

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

HYBRIDATION ET POLYPLOÏDIE DANS LE COMPLEXE D'ESPÈCES *DAPHNIA*
PULEX ET LEURS EFFETS SUR L'ÉVOLUTION D'UN TRANSPOSON.

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN BIOLOGIE EXTENSIONNÉ
DE L'UNIVERSITÉ DU QUÉBEC À MONTRÉAL

PAR
ROLAND VERGILINO

MAI 2011

UNIVERSITÉ DU QUÉBEC À RIMOUSKI
Service de la bibliothèque

Avertissement

La diffusion de ce mémoire ou de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire « *Autorisation de reproduire et de diffuser un rapport, un mémoire ou une thèse* ». En signant ce formulaire, l'auteur concède à l'Université du Québec à Rimouski une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de son travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, l'auteur autorise l'Université du Québec à Rimouski à reproduire, diffuser, prêter, distribuer ou vendre des copies de son travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de la part de l'auteur à ses droits moraux ni à ses droits de propriété intellectuelle. Sauf entente contraire, l'auteur conserve la liberté de diffuser et de commercialiser ou non ce travail dont il possède un exemplaire.

*À mes grands parents,
qui ont élevé mes parents.*

Remerciements

Mes remerciements vont tout d'abord à ma directrice de recherche, France Dufresne, pour m'avoir accueilli au sein de son laboratoire et de m'avoir donné la chance de travailler sur un sujet passionnant. Je la remercie aussi pour avoir su me laisser la liberté nécessaire à l'accomplissement de mes travaux de recherche. Ses commentaires sur les idées, les présentations et les écrits que je lui présentais m'ont aiguillé et ont grandement aidé à la construction de la thèse présentée ici. Merci encore pour tout... j'espère continuer de collaborer avec toi...

Je voudrai aussi remercier mon comité de soutenance de thèse, Dre. Malika Ainouche, Dre. Gesche Winkler et Dre. Béatrix Beisner, respectivement membre externe, membre interne et présidente de jury, pour avoir pris le temps de lire, de corriger et de commenter ce texte qui, je l'espère, ne vous paraîtra pas abscons.

Le passage dans le laboratoire du docteur Teresa Crease a été pour moi un grand plaisir. Je tiens à remercier cette chercheuse qui, par son aide technique, ses conseils et ses commentaires, m'a guidé dans le processus scientifique lors du troisième chapitre. J'espère pouvoir collaborer très prochainement avec elle.

Il me plaît à dire qu'un doctorat représente une part significative d'une vie. Dans mon cas celui-ci à occuper 5 ans de ma vie (ce qui est au dessus de 5% de l'espérance de vie moyenne d'un humain de sexe masculin en 2011). Pendant ces cinq années, j'ai rencontré et côtoyé une foule de personnes. Il sera difficile de remercier toutes les personnes qui m'ont soutenu mais je citerai pêle-mêle et sans ordre de préséance :

- Marie-Claude Gagnon et Jeff avec qui j'ai beaucoup partagé

- Thierry, Nico, Kaven, Sophie, Thomas, Daniel le cobra, Cy«raie»na et la gang des laboratoires C-101, C-103 et C104
- Ma coloc' Véro, Éric, Ariane, Oli, P'tit Gab, Phil et toute la gang d'étudiants gradués avec qui j'ai eu énormément de plaisir au sein du RÉGÉE et du colloque « la Biologie dans tous ses états »
- Pierre, Christian, Julie et Claude qui par leurs commentaires constructifs m'ont permis d'améliorer mes présentations et mes articles...
- Marion et Ghis, Antoine, Sam, Émilie, Xav', Gaël, Francesc' et les potes qui sont loin géographiquement mais proches dans mon cœur...
- Les auxiliaires d'enseignements (Pierre, Robert, Alain et Yves) avec qui j'ai beaucoup appris...
- Pauline, Louise, Danie et Mélanie, la dreamteam du secrétariat.
- Luc avec qui j'ai apprécié de jaser autour d'une bière certains vendredi soir...
- Véro, Mich', Julie avec qui j'ai eu du fun dans le laboratoire
- Bernard Angers et la gang de son laboratoire avec qui c'est toujours passionnant de jaser.
- Les étudiants de « génétique », d'« invertébrés » et d'« outils moléculaires en écologie » pour lesquels j'ai pris un grand plaisir à enseigner...
- Émilie P-W et bien sûr, toute la gang du Baro
- Le personnel de l'UQAR que j'ai parfois cotoyé tard le soir...

Une personne a tout particulièrement compté pour moi les derniers mois de mon doctorat et je tiens tout particulièrement à la remercier pour son soutien et sa compréhension : Merci Tessa.

Mes parents m'ont soutenu et encouragé dans mon cursus scolaire et ils ont été pour moi une source d'immense réconfort lors de ce doctorat. Merci à eux pour m'avoir permis de me rendre jusque là. Ma gratitude est infinie...

Table des matières

Titres	Page
Liste des Figures.....	vi
Liste des Tableaux	viii
Résumé	x
Chapitre I: Introduction générale	1
1.1 Polyploïdie et éléments transposables.....	1
1.1.1 Diversité des génomes	1
1.1.2 Importance des éléments transposables	2
1.1.3 Importance de la polyploïdie chez les eucaryotes.....	10
1.1.4 Interaction entre polyploïdie et éléments transposables	12
1.1.5 Importance d'un modèle animal et objectif de la thèse.....	14
1.2 Le complexe <i>Daphnia pulex</i> : Un modèle animal approprié?	15
1.2.1 Taxonomie et écologie.....	15
1.2.2 Hybridation et polyploïdisation.....	18
1.2.3 Mode de reproduction.....	19
1.2.4 Les éléments transposables chez <i>Daphnia pulex</i>	23
1.3 Objectifs du premier chapitre	24
1.3.1 Etude de la taille de génome	24
1.3.2 Étude de la polyploïdie	25
1.4 Objectifs du deuxième chapitre.....	26

1.4.1	Relations évolutives et spéciation	26
1.4.2	Hybridation dans le complexe <i>Daphnia pulex</i>	28
1.5	Objectifs du troisième chapitre	29
1.5.1	Hybridation et diversité des allèles du transposon <i>Pokey</i>	29
1.5.2	Effet de la polyploïdie sur le transposon <i>Pokey</i>	30
	Chapitre II.....	32
2.0	Résumé.....	33
2.1	Abstract	34
2.2	Introduction.....	35
2.3	Material and methods	37
2.3.1	<i>Daphnia</i> collection	37
2.3.2	Clone identification.....	42
2.3.3	Phylogenetic analyses.....	42
2.3.4	Multilocus genotypes.....	44
2.3.5	Flow cytometry	45
2.4	Results.....	46
2.4.1	Clone identification.....	46
2.4.2	Polyploid diversity and ploidy level.....	49
2.4.3	Genome sizes and ploidy level	53
2.5	Discussion.....	57
2.5.1	Genome size in diploid members of the <i>Daphnia pulex</i> complex..	57
2.5.2	Evolution of polyploidy in the <i>Daphnia pulicaria</i> group.....	57
2.5.3	Genome size evolution of <i>D. tenebrosa</i> polyploids	59

2.5.4	Origins of polyploidy in the <i>pulicaria</i> lineage	60
2.5.5	Importance of triploidy in the <i>D. pulex</i> complex	61
2.6	Acknowledgments	62
	Chapitre III.....	67
3.0	Résumé.....	68
3.1	Abstract	69
3.2	Introduction.....	70
3.3	Material and methods	74
3.3.1	Samples and DNA extraction	74
3.3.2	PCR amplifications	82
3.3.3	Phylogenetic Analyses of <i>ND5</i> and <i>Rab4</i> subunit of the GTPase family.....	85
3.3.4	Tests of congruencies between trees.....	86
3.3.5	Coalescent approach to test the hybridization hypothesis.....	87
3.3.6	Microsatellite analyses	88
3.4	Results.....	89
3.4.1	Diversity of <i>ND5</i> haplotypes.....	89
3.4.2	<i>Rab4</i> gene tree and comparisons with mitochondrial <i>ND5</i> gene tree.....	93
3.4.3	Microsatellite analyses	96
3.5	Discussion.....	101
3.5.1	Genetic relationships between <i>D. pulex</i> and <i>D. pulicaria</i>	101
3.5.2	Polyploidy in the <i>pulicaria</i> clade	102
3.5.3	Genetic relationships among species of the <i>tenebrosa</i> clade.....	103

3.5.4	Genetic relationships between species from <i>D. pulex complex</i> ...	103
3.5.5	Phylogenetic incongruence and lineage sorting in a species complex.....	105
3.6	Conclusion	106
3.7	Acknowledgments.....	106
	Chapitre IV	109
4.0	Résumé.....	110
4.1	Abstract	111
4.2	Introduction.....	112
4.3	Material and methods	116
4.3.1	Daphnia samples.....	116
4.3.2	Transposon display	119
4.3.3	Cloning and sequencing <i>Pokey</i> alleles.....	120
4.3.4	Recombination analyses	121
4.3.5	Phylogenetic analyses and ancestral recombination graphs	122
4.3.6	RFLP analysis	122
4.4	Results.....	123
4.4.1	Quantification of <i>Pokey</i> insertion sites	123
4.4.2	<i>Pokey</i> sequences	128
4.4.3	<i>Pokey</i> alleles in different isolates	138
4.5	Discussion.....	143
4.5.1	Effect of ploidy level on the density of <i>Pokey</i> into the genome....	143
4.5.2	Evolution of <i>Pokey</i> via intragenic recombination	145

4.6 Acknowledgments	148
Chapitre V: Discussion générale.....	153
5.1 Tailles de génome et niveaux de ploïdie d'isolats du complexe <i>Daphnia pulex</i>	155
5.1.1 Polyploïdie chez les isolats du complexe <i>Daphnia pulex</i>	155
5.1.2 Tailles du génome chez les espèces du complexe <i>Daphnia pulex</i>	156
5.1.3 Apport de la méthode	159
5.2 Hybridation, relations génétiques et phylogénétiques entre les espèces du complexe <i>Daphnia pulex</i>	161
5.2.1 Analyse des microsatellites	161
5.2.2 Conflits « gène nucléaire-haplotype mitochondrial ».....	164
5.2.3 Estimation de la part des différents génomes parentaux dans les génomes des hybrides	166
5.3 Évolution du gène <i>Rab4</i> dans le complexe <i>pulex</i>	167
5.4 Évolution de <i>Pokey</i> dans le complexe <i>pulex</i>	171
5.4.1 Effet du niveau de ploïdie les patrons d'insertion du transposon <i>Pokey</i>	171
5.4.2 Recombinaison intra-génique du transposon <i>Pokey</i>	175
5.5 Conclusion	178
Bibliographie	184

Liste des Figures

Figure	Page
1-1 Représentation schématique de différents types d'éléments transposables	5
1-2 Cycle de vie « classique » d'un transposon de classe II	8
1-3 Différents modes de reproduction de l'espèce <i>Daphnia pulex</i>	22
2-1 Bayesian reconstruction of the phylogenetic relationships among individuals of the <i>Daphnia pulex</i> complex	48
2-2 Nuclear DNA content of diploid clones of the <i>Daphnia pulex</i> complex	56
3-1 Phylogenies of ND5 and Rab4 partial sequences from 51 isolates of the <i>Daphnia pulex</i> complex	91
3-2 Principal Coordinate analyses and K-means analyses constructed using different distances on 105 microsatellite genotypes from the <i>Daphnia pulex</i> complex	98
3-3 Phylogenetic networks constructed using the NeighborNet algorithm on 105 microsatellite genotypes from the <i>Daphnia pulex</i> complex.....	99
3-4 Phylogenies of ND5 and Rab4 partial sequences from 51 isolates of the <i>Daphnia pulex</i> complex	107
3-5 Principal Coordinate analyses and K-means analyses constructed using different distances on 105 microsatellite genotypes from the <i>Daphnia pulex</i> complex	108

4-1 Phylogenetic network on the binary matrix obtained with the TE display technique from diploid and polyploid hybrid isolates of the <i>Daphnia pulex</i> complex.	125
4-2 Relation between <i>Pokey</i> insertion number sites and ploidy weighted heterozygosity (H_{pi}).....	127
4-3 Phylogenetic network constructed using the NeighborNet algorithm on the 53 <i>Pokey</i> sequences dataset from isolates of the <i>Daphnia pulex</i> complex	133
4-4 Unrooted Neighbor-Joining trees constructed with fragmented sequences flanked by recombination breakpoints highlighted by GARD analysis on a 53 sequences dataset	134
4-5 Distribution of recombination parameters calculated on the 53 <i>Pokey</i> sequences dataset using kwarg and DnaSP software	136
4-6 One of the possible ancestral recombination graphs constructed on <i>Pokey</i> sequences showing 89 recombination events.....	137
4-7 Phylogenetic network constructed using the NeighborNet algorithm with <i>Pokey</i> alleles encountered in the v1.1 draft genome of <i>Daphnia pulex</i>	142
5-1 Deux inférences phylogénétiques Bayésiennes différentes montrant l'insertion des introns 2a, 2b et 2b' dans le gène <i>Rab4</i>	170

Liste des Tableaux

Tableau	Page
2-1 Mitochondrial lineages (ND5 sequences), geographic origins, and habitats of <i>Daphnia</i> clones used in this study.....	38
2-2 Genetic and cytogenetic characteristics of <i>Daphnia</i> clones	50
2-3 Average 2C-values of <i>Daphnia</i> diploid species.....	55
2-4 Multilocus genotypes of each strains used in our study	63
2-5 Multilocus genotypes of each Argentinean <i>Daphnia</i>	66
3-1 Identification, ploidy, and geographic origins of <i>Daphnia</i> isolates used in the present study.....	75
3-2 Mean divergences (lower-left triangle) and standard errors (upper-right triangle) using the Maximum Composite Likelihood model between groups of ND5 haplotypes	92
3-3 SH-test results for the comparisons between ND5 and <i>rab4</i> phylogenetic trees (see Figure 3-1 and 3-4).....	95
3-4 Mean heterozygosities (\pm standard deviation) of diploids clustered in “pure <i>D. pulex</i> ”, “hybrid <i>D. pulex</i> ” and “ <i>D. pulicaria</i> ” groups (see Figure 3-3).....	100
4-1 Identification, ploidy level, and geographic origins of <i>Daphnia</i> isolates used in the present study.....	117

4-2 Results of the covariance analysis (ANCOVA) on TE display results using the ploidy level and H_{pi} as the independent variables.	126
4-3 Summary of recombination events detected on 53 <i>Pokey</i> sequences from the <i>Daphnia pulex</i> complex using RDP3 software	130
4-4 <i>Pokey</i> RFLP haplotypes amplified from 41 isolates of the <i>Daphnia pulex</i> complex.....	140
4-5 Profiles of <i>Pokey</i> insertions highlighted by the “TE display” technique.....	149
5-1 Tailles de génomes estimées pour quatre espèces appartenant aux genres <i>Daphnia</i> et <i>Ctenodaphnia</i>	158

Résumé

Les différentes composantes des génomes et leur agencement sont issus de divers processus moléculaires. L'activité des éléments transposables (i.e. séquences d'ADN capables de se déplacer et se multiplier dans le génome) et l'augmentation du niveau de ploïdie (i.e. ensemble du nombre de chromosomes du noyau cellulaire) semblent avoir eu des rôles importants dans l'évolution de l'architecture des génomes des eucaryotes pluricellulaires. La polyploïdie se rencontre principalement chez les plantes mais elle se rencontre épisodiquement aussi dans des taxons animaux. De plus plusieurs évènements de doublement de génome semblent avoir eu un impact sur la diversification des vertébrés. Les éléments transposables et la polyploïdie auraient des impacts non négligeables sur la diversification phénotypique des êtres vivants. Leur variation serait alors soumise à des processus sélectifs. Ces deux phénomènes pourraient interagir pour structurer les génomes. Selon différentes théories, l'augmentation du niveau de ploïdie aurait pour conséquence une augmentation du nombre d'insertions d'éléments transposables dans le génome. Les études empiriques principalement basées sur des plantes allopolyploïdes ont mis en évidence une activation des éléments transposables lors de la formation dans les génomes polyploïdes, toutefois la plupart de ces études montrent une perte du nombre d'insertions. Peu d'études se sont focalisées sur l'interaction entre la polyploïdie et les éléments transposables chez les animaux.

Cette thèse de doctorat avait plusieurs objectifs. D'une part, cette thèse devait éclaircir l'histoire évolutive du complexe d'espèces *Daphnia pulex*. D'autre part, à partir de cette description, il était possible de tester l'effet du niveau de ploïdie sur un élément transposable de classe II, *Pokey*, en nature dans le complexe *Daphnia pulex*. *Daphnia pulex* est un petit crustacé d'eau douce. Les espèces de ce complexe se retrouvent dans les étangs et les lacs d'Amérique du Nord et d'Europe. Il existe une variation du mode de reproduction dans les populations (parthénogénétique cyclique ou obligatoire). Des lignées hybrides diploïdes et polyploïdes ont des origines multiples. Ce complexe présente un patron de polyploïdie géographique avec des diploïdes dans les régions tempérées et des polyploïdes dans les régions subarctiques et arctiques. L'importance de l'hybridation et le niveau de ploïdie dans ce complexe ont été réévalués en utilisant différentes techniques de cytogénétique et de génétique des populations. L'impact de la polyploïdie sur la densité de *Pokey* a été évalué en utilisant une technique de « TE display ». L'évolution des séquences de l'élément *Pokey* dans les lignées hybrides a été décrite avec des analyses phylogénétiques.

Les données récoltées grâce à la cytométrie en flux combinées à l'analyse microsatellite ont révélé que la plupart des clones polyploïdes dans le complexe *D. pulex* sont triploïdes et non pas tétraploïdes comme le suggéraient les études antérieures. Les clones triploïdes sont probablement issus d'interactions entre des populations sexuées et asexuées. Diverses analyses sur les données microsatellites ont permis de séparer certaines populations en fonction des espèces auxquelles elles appartiennent et d'un groupe constitué de génotypes hybrides. Ce dernier groupe est composé d'hybrides diploïdes et polyploïdes avec des mitochondries *D. pulex* et certaines lignées avec des mitochondries *D. pulicaria*. Les analyses conjuguées des données microsatellites et d'un gène nucléaire ont permis d'éclaircir les relations évolutives entre un groupe d'espèces où le maintien du polymorphisme ancestral et l'hybridation affectent les relations phylogénétiques observées en n'analysant qu'un seul gène. Cette étude montre que l'hybridation et l'introgession ont joué un rôle important dans l'évolution de ce complexe. Le nombre d'insertions TE a été comparé entre des diploïdes (14) et des polyploïdes (13) hybrides en utilisant une technique de «TE display». Bien que le nombre moyen de sites d'insertion était plus élevé chez les polyploïdes que chez les diploïdes cette différence n'était pas statistiquement significative. De nombreux événements de recombinaisons ont été révélés dans des allèles de *Pokey* dans des lignées diploïdes et polyploïdes du complexe *pulex*. Des analyses phylogénétiques et de recombinaison ont montré que la recombinaison est une force majeure qui façonne l'évolution du transposon *Pokey*.

Cette thèse de doctorat présente donc divers patrons et processus génétiques se déroulant lors de la formation du complexe d'espèces *Daphnia pulex*. L'hybridation et la polyploïdie semblent avoir eu un impact significatif sur la diversification et possiblement sur l'adaptation de populations à différents environnements. L'absence d'augmentation du nombre d'insertions d'éléments transposables pourrait être expliquée par une augmentation de ceux-ci dans les lignées diploïdes hybrides parentes. L'intégration d'une vision écosystémique du génome pourrait permettre de mieux comprendre comment les génomes se structurent au cours de l'évolution. Des études d'écologie évolutive et expérimentales permettraient de répondre à différentes hypothèses proposées lors de cette étude.

Chapitre I: Introduction générale

1.1 Polyploïdie et éléments transposables

1.1.1 Diversité des génomes

Différentes structures composent les génomes des organismes vivants (Gregory, 2005 ; Lynch, 2007). Le génome ne se résume pas qu'à un ensemble de gènes codant pour des protéines. Il contient aussi des introns (i.e. séquences « non codantes » insérées à l'intérieur d'un gène), des séquences répétées non codantes (i.e. microsatellites, minisatellites, etc.), des pseudogènes (i.e. gènes ayant perdu leur capacité à produire des protéines viables), et des éléments mobiles (par exemple des éléments transposables). La proportion de ces diverses structures diffère entre les génomes d'organismes appartenant à différents règnes (Lynch, 2007), voire à différentes espèces. Les eubactéries sont quasiment dépourvues d'introns alors que les génomes des eucaryotes sont très variables et ont de 10^2 à 10^6 introns par génome (Roy et Gilbert, 2006). Chez les eucaryotes, les éléments génétiques mobiles peuvent représenter moins de 3% de la totalité du génome chez la levure à fission (*Schizosaccharomyces pombe*) à près de 80% chez l'orge (*Hordeum vulgare*) (Pidpala, Yatsishina et Lukash, 2008). Différents mécanismes génétiques semblent être à l'origine de la variabilité des structures des génomes (Lynch, 2007). Les mutations ponctuelles ne peuvent pas expliquer, à elles seules, la diversité des génomes des organismes vivants. Un des paramètres qui est le plus intrigant concernant la diversité des génomes est la quantité d'ADN nucléaire par cellule. En effet, cette quantité varie de manière importante à l'intérieur des embranchements phylogénétiques et parfois même entre espèces proches (Gregory, 2005 ; Petrov, 2001). Cette variation ne semble pas être en relation avec

la « complexité » morphologique des organismes, par exemple l'amibe *Amoeba dubia* a une quantité d'ADN nucléaire 200 fois plus grande que celle de l'homme *Homo sapiens sapiens* (Friz, 1968). La variation de la quantité d'ADN nucléaire est fréquemment accompagnée de modifications phénotypiques (Gregory, 2002 ; Gregory et Mable, 2005 ; Levin, 1983 ; Otto et Whitton, 2000 ; Vinogradov, 1995, 1997) telles que des modifications de la taille des noyaux, de la taille des cellules qui les contiennent (Gregory, 2001) et de l'activité enzymatique (i.e. enzymes impliquées dans la production d'acides aminés et au niveau de la surface cellulaire) à l'intérieur de ces dernières (Cavalier-Smith, 1978 ; Weiss, Kukora et Adams, 1975). Ces bouleversements cellulaires ont comme conséquences des altérations de la taille des organismes chez les espèces à croissance indéterminée, de leur physiologie (Cavalier-Smith, 1978 ; Weiss, Kukora et Adams, 1975), de leur taux de croissance et/ou de leur mode de reproduction (Lundmark, 2006 ; Otto et Whitton, 2000). Ces modifications phénotypiques sont le « matériel » sur lequel la sélection naturelle agit (Levin, 1983 ; Otto et Whitton, 2000). Divers mécanismes génétiques sont sous-jacents à l'évolution de la quantité d'ADN nucléaire. Des insertions et des délétions de quelques nucléotides ou de séquences complètes (duplication génique ou réplication d'éléments mobiles par exemple) modifient la taille du génome haploïde (Petrov, 2001 ; SanMiguel *et al.*, 1998 ; Vitte et Panaud, 2005). Elle peut être brusque par l'addition de chromosomes complets (aneuploïdie) ou même de génomes complets (polyploïdie; Levin, 1983 ; Otto et Whitton, 2000).

1.1.2 Importance des éléments transposables

Depuis la découverte des éléments transposables (Figure 1-1) à l'intérieur du génome du maïs (*Zea mays* L.) par Barbara McClintock dans les années 1940 (McClintock, 1984), un nombre croissant d'études montre qu'ils jouent un rôle important dans l'évolution des génomes (Böhne *et al.*, 2008 ; Capy *et al.*, 2000 ; Kidwell et Lisch, 2000 ; McDonald, 1995 ; Pritham, 2009). Ces séquences d'ADN ont

la propriété de se déplacer de manière répliquative (comme un copier-coller dans un traitement de texte) ou non (couper-coller) à l'intérieur des génomes. Par leur action les éléments transposables peuvent contribuer à la variation de la taille des génomes (Kidwell, 2002 ; SanMiguel et Bennetzen, 1998 ; Vitte et Panaud, 2005) et aux variations phénotypiques qui sont observées chez les organismes vivants (Aminetzach, Macpherson et Petrov, 2005 ; González *et al.*, 2008 ; Kidwell et Lisch, 1997 ; Lerman *et al.*, 2003 ; Promislow, Jordan et McDonald, 1999).

1.1.2.1 Diversité des éléments transposables

Chez les eucaryotes, les éléments transposables sont divisés en différentes classes (Figure 1-1; Böhne *et al.*, 2008 ; Wicker *et al.*, 2007). La classe I est constituée d'éléments transposables, communément appelés rétrotransposons, transcrits en ARN par l'ARN polymérase du génome hôte, recodé en ADN via une transcriptase inverse et réinséré dans le génome. Les rétrotransposons sont divisés en deux sous-classes: les rétrotransposons à LTR (Long Terminal Repeat) et les rétrotransposons sans-LTR. Ces derniers se divisent encore en deux groupes, les LINEs (Long Interspersed Elements) et SINEs (Short Interspersed Elements) (Böhne *et al.*, 2008). Les LINEs sont constitués de deux gènes principaux codant pour une reverse transcriptase et une endonucléase. Les SINEs ne codent pour aucune protéine leur permettant de se transposer mais sont évolutivement proches des LINEs. Les transposons (classe II) se déplacent à l'intérieur du génome sans passer par la phase ARN. Cette classe regroupe différents éléments transposables (Wicker *et al.*, 2007). Les transposons de la sous-classe 1 sont des éléments flanqués de séquences terminales inversées répétées (TIR). Ils codent pour une transposase permettant au transposon de s'intégrer à un autre site dans le génome. Des transposons ne codant pour aucune protéine qui leur permettrait de se déplacer dans le génome mais possédant des TIRs (MITEs) ont été rencontrés dans de nombreux organismes (Böhne *et al.*, 2008 ; Casacuberta et Santiago, 2003 ;

Feschotte, Zhang et Wessler, 2002 ; Shan *et al.*, 2005). Les transposons de la sous-classe 2 regroupent des transposons découverts récemment (Hélitrons et Mavericks) et ne possédant pas de gène codant pour une transposase (Kapitonov et Jurka, 2001 ; Pritham, Putliwala et Feschotte, 2007 ; Wicker *et al.*, 2007). Les Hélitrons codent pour une hélicase produisant un simple brin d'ADN qui s'intégrera par la suite à un site cible (Kapitonov et Jurka, 2001). Les transposons Mavericks ou Polintons codent pour plusieurs protéines dont une intégrase et une protéine homologue d'ADN polymérase (Pritham, Putliwala et Feschotte, 2007). Le mécanisme de transposition de ces éléments pourrait impliquer une réplication extrachromosomale par une polymérase et une intégration de l'ADN double-brin par une intégrase (Böhne *et al.*, 2008). Les éléments transposables sont dits autonomes s'ils possèdent les gènes codant pour leur propre transposition et non autonomes si les éléments transposables sont dépendants des enzymes codés en *trans* par un autre élément mobile ou par le génome hôte (Kidwell et Lisch, 2001 ; Shan *et al.*, 2005).

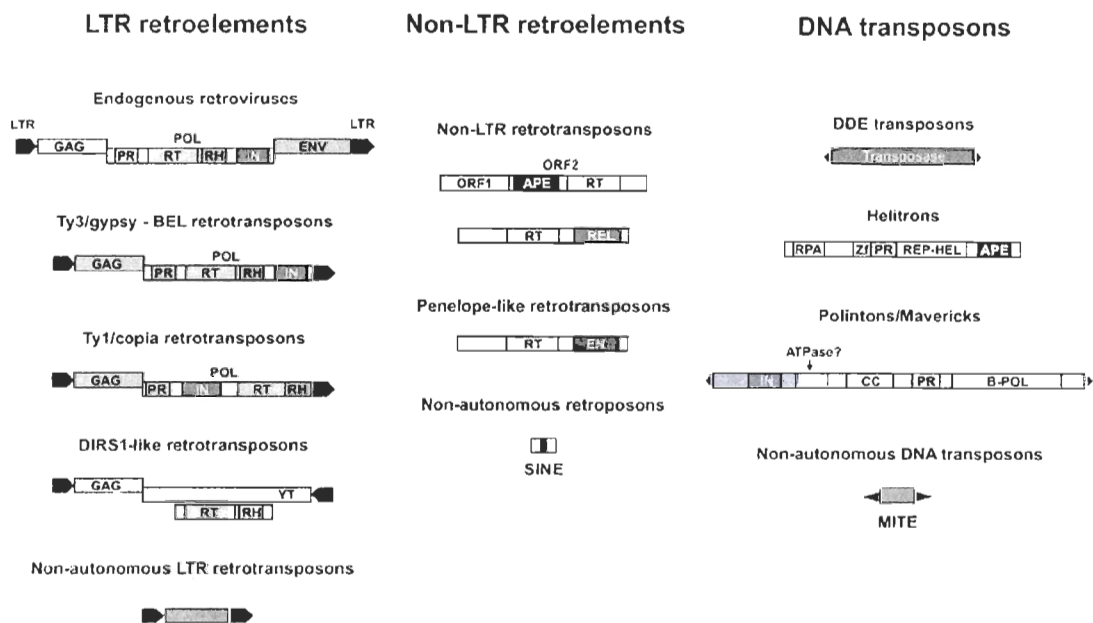


Figure 1-1 Représentation schématique de différents types d'éléments transposables
 APE, endonucléase; B-POL, ADN polymérase de type B; CC, domaine à faisceau d'hélices;
 EN, endonucléase; ENV, enveloppe; IN, intégrase; LTR, répétition terminale longue; ORF,
 cadre ouvert de lecture; PR, protéase; REL, endonucléase ressemblant à un enzyme de
 restriction; REP-HEL, initiateur à réplication circulaire et hélicase; RH, ribonucléase H;
 RPA, domaine de reconnaissance de protéine A de réplication se fixant à de l'ADN simple brin;
 RT, transcriptase inverse; YT, transposase à tyrosine; Zf, zinc finger. (Schéma tiré de Böhne *et al.*, 2008)

1.1.2.2 Éléments transposables et sélection naturelle

Dû à leur prévalence dans le génome du maïs, Barbara McClintock avait supposé un rôle bénéfique des éléments transposables pour l'hôte (Gonzalez et Petrov, 2009). Toutefois, Hickey (1982) a suggéré que la prévalence des éléments transposables pourrait être expliquée par leur seule capacité à se répliquer. Selon différentes études, la distribution des éléments transposables essentiellement dans les parties non codantes du génome de *Drosophila melanogaster* aurait pour cause une sélection purificatrice des insertions aux niveaux des gènes codants (Bartolomé, Maside et Charlesworth, 2002 ; Biémont, 1992 ; Dolgin et Charlesworth, 2008 ; Finnegan, 1992 ; Montgomery, Charlesworth et Langley, 1987). Les éléments transposables ont donc été vus comme de l'ADN parasite amenant une baisse de la fitness de leur hôte (Doolittle et Sapienza, 1980 ; Hickey, 1982 ; Orgel et Crick, 1980 ; Yoder, Walsh et Bestor, 1997). Toutefois des insertions en amont ou dans des gènes codant sont fixées dans certaines populations (Gonzalez et Petrov, 2009). Certaines mutations induites par l'insertion d'éléments mobiles pourraient être adaptatives pour leur hôte (Kidwell et Lisch, 1997 ; Kidwell et Lisch, 2001 ; McDonald, 1995). L'insertion de transposons dans des régions promotrices de certains gènes peuvent modifier leur expression et avoir une forte influence sur les phénotypes des organismes qui les portent (Britten, 1996 ; Kidwell et Lisch, 1997). Chez la drosophile, des éléments transposables sont impliqués dans la réponse à des stimuli extérieurs et auraient des rôles adaptatifs tels que la résistance aux insecticides ou l'adaptation à la température (Gonzalez et Petrov, 2009). Chez la drosophile, les télomères, séquences assurant l'intégrité des chromosomes, seraient protégés par deux rétrotransposons (Pardue *et al.*, 2005). Des éléments transposables « domestiqués », i.e. de nouveaux gènes avec des fonctions bénéfiques pour l'hôte seraient issus d'éléments transposables auraient été rencontrés dans divers génomes séquencés (International Human Genome Sequencing Consortium, 2001 ; Sarkar *et al.*, 2003 ; Volff, 2006).

1.1.2.3 Dynamique et distribution des éléments transposables

Les éléments transposables sont présents dans la quasi-totalité des organismes vivants (Pritham, 2009 ; mais voir Shigenobu *et al.*, 2000) et la proportion du génome qu'ils constituent est variable chez les eucaryotes (Pidpala, Yatsishina et Lukash, 2008). Ils sont une source de variabilité importante, représentant à peu près de 50 à 85 % des mutations spontanées chez *Drosophila melanogaster* (McDonald, 1998).

Le cycle de vie classique d'un transposon de classe II (Figure 1-2) commence par son insertion dans un génome, s'ensuit une phase active où il se déplace et/ou se multiplie et enfin, dans la majorité des cas une mise sous-silence par des phénomènes génétiques ou épigénétiques dépendant de l'hôte et une accumulation de mutations conduisant à une inactivation (Kidwell et Lisch, 2001). Les éléments transposables peuvent se transmettre soit des parents vers leur progéniture (i.e. transfert vertical) soit d'un organisme vers un autre sans lien de parenté (i.e. transfert horizontal). Les mécanismes de transferts horizontaux sont mal connus mais pourraient se faire via des infections virales, bactériennes ou parasitaires (Pace *et al.*, 2008 ; Yoshiyama *et al.*, 2001). Les éléments transposables de classe II, sont plus enclins aux transferts horizontaux que les rétrotransposons. Ceci est probablement dû au fait que les molécules d'ADN sont plus stables chimiquement que les molécules d'ARN (Silva, Loreto et Clark, 2004).

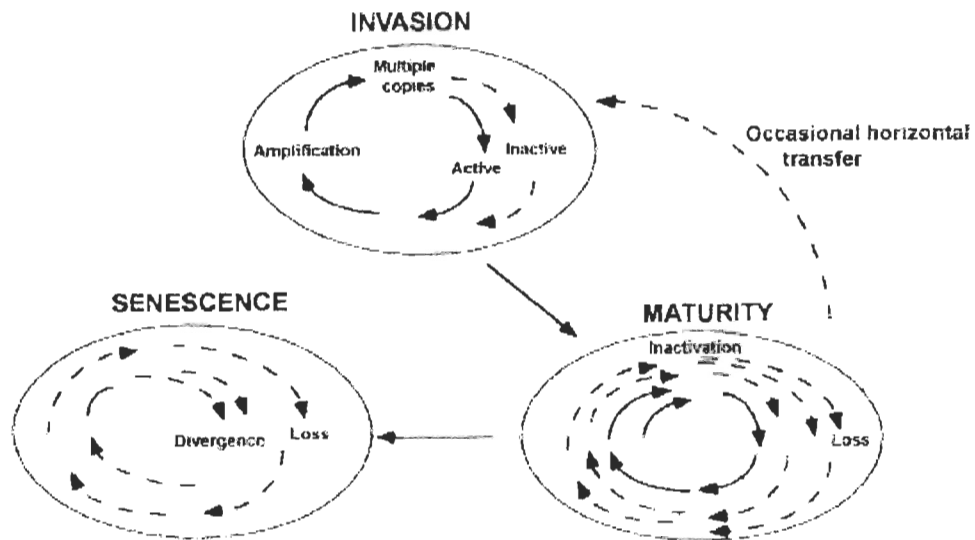


Figure 1-2 Cycle de vie « classique » d'un transposon de classe II
Trois étapes principales sont illustrées; une phase rapide d'invasion où l'amplification du nombre de copies se produit accompagné de mutations qui rendent certains éléments inactifs; une phase de maturité où l'amplification et la perte des transposons sont à l'équilibre; et une étape de sénescence où tous les éléments autonomes ont été perdus, aucun amplification se produit et les transposons non autonomes sont perdus ou divergent. Les éléments autonomes sont représentés par les lignes continues et les éléments non autonomes par des tirets (Figure tirée de Kidwell et Lisch, 2001)

La dynamique des populations hôtes a un impact important sur la dynamique des populations d'éléments transposables (Charlesworth et Charlesworth, 1983 ; Charlesworth, Langley et Sniegowski, 1997 ; Charlesworth et Wright, 2001). Une augmentation du nombre d'insertions d'éléments transposables dans les génomes hôtes est attendue si la taille effective des populations hôtes baisse (Brookfield et Badge, 1997), ou si le niveau d'endogamie (i.e. union entre individus appartenant à un groupe ou à un sous-groupe et qui sont le plus souvent apparentés) augmente dans la population (Wright et Schoen, 1999). Le mode de reproduction a une influence sur la dynamique des éléments transposables et leur densité (i.e. nombre d'insertions) dans le génome (Bestor, 1999 ; Charlesworth et Charlesworth, 1995 ; Schaack *et al.* ; Schaack *et al.*, 2010b ; Wright et Finnegan, 2001 ; Zeyl, Bell et Green, 1996). La plupart des études ont considéré les éléments transposables comme des mutations neutres ou réduisant la « fitness » de leur hôte. Même s'ils sont légèrement délétères pour leur hôte, les éléments transposables devraient se répandre dans les différents génomes hôtes *via* les phénomènes de recombinaison et de fécondation croisée entre individus distincts (i.e. allogamie) dans des populations sexuées (Wright et Finnegan, 2001) alors qu'en absence de transfert horizontal, l'accroissement des éléments transposables serait empêché dans les populations asexuées (Arkhipova et Meselson, 2000 ; Gladyshev, Meselson et Arkhipova, 2008). Des études empiriques ont testé ces hypothèses (Arkhipova et Meselson, 2000 ; Gladyshev, Meselson et Arkhipova, 2008 ; Schaack *et al.*, 2010b ; Valizadeh et Crease, 2008 ; Zeyl, Bell et Green, 1996). Par exemple, les transposons et rétrotransposons se retrouvent dans les génomes des rotifères pour lesquels la reproduction sexuée est connue. Dans les génomes des rotifères Bdelloïdes, dont les formes sexuées sont inconnues, les éléments transposables de classe II sont continuellement intégrés *via* des transferts horizontaux alors que les rétrotransposons ont été éliminés et sont absents (Arkhipova et Meselson, 2000 ; Gladyshev, Meselson et Arkhipova, 2008).

La distribution des éléments transposables à l'intérieur des génomes ne semble pas être aléatoire (Biéumont, 1992 ; Duret, Marais et Biéumont, 2000). Elle

semble être dépendante à la fois du type d'éléments transposables (par exemple, LTR-retroélément vs. transposon à ADN) et des organismes hôtes (Bartolomé, Maside et Charlesworth, 2002 ; Duret, Marais et Biemont, 2000 ; Eickbush et Eickbush, 1995 ; Kim *et al.*, 1998 ; Penton, Sullender et Crease, 2002 ; Rizzon *et al.*, 2002 ; Wright, Agrawal et Bureau, 2003). Cette répartition non aléatoire serait, d'une part la conséquence d'une contre-sélection des insertions au niveau des régions codantes des gènes, et d'autre part, les transposons s'inséreraient préférentiellement, en fonction de la séquence cible, dans certaines parties du génome (« hot spot »). Cette observation a amené Kidwell & Lisch (2001) à proposer que les éléments transposables occupent des niches écologiques spécifiques dans le génome.

1.1.3 Importance de la polyplœidie chez les eucaryotes

La polyplœidie, i.e. la possession de plus de deux ensembles complets de chromosomes, a des implications importantes dans l'écologie et l'évolution des eucaryotes (Ohno, 1970 ; Otto, 2007 ; Otto et Whitton, 2000 ; Soltis et Soltis, 2000 ; Stebbins, 1985). Chez les végétaux, 30% à 80% des angiospermes semblent issus d'évènements de polyplœidisation (Masterson, 1994 ; Soltis et Soltis, 2000). La plupart des fougères (Pteridophyta) semblent polyplœides (Otto et Whitton, 2000 ; Soltis et Soltis, 1999). Chez les animaux, la fréquence des espèces polyplœides semble plus faible (Mable, 2004 ; Otto et Whitton, 2000). Cette plus faible prévalence ne signifie pas que l'importance évolutive de la polyplœidisation est moindre chez les animaux. Des espèces polyplœides ont été répertoriées dans la plupart des embranchements animaux connus (Gregory et Mable, 2005 ; Mable, 2004 ; Otto, 2007 ; Otto et Whitton, 2000). Des espèces polyplœides sont retrouvées chez des amphibiens (Hughes et Hughes, 1993), des poissons à nageoires rayonnées (Hoegg et Meyer, 2005 ; Liu *et al.*, 2009), des annélides (Viktorov, 1997), des crustacés (Dufresne et Hebert, 1995 ; Turgeon et Hebert, 1994) et des insectes

(Lokki et Saura, 1980 ; Saura, Lokki et Suomalainen, 1993). L'évolution des vertébrés semble avoir eu lieu *via* deux évènements de doublement de génome (Ohno, 1970) et peut-être un troisième qui a permis aux poissons à nageoires rayonnées de se diversifier (Hoegg et Meyer, 2005). La levure, *Saccharomyces cerevisiae*, semble avoir subi une duplication de son génome il y a à peu près 100 Millions d'années (Piskur *et al.*, 2006). De plus, cette apparente faible fréquence chez les animaux peut être due à un plus faible effort d'échantillonnage par les zoologistes (Mable, 2004).

La polyploïdie peut résulter de différents phénomènes cellulaires (Ramsey et Schemske, 1998). Un organisme polyploïde peut être soit issu d'un doublement de chromosomes dû à une perturbation de la division cellulaire d'un œuf fécondé (i.e. polyploïdie somatique, Ramsey et Schemske, 1998), soit à la fusion de gamètes contenant plus d'un ensemble complet de chromosomes (gamètes non-réduits, Bretagnolle et Thompson, 1995) ou à de la polyspermie (i.e. fécondation d'un gamète femelle par plus d'un gamète mâle). Les polyploïdes peuvent être soit allopolyploïdes, i.e. issus d'un croisement entre espèces différentes, soit autopolyploïdes, i.e. issus de croisement d'individus de la même espèce ou du doublement de génome d'un même individu par autofécondation ou duplication somatique.

Diverses modifications morphologiques (taille des graines chez les plantes par exemple), physiologiques (taux de développement par exemple), dans l'expression des gènes et dans la structure du génome, sont induites par les évènements de polyploïdisation (Adams *et al.*, 2003 ; Adams et Wendel, 2005a ; Madlung *et al.*, 2005 ; Osborn *et al.*, 2003 ; Otto et Whitton, 2000 ; Salmon, 2005 ; Scheid *et al.*, 1996 ; Soltis et Soltis, 1995, 1999). Une hétérozygotie élevée par rapport aux populations diploïdes apparentées est attendue dans les populations polyploïdes que ce soit chez allopolyploïdes où l'hétérozygotie serait fixée ou que ce soit chez les autopolyploïdes à ségrégation tétrasomique (Soltis et Soltis, 2000). La possession de plusieurs copies du même gène et une forte hétérozygotie permettrait

une meilleure flexibilité phénotypique de l'organisme polyploïde. Ces modifications génomiques et phénotypiques pourraient expliquer le succès des organismes polyploïdes dans des environnements marginaux et hétérogènes (Dufresne et Hebert, 1994 ; Lundmark, 2006 ; Otto et Whitton, 2000). Certains organismes polyploïdes occupent des niches écologiques différentes des espèces diploïdes apparentées (Otto et Whitton, 2000). Les populations polyploïdes ont souvent une plus grande distribution géographique et se retrouvent à de plus hautes latitudes que les populations diploïdes évolutivement proches (Brochmann *et al.*, 2004 ; Levin, 1983 ; Otto et Whitton, 2000). Chez les allopolyploïdes, l'effet de l'hybridation peut se confondre avec l'effet de l'ajout de génome. L'hybridation est un autre phénomène qui pourrait expliquer le succès des organismes polyploïdes dans les milieux extrêmes et hétérogènes (Kearney, 2005). L'hybridation joue un rôle important dans la diversification des animaux et des végétaux (références citées dans Dowling et Secor, 1997). Des modifications de l'expression des gènes se produisent lors d'événements d'hybridation interspécifique (Fontdevila, 2005 ; O'Neill, O'Neill et Graves, 1998 ; Shan *et al.*, 2005). Lundmark (Lundmark, 2006) suggère que l'hybridation n'est qu'un pas évolutif afin de former un génome polyploïde plus adapté à des milieux extrêmes. L'hybridation et la polyploïdisation pourraient être des phénomènes complémentaires dans les processus d'adaptations à des milieux marginaux (Otto et Whitton, 2000). Ils auraient un effet synergique conduisant au succès écologique des individus polyploïdes. Comme le remarque Lundmark (2006), il est difficile de dissocier l'effet de l'augmentation du niveau de ploïdie de celui de l'hybridation, surtout dans le règne animal où les espèces autopolyploïdes sont rares (mais voir Otto & Whitton 2000).

1.1.4 Interaction entre polyploïdie et éléments transposables

Les éléments transposables pourraient jouer un rôle dans le maintien et le succès des organismes polyploïdes (Matzke et Matzke, 1998). De nombreuses

données chez les animaux et les végétaux montrent que les éléments transposables se comportent différemment quand ils sont confrontés à certains stress (Capy *et al.*, 2000 ; Labrador et Fontdevila, 1994 ; Shan *et al.*, 2005 ; Wessler, 1996). Chez les plantes, l'hybridation et la polyploïdisation peuvent avoir pour conséquence l'activation de certains éléments transposables (Kashkush, Feldman et Levy, 2002, 2003 ; Shan *et al.*, 2005). Cette activation serait due, au relâchement dans les génomes d'individus hybrides et/ou polyploïdes de phénomènes épigénétiques mettant sous silence les éléments transposables dans les génomes des individus géniteurs (Ainouche *et al.*, 2009 ; Comai, 2000 ; Madlung *et al.*, 2002 ; Madlung *et al.*, 2005 ; Scheid *et al.*, 1996 ; Shaked *et al.*, 2001).

Différentes théories, prédisent une activation des éléments transposables et/ou une augmentation du nombre de leurs insertions dans les génomes polyploïdes par rapport aux génomes diploïdes des individus apparentés (Comai, 2000 ; Lynch, 2007 ; Matzke et Matzke, 1998 ; McClintock, 1984 ; Parisod *et al.*, 2010). La rencontre de génomes divergents pourraient relâcher les contraintes biochimiques qui mettent sous silence les éléments transposables (hypothèse du "choc génomique"; Comai, 2000 ; McClintock, 1984). La redondance de niches génomiques due à l'addition de génomes (hypothèse de la redondance; Matzke et Matzke, 1998) ou la réduction de taille effective des populations chez les polyploïdes (hypothèse du bottleneck; Lynch, 2007) pourraient avoir pour conséquence une augmentation du nombre d'insertions d'éléments transposables dans les génomes polyploïdes (Parisod *et al.*, 2010). Une activation et une augmentation du nombre d'éléments transposables pourraient avoir un impact important sur le réarrangement du génome polyploïde qui les contient (McClintock, 1984) et pourrait mener à une diversification phénotypique chez les organismes polyploïdes (Comai, 2000 ; Madlung *et al.*, 2002). Cette diversification phénotypique pourrait produire des individus sélectionnés positivement dans des environnements hétérogènes, instables et/ou marginaux (Leitch et Leitch, 2008 ; Otto et Whitton, 2000 ; Wendel, 2000).

Chez les plantes, diverses études ont montré une modification du patron d'expression des éléments transposables après la production d'allopolyploïdes et d'autopolyploïdes (Ainouche *et al.*, 2009 ; de Araujo *et al.*, 2005 ; Hanson *et al.*, 2000 ; Kashkush, Feldman et Levy, 2002 ; Liu et Wendel, 2003 ; Madlung *et al.*, 2002 ; passé en revue dans Parisod *et al.*, 2010) et une certaine instabilité du génome (Madlung *et al.*, 2005). À plus ou moins long terme, le nombre d'insertions d'éléments transposables n'est pas significativement plus grand entre certains polyploïdes et les diploïdes apparentés et ont même tendance à diminuer après l'augmentation du niveau de ploïdie (Parisod *et al.*, 2010).

1.1.5 Importance d'un modèle animal et objectif de la thèse

La quasi-totalité des études portant sur les modifications génétiques et épigénétiques (c'est à dire les modifications d'expression génique héritable ne résultant pas de l'altération des séquences d'ADN) liées au changement de niveau de ploïdie implique des espèces végétales (Adams *et al.*, 2003 ; Adams et Wendel, 2005a, 2005b ; Ainouche *et al.*, 2009 ; Kashkush, Feldman et Levy, 2002, 2003 ; Madlung *et al.*, 2002 ; Madlung *et al.*, 2005 ; Osborn *et al.*, 2003 ; Parisod *et al.*, 2010 ; Scheid *et al.*, 1996 ; Wendel, 2000 ; Zhao *et al.*, 1998). Très peu d'études se sont focalisées sur l'impact de la polyploïdie sur la dynamique et l'évolution des éléments transposables chez les animaux (mais voir Liu *et al.*, 2009). La plupart des études concernant la polyploïdie chez les animaux ont décrit les patrons de distribution des populations présentant différents niveaux de ploïdie (Adamowicz *et al.*, 2002 ; Ghiselli *et al.*, 2007 ; Holloway *et al.*, 2006 ; Johnson, 2006 ; Mergeay *et al.*, 2008 ; Stenberg *et al.*, 2003 ; Tomiuk et Loeschcke, 1992 ; Viktorov, 1997 ; Ward *et al.*, 1994). Une seule étude expérimentale s'est focalisée sur la dynamique des transposons après hybridation et polyploïdisation entre deux espèces de carpes (Liu *et al.*, 2009). Celle-ci présente une augmentation du nombre d'insertions d'éléments transposables dans les génomes de progénitures tétraploïdes fertiles mais pas dans

celles des progénitures triploïdes stériles. Cette étude n'a toutefois pas dissocié l'effet de l'hybridation de l'effet de la polyploïdisation *per se* sur la dynamique et la densité des éléments transposables dans les génomes. En effet, plusieurs études ont montré que les phénomènes d'hybridation pourraient induire une activation des éléments transposables chez les animaux (Fontdevila, 2005 ; O'Neill, O'Neill et Graves, 1998). L'augmentation du nombre d'insertions dans les génomes allopolyploïdes des carpes pourrait être la conséquence de l'hybridation et non de l'addition de génome(s) surnuméraire(s). De plus, les carpes hybrides ont été produites expérimentalement entre des espèces divergentes (genres différents). Il n'est pas évident que ces événements de croisements et de polyploïdisation se produisent en nature. Il est donc important d'étudier d'autres organismes animaux comptant des individus diploïdes et polyploïdes produit naturellement afin de mieux comprendre la dynamique et l'évolution des éléments transposables dans les génomes des animaux polyploïdes. Les données récoltées permettraient de mieux comprendre l'évolution de la structure des génomes chez les animaux.

L'objectif de cette thèse de doctorat est de tester l'effet de l'augmentation du niveau de ploïdie sur le nombre d'insertions d'un élément transposable de classe II en nature chez un organisme animal. Pour cela, diverses lignées diploïdes et polyploïdes provenant du complexe d'espèces *Daphnia pulex* ont été étudiées.

1.2 Le complexe *Daphnia pulex* : Un modèle animal approprié?

1.2.1 Taxonomie et écologie

Le complexe d'espèces *Daphnia pulex* rassemble diverses espèces comportant des lignées hybrides, diploïdes ou polyploïdes (Colbourne *et al.*, 1998). Il est difficile d'identifier taxonomiquement les espèces de ce complexe à partir simplement des caractéristiques morphologiques en raison d'une forte stase

morphologique (ou de la convergence), d'une forte plasticité phénotypique, et de phénomènes d'hybridation interspécifique (Colbourne et Hebert, 1997). Ces espèces se différencient le plus souvent par leur habitat, leur distribution géographique ou des caractères subtils. De nombreuses études se sont penchées sur les relations phylogénétiques et la polyploïdie dans le complexe *D. pulex* (Adamowicz *et al.*, 2002 ; Aguilera *et al.*, 2007 ; Colbourne *et al.*, 1998 ; Colbourne et Hebert, 1996 ; Lehman *et al.*, 1995 ; Mergeay *et al.*, 2008 ; Ward *et al.*, 1994 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b). L'étude de gènes mitochondriaux a éclairci les relations phylogénétiques à l'intérieur du genre *Daphnia* (Colbourne *et al.*, 1998 ; Lehman *et al.*, 1995). Il est toutefois important de prendre en considération le niveau de ploïdie, les phénotypes allozymiques et les caractéristiques écologiques afin d'identifier les différentes espèces du complexe *pulex* (Hebert, 1995). À partir de la diversité nucléotidique au niveau du gène mitochondrial *NADH5*, ce complexe a pu être séparé en trois groupes: le groupe *pulicaria*, le groupe *tenebrosa* et la lignée appelée « *D. pulex* européennes » (Colbourne *et al.*, 1998 ; Marková *et al.*, 2007). Le groupe *pulicaria* est constitué des espèces *D. arenata*, *D. melanica* (toutes les deux ont été caractérisées taxonomiquement seulement récemment et semblent être endémiques de l'Oregon, U.S.A. (Hebert, 1995)), *D. middendorffiana*, *D. pulex* dite panarctique et trois lignées de *D. pulicaria* dites de l'est, de l'ouest et polaire. Le groupe *tenebrosa* est constitué des espèces *D. tenebrosa* et « *D. pulicaria* européenne » (Colbourne *et al.*, 1998). *D. pulex* et *D. pulicaria* semblent être dominantes du point de vue de la distribution géographique en Amérique du Nord (Hebert et Finston, 2001b ; Hebert *et al.*, 1993b). Ces deux espèces sont difficilement distinguables du point de vue morphologique mais sont distinctes au niveau de leur écologie (Hebert, 1995). *D. pulex* vit dans les étangs alors que *D. pulicaria* est trouvé dans les lacs (Hebert, 1995 ; Hebert *et al.*, 1993b ; Heier et Dudycha, 2009) dans l'est de l'Amérique du Nord (Hebert, 1995). Dans l'ouest de l'Amérique du Nord ces deux espèces peuvent partager le même habitat dans les étangs (Brooks, 1957a ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b) mais *D. pulicaria* est dominante dans les lacs (Hebert *et al.*, 1993b). Ces deux espèces sont

identifiées grâce à leur profil allozymique au niveau de la protéine LDH (Lactate déshydrogénase; Hebert *et al.*, 1989 ; Hebert *et al.*, 1993b). Les populations de l'espèce *D. pulex* sont fixées pour la forme allozymique la plus lente (S) alors que celles de *D. pulicaria* sont fixées pour la plus rapide (F) (Hebert *et al.*, 1989). *D. arenata* semble endémique de l'ouest de la chaîne des Cascades dans l'Oregon. Cette espèce est commune dans des étangs de forêt, mais se retrouve parfois en faible densité dans les lacs. Elle n'a été différenciée taxonomiquement qu'en 1995 sur la base de phénotypes allozymiques particuliers (Hebert, 1995). Certains chercheurs considèrent qu'elle ne représente qu'une population particulière de l'espèce *D. pulex* (Li *et al.*, 2009 ; Omilian, Scofield et Lynch, 2008). *D. melanica* est aussi une espèce dont la description taxonomique est récente (Hebert, 1995), elle a été répertoriée dans des étangs de dune (Pfrender, Spitze et Lehman, 2000) et dans des étangs de montagne en Oregon (T. J. Crease communication personnelle). *D. middendorffiana sensu stricto* est largement distribuée en Amérique du Nord arctique (Dufresne et Hebert, 1997 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b). Son habitat est restreint aux étangs peu profonds de toundra où le copépode *Heterocope septentrionalis* est absent. Cette espèce est de nature polyploïde (Colbourne *et al.*, 1998 ; Dufresne et Hebert, 1997). L'espèce *D. tenebrosa* a une distribution holarctique. En Amérique du Nord, sa distribution géographique s'étend de l'Alaska à l'ouest de la Baie d'Hudson (Dufresne et Hebert, 1994, 1995 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b). Son habitat correspond aux étangs et aux lacs sans poisson de la transition forêt boréale/toundra et se retrouve dans les étangs de toundra où *H. septentrionalis* est présent. Elle se différencie de *D. middendorffiana* par ses juvéniles qui possèdent une bosse dorsale (Hebert, 1995). Les lignées *D. pulex* européennes et *D. pulicaria* européennes semblent se retrouver dans les étangs et lacs européens respectivement (Marková *et al.*, 2007) mais peuvent aussi se rencontrer dans les mêmes habitats (T. J. Crease, communication personnelle).

