

Université de Montréal

# La génétique au service de la conservation de la tortue des bois (*Glyptemys insculpta*)

Par

Cindy Bouchard

Département de sciences biologique  
Faculté des arts et des sciences

Thèse présentée  
en vue de l'obtention du grade de *Philosophiae doctor* (Ph. D.)  
en sciences biologiques

Septembre 2020

© Cindy Bouchard, 2020

Université de Montréal  
Département de sciences biologiques, Faculté des arts et des sciences

*Cette thèse intitulée*

**La génétique au service de la conservation de la tortue des bois  
(*Glyptemys insculpta*)**

*Présentée par*

**Cindy Bouchard**

*A été évaluée par un jury composé de*

**Pierre Legendre**

Président-rapporteur

**François-Joseph Lapointe**

Directeur de recherche

**Nathalie Tessier**

Codirectrice

**Simon Joly**

Membre du jury

**Christina Davy**

Examineur externe

# Résumé

La biologie de la conservation est un domaine de recherche en pleine expansion en raison de la perte accélérée de la biodiversité à l'échelle mondiale. Pour mieux comprendre les processus et les menaces au maintien des populations de petite taille et les effets des facteurs anthropiques sur la biodiversité, la génétique est fréquemment utilisée en conservation. Des analyses génétiques peuvent, par exemple, nous informer sur les tendances à long terme, la diversité des populations et les stratégies de reproduction d'une espèce. La tortue des bois (*Glyptemys insculpta*) est une espèce endémique à l'Amérique du Nord qui est *en danger* d'extinction selon l'Union internationale pour la conservation de la nature. Dans le cadre de ma thèse, j'avais comme objectif de caractériser la diversité génétique de cette espèce menacée au Canada. À cet effet, j'ai analysé la génétique des populations de tortues des bois à plusieurs échelles spatiales et temporelles, afin de mieux cerner les processus ayant un impact sur la diversité des populations.

Dans un premier temps, les relations de parentalité ont été reconstruites au sein d'une population de tortues des bois pour estimer la fréquence de paternité multiple et de paternité répétée. Les résultats de mes travaux suggèrent que l'emménagement de sperme chez la femelle et la reproduction multiple avec les mêmes partenaires pour plus d'une saison de reproduction pourraient expliquer ces phénomènes. Ces stratégies de reproduction pourraient dans ce cas être induites par la faible densité de la population à l'étude, ou encore par la fidélité au site d'hibernation où la majorité des événements de copulation ont lieu.

Par la suite, je me suis intéressée à la diversité génétique des populations de tortues de bois. J'ai voulu comprendre les effets de la configuration spatiale des éléments du paysage et les événements de dispersion géographique sur la diversité des populations. À l'aide d'une approche de génétique du paysage, mes analyses montrent que la division des populations par bassins versants explique une large fraction de la diversité génétique interpopulations. Ces résultats confirment également que les bassins versants représentent des unités de gestion propices à la protection des populations de tortues des bois.

Enfin, des analyses de réseaux ont été utilisées pour mieux cerner la dynamique de flux génique entre les populations de la rive nord et de la rive sud du fleuve Saint-Laurent. Plus spécifiquement, la rive nord se caractérise par un réseau robuste de populations isolées, alors que les populations de la rive sud présentent plutôt une structure de métapopulation. En utilisant les réseaux construits à partir de données génétiques, des scénarios hypothétiques furent comparés pour explorer la sélection de populations à l'aide du logiciel BRIDES. Les résultats de ces analyses ont permis de cibler l'importance de certaines populations de tortues des bois pour la connectivité du réseau. L'importance de ces populations n'aurait pu être prédite par les résultats de la diversité et de la différenciation génétique, les indices de centralité et les analyses d'élimination de nœuds.

Grâce à la génétique, cette thèse apporte de nouvelles connaissances sur la tortue des bois, les stratégies de reproduction des différents sexes, le flux génique, la connectivité et l'influence du réseau hydrographique sur la diversité des populations. Ces résultats nous permettent d'avoir une meilleure compréhension des processus affectant la diversité génétique de cette espèce afin de mieux la protéger. Toutes les analyses réalisées pour cette thèse sont directement applicables à l'ensemble des autres espèces longévives avec des générations chevauchantes.

**Mots-clés** : analyse de réseaux, conservation, logiciel BRIDES, bassin versant, connectivité, flux génique, espèce longévive, génétique du paysage, *Glyptemys insculpta*, théorie des graphes, groupement bayésien, parentalité, paternité, priorisation des populations, tortue d'eau douce, tortue des bois.

## Abstract

Conservation biology is a rapidly expanding field of research due to the accelerating loss of global biodiversity. To better understand the processes and threats to the persistence of small populations and the effects of anthropogenic factors on biodiversity, genetic approaches are frequently used in conservation. Genetic analyzes can, for example, inform us about long-term trends, population diversity and reproductive strategies of a species. The wood turtle (*Glyptemys insculpta*) is a species endemic to North America that is *endangered* according to the International Union for the Conservation of Nature. As part of my thesis, my objective was to characterize the genetic diversity of this threatened species in Canada. In order to better understand the impact of reproductive strategy and landscape structure on population diversity, I analyzed the genetics of wood turtle populations at several spatial and temporal scales.

First, parentage relationships were reconstructed in a population of wood turtles to estimate the frequency of multiple and repeated paternity. The results of my work suggest that sperm storage in females and multiple reproduction with the same partners for more than one breeding season could explain these phenomena. These reproduction strategies could in this case be induced by the low density of the study population, or by fidelity to the overwintering site where the majority of copulation events take place.

Subsequently, I assessed the genetic diversity of wood turtle populations. I wanted to understand the effects of the landscape configuration and geographic dispersion events on the diversity of populations. Using a landscape genetics approach, my analyzes show that the division of populations by watershed explains a large fraction of the genetic diversity between populations. These results also confirm that watersheds represent management units conducive to the protection of wood turtle populations.

Finally, network analysis was used to better understand the gene flow dynamics among populations located on the north and south shores of the St. Lawrence River. More

specifically, the north shore is characterized by a robust network of isolated populations, whereas the populations on the south shore present more of a metapopulation structure. Using population graphs, hypothetical scenarios were compared to explore the node selection process using the BRIDES algorithm. The results of these analyzes made it possible to point out specific populations of wood turtles, considering their importance for network connectivity. This could have not been predicted by using genetic diversity and distinctiveness estimates, node-based metrics, and node removal analysis for these populations.

Thanks to genetics, this thesis brings new knowledge on the wood turtle, the reproductive strategies of both sexes, the gene flow, the connectivity and the influence of the hydrographic network on population diversity. These results allow us to have a better understanding of the processes affecting the genetic diversity of this species in order to better protect it. All analyses performed for this thesis are directly applicable to other long-lived species with overlapping generations.

**Key words** : conservation, BRIDES software, watersheds, connectivity, gene flow, long-lived species, landscape genetics, *Glyptemys insculpta*, clustering analysis, parentage analysis, paternity, population prioritization, network analysis, population graph, freshwater turtles, wood turtles.

# Table des matières

Résumé.....	3
Abstract .....	5
Table des matières .....	7
Liste des Tableaux .....	10
Liste des Figures .....	12
Liste des sigles et abréviations.....	14
Remerciements.....	17
Avant-Propos .....	20
<b>Chapitre 1. Introduction.....</b>	<b>21</b>
<b>1.1 La biologie de la conservation.....</b>	<b>21</b>
La priorisation des populations.....	23
<b>1.2 La variation génétique.....</b>	<b>24</b>
Importance de la diversité génétique .....	24
Origine et maintien de la diversité génétique.....	25
Érosion de la diversité génétique.....	27
<b>1.3 La variation génétique intrapopulation .....</b>	<b>28</b>
<b>1.4 La variation génétique interpopulations .....</b>	<b>30</b>
La connectivité des populations .....	31
<b>1.5 La génétique du paysage.....</b>	<b>32</b>
La notion d'échelle.....	33
Les espèces longévives .....	34
Le choix des marqueurs moléculaires .....	34
<b>1.6 Biologie de l'espèce à l'étude .....</b>	<b>36</b>
La conservation des tortues.....	36
La tortue des bois.....	38
Statut de protection .....	41
<b>1.7 Organisation et objectifs de la thèse .....</b>	<b>42</b>
<b>Chapitre 2. Les liens de parentalité chez la tortue des bois .....</b>	<b>46</b>
<b>2.1 Résumé .....</b>	<b>47</b>
<b>2.2 Abstract .....</b>	<b>48</b>
<b>2.3 Introduction .....</b>	<b>49</b>

<b>2.4 Materials and methods</b> .....	<b>52</b>
Sampling .....	52
DNA extraction and amplification.....	52
Population Genetics .....	53
Assignment of candidate mothers and fathers .....	54
Multiple paternity analysis.....	54
Genotype reconstruction of the fathers .....	56
Statistical analysis .....	57
<b>2.5 Results</b> .....	<b>57</b>
<b>2.6 Discussion</b> .....	<b>64</b>
Multiple paternity .....	64
Repeat paternity.....	65
Male reproductive success .....	67
<b>2.7 Acknowledgments</b> .....	<b>70</b>
<b>Chapitre 3. La génétique du paysage de la tortue des bois</b> .....	<b>71</b>
<b>3.1 Résumé</b> .....	<b>72</b>
<b>3.2 Abstract</b> .....	<b>73</b>
<b>3.3 Introduction</b> .....	<b>74</b>
<b>3.4 Materials and methods</b> .....	<b>76</b>
Sampling design.....	76
DNA extraction and amplification.....	77
Microsatellite analysis .....	77
Genetic structure.....	78
Isolation-by-watershed .....	79
Within-watershed genetic diversity and differentiation.....	81
<b>3.5 Results</b> .....	<b>81</b>
<b>3.6 Discussion</b> .....	<b>87</b>
Genetic structure.....	87
Impact of the landscape on genetic variation.....	89
Genetic diversity of watersheds .....	91
Conservation and management.....	92
<b>3.7 Acknowledgments</b> .....	<b>94</b>
<b>Chapitre 4. Les réseaux dynamiques pour étudier la connectivité des populations chez la tortue des bois</b> .....	<b>95</b>
<b>4.1 Résumé</b> .....	<b>96</b>
<b>4.3 Introduction</b> .....	<b>98</b>



<b>4.4 Materials and methods.....</b>	<b>102</b>
Wood turtle population characterization.....	102
Wood turtle population graph.....	102
Node-based metrics of connectivity and node removal .....	103
BRIDES analysis.....	104
BRIDES selection algorithm.....	106
Weighted models and conversation scenarios .....	106
<b>4.5 Results .....</b>	<b>107</b>
Population graph.....	107
Network metrics and node removal.....	108
BRIDES selection procedure.....	111
<b>4.6 Discussion .....</b>	<b>114</b>
<b>4.7 Acknowledgments.....</b>	<b>119</b>
<b>4.8 Supporting Material .....</b>	<b>120</b>
<b>Chapitre 5. Discussion générale .....</b>	<b>122</b>
<b>5.1 La polyandrie et le succès reproducteur des mâles .....</b>	<b>122</b>
<b>5.2 L'isolement par bassins versants et l'expansion des populations .....</b>	<b>124</b>
<b>5.3 Priorisation des populations .....</b>	<b>125</b>
<b>5.4 Application de la génétique pour la conservation de la tortue des bois.....</b>	<b>126</b>
Le soutien au recrutement .....	126
La délimitation des unités désignables et des populations locales.....	127
Les analyses forensiques.....	128
La conservation des espèces longévives .....	128
Recherches futures sur la génétique de la tortue des bois .....	129
<b>5.8 Conclusion .....</b>	<b>130</b>
<b>Bibliographie.....</b>	<b>132</b>
<b>Annexe I. Analyses forensiques .....</b>	<b>155</b>
<b>AI.1 Procédure de décontamination du matériel .....</b>	<b>156</b>
<b>AI.2 Extraction et amplification de l'ADN .....</b>	<b>157</b>
<b>AI.3 Populations analysées .....</b>	<b>157</b>
Analyses statistiques.....	158
<b>AI.4 Résultats et conclusion .....</b>	<b>159</b>
<b>AI.4 Bibliographie .....</b>	<b>162</b>

# Liste des Tableaux

**Tableau 1.1** Description des études portant sur la parentalité et les populations de la tortue des bois à l'aide de marqueurs génétiques. Le sujet abordé dans l'étude de même que le(s) lieu(x) géographique(s), le nombre de sites (Nb sites), le nombre total d'individus échantillonnés (Nb ind), le type et le nombre de marqueurs utilisés y sont présentés..... 41

**Table 2.1** Characteristics of each of the seven loci, including the numbers of alleles, expected ( $H_E$ ) and observed ( $H_O$ ) heterozygosity, Hardy-Weinberg equilibrium (HWE) p-value, probability of detecting multiple paternity (dIndex) at each locus within their clutches (i.e. probability of exclusion), probability of identity (PI), probability of identity for siblings (PISibs), error rate (Error), and the germinal mutation rate (Mutation). \*Significant departure from HWE following a Bonferroni correction..... 59

**Table 2.2** Probabilities and  $\Delta$ -LOD score for females assigned to the seven clutches without a known mother during the 2006 and 2007 nesting seasons, estimated by COLONY and CERVUS, respectively. Critical values for the relaxed and strict confidence levels are 1.77 and 4.92, respectively.  $B\Delta$ -LOD score for female M180 is 6.46 and M181 is 4.44..... 60

**Table 2.3** Repeat paternity results for eight females with clutches in both nesting seasons, and the father Id assigned to their successfully emerged offspring of each clutch. .... 62

**Table 2.4** Percentages of multiple paternity obtained from the minimum allele count, GERUD and COLONY methods, for each season and for both sampling seasons combined. Multiple paternity was confirmed when the consensus of all three methods detected a minimum of two sires in the successfully emerged offspring of each clutch sampled..... 63

**Table 3.1** Categorization of all sites based on post-glacial Expansion Model constructs (Expansion) and the Watershed Model (Watershed). Sample site (ID) and the number of genotyped individuals (N) in each site is included, as well as the total number of alleles (A), number of private alleles (Pa), mean number of alleles per locus (k), allelic richness (Ar), private allelic richness (PAr), inbreeding coefficient ( $F_{IS}$ ), and observed ( $H_O$ ) and expected heterozygosity ( $H_E$ ) for each watershed group..... 80

**Table 3.2** Summary statistics for nine microsatellite loci amplified in wood turtle samples, including the number of individuals successfully amplified at each locus (N). The number of alleles (k), observed ( $H_O$ ) and expected heterozygosity ( $H_E$ ), probability of identity (PI) and probability of identity for siblings (PISibs) are also estimated for each locus..... 82

**Table 3.3** Distance-based redundancy analysis (dbRDA) and partial redundancy analysis (pRDA) using genetic distance Jost ( $D_{EST}$ ) as the dependent variable, and three models as independent variables or conditions: the River Distance Model (RD) produced by spatial eigenfunctions of the truncated river distance matrix, the Expansion Model (EM) by ancestral

post-glacial expansion hypothesis and the Watershed Model (WS) by a watershed isolation model. The constrained and conditional proportions of the variation, the F statistics and adjusted or semi-partial adjusted R-squares ( $R_a^2$ ) are shown for each model. Stars indicate significance levels of the F-statistics. .... 86

**Table 4.1** Sampled population parameters and corresponding node-based metrics, node removal statistics and BRIDES results. Site identification (Id), number of genotyped individuals (N), expected heterozygosity ( $H_E$ ), estimated population size (PSize), element occurrence rank (EO), betweenness centrality (BC), degree centrality (DC), eigenvector centrality (EC), strength, average inverse edge weight (AIEW), clustering coefficient (CC), average path length (APL) estimated after node removal, and selection rate of each node using the stepwise selection procedure (Selection). Values in bold indicate the top three populations within each of the metrics, node removal and the selection process. .... 110

**Table 4.2** Summary of the BRIDES node selection procedure providing the minimum number of nodes required to reach a maximal score, and the number of multiple solutions with the same score for different scenarios and weighted model schemes. A: Node selection based on heterozygosity ( $H_E$ ) values. B: Node selection based on estimated population size. C: Node selection based on element occurrence (EO) values. For each scenario, nodes included in the subgraph represented populations with the largest (Proactive) or smallest (Reactive) parameter values. .... 113

**Table 4.S1** Results of the BRIDES node selection procedure for six scenarios (A, B, C) and five weighted models (1 to 5). For each node, the presence of a population in the subgraph (x) and its selection (1) or not (0) in at least one of the optimal solution are indicated. Scenarios for which no nodes were selected (-) are also shown. .... 121

**Tableau AI.1** Code des populations, nombre d'individus utilisés et nombre d'allèles total obtenu pour les différentes populations du Québec utilisées lors des analyses. .... 158

**Tableau AI.2** Valeurs de probabilités obtenues lors de l'assignation des individus du Québec à leur population d'origine. .... 160

**Tableau AI.3** Valeurs des probabilités pour les individus de Hong Kong lors de l'assignation à chacune des populations du Québec. .... 161

**Tableau AI. A** Profil génétique des individus saisis au poste de douane de Hong Kong ..... 161

# Liste des Figures

<b>Figure 1.1</b>	Représentation de la dépression de consanguinité, l'hétérosis et la dépression d'hybride. Adapté de Quaglietti et al. 2016. ....	27
<b>Figure 1.2</b>	Aire de répartition de la tortue des bois au Canada (COSEPAC 2018).....	39
<b>Figure 1.3</b>	Cadre conceptuel de la thèse présentant les thématiques abordées dans les différents chapitres.....	43
<b>Figure 2.1</b>	Details of the clutches of wood turtles sampled in the Shawinigan population over two successive years (2006 and 2007). ....	58
<b>Figure 2.2</b>	Relationships between all successfully hatched offspring sampled and adults (red circles: mothers; blue squares: fathers, white pentagons: clutches sampled in 2006; gray pentagons: clutches sampled in 2007). Loops indicate individuals that produce offspring with the same partners.....	61
<b>Figure 2.3</b>	Histogram representing the numbers of mates, sired clutches, and emerged offspring for each reconstructed genotypes of putative fathers estimated on successfully hatched offspring of clutches sampled in 2006 and 2007.....	61
<b>Figure 2.4</b>	Boxplots representing the mean number of emerged offspring per father in clutches having multiple fathers versus single fathers (Student t-test, $p = 0.001$ ).....	63
<b>Figure 3.1</b>	Locations of 24 wood turtle sample sites, with watersheds outlined in white. ....	82
<b>Figure 3.2</b>	Estimates of admixture proportions inferred by the most probable $K_{max}$ values obtained from TESS and $K$ values obtained from STRUCTURE using a hierarchical approach for 331 wood turtles sampled in Québec, Ontario and New Brunswick watersheds. Gatineau (GA), Madawaska (MD), L'Assomption (LA), Loup (LO), St. Maurice (SM), Châteauguay (CH), Yamaska (YA), Missisquoi (MI), St. François (SF), Bécancour (BE), Du Chêne (DC), Matapédia (MA). The five different levels represent successive runs of the STRUCTURE algorithm. ....	84
<b>Figure 3.3</b>	Venn diagram presenting the variation partitioning results of the genetic distance (Jost $D_{EST}$ ) between sites, with respect to spatial eigenfunctions provided by the River Distance Model, the ancestral post-glacial Expansion Model and the Watershed Model. The values presented were computed from semi-partial adjusted R-squares of a pRDA analysis. The Venn diagram was produced using the varpart function (R, vegan library). ....	86
<b>Figure 4.1</b>	A) Example of the six types of paths computed in BRIDES by comparing the shortest path between pairs of nodes in subgraph $X_n$ (open circles) and the shortest path between the same pairs or nodes in the complete graph $X_{all}$ including all candidate nodes (blue circles). In $X_{all}$ , the shortest path between any pair of nodes from $X_n$ must include at least one of these	

candidate nodes. B) Each of the three nodes available for the selection process was iteratively added to  $X_n$  to create the subgraphs  $X_{n+1}$  and corresponding scores were computed with respect to the complete graph  $X_{all}$ . The selection process selected node 11 as the best candidate (score = 1) to maximize network connectivity while minimizing the number of additional nodes. The weighted model used for the selection process was (1, 0, 0, 0, 1, 1)...... 105

**Figure 4.2** Population graph of 19 wood turtle populations based on conditional genetic distance (cGD). Redundant edges were pruned when they did not contribute to the overall genetic covariance structure. Circular node sizes are scaled according to the average inverse edge weight (AIEW), where greater AIEW values are reflected through larger circular diameters. Populations located on the northern and southern shores of the St. Lawrence River are identified with gray and white nodes, respectively..... 108

**Figure 4.3** Different population subgraphs based on distinct node selection parameters using a proactive or reactive conservation approach. A: Node selection based on heterozygosity ( $H_E$ ) values. B: Node selection based on estimated population size. C: Node selection based on element occurrence (EO) values. For each scenario, nodes included in the subgraph represented populations with the largest (Proactive) or smallest (Reactive) parameter values. Populations located on the northern and southern shores of the St. Lawrence River are identified with gray and white nodes, respectively..... 111

**Figure 4.4** BRIDES node selection process for scenario Ap and weighting scheme 5 (3, 0, 0, 1, 2, 3). All selection steps are shown, even when multiple solutions had identical scores. In such cases, distinct solutions (i.e. MA1 or MA2) are identified by dashed edges connecting the selected nodes. Maximum score (1.00) was reached with the addition of four candidate nodes. At each step of the procedure, the selected nodes and their corresponding links are presented in blue..... 113

**Figure 4.S1** Results of the BRIDES node selection procedure for two scenarios and three weighted models. In each case, networks with optimal scores while minimizing the number of candidate nodes are presented. Nodes from the subgraph  $X_n$  are depicted by empty circles, whereas candidate nodes and corresponding edges in  $X_{n+1}$  are depicted by blue circles..... 120

**Figure AI.1** Échantillons de sang tels que reçus à l'Université de Montréal pour chacune des tortues des bois..... 156

**Figure AI.2** La figure fut retirée puisqu'il s'agit de données sensibles..... 157

**Figure AI.3** Arbre phylogénétique construit à partir de distances de  $D_{CE}$ . Les valeurs de bootstrap sont inscrites le long des branches. .... 161

# Liste des sigles et abréviations

%	Pourcentage
°C	Degré Celsius
μmol	micromole
ADN	Acide désoxyribonucléique
Ar	allelic richness
AVP	analyse de viabilité des populations
bp	Base pair/Paire de bases
CITES	Convention sur le commerce international de la faune et de la flore sauvages menacées d'extinction
cm	Centimètre
COSEPAC	Comité sur la situation des espèces en péril au Canada
COSEWIC	Committee on the status of endangered wildlife in Canada
CRSNG	Conseil de recherche en sciences naturelles et en génie du Canada
CSBQ	Centre de la science de la biodiversité du Québec
DAI	Direction des affaires internationales de l'Université de Montréal
dbMEM	Distance-based Moran's eigenvector maps
dbRDA	Distance-based redundancy analysis
D <sub>EST</sub>	Allelic differentiation measures
DIC	Deviance information criterion
dIndex	Detection index
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
DU	Designatable unit
ESU	Evolutionarily significant unit
etc.	et cetera
ex.	Exemple
F <sub>IS</sub>	Inbreeding coefficient
F <sub>ST</sub>	Index de fixation/Fixation index
FRQNT	Fonds de recherche du Québec – Nature et Technologies
Ha	Hectare
H <sub>E</sub>	Hétérozygotie attendue /Expected Heterozygosity
H <sub>O</sub>	Hétérozygotie observée/Observed heterozygosity
HWE	Hardy-Weinberg equilibrium
i.e.	<i>id est</i> (c'est-à-dire)
Id	Identity
IUCN	International union for conservation of nature
k	Mean number of alleles per locus
km	Kilomètre

L	Litre
LD	Linkage disequilibrium
LEP	Loi sur les espèces en péril
m	Mètre/metre
MAC	Minimum allele count
MCMC	Markov Chain Monte-Carlo
MFFP	Ministère des Forêts, de la Faune et des Parcs
mg	Milligram
MgCl <sub>2</sub>	Magnesium chloride
mm	Millimètre/millimetre
mmol	Milimole
mtDNA	Mitochondrial DNA
MU	Management unit
N	Number of genotyped individuals
nkat	Nanokatal
ng	Nanogram
NSERC	Natural Sciences and Engineering Research council of Canada
P	Probabilité/probability
Pa	Private alleles
Par	Private allelic richness
PCR	Polymerase chain reaction
PCoA	Principal coordinate analysis
PI	Probability of identity
PIsibs	Probability of identity for siblings
pRDA	Partial redundancy analysis
PVA	Population viability analysis
R <sup>2</sup>	Coefficient of determination
Ra <sup>2</sup>	Adjusted coefficient of determination
RDA	Redundancy analysis
SARA	Species at Risk Act
SD	standard deviation
UD	Unité désignable
UICN	Union internationale pour la conservation de la nature
US	United states
UV	Ultraviolet
v.	Version
vs	Versus

*À Guillaume, Charlie, Henri, Léo, Renaud  
et Nintendo pour son message étonnamment à propos ...*

***Everything not saved will be lost***  
*- Nintendo quit screen message*



# Remerciements

Un énorme merci à mon directeur de thèse François-Joseph Lapointe et à ma codirectrice Nathalie Tessier. Les tortues m'ont menée jusqu'à vous en mai 2011, mais c'est en grande partie pour vous que je suis restée. Merci d'avoir fait de moi la chercheuse que je suis aujourd'hui. François-Joseph, merci pour TOUT. Ton écoute, ton ouverture d'esprit, ton respect, ta confiance, pour tous les vœux de *bonne journée de la femme* et de *bonne fête des Mères*. Merci pour ces beaucoup trop nombreux échanges de courriels à des heures pas d'allure (merci Cathia)! C'était un honneur d'évoluer à tes côtés durant les 9 dernières années. Merci pour cette liberté, d'avoir été là et de ne pas avoir douté de moi. Merci surtout pour ta patience et ta compréhension quand tu t'es rendu compte que je ne l'enverrais jamais à la garderie mon p'tit! Super Nathalie, merci pour ta bonne humeur, ta rigueur et ton efficacité. Les très longues heures passées au laboratoire furent beaucoup plus faciles quand tu étais là. Merci pour toutes ces connaissances que tu m'as enseignées, tant au laboratoire que sur le terrain. Merci de m'avoir transmis ta passion pour la génétique et la conservation. Un énorme merci pour les sorties de terrain et les bonbons. Surtout, merci de m'avoir offert de travailler sur tes projets de tortues.

Merci à tous les membres du LEMEE avec qui j'ai eu la chance de travailler au fil des années : Anita, ma *super ninja oily sister*, merci pour toutes les folies dans le labo et les corrections de mon anglais boiteux. Geneviève, ma belle amie, tes remises en question ont certainement fait de moi une meilleure personne. Merci de m'avoir emmenée au Shakti et convaincue de faire des entraînements avec Guillaume ... tsé! Virginie, incroyable partenaire de stage et super gardienne de Charlie, merci pour tout. Belle Marianne, merci d'apporter de la diversité à nos discussions de laboratoire et pour tous tes conseils de maman. Maïté, merci pour les discussions sur l'accouchement et la naissance. Je te souhaite une superbe aventure avec la maternité. Étienne, merci pour ta participation enthousiaste à mon dernier article de thèse. Merci également à Margaux, Carine, Maria-Camila, Astrid, Jessica et Chantale.

Un grand merci à Pierre Legendre, Simon Joly et Christina Davy d'avoir accepté de participer à mon jury de thèse, à Patrick James et Mohamed Hijri pour votre implication dans mon comité-conseil et mon examen de synthèse et à tous les professeurs et le personnel du Département de sciences biologiques.

Merci aux cinq hommes de ma vie. Mon amoureux, merci d'être toi et d'avoir soutenu notre famille sur tous les fronts durant toutes ses années. Sans toi, je serais loin d'être la femme, la mère et la belle-mère que je suis aujourd'hui. Ce parcours aurait été beaucoup plus tortueux sans ton amour et ton soutien. Merci à mes enfants, Charlinours pour ta bonne humeur, Henrikiki pour tes chansons sans fins, Léo d'amour pour ta curiosité et Renaud coco pour tes discussions sans fin. Sans vous, j'aurais probablement terminé mon doctorat beaucoup plus vite, mais avec un grand déficit de bonheur et d'amour. Désolée pour la grognitude suivant mes nuits trop courtes. Je vous aime.

Merci à ma famille, à ma belle-famille et à mes amies. Cette thèse, elle est aussi pour vous. Merci à mes sœurs, Daisy et Kim, pour votre support sans faille malgré la distance. Vous êtes des femmes incroyables qui me permettent d'être une meilleure personne chaque jour. Pis, on se le cachera pas, nos blagues sont les meilleures. Merci aussi à mes beaux-frères Alex<sup>2</sup>, mes parents, mes beaux-parents, ma belle-sœur, Marie-Aimée (oui tu rentres dans la section famille toi avec), ma filleule et mes six neveux! Merci de votre compréhension à toutes les fois où j'étais sur l'ordinateur au lieu d'être avec vous, de vous intéresser à mes recherches, de lire mes articles, d'aider avec les enfants et de mettre plein de joie dans ma vie. Merci à Vanessa pour ton écoute et tes conseils, t'avais raison, il fallait juste que ça se fasse ! Merci à Marilou pour tous ces matins de captures d'oiseaux et pour ton enthousiasme contagieux : WôûHou ! Merci à la gang de l'observatoire d'oiseaux de McGill et du Shakti rock Gym, sans vous tous, mon doctorat aurait certainement été plus difficile.

Merci à tous les biologistes et techniciens amoureux des tortues qui ont répondu à mes nombreuses questions et m'ont transmis de nombreux détails importants pour mes projets. Merci à tous d'avoir partagé vos connaissances et expériences avec moi : Pierre-André Bernier, Walter Bertacchi, Geneviève Bourget, Yohann Dubois, Sylvain Giguère, Mélanie Laflèche, Philippe Lamarre, Denis Masse, Stéphane Pelletier et Lucie Veilleux. Un énorme merci évidemment à tou.te.s ceux et celles qui ont participé à la récolte d'échantillons de tissus de tortue. Sans votre travail parfois très intense, ma thèse n'aurait jamais pu exister. Je suis extrêmement reconnaissante de tout le travail que vous avez accompli. Un grand merci aux biologistes, techniciens et stagiaires du gouvernement du Québec, de Parcs Canada, de la Corporation Gestion de la Forêt de l'Aigle, de la Corporation de l'Aménagement de la rivière L'Assomption, de l'Écomuseum, et de l'Université de Montréal qui ont participé aux inventaires.

Merci aux nombreux étudiants à qui j'ai eu la chance d'enseigner depuis 2014, vous m'avez beaucoup appris.

Je suis très reconnaissante du financement qui m'a été accordé par le CRSNG et le FRQNT durant mes études doctorales. Merci de m'avoir permis de me concentrer sur ma recherche et pour le prolongement de mon congé de maternité. Merci au CSBQ et à la FAI d'avoir financé en partie mes voyages de formation et mes conférences à l'étranger.

Finalement, merci à plus de 600 tortues des bois qui furent manipulées, avec l'accord du comité d'éthique, dans le cadre de cette thèse.



## Avant-Propos

Malgré les diverses lois assurant sa protection, la collecte illégale et la revente sur le marché noir demeurent des problèmes importants pour la conservation de la tortue des bois. Afin d'éviter que les braconniers utilisent l'information de cette thèse pour leur bénéfice personnel, les localisations géographiques précises des tortues échantillonnées ne sont pas indiquées et les sites d'échantillonnage sont vaguement identifiés sur les cartes géographiques.

(Crédit photo : Fondation Séthy)

# Chapitre 1. Introduction

## 1.1 La biologie de la conservation

La biologie de la conservation est une science multidisciplinaire, développée en réponse à la crise mondiale de la perte de biodiversité et qui intègre l'ensemble des facteurs écologiques, économiques et sociaux qui y sont associés (Soulé 1991). Cette science documente la distribution de la biodiversité à tous ces niveaux, génétique, spécifique et écosystémique, et cherche à comprendre l'impact de l'Homme sur ceux-ci. La biologie de la conservation vise à développer une approche pratique qui prend tout son sens dans un ensemble de valeurs éthiques afin de prévenir l'extinction des espèces, maintenir la diversité génétique au sein de celles-ci, protéger et restaurer les communautés et les fonctions écosystémiques qui leurs sont associées (Soulé 1985, Caughley 1994). Les pressions néfastes auxquelles font face les espèces sont principalement causées par l'expansion rapide des populations humaines et leurs activités. Il s'agit principalement de la destruction, la fragmentation et la dégradation de l'habitat par le développement commercial et résidentiel, l'agriculture, l'exploitation des ressources naturelles, les changements climatiques et l'augmentation de la pollution. Le commerce illégal, les maladies et la présence d'espèces envahissantes exotiques peuvent aussi avoir des conséquences dévastatrices sur la survie des espèces et des populations (Cardillo et al. 2004, McKee et al. 2004, Rosen & Smith 2010, IUCN 2015).

Bien que l'espèce soit généralement considérée comme l'élément central en conservation, les niveaux infraspécifiques, notamment les populations, vont offrir des informations importantes pour leur protection. Les populations sont très souvent porteuses d'adaptations locales importantes permettant à l'espèce d'occuper un habitat spécifique ou de répondre aux changements environnementaux, d'où l'importance de protéger leur diversité (Soulé 1987, Moritz 1994, Fitzpatrick & Keller, 2015). De plus, les populations sont très sensibles à l'extirpation (extinction locale) et leurs tendances démographiques peuvent nous

renseigner sur le déclin d'une espèce (Hughes et al. 1997). Les menaces auxquelles les populations font face peuvent être complètement opposées considérant le contexte économique, social et culturel de la société humaine près de laquelle elles évoluent, requérant par le fait même des actions de conservation spécifique (Mace & Lande 1991, Wood & Gross 2008, Razgour et al. 2017).

C'est au début des années 80 que l'application de la génétique en conservation devint populaire, notamment par la publication de trois livres qui établirent les fondements du domaine (Soulé & Wilcox 1980, Frankel & Soulé 1981, Schonewald-Cox et al. 1983). Les principes de la génétique des populations, soit l'étude des changements dans les fréquences alléliques et génotypiques d'une population, y sont utilisés afin de mieux comprendre les enjeux auxquels font face les populations de petite taille. La structure génétique intrapopulation et interpopulation peut être affectée par de nombreux facteurs, tels que la dérive génique, la taille efficace des populations, le flux génique, la sélection naturelle et le taux de mutation (Manel et al. 2003, Segelbacher et al. 2010). L'avancée des technologies en biologie moléculaire et bio-informatique a certainement favorisé l'utilisation grandissante de la génétique à des fins de conservation. Certaines études de cas ont grandement contribué à établir l'importance de ce domaine d'étude pour la survie des espèces menacées. Un exemple classique est sans doute l'étude de la diversité génétique appauvrie des guépards (*Acinonyx jubatus*) qui représente de 1 à 10% de la diversité mesurée chez les autres espèces de félins (O'Brien et al. 1983). Cette quasi-uniformisation de la variation génétique entraîne de nombreux problèmes tels que des malformations des spermatozoïdes, un taux de mortalité élevé des juvéniles et un système immunitaire affaibli (O'Brien et al. 1985). Dans un environnement en perturbation, cette faible diversité diminue grandement les probabilités d'adaptation et de survie d'une espèce.

Les techniques moléculaires sont maintenant fréquemment utilisées pour diverses applications en conservation. L'objectif commun de l'ensemble de ces applications est la préservation de la diversité qui permettra le maintien d'un processus fondamental: l'évolution (Woodruff 2001). Les principaux aspects abordés dans les études en génétique de la conservation se regroupent en cinq grandes catégories basées sur une étude de Wayne & Morin

(2004). (1) La description et l'identification des individus, des relations de parentalité, de la structure génétique des populations et des relations taxonomiques. (2) La détection et la prédiction des effets de la perte d'habitats, de la fragmentation et de l'isolation géographique. (3) La détection et la prédiction des effets de l'hybridation et de l'introgession. (4) La compréhension des relations entre les adaptations phénotypiques et les caractères génétiques d'un individu ou d'une population dans un environnement donné. (5) La gestion et la réintroduction d'individus élevés en captivité et la restauration des communautés biologiques. Ces connaissances sont jugées essentielles au maintien à long terme des différentes populations dans leur habitat naturel et par le fait même, à la pérennité des espèces au statut précaire.

### **La priorisation des populations**

Puisque les populations évoluent dans des contextes différents et que les ressources financières sont limitées, des choix difficiles s'imposent fréquemment. Quelles populations doit-on protéger en priorité ? Quelles actions de protection seront accomplies et quand ces actions auront-elles lieu ? Ces questions ont notamment poussé de nombreux auteurs à développer des cadres conceptuels permettant de répondre objectivement à de tels enjeux décisionnels (Given & Norton 1993).

Les méthodes les plus utilisées sont certainement celles reposant sur des règles précises (*rule-based method*). Le statut d'une espèce est dans ces cas déterminé à l'aide d'une série de critères associés au risque d'extinction, tels que la taille des populations, le taux de déclin, la fragmentation de l'habitat et la distribution des populations dans l'aire de répartition. Le statut de conservation permet d'identifier les espèces et les populations les plus à risque de disparaître et d'en connaître les causes. L'Union internationale pour la conservation de la nature (UICN) a établi un système de classification des espèces permettant de déterminer la précarité d'une espèce menacée à l'échelle mondiale (Master 1991, IUCN 2015). Ces critères ont servi de référence à l'établissement de nombreuses unités de conservation fréquemment utilisées par les gestionnaires, telles que les unités désignables (UD, COSEWIC 2015), les unités de gestion (MU, Moritz et al. 1994, Palsbøll et al. 2007) et les unités évolutives significatives (ESU, Moritz et al. 1994) dont certaines permettent une protection légale.

Le processus de priorisation des populations, bien qu'il soit fréquemment basé sur les tendances démographiques révélées par l'analyse de viabilité des populations (AVP), gagne à utiliser la génétique qui lui offre une tout autre perspective. Les mesures de diversité génétique intra et interpopulations sont déjà utilisées pour définir et caractériser les unités de conservation (Petit et al. 1998) et de récentes études les appliquent pour créer des cadres conceptuels plus approfondis. Des réseaux phylogénétiques ont, par exemple, permis d'évaluer le niveau de différenciation des populations pour être en mesure de les prioriser en considérant leur contribution relative à la diversité génétique (Volkman et al. 2014). Dans le but d'offrir un outil facilement utilisable par les gestionnaires, Ottewell et al. (2016) ont développé un cadre conceptuel complet où la génétique sert à guider les actions de conservation. Le niveau de diversité génétique et de consanguinité de chaque population de même que le niveau de différenciation entre celles-ci sont utilisés pour établir, par exemple, les unités de gestion et les stratégies de translocation d'individus pour restaurer la diversité.

## **1.2 La variation génétique**

### **Importance de la diversité génétique**

De nombreux facteurs peuvent causer la disparition d'une espèce de son milieu naturel. On les sépare habituellement en deux groupes; les facteurs déterminants (ex. la perte d'habitat, l'exploitation, la collecte illégale et la pollution) et les facteurs aléatoires (ex. la démographie, l'environnement et la génétique). Afin de persister, les espèces doivent être en mesure de s'adapter à ces modifications plus ou moins prévisibles de leur environnement (Frankham & Kingslover 2004, Frankham 2005), Frankel (1970) fut le premier à proposer une relation entre la perte de diversité génétique et l'augmentation des risques d'extinction. Plus récemment, lors d'une méta-analyse comparant 170 espèces en péril à 170 espèces parentes abondantes, Spielmann et al. (2004) ont observé que plus de 77 % des espèces avec un statut précaire possédaient une diversité génétique réduite. La perte de ce potentiel d'adaptation peut être irréversible considérant la vitesse du processus de création de la diversité et la rapidité des



changements anthropiques qui ont pour effet de fragmenter l'habitat en isolant les populations les unes des autres.

La diversité génétique, tant au niveau individuel que populationnel, peut survenir grâce à divers processus tels que les mutations germinales, la recombinaison, la reproduction sexuée et le flux génique. À l'opposé, l'érosion de cette diversité peut survenir lors de croisements entre individus apparentés (dépression de consanguinité) ou très éloignés (dépression d'hybride), ou encore par l'effet de la dérive génique et de la sélection naturelle. Les conséquences néfastes de ces processus causant l'érosion seront amplifiées par la réduction de la taille d'une population (goulot d'étranglement) pouvant entraîner celle-ci dans un vortex d'extinction (Gilpin & Soulé 1986, McGinnity et al. 2003).

