, à cette époque, vus comme des éléments parasites et déstabilisateurs des génomes (Orgel et Crick, 1980) alors un problème persiste : si les transposons remanient tant les génomes, comment se fait-il que l'ordre des gènes (synténie) demeure très conservé ? (Mise en évidence par les études sur les relations entre les génomes de plantes cultivées telles que les *Solanaceae* et les *Graminaceae* [Gale et Devos, 1998]). Des études ultérieures sur les génomes du maïs (2 500 Mb) et du sorgho (750 Mb) montrent que leur divergence (en terme de taille) viendrait essentiellement de l'accumulation de TE dans des zones particulières. **SanMiguel** et collaborateurs publièrent entre 1996 et 1998 la découverte de régions (240 Kb au total) présentes chez le maïs et non chez le sorgho, essentiellement composées (74 %) de rétrotransposons (Bennetzen et al., 1998; SanMiguel et al., 1996).

Actuellement, il est difficile de résumer en quelques lignes, l'ensemble des impacts des TE sur les génomes. Ils sont, par exemple, les vecteurs de *crossing over* inégaux ou illégitimes. De nombreux exemples et détails affluent quant aux rôles potentiels de la transposition, de la rétrotransposition et des réarrangements dirigés par transposons (recombinaisons ectopiques). Par exemple, chez le chromosome Y de diverses espèces de drosophiles, des rétroéléments pourraient induire soit la stérilité (transposons gypsy chez D. hydei [Hochstenbach et al., 1996]), soit la dégénérescence du chromosome Y (transposon TRAM chez D. miranda [Steinemann et Steinemann, 1997]). Les TE pourraient également catalyser le contrôle et la production d'anthocyanine et d'aleurone, la régulation de ces protéines étant effectuée par des transposons de type CACTA, dont font partie les transposons En/Spm (Kunze et Weil 2002). Ironiquement, la fin des années 90 voit l'émergence d'une nouvelle idée : la caractéristique majeure des transposons pourrait ne pas être leur instabilité au sein du génome, mais plutôt leur stabilité. Le fait même de dénombrer de nombreux TE au sein des chromosomes peut paraître intriguant si l'on considère que ces éléments perturbent les génomes. On peut alors se poser la question suivante : par quel moyen tout ceci reste en équilibre, se prévient des recombinaisons, transposition, réarrangements ou délétion ?

1995 : premier génome séquencé
Haemophilus influenza (1,8 Mb)
1996 : Saccharomyces cerevisea (12 Mb)
1997 : Echerichia coli (5 Mb)
1998 : Ceanorhabditis elegans (100 Mb)

1999 : Drosophila melanogaster (180 Mb)2000 : Arabidopsis thaliana (125 Mb)2004 : Populus trichocarpa (480 Mb)2004 : séquences consensuelles deHomo sapiens (3,29 Gb). Séquences partielles dePan troglodytes (3,02 Gb).2006 : transposonsOPHIO ☺

2001 : <u>Linder-Basso D.</u> *Crypt*1 : isolation du premier transposon à ADN chez *Cryphonectria parasitica*.

1987 : <u>Selker E.</u> Mécanismes de « silencing » par R.I.P. (*Repeat-Induced Point mutations*) 1989 : <u>Kindsey JA.</u> *Tad*, isolation du premier rétrotransposon chez *Neurospora crassa*.

1997 : <u>Levis C.</u> *Flipper* : isolation du premier transposon à ADN chez *Botrytis cinerea*.

1984-1989 : Variation des patrons de méthylation entre TE actifs et non actifs.

Figure 6 (suite)

1994 : <u>Kachroo P.</u> *Pot2* : isolation du premier transposon à ADN chez *Magnaporthe grisea*.

1992 : <u>Daboussi MJ.</u> *Fot1* : isolation du premier transposon à ADN chez *Fusarium oxysporum*. 17

La fin des année 90 a vu la publication d'études qui ont amélioré la compréhension des relations entre TE et génome : **Kidwell** publia une étude démontrant la corrélation entre la taille des génomes et la taille des éléments de transposition présents dans ces génomes alors que Fedoroff émettait les trois hypothèses suivantes :



(1) une des caractéristiques particulières des génomes complexes est l'existence de mécanismes épigénétiques qui permettent la redondance, à des niveaux plus ou moins importants, d'éléments dupliqués en tandem ou dispersés. (2) Les éléments de transpositions servent à « moduler » le génome, maintenant dans un état de dynamique structurale qui augmente sa taille et sa complexité. Enfin (3), les forces de modélisation de la recombinaison et de la transposition sont en équilibre, donnant une dynamique au système génomique, en augmentant complexité sa approchant parfois le chaos (Fedoroff, 1999).

<u>Le paradoxe de la valeur C</u> met en évidence l'absence de corrélation entre la taille du génome d'un organisme et sa complexité. En effet, on a tout d'abord pensé que plus un organisme était complexe plus la taille de son génome devait être importante. Cependant, les travaux de cartographie et de séquençage, ont permis de démontrer qu'il en était autrement : il est en effet clair que les génomes des organismes de l'embranchement des conifères ont une taille plus importante que ceux des primates, réduisant ainsi la vision anthropomorphique. Toutefois, bien que possédant moins de portions codantes, le génome de l'homme générerait plus de protéines que celui de bien d'autres organismes.

Figure 7. Génomes et TE, mécanismes épigénétiques et paradoxe de la valeur C (adaptation de Fedoroff, 1999 ; Kidwell et Lisch, 2002)

et de recombinaison de l'ADN. De courtes homologies suffiraient à provoquer la duplication par « patinage » durant la réplication, des *crossing-over* inégaux et l'excision/réparation de l'ADN double brin (Gorbunova et Levy, 1997 ; Liang et al., 1998). Pour d'autres comme Fedoroff, la coexistence de ces éléments dans le génome est le résultat d'une <u>évolution stable</u> (Fig. 7). De plus en plus nombreux sont maintenant les scientifiques qui pensent qu'il est plus légitime de voir l'interaction TE/génome hôte comme une évolution stable entre deux organismes. Deux idées vont dans ce sens : (1) le fait qu'il n'existe pas de génome actuellement séquencé qui soit saturé ou absous de TE et que, (2) de plus en plus de cas montrent que ces derniers font partie intégrante des génomes. De ces dernières sont nés les concepts de *domestication* (élément *P*, *RAG1* et *RAG2* ; [Nouaud et Anxolabehere, 1997 ; Fugmann et al., 2006]) et de présence obligatoire pour la bonne expression de gènes.

I.1.4. Proportions des éléments mobiles dans les génomes

La taille des génomes montre un facteur de 80 000 entre le génome eucaryote le plus grand (parmi les plus grands, *Triticum* sp. 16 000 Mb ou *Pinus* sp., 24 000 Mb) et le plus petit génome procaryote (parmi les plus petits, *H. influenzae*, 1,5 Mb). Bien que la variabilité génomique soit très faible entre le génome des oiseaux, des mammifères et des poissons téléostéens, elle est très forte au sein des invertébrés et des plantes (Kidwell et Lisch, 2002). D'autre part, les éléments de transposition, représentant une proportion très variable des génomes, montrent peu de variation chez les organismes ayant une taille de génome restreinte (les levures) et inversement, une grande variation chez les organismes à forte taille génomique telles que les plantes. Margaret Kidwell a publié, en 2002, une étude intéressante traitant de la relation entre la taille des génomes et celle des éléments de transposition et a remarqué qu'il n'est toutefois pas si aisé d'établir un lien de cause à effet. Du fait de leur proportion dans certains génomes, certains auteurs se sont aventurés en spéculation concernant de possibles relations entre les TE et la valeur C (Fig. 7).

Deux tendances s'opposent : bien que les éléments mobiles expliquent en partie le paradoxe de la valeur C chez les herbacées, il semble toutefois que d'autres structures entrent en compte pour expliquer les variations de taille des génomes (Wendel et Wessler, 2000). Pour

d'autres, la corrélation ne fait aucun doute : des études ont démontré la présence de 1 000 copies de l'élément *mimo* chez *Culex pipiens*. C'est à la fois plus que l'ensemble des MITE mis en évidence chez *Anophele gambiae*, mais très en dessous du nombre trouvé chez *Aedes aegypti* (Bensaadi-Merchermek et al., 1994). Il est possible que le séquençage du génome de certains conifères nous aide à éclaircir ce point, car on s'attend à trouver une très forte proportion de TE dans ces génomes.

I.1.5. Les éléments de transpositions chez les champignons filamenteux

Les recherches sur les éléments mobiles chez les champignons ont vu le jour suite aux travaux effectués sur des champignons pathogènes ayant des répercussions au niveau agricole, médical ou biotechnologique (Daboussi, 1997). On trouve des éléments mobiles (de types 1 et 2 – Fig. 8) au sein des trois principaux ordres de champignons : ascomycètes, basidiomycètes et zygomycètes (Fig. 9). Bien que les premiers travaux aient démontré la présence d'éléments inactifs dans les génomes de *Neurospora crassa* et d'*Aspergillus nidulans* (Galagan et Selker, 2004), les analyses étendues au génome de nombreux champignons ont mis en évidence de nombreux éléments actifs, plusieurs de ces derniers étant impliqués dans les modifications génomiques suivantes :

1- **les variations génomiques engendrées par les TE** : de par leur mobilité, ces derniers génèrent des insertions ou excisions au sein de régions codantes ou potentiellement codantes, créant ainsi des modifications dans l'expression des gènes : l'arrêt de l'expression de gènes si l'insertion est dans la région codante ou la création d'allèles instables si l'insertion est dans des régions promotrices. Les excisions, quant à elles, produisent des empreintes de quelques nucléotides pouvant avoir des répercussions majeures telles que le décalage du cadre de lecture. Enfin, des transpositions aberrantes (impliquant des TIR de plusieurs TE) ont été décrites, notamment chez le transposon *impala* (Hua-Van et al., 2002)

Classe 1 / Classe 2 Rétrotransposons / Transposons à ADN



Figure 8. Synthèse des éléments mobiles découverts chez les champignons (adaptation de Daboussi et Capy, 2003)



Figure 9. Répartition des rétrotransposons et transposons chez les champignons filamenteux

Stress abiotiques	Éléments Mobiles	Références
Température	MAGGY	(Ikeda et al., 2001)
Irradiation	Foxy	(Mes et al., 2000)
Stress oxydatif	Skippy	(Anaya et al., 1996)
Stress biotiques		
Isolation de protoplastes	Tnt1	(Grandbastien, 1998)
Culture tissulaire	Ac et Spm	(Peschke et Phillips, 1991)
Infection pathogène	Tnt1	(Pouteau et al., 1994)
Embryonic cell	Tc1/mariner	(Luo et al., 1998)

Tableau 1. Exemples de stress environnementaux ayant pour conséquence la mobilité desTE

Ces dernières ayant souvent pour conséquence le repositionnement d'un promoteur actif en amont d'un gène. Ce mode aléatoire de transposition semble intéressant, car il faciliterait la création de nouveaux allèles potentiellement actifs.

2les réarrangements génomiques engendrés par les TE : des études ont montré la corrélation entre de forts taux de recombinaison chromosomique (Dobigny et al., 2004) et une forte densité de transposons, résultant certainement de recombinaisons ectopiques entre les différentes copies d'un même élément. Ce mécanisme peut toutefois être intéressant pour un champignon ayant perdu son cycle sexuel. Ces recombinaisons, favorisant le brassage génétique, pourraient entraîner la création de nouveaux allèles (Daboussi et Capy, 2003). De par leur proportion non négligeable dans le génome de certains champignons, on leur attribue un rôle important dans les restructurations génomiques. À ce sujet, Michael Thon a publié un article traitant de la perte de synténie chez M. grisea. Cet article est l'un des premiers à avoir mis en évidence l'impact structurel majeur que peuvent induire les TE chez les champignons filamenteux, en générant des « îles » d'éléments mobiles. Ces dernières contribuent à l'augmentation des réarrangements chromosomiques, des duplications géniques et à l'accroissement de la divergence entre des gènes (Thon et al., 2006). La compréhension du contrôle de ces éléments a mis en lumière deux systèmes de régulation : d'une part, l'application de stress environnementaux (Table 1), qu'ils soient biotiques (chocs génomiques comme les isolations de protoplastes, cultures tissulaire ou infection pathogène) ou abiotiques (variations de température, irradiation, stress oxydatifs) ont permis de mettre en évidence la mobilité de TE (Dobigny et al., 2005). D'autre part, des mécanismes de « silencing » épigénétiques ont été mis en évidence après l'introduction de TE au sein de génomes hôtes. Parmi ces mécanismes, trois d'entre eux ont été distinctement décrits : (1) le mécanisme de RIP (pour Repeat-Induced Point mutations), (2) le mécanisme de MIP (pour Methylation Induced Premeiotically) et (3) le Quelling (Annexe). Ces mécanismes de silencing ont essentiellement pour but de contrôler ou protéger le génome hôte d'une invasion, les TE étant les cibles préférentielles de ces mécanismes détruisant de manière irréversible, ou non, la capacité des TE de se déplacer.

La suppression des réarrangements génomiques via le *MIP*, ou encore la rapide divergence de séquences générée par les mutations *RIP*, semble protéger les génomes des réarrangements chromosomiques délétères. L'ensemble des points décrits ci-dessus permet de définir une relation dynamique entre les TE et les génomes hôtes des champignons se traduisant en terme de <u>perte ou de gain de copies</u>. Les gains sont évidemment dus à des transpositions complètes, les pertes à des transpositions incomplètes, pouvant avoir plusieurs origines telles que la non réinsertion dans le génome hôte (dégradation enzymatique du transpososome, dégradation de l'ARNm par ARN interférence) ou encore des duplications ectopiques ou des recombinaisons non homologues. Enfin les <u>transferts horizontaux</u> de transposons ont été mis en évidence chez de nombreuses espèces telles que chez les plantes (Diao et al., 2006), ou les insectes (Lampe et al., 2003 ; Sanchez-Gracia et al., 2005). Chez les champignons, plusieurs évidences de transfert ont été démontrées, notamment chez *Fusarium* spp. dans lesquelles *Fot1* découvert chez *F. oxysporum*, serait passé par transfert horizontal chez *F. solani* et *Neocosmospora* sp. (Daboussi et al., 2002). Les mécanismes de ces transferts sont inconnus.

I.2. La maladie hollandaise de l'orme

I.2.1. Généralités

Vraisemblablement d'origine asiatique, la maladie hollandaise ou graphiose de l'orme (M.H.O.) est apparue pour la première fois en 1919 en Hollande (d'où son nom) ainsi que dans le nord de la France. Elle s'est ensuite propagée à l'Europe entière. Son apparition en Amérique du Nord (États-Unis et Canada) a été datée des années 1930 et a provoqué de très importants dégâts dans les populations d'orme américain (*Ulmus americana*), une espèce très sensible. Cette maladie est la plus destructrice des maladies inféodées aux ormes plantés ou à l'état sauvage.



(adaptation de Brasier, 1988; Brasier, 1996; Brasier et Kirk, 2000; Brasier et al., 2004)

Figure 10. L'histoire d'une invasion (synthèse effectuée à partir des études de C.M. Brasier)

1910/1919 :	Apparition de la M.H.O. en Hollande
1920-40 :	Première pandémie causée par O. ulmi
a.	(1) : expansion en Europe de l'Ouest et en Europe centrale
	(2) : introduction en Amérique du Nord via les exportations commerciales
	(3) : expansion et introduction en Europe de l'Est et au Moyen-Orient
1950-70 :	Deuxième pandémie causée par O. novo-ulmi
b.	 (4) et (5) : émergence simultanée de deux pandémies causées d'une part en Europe par O. novo-ulmi subsp. novo-ulmi et en Amérique du Nord par O. novo-ulmi subsp. americana. (6) : introduction de la sous-espèce nord américaine, O. novo-ulmi subsp. americana, en Angleterre, toujours par le biais du commerce international
	(7) : expansion de la sous-espèce eurasienne, O. novo-ulmi subsp. novo-ulmi, au Moyen- Orient

En **1993**, une troisième espèce, *Ophiostoma himal-ulmi*, a été mise en évidence dans la région du *Himachal Pradesh* située dans le nord de l'Inde. L'étude génotypique de cette nouvelle espèce a donné naissance à de nouvelles hypothèses quant à l'origine de la M.H.O. (voir page 41).

<u>Remarque</u> : L'ensemble de ces vagues d'expansion et de colonisation des diverses espèces et sous-espèces ont généré de nombreuses zones sympatriques. Ces zones semblent être de nouveaux foyers d'hybrides inter ou intraspécifiques, mettant en lumière la récente spéciation de ces deux espèces, *O. ulmi* et *O. novo-ulmi*.
Les champignons de la maladie hollandaise de l'orme possèdent d'intéressantes caractéristiques comme l'**hétérothallisme** (organisme qui nécessite la présence d'un partenaire de type sexuel compatible lors de la reproduction). Cette reproduction est dirigée par deux idiomorphes de type sexuel (gènes *Mat-1* et *Mat-2*), mais également par des gènes d'incompatibilité végétative (estimation de sept gènes *Vic*). Ces champignons présentent également un dimorphisme **cellules levuriformes / mycélium** (c.f. I.2.5.). Les estimations actuellement supposeraient la présence de 8 000 à 10 000 gènes au sein de génomes haploïdes de champignons ascomycètes (Kupfer et al., 1997) répartis, chez les champignons de la M.H.O., sur 4 à 8 chromosomes et totalisant 30 à 40 millions de nucléotides (Dewar et al., 1997 ; Dewar et Bernier, 1995). Trois espèces distinctes de champignons sont responsables de cette maladie et font partie du complexe des Ophiostomatoïdes (Fig. 11).



Figure 11. Relations phylogénétiques entre les différentes espèces du complexe des Ophiostomatoïdes déterminées à partir de marqueurs RAPD (adaptation Pipe et al., 1995) Les agents responsables de la maladie hollandaise de l'orme

I.2.1.1. Ophiostoma ulmi

La première espèce, présentant le plus faible pouvoir pathogène, est *Ophiostoma ulmi*. On estime qu'elle est à l'origine de la première pandémie survenue en Europe et en Amérique

du Nord des années 1920 aux années 1940. Elle a été recensée pour la première fois en Ohio en 1930 (sur *U. americana*) et s'est étendue par la suite à presque toute l'Amérique du Nord. Elle a été également décelée au Manitoba, en 1975, et en Saskatchewan, en 1981. Elle n'avait pas encore été recensée en Alberta en décembre 1991. Tous les ormes indigènes et introduits sont sujets à la maladie et seules quelques espèces (*U. minor*, *U. procera*) semblent lui résister partiellement.

I.2.1.2. Ophiostoma novo-ulmi

Vers les années 1940-1950, presque simultanément est apparue une nouvelle espèce, en Europe de l'Est (Slovakie) et en Amérique du Nord (région des grands lacs). Plus aggressive et possédant une croissance plus rapide, cette espèce, nommée *Ophiostoma novo-ulmi*, a rapidement colonisé l'ensemble de ces deux continents. Elle est responsable de la deuxième pandémie (toujours en cours actuellement) et a par exemple causé, entre 1971 et 1978, la mort d'environ 22 millions d'ormes en Grande-Bretagne. Cette espèce se divise en deux sous-espèces bien distinctes : la sous-espèce *novo-ulmi* [EAN] et la sous-espèce *americana* [NAN]. Ces deux sous-espèces ont des morphologies, des physiologies (Fig. 12) et des génotypes différents. Au Québec, les premiers arbres infectés ont été découverts en 1944 près de St-Ours, dans le comté de Richelieu. Une fois l'aire de distribution de la maladie identifiée vers 1945, il est devenu apparent que le foyer d'infection était le port de Sorel.



Figures 12. Analyses comparatives de la longueur moyenne des périthèces (en μ m) des champignons de la M.H.O. (adaptation de Brasier et Kirk, 2001)

Un examen ultérieur des premières données sur la distribution a montré que la maladie hollandaise de l'orme avait toutefois été introduite avant les années 1940. L'épidémie survenue au Québec s'est déclarée à plus de 300 kilomètres de la limite nord de l'aire de répartition américaine. On en a déduit qu'il y avait eu une introduction distincte au Canada et différente de celle des États-Unis (très certainement à partir de caisses d'ormes transportées sur des navires en provenance de ports européens).

I.2.1.3. Ophiostoma himal-ulmi

Au cours des mois de septembre et octobre 1993, une expédition scientifique a été menée dans la région de *Himachal Pradesh* située dans le nord de l'Inde et ayant pour but de trouver l'origine de la M.H.O. Cette région du monde a été choisie car elle serait également le berceau des principales variétés d'ormes. Bien qu'aucun symptôme n'ait été observé dans la région, un Ophiostoma morphologiquement similaire à *O. ulmi* et *O. novo-ulmi* sp. a été isolé dans des galeries de scolytes de l'écorce d'*U. wallichiana*. Cette nouvelle espèce, nommée *Ophiostoma himal-ulmi*, est également hétérothallique. Cette dernière possède des caractéristiques physiologiques et morphologiques proches des deux autres espèces bien que toutefois différentes, incluant un type de colonies distinctes, la production de longs périthèces ou encore un très fort taux de cérato-ulmine. Le pouvoir pathogène d'*O. himal-ulmi* varie de moyen à fort sur *U. procera* (Pipe et al., 1997). Enfin bien que des hybrides puissent être obtenus entre *O. himal-ulmi* et *O. ulmi* ou *O. novo-ulmi* sp., une isolation reproductive a pu être mise en évidence, confirmant de ce fait qu'*O. himal-ulmi* est bien une nouvelle espèce (Brasier et Mehrotra, 1995).

I.2.2. Origine de la maladie hollandaise de l'orme

L'origine de la maladie hollandaise de l'orme reste à ce jour inconnue. Toutefois, trois hypothèses sont actuellement émises :

(1) La première hypothèse est basée sur un lien commun ancestral des trois espèces d'Ophiostoma responsables de la M.H.O. (ou un ancêtre commun ayant disparu de nos jours). Du fait de sa dispersion spécifique et de son isolement géographique, *Ophiostoma himal-ulmi* est devenue une espèces qui aurait évolué différemment des deux autres et n'aurait pas de lien spécifique avec les pandémies du 20^{ème} siècle causées par *O. ulmi* et *O. novo-ulmi*. Si une telle hypothèse se vérifiait, cette espèce et ses potentiels hybrides pourraient représenter une nouvelle menace pour les populations d'orme d'Eurasie et d'Amérique du Nord.

- (2) La seconde hypothèse repose sur la ressemblance entre O. himal-ulmi et O. novoulmi : l'espèce O. novo-ulmi aurait alors pu évoluer d'une souche ancestrale d'O. himal-ulmi après son introduction et son expansion en Europe. Cette évolution divergente pourrait facilement s'expliquer par la présence d'ormes plus sensibles en Europe qu'en Asie, la présence de conditions météorologiques différentes ainsi que la présence de la première espèce, O. ulmi, en zone sympatrique (cette dernière pouvant jouer le rôle d'une pression sélective, positive ou négative). Toutefois une question subsiste : comment pourrait avoir lieu l'apparition d'une barrière reproductive entre ces trois espèces ?
- (3) Enfin, la dernière hypothèse mettrait en cause la sélection d'un hybride agressif entre O. himal-ulmi et O. ulmi, qui aurait donné naissance à O. novo-ulmi. Il faudrait, de la même façon qu'énoncé dans l'hypothèse #2, que l'espèce O. himalulmi soit présente en Europe. Cette hypothèse est également appuyée par le fait que les hybrides viables de génération F1 entre O. himal-ulmi et O. ulmi présentent une morphologie proche d'O. novo-ulmi.

I.2.3. <u>Homing endonucléases ou introns mobiles : les éléments répétés actuellement</u> présents chez O. ulmi, O. novo-ulmi sp. et O. himal-ulmi

Les mycètes du complexe *Ophiostoma ulmi* ont souvent pu être distingués grâce à leurs simples caractéristiques morphologiques, leurs taux de croissance, leurs virulences ou encore leurs caractères phénotypiques propres. Cependant, ces caractéristiques n'étant toutefois pas infaillibles (Fig. 12), parfois erronées ou nécessitant une expertise dans le domaine (dus aux hybrides interspécifiques), des marqueurs moléculaires plus fiables ont été développés tels que : les polymorphismes des IGS (*intergenic spacers*) de l'ADN ribosomal nucléaire (Hintz et al., 1993) ou le gène de *cerato-ulmine (CU)* ou encore les RAPD. Les variations observées au niveau de la taille des génomes mitochondriaux sont généralement dues à la présence ou l'absence d'introns de type I ou II ainsi qu'aux variations des régions IGS et autres régions non codantes. Les introns de types I et II sont

considérés comme de potentiels ribosomes qui catalysent leur propre excision (Cech, 1990). Ils peuvent être différentiés à partir de leurs caractéristiques biochimiques et structurales (Gibb et Hausner, 2003). D'une part, les introns de type I codent une maturase possédant deux domaines conservés (ces enzymes étant requis dans le repliement de l'ARN mais ne possédant pas d'activité catalytique propre), ou encore deux motifs d'acides aminés LAGLIDADG en symétrie. Ces protéines semblent faire partie du groupe des « *homing* » endonucléases de type LAGLIDADG qui codent spécifiquement pour des enzymes impliqués dans le phénomène d'« *homing* » introns (littéralement intron voyageur [Lambowitz et al., 1999]). L'« *homing* » est une forme de recombinaison site-spécifique. Cette recombinaison est généralement associée à une conversion génique des zones flanquantes. Ces introns de type I ont été également décrits comme codant une endonucléase de type GIY-YIG (Belfort et Roberts, 1997 ; Saguez et al., 2000). D'autre

part, les introns de type II sont des régions non codantes de l'ARN qui catalysent leur propre excision du pré-ARN et permettent la ligation subséquente des régions codantes flanquantes chez certains protozoaires et eucaryotes. Des introns de type I ont été mis en évidence récemment chez *Ophiostoma ulmi* et *O. novo-ulmi* sp. (Gibb et Hausner, 2005).

I.2.4. Cycle de la maladie (transmission, adhésion, colonisation et invasion)

La maladie touche une multitude d'espèces d'ormes, allant des espèces très sensibles indigènes comme l'orme d'Amérique (*Ulmus americana* L.), l'orme rouge (*U. rubra* Mühl.), l'orme liège (*U. thomasii* Sarg.) aux espèces introduites comme l'orme champêtre (*U. procera* Salisb.) et l'orme de montagne (*U. glabra* Huds.). Il existe des espèces peu sensibles telles qu'*U. minor*, *U. campestris U. pumila* ainsi que de nombreuses espèces d'ormes asiatiques (*U. japonica*, *U. davidiana* var. *japonica* R.). Le champignon se développe dans les vaisseaux conducteurs de la sève brute. Les premiers symptômes de la maladie se manifestent habituellement entre la mi-juin et la mi-juillet. Les feuilles d'une ou plusieurs branches flétrissent, s'enroulent, se dessèchent et prennent une coloration jaunâtre à brunâtre, sans défoliation.



Figure 13. Cycle de la maladie hollandaise de l'orme (Agrios, 2005)

Lorsqu'on coupe une branche atteinte, on observe que le dernier anneau de croissance est brunâtre. Une fois l'écorce retirée, on constate que le bois est couvert de stries brunes. En Amérique du Nord, la maladie est transmise par des insectes du groupe des scolytes et a deux vecteurs principaux : le scolyte de l'orme, Hylurgopinus rufipes (Eichh.) et le petit scolyte européen de l'orme, Scolytus multistriatus (Marsh.) (Fig. 13). Les insectes creusent des galeries sous l'écorce des arbres malades ou morts où ils se reproduisent. Les spores du champignon (phase levuriforme) adhèrent aux corps des scolytes adultes. Au printemps, ces derniers migrent vers d'autres arbres pour se nourrir, soit de l'écorce des ramilles et des rameaux de la cime, dans le cas de S. multistriatus, soit des branches et des jeunes troncs, dans le cas de H. rufipes. Ils contaminent ainsi leurs nouveaux hôtes. Les spores alors libérées dans les galeries des scolytes peuvent pénétrer dans les tissus vasculaires (phase mycélienne). Privés de sève en raison de l'obstruction des vaisseaux du xylème par la multiplication des spores du champignon, l'arbre flétrit et meurt plus ou moins rapidement dépendamment de la taille des vaisseaux obstrués (un vaisseau obstrué de diamètre important causera plus de dommage qu'un vaisseau obstrué de faible diamètre). C'est d'ailleurs en partie dû à cette caractéristique que certains ormes semblent plus résistants que d'autres : U. minor possède des vaisseaux conducteurs de diamètre plus petits mais en plus grandes quantités comparés à U. americana.

De par cette particularité structurale, cette première espèce est moins sensible à une attaque des champignons de la M.H.O. (Solla et Gil, 2002). Enfin, lorsque les ormes sont rapprochés, la maladie peut se transmettre par leurs racines qui se greffent les unes aux autres.

Le **pouvoir pathogène** d'un organisme est sa capacité à provoquer des modifications génotypiques ou phénotypiques diminuant la valeur adaptative d'un individu (Cutler, 1991). Il dépend de son pouvoir *invasif* (capacité à se répandre dans les tissus et à y établir un/des foyers infectieux), et de son pouvoir *toxicogène* (capacité à produire des toxines). Chez *O. novo-ulmi* sp., la pathogénécité est liée au stade du champignon : les cellules levuriformes ayant pour but la dissémination à grande échelle du champignon dans l'arbre (production importante de biomasse aisément mobile dans la sève brute), les cellules mycéliennes ayant pour but la contamination (grâce à leur pouvoir pénétrant, via les appressoriums, qui n'ont toutefois pas été mis en évidence chez les agents de la M.H.O.). Ce dernier stade est considéré comme le stade infectieux de la maladie. Considérée comme <u>polygénique</u>, l'étude du pouvoir pathogène des champignons de la M.H.O. via l'analyse de la descendance d'un croisement intraspécifique

entre deux souches de la sous-espèce européenne (AST27 X H327) a abouti à la cartographie d'un gène, *Pat1* (Et-Touil et al., 1999), impliqué dans le pouvoir pathogène. De nombreux mécanismes physiques sont impliqués dans le pouvoir pathogène d'un organisme (Tucker et Talbot, 2001). Les premiers sont les protéines d'adhésion des spores. Une fois pénétré dans les tissus corticaux de son hôte, une succession de voies de biosynthèse telles que celles des intégrines et hydrophobines sont activées. Les champignons filamenteux produisent également des protéines spécialisées dans le développement des structures aériennes, dans les interactions entre l'hyphe et des composés hydrophobes, connues sous le nom **d'hydrophobines**. Elles jouent également un rôle dans la formation des spores, le développement des fructifications et la formation de l'appressorium. Un exemple bien connu est celui de l'hydrophobine de type II nommée cérato-ulmine (CU) qui est impliquée dans l'adhésion des spores d'*Ophiostoma sp.* Découverte 25 ans auparavant, cette protéine était considérée comme une toxine directement liée au pouvoir pathogène.



(A)

(B)



(A) Migration chromosomique réalisée sur gel PFGE et hybridation Southern avec sondes spécifiques de chacun des chromosomes. (B) Visualisation générale des polymorphismes de longueurs des chromosomes entre CESS16K et FG245Br-O (FG) basée sur la corrélation des hybridations obtenues sur chacun des chromosomes. On voit ici clairement que le chromosome surnuméraire VII chez CESS16K proviendrait du chromosome 1 (Dewar et al., 1995).

En effet, plus une souche était agressive, plus cette dernière produisait de la CU. Cependant lorsque des mutants dont on avait inactivé le gène CU, étaient injectés dans des ormes sains, aucune différence de pathogénicité ne pouvait être mise en évidence (Bowden et al., 1996 ; Bowden et al., 1994). La CU est produite en grandes quantités pendant le développement des spores et contribue à l'hydrophobicité de ces dernières, augmentant ainsi leur dispersion. Ce rôle est conforme à ceux de facteurs de <u>pathogénicité secondaires</u> comme les médiateurs d'adhérence. Il existe enfin d'autres enzymes ont été mis en évidence chez les champignons de la M.H.O. comme les pectinases ou polygalacturonase qui sont des enzymes impliqués dans les dégradations de molécules de défense sécrétés par la plante.

I.2.5. Variabilités intra- et inter-spécifiques = impacts visibles d'éléments mobiles ?

I.2.5.1. Polymorphisme chromosomique de la souche CESS16K

La rupture chromosomique réverse a été mise en évidence chez Ophiostoma novo-ulmi après des études de migrations chromosomiques par K. Dewar (Dewar et Bernier, 1995, Dewar et al. 1997). Tout d'abord, selon les souches étudiées, quatre à huit bandes chromosomiques étaient présentes, se divisant en trois classes distinctes. Toutefois, une exception toutefois, la souche CESS16K (O. novo-ulmi subsp. americana) contient une bande d'ADN unique ayant le même champ de migration qu'un chromosome de Saccharomyces cerevisiae d'une taille de 0,95 Mb; cette bande unique correspondait au plus petit chromosome observé (Fig. 14). Ce type d'observation a été de nombreuses fois décrit chez les champignons sans qu'aucune explication ne soit toutefois trouvée (Ustilago maydis, Kinscherf et Leong, 1988; M. oryzae, Talbot et al., 1993 ou encore C. cinereus, Zolan et al., 1994). Aucune corrélation n'a pu être faite entre un profil chromosomique et une espèce donnée, les profils caryotypiques étant partagés à la fois par les deux espèces. La souche CESS16K a été plus amplement analysée, en raison de cette particularité chromosomique. L'analyse de 16 souches F1 issues du croisement entre FG245Br-O et CESS16K (O. novo-ulmi subsp. americana) a démontré que le chromosome unique est une composante intégrale du génome de CESS16K (dont les gènes sont exprimés normalement). De ce fait, des recombinaisons chromosomiques peuvent se produire par suite de la méïose mettant en évidence, pour la première fois, la plasticité du génome de ce mycète. Enfin, la transmission de ce polymorphisme de longueur du chromosome I chez CESS16K a été vérifiée dans la génération F2 pour déterminer la reproductibilité d'un réarrangement génomique aboutissant à la conversion d'un chromosome de 1,0 Mb en un chromosome de 800 Kb (Fig. 14).

Ces analyses ont permis de démontrer que le chromosome de 1,0 Mb s'apparie avec lui-même ou avec le chromosome de 800 Kb, l'ensemble de la descendance analysée ayant hérité de l'une des deux formes (la forme de 800 Kb n'étant toutefois présente que dans 9 % des cas). Cette modification montre que les processus méiotiques sont responsables de ce PLC (Polymorphisme de longueur chromosomique), et que des souches d'*Ophiostoma novo-ulmi* sp. ayant diverses architectures génomiques peuvent demeurer sexuellement compatibles (Dewar et Bernier, 1993 ; Dewar et Bernier, 1995 ; Dewar et al., 1997).

Liens avec les TE : Les premiers transposons ont été mis en évidence (*Ac* pour *Activator* et *Ds* pour *Dissociation*) grâce à leur capacité de générer des cassures chromosomiques (McClintock, 1951 ; i.e. Historique p. 24). La dénomination *Dissociation* provient de la capacité du transposon à induire une cassure chromosomique à son site d'insertion. Cette cassure chromosomique impliquant la présence d'un autre locus, se révéla être par la suite l'élément *Activator*. La mise en évidence a été observée par cassure de la partie terminale du chromosome 9 du maïs (Raslton et al., 1989 ; Fedoroff, 2002).

I.2.5.2. Variabilité intraspécifique du pouvoir pathogène

Des inoculations en champs sur des ormes moyennement résistants, tels qu'*Ulmus procera* et *U*. sp. X Commelin, ont été effectuées avec la descendance d'un croisement intraspécifique entre AST27, un isolat d'*O. novo-ulmi* subsp *novo-ulmi* modérément virulent et H327, un isolat agressif de la même sous-espèce. Ces dernières ont confirmé les résultats d'une étude précédente (Brasier, 1987) indiquant que la différence de pathogénicité semble dirigée par un gène nucléaire simple. Ce gène potentiel de pathogénicité, nommé *Pat1*, est le premier gène putatif de pathogénicité identifié chez cette espèce. Dans une analyse de ségrégation aléatoire, 10 marqueurs RAPD (ADN polymorphe amplifié aléatoirement) ont été liés à *Pat1* (Fig. 15).



Figure 15. Cartes génétiques des fragments introgressés de l'espèce O. ulmi vers O. novo-ulmi sp. novo-ulmi (adaptation Et-Touil et al., 1999, Paoletti et al., 2005)

Les fragments présentés ci-dessus sont des fragments cartographiés avec des marqueurs RFLP à partir d'une descendance issue d'un croisement *O. ulmi* X *O. novo-ulmi* subsp. *novo-ulmi*. Ce type particulier de cartographie permet de mettre en évidence les régions introgressées d'un génome d'une espèce à une autre. Les fragments (A) et (B) mettent en évidence les introgressions de gènes de compatibilité végétative (*Vic-w1a* et *Vic-w1b*) et de reconnaissance sexuelle (*MAT-1*) alors que le fragment (C) met en évidence l'introgression d'un gène de pathogénécité (*Pat-1*) et une copie d'un transposon à ADN (*OPHIO2-int.*)

Fait intéressant, cinq amplicons RAPD identifiés dans AST27 provenaient d'*O. ulmi* et non d'*O. novo-ulmi* sp. et deux d'entre eux ont été liés à *Pat1*. Ceci suggère que l'allèle *Pat1* conférant une virulence modérée à AST27 aurait pu être acquis d'*O. ulmi* par introgression (ce locus serait situé sur le chromosome II). À ce travail de cartographie se sont ajoutés des marqueurs d'intérêt tels que *COL1*, gène impliqué dans la croissance et le dimorphisme ou encore *chsA*, codant pour une chitine synthétase impliquée dans la composition des parois cellulaires chez les champignons (Hintz, 1999 ; Peirera et al., 2000), ou encore le gène de cérato-ulmine, précédemment catalogué comme gène de pathogénécité (Richards et al., 1993). Des études ultérieures ont finalement mis en évidence que la cérato-ulmine jouait un rôle dans la cohésion des spores sur les scolytes avant la pénétration (Temple et al., 1997). Enfin de nombreux polymorphismes génétiques inter- comme intra-spécifiques ont été mis en évidence comme ceux situé sur le locus *MAT* (codant un idiomorphe de reconnaissance sexuelle) ou les loci *Vic* (codant les gènes d'incompatibilité végétative) (Et-Touil et al., 1999). L'introgression du locus *MAT-1-2*

a été démontrée chez *O. novo-ulmi* à partir d'*O. ulmi*, les hybrides interspécifiques servant de ponts génétiques entre les deux espèces (Paoletti et al., 2006).

Liens avec les TE : Les transposons et leur mobilité (insertions) ont déjà été associés à une perte ou à un gain de pathogénécité chez les champignons, l'exemple le plus connu est celui de *M. oryzae* et des transposons *Pot3* (gain) et *MGR* (perte). En effet, il a été démontré que l'insertion d'une copie de l'élément *Pot3* dans le promoteur du gène *AVR-Pita* générait une variation de la pathogénécité chez le cultivar *Yashiro-mochi* (peut-être en raison de la modification engendrée par l'insertion qui permet au gène *AVR-Pita* de ne pas être reconnu par le gène de résistance *Pita* contenu dans de nombreux cultivars de riz (Kang et al., 2001 ; Zhou et al., 2007).



Division cellulaire et commutation sexuelle chez *Saccharomyces cerevisiae* entre une cellule mère et une cellule fille après deux générations. Ce système permet à la levure du boulanger de pouvoir former, à partir d'une unique cellule, une colonie capable de se reproduire et se multiplier.