1.2.2 Hybridation et polyploïdisation

Les espèces appartenant aux complexes d'espèces (i.e. assemblages d'espèces proches dans un genre ou assemblages de sous-espèces dans une espèce avec de nombreuses sous-espèces; Mayr, 1963) sont souvent évolutivement jeunes et parfois partiellement différenciées, des populations hybrides se retrouvent dans plusieurs complexes d'espèces (passé en revue dans Seehausen, 2004). Le complexe d'espèces *D. pulex* semble s'être diversifié à partir du Pliocène (à peu près 3 millions d'années avant notre ère; Colbourne *et al.*, 1998) et différentes populations hybrides existent dans ce complexe (Adamowicz *et al.*, 2002 ; Dufresne et Hebert, 1994 ; Hebert *et al.*, 1989 ; Innes et Hebert, 1988 ; Weider *et al.*, 1999b). Des populations diploïdes hybrides entre des femelles *D. pulex* et des mâles *D. pulicaria* se rencontrent dans des milieux perturbés (Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Hebert et Finston, 2001b ; Innes et Hebert, 1988). Ces hybrides ont un phénotype [SF] pour l'enzyme LDH (i.e. possession de l'allèle S diagnostique de *D. pulex* et de l'allèle F diagnostique de *D. pulicaria*) et un haplotype diagnostique de l'espèce *D. pulex* au niveau du gène mitochondrial *NADH5*. Des souches diploïdes hybrides viables entre ces deux espèces ont été produites en laboratoire (Heier et Dudycha, 2009 ; Innes et Hebert, 1988).

En Amérique du Nord, le complexe *pulex* tend à avoir un patron de polyploïdie géographique avec des populations polyploïdes dans l'arctique et le bas-arctique et des populations diploïdes dans des environnements plus tempérés (Beaton et Hebert, 1988 ; Weider, 1987). Des lignées polyploïdes ont été observées dans une région tempérée d'Argentine (Adamowicz *et al.*, 2002) et dans certains lacs montagneux de Bolivie (Aguilera *et al.*, 2007 ; Mergeay *et al.*, 2008). Deux espèces du complexe, *D. middendorffiana sensu lato* et *D. tenebrosa*, sont constituées entièrement ou partiellement de populations polyploïdes (Dufresne et Hebert, 1994, 1995, 1997) et ont une distribution autour de l'Arctique. Les lignées polyploïdes du groupe *pulicaria* (*D. middendorffiana sensu lato*) ont des origines hybrides multiples (Dufresne et Hebert, 1994, 1995, 1997) et ont des haplotypes

mitochondriaux issus soit de *D. middendorffiana sensu stricto* soit de l'espèce *D. pulicaria* (Dufresne et Hebert, 1997). Les haplotypes mitochondriaux de l'espèce *D. middendorffiana sensu stricto* étant proches évolutivement des haplotypes mitochondriaux de *D. pulicaria* et *D. pulex* (Colbourne *et al.*, 1998 ; Dufresne et Hebert, 1997 ; Marková *et al.*, 2007). Les lignées polyploïdes que constitue *D. middendorffiana s. s.* auraient émergés lors d'évènements d'hybridation lors du Pléistocène (Colbourne *et al.*, 1998 ; Dufresne et Hebert, 1997). Les lignées polyploïdes ayant des haplotypes mitochondriaux *D. pulicaria* seraient issues de croisements entre des femelles *D. pulicaria* et des mâles *D. pulex* (Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994, 1997). Tout comme les populations diploïdes hybrides, la nature hybride des populations polyploïdes a été inférée principalement par la combinaison du phénotype allozymique de l'enzyme LDH et des données mitochondriales. Des populations polyploïdes existent chez *D. tenebrosa* mais leur nature hybride, si telle est le cas, est inconnue. Les phénotypes allozymiques laissent à penser que d'importants flux géniques existeraient ou ont existé entre les espèces des groupes *tenebrosa* et *pulicaria* (Weider *et al.*, 1999b) mais de plus amples analyses sont nécessaires pour déterminer les espèces parentales impliquées dans la formation des polyploïdes.

1.2.3 Mode de reproduction

La plupart des populations du complexe *pulex* présente une reproduction parthénogénétique cyclique (Figure 1-3A). Pendant la majeure partie de l'année, les femelles se reproduisent d'une manière clonale. En automne, des mâles sont produits et les œufs sont alors fécondés. Cette reproduction sexuée permet la production d'œufs résistants (éphippies) à des conditions climatiques non optimales. Toutefois certaines populations sont strictement asexuées et produisent des éphippies de manière clonale (Figure 1-3B; Hebert *et al.*, 1989 ; Hebert et Crease, 1983 ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b ; Innes et Hebert, 1988 ; Innes et Singleton, 2000). L'origine de l'asexualité chez *D. pulex* semble récente (Paland et Lynch, 2006) et polyphylétique (apparition du régime de reproduction asexuée de

multiples fois et indépendamment dans l'histoire évolutive de *D. pulex*; Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Paland, Colbourne et Lynch, 2005). La capacité à produire des mâles viables et fertiles peut être conservée dans des lignées strictement asexuées (Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Innes et Hebert, 1988 ; Paland, Colbourne et Lynch, 2005). Cette production de mâle est dépendante de l'environnement (Innes, Fox et Winsor, 2000).

La perte de la reproduction sexuée peut être due à différents mécanismes génétiques (Simon *et al.*, 2003). Chez *D. pulex*, le passage à une reproduction strictement asexuée pourrait avoir différentes origines (Innes, Schwartz et Hebert, 1986). L'existence d'un ou de plusieurs allèle(s) dominant chez les femelles et conduisant à la suppression de la méiose pourrait expliquer l'apparition et la transmission de l'asexualité (« asexualité contagieuse »; Crease, Stanton et Hebert, 1989 ; Innes et Hebert, 1988 ; Lynch *et al.*, 2008 ; Paland, Colbourne et Lynch, 2005). Une analyse quantitative de cartographie à partir de marqueurs microsatellites a montré que 4 facteurs génétiques non liés situés sur les chromosomes V, VIII, IX et X pourraient interagir de manière épistatique afin de produire le phénotype de suppression de méiose (Lynch *et al.*, 2008). Ces allèles seraient transmis de manière sexuée par des mâles produits de manière parthénogénétique. Les lignées produites se reproduiraient alors de manière strictement asexuée et pourraient engendrer des mâles. La production de lignées parthénogénétiques obligatoires peut être due à des croisements d'individus génétiquement divergents (i.e. hybridation; Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Innes, Schwartz et Hebert, 1986). La production de mâles chez les lignées strictement clonales peut être retenue que ces lignées soient d'origines hybrides ou qu'elles soient produites par « asexualité contagieuse » (Innes et Hebert, 1988). La répartition géographique des lignées strictement asexuées, hybrides ou non, a été particulièrement bien étudiée en Amérique du Nord (Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b ; Innes, Schwartz et Hebert, 1986). En Ontario et au Michigan, les lignées

de *D. pulex* strictement asexuées hybrides et non hybrides peuvent coexister (Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Hebert *et al.*, 1993b).

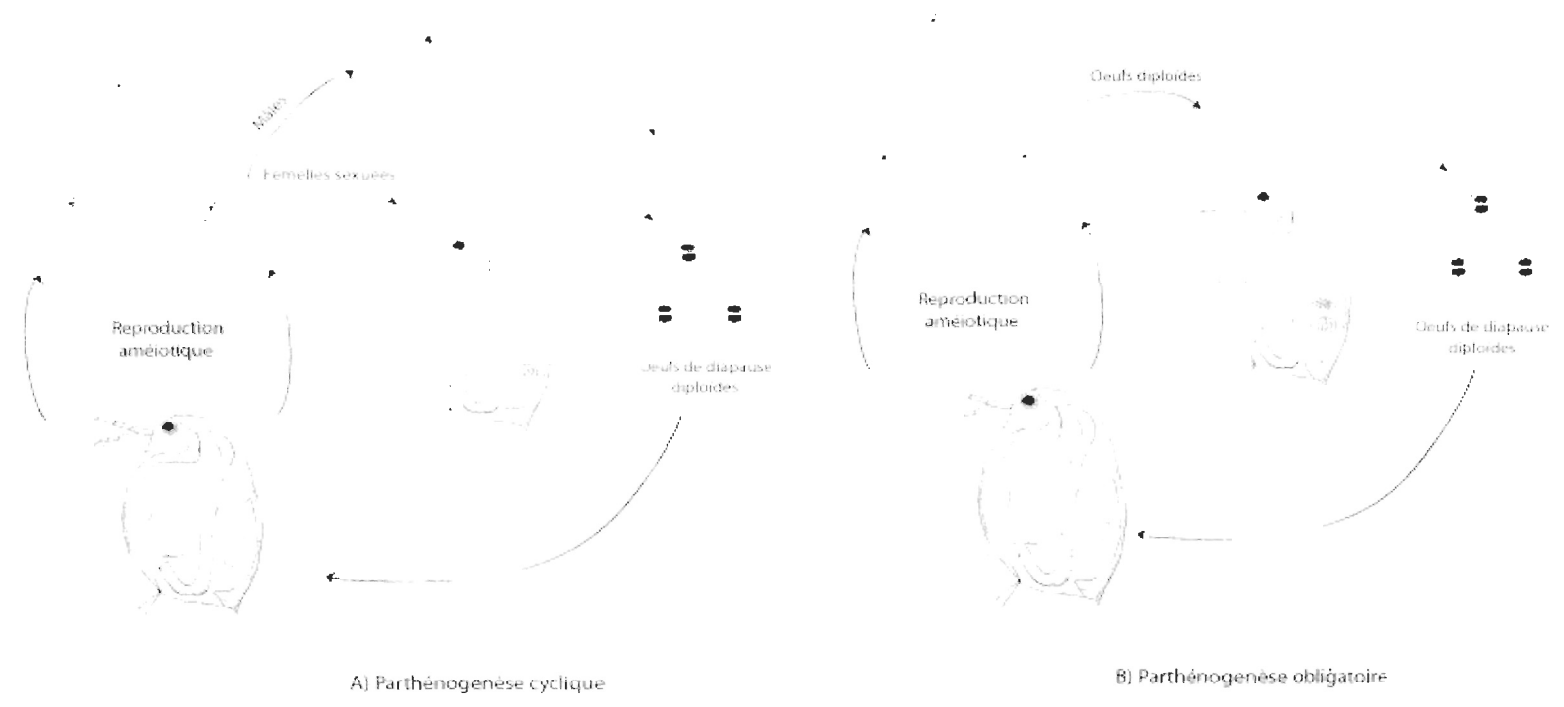


Figure 1-3 Différents modes de reproduction de l'espèce *Daphnia pulex*.
 (Figure adaptée à partir de Dufresne, sous presse)

1.2.4 Les éléments transposables chez *Daphnia pulex*

De nombreux éléments transposables, transposons de classe II et rétrotransposons à LTR, ont été identifiés (Rho *et al.*, 2010 ; Schaack *et al.*, 2010a) dans le génome de *D. pulex*. *Pokey*, un transposon de classe II de la famille *piggyBac*, a été particulièrement bien étudié dans les populations diploïdes (Penton et Crease, 2004 ; Penton, Sullender et Crease, 2002 ; Schaack *et al.*, 2010a ; Schaack *et al.*, 2010b ; Sullender et Crease, 2001a ; Valizadeh et Crease, 2008). *Pokey* est un transposon possédant des répétitions terminales inversées (TIRs) de 16 pb et deux cadres ouverts de lecture (ORF) codant pour une transposase (Penton, Sullender et Crease, 2002; Teresa J. Crease, communication personnelle). Il s'insère préférentiellement dans les gènes répétés codant pour la grande sous-unité 28S de l'ARN ribosomal (Penton, Sullender et Crease, 2002) et se rencontre aussi dans différentes parties du génome (Schaack *et al.*, 2010a ; Schaack *et al.*, 2010b ; Sullender et Crease, 2001a ; Valizadeh et Crease, 2008). Les transposons de classes II sont plus enclins aux transferts horizontaux que les rétrotransposons. Ce mode de transfert entre les lignées phylogénétiques ne semble pas jouer un rôle majeur dans l'évolution de *Pokey* (Penton et Crease, 2004). La comparaison du nombre d'insertions de *Pokey* dans des lignées provenant de populations parthénogénétiques cycliques et obligatoires montre qu'en moyenne le nombre d'insertions de *Pokey* est significativement plus grand chez les parthénogènes cycliques que chez les parthénogènes obligatoires (Schaack *et al.*, 2010b ; Sullender et Crease, 2001a ; Valizadeh et Crease, 2008). La variation des sites d'insertion entre isolats est similaire mais seulement 27% des sites d'insertions sont communs aux populations parthénogènes cycliques et obligatoires (Valizadeh et Crease, 2008). Ces données suggèrent que *Pokey*, comme d'autres éléments transposables de classe II (Schaack *et al.*, 2010b), est actif dans les populations cycliques en nature (Schaack *et al.*, 2010b ; Sullender et Crease, 2001a ; Valizadeh et Crease, 2008).

La grande diversité du nombre de lignées hybrides diploïdes et polyploïdes et

la présence d'éléments transposables connus dans le génome de *Daphnia pulex* justifie l'étude du complexe *pulex* afin de tester l'effet du niveau de ploïdie sur le nombre d'insertions chez une espèce animale (Chapitre VI). Avant de tester cet effet, ma thèse s'est attelée à redéfinir, avec les outils de la génétique, de la génomique et de la cytogénétique mis à notre disposition, d'une part le (ou les) niveau(x) de ploïdie présenté(s) par différentes lignées de ce complexe (Chapitre II), et d'autre part les relations évolutives entre différentes espèces du complexe et les éventuels événements d'hybridation impliquant des lignées diploïdes et polyplloïdes (Chapitre III).

1.3 Objectifs du premier chapitre

1.3.1 Etude de la taille de génome

Diverses études ont tenté d'évaluer la taille de génome haploïde (i.e. la « C-value ») des espèces du complexe d'espèces *Daphnia pulex* (Beaton, 1995 ; Beaton et Hebert, 1988 ; Beaton et Hebert, 1989 ; Korpelainen, Ketola et Hietala, 1997 ; Rasch, 1985). Les « C-value » observées en utilisant différentes méthodes et tissus varient de manière importante. Par exemple pour *D. pulex*, des tailles de génomes haploïdes de 0,24 pg (Rasch, 1985), 0,34 pg (Beaton, 1995) et 0,37 pg (Korpelainen, Ketola et Hietala, 1997) ont été évaluées. Or le séquençage du génome complet de *D. pulex* suggère une « C-value » de 0,20 pg (Colbourne *et al.*, 2011). De plus, alors que toutes les espèces du complexe possèdent 12 paires de chromosomes (Beaton, 1995 ; Beaton et Hebert, 1994), l'étude de Beaton et collaborateurs a montré que les « C-value » des espèces diploïdes du complexe varie de 0,34 à 0,58 pg. (Beaton, 1995 ; Beaton et Hebert, 1989). Cette variation pourrait être due à une mauvaise évaluation du niveau de ploïdie de certains isolats.

Le premier objectif du premier chapitre consiste à préciser les tailles de

génomés haploïdes des espèces diploïdes du complexe d'espèces *Daphnia pulex*. En utilisant la technique de cytométrie en flux et un protocole standardisé adapté à partir de la méthode de Rees et collaborateurs (Rees *et al.*, 2007), les tailles de génomes de différentes lignées de différentes espèces du complexe *pulex* mises en cultures en laboratoire seront évaluées. Cette évaluation de la quantité d'ADN nucléaire de génomes diploïdes permettra de réévaluer le niveau de ploïdie des isolats polyploïdes du complexe d'espèces *Daphnia pulex* (voire plus bas).

1.3.2 Étude de la polyploïdie

Le phénomène de polyploïdie dans le complexe *pulex* a été initialement mis en évidence par l'observation de patrons allozymiques débalancés (coloration plus intense de certains allèles pour des allozymes monomériques par exemple; Weider et Hebert, 1987). La majorité des études sur la polyploïdie a analysé les patrons allozymiques et les tailles de génomes par spectrodensitométrie en utilisant la coloration de Feulgen. Ces analyses sont arrivées à la conclusion que les clones polyploïdes du complexe *pulex* sont tétraploïdes (Adamowicz *et al.*, 2002 ; Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994, 1995, 1997 ; mais voire Ward *et al.*, 1994). La variabilité des valeurs de la quantité d'ADN des noyaux cellulaires observée chez les lignées présumées polyploïdes (de 0,70 à 0,97 pg pour les clones tétraploïdes de *D. tenebrosa*; Dufresne et Hebert, 1995) pourrait être due à une trop faible quantité de cellules observées lors des analyses de densitométrie. De plus, la quantité d'ADN des cellules des lignées polyploïdes a été comparée à celle de *Daphnia pulex* dont l'estimation semble biaisée (0,37 pg; Dufresne et Hebert, 1995).

Le deuxième objectif du premier chapitre consiste à préciser le niveau de ploïdie de différents isolats polyploïdes provenant du complexe d'espèces *Daphnia pulex*. L'application de la technique de cytométrie en flux, et l'utilisation d'une référence externe (i.e. erythrocytes de *Gallus gallus domesticus*), permettraient de

préciser la quantité d'ADN nucléaire dans les cellules des espèces polyploïdes du complexe *D. pulex*. Comparée à la microdensitométrie, où l'absorbance des noyaux des cellules d'exopodites colorés au réactif de Feulgen est mesurée (Beaton, 1995), la cytométrie en flux est une méthode plus simple et plus précise (Korpelainen, Ketola et Hietala, 1997). En raison de l'endopolyploïdie (c'est à dire le doublement de génome dans certaines lignées cellulaires à l'intérieur d'un individu), l'isolement de cellules, de niveau de ploïdie de base, lors de l'analyse par microdensitométrie est rendue difficile chez les espèces du complexe *pulex*. La cytométrie évite une dissection complexe et permet d'évaluer la quantité d'ADN à l'intérieur de noyaux de centaines, voire de milliers, de cellules en une observation. La publication d'une base de données microsattellites (c'est à dire des séquences de quelques nucléotides répétées présentant une forte variabilité dans les populations) chez *Daphnia pulex* (Colbourne *et al.*, 1998 ; Cristescu *et al.*, 2006) peut permettre d'évaluer le niveau de ploïdie des différents clones. Du fait de la nature codominante des marqueurs microsattellites, le nombre d'allèles à un locus variable donné décrit potentiellement le niveau de ploïdie de l'individu qui les porte.

1.4 Objectifs du deuxième chapitre

1.4.1 Relations évolutives et spéciation

Les principales études sur les relations phylogénétiques entre les espèces du complexe *Daphnia pulex* se sont focalisées sur les gènes mitochondriaux (Aguilera *et al.*, 2007 ; Crease *et al.*, 1997 ; Crease, Stanton et Hebert, 1989 ; Dufresne et Hebert, 1994, 1995, 1997 ; Hebert *et al.*, 1989 ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b ; Mergeay *et al.*, 2008 ; Weider et Hobaek, 2003 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b) et/ou les phénotypes allozymiques (Weider *et al.*, 1999a ; Weider *et al.*, 1999b). Les données mitochondriales ont permis d'estimer la divergence entre les groupes *tenebrosa*, *pulicaria* et « *pulex* Européen » lors du

Pliocène (Colbourne *et al.*, 1998). La séparation des différentes espèces *D. pulex* et *D. pulicaria* à l'intérieur du groupe *pulicaria* semble, elle, datée du milieu du Pléistocène (Dufresne et Hebert, 1997 ; Weider et Hobaek, 2003). Les gènes mitochondriaux n'étant hérités que par la lignée maternelle, ils ne permettent pas de mettre en lumière les possibles phénomènes d'hybridation. Le déséquilibre de liaison entre différents gènes mitochondriaux est très important (voire maximal) due à une quasi absence de recombinaison du génome mitochondrial. La phylogénie de gènes ou même de groupes de gènes liés, comme cela peut être le cas pour différents gènes mitochondriaux, peut ne pas correspondre à la phylogénie des espèces. Dans les groupes d'espèces ayant divergé récemment, un polymorphisme hérité de la lignée ancestrale peut avoir été retenue dans les différentes espèces (Pamilo et Nei, 1988 ; Wu, 1991). Dû à leur implication dans les différentes voies métaboliques les allozymes sont soumis à diverses pressions sélectives (Eanes, 1999). Ils sont relativement peu variables et peuvent présenter un fort taux d'homoplasie (i.e. similitude de phénotypes non hérités d'un ancêtre commun). Les allozymes ne sont donc pas appropriés pour étudier des relations phylogénétiques entre des espèces ayant divergées récemment. Afin d'étudier les relations évolutives entre des espèces jeunes, il est important d'utiliser d'une part différents marqueurs nucléaires et d'autre part des marqueurs présentant une forte variabilité. Les études portant sur la diversité nucléotidique à fine échelle de gènes nucléaires dans le complexe *pulex* n'ont pris en compte que peu d'espèces (i.e. le plus souvent *D. pulex*, *D. pulicaria* et *D. arenata*; Omilian et Lynch, 2009 ; Omilian, Scofield et Lynch, 2008 ; Pfrender, Spitze et Lehman, 2000).

Un des objectifs de ce chapitre est de décrire les relations évolutives et phylétiques entre la plupart des espèces du complexe *Daphnia pulex*. Pour ce faire, un gène mitochondrial, un gène nucléaire (*Rab4*) et différents locus microsatellites (donc supposément non soumis à sélection) ont été analysés sur 6 espèces nominales présentes en Amérique du Nord et certaines lignées appartenant à des espèces homonymes en Europe.

1.4.2 Hybridation dans le complexe *Daphnia pulex*

Dans le complexe *pulex*, les relations évolutives entre les différentes espèces ou entre les populations hybrides et leurs espèces « parents » sont encore incertaines, en particulier pour les clones polyploïdes de *D. tenebrosa* (Dufresne et Hebert, 1995, 1997). *D. pulicaria* et *D. middendorffiana* sembleraient contribuer aux différents génomes maternels rencontrés dans les différentes populations de *D. middendorffiana* sensu lato (Colbourne *et al.*, 1998 ; Dufresne et Hebert, 1997). Deux espèces pourraient contribuer paternellement au génome des clones polyploïdes de *D. middendorffiana*: *D. pulex* seulement sur la base de la concordance des patrons allozymiques et *D. melanica* qui partage le caractère « mélanisé » avec *D. middendorffiana* (Dufresne et Hebert, 1997). D'après les données allozymiques et mitochondriales, les clones de *D. middendorffiana sensu lato* sont considérés issus de plusieurs événements d'hybridations (Dufresne & Hebert 1997). En comparant les phénotypes pour divers allozymes, une étude portant sur 850 populations échantillonnées autour de l'Arctique a estimé qu'un fort taux d'introggression existait entre les espèces du groupe *pulicaria* et celles du groupe *tenebrosa* (Weider *et al.*, 1999a ; Weider *et al.*, 1999b). Toutefois comme cité précédemment, les allozymes ne seraient pas des marqueurs suffisamment puissants lors de l'analyse fine de relations phylétiques et de relations d'hybridations entre espèces proches évolutivement. Dans le cas des phénomènes d'hybridations, l'apport d'une partie du génome nucléaire d'une espèce dans une autre espèce aurait pour conséquence une absence de congruence entre les données nucléaires et les données mitochondriales. Les espèces des complexes d'espèces sont supposées avoir divergées relativement récemment du point de vue évolutif (Mayr, 1963). Il est connu que des événements autres que les événements d'hybridations tels que le maintien du polymorphisme ancestral et qui sont liés à l'âge relatif de la formation des espèces peuvent induire une incongruence entre les arbres des différents gènes étudiés (Pamilo et Nei, 1988 ; Wu, 1991). Afin de discerner ces différents événements, il est important d'utiliser plusieurs gènes ayant différents taux de mutations. Des gènes ou des marqueurs génétiques ayant un fort taux de

mutation seraient plus efficaces pour dissocier les événements d'hybridation des événements liés au maintien du polymorphisme dans les populations. Le complexe *pulex* étant relativement jeune évolutivement, les marqueurs microsatellites mis au point pour l'espèce *D. pulex* (Colbourne *et al.*, 2004 ; Cristescu *et al.*, 2006) peuvent être amplifiés dans les différentes espèces du complexe. Ces marqueurs sont théoriquement assez variables et permettrait de décrire les relations évolutives entre les différentes espèces du complexe.

Afin de mettre en relief les phénomènes d'hybridations entre les espèces du complexe *pulex*, deux méthodes seront utilisées : la comparaison d'arbres phylogénétiques entre un gène nucléaire et un gène mitochondrial et une analyse d'ordination couplée à une analyse de groupements sur des génotypes multilocus microsatellites. L'analyse des microsatellites, ayant des taux de mutations importants, devrait résoudre avec une plus grande fidélité les relations évolutives entre les différentes espèces du complexe *pulex* qu'une analyse d'un gène nucléaire codant pour une protéine et donc soumis à sélection. La congruence des arbres phylogénétiques sera testée et une analyse de coalescence sera effectuée afin de tester si l'incongruence observée est due uniquement au maintien du polymorphisme ancestral dans les différents isolats du complexe *pulex* ou si elle peut être due à des événements d'hybridations.

1.5 Objectifs du troisième chapitre

1.5.1 Hybridation et diversité des allèles du transposon *Pokey*

Dû à leur transposition sous une forme ADN, les transposons de classe II sont enclins aux transferts horizontaux entre différents organismes. Il est fréquent que les phylogénies des transposons et des taxons hôtes ne concordent pas (Capy, Anxolabéhère et Langin, 1994 ; Pace *et al.*, 2008 ; Silva, Loreto et Clark, 2004). Le

transposon *Pokey* est atypique par le fait que sa phylogénie correspond globalement à la phylogénie (mitochondriale et nucléaire) de son genre hôte (Penton et Crease, 2004). La seule incongruence correspond à la présence des deux allèles de *Pokey* divergents et présentant des insertions et délétions importantes chez un individu de l'espèce *D. pulicaria* (Penton, Sullender et Crease, 2002) suggérant un phénomène d'introgession, bien qu'il puisse exister d'autres phénomènes expliquant ce patron (Capy, Anxolabéhère et Langin, 1994). Dans des populations constituées d'hybrides entre deux espèces, comme cela peut être le cas pour différentes populations diploïdes et polyploïdes du complexe *pulex*, il serait attendu que les génomes d'individus hybrides contiennent les éléments transposables des espèces parentes. Ce patron ne serait pas respecté si l'élément transposable est soumis à de fréquents transferts horizontaux ou si une des espèces parentes ne possède pas l'élément transposable dans son génome.

Un des objectifs du quatrième chapitre est de décrire les différents allèles de *Pokey* présents dans les génomes diploïdes et polyploïdes du complexe *pulex*. Pour atteindre cet objectif des séquences partielles du transposon de classe II, *Pokey*, s'intégrant dans les gènes codant pour la grande sous-unité 28S de l'ARN ribosomal seront clonées, séquencées et comparées avec des séquences disponibles sur GenBank (Penton et Crease, 2004 ; Penton, Sullender et Crease, 2002).

L'hypothèse testée est que les allèles de *Pokey* présents dans les génomes hybrides correspondent aux allèles de *Pokey* présents chez les espèces parentes.

1.5.2 Effet de la polyploïdie sur le transposon *Pokey*

Comme il a été mentionné dans le paragraphe 1.1.4, différentes théories, prédisent une activation d'éléments transposables et une augmentation du nombre de leurs insertions dans les génomes polyploïdes (Comai, 2000 ; Lynch, 2007 ; Matzke et Matzke, 1998 ; McClintock, 1984 ; Parisod *et al.*, 2010). Chez les

animaux, peu d'études se sont focalisées sur l'effet du niveau de ploïdie sur la dynamique et la densité des éléments transposables. Des études ont montré que l'hybridation de génomes divergents pouvait avoir pour conséquence l'activation et l'augmentation du nombre d'insertions d'éléments transposables chez des espèces de drosophile (Fontdevila, 2005 ; Petrov *et al.*, 1995) ou des wallabies (O'Neill, O'Neill et Graves, 1998). Une étude a mentionnée une activation d'éléments transposables chez des allotétraploïdes de carpes mais cette étude n'a pas été en mesure de différencier l'effet de l'hybridation de l'effet de la polyploïdisation *per se* (Liu *et al.*, 2009).

Diverses populations hybrides diploïdes et polyploïdes sont rencontrées dans le complexe *Daphnia pulex*. L'objectif de la dernière partie de ma thèse de doctorat est donc de comparer le nombre d'insertions de l'élément transposable *Pokey* dans des génomes diploïdes hybrides et polyploïdes hybrides afin de tester indirectement l'effet du niveau de ploïdie sur la dynamique et la densité de ce transposon de classe II. Pour ce faire, une technique de « TE display » (Wright *et al.*, 2001) adapté de l'étude de Valizadeh et Crease (2008) permettra de comptabiliser les sites d'insertions de *Pokey* dans les génomes d'isolats diploïdes hybrides et polyploïdes hybrides. L'analyse d'individus hybrides uniquement permet de discerner l'effet de la ploïdie de l'effet de l'hybridation.

Un nombre d'insertions moyen de l'élément transposable *Pokey* égal entre les lignées hybrides diploïdes et polyploïdes est l'hypothèse nulle qui est testé lors du quatrième chapitre de cette thèse.

Chapitre II

GENOME SIZE EVOLUTION AND POLYPLOIDY IN THE *DAPHNIA PULEX* COMPLEX (CLADOCERA: DAPHNIIDAE)

Roland Vergilino^a, Claude Belzile^b, France Dufresne^{a*}

^a *Département de Biologie, Université du Québec à Rimouski, Québec G5L 3A1, Canada*

^b Institut des sciences de la mer de Rimouski (ISMER), Université du Québec à Rimouski, Rimouski, Québec G5L 3A1, Canada

(Roland Vergilino : conception du projet, échantillonnage, expérimentations, traitement statistique, rédaction de l'article; Claude Belzile : expérimentation et rédaction; France Dufresne : conception du projet, échantillonnage, rédaction, financement)

Paru en 2009 dans la revue *Biological Journal of the Linnean Society* vol. 97, no 1, p. 68-79. Les références citées dans l'article sont reportées à la fin de la thèse.

2.1 Résumé

La taille du génome et le niveau de ploïdie a été estimée par cytométrie en flux et des analyses d'ADN microsatellite pour 49 isolats du complexe d'espèces *Daphnia pulex* provenant de régions géographiques tempérées et subarctiques. Des différences significatives de taille du génome ont été trouvées chez les espèces diploïdes appartenant aux deux groupes génétiquement distincts. Les tailles de génome du groupe *tenebrosa* étaient en moyenne 22% plus grandes que celles du groupe *pulicaria*. Les données récoltées grâce à la cytométrie en flux combinées à l'analyse microsatellite ont révélé que la plupart des clones polyploïdes dans le complexe *D. pulex* sont triploïdes et non pas tétraploïdes comme le suggéraient les études antérieures. L'analyse de séquençage du gène *ND5* de clones a mis en relief des clones triploïdes possédant un haplotype mitochondrial *D. pulex*. Ce résultat était inattendu car les parents maternels des clones polyploïdes hybrides dans le groupe *pulicaria* ont été identifiés comme soit *D. pulicaria* soit *D. middendorffiana sensu stricto*. Les clones triploïdes sont probablement issus d'interactions entre des populations sexuées et asexuées. Différentes interactions dans le groupe *tenebrosa* ont généré des clones tétraploïdes mais ces événements semblent rares.

2.2 Abstract

Genome size and ploidy level were estimated in 49 clones of the *Daphnia pulex* complex from temperate and subarctic locations using flow cytometry and microsatellite DNA analyses. Significant genome size differences were found in diploid species belonging to the two genetically distinct groups (the *pulicaria* and the *tenebrosa* groups), with clones from the *tenebrosa* group having genome sizes 22% larger than those in the *pulicaria* group. Combined flow cytometry and microsatellite DNA analyses revealed that nearly all polyploid clones in the *D. pulex* complex are triploid and not tetraploid as was previously suggested. Sequencing analyses of the ND5 gene to position clones in their respective clades within the *D. pulex* complex have uncovered three triploid clones of *D. middendorffiana* with a *D. pulex* maternal parent. This result was unexpected since *D. pulicaria* has always been identified as the maternal parent of these hybrid polyploid clones. Triploid clones likely owe their origins to interactions between sexual and asexual populations. Further interactions in the *tenebrosa* group have generated tetraploid clones but these events have been rare.

2.3 Introduction

The overall nuclear DNA quantity in cells may evolve in different ways. Genome size (the C-value) alterations can be gradual and involve small-scale insertions and deletions of nucleotides or large-scale alterations such as gene duplications or transposon insertions (Bennetzen, 2002 ; Petrov, 2001). Additions of chromosomes (aneuploidy) or complete genomes (polyploidy) can abruptly increase the DNA quantity contained in the nucleus. Nuclear DNA content increases (whether they arise by polyploidization or not) are frequently accompanied by phenotypic modifications (Gregory, 2002 ; Gregory et Mable, 2005 ; Levin, 1983 ; Otto et Whitton, 2000 ; Vinogradov, 1995, 1997). Polyploids often have a wider geographical range than their diploid relatives and are often more prevalent at high latitudes (Brochmann *et al.*, 2004 ; Levin, 1983 ; Otto et Whitton, 2000). It has been suggested that elevated ploidy could be positively selected in marginal habitats (Brochmann *et al.*, 2004 ; Dufresne et Hebert, 1998 ; Levin, 1983 ; Little et Hebert, 1997 ; Otto et Whitton, 2000 ; Stenberg *et al.*, 2003).

Members of the *Daphnia pulex* complex provide good models to assess the evolutionary consequences of changes in nuclear DNA content as they vary both in genome size and in ploidy levels. Previous studies have revealed a fourfold variation in nuclear DNA content in the *Daphnia pulex* complex, with the largest value arising from polyploidy (Beaton, 1995). While all diploid individuals of this complex possess 24 chromosomes (Beaton, 1995 ; Beaton et Hebert, 1994), their C-values vary from 0.37 ± 0.010 to 0.58 ± 0.013 pg. (Beaton, 1995 ; Beaton et Hebert, 1989). Diploid clones of *D. tenebrosa* have the largest genome size in the genus *Daphnia* (Beaton, 1995 ; Dufresne et Hebert, 1995). The *D. pulex* complex shows a geographical polyploidy pattern (Dufresne et Hebert, 1995, 1997 ; Hobaek, Weider et Wolf, 1993 ; Ward *et al.*, 1994 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b) with polyploid populations found at high latitudes (53°N and higher) and altitudes (Bolivian Andes) and diploid clones being prevalent in temperate regions (Aguilera *et al.*, 2007 ; Beaton et Hebert, 1988 ; Dufresne et Hebert, 1995). An apparent exception to this

pattern is the discovery of two polyploid clones, closely related to the North American arctic clones, in Patagonia (within a latitudinal range of 46–54° S; Adamowicz *et al.*, 2002). Members of polyploid populations of the *D. pulex* complex are obligate apomicts (i.e. parthenogens that develop from unfertilized egg by a mitosis-like cell division) and have multiple hybrid origins (Dufresne et Hebert, 1995, 1997). Two species of the complex, *Daphnia middendorffiana* and *Daphnia tenebrosa*, are composed entirely or partially of polyploid populations (Adamowicz *et al.*, 2002 ; Dufresne et Hebert, 1994, 1995, 1997) and have a circumarctic distribution (Weider *et al.*, 1999a ; Weider *et al.*, 1999b). Polyploid clones of this complex are thought to be tetraploids due to their unbalanced phenotypes at allozyme loci (Dufresne et Hebert, 1994, 1995) and their larger genome sizes (1.7 times the diploid values) as estimated by Feulgen densitometry (Beaton, 1995 ; Dufresne et Hebert, 1995). Combined allozymes and mitochondrial gene sequences have shown that tetraploid clones of *D. middendorffiana* and *D. tenebrosa* result from several events of hybridizations (Dufresne et Hebert, 1997).

Genome size estimates of *D. pulex*, obtained through Feulgen densitometry, often vary depending on the study. For example Rasch, Stanton & Hebert (1983, unpublished, cited in Rasch, 1985) obtained C-values between 0.23 and 0.25 pg for *D. pulex* whereas Beaton & Hebert (1989) and Beaton (1995) found values of 0.37 ± 0.010 pg and 0.38 ± 0.003 pg, respectively. The recent *Daphnia* genome sequencing project suggests that the C-value of *Daphnia pulex* is near 0.20 pg (Cristescu *et al.*, 2006), which is much smaller than values reported for the diploid *D. pulex* clones that served as references for the estimation of ploidy levels (Beaton et Hebert, 1988).

This study aimed to revisit genome size and ploidy level estimates in members of the *D. pulex* complex using alternative techniques. Microsatellite loci are more variable than allozyme loci and more amenable to assess allele numbers and hence ploidy level. The majority of studies on *Daphnia* genome sizes have been carried out with Feulgen staining (Adamowicz *et al.*, 2002 ; Beaton, 1995 ; Beaton et

Hebert, 1988 ; Dufresne et Hebert, 1997). Flow cytometry is an accurate and rapid technique that allows the evaluation of genome size of thousands of cells in the same run (Bennett *et al.*, 2003 ; Galbraith *et al.*, 1983). This study provides new insights into genome size and ploidy levels of 49 clones of five species in the *D. pulex* complex using both flow cytometry and microsatellite markers.

2.4 Material and methods

2.4.1 *Daphnia* collection

Forty-nine *Daphnia* clones were reared in the laboratory for several generations prior to the experiments (Table 1). Each clone was kept in a 500 mL culture flask at a constant temperature and was fed two times per week with the green alga *Selenastrum capricornutum*. Clones K9, K154, K52, K207, K228, K230, K232, K86, K92, K95 and K31 were sampled from rock pools and ponds during the summers of 2006 and 2007 at Kuujjuarapik, Québec, CANADA. Clones A17-200, A35-201, B141-201, B141-202, A, B, C, D, E, F, G and H were sampled in ponds located near Churchill, Manitoba, CANADA during the summers of 2005, 2006 and 2007. Clone Disp1-21 was provided by Dr. David J. Innes. *D. pulicaria* clones C2, Puli 9, Puli 13, Puli 15, Puli 21, Puli 22 were provided by Dr. M Cristescu and Dr. CE Càceres.

Tableau 2-1 Mitochondrial lineages (ND5 sequences), geographic origins, and habitats of *Daphnia* clones used in this study.

Stocks	Identification (for methods see text)	Reproduction	localisation	habitats
CHQ6	pulex	Cyclical parthenogen	Wisconsin, USA	Pond
EB1	pulex	Cyclical parthenogen	Minnesota, USA	Pond
POV1-102	pulex	Cyclical parthenogen	Michigan, USA	Pond
DISP13	pulex	Cyclical parthenogen	Ontario, CAN	Pond
PET	pulex	Obligate parthenogen	Quebec, CAN	Pond
LIS3	pulex	Obligate parthenogen	Ontario, CAN	Pond
WG1	pulex	Obligate parthenogen	Michigan, USA	Pond
CU4-1	pulex	Obligate parthenogen	Quebec, CAN	Pond
A17-200	pulex	Obligate parthenogen	Churchill, Manitoba, CAN	Rock bluff
Metis-3	pulex	Obligate parthenogen	Metis, Quebec, CAN	Pond
Disp1-21	pulex	Obligate parthenogen	Ontario, CAN	Pond
K86	pulex	Obligate parthenogen	Kuujuarappik, Quebec, CAN	Rock bluff
K92	pulex	Obligate parthenogen	Kuujuarappik Quebec, CAN	Rock bluff
K52	pulex	Obligate parthenogen	Kuujuarappik Quebec, CAN	Rock bluff

Stocks	Identification (for methods see text)	Reproduction	localisation	habitats
K9	pulex	Obligate parthenogen	Kuujjuarappik Quebec, CAN	Rock bluff
K154	pulex	Obligate parthenogen	Kuujjuarappik Quebec, CAN	Rock bluff
K230	eastern pulicaria	Obligate parthenogen	Kuujjuarappik Quebec, CAN	Rock bluff
C2	pulicaria	Cyclical parthenogen	Michigan, USA	Lake
Puli9	pulicaria	Cyclical parthenogen	Michigan, USA	Lake
Puli13	pulicaria	Cyclical parthenogen	Michigan, USA	Lake
Puli15	pulicaria	Cyclical parthenogen	Michigan, USA	Lake
Puli21	pulicaria	Cyclical parthenogen	Michigan, USA	Lake
Puli22	pulicaria	Cyclical parthenogen	Michigan, USA	Lake
W4-111-b1	western pulicaria	Cyclical parthenogen	Indiana, USA	Lake
W4-95	western pulicaria	Cyclical parthenogen	Indiana, USA	Lake
K207	western pulicaria	Obligate parthenogen	Kuujjuarappik Quebec, CAN	Rock bluff
A35-201	middendorffiana (clade I)	Obligate parthenogen	Churchill, Manitoba, CAN	Rock bluff
B141-201	middendorffiana (clade I)	Obligate parthenogen	Churchill, Manitoba, CAN	Rock bluff
MIDA20	middendorffiana (clade IIA)	Obligate parthenogen	Churchill, Manitoba, CAN	Rock bluff

Stocks	Identification (for methods see text)	Reproduction	localisation	habitats
B141-202	middendorffiana (clade IIA)	Obligate parthenogen	Churchill, Manitoba, CAN	Rock bluff
K31	middendorffiana (clade IIB)	Obligate parthenogen	Kuujuarappik Quebec, CAN	Rock bluff
K95	middendorffiana (clade III)	Obligate parthenogen	Kuujuarappik Quebec, CAN	Rock bluff
K228	middendorffiana (clade III)	Obligate parthenogen	Kuujuarappik Quebec, CAN	Rock bluff
K232	middendorffiana (clade III)	Obligate parthenogen	Kuujuarappik Quebec, CAN	Rock bluff
CC3	arenata	Cyclical parthenogen	Oregon, USA	coastal pond
OP1	arenata	Cyclical parthenogen	Oregon, USA	coastal pond
Clone A	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond
Clone B	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond
Clone C	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond
Clone D	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond
Clone E	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond
Clone F	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Thermokast pond
Clone G	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond
Clone H	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Pond

Stocks	Identification (for methods see text)	Reproduction	localisation	habitats
Clone 13	tenebrosa	Obligate parthenogen	Churchill, Manitoba, CAN	Rock bluff
Gerber 1a	european pulicaria	Cyclical parthenogen	Central Pyrenees, Catalonia, Spain	Mountain lake
Gerber 3	european pulicaria	Cyclical parthenogen	Central Pyrenees, Catalonia, Spain	Mountain lake
Chmelnice	european pulicaria	Cyclical parthenogen	Chmelnice, Czech republic	Lowland lake
Malá Kuš 5-1	european pulicaria	Cyclical parthenogen	Malá Kuš, Czech republic	Lowland lake

2.4.2 Clone identification

Total DNA was extracted from individual *Daphnia* using 50µl of 10% CHELEX 100 solution (Bio-Rad, Hercules, CA, U.S.A.) or using 50µl of the QuickExtract DNA extraction solution (EPICENTRE® Biotechnologies, Madison, WI, U.S.A.). A fragment of a mitochondrial gene NADH dehydrogenase subunit 5 (ND5) was amplified and sequenced to confirm the identity of the maternal lineage. DpuND5b (5'-GGGGTGTATCTATTAATTCG-3') as reverse primer and DpuND5a (5'-ATAAACTCCAATCAACCTTG-3') or ND5ColbF (5'-AACTTAGTATCACCCAGCAGG-3') as forward primers were used to amplify approximately 850 bp and 700 bp fragments respectively in the *pulicaria* group. Clones C2, Puli 9, Puli 13, Puli 15, Puli 21, Puli 22 were identified through LDH allozyme analysis (e.g. homozygosity of the fast "F" allele, was considered to be of *D. pulicaria*). The primers ND5tene-F (5'-TTTCAACGAAAACCCCTGAC-3') and ND5tene-R (5'-AGGCCCTGCTCTTTCTTTGT-3') were needed to amplify a 600 bp fragment in the *tenebrosa* group. Amplifications were carried out using the following conditions: one cycle of denaturation at 94°C for 3 min, and 40 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min, and extension at 72°C for 2 min and a final extension at 72°C for 15 min. ND5 partial sequences were corrected manually and aligned using the ClustalW module of BioEdit v.7.0.5 (Hall T., unpublished; <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

2.4.3 Phylogenetic analyses

Phylogenetic analyses were carried out to assign clones to their respective lineages and to estimate evolutionary relationships between diploid and polyploid clones in the *Daphnia pulex* complex.

Twenty-nine ND5 sequences (from CU4-1, K9, K154, K52, K207, K228, K230, K232, K86, K92, K95, K31, A17-200, A35-201, B141-201, B141-202, MIDA20,

Disp1-21, W4-95, W4-111-b1, Clones 13, A, B, C, D, E, F, H and Chmelnice clones; accession numbers: FJ591096 to FJ591124) out of the 49 clones analysed for genome size were added to the Marková *et al.*, (2007) data set, yielding a total of 93 distinct haplotypes.

The HKY+G model of DNA substitution with a gamma substitution parameter (α) of 0.3739, transition/transversion ratios of 5.2071, and base frequencies of A=0.1965, C=0.2025, G=0.2173, T=0.3837, were identified by the hierarchical likelihood ratio test performed using MODELTEST 3.07 (Posada et Crandall, 2001) and PAUP* software version 4.0b10 (Swofford, 2003).

To classify polyploid clones into the phylogeny of the complex *Daphnia pulex* containing the different groups defined by Dufresne & Hebert (1995, 1997), Colbourne *et al.* (1998) and Weider *et al.* (1999a), we performed phylogenetic analyses with PAUP* 4.0b10 (Swofford, 2003) and PHYML (Guindon et Gascuel, 2003). Gene-tree reconstruction was accomplished using neighbor-joining (NJ) (Saitou et Nei, 1987), with the respective optimal distance model identified with MODELTEST; and maximum-parsimony (MP) using heuristic searches with the default option in PAUP*. Maximum likelihood (ML) analysis was carried out with PHYML version 2.4.4 (Guindon et Gascuel, 2003), again using the best fit DNA substitution model chosen by the MODELTEST program. Confidence in the partitioning within the ML tree was estimated with 1000 non-parametric bootstrap replicates (Felsenstein, 1985).

Bayesian inference was performed with MrBayes 3.1.1 (Huelsenbeck et Ronquist, 2005 ; Ronquist et Huelsenbeck, 2003) assuming the gamma-HKY model of sequence evolution (two substitution types with gamma-distributed rates across sites) and using the same parameters as those used in the Bayesian phylogenetic analysis of Marková *et al.* (2007).

2.4.4 Multilocus genotypes

Each individual was genotyped at nine microsatellite loci: Dp183 (Wfms0000187), Dp502 (Wfms0000519), Dp512 (Wfms0000529), Dp513 (Wfms0000530), Dp514 (Wfms0000531), Dp514alt (Wfms0000532), Dp519 (Wfms0000537), Dp523 (Wfms0000541), Dp525alt (Wfms0000544) described in Colbourne *et al.* (2004) and Cristescu *et al.*, (2006). Polymerase chain reactions (PCR) were performed in 12- μ l reactions with 10 ng of DNA template, 1 \times PCR buffer with 25 nmol of MgCl₂, 0.5 units of Taq polymerase, 2.5 nmol of each dNTP, 2 pmol of the forward primer end-labeled with fluorescent dyes HEX, FAM or TET (Alpha DNA, Montréal Canada), and 2 pmol of the reverse primer. Each primer pair was used separately and PCR reactions for each locus were not pooled. Thermal cycle programs included an initial denaturation step of 3 min. at 95°C followed by 10 cycles of 35 sec. denaturation at 94°C, 35 sec. at a initial annealing temperature of 55°C for Dp183, Dp502, Dp512, Dp514, Dp514alt, Dp519, Dp523, Dp525alt and 64°C for Dp513 (the annealing temperature was decreased by 1°C every cycle during each of the 6 following cycles), 45 sec. extension at 72°C followed by 30 cycles of 35 sec. denaturation at 94°C, 35 sec. annealing temperature at 48°C or 59°C, and 45 sec. extension at 72°C, with a final extension at 72°C for 10 min. PCR products were denatured at 95°C for 3 minutes and alleles were separated on a denaturing polyacrylamide gel containing 6% acrylamide/bis-acrylamide (19:1) at 50 Watts and 40°C for 2 to 3h. Fluorescent dye-labeled alleles were detected on a FMBIO® III fluorescent scanner (Hitachi Software Engineering America Ltd., San Bruno, CA, U.S.A.) and analyzed using the software ImageAnalysis (v.3.0.0.21 Miraibio, San Francisco, CA, U.S.A.). Ploidy level for each clone was estimated by counting the number of alleles at each locus. The assumption of ploidy level n is rejected when level $n+1$ is met at a given locus. With this method, there is a risk of erroneously rejecting level $n+1$, but this risk decreases with the number of loci used

and the number of alleles per locus. The large number of polymorphic loci used in this study (i.e. 9) should provide an accurate estimate of ploidy level.

2.4.5 Flow cytometry

The genome size of each of our 49 clones was determined using flow cytometry (Galbraith *et al.*, 1983). A stock solution of chicken erythrocytes was prepared for use as an internal standard. Freshly drawn chicken blood was transferred to a sterile tube containing a 20 µl heparin solution (100 U/ml). In a Kontes Dounce tissue grinder, 25 µl of blood was added to 1 ml of modified Galbraith buffer (Bennett *et al.*, 2003 ; Galbraith *et al.*, 1983) and was ground gently for 20 strokes with an “A” pestle. This solution was filtered through a 40 µm mesh and stored at 4°C for less than 1 week.

Five females of each *Daphnia* clone were centrifuged in distilled water at 3000 rpm and washed twice with distilled water. They were subsequently ground gently in a Kontes Dounce tissue grinder for 20 strokes with an “A” pestle in 1 ml of modified Galbraith buffer (Bennett *et al.*, 2003 ; Rees *et al.*, 2007). After filtering twice through a 40 µm mesh, 2 µl of the chicken erythrocyte cell suspension were added to each sample. Cell suspensions were stained with propidium iodide (50 ppm, Invitrogen, Burlington, ON, Canada) at 4°C in the dark for about 10 hours. The measurements were repeated 2-5 times on different individuals over a period of less than one month for the majority of clones.

The nuclear DNA content of each clone was assessed using an Epics Altra flow cytometer (Beckman-Coulter, Fullerton, CA, U.S.A.) with an argon laser emitting 14 mW of light at 488 nm. DNA-PI fluorescence emission was measured at 600-640 nm. Instrument alignment and stability were monitored by adding 5 µL of a solution of red-fluorescing beads (Linear Flow Carmine, Molecular Probes Eugene, OR, U.S.A.). Due to endopolyploidy (somatic cell lines with different ploidy level), only the

first peak, which has an amplitude (e.g. number of nucleus measured) always greater than twice the “endopolyploid peaks”, of the cytograms was taken into account to estimate genome sizes. Nuclear DNA content was calculated from the mean fluorescence intensity (FL, arbitrary units) on the gated data of the first peak as:

$$\text{Total genome size} = (\text{sample FL} / \text{chicken FL}) \times 2.5 \text{ pg}$$

where 2.5 pg corresponds to the diploid genome size of chicken erythrocytes (Rasch, Barr et Rasch, 1971). The monoploid genome size of each clone (i.e. the 1Cx-value corresponding to the 2C-value divided by the ploidy level; (Greilhuber *et al.*, 2005)) was calculated.

Differences between 2C-values obtained for each diploid species were analysed with a non-parametric Kruskal-Wallis test due to unequal sample sizes using SYSTAT v.11 (SYSTAT Software Inc., Evanston, IL, U.S.A.). The repeated measurements from each clone belonging to the same species were averaged. The GT2-method of comparison of means (Hochberg, 1974 cited in (Sokal et Rohlf, 1995)) was performed to determine which pairwise comparisons were significantly different. The GT2-method gives only a threshold level of significance (eg. P-value<0.1, P-value<0.05 or P-value<0.01). Means were considered different when P-value<0.05.

2.5 Results

2.5.1 Clone identification

The Bayesian phylogeny inferred from the ND5 sequences (Figure 2-1) is consistent with the phylogenies of Colbourne *et al.* (1998) and Marková *et al.* (2007). Phylogenetic reconstructions obtained with the ML, NJ and MP approaches showed

similar topologies to the Bayesian tree (results not shown). They consistently recovered all nine evolutionary lineages within the *D. pulex* complex, which fell into two major groups (*pulicaria* and *tenebrosa*) as described by Colbourne *et al.* (1998) and Weider *et al.* (1999a). The *tenebrosa* group consists of European *D. pulicaria* and *D. tenebrosa*. The *D. tenebrosa* lineage included both polyploid (A, B, C, G and H) and diploid (E, D and F) clones whereas European *D. pulicaria* included only diploid clones. The *pulicaria* group consisted of three *D. pulicaria* mitochondrial subgroups (Eastern *D. pulicaria*, Western *D. pulicaria* and Polar *D. pulicaria*, all of which are found in lakes), the *D. middendorffiana* lineage, the Panarctic *D. pulex* lineage, and the *D. melanica* lineage (Figure 2-1). A fifth species, *D. arenata*, belongs to the *pulicaria* group (Colbourne et Hebert, 1997) but their ND5 sequences were not included in our phylogenetic analysis. Polyploid clones of *D. middendorffiana* have multiple origins as described by Dufresne & Hebert (1997). Clones (A35-201 and B141-201) belonged to the *D. middendorffiana* lineage (Figure 2-1). Clones midA20 and B141-202 had mtDNA haplotypes that were closely allied to Polar *D. pulicaria* whereas clone K31 clustered with the western lineage of *D. pulicaria*. Clones (K95, K228 and K232) were closely related to eastern populations of *D. pulicaria*. Finally, three polyploid clones (Disp1-21, K9 and K154) had mtDNA haplotypes that clustered with Panarctic *D. pulex* haplotypes (Figure 2-1).