### **Origine et maintien de la diversité génétique**

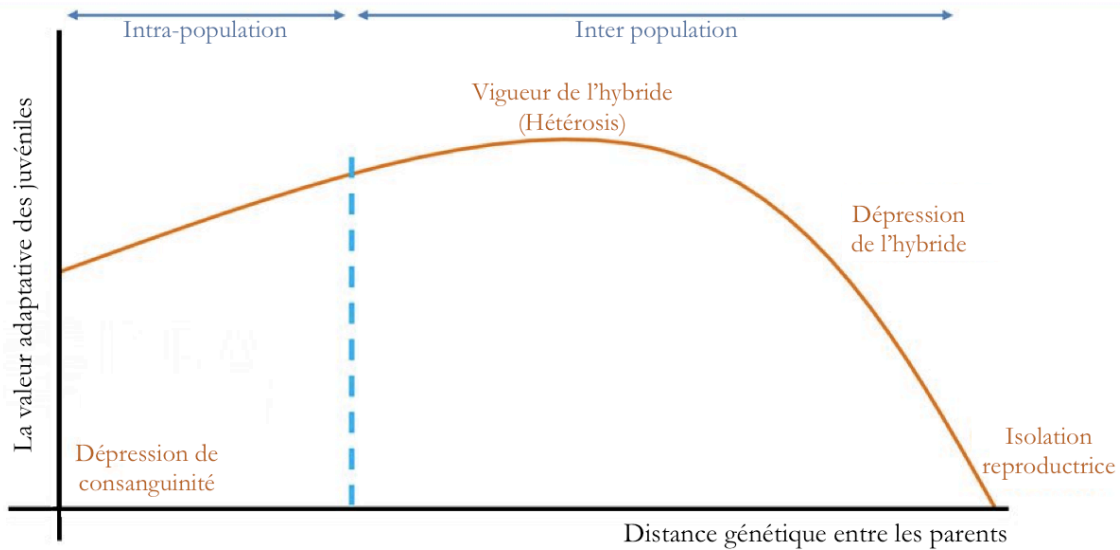
Une mutation germinale est une modification de la séquence de l'ADN dans les cellules germinales qui entraîne la création de nouveau matériel génétique chez les descendants. Les mutations sont causées par plusieurs phénomènes, dont des erreurs lors de la réplication, des éléments génétiques mobiles et plusieurs facteurs de stress comme les éléments radicaux libres et les rayons ultraviolets. Les plus importantes sont celles affectant la valeur adaptative (fitness) des individus, plus particulièrement les mutations létales ou délétères. Les mutations neutres, quant à elles, nous informent principalement sur les différences génétiques accumulées au cours du temps entre les individus, les populations ou les espèces. Ces informations permettent, par exemple, de déterminer des événements passés d'expansion ou de réduction démographique des populations (Storz & Beaumont 2002). Ces mutations n'ont pas d'impact sur la survie des individus (Charlesworth et al. 1993).

Le taux de mutation est particulièrement important puisqu'il nous informe sur la vitesse à laquelle ces mutations s'accumulent au sein d'une espèce pour engendrer la diversité génétique (Baer et al. 2007). Chez les eucaryotes, le taux de mutation moyen est relativement faible et on l'estime à une mutation par locus par 100 000 gamètes par génération. Lorsque la

diversité génétique est diminuée, le temps requis pour sa régénération est très long, généralement de l'ordre du millier de générations, puisque les taux de mutation sont très faibles.

La recombinaison et la reproduction sexuée augmentent le potentiel de la variation génétique créée par les mutations et participent à la création d'un individu unique. Le mélange de la séquence d'ADN entre les copies d'un chromosome lors de la méiose augmente la variation. Par la suite, l'individu héritant d'un chromosome maternel et d'un chromosome paternel pourra bénéficier d'une combinaison unique d'allèles et d'un génotype qui lui est propre.

L'ajout de nouveau matériel génétique dans une population par le flux de gènes survient de manière plus rapide que la création de la diversité par les mutations. Il est causé par le déplacement de gènes d'une population à une autre par l'entremise d'un individu migrant. Pour qu'il y ait flux génique, l'individu doit se déplacer dans une population avoisinante, se reproduire et y laisser des descendants viables et fertiles. Le flux de gènes permet aux populations de petite taille de maintenir un niveau de diversité génétique favorable au maintien de la population dans le temps (Palsbøll et al. 2007). Il peut prévenir la perte de diversité causée par la dérive et les effets néfastes de la dépression de consanguinité. Un gain de performance peut être observé en raison du brassage des allèles de populations distinctes, mais similaires; ce phénomène se nomme l'hétérosis (Figure 1.1). Pour ce faire, la connectivité doit être maintenue entre les habitats occupés par les individus d'une même population et entre les populations formant une métapopulation pour permettre la dispersion des juvéniles et la migration des individus.



**Figure 1.1** Représentation de la dépression de consanguinité, l'hétérosis et la dépression d'hybride. Adapté de Quaglietti et al. 2016.

## Érosion de la diversité génétique

La dérive génique est un processus qui cause des changements dans la fréquence des gènes en raison de phénomènes aléatoires, tels que l'appariement des gamètes. D'une génération à l'autre, ces modifications s'accumuleront de manière indépendante dans les populations isolées. À long terme, il est possible d'observer une uniformisation des allèles causée par la perte ou la fixation de certains d'entre eux. Ce phénomène est grandement accéléré dans les populations de petite taille (Ezard & Travis 2006, Pittman et al. 2011). Il en résulte une diminution rapide (en quelques générations seulement) de la diversité génétique. Des phénomènes démographiques associés à la réduction du nombre d'individus colonisateurs peuvent causer une telle diminution de la diversité. Il s'agit, notamment, du goulot d'étranglement défini comme la perte soudaine d'individus dans le temps (ex. à la suite d'une épidémie ou dans le cas de captures illégales) ou de l'effet fondateur par un nombre restreint d'individus (DiBattista 2008, Vargas-Ramírez et al. 2012). À l'inverse de la dérive, la sélection naturelle est un processus évolutif non aléatoire. En favorisant les probabilités de survie et de reproduction des individus porteurs, un allèle sera de plus en plus représenté dans la population au fil des

générations. À long terme, cette sélection participera à l'érosion de la diversité en favorisant la fixation des allèles porteurs d'adaptations avantageuses et en éliminant les allèles porteurs d'effets délétères.

La consanguinité (ou endogamie) est la reproduction entre individus apparentés (*inbreeding*). La réduction de la taille des populations entraîne l'augmentation des croisements consanguins au sein d'une population. La conséquence principale de l'endogamie, nommée dépression de consanguinité (Figure 1.1), est d'altérer les probabilités de survie des individus d'une population en augmentant la fréquence de phénotypes récessifs délétères (Lande 1998). Les rejetons issus de tels croisements auront notamment une probabilité de survie réduite, une fécondité plus faible et une fréquence de malformations plus élevée que lors de croisements entre individus non apparentés. De nombreux cas de dépression de consanguinité sont répertoriés chez des populations captives ou des populations isolées en milieu naturel (Ralls & Ballou 1983, Frankham 2000). À l'opposé, la dépression d'hybride résultera d'une valeur sélective plus faible chez les descendants issus d'un croisement entre deux individus provenant de populations très différentes (*outbreeding*) (Figure 1.1). Dans le cadre d'une méta-analyse, Frankham (2015) observa que dans 92,9 % des cas à l'étude, des bénéfices furent observés chez les populations consanguines lorsque des croisements avec des individus de populations différentes avaient eu lieu.

### **1.3 La variation génétique intrapopulation**

Par le passé, lors de l'utilisation de la génétique en conservation, l'accent était mis sur une vision globale de la population, délaissant par le fait même des problématiques importantes pour la gestion, telles que les relations de parentalité entre les individus, la paternité et la consanguinité (Wayne & Morin 2004). Or, la stratégie de reproduction d'une espèce a une grande influence sur la taille efficace des populations, notamment en affectant la variance du succès reproducteur et le sexe-ratio (Balloux & Lehmann 2003). Elle a de même un effet certain sur le maintien de la diversité génétique dans le temps. Les stratégies de reproduction et les analyses

de parentalité sont maintenant largement étudiées et les résultats de ces recherches sont d'une grande importance pour la conservation des populations et le maintien de leur diversité génétique (Holman & Kokko 2013).

Ces stratégies sont généralement divisées en trois groupes : la monogamie, la polygamie et la promiscuité sexuelle. La monogamie est présente lorsque deux individus forment un couple exclusif pour toute la durée de leur vie. La polygamie peut être divisée en trois catégories: la polygynie (un mâle se reproduit avec plusieurs femelles), la polyandrie (une femelle se reproduit avec plusieurs mâles) et la polygynandrie (deux ou plusieurs femelles se reproduisent en exclusivité avec deux ou plusieurs mâles) (Thornhill & Alcock 1983). Cette dernière est légèrement différente de la promiscuité où tous les mâles auront la possibilité de se reproduire avec toutes les femelles d'une population (Karl 2008). Plusieurs études démontrent que la polyandrie est commune chez les espèces animales, quoique très variable entre les populations et les espèces (Joseph & Shaw 2011, Taylor et al. 2014).

Pour les femelles, les bénéfices directs de la polyandrie ou de la promiscuité incluent, entre autres, l'augmentation de la contribution parentale et l'accumulation d'une quantité suffisante de sperme pour la fécondation des ovules. Certains coûts peuvent être observés, tels que la propagation d'infections transmises sexuellement et les impacts négatifs associés aux agressions par les mâles et/ou les tentatives de copulation répétées (Holman & Kokko 2013). Dans le cas d'espèces avec des soins parentaux limités, les impacts de la promiscuité sont principalement indirects. Par exemple, il peut s'agir de l'augmentation de la qualité, de la complémentarité et de la variation génétique des jeunes de la nichée. Plusieurs mécanismes sont évoqués comme hypothèses pour augmenter le succès de la nichée de la femelle (Jennions & Petrie 2000). La reproduction multiple permet le *genetic trade-up*, phénomène observé lorsque la femelle a la possibilité de se reproduire de nouveau avec un mâle de plus grande qualité que celui rencontré précédemment. La promiscuité permettrait une plus grande diversité génétique des spermatozoïdes en plus de la compétition et de la sélection spermatique postcopulatoire. La reproduction avec plusieurs mâles est avantageuse dans la mesure où, grâce à la sélection postcopulatoire, la femelle a l'occasion d'évaluer la valeur adaptative du mâle par rapport à la

fécondité (Bowen & Karl 2007). Une dernière hypothèse proposée est la stratégie de minimisation des risques (*genetic bet-hedging*). Cette hypothèse considère qu'une nichée plus diversifiée aura une plus grande probabilité de survie dans un environnement imprévisible.

Chez les mâles, le coût de la production de spermatozoïdes est négligeable. Par contre, ceux associés à la défense du territoire et à l'accès aux femelles seront importants. Considérant la variation dans le succès reproducteur des femelles et dans le succès de nidification, il serait avantageux pour le mâle de maximiser le nombre d'évènements de copulation (Stewart & Dutton 2011). La polygamie et la promiscuité peuvent mener à la paternité multiple où plusieurs mâles seront les pères de jeunes d'une même fratrie. L'impact de la paternité multiple sur la taille efficace de la population ne fait pas consensus dans la littérature. Certains auteurs sont d'avis qu'il y aurait augmentation de la variation du succès reproducteur des mâles par rapport à une stratégie de reproduction monogame (Karl 2008, Theissinger et al. 2009), alors que d'autres sont d'avis contraire (Sugg & Chesser 1994, Pearse & Anderson 2009, Lotterhos 2011). L'effet de la paternité multiple dépend en fait de plusieurs facteurs, tels que la fréquence de la reproduction, la moyenne et la variation dans le nombre de jeunes produits ainsi que la distribution de la paternité dans la nichée (Lotterhos 2011). L'impact de la paternité multiple sera plus important dans les populations de petite taille où elle pourrait aller jusqu'à diminuer les probabilités de dépression de consanguinité et prévenir une perte de diversité génétique (Holman & Kokko 2013).

## **1.4 La variation génétique interpopulations**

Au fil du temps, des différences dans les fréquences alléliques s'accumuleront entre les populations isolées, soit de manière aléatoire, par la dérive génétique, ou par la sélection naturelle. La différenciation entre deux populations pourra s'inscrire dans un spectre allant de populations complètement panmictiques lors d'un libre échange de gènes entre celles-ci jusqu'à l'atteinte de deux espèces distinctes en présence de barrières à la reproduction établies de manière irréversible (Coates et al. 2018). Entre ces deux extrêmes se trouvent plusieurs niveaux

de différenciation où certaines populations pourront dépendre des autres pour leur survie, soit pour le recrutement de juvéniles ou pour l'apport de nouveau matériel génétique. C'est précisément le cas des métapopulations. Une métapopulation est un réseau de populations qui occupent des parcelles d'habitats favorables, isolées par une matrice d'habitats non favorables (Levins 1970). Elles ne sont cependant pas suffisamment interconnectées pour constituer un ensemble panmictique, mais l'ensemble des interactions qu'elles partagent est nécessaire à leur maintien à long terme. Le concept de métapopulation est principalement spatial et démographique. La survie d'une population qui se trouve dans une parcelle est donc influencée par les échanges d'individus avec les parcelles voisines. Ces échanges affecteront les probabilités d'extinction de l'espèce sur cette partie de territoire.

## **La connectivité des populations**

L'estimation de la connectivité directe entre populations est complexe et nécessite des connaissances sur la dispersion des individus, soit par suivi télémétrique ou par capture-marquage-recapture à très grande échelle. Ces données étant rarement disponibles, la connectivité est fréquemment estimée de manière indirecte à l'aide du flux de gènes entre les populations (Lowe & Allendorf 2010). D'un point de vue génétique, une métapopulation est caractérisée par un ensemble de populations spatialement isolées qui possèdent, ou possédaient jusqu'à récemment, un certain flux génique entre elles (Hanski & Gilpin 1991). La distribution de la variation génétique entre les populations est le résultat des effets opposés de la dérive génique (divergence des populations) et du flux génique (uniformisation des populations). La différenciation des populations fut longtemps estimée à l'aide d'un indice de fixation ( $F_{ST}$ ) et des mesures de distances génétiques qui comparaient des paires de populations. Une vision d'ensemble du portrait de la connectivité et de l'isolation des populations à l'étude était alors difficile à obtenir.

La théorie des graphes (ou science des réseaux) est un des outils qui permet désormais d'obtenir une vision globale des relations entre populations (Kool et al. 2013, Garroway et al. 2008). Les réseaux permettent de délaisser le cadre plus strict de la génétique des populations basée sur une comparaison des populations naturelles avec une population idéale hypothétique.

Un graphe est un modèle mathématique représentant un ensemble de nœuds connectés par des liens qui peuvent parfois être pondérés et directionnels. De nombreux indices et propriétés ont été définis et développés dans diverses disciplines pour caractériser les composantes, les nœuds et les liens d'un réseau (Newman 2003). Ces mesures sont de plus utilisées en écologie et en conservation des écosystèmes terrestres et aquatiques, où les réseaux trophiques, la dynamique des métapopulations et le développement de réserves écologiques peuvent être aisément modélisés et visualisés (Treml et al. 2008, Marrotte et al. 2017, Matos et al. 2019).

## **1.5 La génétique du paysage**

La capacité des individus à se déplacer au sein d'un paysage complexe et hétérogène est essentielle au maintien régional des espèces. La génétique du paysage cherche à comprendre la manière dont la structure et la diversité génétique des populations sont affectées par la composition et la configuration du paysage et des variables environnementales (Manel et al. 2003, Segelbacher et al. 2010). Plus spécifiquement, elle s'intéresse à l'interaction entre les composantes biotiques et abiotiques du paysage et les processus de microévolution tels que le flux génique, la dérive génique et la sélection naturelle. Les deux étapes clés en génétique du paysage consistent d'abord en la détection de discontinuités génétiques et ensuite, à la corrélation de ces discontinuités avec les composantes du paysage. Ces connaissances permettent d'établir des mesures de conservation concrètes afin de maintenir la connectivité nécessaire au maintien du potentiel évolutif des espèces. La construction de corridors de migration est un bon exemple d'application qui permet la migration et la dispersion des individus entre deux parcelles d'habitat favorable dans un milieu fragmenté par l'homme (Rudnick et al. 2012). Au cours du siècle dernier, l'expansion rapide des populations humaines et l'augmentation des infrastructures construites par l'Homme ont considérablement modifié l'habitat au sein duquel les espèces évoluent. La perte, la dégradation et la fragmentation de l'habitat engendrées par les activités anthropiques constituent une menace importante pour les espèces qui occupent les régions les plus densément peuplées (Gibbon et al. 2000, Pittman et



al. 2011). Des conséquences négatives pourront être observées sur la dispersion, le taux de mortalité et la diversité génétique, ainsi que sur la distribution et la densité des populations (Templeton et al. 1990, Luque et al. 2012).

## **La notion d'échelle**

Une part importante des analyses au niveau du paysage consiste à considérer des échelles spatiales et temporelles appropriées pour toutes les étapes de l'étude, plus précisément lors de l'élaboration du plan d'échantillonnage, de la sélection des types de marqueurs moléculaires et du choix des analyses statistiques appropriées (Anderson et al. 2010). Une erreur d'adéquation entre l'échelle considérée pendant l'étude et l'échelle à laquelle un processus se manifeste peut entraîner des conclusions inexactes, notamment quant à l'impact de la fragmentation de l'habitat (Prunier et al. 2014, Epps & Keyghobadi 2015). L'échelle spatiale à laquelle les études en génétique du paysage ont lieu est déterminée par les propriétés écologiques et biologiques propres à l'espèce, telles que les capacités de dispersion et de migration des individus (Holderegger & Wagner 2008).

En plus du choix de l'échelle spatiale appropriée, il est important de considérer également l'échelle temporelle à laquelle un processus se manifeste. Plus particulièrement, il s'agit de connaître le délai entre le processus et les conséquences observables sur la structure génétique (Anderson et al. 2010, Segelbacher et al. 2010). Par exemple, selon l'espèce, de 10 à 200 générations peuvent être nécessaires afin que les effets génétiques d'une modification du paysage soient détectables (Landguth et al. 2010). L'historique des populations peut donc avoir un effet très important sur la structure génétique actuelle. Des événements passés comme l'extinction et la recolonisation peuvent masquer des effets du paysage contemporain (ex. la fragmentation par des infrastructures humaines) (Wade & McCauley 1988). Pour cette raison, de plus en plus de chercheurs tentent de dissocier les effets historiques du paysage passé des effets contemporains du paysage actuel sur la structure génétique des populations (voir: Zellmer & Knowles 2009, Anderson et al. 2010, Freedman et al. 2010).

## **Les espèces longévives**

Les espèces longévives, caractérisées par une longue durée de vie et une maturité sexuelle tardive, représentent un défi particulier lors de l'étude des phénomènes récents sur la structure génétique des populations. Pour ces espèces, les délais associés à la réponse génétique peuvent être très longs puisque dans leur cas, quelques générations à peine peuvent facilement se traduire en plusieurs centaines d'années (Kuo & Janzen 2004). Les effets des paysages historiques seront donc principalement détectés lors des analyses génétiques. Les impacts démographiques de la fragmentation et de la perte d'habitat apparaîtront bien avant les signaux génétiques qui devraient être la perte de diversité et la différenciation des populations (Litzgus 2006, Pittman et al. 2011). Malgré tout, l'étude d'une espèce sédentaire, le sphénodon (*Sphenodon punctatus*) (Moore et al. 2008), montre qu'à petite échelle spatiale, la structure génétique peut tout de même refléter les changements récents du paysage.

Malgré la difficulté que représente l'étude des phénomènes récents sur la génétique des espèces longévives, leurs caractéristiques particulières peuvent offrir un avantage pour la biologie de la conservation. Par exemple, à la suite d'un goulot d'étranglement, il est possible que l'intégrité génétique soit préservée durant plusieurs années, contrairement à ce qu'on observe chez des espèces à longévité plus courte où les effets de la dérive génétique seraient plus importants. À l'aide de mesures appropriées, la taille de la population pourra ainsi être rétablie en affectant très peu la diversité génétique.

## **Le choix des marqueurs moléculaires**

Les marqueurs utilisés lors des analyses génétiques auront un impact certain sur l'échelle spatiale et temporelle où le flux génique peut être détecté. Ce choix doit être fait en fonction de leurs taux de mutation et de leur variabilité. Les marqueurs les plus communément utilisés en génétique du paysage sont les microsatellites (Anderson et al. 2010, Storfer et al. 2010), définis comme de courtes séquences composées d'un motif de deux à six nucléotides répétés en tandem. Le nombre de répétitions est hautement variable selon l'individu et le nombre d'allèles par locus peut être très élevé (jusqu'à 50). Le génome eucaryote possède un grand

nombre de locus microsatellites (environ 1.4 % du génome) (Chapuis & Estoup 2007). Ces marqueurs moléculaires grandement variables ont un taux de mutation très élevé (1 mutation/1000 nucléotides) principalement dû au glissement de la polymérase lors de la réplication (Stallings et al. 1991, Chapuis & Estoup 2007). La grande majorité des locus se trouve dans les régions non codantes du génome. Il est donc généralement admis que les microsatellites sont des marqueurs sélectivement neutres qui ne sont pas affectés par le processus de sélection naturelle.

L'utilisation de plusieurs locus microsatellites est nécessaire afin d'augmenter la précision des analyses (Haig 1998), principalement, lorsque le nombre d'individus à l'étude est restreint (<500) ou lorsque l'étude porte sur la configuration spatiale des populations (Anderson et al. 2010, Landguth et al. 2012). Le polymorphisme de ces marqueurs les rend particulièrement performants pour révéler la structure génétique des populations à petite échelle et leur caractère codominant est un avantage lors d'analyse de parentalité (Slatkin 1995, Sunnucks 2000, Guichoux et al. 2011, Defaveri et al. 2013).

Malgré leur vaste utilisation en génétique, plusieurs limites des marqueurs microsatellites doivent être considérées lors de l'analyse des données, telles que les erreurs dues aux grands nombres de manipulations de laboratoire et la présence d'homoplasie et d'allèles nuls. L'homoplasie de taille (*size homoplasy*) est fréquemment observée chez les marqueurs microsatellites (Estoup et al. 2002) : il s'agit de deux copies d'un locus identiques dans leur état, leur taille, mais qui ne sont pas identiques par descendance (*identical-by-descent*). Les modèles de mutation utilisés dans les logiciels d'analyses considèrent les probabilités d'homoplasie entre deux allèles identiques, ce qui permet de diminuer grandement les conséquences sur les résultats (Hoshino et al. 2012). Pour un locus microsatellite donné, un allèle nul est défini comme un allèle qui n'amplifie pas à un niveau détectable lors de la réaction de polymérase en chaîne. Un individu hétérozygote portant un allèle nul sera donc identifié et qualifié incorrectement d'homozygote pour l'allèle visible. L'impact des allèles nuls sera variable en fonction de leur fréquence dans un locus et de la capacité des chercheurs à les détecter et à les corriger lors des analyses. Les allèles nuls peuvent affecter les estimations de la diversité

génétique et mener à de fausses exclusions lors d'analyses de parentalité (Dakin & Avise 2004). Plusieurs causes peuvent expliquer la présence d'allèles nuls; des problèmes d'appariement d'amorces dus à une divergence au niveau des séquences et l'utilisation de l'ADN dégradé en sont des exemples. De nombreux logiciels sont disponibles pour détecter les allèles nuls et corriger les fréquences alléliques des marqueurs microsatellites problématiques (Van Oosterhout et al. 2004, 2006), alors que certaines approches considèrent directement cette source d'erreur lors des analyses génétiques (Jones & Wang 2010).

## **1.6 Biologie de l'espèce à l'étude**

### **La conservation des tortues**

Selon l'Union internationale pour la conservation de la nature, près des deux tiers des espèces de tortues seraient en péril. Selon certains auteurs, les tortues représentent le groupe de vertébrés le plus menacé (Gibbons et al. 2000, Rhodin et al. 2018). Comme plusieurs espèces longévives, elles se caractérisent par un taux de survie élevé des adultes, l'atteinte tardive de la maturité sexuelle et un faible taux de recrutement (Congdon & Gibbons 1990). Depuis 200 millions d'années, leur morphologie unique a permis aux tortues de rayonner dans de nombreux écosystèmes dulcicoles, marins et terrestres et d'y jouer des rôles écologiques importants. En présence d'une communauté de tortues diversifiée et composée de populations importantes, la biomasse peut même surpasser celle de grands mammifères herbivores. Les rôles écologiques des tortues sont très diversifiés, allant du transfert d'énergie dans et entre les écosystèmes, la participation au cycle du calcium, et la dispersion de semences végétales. Certaines espèces bio-indicatrices sont même utilisées pour restaurer les écosystèmes dégradés (Lovich et al. 2018).

Au cours du dernier siècle, la croissance exponentielle des populations humaines et la rapidité des perturbations environnementales ont affecté les processus naturels d'adaptation des espèces. Dans le cas des tortues, les principales menaces à la survie des populations sont la

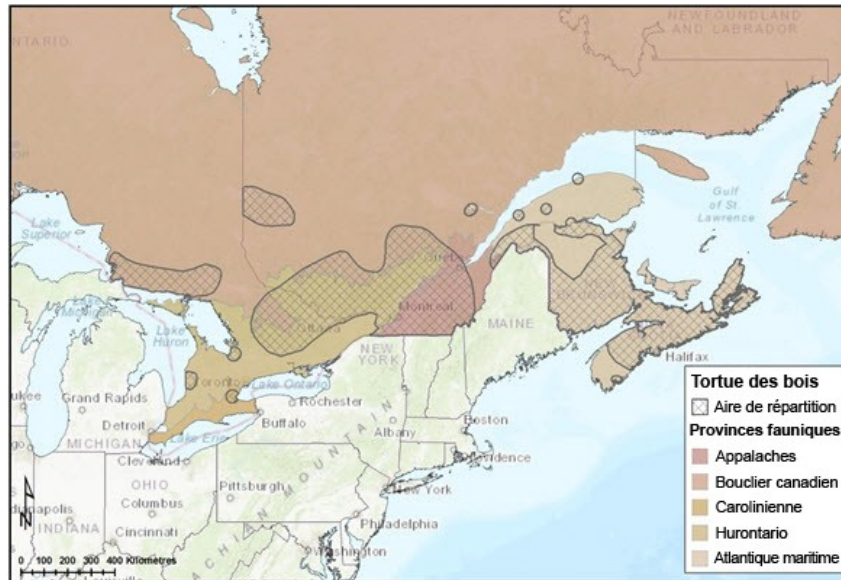
fragmentation et la destruction de l'habitat, l'augmentation du nombre de prédateurs causée par la présence humaine et la collecte illégale. Le déclin des tortues peut parfois être attribué à un effet synergique de plusieurs de ces menaces (Vogt 1981, Traversy et al. 1989, Klemens 2000, Lovich et al. 2018).

Il est souvent difficile pour les biologistes d'estimer les tendances démographiques des populations de tortues en raison de leur grande longévité. Deux causes majeures peuvent expliquer ce phénomène. La première est qu'en l'absence de recrutement, les adultes peuvent persister dans l'habitat pendant plusieurs années, créant ainsi une fausse perception de persistance à long terme (Mortimer 1995). La seconde cause est liée à la densité réduite de certaines populations de tortues qui peuvent perdurer à de faibles densités pendant plusieurs années. Pour les chercheurs qui ont rarement observé des tortues en grand nombre, cette faible densité est devenue la nouvelle norme (Papworth et al. 2009, de Miranda 2017). Au niveau génétique, l'étude des tortues sera d'autant plus complexifiée par le long temps de génération, les générations chevauchantes et la capacité de certains individus à se déplacer sur de très longues distances. Les signaux génétiques, habituellement observés chez les populations de petite taille vivant en milieu fragmenté, sont parfois différents de ce que prédit la théorie (Davy 2013, Davy et al. 2014).

Au Québec, on dénombre huit espèces de tortues d'eau douce et toutes sont en péril. Les populations québécoises se situent à la limite nord de la répartition géographique des tortues en Amérique du Nord, ce qui leur confère un statut particulier pour la conservation des espèces (Lesica & Allendorf 1995). Puisque l'habitat des populations nordiques peut grandement différer de celui des populations situées plus au sud, les populations du Québec peuvent avoir des adaptations spécifiques en plus de diverger génétiquement. Les tortues présentes à nos latitudes se sont adaptées à des températures plus froides, notamment en laissant les jeunes hiberner dans les nids (Storey et al. 1988). Elles sont aussi caractérisées par une taille plus importante à la maturité sexuelle. Dans un contexte de changements climatiques, ces caractères et cette diversité génétique deviennent essentiels afin d'assurer l'adaptation, l'expansion et le maintien de ces espèces à long terme.

## La tortue des bois

Autrefois largement répandues dans le nord-est de l'Amérique du Nord, les populations de tortue des bois (*Glyptemys insculpta*) sont maintenant sporadiquement distribuées (Figure 1.2) et comprennent un nombre restreint d'individus (moins de 50 à 350 tortues; Galois & Bonin 1999, Tessier & Lapointe 2002, Bowen & Gillingham 2004). Comme la majorité des espèces de tortues, leur stratégie est basée sur une survie à long terme des individus matures qui peuvent survivre jusqu'à 50 ans en milieu naturel (Cameron et al. 2002). Le temps de génération est long (estimé entre 25 à 47 ans), la maturité sexuelle est tardive (environ 18 ans) et le recrutement est faible (Van Dijk & Harding 2013). La copulation a lieu principalement à l'automne et au printemps, au moment où les individus se rassemblent près des sites d'hibernation. Cependant, puisque la reproduction peut avoir lieu de mai à septembre et que les femelles ont la capacité d'emmagasiner le sperme, la reproduction entre tous les individus matures d'une population est considérée possible (Walde et al. 2003). Chez les mâles et les femelles, plusieurs événements de copulation sont généralement observés au cours d'une saison de reproduction, soit entre les deux mêmes partenaires ou encore entre partenaires différents (Harding & Bloomer 1979, Walde et al. 2003). Des interactions agressives entre les mâles ont aussi été observées en captivité (Dinkins 1954, Harding & Bloomer 1979) et en milieu naturel (Kaufmann 1992, Walde et al. 2003, Parren 2013). Une hiérarchie de dominance existe au sein des mâles et les vainqueurs des combats occuperaient un rang social supérieur et profiteraient d'un meilleur accès à la copulation avec les femelles (Kaufmann 1992) et d'un succès de reproduction plus élevé (Galbraith 1991). La dominance des mâles serait déterminée par la taille, la maturité et le niveau d'agressivité de l'individu.



**Figure 1.2** Aire de répartition de la tortue des bois au Canada (COSEPAC 2018).

La tortue des bois est la plus terrestre des tortues d'eau douce retrouvées au Québec. Elle occupe des milieux divers tels que les ruisseaux, les rivières et les plaines inondables (Walde 1998). Les distances parcourues au cours d'une année entre le site d'hibernation, le site de ponte et les sites d'alimentation peuvent être supérieures à un kilomètre. Dans la population de Shawinigan, par exemple, les distances maximales parcourues en une seule journée pour trois femelles en déplacement vers le site de nidification ont été estimées à 2.24 km, 2.4 km et 2.94 km (Walde et al. 2007). Les individus font de plus grands déplacements aquatiques ou terrestres le long des cours d'eau, alors que les déplacements perpendiculaires au cours d'eau sont plus courts, en moyenne 0.276 km pour les femelles et 0.107 km pour les mâles (Parren 2013). Les déplacements terrestres rendent la tortue particulièrement vulnérable aux accidents causés par la machinerie agricole, à l'augmentation de la prédation et à la collecte illégale (Litzgus & Brooks 1996, Équipe de rétablissement 2005).

Les approches génétiques sont utilisées fréquemment pour acquérir des connaissances multiples sur la tortue des bois (Tableau 1.1). Par exemple, le sexe d'individus immatures peut être déterminé à l'aide de marqueurs spécifiques qui ciblent les chromosomes (XX/XY)

permettant la détermination sexuelle (Litterman et al. 2017), la présence de tortues près d'un cours d'eau peut être confirmée par un prélèvement d'ADN environnemental (Lacoursière-Roussel et al. 2016, Akre et al. 2019) et la paternité multiple peut être détectée à l'aide de marqueurs génétiques (Galbraith 1991). La première étude sur la génétique des populations québécoises de tortues des bois fut publiée par Tessier et al. (2005) à l'aide de marqueurs microsatellites. Depuis, plusieurs études portant sur la génétique des populations au Canada (Fridgen et al. 2013) ou aux États-Unis (Castellano et al. 2009, Willoughby et al. 2013) furent publiées, certaines ciblant plus spécifiquement l'importance des populations en périphérie de l'aire de répartition de l'espèce (Spradling et al. 2010) et l'effet des routes sur la structure des populations (Robillard et al. 2019) (Tableau 1.1). Ces études ont révélé peu d'impact de la fragmentation plus récente de l'habitat sur la diversité génétique et la différenciation des populations. L'effet des diminutions démographiques observées ne fut pas détecté génétiquement, même lorsque les analyses ciblaient spécifiquement les juvéniles. Une analyse de la phylogéographie fut aussi réalisée à l'aide de marqueurs mitochondriaux pour étudier l'expansion des populations (Amato et al. 2008).



**Tableau 1.1** Description des études portant sur la parentalité et les populations de la tortue des bois à l'aide de marqueurs génétiques. Le sujet abordé dans l'étude de même que le(s) lieu(x) géographique(s), le nombre de sites (Nb sites), le nombre total d'individus échantillonnés (Nb ind), le type et le nombre de marqueurs utilisés y sont présentés.

Référence	Sujet à l'étude	Lieu géographique	Nb sites (Nb ind.)	Type de marqueur (Nb)
Galbraith (1991)	Succès reproducteur des mâles et paternité multiple	Pennsylvanie	1 (80)	Minisatellites (2)
Amato et al. (2008)	Phylogéographie et dispersion postglaciaire	Amérique du Nord	29 (117)	ADNmt (750-pb)
Tessier et al. (2005)	Génétique des populations	Québec	6 (136)	Microsatellites (5)
Castellano et al. (2009)	Génétique des populations	New Jersey et Pennsylvanie	4 (80)	Microsatellites (7)
Spradling et al. (2010)	Génétique d'une population périphérique et isolée	Iowa, Minnesota et Virginie-Occidentale	10 (135)	Microsatellites (9)
Fridgen et al. (2013)	Génétique des populations par structure d'âge	Ontario	4 (78)	Microsatellites (5)
Willoughby et al. (2013)	Génétique des populations et déclin démographique	Michigan	3 (68)	Microsatellites (9)
Robillard et al. (2019)	Effet de l'isolation par une autoroute	New York	4 (66)	Microsatellites (7)

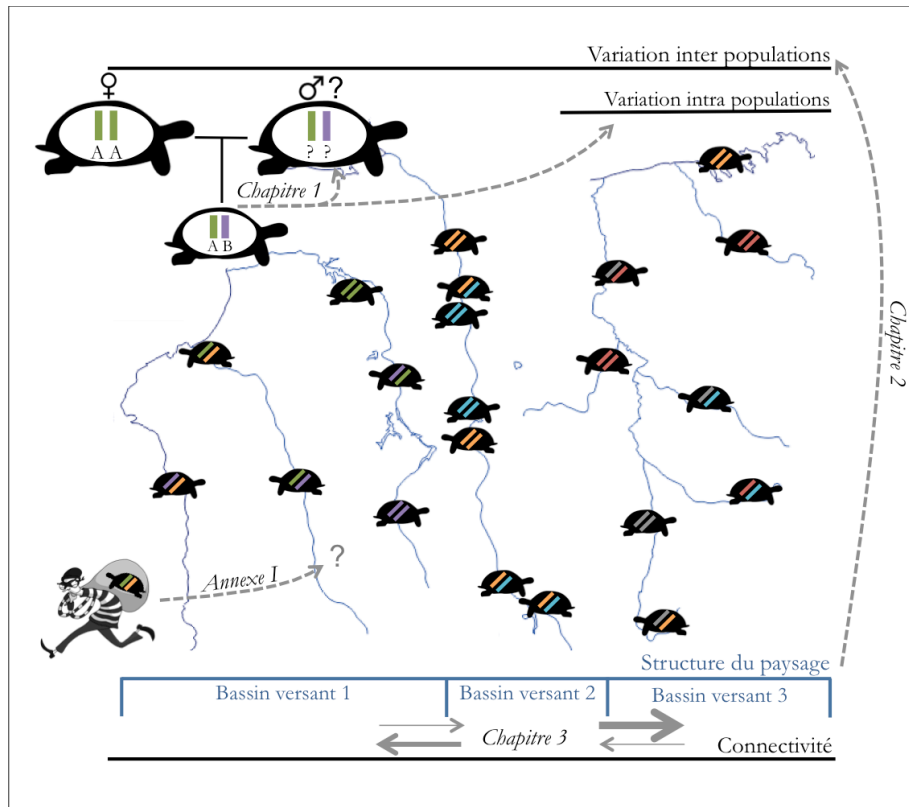
## Statut de protection

Depuis 2011, l'UICN classe la tortue des bois comme espèce *en danger* d'extinction en raison du déclin global de la plupart de ses populations (Van Dijk & Harding 2013). Les menaces l'affectant varient régionalement. Les principales sont la mortalité accidentelle causée par les véhicules routiers et la machinerie agricole, les collectes illégales, la perte et la dégradation de l'habitat pour le développement résidentiel, commercial ou l'exploitation forestière, ainsi que l'augmentation de la prédation (Environnement Canada 2016). Au Canada, la tortue des bois possède le statut d'espèce *menacée* depuis 2007 (COSEPA 2018). Elle est inscrite à l'Annexe I de la Loi sur les espèces en péril (LEP). Au Québec, elle est désignée comme espèce *vulnérable*

(Gazette officielle du Québec 2005). Elle possède différents statuts de protection dans l'ensemble des provinces et états où on la retrouve. Elle est inscrite à l'Annexe II du Traité de la Convention sur le commerce international de la faune et de la flore sauvages menacées d'extinction (CITES).

## **1.7 Organisation et objectifs de la thèse**

Le **Chapitre 1** de cette thèse présente une revue de la littérature portant sur l'utilisation de la génétique dans le domaine de la conservation et plus particulièrement pour la conservation de la tortue des bois. Les avancées dans le domaine de la génétique ont permis d'accumuler des connaissances essentielles pour assurer le maintien à long terme des espèces menacées. Les divers éléments étudiés dans le cadre de ce doctorat sont mis en relation à la Figure 1.3. Je me suis d'abord intéressée à caractériser les liens parentaux au sein d'une population de tortue des bois. Par la suite, des analyses en génétique du paysage m'ont permis de mieux comprendre les effets de la division des bassins versants sur la connectivité des populations et des analyses de réseaux m'ont finalement permis de développer un outil de priorisation des populations en fonction de leur importance dans la connectivité du réseau.



**Figure 1.3** Cadre conceptuel de la thèse présentant les thématiques abordées dans les différents chapitres.

Avec le développement des marqueurs microsatellites, l'échantillonnage peu invasif de tissus et la robustesse des outils statistiques, les analyses de parentalité de la tortue peuvent apporter des informations cruciales pour la conservation de l'espèce. Les objectifs du **Chapitre 2** étaient d'abord (1) d'assigner ou de reconstruire le génotype maternel pour chacune des nichées à l'étude. Les génotypes paternels ont par la suite été reconstruits afin d'estimer (2) la fréquence de la paternité multiple et (3) de la paternité répétée au sein des nichées ainsi que leur impact sur la diversité génétique neutre de la population. Les résultats de ce chapitre ont permis de mieux comprendre les liens de parentalité et le succès reproducteur chez la tortue des bois ainsi que de confirmer la présence de paternité multiple et répétée au sein de cette population. Ces résultats sont des plus importants pour la mise en place de plusieurs actions de conservation concrètes, telle que l'augmentation du nombre d'individus par un soutien au

recrutement. Ils nous informent également sur les relations entre les individus qui permettront de maintenir la diversité génétique de la population.

Dans le **Chapitre 3**, je me suis intéressée à l'impact de la structure du paysage sur la différenciation génétique des groupes d'individus échantillonnés. Bien que de nombreuses études portant sur la caractérisation génétique aient été réalisées, aucune à ce jour n'a étudié directement le lien entre la structure du paysage et la différenciation des populations. Les objectifs de ce volet de ma thèse consistaient d'abord (1) à estimer les effets de la division par bassins versants sur la structure génétique des tortues des bois par des analyses de groupement bayésien et un modèle d'isolation par bassin versant. Par la suite, (2) la diversité génétique et le niveau de consanguinité de chaque groupe ont été évalués, afin (3) de définir des unités de conservation pour une meilleure gestion de l'espèce. Les résultats de ce chapitre ont permis d'observer qu'une partie de la variation génétique interpopulation était expliquée par l'isolation créée par les bassins versants, ce qui permettra de consolider le modèle de gestion par bassin versant préconisée chez cette espèce. Mes résultats ont, en plus, permis d'établir que 10 groupes répondent au critère de distinction génétique, une première étape pour leur désignation en tant qu'unités désignables.