Figure 16. Commutation des gènes de reconnaissance sexuelle chez la levure par l'intermédiaire de rupture chromosomique (inactivation du gène sexuel actif, par exemple MATa) et conversion génique (remplacement et réactivation du nouvel allèle MAT α) (adaptation de Brown, 2006))

I.2.5.3. Mating type switching de la souche VA

Comme il a été décrit précédemment, les trois espèces d'Ophiostoma responsables de la M.H.O. sont hétérothalliques et chacune d'entre elles possède un unique locus MAT bi-allèlique, grâce auquel des individus portant le locus MATA (MAT-1 par convention) peuvent se croiser avec des individus ayant le locus MATB (MAT-2 par convention) (Brasier & Mehrotra 1995). Les gènes de *mating type* (gènes *MAT* ou type sexuel) ont été bien étudiés comme marqueurs évolutifs en raison de leur évolution interspécifique et de leur conservation intra-spécifique. La reproduction est alors possible si une souche d'un type sexuel est en contact avec une souche de l'autre type. Il est bien évident que d'autres mécanismes entrent en jeu lors de cette reproduction, par exemple les gènes d'incompatibilité végétative (gène Vic) impliqués dans les fusions caryotiques et la reconnaissance du non-soi (Paoletti et al., 2007) Des expériences en laboratoire ont permis de mettre en évidence que certaines souches, dont la souche VA (MAT-1), pouvaient se croiser avec d'autres souches du même type sexuel. Une hypothèse aurait été formulée quant à la capacité de cette souche d'effectuer des commutations de type sexuel (ou mating type switching). La première description de ce système a été mise en évidence chez Saccharomyces *cerevisiae* (Fig. 16). Cette dernière est capable de commuter son allèle MATa en allèle MAT α (et inversement) pour s'autoféconder. De ce fait, une cellule haploïde d'un type sexuel peut ainsi générer une colonie de nouvelles cellules qui présentera les deux types sexuels MATa et MATα. Comme il a été décrit précédemment, la commutation sexuelle est le remplacement d'un allèle par un autre. Ce remplacement est possible car les cellules de levure possèdent une copie additionnelle inactivée (par hypo-acétylation des histones) de chacun des allèles MAT nommés HMLa et HMRa (pour Hidden MAT Left a et Hidden MAT Right a) (Brown, 2006). Une endonucléase (HO) permet le clivage de l'allèle actif (par exemple MATa en stade G1 du cycle de reproduction cellulaire de la cellule mère) et l'activation d'un activateur de recombinaison (RE pour *Recombination Enhancer*) permet la conversion génique, initié par la cassure de l'ADN double brin de l'allèle opposé, restituant de ce fait une allèle MAT active et de locus opposé (par exemple MATα) (Haber, 1998).

Liens avec les TE : En 1990, Weinstock et collaborateurs ont mis en évidence que des insertions multiples du rétrotransposon Ty1 dans la région HML α pouvaient activer de manière autonome la commutation de type sexuel chez la levure (Weinstock et al., 1990). Cette équipe a

également corrélé le taux d'insertions du rétrotransposon avec la fréquence des commutations : plus on a d'insertion dans la région HML α et plus il est facile pour la cellule de commuter son allèle MAT**a** en MAT α . Ceci est expliqué par un désarrangement de la structure hétérochromatinienne qui maintient la région HML α silencieuse à toute conversion génique. Il a été démontré que les rétrotransposons *Ty1* peuvent générer des cassures d'ADN double brin. Si ces dernières ont lieu à proximité des zones HML α ou HMR**a**, elles peuvent alors déclencher le processus de conversion génique conduisant à la commutation sexuelle (Garfinkel, 2005).

L'ensemble de ces variations précédemment décrites, qu'elles soient d'ordre phénotypique ou génotypique, nous ont amené à nous poser les questions suivantes : existe-t-il d'autres types de séquences répétées dans le génome des champignons de la maladie hollandaise de l'orme qui pourraient générer ce type de variabilité, notamment les éléments de transposition ? Peut-on relier ces derniers à des polymorphismes observés chez *Ophiostoma* sp. ? Avant de pouvoir dégager de telles relations, il a été nécessaire de mettre en évidence la présence de nouvelles séquences répétées, ce qui a débouché sur l'étude des transposons à ADN.

I.3. Objectifs et hypothèses de recherche

Le premier objectif de cette thèse a été de caractériser des éléments mobiles dans les génomes d'*O ulmi* et *O. novo-ulmi*. D'une part, une première hypothèse tendait plus à vérifier le paradigme énonçant que les TE, de type 1 ou 2, sont présents dans le génome de **tous** les organismes vivants :

«... transposable elements are present in the genomes of all life-forms...» (Watson et al., 2004).

D'autre part, tel qu'expliqué précédemment, la taille des éléments mobiles semblant être corrélée à la taille des génomes (Kidwell et Lisch, 2002), il nous a paru plus judicieux de rechercher des TE de taille restreinte étant donné que les génomes des champignons responsables de la M.H.O. sont d'environ 30 à 40 Mb. Ainsi, la caractérisation de transposons à ADN a été décrite dans le **chapitre II** de cette thèse. Ce premier chapitre nous a également permis de décrire pour la première fois, à partir de séquences amplifiées d'*OPHIO3* chez nos différentes espèces, le mécanisme de silencing RIP (*Repeat induced point mutation*), spécifique aux champignons. L'analyse subséquente des séquences d'*OPHIO3* nous a également permis de développer une technique de visualisation, appelée *Cumulative Transition Score* (CTS), basée sur la pondération des mutations impliqués dans ce mécanisme.

Une fois ces éléments (au nombre de trois et nommés *OPHIO1, OPHIO2* et *OPHIO3*) caractérisés, le **chapitre III** de cette thèse examine la mobilité de deux des trois transposons : *OPHIO1* et *OPHIO2*. La mobilité étant une des caractéristiques majeures des éléments de transposition, il nous a semblé indispensable d'en faire une mise en évidence. Plusieurs choix quant à la visualisation de la mobilité des transposons étaient possibles, mais celui de l'induction via des stress biotiques ou abiotiques a été retenu.

L'observation de l'identité des séquences d'*OPHIO1* et d'*OPHIO2* a montré des caractéristiques particulières. D'une part, bien qu'étant tous deux similaires en de nombreux points tels que la longueur des diverses régions les constituant, la position des codons d'initiation et de terminaison de leur région codante, ces deux copies ne présentent que 75 % d'identité. D'autre

part, le fait que ces deux transposons soient spécifiques de deux espèces distinctes, il est légitime de se demander si les génomes de ces espèces pourraient avoir un rôle dans l'explication de ces variations nucléiques. L'hypothèse que nous avons émise est que des mécanismes induisant des mutations au sein des séquences d'ADN soient faisant l'object d'une possible sélection positive. Si tel est le cas, les fluctuations du ratio ω peuvent être un excellent moyen de les visualiser. Le **chapitre IV** présente une étude de sélection positive mise en évidence au sein de régions particulières des séquences chez différents TE provenant de génomes eucaryotes ou procaryotes.

Chapitre II : Caractérisation des premiers transposons à ADN chez les agents responsables de la maladie hollandaise de l'orme

II.1. Avant propos.

L'étude présentée dans ce chapitre visait à caractériser des transposons à ADN chez les agents de la maladie hollandaise de l'orme. Dans cette optique, nous avons présenté la reconstitution et l'analyse de leur séquence complète, leur étude phylogénétique ainsi que leur répartition. Les trois transposons, nommé *OPHIO1*, *OPHIO2* et *OPHIO3*, serviront ensuite de bases aux études des chapitres suivants. Il est également important de mentionner que le transposon *OPHIO3* a fait l'objet d'une étude particulière due aux mutations de sa séquence. Nous avons pu établir la première analyse chez *Ophiostoma* sp., de *silencing* induit par des mutations de type *Repeatinduced point*. Nous avons pu également établir que des copies, potentiellement actives, peuvent être introgressées dans de nouveaux génomes hôtes grâce à des hybridations interspécifiques. L'essentiel de ces résultats a été présenté à la «96^e Assemblée annuelle de la Société de protection des plantes du Québec », les 17 et 18 juin 2004 à Sherbrooke (Canada) ainsi qu'à l'« APS - CPS – MSA joint meeting », du 29 juillet au 2 août 2006, à Québec (Canada). Cette dernière conférence m'a également donné l'opportunité d'être co-modérateur de la session « the Fungi – Genetics/Molecular Biology/Cell Biology » section « Molecular/Cellular Plant-Microbe Interactions ».

Ces résultats ont été également publiés par Bouvet *et al.* en 2007 dans la revue *Fungal Genetics and Biology* 44: 430-443.

<u>Réalisations et remerciements</u> : V. Jacobi et L. Bernier = corrections/rédaction. C. Feschotte et A. Hua-Van = discussions/suggestions et lecture du manuscrit inital

43

II.2. Résumé.

Les éléments de transposition (TE) sont des constituants fondamentaux des génomes eucaryotes (et procaryotes) et peuvent contribuer à la plasticité et l'évolution de ces derniers en engendrant des cassures chromosomiques, des recombinaisons non homologues, des délétions/duplications de gènes, etc. Dans ce chapitre nous décrivons, pour la première fois, trois transposons à ADN découverts chez les agents de la maladie hollandaise de l'orme, Ophiostoma ulmi et O. novoulmi sp., nommés OPHIO1, OPHIO2 et OPHIO3. Nous avons démontré que les transposons OPHIO, s'inscrivant dans la superfamille des *Tc1/mariner*, famille des *Fot1/pogo*, la plus représentée chez les champignons ascomycètes, possèdent différents patrons de distribution et différentes spécificités d'hôtes. Grâce au transposon OPHIO2, mis en évidence chez O. ulmi, nous avons pu démontrer que les hybrides interspécifiques peuvent jouer le rôle de pont génétique entre espèces et favoriser la transmission de copies de transposons. En effet, AST27, introgressant interspécifique de l'espèce O. novo-ulmi sp. novo-ulmi, a intégré au sein de son génome une copie nommée OPHIO2-int., qui pourrait s'avérer potentiellement active. Le transposon OPHIO3, découvert dans une banque issue d'hybridation soustractive (SSH library), nous a permis de démontrer la présence du mécanisme de « Repeat-Induced Point mutations » (RIP) décrit par Selker en 1987. Ce mécanisme est le premier mécanisme de silencing génétique mis en évidence pour la première fois chez Neurospora crassa. C'est un mécanisme prétranscriptionnel irréversible, qui convertit les dinucléotides CpA en TpA (et par extension, les dinucléotides TpG en TpA sur le brin complémentaire). Cette particularité nous a permis de vérifier l'exactitude de ce mécanisme par le calcul de ratios développés par Margolin, mais nous a également permis de développer une méthode de visualisation des mutations RIP, complémentaire au calcul de ces ratios. Cette technique, appelée CTS pour Cumulative Transition Score, est basée sur la pondération des mutations RIP et permet aisément de visualiser une potentielle préférence mutationnelle (sur le brin matrice ou complémentaire).

II.3. Abstract

Transposable elements (TE) are fundamental components of eukaryotic genomes and can contribute in various ways to genome plasticity and evolution. We describe here the first three DNA transposons in the Dutch elm disease (DED) pathogens *Ophiostoma ulmi* and *O. novo-ulmi*, named *OPHIO1*, *OPHIO2* and *OPHIO3*. We demonstrate that *OPHIO* transposons, which show high homology to *Fot1/pogo* TE within the *Tc1/mariner* superfamily, have different distribution patterns and specificity in the DED fungi and that interspecific hybrids could act as genetic bridges for transmission of TE between closely related fungal species. *OPHIO3* was found to have undergone repeat-induced point mutations (RIP). We have also developed a complementary method to Margolin's ratios based on the computation of cumulative transition scores (CTS) in order to rapidly visualize RIP signatures on individual DNA strands of *OPHIO* transposons and TE found in other ascomycete fungi.

II.4. Introduction

Transposable elements are repetitive and ubiquitous genomic DNA sequences present in most eukaryotic organisms (Craig et al., 2002). They account for 15% of the genome in fruit flies, 45% in humans and more than 80% in barley and *Lilium* species (Feschotte et al., 2002; Goyon et al., 1996). In the fungal phylum Ascomycota, TEs have been estimated to represent about 14% of the genome (Goyon et al., 1996). The principal characteristic of TEs is their ability to move (or transpose) from one chromosomal location to another and to replicate themselves. This characteristic is what distinguishes the autonomous TE from the non-autonomous ones, depending on their ability to synthesize proteins that allow them to transpose. Mobility is the essence of their evolutionary success and is certainly at the origin of implications for the dynamics between TE and host genomes (Bowen and Jordan, 2002). TE are generally believed to be a major driving force of genome evolution and plasticity by generating chromosomal rearrangements and breakage, and promoting ectopic recombination, as well as variability in gene expression (Kidwell and Lisch, 2002). For example, transposon MGR386 was found to decrease virulence in the plant pathogenic fungus Magnaporthe oryzae (formerly M. grisea) (Farman et al., 1996), whereas Pot3 had the opposite effect (Kang et al., 2001). Transposable elements can be divided into two major classes (type I = retrotransposons, type II = DNA transposons) according to their structural

organization and transposition mechanism. As of 2005, 52 Class I and 58 Class II TE families have been described in 22 species or subspecies in the phylum Ascomycota. Of the currently nine recognized superfamilies, only three have been identified in fungi: *Tc1/mariner*, *hAT* and *Mutator*, with an over representation of the *pogo/Fot1* group within the *Tc1/mariner* superfamily (Chalvet et al., 2003; Daboussi and Capy, 2003; Kempken, 1999; Zhang et al., 2001). DNA transposons move from a chromosomal site to another through a DNA intermediate. They demonstrate specific structures, such as Terminal Inverted Repeats (TIRs) that flank an Open Reading Frame (ORF) encoding a transposase enzyme required for mobility and a Target Site Duplication (TSD) recognized during insertion (Richardson et al., 2006).

Following a collaborative initiative of four research teams, the Canadian Ophiostoma Genome Project was launched in 2001 with the aim of understanding functional and structural genomic variations in three ascomycete fungi belonging to the genus *Ophiostoma* (Sordariomycetes, Ophiostomatales, Ophiostomataceae): the saprobe *O. piceae* (Dogra and Breuil, 2004) and the pathogens *O. ulmi* and *O. novo-ulmi* (Bernier et al., 2004). The latter two species were responsible for two devastating pandemics of Dutch Elm Disease (DED) which have resulted in heavy losses in elm populations. The first pandemic, caused by *O. ulmi*, started in northwestern Europe in 1910 and lasted until the late 1960's (Brasier et al., 2004). The second, much more destructive, pandemic is thought to have begun in the 1940's and is due to the highly pathogenic *O. novo-ulmi* subsp. *americana* (formerly known as the North American [NAN] race) and *O. novo-ulmi* subsp. *novo-ulmi* (formerly called the Eurasian [EAN] race) (Brasier et al., 2004; Pipe et al., 1995).

However, although speciation between *O. ulmi* and *O. novo-ulmi* was demonstrated (Brasier et al., 1993), rare interspecific hybrids with reproductive capacities can be found in natural populations (Brasier et al., 1998). *O. ulmi* and *O. novo-ulmi* display extensive chromosome polymorphisms (Dewar and Bernier, 1993, 1995; Dewar et al., 1997). Pathogenicity of these species is believed to be under additive polygenic control (Brasier, 1986) but the molecular bases of this important trait remain unresolved.

The DED fungi are vectored principally by elm bark beetles (*Coleoptera*, *Curculionidea*, *Scolytinae*). Upon emerging from breeding galleries in which they may have acquired *Ophiostoma* spores, young adults fly to the crown of healthy elms where they will feed in twig

crotches. In doing so, the beetles inoculate the pathogen into xylem vessels where it will multiply and induce the blockage of the water-conducting tissue within the tree. Soon after infection, susceptible elm species will display characteristic wilting, curling and yellowing of leaves, and usually die (Lanier and Peacock, 1981). In order to develop a global understanding of the genome of the DED fungi and to supplement ongoing studies on differential gene expression, we explored whether strains of O. ulmi and O. novo-ulmi harbored DNA transposons in their nuclear genome. We describe here the first three transposable elements found in the Ophiostoma genome, termed OPHIO1, OPHIO2 and OPHIO3. These TE show various characteristics of type II transposons, including sequences coding for a transposase, TIRs and TSD. Comparative analysis of Southern profiles allowed us to observe that the three OPHIO transposons differed in their distribution patterns and specificities in the DED fungi. In addition, it would seem that interspecific hybrids may have acted as genetic carriers for TE between O. ulmi and O. novo-ulmi. On the other hand, only OPHIO3 exhibited mutation patterns typical of repeat-induced point (RIP) mutation. RIP is a process that efficiently detects and mutates duplicated sequences during the sexual cycle. The mechanism identifies duplications and introduces C:G to T:A mutations into both copies of the duplicated sequences (Galagan and Selker, 2004; Selker, 2002). The occurrence of RIP silencing is usually deduced from the computation of Margolin's ratios (Margolin et al., 1998). We developed a complementary method, called "cumulative transition score" (CTS), to visualize C:G to T:A transitions in *OPHIO3* and transposase sequences from various ascomycete fungi.

II.5. Materials and Methods

II.5.1. Strains, media and culture conditions

All strains of *O. ulmi* and *O. novo-ulmi* used in this study (Table 2) are kept in the DED collection of the Centre d'étude de la forêt (CEF, Laval University, Quebec). The collection consists of one set of samples kept under liquid nitrogen and one set kept in 15% glycerol at -80 °C.

Table 2. Ophiostoma ulmi and O. novo-ulmi strains used

Strain	Species/subspecies	Origin (reference)*
Н5	O. ulmi	France, Doulogne, J. N. Gibbs, (1)
H200	O. ulmi	Ireland, County Offaly, A. Mangan, (2)
Q412T	O. ulmi	Canada, St-Romuald, QC, (3)
R21	O. ulmi	Romania, Bozovici, C. M. Brasier, 1980 (4)
W9	O. ulmi	UK, Thames, Oxfordshire, J. N. Gibbs, 1971 (5)
785401	O. ulmi	France, Meurthe et Moselle, J. Pinon, 1978
H327	O. novo-ulmi subsp. novo-ulmi	Slovakia, Brezno-Nizke, Tatry, J. Jamnicky, 1979 (5)
AST27	O. novo-ulmi subsp. novo-ulmi	Iran, Assalem Forest, C. M. Brasier, 1977 (4)
AST20	O. novo-ulmi subsp. novo-ulmi	Iran, Assalem Forest, C. M. Brasier, 1977 (5)
CKT11	O. novo-ulmi subsp. novo-ulmi	Iran, Chachkam Forest, C. M. Brasier, 1977 (4)
788601	O. novo-ulmi subsp. novo-ulmi	France, Vienne, J. Pinon, 1978
15401	O. novo-ulmi subsp. novo-ulmi	France, Meurthe et Moselle, J. Pinon, 2001
W2	O. novo-ulmi subsp. americana	UK, Tewkesbury, Gloucestershire, J. N. Gibbs, 1972 (5)
VA	O. novo-ulmi subsp. americana	USA, Virginia, L.R Schreiber and A.M. Townsend, 1975 (6)
MH75	O. novo-ulmi subsp. americana	Canada, Toronto, ON, M. Hubbes 1975 (3)
416R14	O. novo-ulmi subsp. americana	Canada, Quebec City, QC, Laboratory Strain, L. Bernier
FG245	O. novo-ulmi subsp. americana	UK, Crediton, Devon, C. M. Brasier, 1980 (4)
Cess16k	O. novo-ulmi subsp. americana	Canada, Sault Saint-Marie, ON, S. Takai (6)
17270	O. novo-ulmi subsp. americana	Canada, Quebec City, QC, L. Innes, 2003
15927	O. novo-ulmi subsp. americana	Canada, Quebec City, QC, L. Innes, 2003

* Not necessarily the reference of the original description.

(1) Temple et al., 1997.	(4) Brasier, 1986.
(2) Brasier and Kirk., 2001.	(5) Brasier et al., 1998.
(3) Bernier and Hubbes, 1990.	(6) Takai, 1980.

The strains were initially grown and maintained on solid Ophiostoma complete medium (OCM) (Bernier and Hubbes, 1990). Yeast-phase cultures were initiated by seeding 50 ml of liquid OCM in 125 ml Erlenmeyer flasks with 3 mm² agar plugs cut from the edge of actively growing mycelial colonies on agar plates. Cultures were grown at 22 to 24°C with agitation (120 rev/min) for 3 days. Then yeast cells were harvested by centrifugation, washed with 50-mM EDTA solution and lyophilized. Cells were ground to a fine powder under liquid nitrogen in a mortar with pestle and stored at -80°C until DNA extraction or processed immediately.

Name	Sequence 5'> 3'	Annealing temperature (°C)
Transposase_L	ATYHHVCCGCAAAACMG	54.8
Transposase_R	GGAAACCAYTGYTGYTG	52.4
TIR_OPHIO1	TAAGTAATCCACCCTGCTTACCGGACG	67.6
TIR_OPHIO2	TAAGTGATCCACCCTACTTTTCACACACA	62.9
TIR_OPHIO3_5'	CCCTTTTGGGCCACCCCTG	68.6
TIR_OPHIO3_3'	CCCCTTTTAGGCCACCCCCC	68.6
GRLI_OPHIO1_L1	TTCAACTCTTGGCAATTCACTGCTACT	63.1
GRLI_OPHIO1_R1	GCAGGGGCAGAACATGGAAT	62.4
IPCR_OPHIO1_L2	CGCAAATCACTGCTAGAGAGACTGTTA	64.6
IPCR_OPHIO1_R2	TTTTCTTCCGTATATTCCCGCATTA	59.7
IPCR_OPHIO2_L1	ACGCGTCTGTGGTCTGTAATAA	60.8
IPCR_OPHIO2_R1	GAAAATAATCAAAGGAGGGAGTGAT	59.7
IPCR_OPHIO2_L2	GCTGCCGGTGAGTCTGTTACTGCTA	67.9
IPCR_OPHIO2_R2	ACGCGTCCCGGAAAGAAGTGTTATT	64.6
IPCR_OPHIO3_L1	GTCCTTCAACCGCTCGATATCTCTT	64.6
IPCR_OPHIO3_R1	AACCGGCCTACGATTAAGCTATTGG	64.6
IPCR_OPHIO3_L2	TATCGAATTGGAAGAAAATGAAAGG	58
IPCR_OPHIO3_R2	TATAGAGGCATCGTTGGATGTAGAA	61.3
<i>OPHIO3</i> _1414_L	AAACCCTTAAGCGTATACTGATCGAA	61.4
<i>OPHIO3</i> _1414_R	AAGGAAGTTGCGCTTATTCTAGTGA	61.3

Table 3. Primers used for PCR cloning of DNA transposons in Ophiostoma sp.

II.5.2. Genomic DNA extractions, degenerate primers and PCR conditions

High molecular weight fungal genomic DNA was obtained following the protocol of Zolan and Pukkila (Zolan and Pukkila, 1986). DNA was resuspended in Tris-EDTA buffer (10-mM Tris-HCl, 1-mM EDTA, pH 8.0) and stored at -20 °C. Degenerate primers were designed according to Kempken and Kuck (Kempken and Kuck, 1998) and used to PCR amplify conserved motifs identified by comparison of transposase nucleotide sequences from the NCBI database. From the alignment of *Fot1*-like elements, such as *Pot2* (from *Magnaporthe oryzae*: **Z33638**), *Flipper* (from *Botryotinia fuckeliana* [anamorph: *Botrytis cinerea*]: **U74294**) and *Fot4* (from *Fusarium oxysporum*: **AF076632**) chosen for their phylogenetic proximity to *O. ulmi* and *O. novo-ulmi*, we designed degenerate primers Transposase-L and Transposase-R (Table 3). Putative DNA

transposons were amplified by PCR conducted from 10 ng of either *O. ulmi* or *O. novo-ulmi* genomic DNA as follows: thirty cycles of denaturation, annealing and polymerization were conducted at 94°C for 1 min, 55 to 68°C (depending on primers) for 1 min and 72°C for 1 min, respectively. Amplifications were performed with 1U of *Platinum Taq*® (Invitrogen) in a 10µl final volume. Amplified fragments were separated on 1% agarose gels.

II.5.3. Recovery of OPHIO1 by genomic library ligation reaction (GLLR)-mediated PCR

In order to reconstruct the complete sequence of the TE, we first developed a technique for PCR screening of a *O. novo-ulmi* subsp. *novo-ulmi* H327 genomic DNA library. Fungal DNA was partially digested with *Sau3*AI (Roche) and sticky ends were partially filled-in using Klenow enzyme in the presence of dATP and dGTP (Klenow Fill-In Kit, Stratagene). DNAs were size-fractionated by sucrose gradient centrifugation (10% to 40% w/v sucrose gradient) for 16 hours at 20°C and 80,000 x g. Individual fractions were collected and analyzed by agarose gel electrophoresis. DNA fractions from 9 to 23 kb were selected and ligated into the Lambda replacement vector FixII (Stratagene). After synthesis of specific primers GLLR_*OPHI01_L1* and GLLR_*OPHI01_R1* (Table 3) from DNA sequences previously amplified with degenerate primers Transposase-L and Transposase-R, we used 1µl of ligation reaction to generate amplicons by combining either one of these GLLR primers and one of the vector-specific T3 and T7 primers (standard PCR conditions described above with *Platinum Taq* (Invitrogen) in a 10µl final volume). From the sequences thus generated we designed new sets of primers and continued walking on both sides of the genomic DNA clone. The final sequence recovered from *O. novo-ulmi* H327, called *OPHI01-*2170, was 2170 nt long.

II.5.4. Recovery of OPHIO2 and additional copies of OPHIO1 by Inverse-PCR

We adapted an I-PCR protocol (Ochman et al., 1988; modified by Sheehan et al., 2003) in order to amplify unknown DNA adjacent to initial transposon sequences found initially in *O. ulmi* strains with degenerate primers. We designed two primers in opposite directions, I-PCR_Oulmi-L1 and I-PCR_Oulmi-R1 (Table 3) on the first sequence previously amplified in *O. ulmi* with degenerate primers. We selected the hexa-cutter enzyme *Nco*I (New England BioLabs) which cuts once between the I-PCR primers. A sample of 250 ng genomic DNA in a final volume of 10 µl was digested with two different hexa-cutter enzymes (*EcoR*I and *Bgl*II; New England BioLabs) that cut outside the region of interest. Following inactivation of the enzymes, the solution was diluted to 100 µl with sterile water and ligase buffer (final concentration of 1x) and 1 unit of T4 DNA ligase (Invitrogen) were added to give a final reaction volume of 127 µL. The mixture was incubated overnight at room temperature. Thereafter, DNA was precipitated with 95% ethanol, washed with 70% ethanol, air dried and resuspended in 8 µl sterile water. After adding 1 µl of 10X restriction digest buffer, the sample was digested with 1 unit of selected *Nco*I enzyme according to the manufacturer's instructions. A second round of I-PCR with primers I-PCR_Oulmi-L2/I-PCR_Oulmi-R2 allowed us to complete the reconstruction of a transposon called *OPHIO2* in *O. ulmi* strain R21. A similar I-PCR protocol was used with primer pair I-PCR_*OPHIO1_L2* and I-PCR_*OPHIO1_R2* to reconstruct other copies of *OPHIO1* from *O. novo-ulmi* strain H327 (Table 3).

II.5.5. Recovery of a third transposon, *OPHIO3*, from a Suppression Subtractive Hybridization (SSH) library

A 800 nt-long sequence from a SSH cDNA library of *O. novo-ulmi* subsp. *novo-ulmi* strain H327 mycelium grown at low (15 °C) temperature was found to have strong homologies with fungal DNA transposons *Fot1* (E value = $6.4e^{-15}$ with AF443562 from *Fusarium oxysporum* subsp. *solani*) and *Pot2* (E value = $5.7e^{-14}$ with Z33638 from *Magnaporthe oryzae*). Genomic DNA from strain H327 was subjected to I-PCR with primers IPCR_*OPHIO3*_L1/R1 and IPCR_*OPHIO3*_L2/R2 (Table 3) in order to find the complete structure of this putative TE named *OPHIO3*. PCR conditions were the same as described above.

II.5.6. Southern hybridizations

Ophiostoma genomic DNA (2 μ g) was digested overnight with restriction enzyme *Eco*RI (New England BioLabs) which does not cut any of *OPHIO* transposon sequences. Digested DNA was separated in a 2% agarose gel and blotted by capillary transfer onto a Hybond N+ (GE Healthcare Life Sciences) membrane in 4 X SSC buffer. The membrane was treated with 0.4 N NaOH for 45 sec and with a solution of 0.2 M Tris-2 X SSC for 1-2 min. The membrane was then exposed for 3 min at 300 nm to fix the DNA prior to incubation for at least 1 h in prehybridization solution (5 X Denhardt's reagent, 200- μ g/mL herring sperm DNA, 0,5 % SDS and 6 X SSC –final concentrations). The probe TIR3-P^{32dCTP} was prepared with the Rediprime II Random Prime Labelling System (GE Healthcare Life Sciences) and purified with QIAquick Nucleotide Removal Kit (QIAGEN) according to the manufacturer's instructions. The labeled probe was allowed to

hybridize overnight at 72°C. The next day, membranes were washed at 65 °C, first with 2 X SSC-0.1% SDS (30 min) and then with 1 X SSC-0.1% SDS (30 min), followed by a final 1-h wash with 0.5 X SSC-0.1% SDS. The membrane was exposed to BioMax MS x-ray film (Kodak) for 2 to 5 days at -80 °C. Probes used are described in Figure 3, whereas primers used to amplify probes are listed in Table 3.

II.5.7. Sequence, phylogenetic and mapping analyses

DNA sequence homology searches were done using the fASTA algorithm of the European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI) optimized for fungus sequences and BLASTP for amino acid homologies (Altschul et al., 1990; Pearson and Lipman, 1988). Alignments of DNA and amino acid sequences were performed with Bioedit® software (Hall, 1999). To assemble I-PCR-amplified sequences in contigs, we used Seqmerge software of the Genetics Computer Group (GCG) package (U. Wisconsin) (Devereux et al., 1984). TIRs of each *Ophiostoma* transposon were located with REPuter® software for genome comparisons (Kurtz and Schleiermacher, 1999). To generate the phylogeny, the resulting sequences and blocks were aligned and selected with T-COFFEE using default parameters (Notredame et al., 2000). Block selections were processed with SEAVIEW software (Galtier et al., 1996). Using amino acid sequences, optimal group alignments based on the matrix of pairwise comparisons were built. Rooted consensual phylogenetic trees were generated with PAUP 4.0b10 using neighbor-joining method with default parameters (Swofford, 2002).

One copy of *OPHIO2*, named *OPHIO2-int*, was directly mapped by determining its presence or absence in the genome of a population of 40 F₁ progeny recovered previously by Et-Touil et al. (1999) from a laboratory cross between introgressant AST27 (*OPHIO2-int* present) and *O. novo-ulmi* subsp. *novo-ulmi* strain H327 (*OPHIO2-int* absent). Mendelian segregation of alleles at the *OPHIO2-int* locus was verified by the χ^2 test. Recombination between *OPHIO2-int* and other markers identified by Et-Touil et al. (1999) was studied using MAPMAKER for the Macintosh, version 2.0 software with a LOD score value of 4.0 and a θ of 0.25 (Lander et al., 1987). Linkage groups were first identified by two-point analysis and then markers were ordered by multipoint analysis.

All RIP ratios were calculated with COMPSEQ software from the EMBOSS package for dinucleotide frequencies (Rice et al., 2000) after the determination of dinucleotide preferences. All RIP reference ratios (Table 4) were calculated from coding sequences extracted from an *O*.

novo-ulmi subsp. *novo-ulmi* H327 strain EST database (unpubl. data) and totaled 1,123,455 nucleotides. Cumulative transition score (CTS) plots were obtained as follows: for RIPed and non-RIPed copies of a given transposon, $CpX \rightarrow TpX$ and $XpG \rightarrow XpA$ transitions were scored as +1 and -1, respectively. The resulting curves represented the sum of all scores. An increasing slope denoted a preference for RIP mutations on the matrix strand and, inversely, a decreasing slope denoted a preference for RIP mutations on the complementary strand. As the number of RIP transitions increases so does the slope of the curve and vice versa.

II.6. Results

II.6.1. Discovery and characterization of the first three DNA transposons in the fungal genus Ophiostoma

Degenerate primers (Transposase-L and –R) were designed to target *Fot1*-like elements. A 473-nt long fragment, amplified from strains H327 (*O. novo-ulmi* subsp. *novo-ulmi*) and R21 (*O. ulmi*) was sequenced. These sequences were 79% identical to each other and shared strong homologies with DNA transposons in other filamentous fungi. First, using this 473-nt fragment as starting point we were able to reconstitute a 3.4 kb-long sequence from a genomic library ligation reaction of strain H327 by using GLLR-mediated PCR. REPuter® software listed 42-nt long palindromic repeats that could potentially be the TIRs and, thus defined, a 1865 nt-long TE was identified and named *OPHIO1*. Secondly, the same 473-nt fragment found in strain R21 (*O. ulmi*) allowed the recovery for amplification by I-PCR of a sequence of 3150 nt in length. Once again, 42-nt long TIRs were found and delimited the second DNA transposon, named *OPHIO2*, which was 1865 nt in length. *OPHIO1* and *OPHIO2* had 75% identity with each other. They also showed 52 to 57% identity with *Flipper*, *Pot2* and *Fotci* (GenBank accession nos <u>U74294</u>, <u>733638</u> and <u>AF397011</u>) (Daboussi et al., 1992; Kachroo et al., 1994; Levis et al., 1997).



Figure 17. Analyses of OPHIO sequences

(a). Alignment of terminal inverted repeats (TIRs) and target site duplications between *Ophiostoma* spp. DNA transposons and members of the fungal *Fot1* group: *Magnaporthe oryzae Pot2* (Z33638), *Fusarium oxysporum Fot1* (X70186), and the *Botrytis cinerea Flipper* element (U74294). Alignment of the TIR sequences was done by T_COFFEE (Notredame et al., 2000).

(b). Amino acid alignment of the putative product of *OPHIO1*, *OPHIO2* and *OPHIO3* with the conserved DDE catalytic domains of Tc1/mariner transposases (Doak et al., 1994; Feschotte and Mouchès, 2000). Alignment was done with T_COFFEE using default parameters. Amino acid sequences include members of the fungal *Fot1* group: *Pot2*, *Fot1* and *Flipper* elements. Each transposase domain is indicated by a color code: green = HTH_psq (helix-turn-helix, psq domain), red = CENPB (putative DNA binding domain in centromere protein B) and blue = DDE domain (catalytic domain in which the DDD motif is indicated by '*').

Further analysis of the TIRs of *OPHIO1* and *OPHIO2* emphasized a 17/19-nt short direct repetition, a classical motif found in fungal DNA transposons (Daboussi and Capy, 2003). Comparisons by identity percentages showed that TIRs were the domain with the lowest identity (64.3%): their sequences differed in 15 out of 42 nucleotides, with more variation in the 3'-extremity of the TIR-5'. Comparative analysis of the sequences flanking the TIRs of *OPHIO1* and

OPHIO2 showed that they had a "TA" dinucleotide as the TSD, similar to the TSD found in *Fot1*-like elements (Fig. 17A). One ORF was found for each transposon which exhibited marked homology with known transposases (45%, 44% and 38% identity with *Flipper*, *Pot2* and *Fotci*, respectively). Homology-based searches of the protein and domain databases with the *OPHIO* predicted proteins revealed the presence of three putative functional domains (Fig. 17B): 1) a N-terminal region with strong similarity to the so-called CENPB; 2) HTH-psq DNA binding domains; and 3) a C-terminal region with similarity to the so-called DD(35)E domain thought to be responsible for coordinating metal ions needed for catalysis (Doak et al., 1994). The HTH-psq domains were less conserved than the CENPB and DDE domains after translation in amino acid sequence (Fig. 17B).

A third transposon, named OPHIO3, which differed in several aspects from OPHIO1 and OPHIO2, was amplified from a H327 suppression subtractive hybridization library enriched for genes differentially expressed at low temperature (15°C). It showed a high identity (> 60% with E values > $10e^{-10}$) with two known fungal DNA transposons: *Pot2* and *Fotci*. Inverse PCR was used to recover, from strain H327, a 3165-nt sequence harboring a transposon sequence of 1849 nt named OPHIO3 that displayed 52 to 57% identity with Flipper, Pot2 and Fotci, with imperfect 42-nt long TIRs and "TA" TSD similar to the one observed in OPHIO1 and OPHIO2 (Fig. 17A). PCR amplification based on TIR_OPHIO3 primers (Table 3) revealed polymorphism resulting from a deletion of 36 nt near the 3' end. After comparison with OPHIO1 and OPHIO2, a potential ORF was found in OPHIO3 with 47, 42 and 39% identity with sequences encoding transposase in Flipper, Pot2 and Fotci, respectively. However, the ORF of OPHIO3 was disrupted by many STOP codons (n=18). The use of specific primers OPHIO3-1414-L and -R (Table 3) allowed us to confirm the presence of copies of OPHIO3 in all strains tested from both subspecies of O. novo-ulmi and also from O. ulmi. A total of 13 copies of OPHIO3 were sequenced (6 from O. ulmi strains; 4 and 3 copies from strains of O. novo-ulmi subsp. novo-ulmi and americana, respectively).



Figure 18. Phylogenetic relationships of DNA transposons in the phylum ascomycota based on amino acid transposase sequences, using transposon pogo sequence (S20478) to root the tree. Reference transposons and accession numbers of the sequences are indicated for each species. The consensus neighbour joining (NJ) tree is inferred from the alignment produced by ClustalW (Higgins et al., 1994). Numbers near the individual nodes denote BioNJ bootstrap values based on 1000 repetitions (Nodes with values of <50% are not shown. Scale bar: 0.1 substitutions (corrected) per base pair).

Sequences were similar within each taxon but varied among the three taxa. The phylogenetic relationships among the transposons were well delineated (e.g. the clustering of species) and supported by strong bootstrap values in all cases (Fig. 18). The alignment of transposases suggested that 1) phylogeny based on DNA transposons of various species harmonises with species clusters; and 2) *OPHIO* transposons are very closely related, and are phylogenetically closest to *Flipper* and *Pot2* found in *Botrytis cinerea* and in *Magnaporthe oryzae*, respectively. Although *OPHIO3* showed homologies with *OPHIO1* and *OPHIO2*, it is more divergent as demonstrated by its particular position on a separate phylogenetic branch.

We thus propose that *OPHIO1*, *OPHIO2* and *OPHIO3* (GenBank accession nos. <u>DQ649003</u>, <u>DQ649004</u> and <u>DQ649005</u>) be considered as the first three DNA transposable elements found in *Ophiostoma* genomes, since their structural characteristics match those described for DNA transposons found in other fungi.

II.6.2. Distribution of DNA transposons in Ophiostoma ulmi and O. novo-ulmi

According to Southern hybridization experiments, the three transposons were usually present in various copy numbers in each individual fungal genome, although the copy number remained low (Fig. 19). Two hybridization signals (molecular weight between 2000 nt and 3500 nt) were observed for *OPHIO1* with the same pattern in all strains tested. We assume that, based on the length of the probes used (1869 nt), these two bands represent two copies of *OPHIO1*. In addition, two copies of *OPHIO1* were amplified by I-PCR. Congruence between Southern hybridizations and PCR amplifications (Fig. 20) suggested that *OPHIO1* was specific to *O. novo-ulmi* subsp. *novo-ulmi*. By comparison with *OPHIO1*, hybridization profiles indicated that *OPHIO2* occurred in higher copy number and was distributed differently depending on the *O. ulmi* and *O. novo-ulmi* subsp. *novo-ulmi* strains. However, hybridizations on *O. novo-ulmi* subsp. *novo-ulmi* and *O. ulmi* strain W9 were not correlated with PCR amplification, which was likely due to primer mismatch. Hybridizations in *O. novo-ulmi* subsp. *novo-ulmi* subsp. *novo-ulmi* strain AST27. The latter was also the only *O. novo-ulmi* subsp. *novo-ulmi* strain from which *OPHIO2* was PCR-amplified using TIR-specific primers (Fig. 20). Interestingly, this strain is a natural, sexually fertile introgressant.



Figure 19. Distribution of OPHIO transposons in Ophiostoma sp. as determined by Southern hybridization

Genomic DNA (2 µg) of 20 strains of *O. novo-ulmi* subsp. *novo-ulmi*, *O. ulmi* and *O. novo-ulmi* subsp. *americana* (Table 2) was digested with *Eco*RI prior to separation in a 2% agarose gel and blotting. The blots were hybridized with *OPHIO1* (a), *OPHIO2* (b) and a partial fragment (1414-nt) of *OPHIO3* (c) as probe. Final washings were done in 0.5X SSC – 0.1X SDS at 65 °C. Molecular weights (Kb) are indicated in the right margin. Exposure time at - 80°C was 12h.