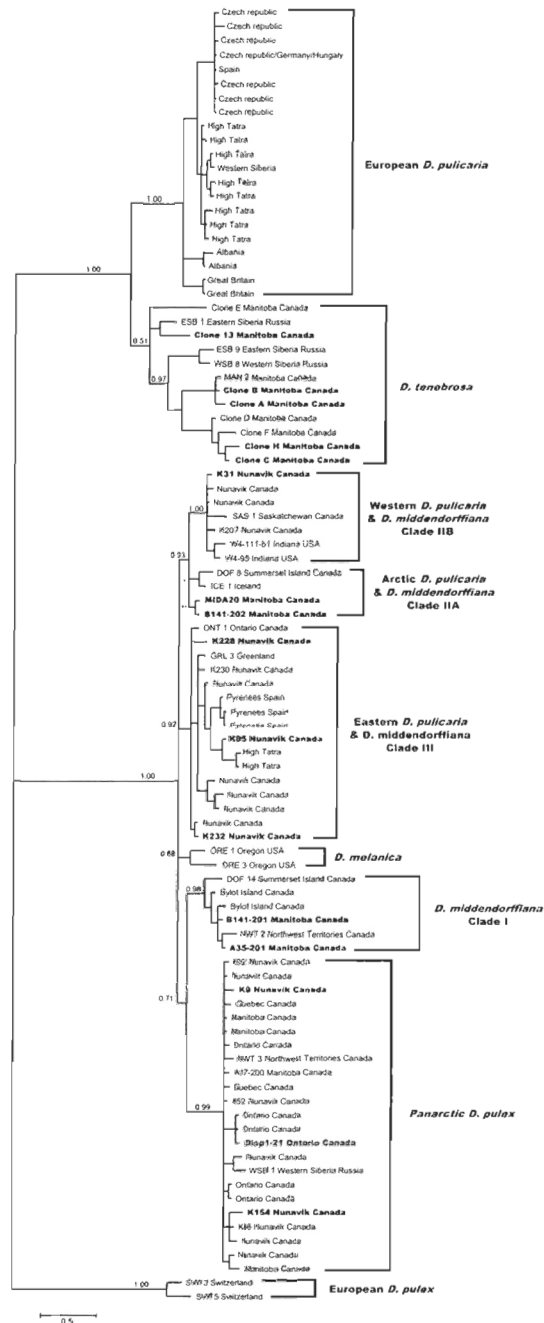


Figure 2-1 Bayesian reconstruction of the phylogenetic relationships among individuals of the *Daphnia pulex* complex

Polyploid and diploid *D. tenebrosa* haplotypes determined in this study were added to the datasets of Colbourne *et al.* (1998) and Marková *et al.* (2007). The tree is rooted through the European *D. pulex* group. Bayesian posterior probabilities are indicated for major groups. Polyploids tested in this study are in bold.

2.5.2 Polyploid diversity and ploidy level

All microsatellite loci used in this study were polymorphic in the 49 clones of the *D. pulex* complex, with 4 to 24 alleles per locus and an average of 11.56 alleles per locus (see supplementary material, Table 2-4). A maximum of three alleles per locus was observed within clones. Seventeen clones of the 49 tested in this study were polyploids (Table 2-2). Due to incongruence between ploidy levels estimated from the first nine microsatellite loci and that from flow cytometric measurements, six additional loci were screened; Dp612 (DQ249459), Dp936 (DQ249392), Dp1005 (DQ249458), DP1155 (DQ249427), Dp1232 (DQ249380), Dp1351 (DQ249397) described in Cristescu *et al.* (2006). These additional loci failed to reveal more than three alleles per locus indicating that these seventeen polyploid clones are at least triploid. In addition to the forty-nine clones tested in this study, we obtained multilocus genotypes at microsatellite loci for another set of 34 polyploid clones of the *pulex* complex and none of them had more than three alleles per locus (data not shown).

Tableau 2-2 Genetic and cytogenetic characteristics of Daphnia clones
 “n” are the number of repeated measurements

Stocks	identification (see text for methods)	ploidy level (inferred from microsatellite data)	n	Nuclear DNA content 2C-value Mean \pm SE (pg)	Monoploid genome size 1Cx-value Mean \pm SE (pg)
CHQ6	<i>pulex</i>	2x	2	0.44 \pm 0.003	0.22 \pm 0.002
EB1	<i>pulex</i>	2x	1	0.44	0.22
POV1-102	<i>pulex</i>	2x	4	0.47 \pm 0.012	0.23 \pm 0.006
DISP13	<i>pulex</i>	2x	3	0.44 \pm 0.015	0.22 \pm 0.007
PET	<i>pulex</i>	2x	1	0.48	0.24
LIS3	<i>pulex</i>	2x	1	0.41	0.20
WG1	<i>pulex</i>	2x	1	0.46	0.23
CU4-1	<i>pulex</i>	2x	7	0.44 \pm 0.009	0.22 \pm 0.004
A17-200	<i>pulex</i>	2x	1	0.48	0.24
Metis-3	<i>pulex</i>	2x	1	0.48	0.24
Disp1-21	<i>pulex</i>	3x	3	0.62 \pm 0.033	0.21 \pm 0.011
K52	<i>pulex</i>	2x	2	0.45 \pm 0.029	0.22 \pm 0.014
K86	<i>pulex</i>	2x	7	0.46 \pm 0.007	0.23 \pm 0.004
K92	<i>pulex</i>	2x	1	0.47	0.23
K9	<i>pulex</i>	3x	2	0.61 \pm 0.002	0.20 \pm 0.001
K154	<i>pulex</i>	3x	3	0.70 \pm 0.044	0.23 \pm 0.014
K230	<i>eastern pulicaria</i>	2x	2	0.44 \pm 0.018	0.22 \pm 0.009
C2	<i>pulicaria</i>	2x	1	0.45	0.23
Puli9	<i>pulicaria</i>	2x	1	0.44	0.22
Puli13	<i>pulicaria</i>	2x	1	0.50	0.25
Puli15	<i>pulicaria</i>	2x	1	0.49	0.24
Puli21	<i>pulicaria</i>	2x	1	0.46	0.23
Puli22	<i>pulicaria</i>	2x	1	0.46	0.23
W4-111-b1	<i>western pulicaria</i>	2x	3	0.49 \pm 0.007	0.25 \pm 0.004
W4-95	<i>western pulicaria</i>	2x	3	0.49 \pm 0.024	0.25 \pm 0.012

Stocks	identification (see text for methods)	ploidy level (inferred from microsatellite data)	n	Nuclear DNA content 2C-value Mean \pm SE (pg)	Monoploid genome size 1Cx-value Mean \pm SE (pg)
K207	western <i>pulicaria</i>	2x	2	0.47 \pm 0.012	0.23 \pm 0.006
A35-201	<i>middendorffiana</i> (Clade I)	3x	1	0.72	0.24
B141-201	<i>middendorffiana</i> (clade I)	3x	1	0.70	0.23
MIDA20	<i>middendorffiana</i> (Clade IIA)	3x	5	0.70 \pm 0.013	0.23 \pm 0.005
B141-202	<i>middendorffiana</i> (clade IIA)	3x	1	0.69	0.23
K31	<i>middendorffiana</i> (Clade IIB)	3x	2	0.76 \pm 0.031	0.25 \pm 0.010
K95	<i>middendorffiana</i> (clade III)	3x	1	0.67	0.22
K228	<i>middendorffiana</i> (clade III)	3x	1	0.65	0.22
K232	<i>middendorffiana</i> (clade III)	3x	1	0.64	0.21
CC3	<i>arenata</i>	2x	1	0.69	0.24
OP1	<i>arenata</i>	2x	5	0.48 \pm 0.011	0.24 \pm 0.005
Clone A	<i>tenebrosa</i>	3x	2	0.89 \pm 0.031	0.30 \pm 0.010
Clone B	<i>tenebrosa</i>	3x	2	0.71 \pm 0.012	0.24 \pm 0.004
Clone C	<i>tenebrosa</i>	3x	1	1.07	0.36
Clone D	<i>tenebrosa</i>	2x	2	0.62 \pm 0.021	0.31 \pm 0.010
Clone E	<i>tenebrosa</i>	2x	4	0.54 \pm 0.018	0.27 \pm 0.009
Clone F	<i>tenebrosa</i>	2x	2	0.57 \pm 0.024	0.28 \pm 0.012
Clone G	<i>tenebrosa</i>	3x	1	1.00	0.33
Clone H	<i>tenebrosa</i>	3x	1	1.03	0.34
Clone 13	<i>tenebrosa</i>	3x	4	0.66 \pm 0.026	0.22 \pm 0.009

Stocks	identification (see text for methods)	ploidy level (inferred from microsatellite data)	n	Nuclear DNA content 2C-value Mean \pm SE (pg)	Monoploid genome size 1Cx-value Mean \pm SE (pg)
Gerber 1a	european <i>pulicaria</i>	2x	1	0.46	0.23
Gerber 3	european <i>pulicaria</i>	2x	1	0.46	0.23
Chmelnice	european <i>pulicaria</i>	2x	5	0.57 \pm 0.022	0.28 \pm 0.011
Malá Kuš 5-1	european <i>pulicaria</i>	2x	1	0.52	0.26

2.5.3 Genome sizes and ploidy level

Nuclear DNA content and monoploid genome size estimates of the 49 clones are presented in Table 2-2. The average monoploid genome size (which is equivalent to the C-value) from different *D. pulex* clones found in the present study (0.23 ± 0.002 pg) is the closest estimate to the value of 0.2035 pg estimated through the *Daphnia pulex* genome sequencing (199×10^6 (bp)/(0.978×10^9), 0.978×10^9 being the number of base pairs contained in 1pg, (see Doležel *et al.*, 2003)). The low standard errors (± 0.001 pg to ± 0.044 pg) for repeated measurements conducted over 1 month on different individuals from the same clone (Table 2-2), also indicate the consistency of flow cytometric genome size estimates. Significant genome size differences were found among diploid species (Figure 2-2, P-value < 0.001, Kruskal-Wallis Test Statistic = 39.66). Average Nuclear DNA content for the five diploid species and significance of test for differences between 2C-values (GT2-method) are presented in the Table 2-3.

D. middendorffiana clones had nuclear DNA content from 0.64 to 0.38 pg (see Table 2-2). Due to the limited number of *D. middendorffiana* clones, no statistical test was applied to compare genome sizes of *D. middendorffiana* clones clustering in different lineages. Triploid clones (A35-201 and B141-201) from *D. middendorffiana sensu stricto*, had an average 2C-value of 0.71 ± 0.006 pg (mean \pm SE). *D. middendorffiana* in the polar *pulicaria* lineage (MIDA20 and B141-202), had an average overall 2C-value of 0.69 ± 0.012 pg. Clone K31 from the western *pulicaria* lineage had a 2C-value of 0.76 ± 0.031 pg. *D. middendorffiana* with eastern *pulicaria* mitochondria (K95, K228 and K232), had the smallest 2C-values (with an average of 0.65 ± 0.007 pg) of all the *D. middendorffiana* clones. The *D. middendorffiana* clones with *D. pulex* mtDNA had 2C-values ranging from 0.61 ± 0.002 to 0.70 ± 0.044 pg (Table 2-2). Nuclear DNA contents showed large variation among the five *D. tenebrosa* polyploid clones with values ranging from 0.71 ± 0.012 to 1.07 pg (Table 2-2). The average 2C-value of Clone 13 (0.66 ± 0.026 pg), known as an introgressed *D. pulex* with *D. tenebrosa* mitochondria (Dufresne et Hebert,

1994) is concordant with 1.5 times the average 2C-value of *D. pulex* diploids (see Table 2-2 and 2-3).

Tableau 2-3 Average 2C-values of *Daphnia* diploid species.

	2C-value					
	Mean \pm SE (pg)	<i>pulex</i>	<i>pulicaria</i>	<i>arenata</i>	<i>tenebrosa</i>	European <i>pulicaria</i>
<i>pulex</i>	0.45 \pm 0.003	-				
<i>pulicaria</i>	0.47 \pm 0.007	ns	-			
<i>arenata</i>	0.48 \pm 0.010	ns	ns	-		
<i>tenebrosa</i>	0.57 \pm 0.017	**	**	**	-	
European <i>pulicaria</i>	0.54 \pm 0.021	**	**	*	ns	-

* 0.01 < p < 0.05, ** p < 0.01, ns p > 0.05

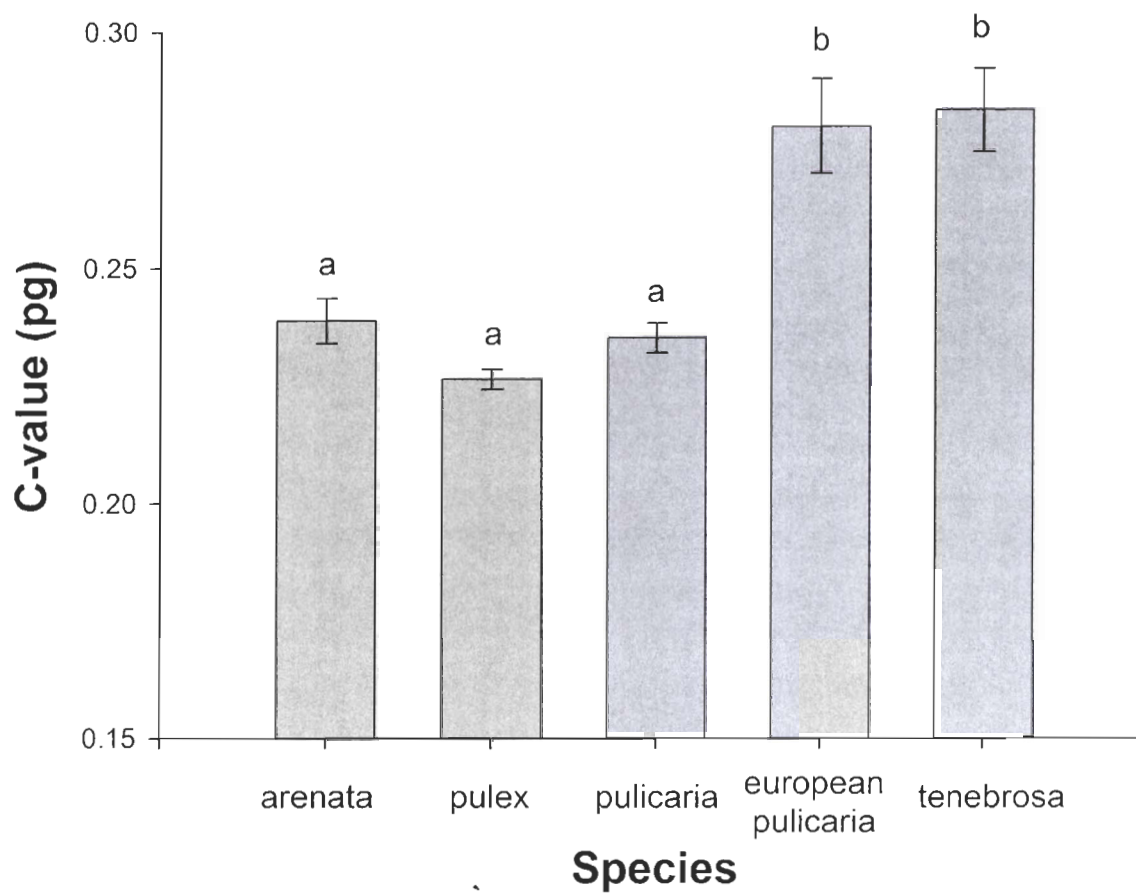


Figure 2-2 Nuclear DNA content of diploid clones of the *Daphnia pulex* complex. Mean 2C-values and standard error are shown. Significant differences in C-value ($P < 0.05$) are indicated by different letters.

2.6 Discussion

2.6.1 Genome size in diploid members of the *Daphnia pulex* complex

Our results have shown that diploid clones belonging to the two major groups in the *D. pulex* complex (*pulicaria* and *tenebrosa*) have significantly different nuclear DNA content. As shown in Figure 2-1, European populations of *Daphnia pulicaria* have mtDNA haplotypes closely related to those of *Daphnia tenebrosa* and a nuclear DNA content (0.54 pg, see Table 2-3) similar in size to the one of diploid clones of *D. tenebrosa* (0.57 pg). Diploid species from the *pulicaria* group all have smaller 2C-values: *D. pulex* (0.45 pg), *D. arenata* (0.48 pg), and *D. pulicaria* (0.47 pg). Clones in the *tenebrosa* group have the largest genomes (not arising as a result of polyploidy) in the whole genus *Daphnia* (Beaton, 1995), which indicates a genome size increase of 22% in this particular group. Larger genome sizes in the *tenebrosa* group may be due to a greater amount of repetitive DNA or to transposons (Petrov, 2001). Future research will aim to understand the mechanisms underlying genome size differences in *Daphnia*.

2.6.2 Evolution of polyploidy in the *Daphnia pulicaria* group

Nuclear DNA content values for polyploid clones of the *pulicaria* group were approximately 1.5X larger than those of diploid members of this group, suggesting a triploid status for these clones and not tetraploid status as was previously inferred (Beaton, 1995 ; Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994, 1997 ; but see Ward *et al.*, 1994). Previous measurements of nuclear DNA contents based on Feulgen staining suggested that polyploids had 1.7 times more DNA in their cells than diploid clones (Beaton, 1995 ; Dufresne et Hebert, 1995). Tetraploid status was inferred by assuming that these clones had lost DNA through genome rearrangements following polyploidy. However, results from the present study

indicate that previous measures of DNA content values, estimated by Feulgen staining, have been overestimated (For *D. pulex* C-values were 0.38 pg Beaton, 1995; 0.37 pg Beaton & Hebert, 1989; 0.23-0.25 pg Rasch, 1985). We report a genome size (1Cx-value in Table 2-2) of 0.23 ± 0.002 pg for *D. pulex* clone. Comparisons of genome size measurements obtained by both flow cytometry and Feulgen imaging in 42 species of moss have shown to be very similar (Voglmayr, 2000). Overestimation of DNA content based on Feulgen image analyses was found in marine crustaceans (Rees *et al.*, 2007). It is possible that the smaller number of nuclei available in certain animal groups (e.g. exopodites cells of the thoracic limbs, epidermal cells and other area underlying the rostrum, the central core area of the limbs and the heart in *Daphnia*) may represent a substantial bias for genome size measures obtained with Feulgen analyses. It is noteworthy that only five individuals were needed for genome size measurements which are much less than the 20-160 used by Korpelainen *et al.* (1997). Propidium iodide rather than ethidium bromide as well as the type of instrument used may explain these differences. Our genome size values, combined with the microsatellite results revealing no more than three alleles per locus in this group, indicate that these clones are triploids and not tetraploids. Polyploid clones belonging to the *Daphnia pulex* complex have recently been found in Bolivia and in Argentina. Argentinean *Daphnia* have been assigned a tetraploid status based on Feulgen staining (Adamowicz *et al.*, 2002) whereas the Bolivian *Daphnia* have also been inferred a tetraploid status but based on allele dosage at microsatellite loci (Aguilera *et al.*, 2007). The analysis of the two Argentinean *Daphnia* revealed a maximum of three alleles at 9 microsatellite loci (supplementary material, Table 2-5). This coupled to the fact that they are very closely related to *D. middendorffiana* clones highly suggests that they are also triploids and not tetraploids. We attempted to obtain ploidy level estimates based on microsatellite allele dosages using two methods: the MAC-PR (Esselink, Nybom et Vosman, 2004) and the method used by Aguilera *et al.* (2007). The former method consists of calculating ratios between the amplitude of the different peaks representing different alleles at one locus for one individual. The obtained ratios are then compared to the

expected ratios under the different ploidy level hypotheses (triploidy, tetraploidy, pentaploidy...) with different allele configurations (for example one locus with three alleles x, y and z in a tetraploid may present different phenotypes xxyz, xyyz and xyzz). In the second method, relative peak sizes were assessed in each individual PCR amplification. To account for differences in amplification intensity between PCRs, allele peak sizes are standardized to the allele with the smallest peak size for each PCR and locus, and relative peak sizes of all alleles per locus are averaged for each genotype. We tested both methods using two different clones (A35-201 and MIDA20). Results for the clone A35-201 indicated one dosage for allele 147, two dosages for allele 145, and one dosage for allele 143 at the Dp519 locus, congruent with a tetraploid status. For the clone MIDA20 at the Dp512 locus, results revealed a phenotype with two dosages of the 128 allele, one of the 135 allele, and two of the 137 alleles, thus indicating a pentaploid status. However the nuclear DNA content of A35-201 and MIDA20 were of 0.72 pg and of 0.70 pg respectively, thus clearly indicating a triploid status. Therefore results based solely on microsatellite allele dosage would have proved erroneous in our case. One important problem with these methods is the hybrid origins of many of these clones. We observed several cases of unequal amplification efficiency at microsatellite loci between the *pulicaria* and the *tenebrosa* groups. Thus clones that are hybrids between these two groups may not possess the same allele dosage and hence may be falsely classified as tetraploids. Therefore, we believe that ploidy level assessments obtained through microsatellite dosage alleles must be corroborated with DNA quantification analyses and/or chromosome analyses when feasible.

2.6.3 Genome size evolution of *D. tenebrosa* polyploids

D. tenebrosa clones never showed more than three alleles per locus, as found in *D. middendorffiana*, indicating a triploid nature. Flow cytometry estimates confirmed the presence of polyploid clones in this species but also revealed an

important variation in 2C-values in these clones (0.71-1.07 pg). The *D. tenebrosa* clones with 2C-values of 0.71 pg and 0.89 pg are likely triploids whereas the three *D. tenebrosa* clones with 2C-values ranging from 1.00 to 1.07 pg are probably tetraploids as they have nuclear DNA content of 1.76 to 1.89 fold the average diploid 2C-value. Polyploidy can be underestimated if based solely on limited microsatellite data. One of the putative triploid clones (clone B), with a 2C-value of 0.71 pg, has only one locus out of nine with three alleles. Clone D has a 2C-value of 0.62 pg but none of its loci presented a three allele pattern suggesting a diploid nature of this clone (screening six additional loci also failed to reveal three alleles). The absence of four alleles in the presumed tetraploid clones could be due to the diploidization of these loci. Similar diploidizations are observed at allozyme loci. Many polyploid clones show unbalanced phenotypes at only a few loci (Dufresne et Hebert, 1995). Moreover, for all alleles to differ at one locus in a tetraploid that arose by the chance union of gametes from the same species, the heterozygosity of the sexual diploid ancestor must be higher than 0.7 (Lokki et Saura, 1980). It is possible that none of the tetraploid clones tested here show four different alleles at a locus due to a low level of heterozygosity in their sexual diploid ancestors.

2.6.4 Origins of polyploidy in the *pulicaria* lineage

In North America, diploid *D. pulex* shows a longitudinal gradient in their mode of reproduction with obligately asexual populations in the eastern Canada, mixed populations in Ontario, and cyclically parthenogenetic populations in western Canada (Hebert et Crease, 1983 ; Hebert *et al.*, 1993b ; Innes, Schwartz et Hebert, 1986). Transitions to obligate parthenogenesis are thought to incur due to meiosis-suppressor genes that disrupt meiosis in females but not in males (Innes et Hebert, 1988). As a result, males carrying these genes spread the 'asexuality genes' in a contagious fashion. Meiosis-suppressor genes are thought to have arisen in an eastern *D. pulex* lineage, hence the geographical distribution of obligate

parthenogenesis in this species (Paland et Lynch, 2006). *D. pulicaria* shows an opposite pattern with obligate asexual populations found in western regions and cyclically parthenogenetic populations in eastern regions (Cerny et Hebert, 1993). These two species hybridize to form diploid apomictic clones in temperate regions and polyploid apomictic clones in subarctic and arctic regions (Crease, Stanton et Hebert, 1989 ; Dufresne et Hebert, 1994, 1997). The direction of hybridization in *D. middendorffiana* is highly asymmetric as the majority of these clones are known to have *D. pulicaria* as maternal parent (Dufresne et Hebert, 1994). We report here for the first time the occurrence of three triploid clones with *D. pulex* mtDNA (Disp1-21, K154 and K9). Their 2C-values ranged from 0.61 to 0.70 pg. Two of them are from Kuujuarapik, QC, Canada whereas the other triploid clone was from southern Ontario, a region that has been extensively sampled (Dufresne et Hebert, 1994, 1995, 1997 ; Ward *et al.*, 1994 ; Weider et Hobaek, 2003 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b). Are there more clones in arctic or subarctic regions with *D. pulex* mtDNA? Why do the majority of polyploid clones analysed so far have *D. pulicaria* mtDNA? These questions deserve additional studies *i.e.* more extensive sampling in eastern subarctic regions and investigation of mitochondrial-nuclear interactions in both types of hybrids.

2.6.5 Importance of triploidy in the *D. pulex* complex

The combined use of different techniques (microsatellite DNA and flow cytometry) have allowed us to rectify the misleading level of ploidy (tetraploidy) previously described with allozymes and Feulgen densitometry (Adamowicz *et al.*, 2002 ; Dufresne et Hebert, 1995, 1997 ; Ward *et al.*, 1994). The fact that most polyploids identified in this study are triploids suggests that asexuality is an important prerequisite for the evolution of polyploidy in *Daphnia*. As most invertebrate and vertebrate triploid clones are closely related to diploid ancestor species (Otto et Whitton, 2000) that have unisexual modes of reproduction (e.g. Thelytokous

parthenogenesis, gynogenesis, hybridogenesis...; as defined by Gregory & Mable, 2005) it seems that unisexuality in diploids is an important premise to the formation of triploid unisexuales.

In plants, polyploidy has been shown to generate novelty by genetic and epigenetic mechanisms (Levin, 1983 ; Salmon, 2005). Much less is known about the effects of polyploidization in animals. Members of the *Daphnia pulex* complex thus provide a useful model to study ecological and evolutionary implications of increase in ploidy levels in animals.

2.7 Acknowledgments

We thank Dr. MC Marinone for providing clones from Argentina, Drs. M Cristescu and C Caceres for providing some eastern *Daphnia pulicaria* and Dr. D Innes for providing polyploid *D. pulex* Disp1-21 from Ontario. The flow cytometry facility at ISMER was funded by the Canada Economic Development agency. We are thankful to Teresa J Crease and three anonymous reviewers who providing useful comments on this paper. This work was supported by a Natural Sciences and Engineering research Council (NSERC) research grants and by a Canadian Foundation for Innovation (CFI) equipment grants to France Dufresne. Roland Vergilino acknowledges a scholarship from Centre d'études nordiques.

Tableau 2-4 Multilocus genotypes of each strains used in our study
 000 denote alleles that we were not able to discern.

Clones	Dp183	Dp502	Dp525a	Dp519	Dp512	Dp514	Dp514a	Dp523	Dp513
CHQ6	093 097	106 106	103 103	151 153	127 137	101 104	127 135	130 143	115 117
EB1	093 097	106 109	103 103	145 153	137 143	101 102	122 130	136 143	115 117
POV1-102	093 093	103 109	103 108	153 155	135 137	100 101	129 130	131 132	100 117
DISP13	093 097	101 109	103 103	145 155	137 145	103 103	129 129	132 137	113 118
PET	093 093	106 109	103 103	147 153	137 143	099 100	129 129	132 133	115 117
LIS3	093 097	106 109	103 103	145 151	138 143	101 102	123 130	135 143	115 117
WG1	093 093	109 109	103 103	147 155	136 137	100 103	126 138	132 134	117 117
CU4-1	093 099	107 109	103 103	147 155	128 137	100 100	130 130	132 140	116 116
A17-200	097 101	109 115	103 103	147 147	128 137	095 103	116 132	133 133	115 115
Metis-3	093 099	106 115	103 103	147 153	137 138	095 102	116 116	130 134	115 115
Disp1-21	093 095 099	109 115 000	102 103 000	145 147 151	128 133 137	100 101 102	132 135 137	133 134 000	112 113 114
K52	093 099	109 115	103 103	147 153	137 137	099 102	132 142	135 135	114 116
K86	093 099	109 115	103 103	147 151	137 137	100 102	130 132	134 136	113 115
K92	093 099	109 115	102 103	147 153	128 137	100 102	128 132	132 143	113 115
K9	097 099 000	106 109 115	103 103 103	145 147 153	136 137 138	099 100 000	132 142 000	133 134 000	113 115 000

Clones	Dp183	Dp502	Dp525a	Dp519	Dp512	Dp514	Dp514a	Dp523	Dp513
K154	099 099 099	115 115 115	103 103 103	145 147 000	136 138 142	099 100 000	126 132 000	132 134 000	115 115 115
K230	099 099	115 115	103 103	145 145	138 138	098 102	132 138	133 134	116 116
C2	093 099	106 115	101 103	145 145	138 138	098 101	132 150	136 137	112 112
Puli9	092 092	112 112	102 102	155 155	133 138	101 101	132 138	139 139	112 112
Puli13	092 092	112 112	103 103	145 151	137 138	101 101	132 132	137 139	112 112
Puli15	092 092	112 112	103 103	155 155	133 138	101 101	132 138	137 139	112 112
Puli21	092 092	112 112	102 102	151 155	132 132	101 104	132 132	139 139	112 112
Puli22	092 092	112 112	103 103	145 151	137 138	101 101	132 132	135 135	112 112
W4-111-b1	099 099	115 115	101 102	145 145	138 142	098 103	132 132	133 133	116 116
W4-95	099 099	115 115	102 102	145 145	136 136	098 099	132 132	133 133	116 116
K207	099 099	109 115	103 103	145 147	137 137	099 102	132 132	133 134	115 116
A35-201	101 101 101	109 115 000	103 103 103	143 145 147	128 128 128	100 102 000	118 138 000	134 134 134	113 115 000
B141-201	099 101 000	109 115 000	103 103 103	145 151 000	128 130 135	102 102 102	118 132 000	132 133 000	113 113 113
MIDA20	099 101 000	109 115 000	103 108 000	145 145 145	128 135 137	100 102 105	118 126 132	132 134 135	112 113 000
B141-202	099 099 099	109 115 000	103 103 103	145 145 145	128 130 137	100 102 105	118 132 000	132 134 136	113 113 113
K31	099 099 099	115 115 115	103 103 103	145 147 000	135 137 138	099 100 102	132 138 000	133 134 000	113 113 113
K95	099 099 099	115 115 115	103 103 103	145 151 000	137 138 000	099 100 000	132 138 000	133 134 000	113 113 113

Clones	Dp183	Dp502	Dp525a	Dp519	Dp512	Dp514	Dp514a	Dp523	Dp513
K228	099 099 099	109 115 000	103 103 103	145 147 000	137 138 000	098 102 000	138 148 000	133 133 133	103 116 000
K232	093 099 000	112 115 000	103 103 103	145 147 149	137 138 000	98 101 102	132 138 148	132 133 134	113 113 113
CC3	093 093	093 117	000 000	157 157	134 143	095 095	134 134	137 137	096 096
OP1	093 093	117 117	108 108	152 154	133 133	095 095	143 143	139 140	096 096
Clone A	093 095 097	106 115 000	108 108 108	145 147 000	128 128 128	098 100 104	127 133 142	134 134 134	113 113 113
Clone B	093 095 101	106 127 000	108 108 108	147 147 147	123 123 123	100 104 000	124 127 000	133 133 133	107 107 107
Clone C	093 095 000	106 115 000	108 108 108	145 148 151	123 128 000	098 100 000	130 132 000	130 132 134	107 113 000
Clone D	095 097	106 115	108 108	145 147	128 128	100 104	124 145	131 135	107 113
Clone E	095 101	112 115	102 108	147 147	128 129	100 104	132 132	134 134	113 115
Clone F	095 095	109 115	108 108	145 147	128 128	097 104	124 124	125 131	107 115
Clone G	093 095 000	106 115 000	000 000 000	142 151 000	000 000 000	100 102 104	115 130 000	129 131 000	103 107 113
Clone H	000 000 000	106 109 000	108 108 108	145 149 151	124 127 128	098 105 107	127 129 132	129 133 134	109 113 000
Clone 13	093 097 000	101 109 000	103 103 103	147 149 151	128 137 000	095 097 103	123 129 148	133 135 136	105 115 000
Gerber 1a	099 099	115 115	108 108	145 145	128 128	100 100	154 154	135 135	112 112
Gerber 3	099 099	115 115	108 108	145 145	128 128	100 100	154 154	134 134	115 115
Chmelnice	099 099	112 112	108 108	145 145	128 130	099 100	153 153	138 139	113 113
MaláKuš5-1	092 092	112 112	108 108	145 149	128 128	100 100	153 153	139 139	113 113

Tableau 2-5 Multilocus genotypes of each Argentinean *Daphnia*.

Clones	Dp183	Dp502	Dp525a	Dp519	Dp512	Dp514	Dp514a	Dp523	Dp513
Argsc1	093 097 099	106 109 118	103 103 103	147 148 151	129 135 137	094 099 101	129 135 153	134 134 134	111 112 113
Argsc4	093 097 099	106 109 124	103 103 103	147 148 151	129 135 137	094 099 101	129 135 150	134 134 134	111 112 113

Chapitre III

RETICULATE EVOLUTION OF THE *DAPHNIA PULEX* COMPLEX AS REVEALED
BY NUCLEAR MARKERS

Roland Vergilino, France Dufresne

*Département de Biologie, Université du Québec à Rimouski, Québec G5L 3A1,
Canada*

(Roland Vergilino : conception du projet, échantillonnage, expérimentations,
traitement statistique, rédaction de l'article; France Dufresne : conception du projet,
échantillonnage, rédaction, financement)

Accepté le 23 décembre 2010 dans la revue *Molecular Ecology*. Les références
citées dans l'article sont reportées à la fin de la thèse.

3.1 Résumé

L'étude des complexes d'espèces est particulièrement intéressante afin de comprendre comment les espèces jeunes maintiennent l'intégrité de leur génome. Le complexe *Daphnia pulex* comprend des espèces dominant les environnements d'eau douce dans l'hémisphère Nord. Des populations de certaines espèces de ce complexe ont transité vers une reproduction parthénogénétique obligatoire. Des études antérieures utilisant des marqueurs mitochondriaux ont révélé la présence de 10 différents lignées dans le complexe. Cette étude est la première à examiner les relations génétiques entre les sept espèces du complexe en utilisant des marqueurs nucléaires (9 microsatellites et un gène codant pour une protéine). Des clones appartenant aux sept espèces du complexe *Daphnia pulex* ont été caractérisés aux niveaux du gène mitochondrial *NADH5* et du locus codant pour la lactate déshydrogénase (Ldh). Des analyses de coordonnées principales (PCO), de K-means, et un réseau phylogénétique sur les données microsatellites ont permis de séparer les clones de *D. pulicaria* européennes, de *D. tenebrosa*, de *D. pulicaria* d'Amérique du Nord, de *D. pulex*, et leurs hybrides en groupes distincts. Le groupe de génotypes hybrides est composé d'hybrides diploïdes et polyploïdes avec des mitochondries *D. pulex* et certains clones avec des mitochondries *D. pulicaria*. En revanche, la phylogénie du complexe *D. pulex* à partir du gène *Rab4* n'était pas précisément résolue, elle présente tout de même des groupes majoritairement constitués d'allèles *D. pulex* et *D. pulicaria*. Le maintien du polymorphisme ancestral et l'hybridation peuvent affecter les relations génétiques à ce locus. Cette étude montre que l'hybridation et l'introgession ont joué un rôle important dans l'évolution de ce complexe.

3.2 Abstract

The study of species complexes is of particular interest to understand how evolutionary young species maintain genomic integrity. The *Daphnia pulex* complex has been intensively studied as it includes species that dominate freshwater environments in the Northern hemisphere and as it is the sole North American complex that shows transitions to obligate parthenogenesis. Past studies using mitochondrial markers have revealed the presence of 10 distinct lineages in the complex. This study is the first to examine genetic relationships among seven species of the complex at nuclear markers (9 microsatellite loci and one protein-coding gene). Clones belonging to the seven species of the *Daphnia pulex* complex were characterized at the mitochondrial NADH dehydrogenase (*ND5*) gene and at the Lactate dehydrogenase (LDH) locus. K-means, principal coordinate analyses, and phylogenetic network analyses on the microsatellite data all separated European *D. pulicaria*, *D. tenebrosa*, North American *D. pulex*, *D. pulicaria*, and their hybrids into distinct clusters. The hybrid cluster was composed of diploid and polyploid hybrids with *D. pulex* mitochondria and some clones with *D. pulicaria* mitochondria. By contrast, the phylogeny of the *D. pulex* complex using *Rab4* was not well resolved but still showed clusters consisting mostly of *D. pulex* alleles and others of *D. pulicaria* alleles. Incomplete lineage sorting and hybridization may obscure genetic relationships at this locus. This study shows that hybridization and introgression have played an important role in the evolution of this complex.

3.3 Introduction

The debate on species concept is one of the oldest in biology (Mallet, 2001 ; Mayr, 1963). Despite the fact that more than 20 species concepts are currently acknowledged (Mayden, 1997), the most widely used is still the Biological Species Concept (BSC) (Mayr, 1963) even though it can not be applied to numerous cases such as parthenogenetic species (Cohan, 1994). Species complexes are of particular interest in the debate of species definition and species delimitation. As species belonging to species complexes are evolutionarily young and sometimes not well differentiated, hybrids are encountered in several complexes (reviewed in Seehausen, 2004). In contrast to previous views (Mayr, 1963), hybridization may have a substantial role in evolution and species diversification (Dowling et Secor, 1997 ; Kearney, 2005 ; Seehausen, 2004). A growing list of species appears to be of hybrid origin (Delmotte *et al.*, 2003 ; Seehausen, 2004). Empirical evidence has shown that hybrids may have higher fitness than their parental species or if not they occupy different niches (Lewontin et Birch, 1966 ; Seehausen, 2004 ; Templeton, 1981). Association of genetic diversity fixed in different parental species within hybrid or introgressed species via transgressive segregation may help facilitate adaptive radiation (Bell et Travis, 2005 ; Rieseberg, Archer et Wayne, 1999). Studying the hybridization process and gene flow within species complexes may help to understand speciation and to disentangle species delimitation problems.

As species in species complexes may have diverged only recently, we might expect that shared genetic polymorphisms may be retained in sister species (incomplete lineage sorting) and blur phylogenetic relationships between them. Hybrids between these species may be difficult to differentiate from pure species as a consequence of incomplete lineage sorting and hybridization superimposed signals. Using genetic markers with high mutation rate we might expect to disentangle speciation and hybridization relationships in species complex more

readily than with markers with low mutation rate. Species complexes with reticulate evolution and multiple transitions to polyploidy are frequent in plants (Wendel, 2000) but less prevalent and less well studied among animals (Mable, 2004). Hence the molecular scrutiny of an animal species complex at the nuclear level could provide important insights into how the genomes of young species maintain their integrity through time.

The *Daphnia pulex* complex has been intensively studied due to its dominance in freshwater habitats in North America and to its propensity for hybridization and its variation in reproductive mode and ploidy level. This complex includes seven species that have been distinguished on the basis of morphological, ecological, and genetic data (Colbourne *et al.*, 1998 ; Hebert, 1995 ; Marková *et al.*, 2007). A portion of the *ND5* gene has proven very reliable in distinguishing the various species of the *D. pulex* complex (Colbourne *et al.*, 1998). Previous studies, based on mitochondrial genes, have revealed the presence of two major groups in this complex. The *pulicaria* group consists of seven different lineages, Eastern *D. pulicaria* (EPC), Western *D. pulicaria* (WPC), Polar *D. pulicaria* (PPC) (these three lineages are known as North-American *D. pulicaria*), Panarctic *D. pulex* (PanPX), *D. melanica*, the hybrid taxon *D. middendorffiana* (MIDD), and *D. arenata*, endemic to Oregon ponds) (Colbourne *et al.*, 1998 ; Colbourne et Hebert, 1996 ; Marková *et al.*, 2007 ; Omilian et Lynch, 2009 ; Omilian, Scofield et Lynch, 2008). The *tenebrosa* group includes two lineages: European *D. pulicaria* (EuroPC) and *D. tenebrosa* (TENE) (Colbourne *et al.*, 1998). Although the mitochondrial gene phylogeny of this complex is relatively well understood, nuclear genetic relationships among species remain unclear. The elucidation of these links gets its importance as the full genome of a clone from the *D. pulex* complex belonging to *D. arenata* has just been sequenced and the relationship of *D. arenata* with other species at the nuclear level is not fully understood. The two dominant species of the complex *D. pulex* and *D. pulicaria* are morphologically cryptic but ecologically distinct species (Hebert, 1995). *D. pulex* lives in ponds and *D. pulicaria* is found in lakes (Hebert, 1995 ; Hebert *et al.*, 1993b ; Heier et Dudycha, 2009) in eastern North America (Hebert, 1995). In

western North-America these two species may share the same habitat in ponds (Brooks, 1957b ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b) but *D. pulicaria* is dominant in lakes (Hebert *et al.*, 1993b). These two morphologically cryptic species can be differentiated at the mitochondrial *ND5* gene in populations north of Maine (Dufresne et Hebert, 1994) and by their Ldh allozyme phenotypes in all populations. The lactate dehydrogenase enzyme (Ldh) variation has been found to be diagnostic (Hebert *et al.*, 1989 ; Hebert *et al.*, 1993b). *D. pulex* is fixed for slower allozymic form (S) whereas the fastest (F) is fixed in *D. pulicaria* (Hebert *et al.*, 1989). *Daphnia* reproduce by cyclic parthenogenesis that is clonal reproduction interrupted by bouts of sexual reproduction. *D. pulex* and *D. pulicaria* both show transitions to obligate parthenogenesis as a result of a mutation that suppress meiosis in females but not in males (Innes, Fox et Winsor, 2000). Obligate parthenogenesis in the *pulex* complex may also be triggered by hybridization (Hebert et Finston, 2001b ; Innes et Hebert, 1988). *D. pulex* and *D. pulicaria* hybridize in nature and are heterozygous for the Ldh phenotype [SF] and appear to occupy different niches than their parental species (Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989 ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b ; Innes et Hebert, 1988). Diploid hybrids always have *D. pulex* mitochondria and have been found to reproduce by obligate parthenogenesis in nature (Hebert et Finston, 2001b ; Innes et Hebert, 1988). Analyses of allozymic variation indicate that introgression was rare or non-existent in areas where the two species co-occured (Brooks, 1957a ; Hebert et Finston, 2001b ; Hebert *et al.*, 1993b ; Innes, Schwartz et Hebert, 1986). Two recent studies show evidence of possible or substantial gene flow between *D. pulex* and *D. pulicaria* (Heier et Dudycha, 2009 ; Omilian et Lynch, 2009). Hybrids produced by crosses between *D. pulex* and *D. pulicaria* have been found to reproduce sexually in the laboratory and may be able to backcross with parental species (Heier et Dudycha, 2009). Omilian and Lynch (2009) have shown that low but significant gene flows exist *in natura* between *D. pulex* and *D. pulicaria*.

Polyploidy has occurred repeatedly in the *D. pulex* complex (Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994, 1995, 1997). One group of polyploids

(collectively known as *D. middendorffiana*) is thought to have arisen from hybridization between *D. pulex* males and *D. pulicaria* females (Beaton & Hebert 1988; Dufresne & Hebert 1994, 1997). A recent study, using microsatellite data, flow cytometry and mitochondrial sequences, has shown that most polyploids of the *Daphnia pulex* complex are triploids (Vergilino, Belzile et Dufresne, 2009) and they reproduce by obligate parthenogenesis. The other group of polyploids has been found in *D. tenebrosa* group, a circumarctic endemic species (Dufresne et Hebert, 1995) and includes both triploid and tetraploid clones (Vergilino, Belzile et Dufresne, 2009). It is not clear whether these polyploids have hybrid origins. A survey based on allozyme data from 850 individuals coming from various populations the circumarctic region suggested the possibility of genetic exchange between the *D. pulicaria* and the *tenebrosa* groups (Weider *et al.*, 1999a ; Weider *et al.*, 1999b). As allozymes show limited polymorphism and might be under selection, genetic exchanges between populations of these groups needs to be confirmed with additional nuclear markers. Clearly nuclear data are lacking in the literature to properly describe evolutionary relationships between species and the origins of hybrids in the *Daphnia pulex* complex.

The purpose of this study was to disentangle evolutionary relationships between seven species (*D. pulex*, North-American *D. pulicaria*, *D. middendorffiana*, *D. melanica*, *D. arenata*, *D. tenebrosa* and European *D. pulicaria*) of the *Daphnia pulex* complex and to assess the hybrid origin of some diploid and polyploid clones belonging to this complex using information from both nuclear and mitochondrial markers.

3.4 Material and methods

3.4.1 Samples and DNA extraction

One hundred and five clones belonging to 7 species of the *Daphnia pulex* complex from different locations were used in this study (Table 3-1). The majority of these clones have been sampled from 2004-2010 and maintained in culture in the laboratory. Species identity of the clones was assessed by combining information on morphology, *ND5* haplotypes, Ldh phenotypes/genotypes, and ecology (whether clones came from ponds or lakes). Total DNA was extracted from individuals using the QuickExtract DNA extraction solution (EPICENTRE® Biotechnologies, Madison, WI, U.S.A.) using 50µl of solution per individual.

Tableau 3-1 Identification, ploidy, and geographic origins of *Daphnia* isolates used in the present study

(availabilities of partial *ND5* sequences, multilocus genotype or partial *Rab4* sequences are represented by crosses in the *ND5*, *MLG* or *Rab4* columns respectively, EuroPX, European *D. pulex*; WPC, Western *D. pulicaria*; MIDD, *D. middendorffiana*; PPC, Polar *D. pulicaria*; PanPX, Panarctic *D. pulex*; EPC, Eastern *D. pulicaria*; PCX, *D. pulicaria* with “*D. pulex* haplotype”; EuroPC, European *D. pulicaria*; TENE, *D. tenebrosa*; ARE, *D. arenata*; MEL, *D. melanica*; ARG; *Daphnia* sp. polyploids from Argentina). Hybrid states [SF] between *D. pulex* [SS] and between *D. pulicaria* [FF] are indicated in the LDH column. * indicated that Ldh phenotypes come from previous allozyme analyses.

code	<i>ND5</i>	<i>MLG</i>	<i>Rab4</i>	ploidy	LDH	Identification	Location
A35_201	x	x	x	3x	SS	MIDD	Churchill, Manitoba, CAN
A17_200	x	x	x	2x	SS	PanPX	Churchill, Manitoba, CAN
Argsc1	x	x		3x/4x		ARG	Patagonia, Argentina
Argsc4	x	x		3x/4x		ARG	Patagonia, Argentina
B141_201	x	x	x	3x	SF	MIDD	Churchill, Manitoba, CAN
B141_202	x	x	x	3x	SF	PPC	Churchill, Manitoba, CAN
BGR3_10		x		3x		PanPX	Kuujjarapik, Québec, CAN
BGR4_5		x		3x		PanPX	Kuujjarapik, Québec, CAN
BGR8_2	x	x		3x		EPC	Kuujjarapik, Québec, CAN
BGR9_7	x	x		3x		PanPX	Kuujjarapik, Québec, CAN
Bohdanec	x	x		2x		EuroPC	Lowland, Czech Republic
BW102	x	x	x	2x	SS*	PanPX	Busey Woods, Illinois, USA
C1_3		x		2x	SF*	PanPX	Churchill, Manitoba, CAN
C2		x		2x	FF*	PCX	Vermillion County, Illinois, USA

code	<i>ND5</i>	<i>MLG</i>	<i>Rab4</i>	ploidy	LDH	Identification	Location
C12_300	x	x	x	3x	SF	WPC	Churchill, Manitoba, CAN
C12_301	x	x	x	2x	SF	PanPX	Churchill, Manitoba, CAN
C88		x		2x	SS	PanPX	Churchill, Manitoba, CAN
Can2_1_11	x	x	x	2x	SS	PanPX	Windsor, Ontario, CAN
Can3_1_1	x	x	x	2x	SF	PanPX	Windsor, Ontario, CAN
Can3_1_20	x	x	x	2x	SF	PanPX	Windsor, Ontario, CAN
CC3	x	x	x	2x		ARE	Creswell Court, Oregon, USA
CHQ6	x	x		2x	SS	PanPX	Chequamegon, Wisconsin, USA
Chmelnice	x	x	x	2x		EuroPC	Lowland, Czech Republic
Clear13	x	x	x	3x	SS*	Introgressed TENE	Churchill, Manitoba, CAN
Clone A	x	x	x	3x		TENE	Churchill, Manitoba, CAN
Clone B	x	x	x	3x/4x		TENE	Churchill, Manitoba, CAN
Clone C	x	x	x	3x		TENE	Churchill, Manitoba, CAN
Clone D	x	x	x	2x		TENE	Churchill, Manitoba, CAN
Clone E	x	x	x	2x		TENE	Churchill, Manitoba, CAN
Clone F	x	x	x	2x		TENE	Churchill, Manitoba, CAN
Clone G		x		3x		TENE	Churchill, Manitoba, CAN
Clone H	x	x		3x		TENE	Churchill, Manitoba, CAN

code	ND5	MLG	Rab4	ploidy	LDH	Identification	Location
Clone J	x	x	x	2x		TENE	Churchill, Manitoba, CAN
CR1		x		2x	SS*	PanPX	Closed Road, Michigan, USA
CU4-1	x	x	x	2x	SS	PanPX	Duchénier IV, Québec, CAN
CZ	x	x		2x		EuroPC	High Tatra Mountain, Slovakia
DP003_003	x		x	n-a		EuroPC	High Tatra Mountain, Slovakia
DP0015_001	x		x	n-a		EuroPC	High Tatra Mountain, Slovakia
DP0024_002	x		x	n-a		EuroPC	Antermoia, Italy
DP0076_002	x		x	n-a		EuroPX	Sweden
DP0077_003	x		x	n-a		EuroPC	Sweden
DP0081_001	x		x	n-a		EuroPX	Lithuania
DP0085_003	x		x	n-a		PanPX	Avigliana, Italy
DISP1_1_24	x	x	x	2x	SS*	PanPX	Disputed Road, Ontario, CAN
DISP1_21	x	x	x	3x	SF*	PanPX	Disputed Road, Ontario, CAN
DISP13	x	x	x	2x	SS	PanPX	Disputed Road, Ontario, CAN
EB1	x	x	x	2x	SS*	PanPX	Eloise Butler, Minnesota, USA
Fence1_1	x	x	x	2x	SF	PanPX	Windsor, Ontario, CAN
FIS	x	x	x	2x	SF	PanPX	Ste-Foy, Québec, CAN
GAR3	x	x	x	2x	SF	PanPX	Ste-Foy, Québec, CAN
Gerber1a	x	x	x	2x		EuroPC	Catalonia, Spain
Gerber1b	x	x		2x		EuroPC	Catalonia, Spain
Gerber3	x	x		2x		EuroPC	Catalonia, Spain
IWR10	x	x	x	2x	SF	PanPX	Michigan, USA

code	ND5	MLG	Rab4	ploidy	LDH	Identification	Location
K9	x	x	x	3x	SF	PanPX	Kuujjarapik, Québec, CAN
K52	x	x	x	2x		PanPX	Kuujjarapik, Québec, CAN
K31	x	x		3x	SF	WPC	Kuujjarapik, Québec, CAN
K86	x	x	x	2x	SF	PanPX	Kuujjarapik, Québec, CAN
K92	x	x	x	2x	SF	PanPX	Kuujjarapik, Québec, CAN
K95	x	x	x	3x	FF	EPC	Kuujjarapik, Québec, CAN
K154	x	x	x	3x	SF	PanPX	Kuujjarapik, Québec, CAN
K207	x	x	x	2x	FF	WPC	Kuujjarapik, Québec, CAN
K228	x	x	x	3x	SF	EPC	Kuujjarapik, Québec, CAN
K230	x	x	x	2x	FF	EPC	Kuujjarapik, Québec, CAN
K232	x	x	x	3x	SF	EPC	Kuujjarapik, Québec, CAN
L1_9	x	x		3x		PanPX	Kuujjarapik, Québec, CAN
L3_11	x	x		2x		EPC	Kuujjarapik, Québec, CAN
L3_12	x	x		2x		EPC	Kuujjarapik, Québec, CAN
LIS3	x	x		2x	SS*	PanPX	Listowel Pond, Ontario, CAN
LP8B7		x		2x	SS*	PanPX	Long Point, Ontario, CAN
Malakus5_1	x	x	x	2x		EuroPC	Lowland, Czech Republic
Metis3	x	x	x	2x	SF	PanPX	Metis , Québec, CAN

code	ND5	MLG	Rab4	ploidy	LDH	Identification	Location
MIDA20	x	x	x	3x	SF	PPC	Churchill, Manitoba, CAN
NFL1		x		2x	SS*	PanPX	North Flatley, Indiana, USA
NOS1		x		2x	SS*	PanPX	Nerves of Steel, Minnesota, USA
OP1	x	x	x	2x		ARE	Oak Patch, Oregon, USA
PET	x	x		2x	SF*	PanPX	Rimouski, Québec, CAN
POVI	x	x		2x	SS*	PanPX	Povi Pond, Michigan, USA
Ppuli1		x		2x		EPC	Sainte Adèle, Québec, CAN
Ppuli2		x		2x		EPC	Sainte Adèle, Québec, CAN
Puli9	x	x	x	2x	FF	PCX	Vermillion County, Illinois, USA
Puli13	x	x		2x	FF	PCX	Vermillion County, Illinois, USA
Puli15	x	x	x	2x	FF	PCX	Vermillion County, Illinois, USA
Puli21	x	x		2x	FF	PCX	Vermillion County, Illinois, USA
Puli22	x	x		2x	FF	PCX	Vermillion County, Illinois, USA
R2_10_2	x	x		2x		PanPX	Kuujjarapik, Québec, CAN
R2_14_4	x	x		2x		PanPX	Kuujjarapik, Québec, CAN
RES4	x	x		2x	SS*	PanPX	Restigouche, New Brunswick, CAN
Rimov	x	x		2x		EuroPC	Lowland, Czech Republic
Slapy2	x	x		2x		EuroPC	Lowland, Czech Republic

code	ND5	MLG	Rab4	ploidy	LDH	Identification	Location
SQR141	x	x	x	2x		MEL	Oregon, USA
ST04_1		x				EPC	Sainte Adèle, Québec, CAN
ST04_2		x		2x		EPC	Sainte Adèle, Québec, CAN
STpuli_1		x		2x		EPC	Sainte Adèle, Québec, CAN
STpuli_2		x		2x		EPC	Sainte Adèle, Québec, CAN
T1_07_8	x	x		2x		EPC	Kuujjarapik, Québec, CAN
T1_07_11	x	x		2x		EPC	Kuujjarapik, Québec, CAN
T1_07_20	x	x		3x		WPC	Kuujjarapik, Québec, CAN
T1_08_14	x	x		3x		PanPX	Kuujjarapik, Québec, CAN
T1_08_20	x	x		3x		PanPX	Kuujjarapik, Québec, CAN
T2_02_4	x	x		2x		EPC	Kuujjarapik, Québec, CAN
T2_07_6	x	x		2x		EPC	Kuujjarapik, Québec, CAN
T2_07_7	x	x		2x		EPC	Kuujjarapik, Québec, CAN
T2_08_4	x	x		2x		EPC	Kuujjarapik, Québec, CAN
TRO3	x	x		2x	SS*	PanPX	Troy II, Maine, USA
U2_7	x	x		3x		PanPX	Kuujjarapik, Québec, CAN
W4_95	x	x	x	2x	FF	WPC	Indiana, USA
W4_111B1	x	x	x	2x	FF	WPC	Indiana, USA
WCH1		x		2x	SS*	PanPX	Winchell, Ohio, USA
WG1		x		2x	SS*	PanPX	West Gull, Michigan, USA

code	<i>ND5</i>	MLG	<i>Rab4</i>	ploidy	LDH	Identification	Location
Winni1_1	x	x		2x	FF	WPC	Winnipeg, Manitoba, CAN
Winni1_2	x	x	x	2x	FF	WPC	Winnipeg, Manitoba, CAN

3.4.2 PCR amplifications

3.4.2.1 *LDH-A* genotypes

A PCR based method was performed to determine the SS, FF and SF genotypes of the *LDH-A* locus of 38 clones of *D. pulex*, *D. pulicaria*, diploid and polyploid hybrids (see Table 2-1). The expression of the *LDH-A* gene appears responsible for the Ldh phenotypes identified using allozymes (Cristescu *et al.*, 2008) hence the *LDH-A* genotypes give the same information as Ldh allozymic phenotypes. In diploids, the [SS] phenotypes represent pure *D. pulex* individuals, [FF] phenotypes represent pure *D. pulicaria* individuals, and [SF] phenotypes represent hybrids between these two species. Specific primers (provided by Dr. TJ Crease) were used to amplify separately two different amplicons of 248 nt. of the *LDH-A* gene. LDH-A-SF (5'-GAGCGATTTAACGTTGCGCCC-3') and LDH-A-SR (5'-GGACGACTTGTGTGTGAATTTG-3') were used to amplify S alleles known to be diagnostic to "pure" *D. pulex*. LDH-A-FF (5'-GAGCGATTTAACGTTGCGCCT-3') and LDH-A-FR (5'-GGACGACTTGTGTGTGAATTTTC-3') were used to amplify F alleles known to be diagnostic to "pure" *D. pulicaria*. For each clones two PCR were performed with the following recipe: PCR conditions are one cycle of denaturation at 94°C for 3 min, 40 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min, and extension at 72°C for 2 min and a final extension at 72°C for 15 min. PCR were repeated twice with a "pure" *D. pulex* clone and a "pure" *D. pulicaria* clone as controls to account for false positives and false negatives.

