COMPARISON OF THE DNA METHYLATION PATTERNS
BETWEEN INTERSPECIFIC *XENOPUS* F1
HYBRIDS AND THEIR PARENTAL
SPECIES

by

ALIE PATRICK KOROMA

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON
December 2010
DEDICATION

This work is dedicated to my four daughters Alizatu Koroma (15), Janet Isatu Koroma (11), Aminata Paulina Koroma (7) and Abibatu Patricia Koroma (7) and, above all, to my loving wife of 16 years Zainabu Koroma
ACKNOWLEDGEMENTS

I am greatly indebted to a lot of people for all the assistant they have provided in one way or another to make this dissertation a success.

I would first like to extend my sincere appreciation to my supervising committee: Drs Paul Chippindale, Pawel Michalak, Esther Betráan, Ellen Pritham and James Robinson. I would like to extend special thanks to Paul Chippindale for stepping up and taking over as my supervising Professor after Pawel Michalak had to leave for another appointment. I would also like to say thank you very much to Pawel Michalak, my initial supervisor, for his guidance and help in funding this research. To the rest of my supervising committee I appreciate your constructive criticisms, help and suggestions on how I should make my work and dissertation document better.

I also want to thank the following people: John Malone, who together with Pawel Michalak, bred the hybrids used in this study; Robert Makowsky and Brian Fontenot for their assistant with the AFLP protocol; Clement Gilbert for his assistance with my effort to produce a chromosome map for the F1 hybrids, Jesse Meik, my office mate, for being a friend and accepting me the way I am and Solomon Macauley for his help with working with Excel spreadsheet to analyze my AFLP data. In addition, I want to express my gratitude to all graduate students past and present who have in one way or another help to make my stay at UTA a memorable one.
I sincerely extend my appreciation to my lab mates: Mersee Madison-Villar and Daina Ma for all the time we have spent together in the lab.

To the Biology Department front office staff Linda, Gloria, Peggy and Paulette I am lost of words on what to say. You have been so helpful and very patient with me. Even though thank you is not enough to express the way I feel, I will say thank you very much.

I would also like to thank the Biology Department faculty especially Drs Campbell, Badon, Formanowicz, and Watson for all their help.

To the staff at the core facility, Amy, Bob and Ray I say thank you. Ray all those extra hours you spent with me making sure I get things right I assure you did not go in vain. I appreciate all of them.

I am very grateful to my parents for teaching me the value of education.

To my four daughters, for all those hours that I spent away from you trying to get things done I am very grateful for your understanding. I love you and appreciate your time.

Behind every successful man there is always a strong, hardworking and loving woman. That honor goes to my wife of 16 years. I don’t know what I will have done without you. Thank you, thank you for your encouragement, support and love.

December 3 2010
ABSTRACT

COMPARISON OF THE DNA METHYLATION PATTERN
BETWEEN INTERSPECIFIC XENOPUS F1 HYBRIDS
AND THEIR PARENTAL SPECIES

Alie Patrick Koroma, Ph.D.
The University of Texas at Arlington, 2010

Supervising Professors: Dr. Paul Chippindale, Dr. Pawel Michalak

In this study the main focus is to compare the methylation patterns of interspecific Xenopus F1 hybrids to the parental species using Methyl-Sensitive Amplification Polymorphisms (MSAPs). Genomic DNA was extracted from the liver and muscle tissues from a representative sample of F1 hybrids and the parental species (Xenopus laevis and X. muelleri).

First, the MSAP markers were very effective in comparing the methylation patterns of hybrids and the parental species. Genetic analysis of 504 liver and 364 muscle MSAP markers revealed that these markers were highly polymorphic. Principal coordinate analysis showed four distinct clusters with the two parental species separate and the F1s in between with respect to levels and patterns of methylation. Most of the
variance among clusters can be explained by within cluster (population) variation. There was more variation in the liver tissues.

Furthermore, Nei’s genetic distances revealed more dissimilarity between the clusters with the muscle tissues compared to liver.

Also, despite the fact that at least 50% of the MSAP markers in the F1 hybrids are derived from the parental species, there is some loss that is biased to both parental markers (i.e. most of the parental fragments not detected in the F1 hybrids come from either the paternal or maternal species depending on the sex of the F1). Fisher’s exact test indicates the proportion of the loss of diagnostic parental fragments in the F1 is significant for the muscle tissue. The hybrids also exhibited unique MSAP markers.

The average fraction of methylated MSAP fragments generated in this study ranges from 70.1% to 78.8% (75.0 ± 4; mean± SD) and 67.0% to 73.8% (69.1 ± 3.2; mean± SD) for liver and muscle tissues respectively.

Furthermore genome sizes of F1 hybrids were intermediate between those of the parents.

Finally undermethylation was observed in liver tissues of male F1 hybrid.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS........................................................................ iv

ABSTRACT......................................................................................... vi

LIST OF FIGURES................................................................................ ix

LIST OF TABLES.................................................................................. xiii

Chapter

1. INTRODUCTION.............................................................................. 1

   Epigenetics and DNA methylation....................................................... 1

   Genome duplication........................................................................... 2

   Conclusion......................................................................................... 4

   *Xenopus*........................................................................................ 6

2. METHYLATION PATTERNS AMONG HYBRIDS AND PARENTS (*Xenopus laevis* and *Xenopus muelleri*) USING AMPLIFIED FRAGMENT LENGTH POLYMORPHISM........... 8

   Introduction..................................................................................... 8

   Methods.......................................................................................... 13

      Specimens.................................................................................... 13

      DNA Extraction........................................................................... 13

      Methylation-Sensitive Amplification Polymorphism (MSAP).... 14

   Data quality/ genotyping error......................................................... 18

   Statistical Analysis.......................................................................... 19

   Analysis of Molecular Variance....................................................... 19
Cluster Analysis/Population Structure………………………….20

Results……………………………………………………………………20

Molecular Markers……………………………………………………20

Cluster Analysis/Population Structure……………………………23

Analysis of Molecular Variance……………………………………27

Genetic Diversity……………………………………………………30

Comparison of MASP methylated fragments between parental and F1 hybrids…………………………………………..34

Discussion……………………………………………………………….43

Population Structure…………………………………………………..43

Genetic Diversity……………………………………………………43

Comparison of MSAP methylated fragments between parental species and F1 hybrids………………………………….45

3. COMPARISON OF GENOME SIZE OF HYBRIDS TO PARENTAL SPECIES (X. l. laevis and X. muelleri) USING FLOW CYTOMETRY……………………………………………………………..49

Methods………………………………………………………………52

Compare the genome size of the parents and the hybrids……….52

Digestion of genomic DNA of the hybrids and the parents ……53

Results………………………………………………………………54

Genome Size…………………………………………………………54

Analysis of genome methylation………………………………….62

Discussion……………………………………………………………..71
Comparison of genome sizes between parental species and their hybrids ........................................71

Genome Methylation.................................................................71

4. DISCUSSION............................................................................73

APPENDICES

   A. CHECKING MSAP FRAGMENT DATA QUALITY ......................79

   B. FORMULAS USED TO PERFORM STATISTICAL CALCULATIONS...............................................82

REFERENCES...............................................................................85

BIOGRAPHICAL INFORMATION......................................................102
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Diagrammatic representation of the AFLP Technique</td>
<td>17</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary graph showing the results of methylated band (alleles) patterns across populations [a) liver and b) muscle tissues]</td>
<td>22</td>
</tr>
<tr>
<td>2.3</td>
<td>Principal Coordinates Analysis (PCoA) plot of MSAFLP for all individual pair-wise genetic distances [liver tissue] - PCoA via Covariance with data standardization</td>
<td>25</td>
</tr>
<tr>
<td>2.4</td>
<td>Principal Coordinates Analysis (PCoA) plot of MSAFLP for all individual pair-wise genetic distances [muscle tissue] - PCoA via Covariance with data standardization</td>
<td>26</td>
</tr>
<tr>
<td>2.5</td>
<td>Pie Chart summarizing the outcome of hierarchical AMOVA, showing the partitioning of total molecular variance within and among all population [a) liver and b) muscle tissues]</td>
<td>29</td>
</tr>
<tr>
<td>2.6</td>
<td>Frequency distribution of random permuted PhiPT (φPT) versus observed PhiPT (φPT) for 9999 Permutations – input as genetic distance matrix for calculation of PhiPT (φPT) [a) liver and b) muscle tissues]</td>
<td>33</td>
</tr>
<tr>
<td>2.7</td>
<td>Liver MSAP fragments observed in individuals from <em>X. laevis, X. muelleri</em> and F1 hybrid males and females</td>
<td>39</td>
</tr>
<tr>
<td>2.8</td>
<td>Muscle MSAP fragments observed in individuals from <em>X. laevis, X. muelleri</em> and F1 hybrid males and females</td>
<td>40</td>
</tr>
<tr>
<td>3.1</td>
<td>Histogram of the mean PI-A measurements</td>
<td>57</td>
</tr>
<tr>
<td>3.2</td>
<td>Flow cytometric histograms of fluorescence intensity for a) <em>X. laevis</em> and b) <em>X. muelleri</em></td>
<td>58</td>
</tr>
<tr>
<td>3.3</td>
<td>Flow cytometric histograms of fluorescence intensity for a) F1 hybrid and b) <em>X. laevis/muelleri</em></td>
<td>59</td>
</tr>
</tbody>
</table>
3.4 Flow cytometric histograms of fluorescence intensity for a) *X. laevis*/F1 hybrid and b) *X. muelleri*/F1 Hybrid…………………60
3.5 Flow cytometric histograms of fluorescence intensity for *X. laevis/muelleri* F1 hybrid……………………………………………….61
3.6 Gel electrophoresis for a) gDNA and b) restriction digest products……………………………………………………………………..64
3.7 Electrophenogram for male muscle tissues (*X. laevis*).........................65
3.8 Electrophenogram for male muscle tissues (*X. muelleri*).....................66
3.9 Electrophenogram for male muscle tissues (F1 hybrid).........................67
3.10 Electrophenogram for male liver tissues (*X. laevis*)............................68
3.11 Electrophenogram for male liver tissues (*X. muelleri*).......................69
3.12 Electrophenogram for male liver tissues (F1 hybrid)...........................70
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Descriptive statistics of AFLP loci from liver and muscle tissues [band frequencies, estimated allele frequencies and estimated heterozygosity by population for binary (diploid) data]</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary statistics for Principal Coordinate Analysis (PCoA) for all individuals [liver and muscle tissues]</td>
<td>24</td>
</tr>
<tr>
<td>2.3</td>
<td>Results of AMOVA performed on all populations identified as clusters [liver and muscle tissues]</td>
<td>28</td>
</tr>
<tr>
<td>2.4</td>
<td>Results of pair-wise population PhiPT ((\phi PT)) and LinPhiPT Values and Estimates of Nm [liver and muscle] -input as genetic distance matrix for calculation of PhiPT ((\phi PT))</td>
<td>31</td>
</tr>
<tr>
<td>2.5</td>
<td>Results of pair-wise population unbiased Nei genetic distance and unbiased Nei genetic identity values as [liver and muscle tissues]</td>
<td>32</td>
</tr>
<tr>
<td>2.6</td>
<td>Primers used in MSAP</td>
<td>36</td>
</tr>
<tr>
<td>2.7</td>
<td>Methylated MSAP fragments observed in parents ((X. laevis) and (X. muelleri)) and hybrids ((X. laevis \times X. muelleri)) generated by EcoRI-HpaII and EcoR1-Msp1 digestion (one selective primer combination) - [liver]</td>
<td>37</td>
</tr>
<tr>
<td>2.8</td>
<td>Methylated MSAP fragments observed in parents ((X. laevis) and (X. muelleri)) and hybrids ((X. laevis \times X. muelleri)) generated by EcoRI-HpaII and EcoR1-Msp1 digestion (one selective primer combination - [muscle Tissues]</td>
<td>41</td>
</tr>
<tr>
<td>2.9</td>
<td>Percentage of Methylated Fragment Liver and Muscle</td>
<td>41</td>
</tr>
</tbody>
</table>
2.10 Fisher’s Exact Test for the proportion of MSAP diagnostic parental fragments present or missing in the F1 for liver and muscle tissue……………………………………………..42