Figure 20. Distribution of OPHIO1 and OPHIO2 transposons in Ophiostoma spp. as determined by PCR

Genomic DNA (5 ng) of 20 strains of *O. novo-ulmi* subsp. *novo-ulmi*, *O. ulmi* and *O. novo-ulmi* subsp. *americana* (Table 2) was subjected to PCR amplifications with primers homologous to TIRs of each transposon (Table 3). Amplicons were separated in a 1% agarose gel and stained with ethidium bromide. The TrackIt 1 Kb Plus DNA Ladder (Invitrogen) was used as molecular weight marker

A meiotic analysis of H327 x AST27 F_1 progeny allowed us to map one copy of *OPHIO2* (Fig. 21), which was located 16 cM from locus *Pat1* thought to be involved in pathogenicity (Et-Touil et al., 1999). Southern hybridization patterns for *OPHIO3* exhibited various profiles and low copy numbers but showed that this transposon, contrary to *OPHIO1* and *OPHIO2*, was present in all strains tested including *O. novo-ulmi* subsp. *americana*.

II.6.3. Signature of silencing by repeat-induced point (RIP) mutations and analysis of variations in RIP-based silencing

A striking feature of the copy of OPHIO3 recovered from the O. novo-ulmi subsp. novo-ulmi H327 EST library was the occurrence of numerous STOP codons. To investigate the possibility of RIP inactivation of OPHIO3, we first calculated RIP indices of various Ophiostoma sequences, including OPHIO1, OPHIO2 and OPHIO3, according to the method described by Margolin et al (Margolin et al., 1998). RIP ratios were calculated for a 1414-nt long fragment of OPHIO3 designated OPHIO3_1414. This fragment excludes 3' nucleotide variations observed in OPHIO3 copies amplified from O. ulmi strains. RIP ratios computed in Table 4 suggested that the original copy of OPHIO3 was indeed RIPed in O. novo-ulmi subsp. novo-ulmi and O. ulmi, as it displayed the pattern of inversed ratios typical of RIP-mutated sequences. Analysis of both strands of the *OPHIO3* sequences showed the following nucleotide preferences: CpA (>60%) > CpT (18-30%) > CpC (< 10%) > CpG (5%) either in O. ulmi or O. novo-ulmi subsp. novo-ulmi. The copies of OPHIO3 found in O. ulmi appeared to be more RIPed than those in O. novo-ulmi subsp. novoulmi, as evidenced by their higher number of transitions (156 vs. 99). On the other hand, results suggested that none of the three copies of OPHIO3 analyzed in O. novo-ulmi subsp. americana was RIPed. Since we amplified both RIPed and non-RIPed copies of OPHIO3 in O. ulmi and O. *novo-ulmi* subsp. *novo-ulmi*, we could assign a score of +1 or -1 to CpX \rightarrow TpX and XpG \rightarrow XpA transitions, respectively and compute cumulative transition scores (CTS). Figure 22 shows CTS plots for OPHIO3 and other mobile elements found in ascomycete fungi for which both RIPed and non-RIPed sequences were available. Contrary to O. ulmi, O. novo-ulmi subsp. novo-ulmi showed a variation in mutation intensity in the central zone of the sequence which corresponds to the catalytic DDE domain (Fig. 22a). In the case of Fotl (Fig. 22b), analyses suggested an alternation between the matrix and the complementary strand, as observed for OPHIO3. Mutated copies of Crawler (Fig. 22c), Flipper (Fig. 22d) and MATE5 and MATE4 (Fig. 22f) appeared to be RIPed predominantly on the matrix strand, whereas RIP-based silencing of MAGGY (Fig. 22e) appeared to affect mostly the complementary strand, with the exception of the region near the 5' end. Likewise, Tad clone2 and clone3 also exhibited RIP modifications of the matrix strand, whereas Tad clone4 was RIPed only on the complementary strand (Fig. 22g with a score of -178). Furthermore, the comparison of cumulative score plots for *Tad* clone1 and clone2, extracted from two distinct contigs (7.40 and 7.248) from the Neurospora crassa database, revealed that they had

the same RIP pattern between 1450nt and 3390nt. Differences in RIP intensity were also observed among the three mutated copies of *Tan1* analyzed (Fig. 22h).

Name	TpA/ApT	(CpA/TpG)/(ApC+GpT)
Pat	1.33	0.38 *
Yeti	1.44	0.40 *
Punt	1.32	0.56 *
Afut l	1.57	0.78 *
OPHIO3-1414 O. ulmi	1.51	0.60 *
OPHIO3-1414 O. novo-ulmi subsp. novo-ulmi	1.49	0.71 *
Fot1 RIP	1.12	0.75 *
OPHIO2	0.94	0.96
Mars2	0.88	1.01
OPHIO1	0.94	1.06
Fot1	0.75	1.07
Impala E	0.79	1.16
OPHIO3-1414 O. novo-ulmi subsp. americana	0.97	1.19
NCBI O. novo-ulmi ^a	0.86	1.20
ESTs $H327 database^b$	0.71	1.25
NCBI O. ulmi ^c	0.84	1.24
Dane	0.92	1.35
Pcc2	0.66	1.38
Guest	0.48	1.39

Table 4. RIP indices of various DNA transposons and retroelements in ascomycete fungi

Repeat-Induced Point (RIP) mutation ratios were calculated from dinucleotide frequencies. All frequencies were obtained with COMPSEQ® software from the EMBOSS package.

(*) indicates RIP-mutated transposons or retrotransposons.

(^a) A set of 24 *O. novo-ulmi* subsp. *novo-ulmi* single gene sequences covering 20759 bp, obtained from the NCBI database.

(^b)A set of more than two thousand ESTs, covering 1123455 bp, obtained from *Ophiostoma novo-ulmi* subsp. *novo-ulmi* strain H327.

(°) A set of 24 O. ulmi single gene sequences covering 22341 bp, obtained from the NCBI database.

Other sources of DNA transposons used are as follows: *Pat* from *Podospora anserina*: Genbank accession no <u>AJ270953</u>; *Yeti* from *Podospora anserina*: Genbank accession no <u>AJ272171</u>; *Punt* from *Neurospora crassa*: Genbank accession no <u>AF181821</u>; *Afut1* from *Aspergillus fumigatus*: Genbank accession no <u>L76086</u>; *Fot1_*RIP from *Fusarium oxysporum*: Genbank accession no <u>AF434909</u>; *Mars2* from *Ascobolus immersus*: <u>X99083</u>; *Fot1* from *Fusarium oxysporum*: Genbank accession no <u>X64799</u>; *impala E* from *Fusarium oxysporum f.* subsp. *Melonis*: Genbank accession no <u>AF363407</u>; *Dane* from *Aspergillus nidulans*: Genbank accession no <u>AF295689</u>; *Pcc2* from *Phanerochaete chrysosporium*: Genbank accession no <u>X98835</u> and *Guest* from *N. crassa*: Genbank accession no <u>AY197334</u>

II.7. Discussion
II.7.1. OPHIO1, OPHIO2 and OPHIO3: three transposons from the Tc1/mariner superfamily found in the DED fungi

The first objective of this study was to verify the presence of DNA transposons in the genome of the DED fungi *O. ulmi* and *O. novo-ulmi*. We used a combination of PCR-based approaches to recover three TE, designated *OPHIO1*, *OPHIO2* and *OPHIO3*, which are the first mobile elements to be reported in Ophiostomatales. Several characteristics of *OPHIO1*, *OPHIO2* and *OPHIO3*, such as "TA"-TSD, direct repeat within TIR sequences, as well as homologies with known transposases harboring HTH-psq and DDE motifs with the (DDD) signature, placed these transposable elements in the Tc1/mariner superfamily of type II elements and confirmed that homology was strongest with members of the *Fot1/pogo* group. The presence of these TE in Ophiostomatales confirms the wide occurrence of the Tc1/mariner superfamily of DNA transposons in filamentous fungi.

Transposon *OPHIO1* exhibited characteristics of a potentially active TE: perfect TIRs, no STOP codon in the ORF coding for the transposase, and identity among all sequences amplified in several strains. Southern analyses suggested the presence of two copies of *OPHIO1* in all strains tested. Furthermore, it appeared to be specific to *O. novo-ulmi* subsp. *novo-ulmi*. PCR analysis on the flanking regions of the two copies of *OPHIO1* showed that they were located at the same position in all strains (data not shown). This observation argues against a recent transposition of *OPHIO1* into the *O. novo-ulmi* subsp. *novo-ulmi* genome. Therefore, it seems that *OPHIO1* is inactive, even though we did not detect evidence for the signature of silencing mechanisms (such as RIP) in the copies of *OPHIO1* we analyzed.

Transposon *OPHIO2* had 75% identity with *OPHIO1* and also had characteristic features of an active TE. However, the distribution of *OPHIO2* in the genome of DED fungi differed strikingly from that of *OPHIO1*. Southern analyses showed that *OPHIO2* occurred in both *O. ulmi* and *O. novo-ulmi* subsp. *novo-ulmi*, and that variations in copy number and location were more pronounced in the former. Indeed, if introgressant strain AST27 were excluded from the analysis, the true *O. novo-ulmi* subsp. *novo-ulmi* strains analyzed had similar profiles. These results are consistent with activity of *OPHIO2* in *O. ulmi* and further suggest that this TE was acquired first by *O. ulmi*, which is believed to have evolved prior to *O. novo-ulmi*, based on the fact that it was detected a few decades before *O. novo-ulmi* in Europe and North America (Brasier et al., 2004).

The third transposon, *OPHIO3*, exhibited several structural differences with *OPHIO1* and *OPHIO2*, including its slightly shorter length and the occurrence of many STOP codons along its sequence. Furthermore, Southern analyses showed that *OPHIO3* occurred in *O. ulmi* as well as in both subspecies of *O. novo-ulmi*. As was observed for *OPHIO2*, hybridization profiles were more variable in *O. ulmi*. Taken together, our results suggest that *OPHIO3* is more ancient than *OPHIO1* and *OPHIO2* which is also supported by the phylogenetic analysis (Fig. 18).

II.7.2. The potential role of interspecific hybrids in TE transmission

The *Ophiostoma* strains that we analyzed included a natural introgressant, strain AST27, whose nuclear genome contains an estimated 10-15% of O. ulmi sequences within an O. novo-ulmi subsp. novo-ulmi background (Et-Touil et al., 1999). Strain AST27 contained the three DNA transposons described in this study. The fact that OPHIO2 is present in AST27 is interesting because the OPHIO2_TIR-primer did not allow us to amplify this transposon in other O. novoulmi subsp. novo-ulmi strains. The reconstituted AST27 OPHIO2 sequence had over 99,1% identity with OPHIO2 found in O. ulmi. The few nucleotide variations observed were only transition mutations and ORFfinder denoted a complete ORF without STOP codons. This copy, named OPHIO2-int., was mapped in a linkage group known to have been introgressed from O. *ulmi* into the AST27 genome and was found to be linked to *Pat1*, thought to be a pathogenicity locus (Et-Touil et al., 1999). Although mapping data suggested that *OPHIO2-int*. was physically too far from Pat1 to have a direct impact on its expression, the fact that a copy of a DNA transposon could be introgressed with a portion of one genome into another genome is noteworthy and supports the view that interspecific hybrids can act as genetic bridges between species. These interspecific hybrids are also thought to be responsible for the acquisition, by O. novo-ulmi, of useful alleles from O. ulmi (Erseck et al., 1995; Paoletti et al., 2006).





The linkage map is based on the analysis of a set of 40 F_1 single-ascospore progeny previously recovered by Et-Touil et al. (1999) from the cross H327 × AST27. Map distances were calculated from multipoint analysis (LOD score of 4.0 and Theta of 0.3) with the software MAPMAKER version 2.0 for Macintosh (Lander et al., 1987). Transposon *OPHIO2*-int was found to be linked to random amplified polymorphic DNA (RAPD) and coding (*Pat1*) markers introgressed from *O. ulmi* into AST27.

Figure 21. Meiotic mapping of OPHIO2 DNA transposon of Ophiostoma novo-ulmi (OPHIO2-

The recovery of transposon *OPHIO3* allowed us to analyze RIP mutations, one of the principal mechanisms of transposon silencing in filamentous fungi (Galagan and Selker, 2004; Selker, 2002; Selker et al., 1987). In *N. crassa*, RIP has been shown to act only during the sexual cycle between fertilization and premeiotic DNA synthesis (Selker, 1997). According to sequencing data, *OPHIO3* copies found in *O. ulmi* and *O. novo-ulmi* subsp. *novo-ulmi* were RIPed, whereas the three copies sequenced from *O. novo-ulmi* subsp. *americana* were not. However, further conclusions on the occurrence of RIP in *O. novo-ulmi* subsp. *americana* must await the analysis of additional copies of *OPHIO3* from this taxon. The fact that *OPHIO3* was found in an expression library is interesting since it was highly altered, with numerous STOP codons. It is known that mechanisms operating in eukaryotes, such as nonsense-mediated mRNA decay (NMD) reported recently in *A. nidulans* (Morozov et al., 2006), are able to eliminate altered RNAs. Perhaps this altered *OPHIO3* transcript was over-expressed as a specific response to low temperature stress, as part of an adaptive response of the genome (Ikeda et al., 2001), without being destroyed beforehand by NMD or a related mechanism.

As mentioned earlier, we did not find evidence for the occurrence of RIP-mutated copies of *OPHIO1* or *OPHIO2* in *O. ulmi* or *O. novo-ulmi* subsp. *novo-ulmi*, although the small number of specimens analyzed prevents us from drawing a definitive conclusion as to their presence or not in natural populations of the Dutch elm disease fungi. However, none of the copies of *OPHIO1* and *OPHIO2* found in introgressant strain AST27 were RIPed. Furthermore, we amplified and sequenced five copies of *OPHIO1* in F1 and backcross progeny strains from the AST27 x H327 cross (Et-Touil et al. 1999). The PCR primers that we used (TIR_*OPHIO1*) targeted the TIR regions, which exhibit a low rate of RIP compared to other domains of TE (Clutterbuck, 2004), and were thus expected to allow amplification of both RIPed and non-RIPed forms of *OPHIO1*. All copies of *OPHIO1* had the same nucleotide sequence and, again, none was RIPed (data not shown). By definition, introgression involves several rounds of sexual crosses, which would thus provide opportunities for RIP-driven methylation.



Figure 22. Cumulative transition scores of RIPed forms of DNA transposons and retroelements found in ascomycete fungi

Figure 22. (suite) CpX \rightarrow TpX transitions were assigned a value of +1, whereas XpG \rightarrow XpA transitions were assigned a value of -1. **a**) *OPHIO3* in *Ophiostoma ulmi* and *O. novo-ulmi* sp. *novo-ulmi* **b**) *Fot1* in *Fusarium oxysporum* (GenBank accession nos

X64799 and AF434909 for the active and RIPed forms, respectively); c) *Crawler* in *Aspergillus oryzae* (GenBank accession nos AB256519 and AB256520 for the active and RIPed forms, respectively); d) *Flipper* in *Botrytis cinerea* (GenBank accession nos U74294 and AAID01000200 for the active and RIPed forms, respectively); e) *MAGGY* in *Magnaporthe oryzae* (GenBank accession nos L35053 and AB024423 for the active and RIPed forms, respectively); f) *MATE* in *Aspergillus nidulans* (GenBank accession nos BK001592 [active]; BK001595 and BK001596 [RIPed]); g) active *Tad* from *Neurospora crassa* strain Adiopodoume (GenBank accession nos L25663) versus RIP-mutated copies found inside the sequenced *N. crassa* genome (GenBank accession nos AABX01000192, AABX01000499, AABX01000464 and AABX01000183); h) active *Tanl* in

Aspergillus niger (GenBank accession no <u>U58946</u>) versus degenerate copies found in *A. oryzae* (GenBank accession nos <u>U58946</u> and <u>DO327732-34</u>). Cumulative scores for mutated copies of *OPHIO3*, *Crawler*, *Flipper* and *Maggy* suggested that one DNA strand was preferentially RIPed. Strand preference was also observed for the two mutated copies of *MATE* and three of the four RIPed copies of *Tad* analyzed; furthermore, the analysis suggested that different copies of the same transposon could exhibit different strand preferences. In the case of RIPed copies of *Tan*, variations in RIP intensity occurred among copies.

One of the possible explanations for our results is that RIP does not occur in the DED fungi and silenced copies of *OPHIO3* found in *O. ulmi* and *O. novo-ulmi* subsp. *novo-ulmi* were RIPed in an ancestral species.

The cumulative transition score (CTS) method presented here adds a visual component to the analysis of RIPed sequences and, thus, complements the dinucleotide preference ratios developed by Margolin et al. (1998). CTS allowed us to: i) plot transition density curves for matrix and complementary strands; and ii) compare directly RIP patterns among copies of various transposons found in ascomycete fungi. The CTS visualization does not take into account dinucleotide preferences, contrary to Margolin's ratios. Therefore, we can analyze any DNA transposon, as long as RIPed and non-RIPed copies are available. Thus, comparisons among different mutated copies for transposons MATE, Tad and Tan highlighted the occurrence of two different phenomena: 1) differences in strand preference (e.g. between MATE4 and MATE5, Fig.22f); and 2) variations in RIP intensity at several points along the sequences (e.g. among RIPed copies of *Tan*, Fig. 22h). Moreover, the fact that two different copies (clone1 and clone2) of Tad (Fig. 22g) at different genome locations had the same curve profile is intriguing because it suggests that either one RIPed copy was partially duplicated, or that these two copies underwent the same RIP mutations. In the latter case, RIP would not occur in random fashion or, alternatively, we have visualized a portion completely saturated with RIP mutations, and thus resulting in similar CTS curves for the two TE. Furthermore, the CTS method provided evidence that RIP mutations occurred either on matrix or complementary strand. This suggests that DNA

was single-stranded, as proposed earlier by Clutterbuck for *MATE* transposons in *Aspergillus nidulans* (Clutterbuck, 2004).

To our knowledge, this is the first study to identify and characterize DNA transposons in ascomycete fungi from the order Ophiostomatales. We demonstrated by Southern analysis that the three *OPHIO* transposons, which belong to the *pogo/Fot1* group within the *Tc1/mariner* superfamily, had different distribution patterns in *O. ulmi* and *O. novo-ulmi* and that interspecific hybrids might serve as genetic bridges for TE. Sequence analysis revealed that at least one of these transposons, *OPHIO3*, bore the signature of RIP silencing. Finally, CTS plots appear to be a useful tool for visualizing and comparing patterns of RIP mutations in ascomycete fungi.

		10	20	30	40	50	60	70	80	90
OPHIO1 OPHIO3_EAN OPHIO3_O.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	GCGAAGCG TCTTAACG TCTCAACG TCTCAGCG TCTAAGCG	AGTCAATGG TATTAATAGG TATTAATAGG TGTCAATGGG TATTAATGGG		II ATGCTATTAG TAAAGATTGA TAAAGATTGA TAGAGATTGG CAAAGATTGA	. GACTTGGTTT AGCATTTTTC AGTATTTTTC GGCATTTTTC AGCATTTTTC	. CCGGCGTCTCC GAGCTTTTAT GAGCTTTTAT GGGCTTCTAT		II.AGATTAAGCAT AGATTACTAAC AATTACTAAC GATTACTAAC GATTACTAAC	. ATTTTGCCGCA ATCCTACTACA ATCCTACTACA ATCCTGCCACA ATCCCACCACA	. . AA AA AA AA
OPHIO1 OPHIO3_EAN OPHIO3_O.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	AACAGATG CACCGGTA CACCGGTA CACCGGTG CACCGGTA	100 . GAATATGGAT TAGTATAAAA TAATATAAAA TAGTATGGAT TAATATGGA	110 IGAGACAGGCI GAGACAAGCCI GAGACAGGCCI GAGACAGGCCCI GAGACAGGCCCI	120 . TCTCTTCGAG TTACAGAAAG TTATAAAAGG TCACAGAAGAG TTGCAGAAGG	130 . TCAAGGGGAT TCTAAGGGCC TCTAGGGGCC TCTAGGGGCC	140 	150 	160 	170 . ACAAAAATTCC .CGTCGAAAGGT .CGTCGAGAAGGT .CGTCGAGAAGGT	180 . . 3G FA FA IG
OPHIO1 OPHIO3_EAN OPHIO3_0.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	AAAAAACA GTAAAGTC GTAAAGTT GTAAAGCC GTAAAGTC	190 . AGTTGGCTCZ AAGCGGCTCC AAGCGGCTCC AAGCGGCTCC	200 ACGCGCCTGGA GTGTATCTAAG GCGTGTTTAAG GCGCGCCTGGG GCGTGTCTAGG	210 . TAACAACATT TGACAGTATT CCAACAACATT TGACAGCATT	220 . TGAATGTGTG TAAATACATC TAAATACATC TGAATGCGTC TGAATGCATC	230 . TCAGCCGCCGC TCTGCCGCCCG TCTGCCGCCCG TCTGCCGCCCG	240 GGTGCGTCTCT GGCTTTATACT GGCTTTATACT GGCTTTGTACT	250 .II CCCCTCCTTTG CCCCTCCTTTA TTCCCCTCCTTTA CCCCCCCTTTA	260 ATTATTTTCA ATTATTTTTG ATTATTTTTG ATTATTTTCG ATTATTTTCG	270 .I. AG DT DT DT CT
OPHIO1 OPHIO3_EAN OPHIO3_0.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	GGAAAAAA ACGAATAC ACGAATAT GCGAATAC GCGAATAC	280 . AGTCCAACAA AGTCTAGCAA AGTCTAGCAA AGTCCAGCAA AGTCCAGCAA	290 	300 	310 . . GAGTCTTTTC AAAAGAATAT AGAAGAATAT GAGAGAATAT GGAAGAATAT	320 . AACTCTTGGC CGTCCTAAA CGTCCCTGGA CGTCCCTGGA	330 CAATTCACTGC AAGTTTACCAA AAGTTTACTAG AAGTTCACCGA	340 . TACTAATAAT CACTCAGACT CACTCAGACT TACTCAGACT	350 . GGGTGGACTGA AGCTAGACAAA AGCTAGACAAA GGCTGGACAAA GGCTGGACAAA	360 • AC AC AC AC AC
OPHIO1 OPHIO3_EAN OPHIO3_O.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	TACGAAAC AACGATAT AACGATAT AACGATAC AACGATAT	370 . GGGGTTGAAA AGCTAAAGAA GGCTAAGGAA AGCTAAAGAA	380 ATGGCTCGAAG STAGCTCGATA STGGCTCGATG STGGCTCGACA STGGCTCGACA	390 . GATGTGTTTAT MAAGTATTTCT MAAGTGTTTCT MAAGTATTTCT	400 . TCCATGTTCT CCTAGAAACG CCCAGGAACG CCCAGGAACG	410 	420 	430 . AGCTAGATTG ATAGCGCCTA ATGGCGCCTG ATGGCGCCTG	440 CTAGTTATTAA CTAATCCTCGA CTAATCCTCGA CTAATCCTCGA	450 . AT AT AT AT AT
OPHIO1 OPHIO3_EAN OPHIO3_O.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	GGCCACGG AGCCATAG GGCTATAG GGCCACGG GGCCATGG	460 . AAGCCATGAG AAGTTATNC2 AAGTTATATA AAGTTATGCC AAGTCATAC2	470 SACCGATGGCT AACCGATGCTT AACCGATGCTT SACCGATGCTT AACCGATGCTT	480 . TTATGAAGTT TCATGCTTAC TTATGCTTAT TCATGCTTAC TCATGCTTAC	490 . ATGCTTCGAG ATACATATAA ATGTATGTAA ATGCATAGAG ATGCATGCAA	500 . AATAACATAT AATAAAGTCT AACAAAGTCT AATAAAGTCT AATAAAGTCT	510 CGCTTCTTAT CGCTTCTTAT CGCTTCTTAT	520 	530 . CACTCATCACZ CACAGCTCTCZ TACAGCTTTCZ CACAGCTCTCZ CACAGCTCTCZ	540 . AC AC AC AC AC
OPHIO1 OPHIO3_EAN OPHIO3_O.ulmi OPHIO3_deRIP_in sil. OPHIO3 in vivo	GTCCTTCA GTCCTTCA GTCCTTCA GTCCTTCA GTCCTTCA	550 	560 ITTAACTATCI IATCTCTTACI IATCTCTTGCI IATCTCTTACI	570 . TCTCGCCTTT TCTCTCCTCT TCTCTCCTCTTT TCTCTCCTCTTCT	580 . GAAAGCATAT CAAGGCCCGT CAAGGCCCGT CAAGGCCCGT	590 . TACAAGAAAG TACCGNGTCT TACCGTGTCT TACCGTGTCT	600 . GAGATTGAGAA GTTTTTCTTC GTTTTTCTTC GTTTTTCTTC	610 . AATCGGCAGT CCTGCCCTCT CCTGCCCTCT CCTGCCCTCT	620 GACGAAGAAGA AAAGAAGAGAA AAAGAAGAGAA GAAGAAG	630 . CT AT AT AT AT

Figure 23. Alignement des séquences d'OPHIO1 et OPHIO3 à partir d'amplifications in vivo sur O. ulmi, O. novo-ulmi subsp. novo-ulmi et d'une séquence prédéterminées in silico (OPHIO3_de_in silico)

69

L'ajout d'une dernière information, non publiée dans cet article, nous a paru pertinent. A partir de ces séquences, nous avons voulu tester l'hypothèse suivante : il existe encore au sein d'un des génomes de nos espèces, des formes non-RIPpées du transposon OPHIO3. Cette approche a consisté à comparer différentes séquences RIPpées d'OPHIO3 et à prédire, in silico, une potentielle séquence non RIPpée, via la réversion des mutations. (Cambareri et al., 1998). La figure 23 présentée page 74, montre l'alignement de séquences partielles de 644nt d'OPHIO3 RIPpée amplifié chez O. novo-ulmi sp. novo-ulmi et O. ulmi avec la séquence partielle d'OPHIO1 amplifié chez O. novo-ulmi sp. novoulmi. Cette séquence, bien qu'ayant seulement 55% d'identité, a été choisie comme référent de transposon potentiellement actif. Nous avons nommé OPHIO3_déRIP_in silico la copie fictive exempte d'une partie des mutations RIP (qui consiste à convertir les nucléotides T : A en nucléotides G :C). Dans notre cas, nous avons décidé de ne pas tenir compte des préférences dinucléotidiques CpA et TpG pour augmenter les chances d'obtenir une séquence qui puisse être amplifiable subséquemment. Sur 375nt non homologue entre les 3 séquences, 90 d'entre eux présentent des transitions C : G vers A : T, et donc potentiellement dues au mécanisme de RIP (plus exactement, 58 mutations sont des transitions de type G vers A (64,5%) et 32 sont des transitions de type C vers T). La conversion des nucléotides A :T en C :G a permis d'éliminer 7 des 11 codons stop présent au sein des séquences OPHIO3. Nous avons pour finir construit des amorces spécifiques situé à chacune des extrémités de la séquences d'OPHIO3_déRIP_in silico. Le résultat de l'amplification et du séquençage de différents fragments a démontré que l'un d'entre eux, nommé OPHIO3_deRIP_in vivo, possédait 84% d'identité avec OPHIO3_deRIP_in silico et 85% d'identité avec OPHIO3 O nu sp. nu. D'une manière générale, 90% des différences entre ces séquences sont dues à des transitions de type RIP ce qui confirme le fait que notre hypothèse a été vérifié et que cette séquence semble être une copie d'OPHIO3 n'ayant pas subi de mutations RIP et pouvant donc être encore active.

II.9. Corrélations et applications préliminaires avec l'étude menée dans ce chapitre sur le RIP chez *Ophiostoma* sp.

Les amorces décrites au *paragraphe II.5.4.* (p. 73) de ce chapitre (*OPHIO3_1414-L* et *OPHIO3_1414-R*) nous ont permis, en faisant varier la température d'hybridation de ces dernières au cours des amplifications PCR, d'amplifier 3 types de séquences spécifiques de chacune des 3 espèces et sous-espèces étudiées à ce moment. La séquence d'*OPHIO3* provenant des souches de *O. novo-ulmi* sp. *americana*, ne présentant pas de mutations RIP, nous a permis de développer par la suite notre technique dite de *Cumulative Transition Score* (CTS). Cette séquence a démontré 99,5% d'identité avec la séquence *OPHIO3_deRIP_in vivo* amplifiée chez la sous espèce *americana*. Ceci qui nous permet de conclure que ces deux séquences semblent être issues d'une même copie ou de copies d'*OPHIO3* non RIPpées, validant le fait qu'aucun des copies d'*OPHIO3* ne semble RIPpée chez cette sous-espèce.



Figure 24. Analyse par CTS de différentes copies d'OPHIO3 isolés chez diverses souches d'Ophiostoma ulmi.

D'une manière générale, 6 profils distincts se dégagent de cette analyse par *CTS*. La comparaison de ces différents profils nous indique que : (1) le profil #4 est celui majoritairement retrouvé au sein des diverses souches testés (pour un total de 40 souches, près de la moitié de ces dernières démontrent le profil #4. (2) Il semble y avoir un gradient d'intensité au niveau du mécanisme de RIP chez *O. ulmi*. En effet l'allure générale de ces courbes indique une augmentation croissante de l'intensité des mutations de type CpA vers TpA et inversement du profil #1 vers le profil #6. Cette mise en évidence nous a fait nous poser une question supplémentaire : peut-on corréler une localisation géographique ou un gradient géographique avec ce gradient d'intensité ?



Figure 25. Localisation géographique des souches d'Ophiostoma ulmi *précédemment analysées par CTS.*

Cette carte indique l'emplacement des souches d'*O. ulmi* isolées pour l'analyse précédente via *CTS*. Les couleurs utilisées reprennent les couleurs précédemment établies : on remarque qu'il n'y a pas de corrélation géographique possible avec le gradient d'intensité de mutations RIP. Toutefois, la Hollande semble être un des pays où les souches semblent indiquer un fort gradient d'intensité. Ceci correspond bien à la première mise en évidence de la maladie hollandaise de l'orme, en Hollande. Toutefois, le profil #6, provenant de souches isolées en Irlande, démontre une intensité supérieure à celles isolés en Hollande.

(nombre de souches appartenant aux profils : #1=1, #2=2, #3=2, #4=12, #5=2, #6=4).

Chapitre III : Impacts des transposons à ADN sur le génome hôte : étude de la mobilité des transposons *OPHIO1* et *OPHIO2*

III.1. Avant propos

L'étude présentée dans ce chapitre porte sur la mobilité des éléments *OPHIO1* et *OPHIO2*. Une partie des ces résultats a été présenté au « *Ceratocystis and Ophiostoma* workshop », du 16 au 18 août 2006 à Moreton Bay Research Station, à Brisbane, en Australie ainsi qu'au « 8th International Mycological Congres » du 20 au 25 août 2006 à Cairns, en Australie. Ces résultats ont été acceptés pour publication dans la revue *Fungal Genetics and Biology* après corrections et modifications.

<u>Réalisations et remerciements</u>: K. Plourde = réalisation des hybridations Southern pour le transposon *OPHIO1* et clonages des produits de la seconde série de RACE-5' effectuée pour vérification, V. Jacobi = production des RACE-5' et corrections/rédaction. L. Bernier = corrections/rédaction

III.2. Résumé

La mobilité des transposons est une caractéristique essentielle de ces éléments génomiques. Elle fait partie des mécanismes qui confèrent aux génomes leur plasticité, que ce soit par la sur ou sous expression des gènes ou par les recombinaisons ectopiques qu'ils engendrent. Les données recueillies lors de cette étude fournissent l'évidence de la mobilité des transposons *OPHIO1* et *OPHIO2*, récemment mis en évidence chez *Ophiostoma ulmi* et *O. novo-ulmi*, agents de la maladie hollandaise de l'orme. Cette mobilité semble, de surcroît, induite par des stress abiotiques. Premièrement, l'analyse des régions TIR-5' et UTR-5' de ces éléments OPHIO indique la présence de deux motifs nucléides reconnus comme étant des sites de fixation de éléments *trans* ainsi qu'un promoteur de réponse aux chocs thermiques (*Hsp promoter*). Après la vérification de l'expression de chacune des deux transposases, nous avons testé deux procédures expérimentales : (1) une copie exogène d'*OPHIO1* a été intégrée par mutagénèse dans la souche *O. novo-ulmi* subsp. *americana* W2, exempte de copies endogènes d'*OPHIO1*. Après exposition

des mutants (W2 :*OPHIO1*) à des chocs abiotiques tels qu'un choc thermique à 55°C durant 1 heure, l'analyse subséquente des survivants a révélé que certains d'entre eux montraient des signes de transposition incomplète des copies d'*OPHIO1* (excision sans réinsertion). (2) La souche AST27 (*O. novo-ulmi* subsp. *novo-ulmi*) possédant une unique copie : *OPHIO2-int.* introgressée naturellement a été soumise à une même série de stress abiotiques. Bien qu'aucun promoteur n'ait pu être identifié au sein de cette séquence, une exposition aux U.V. pendant 2 minutes ainsi que la culture au froid (4°C) de cette souche a permis de mettre en évidence, au sein des survivants, des excisions parfaites d'*OPHIO2-int.* Dans ce cas, le choc thermique n'a eu aucun effet sur la mobilité d'*OPHIO2-int.* L'ensemble de ces résultats nous a permis d'émettre une hypothèse quant à la potentielle invasion interspécifique des transposons OPHIO.

III.3. Abstract

The mobility of transposable elements (TEs) can contribute to genome plasticity, under- or overexpression of genes and ectopic recombination. The data collected in this study provide evidence of stress induced mobility of OPHIO1 and OPHIO2 transposons, recently detected in Ophiostoma ulmi and O. novo-ulmi, the causal agents of Dutch elm disease (DED). The analyses of OPHIO UTRs and TIRs indicated the presence of two potential binding site motifs and a heat shock protein (hsp) promoter which could be involved in the mobility of OPHIO1 following a heat shock stress. The exact position of the hsp promoter was determined by 5' RACE PCR. After confirmation of the expression by RT-PCR of both OPHIO1 and OPHIO2 transposases in the absence of stress factors, we tested two experimental procedures to induce mobility of OPHIO TEs: (1) an exogenous (cloned) copy of OPHIO1 was introduced into the O. novo-ulmi subsp. americana strain W2 (OPHIO1 free strain) to give mutant strain W2:OPHIO1. After exposure of W2:OPHIO1 to a 55°C heat shock treatment, some of the survivors showed signs of incomplete transposition (excision without reinsertion) of OPHIO1. (2) The O. novo-ulmi subsp. novo-ulmi strain AST27, introgressed from O. ulmi and carrying a distinct endogenous copy of OPHIO2 (OPHIO2-int), was subjected to a series of abiotic stress treatments. Although a promoter sequence could not be identified, both exposures to UV light and to a 4°C cold treatment caused perfect excision of OPHIO2-int. In contrast to OPHIO1, heat shock stress did not induce OPHIO2-int. mobility. Taken together, these results allow us to hypothesize a potential interspecific invasion of OPHIO transposons due to their mobility in Ophiostoma spp.

III.4. Introduction

Mobility is a key feature of transposable elements (TEs) and its study is a way to predict the impact of non-Mendelian elements on genomic architecture and long-term evolutionary consequences on host genomes (Bowen and Jordan, 2002; Kidwell and Lisch, 2002). TEs are repetitive mobile DNA sequences and are broadly divided into two major groups depending on the type of mobility. One of these two groups includes DNA transposons, also named class II TEs, which are generally excised from one genomic site and integrated into another by a "cut-andpaste" mechanism (Craig, 2002). Their random insertion can occur at a number of genomic sites because of the short length of target sequences, and consequently, DNA transposons are now considered as a major source of evolutionary changes, generating gene and chromosome evolution through diverse mechanisms, including chromosome breaks, translocation, gene conversion, ectopic recombination, gene duplication and mobility of other DNA fragments (Brookfield, 2002; Daboussi and Capy, 2003). Active DNA transposons code for a transposase enzyme located between the terminal inverted repeats (TIRs), which are palindromic sequences located at each end of DNA transposons. The transposase binds at or near the TIRs to two specific boxes, located in the middle and the 3'end of each TIR, and which seem to play a key role in transposase binding (Feschotte et al., 2005). The transposase then performs a DNA breakage reaction to excise the transposon out of the genomic DNA. Successive chemical reactions occur, such as hydrolysis of phosphodiester bonds and DNA repair of small flanking gaps that lead to short duplication of sequences at the target sites, called target site duplications (TSD) (Feschotte et al., 2005; Mizuuchi and Baker, 2002). The TSD associated with TEs from the *Tc1/mariner* superfamily is usually a TANNNTA footprint, created after duplication of the target site TA generated during the insertion added to a few terminal nucleotides of the TE TIRs (Yang et al., 2006). The reinsertion of a DNA transposon into its host DNA is possible after recognition of a new specific target site and DNA cleavage. A unique TE, PiggyBac, is capable of perfect excision, i.e. excision and reinsertion without target site duplication (Elick et al., 1996).

Complex questions surround the study of TE mobility, one of them being concerned with stimuli involved in excision. Abiotic and biotic stresses have often been related to mobility: in 1986, one of the first experiments linked radiation (gamma or X-rays) to retrotransposon activity (*P* element in *Drosophila melanogaster*) (Margulies et al., 1986; Margulies et al., 1989). Six years later, *Fot1* mobility in *Fusarium oxysporum* (Daboussi et al., 1992) was established by

visualization of a reversible insertion of *Fot1* in the nitrate reductase gene. A similar experimentation was used to demonstrate the mobility of *Restless* transposon induced by an abiotic stress (Kempken and Kuck, 1996). In 1996, Jouan-Dufournel and collaborators demonstrated that mobility of the *copia* and *mdg1 retro*transposon could be induced by a viral infection (Jouan-Dufournel et al., 1996). Concerning thermal stresses, TEs were found to have many interactions with heat shock (*hsp*) genes or their promoters: (1) the *DIRS-1* transposon of *Dictyostelium* sp. carries a *hsp* promoter in its TIRs (Zuker et al., 1984); (2) a heat shock activates the mobility of *MAGGY* retrotransposon in *Magnaporthe oryzae* (Ikeda et al., 2001); (3) a heat shock was demonstrated to increase production of the *erm* transposase and concomitant loss of a portion of this mobile element in *Lactobacillus crispatus* (Stroman et al., 2003); and (4) *P* elements were shown to insert preferentially into *hsp* promoters in *Drosophila* (Walser et al., 2006).

Recently, two DNA transposons named OPHIO1 (DQ649003) and OPHIO2 (DQ649004) were identified in the ascomycete fungi Ophiostoma ulmi and O. novo-ulmi, causal agents of Dutch elm disease (DED). Transposons OPHIO1 and OPHIO2 display characteristics of mobile elements such as perfect identities between TIRs, an unique open reading frame coding for a complete transposase and variations in Southern profiles of reference strains, which suggest an ancestral mobility for OPHIO2 (Bouvet et al., 2007). Validation of their potential autonomous mobility is fundamental for understanding the possible impact of these transposons on their host genomes. In this paper, we report on the detection of two transcription factor binding motifs within TIRs of both OPHIO TEs and the presence, within OPHIO1 TIRs only, of a sequence motif with resemblance to a consensual *hsp* sequence originally found in *D. melanogaster*. These motifs may be implied in the response to heat shock stress. The discovery of the promoter could be an indication that a heat shock stress may induce or increase the OPHIO1 transcription, contrary to OPHIO2. The expression of OPHIO1 and OPHIO2 transposases, in the absence of stress, in mycelium of reference strains was demonstrated by reverse transcription polymerase chain reaction (RT-PCR). To demonstrate that *OPHIO1* mobility could be linked to heat shock stress, we analyzed, by Southern hybridizations, the mobility of an exogenous (introduced) copy of OPHIO1 in a mutant strain of O. novo-ulmi subsp. americana W2. Previously, we had shown that the wild-type W2 strain does not carry a copy of *OPHIO1* (Bouvet et al., 2007). We provide evidence that a 55°C heat shock treatment caused incomplete transposition (excision with out reinsertion) of the introduced *OPHIO1* copy. To induce *OPHIO2* mobility, we applied the same three stresses to the *O. novo-ulmi* subsp. *novo-ulmi* strain AST27, an interspecific hybrid, harboring a unique copy of *OPHIO2* (named *OPHIO2-int*. Bouvet et al., 2007) which has been naturally introgressed from *O. ulmi* to AST27 after several rounds of sexual crosses. Our results suggest that fungus growth at 4°C and exposure to UV light induce the mobility of *OPHIO2-int*. in AST27. Interestingly, mobility of *OPHIO2-int* seems to be correlated with upregulation of the transcript coding for its transposase as determined by semi-quantitative RT-PCR. Thus, environmental stresses induce *OPHIO1* and *OPHIO2* mobility which, in turn, may be a prerequisite for impacts on genome architecture of DED fungi. Mobility of the unique introgressed *OPHIO2-int*. TE suggests that mobile OPHIO TEs may be horizontally transmitted between closely related species by means of interspecific hybrids.