3.4.2.2 Mitochondrial haplotypes

DNA amplifications of a portion of the mitochondrial *ND5* gene were carried out on 92 clones as in Vergilino *et al.* (2009).

3.4.2.3 *Rab4* subunit gene of the GTPase family

A recent study has shown the potential of the *Rab4* subunit gene of the GTPase family to differentiate between *D. pulex*, *D. pulicaria*, and *D. arenata* populations (Omilian *et al.* 2008). As such, this locus should be helpful to disentangle hybrid origins in the *D. pulex* complex. A portion of the *Rab4* subunit gene containing three to four exons, two to three introns and a 3'UTR region (Omilian, Scofield et Lynch, 2008) was amplified in 51 clones belonging to seven species and accounting for the diversity of ploidy and hybrid state of the *pulex* complex. Portions of *Rab4* subunit gene were amplified using the protocol of Omilian *et al.* (2008). As diploid and polyploid hybrids may have multiple alleles of the *Rab4* gene, alleles were separated using the SSCP (Single-Stranded Conformation Polymorphism) method (Orita *et al.*, 1989). PCR products migrated on a 6% non-denaturing polyacrylamide gel (acrylamide/ bisacrylamide 37.5:1) for 16 h at 20 W at 4°C in 0.5x TBE. Alleles were recovered from the gel by cutting each band with a sharp scalpel. DNA bands were extracted from the gel using the modified protocol of the QIAquick® gel extraction kit (QIAGEN Inc., Mississauga, ON, Canada) or the freeze/thaw method. Alleles were reamplified following the protocol of Omilian *et al.* (2008) and sequenced.

3.4.2.4 Microsatellite markers

Each clone was genetically identified at nine microsatellite loci [Dp183,

Dp502, Dp512, Dp513, Dp514, Dp514alt, Dp519, Dp523, Dp525alt; accession numbers: Wfms0000187, Wfms0000519, Wfms0000529, Wfms0000530, Wfms0000531, Wfms0000532, Wfms0000537, Wfms0000541, Wfms0000544 respectively described in Colbourne *et al.* (2004) and Cristescu *et al.* (2006)]. Polymerase chain reactions (PCR) were performed in 12- μ l reactions with 10 ng of DNA template, 1 \times PCR buffer with 25 nmol of MgCl₂, 0.5 units of Taq polymerase, 2.5 nmol of each dNTP, 2 pmol of the forward primer, 2 pmol of the reverse primer. Thermal cycle programs included an initial denaturation step of 3 min. at 95°C followed by 10 cycles of 35 sec. denaturation at 94°C, 35 sec. at a initial annealing temperature of 55°C for Dp183, Dp502, Dp512, Dp514, Dp514alt, Dp519, Dp523, Dp525alt and 64°C for Dp513 (the annealing temperature was decreased by 1°C every cycle during each of the 6 following cycles), 45 sec. extension at 72°C followed by 30 cycles of 35 sec. denaturation at 94°C, 35 sec. annealing temperature at 48°C or 59°C , and 45 sec. extension at 72°C, with a final extension at 72°C for 10 min PCR products were denatured at 95°C for 3 minutes and alleles were separated on a denaturing polyacrylamide gel containing 6% acrylamide/bis-acrylamide (19:1) at 50 Watts and 40°C for 2 to 3h. Fluorescent dye-labeled alleles were detected on a FMBIO® III fluorescent scanner (Hitachi Software Engineering America Ltd., San Bruno, CA, U.S.A.) and analyzed using the software ImageAnalysis (v.3.0.0.21 Miraibio, San Francisco, CA, U.S.A.). Ploidy level for each clone was estimated by counting the number of alleles at each locus. The assumption of ploidy level n is rejected when level n+1 is met at a given locus. With this method, there is a risk of erroneously rejecting level n+1, but this risk decreases with the number of loci used and the number of alleles per locus.

3.4.3 Phylogenetic Analyses of *ND5* and *Rab4* subunit of the GTPase family

Partial *Rab4* and *ND5* sequences were manually aligned using the ClustalW module of BioEdit v.7.0.5 (Hall T., unpublished; <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

Partial *ND5* genes were sequenced from 92 clones (GenBank [accession numbers: HQ434626 to HQ434690 and FJ591096 to FJ591124]). Only 51 mitochondrial sequences were used to unravel mitochondrial haplotype evolutionary relationships in a subset of clones studied at the nuclear level (see Table 3-1). Aligned sequences of a portion of the *ND5* gene were imported into MEGA 4.0 (Tamura *et al.*, 2007) for analyses of sequence divergence using the maximum composite likelihood model of nucleotide substitution.

Two phylogenetic Bayesian inferences were performed using MrBayes v3.1.2 (Ronquist et Huelsenbeck, 2003). A Bayesian inference was conducted and allowed three codon positions (1st, 2nd, and 3rd codon positions) to evolve freely. A model of sequence evolution was identified using Topali v2.5 (Milne *et al.*, 2009) using the hierarchical likelihood ratio test (hLRT) and the Akaike criterion (AIC). The DNA substitution model selected by Topali v2.5 for the Bayesian inference is a Hasegawa, Kishino and Yano model with a Gamma shape (HKY+G) with a gamma substitution parameter (α) of 0.362, a transition/transversion ratio of 13.375, and fixed base frequencies of A=0.185, C=0.192, G=0.215, T=0.408. Markov chain Monte Carlo analyses were run for twelve million generations, sampling from the chain every 100 generations. After determining chain convergence (average standard deviation of split frequencies <0.01), we discarded the initial 25% of trees as the “burn-in period” and a 50% majority-rule consensus tree with posterior probability (PP) values for each node was constructed from the remaining Bayesian trees.

Seventy one sequences (GenBank [accession numbers: HQ434691 to HQ434761]) of a portion of *Rab4* subunit from the fifty one clones tested for partial *ND5* gene were aligned and used to construct phylogenetic bayesian inference.

Eventual recombination events during the *Rab4* gene evolution in the *Daphnia pulex* complex were tested on the dataset without indels using the Phi-test (Bruen, Philippe et Bryant, 2006) implemented in the Splitstree4 version 4.10 software (Huson et Bryant, 2006). All the insertion/deletion polymorphic sites were excluded for phylogenetic inferences. Two different Bayesian phylogenetic inferences were performed using MrBayes v3.1.2 (Ronquist et Huelsenbeck, 2003). First the sequence data were partitioned into one non-coding (introns and UTR) and three codon positions (1st, 2nd, and 3rd codon positions). In a second time, model of sequence evolution were identified using Topali v2.5 (Milne *et al.* 2009). A HKY+G model of DNA substitution with gamma substitution parameter (α) of 0.158, a fixed transition/transversion rate ratio of 3.174, and fixed base frequencies of A=0.313, C=0.168, G=0.189, T=0.331 was selected using both the hLRT and AIC criteria. Markov chain Monte Carlo analyses were run for twenty million generations, sampling from the chain every 100 generations. After determining chain convergence (average standard deviation of split frequencies <0.01), we discarded the initial 25% of trees as the “burn-in period”. A 50% majority-rule consensus tree with posterior probability (PP) values for each node was constructed from the remaining Bayesian trees.

3.4.4 Tests of congruencies between trees

Congruencies between mitochondrial and nuclear gene trees were tested using the incongruence length difference test (ILD; (Farris *et al.*, 1994)) and the SH-test (Shimodaira et Hasegawa, 1999) implemented in PAUP* version 4.0b.10 (Swofford, 2003). As in our dataset an individual may carry more than one *Rab4* allele, we adjusted the numbers of sequences of *ND5* to be comparable to the *Rab4* alleles dataset using the same sequence twice or three times in the first dataset (1+1, 2+1 and 3+1). The ILD test for incongruence tests the null hypothesis of congruence between data sets (Farris *et al.*, 1994). Parsimony analyses are carried

out for data set X (here the *Rab4* data matrix) and data set Y (here the *ND5* data matrix) separately. Then, the lengths of each most parsimonious tree obtained are added ($L_X + L_Y$), and this length is compared to the sum of the lengths ($L_P + L_Q$) of the most parsimonious trees obtained from two data sets, P and Q , of the same size as the original data sets and generated by random partitioning of the original couple of data sets. For W random partitions, S is the number of times when $(L_X + L_Y) < (L_P + L_Q)$, and the null hypothesis of congruence is rejected when the error rate « alpha » called P-value in PAUP*, $(1 - S/(W + 1))$, is below 5%. This indicates that there is more incongruence between the data sets than would be expected from chance alone. The ILD tests were performed with 1,000 replicates of heuristic search with TBR branch swapping, and simple sequence addition. The SH test uses the maximum of the standardized difference of log likelihoods (Diff. -lnL) as the test statistics, and the nonparametric bootstrap with reestimated log likelihoods (RELL) approximation (Kishino, Miyata et Hasegawa, 1990) is used for resampling the log likelihoods.

3.4.5 Coalescent approach to test the hybridization hypothesis

Phylogenetic incongruence between gene trees from independent loci (as nuclear and mitochondrial loci may be) may arise through the effect of hybridization or the effect of incomplete lineage sorting. To disentangle these two effects we used the coalescent approach as in Giessler and Englbrecht (2009). Coalescence simulations were applied to understand the causes of the incongruence observed between the *ND5* and *Rab4* gene trees. We used Mesquite 2.73 (Maddison et Maddison, 2010) to simulate *Rab4* gene trees based on neutral sequence variation. A null distribution of “deep coalescence” scores (number of extra gene lineages for the *Rab4* tree within species tree) was generated from simulated data sets (option: coalescence contained in current tree; mitochondrial *ND5* gene-tree topology as species tree) and the 95% confidence intervals of the null distribution were

calculated. Under the assumption of panmixis the observed «deep coalescence» score (which was derived from the Bayesian phylogenetic inference based on the *Rab4* gene) would be found within the confidence interval. If the observed value (DC) falls outside this interval the model for which incongruence may be solely explained by incomplete lineage sorting is rejected and existence of hybridization events is accepted. We did 1000 simulations with different effective population sizes $N_e=10^4$, 10^5 and 10^6 , assuming a generation of one year.

3.4.6 Microsatellite analyses

To assess the genetic relatedness among different clones of the *Daphnia pulex* complex four distance measures were used. First, alleles of each locus were transformed to produce a binary matrix (0 for absence and 1 for presence). Jaccard's (Jaccard, 1908) and Nei's (Nei, 1978) distances were calculated using Phyltools v.1.32 (Buntjer, 2001) to construct two distance matrices between all the genotypes. These distances were chosen in preference to other distance measures, as it does not use shared absence of an allele as a shared characteristic (Legendre et Legendre, 1998). We therefore judged them to be most appropriate as the present study includes highly polymorphic microsatellite data spanning two ploidy levels. Second, Tomiuk and Loeschke's index (TLI) of genetic similarity (Tomiuk et Loeschcke, 1991) was calculated from microsatellite data. This index is based on mean frequencies of ancestral alleles shared per locus observed and allows comparison of sexual and asexual taxa, including polyploids (Turgeon et Hebert, 1994). Band sharing measure (BSM) was also calculated from non-transformed microsatellite data (Tomiuk, Guldbbrandtsen et Loeschcke, 2009). Calculations of these two last indexes were carried out using the Popdist v.1.2.1 software (Tomiuk, Guldbbrandtsen et Loeschcke, 2009). The resulting similarity index (S) matrices were transformed by the following formula, $D=\sqrt{1-S}$, to produce distance matrices conforming to Euclidian space. Principal Coordinate analyses (PCO; Gower, 1966)

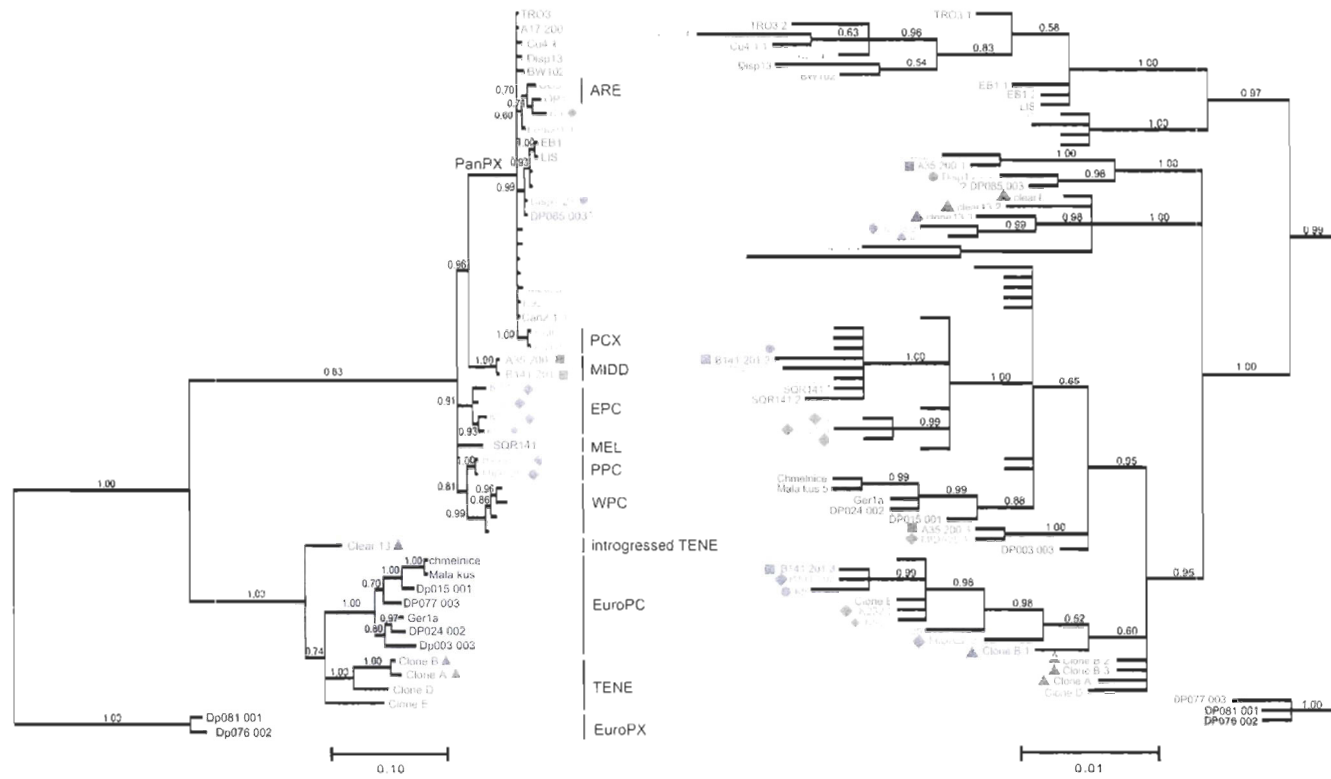
were conducted with the different distance matrices to represent affinities between genotypes using the module `pco` of the R software Version 4.0 (Ihaka et Gentleman, 1996 ; R Development Core Team, 2004). To choose the most probable number of clusters that our distance matrices represent, we used an iterative method, CascadeKM, implemented in the `vegan` package available with the R software. We tested how many clusters best represent each of our distance matrices. For each matrix 2 to 20 clusters were tested. Different criteria, “calinski” (Calinski et Harabasz, 1974) and “ssi” available in the module CascadeKM were used and compared to search the best partitions that represent our data. After choosing the number of clusters that best represent our data, K-means algorithm (Hartigan et Wong, 1979) was applied to the different distance matrices using `kmeans` module implemented in the `stats` package available with the R software. Clusters that were constructed with the K-means analysis were represented on each PCO analyses. Ordination analyses display the relative position of objects compared with each other but does not tell anything about possible evolutionary relations between these objects. As individual genomes are the products of two more or less distantly parental genomes, evolutionary relationships between multilocus genotypes are best represented by a network. We reconstructed a phylogenetic network using the *NeighborNet* algorithm implemented in the *Splitstree4* software version 4.10 (Huson et Bryant, 2006). For this analysis only binary matrix was used as an input.

3.5 Results

3.5.1 Diversity of *ND5* haplotypes

Few differences exist between phylogenetic inferences of the *ND5* gene in the *Daphnia pulex* complex based either on the HKY+G model (Figure 3-1 A) or without model (Figure 3-4 A). The tree constructed with the HKY+G model (Figure 3-1 A) seems to be the one that best explains our data with an average standard

deviation split frequencies of 0.003057 (*versus* 0.004984 for the tree constructed without model). This tree based on partial *ND5* gene (Figure 3-1 A) is similar to the *ND5* gene phylogenies of the *D. pulex* complex constructed previously with different clones (Colbourne *et al.*, 1998 ; Marková *et al.*, 2007). *ND5* haplotypes from two clones of *D. arenata* were included in our study. Three major haplotype groups (*pulicaria*, *tenebrosa* and European *D. pulex*) are revealed in this phylogeny of *D. pulex* complex *ND5* haplotypes (Figure 3-1 A). The average molecular divergence between *pulicaria* and European *D. pulex* clades, *tenebrosa* and European *D. pulex* clades and *pulicaria* and *tenebrosa* clades are 0.252 (± 0.043), 0.223 (± 0.038) and 0.190 (± 0.034) respectively (Table 3-2). The *tenebrosa* group consists of European *D. pulicaria* and *D. tenebrosa* haplotypes. *D. tenebrosa* included both polyploid (A, B and Clear_13) and diploid (E and D) clones whereas European *D. pulicaria* (EuroPC) included only diploid clones (Figure 1 A). The *pulicaria* group consisted of three *D. pulicaria* haplotype lineages (Eastern *D. pulicaria* EPC, Western *D. pulicaria* WPC and Polar *D. pulicaria* PPC, all of which are found in lakes), the *D. middendorffiana* lineage (MIDD), the Panarctic *D. pulex* lineage (PanPX), and the *D. melanica* (MEL) lineage (Figure 3-1 A). *ND5* haplotypes of *D. arenata* (ARE) clones are not differentiated from the PanPX haplotypes with an average molecular divergence of 0.007 (± 0.003) (see Table 3-2). Average molecular divergences within groups are of 0.005 (± 0.002) and 0.008 (± 0.004) for PanPX and ARE respectively. Two clones (Puli9 and Puli15) identified as *D. pulicaria* based on information at *LDH* are grouped with PanPX *ND5* haplotypes.



A) mitochondrial *ND5* gene bayesian phylogenetic inference

B) nuclear *Rab4* gene bayesian phylogenetic inference

Figure 3-1 Phylogenies of *ND5* and *Rab4* partial sequences from 51 isolates of the *Daphnia pulex* complex
 A) Bayesian phylogenetic inference of 51 *ND5* partial sequences constructed using the HKY+G model; B) Bayesian phylogenetic inference of 71 *Rab4* partial sequences constructed using the “three codon position” model (species identifications are indicated on the *ND5* phylogeny, see text ; hybrid diploid *ND5* haplotypes are intermingled with *D. pulex* haplotypes on the *ND5* phylogeny; hybrid diploid *Rab4* alleles are intermingled with *D. pulicaria* alleles on the *Rab4* phylogeny; Symbols indicate polyploids; ? indicates a clone with an unknown hybrid state)

Tableau 3-2 Mean divergences (lower-left triangle) and standard errors (upper-right triangle) using the Maximum Composite Likelihood model between groups of *ND5* haplotypes

	EuroPX	WPC	MIDD	PPC	PanPX	EPC	PCX	EuroPC	TENE	ARE	MEL	ARG
EuroPX	-	0.034	0.036	0.035	0.034	0.034	0.035	0.033	0.032	0.034	0.039	0.035
WPC	0.184	-	0.010	0.006	0.012	0.010	0.014	0.030	0.029	0.013	0.012	0.011
MIDD	0.189	0.029	-	0.011	0.013	0.011	0.014	0.031	0.030	0.014	0.013	0.010
PPC	0.184	0.012	0.029	-	0.012	0.010	0.014	0.030	0.029	0.013	0.012	0.012
PanPX	0.182	0.038	0.040	0.036	-	0.012	0.007	0.031	0.030	0.003	0.015	0.014
EPC	0.175	0.029	0.028	0.027	0.039	-	0.013	0.030	0.029	0.012	0.012	0.012
PCX	0.189	0.048	0.046	0.046	0.017	0.048	-	0.032	0.031	0.008	0.017	0.015
EuroPC	0.181	0.164	0.173	0.163	0.175	0.169	0.185	-	0.014	0.030	0.032	0.030
TENE	0.181	0.160	0.166	0.160	0.170	0.165	0.181	0.065	-	0.030	0.031	0.028
ARE	0.182	0.039	0.041	0.037	0.008	0.039	0.016	0.171	0.172	-	0.016	0.014
MEL	0.203	0.041	0.041	0.037	0.057	0.039	0.068	0.179	0.176	0.059	-	0.014
ARG	0.184	0.037	0.024	0.037	0.048	0.036	0.059	0.163	0.155	0.048	0.050	-

3.5.2 *Rab4* gene tree and comparisons with mitochondrial *ND5* gene tree

The two different consensus phylogenetic inference of the *Rab4* gene in the *Daphnia pulex* complex based either on the HKY+G model (Figure 3-4 B) or without model (Figure 3-1 B) are similar. The tree constructed using the partitioned dataset without model (Figure 3-1 B) seems to be the one that best explains our data with an average split frequencies of 0.002372 (*versus* 0.003190 for the tree constructed with the HKY+G model). The phi-test of recombination event is not statistically significant (P-value=0.054; Bruen *et al.* 2006). A clade of alleles from *D. pulex* and *D. arenata* may be distinguished (Figures 3-1 B and 3-4 B). Another clade includes alleles from various species (Figures 3-1 B and 3-4 B). This last clade may be divided into three principal subclades. Two subclades contain alleles from diploid “pure *D. pulex*”, the introgressed polyploid *D. tenebrosa* clone (Clear_13), “hybrid *D. pulex*” diploid (K92) and hybrid polyploid clones (Disp1-21, A35_301, K228 and K232). Another subclade contains principally alleles from *D. pulicaria* individuals (Figures 3-1 B and 3-4 B). This subclade contains also alleles from, *D. melanica*, *D. tenebrosa* and a clade grouping alleles from European *D. pulicaria*. Alleles from diploid hybrid *D. pulex* and hybrid polyploid clones are not encountered in the *pulex-arenata* clade. Three alleles from European *D. pulex* and European *D. pulicaria* form another clade. For each Bayesian phylogenetic inference the four *Rab4* alleles from *D. arenata* clones are branched together at the base of the *D. pulex* clade (Figures 3-1 B and 3-4 B).

Bayesian phylogenetic inferences from mitochondrial *ND5* partial sequences and single copy nuclear *Rab4* sequences are incongruent (see Figures 3-1 and 3-4). The ILD-test (P-value=0.001; Farris *et al.*, 1994), and the SH-test (Shimodaira et Hasegawa, 1999) are highly significant for the comparisons between both *ND5* phylogenetic trees and the *Rab4* phylogenetic tree constructed without model whereas the test is non-significant between *Rab4* phylogenetic trees (Table 3-3). Observed DC-values (DC=79) for each couple of bayesian phylogenetic tree inference fell outside the 95% confidence interval of the null distribution derived from 1,000 simulations, in all runs for different effective population sizes (confidence

intervals are 189.6-226.2 for $Ne=10^4$, 190.7-223.4 for $Ne=10^5$, and 190.8-223.1 for $Ne=10^6$). Thus incomplete lineage sorting alone can not explain the incongruence among trees and that hybridization event exists in the *Daphnia pulex* complex.

Tableau 3-3 SH-test results for the comparisons between *ND5* and *rab4* phylogenetic trees (see Figure 3-1 and 3-4)

	-ln L	Diff. -ln L	P-values
ND5 tree (without model)	1845.46948	559.08387	<0.0001
ND5 tree (HKY+G model)	1846.59241	560.20680	<0.0001
rab4 tree (HKY+G model)	1303.45967	17.07406	0.5123
rab4 tree (without model)	1286.38561	(Best)	-

3.5.3 Microsatellite analyses

Species identification and ploidy level of each individual genotype are shown on PCO graphs with different forms and colors. Species and hybrids are identified according to their mitochondrial haplotype, *LDH-A* genotype, ecological/morphological characteristics or both. Of the 105 multilocus genotypes analyzed by PCO, 29 of these are polyploid individuals (Table 3-1), six of these polyploids have *D. tenebrosa ND5* haplotypes, two have *D. middendorffiana s.s. ND5* haplotypes, 10 have *D. pulicaria ND5* haplotypes, two are *Daphnia* from Argentina and nine have *D. pulex ND5* haplotypes (Figures 3-2 and supplementary Figure 3-5). Principal Coordinate analyses (PCO) show similar patterns of genotype ordination regardless of the different distances used (Figures 3-2 and 3-5). The first two axes account for 25.08% to 30.28% of the total variability depending on the distance used. The PCO constructed with the Nei's distance matrix has axes with the highest percentages of explanation (Figure 3-2 B). For each distance matrix two different partitions of our data were found depending on the different criterion ("calinski" and "ssi") of the CascadeKM method (see Figures 3-2 and 3-2S). K-means method produces four and eight clusters with Jaccard and Nei distance matrices (Figures 3-2 A, B and 3-5 A, B respectively) and five and seven clusters with TLI and BSM distance matrices (Figures 3-2 C, D and 3-5 C, D respectively). For the "low number clusters", all the K-means analyses clustered *D. tenebrosa*, European *D. pulicaria* and *D. arenata* genotypes together (Figure 3-2 A, B, C, D). In each of these analyses, two clusters represent "pure *D. pulex*" genotypes and hybrid genotypes with *D. pulex* mitochondria. Polyploid *Daphnia* from Argentina and *D. melanica* genotypes are clustered with "pure *D. pulex*" genotypes. The "hybrid genotypes cluster" contains diploid and polyploid hybrids with *D. pulex* mitochondria and some genotypes with *D. pulicaria* mitochondria (Figure 3-2 A, B, C, D). For K-means based on Jaccard and Nei distance matrices a fourth cluster is composed of genotypes identified as *D. pulicaria* and polyploids with *D. pulicaria* and *D. pulex* mitochondria (Figure 3-2 A, B). For K-means based on TLI and BSM distance matrices, five clusters were identified; two clusters represent *D. pulicaria* and

polyploids with *D. pulicaria* and *D. pulex* mitochondria (Figure 3-2 C, D). Diploid genotypes of “pure *D. pulex*”, “hybrid *D. pulex*” and “*D. pulicaria*” clusters respectively have medium, high and relatively low mean heterozygosities (Table 4). Clusters produced by the K-means analyses did not possess diagnostic alleles.

Phylogenetic network based on the binary matrix produced by allele presence/absence (Figure 3-3) shows a slightly different but nonetheless informative pattern. The NeighborNet algorithm clusterizes *D. tenebrosa* clones and European *D. pulicaria* as separate clusters. In the PCO/K-means analyses, diploid and polyploid hybrids of the *pulex-pulicaria* group were somewhat grouped together (see Figures 3-2 and 3-5) but in the phylogenetic analysis they are dispersed in the network (Figure 3-3). Some hybrids branch with “pure *D. pulex*” and others with *D. pulicaria* individuals. Moreover two *D. pulicaria* (Puli21, winni1_2), two polyploids Polar *D. pulicaria* (MIDA20, B141_202) and the two *D. middendorffiana* clones (A35_301, B141_201) appear to branch between *D. pulex* genotypes and the group constituted with *D. arenata*, European *D. pulicaria* and *D. tenebrosa* genotypes (Figure 3-3). Phylogenetic network shows the evolutionary affinities between *D. arenata* and European *D. pulicaria* individuals (Figures 3-2, 3-5 and 3-3) as did the PCO/K-means analyses.

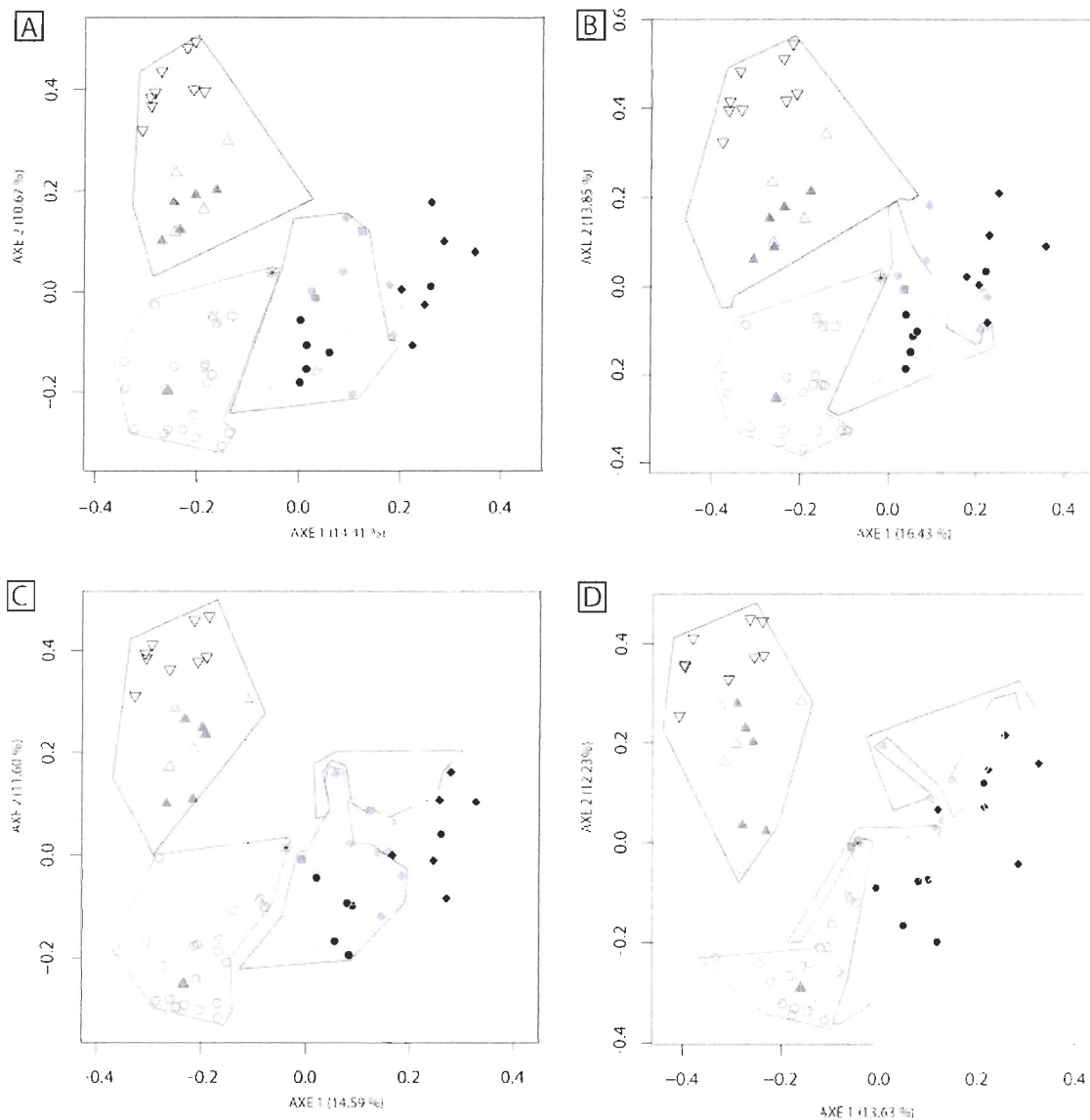


Figure 3-2 Principal Coordinate analyses and K-means analyses constructed using different distances on 105 microsatellite genotypes from the *Daphnia pulex* complex

A) Jaccard's distance; B) Nei's distance; C) Tomiuk and Loeschke's index (TLI); D) Band sharing measure (BSM) (Only the two first axes are represented; no European *D. pulex* isolate was included in this analysis; circles have *D. pulex* mitochondria; diamond have *D. pulicaria* mitochondria; square represent *D. middendorffiana* s.s.; circles with cross are *D. arenata*; triangle have *D. tenebrosa* mitochondria; inverted triangle are European *D. pulicaria*; asterisk represent *D. melanica* isolate and stars are argentinean polyploid clones; black filled symbols represent unknown hybrid state; empty symbols represent putative diploid and filled ones indicate polyploids; polygons represent K-means clusters)

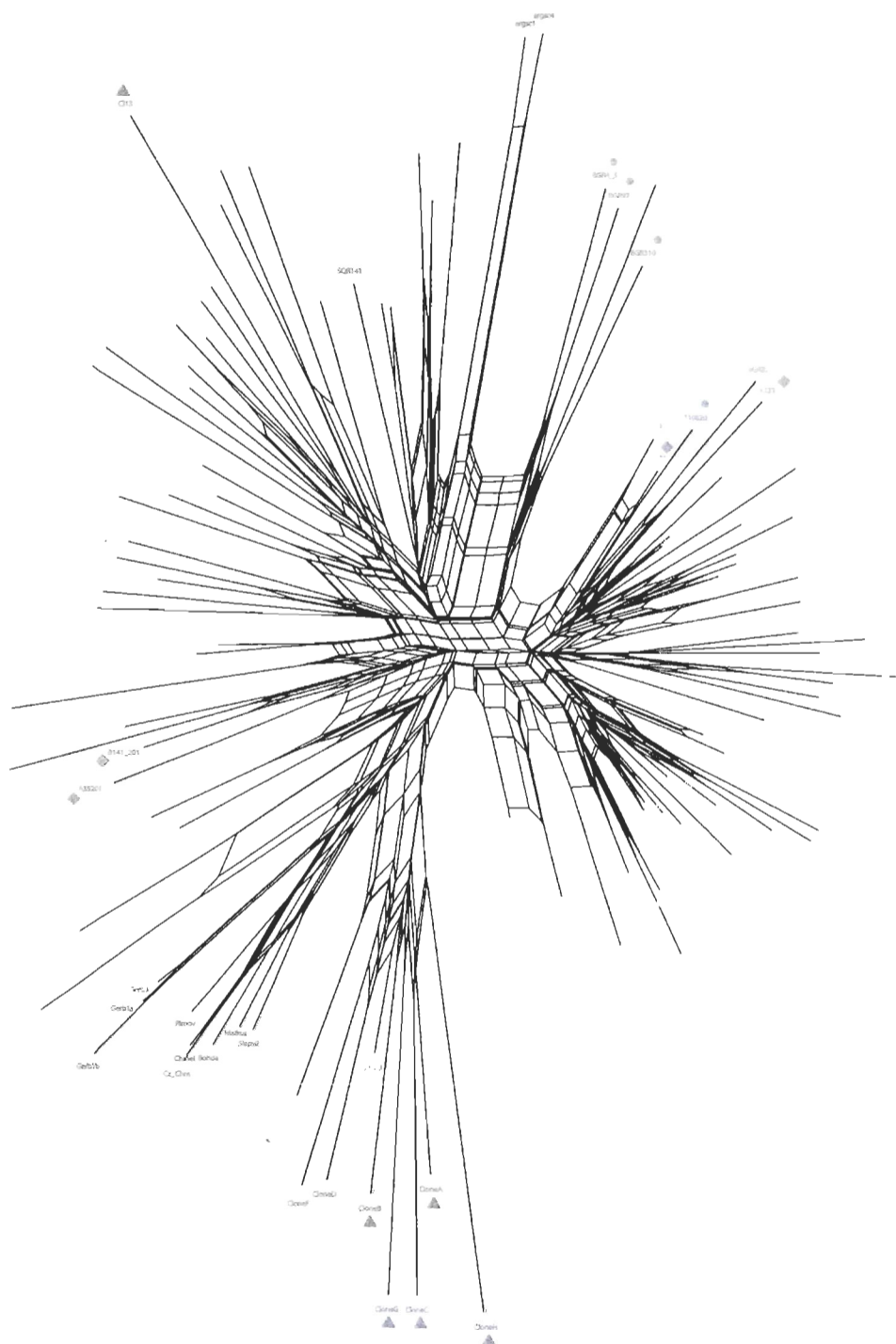


Figure 3-3 Phylogenetic networks constructed using the NeighborNet algorithm on 105 microsatellite genotypes from the *Daphnia pulex* complex (Symbols represent polyploid isolates)

Tableau 3-4 Mean heterozygosities (\pm standard deviation) of diploids clustered in “pure *D. pulex*”, “hybrid *D. pulex*” and “*D. pulicaria*” groups (see Figure 3-3).

	Nei	Jaccard	TLI	BSM
Pure <i>D. pulex</i>	0.62 (\pm 0.18)	0.61 (\pm 0.18)	0.62 (\pm 0.18)	0.62 (\pm 0.19)
Hybrid <i>D. pulex</i>	0.74 (\pm 0.14)	0.69 (\pm 0.20)	0.69 (\pm 0.20)	0.75 (\pm 0.13)
<i>D. pulicaria</i> 1	0.38 (\pm 0.15)	0.35 (\pm 0.14)	0.43 (\pm 0.20)	0.42 (\pm 0.19)
<i>D. pulicaria</i> 2	n-a	n-a	0.33 (\pm 0.17)	0.37 (\pm 0.18)

3.6 Discussion

Species complexes are of particular interest to understand processes underlying speciation as reproductive barriers are still incomplete. This is the first study that examines genetic relationships among species coming from a worldwide geographic range of the *Daphnia pulex* complex using hypervariable nuclear markers. Species of the *Daphnia pulex* complex are very young as they began to diverge less than one million years ago (Colbourne et Hebert, 1996). Furthermore, this complex is the only one in North America that has made transitions to obligate parthenogenesis and that shows high levels of hybridization and polyploidy (Colbourne *et al.*, 1998) thus challenging species delimitation.

3.6.1 Genetic relationships between *D. pulex* and *D. pulicaria*

One of the objectives of this paper was to assess the genetic relationships between *D. pulex*, *D. pulicaria* and their putative hybrids at the nuclear level. These two species are thought to have diverged less than 150,000 years ago (Omilian et Lynch, 2009). Based on research to date, it was not clear if these two young species were really distinct or if selection operating on the Ldh phenotypes was responsible for the fixation of the S allele in pond and the F allele in lake (Crease, Stanton et Hebert, 1989 ; Hebert *et al.*, 1989). The present study reveals that these two species and their hybrids form distinct clusters when using information at nine microsatellite loci. The “hybrid cluster” contains diploid and polyploid hybrids with *D. pulex* mitochondria and some genotypes with *D. pulicaria* mitochondria. In the 4 or 5 group K-means analysis *D. pulicaria* polyploids clustered with *D. pulicaria* diploids and *D. pulex* polyploids clustered with *D. pulex* diploid hybrids. A majority of these clones are heterozygous at *LDH* genotype suggesting that they are hybrids between *D. pulex* and *D. pulicaria* (Hebert *et al.* 1993; Innes *et al.* 1986). The phylogeny of the *Rab4* gene although not fully resolved, still showed clusters consisting mostly of

D. pulex alleles and others of *D. pulicaria* alleles but without hybrid cluster. None of the pure *D. pulex* screened in our study (Figures 3-1 B and 3-4 B) carry *D. pulicaria Rab4* allelic forms. The majority of diploid hybrids with *D. pulex ND5* haplotypes tested in this study are homozygotes for *D. pulicaria Rab4* alleles rather than heterozygotes for *D. pulex* and *D. pulicaria* alleles as we may expect in F1 hybrids. Hybridization followed by gene conversion may produce a homozygosity pattern for some genes in hybrid genomes (Hillis *et al.*, 1991 ; Kovarik *et al.*, 2004 ; Sang et Zhang, 1999). Genomic rearrangements (gene silencing, gene loss and gene conversion) are also known to occur in diploid and polyploid hybrids (Soltis *et al.*, 2008).

3.6.2 Polyploidy in the *pulicaria* clade

Polyploids with *D. pulex* and *D. pulicaria* mitochondria have multiple origins. As polyploids are mainly triploid hybrids (Vergilino, Belzile et Dufresne, 2009) their clustering position may reflect their genetic origins. A recent analysis has shown that pure diploid obligate asexual males may produce diploid sperm as asexual females (hybrid or not) may produce diploid egg (Lynch *et al.*, 2008). Thus different combinations of genomes are possible in triploids. Polyploid individuals with *D. pulicaria* or *D. pulex* mitochondria may contain two copies of *D. pulex* nuclear genomes with one *D. pulicaria* nuclear genome and *vice versa*. This genomic polymorphism may be responsible for adaptation to pond and lake lifestyle in clones of this species complex. In the PCO/K-means analyses, diploid and polyploid hybrids of the *pulex-pulicaria* group were somewhat grouped together but they were dispersed in the network analysis indicating probable multiple hybrid origins. Polyploid *Daphnia* from Argentina are clustered with “pure *D. pulex*” genotypes at microsatellite analyses and with *D. pulicaria* at mitochondrial genome (Adamowicz *et al.*, 2002) suggesting a hybrid origin in these clones.

3.6.3 Genetic relationships among species of the *tenebrosa* clade

In the PCO/K-means and network analyses on the microsatellites dataset, European *D. pulicaria* and *D. tenebrosa* clones are well differentiated from the pulicaria clade. In the 4 or 5 groups K-means analyses *D. tenebrosa* clusterize with European *D. pulicaria*. When K-means clustering analysis is conducted using 7 or 8 groups *D. tenebrosa* and European *D. pulicaria* represent two or three groups. In the *Rab4* gene phylogeny alleles from European *D. pulicaria* and *D. tenebrosa* do not form a single clade. They are mostly grouped with *D. pulicaria* alleles. There is no clear evidence of gene flow between *D. pulicaria* and European *D. pulicaria* in PCO/K-means and phylogenetic network analyses or from the literature (Marková *et al.*, 2007). One *Rab4* allele from European *D. pulicaria* (Dp076 003) branches with alleles from European *D. pulex* clones (Dp076 002 and Dp081 001; see Figures 3-1 B and 3-4 B) suggesting potential gene flow between these two taxa. Past evidence of introgression between the tenebrosa and the pulicaria clades comes from the analysis of a triploid clone (Clear_13) from Churchill, Manitoba that has a *ND5* haplotype closely related to *D. tenebrosa* and *LDH-A* alleles typical of *D. pulex* (Dufresne et Hebert, 1995). Both microsatellite analyses and *Rab4* phylogeny in this study showed that this clone has a nuclear genome closely related to *D. pulex* with no trace of *D. tenebrosa*. There are no clear traces of extensive hybridization events in *D. tenebrosa* polyploids thus suggesting that polyploidy may have arisen through union of unreduced gametes from obligate parthenogens and reduced gametes from cyclic parthenogens.

3.6.4 Genetic relationships between species from *D. pulex* complex

The mitochondrial *ND5* gene has previously been used for *D. pulex* phylogeny and typically resolves it well (Colbourne *et al.*, 1998 ; Vergilino, Belzile et Dufresne, 2009). The *tenebrosa* and *pulicaria* clades are relatively well separated

on PCO, K-means, and the network analyses using microsatellite markers but not as well with *Rab4*. Reasons for these discrepancies are discussed below.

The evolutionary relationships between *D. arenata* and the other species of the complex are particularly interesting. In our K-means analysis and phylogenetic network on microsatellite dataset, *D. arenata* appears to be more closely related to European *D. pulicaria*). In *Rab4* phylogeny, alleles from *D. arenata* branch with pure *D. pulex* alleles. The clones of *D. arenata* studied here have *D. pulex* mitochondria. Some authors have considered *D. arenata* an endemic species from Oregon due to its allozymic patterns (Crease *et al.*, 1997 ; Hebert, 1995). Other authors have considered it as a subpopulation of *D. pulex* (Li *et al.*, 2009 ; Omilian, Scofield et Lynch, 2008) due to its mitochondrial haplotype that is similar to *D. pulex*. Using six nuclear single copy genes including *Rab4* gene Omilian and Lynch (Omilian et Lynch, 2009) have shown that *D. arenata* may be a sister species of the group made of *D. pulex* and *D. pulicaria* species. This last study compared sequences solely from three species; *D. pulex*, *D. pulicaria* and *D. arenata* using *D. obtusa* sequences as an outgroup. Our study includes *Rab4* sequences from all species of the *pulex* complex and variation of nine microsatellite loci from several of these species and show that *D. arenata* may be of hybrid origin between *D. pulex* and another species such as European *D. pulicaria* or *D. tenebrosa*. In the analysis of Omilian *et al.* (2008), different alleles from *D. arenata* branch with *D. pulex* alleles, *D. pulicaria* alleles or are basal to all the *pulex* and *pulicaria* alleles (See Figure 4 in Omilian, Scofield et Lynch, 2008). These findings question the simple grouping of *D. arenata* with *D. pulex* populations.

Daphnia melanica isolate clustered with “pure *D. pulex*” genotypes at microsatellite analyses and with *D. pulicaria* at mitochondrial genome suggesting a hybrid origin in this species.

3.6.5 Phylogenetic incongruence and lineage sorting in a species complex

Phylogenetic trees based on the nuclear *Rab4* gene were not congruent with the tree constructed using the mitochondrial *ND5* partial gene sequences. This incongruence may be due to either incomplete lineage sorting or hybridization. Many methods using gene sequence polymorphisms have been proposed to differentiate patterns of hybridization and incomplete lineage sorting (Addison et Pogson, 2009 ; Buckley *et al.*, 2006 ; Holland *et al.*, 2008 ; Joly, McLenachan et Lockhart, 2009 ; Machado *et al.*, 2002 ; Sang et Zhong, 2000). These methods either involve numerous gene sequence datasets (Addison et Pogson, 2009 ; Buckley *et al.*, 2006 ; Holland *et al.*, 2008 ; Machado *et al.*, 2002 ; Sang et Zhong, 2000) or are dependent on a known phylogeny of the species (Joly, McLenachan et Lockhart, 2009). If speciation events are recent, hybridization and incomplete lineage sorting may be difficult to distinguish as nuclear genes may have not enough time to diverge. Nuclear gene polymorphism has a greater probability to be retained in young species than mitochondrial gene polymorphism. Due to their cytoplasmic inheritance and non-recombinant nature, mitochondrial genes coalesce four times faster than autosomal nuclear genes (Moore, 1995) but if mitochondrial genes are used solely to resolve the evolutionary history of species complex hybridization and introgression may go unnoticed (Moore, 1995). The simple coalescent method used in our study tests if incomplete lineage sorting could alone explain the incongruence among trees. It does not rule out any incomplete lineage sorting events but only tests the occurrence of hybridization. Some *Daphnia* species are known to have relatively high effective population sizes (Haag *et al.*, 2009 ; Omilian et Lynch, 2009) thus it is likely that either nuclear or mitochondrial polymorphism might have been retained in species of the *Daphnia pulex* complex. *Rab4* alleles from European *D. pulicaria* and *D. tenebrosa* group with North-American *D. pulicaria* but their microsatellite genotypes are relatively well separated from this last species. The branching of these alleles with those from North-American *D. pulicaria* are likely due to either stochastic processes such as incomplete lineage sorting or to shared selective

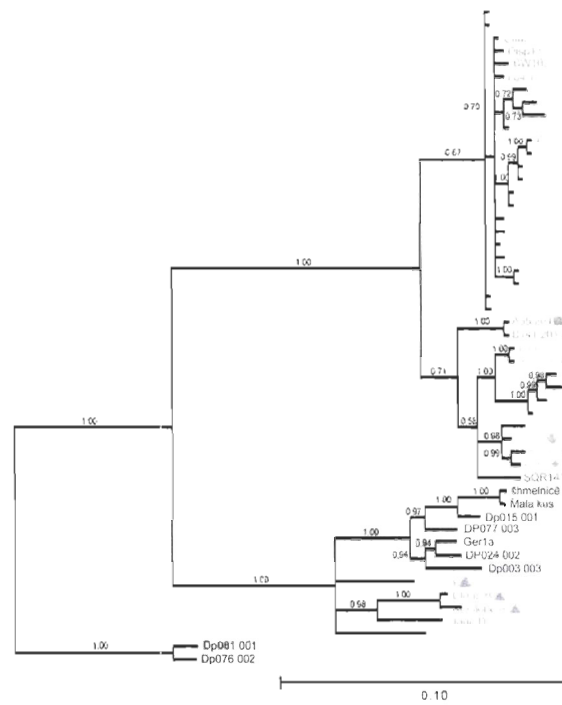
pressures and unlikely due to hybridization events. Further analyses with different nuclear genes are needed to solve this issue.

3.7 Conclusion

To understand the genetic bases of speciation in species complexes such as the *Daphnia pulex* complex, the scrutiny of multiple loci is crucial (Pamilo et Nei, 1988 ; Wu, 1991). Microsatellite loci, as they exhibit high levels of polymorphism and are dispersed throughout the genome, offered better resolution than the single locus nuclear gene, *Rab4*. The *Daphnia pulex* complex is well delineated at the mitochondrial level but has porous boundaries at the nuclear level. Hybridization and introgression have generated numerous clones with different ploidy levels and mitochondrial/nuclear combinations. It is noteworthy that the two other species complexes in *Daphnia* (*D. longispina* and *D. carinata*) are also characterized by hybridization and introgression. Additional markers provided by the newly available genome sequence will help shed lights into forces behind these mini-radiations.

3.8 Acknowledgments

We thank C. Marinone for providing clones from Argentina, M. Cristescu and C. Caceres for providing some *Daphnia pulicaria* S. Markovà for providing some European *Daphnia* sp. and D. Innes for providing *D. pulex* clones from Ontario. We thank three anonymous reviewers for their useful comments and corrections that improve the quality of this paper. This work was supported by a Natural Sciences and Engineering research Council (NSERC) research grants and by a Canadian Foundation for Innovation (CFI) equipment grants to France Dufresne.