3.1 Mean fluorescence (PI-A) for three different measurements…………..55

3.2 Mean fluorescence measurement for all individuals of the same species combined……………………………………………………55

3.3 Fluorescence (PI-A) for the mixed samples……………………………..56

3.4 Summary of regions identified as DNA (smear) for liver and muscle……………………………………………………………………..63
CHAPTER 1

INTRODUCTION

Epigenetics and DNA methylation

Epigenetic studies have come a long way since Waddington first coined the term for the study of “causal mechanisms” by which “the genes of the genotype bring about phenotypic effects” (Haig 2004). This refers to changes in phenotype which occur without any changes in the DNA sequence. These changes are heritable. Today epigenetics is defined as “the sum of all the alterations to the chromatin template that collectively establish and propagate different patterns of gene expression (transcription) and silencing from the same genome” (Allis et al 2007). Epigenetic changes can occur through several different mechanisms: variegation in Drosophila, histone modification, X chromosome inactivation and DNA methylation.

Cytosine methylation is one of the most important epigenetic mechanisms operating in plant and vertebrate genomes. About 60-90% of CpG dinucleotides in vertebrate DNA are methylated at the 5’ position of cytosine (Tucker 2001). The tendency for cytosine to spontaneously deaminate, increases with methylation (Shen et al 1994). Methylated cytosines deaminate to thymines. CpG motifs occur at only about 1/5 of the expected frequency in bulk DNA, suggesting 5MeC deamination to thymine. CpG-rich regions, known as CG islands, are often found in promoters of house-keeping genes where they typically escape methylation (Bird 2002). Methylation of specific CpGs is a
mechanism of gene silencing critical during differentiation (Bird 1993), inactivation of one X chromosome in female eutherian mammals (Chow et al. 2005), and gene imprinting (Holmgren et al. 2001). Changes in methylation patterns contribute to aging (Issa 2003) and cancer (Jones and Baylin 2002).

**Genome duplication**

Gene and genome duplication provided the genetic raw material for adaptation and diversification of organisms at the base of the tree of life (Doolittle and Brown 1994). The ancestors of jawed vertebrates, teleost fishes (Holland and Garcia-Fernández 1996; Vandepoele et al. 2002), many plants (Masterson 1994), and other eukaryotes such as yeast (Wolfe and Shields 1997) likely underwent whole genome duplication. Genome duplications lead to dramatic molecular and cellular reorganizations, some of which can be deleterious: disruptive effects of nuclear and cell enlargement, the propensity of polyploid mitosis and meiosis to produce aneuploid cells, and epigenetic instability (Comai 2005). In humans, chromosome and therefore gene duplication causes severe disorders such as trisomy, triploid and tetraploid individuals very rarely survive until birth (Scarborough et al. 1984; Lopez Pajares et al. 1990). Duplication of dosage sensitive genes may lead to Pelizaeus-Merzbacher disease (Woodward et al. 2005), type CMT1A of Charcot-Marie-Tooth disease (Lupski et al. 1991), and possibly Rett syndrome (Collins et al. 2004).

Ancient polyploidization appears to have been a critical step in eukaryote evolution, and in some groups polyploidization is responsible for major diversification. In the plant kingdom, it is estimated that about 70% of angiosperms (flowering plants) and
about 95% of pteridophytes (ferns) have undergone at least one episode of polyploidization in their evolutionary history (Masterson 1994; Leitch and Bennett 1997). By duplicating genes, polyploidization increases genetic diversity, plasticity, and heterosis, which all contribute to the adaptive potential of polyploids. Widely distributed and invasive common cordgrass (*Spartina anglica*), a recently formed allopolyploid, provides a striking example of rapid adaptation and evolutionary success, in contrast to its non-invasive parental species (Comai 2005). Thompson et al (1991a, b, c) demonstrated that *S. anglica* exhibits significant morphological plasticity in response to environmental change, indicating major fitness advantage. Another example is provided by Arctic flora, with allopolyploid taxa having been particularly effective at invading newly deglaciated areas, presumably because of their increased vigor and resistance to inbreeding (Brochmann et al. 2004). Finally, genome duplications can be selectively advantageous in coping with parasites. For example, Jackson and Tinsley (2003) suggested that the allopolyploid African clawed frog *Xenopus laevis* originated in response to selective pressures from flatworm parasites.

Can polyploidy also buffer against genomic parasites such as viruses and transposable elements (TEs)? It is tempting to assume that unlike diploid genomes that are susceptible to insertional mutagenesis, polyploid genomes with their gene redundancy are relatively tolerant to the deleterious consequences of transposition (Matzke and Matzke 1998). Extra copies of genes compensate for losses or alterations of expression caused by insertions. Therefore, it is expected that mobile elements will tend to proliferate and be retained in polyploid genomes. Indeed, a rough correlation is observed,
as genomes of vertebrates and polyploid plants are densely populated by TEs, whereas genomes of some invertebrates and the diploid *Arabidopsis thaliana* accumulated fewer TEs (Tweedie et al. 1997; Matzke and Matzke 1998).

Allopolyploidization/hybridization can lead to an explosion of transposable element activity (Pikaard 2001). Potential deleterious effects of transposable elements (TEs) on the fitness of their hosts necessitate the evolution of a genomic “immune” system for transposon control. This is critical in the germline where TE activity can produce a substantial mutational burden that would accumulate from generation to generation. It had been assumed previously that such epigenetic phenomena as nucleosomal chromatin formation and DNA methylation have evolved to regulate host gene expression. However, this view has been challenged recently, as it is becoming clear that in addition of being a mechanism of host gene regulation, the primary role of methylation is to regulate transposon activity (Yoder et al. 1997; Martienssen and Colot 2001; Selker 2004). Indeed, the above-mentioned genomes of vertebrates and polyploid plants are characterized by global methylation, whereas the invertebrates and the diploid *Arabidopsis thaliana* show only fractional methylation (Tweedie et al. 1997; Matzke and Matzke 1998).

Conclusion

Increases in gene number that are correlated with major evolutionary transitions have often resulted from whole-genome duplications. For example, using only a subset of paralogous genes duplicated prior to the fish-tetrapod split and plotting the genomic map positions, Dehal and Boore (2005) observed that the global physical organization from
these maps buttress the phenomenon of two distinct genome duplication events early in vertebrate evolution. Perhaps the most intriguing hypothesis explaining the spectacular success of ancient polyploids is that genome duplication buffers against the harmful effects of transposable elements (TEs) – an abundant class of genetic parasites that reside in the germline (Matzke et al. 1999; Wendel 2000). Extra copies of genes may compensate for losses or alterations of gene expression due to TE transposition and in turn facilitate further TE proliferation within genomes. These dramatic genomic reorganizations incurred through polyploidization and TE spread must be concerted with modifications of epigenetic mechanisms of silencing (DNA methylation) and small noncoding RNA machinery (RNAi and piRNA) - the genomic “immune” system against TEs and viruses.

To gain insights into these dynamics, I used *Xenopus* (clawed frogs) and employed Amplified Fragment Length Polymorphism (AFLP) and genome sizing to characterize patterns of global DNA methylation. Since *Xenopus* represent a wide range of ploidy levels (2n – 12n) and are amenable to additional genome size manipulations, they provide a unique system for investigations into the evolution of CpG methylation as a means of genomic silencing in vertebrates. To suppress expression of gene copies, repetitive DNAs and retrotransposons, CpG methylation must be sensitive to duplicated sequences and thereby responsive to genome duplications. Hence, the following specific aims were followed to advance understanding of how epigenetic machinery reacts when challenged with genome size changes: 1) Compare the methylation pattern between hybrids and the parental species (*X. l. laevis* and *X. muelleri*) and 2) Measure the genome
size of the hybrids of *X. l. laevis* and *X. muelleri* and compare them to those of the parents. I hypothesized that the methylation pattern and genome size of the hybrids would be intermediate between those of the parental species.

**Xenopus**

*Xenopus* (the African clawed frogs) have been widely used in research for decades. *Xenopus laevis* has been identified by the National Institute of Health (NIH) as one of five important non-mammalian models of human development and disease (Perry et al 2001). *Xenopus* are characterized genomically by allopolyplidization and range from diploid (n=20) to dodecaplophonic (n=108) numbers of chromosomes (Kobel and Dupasquier 1986). According to Malone et al. (2007) there is an overwhelming preponderance of misexpressed genes in the testis in F1 hybrids of *X. laevis* and *X. muelleri* and they follow a semi-dominant model of expression behavior because hybrid expression is intermediate or additive compared to expression difference between the two species. More genes are differentially expressed between species than between hybrids and each parental species.