III.5. Materials and Methods

III.5.1. Strains, media, culture conditions

O. ulmi strains Q412T and R21, *O. novo-ulmi* subsp. *novo-ulmi* strains AST27, CKT11 and H327 and the *O. novo-ulmi* subsp. *americana* strain W2 were maintained on solid Ophiostoma complete medium (Bernier and Hubbes, 1990). Liquid cultures for yeast-like cells were initiated by inoculating 50 mL of liquid OCM in 125 mL Erlenmeyer flasks with 1 mm² agar plugs cut from the edges of actively growing colonies on solid OCM. Cultures were grown at 22–24 °C with agitation (120 rev/min) for 3 days. For RNA extractions, mycelium tissue was harvested, immediately frozen in liquid nitrogen and ground to a fine powder in a pre-chilled mortar with pestle.

III.5.2. Nucleic acid extractions

Fungal genomic DNA was extracted (Zolan and Pukkila, 1986), resuspended in Tris-EDTA buffer (10-mM Tris-HCl, 1-mM EDTA, pH 8.0) and stored at -20° C. *Ophiostoma* spp. RNA was extracted using a small-scale version of the protocol by Chang et al. (1993): approximately 500 µL of tissue powder was transferred into a 1.5 mL microcentrifuge tube, 750-µL of hot (65°C) CTAB extraction buffer (100mM Tris-HCl, pH 8.0, 25mM EDTA, 2M sodium chloride, 2% CTAB, 2% polyvinylpyrrolidone, 0.5g/L spermidine) was added to each tube, vigorously mixed and placed into a 65°C water bath for 10 minutes with intermittent vortexing. Samples

were extracted three times with chloroform:isoamylalcohol (24:1) and total RNA was precipitated from the final aqueous phase by adding an equal volume of 7.5M lithium chloride, 50mM EDTA. Samples were gently mixed by inverting the tubes 4- to 6-times and incubated at -20°C for a minimum of 1-hour. Total RNA was recovered by centrifugation for 15-minutes at 4°C and 13,000-rpm in a tabletop microcentrifuge, the RNA pellet washed with 80% ethanol, air-dried and resuspended in RNase-free water. Total RNA concentration and quality were determined using a Multiskan spectrophotometer (Thermo Fisher Scientific Inc, Nepean, ON) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). One µg of total RNA was treated with 1U of DNaseI (Invitrogen, Burlington, ON) in a final volume of 10 µl.

Name	Sequence 5'> 3'	Annealing Temp. (°C)	Amplification size (pb)	
TIR_OPHIO1	TAAGTAATCCACCCTGCTTACCGGACG	67.6	1865	
TIR_OPHIO2	TAAGTGATCCACCCTACTTTTCACACACA	62.9	1865	
OPHIO1-internal-L	TTCAACTCTTGGCAATTCACTGCTACT	64	530	
OPHIO1-internal-R	TTGACGAGCTGAATCCAATGGTA	64	-	
OPHIO2-internal-L	GTCGTGGTTTCAGTGCCTTGATA	63	521	
OPHIO2-internal-R	GGATCGAGAGGTTGAAGGAC	60	-	
M13forward-20	GTAAAACGACGGCCAGTG	60	2118 (*)	
M13Reverse	CAGGAAACAGCTATGACCATG	60	-	
IPCR_OPHIO2_L2	GCTGCCGGTGAGTCTGTTACTGCTA	67.9		
IPCR_OPHIO2_R2	ACGCGTCCCGGAAAGAAGTGTTATT	64.6		
Ins-OPHIO2-AST27-L	GCAGGACAACGGCAAGAATG	63	2227 (**)	
Ins-OPHIO2-AST27-R	AGCGCTTTACGGCAGCAAC	62	-	
qRT_OPHIO2-L	CGACCTTCAAGCCGCTATTAAT	61	197	
qRT_OPHIO2-R	TGCCTTGGATAATAATCCAATTC	59	-	
ÔPHIO1-ultra	CGGTTTTGCGCCTTGA	60		
OPHIO2-ultra	CGCGGTTTTGATCCACAG	60		
5'RACE_O1	AAATGACGCGGTTTTGCGCTTGAA	73		
5'RACE_O2	TTCGACGCGTCCCGGAAAGAAGTGT	73		
actin-L	CGTGAGAAGATGACCCAGATTGTCT	65	128	
actin-R	CCATCACCAGAGTCCAGAACGATAC	65	-	

Table 5. Primers used in this study.

(*) The size of the PCR product indicated takes into account the presence of the full-length *OPHIO1* copy (1865 bp length without target site duplications TA). The size of the PCR product without the *OPHIO1* copy is 253 bp and represents pDRIVE plasmid vector sequence. (**) The size of the PCR product indicated takes into account the presence of the full-length

OPHIO2 copy (1865 bp length without target site duplications TA). The size of the PCR product without the *OPHIO2* copy is 360 bp and represents pDRIVE plasmid vector sequence

In order to locate the transcription start sites of *OPHIO1* and *OPHIO2* transcripts we followed the protocol of a commercially available primer extension kit (Promega Corporation, Madison, WI) using Superscript II Reverse Transcriptase (Invitrogen) for the reverse transcription reactions and OPHIO transposase primers (*OPHIO1*-Ultra and *OPHIO2*-Ultra, Table 5 and Figure 26). Primer extension products from the Primer extension system, with a 100-bp DNA ladder (GeneRuler; MBI Fermentas, Burlington, ON) were separated in 8% polyacrylamide gels containing 7 M urea. Primers and DNA size markers were labeled with [γ -32P]ATP. Intensities of hybridization signals were visualized by phosphor autoradiography using an Amersham Biosciences Typhoon 9400 Variable Mode Imager (GE Healthcare Bio-Sciences). For each reaction, about 10 µg of RNA was used.

In order to confirm the location of the consensual *hsp* promoter sequence within the *OPHIO1* TIR, we used a commercial RACE kit (SMARTTM RACE cDNA Amplification Kit, Clontech Laboratories Inc.). RACE primers (see Table 5 for primer sequences and Fig. 26) were designed using v. 0.4.0 of the Primer3 software (http://frodo.wi.mit.edu/; (Rozen and Skaletsky, 2000)) based on the sequence of *OPHIO1* (**DO649003**). Advantage 2 Polymerase Mix (Clontech Laboratories Inc.) was used for all RACE reactions. PCR products of interest were gel purified using Ultrafree-DA columns (Millipore Corporation, Bedford, MA) and modified TAE buffer (Millipore), cloned into the pGEM-T Easy plasmid vector and transformed into chemically competent *E. coli* JM109 cells (Promega Corporation). In cases where the amount of gel-purified RACE PCR product was insufficient for cloning, samples were re-amplified with the Universal Primer Mix (UPM) provided with the RACE kit and the respective gene specific primer, and cloned as described above. Inserts were amplified and sequenced with M13 universal primers.

80



Figure 26. Position of primers on sequences used in the study and indicated in Table 5. Each grey box corresponds to an OPHIO transposon of 1865 bp total length. TIRs are indicated by triangles at the extremities. Black boxes correspond to flanking regions of *OPHIO2-int*. Sizes of amplified sequences are indicated in nucleotides (bp).

III.5.4. Plasmid construction and fungal transformation

The virulent wild-type strain *O. novo-ulmi* subsp. *americana* W2 was used as recipient in all transformation experiments. Protoplasts were prepared following the protocol of Royer et al. (1991). All restriction enzymes used were from New England Biolabs (Pickering, ON). The 1867 bp fragment corresponding to the complete *OPHIO1* sequence was amplified by PCR using TIR_*OPHIO1* primer and *O. novo-ulmi* subsp. *novo-ulmi* H327 genomic DNA as a template. The purified PCR product was cloned into the pDRIVE plasmid vector (Qiagen, Mississauga, ON) and the resulting construct, designated *pD:OPHIO1*, sequenced. The *W2:OPHIO1* mutants were obtained following double transformation of the *O. novo-ulmi* subsp. *americana* W2 strain with 5 μ g each of plasmids pAN7-1 (hygromycin resistance) and *pD:OPHIO1* in a volume of 50 μ l and added to 198 μ l of STC (1-M sorbitol, 50 mM CaCl₂, 50

mM Tris-HCl, pH 8.0) and 2 μ l of protoplasts (1 x 10⁸ cells/mL). Transformation mixtures were kept on ice for 20 min. Then, 2.5 mL of PTC (40% PEG₄₀₀₀, 50 mM Tris-HCl, pH 8.0 and 50 mM CaCl₂) was added and mixed in sequential aliquots of 1 drop, 500 μ l and 2 mL. Samples were kept on ice for an additional 20 min before they were diluted with successive additions of 1 mL, 5 mL and 30 mL of STC. Samples were centrifuged for 10 min at 4800 x g at 4°C, and the resulting pellet was suspended in 1 mL OCM containing 0.6M sucrose. After incubation for 2 h at room temperature, 100 μ l protoplast suspension was mixed with 10 mL OCM containing 0.6 M sucrose and 0.7% agar, and immediately plated onto OCM agar with 0.6 M sucrose and 200 μ g/mL hygromycin B. Hygromycin-resistant colonies appeared after an incubation of 7 days and were transferred twice on selective OCM, followed by successive transfers on non-selective-and selective OCM, to verify the mitotic stability of the hygromycin-resistant phenotype.

III.5.5. PCR, Inverse-PCR, RT-PCR and semi-quantitative RT-PCR

All PCR amplifications were conducted from 10 ng of genomic DNA as follows: an initial denaturation at 94°C for 3 min. followed by 30 cycles of denaturation, annealing and polymerization at 94°C for 1 min, 60 to 68°C (depending on primers) for 1 min and 72°C for 1 min, respectively. All amplifications were performed with 1U of Platinum Tag® DNA polymerase (Invitrogen) in a 10 µl final volume, using primers listed in Table 5. Amplified fragments were separated on 1% agarose gels. Inverse-PCR with primers IPCR OPHIO2-L2 and R2 (Table 5) was used to recover the 5' and 3' flanking region of OPHIO2-int., as described by Bouvet et al. (2007). The efficiency of the OPHIO1 and OPHIO2 transposase primers was evaluated by performing RT-PCR on DNaseI-treated total RNA samples isolated from mycelium of the four strains of Ophiostoma spp. (i.e. Q412T, H327, AST27 and W2). We used SuperScript® II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions with OPHIO transposase primers (OPHIO1-internal-L and -R and OPHIO2-internal-L and -R, Fig. 26). To ensure that the RT-PCR amplicons originated from RNA and did not result from DNA contaminations, control reactions were included in which the reverse transcriptase was omitted. Products from the RT-PCR and control reactions, together with a 1-kp DNA ladder (GeneRuler; MBI Fermentas, Burlington, ON) were separated in 1.5% agarose gels, and visualized with the Gel Doc 1000 Gel Documentation System (BioRad, Mississauga, ON). Semi-quantitative RT-PCR was used to verify some differences in expression of OPHIO transposases after exposure of *Ophiostoma* spp. to UV light or 55°C heat shock, respectively.

DNAse-treated RNA (500 ng) from each survivor was reverse-transcribed (SuperScript® II Reverse Transcriptase, Invitrogen). From each of the resulting cDNA samples we then amplified part of the *O. novo-ulmi actin* gene and the *OPHIO2* transposase gene (see Table 5 for primer sequences). The *actin* gene was used to ensure that cDNAs generated from the various AST27 survivors yielded identical amplification profiles. All PCR amplifications were conducted as follows: an initial denaturation at 94°C for 3 min. and 21, 25 or 30 cycles of denaturation (Prakob and Judelson, 2007), annealing and polymerization at 94°C for 1 min, 62°C for 1 min and 72°C for 1 min, respectively. All amplifications were performed with 1U of *Platinum Taq*® DNA polymerase (Invitrogen) in a 15 μ l final volume. All PCR reactions were repeated once and controls lacking reverse transcriptase were included for each primer set at the maximum number of cycles, to ensure that amplicons were not derived from genomic or contaminating DNA. Products were electrophored in 1.5% agarose, stained with ethidium bromide and photographed. Three repetitions of RT-PCR amplifications after 25 PCR cycles were analyzed. We assessed band intensities visually and verified that differences in expression were consistent between repeats.



Figure 27. Flow diagram and nomenclature of stress treatments used to induce mobility of OPHIO transposons.

Stress treatments applied to the *Ophiostoma novo-ulmi* subsp. *americana* W2 strain and *O. novo-ulmi* subsp. *novo-ulmi* AST27 strain. Three abiotic stress treatments were (heat shock at 55°C, exposure to UV light or growth at 4°C,) used to induce *OPHIO1* and *OPHIO2* mobility. Selected survivors were named as follows: STRAIN_NAME—^{stress-number} as, for example, W2:*OPHIO1*⁻⁵⁵⁻³ which is the third (⁻³) survivor selected after heat stress (⁻⁵⁵) from the reference W2:*OPHIO1* strain. The numbers of selected survivors are indicated for each strain and stress (n=X). The red boxes indicate an absence of mobility contrary to the green boxes which indicate transposon mobility.

Yeast-like cells (10⁶ cells per mL) of *O. novo-ulmi* subsp. *novo-ulmi* strains AST27 and *O. novo-ulmi* subsp. *americana* strain W2:*OPHIO1* were grown on solid OCM. Each treatment, described in the flow diagram (Fig. 27), was done with AST27 and W2:*OPHIO1* and indicates all the stress-strain combinations and the nomenclature used. Depending on the stress applied, 10 to 30 survivors were recovered randomly after 3 days of growth at room temperature on solid OCM medium supplemented with 0.01% deoxycholic acid to limit the radial growth of colonies (Bernier and Hubbes 1990). Total genomic DNA was extracted from survivors after 3 days of agitation in liquid OCM medium at room temperature. The following three abiotic stress

treatments were tested: heat shock (55°C) during 1 hour in controlled incubator, cold stress (4°C) during 5 days; UV exposure (254 nm) for 2 min.

III.5.7. Southern hybridizations and sequence analyses.

The Southern protocol used was identical to that described by Bouvet et al., 2007. All sequences (Centre de recherche du CHUL, Plateforme de séquençage et de génotypage des génomes, Québec City, QC) were obtained on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequence alignments in Figures 28 and 31 were produced with *ClustalW* with default parameters (Higgins et al., 1994). The TATA-box was predicted with TATA prediction software (Milanesi et al., 1996). We used *NeutralNet promoter prediction* (Reese, 2001) to find promoter sequences and *TFsearch* (Heinemeyer et al., 1998) to find binding site motifs of transcription factors (Fig. 28A-28B).

III.6. Results

III.6.1. Analyses of UTRs and TIRs of OPHIO transposons and amplification of OPHIO transcripts

Further analysis of untranslated regions (UTRs) of OPHIO transposons revealed additional characteristics of active TEs. First, the 5'-UTRs displayed an AT-rich region, located about 60 bp upstream from the presumed translational start site, and two motifs, TATATAA at position - 24 bp for *OPHIO1* and TATTAATATA at position -39 bp for *OPHIO2*, which bear a close resemblance to the canonical TATAAA box presented as the putative TATA box (Fig. 28A). The 3'-UTRs of *OPHIO* transposons were found to contain the Cytoplasmic Poly-A Elongation (CPE) motif (Fig. 28B), which is the signal recognized by activated CPE-binding proteins (CPEB) thus triggering poly-A elongation (McGrew et al., 1989). Furthermore, a putative UUAUUUUGCAUUUA motif, only found in *OPHIO1*, looked strongly like the adenylate–uridylate rich element (ARE) motifs (UUAUUUAAA_AUUUA), an important signal in determining mRNA stability and translation (Friedrich et al., 2005) (underlined with a dotted line in Fig. 28B). Taken together, these data suggested that both *OPHIO* transposases could be transcribed.

(A)	
AF454467	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
X03714	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
M11340	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
HSCP	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
J01103>	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
K01966 (1)	~~~~~~~~~~~~~~~~~~~~~~~~~~TAGAACATTCGAGCT
K01966 (2)	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
A15900.1	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
DQ649003 0PHI01	TAA~GTAATCCACCCTGCTTACCGGACGCCCCCCCCGGACGTTTTTTTCCCTTTAATTTTTCATC
DQ649004 0PHI02	TAA~GTGATCCACCCTACTTTTCACACACCCCCTATTCTAATACATTTCCTCTTTATTA
TIR_Fot1	TAA~GTCAAGCACCCATGTAACCG-ACCCCCCT <mark>G</mark> GTAA CGACC
TIR_Pot2	TAACGTTGCGTACCCCCTGTTCGGCACCCCCCTGTTCGGCACCC
TIR Flipper	TAA~GTTGAGTACCCCACTTTCGG-ACCACCCCT <mark>CTTTT</mark> GGACCA
	STID STIR
DQ649003 0PHI01	A <mark>CCAATTCCATCGACGTCGACTTCAAAAAATGGTAA</mark> gtaattcaaaaaaaatcgtacacaacaaatta
DQ649004 0PHI02	A <u>CCAAT</u> TCTACCTAAGTCGACGAAAAAAATCAGTAAGTAATTCAAAAAAAA
DQ649003 0PHI01	<u>tataa</u> ctaactgtattttatag <mark>TA<u>AT</u>GCGGGAATATACGGAA</mark>
DQ649004 0PHI02	TCTAATTAGCC TTATTTATTAGTC<u>AT</u>GCACGATTATACCGAA
	M R E Y T E E N
	MHDYTEED
(B)	
DO64900310PHT01	GACGAAGAGGTGTAGAATGGTGTTTTTTTTTTTTTTTTT
DO64900410PHT02	GACGAAGAGGATTAAAACGGTGTTTTTTTTTTTTTTTTT
2 x 0 10 00 1, 01 11 02	
DQ649003 0PHI01	AAGTTTTTGATTTTTTCTAGTAAAAGAAAATTTTTATCGTC
DQ649004 0PHI02	GAGTTTTTGGATTTTATTACTGAAATGATTTGAATTTGTA

Figure 28. Alignment and analysis of terminal inverted repeats (TIRs) and untranslated regions (UTRs) of OPHIO1 and OPHIO2 transposable elements.

(A) Alignments of the 5'-UTRs and 5'-TIRs of fungal DNA transposons such as *Flipper* U74294, *Pot2* Z33638 and *Fot1* X64799 phylogenetically close to *OPHIO1* (DQ649003) and *OPHIO2* (DQ649004) were produced with ClustalW software (Higgins et al., 1994). The first seven sequences represent heat shock promoter sequences extracted from the NCBI nr database and for each, the accession number is indicated on the left side, the eighth being the canonical *hsp* sequence. The analysis of 5'-TIRs and 5'-UTRs for *OPHIO1* and *OPHIO2* reveals the presence of a canonical *hsp* promoter signature (black box). Dotted line red boxes show the transcription factor binding sites inside the alignment of 5'-TIRs of five DNA transposons. For each ORF of OPHIO transposons, the potential translation initiation codon (ATG) and the putative TATA and CCAAT boxes are underlined. Sequences of 5'RACE PCR products extending into the 5'UTRs are indicated by color coding. Alignment of the genomic DNA and the 5'RACE sequences revealed the presence of a 59 bp long intron immediately upstream of the *OPHIO1* translation initiation codon. A similar intron was not observed for *OPHIO2*. The sequence of the putative intron is indicated in italics.

(**B**) Alignment of the 3'-UTR and 3'-TIRs of *OPHIO1* and *OPHIO2*. Stop codons of each ORF are underlined. The Cytoplasmic Poly-A Elongation (CPE) motif (black box) was present in the two 3'-UTR of OPHIO transposons and can trigger poly-A elongation (McGrew et al., 1989). Furthermore, a putative UUAUUUUUGCAUUUA motif (underlined with a dotted line), only found in *OPHIO1*, demonstrates a high identity with the adenylate–uridylate rich element (ARE) motifs (UUAUUUAAA___AUUUA), involved in mRNA stability and translation (Friedrich et al., 2005).

After validation of the capacity of forward and reverse primers to amplify the expected internal fragment of the *OPHIO* transposases from genomic DNA extracts, these primers were used to verify the presence of *OPHIO1* and *OPHIO2* transcripts in total RNA samples extracted from *O. ulmi* and *O. novo-ulmi* strains. For *OPHIO1* (Fig. 29A), no amplification signals were obtained from RNA extracts from *O. ulmi* strain Q412T and *O. novo-ulmi* subsp. *americana* strain W2, which correlates with the absence of *OPHIO1* transposon in these strains (Bouvet et al., 2007). In contrast, total RNA extracts of *O. novo-ulmi* subsp. *novo-ulmi* strains H327 and AST27 allowed for RT-PCR amplification of *OPHIO1* transposase-specific transcripts. For *OPHIO2*, transposase-specific transcripts were only amplified from the *O. novo-ulmi* subsp. *novo-ulmi* subsp. *novo*

and OPHIO2 transposase enzymes are transcribed.



Figure 29. RT-PCR for OPHIO transposase expression in O. ulmi and O. novo-ulmi.

RT-PCR was performed on samples of total RNA extracted from mycelium of *Ophiostoma ulmi* strain Q412T, *O. novo-ulmi* subsp. *novo-ulmi* strains H327 and AST27 and *O. novo-ulmi* subsp. *americana* strain W2. Amplification products were obtained using *OPHIO1*-specific (Panel A; *OPHIO1*-internal-L and –R primers) and *OPHIO2*-specific (Panel C; *OPHIO2*-internal-L and – R primers). Strain W2 is considered as negative control (C-) due to the absence of *OPHIO1* and *OPHIO2* transposons in *O. novo-ulmi* subsp. *americana*, (Bouvet et al., 2007). Parallel reactions without reverse transcriptase (Panel B) were included to make sure that amplifications were not due to the presence of residual genomic DNA. The Control RNA (C+), provided with the SuperScript II kit (Invitrogen), is an 891-bp, *in vitro* transcribed RNA from the chloramphenicol acetyltransferase (CAT) gene and results in a 500-bp product

Secondly, the detection of many classical structures required for expression suggested that promoters or motifs could be present in 5'-TIR and 5'-UTR. Various software programs were used in order to investigate the presence of promoters upstream from the TATA-box and a putative *hsp* promoter was detected in the 3'-end of the 5'-TIR of *OPHIO1*. This promoter, located 135 bp upstream from the initiation codon, could activate the transcription initiation as already described. An alignment with other *hsp* promoters described in the literature confirmed the idea that this particular promoter bears a close resemblance to the consensual CNNGNNNNTTC promoter (<u>A15900.1</u>) found originally in *D. melanogaster* and the *DIRS-1* transposon in *Dictyostelium discoideum* (Pelham, 1982; Zuker et al., 1984) (Fig. 28A).

TFsearch used for prediction of binding site transcription factors denoted the presence of two potential binding motifs. Firstly, an alignment of TE sequences suggested the presence of the CCCCT motif in *OPHIO1* and *OPHIO2* TIRs but also in other TIRs of DNA transposons such as *Fot1* (from *Fusarium oxysporum*: X64799), *Pot2* (from *Magnaporthe oryzae*: Z33638) and *Flipper* (from *Botrytis cinerea*: U74294), all phylogenetically close to *OPHIO* transposons (Bouvet et al., 2007) (Fig. 28A). Curiously, no other DNA transposon of the *pogo* family (included in the *Tc1-mariner* superfamily described in a review by Daboussi and Capy, 2003) displays this particular motif in its TIRs. Secondly, both OPHIO transposons harbored a potential TC<u>CACCC</u>T motif in TIR-5' (leftmost red dotted box in Figure 28A) located 131-nt and 116-nt upstream from the TATA-box for *OPHIO1* and *OPHIO2*, respectively. The presence of this motif seems unique to *OPHIO1* and *OPHIO2* transposons and has not been reported in the other *Tc1/pogo* transposons in fungi.

III.6.2. RACE PCR and primer extension

5'RACE PCR confirmed positions of the ATG translation initiation codons in both *OPHIO* TEs and of the CCAAT motif. In addition, we identified a 59 bp long intron located immediately upstream of the *OPHIO1* translation initiation codon. The intron, indicated in italics Fig. 28A, presents the characteristics of group1 fungal introns: a conserved GTAATT motif in 5'end and a TAG motif at the 3'end, required for the splicing of the intron (Kempken and Windhofer, 2004). Primer extension experimentation was realized in order to verify the number of transcripts associated with the *OPHIO1* and *OPHIO2* TEs and revealed only a single 159 pb and 148 pb 5'end transcript for *OPHIO1* and *OPHIO2* TEs, respectively (data not show). The 5'-end of each transcript is indicated Fig. 28A in color. The size of the transcript associated to *OPHIO1* suggests that the expression of the *OPHIO1* transposase could in fact be under the control of the putative *hsp* promoter sequence.

III.6.3. *Heat shock induces incomplete transposition of an exogenous OPHIO1 copy in* O. novoulmi *subsp.* americana W2 *mutant strain*

After transformation of *O. novo-ulmi* subsp. *americana* strain W2 with *pD:OPHIO1* to incorporate the complete *OPHIO1* sequence and pAN7-1 to incorporate hygromycin resistance, we recovered transformants (W2:*OPHIO1*-X) with two to five *OPHIO1* insertions (as

determined by Southern hybridization, data not shown) and subjected them to a 55°C heat shock, 4°C cold treatments, or UV exposure. The mitotic stability of the insert was verified by successive subcultures of the selected transformants and Southern hybridizations (data not shown). Mobility was screened by Southern hybridization using a full-length *OPHIO1* PCR product as probe. Out of 50 survivors tested (30 for heat shock, 10 for cold shock and 10 for UV exposure), 6 demonstrated a variation of hybridization patterns (loss of one or more bands) in response to the heat shock treatment indicative of *OPHIO1* mobility (Fig. 30A). We conducted RT-PCR amplifications on selected survivors in order to verify the presence of *OPHIO1* transposase transcripts after mutagenesis. Data confirmed the expression of *OPHIO1* in each survivor except for the negative control (transformed strain *W2:OPHIO1* without the 55°C heat shock stress) (Fig. 30B). This experiment suggests the potential for incomplete transposition of the exogenous *OPHIO1* copy, i.e. excision of the insert from the vector without subsequent insertion into the *O. novo-ulmi* subsp. *americana* genome.



Figure 30. Incomplete transposition of OPHIO1 in O. novo-ulmi subsp. americana mutant strain W2:OPHIO1.

(A) Screening of individuals of the *O. novo-ulmi* subsp. *americana* mutant strain *W2:OPHIO1* for incomplete transposition after exposure to a 55°C heat shock treatment. Southern blot hybridization conditions were similar to those described previously (Bouvet et al., 2007). Lane: 1=W2:OPHIO1; Lanes 2 to 8: $W2:OPHIO1^{-55-A}$ to $W2:OPHIO1^{-55-G}$ which correspond

Lane: 1=W2:*OPHIO1*; Lanes 2 to 8: W2:*OPHIO1*^{-35-A} to W2:*OPHIO1*^{-35-G} which correspond to heat shock survivors obtained from W2:*OPHIO1*. Red stars (*) represent incomplete transposition (excision without reinsertion) illustrated by missing *OPHIO1* hybridizations. The W2 wild type strain served as negative control (C-).

(B) RT-PCR analysis using primers *OPHIO1-internal-L* and -R (see Table 5 for primer sequences) of *OPHIO1* transposase transcripts from total RNA samples extracted from mycelium of the same eight W2:*OPHIO1* stress survivors (1 to 8). *OPHIO1* transposase is not expressed in W2:*OPHIO1* (insertion of *OPHIO1* transposon in W2) without heat shock stress, but expressed after incubation of the fungus culture for 1 hour at 55°C.

III.6.4. Cold shock and UV irradiation induce perfect excision of OPHIO2-int in O. novo-ulmi subsp. novo-ulmi AST27As described recently, a particular copy of OPHIO2, named OPHIO2int., has been introgressed from O. ulmi in strain AST27 (Bouvet et al., 2007). The presence of only one copy of this transposon allowed mapping it inside a cluster linked to *Pat1*, a potential pathogenicity gene (Et-Touil et al. 1999). Based on the presence of OPHIO2 transcripts in AST27 but not in other O. novo-ulmi strains tested, we investigated the mobility of OPHIO2-int. using a distinct approach. Genomic DNA from strain AST27 was subjected to inverse PCR in order to recover the 5' and 3' flanking regions of the insertion which were cloned and sequenced. To corroborate the fact that we amplified flanking regions specific to OPHIO2-int. and not flanking regions of other truncated copies of *OPHIO2*, we mapped once again the copy with primers designed on the flanking regions (named Ins-OPHIO2-AST27-L and -R, Table 5, Fig. 26). Mapping results with these primers on F1 progeny from a H327 X AST27 cross were exactly the same as those obtained from the mapping of *OPHIO2-int*. (Bouvet et al., 2007) (data not shown). The approach used to demonstrate the mobility of OPHIO2-int was based on the presence/absence of this copy in various survivors recovered after UV irradiation, exposure to 4°C or heat shock. PCR amplifications were performed with the flanking region primers. Results indicated that five of the 60 survivors tested yielded a 360 bp long amplicon rather than the expected 2227 bp product (Fig. 31A). Three survivors were obtained after a UV shock (254 nm for 2 minutes) and two after incubation at 4°C for 5 days (2). Cloning and sequencing of AST27⁻ ^{UV3} (*) and AST27⁻⁴⁻³ (**) mutants demonstrated a perfect identity among flanking regions with O. ulmi strains (Q412T and R21), confirming the introgression of this fragment from an ancestral O. ulmi strain and the absence of the expected TAAGTTA footprint after excision of the transposon (Fig. 31B). In this case, a heat shock, as applied on OPHIO1, had no detectable impact on the mobility of OPHIO2-int.



Figure 31.

Figure 31. Perfect excision of OPHIO2-int induced by UV shock and cold stress in O. novo-ulmi subsp. novo-ulmi strain AST27.

(A) PCR analyses of *O. novo-ulmi* subsp. *novo-ulmi* AST27 survivors selected after exposure to UV light, growth at low temperature (4°C), or heat shock at 55°C. For each survivor, two PCR reactions were performed: one reaction using primer *TIR_OPHIO2* amplifying the entire 1865 bp *OPHIO2* sequence and another reaction using primer pair *Ins-OPHIO2-AST27-L* and *–R* located on the flanking region of *OPHIO2-int* insertions and amplifying either a 2227 bp or a 360 bp fragment in the presence or absence of *OPHIO2*. Primer sequences are shown in Table 5. *, ** and *** represent survivors AST27^{-UV3}, AST27⁻⁴⁻³ and AST27⁻⁵⁵⁻⁷, respectively. PCR amplicons from these three survivors were sequenced and compared.

(B) Alignment with *OPHIO2-int*. flanking regions, including a portion of *OPHIO2-int* copy (AST27 as positive control) and the sequences of two 360 bp length amplifications obtained from mutants AST27^{-UV3} (*) and AST27⁻⁴⁻³ (**). The sequence from mutant AST27⁻⁵⁵⁻⁷ (***) was used as negative control, whereas sequences were obtained for the following reference strains for species and subspecies: Q412T for *O. ulmi*, CKT11 and H327 for *O. novo-ulmi* sp. *novo-ulmi* and W2 for *O. novo-ulmi* sp. *americana*. Analyses revealed perfect identities between sequences of 5' and 3' flanking region of AST27, AST27^{-UV3}, AST27⁻⁴⁻³, AST27⁻⁵⁵⁻⁷ and Q412T, which confirmed the introgression of the DNA region from *O. ulmi* into *O. novo-ulmi* subsp. *novo-ulmi*. Target site TA was not duplicated after *OPHIO2-int*. excision in AST27^{-UV3} (*) and AST27⁻⁴⁻³ (**).

III.6.5. Semiquantitative PCR_3.5.

Semiquantitative RT-PCR with actin and OPHIO2 transposase specific primers was used on UV and heat shock treated AST27 survivors, respectively, as well as on AST27 not subjected to these stress treatments (negative control) in order to test whether or not the relative expression levels of the OPHIO2 transposase gene would increase in response to environmental stress. After 21 cycles of PCR, amplification was insufficient to visualize PCR products, whereas no difference in the intensity of amplicons was visually detected after 30 cycles (data not shown). After 25 cycles of PCR, however, we observed a difference in amplification intensities between AST27 survivors treated with UV or heat shock and the corresponding negative control (Fig. 32). Results indicate that: (1) the expression of the actin reference gene did not change in response to environmental stress; (2) the expression of the OPHIO2 transposase gene was more pronounced after UV exposure as compared to heat shock stress; (3) there was no apparent difference between intensities of PCR bands obtained from heat shock survivors and from the corresponding negative control, indicating that the 55°C heat shock treatment did not induce a variation in the expression of the OPHIO2 transposase gene contrary to UV stress; and (4) UV stress induced expression of the OPHIO2 transposase gene in all of the AST27^{-UV} survivors tested. This result is in agreement with the observation of mobility of OPHIO2-int. under UV stress.



Figure 32. Semi-quantitative RT-PCR analysis of Ophiostoma novo-ulmi subsp. novo-ulmi AST27 survivors using primers amplifying the actin gene and the OPHIO2-int transposase gene. Amplicons of OPHIO2 transposase (Tpase) (197 pb) and actin (128 pb) after 25 cycle of PCR amplifications obtained from AST27 survivors after exposure to UV light or to heat shock. RT-PCR was performed on samples of 1µg of total RNA extracted from mycelium of *O. novo-ulmi* subsp. novo-ulmi strain AST27 and reverse transcribed. Amplification products were obtained using *OPHIO2*-specific primers qRT_*OPHIO2*-L and –R and actin-L and -R specific primers (Table 5). The control (C) consists of strain AST27 not previously exposed to any of the above mentioned stress treatments.

III.7. Discussion

III.7.1. An hsp promoter is present in OPHIO1 TIRs

Recent studies on the DED fungi have highlighted the presence of DNA transposons. Two of them, named *OPHIO1* and *OPHIO2*, demonstrate characteristics of active TEs such as *Fot1* or *Flipper* contrary to *OPHIO3*, which was silenced by the RIP mechanism (Bouvet et al., 2007). Because of their potential effects on the genome structure of DED fungi, mobility of *OPHIO* TEs is an important aspect to study. More precisely, an active TE must carry the structural characteristics required for the expression of its own transposase and they must be located inside the 5'-TIR and 5'-UTR. Consequently, we first analyzed *OPHIO1* and *OPHIO2* transposons because their sequences satisfy the requirements for autonomous mobility such as perfect identity between TIRs (Bouvet et al., 2007). A recent study suggested that the pathway of heat shock response was self-regulated by a specific targeting of the *hsp* promoters by transposons (Walser et al. 2006). And inversely, it has been demonstrated that the integration of an *hsp* promoter into the TIRs of TEs can lead to over-expression of the transposase enzyme (Zuker et al., 1983, Zuker et al., 1984). On the basis of these interactions between promoters and TIR-UTRs of DNA transposons, we have investigated and discovered, in the 5'-TIR of *OPHIO1*, a sequence related

to the HSCP promoter described by Pelham (1982) in D. melanogaster. This particular sequence initiates the expression of hsp70 in D. melanogaster. The same promoter, located in TIRs of the DISR-1 transposon in D. discoideum, was also linked to over-expression of the DISR-1 transposase. In our case, we hypothesize that this hsp promoter could activate the transcription of the OPHIO1 transposase gene and then have an impact on OPHIO1 expression. Another striking feature of TIRs found in OPHIO1 and also OPHIO2 and a few Fot1-pogo DNA transposons, is the presence of CCCCT and CACCC transcription factor binding site motifs (Fig. 28A). Ikeda et al. (2001) hypothesized that the presence of the CCCCT motif inside the LTR of the retrotransposon MAGGY was responsible for its autonomous mobility under severe stress conditions. The optimal growth temperatures for O. ulmi and O. novo-ulmi are 28 and 21°C, respectively (Brasier et al. 1981) and we assume that a 55°C heat shock applied during 1 hour on Ophiostoma spp. strains is a severe stress. In this case, the CCCCT motif could have had an impact on *OPHIO1* mobility by binding transcription factors that could have activated the *hsp* promoter found in TIRs and, consequently, activate transposase transcription. The CACCC motif, usually located around 140nt upstream from the TATA box, was reported to be involved in βglobin suppression (Sargent and Lloyd, 2001) and to potentially bind *trans* factors to promote yglobin gene expression (Li et al., 2006). Cis-acting elements such as binding site motifs on TIRs are certainly one of the keys to the mobility of the transposon by catching *trans*-acting components. Because of the fine regulation of the transcription factors expressed during stresses (Dai et al., 2007), TEs that have integrated these motifs could use them to express their own transposase and, in this way, adapt their own mobility to specific stresses.

To verify the results of these genomic DNA sequence analyses concerning the location of some of the *OPHIO1* and *OPHIO2* sequence landmarks, we used 5'RACE PCR. This approach allowed us to confirm the location of the translation initiation codons for both *OPHIO* TEs and the CCAAT motif within the *OPHIO1* 5'UTR. In addition, 5'RACE demonstrated the presence of a 59 bp long intron sequence immediately upstream of the *OPHIO1* translation initiation codon. This intron has the characteristics of group 1 fungal intron sequences: a conserved GTAATT motif at the 5' splice site, which is of great importance for the splicing process and a TAG motif at the 3' splice site that produces the same splicing product as the CAT sequence (Kempken and Windhofer, 2004). A similar feature at the same position was lacking in *OPHIO2*

III.7.2. Expression of OPHIO1 and OPHIO2 transposases

The coding regions of *OPHIO1* and *OPHIO2* transposons have characteristics that suggest that they have the capacity to transcribe and express their respective transposase enzymes, a prerequisite for autonomous mobility. This hypothesis was tested by demonstrating that *OPHIO1* and OPHIO2 transposases were expressed in O. novo-ulmi mycelium. Because of its higher sensitivity for visualization of transcripts of DNA transposons compared to Northern hybridization (Kempken et al., 1998), we used RT-PCR with a high hybridization temperature to amplify OPHIO transcripts of each transposase. Only a limited number of strains were tested, the principal aim being to verify the presence of transposase transcripts and not to study their expression in populations. Our data suggest that OPHIO1 transposase transcripts are present in O. novo-ulmi subsp. novo-ulmi (H327 and AST27), which is correlated with the O. novo-ulmi subsp. novo-ulmi specificity of OPHIO1 demonstrated by Southern hybridization and PCR amplification (Bouvet et al., 2007). In contrast to OPHIO1, OPHIO2 transcripts were present only in O. novoulmi subsp. novo-ulmi strain AST27 and not in O. ulmi strain Q412T. A previous comparison of copies of OPHIO2 amplified from O. ulmi strains Q412T, R21 and H200 and from O. novo-ulmi subsp. novo-ulmi strain AST27 provided evidence for nucleotide mutations, 99.9% of them being C:G to T:A transitions (Bouvet et al., 2007). In fact, 60% of the transitions are CpA to TpA and two of these transitions generate stop codons inside the ORF (CAA \rightarrow TAA position 352 and TGG \rightarrow TAG position 724) except for the copy amplified in AST27 (*OPHIO2-int.*). This kind of transition is similar to repeat induced point (RIP) mutations (Galagan and Selker, 2004), an important silencing mechanism found in fungi, and recently observed in OPHIO3. In fact, this mechanism may be involved in blocking mobile copies of OPHIO2. OPHIO2-int could have been mobile within O. ulmi strains, as suggested by the variability in Southern hybridization profiles for this species (Bouvet et al., 2007). We can hypothesize that past mobility of OPHIO2 activated the RIP mechanism, which could explain the onset of C:G to T:A transitions between different copies. There is a strong possibility that OPHIO2-int was introgressed prior to activation of RIP in O. ulmi. Thus, OPHIO2-int could be the ancestral O. ulmi copy of an active OPHIO2 copy and explain the absence of amplification in Q412T by RT-PCR because of the RIP mutation inside this copy. The expression of OPHIO2-int in AST27 could be explained by a possible lack of the RIP mechanism in O. novo-ulmi or by of the presence of only one copy introgressed in AST27, since the RIP mechanism acts only on two or more copies of 400 bp minimum length and with more than 80% identity (Clutterbuck, 2004). We tried to find the O. ulmi donor of the OPHIO2int. introgressed copy by PCR amplification on 40 O. ulmi strains from a worldwide collection.
Currently, no *O. ulmi* strain has been found to contain the 2227 bp length amplicons of *OPHIO2int* observed in strain AST27.