A) mitochondrial *ND5* gene bayesian phylogenetic inference



B) nuclear *Rab4* gene bayesian phylogenetic inference

Figure 3-4 Phylogenies of ND5 and Rab4 partial sequences from 51 isolates of the *Daphnia pulex* complex

A) Bayesian phylogenetic inference of 51 ND5 partial sequences constructed using the the “three codon position” model; B) Bayesian phylogenetic inference of 71 Rab4 partial sequences constructed using HKY+G model (species identifications are indicated on the ND5 phylogeny, see text; hybrid diploid ND5 haplotypes are intermingled with *D. pulex* haplotypes on the *ND5* phylogeny; hybrid diploid Rab4 alleles are intermingled with *D. pulicaria* alleles on the *Rab4* phylogeny; Symbols indicate polyploids; ? indicates a clone with an unknown hybrid state)

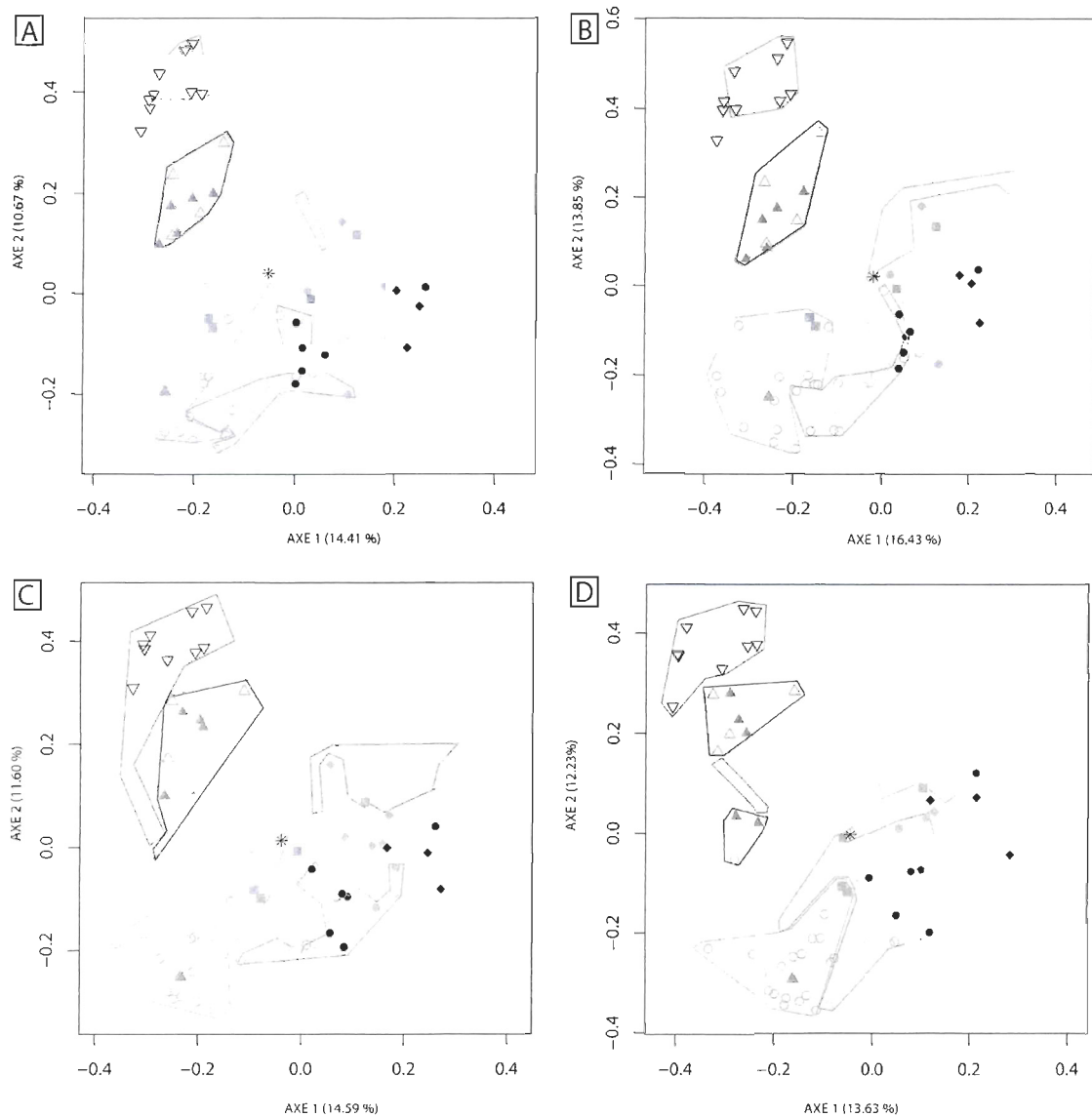


Figure 3-5 Principal Coordinate analyses and K-means analyses constructed using different distances on 105 microsatellite genotypes from the *Daphnia pulex* complex

A) Jaccard's distance; B) Nei's distance; C) Tomiuk and Loeschke's index (TLI); D) Band sharing measure (BSM) (Only the two first axes are represented; no European *D. pulex* isolate was included in this analysis; circles have *D. pulex* mitochondria; diamond have *D. pulicaria* mitochondria; square represent *D. middendorffiana* s.s.; circles with cross are *D. arenata*; triangle have *D. tenebrosa* mitochondria; inverted triangle are European *D. pulicaria*; asterisk represent *D. melanica* isolate and stars are argentinean polyploid clones; black filled symbols represent unknown hybrid state; empty symbols represent putative diploid and filled ones indicate polyploids; polygons represent K-means clusters)

Chapitre IV

EVOLUTION OF A TRANSPOSON IN DAPHNIA HYBRID GENOMES

Roland Vergilino^{1*}, Philippe Desjardins-Proulx¹, Teresa J. Crease², France Dufresne¹

¹ *Département de Biologie, Université du Québec à Rimouski, Québec G5L 3A1, Canada*

² *Department of Integrative Biology, University of Guelph, Ontario, N1G 2W1, Canada*

(Roland Vergilino : conception du projet, échantillonnage, manipulations, traitement statistique, rédaction de l'article; Philippe Desjardins-Proulx: traitement statistique et rédaction; Teresa Crease : conception de l'expérience, soutien technique, rédaction; France Dufresne : conception du projet, échantillonnage, rédaction, financement)

Article en préparation, les références citées dans l'article sont reportées à la fin de la thèse.

4.1 Résumé

Les éléments transposables (TEs) jouent un rôle majeur dans l'évolution des génomes. Leur capacité à se déplacer et/ou se multiplier dans le génome de leur hôte peut avoir impact important sur les phénotypes. Les TEs peuvent avoir des conséquences significatives sur la structure des génomes. La dynamique et la distribution des TEs sont influencées par leur mode de transposition, la disponibilité de niches dans les génomes de leur hôte, et la dynamique des populations de l'hôte. Les théories prédisent que l'augmentation du niveau de ploïdie devrait être suivie par une augmentation du nombre d'insertions de TEs. L'évolution des TEs chez les hybrides polyploïdes a été principalement étudié chez les plantes; peu d'études se sont penchés sur leur incidence dans les génomes d'animaux hybrides et polyploïdes. Le complexe *Daphnia pulex* est un groupe d'espèces où la transition vers l'asexualité a évolué de différentes façons. Des lignées hybrides et polyploïdes ont été produites à plusieurs reprises dans le complexe *D. pulex* et celles-ci se reproduiraient de façon asexuée. *Pokey*, un TE de classe II, a été étudié chez les daphnies et en particulier dans les populations diploïdes non-hybrides de *Daphnia pulex*. Notre étude examine les effets de la polyploïdie sur l'évolution de *Pokey*. Des portions de l'élément *PokeyA* intégrés dans différents génomes hybrides et polyploïdes ont été séquencées et comparées à des données récoltées par des études antérieures. Le nombre d'insertions TE a été comparé entre des diploïdes (14) et des polyploïdes (13) hybrides utilisant une technique de «TE display». Bien que le nombre moyen de sites d'insertion était plus élevé chez les polyploïdes que chez les diploïdes cette différence n'était pas statistiquement significative. De nombreux événements de recombinaisons ont été révélés dans des allèles de *PokeyA* dans des lignées diploïdes et polyploïdes du complexe *pulex*. Des analyses phylogénétiques et de recombinaison ont montré que la recombinaison est une force majeure qui façonne l'évolution du transposon *Pokey*.

4.2 Abstract

Transposable elements (TEs) play a major role in genome evolution. Their capacity to move and/or multiply in the genome of their host may have profound impacts on phenotypes, and may have dramatic consequences on genome structure. The population dynamics and the distributions of TEs are influenced by their mode of transposition, the availability of niches in host genomes, and host population dynamics. Theories predict that polyploidization should be followed by an increase of TE insertions. Evolution of TEs in hybrids and polyploids has been mostly studied in plants; few studies have studied their impacts in hybrid and polyploid animals. Hybrids and polyploids have arisen multiple times in the *Daphnia pulex* complex and are thought to reproduce by obligate parthenogenesis. *Pokey*, a class II TE, has been studied in diploid populations of *D. pulex*. Our study examines the effects of polyploidy on *Pokey* evolution. Portions of *Pokey* elements from various *Daphnia* hybrids and polyploids were sequenced and compared to sequences amplified in a previous study to understand the evolutionary history of *Pokey*. The number of TE insertions was compared between diploid (14) and polyploid (13) hybrids using the “TE display” technique. Although the average number of occupied insertion sites is higher in polyploids than in diploids, this difference is not statistically significant. Evidence of recombination is found in numerous *Pokey* alleles from diploid and polyploid hybrids. Phylogenetic and recombination analyses show that recombination is a major force that shapes *Pokey* evolution and may be useful to understand the evolutionary history of the *D. pulex* complex.

4.3 Introduction

Transposable elements (TEs) are widespread in the living world. Rare are the organisms which lack these genetic elements, which have the ability to move and multiply throughout the genome (Feschotte et Pritham, 2007 ; but see Shigenobu *et al.*, 2000). There are large variations in the proportion of mobile genetic elements and especially transposable elements across genomes (Pidpala, Yatsishina et Lukash, 2008). The ability to move and multiply throughout the genome is dependent of the mode of transposition. Autonomous transposable elements that transpose via an RNA intermediate (Class I TEs) have a copy-and-paste mechanism and are able to multiply throughout the genome. Class II TEs, which transpose *via* a DNA intermediate, use a cut-and-paste mechanism and are less prone to multiply (Böhne *et al.*, 2008). Class II TEs may multiply via the DNA repair mechanism of double strand breaks and copying of homologous sequences after excision (Engels *et al.*, 1990). Non-autonomous TEs such as MITEs or SINEs use the machinery of autonomous TEs elsewhere in the genome to transpose (Feschotte et Pritham, 2007). Reduced availability of genomic niche or host-mediated regulation such as epigenetic factors (e.g. piwiRNA, methylation) may limit transposition (Yoder, Walsh et Bestor, 1997).

Host population dynamics may have a dramatic impact on TE populations (Charlesworth et Charlesworth, 1983 ; Charlesworth et Charlesworth, 1995 ; Charlesworth et Wright, 2001). A decrease of effective host population size (Brookfield et Badge, 1997), or an increase in the level of selfing in host populations (Wright et Schoen, 1999) is expected to lead to an increase in the number of TE insertions. The mode of reproduction also has a substantial impact on TE dynamics and their density in the genome (Schaack *et al.*, 2010a ; Schaack *et al.*, 2010b ; Wright et Finnegan, 2001 ; Zeyl, Bell et Green, 1996). TEs may spread via recombination and out-crossing in sexual populations (Wright et Finnegan, 2001) whereas the spread of TEs is prevented in asexual populations in the absence of horizontal transmission (Arkhipova et Meselson, 2000 ; Gladyshev, Meselson et

Arkhipova, 2008). Empirical studies have tested and are relatively well in accordance with these theoretical assumptions (Arkhipova et Meselson, 2000 ; Gladyshev, Meselson et Arkhipova, 2008 ; Schaack *et al.*, 2010b ; Valizadeh et Crease, 2008 ; Zeyl, Bell et Green, 1996).

More animals than previously thought are known to have evolved via hybridization or polyploidization (Mable, 2004 ; Otto, 2007 ; Seehausen, 2004). These processes may have various impacts on the evolution of TEs. Activations of TEs have been highlighted in hybrid genomes (Fontdevila, 2005 ; O'Neill, O'Neill et Graves, 1998 ; Petrov *et al.*, 1995 ; Shan *et al.*, 2005) and polyploidization may lead to an increase of TE insertions (Comai, 2000 ; Matzke et Matzke, 1998). Bursts of TE activity are thought to have a substantial impact on genome rearrangement (McClintock, 1984) and may lead to phenotypic diversification in hybrids and polyploids (Comai, 2000 ; Madlung *et al.*, 2002). Studying the dynamics of TEs in hybrids and polyploids may allow us to understand the evolution of the genomes and their propensity to adapt to various environments. Three hypotheses have been advanced to explain the burst of TE activity and/or the increase of TE insertions into hybrid/polyploid genomes. The genomic shock hypothesis states that hybridization or polyploidization may represent shocks that activate TEs (McClintock, 1984). This may induce an increase in the number of active TE insertions (Parisod *et al.*, 2010). The redundancy hypothesis considers that the increase in genomic niches available may relax purifying selection acting against TEs (Comai, 2000 ; Matzke et Matzke, 1998) so they are able to multiply in polyploids. The bottleneck hypothesis is population-based hypothesis and states that when a hybrid population is created, effective population size is reduced, which reduces the effect of purifying selection after which TE insertion numbers may increase (Lynch, 2007). Many studies have explored the effects of hybridization and polyploidization on TE dynamics in plants (Ainouche *et al.*, 2009 ; de Araujo *et al.*, 2005 ; Hanson *et al.*, 2000 ; Kashkush, Feldman et Levy, 2002 ; Madlung *et al.*, 2002 ; Parisod *et al.*, 2010 ; Shan *et al.*, 2005), but few studies have studied these effects in animals (Labrador et Fontdevila, 1994 ; but see Liu *et al.*, 2009 ; O'Neill, O'Neill et Graves, 1998).

The *Daphnia pulex* complex has been intensively studied due to its dominance in freshwater habitats in North America, and its variation in reproductive mode and ploidy level. *Daphnia* usually reproduce by cyclic parthenogenesis, which is clonal reproduction interrupted by bouts of sexual reproduction. However, some lineages reproduce by obligate parthenogenesis (Hebert, 1981 ; Hebert et Crease, 1983 ; Hebert et Finston, 2001b ; Innes, Fox et Winsor, 2000 ; Innes, Schwartz et Hebert, 1986). The *D. pulex* complex includes seven species that have been distinguished on the basis of morphological, ecological, and genetic data (Colbourne *et al.*, 1998 ; Hebert, 1995 ; Marková *et al.*, 2007). In North-America, two species of the complex, *D. pulex* and *D. pulicaria*, are dominant in freshwater habitats, and are morphologically similar but ecologically distinct (Hebert, 1995), although they hybridize in nature (Crease, Stanton et Hebert, 1989 ; Hebert et Finston, 2001a ; Hebert *et al.*, 1993a ; Hebert *et al.*, 1989). Diploid hybrids always have *D. pulex* mitochondrial genomes and have been found to reproduce by obligate parthenogenesis in nature based on patterns of allozyme variation (Hebert et Finston, 2001a ; Innes et Hebert, 1988) and laboratory crosses (but see Heier et Dudycha, 2009 ; Innes et Hebert, 1988). Analyses of allozyme variation indicate that introgression is rare or nonexistent in areas where the two species co-occur (Hebert et Finston, 2001a ; Hebert *et al.*, 1993b ; Innes, Schwartz et Hebert, 1986). The *D. pulex* complex shows a geographical polyploidy pattern (Dufresne et Hebert, 1995, 1997 ; Hobaek, Weider et Wolf, 1993 ; Ward *et al.*, 1994 ; Weider *et al.*, 1999a ; Weider *et al.*, 1999b), with polyploid populations found at high latitudes and altitudes and diploid populations (hybrid or not) being prevalent in temperate regions (Aguilera *et al.*, 2007 ; Beaton et Hebert, 1988 ; Dufresne et Hebert, 1995).

Polyploidy has occurred repeatedly in the *D. pulex* complex (Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994, 1995, 1997). One polyphyletic group of polyploids (collectively known as *D. middendorffiana* which we termed *D. middendorffiana sensu lato* in this study) is thought to have arisen from hybridization between *D. pulex* males and *D. pulicaria* females, or females of another species which no longer exists as a cyclic parthenogen (Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994,

1997). A recent study using microsatellite data, flow cytometry and mitochondrial sequences has shown that most polyploids of the *D. pulex* complex reproducing by obligate parthenogenesis are triploids (Vergilino, Belzile et Dufresne, 2009). Triploids that are thought to have arisen from crosses between *D. pulex* females and *D. pulicaria* males are encountered in the Northeast of Quebec and in Ontario (Vergilino, Belzile et Dufresne, 2009 ; Vergilino *et al.*, 2011). Another group of polyploids has been found in *D. tenebrosa*, a circumarctic species (Dufresne et Hebert, 1995) and includes both triploid and tetraploid clones (Vergilino, Belzile et Dufresne, 2009). The *D. pulex* genome has recently been sequenced (Colbourne *et al.*, 2011) and numerous Class II DNA transposons have been identified in it (Schaack *et al.*, 2010a). Previous studies have reported that Class II TE insertion numbers are lower in the genomes of obligately parthenogenetic *D. pulex* than in genomes of lineages that reproduce sexually (Schaack *et al.*, 2010b ; Sullender et Crease, 2001b ; Valizadeh et Crease, 2008) as theoretically predicted if sexual reproduction helps TEs to spread (Wright et Finnegan, 2001). *Pokey*, a class II DNA transposon of the *piggyBac* family, has been extensively studied in diploid populations. It inserts in tandemly repeated rRNA genes (Penton, Sullender et Crease, 2002) and is also encountered in other parts of the genome (Valizadeh et Crease, 2008). Whereas class II DNA transposons are more prone to transfer horizontally between lineages than retrotransposons, horizontal transfer doesn't seem to play a major role in the evolution of *Pokey* as 28S rRNA genes and mitochondrial phylogenies of species in the subgenus *Daphnia* are congruent with the phylogeny of *Pokey* elements (Penton et Crease, 2004). Based on patterns of polymorphism observed among natural populations, previous studies (Sullender et Crease, 2001a ; Valizadeh et Crease, 2008) have suggested that *Pokey* may be active in sexual populations of *D. pulex* but not in obligate parthenogens.

The *D. pulex* complex and *Pokey* represent an interesting model to understand the effect of hybridization and polyploidization on the evolution and dynamics of a class II DNA transposon *in natura*. This study aimed to, 1) determine the effect of ploidy levels on the number of *Pokey* insertion sites and, 2) examine the

impact of hybridization on the evolution of *Pokey* elements inserted in the 28S rRNA gene.

4.4 Material and methods

4.4.1 Daphnia samples

We analyzed 53 isolates representing 6 of the 7 recognized species of the *D. pulex* complex. Only *D. melanica* was not included. Nine *Pokey* sequences (PC2-FF-5_1 [AY630577.1] and PC2-FF-5_1 [AY630578.1] from *D. pulicaria*; AR2-UN-1_1 [AY630585.1] from *D. arenata*; a consensus sequence PX2-SS-CON [AY630581.1] from PX2-SS-5 and PX2-SS-6, PX2-SS-7_1 [AY630579.1] and PX2-SS-8_1 [AY630580.1] from North-American *D. pulex*; EPC2-UN-3_1 [AY630582.1] from European *D. pulicaria*; EPX2-UN-1_1 [AY630583.1] and EPX2-UN-2_1 [AY630584.1] from European *D. pulex*) were obtained from nine isolates studied by Penton and Crease (2004; see Table 4-1). We established parthenogenetic lines (clones) in the laboratory from 45 individual females sampled from nature between 2004 and 2010 (Tableau 4-1), and cultured them using standard techniques (Hebert et Crease, 1980). Genomic DNA was extracted from 10 to 30 individuals weighing approximately 100 mg (wet weight) from each of these clones using the DNeasy Tissue kit (QIAGEN Inc., Mississauga, ON, Canada) according to the supplier protocol. Species identity of these clones was assessed by combining information on morphology, haplotype of the mitochondrial ND5 gene, genotypes of the nuclear *Lactate dehydrogenase* (*Ldh*) gene (described in Vergilino *et al.*, 2011), and ecology (whether clones came from ponds or lakes). Ploidy levels have been assessed previously using microsatellite genotypes and flow cytometry (Vergilino, Belzile et Dufresne, 2009).

Tableau 4-1 Identification, ploidy level, and geographic origins of *Daphnia* isolates used in the present study

Labels of the isolates are composites of their characteristics. The first two letters represent the mitochondrial haplotypes followed by the ploidy level (2 or 3). SS, SF, FF or UN represent *Ldh* genotypes and allow differentiating between *D. pulex*, *pulex-pulicaria* hybrid, *D. pulicaria* or unknown nuclear genotypes respectively. The last number refers to the number of the isolate

Isolates	Mitochondrial haplotype	Ploidy level	Geographic location
PX2-SS-1	North-American <i>D. pulex</i>	2x	Churchill, Manitoba, CAN
PX2-SS-2	North-American <i>D. pulex</i>	2x	Churchill, Manitoba, CAN
PX2-SS-3	North-American <i>D. pulex</i>	2x	Windsor, Ontario, CAN
PX2-SS-4	North-American <i>D. pulex</i>	2x	Disputed Road, Ontario, CAN
PX2-SF-1	North-American <i>D. pulex</i>	2x	Churchill, Manitoba, CAN
PX2-SF-2	North-American <i>D. pulex</i>	2x	Windsor, Ontario, CAN
PX2-SF-3	North-American <i>D. pulex</i>	2x	Windsor, Ontario, CAN
PX2-SF-4	North-American <i>D. pulex</i>	2x	Windsor, Ontario, CAN
PX2-SF-5	North-American <i>D. pulex</i>	2x	Ste-Foy, Québec, CAN
PX2-SF-6	North-American <i>D. pulex</i>	2x	Ste-Foy, Québec, CAN
PX2-SF-7	North-American <i>D. pulex</i>	2x	Michigan, U.S.A.
PX2-SF-8	North-American <i>D. pulex</i>	2x	Kuujjarapik, Québec, CAN
PX2-SF-9	North-American <i>D. pulex</i>	2x	Métis, Québec, CAN
PX2-SF-10	North-American <i>D. pulex</i>	2x	Kuujjarapik, Québec, CAN
PX2-SF-11	North-American <i>D. pulex</i>	2x	Kuujjarapik, Québec, CAN
PX2-SF-12	North-American <i>D. pulex</i>	2x	Métis, Québec, CAN
PX3-SF-1	North-American <i>D. pulex</i>	3x	Kuujjarapik, Québec, CAN
PX3-SF-2	North-American <i>D. pulex</i>	3x	Kuujjarapik, Québec, CAN
PC3-SF-1	Eastern <i>D. pulicaria</i>	3x	Kuujjarapik, Québec, CAN
PC3-SF-2	Eastern <i>D. pulicaria</i>	3x	Kuujjarapik, Québec, CAN
PC3-SF-3	Eastern <i>D. pulicaria</i>	3x	Kuujjarapik, Québec, CAN
PX2-FF-1	North-American <i>D. pulex</i>	2x	Vermillion County, Illinois, USA
PX2-FF-2	North-American <i>D. pulex</i>	2x	Vermillion County, Illinois, USA
PX2-FF-3	North-American <i>D. pulex</i>	2x	Vermillion County, Illinois, USA
PX2-FF-4	North-American <i>D. pulex</i>	2x	Vermillion County, Illinois, USA

Isolates	Mitochondrial haplotype	Ploidy level	Geographic location
PC3-SF-4	Polar <i>D. pulicaria</i>	3x	Churchill, Manitoba, CAN
PC3-SF-5	Polar <i>D. pulicaria</i>	3x	Kuujjarapik, Québec, CAN
PC2-FF-1	Western <i>D. pulicaria</i>	2x	Indiana, U.S.A.
PC2-FF-2	Western <i>D. pulicaria</i>	2x	Indiana, U.S.A.
PC2-FF-3	Western <i>D. pulicaria</i>	2x	Winnipeg, Manitoba, CAN
PC2-FF-4	Western <i>D. pulicaria</i>	2x	Kuujjarapik, Québec, CAN
PC3-SF-6	Western <i>D. pulicaria</i>	3x	Churchill, Manitoba, CAN
PC3-SF-7	Western <i>D. pulicaria</i>	3x	Kuujjarapik, Québec, CAN
MI3-SF-1	<i>D. Middendorffiana s.s.</i>	3x	Churchill, Manitoba, CAN
MI3-SF-2	<i>D. Middendorffiana s.s.</i>	3x	Churchill, Manitoba, CAN
TE2-UN-1	<i>D. tenebrosa</i>	2x	Churchill, Manitoba, CAN
TE2-UN-2	<i>D. tenebrosa</i>	2x	Churchill, Manitoba, CAN
TE2-UN-3	<i>D. tenebrosa</i>	2x	Churchill, Manitoba, CAN
TE3-UN-1	<i>D. tenebrosa</i>	3x	Churchill, Manitoba, CAN
TE3-UN-2	<i>D. tenebrosa</i>	3x	Churchill, Manitoba, CAN
TE3-UN-3	<i>D. tenebrosa</i>	3x/4x	Churchill, Manitoba, CAN
TE3-SS-1	<i>D. tenebrosa</i>	3x	Churchill, Manitoba, CAN
EPC2-UN-1	European <i>D. pulicaria</i>	2x	Lowland, Czech Republic
EPC2-UN-2	European <i>D. pulicaria</i>	2x	Catalonia, Spain
EPC2-UN-3	European <i>D. pulicaria</i>	2x	Grosser Binnense, Germany
EPX2-UN-1	European <i>D. pulex</i>	2x	Grebin, Germany
EPC2-UN-2	European <i>D. pulex</i>	2x	Malente, Germany
PC2-FF-5	<i>D. pulicaria</i>	2x	Humbolt Lake, Saskatchewan, CAN
PX2-SS-5	North-American <i>D. pulex</i>	2x	Vermillion County, Illinois, U.S.A.
PX2-SS-6	North-American <i>D. pulex</i>	2x	Warren County, Indiana, U.S.A.
PX2-SS-7	North-American <i>D. pulex</i>	2x	Essex County, Ontario, CAN
PX2-SS-8	North-American <i>D. pulex</i>	2x	La Tuque, Quebec, CAN
AR2-UN-1	North-American <i>D. pulex</i>	2x	Florence, Oregon, U.S.A.

4.4.2 Transposon display

We used a PCR-based approach called transposon (TE) display (Wright *et al.*, 2001) to test the effect of ploidy level on the abundance and distribution of *Pokey* insertion sites in the genomes of 14 diploid and 13 polyploid isolates from the 53 isolates analyzed in this paper. These 27 isolates were chosen to account for the diversity of diploid and polyploid clones, which have different mitochondrial haplotypes and geographical locations (Tableau 4-1). We followed the TE display protocol of Valizadeh and Crease (2008) but the annealing temperature for each PCR reaction was set at 50 °C to allow amplification of TE inserts in *Daphnia* species with genomes that are divergent from *D. pulex*. To determine if *Pokey* insertion site patterns differ according to the ploidy level of the isolates, we constructed a phylogenetic network from the binary *Pokey* presence/absence profiles using the NeighborNet algorithm implemented in the program, Splitstree4.10 (Huson et Bryant, 2006).

To test the hypothesis that the number of TE insertions increases with ploidy level, we compared the number of TE insertion sites to the heterozygosity of the isolates (diploid and polyploid hybrids) weighted by their ploidy level. Each isolate was genotyped at nine microsatellite loci as in Vergilino *et al.* (2011). Theoretically, triploids may arise from independent hybridization events, and the number of TE insertions in polyploids with different parental genomes may be greater than the number of insertions in polyploids with similar genomes (due to the homozygosity of some TE inserts). If we don't account for the differences in the genomes that constitute polyploids, we may overestimate the effect of ploidy level. TE insertion sites may be more correlated to heterozygosity level than to ploidy level. Thus, we introduce a heterozygosity index (H_{pi}), which takes into account the different ploidy levels of each genotype such that:

$$H_{pi} = (n_3 + 0.5 * n_2) / n_{locus}$$

where n_{locus} is the total number of microsatellite loci analyzed (9), n_3 is the number of

loci with 3 different alleles and n_2 is the number of loci with only 2 different alleles. Genotypes from each diploid isolate that are homozygous for each microsatellite locus have a H_{pi} of 0, diploid isolates that are heterozygous for each locus and triploid isolates with two different alleles at each locus have a H_{pi} of 0.5, and triploid isolates that have three different alleles at each locus have a H_{pi} of 1. Triploids with low heterozygosity values (under 0.5) and/or few numbers of loci with three alleles can thus be compared to diploid hybrids. To disentangle the effect of the addition of different genomes from the effect of increased ploidy level *per se*, we performed ANCOVA (Analysis of Covariance) using R v4.0 software (R Development Core Team, 2004). The “number of TE inserts” is a dependent variable; the ploidy level and H_{pi} are the independent variables. To compare the number of singletons (i.e. transposon insertion sites encountered in only one isolate) between diploid and polyploid clones, a Fisher exact test was performed on a 2x2 contingency table similarly to the approach of Wright *et al.* (2001).

4.4.3 Cloning and sequencing *Pokey* alleles

To examine the impact of hybridization on *Pokey*, *Pokey* alleles inserted in the 28S rRNA gene from various isolates of the *D. pulex* complex were cloned and sequenced. An approximately 1600-bp fragment of *Pokey* alleles in the 28S rRNA gene from 16 isolates (5 diploid hybrids, 5 polyploid hybrids, 3 diploids and 3 polyploids with unknown hybrid state and 1 diploid non-hybrids) of the *D. pulex* complex (Tableau 4-1) were amplified using the internal *Pokey* primer, Pok5026F (5'-TCGAACCTGCAGCCGGACGAATTTGCAG), and the primer 28SR (5'-TCCATTCGTGCGCGTCACTAATTAGATGAC), which is located 46 bp downstream of the *Pokey* TTAA target site. PCR reactions were performed using the Phusion™ high-fidelity PCR kit (Finnzymes, Woburn, MA) according to the supplier protocol. The amplification reactions were performed in an MJ PTC-100 thermal cycler (MJ Research Inc.). The thermocycling profile consisted of 1 cycle of 1 min at 94°C, 35

cycles of 30 sec at 94°C, 30 sec at 55°C and 2 min at 72°C, with a final incubation of 5 min at 72°C. PCR amplicons were cloned using the StrataClone™ Blunt PCR Cloning Kit according to the supplier protocol. For each isolate, 10 to 12 plasmids were purified using the E.Z.N.A. Plasmid Mini Kit II (Omega Bio-tek Inc., Norcross, GA). These *Pokey* alleles were compared to *Pokey* sequences from a previous study which focused on non-hybrid isolates (Penton et Crease, 2004) to ensure that *Pokey* elements analyzed in this study belong to the *Pokey* lineage. Partial *Pokey* sequences were manually aligned using the ClustalW module of BioEdit v.7.0.5 available at <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

4.4.4 Recombination analyses

Potential recombination events among *Pokey* alleles were explored in a dataset of 53 sequences in which ambiguous sites and indels were deleted to prevent identification of false recombination events. The phi-test (Bruen, Philippe et Bryant, 2006) implemented in Splitstree v4.10 (Huson et Bryant, 2006) and GARD, (Kosakovsky Pond *et al.*, 2006) available *via* the DataMonkey website at <http://www.datamonkey.org/> (Pond et Frost, 2005), were used for these analyses. Recombination breakpoints were detected and the identity of parental and recombinant sequences were identified using multiple algorithms: RDP (Martin et Rybicki, 2000); GENECONV (Padidam, Sawyer et Fauquet, 1999); BOOTSCAN (Salminen *et al.*, 1995); MAXIMUM CHI SQUARE (Maynard Smith, 1992); CHIMAERA (Posada et Crandall, 2001) and SISTER SCAN (Gibbs, Armstrong et Gibbs, 2000) implemented in the RDP v3.44 software (Heath *et al.*, 2006 ; Martin, Williamson et Posada, 2005). We use kwarg, a heuristic alternative to the branch and bound method implemented in beagle (Lyngsø, Song et Hein, 2005) to calculate the haplotype lower bound, R_h (Song et Hein, 2005). Kwarg doesn't guarantee that the minimal recombination history will be found but it will return a history with a low number of recombination events under the infinite sites assumption. We ran kwarg

2048 times with the default scoring scheme and recorded the minimal number of recombination events for each iteration. Hudson and Kaplan minimum recombination events, R_m (Hudson et Kaplan, 1985) were calculated with the DnaSP v5 software (Librado et Rozas, 2009) and haplotypebound, a program in the same package as kwarg (Song et Hein, 2005). To determine if recombination has an impact on the transposase protein, a recombination analyses was applied to a reduced dataset containing partial sequences of the transposase ORF using GARD analysis (Kosakovsky Pond *et al.*, 2006) and the phi-test (Bruen, Philippe et Bryant, 2006).

4.4.5 Phylogenetic analyses and ancestral recombination graphs

As all the recombination analyses detected recombination events (see results) we constructed a phylogenetic network using the NeighborNet algorithm implemented in Splitstree v4.10 (Huson et Bryant, 2006). Dendrograms were produced for each fragment separated by recombination breakpoints highlighted by GARD analysis of the 53 sequences using the Neighbor-joining algorithm (Saitou et Nei, 1987). To understand more accurately the process and evolutionary history of recombination in the *Pokey* fragments cloned here, we constructed an ancestral recombination graph using kwarg (Lyngsø, Song et Hein, 2005 ; Song et Hein, 2003, 2005). We chose the simulation with the lowest R_r value to build this graph.

4.4.6 RFLP analysis

Given the intra- and inter-genomic variability of *Pokey* alleles in isolates from which elements were cloned, and the high number of recombination events in these sequences (see results), we aimed to explore the variability of presence/absences of these alleles in different isolates belonging to the *D. pulex* complex. To determine if recombinant *Pokey* alleles are encountered in these isolates we use a restriction

enzyme analysis (i.e. RFLP) on *Pokey* alleles amplified from 28S rRNA genes of 41 isolates (Tableau 4-1). Restriction enzymes that cut at specific sites in some *Pokey* alleles may be used to highlight *Pokey* sequences that have undergone recombination events. We used NebCutter v2.0 (<http://tools.neb.com/NEBcutter2/>) on each of the 53 sequences to choose enzymes that would highlight different pure and recombinant *Pokey* alleles. The restriction enzymes DraI, BspHI and BstEII (NewEngland BiolLabs Inc., Ipswich, MA) are used in a two-step protocol. On the approximately 1600 pb length of the partial *Pokey* sequences amplified in this study DraI, BspHI and BstEII enzymes (New England Biolabs) cut at 380, ~790, ~ 1080 pb. The eight possible conformations may be differentiated on a 3% agarose gel. RFLP are conducted using a two steps proctocol. In the first step DraI and BspHI are used in the same mix using the NEBbuffer4 and digestion is conducted at 37°C during two hours. In the second step NEBbuffer3 for a final 1X concentration, BSA at 1U/μl and BstEII at 1U/μl are added to the solution and digestion is conducted at 60°C during one hour. At each digested PCR products were run on a 3% agarose gel for 2 hours with a fixed voltage of 95 volts. The 100bp VWR® Molecular Weight Marker (VWR International) was used as a size standard. RFLP haplotypes are encoded with letters (see Table 5). Haplotype **a** has been cut with DraI and BstEII. Haplotypes **b**, **c** or **f** have been cut with DraI, BspHI or BstEII respectively. Haplotypes **d** or **g** have been cut with BspHI and BstEII, DraI and BspHI respectively. Haplotype **e** has been cut with DraI, BspHI and BstEII and Haplotype **h** has not been cut by any of these three enzymes.

4.5 Results

4.5.1 Quantification of *Pokey* insertion sites

Twenty-seven *Daphnia* isolates were analyzed for *Pokey* insertion sites, of which 14 are diploids and 13 are polyploids. Of these, six isolates that have *D. tenebrosa* mitochondrial haplotypes and there is no evidence based on microsatellite

genotypes that any of them have a hybrid origin. We used the results of TE display to generate a *Pokey* profile for each isolate indicating the presence (1) or absence (0) of insertion fragments observed (Tableau 4-5). The mean number of *Pokey* insertions in diploid isolates was 17.43 (± 4.64), with values from 7 to 26, whereas the mean number of *Pokey* insertions in triploid isolates was 19.92 (± 4.37), with values from 13 to 28. Overall, 88 different *Pokey* insertion sites were detected (See Tableau 4-5). The 13 polyploid isolates displayed 72 of these *Pokey* insertion sites, of which 14 (19.4%) were found in only one isolate (singletons). The 14 diploid isolates had 68 of these *Pokey* insertion sites, of which 8 (11.8%) were found in only one isolate. In the phylogenetic network based on patterns of *Pokey* insertion sites (Figure 4-1), five of six *D. tenebrosa* isolates branch together (TE3-UN-1, TE3-UN-2, TE3-UN-3, TE2-UN-1 and TE2-UN-3) although *pulex-pulicaria* diploid and polyploid isolates appeared to be distributed without a clear pattern. This pattern may be the consequence of different distributions and dynamics of the *Pokey* transposon in *D. tenebrosa*. We decided to exclude *D. tenebrosa* isolates from the ANCOVA analysis. Without *D. tenebrosa* isolates, 87 *Pokey* insertion sites were detected (Tableau 4-5). The mean number of insertions in diploid hybrid isolates was 18.82 (± 3.60), with values from 14 to 26, whereas the mean number of *Pokey* insertions in polyploid isolates was only 21.50 (± 7.30), with values from 17 to 28. The difference in insertion numbers between diploid and polyploid hybrids is not statistically significant (ANCOVA P-value=0.06485; see Figure 4-2 and Tableau 4-2). The 10 polyploid isolates displayed 68 of the 87 *Pokey* insertion sites, of which 14 (20.6%) were found in only one isolate. Eleven diploid hybrid isolates had 65 of these *Pokey* insertion sites, of which 8 (12.3%) were found in only one isolate. The difference between the number of singletons between diploid and polyploid hybrids was not statistically significant (Fischer exact test, P-value=0.3579).

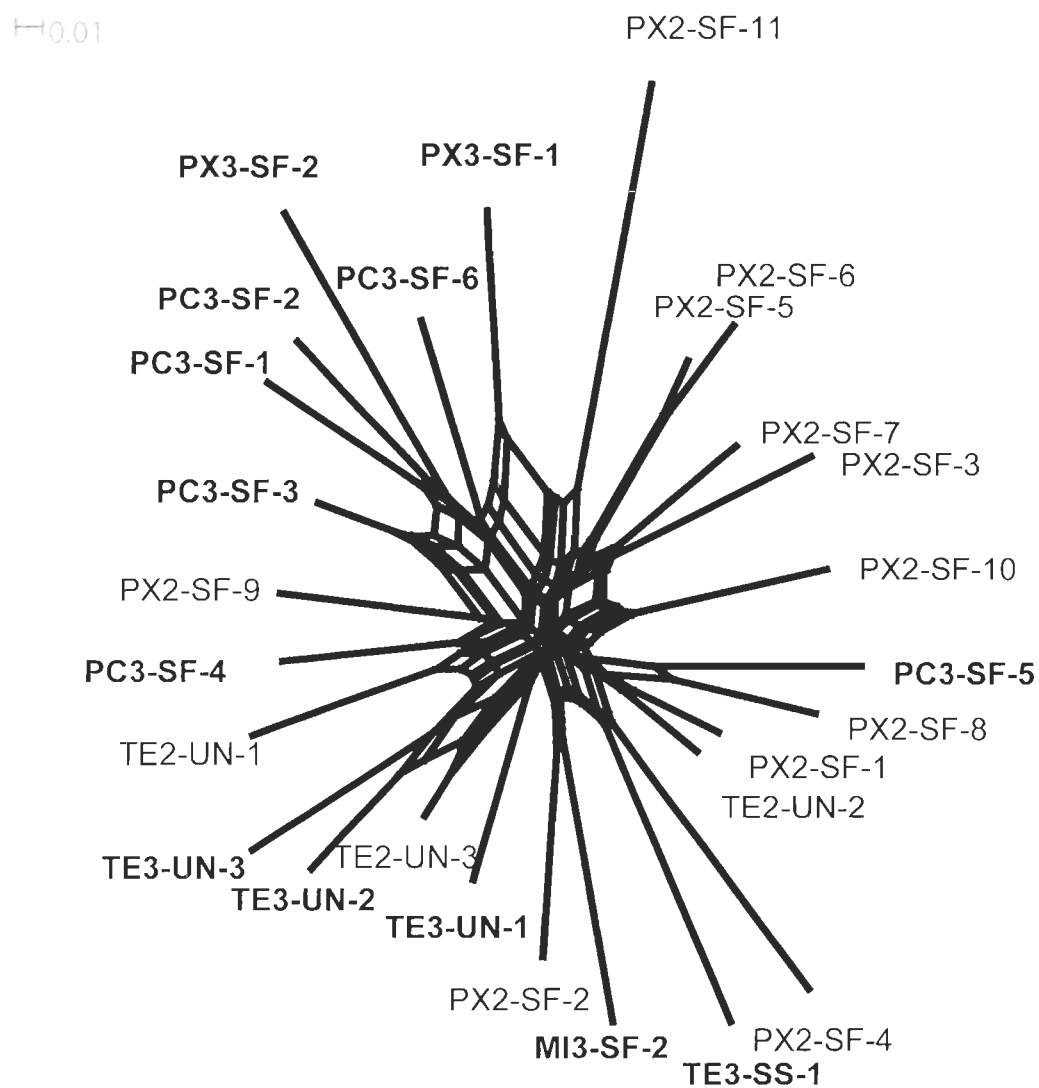


Figure 4-1 Phylogenetic network on the binary matrix obtained with the TE display technique from diploid and polyploid hybrid isolates of the *Daphnia pulex* complex.

Tableau 4-2 Results of the covariance analysis (ANCOVA) on TE display results using the ploidy level and H_{pl} as the independent variables.

	Estimate	Standard Error	t-value	Pr(> t)
(Intercept)	13.737	3.057	4.493	0.000281***
H_{pl}	12.918	7.326	1.763	0.094791
Ploidy	1.812	1.560	1.162	0.260561

Residual standard error: 3.387 on 18 df.; Multiple R^2 : 0.2621; Adjusted R^2 : 0.1801; F-statistic:3.197 on 2 and 18 df.; p-value=0.0648.

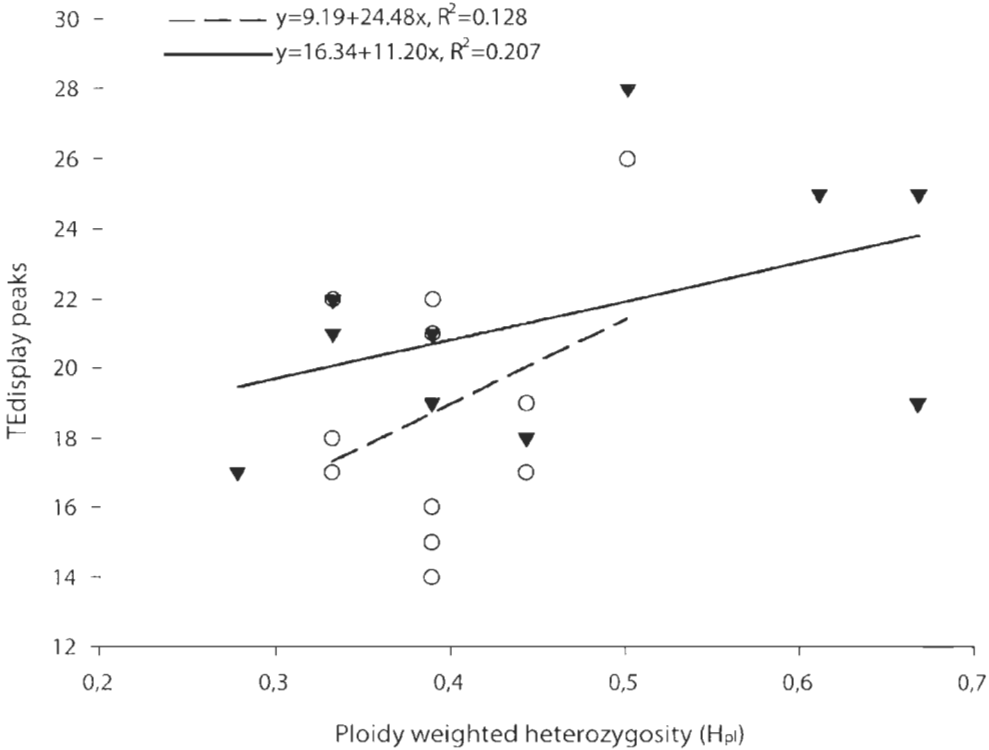


Figure 4-2 Relation between *Pokey* insertion number sites and ploidy weighted heterozygosity (H_{pl})
Empty circles represent diploid isolates and solid inverted triangles represent triploid isolates

4.5.2 *Pokey* sequences

4.5.2.1 Recombination signals

Forty-four sequences from the 3' end of the *Pokey* element (GenBank [accession numbers Xxxxxx to Yyyyyy]) were added to nine *Pokey* alleles previously sequenced from non-hybrid clones for a total of 53 sequences analyzed. These sequences had variable lengths due to multiple indels and varied from 1419 (PX2-SF-8_30) to 1471 (TE3-SS-1_9) nucleotides. These sequences contain two open reading frames (ORFs). The first is not complete and represents the 3'end of the first exon of the *Pokey* transposase (T. Elliott and I. Bigot, unpublished data). The second ORF, which is the second exon of the *Pokey* transposase is complete and is variable in length. Total amino acid (aa) sequence lengths of these two ORFs range from 286 aa in most sequences to 299 aa. One sequence (TE3-UN-1_14) has a 1-nt insertion that causes a frameshift mutation in the first ORF.

All recombination analyses performed on the 53 *Pokey* sequences showed evidence of recombination events with a phi-test (P -value <0.01). Two recombination breakpoints were found (at positions 540 and 820) using GARD analysis of the entire sequences. A minimum of six unique recombination events was detected using the 6 algorithms implemented in RDP3.44 (Tableau 4-3, Figure 4-3). Unique recombination events were detected in sequences that share the same parental sequences and the same recombination breakpoints. Groups of sequences that are the results of these 6 unique recombination events have been named A to F (see Tableau 4-3). Recombination events A and B seem to be similar to the events that are highlighted by the GARD analysis (Tableau 4-3). Some recombination breakpoints may be undetermined due to multiple recombination events that may be entangled. The Hudson-Kaplan minimum number of recombination events (R_m) in our partial *Pokey* sequences dataset is 30 according to the algorithm implemented in DnaSP v5 and 36 using the algorithm implemented in kwarg (Figure 4-5). According to the kwarg algorithm, the values of the haplotype lower bound R_n follow a normal distribution with mean 96.39 and a standard deviation of 3.1 (Figure 4-5). Phi-test (P -

value <0.01), GARD analysis (two recombination breakpoints detected) and algorithms of the RDP3 software detected evidence of recombination events in the reduced dataset representing the 53 ORFs (about 858 to 897 pb. length) encoding *Pokey* transposases.

Tableau 4-3 Summary of recombination events detected on 53 *Pokey* sequences from the *Daphnia pulex* complex using RDP3 software

<i>Recombinant events (number of sequences)</i>	<i>Breakpoints in alignment^a</i>		<i>Parental sequences</i>		<i>algorithms</i>	<i>P-values</i>
	<i>begin</i>	<i>end</i>	<i>minor</i>	<i>major</i>		
A(16)	11 ^b	807	TE3-UN-2_4	PX2-SF-8_30	RDP, GeneConv, Bootscan, Maxchi, Chimaera, SiScan	1.39E-07, 3.23E-04, 4.82E-07, 3.86E-09, 3.93E-09, 3.27E-13
B (2)	1 ^b	547	PC3-SF-1_28	PC3-SF-3_17	Bootscan, Maxchi, Chimaera, SiScan	2.21E-04, 4.67E-06, 3.04E-05, 9.38E-10
C (5)	591	807 ^b	PC3-SF-5_15	AR2-UN-1_1	Maxchi, SiScan	2.38E-05, 7.97E-08
D (13)	824	1113 ^b	PC2-FF-4_3	Unknown ^c	Maxchi, SiScan	6.69E-05, 1.57E-18
E (6)	766	1102 ^b	PX3-SF-1_20	Unknown ^c	Maxchi	8.16E-05
F (1)	494	717 ^b	PC3-SF-3_5	TE3-UN-2_25	SiScan	8.82E-04

^a positions of putative recombinant breakpoints in the alignment; ^b actual breakpoint position is undetermined, this is most likely overprinted by a subsequent recombination event; ^c Only one parent and a recombinant need be in the alignment for a recombination events to be detectable.

4.5.2.2 *Pokey* phylogenetic network and dendrograms

The phylogenetic network using 53 partial *Pokey* sequences showed clear groups of sequences (Figure 4-3). In this phylogenetic network, sequences amplified from different species (TE2-UN and TE3-UN from *D. tenebrosa*, EPX2-UN from European *D. pulex*, PX2-SS from *D. pulex*) are each grouped together (Figure 4-3). All the sequences from the isolate TE3-SS-1 and sequences belonging to triploid isolates from Churchill (PC3-SF-6 and MI3-SF-2) which have *D. pulicaria* or *D. middendorffiana sensu stricto* haplotypes (Vergilino, Belzile et Dufresne, 2009) represent two different groups. Various sequences amplified from *D. pulicaria* (PC2-FF-4 and PC2-FF-5) or *D. pulex hybrid* isolates (PX2-SF-8, PX3-SF-1 and PC3-SF-1) clustered together. Three groups of sequences contain sequences amplified from isolates that belong to different species (Figure 4-3). Two groups contain sequences from *pulex-pulicaria* hybrid clones (PX2-SF-8 and PC3-SF-3), *D. tenebrosa* clones (TE3-UN and TE2-UN), European *D. pulicaria* clones (EPC2-UN). Another group is represented by sequences from *D. arenata* (AR2-UN-1), *D. pulicaria* (PC2-FF-4) and *pulex-pulicaria* hybrid (PX2-SF-8). An allele from a *D. tenebrosa* isolate (TE2-UN-2) groups with sequences from *pulex-pulicaria* hybrid clones.

Dendrograms produced from fragments of the *Pokey* sequences bounded by recombination breakpoints identified by the GARD analysis (Figure 4-4) show some potential origins of *Pokey* alleles from the MI3-SF-1 and PC3-SF-5 isolates (*D. middendorffiana sensu lato* isolates). Some of these alleles seem to have arisen from recombination events between *Pokey* alleles from *D. tenebrosa* isolates (TE3-UN-1, TE3-UN-2, TE3-UN-3 and TE2-UN-1) and *Pokey* alleles from different hybrid (PX2-SF-8) or *D. pulicaria* (PC2-FF-4) isolates. *Pokey* alleles from the TE3-SS-1 isolate (introgressed triploid *D. tenebrosa* with *D. pulex* nuclear genome; see Dufresne et Hebert, 1994 ; Vergilino *et al.*, 2011) seem to have arisen from alleles from *D. pulex* (PX2-SS-CON, PX2-SS-7_1 and PX2-SS-8_1) and *D. tenebrosa* (TE3-UN-1, TE3-UN-2, TE3-UN-3 and TE2-UN-1) with two recombination breakpoints (see Figure 4-4). Some *Pokey* alleles encountered in *D. tenebrosa*

isolates (TE3-UN-1_14, TE2-UN-3_3, TE2-UN-3_20 and TE3-UN-3_13) and European *D. pulicaria* (EPC2-UN-3 and EPC2-UN-2_1) seem to have arisen from recombination between alleles that are most similar to *D. tenebrosa* *Pokey* alleles and to *pulex-pulicaria* *Pokey* alleles. A complex pattern of recombination may have produced the *Pokey* diversity encountered in diploid and polyploid hybrids from the *pulicaria* group (PX2-SF-8, PC2-FF-4, PC3-SF-1, PC2-UN-5_2 and PX3-SF-1) and *D. arenata* (AR2-UN-1). *Pokey* alleles encountered in *D. tenebrosa* (TE2-UN-2_6 and TE2-UN-1_8) and in European *D. pulicaria* (EPC2-UN-1_9) are most closely related to these last alleles.

In the Figure 4-3, we indicated sequences that are putative recombinant sequences (A, B, C, D, E and F; see Tableau 4-3). Some sequences appeared to be the result of two recombination events. The sequence from TE3-UN-3_10 originated from recombination events D and F. The sequences from PX2-SF-8_1 and PX2-SF-8_29 were the result of recombination events E and B. All the *Pokey* alleles from TE3-SS-1 appeared to have undergone recombination events A and C. The shape of the phylogenetic network showed that some additional events might exist between *Pokey* alleles (Figure 4-3). For example, *Pokey* alleles from MI3-SF-1 and PC3-SF-5 appear to be the result of recombination events between *Pokey* alleles from *D. tenebrosa* and the group of alleles comprising PC3-SF-3_5 and EPC2-UN-2_1.