Stancheva et al. (2002) pointed out that the parental genomes of *Xenopus* are methylated before and after the first cleavage of the zygote. This agrees with the conclusions of Reik and Walter (2001) that the outcome of the conflict between parents over the allocation of maternal resources to maintain embryos involve imprinting mechanisms and results in active demethylation of the paternal genome. DNA methylation contributes significantly to the maintenance of transcriptional silencing during the cleavage stages and together with lack of imprinting, may account for the
absence of demethylation and remodeling of methylation patterns in early *Xenopus* embryos (Stancheva et al 2002).

The genus *Xenopus* consists of diverse species separated into two distinct groups, one of which is often placed in the separate genus *Silurana* (*S.*, or *X. tropicalis*, sister to the others and the only diploid species in the group). All of the species are morphologically similar irrespective of the genetic differences. Geographically they are found south of the Sahara in Africa in almost any type of water body. There are some feral populations in North America and Europe as well.

Many *Xenopus* species are parapatric in distribution. Tinsley and Kobel (2002) reported that there have been three recorded cases of interspecific hybrids from contact zones in the wild: between *X. gilli* and *X. l. laevis* (Kobel et al 1981; Picker 1985), *X. borealis* and *X. l. victorianus* (Yager in Tinsley & Kobel 2002) and *X. muelleri* and *X. l. laevis* (Poynton & Broadley 1985).
CHAPTER 2

METHYLATION PATTERNS AMONG HYBRIDS AND PARENTS (Xenopus laevis and Xenopus muelleri) USING AMPLIFIED FRAGMENT LENGTH POLYMORPHISM (AFLP)

Introduction

Polyploidy is an important evolutionary process in plants and animals (Friedman and Hughes 2001; Grant 1981; Gu et al. 2002; Masterson 1994; McLysaght et al. 2002; Otto and Whitton 2000; Soltis and Soltis 1999; Liu and Wendel 2003). Polyploidization can have significant effects on gene expression, development and viability of an organism. Although polyploidization could be viewed as a deleterious event, the prevalence of polyploids especially among plants may be an indication of some evolutionary benefit to increased genome size and massive gene duplication. Hybridization and polyploidy are known to be prominent processes inducing diversification and speciation in plants (Stebbins 1950; Grant 1971; Abbott 1992; Masterson 1994; Rieseberg and Wendel 2004).

Due to the parapatric or allopatric distributions of Xenopus species there have been very few recorded cases in the wild of interspecific hybrids (Kobel 1981). Breeding experiments in a laboratory setting provide a perfect environment to study how polyploidy arose in this genus. According to Evans (2008) duplicate gene evolution and expression in Xenopus provides a unique perspective into some of the earliest genomic
transformations after vertebrate whole genome duplication (WGD) and suggests that functional constraints are relaxed compared to before duplication but still consistently strong for millions of years following WGD. Evans (2008) postulated that transposition and divergent resolution (processes that occur through different genetic mechanisms but have analogous implications for genome structure) were potentially a major catalyst for diversification of clawed frogs.

In *Xenopus* females are the heterogametic sex (ZW). Gametogenesis of F1 hybrids is defective with males being sterile and females producing unreduced eggs- eggs that contain the entire genome of the mother (Kobel 1996; Kobel and Dupasquier 1986). There is variation between individuals and clutches in the proportion of unreduced eggs produced and also these eggs do not always contain the full complement of the maternal chromosomes (Kobel 1996; Kobel and Dupasquier 1986). This is in sharp contrast to Haldane’s rule wherein the heterogametic sex should be sterile.

Polyplloid genome evolution often appears to be accompanied by rapid and biased structural changes (Song et al. 1995; Feldman et al. 1997; Liu et al. 1998a, b; Ozkan et al. 2001) and by activation of transposable elements and epigenetic changes that modulate gene expression (Comai 2000; Comai et al. 2000; Shaked et al. 2001; Kashkush et al. 2002, 2003; Adams et al. 2003; He et al. 2003; Wang et al. 2004). DNA methylation is one of these changes that can affect gene expression and can lead to activity of some of the transposable elements and extra chromosomes in the genome. For example, O’Neill et al. (1998) observed genomic DNA undermethylation in interspecific wallaby hybrids (*Macropus eugenii* x *Wallabia bicolor*) relative to the parents. This lead
them to conclude that deficient methylation and de novo changes in chromosome number in other mammalian hybrids may result in subsequent mobile-element reactivation, in turn facilitating rapid karyotypic reconfigurations in hybrids.

*Xenopus laevis* and *X. muelleri* originated from allopolyploidization following hybridization. According to Kobel (1981) the genus *Xenopus* exhibits clear trends of repeated polyploidization up to comparatively high DNA contents and of subsequent diploidization. Due to the fact that in laboratory settings interspecific hybrids produce polyploid eggs, Kobel (1981) stated that polyploidization most likely results from hybridization between species. This process creates certain zones of conflict in the nucleus of nascent allopolyploids (Jones & Hagarty 2009). Over millions of years these two species have found a way of accommodating the chromosomes from different species by diploidization through homoeologous chromosome pairing (pairing control genes) (Jones & Hagarty 2009) i.e., the chromosomes or genes in species that are related and derived from the same ancestor now coexist in the allopolyploid. According to Jones and Hagarty (2009) when hybrids are first formed conflicts arise due to differences in genome size, genome composition, regulatory mechanisms, cell cycle duration, genetic and epigenetic modifications and all of the factors that contribute to organismal function and success. For these individuals to reproduce accommodations are made in the genome. These may lead to changes such as chromosome loss. In fact methylation changes can be immediately triggered by polyploidization following the first few or during the first few generations after a polyploidization event (Wang et al. 2004a; Paun et al. 2007). Since both polyploidization and hybridization have major effects, hybridization itself may
trigger similar changes. Hence it is possible for these accommodations to be seen as early as the F1 generation. It is therefore very important to study methylation changes in the genomes of F1 hybrids and to see how the evolution of the hybrid genome may have occurred.

According to Malone et al. (2007), gene expression in interspecific hybrid testis (Xenopus laevis x Xenopus muelleri) closely resembles that of X. laevis. The use of additional molecular data could further shed light on the structural changes occurring in the hybrids. Nuclear DNA can be used to examine biparental genetic exchange between species (Fontenot 2009). One method of generating large numbers of genomic DNA markers is Amplified Fragment Length Polymorphism. This method does not require a priori knowledge of any sequence information, in polyploids, during hybridization events and for rapid generation of data (Meudt and Clarke 2007). Although AFLPs are highly reproducible there are always some inconsistencies and issues of interpretation. To reduce error rates, replicates, preferably from different DNA extractions, should be used. It has been estimated that this reduces the error rate per locus to between 2-5% (Meudt and Clarke 2007).

AFLPs are very useful in identifying hybrid individuals (interspecific, or between intraspecific lineages), even in systems where microsatellites have failed to do so (Bensch et al. 2002b). The AFLP technique involves the use of two restriction endonuclease enzymes to digest genomic DNA, which is then followed by two rounds of selective Polymerase Chain Reaction (PCR) amplification of some of the restriction
fragments. Different primers with different selective nucleotides are used to sample the genome. This produces different markers (bands/peaks) that can be scored as presence (score 1) or absent (score 0) (Bensch and Akesson, 2005). AFLP is predominantly a dominant marker technique.

A modification of AFLP (Liu et al. 2001) is used in this study to look at methylation. Two isoschizomeric restriction enzymes (HpaII/MspI), with different sensitivity to methylation but recognizing the same four base segment (5’-CCGG-3’), were used in parallel reactions. The final markers where then compared between the two groups. HpaII is methylation-sensitive whereas its isoschizomer MspI is methylation-insensitive. This method, Methylation-Sensitive AFLP has been applied in other studies such as those of hybridization and polyploidy in Spartina (Poaceae) in Europe (Salmon et al. 2005).

The objectives of this study were to use AFLP data to identify genetic and molecular variation among and between hybrids and parents. Using F1 hybrids between X. laevis and X. muelleri methylation-specific AFLP (MSAP) were employed to determine the difference in methylation patterns between the F1 hybrids and parental types. I hypothesized that the methylation pattern of the F1 hybrids would be intermediate between those of the parental species and addressed the following questions: 1) Is the methylation pattern of F1 hybrids distinct from those of the parental species; 2) What is the genetic diversity of the F1 hybrids compared to the parental species; 3) What is the pattern of inheritance of the MSAP methylated fragments from the parental species
to the F1s; 4) Is there any MSAP fragment loss and if so, what percentage are from the paternal or maternal parental species; 4) and finally is this loss independent or dependent on the parental species.

**Methods**

**Specimens**

There are several species of *Xenopus* being raised in our laboratory. Hybrids have been breed in the lab from *X. laevis* and *X. muelleri*. The parental species are at least five years old and the F1 hybrids are at least four years old. These were used for obtaining tissue samples. The original parents of the F1 hybrids could not be traced, therefore representative samples of the parental species and hybrids from the collection are used for this study. *X. laevis* is the paternal species and *X. muelleri* is the maternal for the F1 hybrids used in this study. Individuals used in this study are *Xenopus laevis* (n = 11), *X. muelleri* (n = 11); F1 hybrids (*X. laevis* x *X. muelleri*, males n = 11 & females n = 9). There a total of eight liver samples used for *X. laevis, X. muelleri and* male F1 Hybrids (*X. laevis* x *X. muelleri*) but only 6 for female F1 hybrids. For the muscle samples four individuals of each group (population) was used for the AFLP analysis.

**DNA Extraction**

QIAGEN DNeasy® Tissue Kits were used to isolate DNA according to the manufacturer’s protocol. Extracted DNA was stored at -20°C. DNA quality was checked after extraction by gel electrophoresis using 1% agarose gel and SYBR-green and
compared to a 1kb ladder. Only DNA that was not degraded and had a distinct band between 10,000 to 20,000 bp was used for further analyses. To quantify the gDNA, it was measured with a Nanodrop spectrophotometer. DNA was extracted from muscle and liver tissues.

*Methylation-Sensitive Amplification Polymorphism (MSAP)*

DNA fragment analysis was done to determine methylation patterns. *HpaII-EcoRI* and *MspI-EcoRI* combination of restriction enzymes were used. The fragments were analyzed using Applied Biosystems DNA Analyzer 3130. Replicates using the same DNA were done to avoid technical bias. Only the stable and repeatable patterns resulting from two independent digestions were retained.