Over the last two decades, our understanding of transposase expression has progressed and elegant studies have begun to elucidate the potential stimuli involved in its initiation. Many biotic and abiotic stresses seem to promote the mobility of the transposon in Table 6 (complementary data). Either in prokaryotic or eukaryotic systems, retroelements or DNA transposons were induced by stresses that can mimic environmental pressure. Their sequence analysis sometimes highlighted the presence of promoters or potential binding sites directly linked with the stress involved in the TE mobility as we have described in *OPHIO1* and *OPHIO2*. TEs and stresses related to their mobility listed in Table 6 suggest a strong influence of the environment on genomic architecture. This theory, first described by McClintock in 1984, proposed a direct relation between genomic response to environmental shocks and mobile elements and gave rise to the "smart cell" concept (Shapiro, 2001). All these results argue in favour of complementary analysis on the mobility of the OPHIO transposases and their inductibility by environmental stresses.

III.7.3. Incomplete transposition of OPHIO1

The phenomenon of autonomous mobility of TEs has been described in the literature (for fungi, see (Migheli et al., 1999): an autonomous TE moves by using its own transposase for excision. However, a TE can also move by catching transposases produced by other TEs (of the same family). Examples are *Guest* TE and their associated MITEs in *N. crassa* (Ramussen et al., 2004) or *Osmar* TE and *Stowaway* MITEs (Miniature TEs) which have been shown to multiply by using transposases of closely related active TEs (Feschotte et al., 2005). Autonomy of *OPHIO1* was tested by introducing copies in *O. novo-ulmi* subsp. *americana* strain W2. Some heat shock survivors of the *W2:OPHIO1* transformants showed a disappearance of *OPHIO1* copies introduced after mutagenesis and presence of *OPHIO1* transposase transcripts. Because of the presence of these transcripts, it seems unlikely that *OPHIO1* would have to depend on transposases expressed by closely related transposons present in the W2 genome. However, the mobility of *OPHIO1* looks strongly like incomplete transposition, i.e. excision from genomic DNA without subsequent re-insertion. This phenomenon was shown previously with *Jittery*, a *MuDR*-like transposon in maize. Thus, *OPHIO1* may behave in a way similar to the *Jittery* and *Restless* TE: elements capable of specifying their own chromosomal removal but not their

reinsertion (Windhofer et al., 2002 ; Xu et al., 2004). Two hypotheses may explain this situation: 1) the reinsertion events, which are rare in biological systems, were not observed in our case due to the small sample size; or 2) the heat shock damages an enzyme that normally suppresses transposon activity, and also damages enzymes required during the reinsertion inside DNA. The absence of *OPHIO1* reinsertion could perhaps explain the lack of variations in Southern profiles described in a previous study (Bouvet et al., 2007).

III.7.4. Interspecific hybrids may shuttle active TE between closely related species

The emergence of interspecific hybrids can help the evolutionary potential of each participating species by combining genes of two (or more) genomes, as evidenced by data obtained for fungal pathogens (Brasier, 2000) and endophytes of plants (Moon et al., 2004). In the case of the Dutch elm disease fungi, it appears that multiple rounds of interspecific crosses resulted in the introgression of a unique copy of OPHIO2 (OPHIO2-int.) from O. ulmi into O. novo-ulmi (Bouvet et al., 2007). The copy shows transcriptional activity and our data further suggest that its mobility can be induced by UV light and growth at low temperature. UV-induced TE mobility has already been demonstrated for *Mutator* (Mu) transposons in maize sperm (Walbot, 1999) as a response to the "UV genomic shock" and was shown to induce IS10 transposition in E. coli as a mechanism to increase survival of cells under severe environmental stresses (Eichenbaum and Livneh, 1998). Cold stress was linked with the mobility of the ISH27 TE in Halobacterium halobium (Pfeifer and Blaseio, 1990), whereas a transposon protein was recently found in a region underlying a quantitative trait locus for cold tolerance in rice seedling (Andaya and Tai, 2006). Semi-quantitative RT-PCR confirmed that UV irradiation increased the transcription of the OPHIO2 transposase gene, although the bioinformatics analyses could not highlight the presence of known promoters in 5'UTR or 5'TIR of OPHIO2-int.

The sequencing of flanking regions of *OPHIO2-int*. revealed neither recombination nor excision footprints but only the initial target TA, which highlighted the mechanism of perfect excision described in the *PiggyBac* element (Fraser et al., 1996; Grossman et al., 2000), the *Mu*-like transposon *Hop* (Chalvet et al., 2003), in *Tn10* and *Tn5* bacterial transposons (Murat et al., 2006) and in *Restless* elements after integration in *Neurospora crassa* (Windhofer et al., 2002). Although no new insertion could be demonstrated in *O. novo-ulmi*, we believe that a combination of introgression and abiotic stresses may act in favor of the invasion of TEs from one species to a closely related one. TE transmissions were already described in other fungi, for

example in *Saccharomyces* spp., where there is strong evidence for horizontal transfer of *Ty2* between *S. cerevisiae* and *S. mikatae* (Liti et al., 2005). According to the literature, a complete transposition seems to be rare within one generation but the rapid succession of new generations could accelerate these transpositions with reinsertions and, by this way, the invasion of new copies. Once *OPHIO2* was introgressed into AST27, the fast growth of *O. novo-ulmi* subsp. *novo-ulmi* could rapidly promote new generations which should help the multiplication of *OPHIO2* copies. Another important question that remains to be addressed is the potential contribution of TEs and interspecific hybrids to speciation (Kidwell and Lisch, 2002). Despite the fact that we were so far unable to generate new insertions in AST27 mutants, we think this hybrid, due to its sexual capacity, will be interesting for future studies to clarify the impact of TEs on genomic architecture and speciation.

During the last twenty years, two key concepts have emerged from genomic and population studies: induction of genomic rearrangements by TEs and contribution of interspecific hybrids to genome evolution. Transposons move and this mobility may be activated during 'genomic shock' as an adaptive mechanism (Madlung and Comai, 2004; McClintock, 1984). Contrary to fixation of stable mutations or rapid repair of DNA damage caused by point mutation or stress denaturation, transposons produce genomic changes by complete transposition long after their activation (Walbot, 1999). Consequently, they could have considerable implications on genomic architecture or gene expression and on species evolution. The present study specifically addresses the question of autonomous mobility of DNA transposons in Ophiostomatoid fungi. Experimental results linked stress to mobility of OPHIO1 and OPHIO2 transposons. In fact, induction of mobility mediated by cis-acting elements in 5'-LTRs of retroelements was previously proposed (Wessler, 1996) and was, globally, in accordance with the genomic shock theory of McClintock (McClintock, 1984). The presence of a promoter and a nucleotide binding site motif in 5'-TIRs of OPHIO transposons suggests the implication of *trans* factors, which are highly enhanced or repressed during stress. A practical application of this work could be the development of OPHIO1 as a tool for mutagenesis driven by an inducible hsp promoter. Concerning OPHIO2, the fact that its mobility can be induced by multiple stresses after being introgressed in a new species is the other main result of this work. In our opinion, these results consolidate our understanding of the dynamics of TE expansion at the intra- and interspecific

levels in fungi and we believe that future studies will correlate interspecific dynamic of TEs and *cis*-acting elements expressed during environmental stresses or after host genome pressures.

Acknowledgements:

G.F.B. thanks Marie-Josée Drouin for proofreading. This study was supported by grants from the Natural Sciences and Engineering Research Council of Canada to L.B.

Complementary data.

2	TE name	Host	Mobility induced by	Reference			
3 4	Prokaryote	karyote					
5	IS1	E. coli	SOS system	Lane et al., 1994			
6	IS5/IS30	E. coli	Carbon starvation	Naas et al., 1995			
7	IS10	E. coli	UV/SOS system	Eichenbaum and Livneh, 1998			
8	ISH27	Halobacterium halobium	Cold shock	Pfeifer and Blaseio, 1990			
9	IS1086	A. eutrophus CH34	Temperature (37°C)	Dong et al., 1992			
10	Mu prophage	E. coli	Carbon starvation	Lamrani et al., 1999			
11	Tn3	E. coli	Temperature (26 to 30° C)	Kretschmer & Cohen, 1979			
12	Tn5/Tn10	E. coli	SOS system	Kuan and Tessman, 1991			
13	Tn4652	Pseudomonas putida	Phenol Starvation	Kasak et al., 1997			
14		-					
15	Eukaryote						
16	412	Drosophila melanogaster	Temperature gradient	Viera & Biémont, 1996			
17	copia, 1731,						
18	297 and <i>mdg</i> -4	Drosophila sp.	Cell culture	Di Franco et al., 1992			
19	DIRS-1	Dictyostelium sp.	Heat shock, cell density	Zuker et al., 1983			
20	Dm-412	Drosophila sp.	Heat shock	Ratner et al., 1992			
21	Fotl	Fusarium oxysporum	Nitrogen starvation	Daboussi et al., 1992			
22	Jordan	Volvox carteri	Nitrogen starvation	Miller et al., 1993			
23	mariner	Drosophila sp.	Temperature following latitude gradient	Giraud & Capy, 1996			
24	MAGGY	Magnaporthe Oryzae	Heat shock, copper and oxidative stresses	Ikeda et al., 2001			
25	MGR586 (PL copy)	Magnaporthe Oryzae	Asexual propagation	Shull and Hamer, 1996			
26	Mutator	Zea mays L.	UV exposition	Walbot, 1999			
27	OPHIO1	Ophiostoma novo-ulmi subsp. novo-ulmi	Heat shock (55°C / 1 hour)	This article			
28	OPHIO2	Ophiostoma ulmi	UV, cold (4°C)	This article			
29	Р	Drosophila willistoni / D. sturtevanti	Temperature following latitude gradient	Silva and Kidwell, 2004			
30	Tam3	Antirrhinum majus	Heat shock (15°C to 25°C)	Robbins et al., 1989			
31	Tnt1	Nicotiana tabacum	Viral infection	Grandbastien et al., 1997			
32	Tos10/Tos17/Tos19	Oryza sativa	Cell culture	Hirochika et al., 1996			
33							
34	Table 6. Various m	able 6. Various mobile elements activated under stress condition in both prokaryotic and eukaryotic organisms.					

In each case, TE name, host genome, stress that induced TE mobility and reference are indicated.

Chapitre IV : Impacts du génome hôte sur les transposons à ADN : mise en évidence de régions positivement sélectionnées au sein des transposases chez les procaryotes et eucaryotes

IV.1. Avant propos

L'étude présentée dans ce chapitre a eu pour but de mettre en évidence des zones de sélection positive au sein des régions codant la transposase des éléments mobiles. Cette partie de ma thèse est la plus exploratoire et a eu pour but d'initier ma recherche de régions soumis à sélection positive au sein de séquences codantes pouvant présenter des particularités biologiques, tels que les transposons à ADN. Ce n'est que très récemment que le Dr. Malik du « Fred Hutchinson Cancer Research Center » à Seattle à initier ces travaux sur les pressions de sélection positive agissant sur certains gènes. En 2004, il publie la mise en évidence d'une adaptation évolutive d'un gène antiviral ancestral (Sawyer et al. 2004), en 2005, il s'attaque à la drosophile (Malik et Henikoff, 2005 ; Vermaak et al., 2005), en 2006, à la levure (Sawyer et Malik, 2006) pour revenir en 2007 à la drosophile (Rodriguez et al., 2007). Ces résultats ont été, pour ma part déterminant et ont influencé l'étude, préliminaire, qui va suivre. Ces résultats seront ultérieurement soumis pour publication.

<u>Réalisations et remerciements</u> : V. Jacobi et L. Bernier = corrections/rédaction.

IV.2. Résumé

Les éléments de transposition (TE) ont la particularité de se déplacer dans un génome, au sein d'une espèce (par le fait même de leur transposition et leur mode de fonctionnement) mais également d'une espèce à une autre (par l'intermédiaire de transferts horizontaux ou d'introgression). Une fois intégrés dans leur nouvel hôte, ils peuvent avoir des impacts sur la structuration génomique via, par exemple, les recombinaisons ectopiques ou encore modifier l'expression des gènes. Réciproquement, des pressions sélectives générées par le génome hôte (HG) peuvent inactiver ou faire diverger les TE en induisant des mutations dans leurs séquences. Nous avons examiné les séquences de plusieurs transposons trouvés

dans les génomes procaryotes ou eucaryotes dans le but de mettre en évidence des régions d'évolution différentielle. Les données utilisées nous ont permis de suivre des épisodes sélectifs fournissant l'évidence que des régions précises semblent soumises à des pressions de sélection positive. Premièrement, le gène codant l'actine a été utilisé comme contrôle négatif et a permis de confirmer l'absence de zones soumises à la sélection positive sur des gènes « métaboliques », quelle que soit l'espèce de champignons analysée. Des séquences des transposases Hermes et Hobo ont été utilisées pour tester, complèter et consolider le modèle afin donner une visualisation en trois dimensions de la localisation de ces sites soumis à une sélection positive. L'ensemble de cette analyse a été ensuite dérivée à partir des transposons de la famille *Fotl/pogo* et a permis de vérifier la présence de régions soumises à des pressions de sélection sur des séquences de copies similaires de TE mais isolées au sein d'espèces proches. Pour finir, nous avons voulu étendre nos analyses à des éléments procaryotiques. Nous avons pour cela utilisé l'élément IS5 qui nous a permis de mettre en évidence que des régions de sélection positive au sein des transposases n'étaient pas ciblées au hasard. L'ensemble de ces résultats pourraient, après démonstration biologique, mettre en évidence que les génomes hôtes auraient une influence sur l'évolution de ces domaines particuliers présents au sein des enzymes requis pour la mobilité des transposons.

IV.3. Abstract

Transposable elements (TE) have the particularity to move inside a genome, in a given species or from one species to another either by horizontal transfer or by introgression. Once in a new host genome, they can have impacts such as ectopic recombination or modification in gene expression. But, reciprocally, host genome (HG) selective pressures can silence TE sequences by inducing mutations. We have examined here the sequences of several transposable elements found in prokaryotic and eukaryotic genomes in order to see the presence of regions under positive selection. The data provide evidence of positively selected regions in each TE tested which have acted in different evolutionary lineages. First, the conserved *actin* gene was used as a negative control for positive selection to confirm the absence of positive selection hot spots on metabolic genes irrespective of the

fungal species tested. *Hermes* and *hobo* transposase sequences were used to consolidate the model and a 3 dimension-visualization of positive selected domains. Secondly, we applied this method to the fungal DNA transposons of the *Tc1/mariner* family to verify the presence of positive selected regions in closely related transposons which evolve in different species. Finally, we extended our study to prokaryotic elements and used the *IS5*-like dataset to reveal that the transposase undergoes positive selection in non random spots throughout the evolution of particular copies. Taken together, these results may demonstrate, after biological experimentations, that host genomes may have an influence on the evolution of particular domains of the enzymes required for the mobility of transposable elements. The identification of these putatively functional sites is important to understand, not only the mechanism of transposition itself but also the dynamics of the evolution of TEs within new species.

IV.4. Introduction

Nowdays, among all the mechanisms involved in the evolution of genomes, the interaction between transposable elements (TEs) and host genomes (HGs) is impossible to circumvent, each protagonist having impacts on the other. TEs are divided into two main classes depending on their mobility: class I or retroelements which move using a copy/paste mechanism via an RNA intermediate and class II or DNA transposons which move using a cut/paste mechanism (Craig, 2002) and each class can induce intra-genomic host variability. Due to their mobility and their random insertions, they contribute to partial or total gene inactivation after insertion into genes or may also place a gene under the control of TE regulatory sequences after insertion upstream of genes. DNA transposons can also generate polymorphisms due to the footprints created after excision. In addition, TEs have the ability to rearrange genomic information, and the resulting DNA rearrangements may be local or associated with large-scale chromosomal modifications (Daboussi and Capy, 2003).

Transposable elements have also developed strategies that enable them to persist. One such strategy is their ability to 'jump' the species barrier and propagate in different species. This

type of transfer, named horizontal transfer, is a well-documented mechanism described in several prokaryotic and eukaryotic systems: *IS5*, *IS200* in proteobacteria (Beuzon et al., 2004), *Fot1* in *Fusarium* spp. (Daboussi et al., 2002), *Tango1* in *Anopheles* spp. (Coy and Tu, 2007), *Hermes* in *Musca domestica* (Sarkar et al., 1997), *Minos* and *P* element in *Drosophila* spp. (de Almeida and Carareto, 2005; Silva and Kidwell, 2000), *mariner* in animals (Robertson et al., 1998) or *MuDr* and *MULE* in plants (Diao et al., 2006). Moreover, there is evidence for introgression of TEs due to interspecific hybrids between species: *OPHIO2-int*. in *Ophiostoma* spp. (Bouvet et al., 2007a), *dTph1* in *Petunia* spp. (Stuurman and Kuhlemeier, 2005), *Pis1* in *Pisum* spp. (Vershinin et al., 2003) or the active MITE *mPing* and *Pong* in wild rice (Shan et al., 2005). Interspecific mobility of TEs by introgression remains a passive transfer but can lead to the introduction of an active element within a new genome in which this element could multiply itself freely (Bouvet et al., 2007b; Shan et al., 2005). Mobility induces the potential expansion of TEs inside their host genome.

Reciproquely, host genomes developed several strategies to eliminate, domesticate or simply coexist with TEs in a stable evolution. Five major silencing mechanisms have been studied: (1) RNA interference (Dawe, 2003), (2) histone tail modifications (Tamaru and Selker, 2001), (3) DNA methylation (Mathieu and Bender, 2004) and (4) Repeat-induced point mutation (RIP) (Galagan and Selker, 2004). (5) The domestication of TEs has been well-documented in three cases: the *P* element in *Drosophila* spp. (by production of P repressor of the transposase, (Nouaud and Anxolabehere, 1997)), *piggyBac* in the genomes of a phylogenetically diverse range of organisms including fungi, plants, insects, crustaceans, urochordates, amphibians, fishes and mammals (Sarkar et al., 2003) and *mPing* MITE in rice cultivars (adaptation to environmental extremes) (Jiang et al., 2003).

Recently, Sawyer and Malik (2006) have developed the 'genetic conflict' model to describe the mechanisms that may govern coevolution between HG and TE whereby host genes are under constant selective pressure to limit the success of mobile elements and mobile elements are under constant selective pressure to evade these limitations. One of the possible signs of these pressures on DNA sequences could be the presence of specific hot spots inside the TE DNA sequence which evolve differently relative to the rest of the

We believe that analysis of TE sequences provides an opportunity to investigate genomic patterns of positive selection in various protein domains found in mobile elements. Many of the previous analyses for positive selection on genes relied on statistical tests based on a traditional pairwise approach, calculating and comparing the rate of non-synonymous (dN) and synonymous (dS) substitution between two sequences averaged over all codons. However, this approach is not appropriate for short sequences since they do not provide sample sizes (numbers of sites and/or substitutions) that are large enough to give significant results. In addition, this approach averages the substitution rate over all amino acid sites in the sequence. Because most amino acids are expected to be under purifying selection and positive selection most likely only affects very few sites, it often loses the power to detect positive selection. In order to improve detection efficiency we combined three strategies: (1) the likelihood ratio tests (LRTs) that allow one to estimate the ω ratio at particular sites rather than averaging it over the whole molecule and (2) the sliding window approach based on maximum parsimony and devised to conservatively predict the presence of selection from alignments, with special attention paid to reducing false positives (Sawyer et al., 2005). Moreover, a 3-D analyzer was used to calculate KA/Ks in 3D window and display the results on the 3D structure which allows the visualization of domains representing either positively, neutral, negatively or saturated selected sites, using the threedimensional (3-D) protein structure (Liang et al., 2006).

sequence.

We have used a succession of analyses on three distinct models to give a comprehensive account of the selection acting at particular domains of enzymes essential for transposon mobility. First, we have tested our model with the *Hermes* transposons in insect to detect and visualize positive selected regions on specific regions of mobile elements. The analysis of dataset suggested that positive selection was applied on the *hATC* domain of the *Hermes* transposase in *Musca domestica*. Secondly, we applied our model on the *Tc1/mariner* TE dataset to highlight positive selected regions and to know if positive selection is acting on mobile elements contrary to host genes. Resuls demonstrated particular spots under positive

selection in transposase and not in actin genes. Finally, we have extended our analysis to the prokaryotic systems and the analyses of *IS5*, an insertion element in proteobacteria, proves that particular positively selected sites may be targeted during the evolution of a TE.

IV.5. Materials and Methods

IV.5.1. Nucleotide sequence accession codes

All reported sequences have been deposited in GenBank. The assigned accession numbers are indicated directly in each phylogenetic tree except for the *Fot1*-like transposons where transposon names are used in the tree and the GenBank accessions listed in the text.

IV.5.2. Phylogenetic analyses

Protein sequences were aligned using T-Coffee (Notredame et al., 2000). The phylogenetic trees were constructed using MEGA (version 4.0) (Tamura et al., 2007) by the Neighborjoining (NJ) method using p-distance estimates, which is thought to be the most reliable method for constructing NJ trees of closely related sequences. The tree was rooted with more divergent sequences depending on the tree and the reliability of each node was assessed with 2000 bootstrap replications. Likelihood ratio tests (LRTs) were performed using codeml with the site-specific models of the PAML3.14 package (Yang, 1997) and the tree constructed as above. The six site-specific models recommended by Anisimova et al. (Anisimova et al., 2002) were tested: M0 (one-ratio), M1 (neutral), M2 (selection), M3 (discrete), M7 (beta), and M8 (beta&). These LRTs indicate whether the substitutions inferred from an alignment are best explained by one of two models of $\omega = dN/dS$, where dN and dS are the non-synonymous and synonymous substitution rates respectively. When parameter estimates under a model allowing positive selection suggested the presence of a number of sites with $\omega > 1$, Bayes inference was used to calculate the posterior probability that a given site is one of those that are selected. The selection operating in different regions of the sequences and within different branches of the phylogenetic tree under study was also estimated using SWAPSC1.01.0 (Fares, 2004) with 2000 simulated data sets generated by Evolver program included in PAML3.15. The simulated datasets were obtained from the M7 model which seems to be the best model to predict the nucleotide substitutions. SWAPSC1.0 uses the differences between the estimated and expected numbers of synonymous and non-synonymous substitutions to evaluate various hypotheses. Firstly, it seeks evidence for regions that have undergone the saturation of synonymous sites, where the number of synonymous substitutions is significantly smaller than expected. In addition, it seeks mutational hotspots: regions where the number of synonymous and non-synonymous nucleotide substitutions are greater than expected under neutrality. Remaining regions where the number of non-synonymous nucleotide substitutions is smaller than expected (or where ω is significantly smaller than the mean ω estimated for the alignment) are identified as under negative selection. Positive selection is inferred where the estimated number of non-synonymous nucleotide substitutions is greater than expected by chance and where ω is significantly greater than 1. Where regions have an estimated number of non-synonymous substitutions greater than expected but $\omega < 1$ or where $\omega > 1$ but there is evidence for saturation of synonymous sites, such regions are said to have accelerated rates of non-synonymous substitutions. Thus SWAPSC1.0 seeks to avoid inferring positive selection where there is insufficient data to support it or where saturation may cause bias. A by-product of the SWAPSC1.0 analysis is substitution rate estimates for all branches of the tree under study, within overlapping 3 codon windows across the alignment.

IV.5.4. Visualization of positive selection

SWAKK (Liang et al., 2006) was used to detect, confirm (with 1-D analyzer) and visualize (with 3-D analyzer) positive selection using a reference protein structure (PDB) file when the 3-D structure of proteins was available. Recent studies further indicate that when the three-dimensional (3D) protein structure is available, it is much more sensitive to detect positive selection if windows in 3D space are employed in the analysis. The SWAKK method is based on pairwise alignment of DNA coding sequences and the PDB file of the

protein encoding by the DNA coding sequence. The alignment is then reversely translated to obtain a codon-based sequence alignment. We have made different genetic code tables available to account for the variability of the genetic code. Each amino acid in the reference structure is represented as the ClÁ atom. SWAKK constructs the 3D windows by placing each amino acid at the center and including all amino acids within a pre-specified distance from the center. For each window, all the corresponding codons within it are extracted to form the sub-alignment, and the dN/dS score is calculated using the PAML3.14 package. Finally, amino acid sites under different selective pressures are colored differently and visualized on the 3D structure using the Chime plug-in component. To increase the robustness of the analysis presented in Fig. 33B-C, we enlarged the window size from 10 Angstroms (default parameter) to 20 Angstroms length.

IV.6. Results

IV.6.1. Visualization of positive selection using the Hermes DNA transposon

Significant evidence of positive selection was sought using two different programs: PAML3.15 and SWAPSC1.0. Following the analyses by Semple et al. (2005), three pairs of PAML site-specific likelihood models were compared that assume variable selective pressure among sites for *actin* and *Tc1/mariner* DNA transposons: M0 (one ratio) versus M3 (discrete), M1 (neutral) versus M2 (selection), and M7 (beta) versus M8 (beta&.) using Maximum likelihood estimates (MLEs) and likelihood ratio tests (LRTs).

Maximum likelihood estimates (MLEs) and likelihood ratio tests (LRTs)

MLE: The probability of observing the data *X*, when viewed as a function of the unknown parameters θ with the data given, is called the *likelihood function*: $L(\theta; X) = f(\theta|X)$. According to the *likelihood principle*, the likelihood function contains all information in the data about parameters θ . The best point estimate of θ is given by the θ that maximizes the likelihood *L* or the log likelihood $l(\theta; X) = \log\{L(\theta; X)\}$. Furthermore, the likelihood curve provides information about the uncertainty in the point estimate.

LRT: Suppose the simpler (null) model has p0 parameters and the more general (alternative) model has p1 parameters, and the (optimal) log likelihood values under the two models are 10 and 11. Then twice the log likelihood difference, $2\Delta I = 2(11 - 10)$, has asymptotically a $\chi 2$ distribution with d.f. = p1 - p0 if the null model is true. So the test statistic $2\Delta I$ can be compared with that $\chi 2$ distribution to test whether the null model is rejected against the alternative model. (Yang, 1997)

Very recently, several 3D structures of TEs such as IS200 (Lee et al., 2006), Hermes (Hickman et al., 2005)or Mos1 (Richardson et al., 2006) became available on the NCBI web site. We included *Hermes* and *Hobo* sequences from five different insects to extend our analysis to invertebrate DNA transposons. First, the LRTs were calculated and showed that model M8 fit the dataset $(LRT(M7:M8)_{Hermes} = 5.9, p < 0.05, see Table 7)$. The Figure 33A represents a phylogenetic tree based on *Hermes* and *Hobo* sequences. Branch in red highlights a positive selected divergence between Musca domestica and the other Muscomorpha tested. The SWAPSC1.0 analysis indicated that three consecutive windows demonstrated positive selected region at the 3'-end of the ORF coding for the Hermes transposase. Figure 33B shows graphs of the distribution of the ω ratios along the sequences by using pairwise comparisons of *Hermes* found in *Musca domestica* and the four other copies found in Drosophila melanogaster, Mamestra brassicae, Bactrocera tryoni and Lucilia cuprina, respectively. For each comparison, we completed the analysis by giving a 3-dimension visualization of the position of the selected sites with a protein databank file available (window size = 10 Angstroms). The positive selected sites are indicated in red and correspond to the ω values higher than 1.2, and the negative selected sites (in blue) to the ω values lower than 0.8. The neutral evolution values are comprised between 0.8 <NE< 1.2 and indicated in grey. Yellow regions show the undetermined position or irrelevant chains The combined results of phylogeny, variation of ω ratios and 3-D visualization seem to indicate the following: first, the more the species have diverged, the more the positive selected regions are present (an increase of peaks and red spots) and secondly, the 3'-end region seems to undergo more important positive pressures.

A conserved domain analysis carried out with BLASTP indicated the occurrence of two domains inside the *Hermes* transposase: the first one, the BED zinc finger domain (in blue, Fig. 33C) is a DNA-binding domain in chromatin-boundary-element-binding proteins located in the 5'-end of the transposase. The second one, the *hAT* family dimerisation domain (named *hATC* domain, in red) is at the 3'-terminus of the transposases of elements belonging to the *Activator* superfamily. The graphic (sliding window = 60-nt) presented in

Fig. 33C shows the mean of the ω ratios of each site and for the four comparisons. This analysis, once again, is possible due to the exact same length of the sequences analyzed and confirmed that the *hATC* domain (3'-end of the transposase) underwent higher positive selection. The SWAKK 3-D visualization with a more stringent analysis window of 20 Angstroms shows clearly that only this region was positively selected (in red, Fig. 33C). Taken together, these results illustrate and that specific regions of the transposase region were positively selected during the evolution of a DNA transposon.



Figure 33.

Figure 33. Three-D visualization of the positively selected site analyses on Hermes in insects

(A) The phylogenetic tree shown is an inferred N-J tree with several orthologous *Hobo-Hermes* transposons present in 5 *Muscomorpha* species. Each sequence has 1590-nt length and represents about 86% of the whole sequence. This phylogeny is largely in agreement with the published phylogeny of the *Hermes* transposons by Robertson (Robertson, 2003). The bar denotes 2% nucleotide divergence and diversity (determined by use of MEGA4.0 software); bootstraps values are indicated below the branches and correspond to 2000 replications. Branch in red in the tree indicate the presence of positively selected regions inside the DNA sequences with ω values higher than 1 (ω >1, p < 0.0001).

(B) Sliding window analyses of 530 amino acid long *Hermes*-like transposases found in 5 *Muscomorpha* species. Regions where dN/dS (ω) > 1.2 are potential sites of positive selection with a 45-nt window size (15aa.). For each comparison, a 3-D visualization of the A chain is shown using SWAKK 3-D analyzer with a 3-D window size of 10 Angstroms (default parameters). **2BW3** file was used to generate the 3-D visualizations of Hermes transposase. Red point represents sites under positive selection (ω > 1.2), blue point the negative ones (ω < 0.8), the grey ones are sites under neutral evolution (0.8 < ω < 1.2) and the yellow ones are undetermined position or irrelevant chain.

(C) The summation fo the 4 graphs of ω ratio mean by site for the comparisons obtained with SWAPSC1.0 (window size = 60-nt) is present and confirmed by SWAKK 1-D analyzer (window size = 60-nt). The 3'-extremity of the transposase sequence exhibits a region containing positive selected sites, which corresponds to the *cAHT* domain. The 3-D visualization of the A chain is shown using SWAKK 3-D analyzer with a 3-D window size of **20** Angstroms and was obtained with the <u>2BW3</u>. Red point represents sites under positive selection ($\omega > 1.2$), blue point the negative ones ($\omega < 0.8$), the grey ones are sites under neutral evolution (0.8 < $\omega < 1.2$) and the yellow ones are undetermined position or irrelevant chain. Under severe constraints, only one red region is present and corresponds to the 5'-extremity to the hATC domain of the *Hermes* transposase.

IV.6.2. Sequence variations and fast-evolving regions in fungal DNA transposons

The starting point of this part came from an observation based on OPHIO transposon sequences. These TEs were recently found in Ophiostoma novo-ulmi (OPHIO1) and O. *ulmi* (*OPHIO2*) and they were subjected to various studies of distribution (Bouvet et al., 2007a) and mobility (Bouvet et al., 2007b). A striking feature is their homology: the same total length (1865-nt) and coding region length (1569-nt) with the same position of start (161-nt) and stop (1732-nt) codons and the same TIRs length (42-nt). In spite of this homology, these two elements have no more than 75.5% identity (Fig. 34A). A rapid analysis showed that the more variable regions of both TEs are the TIRs and the HTHdomain (72.6% of nucleotide identity versus 64.3% of amino-acid identity) and a region located downstream of the catalytic domain of the transposases (74.8% of nucleotide identity versus 62.3% of amino-acid identity) (Fig. 34B). Whether OPHIO1 and OPHIO2 are linked together remains an open question. If they have a common ancestor, could their presence in two different species have induced their sequence variations? In our opinion, these regions of non synonymous variations (higher divergence of amino acid sequences) make these elements interesting candidates for studies on the evolution of particular domains of the transposase. Ascomycete actin genes and fungal Tc1/mariner transposons were chosen to investigate our hypothesis of the existence of evolutionary hot spots within the sequences of these fungus TEs. Two Neighbor-joining (NJ) trees were constructed from fungal amino acid sequences using p-distance estimates and we used respectively 19 and 21 aligned sequences for *actin* genes (816-nt length, precisely 272 amino acid long; Fig. 35A) and *Tc1/mariner* DNA transposons (1149-nt length, precisely 383 amino acid long; Fig. 35B).

Identitiy	Pot2	Flipper	Fot1	OPHIO1	OPHIO2
Pot2	100				
Flipper	46,16	100			
Fot1	40,14	42,52	100		
OPHIO1	47,83	44,56	37,4	100	
OPHIO2	47,71	43,6	37,35	75,51	100



Figure 34. Comparison of Fot1-like TEs and OPHIO transposon sequences

(A) Comparison between various *Fot1*-like TEs used in this study. Accession numbers of each TE are: *Pot2* (from *Magnaporthe oryzae*: gil496853), *Flipper* (from *Botryotinia fuckeliana* [anamorph: *Botrytis cinerea*]: gil1644409), *Fot1* (from *Fusarium oxysporum*: gil2722), *OPHIO1* (from *O. novo-ulmi* subsp. *novo-ulmi*: gil109156270) and *OPHIO2* (from *O. ulmi*: gil109156272). Nucleotide identity is given in percentage.

(B) Structural characteristics of *OPHIO1* and *OPHIO2* transposons found in *Ophiostoma* spp. TDS: target site duplication; TIR: terminal inverted repeats; UTR: untranslated region; ORF: open reading frame. Specific domains of interest are indicated inside the ORF: HTH-psq: helix-turn-helix domain implied in TIR recognition by the transposase; CENPB: Putative DNA-binding domain in centromere proteinB; DDE: DDE superfamily endonuclease domain responsible for the catalytic activity of the transposase. Nucleotide identity (NI) and amino acid identity (AAI) percentages were calculated with compseq software of EMBOSS package (Rice et al., 2000).

Figure 34.

Name P(%)	Model comparison	l(2xDlnL)l	p-value	NEB/BEB
actin	M0 vs. M3	$ 2 \times (-4353,39-(-4294,04)) = 118,68$	<0,0001*	
	M1 vs. M2	$ 2 \times (-4334, 29 - (-4334, 29)) = 0$	ns	
	M7 vs. M8	2 x (-4295, 63-(-3295, 62)) = 0,01	ns	ns/ns
Tc1/mariner	M0 vs. M3	$ 2 \times (-8582,24-(-8501,07)) = 355,77$	<0,00001*	
DNA transposon	M1 vs. M2	$ 2 \times (-8501, 7 - (-8501, 7)) = 0$	ns	
family	M7 vs. M8	$ 2 \times (-8403, 18 - (-8400, 18)) = 6,23$	<0,01	ns/ns
<i>Fot1</i> -like DNA transposons	M7 vs. M8	l2 x (-7907, 5-(-7905,6))l = 3,73	<0,15	0.953/4S* <95%
<i>Aft1</i> -like DNA transposons	M7 vs. M8	$ 2 \times (-4245,01-(-4246,35)) = 0,68$	ns	ns/2S*
IS5	M0 vs. M3	2 x (-11989,7-(-11930, 1)) = 119,23	<0,0001*	ns
	M1 vs. M2	$ 2 \times (-11930, 1 - (-11930, 1)) = 0$	ns	
	M7 vs. M8	$ 2 \times (-11685, 38 - (-11683, 01)) = 4,74$	<0,095	
Hermes	M7 vs. M8	$ 2 \times (-3352,05 - (-3349,10)) = 5,9$	0,05	ns
* M0:M3 comparison,	although cited in the literature, i	s not considered to be a suitable comparison for detecti	on of positive selection.	

Table 7. LRT and Bayesian tests calculated from the transposable element datasets

The first two columns indicate the name of sequences tested and the model comparison used in the LRT tests, respectively. Columns 3 and 4 report results of PAML analysis, where we compared the likelihoods obtained by modeling codon evolution under corresponding models using a multiple alignment. By evaluating twice the difference of the log-likelihoods $(2*\Delta lnL)$ between M7 and M8, under a χ^2 distribution with 2 df, *p* values for positive selection were obtained (n.s. indicate not significant value). The two last columns show the naïve empirical Bayesian (NEB) and the Bayesian empirical Bayesian (BEB) tests estimated from the M8 model. Only significant result was obtained for *Fot1*-like TEs. The values indicated the ω ratio in first and the number of site (4S) under positive selection of BEB tests with a significant (p<95%) in second.

The trees and nucleotide sequences derived from the amino acid sequences were used in the following analyses. We found that clades and global phylogenetic structure of trees reflects the fungal classification published in the literature (Berbee, 2001) with other genes and are well-supported by strong bootstrap values.