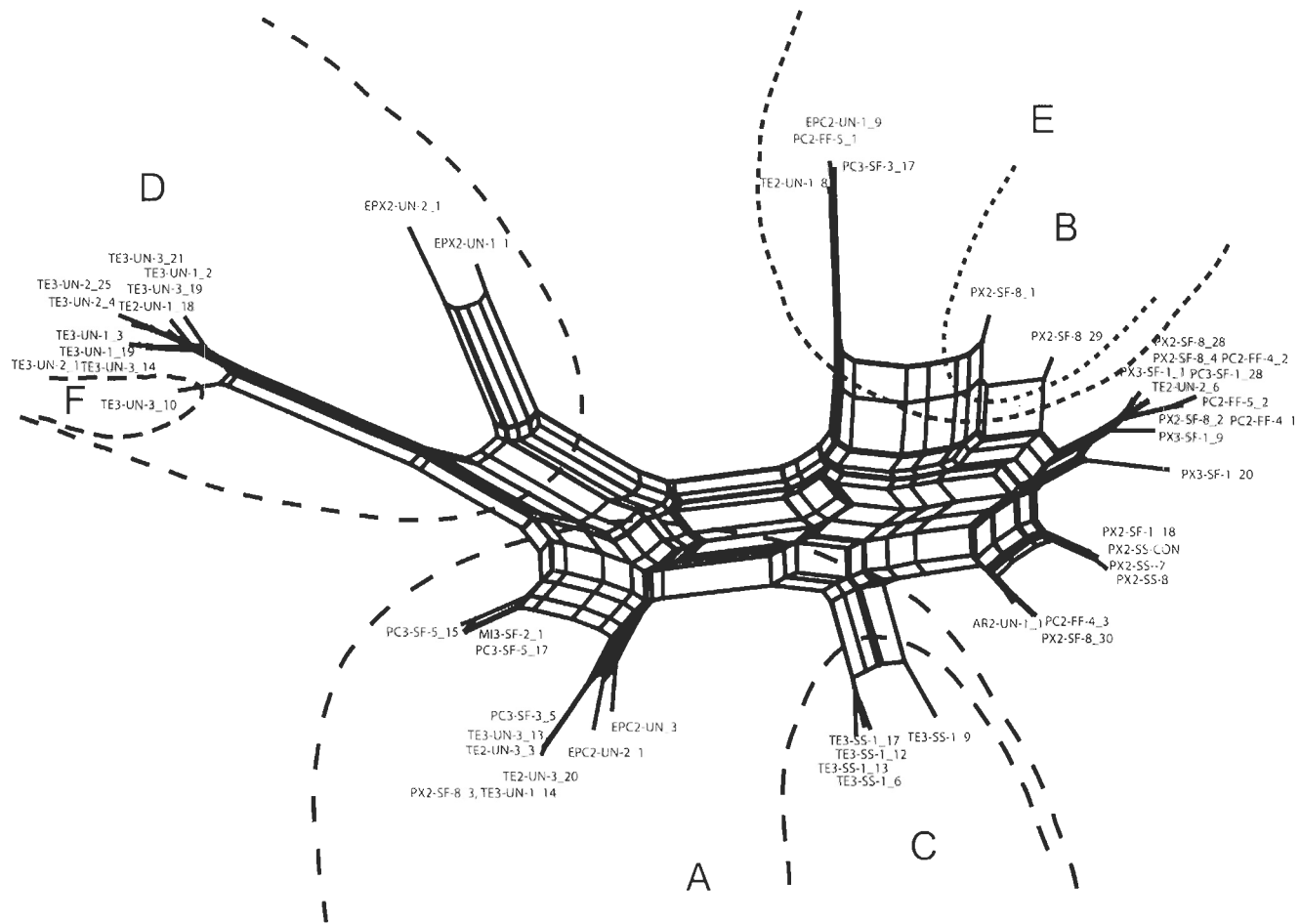


Figure 4-3 Phylogenetic network constructed using the NeighborNet algorithm on the 53 *Pokey* sequences dataset from isolates of the *Daphnia pulex* complex
 Dashed lines represent the set of sequences that have undergone recombination events (capital letters) detected with RDP v3.44 software

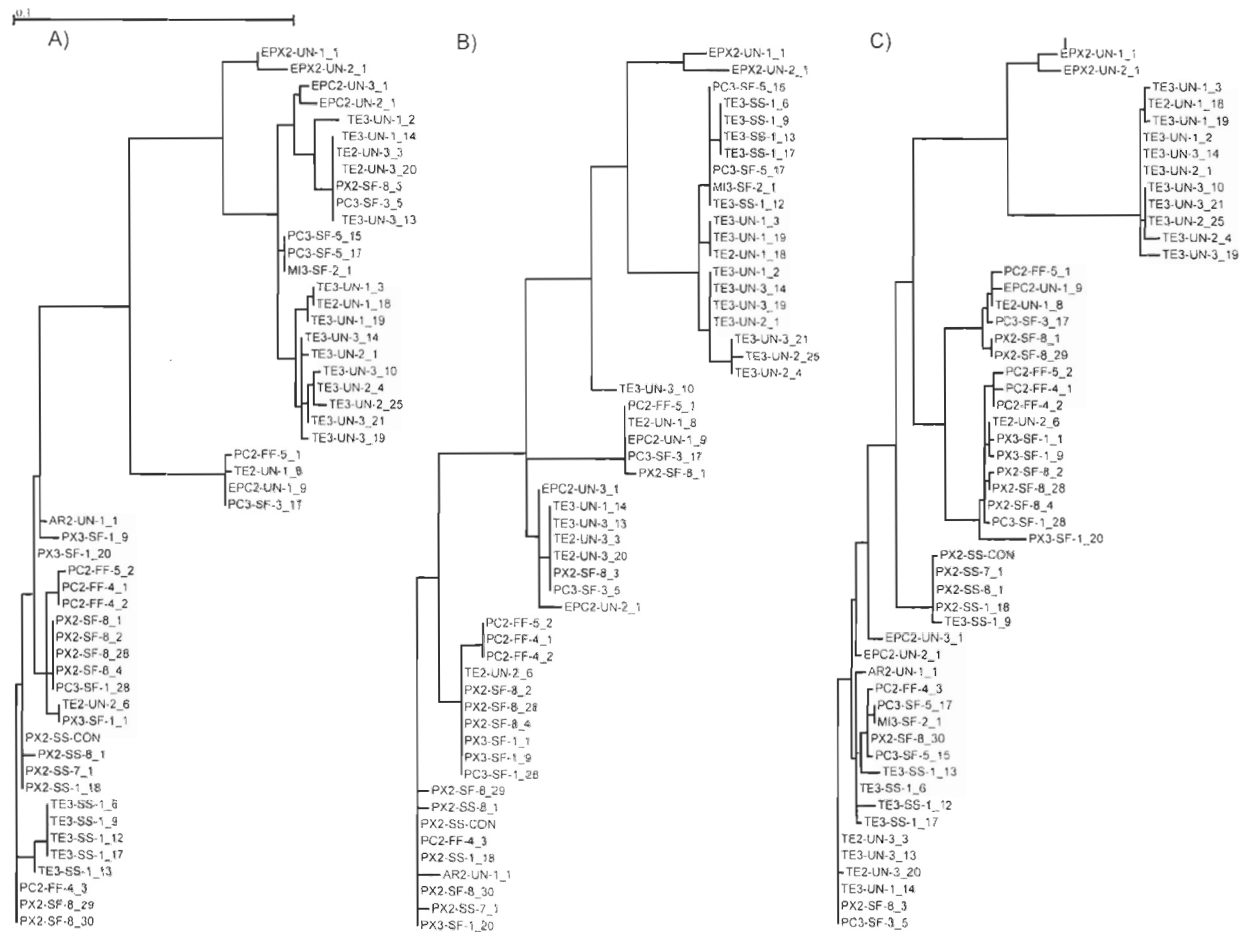


Figure 4-4 Unrooted Neighbor-Joining trees constructed with fragmented sequences flanked by recombination breakpoints highlighted by GARD analysis on a 53 sequences dataset
 A), B) and C) represent trees constructed from each part of recombinant sequences (1~540 nuc., ~540~820 nuc., and ~820~1450 nuc. respectively)

4.5.2.3 Ancestral recombination graphs

The number of recombination events detected by the RDP3 software may be underestimated. Indeed, the estimated values of R_m using DnaSP, R_m using kwarg and R_n are higher (30, 36 and 96.39 ± 3.10 respectively; Figure 4-5) than the six unique recombination events estimated using RDP3 software. Figure 4-6 shows an ancestral recombination graph with 89 recombination events (R_n) using the Song and Hein algorithm (2005). In Figure 4-6, we highlight the evolutionary history of the *Pokey* allele TE3-SS-1_12, which contains 38 recombination events. The Ancestral Recombination Graph shows putative past recombination events that were not detected by the different algorithm of the RDP3 software. For example the evolution of PC2-FF5_2, PC2-FF-4_1 and PC2-FF-4_2 encountered in *D. pulicaria* isolates seem to involve 19 past recombination events. Algorithms provided in the RDP3 software screen for evidence of recombination events using distance and phylogenetic methods (Heath *et al.*, 2006 ; Martin, Williamson et Posada, 2005). They don't reconstruct the recombination events history as do by Ancestral Recombination Graph algorithms (Hudson et Kaplan, 1985 ; Lyngsø, Song et Hein, 2005 ; Song et Hein, 2005) Thus RDP3 analyses may fail to recover past recombination events that may be blurred by more recent recombination events.

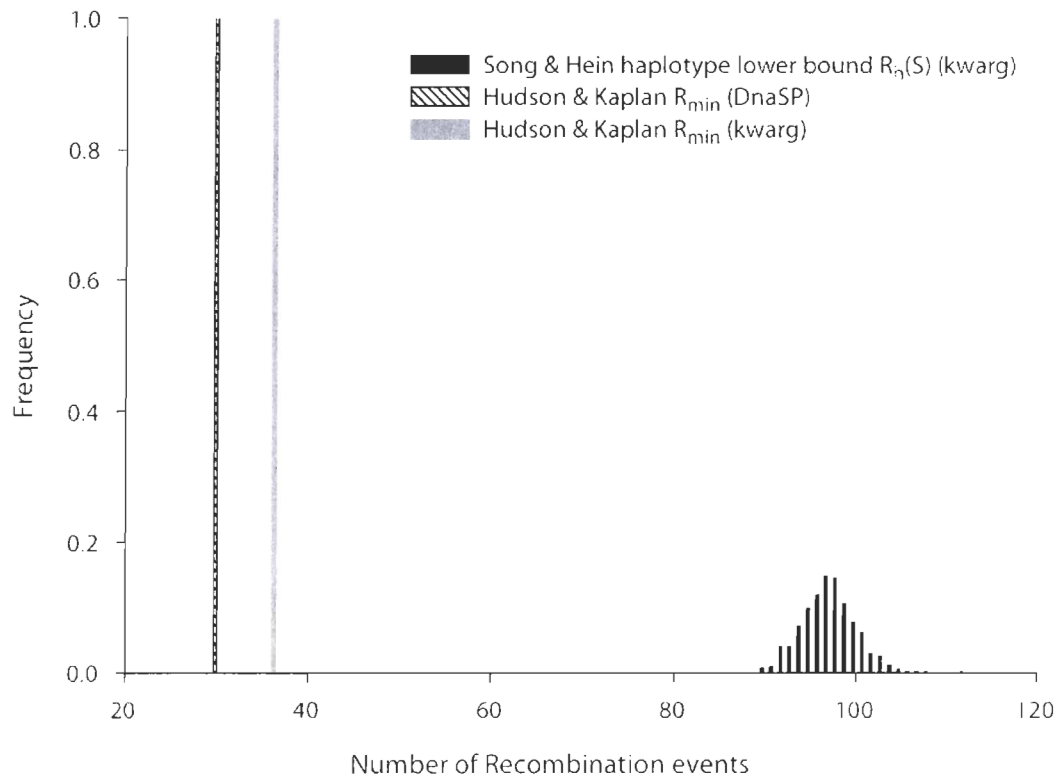


Figure 4-5 Distribution of recombination parameters calculated on the 53 *Pokey* sequences dataset using kwarg and DnaSP software



Figure 4-6 One of the possible ancestral recombination graphs constructed on *Pokey* sequences showing 89 recombination events. Ovals at the bottom of the graph are *Pokey* alleles. Ovals with trifurcation encountered when the history is followed backward in time until the *Pokey* ancestral sequence is reached at the top of the graph represent putative recombination events. Evolutionary history of the *Pokey* allele TE3-SS-1_12 is represented by thicker lines

4.5.3 *Pokey* alleles in different isolates

To test if recombinant *Pokey* alleles are encountered in additional clones (hybrids or not) of the *D. pulex* complex, we used a restriction enzyme analysis (i.e. RFLP) on *Pokey* alleles amplified from the 28S rRNA *gene* from 41 isolates (Tableau 4-4). Eight different haplotypes were found (**a** to **h**; Tableau 4-4). These haplotypes do not represent all the different *Pokey* alleles that may be encountered in each isolate. RFLP haplotypes **a** and **b**, which are found in 33.33% and 38.09% of the isolates, respectively correspond to alleles that are encountered in pure *D. pulex*, *D. pulicaria* and in diploid and polyploid hybrids between them. The RFLP haplotype **b** may also be encountered in *D. tenebrosa* isolates. The RFLP haplotype **c** (encountered in 26.19% isolates) corresponds to the sequences that have undergone the recombination event D or E (Tableau 4-3). The RFLP haplotype, **d** (encountered in 30.95% isolates) corresponds to the sequences that have undergone the recombination event A but not the recombination event C. Haplotype **e** (encountered in just one isolate) corresponds to the sequences that have undergone recombination events A and C. Haplotypes **f** and **g** (encountered in 4 and 1 isolates respectively) correspond to rare alleles of unknown recombination status. Haplotype **h** (encountered in 28.57% isolates) may correspond to the sequences that have undergone recombination events E and/or B. The sequences with recombination event F can not be differentiated from other recombinant D sequences with our RFLP analysis. The different genomes may contain from 1 to 5 different *Pokey* RFLP haplotypes with an average of 1.74 and a standard deviation of 0.96. Some genomes present only one *Pokey* allele amplified from their rRNA genes (see PC3-SF-5, PX2-SS-1, PC3-SF-4, EPC2-UN-1 and EPC2-UN-2 in Tableau 4-4) and this allele is a putative recombinant. The v1.1 genome sequence of *D. pulex* was scanned for our partial sequence of *Pokey* using a homology-based approach. Queries representing *Pokey* elements sequenced in our study were used in TBLASTN searches of the pre-release genome. Sequences from contigs

containing sequences with E-value = 0.0 were aligned with our dataset of 53 sequences to generate a dataset of 58 sequences. Sequences with indels more than 10 nucleotides were omitted in subsequent analyses. A NeighborNet phylogenetic network was constructed and a recombination detection analysis was performed on the larger dataset using RDP3.44 software. The resulting phylogenetic network is shown in Figure 4-7. Recombination events were detected in some of the sequences from the genome sequence. These sequences group with sequences encountered in diploid and polyploid isolates from the *pulicaria* group (PC2-FF-5 and PC3-SF-3), in diploid *D. tenebrosa* (TE2-UN-1) and in European *D. pulicaria* (EPC2-UN-1) which have undergone the recombinant event E (Figure 4-7).

Tableau 4-4 Pokey RFLP haplotypes amplified from 41 isolates of the *Daphnia pulex* complex.

Isolates	RFLP haplotypes							
	a	b	c	d	e	f	g	h
MI3-SF-1	x			x		x		
PX2-SS-1				x				
MI3-SF-2				x		x		
PC3-SF-4				x				
EPC2-UN-1			x					
PC3-SF-6			x					
PX2-SF-1	x							
PX2-SS-2			x	x				
TE3-SS-1		x			x			
TE3-UN-1			x	x				
TE3-UN-2			x	x				
TE3-UN-3			x					
TE2-UN-1			x					x
TE2-UN-2		x						
TE2-UN-3				x				
PX2-SS-3	x							
PX2-SF-2	x							x
PX2-SF-3	x	x						
PX2-SS-4	x							
PX2-SF-5	x	x						
PX2-SF-6	x	x						
EPC2-UN-1				x				
PX2-SF-7	x		x	x			x	x
PX3-SF-2	x							
PX2-SF-8	x	x	x	x				x
PC3-SF-3			x	x				
PX3-SF-1	x	x						x

Isolates	RFLP haplotypes							
	a	b	c	d	e	f	g	h
PC2-FF-4	x	x						
PC3-SF-1		x						
PC3-SF-2		x	x					x
PC3-SF-7						x		x
PX2-SF-9						x		
PC3-SF-5				x				
PX2-SF-12	x							
PX2-FF-1		x						x
PX2-FF-2		x						x
PX2-FF-3		x						x
PX2-FF-4		x						
PC2-FF-2		x						
PC2-FF-1		x						
PC2-FF-3								x
MI3-SF-1	x			x		x		

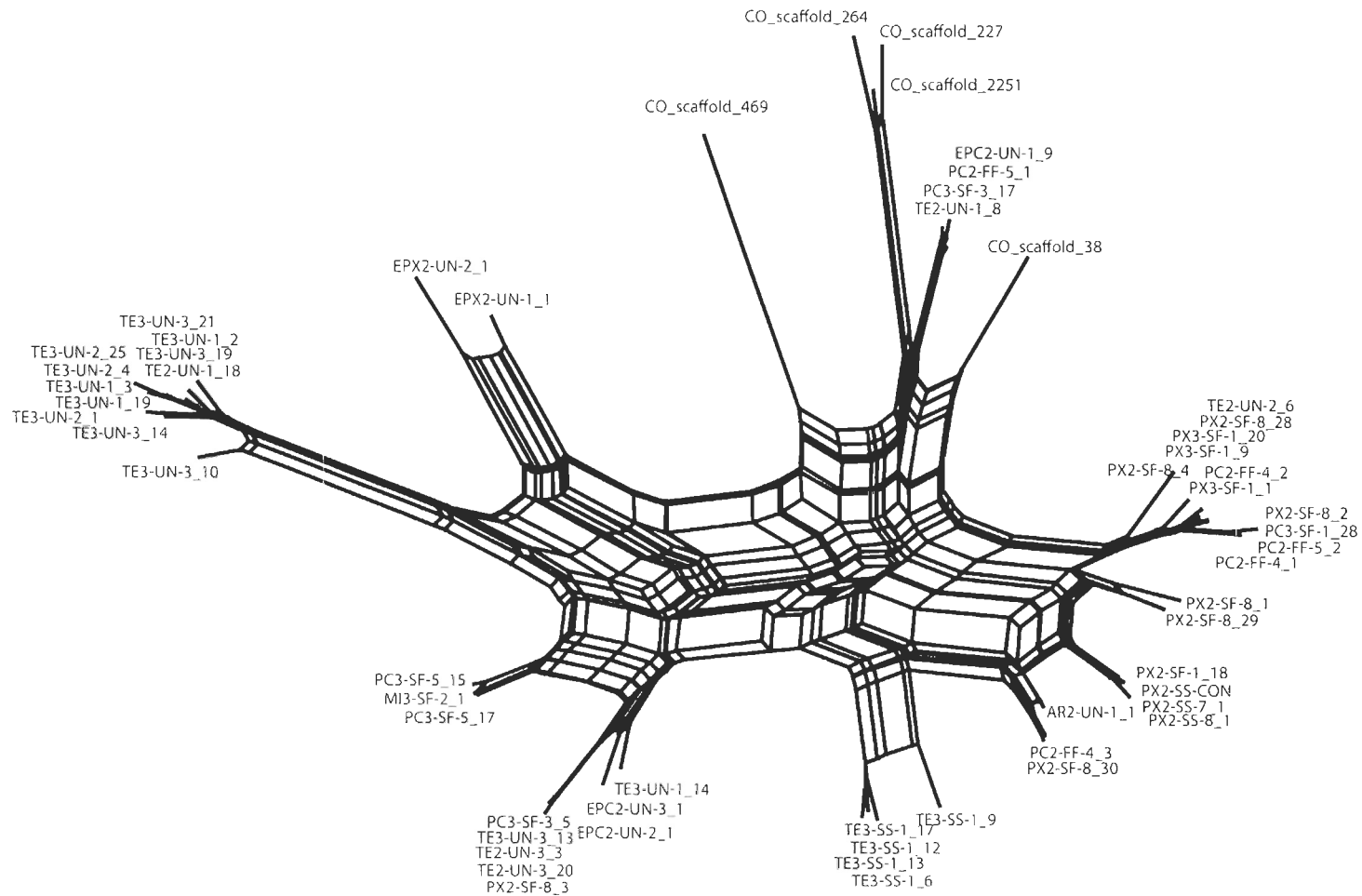


Figure 4-7 Phylogenetic network constructed using the NeighborNet algorithm with *Pokey* alleles encountered in the v1.1 draft genome of *Daphnia pulex*. Sequences named CO_Scaffold_xxx are encountered throughout the v1.1 draft genome of *Daphnia pulex*.

4.6 Discussion

4.6.1 Effect of ploidy level on the density of *Pokey* into the genome

The effect of ploidy level on TE insertion number has been mostly studied in plants including both autopolyploids (Hazzouri *et al.*, 2008) and allopolyploids (Madlung *et al.*, 2002 ; Parisod *et al.*, 2010). Theory predicts that transposon insertions increase with ploidy level (Madlung *et al.*, 2002 ; Matzke et Matzke, 1998 ; Wendel, 2000). Some studies on plants and fishes have shown an activation of TEs in newly synthesized allopolyploids (Liu *et al.*, 2009 ; Parisod *et al.*, 2010) but few studies show an increase in TE insertion sites with ploidy in the short or long term (Parisod *et al.*, 2010). Mobilization of TEs may be family dependent (Parisod *et al.*, 2010), and the regulation ability of the host genome also seems to play a crucial role in controlling mobilization and insertion processes (Madlung *et al.*, 2002 ; Yoder, Walsh et Bestor, 1997). Numerous studies have shown TE mobilization in newly synthesized allopolyploid crops (de Araujo *et al.*, 2005 ; Hanson *et al.*, 2000 ; Kashkush, Feldman et Levy, 2002) and in *Arabidopsis* allopolyploids (Hazzouri *et al.*, 2008 ; Madlung *et al.*, 2002). A recent study on cyprinid fish shows that crosses between two divergent species of carp (*Megalobrama amblycephala* and *Carassius auratus* red var.) generated a burst of activity in a *Tc1*-like retrotransposon in fertile allotetraploid offspring but not in sterile triploid hybrid offspring (Liu *et al.*, 2009). The effects of hybridization *per se* and of the ploidy level increase are entangled in these newly synthesized allopolyploids and hence difficult to separate.

Our study is one of the first to examine the number of TE insertion sites in relation to the ploidy level in natural animal populations. Since the diploid and polyploid isolates we analyzed are both hybrids with the same reproductive mode, they can be compared. No significant differences in insertion number of *Pokey* were found between diploid and polyploid hybrids (18.82 ± 3.60 vs. 21.50 ± 7.30) suggesting that an increase in ploidy level does not lead to an increase in *Pokey* insertions in the long term. Alternatively, it is possible that increases in *Pokey* insertions occur shortly after transitions to higher ploidy levels but that genomic

reorganization results in the loss of TE sequences with time. As the clonal isolates we studied were not experimentally produced, their age is not known. Previous studies have suggested that asexual populations of *D. pulex* have originated some 150,000 years ago (Paland et Lynch, 2006) and that some polyploids from the *D. pulex* complex were produced during the Pleistocene (Dufresne et Hebert, 1997).

The absence of statistical significance between insertion site numbers in diploid and polyploid *Daphnia* isolates may be due to an increase and a subsequent decrease of TEs in polyploid genomes via some ectopic recombination and genome rearrangement events rather than an absence of increase of TEs in polyploids. Parisod et al. (2010) noted that the loss of TEs after polyploidization in both the short and long term is common, whether or not they activate. These losses are thought to be due to genome rearrangement via unequal and ectopic recombination events between TE alleles at non-homologous loci. The slightly greater number of singletons in polyploids (20.6%) than diploids (12.3%) probably reflects *Pokey* insertions in the additional chromosome set those triploids carry.

Bursts of TE activity in hybrids have been highlighted in numerous homoploid (of the same ploidy level) hybrid plants (Ainouche *et al.*, 2009 ; Shan *et al.*, 2005), fruit flies (Fontdevila, 2005 ; Labrador et Fontdevila, 1994 ; Petrov *et al.*, 1995) and wallabies (O'Neill, O'Neill et Graves, 1998). In *Drosophila melanogaster* and *D. virilis*, hybrid dysgenic crosses may lead to bursts of various TE activities (Lewis et Brookfield, 1987 ; Petrov *et al.*, 1995) due to release from control of cytoplasmic repression (Blumenstiel et Hartl, 2005 ; Brookfield, 1991 ; Castro et Carareto, 2004). Valizadeh and Crease (2008) found lower average values in asexual diploid isolates of *D. pulex* (3.27 ± 2.07 , with values from 1 to 8) than in our survey of asexual diploid hybrid isolates (18.82 ± 3.60 , with values from 14 to 26). However, they used a higher annealing temperature so it is possible that the large difference between our results and theirs is due solely to the difference in annealing temperature between the two studies (55°C instead of 50°C). To test this, we performed TE display on nine of our

diploid hybrid isolates using the same annealing temperature used by Valizadeh and Crease (2008) and still found higher values of *Pokey* insertion in our hybrid clones (10.55 ± 5.22 , with values from 3 to 17). There is a substantial variability in the TE insertion numbers in both diploid and polyploid isolates, which may lead to a non-significant difference between these two groups. This variability may be due to the presence or absence of active TEs in parental species, which may influence the subsequent proliferation of TEs in the hybrid progeny. Increases in TE copy number may depend on TE activity and the effectiveness of regulation in the hybrid genome as it has been highlighted in *Drosophila* hybrid dysgenic crosses (Blumenstiel et Hartl, 2005 ; Brookfield, 1991 ; Lewis et Brookfield, 1987 ; Petrov *et al.*, 1995).

4.6.2 Evolution of *Pokey* via intragenic recombination

All the analyses show evidence of recombination events in *Pokey* sequences inserted in the rDNA loci. These recombination signals don't seem to be due to cloning artifacts as some isolates may have just one recombinant *Pokey* allele in their genome (e.g PC3-SF-5 and PC3-SF-4 triploid isolates or EPC2-UN-1, EPC2-UN-2 and PX2-SS-1 diploid isolates). Moreover, increasing the extension time of the PCR cycles to avoid artefactual recombination events during amplification (data not shown) did not lead to any different RFLP results. Numerous studies have focused on the relationships between recombination rates in host genomes and TE distributions (Carr *et al.*, 2002 ; Duret, Marais et Biémont, 2000 ; Langley *et al.*, 1988 ; Rizzon *et al.*, 2003) or dynamics (Schaack *et al.*, 2010a ; Schaack *et al.*, 2010b ; Valizadeh et Crease, 2008). Different theories (Dolgin et Charlesworth, 2008) aim to explain the underlying mechanisms that are responsible for the non-random distribution of TEs in the genome (Biémont, 1992 ; Duret, Marais et Biémont, 2000). The distribution and abundance of TEs may be influenced either by selective forces (Brookfield, 1991 ; Finnegan, 1992 ; Montgomery, Charlesworth et Langley, 1987), the stochastic process of mutation (Biémont, 1992) or the availability of insertion

sites. Models based on the action of selection invoke either the impact of recombination events between TEs inserted in non-homologous loci («ectopic exchange model», Montgomery, Charlesworth, Langley 1987), the deleterious effect of TE insertions in coding (or in regulatory) sequences («deleterious insertion model», Finnegan 1992) or the deleterious effect of transposition activity itself («deleterious transposition model», Brookfield 1991) to account for the deleterious effects of TEs. The «ectopic exchange model» and the «deleterious insertion model» predict a negative correlation between the recombination rate along the chromosome and the TE insertion frequency (Charlesworth, Langley et Sniegowski, 1997 ; Langley *et al.*, 1988) while the mutational model (Biémont, 1992) and the «deleterious transposition model» do not. Empirical data from sequenced genomes show that the relationship between TE density and recombination rate depends on the organism and the TE class or family (Bartolomé, Maside et Charlesworth, 2002 ; Duret, Marais et Biémont, 2000 ; Eickbush et Eickbush, 1995 ; Kim *et al.*, 1998 ; Penton, Sullender et Crease, 2002 ; Rizzon *et al.*, 2002 ; Wright, Agrawal et Bureau, 2003). For example, DNA transposons in *Caenorhabditis elegans* are located preferentially in recombination hotspot regions whereas retrotransposons are not (Duret, Marais et Biémont, 2000). Researchers have suggested a role for recombination in the transposition process to account for this discrepancy between theory and empirical data (Duret, Marais et Biémont, 2000).

Some TEs (*Pokey* in *Daphnia* or *R1* and *R2* in *Drosophila*) are found in rRNA genes, which is a recombination hotspot (Eickbush et Eickbush, 1995 ; Penton, Sullender et Crease, 2002). Using the «ectopic exchange model» as a framework, Zhang et al. (2008) have shown that eukaryotic hosts may tolerate a high load of retrotransposable elements in their rDNA because they generally have many more rDNA copies than the minimum required for rRNA synthesis. Moreover, concerted evolution of rDNA occurs as a result of unequal cross-over and gene conversion between misaligned repeats (Eickbush et Eickbush, 2007), and unequal sister chromatid exchange or intrachromatid exchange seem to be the main processes that are responsible for rDNA copy number variation in *Daphnia* (but see Ambrose et

Crease, 2011 ; Crease, 1995 ; McTaggart *et al.*, 2007). However, Ambrose and Crease (2011) showed that IGS sequences from different *Daphnia* species have undergone numerous interchromatid exchanges. Our results show that recombination has a dramatic impact on the evolution of *Pokey* alleles in the rDNA of *Daphnia*. We detected a minimum of six unique recombination events between *Pokey* alleles from different species using different algorithms, but it seems that more than thirty recombination events have modeled the evolution history of *Pokey* alleles encountered in our study.

Because *Pokey* does not seem to be prone to horizontal transfer (Penton et Crease, 2004 ; Penton, Sullender et Crease, 2002 ; Sullender et Crease, 2001a), recombination between *Pokey* elements that have diverged in the genomes of different species may provide a marker with which to study the history of hybridization events among the members of the *D. pulex* species complex. For example, *Pokey* alleles in hybrid isolates, PX2-SF-8 and introgressed isolate, TE3-SS-1 seem to be recombinants between alleles encountered in their putative parental species (*D. pulex* and *D. pulicaria* for PX2-SF-8, and *D. tenebrosa* and *D. pulex* for TE3-SS-1). As some *Pokey* alleles in *D. middendorffiana sensu lato* (MI3-SF-1 and PC3-SF-5) seem to have arisen by recombination between alleles from *D. tenebrosa* and either *D. pulex* or *D. pulicaria*, it seems possible that some isolates of *D. middendorffiana s. l.* have arisen by hybridization or introgression between these species. *Pokey* alleles from European *D. pulicaria* isolates (EPC2-UN-1_9, EPC2-UN-2_1 and EPC2-UN-3_1) seem to be recombinants between alleles from *D. pulex* or *D. pulicaria* and *D. tenebrosa* (recombinant A and recombinant E). Ambrose and Crease (2011) have shown that IGS segments from some isolates of European *D. pulex* may have originated from North American *D. pulicaria*. As North American *D. pulex* and *D. pulicaria* have been found in Europe (Marková *et al.*, 2007), invasion of *Pokey* alleles from North American species into Europe species via hybridization or introgression is possible.

The location of *Pokey* elements in rDNA loci may be advantageous in the

long term as recombination between transposable elements located in these loci may induce diversification and may allow evading the host suppression machinery (Schaack *et al.*, 2010b). Inter-element sequences exchange may alter element fitness (Sharma, Schneider et Presting, 2008). Such hypothesis has been advanced to explain the activity of CRM1 retroelement in maize plant hybrids (Sharma, Schneider et Presting, 2008). Haplotypes **b** and **h** (which may represent recombinant *Pokey* alleles) are found in different genotypes produced by hybridization or not. Identical recombinant *Pokey* alleles in different genotypes may be the product of the activity of these recombinant *Pokey* alleles. Future studies should test the activity of recombinant *Pokey* alleles to assess if recombination events have a significant effect on the transposition capacity of *Pokey*.

4.7 Acknowledgments

We thank C. Marinone for providing clones from Argentina, M. Cristescu and C. Caceres for providing some *Daphnia pulex* S. Markovà for providing some European *Daphnia* sp. and D. Innes for providing *D. pulex* clones from Ontario. We thank three anonymous reviewers for their useful comments and corrections that improve the quality of this paper. This work was supported by a Natural Sciences and Engineering research Council (NSERC) research grants and by a Canadian Foundation for Innovation (CFI) equipment grants to France Dufresne

Tableau 4-5 Profiles of *Pokey* insertions highlighted by the “TE display” technique
 Numbers at the top of the columns are length of amplified segments

isolates	160	163	166	169	172	175	179	185	190	194	199	203	209	212	215	218	221	226	229	233	238	243	246
PC3-SF-5	1	0	0	0	0	1	1	0	1	0	0	1	1	1	0	0	1	1	0	0	1	0	0
MI3-SF-2	0	1	0	1	0	0	1	0	1	1	0	0	1	0	0	1	1	0	0	0	1	1	0
PC3-SF-4	1	0	0	0	0	1	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1	0	0
PC3-SF-6	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	1	0	1	0	1	0	0
PX2-SF-1	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	1	1	0	0	1	0	1
PX2-SF-2	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0	0	0	0	0	1	1	0	0
PX2-SF-3	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	1	1	0	1	0	1
PX2-SF-4	0	0	1	0	1	0	1	0	1	0	0	0	1	1	0	0	1	0	0	0	1	0	1
PX2-SF-5	0	0	0	0	0	1	0	0	0	1	0	1	1	1	0	0	1	1	0	0	1	0	1
PX2-SF-6	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	1	1	0	0	1	1	1
PX2-SF-7	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	1	0	1
PX3-SF-2	1	0	0	0	0	1	1	1	1	1	0	0	1	1	0	0	1	0	1	0	1	1	0
PX2-SF-10	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	1	1	1	0	1	0	1
PX2-SF-8	0	0	1	0	1	1	1	0	1	0	0	1	1	1	0	0	1	1	0	0	1	0	0
PX2-SF-11	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	1	0
PC3-SF-3	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	0	1	0	1	0	1	0	0
PX3-SF-1	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	1	0	1	0	1	0	0
PC3-SF-1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	1	0	1	1	0
PC3-SF-2	0	0	0	0	0	1	1	1	1	0	0	0	0	1	0	0	1	1	1	0	1	0	0
PX2-SF-9	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	1
TE3-SS-1	0	0	1	0	0	1	1	0	1	0	0	0	0	1	0	1	1	1	0	1	1	0	1
TE3-UN-1	0	0	0	1	0	1	0	1	1	0	0	1	0	1	0	0	1	0	0	0	1	1	0
TE3-UN-2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
TE3-UN-3	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	0
TE2-UN-1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1
TE2-UN-2	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	1	1	0	0	1	0	1
TE2-UN-3	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0

isolates	253	257	260	263	266	269	271	276	279	282	286	289	295	303	306	310	313	317	321	328	332	335	339
PC3-SF-5	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0
MI3-SF-2	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
PC3-SF-4	0	1	0	1	0	1	0	1	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0
PC3-SF-6	0	1	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0
PX2-SF-1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0
PX2-SF-2	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0
PX2-SF-3	0	1	0	0	1	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0
PX2-SF-4	0	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	1	0	1	1	0	0
PX2-SF-5	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0
PX2-SF-6	0	1	0	1	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0
PX2-SF-7	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0
PX3-SF-2	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
PX2-SF-10	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0
PX2-SF-8	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
PX2-SF-11	0	1	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0	0	0	1	1	0	1
PC3-SF-3	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0
PX3-SF-1	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	1
PC3-SF-1	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0
PC3-SF-2	1	1	0	0	1	0	1	1	0	1	0	0	0	0	1	1	1	0	1	1	0	0	0
PX2-SF-9	1	1	0	1	0	0	0	1	1	0	0	0	0	1	0	1	1	0	0	1	0	0	0
TE3-SS-1	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	1
TE3-UN-1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
TE3-UN-2	0	1	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
TE3-UN-3	0	1	0	0	0	1	0	1	0	1	0	0	0	0	1	1	0	1	0	0	1	0	0
TE2-UN-1	0	1	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
TE2-UN-2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
TE2-UN-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

isolafes	342	345	351	360	372	378	381	390	393	398	401	404	406	408	411	415	417	421	424	432	438	441	447
PC3-SF-5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
MI3-SF-2	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
PC3-SF-4	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PC3-SF-6	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
PX2-SF-1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PX2-SF-2	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PX2-SF-3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
PX2-SF-4	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PX2-SF-5	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0
PX2-SF-6	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0
PX2-SF-7	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
PX3-SF-2	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0
PX2-SF-10	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PX2-SF-8	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PX2-SF-11	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
PC3-SF-3	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PX3-SF-1	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0
PC3-SF-1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
PC3-SF-2	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
PX2-SF-9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
TE3-SS-1	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0
TE3-UN-1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
TE3-UN-2	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
TE3-UN-3	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
TE2-UN-1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TE2-UN-2	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
TE2-UN-3	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0

isolates	451	455	463	468	475	478	482	487	497	500	515	524	528	536	561	570	582	586	589	626	647	802
PC3-SF-5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0
MI3-SF-2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PC3-SF-4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PC3-SF-6	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0
PX2-SF-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PX2-SF-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PX2-SF-3	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0
PX2-SF-4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
PX2-SF-5	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
PX2-SF-6	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
PX2-SF-7	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
PX3-SF-2	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0
PX2-SF-10	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PX2-SF-8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PX2-SF-11	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	1	0	0	0	0	0	0
PC3-SF-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PX3-SF-1	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0
PC3-SF-1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
PC3-SF-2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
PX2-SF-9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
TE3-SS-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
TE3-UN-1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TE3-UN-2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
TE3-UN-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TE2-UN-1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
TE2-UN-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TE2-UN-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Chapitre V: Discussion générale

L'objectif final de cette thèse de doctorat était de tester l'effet du niveau de ploïdie sur le nombre d'insertions d'un élément transposable chez une espèce animale. Afin de répondre à cet objectif, nous avons estimé le nombre d'insertions différentes d'un transposon de classe II, *Pokey*, dans les génomes de lignées hybrides diploïdes et polyploïdes du complexe d'espèce *Daphnia pulex*. Ce petit crustacé d'eau douce a été choisi car des individus diploïdes hybrides et polyploïdes hybrides se rencontrent dans différentes populations. Les lignées hybrides, qu'elles soient diploïdes ou polyploïdes ont une reproduction strictement asexuées (Hebert et Crease, 1983 ; Hebert *et al.*, 1993b). Elles peuvent être maintenues en élevage aisément en laboratoire sans que leur bagage génomique s'altère significativement au cours des générations.

Avant de pouvoir quantifier le nombre d'insertions du transposon *Pokey* à l'intérieur de différents isolats (individus mis en élevage) du complexe d'espèces *Daphnia pulex*, il a fallu caractériser les génomes de ces isolats. Afin de ne pas confondre l'effet de l'hybridation sur le nombre d'insertions de *Pokey* dans les génomes diploïdes et polyploïdes, il a fallu comparer des isolats hybrides entre eux. Comme l'hybridation pourrait avoir un effet drastique sur la dynamique et le nombre d'insertions de *Pokey* chez les individus diploïdes (Ainouche *et al.*, 2009 ; Fontdevila, 2005 ; Librado et Rozas, 2009 ; Petrov *et al.*, 1995 ; Shan *et al.*, 2005), les individus non-hybrides ne doivent pas être inclus dans l'analyse. Une bonne caractérisation du niveau de ploïdie et de l'origine des espèces parentes chez les hybrides est importante pour tester l'effet du niveau de ploïdie. Cette caractérisation a été l'objectif des deuxième et troisième chapitres de ma thèse de doctorat.

L'origine des génomes parentaux est aussi utile pour comprendre la

dynamique des transposons dans les génomes hybrides (diploïdes ou polyploïdes). Des éléments transposables ayant évolué dans des espèces différentes ont pu avoir divergé et ont pu acquérir des caractéristiques différentes (autonomie, taux de transposition, sites d'insertions préférentielles...). De plus lors de l'hybridation, il est possible qu'un des individus parents soit dépourvu de l'élément transposable considéré. Un individu hybride diploïde, dont un des parents serait dépourvu d'élément transposable, pourrait ne pas avoir le même fardeau d'éléments transposables qu'un individu dont les deux parents possèdent des éléments transposables. Les individus testés étaient des individus échantillonnés en nature. L'origine et le fardeau en éléments transposables des individus parents étaient inconnus. Des allèles de *Pokey* insérés dans les génomes d'isolats hybrides diploïdes et polyploïdes ont donc été amplifiés et séquencés pour estimer leur origine et évaluer si différents allèles pouvaient être présents à l'intérieur de ces génomes. En utilisant la technique de «TE display» (Wright *et al.*, 2001) mise au point par Valizadeh et Crease (2008) pour l'élément *Pokey* chez les daphnies, nous avons estimé le nombre de signaux d'insertions de ce transposon de classe II dans différents génomes hybrides diploïdes et polyploïdes. L'analyse des séquences de *Pokey* et le nombre d'insertions de ceux-ci sont l'objet du quatrième chapitre de ma thèse de doctorat. Les résultats de ces trois chapitres seront discutés et mis en perspective dans le présent chapitre. De plus, les données accumulées lors de cette thèse de doctorat ont permis de mieux comprendre le système à l'étude sur des points connexes. Ces données seront présentées et brièvement discutées aussi dans ce chapitre. Ce cinquième chapitre est la place de choix pour exposer différentes idées et hypothèses qui découlent du travail ici rapporté. J'espère qu'elles seront explorées et testées dans de futures études. Enfin une conclusion générale ouvrant la voie, je l'espère, à différentes études dans l'analyse écologique du génome clôturera cette thèse.

5.1 Tailles de génome et niveaux de ploïdie d'isolats du complexe *Daphnia pulex*

Le deuxième chapitre de ma thèse s'est focalisé sur la variation du niveau de ploïdie entre différents isolats représentant 6 des 7 différentes espèces du complexe *Daphnia pulex*. Des individus représentant des lignées hybrides et non-hybrides ont été analysés. Afin d'évaluer le niveau de ploïdie, deux techniques différentes ont été utilisées: l'analyse des génotypes multilocus microsatellites a permis une première estimation du niveau de ploïdie de différents individus et l'analyse par cytométrie de flux d'individus de différentes lignées a permis de quantifier le contenu en ADN dans les cellules de lignées diploïdes et polyploïdes et de confirmer les niveaux de ploïdie évalués grâce aux microsatellites.

5.1.1 Polyploïdie chez les isolats du complexe *Daphnia pulex*

Pour évaluer le niveau de ploïdie, nous avons génotypé 49 isolats provenant de différentes localisations géographiques, tempérées et subarctiques, à neuf locus microsatellites. La présence de plus de deux allèles au niveau d'au moins un locus pour un isolat donné était considérée comme un indice que celui-ci était polyploïde. Les isolats étudiés n'avaient pas plus de trois allèles au locus considérés (voire Chapitre II). Seize des 49 isolats étudiés présentaient des génotypes avec plus de deux allèles au moins pour un locus. Le nombre maximal d'allèles par locus de ces seize isolats potentiellement polyploïdes n'excédait pas trois. Les isolats polyploïdes testés dans cette étude sont donc au minimum triploïdes. Toutefois, le nombre d'allèles par locus ne constitue pas une preuve irréfutable du niveau de ploïdie d'un isolat. En effet, un isolat qui a deux allèles par locus peut avoir un niveau de ploïdie supérieur à deux. Si la variabilité génétique n'est pas assez suffisante pour posséder plus de deux allèles par locus, le niveau de ploïdie est sous-estimé. Un isolat ayant

trois allèles à un de ces locus peut ne pas être polyploïde. Cet isolat peut être issu d'une fécondation entre un gamète haploïde et un gamète non-balancé (i.e. gamète possédant deux copies d'un de ses chromosomes). Nous avons donc estimé la quantité d'ADN présent dans les cellules de ces 49 isolats.

5.1.2 Tailles du génome chez les espèces du complexe *Daphnia pulex*

Différentes études ont estimé, par différentes méthodes, la taille de génome d'espèces du genre *Daphnia* (Beaton et Hebert, 1988 ; Korpelainen, Ketola et Hietala, 1997 ; Rasch, 1985). Ces différentes estimations étaient pour la plupart différentes de la taille estimée par le séquençage du génome complet (Colbourne *et al.*, 2011). Les "C-value" (quantité d'ADN contenue dans le génome estimée en picogramme) estimées allaient de 0,24 à 0,37 pg. Seule la valeur obtenue en densitométrie à l'aide de la coloration de Feulgen par Rasch (1985) se rapprochait de la valeur estimée par le séquençage du génome. Une adaptation du protocole de Rees *et al.* (2007) utilisant la technique de cytométrie de flux nous a permis, en utilisant seulement 5 individus entiers et un échantillon de sang de poulet (*Gallus gallus domesticus*) comme référence (Vergilino, Belzile et Dufresne, 2009), d'évaluer la taille de génome de 49 isolats provenant de 6 espèces nominales du complexe *Daphnia pulex*. La taille de génome moyenne des isolats de l'espèce *Daphnia pulex* évaluée par cette technique ($0,23 \pm 0,02$ pg) est la plus proche jusqu'à présent de celle estimée par le projet de séquençage du génome de *Daphnia pulex* (0.2035 pg; Cristescu *et al.*, 2006). Les coefficients de variations des tailles de génomes haploïdes sont assez faibles. La méthode utilisée semble alors être fiable et donner une bonne estimation des tailles de génomes chez les daphnies du complexe *pulex*. Il y a une relativement grande variation dans les tailles de génomes répertoriées pour les genres *Daphnia* et *Ctenodaphnia* dans la littérature (Table 5-1). Cette variation semble due à des problèmes taxonomiques (Colbourne *et al.*, 1998 ; Colbourne et Hebert, 1996 ; Marková *et al.*, 2007) et à l'absence de consensus dans

les méthodes utilisées pour estimer la quantité d'ADN dans les noyaux (Beaton et Hebert, 1988 ; Korpelainen, Ketola et Hietala, 1997 ; Rasch, 1985). Il serait donc intéressant d'utiliser, systématiquement, la technique décrite au chapitre II pour différentes espèces du genre *Daphnia* afin de comparer les tailles de génomes des différentes espèces de ce genre.

Tableau 5-1 Tailles de génomes estimées pour quatre espèces appartenant aux genres *Daphnia* et *Ctenodaphnia*

Espèces	C-values (pg)	technique	Type cellulaire	Espèce(s) standard(s)	références
<i>D. pulex</i>	0,24	Feulgen densitometry	corps entier	<i>G. domesticus</i> <i>O. mykiss</i>	Rasch (1985)
<i>D. pulex</i>	0,34	Feulgen densitometry	exopodite	<i>S. cerevisiae</i> <i>P. sphenops</i> <i>X. laevis</i>	Beaton et Hebert (1989)
<i>D. pulex</i> (Européenne?)	0,37	Cytométrie en flux	corps entier	<i>G. domesticus</i>	Korpelainen et al. (1997)
<i>D. pulex</i>	0,23	Cytométrie en flux	corps entier	<i>G. domesticus</i>	Vergilino et al. (2009)
<i>D. pulicaria</i>	0,40	Feulgen densitometry	exopodite	<i>D. pulex</i> from Beaton et Hebert (1989)	Beaton (1995)
<i>D. pulicaria</i>	0,24	Cytométrie en flux	corps entier	<i>G. domesticus</i>	Vergilino et al. (2009)
<i>D. tenebrosa</i>	0,53	Feulgen densitometry	exopodite	<i>D. pulex</i> from Beaton et Hebert (1989)	Beaton (1995)
<i>D. tenebrosa</i>	0,58	Feulgen densitometry	exopodite	<i>D. pulex</i> from Beaton et Hebert (1989)	Dufresne et Hebert (1995)
<i>D. tenebrosa</i>	0,29	Cytométrie en flux	corps entier	<i>G. domesticus</i>	Vergilino et al. (2009)
<i>C. magna</i>	0,24	Feulgen densitometry	corps entier	<i>G. domesticus</i> <i>O. mykiss</i>	Rasch (1985)
<i>C. magna</i>	0,47	Feulgen densitometry	exopodite	<i>S. cerevisiae</i> <i>P. sphenops</i> <i>X. laevis</i>	Beaton et Hebert (1989)
<i>C. magna</i>	0,48	Feulgen densitometry	exopodite	<i>D. pulex</i> from Beaton et Hebert (1989)	Beaton (1995)
<i>C. magna</i>	0,40	Cytométrie en flux	corps entier	<i>G. domesticus</i>	Korpelainen et al. (1997)

5.1.3 Apport de la méthode

La comparaison des résultats des deux techniques citées précédemment permet d'évaluer le niveau de ploïdie de nombreux isolats. Ces résultats ont permis de mieux caractériser les populations polyploïdes du genre *Daphnia*. Le principal résultat de cette étude a été plutôt inattendu. Des études ultérieures (Adamowicz *et al.*, 2002 ; Aguilera *et al.*, 2007 ; Beaton et Hebert, 1988 ; Dufresne et Hebert, 1994, 1997 ; Mergeay *et al.*, 2008) utilisant différentes techniques proposaient que les individus des populations polyploïdes possédaient 4 ensembles de chromosomes (i.e tétraploïdes). Or, le travail présenté au chapitre II montre que la plupart des isolats testés sont triploïdes. Les résultats supplémentaires présentés au chapitre III étayaient aussi cette dernière proposition. L'idée prédominante était que les individus tétraploïdes se formaient par doublement de génomes après hybridation (Dufresne et Hebert, 1994 ; Mergeay *et al.*, 2008). Le fait que les isolats polyploïdes testés possèdent trois ensembles de chromosomes, et non quatre, modifie le modèle de formation des polyploïdes dans le complexe *pulex*. Il semble que l'hypothèse la plus parcimonieuse quant à la formation des polyploïdes mette en jeu la fécondation de gamètes non-réduits, et donc diploïdes, par des gamètes réduits (i.e. haploïdes). Il existe des indices que des gamètes mâles ou femelles non-réduits puissent être produits par des individus des espèces *D. pulex* et *D. pulicaria*. Chez ces deux espèces, des populations strictement clonales ont été observées en Amérique du Nord (Hebert et Finston, 2001b). Ces populations clonales sont le résultat soit d'évènements d'hybridation soit de la ségrégation d'un gène muté causant la suppression de la méiose chez les femelles et transmis par les mâles (Innes et Hebert, 1988 ; Innes et Singleton, 2000 ; Lynch *et al.*, 2008). Des œufs diploïdes peuvent donc être produits par des femelles hybrides ou non. Il est donc possible que des individus triploïdes aient été produits en nature via la fécondation d'un œuf non-réduit (i.e. diploïde) par du sperme haploïde. De plus, lors d'une expérience visant à identifier et localiser le ou les gènes impliqués dans la suppression de méiose chez *D. pulex*, il a été montré que des gamètes diploïdes pouvaient être produits par des individus mâles (Lynch *et al.*, 2008). Le croisement d'une femelle

sexuée avec un mâle issu d'une femelle se reproduisant de manière clonale a pu produire un individu triploïde (Lynch *et al.*, 2008). Il est donc possible que des individus triploïdes aient été produits en nature via la fécondation d'un ovule haploïde par du sperme diploïde.

La technique de cytométrie en flux a permis de confirmer le niveau de ploïdie de 44 des 49 isolats mis en évidence par leur génotype multilocus. La quantité d'ADN nucléaire de certains isolats (K95, Clone B, Clone G et Clone H) ne concordait pas avec le niveau de ploïdie évalué avec l'analyse des génotypes multilocus. Ces discordances pourraient être dues soit à une sous-estimation du niveau de ploïdie (cf. K95, Clone G et Clone H) par l'analyse des microsatellites comme il a été mentionné précédemment. Une quantité d'ADN nucléaire plus petite qu'attendue relativement au niveau de ploïdie tel qu'il a été observé pour le Clone B peut être due à des évènements de diploïdisation (Li, 1980) ou d'aneuploïdies. Il a été observé la formation de gamètes non-balancés par des mâles supposément porteur du gène suppresseur de méiose (Lynch *et al.*, 2008). Des individus se retrouvant en nature pourraient être issus d'un croisement entre un gamète balancé et un gamète non-balancé. Ces individus se reproduisant d'une manière strictement clonale (i.e. apomictes), leur fitness ne devrait pas être altérée par des phénomènes de disjonctions inégales lors de la méiose comme cela pourrait être le cas pour des individus aneuploïdes se reproduisant de manière sexuée.

L'analyse d'isolats appartenant à différentes espèces du complexe *pulex* a mis en évidence une différence significative entre les valeurs moyennes de taille de génome haploïde (i.e. C-values) des espèces du groupe *pulicaria* et celles du groupe *tenebrosa* (Tableau 2-3). Les C-values moyennes des isolats diploïdes des espèces du groupe *pulicaria* sont systématiquement plus petites que celles des isolats diploïdes des espèces du groupe *tenebrosa*. Les causes de cette différence de tailles de génomes haploïdes ne sont pas connues et restent à être étudiées. Une variation dans la quantité d'ADN non-codant (éléments transposables, introns, séquences répétées...) est possible (Lynch, 2007). Des variations dans le nombre

d'insertions d'éléments transposables (Schaack *et al.*, 2010b ; Valizadeh et Crease, 2008), d'introns (Li *et al.*, 2009 ; Omilian, Scofield et Lynch, 2008) ou de répétitions de gènes codant pour l'ARN ribosomal existent dans les populations de *D. pulex* (Crease, 1995 ; Glass, Moszczyńska et Crease, 2008). Les variations de taille de génome entre espèces peuvent aussi être le résultat d'un (ou de plusieurs) évènement(s) ancien(s) de polyploïdisation dans le groupe *tenebrosa* suivi d'évènement(s) de diploïdisation (Li, 1980). Le génome d'un isolat de *Daphnia pulex* a été entièrement séquencé (Colbourne et al. soumis). Les génomes d'autres isolats provenant de cette même espèce mais aussi de différentes espèces du complexe *pulex* (F. Dufresne, communication personnelle) seront séquencés prochainement. Une analyse de génomique comparative se basant sur le nombre de gènes et le nombre de structures non-codantes, telles que les introns et les éléments transposables, pourrait permettre de tester ces scénarios.

5.2 Hybridation, relations génétiques et phylogénétiques entre les espèces du complexe *Daphnia pulex*

La deuxième partie de ma thèse s'est attelée à clarifier les liens génétiques et phylogénétiques entre des isolats de 7 espèces nominales du complexe *Daphnia pulex*. Les variations de plusieurs marqueurs nucléaires (locus microsatellites et un gène codant une protéine) et un marqueur mitochondrial ont été analysées.

5.2.1 Analyse des microsatellites

Cent-cinq isolats appartenant à 7 espèces du complexe *Daphnia pulex* ont été génotypés à neuf locus microsatellites. Différentes analyses statistiques (des analyses de coordonnées principales, K-means et un réseau phylogénétique) ont été

effectuées afin de connaître les relations génétiques entre ces différentes espèces et les supposés hybrides entre celles-ci. Des populations européennes et Nord-américaines de l'espèce *D. pulicaria* ont été analysées. Les analyses de groupements (K-means) identifient les hybrides diploïdes entre *D. pulex* et *D. pulicaria* (Figure 3-2). Ces hybrides ont des origines multiples comme le montre l'analyse de réseau phylogénétique (Figure 3-3) effectuée sur le même jeu de données. Les individus polyploïdes hybrides entre *D. pulex* et *D. pulicaria* sont séparés dans deux groupes. La plupart de polyploïdes qui possèdent un haplotype mitochondrial *D. pulex* (PanPX) se regroupent avec les isolats diploïdes hybrides. Les isolats ayant des haplotypes mitochondriaux *D. pulicaria* (WPC, EPC et PPC) ou *D. middendorffiana* (MIDD) se regroupent avec les isolats diploïdes de l'espèce *D. pulicaria*. Les isolats des populations européennes et Nord-américaines de *D. pulicaria* sont clairement séparés les uns des autres. Cette analyse confirme que ces populations représentent deux entités taxonomiques différentes comme des analyses se basant sur des haplotypes mitochondriaux le laissent supposer (Colbourne *et al.*, 1998 ; Marková *et al.*, 2007). Les isolats des populations de *D. arenata* sont regroupés avec des isolats de populations du groupe *tenebrosa*. Dû à la possession d'haplotypes mitochondriaux diagnostiques de l'espèce de *D. pulex*, *D. arenata* est considérée comme une population de *D. pulex* dans certaines publications (Li *et al.*, 2009 ; Omilian et Lynch, 2009 ; Omilian, Scofield et Lynch, 2008).