Methylation-sensitive AFLP data was obtained using a modified protocol according to methods of Makowsky et al (2009). The isoschizomers *HpaII* and *MspI* were used instead of *MseI*. Both isoschizomers recognize the same tetranucleotide sequence 5’ –CCGG- 3’ but with different sensitivity to the cystosine methylation. *HpaII* will not cleave if internal cystosine is methylated and *MspI* is only sensitive to methylation of the external cystosine on both strands. If either the external cystosine or a fully methylated CpG site is encountered then the methylation patterns of the two are indistinguishable. *EcoRI*, a six bp cutter, was used in combination with the two restriction enzymes. *EcoRI-HpaII* amplification products compared to *EcoRI-MspI* amplification products will allowed detection of methylation changes. If a fragment is present in both digestions, then the CpG site is unmethylated. However if it is absent in *EcoRI-HpaII* digestion but it is present in *EcoRI-MspI* digestions it is an indication that the internal
cystosine is methylated. On the contrary, if it is present in EcoRI-HpaII but absent in EcoRI-MspI digestions then is an indication of methylation of the external cystosine.

Using two parallel reactions (EcoRI/HpaII & EcoRI/MspI) genomic DNA was digested at 37°C for 6 hours followed by ligation using 75 µM of EcoRI adapter and HpaII/MspI adapter using T4 ligase enzyme for 12 hours at 16°C. 2 µl of the restriction digest and ligated product where used in the preselective amplification reaction using the following PCR conditions 1) initial touchdown procedure with 1 minute at 72°C, 50 seconds at 94°C, 1 minute at 50°C, and 1 minute at 72°C, 2) 24 cycles of denaturation at 95°C for 15 seconds, annealing at 60°C for 1 minute and extension at 72°C for 1 minute, and 3) a final extension at 72°C for 7 minutes. The PCR conditions for the selective amplification were as follows: 1) initial touchdown procedure with 50 seconds at 94°C, 1 minute at 57°C, and 2 minutes at 72°C, and 2) 19 cycles of denaturation at 94°C for 15 seconds, annealing at 56°C for 1 minute and extension at 72°C for 1 minute. Preselective amplification was obtained using [EcoRI + 1] and [HpaII/MspI + 1] primers and the selective amplification is done using [EcoRI + 2] (FAM-labeled) and [HpaII/MspI +2] (Table 2.6). The FAM-labeled EcoRI primer was used to enable visualization by the machine. Selective PCR products were purified using a standard ethanol cleanup protocol and sequenced using an ABI 3130 capillary sequencer with a GenScan Rox 400 HD size standard (Fig 2.1).

The AFLP fragments were scored using Genemarker® program. To minimize complications arising from homoplasy only fragments greater than or equal to 100 base pairs were included in the analysis. Also only peaks with a threshold value greater than
100 will be used and the fluorescence peaks of each locus were verified for each individual. The remaining fragments were scored and assembled into absent (0) or presence (1) fragments to be used for statistical analysis (Appendix A).
Figure 2.1 Diagrammatic representation of the AFLP Technique
To ensure a good quality dataset several precautions were taken. The precautions suggested by Bonin et al. (2004) were implemented with modifications. The reproducibility of AFLP markers was tested by running each sample in duplicate (all the steps from restriction digests to data scoring). The scored data was then grouped into two categories by population (X. laevis, X. muelleri, F1 male hybrids, and F1 female hybrids): ECOR1-HpaII and ECOR1-Msp1. The markers that were duplicated per sample were kept. If only one duplicate is present then all the individuals within the population were looked at and if at least one individual has the fragment duplicated or two-thirds of all the individuals in the population have one duplicate the fragment is not excluded. Secondly the individuals with two duplicate the fragment is scored as one (1) and with only one duplicate or none the fragment is scored a zero (0). The genotyping error rate per individual was calculated as the ratio of the bands (fragments) with only one duplicate to scored bands/fragment (with two duplicates). The genotyping error rate ranges from 1% to 20% per individual (Appendix A). For the analysis ECOR1-HpaII fragments were compared with ECOR1-Msp1 fragments for each individual. Methylated fragments were represented by 1 and nonmethylated fragments represented by zero (0) (Appendix A).
**Statistical Analysis**

The binary data was analyzed using the program GENALEX 6.3 (Peakall and Smouse, 2006). According to Kosman and Leonard (2005) band-based and allele-frequency based approaches exist to extract statistical information from AFLP data. Two types of genetic distances were calculated: genetic distances and frequency-based distance (Appendix B). A pair-wise genetic distance matrix for the binary data was generated producing a profile of individual genetic distances based on discrete values of presence and absence. This is a Euclidean distance matrix. The genetic distances matrices enable further genetic analysis – AMOVA and Principal Coordinate Analysis (PCoA).

It is assumed that AFLP profiles, being multilocus, each band correspond to a locus with two alleles. The presence band represents the homozygote dominant and heterozygote dominant and the absence band correspond to the recessive homozygote. Based on this assumption an estimate of allele frequencies can be calculated in GENALEX using a frequency base genetic distance. The procedure follows Lynch and Mulligan (1994) and assumes complete outcrossing.

**Analysis of Molecular Variance**

The individual genetic distance matrices were further analyzed by analysis of molecular variance (AMOVA) approach, which follows the methods of Excoffer et al (1992). The genetic variation is partitioned into two levels: within and among population. This variation is summarized as total variance and $\phi$-statistic or F-statistic. The statistical significance was tested through random permutation, which was set at 9999.
Using the genetic distance as a starting point Principal Coordinate Analysis was conducted. This a multivariate technique that allows the graphing of major patterns within a multivariate data set. Principal Coordinate Analysis was used instead of Principal Component Analysis because it is more appropriate when using binary presence and absence data. Also in this case the use of Euclidean distances between observations makes little sense, since such data may give counterintuitive results. The first two axes were used in plotting the patterns of individual population genetic distances.

Results

The AFLP technique was carried out for two different tissues: liver and muscle. The liver and muscle from 30 and 16 individuals respectively was used in the analysis.

Molecular Markers

There were 364 and 504 significant methylated bands combined corresponding to muscle (16 individuals) and liver tissues (30 individuals) respectively. From the 504 liver makers 207 are found in the six F1 females, 341 in the eight F1 males, 350 in the eight *X. laevis* and 333 in eight *X. muelleri*. For the muscle samples, 188 were in four F1 females, 208 in four F1 males, 187 in four *X. laevis* individuals and 180 in four *X. muelleri* individuals. All of these bands occur at a frequency $\geq 5\%$ (Table 2.1)
<table>
<thead>
<tr>
<th>Pop</th>
<th>n</th>
<th>No. of Loci</th>
<th>Na (SE)</th>
<th>Ne (SE)</th>
<th>I (SE)</th>
<th>He (SE)</th>
<th>Uhe (SE)</th>
<th>Percent Polymorphic Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Liver</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1-female</td>
<td>6.000</td>
<td>277</td>
<td>1.091</td>
<td>1.234</td>
<td>0.241</td>
<td>0.152</td>
<td>0.166</td>
<td>54.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.044</td>
<td>0.013</td>
<td>0.011</td>
<td>0.007</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>F1-male</td>
<td>8.000</td>
<td>341</td>
<td>1.343</td>
<td>1.270</td>
<td>0.282</td>
<td>0.176</td>
<td>0.188</td>
<td>66.67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.013</td>
<td>0.011</td>
<td>0.007</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>X. laevis</td>
<td>8.000</td>
<td>350</td>
<td>1.381</td>
<td>1.303</td>
<td>0.302</td>
<td>0.192</td>
<td>0.205</td>
<td>68.65%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.041</td>
<td>0.014</td>
<td>0.011</td>
<td>0.008</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>X. muelleri</td>
<td>8.000</td>
<td>333</td>
<td>1.321</td>
<td>1.242</td>
<td>0.265</td>
<td>0.163</td>
<td>0.174</td>
<td>66.07%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.012</td>
<td>0.010</td>
<td>0.007</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>7.500</td>
<td>504</td>
<td>1.284</td>
<td>1.262</td>
<td>0.272</td>
<td>0.171</td>
<td>0.183</td>
<td>63.89%</td>
</tr>
<tr>
<td>(SE)</td>
<td></td>
<td></td>
<td>0.019</td>
<td>0.021</td>
<td>0.006</td>
<td>0.005</td>
<td>0.004</td>
<td>3.29%</td>
</tr>
<tr>
<td><strong>Muscle</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1-female</td>
<td>4.000</td>
<td>188</td>
<td>0.956</td>
<td>1.259</td>
<td>0.232</td>
<td>0.154</td>
<td>0.176</td>
<td>43.96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.050</td>
<td>0.018</td>
<td>0.014</td>
<td>0.010</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>F1-male</td>
<td>4.000</td>
<td>208</td>
<td>1.030</td>
<td>1.274</td>
<td>0.244</td>
<td>0.162</td>
<td>0.185</td>
<td>45.88%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.049</td>
<td>0.019</td>
<td>0.015</td>
<td>0.010</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>X. laevis</td>
<td>4.000</td>
<td>187</td>
<td>0.953</td>
<td>1.249</td>
<td>0.228</td>
<td>0.150</td>
<td>0.171</td>
<td>43.96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.050</td>
<td>0.018</td>
<td>0.014</td>
<td>0.010</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>X. muelleri</td>
<td>4.000</td>
<td>180</td>
<td>0.926</td>
<td>1.230</td>
<td>0.217</td>
<td>0.142</td>
<td>0.162</td>
<td>43.13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.051</td>
<td>0.017</td>
<td>0.014</td>
<td>0.009</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4.000</td>
<td>364</td>
<td>0.966</td>
<td>1.253</td>
<td>0.230</td>
<td>0.152</td>
<td>0.174</td>
<td>44.23%</td>
</tr>
<tr>
<td>(SE)</td>
<td>0.000</td>
<td></td>
<td>0.025</td>
<td>0.009</td>
<td>0.007</td>
<td>0.005</td>
<td>0.006</td>
<td>0.58%</td>
</tr>
</tbody>
</table>

Na = No. of Different Alleles, Ne = No. of Effective Alleles = 1 / (p^2 + q^2), I = Shannon's Information Index = -1 * (p * Ln (p) + q * Ln(q)), He = Expected Heterozygosity = 2 * p * q, UHe = Unbiased Expected Heterozygosity = (2N / (2N-1)) * He Where for Diploid Binary data and assuming Hardy-Weinberg Equilibrium, q = (1 - Band Freq. thư 0.5 and p = 1 - q.
Figure 2.2 Summary graph showing the results of methylated band (alleles) patterns across populations [a) Liver and b) Muscle Tissues]
Cluster Analysis/Population Structure.