For each case (actin genes and Tc1/mariner family), we calculated the likelihood ratio tests (LRTs) of the three pair model comparisons (Table 7). For both the actin gene or the *Tc1/mariner* transposons, the M3 model (allowing variation in ω between two site classes) was a significantly better fit to the data than M0 (allowing no variation in ω) with the LRT statistic (for example, $2\Delta \ln(M0:M3)_{Tc1/mariner} = 415.62$, p < 0.0001 with 2 degrees of freedom for *Tc1/mariner* TEs). However, this comparison is essentially a test of variability in the ω ratio among sites and does not constitute a rigorous test of positive selection (Semple et al., 2003), the M3 model giving an over-estimation of the ω ratio and consequently many false positives. However, the LRT_{M0:M3} is already interesting because it demonstrated a 3-fold increase of the ω ratios in *Tc1/mariner* DNA transposons sequences contrary to the actin genes $(|2\Delta \ln(M0:M3)|_{Tc1/mariner} = 355.77 \gg |2\Delta \ln(M0:M3)|_{actin} =$ 118.68), so a higher variability of the *Tc1/mariner* TE ω ratios contrary to the *actin* genes. Concerning the comparison between M1 (neutral) and M2 (positive selection), the LRT failed to show a significant difference in fit to the data either for TEs or actin sequences. The last comparison suggested was the M7 model versus M8. The M7 model assumes a beta distribution for ω over 10 categories of sites. The beta distribution is limited to values between 0 and 1 providing the most flexible null hypothesis, and most stringent test, for testing positive selection. Model M8 (or Beta&) adds another site class to the M7 model, within which ω is estimated from the data. An LRT showed that the M8 model allowing positive selection was a significantly better fit to the data than M7: $LTR(M7:M8)_{Tc1/mariner} =$ 6.23, p < 0.001 with 2 degrees of freedom for *Tc1/mariner* TEs contrary to the *actin* gene (Table 7). However, neither the subsequent naive empirical (NEB) nor Bayes empirical Bayes (BEB) procedures denote positive selection (95% threshold). These results are expected for this data set and fit the fact that transposable elements must undergo some positive selection hot spots contrary to host genes. It is known that these LRTs suffer from a lack of power to detect significant effects (Anisimova et al., 2002)one of the M2 and M8 models suggested identified sites under positive selection (i.e. at greater than 95% confidence for NEB and BEB tests). A subsequent analysis was carried out by splitting the *Tc1/mariner* data set in two groups: the *Aft1*-like transposon group from *Aspergillus* spp. and the *Fot1*-like group. By using the same LRTs, results demonstrated no evidence for positive selection for *Aft1*-like contrary to *Fot1*-like transposon dataset: LTR(M7:M8)_{*Fot1*-like = 6.23, p < 0.001, df = 2 (Table 7). However, the M8 model suggested that a small proportion of sites were under positive selection for the added class ($\omega = 4.38$ for *Aft1*-like and $\omega = 9.53$ for *Fot1*-like TEs). In Bayesian tests carried out with M8 model, *Fot1*-like elements exhibited one specific positive selection site (i.e. at greater than 95% confidence) for NEB and 4 sites (indicated 4S* in Table 7) for BEB. However, only 2 sites (indicated 2S* in Table 7) were detected for *Aft1*-like elements with the BEB test.}

The SWAPSC1.0 sliding window analyses of all actin gene and TE data broadly reflected the LRT results. Globally, neither positive selected sites, acceleration of non-synonymous substitution regions nor hot spots can be detected with the actin dataset and only 0.25% of the sites are subjected to negative selection (Fig. 35A). The *Tc1/mariner* transposon dataset showed three results: first, the global dataset including 21 sequences of Tcl/mariner transposons demonstrate a small number of selected sites under positive selection: about 0.12% of the sites were estimated to be subjected to positive selection, 0.09% demonstrated an acceleration of non-synonymous substitutions and 0.19% of the sites demonstrated negative selection. Only the branches identified as under positive selection are indicated (in red) in Figure 35B. Secondly, for Fot1-like transposons, positive selected branches were highlighted between Nht1 (gil2439967) in Nectria heamatococca and Fot1 (gil2722) in Fusarium oxysporum and between OPHIO1 (gil109156270) in O. novo-ulmi and OPHIO2 (gil109156272) in O. ulmi. This result must be linked with the first observation showing some variation spots in amino acid sequences between OPHIO1 and OPHIO2 sequences. For Aft1-like transposons, we detected positive selection on Aft1 (gil118505101) and Afu (gil66848978) transposons in Aspergillus fumigatus, and obtained evidence that the Aft1 transposon had undergone positive selection between Aspergillus oryzae and A. fumigatus. We decided to split the global analysis into 2 groups represented by the *Fot1*-like transposons (Fig. 36A) and the Aft1-like transposons (Fig. 36B). Although the dataset is less important in terms of the number of sequences, we obtained the same result as described in Fig. 36B. We completed the split analyses with a 1 dimension analyzer of SWAKK software that compared pairwise sequences. The results of five pairwise sequence comparisons were obtained:

(A) Actin genes



Figure 35. Phylogenetic trees relating to actin genes and Tc1/mariner DNA transposase family proteins in ascomycete fungi

(A) The phylogenetic tree shown is a consensus neighbor-joining phylogenetic tree of *actin* genes derived from 19 fungal species. Each sequence has 816-nt length and each represents about 70.8% of the whole sequence. This phylogeny is largely in agreement with the accepted phylogeny of fungal host species, indicating that the gene was acquired by strict vertical inheritance. The bar denotes 0.02% nucleotide divergence and diversity

(determined by use of MEGA4.0 software); bootstraps values are indicated below the branches and correspond to 2000 replications.



Figure 35. Phylogenetic trees relating to actin genes and Tc1/mariner DNA transposase family proteins in ascomycete fungi

(B) A consensus neighbor-joining phylogenetic tree of *Tc1/mariner* orthologous sequences based on the coding sequence of the transposase. Each sequence has 1068-nt length and each represents about 70% of the whole sequence. This phylogeny is largely in agreement with the published phylogeny *Tc1/mariner* transposons by Bouvet et al. (Bouvet et al., 2007a). The bar denotes 1% nucleotide divergence and diversity (determined by use of MEGA4.0 software); bootstrap values are indicated below the branches and correspond to 2000 replications. Branches in red in the tree indicated the presence of positive selected regions inside the DNA sequences with ω values upper than 1 (ω >1, p < 0.0001).



Figure 36. Split analyses of Tc1/mariner TEs divided in two distinct groups: the Fot1-like and Aft1-like groups

(A) A consensus neighbor-joining phylogenetic tree of *Fot1*-like orthologous sequences based on the coding sequence of the transposase. Each sequence has 1068-nt length and represents about 70% of the whole sequence. The bar denotes 0.05% nucleotide divergence and diversity (determined by use of MEGA4.0 software); bootstraps values are indicated below the branches and correspond to 2000 replications. Branches in red in the tree indicate the presence of positively selected regions inside the DNA sequences with ω values higher than 1 (ω >1, p < 0.0001). In parallel, we present the graph of 5 pairwise *Fot1*-like and *Aft1*-like transposase comparisons extracted from SWAPSC1.0 analyses (window size = 45-nt): (a) *Pot2* copies isolated from the strains 70-15 clone 13H23 (gil496853) and clone 72H05 (gil94540563) without detectable positive selection. (b) *Nht1* from *Nectria heamatococca* (gil2439967) and *Fot1* from *Fusarium oxysporum* (gil2722). The blue curve was produced with SWAKK software and the orange curve with SWAPSC1.0. (c) *OPHIO1* (gil109156270) from *Ophiostoma novo-ulmi* subsp. *novo-ulmi* and *OPHIO2* (gil109156272) from *O. ulmi*; (d) (gil118505101 from *A. fumigatus* and gil83767101 from *A. oryzae*) and (e) *Aft1* (gil66843917) and *Afu* (gil66848978) two transposons from *A. fumigatus*. Peaks higher than 1 (ω >1, p < 0.0001) are considered as positively selected regions.



Figure 36 (suite). Split analyses of Tc1/mariner TEs divided in two distinct groups: the Fot1-like and Aft1-like groups (B) The same analysis based on N-J phylogenetic tree constructed with *Aft1*-like transposase sequences is presented.

(a) Pot2 copies isolated from the Magnaporthe oryzae strains 70-15 clone 13H23 (gil496853) and clone 72H05 (gil94540563) without detectable positive selection. This result was expected due to the clonality of the host genome and can be considered as negative control. (b) Nht1 and Fot1 which have 52% of nucleotide identity, demonstrate one spot of positive selection located just after the catalytic DDE domain. For this pair, we compared results extracted from SWAKK (blue curve) to those obtained with SWAPSC1.0 (orange curve) and the two curves exhibited the same global pattern. However, since SWAPSC1.0 does not take into account the gaps in the alignment, either in the global or in split analyses, we are unable to compare the same length, which explains the length variation between the orange and blue curves. (c) OPHIO1 from Ophiostoma novo-ulmi subsp. novo-ulmi and OPHIO2 from O. ulmi. These two TEs seem to have undergone positive selection, given the bifurcation to *Fot1* transposons, as demonstrated by the SWAPSC1.0 analysis. In this particular case, the pairwise comparison over the whole length of the OPHIO sequences could be realized due to the exact same length of their open reading frames (ORFs). Two positive spots were detected located in the HTH domain for the first and after the catalytic domain for the second. Concerning the Aft1-like transposons, we compared (d) two copies of Aft1 from two different host genomes, (gil118505101 from A. fumigatus and gil83767101 from A. oryzae) and (e) Aft1 and Afu from A. fumigatus. The results were consistent with those obtained with SWAPSC1.0 and highlighted spots of positive selection at the beginning and end of the catalytic domains, respectively.

IV.6.3. Prokaryotic insertion sequences (IS) have already undergone positive selection

By extension of our hypothesis, we want to analyze insertion sequences (*IS*) extracted from prokaryotic genomes. However, only few interesting candidates, such as *IS5*-like sequences, could be used for the positive selection analysis, many *IS* sequences are too short or have a restricted distribution among species. We chose the *IS5* for the following two reasons: a correct length sequence for SWAPSC1.0 analysis (about 270AA) and a repartition between more than 80 species (Chandler and Mahillon, 2003). The LRTs for the three pairs of models were calculated and, once again, the M8 seemed to be the best model to fit with the dataset (LRT = 4.74, p < 0.095, see Table 7). About 1.33% of the codons were subjected to positive selection, 1.79% to an acceleration of non synonymous

125

mutations and 2.84% of hot spot sites which indicated regions where the number of synonymous and non-synonymous nucleotide substitutions are greater than expected under neutrality and hence hot spots. Figure 37A showed the branches under positive selection (in red). This result indicated that not all the *IS5*-like sequences tested have undergone positive selection although they result from various host genomes. But interestingly, some particular copies such as IS5-like copy (gil150378263) found in the α -proteobacterium Sinorhizobium medicae and the group formed by the copies isolated from Shewanella, Alteromonas and Psychromonas (y-proteobacteria), seem to have undergone positive selection for more extended periods of time than other copies. Because SWAPSC1.0 compares the DNA sequence of each IS5 tested with an ancestral sequence and then compares each ancestral sequence with more ancestral ones to arrive at the root of the tree, we can have a dynamic vision of the positively selected mutations of IS5 in Sinorhizobium medicae (gil150376263). A separate analysis of each branch (Fig. 37B, graph 1 to 4) demonstrated that various spots of positive selection were present and the cumulative analysis of the 4 graphs gives a dynamic vision (Fig. 37B, graph 5). We calculated the mean of ω ratios of each site in order to show the global fluctuation during the four event of speciation (one by branch).





(A) A consensus neighbor-joining phylogenetic tree of *IS5* orthologous sequences based on the coding sequence of the transposase. Each sequence has 798-nt length and represents about 99% of the whole sequence. This phylogeny is largely in agreement with the published phylogeny of *IS5* insertion sequences by Redder et al. (Redder et al., 2003). The bar denotes 0.1% nucleotide divergence and diversity (determined by use of MEGA4.0 software); bootstrap values are indicated below the branches and correspond to 2000 replications. Branches in red in the tree indicate the presence of positively selected regions inside the DNA sequences with ω values higher than 1 (ω >1, p < 0.0001).



Figure 37 (suite). Positively selected site analysis on IS5 in proteobacteria

(B) SWAPSC1.0 analysis of 4 successive branches under positive selection in α -proteobacteria. For each branch (from (1) in dark blue to (4) in shiny blue), we present the graph extracted from the SWAPSC1.0 analysis. The graphs (5) and (6) want to give a dynamic vision of positively selected regions after 4 events of speciation. They show a mean of the ω ratios by site which denotes the presence of three (circle 1 to 3) non random positively selected sites during the evolution of the *IS5* transposase.

Three specific regions (red circles) of positive selected sites were highlighted along the evolution of the *IS5* copies found in α -proteobacteria. Interestingly, the same spots were found for the cumulative analysis of *IS5* copies found in the γ -proteobacteria (Fig. 37B, graph 6).

127

IV.7. Discussion

The present data demonstrated evidence for positive (and negative) selection in different branches of mobile elements in various host genomes. Positive selection analysis requires robust statistical analyses of datasets meeting necessary important criteria. The first criterion is the length of sequences used for the analysis which must exceed 200 amino acids or 600 nucleotides. It seems to be possible to use the different softwares with sequences shorter than 200 amino acids by constituting an important sequence dataset (over 20 sequences) and performing a large number of simulation tests (over 2000). These two points were an obstacle for the analysis of positive selection on TEs and limit the analysis to few TEs contrary to their important proportion in genomes. It is obvious that the use of LRT may have little power to detect the positive selected sites due to an analysis of ω ratios on the whole sequence. If, as we supposed, only few specific regions have evolved differently from the others after positive selection, a combined approach using a sliding window analysis (SWAPSC1.0 and SWAKK) is more appropriate to localize the sites.

Copies inside a host genome can be considered as particular paralogs, and positive selection may play a major role in their divergence (Kondrashov et al., 2002). We think that transposable elements may undergo positive selection on particular regions of their sequences. We want to demonstrate that two ancestral copies of a same TE in two different genomes may face pressures generated by the host genome and consequently, may evolve differently. These ideas are consistent with the recent hypothesis of the 'genetic conflict' between TEs and host genes.

In a first time, we have tested and validated our successive analyses with the *Hermes/Hobo* transposons in insects. These two transposons are considerate as closely related due to the possibility of these TEs to make cross-mobilization. The analyses produced on *Hermes/hobo* transposases highlighted the particular evolution of the 3'-end domain of the transposase (Fig. 33C). The region contains the *hATC* domain, a dimerization domain (Michel et al., 2003) which is essential for enzyme activity and also plays a role in the formation of inactive aggregates (Weil and Kunze, 2000). A mutation in this dimerization domain domain would allow the new transposase to escape inactivation if the new monomer failed

to heterodimerize with the original one, but was still able to form active homodimers. Consequently, the new transposase would contribute disproportionately to total activity because it would escape the aggregation inactivation mechanism of the original transposase. A modification of the *hATC* domain could generate the ability to create new homodimers, which would represent an evolutionary interesting strategy for TEs to fight against host genome pressures. Whether or not this latter domain has an importance for the evolution of *Hermes* sequences, it is clear that it evolved differently in Musca contrary to other species. The Southern hybridization published in 1993 (Atkinson et al., 1993) demonstrated a lower number of *Hermes* copies in the *Florida Drosophila domestica* strain compared to the *Maryland D domestica* strain. Unfortunately, in the absence of the complete *hobo* DNA sequence from the *Florida* strain, we are not able to make a link between a potential modification/absence of the *hATC* and the less significant number of copies which denote reduction of mobility.

Finally, the 3-D visualizations with an analysis window of 10 Angstroms clearly showed an abundance of PS sites in the *Hermes* transposase found in *M. domestica* and *L. cuprina* contrary to *B. tryoni, M. brassicae* and finally *D. melanogaster*. The more a copy was found in two distant species, the more the divergence in term of sequence variations involved positive selection (Fig. 33B). The increase of the divergence was a predicted result but not the accumulation of positively selected sites during *Hermes*-like TE evolution.

In a second time, we have applied these analyses to fungal TEs. A simple analysis of nucleotide compostion of TEs could reveal some interesting features. For example, the striking feature of *OPHIO1* and *OPHIO2* is the divergence in their homology (75.5% of identity) (Daboussi et Capy, 2003). Moreover, these two transposons were isolated respectively from *O. novo-ulmi subsp. novo-ulmi* and *O. ulmi*, the causal agents of the Dutch elm disease (Bouvet et al., 2007a) that have diverged recently (Brasier et al., 2004). This particular feature remains interesting to analyze positive selection site on TEs. Several studies were recently published on positive selection analysis on genes known to undergo high rate of variations, such as the *NBS-LRR* genes involved in plant defense mechanisms (Mondragón-Palomino et al., 2002), envelope proteins in viruses (Joos et al., 2007; Malik

and Henikoff, 2005; Sawyer et al., 2005), or mobile elements (Betrán and Long, 2003; Sawyer and Malik, 2006). However, it is now obvious that some transposons have evolved from common ancestral TEs in the same or diverse species.

Globally, all the fungal TEs tested were subjected to positive selection at one moment or another of their evolution contrary to the *actin* host genome. In the case of the *Fot1-like* TEs (Fig. 36A), two potential regions, one located in the HTH domain and the other at the end of the catalytic domains, presented a divergent evolution from the rest of the transposase sequences. It is possible that the HTH binding site domain may evolve with variations of its associated binding site in TIRs. These variations could be induced by random mutations (or not) but it is more difficult to put forward an assumption on the second domain, due to its unknown function.

Split analyses of fungal TEs reveal interesting points: first, Pot2 paralogs were used to verify the potential influence of the telomere activity on copy of TE present in this region. It has been demonstrated previously that telomerase and sometimes DNA transposons with telomerase activities (such as Het-A and TART (Pardue and DeBaryshe, 2003) have an impact on telomere regeneration. In this case, no positively selected site is present. Secondly, particular *Tc1/mariner* TE copies have undergone positive selection: *Fot1* vs. Nht1, OPHIO1 vs. OPHIO2 and Aft1 vs. Afu. In these three cases, it seems possible that an ancestral copy could evolve and generate two different copies due to the selective pressures of the host genes. Consider the case of OPHIO transposons, with OPHIOancestor, the ancestral copy of *OPHIO1* and *OPHIO2*. According to the phylogeny, an ancestral species became O. ulmi (according to the emergence of the various Dutch elm disease pandemics in Europe (Brasier et al., 2004)). OPHIO_{ancestor} could have evolved in OPHIO2 (only present in O. ulmi - Bouvet et al., 2007) and the copies of OPHIO_{ancestor} present in O. novo-ulmi evolved to become OPHIO1. We can hypothesize that the positive selected regions (HTH and catalytic region) may have evolved differently due to the variations of host pressures between O. ulmi and O. novo-ulmi. Host genomes fight against mobile elements by inducing by mechanisms such as RIP (Galagan and Selker, 2004). We known that RIP intensity seemed different between O. ulmi and O. novo-ulmi (Bouvet et al., 2007a) which perhaps influenced the evolution of OPHIO1 and OPHIO2. If mutations appear in the binding site domain of the transposase (located at the 3' extremity of the TIRs) and change the DNA recognition sequence, the HTH binding site domain could evolve more rapidly than the other regions. An absence of transposase fixation induces an absence of TE mobility. *Aft1*-like analyses reveal also different positively selected regions, either between two copies of the same *Aft1* elements under two different genome pressures (*A. fumigatus* vs. *A. oryzae*).

We wanted to extend our study to prokaryotic systems by studying the insertion sequences (IS). They are indissociable from the host genome evolution simply due to the capacity of *IS* sequences to capture resistance genes to form a bacterial transposon (*Tn*). For example, Tn5 is constituted of two nearly identical inverted copies of IS50 (IS5 family) which bracket 3 genes encoding resistance to kanamycin, bleomycin and streptomycin (Reznikoff, 2003). Then, Tn can be transmitted from one bacterium to another (for review see (Redder et al., 2003). We analyzed various IS5-like sequences to demonstrate the presence of hot spots of positive selection. Globally, it is more difficult to find these spots, perhaps due to the particular status of IS and Tn for the bacteria and the small size of their genomes. Indeed, strong HG pressures on these elements could have fatal repercussions, such as the freezing of the Tn which in turn could influence the bacterial survival rate (absence of resistance transmissions) or more simply, the destabilization of the genome. However, because we analyzed a large spectrum of bacteria (from the α -proteobacteria to the γ proteobacteria), we were able to find branches with positive selection. Interestingly, particular copies have undergone positive selection at each event of speciation (for IS5-like in *Sinorhizobium medicae* for example). The cumulative analysis of the four graphs could give a dynamic vision of positive selected regions along the IS5 evolution and demonstrates the presence of three domains under positive selection in which mutations do not occur randomly. If complete random mutations were applied on the ancestral sequences, no particular domains can be highlighted. However, due to the current lack of information, we cannot propose a hypothesis on the role of each site but they certainly have an impact on the transposase either on the binding site or the catalytic domains. Nevertheless, the fact that we found the same three spots after analyzing the *Shewanella* spp. and other γ - proteobacteria suggests that these spots have a particular status inside the *IS5* transposase domains.

Obviously, many TEs could be added to these analyses, including retroelements, rolling circle elements or other mobile elements such as transcription factors. We performed a complementary analysis on 16 *R2R3-Myb* transcription factor orthologous genes present in conifer genomes of *Picea glauca* and *Pinus taeda*. In spite of the limited length of the sequences (549nt), we were able to prove that particular *Myb* sequences have undergone positive selection on *R2R3* fixation domains (Bouvet, unpubl. data). However, this result is in agreement with those that provide evidence of particular mechanisms for evolutionary divergence in R2R3 *Myb* in maize (Dias et al. 2003).

The present study revealed the presence of punctual positively selected regions inside transposase sequences in both prokaryotic and eukaryotic genomes and we think that this evolution could result in part from host pressures on transposon DNA sequences. This idea is confirmed by the particular evolution of the *hATC* domain of the *Hermes* transposase and enriched with the 3-D visualization. Moreover, the prokaryotic TE analysis demonstrated the non-random distribution of PS mutations and suggests the occurrence of specific pressures on particular domains of the transposase.

However, some experiments will have to be designed to specifically address in more detail the reciprocal influences between TEs and their respective HGs. This preliminary study of positive selected regions certainly needs biological confirmations, as for example, the clonal multiplication of the AST27 in levuriforme stage. AST27 is interesting for many reasons, but one of them seems to be more important: the presence of the introgressed copy of *OPHIO2, OPHIO2-int*. With this particularity, AST27 could become a interesting model for understanding the evolution of an active transposon after being introgressed and, in parallel, could be a good candidate to evaluate the power of genomic pressures on this kind of mobile element.

Chapitre V : Conclusions et schéma de synthèse

La maladie hollandaise de l'orme a eu un impact très important sur les espèces forestières ligneuses du genre Ulmus et reste à ce jour, après la brûlure du châtaignier au début des années 1900, la pandémie la plus importante en milieu forestier (Wingfield, 2006). À ce titre, cette maladie a été étudiée depuis les années 1970, tant au niveau phénotypique que génotypique. La création du projet « Genome Ophiostoma», en 2004, donna lieu à de plus amples recherches sur la fonctionnalité des génomes d'Ophiostoma ulmi, et d'O. novo-ulmi, les agents responsables de cette maladie. Les objectifs à court terme étaient au nombre de quatre, essentiellement basés sur l'expression et l'architecture des gènes ainsi que sur la production d'une carte génétique saturée. C'est dans l'optique de comprendre l'architecture génomique des champignons de la maladie hollandaise de l'orme, que nous avons intégré un volet sur les éléments mobiles. Les 50 dernières années ont vu « l'avènement » de l'étude de ces éléments comme instigateurs de cassures chromosomiques, recombinaisons génomiques, inactivateurs ou duplicateurs ainsi sur/sous expresseurs de gènes. Ces éléments sont maintenant considérés, à juste titre, comme l'un des moteurs-clés de l'évolution des génomes. Enfin, au fur et à mesure de leurs études, ils sont devenus de puissants outils tant pour l'analyse de populations que pour la mutagénèse. Il nous a donc semblé intéressant d'analyser l'impact des transposons à ADN sur les génomes des Ophiostoma impliqués dans la M.H.O., potentiellement les plus représentés chez les champignons phytopathogènes. Cette thèse avait trois objectifs majeurs :

1- De découvrir et caractériser des <u>transposons à ADN</u> (*OPHIO transposons*) ainsi que d'étudier leurs <u>caractéristiques propres</u> et leur <u>répartition</u> (**Chapitre II**). Il est évident que ces transposons ne représentent que le sommet de l'iceberg. Le séquençage du génome d'*Ophiostoma novo-ulmi* subsp. *novo-ulmi* permettrait de quantifier la proportion exacte de ce type d'éléments dans le génome. D'une manière générale, ce tandem *Ophiostoma* sp./transposons à ADN s'est révélé scientifiquement intéressant à différents niveaux. Il nous a permis notamment de vérifier le paradigme du premier chapitre et d'enrichir la super-famille *Tc1/mariner* décrite par Daboussi et Capy en 2003. Bien que ces éléments ne

133
possèdent que 50 % d'identité avec les autres transposons retrouvés chez

134

Fusarium oxysporum, Magnaporthe oryzae, Botrytis cinerea, etc., des zones conservées se retrouvent au sein de domaines particuliers tels que les motifs CCCCT ou CCACCC des TIR ou encore le motif protéique LLILDGHGSH chez l'ensemble des transposases analysées au cours de cette thèse. Ce type de caractéristiques nous permet d'émettre l'hypothèse de la présence d'un élément ancestral commun à l'ensemble de ces mycètes.

- 2- De décrire et de démontrer la <u>potentielle mobilité de ces éléments OPHIO</u> (Chapitre III). La capacité de se déplacer de façon autonome de ces éléments est l'un des principaux pouvoirs évolutifs exercés sur les génomes. Dans ce quatrième volet, diverses questions étaient abordées, telles que la vérification de la mobilité propre de ces éléments, de l'autonomie et de l'induction de la transposition via des stress abiotiques. Le principal objectif a été de vérifier si un stress environnemental pouvait avoir un impact sur la mobilité d'un transposon à ADN. Derrière cette dernière question, la notion plus générale de l'impact des stress environnementaux sur l'architecture des génomes était abordée.
- 3- D'illustrer <u>l'impact des génomes hôtes sur les éléments mobiles</u> (Chapitre IV). Les transposons OPHIO1 et OPHIO2, de par leurs caractéristiques structurales, ont servi de point de départ à une étude de sélection positive sur des domaines particuliers des transposases. Nous avons voulu démontrer qu'en tant que marqueurs non mendéliens, les éléments mobiles des génomes font l'objet de pressions sélectives différentes de gènes constitutifs. Il est clairement ressorti de cette étude que des régions particulières, non aléatoirement choisies, subissent une évolution différente du reste de la séquence en fonction de leur génome hôte. Le transposon OPHIO3 a également subi l'influence du génome hôte, par l'intermédiaire de mutations RIP. L'analyse des séquences d'OPHIO3 nous a permis de concevoir une méthode de visualisation simple de ces dernières (Chapitre II).



Figure 38. Comparaison des transposons OPHIO1 et OPHIO2 par visualisation CTS

La visualisation a été effectuée à partir de la méthode de *Cumulative Transition Score (CTS)* développée dans le second chapitre de cette thèse. On a codé ici toutes les Transitions (C :G $\leftarrow \rightarrow$ T :A) par +1 et les transversions par -1. Globalement, 56 % des mutations sont des Transitions avec 51 % des mutations *OPHIO1* : C-G \rightarrow *OPHIO2* : T-A, ce qui ne permet pas de conclure à une éventuelle évolution de *OPHIO1* vers *OPHIO2* (et inversement) due aux mutations engendrées par le phénomène de *RIP*. Bien que le rapport Transitions/transversions soit équilibré, on dénote l'apparition de zones dans lesquelles les Transitions (rectangle rouge) ou les transversions (rectangle vert) sont plus prépondérantes. C'est, par exemple, le cas des TIR (5' ou 3' matérialisés par les deux ∇) où l'on observe une majorité de Transitions.

<u>Avant-propos</u> : De nombreux mécanismes ont été mis en place par les génomes afin de limiter la prolifération d'éléments mobiles ubiquistes et sont indiqués dans l'annexe

V.1. SYNTHESE # 1 = impacts des transposons OPHIO sur le génome d'Ophiostoma sp.

Les transposons, comme nous l'avons décrit à maintes reprises dans cette thèse, ont un rôle dans la structure des génomes et cela à différents niveaux : (1) par leur multiplication, générant des blocs importants influençant la taille du génome hôte (Bowen et al., 2002), (2) par les recombinaisons qu'ils peuvent créer, modulant le génome (Zhang et al., 1999), (3) par la sur/sous expression due aux positionnements de copies en amont de gènes exprimés ou encore (4) par l'intégration de séquences régulatrices au sein de ces copies (Jordan et al., 2003). Comme nous venons de le voir précédemment, la mobilité est l'un des impacts majeurs des transposons à ADN. Il faut toutefois distinguer deux types de mobilité : intragénomique et inter-génomique. La mobilité intra-génomique est mise en exergue par des duplications d'éléments et par la présence/absence de loci. Les transferts horizontaux de transposons ou leurs introgressions forment la majeure partie des évènements de mobilité inter-génomique. Ils ont été démontrés de nombreuses fois (Rosewich et al., 2000 ; Lampe et al., 2003 ; Lohe et al., 1995), et jouent à juste titre, un rôle important dans l'architecture des génomes.

Chez les champignons de la maladie hollandaise de l'orme, plusieurs cas de transferts horizontaux sont actuellement publiés, notamment par l'introgression de gènes de reconnaissance *Mat-1/Mat-2* sexuelle (Paoletti, 2006a et 2006b), gène *Vic*, gène *Pat1* (Et Touil et al., 1999) de pathogénécité ou encore copie d'*OPHIO2* via la présence, dans le milieu naturel, d'hybrides interspécifiques (Brasier, 2000). Un transposon, tel que *OPHIO2*, devient alors un outil intéressant pour comprendre la dynamique de ces éléments au sein des espèces.

V.2. SYNTHESE # 2 = impacts du génome d'Ophiostoma sp. sur les transposons OPHIO

V.2.1. Les mutations RIP chez les champignons de la M.H.O.

L'article du **chapitre II** décrit ce mécanisme de *RIP* chez les champignons de la maladie hollandaise de l'orme : la découverte de mutations sur *OPHIO3* nous a permis de conclure que ce mécanisme avait été présent chez *Ophiostoma* spp. On a observé que les patrons de mutations *RIP*, identiques chez la sous-espèce européenne (EAN), affichaient des variations chez l'espèce *O. ulmi*. Les différences ainsi mises en évidence entre espèces et sous-espèces semblent se corréler parfaitement à l'ordre d'apparition de l'agent de la M.H.O.

RIP intensity OPHIO3 O. ulmi >> RIP intensity OPHIO3 O. novo-ulmi EAN >> RIP intensity OPHIO3 O. novo-ulmi NAN.

Deux scénarios sont possibles : soit le *RIP* chez *O. ulmi* est plus ancien que celui chez *O. novo-ulmi* subsp. *novo-ulmi* et que chez *O. novo-ulmi* subsp. *americana*, soit le mécanisme de *RIP* est plus « actif » chez *O. ulmi* que chez *O. novo-ulmi* spp. Si l'on émet l'hypothèse que le même nombre de mutations est généré à chaque activation du mécanisme et que l'on tient compte du fait qu'il y ait de nombreuses formes variables d'*OPHIO3* chez *O. ulmi*, alors l'ancestralité de *O. ulmi* ne fait aucun doute.

La visualisation de ces variations par *CTS* a démontré qu'il y avait une accumulation de mutations au niveau de la zone centrale de la transposase, plus particulièrement dans le domaine DD(35)E (Fig. 38). La soustraction des courbes indique clairement une différence d'intensité dans cette région. L'inactivation de cette région pourrait être la meilleure option pour immobiliser rapidement un élément mobile



Figure 39. Courbe CTS d'OPHIO1 par rapport à OPHIO2 (en rouge) et inversement (en vert). vert)

Nous avons voulu démontrer par cette étude si l'un des deux transposons avait pu évoluer de l'autre par l'intermédiaire du mécanisme de *RIP*. (A) Pour cela nous avons pris *OPHIO1* comme référence et lui avons comparé *OPHIO2* comme s'il s'agissait d'une copie d'*OPHIO1* RIPpée (courbe rouge) et inversement (courbe verte). (B) La courbe soustractive, en bleu, est la résultante de la soustraction des valeurs de la courbe verte. En montrant une alternance (sinusoïdale), cette courbe démontre qu'aucun des deux TE n'a évolué de l'autre. Si tel avait été le cas, nous aurions dû avoir une droite ($\pm ax + b$)

Enfin, de cette analyse, il ressort que la région terminale de la région codante semble fortement soumise à des transitions. Le mécanisme de RIP pourrait-il alors avoir eu pour effet de générer des mutations qui auraient alors été positivement sélectionnées dans cette zone précise ? La figure 39 présente une analyse de soustraction de courbes CTS (courbe B en bleu). Dans cette dernière analyse, nous avons soustrait la courbe CTS verte obtenue en comparant *OPHIO1* (référence) versus *OPHIO2* (transition C :T +1, transition G :A -1) à la courbe rouge obtenue en comparant *OPHIO2* (référence) versus *OPHIO1* (transition T :C +1, transversion A :G -1). Le but étant de mettre en évidence une prépondérance de transitions de l'un des transposons vers l'autre et donc, de voir si le RIP (responsable de ce type de transition) pouvait avoir eu un impact sur l'un ou l'autre des transposons et généré le second. La relative alternance mise en évidence par cette figure (comparable à une sinusoïde) nous indique qu'il ne semble pas y avoir de région pouvant provenir de l'un des transposons qui, par évolution due aux mutations RIP, aurait généré le second. Toutefois, l'extrémité 3' (zone indiquée en vert clair) semble démontrer qu'il y a plus de transitions de *OPHIO1* vers *OPHIO2*.

V.2.2. Les zones de sélection positive au sein des transposases

Le **chapitre IV** a eu pour but de démontrer qu'en tant que marqueurs non mendéliens, les éléments mobiles des génomes font l'objet de pressions sélectives différentes de gènes hôtes. Cette problématique, née de l'observation des séquences d'*OPHIO1* et *OPHIO2*, s'inscrit parfaitement dans le cadre plus général de la théorie de « conflits génétiques » qui opposent les éléments mobiles et les gènes hôtes (Sawyer et Malik, 2006). De cette étude sont clairement ressortis les points suivants :

(1) des régions particulières des transposases subissent une évolution différente du reste de la séquence. Cette évolution différencielle varie également de celle des gènes hôtes « constitutifs » tel que l'actine. De l'étude des transposons *Fot1*-like, il semblerait, d'une part, que des régions positivement sélectionnées soient préférentiellement localisées au sein des domaines de fixations et des domaines catalytiques. Nous retrouvons ces zones de sélection au niveau de copies qui, par le passé, auraient été introduites d'une espèce à une autre, par transferts horizontaux ou introgressions (cas pour OPHIO1 et OPHIO2 [introgression] et Nht1 et Fot1 [transfert horizontal]). D'autre part, l'étude réalisée sur le transposon Hermes illustre parfaitement que des régions particulières ayant un rôle précis au sein de la transposase, pourraient être ciblées. C'est le cas du domaine hATC 3'-terminal de la transposase Hermes. Il semblerait que ce domaine ait fait l'objet d'une sélection positive entre différentes espèces d'insectes, la visualisation en 3-dimensions des différents domaines subissant des effets sélectifs, l'illustre parfaitement. Des mutations au sein de ce domaine permettraient, si elles ne sont pas dégénératrices, d'augmenter la capacité de la transposase à former des hétérodimères de protéines. Plus il existera d'hétérodimères actifs et plus la transposase aura la possibilité d'échapper aux mécanismes de « silencing » induits par le génome hôte (Weil et Kunze, 2000). (2) Nous avons enfin remarqué que plus les espèces comparées ont divergé tôt dans l'histoire évolutive du genre et plus les zones de sélection positive étaient présentes. (3) L'étude de l'IS5 a également permis de mettre en évidence ces mêmes résultats mais de plus, a permis, par l'étude cumulative de quatre branches successives ayant montré des zones de sélection positive, d'établir que ces régions ne semblaient pas être ciblées aléatoirement lors de l'évolution. Les mécanismes générant la sélection positive sur ces régions particulières semblent agir de façon précise sur les mêmes régions.

V.3. SCHEMA de SYNTHÈSE

La présentation du schéma de synthèse ci-contre va permettre au lecteur de cette thèse d'avoir une vision d'ensemble des diverses interactions entre les éléments mobiles et le génome hôte. Dans ce schéma sont repris les divers mécanismes les plus importants ainsi que l'idée générale émise par N. Fedorov qui sous-entend que les transposons et leurs génomes hôtes évoluent en opportunistes réciproques plutôt que vers une relation pathogène/hôte (Fig. 40).



Figure 40. Schéma de synthèse

Chapitre VI : Perspectives de ces recherches – projets

VI.1. Perspectives - projet #1 = les transposons OPHIO, outils d'analyses

- 1- Origine des transposons OPHIO : une des premières questions directement reliées aux transposons de la M.H.O. concerne l'origine de ces éléments et peut-être indirectement l'origine de la maladie hollandaise. Diverses hypothèses ont été émises dans le but d'expliquer l'apparition de l'espèce O. novo-ulmi. et plusieurs scénarios, dont celui de l'hybridation entre une souche d'O. ulmi et une souche d'une autre espèce d'Ophiostoma, sont envisageables. Cette dernière espèce pourrait potentiellement être O. quercus ou O. piceae. Ces deux espèces ont été retenues d'une part, pour leur proximité phylogénétique par rapport aux champignons de la M.H.O. et d'autre part, pour la relative identité de leurs gènes. Il serait alors intéressant de voir si les transposons OPHIO auraient une origine de type O. quercus/O. piceae et auraient été acquis après transfert horizontal via des hybrides interspécifiques, à un moment donné de l'histoire des Ophiostomatoïdes. L'enjeu serait de mettre en évidence la présence d'un des trois transposons au sein de populations d'O. quercus et O. picea isolées dans des zones sympatriques avec les champignons de la M.H.O. OPHIO3 serait potentiellement le meilleur candidat, sa présence ayant été démontrée chez les deux espèces présentées en Europe et en Amérique du Nord. Il semble que l'analyse de populations européennes serait plus appropriée pour mettre en évidence une telle présence.
- 2- <u>Variation de l'expression d'un gène potentiellement relié à la pathogénécité</u> : une des copies d'*OPHIO2* est située en amont d'un gène codant une sérine/thréonine protéine kinase, protéine pouvant être impliquée dans le pouvoir pathogène des champignons (Xu, 2000 et pour article de revue Dickman et Yarden, 1999). Il serait intéressant de voir :

(a) si ce gène a été dupliqué chez certaines souches d'O. *ulmi* via des recombinaisons ectopiques *d'OPHIO2*, cette duplication pouvant engendrer soit

une surexpression du gène, soit un effet contraire si la duplication est en anti-sens (co-suppression naturelle).

(b) si cette copie présente en amont pourrait jouer le rôle de sur ou sous expresseur de ce gène.

3- Le mécanisme de RIP : est-il encore actif ou non chez l'une des trois espèces ? En effet, l'analyse des empreintes de RIP découvertes et analysées dans le 2^e chapitre démontre que ce mécanisme a été actif par le passé, mais ne démontre pas qu'il est encore actuellement actif. L'analyse de copies présentes en F1, F2 et F3 du croisement entre AST27 et H327 ne montre pas d'apparition de mutations RIP. Cependant, ce croisement reste intra-spécifique ce qui ne permet pas de confirmer que le mécanisme n'est plus actif chez O. ulmi ni chez O. himal-ulmi. L'analyse de nouveaux croisements pourrait s'avérer intéressante combinée à l'analyse de l'expression de méthyltransfèrases et déaminases présente dans la banque d'ESTs d'O. novo-ulmi.

VI.2. Perspectives - projet #2 = la répartition des transposons OPHIO au sein de populations naturelles d'Ophiostoma spp.

Nous pourrions nous servir des TE comme marqueurs non mendéliens afin de voir si un lien pouvait être fait entre la répartition géographique et la variation du nombre de copies de ces trois éléments. Nous avons également voulu répondre aux questions suivantes :

- <u>Concernant les trois éléments</u>, donnent-ils une vision globale corrélable avec des marqueurs génomiques ?
- <u>Volet *OPHIO1*</u> : Existe-t-il une variabilité au sein des souches d'*O. novo-ulmi* subsp. *novo-ulmi* et est-il réellement spécifique de cette sous espèce ? Peut-on mettre en évidence d'autres souches ayant introgressé cet élément ?
- <u>Volet *OPHIO2*</u> : Peut-on voir une expansion du nombre de copies dans certaines souches particulières d'*O. ulmi* ou un profil corrélable avec un isolement géographie ?
- <u>Volet *OPHIO3*</u> : Les problématiques sont différentes : Milgroom et Brasier, en 1997, ont décrit que le développement d'une nouvelle pandémie après introduction pouvait s'effectuer en trois phases : (1) multiplication de la souche introduite par clonalité

(patrimoine génétique identique), (2) compétition entre cette nouvelle population et les populations endémiques générant des hybrides (si possible) ou émergence de nouveaux allèles due à des facteurs internes (exemple des mycovirus) et (3) expansion de nouvelles populations hétérogènes provenant de la souche introduite au niveau des zones d'expansion de la pandémie (patrimoine génétique plus hétérogène – Brasier et al.,). *OPHIO3* pourrait nous permettre de vérifier ou non ce type de structurations génomiques avec toutefois une vision non mendélienne. Les problématiques décrites cidessus pourraient être « doublées » par une analyse, au sein des mêmes populations, de marqueurs nucléaires soumis à fortes variations dues aux pressions sélectives : les gènes de compatibilité végétative (*Vic* genes). Deux de ces gènes *Vic* ont été clonés et séquencés (Bouvet, non publié).