L'incongruence entre les génotypes microsatellites et l'haplotype mitochondrial pourrait être dû à trois scénarios hypothétiques différents. Les marqueurs microsatellites sont connus pour évoluer de manière neutre dans la plupart des cas (Jarne et Lagoda, 1996). Les populations de *D. arenata* sont endémiques de l'Oregon et auraient pu être isolées lors de la glaciation du Pléistocène (Straughan et Lehman, 2000). Des processus stochastiques auraient pu avoir une grande importance dans la fixation de certains allèles dans ces populations. Par simple hasard et/ou par homoplasie, les génotypes microsatellites des isolats de *D. arenata* auraient pu ressembler à ceux de populations ayant divergées lors du Pliocène telles

que des populations appartenant aux espèces du groupe *tenebrosa* (Colbourne *et al.*, 1998). Il est possible de tester cette hypothèse en séquençant les allèles analysés ici et présents chez *D. arenata*, *D. pulex* et « *D. pulicaria* européenne ». Si les séquences flanquantes des allèles de *D. arenata* se rapprochent plus des séquences homologues provenant d'individus « *D. pulicaria* européenne » que de celles des allèles de *D. pulex* panarctique, l'hypothèse d'homoplasie pourrait être rejetée. Les deux hypothèses suivantes pourraient alors être testées. Le maintien d'une certaine quantité de polymorphisme ancestral hérité de l'ancêtre commun du complexe *pulex* représenterait un deuxième scénario possible (Pamilo et Nei, 1988 ; Wu, 1991). Les espèces du complexe sont relativement jeunes d'un point de vue évolutif. Les espèces à l'intérieur des groupes *tenebrosa* et *pulicaria* auraient divergé au cours du Pléistocène (de 1,2 à 2,2 millions d'années avant notre ère; Colbourne *et al.*, 1998 ; Dufresne et Hebert, 1997 ; Weider et Hobaek, 2003) et ces deux groupes auraient divergé lors du Pliocène (à peu près 3 millions d'années avant notre ère; Colbourne *et al.*, 1998). Il est donc possible qu'une certaine quantité de polymorphisme soit conservée dans les différentes espèces du complexe. Des espèces différentes pourraient partager en partie un bagage génétique commun. Certaines populations comme celles de *D. arenata* auraient conservé des allèles communs avec des espèces ayant divergé antérieurement alors que d'autres espèces comme *D. pulex* et *D. pulicaria* auraient perdues ces allèles. Un troisième scénario impliquerait des phénomènes d'hybridations et/ou d'introgessions entre une population de *D. pulex* (donnant le génome mitochondrial) et une population, peut-être éteinte maintenant, ayant des affinités phylogénétiques avec les populations de *D. pulicaria* Européennes. Ces deux populations auraient pu se rencontrer et produire des hybrides dans le refuge de la Béringie et migrer vers l'Oregon lors de la déglaciation après le Pleistocène. La Béringie étant un des principaux refuges glaciaires lors des événements du Pléistocène (Weider et Hobaek, 2003), des groupes d'espèces ayant divergé lors de la période de glaciation précédente (Pliocène) auraient pu se retrouver dans la même localité et échanger du matériel génétique pour produire une population hybride voire introgressée. Les

deux populations auraient aussi pu se rencontrer *in situ* en Oregon. En effet, il semble que l'Oregon soit un refuge probable pour certaines populations Nord-Américaines du complexe *pulex* lors des ères glaciaires du Pléistocène (Straughan et Lehman, 2000). De plus, des phénomènes de glaciations et déglaciations de lacs lors des cycles glaciaires du Pléistocène auraient pu provoquer l'isolement et la rencontre de populations divergentes. Les seules données microsatellites ne peuvent pas nous permettre de tester lesquelles de ces deux dernières hypothèses (« maintien du polymorphisme ancestral » ou « hybridation ») sont à rejeter. D'autres observations (voire plus bas) sur différents marqueurs nucléaires permettent de rejeter certaines de ces hypothèses et tendent à montrer que *D. arenata* aurait un statut taxonomique bien particulier au sein du complexe *pulex*.

5.2.2 Conflits « gène nucléaire-haplotype mitochondrial »

Les phylogénies comparées du gène nucléaire (codant pour la sous-unité protéique Rab4 de la famille des GTPases) et des séquences partielles du gène *NADH5* de 51 isolats, représentant les différentes espèces du complexe *pulex* et les différents niveaux de ploïdie qui s'y rencontrent sont non concordantes (Figures 3-1 et 3-4). Cette absence de concordance peut être due soit à des phénomènes d'hybridation soit à des phénomènes de maintien de polymorphisme dans les différentes espèces du complexe. Un test utilisant une analyse de coalescence (Kingman, 1982) a permis de rejeter l'hypothèse que le phénomène de maintien du polymorphisme était le seul responsable de l'absence de concordance entre les phylogénies mitochondriales et nucléaires. Il est donc permis de penser que des phénomènes d'hybridations aient provoqué le patron observé. Toutefois il semble comme nous avons pu le voir dans l'analyse de microsatellites que du polymorphisme ancestral ait été retenu dans les différentes espèces. Par exemple, la quasi-totalité des séquences amplifiées des isolats des espèces *D. pulicaria* européennes et *D. tenebrosa* se regroupent avec les séquences de l'espèce *D.*

pulicaria Nord-Américaine. Toutefois les analyses microsatellites montrent qu'il n'y a pas ou peu de flux génique entre ces différentes espèces (Figures 3-2 et 3-5). Il est possible que la similarité entre les allèles du gène *Rab4* soient dues à une absence de variabilité des séquences entre espèces divergentes et non pas à des phénomènes d'hybridation. Il est intéressant d'observer que la plupart des individus hybrides entre *D. pulex* et *D. pulicaria* (autant diploïdes que triploïdes) ne présentent pas deux allèles. La plupart des hybrides sont homozygotes pour un allèle similaire aux allèles amplifiés chez les isolats de l'espèce *D. pulicaria* (Figures 3-1 B et 3-4 B). Un biais dans l'échantillonnage des isolats hybrides ne portant que l'allèle *pulicaria* pourrait expliquer ce patron. Toutefois, considérant le nombre important d'isolats hybrides échantillonnés, il est peu probable que nous n'ayons échantillonné que des hybrides avec des allèles *pulicaria* seulement par chance. Une hybridation suivie d'une conversion génique, pourrait expliquer ce patron (Hillis *et al.*, 1991 ; Kovarik *et al.*, 2004 ; Sang et Zhang, 1999). Des réarrangements génomiques (par exemple une mise sous silence de gènes, translocations, conversions géniques, etc.) suivent fréquemment les événements d'hybridations (Soltis et Soltis, 1995, 1999). Des pressions de sélections différentes rencontrées par *D. pulex* et les autres membres du complexe *pulex* pourraient expliquer ce patron. En effet, la protéine Rab4 semble être impliquée dans le trafic membranaire au niveau de voies de biosynthèse et d'endocytose (Van Der Sluijs *et al.*, 1991), Une sélection directionnelle aurait pu conduire à la fixation d'allèles spécifiques à *D. pulex* et ne ressemblant pas aux allèles présents chez les populations de *D. pulicaria* et les populations hybrides. Une étude du polymorphisme synonyme et non-synonyme du gène *Rab4* dans le complexe *pulex* permettrait de tester cette hypothèse.

5.2.3 Estimation de la part des différents génomes parentaux dans les génomes des hybrides

Les analyses du troisième chapitre de cette thèse ont examiné les relations évolutives entre les différentes espèces du complexe *pulex*. Notamment l'analyse des microsatellites a mis en relief la position particulière de l'espèce *D. arenata* au sein du complexe *pulex*, la différenciation génétique entre *D. pulicaria* Nord-Américaine et *D. pulicaria* européenne et les origines multiples de divers isolats hybrides, diploïdes ou polyploïdes. Les analyses de groupements faites à partir des données microsatellites (Figure 3-2) montrent que l'hybridation et une certaine introgression existent entre des espèces du complexe *pulex*. À part pour un isolat polyploïde introgressé avec un haplotype mitochondrial *D. tenebrosa* et un génome nucléaire de *D. pulex*, des analyses du gène nucléaire *Rab4* et des microsatellites n'ont pas été en mesure de préciser si les autres isolats polyploïdes de *D. tenebrosa* sont de nature hybrides. Soit ces isolats sont issus de croisement entre individus de la même espèce soit les traces des événements d'hybridation se sont estompées au cours de l'évolution.

Les isolats polyploïdes du complexe *pulex* sont pour la plupart triploïdes et seraient issus de fécondations de gamètes non-réduits par des gamètes réduits ou inversement. L'analyse de K-means a permis de grouper les isolats triploïdes possédant des haplotypes mitochondriaux *pulicaria* avec les individus diploïdes de cette espèce et les hybrides diploïdes et triploïdes possédant des haplotypes mitochondriaux *pulex*. Les isolats triploïdes possédant l'haplotype mitochondrial d'une espèce sembleraient posséder deux copies de génome provenant de la lignée maternelle et une copie provenant de la lignée paternelle. Ce patron est attendu si les gamètes maternels sont non-réduits. Toutefois ces analyses sont qualitatives et demandent un assez grand nombre de données et des analyses statistiques assez complexes. De plus, certains isolats triploïdes ne se groupent pas avec les individus diploïdes de même haplotype mitochondrial (Figure 3-2). Il est possible que des mâles *Daphnia* produisent des gamètes non-réduits (Lynch *et al.*, 2008). Il serait

intéressant de mettre au point une technique simple de quantification des différents génomes *pulex* et *pulicaria* pour les triploïdes hybrides du groupe *pulicaria*. Par exemple, une méthode de PCR quantitative (qPCR) en utilisant les amorces spécifiques du gène *LDH-A* utilisées dans le chapitre III permettrait de quantifier le nombre d'allèles S et le nombre d'allèles F (diagnostiques des espèces *D. pulex* et *D. pulicaria* respectivement) dans chaque génome triploïde.

5.3 Évolution du gène *Rab4* dans le complexe *pulex*

Un article de Omilian et collaborateurs publié en 2008 dans la revue *Proceedings of the National Academy of Science of the United State of America* traitant du gain et de la perte d'introns dans le gène *Rab4* chez *D. pulex* et *D. pulicaria* (Omilian, Scofield et Lynch, 2008) a permis de comprendre le patron phylogénétique que nous observions avec un jeu de données réduit. Cet article a aussi attiré notre curiosité sur l'évolution des introns à spliceosome. Omilian et collaborateurs (2008) ont montré que deux introns (2a et 2b) se sont insérés au niveau de l'exon 2, le fractionnant en deux, à l'intérieur du gène *Rab4* dans la même population de *D. pulex* (une population de *D. arenata* pour être exact) en Oregon. Ces insertions, se faisant aux niveaux de sites nucléotidiques voisins, semblent indépendantes l'une de l'autre. Ces introns n'étaient pas rencontrés dans d'autres populations de *D. pulex* et de *D. pulicaria* échantillonnées. Un article comparant les patrons de présence/absence d'introns dans de multiples génomes provenant de diverses populations d'Amérique du Nord de *D. pulex* a fait suite (Li *et al.*, 2009). Pour Omilian et ses collaborateurs (2008), la fréquence importante des allèles portant ces introns potentiellement délétères dans la population d'Oregon et leur absence dans les nombreuses populations échantillonnées est due à un relâchement de la sélection suite à une réduction de la taille effective de cette population.

L'analyse effectuée sur des isolats provenant de différentes régions du globe et considérant toutes les espèces connues du complexe *pulex* a permis de mettre en évidence un autre intron (2b'; Figure 5-1). Cet intron s'insère au même site que l'intron 2b rencontré précédemment, il diffère de l'intron 2b d'une insertion de 5 paires de bases. Il se rencontre dans des allèles échantillonnés dans des isolats des espèces divergentes de *D. pulex* européennes (DP0076 et DP0081) et *D. pulicaria* européennes (DP0077) échantillonnées en Suède (Tableau 3-1). Des allèles portant l'intron 2b ont aussi été amplifiés dans des isolats diploïdes et polyploïdes hybrides provenant d'Ontario (Disp1-21), du Manitoba (A35-200), du Nord du Québec (K92) et d'Italie (DP0085) (Figure 5-1). Les inférences phylogénétiques Bayésiennes effectuées avec le logiciel MrBayes (Huelsenbeck et Ronquist, 2005) sur 101 séquences montrent que les allèles portant l'intron 2b se regroupent qu'ils aient été amplifiés d'individus d'Oregon, du Canada ou d'Europe. Les deux arbres (sans modèle ou avec un modèle HKY) sont discordants concernant le positionnement de l'allèle portant l'intron 2a. Dans l'arbre construit avec le modèle HKY (Figure 5-1A), les allèles portant l'intron 2a constituent un groupe externe à tous les allèles rencontrés dans le complexe *pulex*. Dans l'arbre construit sans modèle (Figure 5-1B), les allèles portant l'intron 2a représentent un groupe frère des allèles retrouvés chez les individus diploïdes *D. pulex*. L'arbre construit avec le modèle HKY semble mieux représenter les données que l'arbre construit sans modèle (déviations standards : 0.0031 vs. 0.0047 respectivement). Il est peu probable que l'intron 2b se soit inséré 6 fois indépendamment dans des allèles identiques portés par des individus génotypiquement différents. L'allèle portant l'intron 2b était donc présent dans les populations du complexe *pulex* au niveau du Pléistocène. Il est possible que les allèles portant les introns 2a et 2b' soient des reliquats de polymorphisme ancestral. L'allèle portant intron 2a aurait été fixé par dérive dans la population de *pulex* en Oregon mais son origine semble bien plus ancienne que le laissait supposé l'étude d'Omilian et ses collaborateurs qui ne se basait que sur trois espèces (Omilian, Scofield et Lynch, 2008).

En utilisant des données de présence/absence d'introns dans d'autres gènes,

Li et ses collaborateurs (Li *et al.*, 2009) évaluent ainsi un taux d'apparition des introns à spliceosome de l'ordre de $1,2 \times 10^{-5}$ par site codant par millions d'années et par génération. Ce taux est bien supérieur au taux de moins de 4×10^{-6} par millions d'années et par génération préalablement évalué par différentes études (Lynch, 2007 ; Roy et Gilbert, 2005). Toutefois la présence d'introns surnuméraires se retrouve principalement dans la population d'Oregon échantillonné par Omilian et ses collaborateurs (2008). Sachant que cette population correspond à la population de *D. arenata*, le patron d'insertions des introns pourrait être un artefact dû à une introgression d'un génome nucléaire d'une autre espèce. Selon notre analyse et les résultats obtenus avec les microsatellites, le génome nucléaire de *D. arenata* aurait des similarités avec les génomes nucléaires de populations de *D. pulicaria* européennes ou *D. pulex* européennes. Il est donc possible que l'espèce *D. arenata* pour laquelle un individu a servi au séquençage du génome de *D. pulex* représente une espèce hybride entre *D. pulex* et une espèce (peut-être éteinte maintenant) proche évolutivement de *D. pulicaria* européennes. Le taux d'insertions des introns dans les génomes calculé par Li et al. (2009) serait dans ce cas sûr-estimé.

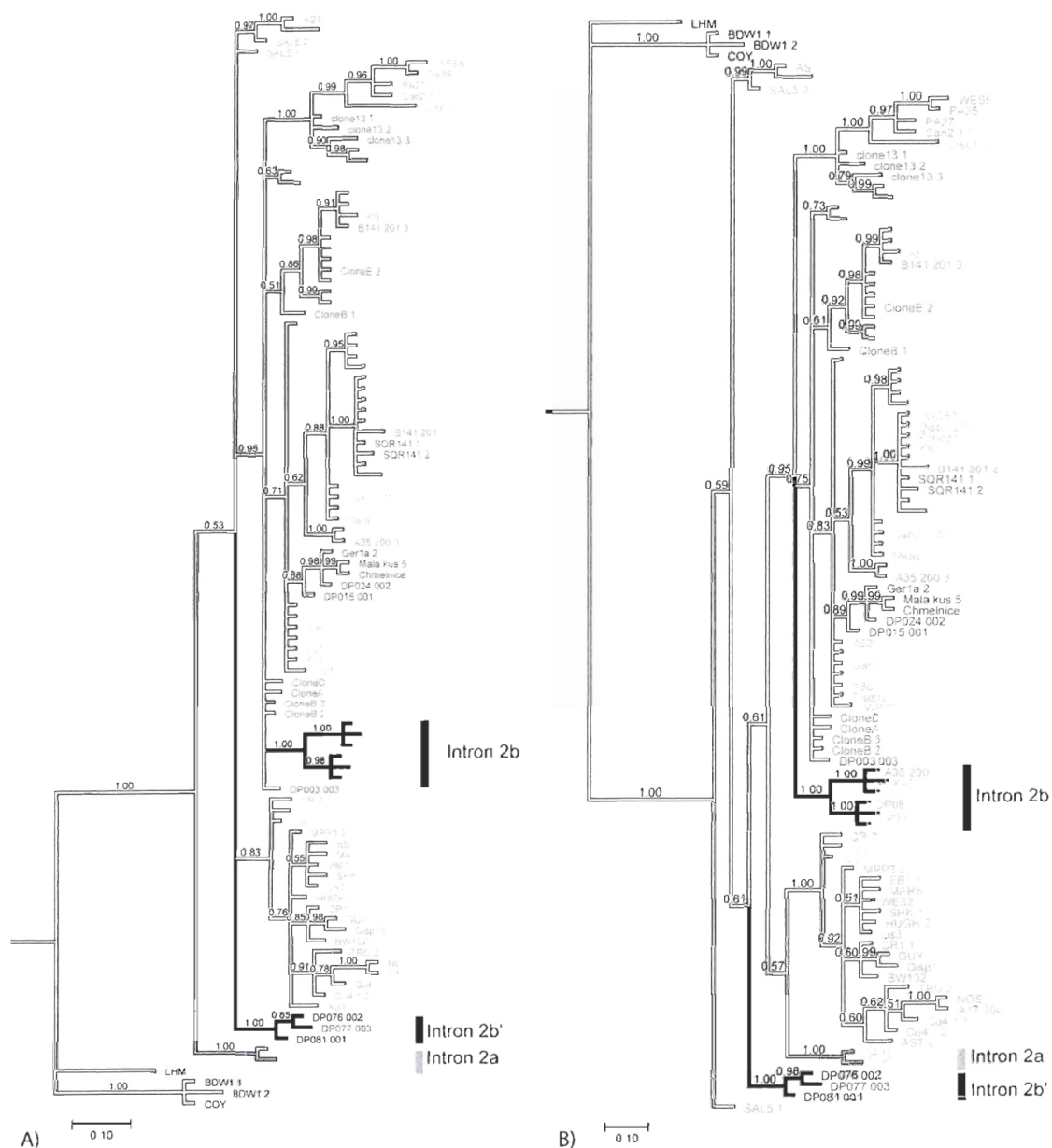


Figure 5-1 Deux inférences phylogénétiques Bayésiennes différentes montrant l'insertion des introns 2a, 2b et 2b' dans le gène *Rab4*

A) Phylogénie Bayésienne utilisant un modèle d'évolution HKY (déviation standard = 0.003105) B) Phylogénie Bayésienne utilisant sans modèle d'évolution particulier (déviation standard = 0.004662)

5.4 Évolution de *Pokey* dans le complexe *pulex*

La quasi-totalité des polyploïdes sont de nature hybride. La polyploïdie dans le complexe *pulex* est liée à des phénomènes d'hybridation. Le nombre de transposons insérés dans un génome dépend de plusieurs facteurs. La dynamique et la densité des éléments transposables dans le génome d'un individu dépend de leurs présences dans les génomes des espèces parentes, de l'activité de l'élément transposable considéré, des niches génomiques disponibles et des possibles mécanismes génétiques et épigénétiques qui pourraient réguler la transposition (Venner, Feschotte et Biémont, 2009). L'objectif du quatrième chapitre était de tester si les génomes ayant un niveau de ploïdie élevé (comparé aux génomes diploïdes), donc un nombre de niches disponibles élevées, possédaient un plus grand nombre d'insertions de l'élément *Pokey*. Afin de tenir compte de l'effet de l'hybridation, nous avons analysé le nombre d'insertions du transposon *Pokey* dans des isolats hybrides diploïdes et triploïdes.

5.4.1 Effet du niveau de ploïdie les patrons d'insertion du transposon *Pokey*

L'absence de résultats significatifs dans la comparaison des densités d'insertions de *Pokey* entre diploïdes et triploïdes hybrides provenant du complexe *pulex* pourrait être expliquée par plusieurs scénarios.

5.4.1.1 Augmentation du niveau de ploïdie et diploïdisation subséquente

Le résultat observé est comparable à des observations faites chez des plantes allopolyploïdes (Parisod *et al.*, 2009 ; Petit *et al.*, 2007). L'activation de certains éléments transposables, la modification de l'expression des gènes, et de nombreux réarrangements sont des phénomènes fréquents lors des événements de

polyploïdisation (passé en revue dans Parisod *et al.*, 2010). Des études portant sur des plantes allopolyploïdes d'origines diverses (formées « récemment » ou « anciennement ») souligne qu'une augmentation du nombre d'insertions d'éléments transposables par rapport aux génomes diploïdes apparentés est rare (Parisod *et al.*, 2010). Une perte du nombre d'insertions d'éléments transposables a été observée chez diverses plantes allopolyploïdes (Parisod *et al.*, 2010). Des phénomènes de diploïdisations sont fréquemment observés au cours des générations après des phénomènes d'augmentation du niveau de ploïdie (Gerstein, 2006 ; Leitch et Bennett, 2004 ; Wolfe, 2001). Les lignées polyploïdes analysées ici ont été échantillonnées en nature. Il n'a pas été possible de dater la formation des lignées polyploïdes. Selon la diversité mitochondriale Il a été estimé que l'espèce polyploïde *D. middendorffiana sensu stricto* pourrait s'être formée lors du Pléistocène (Dufresne et Hebert, 1997 ; Weider et Hobæk, 2003). Donc l'absence de différence entre diploïdes et polyploïdes hybrides pourrait être due à des phénomènes de diploïdisation plus ou moins rapides après la formation des polyploïdes.

5.4.1.2 Absence d'activation des éléments transposables dans les génomes polyploïdes

Il est aussi possible que le nombre d'insertions d'éléments transposables n'aient pas augmenté avec le niveau de ploïdie car l'élément *Pokey* ne s'est pas déplacé. En effet les différentes analyses sur les plantes polyploïdes montrent que toutes les familles d'éléments transposables ne sont pas activées lors d'évènements de polyploïdisation. Les transposables de classe II se déplacent sous une forme d'ADN (double brin pour les transposons à transposase). Ils ont donc potentiellement une capacité à se répliquer moins grande que les rétrotransposons. Toutefois une augmentation du nombre d'insertions des transposons peut se produire lors des phénomènes de transpositions (Engels *et al.*, 1990). Par exemple,

un élément inséré à un locus de manière homozygote en s'excisant (pour s'insérer dans un autre locus) provoque une cassure double brin. Celle-ci peut être réparée par la machinerie moléculaire de l'hôte en prenant comme modèle le transposon homologue restant au locus initial. Les transposons sont alors répliqués via la machinerie moléculaire de l'hôte (Engels *et al.*, 1990). *Pokey* semble inactif dans les populations asexuées de *D. pulex* (Sullender et Crease, 2001a ; Valizadeh et Crease, 2008). Si *Pokey* a accumulé des mutations dans les lignées hybrides asexuées, il est possible qu'il soit inactif dans ces lignées et qu'il n'ait pas la capacité de s'autoexciser et donc de se déplacer chez les polyploïdes. D'autres éléments actifs présents dans le génome de *D. pulex* (Rho *et al.*, 2010 ; Schaack *et al.*, 2010a) auraient pu se déplacer et se répliquer lors de la polyploïdisation. Une analyse sur différents éléments transposables disponibles dans le génome de *D. pulex* (Rho *et al.*, 2010 ; Schaack *et al.*, 2010a) permettrait de tester cette hypothèse.

5.4.1.3 Saturation des niches lors de l'hybridation

L'hybridation *per se* entre espèces semblerait avoir un impact dramatique sur la densité des éléments transposables dans le génome (Fontdevila, 2005 ; Petrov *et al.*, 1995 ; Shan *et al.*, 2005). Les triploïdes testés dans notre étude pourraient être issus d'un croisement entre des individus hybrides asexués produisant des gamètes non-réduits et des individus sexués non-hybrides. Dans une première phase de notre étude nous n'avons comparé que des isolats hybrides (diploïdes et polyploïdes) entre eux. Une comparaison de quelques isolats diploïdes hybrides avec des isolats non-hybrides a montré que la densité de transposons *Pokey* était en moyenne trois fois plus importante dans des isolats hybrides que dans les isolats non-hybrides (10.55 ± 5.22 vs. 3.27 ± 2.07). Toutefois certains individus hybrides ne présentent pas une densité plus importante que les individus non hybrides. L'activation et l'augmentation du nombre d'insertions ne semblent pas être

homogènes à travers les génomes hybrides. Cette variabilité pourrait être due à la capacité de se transposer des différents allèles de *Pokey* hérités des espèces parentes et aussi à l'efficacité des phénomènes génétiques et épigénétiques qui régulent les éléments transposables chez les hybrides.

Si l'augmentation du nombre d'insertions d'éléments transposables se fait lors de l'hybridation, le génome pourrait être saturé par les insertions. Les éléments transposables auraient envahi les différentes niches génomiques potentielles après hybridation. Le nombre d'insertions entre diploïdes hybrides et polyploïdes hybrides serait alors non-significativement différent. Une comparaison de la dynamique de différents éléments transposables dans des isolats diploïdes purs et hybrides produits par croisements contrôlés en laboratoire serait une expérience qui permettrait de tester partiellement cette hypothèse.

Plus de génomes polyploïdes que de génomes diploïdes possèdent des sites d'insertions de *Pokey* non-partagés (avec les autres génomes). Cette différence, qui est toutefois non significative, pourrait refléter simplement l'addition de chromosomes avec des insertions non-partagées par les autres ensembles de chromosomes constituant le génome triploïde. Elle pourrait aussi refléter une plus grande disponibilité des niches génomiques potentielles dans les génomes polyploïdes. En effet, l'hypothèse de la redondance des génomes (Matzke et Matzke, 1998) prédit une augmentation du nombre d'insertions avec le niveau de ploïdie causée par un plus grand nombre de niches (locus) disponibles. Ces niches peuvent être des locus homologues des locus déjà occupés par les éléments transposables chez les individus diploïdes parents. Elles peuvent donc représenter de nouvelles niches, c'est-à-dire des locus où l'insertion d'éléments transposables était impossible chez les individus diploïdes car sous forte contrainte sélective. L'addition de génome aurait relâché ces contraintes sélectives et aurait permis l'insertion des éléments transposables dans de nouvelles niches génomiques. Un plus grand nombre de clones diploïdes et de clones triploïdes et l'analyse des

séquences flanquant les insertions chez les diploïdes et les polyploïdes permettrait de tester cette hypothèse.

5.4.2 Recombinaison intra-génique du transposon *Pokey*

L'analyse des séquences de *Pokey* insérées dans les gènes répétés codant pour la grande sous-unité 28S de l'ARN ribosomal a montré que la grande diversité de ces séquences était due à des événements de recombinaisons entre des allèles divergents appartenant de *Pokey* retrouvés dans diverses espèces du complexe *pulex*. Le nombre important d'évènements de recombinaisons qui a façonné l'évolution de la lignée *Pokey* pourrait être la conséquence des propriétés des gènes où ce transposon s'insère préférentiellement (les gènes répétés codant pour la grande sous-unité 28S de l'ARNr; Penton, Sullender et Crease, 2002 ; Sullender et Crease, 2001a). Les locus fortement répétés comme les gènes codant pour les sous-unités de l'ARN ribosomal seraient des niches génomiques idéales pour les éléments transposables (Eickbush et Eickbush, 2007 ; Zhang, Eickbush et Eickbush, 2008). Un grand nombre de sites d'insertions permettrait aux éléments transposables de s'insérer de multiples fois dans le génome sans trop altérer la fitness de leur hôte (Zhang, Eickbush et Eickbush, 2008). Ces gènes répétés correspondent aussi à une région où le taux de recombinaison est fort (Crease, 1995 ; Eickbush et Eickbush, 2007 ; McTaggart *et al.*, 2007). La sélection naturelle agit sur les éléments transposables à différents niveaux (Hickey, 1982), au niveau de l'élément, augmentant sa valeur d'aptitude (sa capacité à se transposer ou à se transférer à d'autres organismes par exemple) et au niveau de l'hôte afin qu'un trop grand nombre d'insertions n'altère pas dramatiquement la fitness du génome qui le porte. Si les éléments transposables ont un effet principalement délétère sur leur hôte, la sélection naturelle agissant sur l'hôte va avoir pour conséquence une inactivation des éléments transposables au cours de l'évolution. Ces deux niveaux de sélections sont donc en conflit. L'effet de la sélection naturelle au niveau de

l'élément transposable a une conséquence opposée à celle au niveau de l'hôte. L'hypothèse d'Eickbush et ses collaborateurs (2007) tente d'expliquer la distribution des éléments transposables dans les gènes codant pour l'ARN ribosomal en fonction du niveau de sélection de l'hôte. Des régions préférentielles d'insertions à forts taux de recombinaison auraient pu être sélectionnées au niveau du transposon. Un fort taux de recombinaison pourrait être sélectionné afin de permettre une diversification et/ou de permettre de purger les mutations qui inactivent les éléments transposables. L'hypothèse avancée ici n'est pas opposée à l'hypothèse d'Eickbush et ses collaborateurs (Eickbush et Eickbush, 2007 ; Zhang, Eickbush et Eickbush, 2008) et permettrait d'expliquer pourquoi certains éléments transposables sont insérés dans des régions à fort taux de recombinaisons (Duret, Marais et Biemont, 2000).

5.4.2.1 Purge des mutations délétères

Une analyse d'enzymes de restriction a permis de montrer que des allèles de *Pokey* possiblement recombinés se retrouvent dans différents génomes hybrides et non-hybrides (Tableau 4-4). Il serait donc possible que des éléments recombinés soient actifs et qu'ils aient envahi des génomes non-hybrides. Chez *Caenorhabditis elegans*, les transposons à ADN sont situés préférentiellement dans les régions à fort taux de recombinaison alors que les rétrotransposons ne le sont pas (Duret, Marais et Biemont, 2000). Il a été supposé un rôle de la recombinaison dans le processus de transposition pour tenir compte de cet écart entre les théories et les données empiriques. Dans une étude sur des éléments mobiles qui s'insèrent dans les génomes de différentes entérobactéries, Lawrence et ses collaborateurs (1992) ont montré que certaines séquences de *IS1* et *IS3* ont été produites par recombinaison entre des séquences divergentes. Par analogie avec les phénomènes de reproduction sexuée et de purge des allèles délétères dans les populations sexuées, Il est possible que les phénomènes de recombinaison

interélément permettent de purger les mutations délétères qui s'accumulent naturellement dans les séquences du transposon *Pokey* et qui inactivent ceux-ci.

5.4.2.2 Diversification et production de MITES

À partir de l'analyse de deux génomes (souches N2 et NL917) de *Caenorhabditis elegans*, Fischer et ses collaborateurs (2003) ont observé la présence d'éléments transposables Tc1 chimériques (issus de la recombinaison de différentes copies de Tc1) actifs dans le génome de la souche NL917. Cette étude propose que des éléments transposables Tc1 recombinés puissent être produits via le processus de réparation des cassures d'ADN double brin après l'excision du transposon. Ces auteurs concluaient que la recombinaison entre des copies ayant accumulé des mutations pourrait conduire à une évolution rapide et à une diversification des éléments Tc1 (Fischer, Wienholds et Plasterk, 2003). Une hypothèse similaire a été émise pour expliquer l'activité de rétrotransposons recombinés de la sous-famille CRM1 dans des plants hybrides de maïs (Sharma, Schneider et Presting, 2008). Des expériences en laboratoire ont montré que des mutations ponctuelles ou des réarrangements plus conséquents dans différents éléments transposables pouvaient avoir pour conséquence une augmentation du taux de transposition de ceux-ci (Beall, Mahoney et Rio, 2002 ; Han et Boeke, 2004 ; Lampe *et al.*, 1999). Les cadres ouverts de lecture (ORFs) séquencés dans notre étude montrent de la variation qui se reflète en partie sur les séquences en acide aminés traduites (données non montrées). Une analyse de présence d'évènements de recombinaisons a aussi révélé que la recombinaison avait un impact sur les ORFs et donc sur les séquences en acides aminés des protéines de transposases. Il est donc possible que les éléments *Pokey* recombinés retrouvés dans les différents génomes aient des propriétés et des capacités de transpositions différentes de celles des éléments *Pokey* dont ils sont issus.

Les différentes séquences de *Pokey* analysées lors de cette étude contenaient un grand nombre de délétions (ou alternativement d'insertions). Il est possible que les éléments transposables dégénèrent au cours de l'évolution soit pour produire des éléments non-autonomes et non-activables soit pour produire des éléments non-autonomes mais qui sont capables de se déplacer si des transposons actifs de la même famille sont présents dans le génome (MITEs; Feschotte, Zhang et Wessler, 2002). Pour qu'un transposon comme *Pokey* soit autonome il faut que les séquences terminales inversées répétées (TIR), séquences essentielles à l'attachement de la transposase au transposon soient fonctionnelles. De plus, il faut que le gène codant pour la protéine transposase soit fonctionnel. La recombinaison inter-éléments pourrait, en réarrangeant les éléments transposables autonomes par des inversions ou des délétions de nucléotides à l'intérieur des gènes codant pour la transposase, être une source de production de MITEs. Les MITEs sont connus pour constituer une part importante des génomes de plantes et d'animaux (Casacuberta et Santiago, 2003 ; Feschotte, Zhang et Wessler, 2002 ; Shan *et al.*, 2005). Leur origine est fréquemment associée à des transposons homologues présents dans les génomes (Feschotte, Zhang et Wessler, 2002). La recombinaison inter-éléments au moment de phénomènes d'hybridation serait potentiellement un mode d'origine pour ces éléments non-autonomes.

Il serait intéressant de caractériser l'activité des éléments *Pokey* issus d'évènements de recombinaison afin de tester le modèle de diversification et de purge de mutations délétères par recombinaison.

5.5 Conclusion

Il est donc possible de dresser un tableau général du système du complexe d'espèces *Daphnia pulex* afin de mettre en place un modèle d'étude pour mieux comprendre l'évolution et l'écologie d'éléments transposables dans différents

génomés. Diverses lignées hybrides diploïdes et polyploïdes ont été produites entre des espèces écologiquement distinctes dans ce complexe d'espèces au cours de l'évolution. Certaines lignées polyploïdes semblent être issues de croisements entre individus provenant de la même espèce (i.e. autopolyploïdes). Les lignées diploïdes hybrides peuvent être produites en milieux contrôlés (Heier et Dudycha, 2009 ; Innes et Hebert, 1988). Les lignées polyploïdes sont pour la plupart triploïdes et seraient potentiellement issus de croisements entre des gamètes non-réduits (probablement produits par des lignées hybrides) et des gamètes réduits. Certaines lignées triploïdes auraient été produites depuis le Pléistocène (*D. middendorffiana sensu stricto*, *D. pulicaria* polyploïdes et *D. pulex* polyploïdes) alors que d'autres ont une origine évolutive potentiellement plus ancienne (*D. tenebrosa* polyploïdes). La densité de l'élément transposable *Pokey* semble ne pas différer entre les génomes hybrides diploïdes et polyploïdes. Cette différence peut-être due à plusieurs phénomènes (i.e. absence d'activation, activation mais perte de séquences suite à une diploïdisation...). Certains indices laissent à penser que le nombre d'insertions de *Pokey* aurait augmenté lors d'évènements d'hybridations. La recombinaison inter-éléments semble être une force majeure dans l'évolution du transposon *Pokey*. De plus, des allèles recombinés de l'élément *Pokey* se retrouveraient en nature dans des génomes hybrides et non-hybrides. Les éléments *Pokey* recombinés pourraient alors être actifs en nature comme cela a été supposé pour d'autres éléments transposables dans des génomes issus d'hybridation (Sharma, Schneider et Presting, 2008).

Depuis Kidwell et Lisch (1997), différents auteurs (Brookfield, 2005 ; Le Rouzic, Dupas et Capy, 2007 ; Leonardo et Nuzdhin, 2002 ; Venner, Feschotte et Biéumont, 2009 ; Yoder, Walsh et Bestor, 1997) ont comparé l'architecture des génomes à la structuration des écosystèmes. Différentes théories spécifiques à l'écologie systémique, telles que la théorie de partitionnement des niches (Chesson, 2000) ou la théorie neutre en écologie des communautés (Bell, 2000 ; Hubbell, 2005) par exemple, pourraient être adaptées à l'évolution des génomes (Venner, Feschotte et Biéumont, 2009). Des modèles théoriques prenant en

considération des interactions de coopération, d'hôte/parasite ou de compétitions entre éléments transposables (Abrusán et Krambeck, 2006 ; Le Rouzic, Dupas et Capy, 2007) ont été construits afin de comprendre la variation de la dynamique et de la distribution d'éléments transposables ayant des caractéristiques différentes. Une grande quantité de données sur la distribution des éléments transposables dans les différents génomes (Bartolomé, Maside et Charlesworth, 2002 ; Biemont *et al.*, 2003 ; Dolgin et Charlesworth, 2008 ; Duret, Marais et Biemont, 2000 ; Hua-Van *et al.*, 2005 ; Pidpala, Yatsishina et Lukash, 2008 ; Rizzon *et al.*, 2002) et leur dynamique (Biémont, 1992 ; Hanson *et al.*, 2000 ; Hua-Van *et al.*, 2005 ; Valizadeh et Crease, 2008 ; Zeyl, Bell et Green, 1996) dans les populations hôtes est disponible. Toutefois les données récoltées par le séquençage de génomes complets ont rarement été analysées à la lumière de théories écosystémiques. La plupart des études sur la distribution des éléments transposables n'ont pris en considération que l'effet des transposons sur la fitness de leur hôte sans toutefois arriver à un consensus (Bartolomé, Maside et Charlesworth, 2002 ; Biémont, 1992 ; Biemont *et al.*, 2003 ; Charlesworth et Charlesworth, 1983 ; Charlesworth, Langley et Sniegowski, 1997 ; Dolgin et Charlesworth, 2008 ; Langley *et al.*, 1988). Peu d'études empiriques ont considéré les différents éléments transposables comme des structures, ayant des phénotypes propres à leur séquence (e.g. préférence de site d'insertion, mode de transposition, fitness...), en interactions entre-elles et avec les composantes biochimiques du génome afin comprendre leur distribution et leur dynamique dans les génomes. Une vision écosystémique des génomes et des différents processus qui les produisent pourrait permettre de mieux comprendre la distribution des différents éléments transposables dans les génomes. Des études expérimentales seraient nécessaires pour permettre de mieux comprendre la dynamique des éléments transposables lors de la formation de génomes de différentes natures (e.g. hybridations, mis en culture de tissus, etc.). Des études ont montré une activation des éléments transposables lors de différentes perturbations (Capy *et al.*, 2000). Des modifications de la dynamique des éléments transposables (e.g. activation, perte d'insertions...) peu de temps après l'hybridation et/ou la

polyploïdisation ont été observées chez différents organismes (Fontdevila, 2005 ; Hazzouri *et al.*, 2008 ; Kashkush, Feldman et Levy, 2002 ; Liu *et al.*, 2009 ; Madlung *et al.*, 2005 ; O'Neill, O'Neill et Graves, 1998 ; Parisod *et al.*, 2009 ; Petrov *et al.*, 1995 ; Salmon, 2005 ; Shan *et al.*, 2005 ; Zhao *et al.*, 1998). En reprenant l'analogie d'un écosystème génomique, cette dynamique du génome pourrait être comparée à de la résilience¹. En effet, après une perturbation (hybridation par exemple), le génome semble, via de possibles processus médiés par la sélection naturelle, être restructuré et perd des insertions d'éléments transposables (Kashkush, Feldman et Levy, 2002 ; Madlung *et al.*, 2005 ; Parisod *et al.*, 2010 ; Soltis et Soltis, 2000). Les phénomènes de diploïdisation sont connus chez des organismes polyploïdes et paléopolyploïdes (Bennett, 2004 ; Feldman *et al.*, 1997 ; Ferris et Whitt, 1979 ; Leitch et Bennett, 2004 ; Li, 1980 ; Wolfe, 2001). Il a été montré qu'une augmentation du niveau de ploïdie chez la levure (*Saccharomyces cerevisiae*) était suivie d'une rediploïdisation et que la quantité d'ADN la plus stable dans un milieu non-sélectif se rapprochait de la quantité d'ADN de la souche diploïde observée en nature (Gerstein, 2006). La rapidité avec laquelle le génome de *S. cerevisiae* revenait à une quantité d'ADN « diploïde » semblait dépendre de la pression de sélection appliquée (Gerstein, 2006). En continuant cette analogie, il serait possible qu'un génome donné fasse preuve de résistance, dans le sens écologique du terme. Un génome dans lequel les phénomènes épigénétiques seraient assez importants pour limiter les processus de transposition et dont la structure serait stable (son intégrité) même après une perturbation (environnementale ou génomique), serait dit résistant alors qu'un génome dont la structure serait fortement altérée après perturbation serait dit non-résistant. Une relation entre stabilité du génome et

¹ Cette analogie m'a été proposée par Olivier Deshaies, étudiant à la maîtrise en écologie forestière à l'UQAR.

diversité des structures (e.g. séquences codantes, introns, éléments transposables...), telle qui a été proposée en écologie des communautés (Pimm, 1984 ; Tilman, 1996), pourrait être testée en utilisant une approche combinée de génomique et d'évolution expérimentale. Des génomes plus ou moins complexes seraient soumis à des perturbations environnementales ou génomiques, la stabilité de leur structure serait testée par des techniques génomiques (e.g. séquençage, « TE display »). Si la stabilité des génomes est dépendante de la pression de sélection à laquelle est soumis l'individu qui la porte, cette théorie pourrait expliquer le maintien de génomes polyploïdes dans des environnements perturbés. Des génomes complexes (forte hétérozygotie et grand nombre de structures génétiques) seraient sélectionnés positivement dans des milieux perturbés. Cette théorie est similaire à la théorie qui prédit une corrélation entre la stabilité phénotypique (robustesse génétique) et la complexité du génome (de Visser *et al.*, 2003).

Le génome de l'une des espèces constituant le complexe *pulex* a été dernièrement séquencé (Colbourne *et al.*, soumis) et un grand nombre d'éléments transposables a été mis en relief (Rho *et al.*, 2010 ; Schaack *et al.*, 2010a). Les daphnies sont des organismes zooplanctoniques faciles à élever et il est possible de provoquer des perturbations génomiques et environnementales dans des lignées élevées en laboratoire. Des populations parthénogénétiques cycliques et obligatoires se rencontrent en nature. Des croisements peuvent être effectués entre des populations d'une même espèce et ou entre des espèces écologiquement distinctes (i.e. *D. pulex* et *D. pulicaria*) pour produire des hybrides diploïdes (et potentiellement polyploïdes). Dû à un mode de reproduction parthénogénétique cyclique, des lignées sexuées peuvent être maintenues par reproduction clonale en laboratoire. Cette propriété permet de produire des lignées d'accumulation de mutations (Omilian *et al.*, 2006 ; Schaack *et al.*, 2010a). Les lignées hybrides peuvent aussi être entretenues de manière clonale en laboratoire. Les séquences de l'élément transposable *Pokey* sont fortement diversifiées via probablement des processus de recombinaison dans le complexe *Daphnia pulex*. Il serait alors intéressant d'étudier l'effet de la recombinaison sur la fitness et la diversification des transposons et sur

les différentes interactions que cet élément a avec les processus biochimiques qui régulent l'expression de ce transposon. Les différents types d'éléments transposables rencontrés dans le génome (Rho *et al.*, 2010 ; Schaack *et al.*, 2010a) ont différentes caractéristiques (i.e. autonomes vs. Non-autonomes, transposon vs, retrotransposons). Cette diversité permettrait de tester les effets des caractéristiques des éléments transposables sur leur distribution et leur dynamique dans les génomes dans des populations sexuées pures mais aussi dans des génomes perturbés par des facteurs environnementaux ou des facteurs génomiques.

Bibliographie

- Abrusán, G., et H.-J. Krambeck. 2006. «Competition may determine the diversity of transposable elements». *Theoretical Population Biology*, vol. 70, no 3, p. 364-375.
- Adamowicz, S. J., T. R. Gregory, M. C. Marinone et P. D. N. Hebert. 2002. «New insights into the distribution of polyploid *Daphnia*: the Holarctic revisited and Argentina explored». *Molecular Ecology*, vol. 11, p. 1209-1217.
- Adams, K. L., R. Cronn, R. Percifield et J. F. Wendel. 2003. «Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no 8, p. 4649-4654.
- Adams, K. L., et J. F. Wendel. 2005a. «Novel patterns of gene expression in polyploid plants». *Trends in Genetics*, vol. 21, no 10, p. 539-543.
- , 2005b. «Polyploidy and genome evolution in plants». *Current Opinion in Plant Biology*, vol. 8, no 2, p. 135-141.
- Addison, J. A., et G. H. Pogson. 2009. «Multiple gene genealogies reveal asymmetrical hybridization and introgression among stronglylocotritid sea urchins». *Molecular Ecology*, vol. 18, no 6, p. 1239-1251.
- Aguilera, X., J. Mergeay, A. Wollebrants, S. Declerck et L. De Meester. 2007. «Asexuality and polyploidy in *Daphnia* from the tropical Andes». *Limnology and Oceanography*, vol. 52, no 5, p. 2079-2088.

- Ainouche, M., P. Fortune, A. Salmon, C. Parisod, M. A. Grandbastien, K. Fukunaga, M. Ricou et M. T. Misset. 2009. «Hybridization, polyploidy and invasion: lessons from *Spartina* (Poaceae)». *Biological Invasions*, vol. 11, no 5, p. 1159-1173.
- Ambrose, C., et T. Crease. 2011. «Evolution of the nuclear ribosomal DNA intergenic spacer in four species of the *Daphnia pulex* complex». *BMC Genetics*, vol. 12, no 1, p. 13.
- Aminetzach, Y. T., J. M. Macpherson et D. A. Petrov. 2005. «Pesticide Resistance via Transposition-Mediated Adaptive Gene Truncation in *Drosophila*». *Science*, vol. 309, no 5735, p. 764-767.
- Arkhipova, I., et M. Meselson. 2000. «Transposable elements in sexual and ancient asexual taxa». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no 26, p. 14473-14477.
- Bartolomé, C., X. Maside et B. Charlesworth. 2002. «On the Abundance and Distribution of Transposable Elements in the Genome of *Drosophila melanogaster*». *Molecular Biology and Evolution*, vol. 19, no 6, p. 926-937.
- Beall, E. L., M. B. Mahoney et D. C. Rio. 2002. «Identification and Analysis of a Hyperactive Mutant Form of *Drosophila* P-Element Transposase». *Genetics*, vol. 162, no 1, p. 217-227.
- Beaton, M. J. 1995. «Patterns of endopolyploidy and genome size variation in *Daphnia*». Guelph, Ontario, University of Guelph.
- Beaton, M. J., et P. D. N. Hebert. 1988. «Geographical Parthenogenesis and Polyploidy in *Daphnia pulex*». *The American Naturalist*, vol. 132, no 6, p. 837-845.

- , 1989. «Miniature genomes and endopolyploidy in cladoceran crustaceans». *Genome*, vol. 32, no 6, p. 1048-1053.
- , 1994. «Variation in Chromosome Numbers of *Daphnia* (Crustacea, Cladocera)». *Hereditas*, vol. 120, no 3, p. 275-279.
- Bell, G. 2000. «The Distribution of Abundance in Neutral Communities». *The American Naturalist*, vol. 155, no 5, p. 606-617.
- Bell, M. A., et M. P. Travis. 2005. «Hybridization, transgressive segregation, genetic covariation, and adaptive radiation». *Trends in Ecology & Evolution*, vol. 20, no 7, p. 358-361.
- Bennett, M., I. Leitch, J. Price et S. Johnston. 2003. «Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~175 Mb) Using Flow Cytometry Show Genome Size in *Arabidopsis* to be ~157 Mb and thus ~25 % Larger than the *Arabidopsis* Genome Initiative Estimate of ~125 Mb». *Annals of Botany*, vol. 91, no 5, p. 547-557.
- Bennett, M. D. 2004. «Perspectives on polyploidy in plants - ancient and neo». *Biological Journal of the Linnean Society*, vol. 82, no 4, p. 411-423.
- Bennetzen, J. L. 2002. «Mechanisms and rates of genome expansion and contraction in flowering plants». *Genetica*, vol. 115, no 1, p. 29-36.
- Bestor, T. 1999. «Sex brings transposons and genomes into conflict». *Genetica*, vol. 107, no 1, p. 289-295.
- Biémont, C. 1992. «Population genetics of transposable DNA elements». *Genetica*, vol. 86, no 1, p. 67-84.

- Biemont, C., C. Nardon, G. Deceliere, D. Lepetit, C. Loevenbruck et C. Vieira. 2003. «Worldwide distribution of transposable element copy number in natural populations of *Drosophila simulans*». *Evolution*, vol. 57, no 1, p. 159-167.
- Blumenstiel, J. P., et D. L. Hartl. 2005. «Evidence for maternally transmitted small interfering RNA in the repression of transposition in *Drosophila virilis*». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no 44, p. 15965-15970.
- Böhne, A., F. Brunet, D. Galiana-Arnoux, C. Schultheis et J.-N. Volff. 2008. «Transposable elements as drivers of genomic and biological diversity in vertebrates». *Chromosome Research*, vol. 16, no 1, p. 203-215.
- Bretagnolle, F., et J. D. Thompson. 1995. «Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants». *New Phytologist*, vol. 129, no 1, p. 1-22.
- Britten, R. J. 1996. «DNA sequence insertion and evolutionary variation in gene regulation». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no 18, p. 9374-9377.
- Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A. C. Scheen et R. Elven. 2004. «Polyploidy in arctic plants». *Biological Journal of the Linnean Society*, vol. 82, no 4, p. 521-536.
- Brookfield, J. 1991. «Models of Repression of Transposition in P-M Hybrid Dysgenesis by P Cytotype and by Zygotically Encoded Repressor Proteins». *Genetics*, vol. 128, no 2, p. 471-486.
- Brookfield, J., et R. Badge. 1997. «Population genetics models of transposable elements». *Genetica*, vol. 100, no 1, p. 281-294.

- Brookfield, J. F. Y. 2005. «The ecology of the genome - Mobile DNA elements and their hosts». *Nature Reviews Genetics*, vol. 6, no 2, p. 128-136.
- Brooks, J. L. 1957a. «The species problem in freshwater animals». In *The species problem*, E Mayr, p. 81-123. Washington, D. C.: american association for the advancement of science.
- , 1957b. *The systematics of North American Daphnia*. Coll. «Memoirs of the Connecticut Academy of Arts & Sciences». New Haven: Connecticut Academy of Arts & Sciences p.
- Bruen, T. C., H. Philippe et D. Bryant. 2006. «A simple and robust statistical test to detect the presence of recombination». *Genetics*, p. genetics.105.048975.
- Buckley, T. R., M. Cordeiro, D. C. Marshall et C. Simon. 2006. «Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (*Maoricicada Dugdale*)». *Systematic Biology*, vol. 55, p. 411-425.
- Buntjer, J. B. 2001. *Phylogenetic computer tools (PhylTools). Version 1.32 for Windows*. Wageningen: Laboratory of Plant Breeding: Wageningen University.
- Calinski, T., et J. Harabasz. 1974. «A Dendrite Method for Cluster Analysis». *Communications in Statistics*, vol. 3, p. 1-27.
- Capy, P., D. Anxolabéhère et T. Langin. 1994. «The strange phylogenies of transposable elements: are horizontal transfers the only explanation?». *Trends in Genetics*, vol. 10, no 1, p. 7-12.

- Capy, P., G. Gasperi, C. Biemont et C. Bazin. 2000. «Stress and transposable elements: co-evolution or useful parasites?». *Heredity*, vol. 85, no 2, p. 101-106.
- Carr, M., J. Soloway, T. Robinson et J. Brookfield. 2002. «Mechanisms regulating the copy numbers of six LTR retrotransposons in the genome of <small>Drosophila melanogaster</small>». *Chromosoma*, vol. 110, no 8, p. 511-518.
- Casacuberta, J. M., et N. Santiago. 2003. «Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes». *Gene*, vol. 311, p. 1-11.
- Castro, J. P., et C. M. A. Carareto. 2004. «*Drosophila melanogaster* Transposable Elements: Mechanisms of Transposition and Regulation». *Genetica*, vol. 121, no 2, p. 107-118.
- Cavalier-Smith, T. 1978. «Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox». *J Cell Sci*, vol. 34, no 1, p. 247-278.
- Cerny, M., et P. Hebert. 1993. «Genetic diversity and breeding system variation in *Daphnia pulicaria* from North-American Lakes». *Heredity*, vol. 71, p. 497 - 507.
- Charlesworth, B., et D. Charlesworth. 1983. «The population dynamics of transposable elements». *Genetics Research*, vol. 42, no 01, p. 1-27.
- Charlesworth, B., C. H. Langley et P. D. Sniegowski. 1997. «Transposable Element Distributions in *Drosophila*». *Genetics*, vol. 147, p. 1993-1995.

- Charlesworth, D., et B. Charlesworth. 1995. «Transposable elements in inbreeding and outbreeding populations». *Genetics*, vol. 140, no 1, p. 415-417.
- Charlesworth, D., et S. I. Wright. 2001. «Breeding systems and genome evolution». *Current Opinion in Genetics & Development*, vol. 11, no 6, p. 685-690.
- Chesson, P. 2000. «Mechanisms of Maintenance of Species Diversity». *Annual Review of Ecology and Systematics*, vol. 31, p. 343-366.
- Cohan, F. M. 1994. «The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes». *American Naturalist*, vol. 143, no 6, p. 965-986.
- Colbourne, J. K., T. J. Crease, L. J. Weider, P. D. N. Hebert, F. Dufresne et A. Hobaek. 1998. «Phylogenetics and evolution of a circumarctic species complex (Cladocera : *Daphnia pulex*)». *Biological Journal of the Linnean Society*, vol. 65, no 3, p. 347-365.
- Colbourne, J. K., et P. D. N. Hebert. 1996. «The systematics of North American *Daphnia* (Crustacea: Anomopoda): A molecular phylogenetic approach». *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, vol. 351, no 1337, p. 349-360.
- . 1997. «Evolutionary origins of phenotypic diversity in *Daphnia*». In *Molecular Evolution and Adaptive Radiation*, T.K. Givnish et K.J. Sytsma, p. 163-188: Cambridge University Press.
- Colbourne, J. K., M. E. Pfrender, D. Gilbert, W. K. Thomas, A. Tucker, T. H. Oakley, S. Tokishita, A. Aerts, G. J. Arnold, M. K. Basu, D. J. Bauer, C. E. Cáceres, L. Carmel, C. Casola, J.-H. Choi, J. C. Detter, Q. Dong, S. Dusheyko, B. D. Eads, T. Fröhlich, K. A. Geiler-Samerotte, D. Gerlach, P. Hatcher, S. Jogdeo, J. Krijgsveld, E. V. Kriventseva, D. Kültz, C. Laforsch, E. Lindquist, J. Lopez,

J. R. Manak, J. Muller, J. Pangilinan, R. P. Patwardhan, S. Pitluck, E. J. Pritham, A. Rechtsteiner, M. Rho, I. B. Rogozin, O. Sakarya, A. Salamov, S. Schaack, H. Shapiro, Y. Shiga, C. Skalitzky, Z. Smith, A. Souvorov, W. Sung, Z. Tang, D. Tsuchiya, H. Tu, H. Vos, M. Wang, Y. I. Wolf, H. Yamagata, T. Yamada, Y. Ye, J. R. Shaw, J. Andrews, T. J. Crease, H. Tang, S. M. Lucas, H. M. Robertson, P. Bork, E. V. Koonin, E. M. Zdobnov, I. V. Grigoriev, M. Lynch et J. L. Boore. 2011. «The Ecoresponsive Genome of *Daphnia pulex*». *Science*, vol. 331, no 6017, p. 555-561.

Colbourne, J. K., B. Robison, K. Bogart et M. Lynch. 2004. «Five hundred and twenty eight microsatellite markers for ecological genomic investigations using *Daphnia*». *Molecular Ecology Notes*, vol. 4, no 485-490.

Comai, L. 2000. «Genetic and epigenetic interactions in allopolyploid plants». *Plant Molecular Biology*, vol. 43, no 2, p. 387-399.