The results of the Principal coordinate analysis (PCoA) analysis of the individual pair-wise genetic distances are shown in Figs 2.3 and 2.4 for liver and muscle tissues respectively. Only the first two axes were used in plotting the PCoA. In both cases four distinct clusters were partitioned. The hybrids were clustered between the parents (*X. laevis* and *X. muelleri*). The clusters are more distinctly separated with the muscle tissues than the liver tissues. From the PCoA analysis 64.11% (liver) and 68.6% (muscle) of the variance can be explained by the first three eigenvalues (Table 2.2). With one exception all the liver samples from *X. muelleri* form a distinct cluster. The F1 female and male hybrids are spread in-between the two parental clusters. The F1 males are located in two clusters above and below the F1 females. However for the muscle tissue the clusters are distinct and less discrete for the liver tissue. This is probably due to more variation in the methylated fragments amongst individual liver tissue. The general pattern is however the same for liver and muscle tissue.
Table 2.2 Summary Statistics for Principal Coordinate Analysis (PCoA) for all individuals [Liver and Muscle Tissues]

<table>
<thead>
<tr>
<th></th>
<th>1st Eigenvalue (%) Variance</th>
<th>1st Eigenvalue (%) Variance</th>
<th>1st Eigenvalue (%) Variance</th>
<th>Cumulative % variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver</td>
<td>44.643 (25.68)</td>
<td>39.714 (22.84)</td>
<td>23.341 (15.59)</td>
<td>64.11</td>
</tr>
<tr>
<td>Muscle</td>
<td>48.601 (31.14)</td>
<td>35.823 (22.95)</td>
<td>22.643 (14.51)</td>
<td>68.59</td>
</tr>
</tbody>
</table>
Pop 1 = F1-females (X. laevis x X. mulleri)
Pop 2 = F1-males (X. laevis x X. mulleri)
Pop 3 = X. laevis
Pop 4 = X. mulleri

Figure 2.3 Principal Coordinates Analysis (PCoA) plot of MSAFLP for all individual pair-wise genetic distances [Liver Tissue] - PCoA via Covariance with data standardization
Pop 1 = F1-females (X. laevis x X. muelleri)
Pop 2 = F1- males (X. laevis x X. muelleri)
Pop 3 = X. laevis
Pop 4 = X. muelleri

Figure 2.4 Principal Coordinates Analysis (PCoA) plot of MSAFLP for all individual pair-wise genetic distances [Muscle Tissue] - PCoA via Covariance with data standardization
**Analysis of Molecular Variance**

An AMOVA performed on these four clusters was significant (P < 0.0001) for both liver and muscle tissues. This information is summarized in Table 2.3 and Fig 2.4. With liver tissues 10% and 90% of the genetic variance accounts for among and within population (P < 0.0001) and for muscle tissues 34% and 66% of the genetic variance accounts for among and within population (P < 0.0001). Out of 504 methylated liver loci, F1 hybrid females (*X. laevis* x *X. muelleri*) have 54.17% polymorphic loci (methylated), F1 hybrid males (*X. laevis* x *X. muelleri*) have 66.67% polymorphic loci, *X. laevis* 68.65% and *X. muelleri* 66.07% polymorphic loci (Table 2.1). Out of 364 methylated muscle loci 43.96%, 45.88%, 43.96%, 43.13% are polymorphic for F1 hybrid females, F1 hybrid males, *X. laevis*, and *X. muelleri* respectively. With the exception of the F1 hybrid females – F1 hybrid males liver tissues all pair-wise population F-statistic are significant for both the muscle and liver tissues.
Table 2.3 Results of AMOVA performed on all populations identified as clusters [Liver and muscle Tissues]

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Source of variation</th>
<th>df</th>
<th>Variance component</th>
<th>% variance</th>
<th>φ-Statistic</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver</td>
<td>Among</td>
<td>3</td>
<td>7.715</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>26</td>
<td>67.755</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>29</td>
<td>75.470</td>
<td>1</td>
<td>0.102</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Muscle</td>
<td>Among</td>
<td>3</td>
<td>22.781</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>12</td>
<td>44.167</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>15</td>
<td>66.948</td>
<td>100</td>
<td>0.34</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Probability, P (rand<=data), for PhiPT is based on 9999 permutation across the full data set
Figure 2.5 Pie Chart summarizing the outcome of hierarchical AMOVA, showing the partitioning of total molecular variance within and among all population [a) Liver and b) Muscle Tissues]
Genetic Diversity

The estimate for the methylated band frequency, Nei genetic diversity and heterozygosity is presented in Tables 2.4 and 2.5. All the methylated bands/loci used for the analysis for all the four different populations have a frequency $\geq 5\%$. For the liver tissue F1 females have 10, F1 males 21, *X. laevis* 36 and *X. muelleri* 29 unique/ private bands (Fig 2.2). The unbiased expected heterozygosity for F1 hybrid females, F1 hybrid males, *X. laevis* and *X. muelleri* is as follows $0.166\pm0.008$, $0.188\pm0.008$, $0.205\pm0.008$ and $0.174\pm0.007$ respectively. 33 methylated bands/loci for F1 hybrid females, 71 bands for F1 hybrid males, 75 bands for *X. laevis* and 71 bands for *X. muelleri* frequencies $\geq5\%$ are present in 25% to 50% of the population. On the other hand muscle tissues showed unbiased expected heterozygosity for F1 hybrid females, F1 hybrid males, *X. laevis* and *X. muelleri* as follows $0.176\pm0.011$, $0.185\pm0.012$, $0.171\pm0.011$ and $0.162\pm0.011$ respectively. 21 of the 188 methylated bands/loci for F1 hybrid females, 71 of the 208 bands for F1 hybrid males, 75 of the 187 bands for *X. laevis* and 71 of the 180 bands for *X. muelleri* with frequencies $\geq5\%$ are present in 25% to 50% of the population. The mean expected unbiased Nei genetic distance across all populations is 0.169 with the muscle tissues and 0.033 for liver tissues (Table 2.5).
Table 2.4 Results of Pair wise Population PhiPT ($\phi$PT) and LinPhiPT Values and Estimates of Nm [liver and muscle] - Input as Genetic Distance Matrix for calculation of PhiPT ($\phi$PT)

### Liver

<table>
<thead>
<tr>
<th>Pop1</th>
<th>Pop2</th>
<th>PhiPT</th>
<th>LinPhiPT</th>
<th>#Pop1</th>
<th>#Pop2</th>
<th>P (rand $\geq$ data)</th>
<th>No. PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-female</td>
<td>F1-male</td>
<td>0.041</td>
<td>0.042</td>
<td>6</td>
<td>8</td>
<td>0.097</td>
<td>9999</td>
</tr>
<tr>
<td>F1-female $X.\ laevis$</td>
<td>0.119</td>
<td>0.135</td>
<td>6</td>
<td>8</td>
<td>0.001</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>F1-male $X.\ laevis$</td>
<td>0.108</td>
<td>0.121</td>
<td>8</td>
<td>8</td>
<td>0.002</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>F1-female $X.\ muelleri$</td>
<td>0.071</td>
<td>0.077</td>
<td>6</td>
<td>8</td>
<td>0.022</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>F1-male $X.\ muelleri$</td>
<td>0.090</td>
<td>0.099</td>
<td>8</td>
<td>8</td>
<td>0.005</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>$X.\ laevis$ $X.\ muelleri$</td>
<td>0.155</td>
<td>0.183</td>
<td>8</td>
<td>8</td>
<td>0.001</td>
<td>9999</td>
<td></td>
</tr>
</tbody>
</table>

### Muscle

<table>
<thead>
<tr>
<th>Pop1</th>
<th>Pop2</th>
<th>PhiPT</th>
<th>LinPhiPT</th>
<th>#Pop1</th>
<th>#Pop2</th>
<th>P (rand $\geq$ data)</th>
<th>No. PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-female</td>
<td>F1-male</td>
<td>0.193</td>
<td>0.238</td>
<td>4</td>
<td>4</td>
<td>0.027</td>
<td>9999</td>
</tr>
<tr>
<td>F1-female $X.\ laevis$</td>
<td>0.412</td>
<td>0.700</td>
<td>4</td>
<td>4</td>
<td>0.030</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>F1-male $X.\ laevis$</td>
<td>0.384</td>
<td>0.624</td>
<td>4</td>
<td>4</td>
<td>0.029</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>F1-female $X.\ muelleri$</td>
<td>0.294</td>
<td>0.417</td>
<td>4</td>
<td>4</td>
<td>0.029</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>F1-male $X.\ muelleri$</td>
<td>0.335</td>
<td>0.503</td>
<td>4</td>
<td>4</td>
<td>0.029</td>
<td>9999</td>
<td></td>
</tr>
<tr>
<td>$X.\ laevis$ $X.\ muelleri$</td>
<td>0.382</td>
<td>0.618</td>
<td>4</td>
<td>4</td>
<td>0.030</td>
<td>9999</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.5 Results of Pair wise Population Unbiased Nei Genetic Distance and Unbiased Nei Genetic Identity Values as Table [liver and muscle tissues]

Liver

<table>
<thead>
<tr>
<th>Pop1</th>
<th>Pop2</th>
<th>UNei GD</th>
<th>UNei ID</th>
<th>#Pop1</th>
<th>#Pop2</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-female</td>
<td>F1-male</td>
<td>0.017</td>
<td>0.983</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>F1-female</td>
<td><em>X. laevis</em></td>
<td>0.028</td>
<td>0.972</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>F1-male</td>
<td><em>X. laevis</em></td>
<td>0.029</td>
<td>0.972</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>F1-female</td>
<td><em>X. muelleri</em></td>
<td>0.016</td>
<td>0.984</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>F1-male</td>
<td><em>X. muelleri</em></td>
<td>0.021</td>
<td>0.979</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><em>X. laevis</em></td>
<td><em>X. muelleri</em></td>
<td>0.033</td>
<td>0.968</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Muscle

<table>
<thead>
<tr>
<th>Pop1</th>
<th>Pop2</th>
<th>UNei GD</th>
<th>UNei ID</th>
<th>#Pop1</th>
<th>#Pop2</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-female</td>
<td>F1-male</td>
<td>0.017</td>
<td>0.983</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>F1-female</td>
<td><em>X. laevis</em></td>
<td>0.028</td>
<td>0.972</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>F1-male</td>
<td><em>X. laevis</em></td>
<td>0.029</td>
<td>0.972</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>F1-female</td>
<td><em>X. muelleri</em></td>
<td>0.016</td>
<td>0.984</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>F1-male</td>
<td><em>X. muelleri</em></td>
<td>0.021</td>
<td>0.979</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><em>X. laevis</em></td>
<td><em>X. muelleri</em></td>
<td>0.033</td>
<td>0.968</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>
Figure 2.6 Frequency Distribution of Random permuted PhiPT (\(\phi_{PT}\)) versus Observed PhiPT (\(\phi_{PT}\)) for 9999 Permutations – input as Genetic Distance Matrix for Calculation of PhiPT (\(\phi_{PT}\)) [a) liver and b) muscle issues]
**Comparison of MSAP methylated fragments between parental and F1 hybrids**

The MSAP fragments found in the parental species compared to the F1 hybrids are summarized in Tables 2.7 and 2.8. A total of 741 liver and 546 muscle MSAP fragments were generated from one selective primer combination (Table 2.6). 504 and 364 fragments out of the total generated were methylated from the liver and muscle tissues respectively and 237 liver and 182 muscle fragments were unmethylated (Table 2.9).