VI.3. Perspectives - projet #3 = la mobilité des transposons OPHIO

Le chapitre IV de cette thèse a initié l'étude de la mobilité des transposons OPHIO. Pour cela, nous avons voulu corréler la mobilité des ces éléments avec une possible induction via des stress biotiques ou abiotiques, pouvant se retrouver dans le milieu naturel. Cette étude avait pour but de valider un possible rôle actif de l'environnement dans la structuration des génomes, en activant la mobilité des TE. Ces résultats pourraient être mis en parallèle avec les modifications épigénétiques qui sont, par définition, des modifications de l'expression des gènes qui sont transmissibles lors de la mitose et/ou la méiose. Ils ne modifient toutefois pas la séquence de l'ADN. Un transposon peut modifier l'expression de gènes (rupture de l'expression si insertion au sein de la région codante ou encore sur/sous expression si insertion dans la région promotrice du gène), tout en étant transmis verticalement (si la copie insérée n'est plus mobile par la suite). Ce troisième projet s'intègrerait alors dans un volet global pouvant être intitulé : « Transposons OPHIO et leurs impacts épigénétiques chez Ophiostoma sp. ». Cette thématique est directement en accord avec les dernières synthèses de McClintock en 1984 et sa théorie de chocs génomiques. Très récemment, Slotkin et Martienssen ont publié un article de synthèse concernant le possible lien entre « épigénétisme et TE » (Slotkin et Martienssen, 2007) et plus précisément entre stress génomique, élément de transposition et réponse adaptative. En

145

effet, les TE sont « éteints ou allumés » par des stress environnementaux. Dans les deux cas (*silencing* par méthylation ou réactivation via des stress environnementaux), l'activité de ces TE produirait une diversité génétique accrue qui, au terme d'une stratégie adaptative à long terme, serait bénéfique pour l'évolution des espèces en réponse à des stress spécifiques.

<u>Une première thématique</u> serait de générer des survivants à partir de l'hybride interspécifique AST27 via des stress <u>successifs</u> aux UVs et au froid (4°C). Le but initial serait de mettre en évidence cette fois-ci, des excisions avec réinsertions et de ce fait, une augmentation du nombre de copies après des stress successifs. Les survivants stressés seraient à nouveau soumis au même stress, et cela plusieurs fois de suite. Ceci est à mettre en lien avec les résultats mis en évidence dans le chapitre IV qui démontraient l'induction de la mobilité d'une copie. Les avantages d'*Ophiostoma* sp. pour ce modèle seraient son taux de croissance très rapide (2-3 jours après induction du stress) ainsi que sa capacité de ces derniers à générer rapidement de la biomasse. Ces deux caractéristiques font d'*Ophiostoma* sp. un système eucaryote possédant une productivité proche des bactéries (système procaryote). On pourrait alors mettre en lien l'introgression de copies via des hybrides interspécifiques et la réponse adaptative de ces derniers à des stress via l'acquisition de TE afin de décrire un système invasif de colonisation d'espèces proches.

<u>Une seconde thématique</u> serait de quantifier le temps de réponse (par qPCR) de l'expression de la transposase d'*OPHIO1* et *OPHIO2* après induction par un stress (choc thermique pour *OPHIO1* et choc au UV et croissance à 4°C pour *OPHIO2*).

- Pour *OPHIO1*, la comparaison entre la quantité de transposase et la quantité de protéines de choc thermique, après stress à 55°C (présentes dans la banque d'expression développer chez H327) pourrait être effectuée facilement.
- Pour OPHIO2, le système est plus complexe car aucun promoteur n'a pu être mis en évidence dans les régions ITR-5' ou UTR-5' de ce transposon. Toutefois, une étude intéressante serait de quantifier, en parallèle de

l'expression de la transposase *OPHIO2*, celle de facteurs de transcription ou encore des protéines de réponses à des stress sévères (hors protéines *Hsp*).

Cette dernière hypothèse a déjà été décrite : « The correlated response of the transposable elements could be due to the presence of fixation sites of transcription activators in their regulatory regions » (Capy et al. 1999) et démontrée pour la région *LTR*-U3 chez *Tnt1*, rétrotransposon activé après infection pathogène mais également après l'action de multiples stress biotiques et abiotiques (Grandbastien et al., 1997).

VI.4. Épilogue

Les opportunités de projets de recherche sur le modèle *Ophiostoma sp.*/transposons OPHIO sont, comme énumérées précédemment, multiples et variées. Il est fort probable que cela tienne essentiellement à trois caractéristiques majeures de ce modèle :

(1) les 50 dernières années de recherches entreprises sur les champignons de la maladie hollandaise de l'orme nous permettent actuellement d'avoir de solides connaissances, tant physiologiques, génétiques que techniques. (2) La seconde caractéristique est directement liée aux propriétés biologiques des mycètes : Ophiostoma ulmi et O. novo-ulmi sp. possèdent d'incroyables capacités à croître, à générer de grandes quantités de biomasse et à supporter de nombreux stress biotiques ou abiotiques, parfois très sévères. L'ensemble de ces caractéristiques en font un outil biologique très puissant pour un système eucaryote, au même titre que les bactéries le sont pour un système procaryote. (3) Enfin, la troisième caractéristique et non la moindre, met en évidence, d'une part, des similitudes génomiques au sein des génomes de ces champignons mais également la possibilité de générer des hybrides interspécifiques fertiles ou de les trouver en nature et, d'autre part, la présence d'une barrière de reproduction. Ces caractéristiques permettent suggérer que ces mycètes sont en cours de spéciation. Cette dernière remarque me semble importante du fait qu'il n'existe pas, actuellement, de modèle biologique eucaryote simple d'approche et relativement compétitif, qui puisse permettre l'étude « en temps réel » d'une spéciation. Il va être alors intéressant de continuer l'étude en mixant les impacts réciproques des transposons et du génome hôte avec la dimension évolutive que la spéciation peut engendrer. Le séquençage d'Ophiostoma novo-ulmi sp. novo-ulmi serait un atout majeur pour la poursuite de l'étude de ce système. Il nous permettrait en effet de faire le point sur les différents types d'éléments mobiles existant au sein de ces génomes : leur répartition, impacts, copies cryptiques etc. Enfin, une dernière perspective serait, à plus long terme, le séquençage d'Ophiostoma ulmi. Il pourrait permettre l'étude des différents facteurs impliqués dans une spéciation. Ces derniers, qu'ils soient des gènes ou îlots de gènes, pourraientt être susceptibles de se différencier dans le but de générer la nouvelle espèce ou de généere des forces impliquées dans cette différenciation tels que les éléments de transposition.

148

Bibliographie

- Achaz, G., Etude de la dynamique des génomes : les répétitions intrachromosomiques. Thèse de Doctorat Spécialité Génétique. Université Pierre et Marie Curie, Paris, 2002, pp. 230.
- Agrios, G. N., 2005. Agrios, G.N. Plant Pathology, 5th. Edition. Elsevier Academic Press: San Diego, USA. pp. 305.
- Aleshkin, G. I., Kadzhaev, K. V., Markov, A. P., 1998. High and low UV-dose responses in SOS-induction of the precise excision of transposons *Tn1*, *Tn5* and *Tn10* in *Escherichia coli*. Mutat. Res. 401, 179–191.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.
- Anaya, N., Roncero, M. I., 1996. Stress induced rearrangement of *Fusarium* retrotransposon sequences. Mol. Gen. Genet. 253, 89–94.
- Andaya, V. C., Tai, T. H., 2006. Fine mapping of the *qCTS12* locus, a major QTL for seedling cold tolerance in rice. Theor. Appl. Genet. 113, 467-475.
- Anisimova, M., Bielawski, J. P., Yang, Z., 2002. Accuracy and power of bayes prediction of amino acid sites under positive selection. Mol. Biol. Evol. 19, 950-958.
- Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst. Biol. 55, 539-552.
- Atkinson, P. W., Warren, W. D., O'Brochta, D. A., 1993. The *hobo* transposable element of Drosophila can be cross-mobilized in houseflies and excises like the *Ac* element of maize. Proc. Natl. Acad. Sci. U. S. A. 90, 9693-9697.
- Baltimore, D., 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature 226, 1209-1211.
- Bennetzen, J. L., SanMiguel, P., Chen, M., Tikhonov, A., Francki, M., Avramova, Z., 1998. Grass genomes. Proc. Natl. Acad. Sci. U. S. A. 5, 1975-1978.
- Bensaadi-Merchermek, N., Salvado, J. C., Mouches, C., 1994. Mosquito transposable elements. Genetica 93, 139-148.
- Bernier, L., Hubbes, M., 1990. Mutations in *Ophiostoma ulmi* induced by N-methyl-N'nitro-N-nitrosoguanidine. Can. J. Bot. 68, 225-231.

- Bernier, L., Breuil, C., Hintz, W. E., Horgen, P. A., Jacobi, V., Dufour, J., Aoun, M., Bouvet, G., Kim, S. H., Diguistini, S., Tanguay, P., Eades, P., Burgess, S., de la Bastide, P., Pinchback, M., Tadesse, Y., 2004. The Canadian *Ophiostoma* genome project. Invest. Agrar.: Sist. Recur. For. 13, 105-117
- Betrán, E., Long, M., 2003. *Dntf-2r*, a young Drosophila retroposed gene with specific male expression under positive Darwinian selection. Genetics 164, 977-988.
- Beuzon, C. R., Chessa, D., Casadesus, J., 2004. *IS200*: an old and still bacterial transposon. Int. Microbiol. 7, 3-12
- Bingham, P. M., Kidwell, G. M., Rubin, G. M., 1982. The molecular basis of *P-M* hybrid dysgenesis: the role of the *P* element, a P-strain-specific transposon family. Cell 29, 995-1004.
- Bouhouche, K., Zickler, D., Debuchy, R., Arnaise, S., 2004. Altering a gene involved in nuclear distribution increases the repeat-induced point mutation process in the fungus *Podospora anserina*. Genetics 167, 151–159.
- Bouvet, G. F., Jacobi, V., Bernier, L., 2007a. Characterization of three DNA transposons in the Dutch elm disease fungi and evidence of repeat-induced point (RIP) mutations. Fung. Genet. Biol. 44, 430-443.
- Bouvet, G. F., Pourde, K. V., Jacobi, V., Bernier, L., 2007b. Stress-induced mobility of OPHIO1 and OPHIO2, DNA transposons of the Dutch elm disease fungi. Fung. Genet. Biol. (accepted for publication in Fungal Genet. Biol.)
- Bowden C. G., Smalley E., Guries R. P., Hubbes M., Temple B., Horgen P. A., 1996. Lack of association between cerato-ulmin production and virulence in *Ophiostoma novo-ulmi*. Mol. Plant Microbe Int. 9(7):556-64.
- Bowden C. G., Hintz W. E., Jeng R., Hubbes M., Horgen P. A., 1994 Isolation and characterization of the cerato-ulmin toxin gene of the Dutch elm disease pathogen, *Ophiostoma ulmi*. Curr. Genet. 25(4):323-9.
- Bowen, N. J., Jordan, I. K., 2002. Transposable elements and the evolution of eukaryotic complexity. Curr. Issues Mol. Biol. 3, 65-76.
- Brasier, C. M., 1986. A comparison of pathogenicity and cultural characteristics in the EAN and NAN aggressive sub-groups of *Ophiostoma ulmi*. Trans. Brit. Mycol. Soc. 87, 1-13.

- Brasier, C. M., 1987. Some genetical aspects of necrotrophy with special reference to *Ophiostoma ulmi*. In: P. R. Day, G. J. Jellis, (Eds.), Genetics and plant pathogenesis. Blackwell Scientific Publications, Oxford. pp. 297-310.
- Brasier, C. M., 1988. Rapid changes in genetic structure of epidemic populations of *Ophiostoma ulmi*. Nature 322, 538-541.
- Brasier, C. M., Bates, M. R., Charter, N. W., Buck, K. W., 1993. DNA polymorphism, perithecial size and molecular aspects of D factors in *Ophiostoma ulmi* and *O. novoulmi*. In: Sticklen, M.B., Sherald, J. L., (Eds.), Dutch elm disease research. Cellular and molecular approaches. Springer-Verlag, New York, pp. 308-321.
- Brasier, C. M., Mehrotra, M. D., 1995. *Ophiostoma hima-ulmi* sp. *nov.*, a new species of Dutch elm disease fungus endemic to the Himalayas. Mycol. Res. 2, 205-215.
- Brasier, C. M., 1996. Low genetic diversity of the *Ophiostoma novo-ulmi* population in North America. Mycologia 88, 951-964.
- Brasier, C. M., Kirk, S., Pipe, N., Buck, K. W., 1998. Rare interspecific hybrids in natural populations of the dutch elm disease pathogens *Ophiostoma ulmi* and *O. novo-ulmi*. Mycol. Res. 102, 45-57.
- Brasier, C., 2000. The rise of the hybrid fungi. Nature 405, 134-135.
- Brasier, C. M., Kirk, S. A., 2000. Survival of clones of NAN *Ophiostoma novo-ulmi* around its probable centre of appearance in North America. Mycol. Res. 104, 1322-1332.
- Brasier, C. M., Kirk, S. A., 2001. Designation of the EAN and NAN races of *Ophiostoma novo-ulmi* as subspecies. Mycol. Res. 105, 547-554.
- Brasier, C. M., Buck, K., Paoletti, M., Crawford, L., Kirk, S., 2004. Molecular analysis of evolutionary changes in populations of *Ophiostoma novo-ulmi*. Invest. Agrar.: Sist. Recur. For. 13, 93-103
- Brookfield, J. F. Y., 2002. Evolutionary genetics: mobile DNAs as sources of adaptive change? Curr. Biol. 14, R344-R345.
- Brown, T. A., 2006. Genomes. 3rd edition. Garland Science. University of Manchester, UK. 750 pages.
- Brown P. O., Bowerman B., Varmus H. E., Bishop J. M. 1989. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. Proc. Natl. Acad. Sci. U. S. A. 86(8):2525-2529

- Capy, P., Gasperi, G., Biemont, C., Bazin, C., 2000. Stress and transposable elements: coevolution or useful parasites? Heredity 85, 101-106.
- Catteruccia, F., Nolan, T., Blass, C., Müller, H.-M., Crisanti, A., Kafatos, F. C., Loukeris, T. G., 2000. Toward *Anopheles* transformation: *Minos* element activity in anopheline cells and embryos. Proc. Natl. Acad. Sci. U. S. A. 97, 2157–2162.
- Cech, T. R., 1990. Nobel lecture. Self-splicing and enzymatic activity of an intervening sequence RNA from Tetrahymena. Biosci. Rep. 10, 239-261.
- Chalvet, F., Grimaldi, C., Kaper, F., Langin, T., Daboussi, M. J., 2003. Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. Mol. Biol. Evol. 20, 1362-1375.
- Chandler, M., Mahillon, J., 2003. Insertion sequences revisited. In: N. L. e. a. Craig, (Ed.), Mobile DNA II. ASM press, Washington D. C., pp. 305-366.
- Chang, S., Puryear, J., Cairney, J., 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11, 113-116.
- Chicas, A., Cognoni, C., Macina, G., 2004. RNAi-dependent and RNAi-independent mechanisms contribute to the silencing of RIPed sequences in *Neurospora crassa*. Nucleic Acids Res. 32, 4237-4243.
- Clutterbuck, A. J., 2004. MATE transposable elements in *Aspergillus nidulans*: evidence of repeat-induced point mutation. Fungal Genet. Biol. 41, 308-316.
- Coen, E. S., Robbins, T. P., Almeida, J., 1989. Consequences and mechanisms of transposition in *Antirrhinum majus*. In: D. E. Berg, M. M. Howe, (Eds.), Mobile DNA. American Society for Microbiology, Washington DC, pp. 413-436.
- Comfort, N. C., 2001. From controlling elements to transposons: Barbara McClintock and the nobel prize. Trends Genet. 17, 475-478.
- Coy, M. R., Tu, Z., 2007. Genomic and evolutionary analyses of *Tango* transposons in *Aedes aegypti*, *Anopheles gambiae* and other mosquito species. Insect Mol. Biol. 16, 411-421.
- Craig, N. L., 2002. Mobile DNA: an introduction. In: N. L. Craig, et al., (Eds.), Mobile DNAII. ASM press, Washington, D. C., pp. 3-11.
- Cutler, J. E., 1991. Putative virulence factors of *Candida albicans*. Ann. Rev. Microbiol. 45, 187-218.

- Daboussi, M. J., 1997. Fungal transposable elements and genome evolution. Genetica 100, 253-260.
- Daboussi, M. J., Capy, P., 2003. Transposable elements in filamentous fungi. Annu. Rev. Microbiol. 57, 275-299.
- Daboussi, M. J., Daviere, J. M., Graziani, S., Langin, T., 2002. Evolution of the *Fot1* transposons in the genus *Fusarium*: discontinuous distribution and epigenetic inactivation. Mol. Biol. Evol. 19, 510-520.
- Dai, X., Xu, Y., Ma, Q., Xu, W., Wang, T., Xue, Y., Chong, K., 2007. Over-expression of an R1R2R3 MYB gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis. Plant Physiol. 143, 1739-1751.
- Dawe, R. K., 2003. RNA interference, transposons and the centromere. Plant Cell 15, 297-301.
- de Almeida, L. M., Carareto, C. M., 2005. Multiple events of horizontal transfer of the *Minos* transposable element between Drosophila species. Mol. Phylogenet. Evol. 35, 583-594.
- Deininger, P. L., Batzer, M. A., 1999. *Alu* repeats and human disease. Mol. Genet. Metab. 67, 183-193.
- Del Sorbo G., Scala F., Parrella G., Lorito M., Comparini C., Ruocco M., Scala A., 2000. Functional expression of the gene cu, encoding the phytotoxic hydrophobin ceratoulmin, enables *Ophiostoma quercus*, a nonpathogen on elm, to cause symptoms of Dutch elm disease. Mol. Plant Microbe Interact. 13(1):43-53.
- Devereux, J., Haeberli, P., Smithies, O., 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12, 387-395.
- Dewar, K., Bernier, L., 1993. Electrophoretic karyotypes of the elm tree pathogen *Ophiostoma ulmi (sensu lato)*. Mol. Genet. Genomics 238, 43-8.
- Dewar, K., Bernier, L., 1995. Inheritance of chromosome-length polymorphisms in *Ophiostoma ulmi (sensu lato)*. Curr. Genet. 27, 541-549.
- Dewar, K., Bousquet, J., Dufour, J., Bernier, L., 1997. A meiotically reproducible chromosome length polymorphism in the ascomycete fungus *Ophiostoma ulmi* (*sensu lato*). Mol. Genet. Genomics 255, 38-44.

- Dickman M. B., Yarden, O. 1999. Serine/Threonine protein kinases and phosphatases in filamentous fungi. Fungal Genet. Biol. 26, 99-117.
- Di Franco, C., Galuppi, D., Junakovic, N., 1992. Genomic distribution of transposable elements among individuals of an inbred Drosophila line. Genetica 86, 1-11.
- Diao, X., Freeling, M., Lisch, D., 2006. Horizontal transfer of a plant transposon. PLoS Biol. 4, e5 doi:10.1371/journal.pbio.0040005.
- Dias, A. P., Braun, E. L., McMullen, M. D., Grotewold, E., 2003. Recently duplicated maize R2R3 *Myb* genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. Plant Physiol. 131, 610–620.
- Doak, T. G., Doerder, F. P., Jahn, C. L., Herrick, G., 1994. A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc. Natl. Acad. Sci. U. S. A. 91, 942-946.
- Dobigny, G., Aniskin, V., Granjon, L., Cornette, R., Volobouev, V., 2005. Recent radiation in West African *Taterillus (Rodentia, Gerbillinae)*: the concerted role of chromosome and climatic changes. Heredity 95, 358–368.
- Dobigny, G., Ozouf-Costaz, C., Waters, P. D., Bonillo, C., Coutanceau, J. P., Vitaly, V., 2004. *LINE-1* amplification accompanies explosive genome repatterning in rodents. Chromosome Res. 12, 787–793.
- Dogra, N., Breuil, C., 2004. Suppressive subtractive hybridization and differential screening identified genes differentially expressed in yeast and mycelial forms of *Ophiostoma piceae*. FEMS Microbiol. Lett. 238, 175-181.
- Doherty, M., Coutts, R. H., Brasier, C. M., Buck, K. W., 2006. Sequence of RNAdependent RNA polymerase genes provides evidence for three more distinct mitoviruses in *Ophiostoma novo-ulmi* isolate. Virus Genes 33, 41-44.
- Dong, Q., Sadouk, A., van der Lelie, D., Taghavi, S., Ferhat, A., Nuyten, J. M., Borremans,
 B., Mergeay, M., Toussaint, A., 1992. Cloning and sequencing of *IS1086*, an *Alcaligenes eutrophus* insertion element related to *IS30* and *IS4351*. J. Bacteriol. 174, 8133-8138.
- Eichenbaum, Z., Livneh, Z., 1998. UV light induces *IS10* transposition in *Escherichia coli*. Genetics 149, 1173-1181.

- Eickbush, T. H., 1994. Origin and relationships of retroelements. In: S. S. Morse, (Ed.), The evolutionary biology of viruses. Raven Press, New-York, pp. 121-157.
- Elick, T. A., Bauser, C. A., Fraser, M. J., 1996. Excision of the *piggyBac* transposable element in vitro is a precise event that is enhanced by the expression of its encoded transposase. Genetica 98, 33-41.
- Emmons, S. W., Yesner, L., Ruan, K. S., Katzenberg, D., 1983. Evidence for a transposon in *Caenorhabditis elegans*. Cell 32, 55-65.
- Engels, W. R., 1996. *P* elements in Drosophila. Curr. Top. Microbiol. Immunol. 204, 103-123.
- Ersek, T., English, J. T., Schoelz, T. E., 1995. Creation of species hybrids of Phytophthora with modified host ranges by zoospore fusion. Phytopathology 85, 1343-1347.
- Esperença Vieira, C. A., Comportement de rétrotransposons (*copia, mdg1, gypsy* et 412) dans les populations de *Drosophila melanogaster* et *D. simulans*. Thèse de Doctorat. Université Claude Bernard LYON 1, Lyon, 1996, pp. 121.
- Et-Touil, A., Brasier, C. M., Bernier, L., 1999. Localization of a pathogenicity gene in Ophiostoma novo-ulmi and evidence that it may be introgressed from O. ulmi. Mol. Plant Microbe In. 12, 6-15.
- Evans, J. P., Palmiter, R. D., 1991. Retrotransposition of a mouse *L1* element. Proc. Natl. Acad. Sci. U. S. A. 88, 8792-8795.
- Fares, M. A., 2004. SWAPSC: sliding window analysis procedure to detect selective constraints. Bioinformatics 20, 2867-2868.
- Farman, M. L., Taura, S., Leong, S. A., 1996. The *Magnaporthe grisea* DNA fingerprinting probe MGR586 contains the 3' end of an inverted repeat transposon. Mol. Genet. Genomics 251, 675-81.
- Fedoroff, N., 2002b. Control of Mobile DNA. In: Craig N. L.et al., (Ed.), Mobile DNA II. ASM press, Washington D.C., pp. 997-1007.
- Fedoroff, N., Wessler, S., Shure, M., 1983. Isolation of the transposable maize controlling elements *Ac* and *Ds*. Cell 35, 235-242.
- Fedoroff, N. V., 1999. Transposable elements as a molecular evolutionary force Ann. N. Y. Acad. Sci. 18, 251-264.

- Feschotte, C., Mouches, C., 2000. Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a *pogo*-like DNA transposon. Mol. Biol. Evol. 17, 730-737.
- Feschotte, C., Jiang, N., Wessler, S. R., 2002. Plant transposable elements: where genetics meets genomics. Nature Genet. 3, 329-341.
- Feschotte, C., Osterlund, M. T., Peeler, R., Wessler, S. R., 2005. DNA binding specificity of rice *mariner*-like transposases and interactions with Stowaway MITEs. Nucleic Acids Res. 33, 2153-2165.
- Feschotte, C., Zhang, X., Wessler, S. R., 2002. Miniature inverted-repeat transposable elements and their relationship to established DNA transposons. In: N. L. Craig, et al., (Eds.), Mobile DNA II. ASM press, Washington D.C., pp. 1147-1158.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage,
 A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., et al., 1995.
 Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 Genome Biol. 269, 496-512.
- Fraser, C. J., Ciszczon, T., Elick, T., Bauser, C., 1996. Precise excision of TTAA-specific *lepidopteran* transposons *piggyBac* (*IFP2*) and *tagalong* (*TFP3*) from the baculovirus genome in cell lines from two species of *Lepidoptera*. Insect Mol. Biol. 5, 141–151.
- Friedman, R., Hughes, A. L., 2001. Gene duplication and the structure of eukaryotic genomes. Genome Res. 11, 373-381.
- Friedrich, M., Grahnert, A., Hauschildt, S., 2005. Analysis of the 30 UTR of the *ART3* and *ART4* gene by 30 inverse RACE-PCR DNA Sequence. DNA Seq. 16, 53-57.
- Fugmann, S. D., Messier, C., Novack, L. A., Cameron, R. A., Rast, J. P., 2006. An ancient evolutionary origin of the *Rag1/2* gene locus. Proc. Natl. Acad. Sci. U. S. A. 103, 3728-3733.
- Galagan, J. E., Selker, E. U., 2004. RIP: the evolutionary cost of genome defense. Trends Genet. 20, 417-423.
- Gale, M. D., Devos, K. M., 1998. Plant comparative genetics after 10 years. Science 282, 656-659.

- Galtier, N., Gouy, M., Gautier, C., 1996. Seaview and Phylo_win: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543-548.
- Garfinkel, D. J., 2005. Genome evolution mediated by *Ty* elements in Saccharomyces. Cytogenet. Genome Res. 110, 63-69.
- Gibb, E. A., Hausner, G., 2003. A group I intron-like sequence in the nuclear small ribosomal subunit gene of the ophiostomatoid fungus *Gondwanamyces proteae*. Mycol. Res. 107, 1442-1450.
- Gibb, E. A., Hausner, G., 2005. Optional mitochondrial introns and evidence for a homingendonuclease gene in the mtDNA rnl gene in *Ophiostoma ulmi s. lat.* Mycol. Res. 109, 1112-1126.
- Giraud, T., Capy, P., 1996. Somatic activity of the *mariner* transposable element in natural populations of *Drosophila simulans*. Proc. Biol. Sci. 263, 1481-1486.
- Goodwin T.J., Butler M.I., Poulter R.T., 2003. Cryptons: a group of tyrosine-recombinaseencoding DNA transposons from pathogenic fungi. Microbiology 149:3099-3109.
- Gorbunova, V., Levy, A. A., 1997. Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Res. 25, 4650-4657.
- Goyon, C., Rossignol, J. L., Faugeron, G., 1996. Native DNA repeats and methylation in *Ascobolus*. Nucleic Acids Res. 24, 3348–3356.
- Grandbastien, M.-A., Lucas, H., Morel, J.-B., Mhiri, C., Vernhettes, S., Casacuberta, J. M., 1997. The expression of the tobacco *Tnt1* retrotransposon is linked to the plant defense responses. Genetica 100, 241-252.
- Grandbastien, M. A., 1998. Activation of plant retrotransposons under stress conditions. Trends Plant Sci. 3, 181-187.
- Gregory, T. R., Hebert, P. D. N., 1999. The modulation of DNA content: proximate causes and ultimate consequences. Genome Res. 9, 317-324.
- Grossman, G. L., Rafferty, C. S., Fraser, M. J., Benedict, M. Q., 2000. The *piggyBac* element is capable of precise excision and transposition in cells and embryos of the mosquito, *Anopheles gambiae*. Insect Biochem. Mol. Biol. 30, 909-14.
- Haber, J. E., 1998. Mating-type gene switching in *Saccharomyces cerevisiae*. Annu. Rev. Genet. 32, 561-599.

- Hagemann S., Pinsker W., 2001. Drosophila P transposons in the human genome? Mol. Biol. Evol. 18(10):1979-82.
- Hall, B. G., 1999. Transposable elements as activators of cryptic genes in *E. coli*. Genetica 107, 181-187.
- Handler, A. M., Gomez, S. P., 1997. *P* element excision in *Drosophila* is stimulated by gamma irradiation in transient embryonic assays. Genet. Res. 70, 75-78.
- Hawkins, J. S., Kim, H., Nason, J. D., Wing, R. A., Wendel, J. F., 2006. Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*. Genome Res. 16, 1252-1261.
- Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolondny, N. L., Kolchanov, N. A., 1998. Databases on Transcriptional Regulation: TRANSFAC, TRRD, and COMPEL. Nucleic Acids Res. 26, 263-370.
- Hickman, A. B., Perez, Z. N., Zhou, L., Musingarimi, P., Ghirlando, R., Hinshaw, J. E., Craig, N. L., Dyda, F., 2005. Molecular architecture of a eukaryotic DNA transposase. Nat. Struct. Mol. Biol. 12, 715-721.
- Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G., Gibson, T. J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680
- Hintz W. E., Jeng R. S., Yang D. Q., Hubbes M. M., Horgen P. A. 1993. A genetic survey of the pathogenic fungus *Ophiostoma ulmi* across a Dutch elm disease front in western Canada. Genome. 36(3):418-426
- Hintz, W. E., 1999. Sequence analysis of the chitin synthase, a gene of the Dutch elm disease pathogen *Ophiostoma novo-ulmi* indicates a close association with the human pathogen *Sporothrix schenckii*. Gene 237, 215-221.
- Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H., Kanda, M., 1996. Retrotransposons of rice involved in mutations induced by tissue culture. Proc. Natl. Acad. Sci. U. S. A. 93, 7783-7788.

- 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.
- Hochstenbach, R., Harhangi, H., Schouren, K., Bindels, S., Suijkerbuijk, R., Hennig, W., 1996. Transcription of gypsy elements in a Y-chromosome male fertility gene of Drosophila hydei. Genetics 142, 437-446.
- Hua-Van A., Langin T., Daboussi M. J. 2002. Aberrant transposition of a *Tc1-mariner* element, *impala*, in the fungus Fusarium oxysporum. Mol. Genet. Genomics. 267(1):79-87.
- Ikeda, K., Nakayashiki, H., Takagi, M., Tosa, Y., Mayama, S., 2001. Heat shock, copper sulfate and oxidative stress activate the retrotransposon *MAGGY* resident in the plant pathogenic fungus *Magnaporthe grisea*. Mol. Genet. Genomics 266, 318-25.
- Jacob, F., Monod, J., 1959. Genes of structure and genes of regulation in the biosynthesis of proteins. C. R. Hebd. Seances Acad. Sci. 249, 1282-1284.
- Jacobson J. W., Medhora M. M., Hartl D. L., 1986. Molecular structure of a somatically unstable transposable element in Drosophila. Proc. Natl. Acad. Sci. U. S. A. 83(22):8684-8688
- Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S. R., McCouch, S. R., Wessler, S. R., 2003. An active DNA transposon family in rice. Nature 421, 163-167.
- Joos, B., Fischer, M., Schweizer, A., Kuster, H., Boni, J., Wong, J. K., Weber, R., Trkola, A., Gunthard, H. F., 2007. Positive in vivo selection of the *HIV-1* envelope protein gp120 occurs at surface-exposed regions. J. Infect. Dis. 196, 313-20.
- Jouan-Dufournel, I., Cosset, F. L., Contamine, D., Verdier, G., Biémont, C., 1996. Transposable elements behavior following viral genomie stress in *Drosophila melanogaster* inbred line. J. Mol. Evol. 43, 19-27.
- Kachroo, P., Leong, S. A., Chattoo, B. B., 1994. *Pot2*, an inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*. Mol Gen. Genet. 245, 339-348.
- Kang, S., Lebrun, M. H., Farrall, L., Valent, B., 2001. Gain of virulence caused by insertion of a *Pot3* transposon in a *Magnaporthe grisea* avirulence gene. Mol. Plant Microbe In. 14, 671-674.

- Kankel, M. W., Ramsey, D. E., Stokes, T. L., Flowers, S. K., Haag, J. R., Jeddeloh, J. A., Riddle, N. C., Verbsky, M. L., Richards, E. J., 2003. Arabidopsis *MET1* Cytosine methyl-transferase mutants. Genetics 163, 1109-1122.
- Kapitonov VV, Jurka J 2006. Self-synthesizing DNA transposons in eukaryotes. Proc. Natl. Acad. Sci. U. S. A. 103(12):4540-4545.
- Kasak, L., Horak, R., Kivisaar, M., 1997. Promoter-creating mutations in *Pseudomonas putida*: a model system for the study of mutation in starving bacteria. Proc. Natl. Acad. Sci. U. S. A. 94:546-551.
- Kempken, F., 1999. Fungal transposons: from mobile elements towards molecular tools. Appl. Microbiol. Biotechnol. 52, 756-760.
- Kempken, F., Windhofer, F., 2004. Alternative splicing of transcripts of the transposon *Restless* is maintained in the foreign host *Neurospora crassa* and can be modified by introducing mutations at the 5' and 3' splice sites. Curr. Genet. 46, 59-65.
- Kempken F. and Kück U., 1998. Transposons in filamentous fungi--facts and perspectives. Bioessays 20(8):652-659
- Kidwell, M. G., Lisch, D. R., 2002. Transposable elements as sources of genomic variations. In: N. L. Craig et al., (Eds.), Mobile DNAII. ASM press, Washington, D. C., pp. 59-92.
- Kidwell, M. G., Lisch, D. R., 1998. Transposons unbound. Nature 393, 22-23.
- Kinscherf T. G., Leong S. A., 1988. Molecular analysis of the karyotype of *Ustilago* maydis. Chromosoma. 96(6):427-433
- Kondrashov, F. A., Rogozin, I. B., Wolf, Y. I., Koonin, E. V., 2002. Selection in the evolution of gene duplications. Genome Biol. 3(2):RESEARCH0008.
- Kretschmer, P. J., Cohen, S. N., 1979. Effect of temperature on translocation frequency of the *Tn3* element. J. Bacteriol. 139, 515-519.
- Kuan, C. T., Tessman, I., 1991. LexA protein of Escherichia coli represses expression of the Tn5 transposase gene. J. Bacteriol. 173, 6406-6410.
- Kunze, R., Weil, C. F., 2002. The *hAT* and CACTA superfamilies of plant transposons. In:N. L. Craig et al., (Eds.), Mobile DNAII. ASM press, Washington, D. C., pp. 565-610.

- Kupfer, D. M., Reece, C. A., Clifton, S. W., Roe, B. A., Prade, R. A., 1997. Multicellular ascomycetous fungal genomes contain more than 8000 genes. Fungal Genet. Biol. 21, 364-372.
- Kurtz, S., Schleiermacher, C., 1999. REPuter: fast computation of maximal repeats in complete genomes. Bioinformatics 15, 426-427
- Lalo, D. S., Stettler, S., Mariotte, P., Slonimski, P., Thuriaux, P., 1993. Two yeast chromosomes are related by a fossil duplication of their centromeric regions. C. R. Acad. Sci. III 316, 367-373.
- Lampe, D. J., Witherspoon, D. J., Soto-Adames, F. N., Robertson, H. M., 2003. Recent horizontal transfer of Mellifera subfamily *mariner* transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. Mol. Biol. Evol. 20, 554-562.
- Lamrani, S., Ranquet, C., Gama, M. J., Nakai, H., Shapiro, J. A., Toussaint, A., Maenhaut-Michel, G., 1999. Starvation-induced *Mucts62*-mediated coding sequence fusion: a role for *ClpXP*, *Lon*, *RpoS* and *Crp*. Mol. Microbiol. 32, 327-343.
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E., Newburg, L., 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174-181.
- Lane, D., Cavaille, J., Chandler, M., 1994. Induction of the SOS response by *IS1* transposase. J. Mol. Biol. 242, 339-350.
- Lanier, G. N., Peacock, J. W., 1981. Vectors of the pathogen. In: Stipes R. J., Campana, R.J. (Eds.), Compendium of elm diseases. The American Phytopathological Society St. Paul, Minnesota, pp. 14-16.
- Lee, H. H., Yoon, J. Y., Kim, H. S., Kang, J. Y., Kim, K. H., Kim, D. J., Ha, J. Y., Mikami, B., Yoon, H. J., Suh, S. W., 2006. Crystal structure of a metal ion-bound *IS200* transposase. J. Biol. Chem. 281, 4261-4266.
- Lelivelt, M. J., Kawula, T. H., 1995. *Hsc66*, an *Hsp70* homolog in *Escherichia coli*, is induced by cold shock but not by heat shock. J. Bacteriol. 177, 4900-4907.
- Levis, C., Fortini, D., Brygoo, Y., 1997 *Flipper*, a mobile *Fot1*-like transposable element in *Botrytis cinerea*. Mol. Gen. Genet. 254, 674-680

- Li, Q., Fang, X., Olave, I., Han, H., Yu, M., Xiang, P., Stamatoyannopoulos, G., 2006. Transcriptional potential of the g-globin gene is dependent on the CACCC box in a developmental stage-specific manner. Nucleic Acids Res. 34, 3909-3916.
- Liang, F., Han, M., Romanienko, P. J., Jasin, M., 1998. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. 95, 5172-5177.
- Liang, H., Zhou, W., Landweber, L. F., 2006. SWAKK: a web server for detecting positive selection using a sliding window substitution rate analysis. Nucleic Acids Res. 34, W382-384.
- Liti, G., Peruffo, A., James, S. A., Roberts, I. N., Louis, E. J., 2005. Inferences of evolutionary relationships from a population survey of LTR-retrotransposons and telomeric- associated sequences in the Saccharomyces *sensu stricto* complex. Yeast 22, 177–192.
- Luo, G., Ivics, Z., Izsák, Z., Bradley, A., 1998. Chromosomal transposition of *Tc1/mariner*like element in mouse embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 95, 10769-10773.
- Madlung, A., Comai, L., 2004. The effect of stress on genome regulation and structure. Ann. Bot. 94, 481–495.
- Mahillon, J., Chandler, M., 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62, 725-774.
- Malik, H. S., Henikoff, S., 2005. Positive selection of *Iris*, a retroviral enveloppe-derivated host gene in *Drosophila melanogaster*. Plos Genet. 1, e44.
- Margolin, B. S., Garrett-Engele, P. W., Stevens, J. N., Fritz, D. Y., Garrett-Engele, C., Metzenberg, R. L., Selker, E. U., 1998. A methylated *Neurospora* 5S rRNA pseudogene contains a transposable element inactivated by repeat-induced point mutation. Genetics 149, 1787-1797.
- Margulies, L., Briscoe, D. I., Wallace, S. S., 1986. The relationship between radiationinduced and transposon induced genetic damage during Drosophila oogenesis. Mutat. Res. 162, 55-68.

- Margulies, L., Griffith, C. S., Dooley, J. C., Wallace, S. S., 1989. The interaction between transposon mobility and X rays in Drosophila: hybrid sterility and chromosome loss. Mutat. Res. 215, 1-14.
- Mathieu, O., Bender, J., 2004. RNA-directed DNA methylation. J. Cell Sci. 117, 4881-4888.
- McClintock, B., 1951. Chromosome organization and genic expression. Cold Spring Harbor Symp. Quant. Biol. 16, 13-47.
- McClintock, B., 1956. Controlling elements and the gene. Cold Spring Harb. Symp. Quant. Biol. 21, 197-216.
- McClintock, B., 1984. The significance of responses of the genome to challenge. Science 226, 792-801.
- McDonald, J. F., Matyunina, L. V., Wilson, S., Jordan, I. K., Bowen, N. J., Miller, W. J., 1997. LTR retrotransposons and the evolution of eukaryotic enhancers. Genetica 100, 3-13.
- McGrew, L. L., Dworkin-Rastl, E., Dworkin, M. B., 1989. Poly(A) elongation during *Xenopus* oocyte maturation is required for translational recruitment and is mediated by a short sequence element. Genes Dev. 3, 803-815.
- Mes, J. J., Haring, M. A., Cornelissen, B. J., 2000. Foxy: an active family of short interspersed nuclear elements from Fusarium oxysporum. Mol. Gen. Genet. 263, 271–280.
- Michel, K., O'Brochta, D. A., Atkinson, P. W., 2003. The C-terminus of the *Hermes* transposase contains a protein multimerization domain. Insect Biochem. Mol. Biol. 33, 959-970.
- Migheli, Q., Lauge, R., Daviere, J. M., Gerlinger, C., Kaper, F., Langin, T., Daboussi, M.J., 1999. Transposition of the autonomous *Fot1* element in the filamentous fungus *Fusarium oxysporum*. Genetics 151, 1005-1013.
- Milanesi, L., Muselli, M., Arrigo, P., 1996. Hamming Clustering method for signals prediction in 5' and 3' regions of eukaryotic genes. Comput. Applic. Biosci. 12, 399-404.
- Milgroom, M. G., Brasier, C. M., 1997. Potential diversity in vegetative compatibility types of *Ophiostoma novo-ulmi* in North America. Mycologia 89, 722-726.