Les analyses de réseaux permettent d'estimer la connectivité d'un système et l'importance relative de chacun des nœuds pour maintenir cette connectivité. Ce classement peut par la suite aider les biologistes qui désirent établir une priorisation des populations à protéger. Les objectifs du **Chapitre 4** étaient d'abord (1) de présenter un nouveau processus de sélection de nœuds, ajouté au logiciel BRIDES (Lord et al. 2016), permettant de sélectionner des nœuds candidats en fonction de leur rôle pour la connectivité d'un réseau. Par la suite, (2) les indices de centralité et les analyses d'élimination des nœuds ont été utilisés pour comparer les performances de BRIDES et estimer le flux de gènes entre les populations de tortue des bois. BRIDES est un nouvel outil qui permet une analyse de la dynamique des réseaux par l'estimation de six types de chemins affectant la connectivité. Les résultats ont permis d'observer la robustesse de la sélection de BRIDES pour un même scénario analysé. Les indices de centralité ont permis de mieux comprendre les dynamiques de flux de gènes, en dépit du

classement des populations très variable en fonction de l'indice choisi. Il s'agit, à ma connaissance de la première étude portant sur l'utilisation des réseaux pour l'analyse de la génétique des populations de tortues.

Cette thèse comporte également une **Annexe I** qui a été réalisée dans le contexte de mon doctorat et qui présente une application concrète de la génétique en biologie de la conservation. En 2013, au tout début de mes études doctorales, deux tortues des bois furent saisies aux douanes de Hong Kong. Dans le but de les assigner à leur population d'origine, des échantillons de sang furent prélevés et envoyés à l'Université de Montréal à des fins de comparaisons avec les échantillons récoltés au Québec. Les tortues saisies ont donc été comparées à neuf populations de tortues des bois représentant 172 individus analysés à l'aide de huit locus microsatellites. J'ai réalisé les analyses génétiques et statistiques et rédigé le rapport technique présentant les résultats des analyses. La littérature grise fait partie intégrante des études en conservation. Plusieurs de ces documents et de ces projets ne seront, malheureusement, jamais publiés et deviennent, par le fait même, très difficiles d'accès. Je croyais donc important de partager ce projet en l'annexant à ma thèse.

## Chapitre 2. Les liens de parentalité chez la tortue des bois

### Paternity Analysis of Wood Turtles (*Glyptemys insculpta*) reveals complex mating patterns

Cindy Bouchard<sup>1</sup>, Nathalie Tessier<sup>2</sup>, François-Joseph Lapointe<sup>1</sup>

1. Département de sciences biologiques, Université de Montréal, Montréal, QC, Canada

2. Ministère des Forêts, de la Faune et des Parcs, Longueuil, QC, Canada

Publié dans : Journal of Heredity, 109(4): 405-415.  
Copyright The American Genetic Association 2017  
DOI: 10.1093/jhered/esx103

*Format modifié pour les besoins de la thèse*

## 2.1 Résumé

Les stratégies de reproduction d'une espèce sont d'une grande importance, car en plus d'influencer le succès reproducteur des mâles et des femelles, elles peuvent affecter l'isolement reproducteur. La tortue des bois (*Glyptemys insculpta*) est une espèce de tortue d'eau douce classée *en danger* par l'Union internationale pour la conservation de la nature. Considérant son statut de conservation et le manque d'information disponible sur les liens de parentalité chez cette espèce, nous avons effectué des analyses à l'aide de marqueurs microsatellites pour étudier le mode de reproduction de la population de tortues des bois de la rivière Shawinigan (Québec). Nous avons échantillonné 38 nids durant deux saisons de reproduction (14 en 2006 et 24 en 2007) pour un total de 248 tortues analysées avec sept locus microsatellites. La reconstruction du génotype des pères a révélé que le succès reproducteur varie grandement entre les mâles. Une corrélation positive fut observée entre le nombre de femelles et de nids où la participation d'un mâle est détectée et le succès reproducteur du père. La fréquence de la paternité multiple est estimée à 37% des nids échantillonnés grâce à un consensus de trois méthodes d'estimation. Une corrélation positive fut observée entre la diversité génétique d'une nichée et le nombre de pères y participant. La paternité répétée, quant à elle, fut observée chez 88% des femelles dont les nids avaient été échantillonnés en 2006 et en 2007. Ce pourcentage suggère soit un usage fréquent de l'emménagement de sperme chez la femelle, soit des événements de copulation avec le même partenaire durant plus d'une saison de reproduction.

## 2.2 Abstract

Mating system characteristics are of great importance as they may influence male and female reproductive success and reproductive isolation. The wood turtle (*Glyptemys insculpta*) is a terrestrial freshwater species listed as endangered by the International Union for Conservation of Nature. Considering its conservation status and the paucity of information currently available on parentage relationships for the species, we performed a microsatellite analysis to study the mating system of wood turtles in the Shawinigan River (Québec). We sampled 38 clutches over two years (14 in 2006 and 24 in 2007), for a total of 248 hatched offspring genotyped with seven microsatellite loci. The reconstructed genotypes of the fathers revealed that reproductive success in the sampled clutches varied greatly between males and are positively correlated with the number of mates and clutches sired. Frequency of multiple paternity was estimated at 37% through a consensus of three different estimation methods. Positive correlation was observed between the genetic diversity of clutches and the number of fathers. Repeat paternity, however, was observed in 88% of the clutches by the same female in successive years, which suggests either a frequent use of sperm storage, or re-mating with the same partner in successive years.

**Key words :** COLONY, conservation, GERUD, microsatellites, polyandry, repeat paternity



## 2.3 Introduction

Based on inequalities in the cost of gamete production, sexual selection theory (Bateman 1948) suggests that male reproductive success is constrained by the number of mates, where females represent a resource to compete for, whereas, female reproductive success is limited by the amount of resource allocated to gamete production (Trivers 1972, Price 1984, Rowe et al. 1994). Since a single copulation event is usually sufficient to fertilize all of the gametes produced (Trivers 1972), females might not be expected to actively search for several partners. However, multiple mating by both males and females are common in natural populations of vertebrates (Larila & Seppä 1998, Avise et al. 2002, Griffith et al. 2002, Pearse et al. 2002, Kitchen et al. 2006). The benefits of multiple mating on female fitness have been extensively studied. Such direct benefits include paternal contribution to egg production and parental care, nuptial gifts, and nesting protection (Reynolds 1996). In the case of a species with limited parental care, like turtles, the impact of polyandry is mostly indirect (Jennions & Petrie 2000, Pearse & Avise 2001). Such indirect benefits may arise from increased genetic quality and the genetic diversity of the offspring produced by several mechanisms that enhance female fitness (Jennions & Petrie 2000). Genetic trade-up, post-copulatory sperm competition (Curtisinger 1991, Parker 1998), cryptic female choice (Zeh & Zeh 1996, Tregenza & Wedell 2000), and genetic bet-hedging (Yasui 2001) are all possible due to multiple mating events.

Polyandry is widespread in reptiles, and multiple paternity is commonly observed in many species (Uller & Olsson 2008). Multiple paternity may either occur when a female mates with more than one male during a reproductive cycle, or when multiple matings are coupled with sperm storage across reproductive cycles. In freshwater turtles, multiple paternity was first confirmed in a population of wood turtles (*Glyptemys insculpta*) (LeConte 1830) in Pennsylvania (Galbraith 1991). Since then, the mating systems of turtles have been widely studied, and the frequency of clutches with multiple paternity has been found to vary greatly among species as well as among populations of the same species (McTaggart 2000, Pearse et al. 2001, Pearse et al. 2002, Refsnider 2009, Davy et al. 2011, Banger et al. 2013, McGuire et al. 2015).

Multiple paternity within a clutch is also favored by sperm storage in the female oviduct, which allows temporal asynchrony between mating events and fertilization events (Birkhead & Møller 1998). The females of certain freshwater turtle species may store fertile sperm for up to four years (Gist & Jones 1989), and therefore, may use sperm from one mating event to fertilize gametes in successive clutches (within or between reproductive seasons), resulting in repeated paternity (Fitzsimmons 1998, Pearse et al. 2002, Roques et al. 2006). The genetic variability of the offspring may then be temporally lagged, and this process could sustain genetic diversity within a population (Davy et al. 2011). However, sperm storage cannot be easily discriminated from intentional or non-intentional re-mating with the same partner.

Mating systems are of great importance for species conservation and for the implementation of a management plan (McGuire et al. 2013). Multiple paternity and repeat paternity affect the effective population size (Sugg & Chesser 1994, Theissinger et al. 2009) as well as the coancestry of a cohort (Scribner et al. 1993). Together, promiscuity and sperm storage may reduce the loss of genetic variation in small or declining populations by increasing genetic diversity within clutches (Calsbeek et al. 2007). The same processes also create variation in male reproductive success. Furthermore, high within-clutch genetic variation may increase the chance that some hatchlings will survive each year (Kaplan & Cooper 1984, Yasui 1998).

The wood turtle is a terrestrial freshwater turtle that inhabits the Great Lakes region. The International Union for Conservation of Nature (IUCN) has listed the species as endangered since 2013 (van Dijk & Harding 2013). It is also currently listed as endangered or threatened in most of the US states and Canadian provinces where it occurs (COSEWIC 2007). Wood turtles mate throughout the active season (April to September), mostly in the spring and the fall (Kaufmann 1992, Walde et al. 2003), when they aggregate near hibernation sites sometimes located several kilometers away from the nesting sites (Brooks & Brown 1992). They reach sexual maturity at 11-22 years (Brooks et al. 1992, Walde et al. 2003). The females usually lay one clutch per year (average of 8-12 eggs), but not every year (Powell 1967, Walde 1998). The mortality of embryos varies between 20 and 100% depending, for example, on nesting site quality and predation levels (Brooks & Brown 1992). During the nesting season,

aggressive interactions have been observed among individuals of the same sex and also between sexes (Kaufmann 1992, Walde et al. 2003, Parren 2013). The size, sexual maturity and level of aggressiveness of individuals may determine social hierarchies, with dominant males benefiting from extended copulations with females (Kaufmann 1992) and exhibiting higher reproductive success (Galbraith 1991). Multiple paternity has, in fact, been detected in 50% of wood turtle nests using DNA multilocus fingerprinting (Galbraith 1991).

With the development of microsatellite markers (Gerber et al. 2000), non-harmful tissue sampling, and robust statistical tools (Jones & Ardren, 2003, Jones et al., 2010), the potential of parentage analysis has significantly increased since Galbraith's 1991 study. Considering the conservation status of wood turtles and the paucity of information currently available on its mating system (Pearse & Avise 2001), we carried out parentage analyses on successfully emerged hatchlings over two successive years using microsatellite markers. Our main objectives were (1) to estimate the minimum number of males typically involved to sire a clutch (i.e. multiple paternity), (2) to identify clutches in adjacent years sired by the same males (i.e. repeat paternity) and (3) to quantify the sources of variation in the number of emerged hatchlings of each father (i.e. variation in male reproductive success).

## **2.4 Materials and methods**

### **Sampling**

Hatchlings were sampled directly from a nesting site adjacent to the Shawinigan River in the Mauricie region (Québec, Canada) in 2006 and 2007. The population size was estimated to be between 191 and 285 individuals (Bourgeois et al. 2009, Walde et al. 2003). The area was observed daily in June for signs of nesting activity. All clutches were protected within an enclosure until mid-August, where nests were visited daily until the hatchling period began, and up to 8 times per day until the end of the hatchling period. Tissue samples were collected from all emerged hatchlings through a superficial excision (2 mm x 1 mm) at the posterior edge of the shell. All tissue samples were stored in 95% ethanol. Whenever possible, genetic sampling of the females was performed at time of nesting.

Adult turtles were also observed opportunistically on the nesting site, or during annual surveys of the population conducted between 1996 and 2015. Captured individuals were marked (Cagle 1939), measured, weighed, aged, sexed (see Walde et al. 2003), and forelimb skin samples (1 mm x 3 mm) were collected for genotyping. From these individuals, a pool of candidate parents expected to be sexually mature in 2006 and captured at less than 6 km from the nesting site was selected for parentage analysis. The use of such candidate parents increases consistently the precision of parentage analysis (Jones et al. 2010).

### **DNA extraction and amplification**

Total genomic DNA was extracted using PureLink™ Genomic DNA extraction kits (Invitrogen) following the manufacturer's protocol. Genotypic profiles were obtained for seven microsatellite loci; GmuD16, GmuB21, GmuD28, GmuD40, GmuD87, GmuD88, and GmuD93 (King & Julian 2004). Polymerase chain reactions were carried out using a Thermo cycler 9790 (Biosystems) using the cycling parameters and procedures described in King and Julian (2004). Fragment analysis was performed on PCR products using an ABI-3730XL DNA analyzer at the McGill University Genome Quebec Innovation Center. All alleles were manually scored three times using GeneMarker (SoftGenetics LLC, PA), and 15% percent of

all individuals were independently amplified and genotyped twice to confirm the observed genotypes. Any samples displaying unexpected alleles were re-amplified and reanalyzed to confirm their unique genotype. Hatchlings were excluded from the analysis if two loci could not be scored.

## **Population Genetics**

The genotypes of all candidate parents and nesting females were used to provide information on each locus and to estimate allele frequencies for the populations. We assessed whether each locus conformed to the Hardy-Weinberg equilibrium (HWE), and linkage disequilibrium was tested for each pair of loci using the Markov Chain Monte-Carlo method in GENEPOP 3.1 (Raymond & Rousset 1995a, Rousset 2008). The data were also analyzed for allelic stutter, large allele dropouts, and null alleles using MICROCHECKER (Van Oosterhout et al. 2004). A sequential Bonferroni correction was applied to all multiple pairwise comparisons (Rice 1989). The probability of identity (PI) and observed and expected heterozygosity ( $H_O$  and  $H_E$ ) for each locus were estimated with GenAlEx v6.5 (Peakall & Smouse 2006, 2012). We also calculated PISibs, the probability of identity taking into account the possibility that close relatives were sampled (Taberlet & Luikart 1999, Waits et al. 2001). The probability of detecting the presence of alleles from more than a single putative father was calculated as the detection index (dIndex) (Westneat et al. 1987). This method relies on allele frequencies within the population and the genotype arrangements of all possible mating within a representative sample of the population (FitzSimmons 1998). The genotypic correspondence between the mothers sampled at time of nesting and their offspring allowed us to estimate a genotype error rate and a germinal mutation rate for each locus. Likely germinal mutations were confirmed if both alleles did not match any of the maternal alleles, and if the offspring alleles were 4bp larger or smaller than expected (FitzSimmons 1998). All germinal mutations were corrected by the suspected mutated maternal allele prior to the parentage analysis.

## **Assignment of candidate mothers and fathers**

In instances where the mother of a clutch cannot be sampled during nesting activities, assignment of a candidate female as the mother was performed based on the concordance between two parentage programs. COLONY v2.0.6.1 (Jones & Wang 2010) computes prior probabilities derived from clutch-level data (e.g. kinship maternity group) to identify offspring from the same clutch and their probable mother. These assignments were further confirmed with CERVUS v3.0.1 (Kalinowski et al. 2007), a categorical allocation program that relies on likelihood-based approaches to select parents (Jones et al. 2010). Individual exclusion probabilities were first calculated for all possible female-offspring pairings on an individual-based level. The candidate females whose genotypes mismatched at one or more loci were excluded as potential mothers. When more than one female could not be excluded, a log odds (LOD) score was calculated to estimate the likelihood ratio of each candidate female as the true mother compared to an arbitrary mother. The difference in LOD scores between the two most likely candidates ( $\Delta$ ) provides an indication of the reliability of the assignment. By simulation, a critical  $\Delta$ -LOD score was computed at 80% and 95% confidence levels, and the mean  $\Delta$ -LOD score for each clutch was estimated. A female was assigned to a nest only when (1) COLONY provided a 95% confidence level, (2) CERVUS assigned all offspring from that nest to the same mother and (3) when the mean  $\Delta$ -LOD score for all offspring in the nest was positive. In cases where two putative mothers were assigned to the same nest, the female with the highest mean  $\Delta$ -LOD score was selected. Candidate fathers were then assigned following the same procedure. All immature individuals from the pool of candidate parents were used either as candidate mother or father.

## **Multiple paternity analysis**

Multiple paternity and estimates of the minimum number of fathers contributing to each nest were computed using three distinct approaches: (1) a method of Minimum Allele Counts (MAC) (FitzSimmons 1998, Myers & Zamudio 2004), (2) GERUD v2.0 (Jones 2005) and (3) COLONY v2.0.6.1. For each clutch with at least three emerged hatchlings, multiple paternity was inferred using a consensus approach (Theissinger et al. 2009), estimating the number of

fathers based on the majority results. Clutches with only one or two hatchlings were excluded from multiple paternity analyses.

First, we applied the minimum allele count method for the multiple paternity analysis. When the maternal genotypes contributing to each clutch were known (either at time of nesting or genetically assigned from the pool of candidate parents), the minimum number of fathers was determined by subtracting the maternal alleles from the hatchling genotypes. Alleles that were absent in the maternal genotype were then identified as paternal alleles. To provide conservative estimates, it is commonly assumed that all fathers are heterozygous at each locus. In cases where heterozygous offspring had a genotype identical to that of the mother, one of the two alleles was considered to have been inherited from the father. When the minimum number of paternal alleles required to match the observed genetic variation in a clutch was greater than two, multiple paternity was assumed for that locus. To account for mutations, which may overestimate the number of paternal alleles, multiple sires were inferred only when at least two loci provided evidence for multiple paternity. For nests with unknown mothers, the numbers of fathers were estimated based on all alleles found in the clutch divided by two. Mutation and genotyping errors were accounted for as previously described.

The number of sires for each sampled clutch was also inferred with the GERUD v2.0 software. GERUD is a parental reconstruction program, which performs an exhaustive algorithm that tests all possible parental genotypes against the progeny's array to find the minimum number of genotypes explaining that array. In the case of several possible father genotypes, GERUD computes a ranking of the solutions by likelihood. The likelihood score is based on the segregation of paternal alleles, their deviations from Mendelian expectations and the population allele frequencies. These analyses were performed using the seven microsatellite loci and setting the parameter for the maximum number of possible fathers to four as suggested by Theissinger et al. (2009). For females with clutches sampled in successive years (2006 and 2007), each nest was run separately and pooled, to determine the minimum number of fathers. Since GERUD cannot estimate paternity with datasets containing mutations, genotyping errors or missing data, all loci missing in the maternal genotypes were

excluded from the analyses. All offspring lacking maternal alleles at one locus only were corrected by one of the maternal alleles (Theissinger et al. 2009, Ekanayake et al. 2013), whereas those with an extra paternal allele for only one locus were either corrected (when a mutation was suspected), or eliminated (Duran et al. 2015). Mutation was suspected when only one offspring at only one locus had a supplementary allele of 4pb larger or smaller than the other offspring. When the maternal genotype was not available, GERUD determined all possible mother genotypes and all possible combinations of mother-father genotypes before ranking them based on likelihood scores, as previously described.

COLONY was also applied to assign sibship and parentage based on a maximum-likelihood model. Offspring were first clustered into full sibs and half-sibs and the genotype of putative parents was then constructed when it was not available at time of sampling or assigned to candidate parents. The default parameters were used to run COLONY, including a single long-length run with full likelihood analysis by assuming polygamy for both females and males, and by setting a medium prior distribution to an average of six offspring for both females and males. Three runs were performed to detect multiple paternity: one for the nesting season of 2006, one for the nesting season of 2007, and one for both nesting seasons combined. COLONY also allows accounting for two classes of errors in the computations: null alleles and allelic dropout (class I); and typing errors and mutations (class II). A minimal error rate of 0.04 was assumed for the allelic dropout for all loci. For GmuB21 and GmuD93, the class I error was set to 0.0979 and 0.0664 respectively, following detection of null alleles by MICROCHECKER. Genotyping error rates were estimated for each locus using mother-offspring genotype comparisons and PCR re-amplifications.

## **Genotype reconstruction of the fathers**

The minimum number of males required to explain all offspring genotypes was estimated with COLONY. Iterative runs were thus used for the identification of paternal half-sibs, which would identify males that sired offspring with more than one female, in more than one season. In a first run, all nests showing a consensus of single paternity were determined using the same method as previously described, and corresponding clutches were fixed in successive runs until



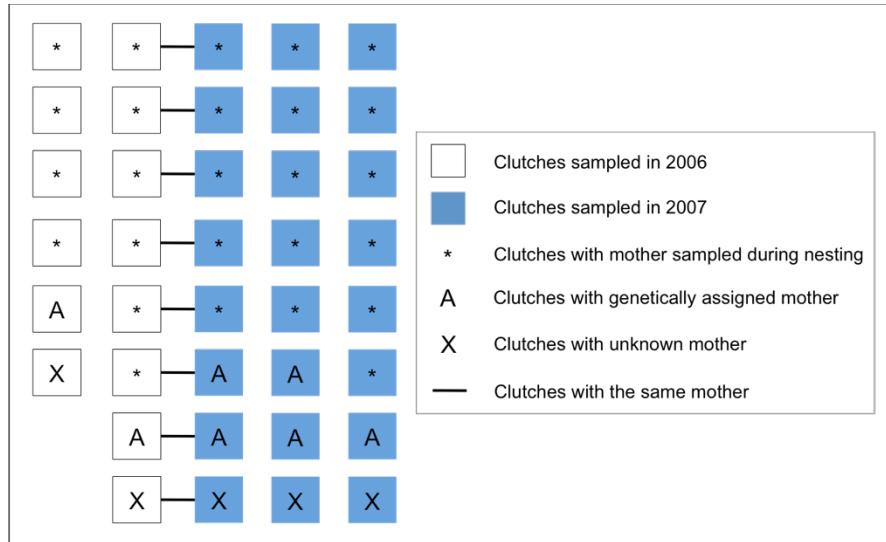
all nests had been assigned to a single father or to multiple fathers. Paternal genotypes were reconstructed from different COLONY runs, and repeat paternity was confirmed when the same male sired offspring in different clutches with the same female.

## **Statistical analysis**

The relationship between the genetic diversity of clutches and the number of fathers was tested using partial Kendall's rank correlation, while controlling for the genetic diversity of the mother. Only clutches with known mothers (i.e. sampled during nesting or assigned from candidate parents) were used for this analysis, and mean observed heterozygosity ( $H_O$ ) values were estimated with GenAlEx v6.5 (Peakall & Smouse 2006, 2012). To test whether male reproductive success was higher in clutches with simple paternity compared to clutches with multiple paternity, the clutches were split into two groups. A Student's t-test (999 permutations) was then computed to test for differences between the means of emerged hatchlings in the corresponding groups. Finally, Kendall's rank correlations were used to assess male reproductive success by testing the relationship between the number of offspring and the number of mates and clutches sired by each father. Statistical analysis was conducted with R v3.3.2 (R core team, 2016) using the "ppcor" package (Kim 2015).

## **2.5 Results**

A total of 38 clutches (14 in 2006, 24 in 2007) were sampled (Figure 2.1). Hatchling success was 76% in 2006, and 70% in 2007. An average of 7 hatchlings per clutch (range of 1-14, SD = 3.24, variance = 14.7) were genotyped, representing 92% of the successfully hatched offspring. Five offspring were excluded from the analysis because of amplification failure. Therefore, parentage analysis was run for 248 successfully emerged hatchlings, 20 mothers sampled at time of nesting and 64 other adults collected from this population to serve as a pool of candidate parents (10 males, 47 females, 7 immature individuals at time of capture).



**Figure 2.1** Details of the clutches of wood turtles sampled in the Shawinigan population over two successive years (2006 and 2007).

The number of alleles for all seven loci ranged from 7 to 14 (Table 2.1). Deviations from HWE were detected at two loci, GmuB21 and GmuD93, likely due to null alleles. The larger error rates found at these loci (0.04 and 0.05, respectively) confirm these results. There were two hatchlings containing a likely germinal mutation of the maternal allele. Values of dIndex ranged from 0.56 to 0.77 at each locus, for a global detection rate of 0.99. The probability of identity (PI) was  $4.4 \times 10^{-10}$ , and PIsibs was 0.0004. Overall, the seven loci allowed us to accurately detect the relationship between siblings.

**Table 2.1** Characteristics of each of the seven loci, including the numbers of alleles, expected ( $H_E$ ) and observed ( $H_O$ ) heterozygosity, Hardy-Weinberg equilibrium (HWE) p-value, probability of detecting multiple paternity (dIndex) at each locus within their clutches (i.e. probability of exclusion), probability of identity (PI), probability of identity for siblings (PISibs), error rate (Error), and the germinal mutation rate (Mutation). \*Significant departure from HWE following a Bonferroni correction.

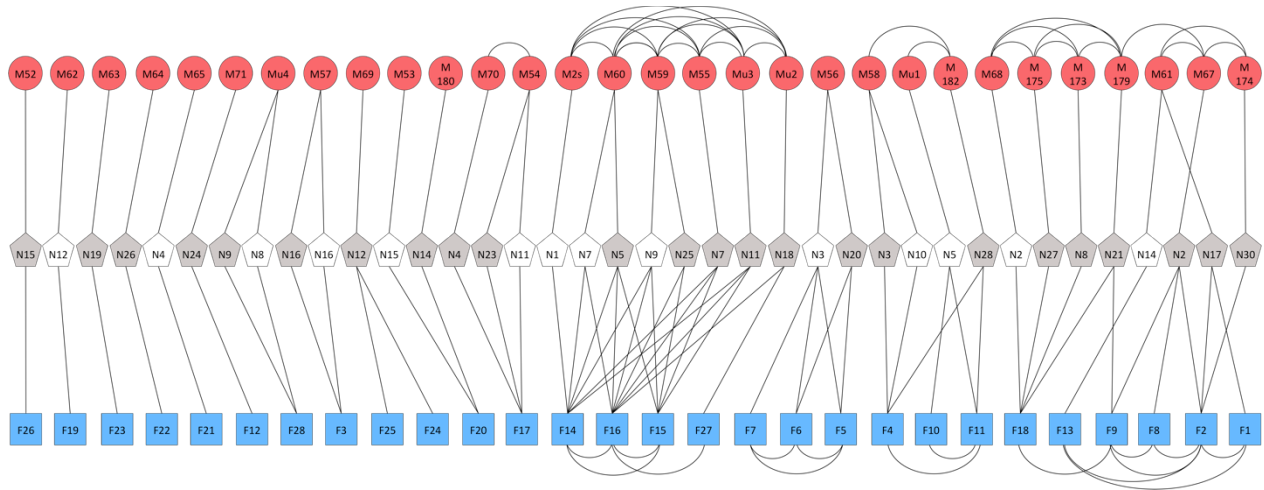
Locus	No. Alleles	$H_E$	$H_O$	HWE	dIndex	PI	PISibs	Error	Mutation
GmuD16	14	0.90	0.92	0.05	0.77	0.019	0.306	0.0097	0.0049
GmuB21	7	0.79	0.62	0.004*	0.56	0.079	0.376	0.0583	0
GmuD28	8	0.85	0.89	0.38	0.58	0.079	0.377	0	0
GmuD40	9	0.83	0.77	0.09	0.60	0.055	0.353	0.0097	0
GmuD87	10	0.78	0.81	0.03	0.57	0.081	0.381	0.0049	0
GmuD88	10	0.85	0.85	0.34	0.69	0.040	0.337	0.0097	0
GmuD93	11	0.81	0.69	0.003*	0.66	0.062	0.363	0.0437	0.0049
Total	69	-	-	-	0.99	$4.4 \times 10^{-10}$	0.0004	-	-

COLONY and CERVUS allowed us to confirm the identification of the 20 mothers sampled at time of nesting to their 25 corresponding nests, and to further assign 6 females from the pool of candidate parents to 7 clutches for which the mother was unknown at time of nesting (Figure 2.1, Table 2.2). Based on these assignments, we ended up with 33 clutches with known mothers, and 5 with unknown mothers. Hatchlings from one clutch per season in both years of the study were analyzed for eight females. None of the 10 males and the 7 immature individuals sampled in the population was successfully assigned to any offspring.

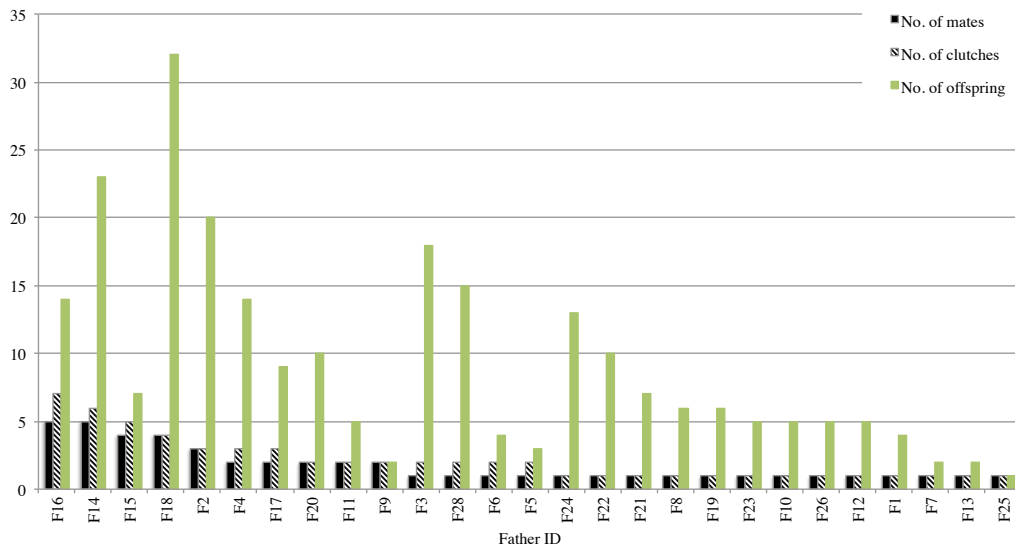
**Table 2.2** Probabilities and  $\Delta$ -LOD score for females assigned to the seven clutches without a known mother during the 2006 and 2007 nesting seasons, estimated by COLONY and CERVUS, respectively. Critical values for the relaxed and strict confidence levels are 1.77 and 4.92, respectively. B $\Delta$ -LOD score for female M180 is 6.46 and M181 is 4.44.

Clutch Id	Assigned female	$\Delta$ -LOD <sup>a</sup>	Probability
1-2006	2SHA	0.67	1.00
11-2006	M54	0.83	0.96
30-2007	M174	0.71	1.00
14-2007	M180	0.95	1.00
24-2007 <sup>b</sup>	M71	3.12	1.00
23-2007	M54	0.79	1.00
20-2007	M56	4.93	1.00

COLONY determined the estimated number of fathers to be 28 (17 males for 14 clutches in 2006, and 23 males for 24 clutches in 2007), with 12 of those fathers identified as identical for both nesting seasons (Figure 2.2). The male to female sex ratio of reproductive individuals was therefore estimated at 1:0.824 in 2006, 0.958:1 in 2007, and 0.933:1 for both nesting seasons. The reproductive success of males ranged from 1 to 32 offspring (mean of 8.86 per male, variance = 53.5) and showed that some males contributed more than others to the clutches (Figure 2.3). Ten males produced successfully emerged offspring with more than one female, and 14 males produced offspring in more than one clutch (e.g. with the same female for two nesting seasons). The mean number of mates per male was 1.24 in 2006, and 1.78 in 2007. On average, each male had 1.75 mates for both years, and produced offspring in 2.11 different clutches. Male reproductive success was positively correlated with the number of mates (Kendall's Tau = 0.43,  $p = 0.006$ ) and the number of clutches sired (Kendall's Tau = 0.45,  $p = 0.003$ ).



**Figure 2.2** Relationships between all successfully hatched offspring sampled and adults (red circles: mothers; blue squares: fathers, white pentagons: clutches sampled in 2006; gray pentagons: clutches sampled in 2007). Loops indicate individuals that produce offspring with the same partners.



**Figure 2.3** Histogram representing the numbers of mates, sired clutches, and emerged offspring for each reconstructed genotypes of putative fathers estimated on successfully hatched offspring of clutches sampled in 2006 and 2007.

Eight females were found to have repeated clutches in both 2006 and 2007 (Table 2.3), from them only one female (M61) had offspring from different males between the two years. Repeat paternity was observed for all other seven females. Among those, four mothers (M54, M57, M58, and Munk4) produced viable offspring with a single male (single repeat paternity), while three mothers (M56, M59 and M60) had offspring with different males (multiple repeat paternity). The same three fathers were identified in the nests of females M59 and M60.

**Table 2.3** Repeat paternity results for eight females with clutches in both nesting seasons, and the father Id assigned to their successfully emerged offspring of each clutch.

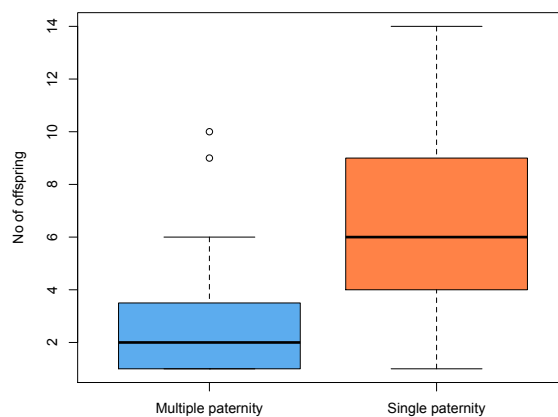
<b>Mother Id</b>	<b>Clutch Id</b>	<b>N</b>	<b>Father Id</b>
M61	14-2006	2	F13
	17-2007	7	F1, F2
M58	10-2006	7	F4
	3-2007	4	
M57	16-2006	10	F3
	16-2007	8	
M54	11-2006	1	F17
	23-2007	4	
Munk4	8-2006	9	F28
	9-2007	6	
M60	7-2006	2	F14, F16
	5-2007	10	F14, F15, F16
M59	9-2006	8	F14, F15, F16
	25-2007	3	F15, F16
M56	3-2006	6	F5, F6, F7
	20-2007	3	F5, F6

Multiple paternity was detected with all three methods (minimum allele count, GERUD and COLONY) for 26.7%, 36.7% and 50% of the clutches sampled, respectively, and a consensus of 36.7% (Table 2.4). The results provided by the MAC method and GERUD were quite similar with the minimum number of fathers per clutch ranging from 1 to 3. Results of

COLONY ranged from 1 to 6 per clutches, identifying multiple paternity in 10 of the clutches for which the other two methods assigned a single father. The average number of mates per female was 1.31 in 2006, 1.39 in 2007, and 1.43 for both years combined. The number of fathers in a clutch was positively correlated with the mean genetic diversity of the successfully emerged offspring, when controlling for the mother's genetic diversity (Partial Kendall's Tau = 0.27,  $p = 0.033$ ). Male reproductive success was significantly lower in clutches with multiple fathers compared to clutches with a single father (Student's t-test,  $p = 0.001$ ) (Figure 2.4).

**Table 2.4** Percentages of multiple paternity obtained from the minimum allele count, GERUD and COLONY methods, for each season and for both sampling seasons combined. Multiple paternity was confirmed when the consensus of all three methods detected a minimum of two sires in the successfully emerged offspring of each clutch sampled.

Season	Allele count	GERUD	COLONY	Consensus
2006	23.08	23.08	46.15	23.08
2007	30.43	39.13	65.22	39.13
Combined	26.67	36.67	50.00	36.67



**Figure 2.4** Boxplots representing the mean number of emerged offspring per father in clutches having multiple fathers versus single fathers (Student t-test,  $p = 0.001$ ).

## 2.6 Discussion

### Multiple paternity

Multiple paternity is frequently estimated in parentage studies as it can be measured directly through kinship data with minimum details about the identity of parents. In this study, we confirmed that multiple paternity is present in at least 37% of the clutches of wood turtles sampled. Polyandry and multiple paternity have been detected in most of the turtle species studied to date (reviewed in Uller & Olsson 2008) but results of parentage analysis cannot easily be compared among species (Davy et al. 2011). Sampling design varies greatly from one study to the next, which makes comparison based on different numbers of sampled clutches and different microsatellite loci difficult (Veliz et al. 2017). For example, the probability to detect multiple paternity increases with the number of clutches analyzed (Pearse et al. 2002), but it is less likely to be detected in small clutches of fewer than five individuals (McCracken et al. 1999). In the present study, multiple paternity was only observed in two of the smaller clutches (< 4 offspring) tested. After removal of these ten clutches from the analysis (results not shown), the occurrence of multiple paternity decreased from 37% to 28%.

Observed variation in the estimated minimum number of fathers in a clutch could also be attributable to the methods employed. For example, COLONY may provide an overestimate with respect to MAC and GERUD. It has been shown that COLONY performs better with five to seven loci, and for genotypic data with  $H_E$  values larger than 0.84 (Sefc & Koblmüller 2009). When the genetic markers do not meet these criteria, the number of parents may be overestimated (Jones et al. 2007), whereas GERUD appears to be more accurate (Jones 2005). This may be the case here, considering that only 3 of the 7 loci analyzed had  $H_E$  values larger than 0.84.

Our results on multiple paternity contribute to a body of literature on paternity analysis in reptiles. Several distinct hypotheses have been proposed to better understand the benefits of multiple fathers in a clutch (Jennions & Petrie 2000, Uller & Olsson 2008). For one, multiple paternity increases the genetic diversity of clutches (Yasui 1998). As suggested here, the



number of fathers who participated towards a clutch is positively correlated with the neutral genetic diversity of the clutches. This positive relationship has been confirmed in lizards for which clutches with only sibs were less diversified than clutches with half-sibs (Calsbeek et al. 2007). This reproductive strategy is predicted to be preferred in unstable and heterogeneous habitats, where it is difficult for a female to predict the best mate for the future. As such, females can minimize the variance in male quality by mating with multiple partners, a strategy known as ‘genetic bet-hedging’ (Watson 1991), or ‘diversified bet-hedging’ (Kaplan and Cooper 1984). By producing a genetically larger range of offspring in a clutch, the probabilities that some of them will survive increase (Yasui 1998). As we used microsatellite markers to measure neutral genetic diversity, however, variation in levels of adaptive genetic diversity (Holderegger et al. 2006) or genomic diversity (Väli et al. 2008) could not be assessed. The observed correlation between the number of fathers and the genetic diversity of a clutch may then provide negligible benefits for the female wood turtles.

## **Repeat paternity**

Our study observed for the first time repeat paternity between clutches in adjacent years for the wood turtle. This reproductive strategy appears to be common in the population under study, as our results identified a very high frequency of repeat paternity for 87.5% (7/8) of the females with nests sampled in successive years. This trend is consistent with what has been observed in other freshwater turtles and it may be due either to sperm storage, or re-mating with the same partner in successive years (Phillips et al. 2013). For example, a four-year study detected repeat paternity in 38% of the clutches in the Painted turtle (*Chrysemys picta*) (McGuire et al. 2014). In Blanding’s turtles (*Emydoidea blandingii*), repeat paternity was also observed 69.9% of sequential clutches over eight years (McGuire et al. 2015).

The storage of viable sperm for over one year has been documented for many species of freshwater turtles (Ewing 1943, Gist & Jones 1989, Pearse et al. 2002, Roques et al. 2006, McGuire et al. 2011), but was never observed in the wood turtle. In addition to favoring multiple paternity, sperm storage would allow for post-copulatory selection such as sperm competition and female cryptic choice (Stockley 1997, Jennions & Petrie 2000). The sperm of

high quality males or more compatible males may then be used for several years while allowing the female to avoid the potential cost of re-mating caused by aggressive interactions (Evans 1961, Uller & Olsson 2008, Ernst & Lovich 2009, Liu et al. 2013). Furthermore, the temporal asynchrony between the mating and fertilization events allowed by long-termed sperm storage can be an advantage for female wood turtles, in particular for those that may not nest on a yearly basis. Fertilization is then assured for such females in case of low mate encounter rates (Olsson et al. 1994, Phillips et al. 2013).

Population density can affect repeat paternity as the potential lack of encounters with mates may favor the use of sperm storage or the probability to re-mate with the same partner (Gist & Congdon 1998). The density of the Shawinigan population is estimated at 0.44 adults/ha (Walde et al. 2003), which is quite low compared to other populations of wood turtles. Namely, population density is estimated at 4.41 adults/ha in Pennsylvania (Ernst 2001), and at 12.5 adults/ha in New Jersey (Harding & Bloomer 1979). In spite of this density, frequent observations of multiple mating have been confirmed in the Shawinigan population (Walde et al. 2003), including a female that bred with five males. Repeat mating events between the same individuals in separate years have also been confirmed. Our results highlight the presence of a number of small clusters of interconnected males and females (Figure 2.2). Such clusters could suggest strong mate choice or potential long-term preferences for re-mating between some individuals. Because pair-bonding has not been observed in freshwater turtles (Galbraith et al. 1993, Pearse & Avise 2001, Roques et al. 2006), re-mating between the same partners is deemed to be unintentional. On the basis of above behavioral observations and the fact that 77% of mating and courting events are observed in the fall when turtles congregate in aquatic habitats (Harding & Bloomer 1979, Walde et al. 2003), it seems likely that repeat paternity may be the result of re-mating events with the same partner. As already observed in snapping turtles (*Chelydra serpentina*) (Brown & Brooks 1994), wood turtles may return to the same overwintering sites, which also increases the probability of mating with the same partner over successive years (Bloomer 1978, Harding & Bloomer 1979).