Of the 504 methylated liver fragments 462 were parental. Of this 215 (46.5%) were unique to either of the parental species. Both sexes share 67 (31.2%) of the diagnostic parental fragments (fragments that are specific to either parental species). A total of 99 (46%) of the diagnostic parental fragments are missing from both sexes. Both sexes share 28.8% of the diagnostic parental fragments. Both sexes possess a total of 42 unique/private methylated fragments: 11 are found in both, 21 in males and 10 in females. The missing methylated diagnostic fragments are biased toward *X. laevis* for the male F1 hybrid (23.5%) and for the females they are biased toward *X. muelleri* (25%) (Table 2.7, Fig 2.7). The proportion of methylated liver fragments from this analysis was estimated as 77.8% F1 female, 78.8% F1 males, 73.5% *X. laevis* and 70.1% *X. muelleri* (Table 2.9).

MSAP for the muscle tissues generated a total of 364 methylated fragments/loci. Of these 289 were from the parental species and 211 (73%) were unique to either parent. A total of 108 (51.2%) of the fragments are missing from both sexes and only 63 (29.9%) are present. A total of 30 (14.2%) and 48 (22.7%) of the parental species fragments are
missing from the males and females respectively. The missing fragments were biased towards *X. laevis* for both the males (10.9%) and the females (24.5%) (Table 2.8, Fig 2.8). For this study the muscle MSAP fragments generated were estimated to be 68.6% (F1 female), 73.8% F1 male, 67.0% (*X. laevis*) and 67.2% (*X. muelleri*) methylated (Table 2.9)

Using Fisher’s exact test the significance of the proportion of missing fragments was tested. For both the muscle and the liver the data was separated into two sets: Parental diagnostic fragments present in both sexes vs. missing in both sexes and parental diagnostic fragments missing in one sex but present in the other (Table 2.10). For the liver the proportion of parental diagnostic fragments that are present in both sexes and missing in both is not significant (*P* = 0.86303; 2-tailed test) whereas the proportions that are missing in one sex is significant (*P* = 0.02846; 2-tailed test). However for the muscle tissue the proportion of the parental fragments present in both sexes vs. missing in both is not significant (*P* = 0.07074; 2-tailed) and the proportion of parental diagnostic fragments missing in one sex is also not significant (*P* = 0.40245; 2-tailed test) (Table 2.10).
<table>
<thead>
<tr>
<th>Table 2.6 Primers used in MSAP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adaptors</strong></td>
</tr>
<tr>
<td><em>ECOR1</em></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><em>HpaII</em> (METHADPR)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Preselective Primers</strong></td>
</tr>
<tr>
<td><em>ECOR1preG</em></td>
</tr>
<tr>
<td><em>HpaII</em> (Fluorescent primer)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Selective Primers</strong></td>
</tr>
<tr>
<td><em>ECOR1</em> (Fluorescent primer)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><em>HpaII</em> (METHADPR)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Table 2.7 Methylated MSAP fragments observed in the parents (*X. laevis* and *X. muelleri*) and in the hybrid (*X. laevis x X. muelleri*) generated by EcoRI-HpaII and EcoRI-Msp1 digestion (one selective primer combination) - [Liver]

<table>
<thead>
<tr>
<th></th>
<th>Parental fragments in both F1 sexes</th>
<th>Parental fragments missing in both F1 sexes</th>
<th>Parental fragments missing in only one F1 sex Male</th>
<th>Parental fragments missing in only one F1 sex Female</th>
<th>Fragments shared only by Hybrids</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common parental fragments</strong></td>
<td>247</td>
<td>158 (64%)</td>
<td>34 (13.8%)</td>
<td>20 (8.1%)</td>
<td>35 (14.2%)</td>
</tr>
<tr>
<td><strong>Diagnostic fragments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From <em>X. laevis</em></td>
<td>115</td>
<td>36 (31.3%)</td>
<td>36 (31.3%)</td>
<td>27 (23.5%)</td>
<td>16 (13.9%)</td>
</tr>
<tr>
<td>From <em>X. muelleri</em></td>
<td>100</td>
<td>31 (31%)</td>
<td>29 (29%)</td>
<td>15 (15%)</td>
<td>25 (25%)</td>
</tr>
<tr>
<td><strong>Total number of diagnostic fragments between parents</strong></td>
<td>215</td>
<td>67 (31.2%)</td>
<td>99 (46%)</td>
<td>62 (28.8%)</td>
<td>76 (35.3%)</td>
</tr>
<tr>
<td><strong>Private/Unique fragments</strong></td>
<td>-</td>
<td></td>
<td></td>
<td>21</td>
<td>10</td>
</tr>
</tbody>
</table>
Table 2.8 Methylated MSAP fragments observed in the parents (*X. laevis* and *X. muelleri*) and in the hybrid (*X. laevis* x *X. muelleri*) generated by EcoRI-HpaII and EcoR1-Msp1 digestion (one selective primer combination) - [Muscle Tissues]

<table>
<thead>
<tr>
<th>Parental fragments</th>
<th>Parental fragments missing in both F1 sexes</th>
<th>Parental fragments missing in only one F1 sex</th>
<th>Fragments shared only by Hybrids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common parental fragments</td>
<td>78</td>
<td>40 (51.3%)</td>
<td>20 (25.6%)</td>
</tr>
<tr>
<td>Diagnostic fragments From <em>X. laevis</em></td>
<td>110</td>
<td>24 (21.8%)</td>
<td>47 (42.7%)</td>
</tr>
<tr>
<td>From <em>X. muelleri</em></td>
<td>101</td>
<td>39 (38.6%)</td>
<td>41 (40.6%)</td>
</tr>
<tr>
<td>Total number of diagnostic fragments between parents</td>
<td>211</td>
<td>63 (29.9%)</td>
<td>108 (51.2%)</td>
</tr>
<tr>
<td>Private/Uniq fragments</td>
<td>21</td>
<td>18</td>
<td>36</td>
</tr>
</tbody>
</table>
Figure 2.7 Liver MSAP fragments observed in individuals from *X. laevis, X. muelleri* and F1 hybrid males and females
Figure 2.8 Muscle MSAP fragments observed in individuals from *X. laevis*, *X. muelleri* and F1 hybrid males and females.

**Legend: Fragments Origin**

- **X. laevis**
- **X. muelleri**
- Common *X. laevis and X. muelleri*
- Unique F1 Male fragments
- Unique F1 Female fragments
- Unique F1 common fragments

**Methylated Fragments**

**Unmethylated Fragment**
Table 2.9 Percentage of Methylated Fragment for Liver and Muscle

Liver

<table>
<thead>
<tr>
<th>Cluster/Population</th>
<th>Nonmethylated</th>
<th>Methylated</th>
<th>Total</th>
<th>% Methylated fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 female</td>
<td>79</td>
<td>277</td>
<td>356</td>
<td>77.8</td>
</tr>
<tr>
<td>F1 male</td>
<td>92</td>
<td>341</td>
<td>433</td>
<td>78.8</td>
</tr>
<tr>
<td>X. laevis</td>
<td>126</td>
<td>350</td>
<td>476</td>
<td>73.5</td>
</tr>
<tr>
<td>X. muelleri</td>
<td>142</td>
<td>333</td>
<td>475</td>
<td>70.1</td>
</tr>
</tbody>
</table>

Muscle

<table>
<thead>
<tr>
<th>Cluster/Population</th>
<th>Nonmethylated</th>
<th>Methylated</th>
<th>Total</th>
<th>% Methylated fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 female</td>
<td>86</td>
<td>188</td>
<td>274</td>
<td>68.6</td>
</tr>
<tr>
<td>F1 male</td>
<td>74</td>
<td>208</td>
<td>282</td>
<td>73.8</td>
</tr>
<tr>
<td>X. laevis</td>
<td>92</td>
<td>187</td>
<td>279</td>
<td>67.0</td>
</tr>
<tr>
<td>X. muelleri</td>
<td>88</td>
<td>180</td>
<td>268</td>
<td>67.2</td>
</tr>
</tbody>
</table>
Table 2.10 Fisher’s Exact Test for the proportion of MSAP diagnostic parental fragments present or missing in the F1 for Liver and Muscle tissue

### Liver

<table>
<thead>
<tr>
<th></th>
<th>Parental fragments in both F1 sexes</th>
<th>Parental fragments missing in both F1 sexes</th>
<th>Parental fragments missing in male F1</th>
<th>Parental fragments missing in female F1</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>X. laevis</em></td>
<td>36</td>
<td>36</td>
<td>0.86303</td>
<td></td>
<td>0.02846</td>
</tr>
<tr>
<td><em>X. muelleri</em></td>
<td>31</td>
<td>29</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Muscle

<table>
<thead>
<tr>
<th></th>
<th>Parental fragments in both F1 sexes</th>
<th>Parental fragments missing in both F1 sexes</th>
<th>Parental fragments missing in male F1</th>
<th>Parental fragments missing in female F1</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>X. laevis</em></td>
<td>24</td>
<td>47</td>
<td>0.07074</td>
<td></td>
<td>0.40245</td>
</tr>
<tr>
<td><em>X. muelleri</em></td>
<td>39</td>
<td>41</td>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

Analysis of the AFLP data shows some interesting patterns with regards to the relationship between hybrids and their parental species (*X. laevis* and *X. muelleri*). The patterns from the genetic variation at 504 methylated fragments/loci and 364 methylated loci from liver and muscle tissues are summarized below.