- Miller, S. M., Schmitt, R., Kirk, D. L., 1993. *Jordan*, an active Volvox transposable element similar to higher plant transposons. Plant Cell 5, 1125-1138.
- Mitchell, A. G., Brasier, C. M., 1994. Contrasting structure of European and North American populations of *Ophiostoma ulmi*. Mycol. Res. 98, 576-582.
- Mitrovich, Q. M., Anderson, P., 2005. mRNA surveillance of expressed pseudogenes in *C. elegans*. Curr. Biol. 15, 963–967.
- Mizuuchi, K., 1992. Polynucleotidyl transfer reactions in transpositional DNA recombination. J. Biol. Chem. 267, 21273-21276.
- Mizuuchi, K., Baker, T. A., 2002. Chemical mechanisms of mobilizing DNA. In: N. L. Craig et al., (Eds.), Mobile DNAII. ASM press, Washington, D. C., pp. 12-23.
- 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 Res. 12, 1305-1315.
- Monroy, F., Sheppard, D. C., 2005. *Taf1*: a class II transposon of *Aspergillus fumigatus*. Fungal Genet. Biol. 42, 638-645.
- Moon, C. D., Craven, K. D., Leuchtmann, A., Clement, S. L., Schardl, C. L., 2004.
 Prevalence of interspecific hybrids amongst asexual fungal endophytes of grasses.
 Mol. Ecol. 13, 1455–1467.
- Morozov, I. Y., Negrete-Urtasun, S., Tilburn, J., Jansen, C. A., Caddick, M. X., Arst, H. N., 2006. Nonsense-mediated mRNA Decay Mutation in *Aspergillus nidulans*. Eukaryotic Cell. Ahead of print on 8 September 2006. EC.00220-06.
- Mujacic, M., Cooper, K. W., Baneyx, F., 1999. Cold-Inducible cloning vectors for low temperature protein expression in *Escherichia coli:* Application to the production of a toxic and proteolytically sensitive fusion protein. Gene 238, 325-332.
- Munoz-Dorado, J., Kondo, K., Inouye, M., Sone, H., 1994. Identification of *cis-* and *trans*acting elements involved in the expression of cold shock-inducible *TIP1* gene of yeast *Saccharomyces cerevisiae*. Nucleic Acids Res. 22, 560-568.
- Murat, D., Bance, P., Callebaut, I., Dassa, E., 2006. ATP hydrolysis is essential for the function of the Uup ATP-binding cassette ATPase in precise excision of transposons. J. Biol. Chem. 281, 6850-6859.

- Naas, T., Blot, M., Fitch, W. M., Arber, W., 1995. Dynamics of *IS*-related genetic rearrangements in resting *Escherichia coli* K-12. Mol. Biol. Evol. 12, 198-207.
- Newcombe, G., 2003. Native *Venturia inopina* sp. nov., specific to *Populus trichocarpa* and its hybrids. Mycol Res. 107, 108-116.
- Newcombe, G., Stirling, B., McDonald, S., Bradshaw, H. D., 2000. *Melampsora* x *columbiana*, a natural hybrid of *M. medusae* and *M. occidentalis*. Mycol. Res. 104, 261-274.
- Notredame, C., Higgins, D., Heringa, J., 2000. T-Coffee: A novel method for multiple sequence alignments. J. Mol. Biol. 302, 205-217
- Nouaud, D., Anxolabehere, D., 1997. *P* element domestication: a stationary truncated *P* element may encode a 66-kDa repressor-like protein in the *Drosophila montium* species subgroup. Mol. Biol. Evol. 14, 1132-1144.
- Ochman, H., Gerber, A. S., Hartl, D. L., 1988. Genetic applications of an inverse polymerase chain reaction. Genetics 120, 621-623
- Orgel, L. E., Crick, F. H. C., 1980. Selfish DNA: the ultimate parasite. Nature 284, 604-607.
- Paoletti, M., Buck, K. W., Brasier, C. M., 2006. Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote *Ophiostoma novo-ulmi*. Mol. Ecol. 15, 249-262.
- Paoletti, M., Seymour, F. A., Alcocer, M. J., Kaur, N., Calvo, A. M., Archer, D. B., Dyer,P. S., 2007. Mating type and the genetic basis of self-fertility in the model fungus *Aspergillus nidulans*. Curr. Biol. 17, 1384-1389.
- Pardue, R. W., DeBaryshe, P. G., 2003. Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. Annu. Rev. Genet. 37, 485-511.
- Pearson, W. R., Lipman, D. J., 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. U. S. A. 85, 2444-2448.
- Peirera, V., Royer, J. C., Hintz, W. E., Field, D., Bowden, C., Kokurewicz, K., Hubbes, M., Horgen, P. A., 2000. A gene associated with filamentous growth in *Ophiostoma novo-ulmi* has RNA-binding motifs and is similar to a yeast gene involved in mRNA splicing. Curr. Genet. 37, 94-103.

- Pelham, H. R., 1982a. A regulatory upstream promoter element in the Drosophila Hsp70
- Peschke, V. M., Phillips, R. L., 1991. Activation of the maize transposable element suppressor-mutator (*Spm*) in tissue culture. Theor. Appl. Genet. 81, 90–97.

heat-shock gene. Cell 30, 517-528.

- Pfeifer, F., Blaseio, U., 1990. Transposition burst of the *ISH27* insertion element family in *Halobacterium halobium*. Nucleic Acids Res. 18 6921–6925.
- Pipe, N. D., Buck, K. W., Brasier, C. M., 1997. Comparison of the *cerato-ulmin* (cu) gene sequences of the Himalayan Dutch elm disease fungus *Ophiostoma himal-ulmi* with those of *O. ulmi* and *O. novo-ulmi* suggests that the cu gene of *O. novo-ulmi* is unlikely to have been acquired recently from *O. himal-ulmi*. Mycol. Res. 101 415-421.
- Pipe, N., Buck, K. W., Brasier, C. M., 1995. Molecular relationships between *Ophiostoma ulmi* and the EAN and NAN races of *O. novo-ulmi* determined by RADP markers. Mycol. Res. 99, 653-658.
- Plasterk, R. H., Izsvak, Z., Ivics, Z., 1999. Resident aliens: the *Tc1/mariner* superfamily of transposable elements. Trends Genet. 15, 326-332.
- Plasterk, R. H. A., Luenen, H. G. A. M., 2003. The *Tc1/mariner* family of transposable elements. In: N. L. Craig, e. al., (Eds.), Mobile DNA II. ASM Press, Washington D. C., pp. 519-532.
- Pouteau, S., Grandbastien, M. A., Boccara, M., 1994. Microbial elicitors of plant defense responses activate transcription of a retrotransposon. Plant J. 5, 535-542.
- Prakob, W., Judelson, H. S., 2007. Gene expression during oosporogenesis in heterothallic and homothallic *Phytophthora*. Fungal Genet. Biol. 44, 726-739.
- Raslton, E., English, J., Dooner, H. K., 1989. Chromosome-breaking structure in maize involving a fractured *Ac* element. Proc. Natl. Acad. Sci. U. S. A. 86, 9451-9455.
- Ramussen, J. P., Taylor, A. H., Ma, L. J., Purcell, S., Kempken, F., Catcheside, D. E., 2004. *Guest*, a transposable element belonging to the *Tc1/mariner* superfamily is an ancient invader of Neurospora genomes. Fungal Genet. Biol. 41, 52-61.
- Ratner, V. A., Zabanov, S. A., Kolesnikova, O. V., Vasilyeva, L. A., 1992. Induction of the mobile genetic element *Dm-412* transpositions in the Drosophila genome by heat shock treatment. Proc. Natl. Acad. Sci. U. S. A. 89, 5650-5654.

- Redder, P., Faguy, D., Brügger, K., She, K., Garret, R. A., 2003. Achaeal mobile DNA. In: Craig N. L. et al. (Ed.), Mobile DNA II. ASM Press, Washington D. C., pp. 1060-1073.
- Reese, M. G., 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput. Chem. 26, 51-56.
- Reznikoff, W. S., 2003. Tn5 transposition. In: Craig N. L. et al. (Ed.), Mobile DNA II. ASM Press, Washington D. C., pp. 403-422.
- Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 16, 276-277.
- Richards, W. C., 1993. Cerato-ulmin: a unique wilt toxin of instrumental significance in the development of Dutch elm disease. In: M. Sticklen, J. Sherald, (Eds.), Dutch Elm Disease Research: Cellular and Molecular Approaches. Springer-Verlag, New York, pp. 89-151.
- Richardson, J. M., Dawson, A., O'Hagan, N., Taylor, P., Finnegan, D. J., Walkinshaw, M.
 D., 2006. Mechanism of *Mos1* transposition: insights from structural analysis.
 EMBO J. 25, 1324-1334.
- Robertson, H. M., 2003. Evolution of DNA transposons in eukaryotes. In: Craig N. L. et al. (Ed.), Mobile DNA II. ASM Press, Washington D. C., pp. 1093-1110.
- Robertson, H. M., Soto-Adames, F. N., Walden, K. O., Avancini, R. M. P., Lampe, D. J., 1998. The *mariner* transposons of animals: Horizontally jumping genes. In: M. Syvanen, C. I. Kado, (Eds.), Horizontal gene transfer. Chapman and Hill, New York, pp. 268–284.
- Rodriguez, M. A., Vermaak, D., Bayes, J. J., Malik, H. S., 2007. Species-specific positive selection of the male-specific lethal complex that participates in dosage compensation in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 104:15412–15417.
- Royer, J. C., Dewar, K., Hubbes, M., Horgen, P. A., 1991. Analysis of a high frequency transformation system for *Ophiostoma ulmi*, the causal agent of Dutch elm disease. Mol. Gen. Genet. 225, 168-176.
- Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132, 365-386.

- Saguez, C., Lecellier, G., Koll, F., 2000. Intronic GIY-YIG endonuclease gene in the mitochondrial genome of *Podospora curvicolla*: evidence for mobility. Nucleic Acids Res. 28, 1299-1306.
- Sanchez-Gracia, A., Maside, X., Charlesworth, B., 2005. High rate of horizontal transfer of transposable elements in Drosophila. Trends Genet. 21, 200-203.
- SanMiguel, P., Tikhonov, A., Jin, Y. K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z., Bennetzen, J. L., 1996 Nested retrotransposons in the intergenic regions of the maize genome. Science 274, 765-768.
- Sargent, T. G., Lloyd, J. A., 2001. The human y-globin TATA and CACCC elements have key, distinct roles in suppressing B-Globin gene expression in embryonic/fetal development. J. Biol. Chem. 276, 41817–41824.
- Sarkar, A., Coates, C. J., Whyard, S., Willhoeft, U., Atkinson, P. W., O'Brochta, D. A., 1997. The *Hermes* element from *Musca domestica* can transpose in four families of cyclorrhaphan flies. Genetica 99, 15-29.
- Sarkar, A., Sim, C., Hong, Y. S., Hogan, J. R., Fraser, M. J., Robertson, H. M., Collins, F. H., 2003. Molecular evolutionary analysis of the widespread *piggyBac* transposon family and related "domesticated" sequences. Mol. Genet. Genomics 270, 173-180.
- Sawyer, S. L., Malik, H. S., 2006. Positive selection of yeast nonhomologous end-joining genes and a retrotransposon conflict hypothesis. Proc. Natl. Acad. Sci. U. S. A. 103, 17614-17619.
- Sawyer, S. L., Wu, L. I., Emerman, M., Malik, H. S., 2005. Positive selection of primate *TRIM5*-alpha identifies a critical species-specific retroviral restriction domain. Proc. Natl. Acad. Sci. U. S. A. 102, 2832-2937.
- Sawyer, S.L., Emerman M., Malik H. S., 2004. Ancient adaptive evolution of the primate antiviral DNA editing enzyme, APOBEC3G. PLoS Biology 2(9): E275.
- Selker, E. U., Cambareri, E. B., Jensen, B. C., R., H. K., 1987. Rearrangement of duplicated DNA in specialized cells of *Neurospora*. Cell, 51, 741-752.
- Selker, E. U., 1997. Epigenetic phenomena in filamentous fungi: useful paradigms or repeat-induced confusion? Curr. Genet. 13, 296-301.
- Selker, E. U., 2002. Repeat-induced gene silencing in fungi. Adv. Genet. 46, 439-450.

- Semple, C., Maxwell, A., Gautier, P., Kilanowski, F. M., Eastwood, H., Barran, P. E., Dorin, J. R., 2005. The complexity of selection at the major primate β-defensin locus. BMC Evolutionary Biology 5, doi:10.1186/1471-2148-5-32.
- Semple, C. A., Rolfe, M., Dorin, J. R., 2003. Duplication and selection in the evolution of primate beta-defensin genes. Genome Biol. 4, R31.
- Sentry J. W. and Symth D. R. (1985). A family of repeated sequences dispersed through the genome of *Lilium henryi*. Chromosoma 92, 149-155
- Shan, X., Liu, Z., Dong, Z., Wang, Y., Chen, Y., Lin, X., Long, L., Han, F., Dong, Y., Liu, B., 2005. Mobilization of the active MITE transposons *mPing* and *Pong* in rice by introgression from wild rice (*Zizania latifolia Griseb.*). Mol. Biol. Evol. 22, 976-990.
- Shapiro, J. A., 2001. The discovery and significance of mobile genetic elements. In: D. J. Sherratt, (Ed.), Mobile genetic elements. Oxford University Press, New York, pp. 1-17.
- Shapiro, J. A., Adhya, S. L., 1969. The galactose operon of *E. coli* K-12. II. A deletion analysis of operon structure and polarity. Genetics 62, 249-264.
- Sheehan, B. J., Bosse, J. T., Beddek, A. J., Rycroft, A. N., Kroll, J. S., Langford, P. R., 2003. Identification of *Actinobacillus pleuropneumoniae* genes important for survival during infection in its natural host. Infect. Immun. 71, 3960-3970
- Shiu, P. K. T., Raju, N. B., Zickler, D., Metzenberg, R. L., 2001. Meiotic Silencing by Unpaired DNA. Cell 107, 905-916.
- Shull, G. H., 1911. Experiments with Maize. Botanical Gazette 52, 480-485.
- Shull, V., Hamer, J. E., 1996. Rearrangements at a DNA-fingerprint locus in the rice blast fungus. Curr. Genet. 30, 263–271.
- Silva, J. C., Kidwell, M. G., 2000. Horizontal transfer and selection in the evolution of *P* elements. Mol. Biol. Evol. 17, 1542–1557.
- Silva, J. C., Kidwell, M. G., 2004. Evolution of *P* elements in natural populations of *Drosophila willistoni* and *D. sturtevanti*. Genetics 168, 1323-1335.
- Singer, M. F., 1994. From genomic juck to human disease. P. A. P. S 138, 11-24.
- Slotkin, R. K., Martienssen, R., 2007. Transposable elements and the epigenetic regulation of the genome. Nat. Genet. 8, 272-285.

- Smyth, D. R., Kaltisis, P., Joesph, J., Sentry, J. W., 1989. Plant retrotransposon from *Lilium henryi* related to *Ty3* of yeast and *gypsy* group of Drosophila. Proc. Natl. Acad. Sci. U. S. A. 86, 5013-5019.
- Solla, A., Gil, L., 2002. Xylem vessel diameter as a factor in resistance of *Ulmus minor* to *Ophiostoma novo-ulmi*. Forest Pathol. 32, 123.
- Steinemann, M., Steinemann, S., 1997. The enigma of Y chromosome degeneration: *TRAM*, a novel retrotransposon is preferentially located on the neo-Y chromosome of *Drosophila miranda*. Genetics 145, 261-266.
- Stoddard, S. F., Howe, M. M., 1989. Localization and regulation of bacteriophage *Mu* promoters. J. Bacteriol. 171, 3440-3448.
- Stroman, P., Müller, C. C., Sorensen, I., 2003. Heat shock treatment increases the frequency of loss of an erythromycin resistance-encoding transposable element from the chromosome of *Lactobacillus crispatus* CHCC3692. Appl. Environ. Microb. 69, 7173-7180.
- Stuurman, J., Kuhlemeier, C., 2005. Stable two-element control of *dTph1* transposition in mutator strains of Petunia by an inactive *ACT1* introgression from a wild species. Plant J. 41, 945-955.
- Surzycki, S. A., Belknap, W. R., 2000. Repetitive-DNA elements are similarly distributed on *Caenorhabditis elegans* autosomes. Proc. Natl. Acad. Sci. U. S. A. 97, 245-249.
- Swofford, D. L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4 beta. Sinauer Associates, Sunderland, Massachusetts.
- Takai, S., 1980. Relationships of the production of the toxin, cerato-ulmin, to synnemata formation, pathogenicity, mycelial habit and growth of *Ceratocystis ulmi* isolates. Can. J. Bot. 58, 658-662
- Talbot N. J., Ebbole D. J., Hamer J. E., 1993. Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea*. Plant Cell. 5(11):1575-1590.
- Tamaru, H., Selker, E. U., 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature 414, 277-283.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 10, 1093/molbev/msm092.
- Temin, H. M., Mizutani, S., 1970. RNA-dependent DNA polymerase in virions of *Rous* sarcoma virus. Nature 226, 1211-1213.
- Temple B, Horgen PA, Bernier L, Hintz WE., 1997. Cerato-ulmin, a hydrophobin secreted by the causal agents of Dutch elm disease, is a parasitic fitness factor. Fungal Genet Biol. 22(1):39-53
- Thon, M. R., Pan, H., Diener, S., Papalas, J., Taro, A., Mitchell, T. K., Dean, R. A., 2006. The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus *Magnaporthe oryzae*. Genome Biol. 7, R16.1-R16.9.
- Tucker, S. L., Talbot, N. J., 2001. Surface attachment and pre-penetration stage development by plant pathogenic fungi. Annu. Rev. Phytopathol. 39, 385-417.
- Tudor, M., Lobocka, M., Goodell, M., Pettitt, J., O'Hare, K., 1992. The *pogo* transposable element family of *Drosophila melanogaster*. Mol. Gen. Genet. 232, 126-134.
- Vasina, J. A., Baneyx, F., 1996. Recombinant protein expression at low temperature under the transcriptional control of the major *E. coli* cold shock promoter *cspA*. Appl. Environ. Microbiol. 62, 1444-1447.
- Vastenhouw, N. L., Plasterk, R. H. A., 2004. RNAi protects the *Caenorhabditis elegans* germline against transposition. Trends Genet. 20, 314-319.
- Vermaak, D. Henikoff, S., Malik, H. S., 2005. Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila melanogaster. PLOS Genetics 1(1):e9.
- Vershinin, A. V., Allnutt, T. R., Knox, M. R., Ambrose, M. J., Ellis, T. H., 2003. Transposable elements reveal the impact of introgression, rather than transposition, in *Pisum* diversity, evolution, and domestication. Mol. Biol. Evol. 20, 2067-2075.
- Viera, C., Biémont, C., 1996. Geographical variation in insertion site number of retrotransposon *412* in *Drosophila simulans*. J. Mol. Evol. 42, 443-451.
- Walbot, V., 1992. Reactivation of Mutator transposable elements of maize by ultraviolet light. Mol. Gen. Genet. 234, 353-360.
- Walbot, V., 1999. UV-B damage amplified by transposons in maize. Nature 397, 398-399.

- Walbot, V., Rudenko, G. N., 2002. *MuDR/Mu* transposable elements of Maize. In: Craig N.L. et al., (Ed.), Mobile DNA II. ASM Press, Washington D.C., pp. 533-564.
- Walser, J. C., Chen, B., Feder, M. E., 2006. Heat-shock promoters: Targets for evolution by *P* transposable elements in *Drosophila*. PLoS Genet. 2, e165.
- Watson, J. D., Baker, T. A., Bell, S. P., Gann A., Levine M., Losick R., 2004. The molecular biology of the gene. Cumming B (Ed.) 5th edition, 830 pp.
- Weil, C. F., Kunze, R., 2000. Transposition of maize Ac/Ds transposable elements in the yeast Saccharomyces cerevisiae. Nat. Genet. 26, 187-190.
- Weinstock, K. G., Mastrangelo, M. F., Burkett, T. J., Garfinkel, D. J., Strathern, J. N., 1990. Multimeric arrays of the yeast retrotransposon *Ty*. Mol. Cell. Biol. 10, 2882-2892.
- Wendel, J. F., Wessler, S. R., 2000. Retrotransposon-mediated genome evolution on a local ecological scale. Proc. Natl. Acad. Sci. U. S. A. 97, 6250-6252.
- Wessler, S. R., 1996. Plant retrotransposons: turned on by stress. Curr. Biol. 6, 959–961.
- Wichman, H. A., Van den Bussche, R. A., Hamilton, M. J., Baker, R. J., 1992. Transposable elements and the evolution of genome organization in mammals. Genetica 86, 287-293.
- Windhofer, F., Hauck, K., Catcheside, D. E. A., Kürk, U., Kempken, F., 2002. Ds-like Restless deletion derivatives occur in Tolypocladium inflatum and two foreign hosts, Neurospora crassa and Penicillium chrysogenum. Fungal Genet. Biol. 35, 171-182.
- Wingfield, M., Forest Fungi in a Changing World. 8th International Mycological Congres, Cairn, Australia, 2006.
- Xu, J. R., 2000. MAP kinase in fungal pathogens. Fungal Genet. Biol. 31, 137-152.
- Xu, Z., Yan, X., Maurais, S., Fu, H., O'Brien, D. G., Mottinger, J., Dooner, H. K., 2004. *Jittery*, a mutator distant relative with a paradoxical mobile behavior: excision without reinsertion. Plant Cell 16, 1105–1114.
- Yang, G., Weil, C. F., Wessler, S. R., 2006. A rice *Tc1/Mariner*-like element transposes in yeast. Plant Cell 18, 2469-2478.
- Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555-556.

- Zhang, X., Feschotte, C., Zhang, Q., Jiang, N., Eggleston, W. B., Wessler, S. R., 2001. P instability factor: an active maize transposon system associated with the amplification of *Tourist*-like MITEs and a new superfamily of transposases. Proc. Natl. Acad. Sci. U. S. A. 98, 12572-12577.
- Zhou, E., Jia, Y., Singh, P., Correll, J. C., Lee, F. N., 2007. Instability of the Magnaporthe oryzae avirulence gene AVR-Pita alters virulence. Fungal Genet. Biol. doi:10.1016/j.fgb.2007.02.003.
- Zolan, M., Pukkila, P., 1986. Inheritance of DNA methylation in *Coprinus cinereus*. Mol. Cell. Biol. 6, 195-200.
- Zolan M. E., Heyler N. K., Stassen N. Y. 1994. Inheritance of chromosome-length polymorphisms in Coprinus cinereus. Genetics. 137(1):87-94.
- Zuker, C., Cappello, J., Lodish, H. F., George, P., Chung, S., 1984. *Dictyostelium* transposable element *DIRS-1* has 350-base-pair inverted terminal repeats that contain a heat shock promoter. Proc. Natl. Acad. Sci. U. S. A. 81, 2660-2664.
- Zuker, C., Cappello, J., Chisholm, R. L., Lodish, H. F., 1983. A repetitive Dictyostelium gene family that is induced during differentiation and by heat shock. Cell 34, 997-1005.

<u>Annexes</u>

1- Les mécanismes de silencing.

Ces mécanismes sont divisés en deux grands groupes par ordre d'apparition dans les cascades de régulation : A) les mécanismes transcriptionels : (1) la méthylation des histones (il s'agit de modifications en N-terminal de la queue des histones, à savoir des acétylation, méthylation ou phosphorylation permettant ou non l'enroulement de l'ADN autour de l'histone et donc conduisant soit l'expression ou non d'un gène), (2) les mutations RIP pour repeat-induced point (il s'agit de la transformation préférentielle des dinucléotides CpA en groupement TpA par 5'-méthylation et déamination des cystéines), (3) le silencing méiotique par non appariement d'ADN (il s'agit du silencing de l'ensemble des copies d'un gène si une de ces dernières ne se trouvent pas appariée à son homologue durant la méiose) et B) les mécanismes post-transcriptionels comme (4) l'ARN interférence (il s'agit, d'une manière générale, de l'ensemble des mécanismes impliquant la dégradation des ARNm dans la cellule) incluant le Quelling (il s'agit de l'insertion d'un court segment homologue d'une région transcrite entraînant le blocage de l'information transcrite) et la co-suppression (ou l'incorporation de manière naturelle, d'un transgène en anti-sens, générant une non expression de son endogène actif correspondant) et enfin (5) la domestication des éléments mobiles (il s'agit de détourner l'utilisation de la fonction propre d'un transposon dans le but de l'autoréguler). Parmi les mécanismes découverts chez les champignons filamenteux, le mécanisme de RIP a été le premier découvert et par conséquent le mieux étudié (Bouhouche et al., 2004). Il a été mis en évidence par l'introduction de transgènes chez Neurospora crassa et génère des transitions de cytosine en thymine via l'intermédiaire de deux enzymes (5'-methyl transférase et déaminases). Les particularités de ce mécanisme sont les suivantes : il agirait sur des séquences de plus de 400-nt, identiques à plus de 85% et ayant au moins 2 copies dans le génomes, dans des zones non protégées (référence aux zones NOR - nucleolus organizer region - contenant les ADN ribosomiques). RIP a un très fort impact sur la suppression de création de nouveaux gènes ou de gènes partielles via les duplications. En effet il a été prouvé que 80 % des copies dupliquées vont intégrer à 99,5 % une mutation due au RIP qui va entraîner l'apparition d'un codon stop dans leur cadre de lecture et ainsi stopper la traduction.

Cinq principaux mécanismes de défense contre les séquences répétées chez les champignons :



(Kankel et al., 2003); (Mathieu et al., 2004); (Shiu et al., 2001); (A) (Chicas et al., 2004); (Shiu et al., 2004

2- Séquences des ORF putatives pour OPHIO1 (DQ649003)et OPHIO2 (DQ649004) et

leur traduction en acides aminés.

ORF OPHIO1 ATGCGGGAATATACGGAAGAAAACGTTATCGCCGCTGTCCAAGCGGTTCAAAATGGCACCTCTTACCGGAAGGCTTCGAAACAGTACGGT MREY TEENVIA VQA VQNGTSYRKA ORF OPHIO2 AT GCAC GATTATAC CGAAGAC GACCTTCAAGC CGCTATTAATGC GGT CAAAAATAACAC TTC TTTC CGGGACGC GTC GAAA CAGTACGGG M H D Y T E D D L Q A A I N A V K N N T S F R D A S K Q Y LTRVO P VST Т G KP RHLAF 0 HL OKL V P F S T I P S R V C G S K P R S L A F Q S F OKLSEA E TRLA 3 WICTO G LGRP P THA OVKA Т SL ORF OPHIO2 GAGACCCGTTTAGCGAATTGGATTATTATCCAAGGCACTTTAGGACGTCCTCCTACCCACTCTCAGGTCAAGGCCGTTGCGGCGTCTTTA т Е TRLAN WI IIO G LGR P P THS 0 v KA VAA S L SPDENILG KNWLHSF LRRN SI OPHIO2 AT CGGA GCAT CAGC GGG CGAA AAAGTCATTGG AAAGAACT GGCTAGA TGCC TTTTTGCG CAGAAAT CCGT CGAT AAA GGT C CAAAGAAG C ORF I G A S A G E K V I G K N W L D A F LRRNP SIK U ORF OPHIO1 AAATCGATCGATGCGAAGCGAGTCAATGGTGCGTCGACTGATGCTATTAGGACTTGGTTTCGGCGTCTCGATATTCCAGAGATTAAGCAT K S T D A K B V N G A S T D A T B T M F B B L D T P F T K H ORF OPHIOZ AAATCAATTGATTCTAAGCGGGTCAATGGTGCCTCGACTGATATAATCAAGTCGTGGTTTCAGTGCCTTGATATACCTGAAAATCAAAAAC STDSKRVNGASTDTTKSMF OCTDTPETKN K ORF OPHIO1 ATTTTGCCGCAAAACAGATGGAATATGGATGAGACAGGCTTCTCTTCGAGTCAAGGGGATCCTATATATGTGCCTTGGGACAGCCGAAAAG WNMDET QG D R I ORF OPHIOZ ATTCTACCACAAAACAGATGGAACATGGACGAGACGGGCTTCTCGATGGGCCAGGGAGATCCTTTATACGTCCTCGGAAACTGCCCAAAAG O N R W N M D E T G F S M G OGD P L L K T R K K O V G S R A M T S T TECVSA TP . A G A 9 T. D Т. VECISAT T KIRKKOMGSRAW T ST GKS L P P L ORF OPHIO1 ATTATTTTCAAGGGAAAAAAAGTCCAACAACAATGGTTTCCGACTGATTTGAGTCTTTTCAACTCTTGGCAATTCACTGCTACTAATAAT K G K K V O O O M F N S M P TDLS L F TTF OF N N ORF OPHIOZ ATTATTTTCAAGGGAAAATCGGTCCAGCAGCAGCGGTTTCCAGCCGGATCTAAATCCTTACAGCTCTTGGCAATTCACAGCTACCAATAAT IIFKGKSVQQQWFPADLNPYSSWQFTATNN ORF OPHIO1 GGGTGGACTGACTACGAAAACGGGGTTGAAATGGCTCGAAGATGTGTTTATTCCATGTTCTGCCCCTGCGCGCCCTTCAGAAGCTAGATTG D E т G LKW LED 37 D C S D R P ORF OPHIOZ GGGTGGTCTGACAACGAGACGGGCTTGAAGTGGCTTAAAGATATCTTTATTCCATGTTCTACTCCTATGCGTCCTTCCGACGCTAGATTG G W S D N E T G L K W L K D I F I P C S T P M R P S D A т. ORF OPHIO1 CTAGTTATTGATGGCCACGGAAGCCATGAGACCGATGGCTTTATGAAGTTATGCTTCCGAGAATAACATATTCTTGCTTTCTTACCTTCT s н Е Т D MK C E D H G L N I L I F N LL L ORF OPHIO2 ${\bf TT} {\bf GGTCCTT} {\bf GAT} {\bf GGCA} {\bf TGGA} {\bf AGCCACGATT} {\bf CAGACGAATTTAT} {\bf GAAGCT} {\bf TGGT} {\bf GAATAATAT} {\bf CT} {\bf ATTT} {\bf ACTTTTC} {\bf TT} {\bf CC} {\bf TGCC} {\bf CT} {\bf CC} {\bf CC$ L V L D G H G S H D S D E F M K L C F E N N I Y L L F L P 8 SHVL OPLDLT т F SP LKA KKEI EKT v н SY VL OP LDLT F S LLKGY F RKEME K 37 ORF OPHIO1 GACGAAGAAGCTGTTATATGTAACAAAAGAACTTTTCTAAAGACGTATGCAGCTGCACGTACGGCTGCTTTGACAGCAAAAAATATCAAA EEAVICNKRT FLKT Y A B T T. T. A. K. M TK D A A A Δ. ORF OPHIOZ GACGACGCGTCTGTGGTCTGTAATAAAAGGACTTTCATAAAGAGCTATTCTATAGCTCGTACAGCTGCATTAACGTCTCAAAAACATCAGA D D A S V V C N K R T F I K S Y S I A R T A A L T S Q N I R ORF OPHIO1 ATGGGATGGCGTACGGCTGGTTTATGGCCGAGAAACGTTATGCGGCCATTAATGAGTCCTTTTGTTCTCTCAGTCGACAACCAGCCATCG LWPRNVMRP м WR A G L M S P ਾਜ VL 9 U DNOP ORF OPHIO2 AGCGGCTGGCGTACAGCCGGGTTATGGCCTAGAAACGTTCTGCGGCCACTTATGAGCCCCTTTCGTCCTTTCAGACGACAACCAGCCATCG T A G LW P R N VLR P L M S P F υ ь s D D PNETNTIG SSRAI DSTF TP. P TP . TP. F 8 V T T INI E T P VOTN TIRF S SSRAI SE Y PIR T Y P S V ToT E ORF OPHIO1 CCAAAGAAGTCGACAGATTGGTCTCAACAAGTACACAAATTCAGCCAACAAGAAGGGACGTCTACGTCGCCGCCTACTTCTTCCAGAAA ToT SQ Q v H K F E K S T R ORF OPHIO2 F K K S T D L S R Q V Q Q F N Q Q E K T S I S S R H F F QK ORF OPHIO1 CTTCGAAAGGGATTCGATGAAAAGCTAGACCTTTTAGCTAGTCAAAAACGAAAGATCGAAGACTTAGAGGGCTATCGTTGATACTTTACGG L B K G FDEKLDLLA SOKRKIEDLEA TVD T. B ORF OPHIOZ ATTAGAAAAGGATTCGATGAAAAGTTAGACCTTTTAGCTAGTCAAAAACGACGAATCGAAGAGCTAGAAGCAATTGTCGACGGTCTACAA RKGFDEKLDLLASQKRRIEELEAIVDGL ORF OPHIO1 CCTAGAAAGAAGAAGAAGGCAGTAGTGACCACTCCAAATTCGAAGTTGCTACTATCTGGCATGTCCAGCGTACGCAAATCACTGCTAGAGAG v v P I W R KKKA T N S K T H v QRT RE ORF OPHIO2 GTCAGGAAAAAAAAAAGAAGGTGACTATTAGTCCAAATTCAAAGTTTGCTACCATCTGGCACGTACAGCGGTCGCAAATCGCTGCCGGTGAG RKKKKVTISP NSKF ATIWHV ORS O I A A E ORF OPHIO1 ACTGTTACTGCTCTAGATAATAGTAGTGACGAAGAGGTGTAG V T A L D N S S D E E V ORF OPHIO2 TCTGTTACTGCTATTGAAGGTAGTAGCGACGAAGAGGATTAA v TAIEG SSDEED

3- Séquences nucléotiques et conversion en acides aminés de l'ORF putative pour le transposon *OPHIO3* (<u>DQ649005</u>) amplifiées d'ADN génomique de la souche H327. Le putative codon d'initiation a été indiqué à l'aide d'un carré noir.

OPHIO3	AGT CGGTGGGGCCCCTTTTGGGCCACCCCTGTTTTGGGCCACCATTAATAATAAGGCGGTATACAAATTAGGTTAGATAAATGGCAATG
ОРНТОЗ	₩ 3.5 8 ₩ ₩ C (3.5 X 3.5 2 ₩ C (3.5 5 ₩ C (3.5 5 ₩ ₩ ₩ ₩ ₩ C 3.5 5 ₩ ₩ ₩ ₩ 6 € 5 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5 5 € 5
0111100	* I R I K I V K I K F L K I M K I Y D F L H P T M P L Y T E
OPHIO3	ACGGATGTGGCCAATGCCCTGGCTGACGTAGATCGAGGTGTATTTTTTNGAGTAGTAGCTAAAAGCTATAACGTACCCAGAAGTACTCTT
	T D V A N A L A D V D R G V F F X V V A K S Y N V P R S T L
OPHIO3	TAAGCCCGTTAAAAATGGACGTTAAGGCTATAAAGTAGGGGTAGAGTCCTTATAAATCCTCTATACGCCTTGAAGCTGTGCTTGTAGAG
	* A R * N G R * G Y K V G V E S L * I L S I R L E A V L V E
OPHIO3	TGGGTGACGGTATAAGCATCGCTAGGCATGGCTTTTACATACGGCTAAATGAGAGACGTNGTAGAGCATTTTTTGACTGTTAATGGCAG
	WVTV * A S L G M A F T Y G * M R D X V E H F L T V N G R
OPHIO3	CCGTAAAAGCTGGGAAAGACTTGGTAAGAAGCGTTTTTAAAACGTAATCCAGCTATTAAGACCCTTAAGCGTGTATTAGTCGAATCTCA/
	P*KLGKTW*EAFLKRNPAIKTLKRVLVESQ
OPHIO3	CGTATTAATAGCGCTTCAAAACCAAAGATTGAAGCATTTTTCGAGCTTTTATAAAACGAAGCAATTACTAATATCCCACTACAATACCGG
	RINSASKPKIEAFFELL*NEAITNIPLQYR
OPHIO3	TATAATATAAATGAGACAAGCCTTACAGAAAGTCTAAGGGCCCAATAGCTTAATCGTAGGCCGGTTAAACAAAC
	YNINETSLTESLRANSLIVGRLNKRRKVVK
OPHIO3	T CAAGCGGCTCGT GTATCTAAGTAACAA CATTTAAATACATCTCTGCCGCCGGCTTTAACCTCTCTCT
	S S G S C I * V T T F K Y I S A A G F I L S P L I I F A T N
OPHTO3	<u>ᲞᲝᲑᲚᲝᲐᲚᲝᲐᲚᲝᲐᲜᲐᲝᲐᲚᲝᲝᲚᲝᲐᲚᲐᲝᲐᲐᲐᲐᲚᲐᲐᲜᲐᲐᲝᲐᲚ</u> ᲝᲝᲚᲝᲚᲝᲚᲝᲚᲝᲐᲚᲐᲐᲐᲐᲐᲚᲝᲝᲐᲚᲐᲐᲜᲐᲝᲚᲐᲚᲐᲐᲜᲐᲚᲐᲚᲐᲚᲐᲜᲐᲝᲐᲐᲜᲐᲝᲐᲐᲜᲐᲝᲐᲐᲜᲐ
0111100	TV*HOYFLLDIKEYRP*KFTNTOTS*TNND
ODWTO2	
OFHIOS	TARE TARGET AND THE TOTAL AND THE TOTAL AND THE TARGET AND THE TAR
ODMTO2	
OPHIOS	A GRANT TATACACCERTS TTATA TATACATATA ARA ARA STORES TO THE THE CONCERNMENT AND A STORES AND A S
ODUTO2	
OPHIOS	
	Q E D T S T F S E D K A K T K V C F S S D E S K B B N S N
OPHIO3	
	"NERNFLACIRBARTOALSATNIESG" RAA
OPHI03	GGCTTATGGCCGTCAGTCTTCGCAAAGCTTTAGCGAATCCATTTGTTATCATAGAGACACTAAACGACCTTTTAGATAATCGTCAAGA
	G L W P V S L K K A L A N P F V P L L N D L L D N K Q D
OPHIO3	CAGCCTGAAATGCCTATCACACCGAGAAAAACTCGTTCTCAAGCAATCGCTTTCCAGACGCCGTTATCTAGCGTTCAAGTCCGTCGACA
	Q P E M P I T P R K T R S Q A I A F Q T P L S S V Q V R R H
OPHIO3	GTCTCTCAAATCCAATACGGAGATTACGAAAAGCTTACGGAGACTTTTAGCGAAGACTAGCAATGCCTTAGACGTTAAAACAGGTGAAAT
	V S Q I Q Y G D Y E S L R R L L A K T S N A L D V K T G E I
OPHIO3	ACCGTCTTACAGCGCCAAATCAAAGCCTTAGAGGCCCAGCTAGCCTCTTATAAAAATACAAAAAAAA
	T V L Q R Q I K A L E A Q L A S Y K N T K K K E V R P D P N
OPHIO3	ACTAGGTTCGTCACAATCGAAGACGTACAGACGGTCAGAAATGCTATCGAAGAGAAAATGGAAAGGACGGCTCGGGAGATAGCAGAC
	TRFVTIEDVQTVRNAIELEENERTAREIAD
OPHIO3	GGTGATATAGCCGATTTTTTGTACACTTTTGATAGGTTTACGGGGGCTGAAGTGCCGTTGTAGTTCTTAATACATGTTTTTTTT
	G D I A D F L Y T F D R F T G A E V P L *
OPHIO3	TTATAAGTGGCCTAGAAGGGGGGTGGCCTAAAAGGGGCCCCACCGACT