Considering the life history traits of wood turtles in combination with our results, we believe that the use of sperm storage is the best hypothesis to explain in large part the high frequency of repeat paternity observed. The reconstructed genotypes from successive clutches suggest that three of the seven females would have had to mate with the same multiple males (two or three) for at least two years to support the alternative re-mating hypothesis. Our results identified a single female who mated with completely different males in successive clutches, implying separate reproductive events with at least three males in two reproductive seasons. Yet, on the basis of genetic parentage analysis alone, distinction between the two hypotheses remains problematic. In long-term studies, the use of sperm storage and repeat mating with the same partner may possibly be discriminated with the decreasing fertility of stored sperm over time (Cuellar 1966, Jun-Yi 1982, for an example see McGuire et al. 2013) as a result of sperm depletion or declining sperm viability (Goin et al. 1978, Gist & Jones 1987, Palmer et al. 1998).

### **Male reproductive success**

The number of emerged offspring sired per male varies according to their number of mates, clutches per mates (i.e. repeat paternity), and the proportion of offspring sired within a clutch. Our results exhibited a large variation in male reproductive success over two years. This variation was positively correlated with the number of mates and the number of clutches sired. The higher correlation observed between the number of clutches sired and the number of mates highlight the importance of repeat paternity in male reproductive success. Yet, our results also revealed a significant difference in male reproductive success in clutches with multiple paternity or single paternity. Clutches with multiple paternity resulted in lower numbers of emerged offspring per father. For a given male, the best strategy to increase its reproductive success would be to (1) mate with a high number of females (i.e. multiple mating) or re-mate with the same females (i.e. repeat paternity); (2) mate with a female that will use long-term sperm storage (i.e. repeat paternity); and (3) be the only sire contributing to a clutch (i.e. single paternity).

As some males contributed more than others to the sampled clutches (Figure 2.3), the hierarchical dominance observed in Pennsylvania (Galbraith 1991, Kaufman 1992) may also be present in the Shawinigan population. These social behaviors would be more frequent near overwintering sites since the population density is usually quite low during the summer and the probability of encounters between more than two individuals is very unlikely. Kaufman (1992) observed that males with higher dominance ranks have better reproductive success, but we obviously could not establish a ranking order solely based on the reconstructed genotypes. Behavioral observations such as male and female sizes, level of aggressiveness, and the number of mating events, are required to correlate hierarchical dominance with mating success and to better understand its impact on male reproductive success. Besides, since postcopulatory sexual selection affects the reproductive success of males (Birkhead & Pizzari 2002), actual numbers of mating events are not always directly correlated with numbers of sired successfully emerged offspring.

Clutches size, nesting frequency, quality of nesting sites, and sperm storage may also increase variation in male reproductive success. Some studies on freshwater turtles combining long-term population monitoring and parentage analysis contribute to discriminating the relative contribution of number of mates and mate quality to male reproductive success (Pearse & Avise 2001, Pearse et al. 2001, 2002, McGuire et al. 2011, 2014, 2015). Namely, McGuire et al. (2015) observed that the clutch size and egg width were correlated with the female body size in Blanding's turtles. Males that mate with bigger females would then be favored.

Egg incubation in nature is causing a lot of uncertainties for parentage analysis. The number of emerged hatchlings may not accurately represent the actual contribution of a male to a clutch, as some males may contribute to clutches with low survival rates. A long-term monitoring of the principal nesting site in the Shawinigan population was carried out by Masse et al. (2014), allowing for a better estimation of nesting success. The principal causes for nesting failure were the poor climatic conditions (10% of the clutches of 2007), the presence of parasites (9% of the clutches in 2006), infertility (12% in 2006, 11% in 2007), and embryonic mortality (14% in 2006 and 16% in 2007) (Masse et al. 2014). As such, the results of our

paternity analysis are clearly affected by numerous other factors not directly related to male quality.

In spite of a conservative approach to determine the number of fathers and their genotypes, some potential sources of error may affect the interpretation of the results. For instance, the reconstructed genotypes of the fathers were extremely similar in some of the clutches, which suggests that some fathers could be close relatives. As argued by Peare and Parker (1996), green turtles (*Chelonia mydas*) nesting closer to each other tend to be more related. Wood turtles may live up to 50 years in the wild (Cameron et al. 2002), and some mature individuals could be related, namely as father and son. However, considering that only one out of 100 offspring is likely to reach sexual maturity (Klemens 2000), the probability that two brothers or cousins sired offspring from the same female is extremely unlikely. These similar genotypes may also be caused by genotyping errors or germinal mutations in the offspring. Instead of assigning a clutch to a single father, a mutation would likely introduce an additional allele, and a second father would be assigned to match the genotypes of the offspring. To circumvent this problem, genotyping error rates were estimated for each of the COLONY runs, and 15% of all individuals were amplified and genotyped twice to confirm the observed genotypes.

Based on the occurrences of multiple paternity and the extent of repeat paternity, our results may be used to provide recommendations for the conservation of the wood turtle in Canada. Namely, considering the relative contributions of males and females to the next generation will be crucial to foster recruitment of new individuals and the long-term survival of endangered populations. Our results showed sex ratio close to 1:1 for both nesting seasons. Yet, the number of males involved in reproductive output is similar to that of females, although a higher variance in reproductive success is observed for the males, probably due to repeat paternity. When higher variance is consistently observed over successive generations, the disproportionate male contribution to the population can result in a lower effective population size (Hedrick 2005, Karl 2008). On the other hand, multiple paternity, by increasing the number of males contributing to offspring, is frequently considered to have a positive impact

on effective population size (Sugg & Chesser 1994, Pearse and Anderson 2009, Lotterhos 2011). Because both multiple paternity and repeat paternity may be the result of sperm storage yet have opposite effects on male and female reproductive success, further studies will be needed. Moreover, to better understand the long-term effect of repeat paternity in the mating system of wood turtles, we suggest conducting a future study on hibernation site fidelity. Because site fidelity may lead to reproductive isolation, it is of great interest for the species conservation and management.

## 2.7 Acknowledgments

We would like to thank D. Masse (Parks Canada), Y. Robitaille, C. Jutras (Ministère des Forêts, de la Faune et des Parcs du Québec), and D. Pouliot for providing samples. The sampling protocol was approved by the Ministère des Forêts, de la Faune et des Parcs du Québec. Improvements of earlier drafts of the manuscript are the results of comments from A. Rogic and the comments of three anonymous reviewers. This study was supported by a Natural Sciences and Engineering Research Council grant RGPIN/05219 to F.-J. Lapointe and a research grant from Parks Canada.

**COMPLIANCE WITH ETHICAL STANDARDS** The authors declare no conflicts of interest.

*Data Availability* – We have deposited the primary data underlying these analyses as follows:

– Microsatellites genotypes: Dryad

# Chapitre 3. La génétique du paysage de la tortue des bois

## Watershed influence wood turtle's (*Glyptemys insculpta*) genetic structure

Cindy Bouchard<sup>1</sup>, Nathalie Tessier<sup>2</sup>, François-Joseph Lapointe<sup>1</sup>

1. Département de sciences biologiques, Université de Montréal, Montréal, QC, Canada

2. Ministère des Forêts, de la Faune et des Parcs, Longueuil, QC, Canada

Publié dans: Conservation Genetics, 20(3), 653-664.

Copyright Springer Nature B.V. 2019

DOI 10.1007/s10592-019-01169-1

*Format adapté pour les besoins de la thèse*

### 3.1 Résumé

La tortue des bois (*Glyptemys insculpta*) est une espèce de tortue d'eau douce endémique à l'est de l'Amérique du Nord, présentement considérée comme espèce *en danger* par l'Union internationale pour la conservation de la nature. Les populations locales de tortue des bois sont considérées comme des unités distinctes pour la gestion de l'espèce. Elles sont définies comme un ensemble d'individus pouvant se reproduire entre eux et occupant un même bassin versant. Cependant, aucune étude à ce jour n'a pu vérifier si les bassins versants représentent une unité adéquate pour expliquer la différenciation des populations locales de tortue des bois. L'objectif principal de cette étude visait à caractériser génétiquement les tortues des bois dans la partie nord de leur aire de distribution, afin de tester l'hypothèse d'isolation par bassin versant. Le génotypage de 331 individus à l'aide de neuf locus microsatellites fut réalisé sur des tortues échantillonnées dans 24 cours d'eau provenant de 12 bassins versants. La diversité génétique était similaire entre les sites d'un même bassin versant et l'hétérozygotie observée entre tous les sites variait de 0.677 à 0.754. Des analyses canoniques de redondance partielle ont ensuite confirmé que l'isolation par bassin versant expliquait significativement 18% du total de la variation génétique observée entre les sites, alors que les modèles d'autocorrélation spatiale et d'expansion postglaciaire n'étaient pas significatifs. Les analyses de groupement ont révélé une structure génétique substantielle qui sépare les individus échantillonnés en 10 groupes imbriqués hiérarchiquement. Ces différents groupes répondent ainsi au critère de distinction utilisé lors de l'établissement des unités désignables, et nous recommandons leur considération afin que le critère d'importance soit évalué.



## 3.2 Abstract

The wood turtle (*Glyptemys insculpta*) is a freshwater species endemic to eastern North America and is currently listed as endangered by the International Union for Conservation of Nature. Wood turtle local populations are considered units for the species recovery, and are defined as discrete interbreeding populations in a distinct watershed. However, there are no studies to date supporting watershed boundaries as an appropriate unit to represent local population differentiation. The main objective of this paper was to genetically characterize wood turtles from a northern portion of their range, and test an isolation-by-watershed hypothesis, the first of its kind at such a large geographical scale. Turtles were sampled in 24 watercourses from 12 different watersheds for a total of 331 individuals, each genotyped for nine microsatellite loci. Within each watershed, genetic diversity was similar between sites, and observed heterozygosity ranged from 0.677 to 0.754. Partial redundancy analyses then confirmed that watershed isolation contributed to 18% of the total observed genetic variation, while spatial autocorrelation and the post-glacial expansion model did not significantly explain any variation. Clustering methods revealed substantial spatial genetic structure, with sampled groups falling into ten nested hierarchical clusters. We recommend further consideration of these ten clusters to determine if they meet the evolutionary significance criterion to become designatable units, whose genetic structures were influenced by watershed structure.

**Key words** : Clustering, conservation genetics, freshwater turtles, landscape genetics, microsatellite, partial redundancy analysis, watershed.

### 3.3 Introduction

Landscape composition and configuration could impact the genetic differentiation of a species by isolating or restraining the movements of individuals between habitat patches. In addition, several natural processes and conditions, such as climate change, glacier retreat, and river networks are affecting the species geographical ranges, creating in turn distinct genetic patterns of population differentiation (Fahrig 2003, Manel et al. 2003). Habitat modification and loss caused by recent anthropogenic activities may also have an important effect on a species' dispersal ability, mortality rate, genetic diversity, population distribution and density (Templeton et al. 1990, Luque et al. 2012). All of these processes are operating at different rates from an evolutionary perspective. Temporal scale is thus a crucial factor to consider in genetic studies since an estimated 10 to 200 generations are required to detect any genetic variations caused by landscape modifications that affect gene flow (Landguth et al. 2010). In long-lived organisms, this implies that genetic differentiation among populations may take hundreds of years to appear, causing an important temporal lag between the phenomenon and its genetic response (Anderson et al. 2010, Segelbacher et al. 2010). This has been observed in small, isolated populations of long-lived species that exhibit high levels of genetic diversity, which did not reflect recent demographic bottleneck events (Kuo & Janzen 2004, Pittman et al. 2011, Davy and Murphy 2014). As a consequence, contemporary landscape modifications are often indistinguishable from past colonization events in population genetic studies (Wade & McCauley 1988). The field of landscape genetics (Manel et al. 2003) was specifically created to dissociate the confounding effects of past and present landscapes at small and large spatial scales (Zellmer & Knowles 2009, Anderson et al. 2010, Manel et al. 2013), thus allowing researchers to explore alternative hypotheses using novel approaches (Chikhi et al. 2010, Paz-Vinas et al. 2013).

According to the International Union for Conservation of Nature (IUCN) Red List of Threatened Species, more than two thirds of turtle species and subspecies were listed as threatened in 2000 (Gibbons et al. 2000, Turtle Conservation Fund 2002). Although few species have been added to the list since, the status of many turtles has nonetheless changed

from vulnerable to critically endangered ([www.iucnredlist.org](http://www.iucnredlist.org)). Reduced genetic diversity has been a criterion used by the IUCN to determine conservation priorities (Willoughby et al. 2015), and conservation genetics can provide crucial support towards management plans for endangered species (Alacs et al. 2007, Shaffer et al. 2015). However, life-history traits specific to turtles, including delayed sexual maturity, high mortality of juveniles, extended life spans, and overlapping generations, are affecting the power of genetic analyses (Kuo & Janzen 2004).

The wood turtle (*Glyptemys insculpta*) is a freshwater species endemic to eastern North America (Lat 38.00000 to 48.00000), with a patchy distribution in Canada. It is currently listed as endangered by the IUCN (van Dijk & Harding 2011), and is considered endangered or threatened in most of the US states and Canadian provinces where it occurs. Habitat fragmentation, destruction and alteration, collection for the pet trade, vehicle collisions, and increasing predation are the principal threats affecting this species (COSEWIC 2007). Wood turtles are found near rivers and streams with a moderate current, and the terrestrial habitats most commonly used consist of riparian areas with patchy forest cover (Arvisais et al. 2004). Wood turtles tend to move along streams, rarely more than 300 m away from the watercourse (Harding & Bloomer 1979, Arvisais et al. 2002). Maximum travel distances of up to 8.3 km in one year have been recorded (Daigle 1997), while some females have been known to travel as much as 2 km in one day during the nesting season (Walde et al. 2007). The wood turtle's longevity exceeds 50 years (Lovich et al. 1990), and generation time is estimated at 36-47 years (van Dijk & Harding 2011).

The genetic structure of wood turtle populations based on neutral markers was initiated by Tessier et al. (2005), and has since been the focus of several other studies based on small numbers of sampling sites and/or individuals (Amato et al. 2008, Castellano et al. 2009, Spradling et al. 2010, Fridgen et al. 2013, Willoughby et al. 2013). These studies revealed that recent fragmentation events and population decline slightly affected the genetic diversity and the genetic differentiation of the species. Additional data on endangered wood turtle populations are sorely needed to implement conservation plans, monitor head start programs and target populations at higher risk. Wood turtle local populations are currently considered

units for wood turtle recovery, and are defined as a discrete interbreeding population in a distinct watershed (Environment Canada 2016). There is no study to date, however, to support this definition.

The main objective of this paper was to assess the genetic structure of wood turtles from a northern part of their species range (Lat 45.00000 to 48.00000) to help further define meaningful units for conservation. We tested the effects of watersheds on their spatial genetic structure using clustering analyses based on neutral genetic markers, and an isolation-by-watershed model. We also assessed genetic diversity within each cluster. Finally, we discussed the implications of our results towards the conservation and management practices critically required for this highly endangered species.

### **3.4 Materials and methods**

#### **Sampling design**

Wood turtles have a discontinuous range in Canada from Nova Scotia through to New Brunswick, and are also found in southern Québec and Ontario (Ernst & Lovich 2009) where populations are considered to be at the northernmost limit of their range. We visited 24 sampling sites from 12 watersheds to help provide a thorough overview of the species distribution in Canada, including 22 in Québec, one in New Brunswick and one in Ontario (Figure 3.1). In Québec and New Brunswick, turtles were observed opportunistically or during annual surveys conducted since 2000. Captured individuals were marked as previously described (Cagle 1939), measured, weighed, aged, sexed (see Walde et al. 2003), and forelimb skin samples (1 mm × 3 mm) were collected for DNA analysis. Samples from three sites (SF1, GA3 and DC1) were the same individuals collected by Tessier et al. (2005) in Tomifobia, Forêt de l'Aigle, and Des Chênes, respectively. In Ontario, blood samples were collected in 2000 by government biologists in the Madawaska watershed (Algonquin Provincial Park).

## **DNA extraction and amplification**

Total genomic DNA was extracted using PureLink™ Genomic DNA extraction kits (Invitrogen) following the manufacturer's protocol. Genotypic profiles were obtained for nine microsatellite loci: GmuD16, GmuB21, GmuD28, GmuD40, GmuD55, GmuD87, GmuD88, GmuD93, and GmuD95 (King & Julian 2004). Polymerase chain reactions (PCR) were carried out using a Thermo cycler 9790 (Biosystems) in 18 µl volumes containing 25-100 ng of genomic DNA, 0.5 U (1 U ≈ 16.67 nkat) Taq DNA polymerase (Feldan Inc.), 2 mmol/L MgCl<sub>2</sub>, 2.5 mmol/L of each dNTP, and 0.16 µmol/L of each primer (forward primer 5' end labeled with a FAM or HEX fluorescent dye group), using the same cycling parameters as King and Julian (2004). Amplification success was determined by staining PCR products with SYBR Green nucleic acid gel stain (Invitrogen Inc., Burlington, Canada), running them on a 2% agarose gel, and revealing the bands under UV light. All PCR products were genotyped using an ABI-3730XL DNA analyzer with Liz\_500 fluorescent size standard (Thermo Fisher Scientific Inc. NYSE:TMO) at the McGill University Génome Québec Innovation Center. All alleles were manually scored three times using GeneMarker (SoftGenetics LLC, PA), and 15% percent of all individuals were independently amplified and genotyped twice to confirm observed genotypes. Samples displaying unexpected alleles (i.e. unique or rare alleles for the population or the species) were re-amplified and reanalyzed to confirm the genotypes.

## **Microsatellite analysis**

We used GENEPOP 4.0 (Raymond & Rousset 1995a, Rousset 2008) to test each locus for departures from Hardy-Weinberg equilibrium (HWE), and pairs of loci were tested for linkage disequilibrium (LD) using the Markov Chain Monte-Carlo (MCMC) method. Each locus was also analyzed for allelic stutter, large allele dropout and null alleles using MICROCHECKER 2.2 (Van Oosterhout et al. 2004). The probability of identity (PI), the probability of identity taking into account the possibility that close relatives were sampled (PISibs) (Taberlet & Luikart 1999, Waits et al. 2001), and the observed ( $H_O$ ) and expected heterozygosity ( $H_E$ ) for each locus were estimated with GenAlEx v 6.5 (Peakall & Smouse 2006, 2012). A sequential Bonferroni correction was applied to multiple pairwise comparisons (Rice 1989).

## Genetic structure

Two Bayesian clustering algorithms were used to explore the genetic structure of our sampled wood turtles. The relative probability of a series of models ranging from 1 to 15 populations ( $K$ ) was first assessed using a hierarchical approach with STRUCTURE v.2.3.4 (Pritchard et al. 2000). Length of burn-in period was set to 100 000, and the number of MCMC replicates to 1 100 000. Ten runs were performed for each  $K$  using the LOCPRIOR function to account for sampling information. All runs were compiled with STRUCTURE HARVESTER v 0.6.94 (Earl & vonHoldt 2012). The optimal number of clusters was determined by the log probability of the data for a distinct  $K$  (Pritchard et al. 2000), or by the delta  $K$ , which is based on the rate of change in the log probability of the data between successive  $K$  values (Evanno et al. 2005). To implement a hierarchical analysis of STRUCTURE, all individuals were used in a first run, and STRUCTURE was then run again, separately, on each cluster to see how it breaks down. This procedure was repeated iteratively until further structure was identified, or until large enough groups to support a STRUCTURE run were available (i.e. <20 individuals per putative cluster). In cases where a group of individuals were admixed between clusters, these individuals were run successively with each cluster.

TESS v.2.3.1 (Chen et al. 2007) was also used to explore clustering with respect to spatial information. An admixture model was assumed (Durand et al. 2009), and increasing values of  $K_{\max}$  (maximum number of groups in the dataset) ranging from 2 to 15 were considered, with 10 runs for each  $K_{\max}$ . Each run included 50,000 sweeps with a burn-in of 10,000 sweeps. Spatial coordinates were used to calculate great circle distances between sampling sites located on separate shores of the St. Lawrence River, including 129 dummy specimens (i.e., individuals with no genetic information) from the Québec database to increase the precision of the computations. The deviance information criterion (DIC) was applied to determine the most likely value of  $K_{\max}$  (Chen et al. 2007). CLUMPP v.1.1.2 (Jakobsson & Rosenberg 2007) was used to combine the results of the independent runs, and results were visualized with R (R Development Core Team 2016), library TESS3r (Caye et al. 2018).

## Isolation-by-watershed

Jost (2008) distances ( $D_{EST}$ ) were computed between sites with SMOGDv1.2.5 (Crawford 2010) using a square root transformation prior to computing a principal coordinate analysis (PCoA). Eigenvectors with positive eigenvalues were then included in a response matrix for a distance-based redundancy analysis (dbRDA, Legendre & Anderson 1999), computed with the `rda` function (R, `vegan` library) (Oksanen et al. 2017). To model watershed structures (Watershed Model), a series of dummy variables [0,1] were coded according to the classification of rivers by the Québec Government (Table 3.1). Two separate ecological processes were also accounted for in the analysis to dissociate the confounding effects of past and present events; the first one representing spatial autocorrelation between sites (River Distance Model), and the second representing a post-glacial expansion signal (Expansion Model). Distance-based Moran's eigenvector maps (dbMEM) (Borcard & Legendre 2002) were used to code for the River Distance Model using the distances between sites (in km) along the river network, whereas colonization events proposed by Amato et al. (2008) were used to code for the Expansion Model (Table 3.1). A dbRDA was first conducted based on the Watershed Model, followed by three partial redundancy analysis (pRDA), either controlling for the River Distance Model, the Expansion Model, or both models at the same time. Adjusted  $R^2$  ( $Ra^2$ ) and semi-partial  $Ra^2$  values (Legendre et al. 2011) were computed for each test to provide estimates of explained variation with respect to the total variation in our data. To further explore the proportion of variation explained by each model, variation partitioning of the genetic distance matrix according to the three distinct models was also performed (Peres-Neto et al. 2006). For all tests, significance was assessed using 1000 permutations (Anderson & Legendre 1999).

**Table 3.1** Categorization of all sites based on post-glacial Expansion Model constructs (Expansion) and the Watershed Model (Watershed). Sample site (ID) and the number of genotyped individuals (N) in each site is included, as well as the total number of alleles (A), number of private alleles (Pa), mean number of alleles per locus (k), allelic richness (Ar), private allelic richness (PAr), inbreeding coefficient ( $F_{IS}$ ), and observed ( $H_O$ ) and expected heterozygosity ( $H_E$ ) for each watershed group.

Expansion	Watershed	ID	N	A	Pa	k	Ar	PAr	$F_{IS}$	$H_O$	$H_E$
Western	Gatineau (GA)	GA1	3								
		GA2	9								
		GA3	31	88	4	9.78	7.12	0.25	0.078	0.677	0.728
		GA4	7								
		GA5	15								
	Madawaska (MD)	MD1	18	64	2	7.11	6.64	0.27	0.015	0.759	0.724
	L'Assomption (LA)	LA1	11	50	0	5.56	5.56	0.09	-0.006	0.739	0.701
	Loup (LO)	LO1	21	71	2	7.89	7.22	0.28	0.057	0.730	0.754
	St. Maurice (SM)	SM1	56	71	1	7.89	6.47	0.06	0.038	0.749	0.771
	Châteauguay (CH)	CH1	2	19	1	2.11	2.11	0.18	0.053	0.500	0.389
Yamaska (YA)	YA1	2	22	0	2.45	2.44	0.06	-0.238	0.722	0.472	
Eastern	Missisquoi (MI)	MI1	20								
		MI2	20	90	4	10	7.99	0.35	0.044	0.724	0.749
		MI3	6								
	St. François (SF)	SF1	6								
		SF2	1	80	4	8.89	7.83	0.4	0.045	0.754	0.772
		SF3	5								
		SF4	13								
	Bécancour (BE)	BE1	24	49	0	5.45	5.08	0.09	-0.059	0.685	0.634
	Du Chêne (DC)	DC1	17	77	2	8.56	8.14	0.31	0.076	0.703	0.736
	Matapédia (MA)	MA1	9								
MA2		3	82	5	9.11	7.45	0.42	0.071	0.743	0.757	
MA3		25									
MA4		7									

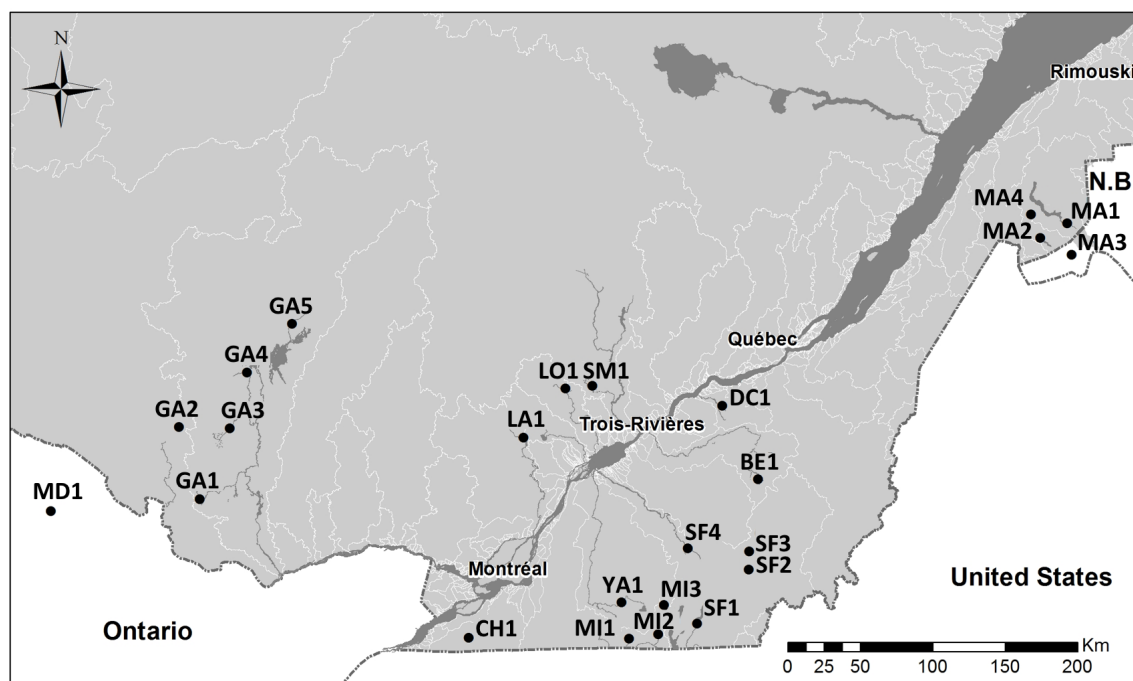


## **Within-watershed genetic diversity and differentiation**

Within each watershed, observed heterozygosity ( $H_O$ ), expected heterozygosity ( $H_E$ ) and the number of private alleles were estimated using GenAlEx v6.5 (Peakall & Smouse 2006, 2012). Allelic richness ( $A_r$ ) and private allelic richness ( $PA_r$ ) were adjusted to account for sample size using HP-RARE v.1.0 (Kalinowski 2004, 2005). Within watershed inbreeding coefficients ( $F_{IS}$ ) were computed using ARLEQUIN 3.0 (Excoffier et al. 2005). Significant differences in allele frequencies between watersheds were assessed with Fisher's exact tests using GENEPOP 4.0 (Raymond & Rousset 1995a, Rousset 2008).

## **3.5 Results**

A total of 331 turtles were sampled and genotyped from 24 different sites (Figure 3.1), with numbers of individuals per site ranging from 1 to 56 (Table 3.1). Individuals were sampled from twelve watersheds, eleven in Québec and New Brunswick and one in Ontario. Nine polymorphic loci were successfully amplified for each individual (Table 3.2), and the number of alleles per locus ranged from 4 to 33 (mean 15.89), for a total of 143 alleles.  $PI$  and  $PI_{sibs}$  decreased to  $< 0.01$  with inclusion of 2 loci and 5 loci respectively.  $H_O$  values ranged from 0.311 at locus GmuD55 to 0.850 at locus GmuD40 (mean = 0.720, standard error = 0.061). Of the 432 total pairwise comparisons, linkage disequilibrium was significant in eight pairs of loci from five different watersheds, but this was not consistent across loci. Moreover, linkage disequilibrium was not detected in previous studies using the same loci (Tessier et al. 2005, Castelanno et al. 2009, Spradling et al. 2010, Willoughby et al. 2013). All loci were thus considered independent for downstream analyses. Most loci were at HWE, except for GmuD95. Deviance from HWE may be explained by inbreeding depression, genetic drift or null alleles, as revealed by MICROCHECKER. No differences were observed when locus GmuD95 was omitted (results not shown) and it was thus retained in the data set for subsequent analyses. No scoring errors due to stuttering or evidence of large allele dropouts were found in the entire data set.

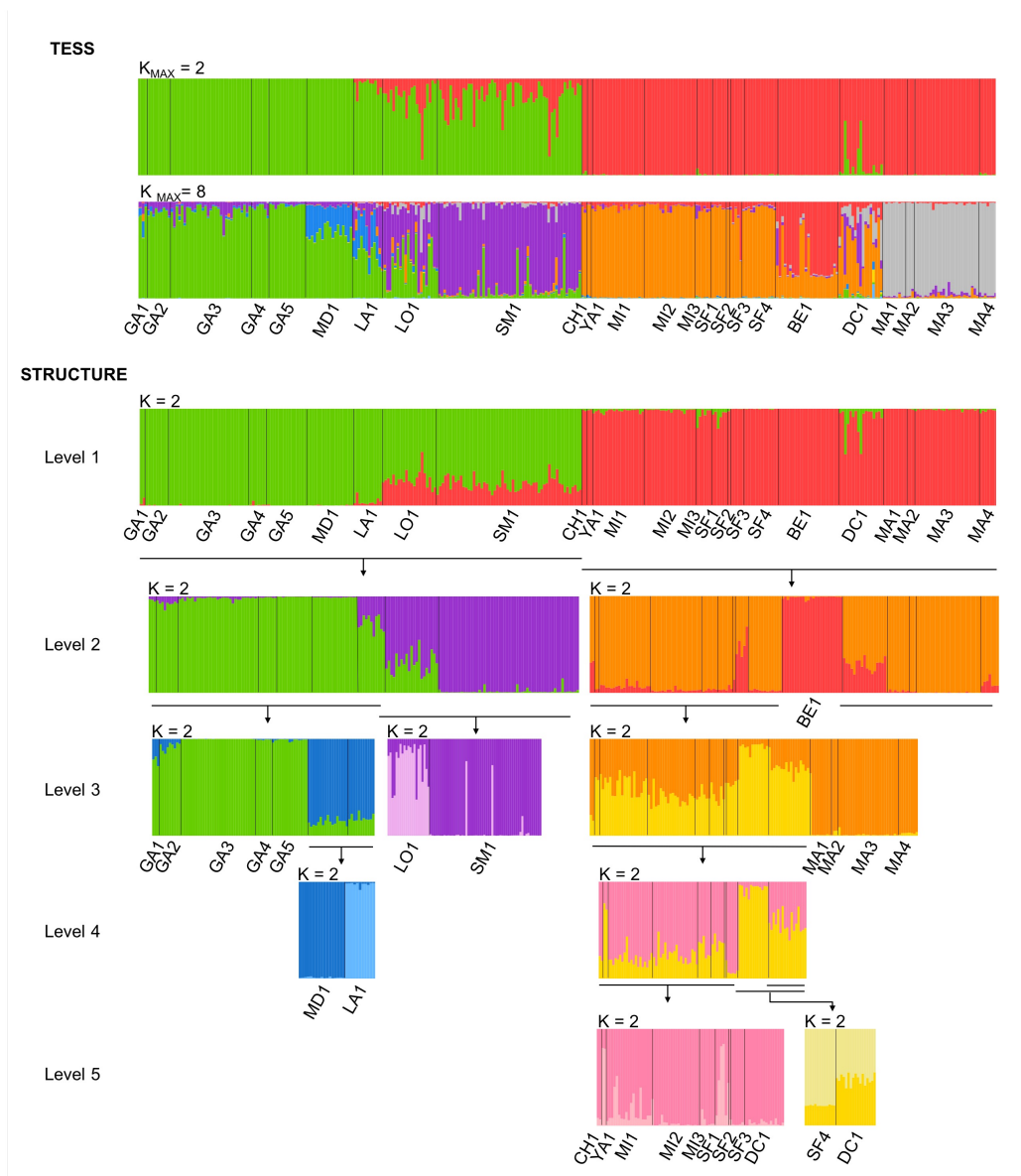


**Figure 3.1** Locations of 24 wood turtle sample sites, with watersheds outlined in white.

**Table 3.2** Summary statistics for nine microsatellite loci amplified in wood turtle samples, including the number of individuals successfully amplified at each locus (N). The number of alleles (k), observed ( $H_o$ ) and expected heterozygosity ( $H_e$ ), probability of identity (PI) and probability of identity for siblings (PIsibs) are also estimated for each locus.

Locus	N	k	$H_o$	$H_e$	PI	PIsibs
Gd16	330	16	0.840	0.903	0.017	0.303
Gd21	331	7	0.740	0.810	0.063	0.361
Gd28	306	15	0.786	0.857	0.035	0.330
Gd40	331	16	0.850	0.889	0.023	0.311
Gd55	326	4	0.311	0.540	0.310	0.557
Gd87	330	20	0.777	0.845	0.039	0.337
Gd88	331	17	0.847	0.911	0.015	0.298
Gd93	330	35	0.799	0.934	0.008	0.285
Gd95	326	13	0.530	0.842	0.043	0.340

Results of the hierarchical STRUCTURE runs and TESS analyses were quite similar (Figure 3.2). The mean DIC in TESS analysis gradually decreased, starting at  $K_{\max} = 2$ , with small inflections at  $K_{\max} = 8$  and 10, but no new cluster for  $K_{\max} > 8$ . For the hierarchical STRUCTURE analyses, five runs were required to split the data into distinct clusters. Delta K values (Evanno et al. 2005) revealed a peak at  $K = 2$  for every successive run, the first one splitting the sites into two clusters corresponding to the northern and southern shores of the St. Lawrence River. Following this split, individuals from the northern shore were analyzed separately, and three runs were used to subdivide all individuals into five clusters. For the southern shore, four additional runs were used to isolate at each level individuals sampled at a single site or watershed from all other individuals. Namely, the Bécancour watershed was isolated at level 2, whereas the Matapédia watershed was isolated at level 3. At level 4, genetic differentiation was observed between individuals from the SF4 site and the other individuals. However, individuals from Du Chêne watershed were admixed between the two clusters at this level, and they were thus run twice separately at level 5; first with individuals sampled from the SF4 site, where it showed different levels of admixture, and finally with the remnant individuals where no further differentiation was detected. Based on its high level of admixture, Du Chêne watershed was deemed different, and thus considered as a distinct cluster. For STRUCTURE and TESS, the five clusters observed on the north shore of the St. Lawrence River corresponded exactly to five watersheds from which wood turtles have been sampled: Gatineau, Madawaska, L'Assomption, Loup and St. Maurice. On the south shore, watershed divisions were not as obvious; some clusters contained more than one watershed (e.g. cluster of watersheds Châteauguay, Yamaska, Missisquoi, and site SF1-2-3) while others only contained one sample site (e.g. sample site SF4) from one watershed.



**Figure 3.2** Estimates of admixture proportions inferred by the most probable  $K_{max}$  values obtained from TESS and  $K$  values obtained from STRUCTURE using a hierarchical approach for 331 wood turtles sampled in Québec, Ontario and New Brunswick watersheds. Gatineau (GA), Madawaska (MD), L'Assomption (LA), Loup (LO), St. Maurice (SM), Châteauguay (CH), Yamaska (YA), Missisquoi (MI), St. François (SF), Bécancour (BE), Du Chêne (DC), Matapédia (MA). The five different levels represent successive runs of the STRUCTURE algorithm.

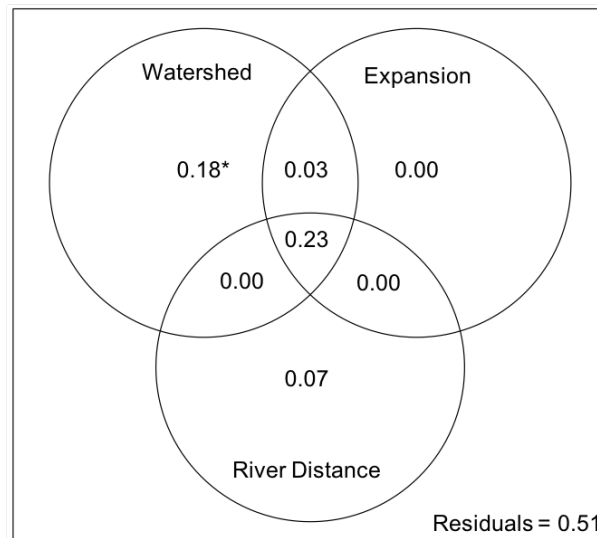
$D_{EST}$  values between sample sites ranged from 0.000012 (SF1-MI1) to 0.7654 (SF3-GA5). The PCoA of the corresponding distance matrix produced 20 positive eigenvalues, for a total inertia of 0.2713. Sample sites were attributed to twelve distinct watersheds (eleven in Québec, one in Ontario) (Table 3.1), and eleven dummy variables were thus used to code the Watershed Model. Only one dummy variable was used to code the Expansion Model, where sample sites were distributed into two groups. The distances among sites along the river network varied between 21.1 km (MI2-MI3) and 1347.3 km (MA3-GA5). The longest link of the minimum spanning tree was 405.8 km, which was set as the threshold value when truncating the distance matrix and creating the River Distance Model.

The pRDA used to test for watershed isolation while controlling for the other two models was significant (F-statistic = 1.6257,  $p = 0.003$ ) (Table 3.3). The variation partitioning showed the Watershed Model to explain a substantial part of the observed genetic differentiation (semi-partial  $Ra^2 = 0.1766$ ), compared to the Expansion Model and the River Distance Model (Figure 3.3). The River Distance Model accounted for only a small fraction of the genetic differentiation not explained by the other two models (semi-partial  $Ra^2 = 0.07$ ), and this result was not significant (F-statistic = 1.518,  $p = 0.12$ ). On the other hand, the Expansion Model uniquely explained none of the genetic variation (semi-partial  $Ra^2 = 0.00$ ). The intersection among the three models explained a substantial part of the variation (semi-partial  $Ra^2 = 0.23$ ).

**Table 3.3** Distance-based redundancy analysis (dbRDA) and partial redundancy analysis (pRDA) using genetic distance Jost ( $D_{EST}$ ) as the dependent variable, and three models as independent variables or conditions: the River Distance Model (RD) produced by spatial eigenfunctions of the truncated river distance matrix, the Expansion Model (EM) by ancestral post-glacial expansion hypothesis and the Watershed Model (WS) by a watershed isolation model. The constrained and conditional proportions of the variation, the F statistics and adjusted or semi-partial adjusted R-squares ( $Ra^2$ ) are shown for each model. Stars indicate significance levels of the F-statistics.

Model	Condition	Constrained	Conditional	F	$Ra^2$
Watershed (WS)	-	0.7101	NA	2.4491**	0.4201
	RD	0.42947	0.38577	1.6906**	0.203
	EM	0.42993	0.2874	1.6035**	0.1666
	RD + EM	0.37545	0.43979	1.6257**	0.1766

\*\*P < 0.01



**Figure 3.3** Venn diagram presenting the variation partitioning results of the genetic distance (Jost  $D_{EST}$ ) between sites, with respect to spatial eigenfunctions provided by the River Distance Model, the ancestral post-glacial Expansion Model and the Watershed Model. The values presented were computed from semi-partial adjusted R-squares of a pRDA analysis. The Venn diagram was produced using the varpart function (R, vegan library).

Within each watershed,  $H_O$  values were similar between most sites and ranged from 0.677 to 0.754 (Table 3.1), with the exception of site CH1. Several watersheds had private alleles, but most were located on the southern shore of the St. Lawrence River (Table 3.1). Allelic richness ( $A_r$ ) ranged from 2.11 to 8.14, and private allelic richness ( $PA_r$ ) ranged from 0.06 to 0.42. Inbreeding coefficients ( $F_{IS}$ ) for certain watersheds were negative, suggesting that they are less related than expected under HWE (e.g. watershed L'Assomption and Bécancour), while others appeared to be more related than expected under HWE (e.g. watershed Gatineau and Madawaska). Fisher's exact tests were highly significant for most watershed pairs, except between sites CH1 and YA1, whose low sample sizes affected the statistical power of this test ( $p = 0.623$ ). No differences were observed when sites CH1 and YA1 were omitted from the analyses (results not shown), and they were thus kept in the data set for subsequent analyses.

### **3.6 Discussion**

Our genetic analysis of the wood turtle is the first of its kind to incorporate a large number of sample sites over a relatively small geographical scale. We identified significant genetic structuring, largely due to the isolation of individuals by watersheds. Our sampling covered a significant portion of the species' range within Québec, but also included samples from both Ontario and New Brunswick. This range allowed us to increase the precision of our results, which would have otherwise been altered with a smaller number of sampled sites (Koen et al. 2013).