**Population Structure**

Four distinct clusters were identified from the PCoA scatter plot. The parents produce 2 separate clusters with the hybrids in between. The clusters were less discrete with the liver samples compared to the muscle tissues. This is an indication that there is more variation in methylated fragments with the liver tissues. This clustering of the hybrids between the two parental species is not surprising, since the hybrids are considered as having intermediate genotypes.

**Genetic Diversity**

The percentage of polymorphic methylated fragments ranges from 43.13% to 45.88%, and 54.17% to 68.65%, with mean 44.23%±0.58% and 63.89% ±3.29% for muscle and liver tissues, respectively. All the four groups exhibited a high degree of polymorphic methylated fragments for both tissues. There is very little difference in the expected heterozygosity amongst the F1 hybrids and the parents for the muscle (Table 2.2). With the exception of *X. laevis* (0.205±0.008) the values for the liver which range from 0.166±0.008 to 0.188±0.008 show also very little difference. This is unexpected because the level of heterozygosity is expected to be high in newly formed species/nuclei
compared to the older more established species. The F1 hybrid with a new genome formed via hybridization of species that have separated over 60 million years ago might be expected to have much higher heterozygosity compared to the parental species. This therefore requires further investigation. Formation of a nascent nucleus by interspecific hybridization produces a nucleus in chaos with different chromosomes being eliminated or silenced. The remaining ones are expected to be in flux resulting in a higher heterozygosity. It did not seem to be the case for the hybrids. One possible explanation is that homologous chromosomes (alleles) from these two parental species that that were probably derived from the same ancestors were able to pair up in the hybrid hence reducing the heterozygosity to a level that is the same as the parental species. Also the similar level of heterozygosity is not unexpected since AFLP markers are dominant markers. Both the heterozygous and the homozygous condition are represented by the presence of an allele and the homozygous recessive is represented by the absence (0). It is in the expected unbiased Nei genetic distance (UNei GD) that some interesting relations emerged that pose further questions (Table 6). There is very little genetic dissimilarity for the methylated fragments between groups within liver. This was probably one of the reasons why there was so much variation and the difficulty of resolving the clusters with the PCoA. The DNA in the liver probably contains highly conserved sequences and therefore these same sequences were inherited from by the hybrid from the parental species. There were therefore very little epigenetic changes in the F1 hybrid liver DNA. The parents have adapted over time to live with these homologous chromosomes. It is
probably advantageous for the F1 hybrid to continue to maintain this same methylation pattern.

The situation is completely reversed when the expected unbiased Nei genetic distance (UNei GD) is calculated for the muscle tissues. The sample size is much smaller so the results were looked at with caution. With that said, the amount of molecular dissimilarity found is much greater between the F1 hybrid sexes and between the F1 and the parental species. For the parental species it is expected because they have separated over a long period of time. Even though this study did not investigate if there is any direct correlation between muscle tissue DNA methylation and the deformities seen in the F1 hybrids, it is something to be looked at as a result of this finding. Furthermore the influence of the female sex chromosome in the methylation pattern of the F1 female can play a role in this. This is something that requires further investigation since in *Xenopus* the females are heterogametic.

*Comparison of MSAP methylated fragments between the parental species and F1 hybrids.*

Despite the fact that in both tissues the F1 hybrids derived a significant portion of the methylated fragments from the parents (males 64.5% and 52.2% of the total parental methylated liver and muscle fragments; females 61.5% of and 45.9% of muscle methylated fragments), there was still some fragment loss associated with hybridization (30.9% liver and 51.6% muscle for males; 37.4% liver and 60.1% muscle for females). Fragment loss can be due to either the nature of the AFLP markers or the phenomenon of “genome shock” which occur when two or more diverse genomes are brought together.
within the same nucleus as does occur in hybridization. It may also be possible that sequence divergence between *X. laevis* and *X. muelleri* may play a role in fragment loss. However what percentage of this loss is due to sequence divergence is difficult to estimate since there is currently no information about divergence at the genome-wide level.

All the four clusters generated over 50% methylated fragments. The F1 males generated the most with 78.8% liver methylated fragments and 73.8% muscle methylated fragments (Table 2.9). The average methylated fragment was 75.0% ± 4 and 69.1% ± 3.2 for liver and muscle tissues respectively. It should however be noted that the ability of restriction enzyme digestion is limited only to CpG methylation within 5’-CCGG-3’ sequences. Even with these sequences, these enzymes do not cut every time. As a result of this the ability to measure total cytosine methylation within the genome is limited. This was the also the shortcoming of Luminometric Methylation Assay (LUMA) reported by Pilsner et al. (2010). A much greater degree of accuracy in measuring total genome methylation is achieved with HPLC or mass spectroscopy. This involves digesting the gDNA with DNase into mononucleotides then followed by filtration before running them on a HPLC-MS. The gDNA be first treated with RNase to remove RNA contamination. This is then followed by phenol:chloroform:isoamyl extraction to clean up the sample before digesting with DNase. However what MSAP indicates is a methylation pattern across different clusters/populations that can be used to differentiate them. Furthermore with more selective primer combinations generating a larger number of MSAP fragments a better judgment can be reached as to the state of methylation
within the genome. In Pilsner et al. (2010) they found the average methylation of brain stem DNA (i.e. CpG sites) among polar bears was 57.9%, ranging from 42% to 72.4%. This is in line with the results obtained in this study. However because of the limitation already stated and the small data set and use of only one selective primer combination one cannot state with certainty that one cluster’s genome is more methylated than the other. However it gives a sense of the methylation pattern across the *Xenopus* species used in this study. Also based on this result one can expect a large percentage of methylation. This supports the hypothesis that DNA methylation helps in gene regulation, controlling transposon and associated activities. With allopolyploidization creating new nascent genomes, the genome restores order by silencing some of the duplicated genes. And one way of achieving this is through DNA methylation.

The AFLP markers are dominant multilocus markers. Both heterozygous and homozygous situation are represented by the presence of a marker (1) and the recessive by the absence (0). It is justifiable to assume that the loss is a reflection of the segregating heterozygosity in the former parental species, whereby one of the hybrids received the recessive genotypic state (Salmon et al. 2005).

In the case of “genome shock” conflicts arose between the two diverse genomes. According to Jones and Hegarty (2009) this is due to differences in genome size, genome composition, regulatory mechanisms, cell cycle duration, genetic and epigenetic modifications and indeed all aspects that contribute to harmony of the diploid genome. The nascent nucleus tries to restore order in one or more epigenetic mechanisms such as chromosome elimination, sequence loss, repatterning of transposable elements and
epigenetic changes like modification of methylation patterns. Widespread modification of methylation patterns was found in allopolyploids of *A. suecica* (Madlung et al. 2002 reported in Jones and Hegarty 2009). Song et al. (1995) reported that in synthetic allopolyploid *Brassica* there is sequence elimination directed at the paternal nuclear genome. Salmon et al. (2005) showed that sequence elimination in hybrid genome was directed at the maternal genome. The polyploid hybrids of *Spartina* used in this study were formed by hybridization 150 years ago in Europe.

In this study the loss of methylated parental fragments was biased towards both paternal and maternal genomes depending on the sex of the hybrid for both the liver and muscle tissues. However, only the proportion of fragment loss for the liver was statistically significant. This loss can be explained by either the structural changes such as those detected by standard AFLP analysis (mutations in the restriction sites) or from methylation changes undetected by MSAP (e.g. hypermethylation) (Salmon et al. 2005) or hypomethylation. At this junction one can only speculate as to the correct reason for fragment loss in the F1 hybrids.

Stancheva et al. (2002) reported that regulation of gene expression during development of *Xenopus* embryos is accompanied by a remodeling of methylation patterns at specific loci. This can explain the appearance of unique MSAP markers in the F1 hybrids.
Polyploidy can lead to transposon remobilization in the genome, which can change gene order and even genome size in hybrids (Ungerer et al 2006). Hybridization can lead to either an increase or decrease in genome size in comparison to the parents. Ungerer et al (2006) found that each of the three hybrid taxon of the genus *Helianthus* have at least 50% larger genome in comparison to either of the parental species, and this difference is at least partly accounted for by mobile elements. This is in addition to novel karyotypic rearrangements found in the hybrid sunflowers. Genome size is very important in affecting many basic ontogenetic and physiological processes e.g. rate of development, metabolic rate, and duration of cell division or complexity of life cycles (Horner and Macgregor 1983; Vinogradov 1995, 1997; Jockusch 1997; Gregory 2002; Kozlowski et al. 2003). Since cell size and genome size are positively correlated, it is not exactly clear whether the above characteristics are affected by genome size (GS) or cell size (CS).

There are differing theories that try to explain the correlation seen between cell size and genome size. The nucleoskeletal theory proposed by Cavalier-Smith (1978) predicts that nuclear size (presumably directly correlated to DNA amount) is secondarily selected after changes in cell size to meet the demands of a cell balanced growth and
function. Despite the fact that changes in genome size could sometimes precede changes in cell size, the two should still co-evolve to reach an optimal karyoplasmic ratio (Cavalier-Smith, 2005). According to the alternative scenario, the mutation pressure theory, the relationship between genome size and cell size is purely coincidental. The mutation pressure theory is criticized for its inability to explain GS-CS correlations (Gregory 2001b; Cavalier-Smith 2005). Starostova et al (2008) in their study using geckos came down on the side of the mutation pressure theory. Their study matches the claim that junk or selfish DNA can accumulate in the nucleus as long as it does not negatively affect the fitness of an organism (Pagel and Johnstone, 1992) as proposed by the mutation pressure theory. Furthermore genome size increases due to an accumulation of slightly deleterious extra-DNA, which is effectively eliminated by natural selection only in populations with effectively large sizes (Lynch 2002, 2006; Yi and Streelman 2005). For organisms that have small populations (with large cells and consequently small metabolic rate) random genetic drift causes fixations and accumulations of these slightly deleterious elements (Starostova et al 2008). Since Starostova et al (2008) worked on only a small clade of geckos, they did not explicitly make the claim that non-correlation of evolutionary changes in cell size and genome size is universal.

Tymowska and Fischberg (1973) pointed out that in the Xenopus genus changes in DNA content occurred without changes in chromosome numbers in some species whereas others became polyploid. Allopolyploidization, which occurred during the evolution of the Xenopus genus, can lead to ‘permanent heterozygosity’ in which one set of alleles from each progenitor is transmitted in each gamete (Pikaard, 2001). This might
end up giving the resulting allopolyploid a significant advantage compared to its progenitor, allowing inbreeding or self-fertilization without suffering inbreeding depression (Pikaard 2001).