Crease, T. J. 1995. «Ribosomal DNA Evolution at the Population Level: Nucleotide Variation in Intergenic Spacer Arrays of *Daphnia pulex*». *Genetics*, vol. 141, no 4, p. 1327-1337.

Crease, T. J., S.-K. Lee, S.-L. Yu, K. Spitze, N. Lehman et M. Lynch. 1997. «Allozyme and mtDNA variation in populations of the *Daphnia pulex* complex from both sides of the Rocky Mountains». *Heredity*, vol. 79, no 3, p. 242-251.

Crease, T. J., D. J. Stanton et P. D. N. Hebert. 1989. «Polyphyletic Origins of Asexuality in *Daphnia pulex*. II. Mitochondrial-DNA Variation». *Evolution*, vol. 43, no 5, p. 1016-1026.

Cristescu, M., D. Innes, J. Stillman et T. Crease. 2008. «D- and L-lactate dehydrogenases during invertebrate evolution». *BMC Evolutionary Biology*, vol. 8, no 1, p. 268.

- Cristescu, M. E. A., J. K. Colbourne, J. Radivojc et M. Lynch. 2006. «A micro satellite-based genetic linkage map of the waterflea, *Daphnia pulex*: On the prospect of crustacean genomics». *Genomics*, vol. 88, no 4, p. 415-430.
- de Araujo, P. G., M. Rossi, E. M. de Jesus, N. L. Saccaro, D. Kajihara, R. Massa, J. M. de Felix, R. D. Drummond, M. C. Falco, S. M. Chabregas, E. C. Ulian, M. Menossi et M. A. V. Sluys. 2005. «Transcriptionally active transposable elements in recent hybrid sugarcane». *The Plant Journal*, vol. 44, no 5, p. 707-717.
- de Visser, J. A. G. M., J. Hermisson, G. P. Wagner, L. A. Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, G. Gibson, T. F. Hansen, D. Krakauer, R. C. Lewontin, C. Ofria, S. H. Rice, G. v. Dassow, A. Wagner et M. C. Whitlock. 2003. «Perspective: evolution and detection of genetic robustness». *Evolution*, vol. 57, no 9, p. 1959-1972.
- Delmotte, F., B. Sabater-Munoz, N. Prunier-Leterme, A. Latorre, P. Sunnucks, C. Rispe et J. C. Simon. 2003. «Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species». *Evolution*, vol. 57, no 6, p. 1291-1303.
- Doležel, J., J. Bartoš, H. Voglmayr et J. Greilhuber. 2003. «Nuclear DNA content and genome size of trout and human.». *Cytometry Part A*, vol. 51A, no 2, p. 127-128.
- Dolgin, E. S., et B. Charlesworth. 2008. «The Effects of Recombination Rate on the Distribution and Abundance of Transposable Elements». *Genetics*, vol. 178, no 4, p. 2169-2177.
- Doolittle, W. F., et C. Sapienza. 1980. «Selfish genes, the phenotype paradigm and genome evolution». *Nature*, vol. 284, no 5757, p. 601-603.

- Dowling, T. E., et C. L. Secor. 1997. «The role of hybridization in diversification of animals». *Annual Review of Ecology & Systematics*, vol. 28, p. 593.
- Dufresne, F. sous presse. «The History Of The *Daphnia Pulex* Complex: Asexuality, Hybridization, And Polyploidy». In *Intraspecific DNA variation in Crustacea*, C. Held, C. Schubart et Koenemann S. Boca Raton, FL.: CRC Press, Taylor & Francis Group.
- Dufresne, F., et P. D. N. Hebert. 1994. «Hybridization and Origins of Polyploidy». *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 258, no 1352, p. 141-146.
- , 1995. «Polyploidy and clonal diversity in an arctic cladoceran». *Heredity*, vol. 75, no 1, p. 45-53.
- , 1997. «Pleistocene glaciations and polyphyletic origins of polyploidy in an arctic cladoceran». *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 264, no 1379, p. 201-206.
- , 1998. «Temperature-related differences in life-history characteristics between diploid and polyploid clones of the *Daphnia pulex* complex». *Ecoscience*, vol. 5, no 4, p. 433-437.
- Duret, L., G. Marais et C. Biemont. 2000. «Transposons but Not Retrotransposons Are Located Preferentially in Regions of High Recombination Rate in *Caenorhabditis elegans*». *Genetics*, vol. 156, no 4, p. 1661-1669.
- Eanes, W. F. 1999. «Analysis of selection on enzyme polymorphisms». *Annual Review of Ecology and Systematics*, vol. 30, no 1, p. 301-326.

- Eickbush, D. G., et T. H. Eickbush. 1995. «Vertical Transmission of the Retrotransposable Elements R1 and R2 During the Evolution of the *Drosophila melanogaster* Species Subgroup». *Genetics*, vol. 139, no 2, p. 671-684.
- Eickbush, T. H., et D. G. Eickbush. 2007. «Finely Orchestrated Movements: Evolution of the Ribosomal RNA Genes». *Genetics*, vol. 175, no 2, p. 477-485.
- Engels, W. R., D. M. Johnson-Schlitz, W. B. Eggleston et J. Sved. 1990. «High-frequency P element loss in *Drosophila* is homolog dependent». *Cell*, vol. 62, no 3, p. 515-525.
- Esselink, G. D., H. Nybom et B. Vosman. 2004. «Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting—peak ratios) method». *TAG Theoretical and Applied Genetics*, vol. 109, no 2, p. 402-408.
- Farris, J. S., M. Källersjö, A. G. Kluge et C. Bult. 1994. «Testing Significance of Incongruence». *Cladistics*, vol. 10, no 3, p. 315-319.
- Feldman, M., B. Liu, G. Segal, S. Abbo, A. A. Levy et J. M. Vega. 1997. «Rapid elimination of low-copy DNA sequences in polyploid wheat: A possible mechanism for differentiation of homoeologous chromosomes». *Genetics*, vol. 147, no 3, p. 1381-1387.
- Felsenstein, J. 1985. «Confidence Limits on Phylogenies: An Approach Using the Bootstrap». *Evolution*, vol. 39, no 4, p. 783-791.
- Ferris, S. D., et G. S. Whitt. 1979. «Evolution of the differential regulation of duplicate genes after polyploidization ». *Journal of Molecular Evolution*, vol. 12, no 4, p. 267-317.

- Feschotte, C., X. Zhang et S. Wessler. 2002. «Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons». In *Mobile DNA II*, N.L. Craig, R. Craigie, M. Gellert et A.M. Lambowitz, p. 1147–1158. Washington, DC: American Society of Microbiology Press.
- Feschotte, C. d., et E. J. Pritham. 2007. «DNA Transposons and the Evolution of Eukaryotic Genomes». *Annual Review of Genetics*, vol. 41, no 1, p. 331-368.
- Finnegan, D. J. 1992. «Transposable elements». In *The genome of Drosophila melanogaster*, Dan L. Lindsley et Georgianna G. Zimm, p. 1096-1107. New York: Academic Press.
- Fischer, S. E. J., E. Wienholds et R. H. A. Plasterk. 2003. «Continuous Exchange of Sequence Information Between Dispersed Tc1 Transposons in the Caenorhabditis elegans Genome». *Genetics*, vol. 164, no 1, p. 127-134.
- Fontdevila, A. 2005. «Hybrid genome evolution by transposition». *Cytogenetic and Genome Research*, vol. 110, no 1-4, p. 49-55.
- Friz, C. T. 1968. «The nitrogen content of the water soluble proteins of three species of the free-living amoebae». *Life Sciences*, vol. 7, no 12, Part 1, p. 583-585.
- Galbraith, D. W., K. R. Harkins, J. M. Maddox, N. M. Ayres, D. P. Sharma et E. Firoozabady. 1983. «Rapid Flow Cytometric Analysis of the Cell Cycle in Intact Plant Tissues». *Science*, vol. 220, no 4601, p. 1049-1051.
- Gerstein, A. C., Chun, H-J. E., Grant, A., Otto, S.P. 2006. «Genomic Convergence toward Diploidy in Saccharomyces cerevisiae». *PLoS genetics*, vol. 2, no 9, p. e145.

- Ghiselli, F., L. Milani, V. Scali et M. Passamonti. 2007. «The *Leptynia hispanica* species complex (Insecta Phasmida): polyploidy, parthenogenesis, hybridization and more». *Molecular Ecology*, vol. 16, no 20, p. 4256-4268.
- Gibbs, M. J., J. S. Armstrong et A. J. Gibbs. 2000. «Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences». *Bioinformatics*, vol. 16, no 7, p. 573-582.
- Giessler, S., et C. C. Englbrecht. 2009. «Dynamic Reticulate Evolution in a *Daphnia* Multispecies Complex». *Journal of Experimental Zoology Part a-Ecological Genetics and Physiology*, vol. 311A, no 7, p. 530-U533.
- Gladyshev, E. A., M. Meselson et I. R. Arkhipova. 2008. «Massive Horizontal Gene Transfer in Bdelloid Rotifers». *Science*, vol. 320, no 5880, p. 1210-1213.
- Glass, S. K., A. Moszczyńska et T. J. Crease. 2008. «The effect of transposon Pokey insertions on sequence variation in the 28S rRNA gene of *Daphnia pulex*». *Genome*, vol. 51, p. 988-1000.
- González, J., K. Lenkov, M. Lipatov, J. M. Macpherson et D. A. Petrov. 2008. «High Rate of Recent Transposable Element Induced Adaptation in *Drosophila melanogaster*». *PLoS Biol*, vol. 6, no 10, p. e251.
- Gonzalez, J., et D. A. Petrov. 2009. «The adaptive role of transposable elements in the *Drosophila* genome». *Gene*, vol. 448, no 2, p. 124-133.
- Gower, J. C. 1966. «Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis». *Biometrika*, vol. 53, no 3/4, p. 325-338.

- Gregory, T. R. 2001. «The Bigger the C-Value, the Larger the Cell: Genome Size and Red Blood Cell Size in Vertebrates». *Blood Cells, Molecules, and Diseases*, vol. 27, no 5, p. 830-843.
- , 2002. «A bird's-eye view of the c-value enigma: genome size, cell size, and metabolic rate in the class aves». *Evolution*, vol. 56, no 1, p. 121-130.
- , 2005. «Synergy between sequence and size in large-scale genomics». *Nature Reviews Genetics*, vol. 6, p. 699-708.
- Gregory, T. R., et B. K. Mable. 2005. «Polyploidy in animals». In *The Evolution of the Genome.*, T.R. Gregory, p. 427-517: Elsevier Academic Press.
- Greilhuber, J., J. Doležal, M. A. Lysák et M. D. Bennett. 2005. «The Origin, Evolution and Proposed Stabilization of the Terms "Genome Size" and "C-Value" to Describe Nuclear DNA Contents». *Annals of Botany*, vol. 95, no 1, p. 255-260.
- Guindon, S. p., et O. Gascuel. 2003. «A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood». *Systematic Biology*, vol. 52, no 5, p. 696-704.
- Haag, C. R., S. J. McTaggart, A. Didier, T. J. Little et D. Charlesworth. 2009. «Nucleotide Polymorphism and Within-Gene Recombination in *Daphnia magna* and *D. pulex*, Two Cyclical Parthenogens». *Genetics*, vol. 182, no 1, p. 313-323.
- Han, J. S., et J. D. Boeke. 2004. «A highly active synthetic mammalian retrotransposon». *Nature*, vol. 429, no 6989, p. 314-318.

- Hanson, R. E., M. N. Islam-Faridi, C. F. Crane, M. S. Zwick, D. G. Czeschin, J. F. Wendel, T. D. McKnight, H. J. Price et D. M. Stelly. 2000. «Ty1-copia-retrotransposon Behavior in a Polyploid Cotton». *Chromosome Research*, vol. 8, no 1, p. 73-76.
- Hartigan, J. A., et M. A. Wong. 1979. «Algorithm AS 136: A K-Means Clustering Algorithm». *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no 1, p. 100-108.
- Hazzouri, K. M., A. Mohajer, S. I. Dejak, S. P. Otto et S. I. Wright. 2008. «Contrasting Patterns of Transposable-Element Insertion Polymorphism and Nucleotide Diversity in Autotetraploid and Allotetraploid Arabidopsis Species». *Genetics*, vol. 179, no 1, p. 581-592.
- Heath, L., E. van der Walt, A. Varsani et D. P. Martin. 2006. «Recombination Patterns in Aphthoviruses Mirror Those Found in Other Picornaviruses». *J. Virol.*, vol. 80, no 23, p. 11827-11832.
- Hebert, P., et T. Finston. 2001a. «Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group from the United States and Mexico». *Heredity*, vol. 87, p. 153 - 161.
- Hebert, P., S. Schwartz, R. Ward et T. Finston. 1993a. «Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group .1. Breeding systems of Canadian populations». *Heredity*, vol. 70, p. 148 - 161.
- Hebert, P. D. N. 1981. «Obligate Asexuality in *Daphnia*». *The American Naturalist*, vol. 117, no 5, p. 784-789.
- . 1995. *The Daphnia of North America : an illustrated fauna*. Guelph, Ontario: University of Guelph

- Hebert, P. D. N., M. J. Beaton, S. S. Schwartz et D. J. Stanton. 1989. «Polyphyletic Origins of Asexuality in *Daphnia pulex*. I. Breeding-System Variation and Levels of Clonal Diversity». *Evolution*, vol. 43, no 5, p. 1004-1015.
- Hebert, P. D. N., et T. Crease. 1983. «Clonal diversity in populations of *Daphnia pulex* reproducing by obligate parthenogenesis». *Heredity*, vol. 51, no 1, p. 353-369.
- Hebert, P. D. N., et T. J. Crease. 1980. «Clonal Coexistence in *Daphnia pulex* (Leydig): Another Planktonic Paradox». *Science*, vol. 207, no 4437, p. 1363-1365.
- Hebert, P. D. N., et T. L. Finston. 2001b. «Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group from the United States and Mexico». *Heredity*, vol. 87, p. 153-161.
- Hebert, P. D. N., S. S. Schwartz, R. D. Ward et T. L. Finston. 1993b. «Macrogeographic patterns of breeding system diversity in the *Daphnia pulex* group. I. Breeding systems of Canadian populations». *Heredity*, vol. 70, no 2, p. 148-161.
- Heier, C. R., et J. L. Dudycha. 2009. «Ecological speciation in a cyclic parthenogen: Sexual capability of experimental hybrids between *Daphnia pulex* and *Daphnia pulicaria*». *Limnology and Oceanography*, vol. 54, no 2, p. 492-502.
- Hickey, D. A. 1982. «Selfish DNA: a sexually-transmitted nuclear parasite». *Genetics*, vol. 101, no 3-4, p. 519-531.
- Hillis, D. M., C. Moritz, C. A. Porter et R. J. Baker. 1991. «Evidence for biased gene conversion in concerted evolution of ribosomal DNA». *Science*, vol. 251, no 4991, p. 308-310.

- Hobaek, A., L. J. Weider et H. G. Wolf. 1993. «Ecological genetics of Norwegian *Daphnia*. III. Clonal richness in an Arctic apomictic complex». *Heredity*, vol. 71, no 3, p. 323-330.
- Hoegg, S., et A. Meyer. 2005. «Hox clusters as models for vertebrate genome evolution». *Trends in Genetics*, vol. 21, no 8, p. 421-424.
- Holland, B. R., S. Benthin, P. J. Lockhart, V. Moulton et K. T. Huber. 2008. «Using supernetworks to distinguish hybridization from incomplete lineage sorting». *BMC Evolutionary Biology*, vol. 8, no 202.
- Holloway, Alisha K., David C. Cannatella, H. C. Gerhardt et David M. Hillis. 2006. «Polyploids with Different Origins and Ancestors Form a Single Sexual Polyploid Species». *The American Naturalist*, vol. 167, no 4, p. E88-E101.
- Hua-Van, A., A. Le Rouzic, C. Maisonhaute et P. Capy. 2005. «Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences». *Cytogenetic and Genome Research*, vol. 110, no 1-4, p. 426-440.
- Hubbell, S. P. 2005. «Neutral theory in community ecology and the hypothesis of functional equivalence». *Functional Ecology*, vol. 19, no 1, p. 166-172.
- Hudson, R. R., et N. L. Kaplan. 1985. «Statistical properties of the number of recombination events in the history of a sample of dna sequences». *Genetics*, vol. 111, no 1, p. 147-164.
- Huelsenbeck, J. P., et F. Ronquist. 2005. «Bayesian Analysis of Molecular Evolution Using MrBayes». In *Statistical Methods in Molecular Evolution*, R. Nielsen, p. 183-232: Springer.

- Hughes, M. K., et A. L. Hughes. 1993. «Evolution of Duplicate Genes in a Tetraploid Animal, *Xenopus laevis*». *Molecular Biology and Evolution*, vol. 10, no 6, p. 1360-1369.
- Huson, D. H., et D. Bryant. 2006. «Application of Phylogenetic Networks in Evolutionary Studies». *Molecular Biology and Evolution*, vol. 23, no 2, p. 254-267.
- Ihaka, R., et R. Gentleman. 1996. «R: a language for data analysis and graphics.». *Journal of Computational and Graphical Statistics*, vol. 5, no 3, p. 299-314.
- Innes, D. J., C. J. Fox et G. L. Winsor. 2000. «Avoiding the cost of males in obligately asexual *Daphnia pulex* (Leydig)». *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 267, no 1447, p. 991-997.
- Innes, D. J., et P. D. N. Hebert. 1988. «The Origin and Genetic Basis of Obligate Parthenogenesis in *Daphnia pulex*». *Evolution*, vol. 42, no 5, p. 1024-1035.
- Innes, D. J., S. S. Schwartz et P. D. N. Hebert. 1986. «Genotypic diversity and variation in mode of reproduction among populations in the *Daphnia pulex* group». *Heredity*, vol. 57, no 3, p. 345-355.
- Innes, D. J., et D. R. Singleton. 2000. «Variation in allocation to sexual and asexual reproduction among clones of cyclically parthenogenetic *Daphnia pulex* (Crustacea: Gladocera)». *Biological Journal of the Linnean Society*, vol. 71, no 4, p. 771-787.
- International Human Genome Sequencing Consortium. 2001. «Initial sequencing and analysis of the human genome». *Nature*, vol. 409, no 6822, p. 860-921.

Jaccard, P. 1908. «Nouvelles recherches sur la distribution florale.». *Bulletin de la société Vaudoise des Sciences Naturelles*, vol. 44, p. 223-270.

Jarne, P., et P. J. L. Lagoda. 1996. «Microsatellites, from molecules to populations and back». *Trends in Ecology & Evolution*, vol. 11, no 10, p. 424-429.

Johnson, S. G. 2006. «Geographic ranges, population structure, and ages of sexual and parthenogenetic snail lineages». *Evolution*, vol. 60, no 7, p. 1417-1426.

Joly, S., Patricia A. McLenachan et Peter J. Lockhart. 2009. «A Statistical Approach for Distinguishing Hybridization and Incomplete Lineage Sorting». *The American Naturalist*, vol. 174, no 2, p. E54-E70.

Kapitonov, V. V., et J. Jurka. 2001. «Rolling-circle transposons in eukaryotes». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no 15, p. 8714-8719.

Kashkush, K., M. Feldman et A. A. Levy. 2002. «Gene loss, silencing and activation in a newly synthesized wheat allotetraploid». *Genetics*, vol. 160, no 4, p. 1651-1659.

----- . 2003. «Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat». *Nature Genetics*, vol. 33, no 1, p. 102-106.

Kearney, M. 2005. «Hybridization, glaciation and geographical parthenogenesis». *Trends in Ecology & Evolution*, vol. 20, no 9, p. 495-502.

Kidwell, M. G. 2002. «Transposable elements and the evolution of genome size in eukaryotes». *Genetica*, vol. 115, no 1, p. 49-63.

- Kidwell, M. G., et D. Lisch. 1997. «Transposable elements as sources of variation in animals and plants». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no 15, p. 7704-7711.
- Kidwell, M. G., et D. R. Lisch. 2000. «Transposable elements and host genome evolution». *Trends in Ecology & Evolution*, vol. 15, no 3, p. 95-99.
- , 2001. «Perspective: transposable elements, parasitic DNA, and genome evolution». *Evolution*, vol. 55, no 1, p. 1-24.
- Kim, J. M., S. Vanguri, J. D. Boeke, A. Gabriel et D. F. Voytas. 1998. «Transposable Elements and Genome Organization: A Comprehensive Survey of Retrotransposons Revealed by the Complete *Saccharomyces cerevisiae* Genomeâ€™s Sequence». *Genome Research*, vol. 8, no 5, p. 464-478.
- Kingman, J. F. C. 1982. «The coalescent». *Stochastic Processes and their Applications*, vol. 13, no 3, p. 235-248.
- Kishino, H., ., T. Miyata et M. Hasegawa. 1990. «Maximum likelihood inference of protein phylogeny and the origin of chloroplasts». *Journal of Molecular Evolution*, vol. 3, no 2, p. 151-160.
- Korpelainen, H., M. Ketola et J. Hietala. 1997. «Somatic polyploidy examined by flow cytometry in *Daphnia*». *Journal of Plankton Research*, vol. 19, no 12, p. 2031-2040.
- Kosakovskiy, S. L., D. Posada, M. B. Gravenor, C. H. Woelk et S. D. W. Frost. 2006. «GARD: a genetic algorithm for recombination detection». *Bioinformatics*, vol. 22, no 24, p. 3096-3098.

- Kovarik, A., R. Matyasek, K. Y. Lim, K. Skalicka, B. Koukalova, S. Knapp, M. Chase et A. R. Leitch. 2004. «Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids». *Biological Journal of the Linnean Society*, vol. 82, no 4, p. 615-625.
- Labrador, M., et A. Fontdevila. 1994. «High transposition rates of *Osvaldo*, a new *Drosophila buzzatii* retrotransposon». *Molecular and General Genetics MGG*, vol. 245, no 6, p. 661-674.
- Lampe, D. J., B. J. Akerley, E. J. Rubin, J. J. Mekalanos et H. M. Robertson. 1999. «Hyperactive transposase mutants of the Himar1 mariner transposon». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no 20, p. 11428-11433.
- Langley, C. H., E. Montgomery, R. Hudson, N. Kaplan et B. Charlesworth. 1988. «On the role of unequal exchange in the containment of transposable element copy number.». *Genetical Research*, vol. 52, p. 223-235.
- Lawrence, J. G., H. Ochman et D. L. Hartl. 1992. «The Evolution of Insertion Sequences Within Enteric Bacteria». *Genetics*, vol. 131, no 1, p. 9-20.
- Le Rouzic, A., S. Dupas et P. Capy. 2007. «Genome ecosystem and transposable elements species». *Gene*, vol. 390, no 1-2, p. 214-220.
- Legendre, P., et L. Legendre. 1998. *Numerical ecology*, 2nd English edition. Amsterdam: Elsevier Science, 853 p.
- Lehman, N., M. E. Pfrender, P. A. Morin, T. J. Crease et M. Lynch. 1995. «A hierarchical molecular phylogeny within the genus *Daphnia*». *Molecular Phylogenetics and Evolution*, vol. 4, no 4, p. 395-407.

- Leitch, A. R., et I. J. Leitch. 2008. «Genomic Plasticity and the Diversity of Polyploid Plants». *Science*, vol. 320, no 5875, p. 481-483.
- Leitch, I. J., et M. D. Bennett. 2004. «Genome downsizing in polyploid plants». *Biological Journal of the Linnean Society*, vol. 82, no 4, p. 651-663.
- Leonardo, m. E., et S. V. Nuzdhin. 2002. «Intracellular battlegrounds: conflict and cooperation between transposable elements». *Genetics Research*, vol. 80, p. 7.
- Lerman, D. N., P. Michalak, A. B. Helin, B. R. Bettencourt et M. E. Feder. 2003. «Modification of Heat-Shock Gene Expression in *Drosophila melanogaster* Populations via Transposable Elements». *Molecular Biology and Evolution*, vol. 20, no 1, p. 135-144.
- Levin, D. A. 1983. «Polyploidy and Novelty in Flowering Plants». *The American Naturalist*, vol. 122, no 1, p. 1-25.
- Lewis, A. P., et J. F. Y. Brookfield. 1987. «Movement of *Drosophila melanogaster* transposable elements other than P elements in a P-M hybrid dysgenic cross». *Molecular and General Genetics MGG*, vol. 208, no 3, p. 506-510.
- Lewontin, R. C., et L. C. Birch. 1966. «Hybridization as a Source of Variation for Adaptation to New Environments». *Evolution*, vol. 20, no 3, p. 315-336.
- Li, W.-H. 1980. «Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes.». *Genetics*, vol. 95, no 1, p. 237-258.

- Li, W. L., A. E. Tucker, W. Sung, W. K. Thomas et M. Lynch. 2009. «Extensive, Recent Intron Gains in *Daphnia* Populations». *Science*, vol. 326, no 5957, p. 1260-1262.
- Librado, P., et J. Rozas. 2009. «DnaSP v5: A software for comprehensive analysis of DNA polymorphism data». *Bioinformatics*, vol. 25, p. 1451-1452.
- Little, T. J., et P. D. N. Hebert. 1997. «Clonal diversity in high arctic ostracodes». *Journal of Evolutionary Biology*, vol. 10, p. 233-252.
- Liu, B., et J. F. Wendel. 2003. «Epigenetic phenomena and the evolution of plant allopolyploids». *Molecular Phylogenetics and Evolution*, vol. 29, no 3, p. 365-379.
- Liu, D., C. You, S. Liu, L. Liu, W. Duan, S. Chen, J. Yan et Y. Liu. 2009. «Characterization of a Novel Tc1-Like Transposon From Bream (*Cyprinidae*, *Megalobrama*) and Its Genetic Variation in the Polyploidy Progeny of Bream-Red Crucian Carp Crosses». *Journal of Molecular Evolution*, vol. 69, no 4, p. 395-403.
- Lokki, J., et A. Saura. 1980. «Polyploidy in insect evolution». In *Polyploidy: Biological Relevance*, W.H. Lewis, p. 277-312. New York: Plenum.
- Lundmark, M. 2006. «Polyploidization, hybridization and geographical parthenogenesis». *Trends in Ecology & Evolution*, vol. 21, no 1, p. 9-9.
- Lynch, M. 2007. *The origins of genome architecture*. Sunderland, MA, USA: Sinauer Associates, Inc. p.

- Lynch, M., A. Seyfert, B. Eads et E. Williams. 2008. «Localization of the genetic determinants of meiosis suppression in *Daphnia pulex*». *Genetics*, vol. 180, no 1, p. 317-327.
- Lyngsø, R., Y. Song et J. Hein. 2005. «Minimum Recombination Histories by Branch and Bound». In *Algorithms in Bioinformatics*, Rita Casadio et Gene Myers, p. 239-250: Springer Berlin / Heidelberg. En ligne. <http://dx.doi.org/10.1007/11557067_20>.
- Mable, B. K. 2004. «'Why polyploidy is rarer in animals than in plants': myths and mechanisms». *Biological Journal of the Linnean Society*, vol. 82, no 4, p. 453-466.
- Machado, C. A., R. M. Kliman, J. A. Markert et J. Hey. 2002. «Inferring the History of Speciation from Multilocus DNA Sequence Data: The Case of *Drosophila pseudoobscura* and Close Relatives». *Mol Biol Evol*, vol. 19, no 4, p. 472-488.
- Maddison, W. P., et D. R. Maddison. 2010. *Mesquite: a modular system for evolutionary analysis Version 2.73*: <http://mesquiteproject.org> p.
- Madlung, A., R. W. Masuelli, B. Watson, S. H. Reynolds, J. Davison et L. Comai. 2002. «Remodeling of DNA Methylation and Phenotypic and Transcriptional Changes in Synthetic Arabidopsis Allotetraploids». *Plant Physiol.*, vol. 129, no 2, p. 733-746.
- Madlung, A., A. P. Tyagi, B. Watson, H. M. Jiang, T. Kagochi, R. W. Doerge, R. Martienssen et L. Comai. 2005. «Genomic changes in synthetic Arabidopsis polyploids». *Plant Journal*, vol. 41, no 2, p. 221-230.
- Mallet, J. 2001. «Species, Concept of». In *Encyclopedia of Biodiversity*, S A Levin, p. 427-440: Academic Press.

- Marková, S., F. Dufresne, D. J. Rees, M. Cerný et P. Kotlík. 2007. «Cryptic intercontinental colonization in water fleas *Daphnia pulicaria* inferred from phylogenetic analysis of mitochondrial DNA variation». *Molecular Phylogenetics and Evolution*, vol. 44, no 1, p. 42-52.
- Martin, D., et E. Rybicki. 2000. «RDP: detection of recombination amongst aligned sequences». *Bioinformatics*, vol. 16, no 6, p. 562-563.
- Martin, D. P., C. Williamson et D. Posada. 2005. «RDP2: recombination detection and analysis from sequence alignments». *Bioinformatics*, vol. 21, no 2, p. 260-262.
- Masterson, J. 1994. «Stomatal Size in Fossil Plants: Evidence for Polyploidy in Majority of Angiosperms». *Science*, vol. 264, no 5157, p. 421-424.
- Matzke, M. A., et A. J. M. Matzke. 1998. «Polyploidy and transposons». *Trends in Ecology & Evolution*, vol. 13, no 6, p. 241-241.
- Mayden, R. L. 1997. «A hierarchy of species concepts: the denouement in the saga of the species problem». In *Species: the Units of Biodiversity*, M F Claridge, H A Dawah et M R Wilson, p. 381–424. London: Chapman & Hall Ltd.
- Maynard Smith, J. 1992. «Analyzing the mosaic structure of genes». *Journal of Molecular Evolution*, vol. 34, no 2, p. 126-129.
- Mayr, E. 1963. *Animal species and evolution*. Cambridge, MA: Harvard University Press p.

- McClintock, B. 1984. «The significance of responses of the genome to challenge». *Science*, vol. 226, no 4676, p. 792-801.
- McDonald, J. F. 1995. «Transposable elements: possible catalysts of organismic evolution». *Trends in Ecology & Evolution*, vol. 10, no 3, p. 123-126.
- , 1998. «Transposable elements, gene silencing and macroevolution». *Trends in Ecology and Evolution*, vol. 13, no 3, p. 94-95.
- McTaggart, S. J., J. L. Dudycha, A. Omilian et T. J. Crease. 2007. «Rates of Recombination in the Ribosomal DNA of Apomictically Propagated *Daphnia obtusa* Lines». *Genetics*, vol. 175, no 1, p. 311-320.
- Mergeay, J., X. Aguilera, S. Declerck, A. Petrusek, T. Huyse et L. De Meester. 2008. «The genetic legacy of polyploid Bolivian *Daphnia*: the tropical Andes as a source for the North and South American *D. pulicaria* complex». *Molecular Ecology*, vol. 17, no 7, p. 1789-1800.
- Milne, I., D. Lindner, M. Bayer, D. Husmeier, G. McGuire, D. F. Marshall et F. Wright. 2009. «TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops». *Bioinformatics*, vol. 25, no 1, p. 126-127.
- Montgomery, E., B. Charlesworth et C. H. Langley. 1987. «A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*». *Genetical Research*, vol. 49, p. 31-41.
- Moore, W. S. 1995. «Inferring Phylogenies from mtDNA Variation: Mitochondrial-Gene Trees Versus Nuclear-Gene Trees». *Evolution*, vol. 49, no 4, p. 718-726.

- Nei, M. 1978. «Estimation of average heterozygosity and genetic distance from a small number of individuals». *Genetics*, vol. 89, no 583-590.
- O'Neill, R. J. W., M. J. O'Neill et J. A. M. Graves. 1998. «Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid». *Nature*, vol. 393, p. 68-72.
- Ohno, S. 1970. *Evolution by Gene Duplication*. New York: Springer-Verlag p.
- Omilian, A. R., M. E. A. Cristescu, J. L. Dudycha et M. Lynch. 2006. «Ameiotic recombination in asexual lineages of *Daphnia*». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no 49, p. 18638-18643.
- Omilian, A. R., et M. Lynch. 2009. «Patterns of Intraspecific DNA Variation in the *Daphnia* Nuclear Genome». *Genetics*, p. genetics.108.099549.
- Omilian, A. R., D. G. Scofield et M. Lynch. 2008. «Intron presence-absence polymorphisms in *Daphnia*». *Molecular Biology and Evolution*, vol. 25, no 10, p. 2129-2139.
- Orgel, L. E., et F. H. C. Crick. 1980. «Selfish DNA: the ultimate parasite». *Nature*, vol. 284, no 5757, p. 604-607.
- Orita, M., Y. Suzuki, T. Sekiya et K. Hayashi. 1989. «Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction.». *Genomics*, vol. 5, no 4, p. 874-879.
- Osborn, T. C., J. C. Pires, J. A. Birchler, D. L. Auger, Z. J. Chen, H. S. Lee, L. Comai, A. Madlung, R. W. Doerge, V. Colot et R. A. Martienssen. 2003.

- «Understanding mechanisms of novel gene expression in polyploids». *Trends in Genetics*, vol. 19, no 3, p. 141-147.
- Otto, S. P. 2007. «The Evolutionary Consequences of Polyploidy». *Cell*, vol. 131, no 3, p. 452-462.
- Otto, S. P., et J. Whitton. 2000. «Polyploid incidence and evolution». *Annual Review of Genetics*, vol. 34, p. 401-437.
- Pace, J. K., C. m. Gilbert, M. S. Clark et C. d. Feschotte. 2008. «Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods». *Proceedings of the National Academy of Sciences*, vol. 105, no 44, p. 17023-17028.
- Padidam, M., S. Sawyer et C. M. Fauquet. 1999. «Possible Emergence of New Geminiviruses by Frequent Recombination». *Virology*, vol. 265, no 2, p. 218-225.
- Paland, S., J. K. Colbourne et M. Lynch. 2005. «Evolutionary history of contagious asexuality in *Daphnia pulex*». *Evolution*, vol. 59, no 4, p. 800-813.
- Paland, S., et M. Lynch. 2006. «Transitions to asexuality result in excess amino acid substitutions». *Science*, vol. 311, no 5763, p. 990-992.
- Pamilo, P., et M. Nei. 1988. «Relationships between gene trees and species trees». *Mol Biol Evol*, vol. 5, no 5, p. 568-583.
- Pardue, M. L., S. Rashkova, E. Casacuberta, P. G. DeBaryshe, J. A. George et K. L. Traverse. 2005. «Two retrotransposons maintain telomeres in *Drosophila*». *Chromosome Research*, vol. 13, no 5, p. 443-453.

- Parisod, C., K. Alix, J. Just, M. Petit, V. Sarilar, C. Mhiri, M. Ainouche, B. Chalhoub et M. A. Grandbastien. 2010. «Impact of transposable elements on the organization and function of allopolyploid genomes». *New Phytologist*, vol. 186, no 1, p. 37-45.
- Parisod, C., A. Salmon, T. Zerjal, M. Tenaillon, M.-A. Grandbastien et M. Ainouche. 2009. «Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in *Spartina*». *New Phytologist*, vol. 184, no 4, p. 1003-1015.
- Penton, E. H., et T. J. Crease. 2004. «Evolution of the transposable element Pokey in the ribosomal DNA of species in the subgenus *Daphnia* (Crustacea : Cladocera)». *Molecular Biology and Evolution*, vol. 21, no 9, p. 1727-1739.
- Penton, E. H., B. W. Sullender et T. J. Crease. 2002. «Pokey, a new DNA transposon in *Daphnia* (Cladocera : Crustacea)». *Journal of Molecular Evolution*, vol. 55, no 6, p. 664-673.
- Petit, M., K. Lim, E. Julio, C. Poncet, F. Dorlhac de Borne, A. Kovarik, A. Leitch, M.-A. Grandbastien et C. Mhiri. 2007. «Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (<i>Nicotiana tabacum&/i>)». *Molecular Genetics and Genomics*, vol. 278, no 1, p. 1-15.
- Petrov, D. A. 2001. «Evolution of genome size: new approaches to an old problem». *Trends in Genetics*, vol. 17, no 1, p. 23-28.
- Petrov, D. A., J. L. Schutzman, D. L. Hartl et E. R. Lozovskaya. 1995. «Diverse transposable elements are mobilized in hybrid dysgenesis in *Drosophila virilis*». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no 17, p. 8050-8054.

- Pfrender, M. E., K. Spitze et N. Lehman. 2000. «Multi-locus genetic evidence for rapid ecologically based speciation in *Daphnia*». *Molecular Ecology*, vol. 9, no 11, p. 1717-1735.
- Pidpala, O., A. Yatsishina et L. Lukash. 2008. «Human mobile genetic elements: Structure, distribution and functional role». *Cytology and Genetics*, vol. 42, no 6, p. 420-430.
- Pimm, L. S. 1984. «The complexity and stability of ecosystems». *Nature*, vol. 307, no 5949, p. 321-325.
- Piskur, J., E. Rozpedowska, S. Polakova, A. Merico et C. Compagno. 2006. «How did *Saccharomyces* evolve to become a good brewer?». *Trends in Genetics*, vol. 22, no 4, p. 183-186.
- Pond, S. L. K., et S. D. W. Frost. 2005. «Datamonkey: rapid detection of selective pressure on individual sites of codon alignments». *Bioinformatics*, vol. 21, no 10, p. 2531-2533.
- Posada, D., et K. A. Crandall. 2001. «Evaluation of methods for detecting recombination from DNA sequences: Computer simulations». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no 24, p. 13757-13762.
- Pritham, E. J. 2009. «Transposable Elements and Factors Influencing their Success in Eukaryotes». *Journal of Heredity*.
- Pritham, E. J., T. Putliwala et C. Feschotte. 2007. «Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses». *Gene*, vol. 390, no 1-2, p. 3-17.

- Promislow, D. E. L., I. K. Jordan et J. E. McDonald. 1999. «Genomic demography: a life-history analysis of transposable element evolution». *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 266, no 1428, p. 1555-1560.
- R Development Core Team. 2004. «R: A language and environment for statistical computing.».
- Ramsey, J., et D. W. Schemske. 1998. «Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants». *Annual Review of Ecology and Systematics*, vol. 29, p. 467-501.
- Rasch, E. M. 1985. «DNA "standards" and the range of accurate DNA estimates by Feulgen absorption microspectrophotometry.» *Progress in clinical and biological research*, vol. 196, no 137-166.
- Rasch, E. M., H. J. Barr et R. W. Rasch. 1971. «The DNA content of sperm of *Drosophila melanogaster*». *Chromosoma*, vol. 33, no 1, p. 1-18.
- Rees, D. J., F. Dufresne, H. I. n. GIÃ©met et C. Belzile. 2007. «Amphipod genome sizes: first estimates for Arctic species reveal genomic giants». *Genome*, vol. 50, no 2, p. 151-158.
- Rho, M., S. Schaack, X. Gao, S. Kim, M. Lynch et H. Tang. 2010. «LTR retroelements in the genome of *Daphnia pulex*». *BMC Genomics*, vol. 11, no 1, p. 425.
- Rieseberg, L. H., M. A. Archer et R. K. Wayne. 1999. «Transgressive segregation, adaptation and speciation». *Heredity*, vol. 83, no 4, p. 363-372.

- Rizzon, C., G. Marais, M. Gouy et C. Biemont. 2002. «Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome». *Genome Res*, vol. 12, no 3, p. 400-407.
- Rizzon, C., E. Martin, G. Marais, L. Duret, L. Segalat et C. Biemont. 2003. «Patterns of Selection Against Transposons Inferred From the Distribution of Tc1, Tc3 and Tc5 Insertions in the mut-7 Line of the Nematode *Caenorhabditis elegans*». *Genetics*, vol. 165, no 3, p. 1127-1135.
- Ronquist, F., et J. Huelsenbeck. 2003. «MRBAYES 3: Bayesian phylogenetic inference under mixed models». *Bioinformatics*, vol. 19, p. 1572 - 1574.
- Roy, S. W., et W. Gilbert. 2005. «Rates of intron loss and gain: Implications for early eukaryotic evolution». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no 16, p. 5773-5778.
- , 2006. «The evolution of spliceosomal introns: patterns, puzzles and progress». *Nature Reviews Genetics*, vol. 7, p. 211-221.
- Saitou, N., et M. Nei. 1987. «The neighbor-joining method: a new method for reconstructing phylogenetic trees». *Molecular Biology and Evolution*, vol. 4, no 4, p. 406-425.
- Salminen, M. O., J. K. Carr, D. S. Burke et F. E. McCutchan. 1995. «Identification of Breakpoints in Intergenotypic Recombinants of HIV Type 1 by Bootscanning». *AIDS Research and Human Retroviruses*, vol. 11, no 11, p. 1423-1425.
- Salmon, A. 2005. «Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina*(Poaceae)». *Molecular Ecology*, vol. 14, no 4, p. 1163.

- Sang, T., et D. M. Zhang. 1999. «Reconstructing hybrid speciation using sequences of low copy nuclear genes: Hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies». *Systematic Botany*, vol. 24, no 2, p. 148-163.
- Sang, T., et Y. Zhong. 2000. «Testing Hybridization Hypotheses Based on Incongruent Gene Trees». *Systematic Biology*, vol. 49, no 3, p. 422-434.
- SanMiguel, P., et J. L. Bennetzen. 1998. «Evidence that a Recent Increase in Maize Genome Size was Caused by the Massive Amplification of Intergene Retrotransposons». *Annals of Botany*, vol. 82, no suppl 1, p. 37-44.
- SanMiguel, P., B. S. Gaut, A. Tikhonov, Y. Nakajima et J. L. Bennetzen. 1998. «The paleontology of intergene retrotransposons of maize. ». *Nature Review Genetics*, vol. 20, no 1, p. 43-53.
- Sarkar, A., C. Sim, Y. S. Hong, J. R. Hogan, M. J. Fraser, H. M. Robertson et F. H. Collins. 2003. «Molecular evolutionary analysis of the widespread piggyBac transposon family and related "domesticated" sequences». *Molecular Genetics and Genomics*, vol. 270, no 2, p. 173-180.
- Saura, A., J. Lokki et E. Suomalainen. 1993. «Origin of Polyploidy in Parthenogenetic Weevils». *Journal of Theoretical Biology*, vol. 163, no 4, p. 449-456.
- Schaack, S., E. Choi, M. Lynch et E. Pritham. 2010a. «DNA transposons and the role of recombination in mutation accumulation in *Daphnia pulex*». *Genome Biology*, vol. 11, no 4, p. R46.
- Schaack, S., E. J. Pritham, A. Wolf et M. Lynch. 2010b. «DNA transposon dynamics in populations of *Daphnia pulex* with and without sex». *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no 1692, p. 2381-2387.

- Scheid, O. M., L. Jakovleva, K. Afsar, J. Maluszynska et J. Paszkowski. 1996. «A change of ploidy can modify epigenetic silencing». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no 14, p. 7114-7119.
- Seehausen, O. 2004. «Hybridization and adaptive radiation». *Trends in Ecology & Evolution*, vol. 19, no 4, p. 198-207.
- Shaked, H., K. Kashkush, H. Ozkan, M. Feldman et A. A. Levy. 2001. «Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat». *Plant Cell*, vol. 13, no 8, p. 1749-1759.
- Shan, X., Z. Liu, Z. Dong, Y. Wang, Y. Chen, X. Lin, L. Long, F. Han, Y. Dong et B. Liu. 2005. «Mobilization of the Active MITE Transposons mPing and Pong in Rice by Introgression from Wild Rice (*Zizania latifolia* Griseb.)». *Molecular Biology and Evolution*, vol. 22, no 4, p. 976-990.
- Sharma, A., K. L. Schneider et G. G. Presting. 2008. «Sustained retrotransposition is mediated by nucleotide deletions and interelement recombinations». *Proceedings of the National Academy of Sciences*, vol. 105, no 40, p. 15470-15474.
- Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki et H. Ishikawa. 2000. «Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS». *Nature*, vol. 407, no 6800, p. 81-86.
- Shimodaira, H., et M. Hasegawa. 1999. «Multiple comparisons of log-likelihoods with applications to phylogenetic inference». *Molecular Biology and Evolution*, vol. 16, no 8, p. 1114-1116.

- Silva, J., E. Loreto et J. Clark. 2004. «Factors that affect the horizontal transfer of transposable elements». *Current issues in molecular biology*, vol. 6, no 1, p. 57-71.
- Simon, J.-C., F. Delmotte, C. Rispe et T. J. Crease. 2003. «Phylogenetic relationships between parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals». *Biological Journal of the Linnean Society*, vol. 79, no 1, p. 151-163.
- Sokal, R. R., et F. J. Rohlf. 1995. *Biometry: the principles and practice of statistics in biological Research*, 3e. New York: W. H. Freeman and Co., 887 p.
- Soltis, D. E., E. V. Mavrodiev, J. J. Doyle, J. Rauscher et P. S. Soltis. 2008. «ITS and ETS sequence data and phylogeny reconstruction in allopolyploids and hybrids». *Systematic Biology*, vol. 33, no 1, p. 7-20.
- Soltis, D. E., et P. S. Soltis. 1995. «The Dynamic Nature of Polyploid Genomes». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no 18, p. 8089-8091.
- , 1999. «Polyploidy: recurrent formation and genome evolution». *Trends in Ecology & Evolution*, vol. 14, no 9, p. 348-352.
- Soltis, P. S., et D. E. Soltis. 2000. «The role of genetic and genomic attributes in the success of polyploids». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no 13, p. 7051-7057.
- Song, Y. S., et J. Hein. 2003. «Parsimonious Reconstruction of Sequence Evolution and Haplotype Blocks». In *Algorithms in Bioinformatics*, Gary Benson et Roderic Page, p. 287-302: Springer Berlin / Heidelberg. En ligne. <http://dx.doi.org/10.1007/978-3-540-39763-2_22>.

- . 2005. «Constructing Minimal Ancestral Recombination Graphs». *Journal of Computational Biology*, vol. 12, no 2, p. 147-169.
- Stebbins, G. L. 1985. «Polyploidy, Hybridization, and the Invasion of New Habitats». *Annals of the Missouri Botanical Garden*, vol. 72, no 4, p. 824-832.
- Stenberg, P., M. Lundmark, S. Knutelski et A. Saura. 2003. «Evolution of Clonality and Polyploidy in a Weevil System». *Molecular Biology and Evolution*, vol. 20, no 10, p. 1626-1632.
- Straughan, D., et N. Lehman. 2000. «Genetic differentiation among Oregon lake populations of the *Daphnia pulex* species complex». *Journal of Heredity*, vol. 91, no 1, p. 8-17.
- Sullender, B. W., et T. J. Crease. 2001a. «The behavior of a *Daphnia pulex* transposable element in cyclically and obligately parthenogenetic populations». *Journal of Molecular Evolution*, vol. 53, no 1, p. 63-69.
- . 2001b. «The Behavior of a *Daphnia pulex*; Transposable Element in Cyclically and Obligately Parthenogenetic Populations». *Journal of Molecular Evolution*, vol. 53, no 1, p. 63-69.
- Swofford, D. 2003. *PAUP* - Phylogenetic analysis using Parsimony and other methods, version 4.0*: sunderland: Sinauer p.
- Tamura, K., J. Dudley, M. Nei et S. Kumar. 2007. «MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0». *Mol Biol Evol*, vol. 24, no 8, p. 1596-1599.

- Templeton, A. R. 1981. «Mechanisms of Speciation - A Population Genetic Approach». *Annual Review of Ecology and Systematics*, vol. 12, no 1, p. 23-48.
- Tilman, D. 1996. «Biodiversity: Population Versus Ecosystem Stability». *Ecology*, vol. 77, no 2, p. 350-363.
- Tomiuk, J., B. Guldbrandtsen et V. Loeschcke. 2009. «Genetic similarity of polyploids: a new version of the computer program POPDIST (version 1.2.0) considers intraspecific genetic differentiation». *Molecular Ecology Resources*, vol. 9, no 5, p. 1364-1368.
- Tomiuk, J., et V. Loeschcke. 1991. «A New Measure of Genetic Identity between Populations of Sexual and Asexual Species». *Evolution*, vol. 45, no 7, p. 1685-1694.
- , 1992. «Evolution of parthenogenesis in the *Otiorhynchus scaber* complex». *Heredity*, vol. 68, p. 391-397.
- Turgeon, J., et P. D. N. Hebert. 1994. «Evolutionary Interactions Between Sexual and All-Female Taxa of Cyprinotus (Ostracoda: Cyprididae)». *Evolution*, vol. 48, no 6, p. 1855-1865.
- Valizadeh, P., et T. Crease. 2008. «The Association Between Breeding System and Transposable Element Dynamics in *Daphnia Pulex*». *Journal of Molecular Evolution*, vol. 66, no 6, p. 643-654.
- Van Der Sluijs, P., M. Hull, A. Zahraoui, A. Tavitian, B. Goud et I. Mellman. 1991. «The small GTP-binding protein rab4 is associated with early endosomes». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no 14, p. 6313-6317.

- Venner, S., C. Feschotte et C. Biéumont. 2009. «Dynamics of transposable elements: towards a community ecology of the genome». *Trends in Genetics*, vol. 25, no 7, p. 317-323.
- Vergilino, R., C. Belzile et F. Dufresne. 2009. «Genome size evolution and polyploidy in the *Daphnia pulex* complex (Cladocera: Daphniidae)». *Biological Journal of the Linnean Society*, vol. 97, no 1, p. 68-79.
- Vergilino, R., S. Markova, M. Ventura, M. Manca et F. Dufresne. 2011. «Reticulate evolution of the *Daphnia pulex* complex as revealed by nuclear markers ». *Molecular Ecology*, vol. 20, no 6, p. 1191-1207.
- Viktorov, A. G. 1997. «Diversity of polyploid races in the family Lumbricidae». *Soil Biology & Biochemistry*, vol. 29, no 3-4, p. 217-221.
- Vinogradov, A. E. 1995. «Nucleotypic Effect in Homeotherms: Body-Mass-Corrected Basal Metabolic Rate of Mammals is Related to Genome Size». *Evolution*, vol. 49, no 6, p. 1249-1259.
- , 1997. «Nucleotypic Effect in Homeotherms: Body-Mass Independent Resting Metabolic Rate of Passerine Birds is Related to Genome Size». *Evolution*, vol. 51, no 1, p. 220-225.
- Vitte, C., et O. Panaud. 2005. «LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model.». *Cytogenetics and Genome Research*, vol. 110, no 1-4, p. 121-129.
- Voglmayr, H. 2000. «Nuclear DNA Amounts in Mosses (Musci)». *Annals of Botany*, vol. 85, no 4, p. 531-546.

- Volff, J.-N. 2006. «Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes». *Bioessays*, vol. 28, no 9, p. 913-922.
- Ward, R. D., M. A. Bickerton, T. Finston et P. D. N. Hebert. 1994. «Geographical cline in breeding systems and ploidy levels in European populations of *Daphnia pulex*». *Heredity*, vol. 73, no 5, p. 532-543.
- Weider, L. J. 1987. «Life history variation among low-arctic clones of obligately parthenogenetic *Daphnia pulex*: a diploid-polyploid complex». *Oecologia*, vol. 73, no 2, p. 251-256.
- Weider, L. J., et P. D. N. Hebert. 1987. «Ecological and Physiological Differentiation Among Low-Arctic Clones of *Daphnia Pulex*». *Ecology*, vol. 68, no 1, p. 188-198.
- Weider, L. J., et A. Hobaek. 2003. «Glacial refugia, haplotype distributions, and clonal richness of the *Daphnia pulex* complex in arctic Canada». *Molecular Ecology*, vol. 12, no 2, p. 463-473.
- Weider, L. J., A. Hobaek, J. K. Colbourne, T. J. Crease, F. Dufresne et P. D. N. Hebert. 1999a. «Holarctic phylogeography of an asexual species complex I. Mitochondrial DNA variation in arctic *Daphnia*». *Evolution*, vol. 53, no 3, p. 777-792.
- Weider, L. J., A. Hobaek, P. D. N. Hebert et T. J. Crease. 1999b. «Holarctic phylogeography of an asexual species complex - II. Allozymic variation and clonal structure in Arctic *Daphnia*». *Molecular Ecology*, vol. 8, no 1, p. 1-13.
- Weiss, R. L., J. R. Kukora et J. Adams. 1975. «The relationship between enzyme activity, cell geometry, and fitness in *Saccharomyces cerevisiae*». *Proceedings of the National Academy of Science U.S.A.*, vol. 72, no 2, p. 794-798.

- Wendel, J. F. 2000. «Genome evolution in polyploids». *Plant Molecular Biology*, vol. 42, no 1, p. 225-249.
- Wessler, S. R. 1996. «Plant retrotransposons. Turned on by stress». *Current Biology*, vol. 6, no 959-961.
- Wicker, T., F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel et A. H. Schulman. 2007. «A unified classification system for eukaryotic transposable elements». *Nat Rev Genet*, vol. 8, no 12, p. 973-982.
- Wolfe, K. H. 2001. «Yesterday's polyploids and the mystery of diploidization». *Nat Rev Genet*, vol. 2, no 5, p. 333-341.
- Wright, S., et D. Finnegan. 2001. «Genome evolution: Sex and the transposable element». *Current Biology*, vol. 11, no 8, p. R296-R299.
- Wright, S., et D. Schoen. 1999. «Transposon dynamics and the breeding system». *Genetica*, vol. 107, no 1, p. 139-148.
- Wright, S. I., N. Agrawal et T. E. Bureau. 2003. «Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana*». *Genome Research*, vol. 13, no 8, p. 1897-1903.
- Wright, S. I., Q. H. Le, D. J. Schoen et T. E. Bureau. 2001. «Population Dynamics of an Ac-like Transposable Element in Self- and Cross-Pollinating *Arabidopsis*». *Genetics*, vol. 158, no 3, p. 1279-1288.

- Wu, C. I. 1991. «Inferences of Species Phylogeny in Relation to Segregation of Ancient Polymorphisms». *Genetics*, vol. 127, no 2, p. 429-435.
- Yoder, J. A., C. P. Walsh et T. H. Bestor. 1997. «Cytosine methylation and the ecology of intragenomic parasites». *Trends in Genetics*, vol. 13, no 8, p. 335-340.
- Yoshiyama, M., Z. Tu, Y. Kainoh, H. Honda, T. Shono et K. Kimura. 2001. «Possible Horizontal Transfer of a Transposable Element from Host to Parasitoid». *Molecular Biology and Evolution*, vol. 18, no 10, p. 1952-1958.
- Zeyl, C., G. Bell et D. M. Green. 1996. «Sex and the spread of retrotransposon Ty3 in experimental populations of *Saccharomyces cerevisiae*». *Genetics*, vol. 143, no 4, p. 1567-1577.
- Zhang, X., M. T. Eickbush et T. H. Eickbush. 2008. «Role of Recombination in the Long-Term Retention of Transposable Elements in rRNA Gene Loci». *Genetics*, vol. 180, no 3, p. 1617-1626.
- Zhao, X. P., Y. Si, R. E. Hanson, C. F. Crane, H. J. Price, D. M. Stelly, J. F. Wendel et A. H. Paterson. 1998. «Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton». *Genome Research*, vol. 8, no 5, p. 479-492.