#### **Genetic structure**

Our results revealed important genetic structuring of the sampled individuals. Bayesian clustering analyses first showed a hierarchical separation into two clusters, consistent with previous observations of wood turtle populations in Québec. Tessier et al. (2005) discovered a dichotomous separation of the data set reflecting the northern and southern shores of the St. Lawrence River, which itself acted as a dispersal barrier. This division was also found to

correspond to two post-glacial colonization routes from a southeastern refuge (near the southern Appalachians), as revealed by mtDNA analysis (Amato et al. 2008). The maternal inheritance and lower effective population size of mtDNA both could affect the spatial structure and time to fixation of allelic variation. Therefore, the structure observed with microsatellite data is more likely to reflect an ancient barrier to dispersal created by the St. Lawrence River, rather than post-glacial colonization. Yet, certain individuals from sites located on the north shore (i.e. SM1, LO1, and LA1) exhibited an ancestral genetic signature similar to what was found on the south shore. This may be explained by an initial northeastern expansion, followed by a second western movement. A similar genetic signature was also observed in a recent study on the Canada lynx (*Lynx canadensis*), where the St. Lawrence River acted as a barrier to gene flow, which only becomes permeable when ice and snow cover allowed individuals to cross (Koen et al. 2015). However, based on the limited dispersal abilities of wood turtles (Daigle 1997), and the distance along the river network separating sites on both shores ( $\geq 200$  km between CH1 and LA1), unusual long-distance travel and the crossing of the St. Lawrence River are unlikely. Thus, we postulate that the movements between the two shores may also have been caused by anthropogenic displacement. It is well known that wood turtles are subjected to illegal trade (COSEWIC 2007), with 29 documented individuals in the Matapédia watershed alone exhibiting signs of captivity (Bertacchi, pers. comm.). Poaching is a serious threat to wood turtle survival (COSEWIC 2007, Environment Canada 2016). It may not only affect the number of reproductive adults within each site, but may also cause outbreeding depression if collected individuals are released outside of their native habitat (Lynch 1991).

The clustering analysis also revealed a finer genetic structure within each shore of the St. Lawrence River, which may reflect the life-history traits of wood turtles. This species demonstrates limited dispersal abilities ( $\leq 12$  km), homing behaviors of up to 8 km (Harding & Bloomer 1979), fidelity to nesting sites (Walde et al. 2007), and a stable home range (COSEWIC 2007), all factors which may increase isolation and foster genetic differentiation between watersheds (Reid et al. 2017). This phenomenon was observed in a study of three co-



occurring species of freshwater turtles, which demonstrated that populations of the terrestrial Blanding's turtle (*Emydoidea blandingii*) were highly isolated compared to the aquatic painted turtle (*Chrysemys picta*) and snapping turtle (*Chelydra serpentina*) (Reid et al. 2017). The large geographical extent of our study area may also have contributed towards the unique genetic divergence we were able to observe in this study (Manel et al. 2003, Schwartz & McKelvey 2009), as compared to previous studies on wood turtles. The longest river distance measured between our most distant sites (MD1 and MA3) was 1340 km, compared to 43.3 km in a study in Pennsylvania (Castellano et al. 2009) and 120 km in Michigan (Willoughby et al. 2013).

### **Impact of the landscape on genetic variation**

The hierarchical structure of a freshwater landscape offers multiple scales of connectivity for aquatic and semi-aquatic species. At a finer scale, the dendritic system will act as a network of dispersal corridors. Without any physical barriers, gene flow will thus occur naturally among sites located within the same watershed, and isolation by distance will prevail. At a larger scale, different watersheds may be isolated from one another, creating a genetic structure of populations. Population differentiation will then be greater for individuals sampled in different watershed than within the same watershed. Among-watershed genetic differentiation has been observed for several aquatic species, mainly fish (Beneteau et al. 2009, Eaton et al. 2018). For semi-aquatic species, including turtles, the watershed structure impacts differently the genetic structure according to the dispersal ability of the species under study (e.g. Souza et al. 2002, Mockford et al. 2005). Namely, landscape configuration appears to be less important for the genetic differentiation of less vagile species, compared to site-based factors. For example, the genetic structure of the Slimy Salamander (*Plethodon glutinosus*) is significantly driven by environmental variables (climate and vegetation) rather than isolation and spatial factors (Smith et al. 2018). Yet, our statistical results are consistent with the isolation-by-watershed hypothesis, and is the first study to confirm the effects of watershed divisions on wood turtle genetic structure. We also determined that the movements of individuals between and along streams were quite limited. The dispersal abilities, life-history traits, and habitat selection of

wood turtles thus favor the isolation of individuals within a river network, resulting in genetic structuring by watershed (Compton et al. 2002).

The River Distance and Expansion landscape models did not explain a significant portion of the genetic divergence observed. A small fraction of the partial model was associated with the River Distance Model, often used to model isolation-by-distance and relationships among sample sites on a smaller scale, however the association was not significant. River distances could explain some of the variation observed in sites located close to each other, more so than watershed divisions. For example, Willoughby et al. (2013) observed that sites separated by forested land and roads were not genetically distinct from each other, even though they were located in different watersheds. Migrants traveling between streams may have been responsible for this genetic similarity. Stochastic events, such as seasonal floods, may also displace wood turtles between habitats that are not usually connected by a river network (Jones & Sievert 2009), in turn influencing the genetic structure within and among watersheds.

The genetic structuring of sites located on the south shore of the St. Lawrence River may be explained by additional landscape features. For example, the SF4 site from the St. François watershed is highly divergent from all other sites within the same watershed (i.e. SF1, SF2, and SF3), even if the river distances between them are similar to those observed in other watersheds. Our landscape models, however, only considered watersheds, post-glacial colonization, and river distances. Incorporating altitude, vegetation cover, and habitat suitability may increase the explanatory power of our observations in the south shore (Marchand & Litvaitis 2004). These factors can altogether alter movements between sites, even for similar distances, since the type of vegetation cover can facilitate or impede the dispersion of individuals. Altitudinal differences between sites can also act as a natural barrier. Finally, anthropogenic factors may explain some of the observed differences by impacting population sizes, even when they are quite recent in the turtle landscape and invisible in the genetic signal. For example, the SF4 site is separated from other sampled localities by a large city, a network of highways (Gibbs & Shriver 2002), and agricultural practices (Daigle & Jutras 2005, Saumure

et al. 2007), all of which negatively impacting wood turtle movements and adult mortality rate (Environment Canada 2016).

In spite of the large sampling effort devoted to this study, difference in sample size among sites and watersheds could possibly affect the significance of our analyses. For one, unbalanced sampling designs can influence measures of genetic differentiation within and among watersheds (Balloux & Lugon-Moulin 2002, Ryman et al. 2006). Although we have run all analyses twice (i.e. with and without Châteauguay and Yamaska watersheds), and found no qualitative differences between results, the reduced sample sizes of other sites may still influence the measures of genetic differentiation. For example, all sites sampled within the St. François watershed had sample sizes ranging from 1 to 13 individuals, and this could explain the lack of genetic differentiation observed between SF1-2-3 and the Missisquoi, Yamaska, and Châteauguay watersheds. Considering the significant pRDA analysis based on the Watershed Model, the observed differentiation in the watershed with larger sampling size, and the ecology of wood turtles, it may be possible that every watershed should be divergent from others.

### **Genetic diversity of watersheds**

Genetic diversity was quite high within each watershed, an observation commonly associated with long-lived organisms (Pittman et al. 2011, Vargas-Ramirez et al. 2012). Genetic diversity may remain high, even for species where recent habitat degradation and fragmentation have occurred (Kuo & Janzen 2004, Cureton et al. 2014). The temporal lag observed between a phenomenon and its genetic consequence in long-lived organisms (Anderson et al. 2010) may be advantageous for the wood turtle by allowing more time for conservationists to implement management plans before a loss of diversity becomes a concern (Marsack & Swanson 2009).

In spite of different colonization processes operating in the north and south shores (see Amato et al. 2008), combined with anthropogenic disturbances causing important declines in populations in the south (Saumure & Bider 1998, Daigle & Jutras 2005), no significant differences in genetic diversity were observed between watersheds located on separate shores of the St. Lawrence River. Contrary to Tessier et al. (2005), our results exhibited lower genetic

diversity in watersheds undergoing higher isolation (Gatineau, Bécancour, and Du Chêne). Considering the long generation time of wood turtles, these lower levels of genetic diversity may be explained by a founder effect immediately following post-glacial colonization, rather than by more recent anthropogenic factors. Higher isolation between sites could also contribute to the reduced genetic diversity due to the absence of introduced novel variation via gene flow (Segelbacher et al. 2003).

## **Conservation and management**

Delimitations of a species conservation units below the species level are often required for their legal protection and the implementation of a sound management strategy. Evolutionary significant units (ESU) are most commonly used to account for reciprocal monophyly of mitochondrial DNA variation (Moritz 1994). On the other hand, management units (MU) are usually based on populations with limited gene flow and divergent allele frequencies (Moritz 1994, Palsbøll et al. 2007). The challenge of defining conservation units for turtle species was raised by Jensen et al. (2014), who argued that ESUs may not be suitable because the slow rate of mtDNA evolution may not reflect their adaptive capabilities, as suggested by Avise et al. (1992). The same concerns have already been raised for wood turtles by Amato et al. (2008), who showed that mtDNA lineages were not distinct enough to be applied to conservation units. At a finer scale, MUs are generally preferable for conservation purposes, although different methods may produce different results within the same data (Palsbøll et al. 2007). Our estimates of  $F_{ST}$  and Fisher's exact tests between pairs of samples sites (results not shown) suggest that each site should be identified as a distinct MU. However, as pointed out by Jensen et al. (2014), independently managing these numerous MUs would prove challenging.

In Canada, designatable units (DU) are identified by COSEWIC to define distinct populations or subspecies of wildlife that require protection under the Species at Risk Act (SARA). It is a more flexible concept, based on the joint criteria of a population discreteness and evolutionary significance. Our study is a first step towards recognizing DUs for wood turtles in Canada as the discreteness criterion (COSEWIC 2015) may be addressed directly by our results; the evidence of genetic distinctiveness by neutral genetic markers. Based on the

significant isolation-by-watershed model, as well as the results of the clustering analysis we have identified a total of ten units, including nine in Québec (all sites from watersheds Gatineau, L'Assomption, Loup, St. Maurice, Bécancour, Du Chêne, and Matapédia, site SF4, and finally the watershed Missisquoi including sites SF1, SF2, and SF3) and one in Ontario (Madawaska). All individuals from New Brunswick are considered to be part of the Matapédia watershed. A larger sample size would, however, be required to distinguish Châteauguay and Yamaska watersheds from other DUs, and to better understand the relationships between sites SF1-2-3, SF4, and the Missisquoi watershed. For conservation purposes, we thus recommend the further consideration of these ten units to determine if they meet the evolutionary significance criterion to become DUs.

In spite of the genetic similarities between some watersheds (e.g. Missisquoi and St. François), we believe that local populations should still be considered unique within each watershed (Environment Canada 2016). Even if this implies splitting genetic clusters for management purposes, it will allow for a conservative approach to avoid genetic incompatibility, which may cause outbreeding depression (Meffe & Carroll 1997). The isolation of individuals within the river network, resulting in a significant genetic structuring by watershed, is of great importance because watershed protection is already independently managed by local non-governmental organizations. Wood turtle conservation may then easily fit within their boundaries.

While recognizing that genetic studies are an essential part of conservation plans for endangered species (Alacs et al. 2007, Shaffer et al. 2015), we also recognize that such analyses may alter the process in cases where urgent actions are needed (DeSalle & Amato 2004). We believe that our comprehensive genetic analysis of the wood turtle can be readily implemented into their management. We recommend that offspring from nearby nesting sites in the same genetic cluster could be used for head-start programs. Similarly, translocation strategies to manage populations at risk (Storfer 1999) can be applied as well. Finally, watershed markers may be key factors to use towards identifying with high assignment success the origin of poached individuals.

### 3.7 Acknowledgments

We would like to thank numerous individuals for providing wood turtle samples: W. Bertacchi, G. Bourget, M. Bruno, C. Daigle, Y. Dubois, A. Dumont, M. Dumont, S. Gagnon, P. Gaudet, S. Giguère, C. Greaves, J. Jutras, P. Labont., M. Laflèche, M. Toner, M. Leclerc, D. Masse, S. Paradis, Y. Robitaille, D. Rodrigue, K. Smith, D. St-Hilaire, and C. Trochue. We are also grateful to several biologists, technicians, and volunteers from the Québec Government, Parks Canada, Corporation Gestion de la Forêt de l'Aigle, Corporation de l'Aménagement de la Rivière l'Assomption, Ecomuseum, and Université de Montréal for helping with field sampling. We would also like to thank G. Bourget, Y. Dubois, S. Giguère, D. Masse, and S. Pelletier for providing additional details of wood turtle locations, L. Veilleux for producing the map, and A. Rogic for editing earlier drafts of this manuscript. Finally, we are thankful for the thoughtful comments and suggestions provided by two anonymous reviewers. This study was supported in part by a Natural Sciences and Engineering Research Council grant (No. 0155251, F.-J. Lapointe) Fondation de la faune du Québec, Parks Canada, Fédération canadienne de la faune, and Faune-Nature Québec program.

**COMPLIANCE WITH ETHICAL STANDARDS** The authors declare no conflicts of interest.

# **Chapitre 4. Les réseaux dynamiques pour étudier la connectivité des populations chez la tortue des bois**

**Applying novel connectivity networks on wood turtle populations to provide comprehensive conservation management strategies for species at risk**

Cindy Bouchard<sup>1\*</sup>, Étienne Lord<sup>1</sup>, Nathalie Tessier<sup>2</sup>, François-Joseph Lapointe<sup>1</sup>

<sup>1</sup> Département de sciences biologiques, Université de Montréal, Montréal, QC, Canada

<sup>2</sup> Ministère des Forêts, de la Faune et des Parcs, Longueuil, QC, Canada

*Article en préparation*

## 4.1 Résumé

La diversité génétique intra et interpopulation est fréquemment utilisée lors du processus de priorisation des populations afin de les classer en fonction de leur vulnérabilité ou de leur importance. Néanmoins, la connectivité et le flux génique sont rarement considérés dans les cadres conceptuels de priorisation. En utilisant comme exemple un réseau construit à l'aide de données génétiques pour la tortue des bois (*Glyptemys insculpta*), nous utilisons le logiciel BRIDES pour évaluer la connectivité du réseau dans son ensemble plutôt que d'étudier les nœuds individuellement. BRIDES permet de comparer les plus courts chemins entre une paire de nœuds d'un sous-graphe et la même paire de nœuds dans le réseau complet. Les objectifs principaux de cette étude consistaient à (1) présenter un nouveau processus de sélection de nœuds à l'aide du logiciel BRIDES pour assister les biologistes œuvrant en conservation dans le processus de priorisation et à (2) utiliser les indices des réseaux pour comparer les résultats de BRIDES et estimer le flux de gènes entre les populations de tortues des bois. Nous avons construit six sous-graphes et estimé le nombre optimal de nœuds à ajouter pour maximiser la connectivité du réseau à l'aide d'un algorithme de sélection pas à pas. Nos résultats montrent une grande robustesse du processus de sélection implanté pour un scénario donné, alors que des incohérences furent observées dans le classement fait à l'aide des indices de réseaux. Certaines populations furent sélectionnées à de multiples reprises lors de l'analyse des divers scénarios. L'importance de celles-ci n'aurait pu être prédite par les résultats de la diversité et de la différenciation génétique, les indices de centralité et les analyses d'élimination de nœuds. Contrairement aux indices de centralité qui évaluent des réseaux statiques, BRIDES permet l'analyse de réseaux dynamiques. À notre connaissance, il s'agit de la première étude utilisant la théorie des graphes pour l'analyse de la génétique des populations de tortues. Notre étude montre que les réseaux peuvent traiter des dynamiques de flux de gènes complexes impliquées dans la résilience des populations à l'extinction locale. BRIDES offre un complément intéressant aux indices de centralité pour comprendre les processus globaux et ainsi mieux aborder la crise mondiale de la perte de biodiversité.



## 4.2 Abstract

Genetic diversity within and among populations is frequently used in prioritization processes to rank populations based on their vulnerability or distinctiveness, however, connectivity and gene flow are rarely considered within these frameworks. Using a wood turtle (*Glyptemys insculpta*) population graph, we introduce BRIDES as a new tool to evaluate populations for conservation purpose without focusing solely on individual nodes. BRIDES characterizes different types of shortest paths among the nodes of a subgraph and compares the shortest paths among the same nodes in a complete network. The main objectives of this study were to (1) introduce a BRIDES selection process to assist conservation biologists in the prioritization of populations, and (2) use centrality indices and node removal statistics to compare BRIDES results and assess gene flow among wood turtle populations. We constructed six population subgraphs and used a stepwise selection algorithm to choose the optimal number of additional nodes required to maximize network connectivity under different weighting schemes. Our results demonstrate the robustness of the BRIDES selection process for a given scenario, while inconsistencies were observed among node-based metrics. Results showed multiple selection of certain wood turtle populations, which could have not been predicted following the genetic diversity and distinctiveness estimation, node-based metrics and node removal analysis. Contrary to centrality measures focusing on static networks, BRIDES allowed for the analysis of evolving networks. To our knowledge, this study is the first to apply graph theory for turtle conservation genetics. We show that population graphs can reveal complex gene flow dynamics and population resiliency to local extinction. As such, BRIDES offers an interesting complement to node-based metrics and node removal to better understand the global processes at play when addressing population prioritization frameworks.

**Key words** : BRIDES software, connectivity, conservation genetics, gene flow, *Glyptemys insculpta*, graph theory, population prioritization, Wood turtles.

### 4.3 Introduction

The conservation of units below the species level, commonly populations, is widely used in wildlife management, and provides critical information to support species survival and evidence of within-species differences. Populations show important local adaptive variations that are essential for species adaptation to changing environments (Soulé 1987, Moritz 1994, Fitzpatrick & Keller 2015). Furthermore, they exhibit greater sensitivity to extirpation, which may inform of the general trend at the species level (Hughes et al. 1997). The threats experienced by populations require distinct management practices that must take into account the economic, social and cultural aspects of the society living in closest proximity to them (Mace & Lande 1991, Wood & Gross 2008, Razgour et al. 2017). Rule-based methods have been developed by the International Union for Conservation of Nature (IUCN) to categorize populations based on the level of threat they face (Millsap et al. 1990, Master 1991, IUCN 1994). This multiple criteria system has been used to categorize species at risk and aid-governmental and non-governmental organizations to define conservation units, including Management Units (MUs, Moritz et al. 1994, Palsbøll et al. 2007), Evolutionary Significant Units (ESUs, Moritz et al. 1994) and Designable Units (DUs). The latter is used by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) to define populations or subspecies of wildlife that require protection under the Species at Risk Act (SC 2002, c 29).

Considering that two populations may be ranked in the same category even when faced with entirely different threats, and the limited financial resources available for conservation biology, prioritizing populations arose as an essential concept to target the most effective conservation actions. Multiple prioritization frameworks have been developed to provide objective criteria during decision-making processes, and either estimate species extinction risks (Mace & Lande 1991) or prioritize threatened populations and determine the management actions required to support them (Given & Norton 1993). Prioritization frameworks are often reactive, targeting highly vulnerable populations in need of immediate management actions to avoid reducing genetic diversity and local or global species extinction events (Brooks et al. 2006). However, for species with long generation times, it may be beneficial to react in a

proactive fashion and prioritize populations that exhibit low vulnerability to protect their global genetic diversity as a reservoir (Crandall et al. 2000, Moritz 2002, Mace & Purvis 2008).

Population viability analysis (PVA) is frequently used to evaluate extinction risks given its comprehensive and quantitative basis (Coulson et al. 2001). To complement such models, or in cases where data are unsuitable, the relative importance of each population can also be estimated to measure the impact of local extinction on species survival. A series of binary questions have been proposed by Allendorf et al. (1997) for the Pacific salmon (*Oncorhynchus* spp.) and adapted for all freshwater fish by Clarkson et al. (2012) to score the biological consequence of local population extinctions for the species as a whole. Although these rankings are intended to represent objective criteria, the list of questions may overlook important features that deserve conservation attention.

Measures of genetic diversity within and among populations are frequently applied to prioritize populations and define conservation units based on past bottleneck events, effective population size and population distinctiveness (Petit et al. 1998, Ottewill et al. 2016). To bridge the gap between researchers and practitioners, Ottewill et al. (2016) developed a simple framework using population differentiation, genetic diversity and inbreeding coefficients. The quantification and ranking of these parameters can be of great importance to estimate population isolation, predict their trends and target the most informative genetic processes for management purposes. Although many of these estimators allow us to understand the distinctiveness and diversity of a population to predict its persistence, only few consider the dynamic structure underlying the population network. Some demographic studies have accounted for population connectivity by using a metapopulation model in their prioritization framework. However, the data requirements of complex metapopulation models greatly limit their application in conservation studies, requiring a short-term decision-making time frame. For example, the incidence function model of metapopulation dynamics (Hanski 1994) requires estimates of patch colonization and extinction rates that are obtained through longitudinal surveys of all available habitat patches (Moilanen et al. 1998). Yet, connectivity among isolated habitat patches is essential to several ecological processes and should be

considered in population prioritization models. Seasonal migration, dispersal of offspring, recolonization following local extinction events, range shifts caused by climate change, and gene flow for the transmission of favourable alleles may all be distorted if change occurs within the connectivity pattern (Hoban 2018, Razgour 2018).

Gene flow is frequently cited as an evolutionary force that maintains connectivity among populations (Crooks & Sanjayan 2006, Mills 2007), but connectivity in itself can potentially operate at different time scales other than demographic processes, such as dispersion and recolonization events (Lamy et al. 2012, Creech et al. 2014). As such, gene flow should also be considered in population prioritization as it offers complementary insights into population connectivity (Lamy et al. 2012, Gubili et al. 2017). Theoretical and empirical studies support the use of gene flow as a measure of functional connectivity (Hedgecock et al. 2007, Lowe & Allendorf 2010, Creech et al. 2014, Row et al. 2018), and graph theory is a valuable tool emphasizing the relevance of connectivity in conservation planning (Kool et al. 2013). The processes underlying the structure of complex networks have been efficiently characterized in many disciplines (Newman 2003) and are increasingly applied in ecology and conservation where food webs, metapopulation dynamics or reserve networks can be modelled (Matos et al. 2019). More precisely, population graphs have been developed by Dyer and Nason (2004) as a network depicting populations (the nodes of the graph) connected by their genetic covariance (the edges of the graph). Gene flow has since been modelled for many plant and animal species in marine, freshwater and terrestrial ecosystems using networks (see Garroway et al. 2008, Treml et al. 2008, Marrotte et al. 2017). For management purposes, population prioritization can be facilitated by the ranking of several node-based centrality indices (Jacoby & Freeman 2016), or by using node removal to assess their relative impact on network connectivity (Creech et al. 2014). Population prioritization has also been studied with phylogenetic networks (Volkman et al. 2014), for which genetic distinctiveness was used to score each population. Yet, the choice of relevant network metrics is crucial for the prioritization process using such methods (Jordán et al. 2003, Laita et al. 2011, Cross et al. 2018).

The wood turtle (*Glyptemys insculpta*) is a threatened freshwater turtle species and listed as endangered, according to the IUCN red list, following an overall decrease in population trends (van Dijk & Harding 2011). It is a semi-terrestrial, long-lived species with longevity that may exceed 50 years (Lovich et al. 1990), delayed sexual maturity at 11-22 years (Brooks et al. 1992, Walde et al. 2003), and a long generation time estimated at 36-47 years (van Dijk & Harding 2011). Canadian wood turtle populations are protected under the Species at Risk Act, but financial support is lacking to ensure the persistence of all populations. In cases where wildlife managers have to decide which populations to protect, a sound prioritization framework is required. Yet, considering the cost and short duration of radiotelemetry studies, the temporal and spatial variation in nest predation, and the difficulty of monitoring hatchlings and juveniles, PVA is not easily applicable to wood turtle populations (Jones et al. 2015). Population genetics have been previously used to characterize the population structure of wood turtles (Tessier et al. 2005, Castellano et al. 2009, Fridgen et al. 2013, Willoughby et al. 2013) and the importance of peripheral populations for genetic diversity (Spradling et al. 2010). In addition, landscape genetics has revealed the importance of the watershed structure for wood turtle population differentiation in Canada (Bouchard et al. 2019). None of these previous studies, however, have relied on graph theory as a statistical tool to analyze connectivity and prioritize populations.

Although many studies highlight the importance of connectivity and gene flow for the maintenance of diversity and the persistence of populations in a fragmented landscape, few frameworks consider connectivity as an important feature of the prioritization ranking. In this paper, we introduce the use of BRIDES (Lord et al. 2016) as a new tool to evaluate whole population graphs for conservation purposes, rather than focusing solely on individual patches. Thus, the main objectives of our study were (1) to demonstrate the use of the BRIDES selection algorithm to assist conservation biologists and managers during decision-making processes, and (2) to use graph theory to assess gene flow among wood turtle populations. To do this, we used a dataset of 19 wood turtle populations collected in eastern Canada and genotyped by Bouchard et al. (2019). Centrality indices and node removal statistics were first

estimated from a population graph. We then constructed six population subgraphs based on three different criteria to protect populations in a proactive or reactive fashion, and using BRIDES, selected the optimal number of additional nodes required to maximize network connectivity under different weighting schemes. The results of the three methods were then applied within the perspective of wood turtle conservation.

## **4.4 Materials and methods**

### **Wood turtle population characterization**

The original data set from Bouchard et al. (2019) included 24 sites, but five sites were excluded due to insufficient sample size ( $N < 6$ ). As a result, 327 wood turtles sampled at 19 different sites and genotyped for nine microsatellite loci as described in Bouchard et al. (2019) were used in the present study. Using microsatellite data, expected heterozygosity ( $H_E$ ) was estimated using the *diveRsity* package with R version 3.5.1 (Keenan et al. 2013). Element occurrence (EO) ranks were used to estimate the persistence probability of a population for a defined period of time (20-100 years), while considering that current environmental conditions prevailed during that time (NatureServe 2002). These EO ranks vary from excellent viability (A) to poor viability (D), and some rank combinations can be used in cases of uncertain estimates. All EO ranks and estimated population sizes were provided by wildlife biologists at the Québec Ministry of Forests, Wildlife and Parks.

### **Wood turtle population graph**

We constructed a population graph based on Dyer and Nason's (2004) conditional genetic distance (cGD) and pruning method using R v. 3.5.1 (R core team 2018) *popgraph* (Dyer 2014) and *gstudio* (Dyer 2016) packages. In this network, nodes representing different sample sites are connected to each other by undirected edges weighted by the genetic distances between them. However, as some edges do not adequately describe overall among-population distances,

the network was pruned by removing redundant edges that did not contribute to the overall genetic covariance of the network, as described by Dyer & Nason (2004).

### **Node-based metrics of connectivity and node removal**

To evaluate the importance of each node with respect to network connectivity, we used two current methods: node-based metrics (Koen et al. 2016, Cross et al. 2017) and the node removal approach (Garroway et al. 2008, Zero et al. 2017). Several node metrics were computed to estimate their contribution to gene flow and importance in graph connectivity. Degree centrality (DC) is defined as the number of edges connected to a node, whereas betweenness centrality (BC) is the number of shortest paths upon which the node lies. Based on immediate neighbour connections, the eigenvector centrality (EC) may also provide information on both direct and indirect levels of node connectivity. The clustering coefficient (CC) is the probability that two nodes connected to a neighboring node are also connected together. Computing strength as the sum of all edge weights connected to a given node is another centrality measure of interest for conservation biology, since the average inverse edge weight (AIEW) is positively correlated with the number of migrants to and from a focal node (Koen et al. 2016).

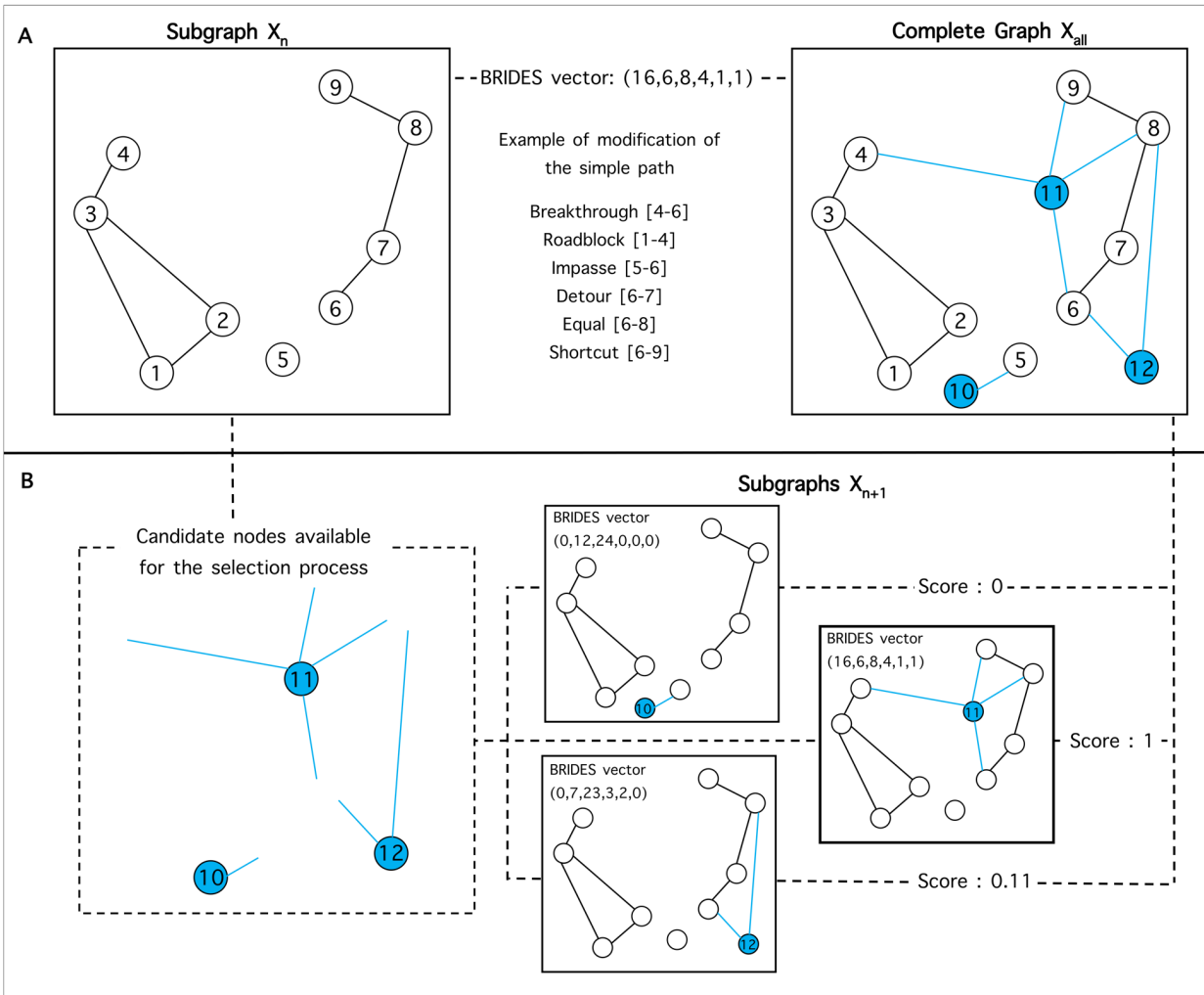
The node removal approach was also used to estimate the effect of local extirpation on the connectivity of the network and to obtain an overview of the resilience of the network to node loss (Garroway et al. 2008). In this study, we simulated the effect of local extirpation by removing a node and its associated edges from the graph and measuring the average path length in the resulting graph. The process was repeated iteratively by removing each node, one at a time. Differences in average path lengths were estimated by comparing the reduced and complete networks, with larger differences being associated with more important nodes for network connectivity. A cut node is a node whose deletion, along with incident edges, results in a disconnected graph. Node-based metrics and node removal statistics were used to estimate gene flow between wood turtle populations and were compared with results of the BRIDES selection procedure described below.

## BRIDES analysis

Network connectivity was explored using BRIDES v1.2 (Lord et al. 2016), a software that allows the characterization of dynamic networks through the comparison of shortest paths among pairs of nodes of a subgraph with their respective paths among the same nodes in a complete graph (Figure 4.1A). Both the complete graph and the subgraph are defined by the users and selected as input networks for the BRIDES algorithm. When applied to population genetic data, the paths identified by BRIDES may thus represent actual gene flow between populations, or past gene flow between remnant populations or metapopulations relative to neutral genetic markers.

Precisely, BRIDES is a polynomial-time algorithm used to characterize different types of paths between a subgraph  $X_n$  and a complete graph  $X_{all}$ , where all nodes in  $X_n$  are a subset of the nodes in  $X_{all}$ . The shortest paths between pairs of nodes in  $X_n$  are first computed. Then, the lengths of the paths among the same pairs of nodes are recomputed in  $X_{all}$  by forcing these paths to include at least one node present in the complete graph  $X_{all}$  but absent in the subgraph  $X_n$ . The comparison of these corresponding path lengths thus allows six types of paths to be distinguished (Figure 4.1A): a *breakthrough* (B) is a path that is impossible in  $X_n$  but possible in  $X_{all}$  (path 4-6); a *roadblock* (R) is a path that is possible in  $X_n$  but impossible in  $X_{all}$  (path 1-4); an *impasse* (I) is a path that is impossible in both  $X_n$  and  $X_{all}$  (path 5-6); a *detour* (D) is a path that is shorter in  $X_n$  than in  $X_{all}$  (path 6-7); an *equal* (E) path has the same length in  $X_n$  and  $X_{all}$  (path 6-8); a *shortcut* (S) is a path that is longer in  $X_n$  than in  $X_{all}$  (path 6-9). The results are summarized in a vector [B, R, I, D, E, S] containing the numbers of each type of paths estimated by the algorithm (see Figure 4.1A).





**Figure 4.1** A) Example of the six types of paths computed in BRIDES by comparing the shortest path between pairs of nodes in subgraph  $X_n$  (open circles) and the shortest path between the same pairs of nodes in the complete graph  $X_{all}$  including all candidate nodes (blue circles). In  $X_{all}$ , the shortest path between any pair of nodes from  $X_n$  must include at least one of these candidate nodes. B) Each of the three nodes available for the selection process was iteratively added to  $X_n$  to create the subgraphs  $X_{n+1}$  and corresponding scores were computed with respect to the complete graph  $X_{all}$ . The selection process selected node 11 as the best candidate (score = 1) to maximize network connectivity while minimizing the number of additional nodes. The weighted model used for the selection process was (1, 0, 0, 0, 1, 1).

## BRIDES selection algorithm

We implemented an iterative selection algorithm in BRIDES to determine which candidate node should be added to  $X_n$  to maximize the connectivity of  $X_{n+1}$  (Figure 4.1B). Given that each type of path may have a different impact on the choice of candidate nodes, distinct weighted models could be defined by the user for the selection algorithm. For example, one may consider only *breakthroughs*, *equals* and *shortcuts* to maximize network connectivity. In that case, the corresponding model would be (1, 0, 0, 0, 1, 1). At each step of the algorithm, a selection score is then computed as the sum of weighted numbers of path types in the BRIDES vector. The maximum score that can be achieved is estimated by including all candidate nodes in the subgraph  $X_n$ . For comparison purposes, this maximal score is set to one, and suboptimal solutions are scaled by dividing their scores by the maximum value (see example in Figure 4.1B).

## Weighted models and conversation scenarios

In the present study, five distinct weighting schemes were used for the six types of paths (B, R, I, D, E, S) in order to select optimal solutions under different scenarios, with positive weights attributed to preferred types of paths, negative weights attributed to rejected types of paths, and null weights (zero) attributed to types of paths ignored in the final score computation. The first weighted model (0, 0, 0, 0, 0, 1) only considered *shortcuts*, whereas the second (3, 0, 0, 0, 0, 1) and third (1, 0, 0, 0, 0, 3) models also accounted for *shortcuts* and *breakthroughs*, but with distinct weights. The fourth model (1, -1, 0, -1, 0, 1) considered the negative impact of *roadblocks* and *detours* with respect to *breakthroughs* and *shortcuts*. For the fifth model (3, 0, 0, 1, 2, 3), we used an ecological interpretation of BRIDES in which *breakthroughs*, *detours*, *equals* and *shortcuts* were all considered to increase the connectivity of the network, but with different weights.

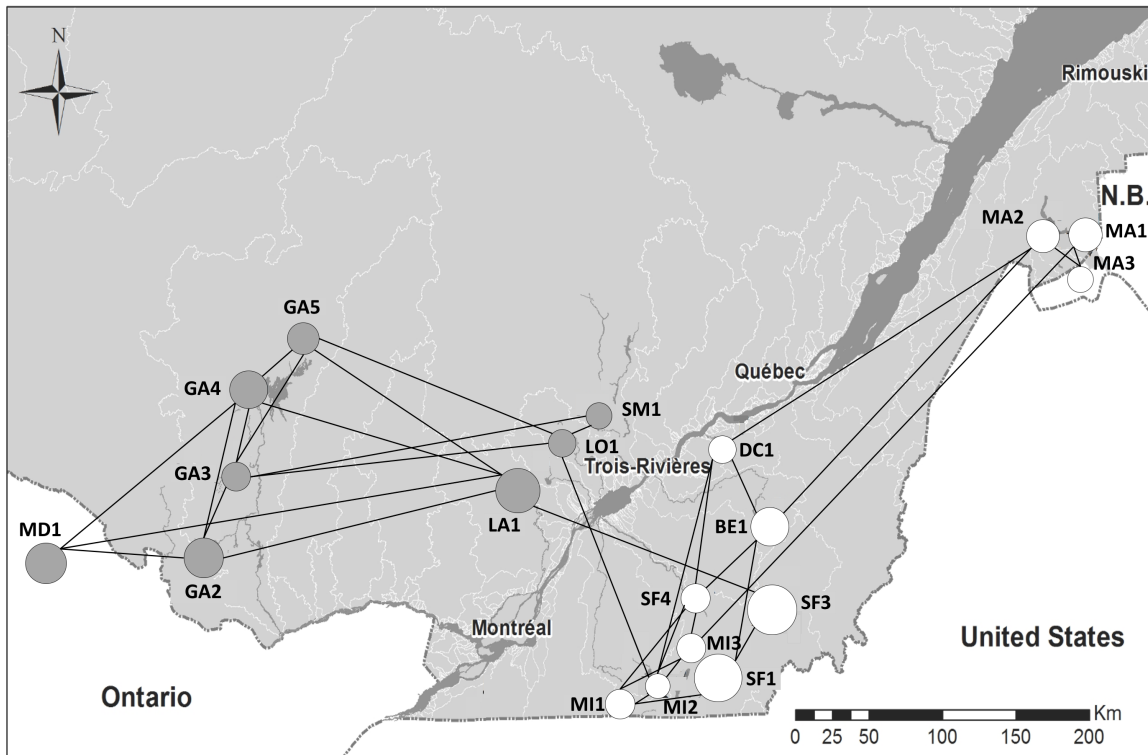
In addition to the five weighting schemes, six scenarios encompassing three different parameters important for management purposes were considered: genetic diversity ( $H_E$ ), estimated population size, and persistence probability (EO ranks). These scenarios were used

to generate different subgraphs  $X_n$  onto which candidate nodes were added in turn to construct subgraphs  $X_{n+1}$ . For each parameter, pairs of networks (subgraphs) were generated by selecting nodes with either the largest or smallest values, representing nodes that prioritize a proactive (P) or a reactive (R) conservation perspective, respectively. For the first pair of networks, populations with higher (Scenario A<sub>P</sub>) and lower (Scenario A<sub>R</sub>) heterozygosity values ( $H_E$ ) were selected. For the next pair of networks, nodes with higher (Scenario B<sub>P</sub>) and lower (Scenario B<sub>R</sub>) estimates of population sizes were selected. Finally, EO ranks were used to select nodes for the fifth (Scenario C<sub>P</sub>) and sixth (Scenario C<sub>R</sub>) networks based on their persistence probability for a period of time. The five weighting models were then applied to the six scenarios to examine their impact on the node selection process.

## 4.5 Results

### Population graph

The wood turtle population graph constructed with cGD resulted in 19 nodes and 37 edges (19% of the edges in a saturated network) after the pruning procedure. Eight nodes were located on the north shore and eleven nodes on the south shore of the St. Lawrence River (Figure 4.2). Sample sizes and estimated population sizes of different sites ranged respectively from 6 to 56 and 6 to 2000 individuals per node (Table 4.1).



**Figure 4.2** Population graph of 19 wood turtle populations based on conditional genetic distance (cGD). Redundant edges were pruned when they did not contribute to the overall genetic covariance structure. Circular node sizes are scaled according to the average inverse edge weight (AIEW), where greater AIEW values are reflected through larger circular diameters. Populations located on the northern and southern shores of the St. Lawrence River are identified with gray and white nodes, respectively.

### Network metrics and node removal

The population graph was analyzed with different node-based metrics to determine which sites contributed most to network connectivity (Table 4.1). One node (GA3) ranked top in all metrics, whereas two nodes ranked at the very bottom (MA1 and MA2). Above all, these results illustrated a clear separation between sites located on the northern and southern shores of the St. Lawrence River, and the importance of the two edges connecting them. Several nodes with high betweenness centrality values may act as stepping-stones or bridges (Garroway et al.

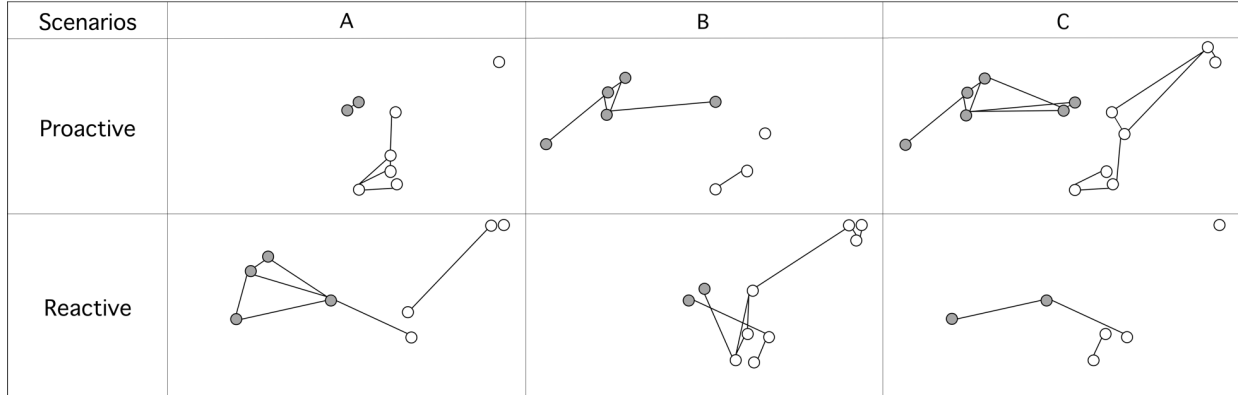
2008). Namely, four of the nodes with large betweenness values (MI2, LO1, LA1, SF3) connected the two shores together. In contrast, nodes with a high-clustering coefficient may act as an anchor of interconnected groups of nodes. The maximum value of the clustering coefficient was obtained for two nodes on the north shore (SM1 and MD1), whereas four nodes on the south shore had null values. To further identify nodes that contributed to both direct (immediate connection) and indirect (connection to its neighbors) network connectivity, node eigenvector centrality was computed. Once again, our results show opposite patterns for nodes located on distinct shores, with those on the northern shore exhibiting much larger eigenvector centrality values compared to nodes in the south. Finally, AIEW was calculated to provide estimates of the numbers of migrants to and from a node. Nodes with higher values ( $AIEW > 0.065$ ) were all located on the south shore of the St. Lawrence River (Table 4.1). Average path length of the complete network was 2.708. While the removal of a single node (MI2) increased the value up to 3.353 (Table 4.1), the removal of six individual nodes (MA3, MD1, GA2, GA4, MA24, SM1) provided results close to that of the complete graph (varying between 2.595 and 2.706), meaning less contributions to network connectivity. There were no cut nodes in our population graph.

**Table 4.1** Sampled population parameters and corresponding node-based metrics, node removal statistics and BRIDES results. Site identification (Id), number of genotyped individuals (N), expected heterozygosity ( $H_E$ ), estimated population size (Psize), element occurrence rank (EO), betweenness centrality (BC), degree centrality (DC), eigenvector centrality (EC), strength, average inverse edge weight (AIEW), clustering coefficient (CC), average path length (APL) estimated after node removal, and selection rate of each node using the stepwise selection procedure (Selection). Values in bold indicate the top three populations within each of the metrics, node removal and the selection process.

Shore	Id	N	$H_E$	Psize	EO	BC	DC	EC	Strength	AIEW	CC	APL	Selection
North	<b>GA2</b>	12	0.65	75	BC	3.483	4	<b>0.85</b>	82.002	0.049	<b>0.67</b>	2.673	0.21
	<b>GA3</b>	31	0.72	2000	A	18.3	5	0.55	74.951	0.068	0.4	2.778	0.22
	<b>GA4</b>	7	0.63	250	B	5.383	5	<b>0.94</b>	<b>98.256</b>	0.051	0.6	2.68	0.00
	<b>GA5</b>	15	0.62	2000	AB	12.22	4	0.58	64.178	0.064	0.5	2.745	0.07
	<b>LA1</b>	11	0.701	50	BC	25.05	5	<b>1</b>	<b>114.731</b>	0.044	0.4	2.948	0.43
	<b>LO1</b>	21	0.754	50	B	<b>47.18</b>	4	0.24	56.732	0.071	0.33	<b>3.281</b>	0.50
	<b>MD1</b>	18	0.724	100	B	0	3	0.76	64.692	0.047	<b>1</b>	2.627	0.00
	<b>SM1</b>	56	0.771	250	AB	0	2	0.13	26.862	<b>0.074</b>	<b>1</b>	2.68	0.00
South	<b>BE1</b>	24	0.634	100	B	22.6	5	0.09	<b>96.79</b>	0.053	0.2	2.843	<b>1.00</b>
	<b>DC1</b>	17	0.736	50	AB	13.23	4	0.05	55.792	0.073	0.5	2.765	0.42
	<b>MA1</b>	9	0.68	20	BC	9.4	3	0.04	51.131	0.06	0	2.706	<b>0.93</b>
	<b>MA2</b>	10	0.625	30	AB	7.5	3	0.03	46.923	0.065	0.33	2.68	0.31
	<b>MA3</b>	25	0.75	25	AB	0.7	2	0.01	26.783	<b>0.075</b>	0	2.595	0.17
	<b>MI1</b>	20	0.73	500	AB	8.45	4	0.08	60.582	0.07	0.5	2.778	0.08
	<b>MI2</b>	20	0.71	30	BC	<b>49.63</b>	5	0.08	64.001	<b>0.079</b>	0.4	<b>3.353</b>	0.78
	<b>MI3</b>	6	0.74	100	B	12	4	0.05	59.394	0.068	0.5	2.77	0.23
	<b>SF1</b>	6	0.74	25	AB	<b>27.82</b>	3	0.17	72.665	0.041	0	<b>2.987</b>	<b>1.00</b>
	<b>SF3</b>	6	0.61	6	C	23.82	2	0.38	50.147	0.04	0	2.889	0.36
<b>SF4</b>	13	0.73	25	BC	5.233	5	0.06	73.463	0.069	0.5	2.732	0.62	

## BRIDES selection procedure

Using the BRIDES selection process, we explored which nodes were the most important for network connectivity using six subgraphs that were constructed from the complete population graph including 19 wood turtle populations following different conservation scenarios. Depending on the criteria selected, these networks contained between 6 and 13 nodes (Figure 4.3). Networks based on genetic heterozygosity included populations with  $H_E$  values greater than 0.725 (Scenario A<sub>P</sub>) or lower than 0.71 (Scenario A<sub>R</sub>). Those based on estimated population sizes were created by including populations with more (Scenario B<sub>P</sub>) or less than (Scenario B<sub>R</sub>) 75 individuals. Networks based on element occurrence included populations with EO ranks greater than BC (Scenario C<sub>P</sub>) or lower than B (Scenario C<sub>R</sub>) (Table 4.1). The five weighting models were then used to examine their impact on the node selection procedure. The identity of the selected nodes was very similar (90%) regardless of the weighted model selected for a given scenario (Figure 4.S1, Table 4.S1).



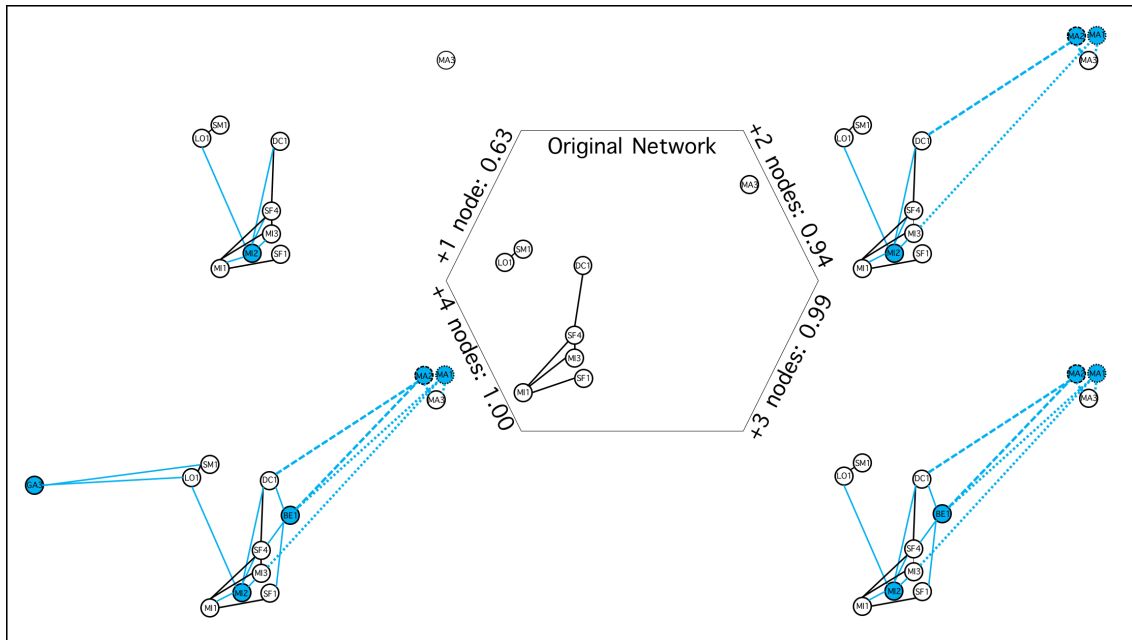
**Figure 4.3** Different population subgraphs based on distinct node selection parameters using a proactive or reactive conservation approach. A: Node selection based on heterozygosity ( $H_E$ ) values. B: Node selection based on estimated population size. C: Node selection based on element occurrence (EO) values. For each scenario, nodes included in the subgraph represented populations with the largest (Proactive) or smallest (Reactive) parameter values. Populations located on the northern and southern shores of the St. Lawrence River are identified with gray and white nodes, respectively.

Results of the node selection process revealed that regardless of the weighting schemes and conservation scenario, maximum scores were obtained by adding from one to six nodes to the subgraph  $X_n$  (Table 4.2). A positive relationship was observed between the number of candidate nodes required to reach the maximum score and the complexity of the weighting scheme, with higher complexity models allowing for more diverse path types, but also requiring more candidate nodes (Kendall Tau = 0.634,  $p < 0.001$ ) (see Figure 4.S1 for examples). Namely, for the first weighting scheme that only accounted for *shortcuts* (complexity of 1), three out of six scenarios returned a null solution as no *shortcuts* were possible by adding candidate nodes to the subgraph. In three other cases, one or two candidate nodes (mean = 1.33) were required to reach a maximal score. Models 2 and 3 considered *shortcuts* as well as *breakthroughs* (complexity of 2), but with different weighting schemes, and the number of nodes required to reach the maximum score varied from one to three (mean = 2.00). Model 4 contained contrasting *roadblocks* and *detours* with respect to *breakthroughs* and *shortcuts* (complexity of 4), and two to five candidate nodes (mean = 3.33) were required to reach a maximal score. Finally, Model 5 accounted for *breakthroughs*, *detours*, *equals*, and *shortcuts* (complexity of 4), and three to six nodes (mean = 4.5) were required to reach a maximal score. In some cases, multiple solutions with equal scores were available at any given step of the procedure. To represent these equally multiple solutions, we depicted alternative paths with different types of edges in corresponding graphs (Figure 4.4). This redundancy was observed in the results from five of the six subgraph  $X_n$  selections for at least one step of the node selection procedure (Table 4.2).



**Table 4.2** Summary of the BRIDES node selection procedure providing the minimum number of nodes required to reach a maximal score, and the number of multiple solutions with the same score for different scenarios and weighted model schemes. A: Node selection based on heterozygosity ( $H_E$ ) values. B: Node selection based on estimated population size. C: Node selection based on element occurrence (EO) values. For each scenario, nodes included in the subgraph represented populations with the largest (Proactive) or smallest (Reactive) parameter values.

Scenario	$A_P$					$A_R$					$B_P$					$B_R$					$C_P$					$C_R$														
Weighted Model	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5					
No. of nodes	1	3	3	3	4	-	1	1	3	6	-	3	3	5	5	1	1	1	2	3	2	2	2	5	5	-	2	2	2	4	1	2	2	2	2	-	1	1	1	1
No. of solutions	1	2	2	2	2	-	1	1	1	4	-	5	5	5	4	1	1	1	1	1	2	1	1	2	2	-	1	1	1	1	1	1	1	1	1	-	1	1	1	1



**Figure 4.4** BRIDES node selection process for scenario  $A_P$  and weighting scheme 5 (3, 0, 0, 1, 2, 3). All selection steps are shown, even when multiple solutions had identical scores. In such cases, distinct solutions (i.e. MA1 or MA2) are identified by dashed edges connecting the selected nodes. Maximum score (1.00) was reached with the addition of four candidate nodes. At each step of the procedure, the selected nodes and their corresponding links are presented in blue.

## 4.6 Discussion

Our study shows that population graphs provide clarity on complex gene flow dynamics and population resiliency to local extinction events. We also show that the new selection algorithm for population prioritization that we implemented in BRIDES produces consistent results when selecting populations, regardless of the weighting schemes considered. As such, it offers an interesting alternative to node-based metrics and provides an understanding of the global processes at play when addressing biodiversity connectivity.

### Properties of the BRIDES selection procedure

BRIDES can rapidly assess the impact of multiple conservation scenarios without considering *a priori* models of population dynamics. The selection process thus allows the user to test different scenarios and weighting schemes to decide which one is best for population prioritization. When applied jointly with PopGraph, BRIDES will select the minimum number of nodes required to optimize gene flow and network connectivity. Moreover, BRIDES can identify multiple solutions with identical scores, thus allowing for alternative node combinations in the prioritization process. Its node selection process is robust for a given subgraph. In addition, BRIDES is highly flexible and provides the user with the opportunity to determine which types of paths need to be accounted for, and how they should be weighted to compute scores.

The BRIDES selection algorithm ranks all nodes based on their contribution to path lengths in the subgraph  $X_n$  with respect to the complete network  $X_{all}$  (Figure 4.1). Based on hypothetical conservation scenarios, our results highlighted the relative importance of some candidate nodes in the complete network, when their addition increased the numbers of *shortcuts* or *breakthroughs*. However, node selection can vary depending on the nodes chosen for the construction of subgraph  $X_n$ . When divergent scenarios are considered, BRIDES provides an objective criterion to guide the decision-making process. Yet, considering the high variation observed in the ranking given the subgraph used, the prioritization process should account for sensitivity to missing data and subgraphs must be chosen carefully, based on questions of

interest. Our results highlight the need to further evaluate that potential source of uncertainty in network analysis, especially when applied to conservation biology where the impact of missing nodes on path lengths and component connectivity has yet to be assessed (Koen et al. 2013, Naujokaitis-Lewis et al. 2013).

Important ranking information is largely provided by node-based metrics, as nodes can be compared across a multitude of indices operating at different scales and patterns of connectivity. Cross et al. (2018) argue that this flexibility can be beneficial when top-ranking nodes are first targeted by conservation actions. Yet, as pointed out by Creech et al. (2014), the choice of a biologically relevant metric to quantify connectivity has a direct effect on the prioritization rankings, and our results revealed such inconsistencies among node-based metrics (Table 4.1). Similarly, node removal is highly sensitive, and it may produce inconsistent outcomes in habitat connectivity analysis (Jordán et al. 2003, Laita et al. 2011). In agreement with previous studies, we trust that network metrics are valuable tools to understand complex patterns of connectivity and gene flow. On the other hand, we believe that node-based metrics should not be applied as the only prioritization criteria. Contrary to centrality measures focusing on static networks, BRIDES allows the analysis of dynamic networks onto which candidate nodes and edges are added. As shown in our results, the nodes more frequently selected by BRIDES were not always consistent with those selected by node-based metrics and node removal (Table 4.1). Furthermore, BRIDES identifies six different types of paths, each one capturing specific properties of path lengths within evolving networks. For example, *breakthroughs* and *shortcuts* are indicative of increased connectivity in a complete network and measure distinct processes, making them crucial for conservation planning. In contrast, *impasses* and *roadblocks* are detrimental to network connectivity and are to be avoided in the prioritization process. Redundancy in the network can also be considered with *equal* paths and *detours* representing alternative routes for dispersal and gene flow. Yet, as it targets nodes based on their connectivity, BRIDES might miss some important peripheral or isolated nodes. For example, isolated nodes in a subgraph may remain disconnected when their selection leads to more *detours* and *roadblocks* than the selection of a central node (see Figure 4.S1, Scenario C<sub>R</sub>).

For that reason, BRIDES may be used jointly with complementary tools, such as phylogenetic networks developed by Volkman et al. (2014) to account for genetic distinctiveness and identify peripheral populations.

## **Prioritization process of wood turtle populations**

In the context of wood turtle conservation genetics, the results we obtained with population graphs are consistent with previous landscape genetics analyses (Bouchard et al. 2019), however, relationships among populations are much more detailed. As opposed to pairwise methods, population graphs account for genetic covariance among all populations simultaneously (Dyer & Nason 2004, Dyer et al. 2010). It is already known that the St. Lawrence River acts as a barrier to gene flow among wood turtle populations located on opposite shores (Tessier et al. 2005, Amato et al. 2008, Bouchard et al. 2019), but the population graph revealed which populations and corresponding links are more connected across the river (Figure 4.2). The network also depicted a detailed account of the interaction patterns between populations located on the two shores. We observed four nodes linking the two shores, which could be viewed as stepping-stones or bridges between populations (Garroway et al. 2008). The results obtained after node removal also support the interconnection among populations on the north shore, except for two nodes linking the southern shore that affect average path lengths. This implies that gene flow among north shore populations makes them less susceptible to local extinctions. On the other hand, populations on the south shore are more loosely connected, except for some sites exhibiting high values of AIEW. Metapopulation dynamics should thus be further investigated for these sites since AIEW is an accurate measure of reproductive success and the number of migrants between nodes (Koen et al. 2016).

The BRIDES algorithm was applied in this study to demonstrate how it may be used to select wood turtle populations with either a reactive or proactive prioritization procedure. To do so, we considered six hypothetical scenarios and five different weighting schemes, and applied a node selection procedure to prioritize populations. For example, the first scenario

only included populations with large heterozygosity values. Even when neutral genetic diversity is not correlated to functional diversity (Hedrick 2001), heterozygosity may still identify well-connected sites, large population sizes despite isolation, habitat fragmentation, and perturbation. In a proactive conservation plan, such populations with low vulnerability may be selected first to protect the genetic variability of the species before it erodes (Crandall et al. 2000, Moritz 2002, Mace & Purvis 2008).

Regardless of their genetic diversity, the topological relationships among nodes within the network may be of importance for connectivity (Dyer 2015). Peripheral populations associated with MA sites are thus of great interest to quantifying functional connectivity. Although MA3 presents high neutral genetic diversity, all MA sites are clustered together in Bouchard et al. (2019). Likewise, node-based metrics for these populations were in the low to mid-range, except for MA3, which exhibited a large AIEW value. Removal of these three nodes had no impact on the average path length of the network. These results are in stark contrast with those obtained with BRIDES, where the MA1 was very likely to be selected to maximize network connectivity. Such a discrepancy illustrates that node-based metrics may overlook important nodal features that BRIDES actually considers during its selection process.

In Québec, riverbeds and riverbanks which are part of wood turtle habitats are protected in public land, but stronger management actions are still needed to ensure long-term survival of populations (Équipe de rétablissement 2005). To do so, one can apply BRIDES to prioritize populations in a reactive fashion. That is that populations with the smallest sizes – those needing high priority management actions such as nest protection from predators or a head-starting program – would first be selected to construct a subgraph (Scenario B<sub>R</sub>, Figure 4.3, Figure 4.S1). As the resulting subgraph may be disconnected and these small populations may face extirpation, one could also adopt a proactive perspective to protect larger populations that support them. More precisely, large turtle populations, as a source of genetically similar individuals, may add redundancy and resilience to population extirpation in the network. Namely, BRIDES can be applied to select these candidate populations based on path lengths and other connectivity indices (Weighted model 5, Figure 4.S1). In this particular case,

populations BE1, MI3 and GA5 (Figure 4.S1, weighted model 5) would be selected by BRIDES to guide the complex decision-making process of population prioritization.

Subgraph constructions in our study were based on proactive and reactive procedures to demonstrate the method, but it is entirely up to the user to select populations in the subgraph. The subgraphs could also be based on characteristics independent of the species, such as land ownership or jurisdictions involved allowing for nesting sites or for hibernacula that can be easily protected and managed. Because the choice of populations in the subgraph greatly influences the subsequent selection of candidate nodes, we suggest to make these choices based on practical conservation criteria, especially when BRIDES is applied for population prioritization.

Uncertainties about genetic connectivity, genetic differentiation, isolation by distance, and population assignment may all be addressed with a single heuristic approach: the network. Moreover, many levels of connectivity can be considered at a different time and spatial scales when using evolving networks, especially those analyzed with BRIDES. Network analysis also offers an alternative model to population genetic analysis when populations are not considered to be at a mutation-drift balance and depart from the Hardy-Weinberg equilibrium (Neigel 1997, Rozenfel et al. 2008, Dyer 2015). Yet, networks are not free of *a priori* assumptions that are associated with the choice of a genetic distance, the pruning procedure, and the criteria selected for the construction of the subgraph. In the present case, subgraphs were created by removing nodes and their corresponding edges from the complete population graph among 19 wood turtle populations. Considering that the graph pruning method is based on the removal of redundant edges that do not contribute to the overall genetic covariance of the network, population graphs should be re-estimated every time a node is removed.

Given the specific, historic life traits of long-lived organisms with low recruitment and late sexual maturity, the genetic characterization of recent perturbation events can be challenging (Kuo & Janzen 2004). As network topology stabilizes rapidly following a perturbation event, network analysis could be part of the solution to detect more precisely the

genetic signal (Dyer 2007, Dyer et al. 2010, Noutsos et al. 2014). BRIDES represents an interesting avenue to study trophic networks or to compare various scenarios during the design of protected area networks. Managers could benefit from using such tools to understand the impact of the addition or removal of a species within a food web, whether it be through the impact of the introduction of an invasive exotic species or the selection of a biocontrol agent. Namely, a scenario with fewer *breakthroughs* or *shortcuts* could be preferred to limit the impact of introduced species on native species. Then, BRIDES could also be applied to optimize the connectivity among isolated habitat patches, with respect to historical levels of connectivity, when protected areas are designed. In a time where changes are accelerating and greater emphasis is given to temporal scales when modeling climate change, population graphs and evolving networks may provide a better understanding of the changing ecosystem, trophic networks and population structure of species.

## 4.7 Acknowledgments

We would like to thank Philippe Lamarre from ministère des Forêts, de la Faune et des Parcs for information and explanation about populations size estimation and element occurrence ranks of the wood turtles in Quebec. Improvements of earlier drafts of the manuscript are the results of comments from C. Davy, S. Joly, P. Legendre and A. Rogic. This study was supported in part by a Natural Sciences and Engineering Research Council grant to F.-J. Lapointe (No. 0155251) and NSERC scholarship to C. Bouchard.

**COMPLIANCE WITH ETHICAL STANDARDS** The authors declare no conflicts of interest.





**Table 4.S1** Results of the BRIDES node selection procedure for six scenarios (A, B, C) and five weighted models (1 to 5). For each node, the presence of a population in the subgraph (x) and its selection (1) or not (0) in at least one of the optimal solution are indicated. Scenarios for which no nodes were selected (-) are also shown.

Scenarios		AP					AR					BP					BR					CP					CR									
Weighted model		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Shore	<b>Node Id</b>																																			
	<b>GA2</b>	0	0	0	0	0	x	x	x	x	x	-	0	0	1	1	0	0	0	0	0	0	0	0	1	1	x	x	x	x	x	-	0	0	0	1
	<b>GA3</b>	0	0	0	0	1	-	0	0	1	1	x	x	x	x	x	0	0	0	0	0	x	x	x	x	x	-	0	0	0	1	-	0	0	0	1
	<b>GA4</b>	0	0	0	0	0	x	x	x	x	x	x	x	x	x	x	0	0	0	0	0	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>GA5</b>	0	0	0	0	0	x	x	x	x	x	x	x	x	x	x	0	0	0	0	1	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>LA1</b>	0	0	0	0	0	x	x	x	x	x	-	1	1	1	1	x	x	x	x	x	0	0	0	1	1	x	x	x	x	x	-	0	0	0	1
	<b>LO1</b>	x	x	x	x	x	-	0	0	0	1	-	1	1	1	1	x	x	x	x	x	x	x	x	x	x	-	0	0	0	1	-	0	0	0	1
	<b>MD1</b>	0	0	0	0	0	-	0	0	0	0	x	x	x	x	x	0	0	0	0	0	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>SM1</b>	x	x	x	x	x	-	0	0	0	0	x	x	x	x	x	0	0	0	0	0	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
South	<b>BE1</b>	1	1	1	1	1	x	x	x	x	x	x	x	x	x	x	1	1	1	1	1	x	x	x	x	x	-	1	1	1	1	-	1	1	1	1
	<b>DC1</b>	x	x	x	x	x	-	0	0	0	1	-	1	1	1	1	x	x	x	x	x	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>MA1</b>	0	1	1	1	1	x	x	x	x	x	-	1	1	1	1	x	x	x	x	x	1	1	1	1	1	x	x	x	x	x	-	0	0	0	0
	<b>MA2</b>	0	1	1	1	1	x	x	x	x	x	-	0	0	0	0	x	x	x	x	x	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>MA3</b>	x	x	x	x	x	-	0	0	1	1	-	0	0	0	0	x	x	x	x	x	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>MI1</b>	x	x	x	x	x	-	0	0	0	1	x	x	x	x	x	0	0	0	0	0	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>MI2</b>	0	1	1	1	1	-	0	0	0	1	-	1	1	1	1	x	x	x	x	x	1	1	1	1	1	x	x	x	x	x	-	0	0	0	0
	<b>MI3</b>	x	x	x	x	x	-	0	0	0	1	x	x	x	x	x	0	0	0	1	1	x	x	x	x	x	-	0	0	0	0	-	0	0	0	0
	<b>SF1</b>	x	x	x	x	x	-	1	1	1	1	-	1	1	1	1	x	x	x	x	x	x	x	x	x	x	-	1	1	1	1	-	1	1	1	1
	<b>SF3</b>	0	0	0	0	0	x	x	x	x	x	-	1	1	1	0	x	x	x	x	x	0	0	0	1	1	x	x	x	x	x	-	0	0	0	0
<b>SF4</b>	x	x	x	x	x	-	0	0	0	1	-	1	1	1	1	x	x	x	x	x	1	0	0	1	1	x	x	x	x	x	-	0	0	0	0	

## Chapitre 5. Discussion générale

Cette thèse avait comme objectif principal d'utiliser la génétique au service de la conservation de la tortue des bois. À cet effet, j'ai étudié la génétique de cette espèce à plusieurs échelles spatiales et temporelles afin de mieux cerner les processus qui affectent la diversité génétique intra et interpopulation. En plus d'une synthèse des principaux résultats de ma thèse, je présente ici des exemples d'actions concrètes qui ont été guidées par mes travaux.

### 5.1 La polyandrie et le succès reproducteur des mâles

Les analyses de parentalité de la tortue des bois ont permis une meilleure compréhension des stratégies de reproduction de cette espèce et du succès reproducteur des mâles de la population de Shawinigan (**Chapitre 2**). La contribution des mâles à la reproduction est très variable et se trouve grandement influencée par la paternité multiple au sein d'une nichée et la paternité répétée entre les nichées consécutives d'une femelle. Selon un consensus de trois méthodes d'estimation, la paternité multiple est observée dans au moins 37% des nichées de tortues des bois à l'étude. La paternité répétée, observée pour la première fois chez cette espèce, fut détectée dans les nichées de sept des huit femelles dont les nids avaient été échantillonnés lors des deux saisons d'échantillonnage.

Deux hypothèses sont présentées pour expliquer la présence de ces stratégies de reproduction chez les femelles de la tortue des bois. La première stipule qu'il existe des événements de copulation multiple et répétée, tels qu'observés chez les femelles de la population à l'étude (Walde et al. 2003). Il est ainsi possible qu'une femelle se reproduise avec plusieurs partenaires au long de la saison de reproduction et ces partenaires pourraient être les mêmes d'une année à l'autre. La deuxième hypothèse repose sur l'emménagement de sperme, ce qui permet la désynchronisation entre les événements de copulation et de fertilisation des

œufs (Birkhead & Møller 1998). Le sperme est emmagasiné dans les oviductes de la femelle lors de la copulation et pourrait, selon l'espèce, être utilisé pour la fécondation jusqu'à 4 ans suivant celle-ci (Gist & Jones 1989). Ce phénomène est cependant difficile à établir avec certitude chez la tortue des bois uniquement avec les résultats du **Chapitre 2**. Néanmoins, cette hypothèse est soutenue en raison de la faible densité de la population à l'étude (0.44 adulte/ha), ce qui diminuerait la probabilité de rencontres entre partenaires sexuels (Gist & Congdon 1998). Cependant, puisque les événements de reproduction sont plus fréquents lors de l'agrégation d'individus près des sites d'hibernation (Brooks & Brown 1992, Kaufmann 1992, Walde et al. 2003), il est possible que mes résultats révèlent plutôt l'existence de groupes distincts de mâles et de femelles interconnectés. Une étude portant spécifiquement sur la fidélité aux sites d'hibernation et le comportement reproducteur des mâles à ces sites apporterait des détails supplémentaires sur la paternité multiple et répétée.

Selon mes résultats, les impacts de la stratégie de reproduction des tortues des bois sur la diversité génétique des populations se divisent en deux catégories. Tout d'abord, une relation directe sera observée en présence de paternité multiple dans une nichée. Des études portant sur les reptiles proposent diverses hypothèses pour expliquer les avantages de la polyandrie pour les femelles d'espèces n'offrant pas de soins parentaux (Jennions & Petrie 2000, Uller & Olsson 2008). Une des plus intéressantes, confirmée par mes résultats, est l'augmentation de la diversité génétique de la nichée (Yasui 1998). Cette augmentation permettrait d'améliorer les probabilités de survie de la nichée dans un environnement imprévisible et hétérogène. Or, il est important de noter que mes résultats sont basés sur des marqueurs génétiques neutres qui ne représentent pas nécessairement l'ensemble de la diversité génomique des jeunes (Holderegger et al. 2006). Indirectement, la variation du succès reproducteur des mâles, en modifiant la taille efficace de la population, aura aussi des conséquences sur le maintien de la diversité génétique intrapopulation (Holman & Kokko 2013). Mes résultats montrent une grande variation dans le succès des mâles pour les deux années à l'étude où seuls quelques mâles, les supergénérateurs, produisent une majorité des jeunes. Cette tendance, observée sur plusieurs générations, pourrait avoir un effet négatif sur la taille efficace de la population

(Hedrick 2005, Karl 2008), contrairement à la paternité multiple qui a généralement un effet plutôt positif (Sugg & Chesser 1994, Pearse & Anderson 2009, Lotterhos 2011).

## 5.2 L'isolement par bassins versants et l'expansion des populations

Les analyses de génétique du paysage (**Chapitre 3**) ont permis d'étudier la relation entre la configuration du paysage de la tortue des bois et la variation génétique interpopulations. Des détails supplémentaires sur la connectivité et le flux de gènes entre populations et leur résistance à l'extinction locale furent précisés par les analyses de réseaux du **Chapitre 4**.

Une barrière au flux génique fut observée dans les analyses de groupement bayésien (**Chapitre 3**) et les analyses de réseaux (**Chapitre 4**) qui divisent les populations de tortues en deux groupes distincts correspondant à chacune des deux rives du fleuve Saint-Laurent. Cette dichotomie fut également observée lors d'études précédentes sur la tortue des bois (Tessier et al. 2005, Amato et al. 2008) et pourrait, selon mes résultats, refléter une barrière à la dispersion des individus. L'hypothèse postulée pour expliquer la structure observée est un mouvement d'expansion des individus vers le nord, suivi par un second mouvement vers l'ouest, expliquant la relation entre quatre populations (LA1-SF3 et LO1-MI2) observée lors des analyses de réseaux (**Chapitre 4**). De ces sites, les individus de la rive nord se seraient par la suite déplacés vers l'ouest, en Outaouais (Sites GA) et les individus de la rive sud vers le nord-est, au Bas-Saint-Laurent (Sites MA).

Bien qu'à grande échelle le paysage d'eau douce puisse s'imposer comme une barrière, la structure hiérarchique qu'il représente offre des opportunités de connectivité à plusieurs niveaux. Localement, le système dendritique agit comme un réseau de corridors de dispersion alors qu'à plus grande échelle, les bassins versants se trouvent isolés, créant une barrière au flux de gènes. En raison de sa capacité de dispersion, de la stabilité de son domaine vital, de la fidélité aux sites de ponte et de la sélection de l'habitat, les tortues des bois sont très souvent

isolées dans les réseaux de rivières, résultant en une différenciation génétique des populations. Les résultats du **Chapitre 3** confirment l'impact significatif de ce modèle d'isolement par bassins versants sur la structure génétique de la tortue des bois. Les analyses de groupement bayésien (**Chapitre 3**) ont révélé une structure hiérarchique au sein de chacune des rives et les analyses de réseaux (**Chapitre 4**) ont permis de préciser la relation entre ces populations. Chacun des bassins versants de la rive nord forme un groupe génétiquement distinct, démontrant un faible niveau de différenciation entre les populations qui y habitent. Les relations entre les populations de la rive sud sont plus complexes et certaines populations semblent avoir une plus grande importance dans la connectivité du réseau. Les populations les plus au sud et qui vivent en milieu densément peuplé ne montrent pas de diminution de la diversité génétique. Par contre, les populations les plus isolées ont une diversité génétique plus faible, ce qui s'explique soit par l'absence de flux génique avec les populations les plus centrales, soit par la colonisation de l'habitat par un petit nombre d'individus lors de l'expansion des populations (effet fondateur). Considérant le décalage temporel observé entre un phénomène et sa réponse au niveau génétique (Landguth et al. 2010), ces résultats seraient le reflet de la diversité génétique passée des populations de tortues des bois.

### 5.3 Priorisation des populations

La procédure de sélection de nœuds du logiciel BRIDES a permis d'identifier quels nœuds ajouter à un sous-réseau afin d'en maximiser la connectivité par rapport à un réseau complet qui contient tous les nœuds possibles. Lorsque plusieurs scénarios de conservation proactifs et réactifs divergents furent comparés (**Chapitre 4**), BRIDES a permis d'identifier de multiples solutions de remplacement. Contrairement au classement des populations par les indices de centralité, les résultats de BRIDES sont robustes pour un même réseau. Six différents types de chemin peuvent être analysés simultanément, chacun représentant une propriété spécifique des chemins possibles au sein de réseaux dynamiques. Ces types de chemins permettent de quantifier différents processus favorisant l'augmentation ou la diminution de la connectivité

en plus de mesurer la redondance des routes. BRIDES a particulièrement permis d'identifier certains nœuds qui n'auraient pas été ciblés par des analyses plus classiques telles que la diversité génétique, les groupements bayésiens et les indices de réseau. C'est le cas, par exemple, de la population MA1, en périphérie de l'aire d'étude, qui est fréquemment sélectionnée par BRIDES pour maximiser la connectivité du réseau et très peu par les analyses mentionnées précédemment. Concrètement, BRIDES pourrait être utilisé pour l'ajout de populations de grande taille dans un sous-réseau de populations précaires. La connectivité des populations est un élément crucial à considérer dans la sélection de ces populations de grande taille puisqu'elle permet de maximiser la résilience du réseau. De plus, ces populations pourraient servir de sources pour le maintien de la diversité génétique.

## **5.4 Application de la génétique pour la conservation de la tortue des bois**

Une des spécificités de la biologie de la conservation est sans aucun doute sa nature appliquée. Dans le cadre de ma thèse, les résultats de mes analyses ont notamment été utilisés pour trouver ou guider des actions concrètes afin de protéger la tortue des bois et sa diversité. Les méthodes que j'applique à la tortue des bois pourraient également s'appliquer à toute autre espèce longévive.

### **Le soutien au recrutement**

Dans la région du Bas-Saint-Laurent, des actions de conservation furent menées, en 2012, dans le but de soutenir le recrutement de la population de tortue des bois, d'en augmenter rapidement les effectifs et d'en assurer la pérennité (Laflèche et al. 2013, 2017). Les œufs de plusieurs femelles furent notamment prélevés dans des populations génétiquement similaires (Bouchard et al. 2013a) et localisées dans le même bassin versant (**Chapitre 3**). Les jeunes furent relâchés dans leur habitat naturel, un ou deux ans après l'éclosion des œufs. Ce délai offre une période de croissance plus longue aux juvéniles, en plus de diminuer le taux de

mortalité causé par la prédation des jeunes tortues (Glowacki & Kuhns 2010, Paterson et al. 2012). Guidés par la théorie de la taille efficace et les analyses de parentalité (résultats pour la population de Shawinigan; **Chapitre 2**), certaines recommandations ont été faites afin de maximiser la diversité génétique des jeunes à prélever. (1) Considérant la fréquence élevée de paternité répétée et la fidélité au site de ponte, nous avons recommandé de prélever des nids de femelles provenant de plusieurs sites de pontes. (2) De plus, nous avons recommandé de prélever le même nombre d'œufs dans chaque nid afin de minimiser les variations du succès reproducteur. (3) Finalement, nous avons aussi recommandé d'augmenter le nombre de sites de ponte afin d'obtenir des jeunes provenant d'un maximum de sites. Les résultats du **Chapitre 3** ont également permis de déterminer que les sites provenant du même bassin versant sont génétiquement similaires et qu'ils peuvent être utilisés lors des stratégies d'augmentation des populations.

### **La délimitation des unités désignables et des populations locales**

À ce jour, le COSEPAC (2018) considère l'ensemble des tortues des bois du Canada comme faisant partie d'une seule et même grande population. Les analyses par bassin versant combinées aux groupements bayésiens (**Chapitre 3**) suggèrent cependant la présence de dix groupes de tortues génétiquement distincts parmi les populations à l'étude. Ces dix groupes répondent au critère de distinction génétique, la première étape pour leur qualification en tant qu'*unités désignables*. À la lueur de mes résultats, il semble impératif de réévaluer la situation de la tortue des bois au Canada.

Afin de planifier le rétablissement de l'espèce, les unités utilisées à des fins de gestion sont les populations locales. Elles sont définies, d'après les données écologiques, comme un groupe d'individus se reproduisant entre eux et occupant un même bassin hydrographique (Environnement Canada 2016). Les résultats des analyses de génétique du paysage (**Chapitre 3**) ont permis de confirmer le rôle des bassins versants pour la différenciation des populations. En dépit d'une similarité génétique entre les tortues de certains bassins versants, mes résultats permettent de consolider la gestion par bassin hydrographique préconisée chez cette espèce.

## Les analyses forensiques

Lorsqu'une espèce est inscrite en annexe de la CITES, il est interdit d'en faire le commerce. En cas de braconnage ou de vente illicite, des analyses forensiques utilisant des marqueurs génétiques (*forensic DNA analysis*) peuvent être utiles pour identifier les individus capturés illégalement. Ces analyses peuvent fournir des indices sur la provenance des individus prélevés, ce qui facilitera leur réintroduction dans leur habitat naturel (Lyengar 2014). Pour ce faire, les profils génétiques des individus capturés sont habituellement comparés à des populations sources potentielles, grâce à des analyses d'assignement et d'exclusion (Alac et al. 2010). Deux tortues des bois furent saisies aux douanes de Hong Kong en 2013 (**Annexe I**, Bouchard et al. 2013b) et le génotype de ces individus a donc été comparé aux profils des populations québécoises afin d'identifier leur provenance. Le haut taux de succès d'assignement aux divers bassins versants et le génotypage des tortues avec l'ensemble des locus résultant du **Chapitre 3** montrent que cette base de données génétique est robuste lors de telles analyses.

## La conservation des espèces longévives

Les méthodes employées dans le cadre de ma thèse pourront être utilisées lors de l'étude d'espèces longévives afin d'obtenir des résultats d'une plus grande précision. Par exemple, les modèles et les analyses canoniques de redondance partielle utilisés au **Chapitre 3** permettent de considérer et de contrôler la variable temporelle et l'autocorrélation spatiale chez ces espèces. Les analyses hiérarchiques de groupement bayésien favorisent aussi une exploration plus fine de la structure génétique. Les analyses de réseaux et le logiciel BRIDES (**Chapitre 4**) présentent des options robustes aux analyses traditionnelles de génétique des populations pour les espèces avec des générations chevauchantes et les populations qui ne respectent pas l'équilibre mutation-dérive et l'équilibre d'Hardy-Weinberg. Suite à une perturbation, la topologie des réseaux se stabilise rapidement (Dyer 2007, Dyer et al. 2010, Noutsos et al. 2014). Ces analyses pourraient ainsi identifier plus aisément les signaux génétiques causés par des perturbations anthropiques, contrairement à certains tests classiques en génétique des populations qui détectent difficilement le déclin récent des populations de tortues (Davy et al. 2014).



Lorsque les impacts d'un évènement récent sont analysés, un délai temporel s'applique fréquemment entre le moment de la perturbation et le signal génétique au sein des populations (Landguth et al. 2010). Chez les espèces longévives, ce délai de quelques générations peut facilement atteindre une durée de plus de 100 ans. Bien qu'il ajoute de la complexité aux analyses, il permet un maintien de la diversité génétique chez ces espèces, malgré une perturbation importante (Marsack & Swanson 2009). La présence d'emmagasinement de sperme favorisera aussi ce maintien et ce délai, puisqu'un mâle pourra produire des descendants après sa mort (Davy et al. 2011).

### **Recherches futures sur la génétique de la tortue des bois**

Ma thèse couvre un large éventail d'objectifs liés à l'utilisation de marqueurs génétiques neutres, tels que l'estimation de la parentalité, la caractérisation de la diversité génétique intra et interpopulations en relation avec la structure du paysage en plus de la priorisation des populations. Par contre, depuis la publication du **Chapitre 2**, de nouveaux marqueurs génétiques furent développés pour déterminer le sexe des individus immatures (Litterman et al. 2017). L'utilisation de tels marqueurs sur les jeunes tortues à l'étude aurait pu révéler le sexe-ratio de la cohorte, un paramètre important en biologie de la conservation. De même, de récentes analyses d'ADN environnemental ont permis la détection de la tortue des bois dans certains cours d'eau, sans qu'aucun individu ait été capturé (Lacoursière-Roussel et al. 2016).

De nombreuses études portant sur la génétique de la tortue des bois ont été réalisées à l'aide de marqueurs neutres tels que les microsatellites. Afin d'acquérir de nouvelles connaissances, les études futures devraient considérer des approches génomiques ou, à tout le moins, utiliser des marqueurs caractérisant la diversité adaptative des populations. Les adaptations locales des populations les plus nordiques, telles que la taille plus importante à la maturité sexuelle et l'hibernation des jeunes dans les nids (Storey et al. 1988), pourraient de cette façon être mieux comprises et estimées. L'étude des populations situées en périphérie de l'aire de distribution de l'espèce pourrait aussi bénéficier d'une approche génomique (Spradling et al. 2010). À long terme, une meilleure compréhension de ces adaptations pourrait appuyer

la prise de décisions des gestionnaires et ainsi favoriser le maintien de l'espèce en ces temps de changements environnementaux importants.

## 5.8 Conclusion

Dans de nombreuses cultures, les tortues sont associées à des symboles de stabilité, de sécurité, de sagesse et de longévité. Par exemple, les légendes de certains peuples autochtones représentent une tortue géante portant le monde sur sa carapace. Chaque espèce de tortue possède une valeur intrinsèque qui lui est conférée par ses adaptations uniques. Les premiers fossiles de tortue telle qu'on se les imagine, avec un plastron et une carapace, datent approximativement de la fin du Trias, il y a 215 millions d'années (Ernst & Lovich 2009, Joyce et al. 2013). Une radiation spectaculaire leur a permis de coloniser les écosystèmes terrestres, dulcicoles et marins, en dépit de leur anatomie particulière. Bien que plusieurs espèces soient désormais éteintes, les tortues ont survécu aux dinosaures et à la grande extinction massive de la fin du Crétacé où l'âge des reptiles prit fin. Aujourd'hui, ces mêmes caractéristiques qui firent leur succès font d'elles des espèces extrêmement vulnérables aux changements environnementaux d'une rapidité sans précédent causé par les activités anthropiques et la surexploitation (Mali et al. 2014).

Puisque les activités anthropiques sont en grande partie responsables de leur déclin (Lovich et al. 2018), des efforts de conservation sont impératifs pour renverser cette tendance. La brèche entre les connaissances et les objectifs des chercheurs et des gestionnaires doit être réduite afin de faciliter le développement et l'utilisation d'outils concrets pour la survie des espèces (Sunderland et al. 2009, Ottewell et al. 2016). À mon avis, il est essentiel que les recherches en biologie de la conservation répondent aux objectifs des gestionnaires du territoire et des décideurs qui possèdent les moyens légaux et financiers pour soutenir les populations d'espèces menacées. Tout au long de mes études doctorales, j'ai toujours voulu m'assurer que mes résultats pouvaient être utiles à l'équipe de rétablissement de la tortue des bois et j'ai tenu compte du plan de rétablissement pour fixer mes objectifs de recherche. Je me

suis plus spécifiquement efforcée de développer, d'utiliser et de promouvoir des analyses qui facilitent l'étude d'espèces longévives comme les tortues.

La communication est essentielle afin d'assurer un suivi à long terme des objectifs de conservation. Malheureusement, plusieurs des biologistes et techniciens qui participent aux travaux sur le terrain ne voient que très rarement les résultats des analyses effectuées sur ce petit bout d'écaille déposé au fond d'un tube d'éthanol. La transmission des connaissances, par le biais de publications scientifiques et de conférences, est une étape cruciale pour que la recherche théorique soit utilisée de façon concrète par les décideurs. Il incombe aux chercheurs de s'impliquer auprès des gestionnaires pour que l'échange de connaissances soit juste et équitable. Pour atteindre cet objectif, il me semble essentiel de promouvoir les recherches et l'enseignement de la biologie de la conservation comme discipline appliquée afin d'assurer la pertinence de sa pratique.

# Bibliographie

- Akre T.S., Parker L.D., Ruther E., Maldonado J.E., Lemmon L., McInerney N.R. (2019) Concurrent visual encounter sampling validates eDNA selectivity and sensitivity for the endangered wood turtle (*Glyptemys insculpta*). PloS One. 14(4):e0215586
- Alacs E.A., Janzen F.J., Scribner K.T. (2007) Genetic issues in freshwater turtle and tortoise conservation. Chelonian Research Monographs. 4:107-123
- Alacs E.A., Georges A., FitzSimmons N.N., Robertson J. (2010) DNA detective: a review of molecular approaches to wildlife forensics. Forensic Science, Medicine, and Pathology. 6(3):180-194
- Allendorf F.W., Bayles D., Bottom D.L. Currens K.P., Frissell C.A., Hankin D., Lichatowich J.A., Nehlsen W., Trotter P.C., Williams T.H. (1997) Prioritizing pacific salmon stocks for conservation. Conservation Biology. 11:140-152
- Amato M.L., Brooks R.J., Fu J. (2008) A phylogeographic analysis of populations of the wood turtle (*Glyptemys insculpta*) throughout its range. Molecular Ecology. 17(2):570-581
- Anderson C.D., Epperson B.K., Fortin M.-J., Holderegger R., James P.M.A., Rosenberg M.S., Scribner K.T., Spear S. (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. Molecular Ecology. 19(17):3565-3575
- Anderson M.J., Legendre P. (1999) An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. Journal of Statistical Computation and Simulation. 62(3):271-303
- Anderson P. (1965) The reptiles of Missouri. University of Missouri Press, Columbia. 330 p.
- Arvais M., Bourgeois J.-C., Levesque E., Daigle C., Masse D., Jutras J. (2002) Home range and movements of a Wood Turtle (*Clemmys insculpta*) population at the northern limit of its range. Canadian Journal of Zoology. 80:402-408
- Arvais M., Levesque E., Bourgeois J.-C., Daigle C., Masse D., Jutras J. (2004) Habitat selection by the Wood Turtle (*Clemmys insculpta*) at the northern limit of its range. Canadian Journal of Zoology. 82:391-398
- Avise J.C., Bowen B.W., Lamb T., Meylan A.B., Bermingham E. (1992) Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. Molecular Biology and Evolution. 9:457-473
- Avise, J.C., Jones A.G., Walker D., DeWoody J.A. (2002) Genetic mating systems and reproductive natural histories of fishes: lessons for ecology and evolution. Annual Review of Genetics. 36:19-45
- Baer C.F., Miyamoto M.M., Denver D.R. (2007) Mutation rate variation in multicellular eukaryotes: causes and consequences. Nature Reviews Genetics. 8(8):619-631
- Balkenhol N., Waits L.P., Dezzani R.J. (2009) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecology. 90(5):818-830

- Balloux F., Lehmann L. (2003) Random mating with a finite number of matings. *Genetics*. 165(4):2313-2315
- Banger N., Blouin-Demers G., Bulté G., Lougheed S.C. (2013) More sires may enhance offspring fitness in Northern Map Turtles (*Graptemys geographica*). *Canadian Journal of Zoology*. 91(8):581-588
- Bateman A.J. (1948) Intra-sexual selection in *Drosophila*. *Heredity*. 2:349-368
- Birkhead T.R., Møller A.P. (1998) Sperm competition and sexual selection. Academic Press, New York. 826 p.
- Birkhead T.R., Pizzari T. (2002) Postcopulatory sexual selection. *Nature Reviews Genetics*. 3(4): 262-273
- Bloomer T.J. (1978) Hibernacula congregating in the *Clemmys* genus. *Journal of the Northern Ohio Association of Herpetologists*. 4:37-42
- Borcard D., Legendre P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*. 153:51-68
- Bouchard C., Tessier N., Bertacchi W., Lapointe F.-J. (2013a). Caractérisation génétique de la population de tortues des bois (*Glyptemys insculpta*) du Bas-Saint-Laurent. Université de Montréal. Rapport présenté au ministère des Ressources naturelles, région du Bas-Saint-Laurent, Montréal. 32 p.
- Bouchard C., Tessier N., Lapointe F.-J. (2013b). Comparaison des profils génétiques de deux tortues des bois (*Glyptemys insculpta*) saisies aux douanes de Hong Kong avec les populations du Québec. Université de Montréal. Rapport présenté à Environnement Canada, Montréal. 10 p.
- Bouchard C., Tessier N., Lapointe F.-J. (2018) Paternity analysis of wood turtles (*Glyptemys insculpta*) reveals complex mating patterns. *Journal of Heredity*. 109(4):405-415
- Bouchard C., Tessier N., Lapointe F.-J. (2019) Watersheds influence the wood turtle's (*Glyptemys insculpta*) genetic structure. *Conservation Genetics*. 20(3):653-664
- Bourgeois J.C., Masse D., Tessier N., Lapointe F.-J. (2009) Plan de conservation de la population de tortues des bois (*Glyptemys insculpta*) du bassin versant de la rivière Shawinigan. Ministère des Ressources naturelles et de la Faune, Québec et Parcs Canada. 41 p.
- Bowen K.D., Gillingham J.C. (2004) R9 species Conservation assessment for wood turtle - *Glyptemys insculpta* (LeConte, 1830). 39 p.
- Bowen B.W., Karl S.A. (2007) Population genetics and phylogeography of sea turtles. *Molecular Ecology*. 16 (23):4886-4907
- Brooks R.J., Brown G.P. (1992) Population biology of the Wood Turtle, *Clemmys insculpta*, in the Madawaska River Drainage. Progress report and proposal to the Ontario Ministry of Natural Resources, Algonquin District.
- Brooks R.J., Shilton C.M., Brown G.P., Quinn N.W.S. (1992) Body size, age distribution, and reproduction in a northern population of Wood Turtles (*Clemmys insculpta*). *Canadian Journal of Zoology*. 70:462-469

- Brooks T.M., Mittermeier R.A., da Fonseca G.A., Gerlach J., Hoffmann M., Lamoreux J.F., Mittermeier C.G., Pilgrim J.D., Rodrigues A.S. (2006) Global biodiversity conservation priorities. *Science*. 313(5783):58-61
- Brown, G.P., Brooks R.J. (1994) Characteristics of and fidelity to hibernacula in a northern population of Snapping Turtles, *Chelydra serpentina*. *Copeia*. 1994(1):222-226
- Cagle F.R. (1939) A system of marking turtles for future identification. *Copeia* 1939(3):170-173
- Calsbeek R., Bonneaud C. (2008) Postcopulatory fertilization bias as a form of cryptic sexual selection. *Evolution*. 62(5):1137-1148
- Calsbeek R., Bonneaud C., Prabhu S., Manoukis N., Smith T.B. (2007) Multiple paternity and sperm storage lead to increased genetic diversity in Anolis lizards. *Evolutionary Ecology Research*. 9(3):495-503
- Cameron M., Goodenough N., McNichols K., Wesley P. (2002) Demography, home range and habitat utilization of Wood Turtles (*Clemmys insculpta*) in the Algoma District. 2002 Project Report for Ontario Ministry of Natural Resources, Sault Ste. Marie District. 26 p.
- Cardillo M., Purvis A., Sechrest W., Gittleman J.L., Bielby J., Mace G.M. (2004) Human population density and extinction risk in the world's carnivores. *PLoS Biology*. 2(7):e197
- Castellano C.M., Behler J.L., Amato G. (2009) Genetic diversity and population genetic structure of the wood turtle (*Glyptemys insculpta*) at Delaware Water Gap National Recreation Area, USA. *Conservation Genetics*. 10:1783-1788
- Caughley G. (1994) Directions in conservation biology. *Journal of Animal Ecology*. 63:215-244
- Cavalli-Sforza L.L., Edwards W.F. (1967) Phylogenetics analysis: models and estimation procedures. *Evolution*. 3:550-570
- Caye K., Jay F., Michel O., Francois O. (2018) Fast inference of individual admixture coefficients using geographic data. *The Annals of Applied Statistics*. 12:586-608
- Chapuis M.-P., Estoup A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*. 24(3):621-631
- Charlesworth B., Morgan M.T., Charlesworth D. (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*. 134(4):1289-1303
- Chen C., Durand E., Forbes F., François O. (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes*. 7:747-756
- Chikhi L., Sousa V.C., Luisi P., Goossens B., Beaumont M.A. (2010) The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics*. 186:983-995
- Clarkson R.W., Marsh P.C., Dowling, T.E. (2012) Population prioritization for conservation of imperilled warmwater fishes in an arid-region drainage. *Aquatic Conservation: Marine and Freshwater Ecosystems*. 22(4):498-510

- Coates D.J., Byrne M., Moritz C. (2018) Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*. 6:165
- Compton B.W., Rhymer J.M., McCollough J. (2002) Habitat selection by Wood Turtles (*Clemmys insculpta*). *Herpetological Review*. 33(3):166
- Congdon J.D., Gibbons J.W. (1990) The evolution of turtle life histories. In: *Life History and Ecology of the Slider Turtle* (ed Gibbons J.W.). Smithsonian Institution Press. Washington, D.C. 45-54 p.
- COSEPAC. (2007) Évaluation et Rapport de situation du COSEPAC sur la tortue des bois (*Glyptemys insculpta*) au Canada - Mise à jour. Comité sur la situation des espèces en péril au Canada. Ottawa. vii + 47 p.
- COSEWIC. (2007) COSEWIC assessment and update status report on the Wood Turtle (*Glyptemys insculpta*) in Canada. Committee on status of Endangered Wildlife in Canada. Ottawa
- COSEWIC. (2015) Guidelines for recognizing designatable units below the species level (approved by COSEWIC November 2015). <http://cosewic.ca/index.php/fr/rapports/preparation-rapports-situation/lignes-directrices-reconnaitre-unites-designables> Access 21 November 2017
- COSEPAC. (2018) Évaluation et Rapport de situation du COSEPAC sur la tortue des bois (*Glyptemys insculpta*) au Canada. Comité sur la situation des espèces en péril au Canada. Ottawa. xiii + 51 p.
- Crawford N.G. (2010) SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources*. 10:556-557
- Creech T.G., Epps C.W., Monello R.J., Wehausen J.D. (2014) Using network theory to prioritize management in a desert bighorn sheep metapopulation. *Landscape Ecology*. 29(4):605-619
- Crooks K.R., Sanjayan M.A. (2006) *Connectivity conservation*. Cambridge University Press, Cambridge. 712 p.
- Cross T.B., Schwartz M.K., Naugle D.E., Fedy B.C., Row J.R., Oyler-McCance S.J. (2018) The genetic network of greater sage-grouse: Range-wide identification of keystone hubs of connectivity. *Ecology and Evolution*. 8(11): 5394-5412
- Cuellar O. (1966) Delayed fertilization in the lizard *Uta stansburiana*. *Copeia*. 1966(3):549-552
- Cureton J.C., Janis M, Lutterschmidt W.I., Randle C.P., Ruthven D.C., Deaton R. (2014) Effects of urbanization on genetic diversity, gene flow, and population structure in the ornate box turtle (*Terrapene ornata*). *Amphibia-Reptilia*. 35(1):87-97
- Curtsinger J.W. (1991) Sperm competition and the evolution of multiple mating. *The American Naturalist*. 138:93-102
- Daigle C. (1997) Size and characteristics of a Wood Turtle, *Clemmys insculpta*, population in southern Québec. *Canadian Field-Naturalist*. 111(3):440-444
- Daigle C., Jutras J. (2005) Quantitative evidence of decline in a southern Québec wood turtle (*Glyptemys insculpta*) population. *Journal of Herpetology*. 39:130-132

- Dakin E.E., Avise J.C. (2004) Microsatellite null alleles in parentage analysis. *Heredity*. 93(5):504-509
- Davy C.M. (2013) Conservation genetics of freshwater turtles. PhD dissertation, University of Toronto, ON, Canada. 163 p.
- Davy C.M., Edwards T., Lathrop A., Bratton M., Hagan M., Henen B., Nagy K.A., Stone J., Hillard L.S., Murphy R.W. (2011) Polyandry and multiple paternities in the threatened Agassiz's desert tortoise, *Gopherus agassizii*. *Conservation Genetics*. 12(5):1313-1322
- Davy C.M., Murphy R.W. (2014) Conservation genetics of the endangered Spotted Turtle (*Clemmys guttata*) illustrate the risks of "bottleneck tests". *Canadian Journal of Zoology*. 92(2):149-162
- DeFaveri J., Viitaniemi H., Leder E., Merilä J. (2013) Characterizing genic and nongenic molecular markers: comparison of microsatellites and SNPs. *Molecular Ecology Resources*. 13(3):377-392
- DeSalle R., Amato G. (2004) The expansion of conservation genetics. *Nature reviews. Genetics* 5(9):702
- DiBattista J.D. (2008) Patterns of genetic variation in anthropogenically impacted populations. *Conservation Genetics*. 9(1):141-156
- Dinkins A. (1954) A brief observation on male combat in *Clemmys insculpta*. *Herpetologica*. 10(1):20
- Duchesne P., Turgeon J. (2012) FLOCK provides reliable solutions to the "number of populations" problem. *Journal of Heredity*. 103(5):734-743
- Duran N., Dunbar S.G., Escobar III R.A., Standish T.G. (2015) High frequency of multiple paternity in a solitary population of olive ridley sea turtles in Honduras. *Journal of Experimental Marine Biology and Ecology*. 463:63-71
- Durand E., Jay F., Gaggiotti O.E., François O. (2009) Spatial inference of admixture proportions and secondary contact zones. *Molecular Biology and Evolution*. 26:1963-1973
- Dyer R.J. (2007) The evolution of genetic topologies. *Theoretical Population Biology*. 71:71-78
- Dyer R.J. (2012) The gstudio package. Virginia Commonwealth University, Virginia.
- Dyer R.J. (2014) Popgraph: This is an R package that constructs and manipulates population graphs, R Package version 1.4.
- Dyer R.J. (2015) Population graphs and landscape genetics. *Annual Review of Ecology, Evolution, and Systematics*. 46:327-342
- Dyer R.J. (2016) gstudio: Tools related to the spatial analysis of genetic marker data. R package version 1.5.0.
- Dyer R.J., Nason J.D. (2004) Population graphs: the graph theoretic shape of genetic structure. *Molecular Ecology*. 13(7):1713-1727
- Dyer R.J., Nason J.D., Garrick R.C. (2010) Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*. 19:3746-3759



- Earl D.A., vonHoldt B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. 4(2):359-361
- Ekanayake E.M.L., Kapurusinghe T., Saman M.M, Rathnakamura D.S., Samaraweera P., Ranawana K.B., Rajakaruna R.S. (2013) Paternity of green turtle (*Chelonia mydas*) clutches laid in Kosgoda, Sri Lanka. *Herpetological Conservation and Biology*. 8(1):27-36
- Environment Canada (2016) Recovery strategy for the Wood Turtle (*Glyptemys insculpta*) in Canada. Species at Risk Act Recovery Strategy Series. Environment Canada, Ottawa. 55 p.
- Epps C.W., Keyghobadi N. (2015) Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Molecular Ecology*. 24:6021-6040
- Équipe de rétablissement de cinq espèces de tortues au Québec pour les années 2005 à 2010: la tortue des bois (*Glyptemys insculpta*), la tortue géographique (*Graptemys geographica*), la tortue mouchetée (*Emydoidea blandingii*), la tortue musquée (*Sternotherus odoratus*) et la tortue ponctuée (*Clemmys guttata*). (2005) Ministère des Ressources Naturelles et de la Faune, Québec. 57 p.
- Ernst C.H. (2001) Some ecological parameters of the Wood Turtle, *Clemmys insculpta*, in southeastern Pennsylvania. *Chelonian Conservation and Biology*. 4:94-99
- Ernst C.H., Lovich J.E. (2009) Turtles of the United States and Canada. Second edition. Johns Hopkins University Press, Baltimore. 827 p.
- Erős T., Schmera D., Schick R.S. (2011) Network thinking in riverscape conservation: a graph-based approach. *Biological Conservation*. 144(1):184-192
- Estoup A., Jarne P., Cornuet J.-M. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*. 11(9):1591-1604
- Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*. 14(8):2611-2620
- Evans L.T. (1961) Aquatic courtship of the wood turtle, *Clemmys insculpta*. *American Zoologist* 1:353
- Ewing H.E. (1943) Continued fertility in female box turtles following mating. *Copeia* 1943:112-114
- Excoffier L., Laval G., Schneider S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*. 1:47-50
- Ezard T.H.G., Travis J.M.J. (2006) The impact of habitat loss and fragmentation on genetic drift and fixation time. *Oikos*. 114(2):367-375
- Fitzpatrick M.C., Keller S.R. (2015) Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*. 18:1-16
- Fitzsimmons N.N. (1998) Single paternity of clutches and sperm storage in the promiscuous green turtle (*Chelonia mydas*). *Molecular Ecology*. 7(5):575-584
- François O., Durand E. (2010) Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources*. 10(5):773-784

- Frankel O.H. (1970) Variation - the essence of life. *Proceedings of the Linnean Society of New South Wales*. 95:158-196
- Frankel O.H., Soulé M.E. (1981) *Conservation and evolution*. Cambridge University Press, Cambridge. 358 p.
- Frankham R. (2000) Modeling problems in conservation genetics using laboratory animals. In: *Quantitative Methods for Conservation Biology* (eds Ferson S., Burgman M.). Springer, New York. 259-273 p.
- Frankham R. (2005) Genetics and extinction. *Biological Conservation*. 126(2):131-140
- Frankham R. (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*. 24:2610-2618
- Frankham R., Kingslover J.G. (2004) Response to environmental change: adaptation or extinction. In: *Evolutionary Conservation Biology* (eds Ferrier R., Dieckman U., Couvet D.). Cambridge University Press, Cambridge. 85-100 p.
- Freedman A.H., Thomassen H.A., Buermann W., Smith T.B. (2010) Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology*. 19:3773-3788
- Frigden C., Finnegan L., Reaume C., Cebek J., Trotter J., Wilson P. (2013) Conservation of wood turtle (*Glyptemys insculpta*) populations in Ontario, Canada. *Herpetological Conservation and Biology*. 8(2):351-358
- Galbraith D.A. (1991) Studies of mating systems in wood turtles (*Clemmys insculpta*) and snapping turtles (*Chelydra serpentina*) using DNA fingerprinting. PhD dissertation, Queen's University, Kingston, Ontario.
- Galbraith D.A., White B.N., Brooks R.J., Boag P.T. (1993) Multiple paternity in clutches of snapping turtles (*Chelydra serpentina*) detected using DNA fingerprints. *Canadian Journal of Zoology*. 71(2):318-324
- Galois P., Bonin J. (1999) Rapport sur la situation de la tortue des bois (*Clemmys insculpta*) au Québec. Ministère de l'Environnement et de la faune, Direction de la faune et des habitats. 45 p.
- Garroway C.J., Bowman J., Carr D., Wilson P.J. (2008) Applications of graph theory to landscape genetics. *Evolutionary Applications*. 1(4):620-630
- Gazette officielle du Québec. (2005) Règlement modifiant le Règlement sur les espèces fauniques menacées ou vulnérables et leurs habitats. 137:705-706
- Gerber S., Mariette S., Streiff R., Bodénès C., Kremer A. (2000) Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology*. 9(8):1037-1048
- Gibbon J.W., Scott D.E., Ryan T.J., Buhlmann K.A., Tuberville T.D., Metts B.S., Greene J.L., Milla T., Leiden Y., Poppy S., Winne C.T. (2000) The global decline of reptiles, déjà vu amphibians: reptile species are declining on a global scale. Six significant threats to reptile populations are habitat loss and degradation, introduced invasive species, environmental pollution, disease, unsustainable use, and global climate change. *BioScience*. 50(8):653-666

- Gibbs J.P., Shriver W.G. (2002) Estimating the effects of road mortality on turtle populations. *Conservation Biology*. 16(6):1647-1652
- Gibbs J.P., Amato G.D. (2000) Genetics and demography in turtle conservation. In: *Turtle Conservation* (ed Klemens M.W.). Smithsonian Institution Press, Washington, D.C. 207-217 p.
- Gilpin M.E., Soulé M.E. (1986) Minimum viable population: process of species extinction. In: *Conservation Biology: The Science of Scarcity and Diversity* (ed Soulé M.E.) Sinauer Associates, Sunderland, Mass. 19-34 p.
- Gist D.H., Congdon J.D. (1998) Oviductal sperm storage as a reproductive tactic of turtles. *Journal of Experimental Zoology*. 282:526-534
- Gist D.H., Jones J.M. (1987) Storage of sperm in the reptilian oviduct. *Scanning Microscopy*. 1:1839-1849
- Gist D.H., Jones J.M. (1989) Sperm storage within the oviduct of turtles. *Journal of Morphology*. 199(3):379-384
- Given D.R., Norton D.A. (1993) A multivariate approach to assessing threat and for priority setting in threatened species conservation. *Biological Conservation*. 64:57-66
- Glowacki G., Kuhns A.R. (2010). Recovery of the Blanding's turtle (*Emydoidea blandingii*) at Spring Bluff Nature Preserve, Lake County Forest Preserves. Spring Edition: Department of Natural Resources. 39 p.
- Goin C.J., Goin O.B., Zug G.R. (1978) *Introduction to herpetology*. Freeman, New York. 378 p.
- Griffith S., Owens I., Thuman K. (2002) Extrapair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology*. 11:2195-2212
- Gubili C., Mariani S., Weckworth B.V., Galpern P., McDevitt A.D., Hebblewhite, M., Nikel B., Musiani M. (2017) Environmental and anthropogenic drivers of connectivity patterns: a basis for prioritizing conservation efforts for threatened populations. *Evolutionary Applications*. 10:199-211
- Guichoux E., Lagache L., Wagner S., Chaumeil P., Léger P., Lepais O., Lepoittevin C., Malausa T., Revardel E., Salin F., Petit R.J. (2011) Current trends in microsatellite genotyping. *Molecular Ecology Resources*. 11(4):591-611
- Guillot G., Mortier F., Estoup A. (2005) Geneland: a program for landscape genetics. *Molecular Ecology Notes*. 5:712-715
- Haig S. (1998) Molecular contributions to conservation. *Molecular Techniques in Ecology*. 79(2):413-425
- Hanski I., Gilpin M.E. (1991) Metapopulation dynamics: brief history and conceptual domain *Biological Journal of the Linnean Society*. 42(1-2):3-16
- Hanski I. (1998) Metapopulation dynamics. *Nature*. 396(6706):41-49
- Harding J.H., Bloomer T.J. (1979) The wood turtle, *Clemmys insculpta*: a natural history. *Bulletin of the New York Herpetological Society*. 15:9-26

- Harrison S., Taylor A.D. (1997) Empirical evidence for metapopulation dynamics. In: *Metapopulation Biology: Ecology, Genetics, and Evolution* (eds Hanski I., Gilpin M.E.). Academic Press, San Diego. 27-42 p.
- Hedgecock D., Barber P.H., Edmands S. (2007) Genetic approaches to measuring connectivity. *Oceanography* 20(3):70-79
- Hedrick P.W. (2001) Conservation genetics: where are we now?. *Trends in Ecology and Evolution*. 16(11):629-636
- Hedrick P. (2005) Large variance in reproductive success and the  $N_e/N$  ratio. *Evolution*. 59(7):1596-1599
- Hoban S. (2018) Integrative conservation genetics: prioritizing populations using climate predictions, adaptive potential and habitat connectivity. *Molecular Ecology Resources*. 18(1):14-17
- Holderegger R., Kamm, U., Gugerli F. (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*. 21(6):797-807
- Holderegger R., Wagner H.H. (2008) Landscape genetics. *BioScience*. 58:199-207
- Holman L., Kokko H. (2013) The consequences of polyandry for population viability, extinction risk and conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 368(1613):20120053
- Hoshino A.A., Bravo J.P., Morelli K.A., Nobile P.M. (2012) Microsatellites as tools for genetic diversity analysis. INTECH Open Access Publisher. doi: 10.5772/35363
- Howell H.J., Legere Jr R.H., Holland D.S., Seigel R.A. (2019) Long-term turtle declines: protected is a verb, not an outcome. *Copeia*. 107(3):493-501
- Hughes J.B., Daily G.C., Ehrlich P.R. (1997) Population diversity: its extent and extinction. *Science* 280(5338):689- 692.
- The IUCN Red List of Threatened Species. 2015 Version 2015-4. <<http://www.iucnredlist.org>>. Downloaded on 19 November 2015
- Jacoby D.M.P., Freeman R. (2016) Emerging network-based tools in movement ecology. *Trends in Ecology and Evolution*. 31:301-314
- Jakobsson M., Rosenberg N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 23:1801-1806
- Jennions M.D., Petrie M. (2000) Why do female mate multiply? A review of the genetic benefits. *Biological Reviews of the Cambridge Philosophical Society*. 75(1):21-64
- Jensen E.L., Govindarajulu P., Russello M.A. (2014) When the shoe doesn't fit: applying conservation unit concepts to western painted turtles at their northern periphery. *Conservation Genetics* 15(2):261-274
- Jones A.G. (2005) GERUD 2.0: a computer program for the reconstruction of parental genotypes from half-sib progeny arrays with known or unknown parents. *Molecular Ecology Notes*. 5(3):708-711

- Jones A.G., Ardren W.R. (2003) Methods of parentage analysis in natural populations. *Molecular Ecology*. 12(10):2511-2523
- Jones A.G., Small C.M., Paczolt K.A., Ratterman N.L. (2010) A practical guide to methods of parentage analysis. *Molecular Ecology Resources*. 10(1):6-30
- Jones B., Grossman G.D., Walsh D.C.I., Porter B.A., Avise J.C., Fiumera A.C. (2007) Estimating differential reproductive success from nests of related individuals, with application to a study of the mottled sculpin, *Cottus bairdi*. *Genetics*. 176:2427-2439
- Jones M.T., Sievert P.R. (2009) Effects of stochastic flood disturbance on adult wood turtles, *Glyptemys insculpta*, in Massachusetts. *The Canadian Field-Naturalist*. 123(4):313-322
- Jones O., Wang J. (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*. 10(3):551-555
- Jones M.T., Willey L.L., Sievert P.R., Akre T.S.B. (2015) Status and Conservation of the Wood Turtle in the Northeastern United States. Report submitted to the Northeast Association of Fish and Wildlife Agencies for Regional Conservation Needs Grant 2011–02. <https://rcngrants.org/sites/default/files/datasets/RCN2011-02v2.pdf>
- Jordán F., Báldi A., Orci K.M., Racz I., Varga Z. (2003) Characterizing the importance of habitat patches and corridors in maintaining the landscape connectivity of a *Pholidoptera transylvanica* (Orthoptera) metapopulation. *Landscape Ecology*. 18(1):83-92
- Joseph J., Shaw P.W. (2011) Multiple paternity in egg clutches of hawksbill turtles (*Eretmochelys imbricata*). *Conservation Genetics*. 12(2):601-605
- Jost L. (2008)  $G_{ST}$  and its relatives do not measure differentiation. *Molecular Ecology*. 17(18):4015-4026
- Joyce, W. G., Schoch, R. R., Lyson, T. R. (2013). The girdles of the oldest fossil turtle, *Proterochersis robusta*, and the age of the turtle crown. *BMC Evolutionary Biology*. 13(1):1-9.
- Jun-Yi L. (1982) Sperm retention in the lizard *Chamaeleo boehmei*. *Copeia*. 1982(2):488-489
- Kalinowski S.T. (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics*. 5:539-543
- Kalinowski S.T. (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes*. 5:187-189
- Kalinowski S.T., Taper M.L., Marshall T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*. 16(5):1099-1106
- Kaplan R.H., Cooper W.S (1984) The evolution of developmental plasticity in reproductive characteristics: an application of the "adaptive coin-flipping" principle. *The American Naturalist*. 123(3):393-410
- Karl S.A. (2008) The effect of multiple paternity on the genetically effective size of a population. *Molecular Ecology*. 17(18): 3973-3977

- Kaufmann J.H. (1992) The social behavior of wood turtles, *Clemmys insculpta*, in central Pennsylvania. Herpetological Monographs. 6:1-25
- Keenan K., McGinnity P., Cross T.F., Crozier W.W., Prodöhl P.A. (2013) diveRcity: an R package for the estimation of population genetics parameters and their associated errors. Methods in Ecology and Evolution. 4(8):782-788
- Kim S. (2015) ppcor: an R package for a fast calculation to semi-partial correlation coefficients. Communications for Statistical Applications and Methods. 22(6):665-674
- King T.L., Julian S.E. (2004) Conservation of microsatellite DNA flanking sequence across 13 Emydid genera assayed with novel bog turtle (*Glyptemys mublenbergii*) loci. Conservation Genetics. 5(5):719-725
- Kitchen A.M., Gese E.M., Waits L.P., Karki S.M., Schauster E.R. (2006) Multiple breeding strategies in the swift fox, *Vulpes velox*. Animal Behaviour. 71:1029-1038
- Klemens M.W. (2000) Turtle conservation. Smithsonian Institution Press, Washington, DC. 334 p.
- Koen E.L., Bowman J., Wilson P.J. (2016) Node-based measures of connectivity in genetic networks. Molecular Ecology Resources. 16(1):69-79
- Koen E.L., Bowman J., Garroway C.J., Wilson P.J. (2013) The sensitivity of genetic connectivity measures to unsampled and under-sampled sites. PLoS One. 8(2):e56204
- Kool J.T., Moilanen A., Treml E.A. (2013) Population connectivity: recent advances and new perspectives. Landscape Ecology. 28(2):165-185
- Kuo C.H., Janzen F.J. (2004) Genetic effects of a persistent bottleneck on a natural population of ornate box turtles (*Terrapene ornata*). Conservation Genetics. 5(4):425-437
- Lacoursière-Roussel A., Dubois Y., Normandeau E., Bernatchez L. (2016) Improving herpetological surveys in eastern North America using the environmental DNA method. Genome. 59(11):991-1007
- Laflèche M., Bertacchi W., Maisonneuve C. (2013) Suivi de la sous-population de tortues des bois de la rivière aux Bouleaux. Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs, Direction régionale du Bas-Saint-Laurent - Secteur de la faune. Rimouski. 40 p.
- Laflèche M., Bertacchi W., Maisonneuve C., Bourget G. (2017) Programme de repeuplement de la population de tortues des bois (*Glyptemys insculpta*) de la région du Témiscouata/Madawaska. Bilan des activités 2012-2014. Ministère des Forêts, de la Faune et des Parcs. Direction de la gestion de la faune du Bas-Saint-Laurent. 42 p.
- Laita A., Kotiaho J., Mönkkönen M. (2011) Graph-theoretic connectivity measures: what do they tell us about connectivity? Landscape Ecology. 26(7):951-967
- Lamy T., Pointier J.P., Jarne P., David P. (2012) Testing metapopulation dynamics using genetic, demographic and ecological data. Molecular Ecology. 21(6):1394-1410
- Landguth E.L., Cushman S.A., Schwartz M.K., McKelvey K.S., Murphy M., Luikart G. (2010) Quantifying the lag time to detect barriers in landscape genetics. Molecular Ecology. 19(19):4179-4191



- Landguth E.L., Fedy B.C., Oyler-McCance S.J., Garey A.L., Emel S.L., Mumma M., Wagner H.H., Fortin M.-J., Cushman S.A. (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*. 12:276-284
- Langella O. (2002) Population 1.2.30. Copyright (C) 1999, Olivier Langella, CNRS-UPR9034. Disponible à: <http://bioinformatics.org/~tryphon/populations/>
- Laurila A., Seppä P. (1998) Multiple paternity in the common frog: genetic evidence from tadpole kin groups. *Biological Journal of the Linnean Society*. 63:221-232
- Legendre P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*. 74(6):1659-1673
- Legendre P., Anderson M.J. (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*. 69(1):1-24
- Legendre P., Lapointe F.-J., Casgrain P. (1994) Modeling brain evolution from behavior: a permutational regression approach. *Evolution*. 48(5):1487-1499
- Legendre P., Oksanen J., ter Braak C.J.F. (2011) Testing the significance of canonical axes in redundancy analysis. *Methods in Ecology and Evolution* 2:269-277
- Lesica P., Allendorf F.W. (1995) When are peripheral populations valuable for conservation? *Conservation Biology*. 9:753-760
- Levins R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological of Society America*. 15:237-240
- Levins R. (1970) Extinction. *Lecture Notes in Mathematics*. 2:75-107
- Lindenmayer D., Scheele B. (2017) Do not publish. *Science*. 356:800-801
- Literman R., Radhakrishnan S., Tamplin J., Burke R., Dresser C., Valenzuela N. (2017) Development of sexing primers in *Glyptemys insculpta* and *Apalone spinifera* turtles uncovers an XX/XY sex-determining system in the critically-endangered bog turtle *Glyptemys mublenbergii*. *Conservation Genetics Resources*. 9(4):651-658
- Litzgus J.D., Brooks J.R. (1996) Status report on the wood turtle, *Clemmys insculpta*, in Canada. COSEWIC status report. Ottawa, Ontario. 76 p.
- Litzgus J.D., Mousseau T.A. (2006) Geographic variation in reproduction in a freshwater turtle (*Clemmys guttata*). *Herpetologica*. 62(2):132-140
- Liu Y.X., Davy C.M, Shi H.T., Murphy R.W. (2013) Sex in the half-shell: a review of the functions and evolution of courtship behavior in freshwater turtles. *Chelonian Conservation and Biology*. 12(1):84-100
- Lord E., Le Cam M., Baptiste É., Méheust R., Makarenkov V., Lapointe F.-J. (2016) BRIDES: a new fast algorithm and software for characterizing evolving similarity networks using Breakthroughs, Roadblocks, Impasses, Detours, Equals and Shortcuts. *PloS One*. 11(8):e0161474
- Lotterhos K.E. (2011) The context-dependent effect of multiple paternity on effective population size. *Evolution*. 65(6):1693-1706

- Lovich J.E., Ernst C.H., McBreen J.F. (1990) Growth, maturity, and sexual dimorphism in the wood turtle, *Clemmys insculpta*. Canadian Journal Zoology. 68:672-677
- Lovich J.E., Ennen J.R., Agha M., Gibbons J.W. (2018) Where have all the turtles gone, and why does it matter? BioScience. 68:771-781
- Lowe W.H., Allendorf F.W. (2010) What can genetics tell us about population connectivity? Molecular Ecology. 19:3038-3051
- Lowe A.J., Smyth A.K., Atkins K., Avery R., Belbin L., Brown N., Budden A.E., Gioia P., Guru S., Hardie M., Hirsch T., Hobern D., La Salle J., Loarie S.R., Miles M., Milne D., Nicholls M., Rossetto M., Smits J., Sparrow B., Terrill G., Turner D., Wardle G.M. (2017) Publish openly but responsibly. Science. 357:141.
- Luque S., Saura S., Fortin M.-J. (2012) Landscape connectivity analysis for conservation: insights from combining new methods with ecological and genetic data. Landscape Ecology. 27:153-157
- Lynch M. (1991) The genetic interpretation of inbreeding depression and outbreeding depression. Evolution. 45(3):622-629
- MacArthur R.H., Wilson E.O. (1967) The theory of island biogeography. Vol. 1. Princeton University Press, Princeton. 224 p.
- Mace G.M., Lande R. (1991) Assessing extinction threats: toward a reevaluation of IUCN threatened species categories. Conservation Biology. 5:148-157
- Mali I., Vandeweghe M.W., Davis S.K., Forstner M.R.J. (2014) Magnitude of the freshwater turtle exports from the US: Long term trends and early effects of newly implemented harvest management regimes. PLoS One. 9:e86478
- Manel S., Berthier P., Luikart G. (2002) Detecting wildlife poaching: identifying the origin of individuals with bayesian assignment tests and multilocus genotypes. Conservation Biology. 16(3):650-659
- Manel S., Schwartz M.K., Luikart G., Taberlet P. (2003) Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution. 18(4):189-197
- Manel S., Holderegger R. (2013) Ten years of landscape genetics. Trends in Ecology and Evolution. 28(10):614-621
- Marchand M.N., Litvaitis J.A. (2004) Effects of habitat features and landscape composition on the population structure of a common aquatic turtle in a region undergoing rapid development. Conservation Biology. 18(3):758-767
- Marrotte R.R., Bowman J., Brown M.G., Cordes C., Morris K.Y., Prentice M.B., Wilson P.J. (2017) Multi-species genetic connectivity in a terrestrial habitat network. Movement Ecology. 5(21):1-11
- Marsack K., Swanson B.J. (2009) A genetic analysis of the impact of generation time and road-based habitat fragmentation on eastern box turtles (*Terrapene c. carolina*). Copeia. 4:647-652
- Masse D., Robitaille Y. (2014) Bilan du suivi de la nidification de la tortue des bois (*Glyptemys insculpta*), population de la rivière Shawinigan (1994-2014) et efficacité du programme de protection



des nids (1994-2014). Québec, Canada. Rapport final présenté à Parcs Canada et au ministère des Forêts, de la Faune et des Parcs du Québec. 72 p.

- Master L.L. (1991) Assessing threats and setting priorities for conservation. *Conservation Biology*. 5:559-563
- Matos C., Petrovan S.O., Wheeler P.M., Ward A.I. (2019) Landscape connectivity and spatial prioritization in an urbanising world: a network analysis approach for a threatened amphibian. *Biological Conservation*. 237:238-247
- McCracken G.F., Burghardt G.M., Houts S.E. (1999) Microsatellite markers and multiple paternity in the garter snake *Thamnophis sirtalis*. *Molecular Ecology*. 8:1475-1479
- McGinnty P., Prodöhl P., Ferguson A., Hynes R., Maoiléidigh N.O., Baker N., Cotter D., O’Hea B., Cooke D., Rogan G., Taggart J., Cross T. (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 270(1532): 2443-2450
- McGuire J.M., Congdon J.D., Scribner K.T., Capps J.D. (2011) Variation in female reproductive quality and reproductive success of male Midland Painted Turtles (*Chrysemys picta marginata*). *Canadian Journal of Zoology*. 89(11):1136-1145
- McGuire J.M., Scribner K.T., Congdon J.D. (2013) Spatial aspects of movements, mating patterns, and nest distributions influence gene flow among population subunits of Blanding’s turtles (*Emydoidea blandingii*). *Conservation Genetics*. 14(5):1029-1042
- McGuire J.M., Congdon J.D., Scribner K.T., Nagle R.D. (2014) Female reproductive qualities affect male painted turtle (*Chrysemys picta marginata*) reproductive success. *Behavioral Ecology and Sociobiology*. 68(10):1589-1602
- McGuire J.M., Congdon J.D., Kinney O.M., Osentoski M., Scribner K.T. (2015) Influences on male reproductive success in long-lived Blanding’s Turtles (*Emydoidea blandingii*). *Canadian Journal of Zoology*. 93(6):487-497
- McKee J.K., Sciulli P.W., Fooce C.D., Waite T.A. (2004) Forecasting global biodiversity threats associated with human population growth. *Biological Conservation*. 115(1):161-164
- McTaggart S.J. (2000) Good genes or sexy sons? Testing the benefits of female choice in the painted turtle, *Chrysemys picta*. MSc thesis, University of Guelph, Guelph, Ontario, Canada. 64 p.
- Meffe G.K., Carroll C.R. (1997) Genetics: conservation of diversity within species. In: *Principles of Conservation Biology* (eds Meffe G.K., Carroll C.R.) Sinauer Associates, Sunderland, Mass. 161-201 p.
- Mills L. (2007) *Conservation of wildlife populations: demography, genetics, and management*. Blackwell Publishing, Malden. 342 p.
- de Miranda, E.B.P. (2017). The plight of reptiles as ecological actors in the tropics. *Frontiers in Ecology and Evolution*. 5:159
- Moilanen A., Smith A.T., Hanski I. (1998) Long-term dynamics in a metapopulation of the American pika. *The American Naturalist*. 152(4):530-542

- Moilanen, A. (2011). On the limitations of graph-theoretic connectivity in spatial ecology and conservation. *Journal of Applied Ecology*. 48(6):1543-1547
- Moore J.A., Miller H.C., Daugherty C.H., Nelson N.J. (2008) Fine scale genetic structure of a long-lived reptile reflects recent habitat modification. *Molecular Ecology*. 17:4630-4641
- Moritz C. (1994) Defining “Evolutionarily Significant Units for conservation”. *Trends in Ecology and Evolution*. 9:373-375
- Mortimer J.A. (1995) Teaching critical concepts for the conservation of sea turtles. *Marine Turtle Newsletter*. 71:1-4
- Myers E.M., Zamudio K.R. (2004) Multiple paternity in an aggregate breeding amphibian: the effect of reproductive skew on estimates of male reproductive success. *Molecular Ecology*. 13(7):1951-1963
- NatureServe. 2002. Element Occurrence Data Standard. NatureServe, Arlington, Virginia, USA. [http://downloads.natureserve.org/conservation\\_tools/element\\_occurrence\\_data\\_standard.pdf](http://downloads.natureserve.org/conservation_tools/element_occurrence_data_standard.pdf)
- Naujokaitis-Lewis I., Rico Y., Lovell J., Fortin M.-J., Murphy M. (2013) Implications of incomplete networks on estimation of landscape genetic connectivity. *Conservation Genetics*. 14:287-298
- Neigel J.E. (1997) A comparison of alternative strategies for estimating gene flow from genetic markers. *Annual Review of Ecology and Systematics*. 28:105-128
- Noutsos C., Borevitz J.O., Hodges S.A. (2014) Gene flow between nascent species: genotypic and phenotypic differentiation within and between *Aquilegia formosa* and *A. pubescens*. *Molecular Ecology* 23:5589-98
- O'Brien S.J., Wildt D.E., Goldman D., Merrill C.R., Bush M. (1983) The cheetah is depauperate in genetic variation. *Science*, 221(4609):459-462
- O'Brien S.J., Roelke M.E., Marker L., Newman A., Winkler C.A., Meltzer D., Colly L., Evermann J.F., Bush M., Wildt, D. E. (1985) Genetic basis for species vulnerability in the cheetah. *Science*. 227(4693):1428-1434
- Oksanen J.F., Blanchet G., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin P.R., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.H., Szoecs E., Wagner H. (2017) *vegan: Community Ecology Package*. R package version 2.4-2.
- Olsson M., Madsen T., Shine R., Gullberg A., Tegelström H. (1994) Rewards of promiscuity. *Nature*. 372(6503):230
- Ottewell K.M., Bickerton D.C., Byrne M., Lowe A.J. (2016) Bridging the gap: a genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. *Diversity and Distributions*. 22(2):174-188
- Paetkau D., Slade R., Burden M., Estoup A. (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation based exploration of accuracy and power. *Molecular Ecology*. 13:55-65

- Palmer K.S., Rostal D.C., Grumbles J.S., Mulvey M. (1998) Long-term sperm storage in the desert tortoise (*Gopherus agassizii*). *Copeia*. 1998(3):702-705
- Palsbøll P.J., Berube M., Allendorf F.W. (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution*. 22(1):11-16
- Papworth S.K., Rist J., Coad L., Milner-Gulland E.J. (2009) Evidence for shifting baseline syndrome in conservation. *Conservation Letters*. 2:93-100
- Parker G.A. (1998) Sperm competition and the evolution of ejaculates: towards a theory base. In: *Sperm Competition and Sexual Selection* (eds Birkhead T.R., Møller A.P.). Academic Press, London. 783-826 p.
- Parren S.G. (2013) A twenty-five year study of the wood turtle (*Glyptemys insculpta*) in Vermont: movements, behavior, injuries, and death. *Herpetological Conservation and Biology*. 8(1):176-190
- Paterson J.E., Steinberg D.B., Litzgus J.D. (2012) Revealing a cryptic life-history stage: differences in habitat selection and survivorship between hatchlings of two turtle species at risk (*Glyptemys insculpta* and *Emydoidea blandingii*). *Wildlife Research* 39(5):408-418.
- Paz-Vinas I., Quemere E., Chikhi L., Loot G., Blanchet S. (2013) The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Molecular Ecology*. 22:3279-3291
- Pearce T., Parker P.G. (1996) Local genetic structure within two rookeries of *Chelonia mydas* (the green turtle). *Heredity*. 77(6):619-628
- Peakall R.O.D., Smouse P.E. (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 6(1):288-295
- Peakall R.O.D., Smouse P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*. 28:2537-2539
- Pearse D.E., Anderson E.C. (2009) Multiple paternity increases effective population size. *Molecular Ecology*. 18(15):3124-3127
- Pearse D.E., Avise J.C. (2001) Turtle mating systems: behavior, sperm storage, and genetic paternity. *Journal of Heredity*. 92(2):206-211
- Pearse D.E., Janzen F.J., Avise J.C. (2001) Genetic markers substantiate long-term storage and utilization of sperm by female painted turtles. *Heredity*. 86(3):378-384
- Pearse D.E., Janzen F.J., Avise J.J. (2002) Multiple paternity, sperm storage, and reproductive success of female and male painted turtles (*Chrysemys picta*) in nature. *Behavioral Ecology and Sociobiology*. 51(2):164-171
- Peres-Neto P.R., Legendre P., Dray S., Borcard D. (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*. 87(10):2614-2625
- Petit R.J., El Mousadik A., Pons O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*. 12(4):844-855

- Phillips K.P., Jorgensen T.H., Jolliffe K.G., Jolliffe S.M., Henwood J., Richardson D.S. (2013) Reconstructing paternal genotypes to infer patterns of sperm storage and sexual selection in the hawksbill turtle. *Molecular Ecology*. 22(8):2301-2312
- Piry S., Alapetite A., Cornuet J.-M., Paetkau D., Baudoin L. et Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*. 95:536-539
- Pittman S.E., King T.L., Faurby S., Dorcas M.E. (2011) Demographic and genetic status of an isolated population of bog turtles (*Glyptemys mublenbergii*): implications for managing small populations of long-lived animals. *Conservation Genetics*. 12(6):1589-1601
- Powell C.B. (1967) Female sexual cycles of *Chrysemys picta* and *Clemmys insculpta* in Nova Scotia. *Canadian Field-Naturalist*. 81(2):134-140
- Price T.D. (1984) The evolution of sexual size dimorphism in Darwin's finches. *The American Naturalist*. 123(4):500-518
- Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics*. 155(2):945-959
- Proulx S.R., Promislow D.E.L., Phillips P.C. (2005) Network thinking in ecology and evolution. *Trends in Ecology and Evolution*. 20(6):345-353
- Prunier J.G., Kaufmann B., Léna J.-P., Fenet S., Pompanon F., Joly P. (2014) A 40-year-old divided highway does not prevent gene flow in the alpine newt *Ichthyosaura alpestris*. *Conservation Genetics*. 15(2):453-468
- Pulliam H.R. (1988) Sources, sinks, and population regulation. *The American Naturalist*. 132(5):652-661
- Quaglietti B., Pekas A., Wackërs F., Crochard D., Fauvergue X., Amouroux P., Palero F., Malausa T. (2016) Inbreeding and outbreeding in biological control agents: do they actually matter? International Congress of Entomology, Orlando, Florida.
- R Development Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ralls K, Ballou J. (1983) Extinction: lessons from zoos. In: *Genetics and Conservation: a Reference for Managing Wild Animal and Plant Populations* (eds Schonewald-Cox C.M., Chambers S.M., MacBryde B., Thomas W.L.). Benjamin/Cummings, Menlo Park. 164-184 p.
- Rannala B., Mountain J.L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*. 94(17):9197-9201
- Raymond M., Rousset F. (1995a). GENEPOP (4.0.10): population genetics software for exact test and ecumenism. *Journal of Heredity* 86(3):248-249
- Raymond M., Rousset F. (1995b). An exact test for population differentiation. *Evolution* 49(6):1280-1283

- Razgour O., Taggart J.B., Manel S., Juste J., Ibanez C., Rebelo H., Alberdi A., Jones G., Park K. (2018) An integrated framework to identify wildlife populations under threat from climate change. *Molecular Ecology Resources*. 18(1):18-31
- Refsnider J.M. (2009) High frequency of multiple paternity in Blanding's turtle (*Emys blandingii*). *Journal of Herpetology*. 43(1):74-81
- Reid B.N., Mladenoff D.J., Peery M.Z. (2017) Genetic effects of landscape, habitat preference and demography on three co-occurring turtle species. *Molecular Ecology* 26(3):781-798
- Reynolds J.D. (1996) Animal breeding systems. *Trends in Ecology and Evolution*. 11:68-73
- Rhodin A.G.J., Stanford C.B., Van Dijk P.P, Eisemberg C., Luiselli L., Mittermeier R.A., Hudson R., Horne B.D., Goode E. V., Kuchling G., Walde A., Baard E.H.W., Berry K.H., Bertolero A. ... Vogt R.C. (2018) Global conservation status of turtles and tortoises (order Testudines). *Chelonian Conservation and Biology*.17:135-161
- Rice W.R. (1989) Analyzing tables of statistical tests. *Evolution*. 43(1):223-225
- Ricklefs R.E., Miller G.L. (2000) *Ecology* (4th ed). Freeman, New York. 833 p.
- Rioux Paquette S., Behncke S.M., O'Brien S.H., Brenneman R.A., Louis E.E., Lapointe F.-J. (2007) Riverbeds demarcate distinct conservation units of the radiated tortoise (*Geochelone radiata*) in southern Madagascar. *Conservation Genetics*. 8:797-807
- Robillard A.J., Robinson S., Bastiaans E., Vogler D. (2019) Impacts of a highway on the population genetic structure of a threatened freshwater turtle (*Glyptemys insculpta*). *Amphibian and Reptile Conservation*. 13(2):267-275
- Rogic A., Tessier N., Lapointe F.-J., Millien V. 2013 Genetic structure of the white-footed mouse in the context of the emergence of Lyme disease in southern Québec. *Ecology and Evolution*. 3(7):2075-2088
- Roques S., Díaz-Paniagua C., Portheault A., Pérez-Santigosa N., Hidalgo-Vila J. (2006) Sperm storage and low incidence of multiple paternity in the European pond turtle, *Emys orbicularis*: a secure but costly strategy? *Biological Conservation*. 129(2):236-243
- Rosen G.E., Smith K.F. (2010) Summarizing the evidence on the international trade in illegal wildlife. *EcoHealth*. 7(1):24-32
- Rousset F. (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*. 8(1):103-106
- Rousset F., Ferdy J.-B. (2014) Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography*. 37(8):781-790
- Rowe L., Arnqvist G., Sih A., Krupa J. (1994) Sexual conflict and the evolutionary ecology of mating patterns: water striders as a model system. *Trends in Ecology and Evolution*. 9(8):289-293
- Rozenfeld A.F., Arnaud-Haond S., Hernández-García E., Eguíluz V.M., Serrão E.A., Duarte C.M. (2008) Network analysis identifies weak and strong links in a metapopulation system. *Proceedings of the National Academy of Sciences of the United States of America*. 105(48):18824-18829

- Rudnick D., Ryan S.J., Beier P., Cushman S.A., Dieffenbach F., Epps C., Gerber L.R., Hartter J., Jenness J., Kintsch J., Merenlender A.M., Perki R.M., Perziosi D.V., Trombulack S.C. (2012) The role of landscape connectivity in planning and implementing conservation and restoration priorities. *Issues in Ecology*. Report No. 16. Ecological Society of America. Washington, DC. 1-20 p.
- Saumure R.A., Bider J.R. (1998) Impact of agriculture development on a population of wood turtles (*Clemmys insculpta*) in southern Québec, Canada. *Chelonian Conservation and Biology*. 3:37-45
- Saumure R.A., Herman T.B., Titman R.D. (2007) Effects of haying and agricultural practices on a declining species: the North American wood turtle, *Glyptemys insculpta*. *Biological Conservation*. 135(4):565-575
- Schonewald-Cox C. (1983) Guidelines to management: a beginning attempt. In: *Genetics and Conservation: a Reference for Managing Wild Animal and Plant Populations* (eds Schonewald-Cox C.M., Chambers S.M., MacBryde B., Thomas W.L.). Benjamin/Cummings, Menlo Park. 414-445 p.
- Schwartz M.K., McKelvey K.S. (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*. 10:441-452
- Scribner K.T., Congdon J.D., Chesser R.K., Smith M.H. (1993) Annual differences in female reproductive success affect spatial and cohort-specific genotypic heterogeneity in painted turtles. *Evolution*. 47(5):1360-1373
- Seftic K.M., Koblmüller S. (2009) Assessing parent numbers from offspring genotypes: the importance of marker polymorphism. *Journal of Heredity*. 100:197-205
- Segelbacher G., Cushman S.A., Epperson B.K., Fortin M.-J., François O., Hardy O.J., Holderegger R., Taberlet P., Waits L.P., Manel S. (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*. 11:375-385
- Shaffer H.B., Gidiş M., McCartney-Melstad E., Neal K.M., Oyamaguchi H.M., Tellez M., Toffelmier E.M. (2015) Conservation genetics and genomics of amphibians and reptiles. *Annual Review of Animal Biosciences*. 3(1):113-138
- Slatkin M. (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*. 139:457-462
- Sorenson L., McDowell J.R., Knott T., Graves J.E. (2013) Assignment test method using hypervariable markers for blue marlin (*Makaira nigricans*) stock identification. *Conservation Genetics Resources*. 5:293-297
- Soulé M.E. (1985) What is conservation biology? *BioScience*. 35(11):727-734
- Soulé, M.E. (1987) *Viable populations for conservation*. Cambridge University Press, Cambridge. 189 p.
- Soulé M.E. (1991) Conservation: tactics for a constant crisis. *Science* 253:744-750
- Soulé M.E., Wilcox B.A. (1980) *Conservation biology: an evolutionary-ecological perspective*. Sinauer Associates, Sunderland, Mass. 395 p.



- Spielman D., Brook B.W., Frankham R. (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America*. 101(42):15261-15264
- Spradling T.A., Tamplin J.W., Dow S.S., Meyer K.J. (2010) Conservation genetics of a peripherally isolated population of the wood turtle (*Glyptemys insculpta*) in Iowa. *Conservation Genetics* 11(5):1667-1677
- Stallings R.L., Ford A.F., Nelson D., Torney D.C., Hildebrand C.E., Moysis R.K. (1991) Evolution and distribution of (GT)<sub>n</sub> repetitive sequences in mammalian genomes. *Genomics*. 10:807-815
- Stewart K.R., Dutton P.H. (2011) Paternal genotype reconstruction reveals multiple paternity and sex ratios in a breeding population of leatherback turtles (*Dermochelys coriacea*). *Conservation Genetics*. 12(4):1101-1113
- Stith B.M., Fitzpatrick J.W., Woolfenden G.E., Pranty B. (1996) Classification and conservation of metapopulations: a case study of the Florida scrub jay. In: *Metapopulations and Wildlife Conservation* (ed McCullough D.R.). Island Press, Washington, D.C. 187-215 p.
- Stockley P. (1997) Sexual conflict resulting from adaptations to sperm competition. *Trends in Ecology and Evolution*. 12(4):154-159
- Storey K.B., Storey J.M., Brooks S.P., Churchill T.A., Brooks R.J. (1988) Hatchling turtles survive freezing during winter hibernation. *Proceedings of the National Academy of Sciences of the United States of America*. 85(21):8350-8354
- Storfer A. (1999) Gene flow and endangered species translocations: a topic revisited. *Biological Conservation*. 87(2):173-180
- Storz J.F., Beaumont M.A. (2002) Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution*. 56(1):154-166
- Sugg D.W., Chesser R.K. (1994) Effective population sizes with multiple mating. *Genetics*. 137:1147-1155
- Sunnucks P. (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution*. 15(5):199-203
- Sunderland T., Sunderland-Groves J., Shanley P., Campbell B. (2009) Bridging the gap: how can information access and exchange between conservation biologists and field practitioners be improved for better conservation outcomes? *Biotropica*. 41(5):549-554
- Taberlet P., Luikart G. (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*. 68(1-2):41-55
- Taylor M.L., Price T.A.R., Wedell N. (2014) Polyandry in nature: a global analysis. *Trends in Ecology and Evolution*. 29(7):376-383
- Templeton A.R., Shaw K., Routman E., Davis S.K. (1990) The genetics consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden*. 77(1):12-27

- Tessier N., Lapointe F.-J. (2002) Caractérisation génétique et conservation des populations de la tortue des bois (*Clemmys insculpta*) au Québec. Université de Montréal, Département de sciences biologiques, Montréal. 46 p.
- Tessier N., Rioux Paquette S., Lapointe F.-J. (2005) Conservation genetics of the wood turtle (*Glyptemys insculpta*) in Québec, Canada. *Canadian Journal of Zoology*. 83(6):765-772
- Theissinger K., FitzSimmons N.N., Limpus C.J., Parmenter C.J., Phillott A.D. (2009) Mating system, multiple paternity and effective population size in the endemic flatback turtle (*Natator depressus*) in Australia. *Conservation Genetics*. 10(2):329-346
- Thornhill R., Alcock J. (1983) The evolution of insect mating systems. Harvard University Press, Cambridge. 547 p.
- Traversy N., McNicoll R., Lemieux R. (1989) Les populations de rats laveurs du sud-ouest du Québec. Direction gestion des espèces et des habitats, ministère du Loisir, de la Chasse et de la Pêche du Québec. 114 p.
- Tregenza T., Wedell N. (2000) Genetic compatibility, mate choice and patterns of parentage: invited review. *Molecular Ecology*. 9(8):1013-1027
- Treml E.A., Halpin P.N., Urban D.L., Pratson L.F. (2008) Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology*. 23(1):19-36
- Trivers R.L. (1972) Parental investment and sexual selection. In: *Sexual Selection and the Descent of Man* (ed Trivers R.L.). Aldine Transaction, Chicago. 136-179 p.
- Turtle Conservation Fund (2002) A global action plan for conservation of tortoises and freshwater turtles. Strategy and Funding Prospectus 2002-2007. Washington DC: Conservation International and Chelonian Research Foundation. 34 p.
- Uller T., Olsson M. (2008) Multiple paternity in reptiles: patterns and processes. *Molecular Ecology*. 17(11):2566-2580
- Urban D., Keitt T. (2001) Landscape connectivity: a graph-theoretic perspective. *Ecology*. 82(5):1205-1218
- Väli Ü., Einarsson A., Waits L., Ellegren H. (2008) To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology*. 17(17):3808-3817
- van Dijk P.P., Harding J. (2011) *Glyptemys insculpta*. (errata version published in 2016) The IUCN Red List of Threatened Species 2011: e.T4965A97416259. Downloaded on 15 June 2017.
- van Dijk P.P., Harding L. (2013) *Glyptemys insculpta*. The IUCN Red List of Threatened Species Version 2013: e.T4965A11102820. Available at <http://www.iucnredlist.org> Archived by WebCite at <http://www.webcitation.org/6hbTlh1Cw> on 18 may 2016.
- Van Dijk K.J., Bricker E., van Tussenbroek B.I., Waycott M. (2018) Range-wide population genetic structure of the Caribbean marine angiosperm *Thalassia testudinum*. *Ecology and Evolution*. 8(18):9478-9490



- Van Oosterhout C., Hutchinson W.F., Wills D.P., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*. 4(3):535-53
- Van Oosterhout C., Weetman D., Hutchinson W.F. (2006) Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes*. 6(1):255-256
- van Strien M.J. (2017) Consequences of population topology for studying gene flow using link-based landscape genetic methods. *Ecology and Evolution*. 7(14):5070-5081
- Vargas-Ramirez M., Stuckas H., Castaño-Mora O.V., Fritz U. (2012) Extremely low genetic diversity and weak population differentiation in the endangered Colombian river turtle *Podocnemis lenyana* (Testudines: Podocnemididae). *Conservation Genetics*. 13:65-77
- Veliz D., Duchesne P., Rojas-Hernandez N., Pardo L.M. (2017) Statistical power to detect multiple paternity in populations of highly fertile species: how many females and how many offspring should be sampled? *Behavioral Ecology and Sociobiology* 71(1):1-6
- Vogt R.C. (1981) Food partitioning in three sympatric species of map turtle genus *Graptemys* (Testudinata, Emydidae). *The American Midland Naturalist*. 105(1):102-110
- Volkman L., Martyn I., Moulton V., Spillner A., Mooers A.O. (2014) Prioritizing populations for conservation using phylogenetic networks. *PLoS One*. 9(2):e88945
- Wade M.J., McCauley D.E. (1988) Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution*. 42:995-1005
- Waits L.P., Luikart G., Taberlet P. (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*. 10(1):249-256
- Walde A.D. (1998) Ecology of the wood turtle, *Clemmys insculpta*, Québec, Canada. MSc thesis, Department of Natural Resource Science, McGill University, Montréal, Québec, Canada. 95 p.
- Walde A.D., Bider J.R., Daigle C., Masse D., Bourgeois J.-C., Jutras J., Titman R.D. (2003) Ecological aspects of a Wood Turtle, *Glyptemys insculpta*, population at the northern limit of its range in Québec. *Canadian Field Naturalist*. 117(3):377-388
- Walde A.D., Bider R.J., Masse D., Saumure R.A., Titman R.D. (2007) Nesting ecology and hatching success of the wood turtle, *Glyptemys insculpta*, in Québec. *Herpetological Conservation and Biology*. 2(1):49-60
- Watson P.J. (1991) Multiple paternity as genetic bet-hedging in female sierra dome spiders, *Linyphia litigiosa* (Linyphiidae). *Animal Behaviour*. 41(2):343-360
- Wayne R.K., Morin P.A. (2004) Conservation genetics in the new molecular age. *Frontiers in Ecology and the Environment*. 2(2):89-97
- Westneat D.F., Frederick P.C., Wiley R.H. (1987) The use of genetic markers to estimate the frequency of successful alternative reproductive tactics. *Behavioral Ecology and Sociobiology*. 21(1):35-45

- Willoughby J.R., Sundaram M., Lewis T.L., Swanson B.J. (2013) Population decline in a long-lived species: the Wood Turtle in Michigan. *Herpetologica* 69(2):186-198
- Willoughby J.R., Sundaram M., Wijayawardena B.K., Kimble S.J., Ji Y., Fernandez N.B., Antonides J.D., Lamb M.C., Marra N.J., DeWoody J.A. (2015) The reduction of genetic diversity in threatened vertebrates and new recommendations regarding IUCN conservation rankings. *Biological Conservation*. 191:495-503
- Wood C.C., Gross M.R. (2008) Elemental conservation units: communicating extinction risk without dictating targets for protection. *Conservation Biology*. 22 36-47
- Woodruff D.S. (2001) Populations, species, and conservation genetics. *Encyclopedia of biodiversity*. 811-829.
- Wright S. (1943) Isolation by distance. *Genetics*. 28(2):114
- Yasui Y. (1998) The “genetic benefits” of female multiple mating reconsidered. *Trends in Ecology and Evolution*. 13(6):246-250
- Yasui Y. (2001) Female multiple mating as a genetic bet-hedging strategy when mate choice criteria are unreliable. *Ecological Research*. 16(4):605-616
- Zeh J.A., Zeh D.W. (1996) The evolution of polyandry I: intragenomic conflict and genetic incompatibility. *Proceedings of the Royal Society of London B: Biological Sciences*. 263:1711-1717
- Zellmer A.J., Knowles L.L. (2009) Disentangling the effects of historic vs. contemporary landscape structure on population genetic divergence. *Molecular Ecology*. 18:3593-3602
- Zero V.H., Barocas A., Jochimsen D.M., Pelletier A., Giroux-Bougard X., Trumbo D.R., Castillo J.A., Mack D.E., Linnell M.A., Pigg R.M., Hoisington-Lopez J., Spear S.F., Murphy M.A., Waits L.P. (2017) Complementary network-based approaches for exploring genetic structure and functional connectivity in two vulnerable, endemic ground squirrels. *Frontiers in Genetics*. 8(81):1-13

# **Annexe I. Analyses forensiques**

**Comparaison des profils génétiques de deux tortues des bois (*Glyptemys insculpta*) saisies aux douanes de Hong Kong avec les populations du Québec**

Cindy Bouchard, Nathalie Tessier, François-Joseph Lapointe

Département de sciences biologiques, Université de Montréal, Montréal, Canada

Rapport technique présenté à Environnement Canada  
Juillet 2013

*Format adapté pour les besoins de la thèse*

Suite à la demande d'Environnement Canada, deux échantillons de sang prélevés sur chacune des tortues des bois interceptées à Hong Kong ont été utilisés pour vérifier leur appartenance aux populations de tortue des bois du Québec. Les échantillons de sang, prélevés à Hong Kong, étaient séchés sur un papier buvard et préservés dans un sac de plastique avec des sachets de billes déshydratantes (Figure AI.1). Les analyses génétiques ont eu lieu au Laboratoire d'écologie moléculaire et d'évolution du Professeur François-Joseph Lapointe à l'Université de Montréal.



**Figure AI.1** Échantillons de sang tels que reçus à l'Université de Montréal pour chacune des tortues des bois

### **AI.1 Procédure de décontamination du matériel**

Plusieurs mesures ont été prises avant les manipulations afin d'éviter la contamination des échantillons. Ainsi, tout le matériel jetable (tubes et embouts) a été stérilisé à l'autoclave avant leur utilisation. De plus, les paillasses de laboratoire et les pipettes ont été désinfectées avec de l'éthanol 95% avant chaque manipulation et les pipettes ont été traitées sous lumière UV pendant 20 minutes, pour détruire toute trace d'ADN. Des gants étaient portés en tout temps par le personnel du laboratoire.

## AI.2 Extraction et amplification de l'ADN

L'ADN a été extrait à l'aide du *DNeasy blood and tissue kit* (Qiagen) en suivant le protocole fourni par la compagnie pour le sang séché. Seulement la moitié de la goutte de sang fut utilisée. Le papier a été découpé en plusieurs morceaux et déposé dans un tube 1,5 ml. La digestion s'est faite dans un volume total de 360 µl pour une durée d'une heure à 55°C. Deux éluions d'ADN ont été obtenues, soit pour un volume de 150 µl et 75 µl de tampon, afin d'obtenir le maximum d'ADN.

Le profil génétique des deux individus a été obtenu à l'aide de locus microsatellites créés pour la tortue de Muhlenberg (*Glyptemys mublenbergii*) (King & Julian 2004). Huit marqueurs ont été sélectionnés; GmuD16, GmuB21, GmuD28, GmuD40, GmuD55, GmuD87, GmuD88 et GmuD90. Ces différents marqueurs ont déjà été utilisés par le passé pour caractériser génétiquement les populations de tortue des bois du Québec (Tessier et al. 2005). Une fois l'amplification des marqueurs génétiques obtenue, le produit RPC a été envoyé à Génome Québec pour être analysé par un séquenceur automatique de type ABI-3730X. La taille des allèles a ensuite été déterminée avec le logiciel GENEMARKER 2.2.0.

## AI.3 Populations analysées

Afin de déterminer si les tortues interceptées aux douanes de Hong Kong proviennent du Québec, leur profil génétique a été comparé à celui de neuf populations de la province (Tableau AI.1, Figure AI.2)

**Figure AI.2** La figure fut retirée puisqu'il s'agit de données sensibles

**Tableau AI.1** Code des populations, nombre d'individus utilisés et nombre d'allèles total obtenu pour les différentes populations du Québec utilisées lors des analyses.

Code	Nb d'individus	Nb d'allèles
FA	36	55
GA	14	32
SH	25	51
RP	24	41
DC	18	62
MA	31	70
SU	7	42
TOM	6	48
MI	11	56

## Analyses statistiques

Dans un premier temps, certaines analyses ont été effectuées afin de s'assurer que les conditions requises pour effectuer les tests d'assignation étaient respectées. Il s'agissait de vérifier l'équilibre d'Hardy-Weinberg et le déséquilibre de liaison pour chacun des locus. Des tests exacts de Fisher (Raymond & Rousset 1995a) ont également été calculés. Une différence statistique significative entre les paires de populations indique qu'elles sont génétiquement distinctes. Ainsi, elles pourront être considérées comme indépendantes lors des analyses suivantes. Tous ces tests statistiques ont été réalisés à l'aide du logiciel GENEPOP version 3.4 (Raymond & Rousset 1995b).

Un test d'assignation a permis de calculer la probabilité selon laquelle les individus de Hong Kong pourraient appartenir aux populations québécoises de tortues des bois. Les analyses ont été effectuées à l'aide du logiciel GENECLASS2 (Piry et al. 2004) selon un algorithme bayésien partiel de Rannala & Mountain (1997). La méthode de rééchantillonnage de Paetkau et al. (2004) a été utilisée avec 10 000 génotypes simulés pour chaque population, tels que mentionnés dans Manel et al. (2002). La méthode de rééchantillonnage permet une plus grande précision et une erreur plus faible lors des analyses. Le génotype des deux individus de Hong Kong a été comparé à chacune des populations de manière indépendante. Ainsi, il est

possible que l'individu sélectionné ne soit affecté à aucune des populations présentes dans les analyses. Dans ce cas, l'individu montrera une faible probabilité d'appartenir à l'une ou l'autre des populations présentes. La méthode utilisée est plus précisément un test d'exclusion. Les tests d'exclusion présentent un avantage certain puisqu'il n'est pas nécessaire que la population d'origine des individus se retrouve dans les données pour que les analyses soient valables (Manel et al. 2002, Sorenson et al. 2013).

Afin de connaître l'efficacité et la précision du test pour nos données, 14 individus sélectionnés au hasard et provenant des différentes populations du Québec ont été retirés de leur population d'origine, afin d'être testés pour leur appartenance. Ainsi les valeurs obtenues lors de l'assignation de ces individus nous indiquent l'étendue des résultats possible lorsqu'un individu est reclassé dans la bonne population. Les deux individus provenant de Hong Kong ont ensuite été testés avec chacune des populations du Québec et les valeurs de probabilités obtenues ont été comparées aux valeurs obtenues précédemment par les individus du Québec avec leur population d'origine.

De plus, les relations entre les populations ont été illustrées en calculant les distances génétiques. La distance  $D_{CE}$  (Cavalli-Sforza & Edward 1967) tient compte de la variance allélique. Des matrices de distances entre les populations sont obtenues et représentées sous la forme d'un arbre phylogénétique. Les estimations de  $D_{CE}$  et les valeurs de bootstrap sur chaque branche de l'arbre ont été obtenues à l'aide du logiciel Populations 3.5 (Langella 1999). Les deux individus de Hong Kong ont été groupés en une population distincte.

## **AI.4 Résultats et conclusion**

Les locus montrent un bon polymorphisme variant de 5 à 24 allèles. Les microsatellites sont en équilibre d'Hardy-Weinberg et aucun déséquilibre de liaison n'est observé. Deux populations n'ont pas démontré de différence significative entre elles dans leur profil génétique lors des tests exacts de Fisher. Les populations SU et TOM ont donc été regroupées pour les

tests d'assignation. Les tests d'assignation des individus du Québec montrent un bon succès de reclassement, dont les valeurs de probabilité se situent entre 0,568 et 0,979 pour une valeur moyenne de 0,8 (Tableau AI.2). Les tests d'assignation des individus de Hong Kong montrent quant à eux de très faibles valeurs d'appartenance variant de 0 à 0,165 (Tableau AI.3). On peut donc penser que ces deux tortues ne sont pas originaires du Québec, même si ce n'est pas l'ensemble des populations présentes au Québec qui ont fait l'objet d'une caractérisation génétique.

L'arbre phylogénétique montre clairement une discrimination entre les populations de tortue des bois de la rive nord et de la rive sud du fleuve Saint-Laurent. On note toutefois que les embranchements sont faiblement supportés par de petites valeurs de bootstrap pour les populations situées sur la rive sud du fleuve Saint-Laurent. Ces résultats ne permettent pas de conclure sur l'origine des individus interceptés aux douanes, mais il est intéressant de noter une certaine proximité génétique entre les individus retrouvés à Hong Kong et la population de Sutton ou Tomifobia (SU-TOM), toutes deux situées dans le sud du Québec près de la frontière américaine (Figure AI.3). Dans une prochaine étape d'identification de la provenance de ces deux individus, nous devrions privilégier les comparaisons avec les populations de tortues des bois situées aux États-Unis comme dans les états du Vermont et de New York. On ne peut toutefois exclure l'appariement de ces deux individus avec une population du sud de l'Ontario, car aucune comparaison génétique n'a été réalisée. De plus, on ne peut déterminer si les deux individus proviennent d'une même population ou de populations différentes.

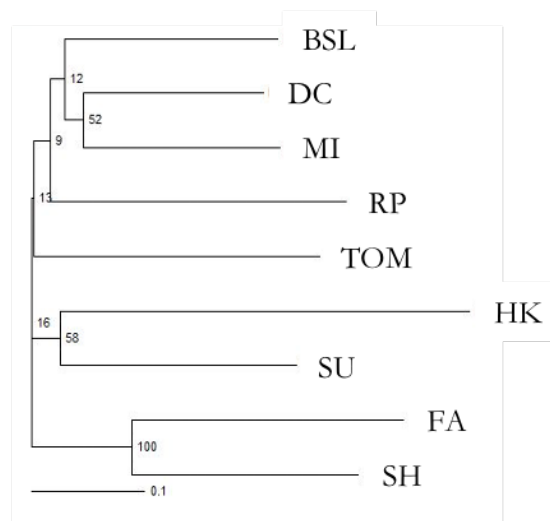
**Tableau AI.2** Valeurs de probabilités obtenues lors de l'assignation des individus du Québec à leur population d'origine.

<b>Individus</b>	<b>Valeur de probabilité</b>	<b>Individus</b>	<b>Valeur de probabilité</b>
FA-20	0.820	RP-16	0.907
FA-25	0.981	DC-13	0.979
NO-5	0.756	DC-19	0.599
NO-9	0.599	BSL-7	0.715
SH-94	0.714	BSL-8	0.932
SH-176	0.714	SU-110	0.974
RP-9	0.568	MI-23	0.940



**Tableau AI.3** Valeurs des probabilités pour les individus de Hong Kong lors de l'assignation à chacune des populations du Québec.

Individus	FA	NO	SH	RP	DC	BSL	SU-TOM	MI
<b>K6052</b>	0	0.001	0	0.003	0.165	0.05	0.159	0.005
<b>K6053</b>	0	0	0	0.002	0.035	0.018	0.096	0.038



**Figure AI.3** Arbre phylogénétique construit à partir de distances de  $D_{CE}$ . Les valeurs de bootstrap sont inscrites le long des branches.

**Tableau AI. A** Profil génétique des individus saisis au poste de douane de Hong Kong

Individu	GmuB21	GmuD87	GmuD40	GmuD88	GmuD28	GmuD16	GmuD55	GmuD90
<b>K6052</b>	178178	210214	212216	170184	244264	148156	110110	194202
<b>K6053</b>	182186	196210	196212	150162	240264	148156	110110	194198

## AI.4 Bibliographie

- Cavalli-Sforza L.L., Edwards W.F. (1967) Phylogenetics analysis: models and estimation procedures. *Evolution*. 3:550-570
- King T.L., Julian S.E. (2004) Conservation of microsatellite DNA flanking sequence across 13 Emydid genera assayed with novel bog turtle (*Glyptemys muhlenbergii*) loci. *Conservation Genetics*. 5(5):719-725
- Langella O. (2002) Population 1.2.30. Copyright (C) 1999, Olivier Langella, CNRS-UPR9034. Disponible à: <http://bioinformatics.org/~tryphon/populations/>
- Manel S., Berthier P., Luikart G. (2002) Detecting wildlife poaching: identifying the origin of individuals with bayesian assignment tests and multilocus genotypes. *Conservation Biology*. 16(3):650-659
- Paetkau D., Slade R., Burden M., Estoup A. (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation based exploration of accuracy and power. *Molecular Ecology*. 13:55-65
- Piry S., Alapetite A., Cornuet J.-M., Paetkau D., Baudoin L. et Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*. 95:536-539
- Rannala B., Mountain J.L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*. 94(17):9197-9201
- Raymond M., Rousset F. (1995a). An exact test for population differentiation. *Evolution* 49(6):1280-1283
- Raymond M., Rousset F. (1995b). GENETPOP (4.0.10): population genetics software for exact test and ecumenism. *Journal of Heredity* 86(3):248-249
- Sorenson L., McDowell J.R., Knott T., Graves J.E. (2013) Assignment test method using hypervariable markers for blue marlin (*Makaira nigricans*) stock identification. *Conservation Genetics Resources*. 5:293-297
- Tessier N., Rioux Paquette S., Lapointe F.-J. (2005) Conservation genetics of the wood turtle (*Glyptemys insculpta*) in Québec, Canada. *Canadian Journal of Zoology*. 83(6):765-772