Flow cytometry is a high-throughput analytical tool that simultaneously detects and quantifies multiple optical properties (fluorescence, light scatter) of single particles, usually cells or nuclei labeled with fluorescence probes, as they move in a narrow liquid stream through a powerful beam of light (Kron et al. 2007). Flow cytometry is used to estimate DNA quantity of a cell or the size of the genome. Genome size refers to either the DNA content of a single chromosome set (the base number) of an organism (“monoploid” or basic genome size; Cx-value) or to the DNA content of the unreplicated haploid genome (“holoploid genome size”; C-value) (Greilhuber et al. 2005; Leitch and Bennet 2004). In polyploids the Cx-value and the C-value are not equivalent since the haploid state will contain more than the single homologous chromosome. DNA flow cytometry has been extensively applied in biomedical research to detect aneuploidy (Kawara et al. 1999), apoptosis (Vermes et al. 2000) and monitor cell cycle kinetics and its perturbations (Rabinovitch 1994), because very large populations of cells can be measured in a short time (Dolezel and Bartos 2005). Samples for use in flow cytometry are easy to prepare and the presence of subpopulations can be detected (Shapiro 2003). However, the downside to this technology is the lack of direct visual observation of the nuclei being studied and the cost of the instrument (Bennett and Leitch 2005a; Greilhuber et al. 2007). This can be overcome with the proper use of standards and attention to
operating conditions (reviewed in Greilhuber et al. 2007) to produce remarkable precise estimates that are sensitive to small variations.

This study aims the following:

1. Compare the genome size of the parental and the hybrids using a Flow Cytometer.
2. Digest the genomic DNA of the hybrids and the parents with the methyl-sensitive HpaII and its isochizomer MspI. Compare the digested samples with the genomic DNA using gel electrophoresis or a DNA Agilent Bioanalyzer. This will show whether the hybrids are undermethylated or hypermethylated.

I hypothesized that the genome size will be intermediate between the parental species whereas the hybrids will be hypermethylated.

Methods

*Compare the genome size of the parents and the hybrids.*

The protocol used for the preparation the samples for the flow cytometry was done according to method proposed by Rodgers (2006). The protocol describes a quantitative measurement of DNA in solid tissues samples using either propidium iodide (PI) or DAPI staining, which is then followed by flow cytometry. Fragmented fresh muscle tissue about 2-3 mm³ was placed in 1 ml of citrate buffer in a 60 mm Petri dish thoroughly squashed with a micropipette, and run seven times through a hypodermic syringe. The liquid was transferred to a culture tube and stained with 800 µl of DNA lysis stain and 20 µl of boiling RNase A. It was left in the dark for 20 minutes and then filtered using a 30 µm pore filter. The resulting suspension was then analyzed with a BD LSR II
flow cytometer (BD Bioscience, San Jose, California) equipped with a 480-nm air cooled argon laser. The FACS Diva™ software (BD Bioscience, San Jose, California) was used to calculate the instrument. The following runs were made:

1. The samples were analyzed individually (n = 2 for each species)
2. The samples were mixed together as follows (X. laevis-X. muelleri, X. laevis – F1 Hybrid, X. muelleri – F1 Hybrid and X. laevis-X. muelleri-F1 Hybrid) resulting in three mixtures and these were analyzed.

The genomic size for X. muelleri and X. laevis using erythrocytes is known. The goal of this experiment was to determine relative genome sizes of the hybrid compared to either parental species. No internal reference was used.

*Digest the genomic DNA of the hybrids and the parents*

Whole genomic DNA from the hybrids and parental species was digested at 37 °C for 60 minutes using two restriction enzymes *Hpa*II (methyl-sensitive) and its methyl-insensitive isoschizomer *Msp*I in parallel digestions followed by heat inactivation at 60 °C and 80 °C for 20 minutes. The digested samples were then filtered and compared to the whole genomic DNA using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, California). A 7500 bp assay chip was used. Less than 50 ng/µl of DNA was used. The standard chip loading protocol from the manufacturing was used with one modification: after the DNA was loaded in was pipetted up and down to obtain a uniform mixture with the marker before vortexing for one minute at the highest setting (Panaro et al. 2000). The Biosizing software (2100 expert), ver B.02.07 was used to analyze the data.
The methylation status of the parental and F1 hybrid genomic DNA was determined by comparing the electrophenograms and looking for smear patterns. A smear in the \textit{MspI} lane is an indication of complete digestion whilst a band of uncut DNA at the top (no smear peaks in the electrophenogram) is an indication of extensive methylation at the 5’-CCGG-3’ sites. Smearing in both lanes is an indication of heavy undermethylation (O’Neill et al. 1998).

**Results**

*Genome Size*

The mean fluorescence (PI-A) of the samples is shown in Table 3.1. The mean fluorescence (in arbitrary units) for \textit{X. laevis} (n = 2) is 70445, \textit{X. muelleri} (n=2) is 799332.17 and for hybrids (n = 2) is 79115.5. From the results the peak of \textit{X. l. laevis} is located at a lower fluorescence when compared to \textit{X. muelleri}. The peak of the hybrid appears to fall between the peaks of \textit{Xenopus l. laevis} and \textit{X. muelleri}. However it appears to be closer to \textit{X. muelleri} since the peaks of the hybrid and \textit{X. muelleri} merge and only show one peak when the two samples are combined. The genome size of \textit{X. muelleri} is 113.5\% of \textit{X. laevis} whereas that of the F1 hybrid is 112.3\%.
Table 3.1 Mean Fluorescence (PI-A) for three different measurements

<table>
<thead>
<tr>
<th>Samples</th>
<th>Mean Flourescence (PI-A)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>X. laevis-3 (female)</td>
<td>71599</td>
</tr>
<tr>
<td>X. laevis-4 (female)</td>
<td>68084</td>
</tr>
<tr>
<td>X. muelleri-3 (female)</td>
<td>80121</td>
</tr>
<tr>
<td>X. muelleri-4 (male)</td>
<td>78879</td>
</tr>
<tr>
<td>Hybrid-3 (X. laevis x X. muelleri) -female</td>
<td>80433</td>
</tr>
<tr>
<td>Hybrid-4 (X. laevis x X. muelleri) -male</td>
<td>77183</td>
</tr>
</tbody>
</table>

Table 3.2 Mean fluorescence measurement for all individuals of the same species combined

<table>
<thead>
<tr>
<th>Specimen</th>
<th>No. of specimens</th>
<th>DNA content (Mean arbitrary units)</th>
<th>DNA content (% of X. laevis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X. Laevis</td>
<td>2</td>
<td>70448.5</td>
<td>100</td>
</tr>
<tr>
<td>X. muelleri</td>
<td>2</td>
<td>79932.17</td>
<td>113.5</td>
</tr>
<tr>
<td>F1 (X. laevis x X. muelleri)</td>
<td>2</td>
<td>79115.5</td>
<td>112.3</td>
</tr>
</tbody>
</table>
Table 3.3 Fluorescence (PI-A) for the mixed samples

<table>
<thead>
<tr>
<th></th>
<th>Peak 1</th>
<th>Peak 1</th>
<th>Peak 2</th>
<th>Peak 2</th>
<th>Peak 1</th>
<th>Peak 2</th>
<th>Peak 1</th>
<th>Peak 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X. laevis-X. muelleri</td>
<td>68603</td>
<td>81039</td>
<td>63839</td>
<td>80796</td>
<td>68390</td>
<td>80628</td>
<td>68454</td>
<td>80821</td>
</tr>
<tr>
<td>X. laevis-Hybrid</td>
<td>65694</td>
<td>75655</td>
<td>66737</td>
<td>76466</td>
<td>67014</td>
<td>76567</td>
<td>66481.67</td>
<td>76229.3</td>
</tr>
<tr>
<td>X. muelleri-Hybrid</td>
<td>0</td>
<td>79774</td>
<td>0</td>
<td>79878</td>
<td>0</td>
<td>80027</td>
<td>0</td>
<td>79893</td>
</tr>
<tr>
<td>X. laevis-X. muelleri-Hybrid</td>
<td>67813</td>
<td>79503</td>
<td>68289</td>
<td>79964</td>
<td>68197</td>
<td>79984</td>
<td>68099.67</td>
<td>79817</td>
</tr>
</tbody>
</table>
Figure 3.1 Histogram of the mean PI-A measurements
Figure 3.2 Flow cytometric histograms of fluorescence intensity for a) *X. laevis* and b) *X. muelleri*
Figure 3.3 Flow cytometric histograms of fluorescence intensity for a) F1 hybrid and b) *X. laevis/muelleri*
Figure 3.4 Flow cytometric histograms of fluorescence intensity for a) *X. laevis*/F1 hybrid and b) *X. muelleri*/F1 Hybrid
Figure 3.5 Flow cytometric histograms of fluorescence intensity for *X. laevis/muelleri* F1 hybrid
The analysis genome methylation

A summary of the peaks obtained after digestion and compared to the genomic DNA are shown in Table 3.4 and Figure 3.7 to 3.12. Regions of smear were identified and classified into regions. The difference in smear pattern between \textit{HpaII} digestion and \textit{MspI} digestion were compared to the gDNA. With the exception of the \textit{X. laevis} muscle none of the \textit{HpaII}/\textit{MspI} digestion for the \textit{X. muelleri} and the F1 hybrid has any smear pattern. The smear pattern obtained for the \textit{X. laevis} was ruled out because there was smear in the electrophenogram of the gDNA which overlap with that of the \textit{HpaII} and \textit{MspI} digestion. This may therefore indicate degraded DNA around the 10,000 bp. For the liver tissue there was no smearing for the \textit{X. laevis} and \textit{X. muelleri} but with the F1 liver there was smearing in both the \textit{HpaII} and the \textit{MspI} digestion. However the smearing for the \textit{HpaII} was excluded because it overlapped with the smear of the gDNA. The smearing in \textit{MspI} may be an indication of undermethylation.
Table 3.4 Summary of Regions identified as DNA smear for Liver and muscle

### Liver

<table>
<thead>
<tr>
<th></th>
<th>XL-2-110-male liver</th>
<th>XM-3-809 male liver</th>
<th>F1-B-410 male liver</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gDNA</td>
<td>HpaII</td>
<td>MspI</td>
</tr>
<tr>
<td>From (s)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>To (s)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Area</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>% of Total</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average Size (bp)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Size distribution</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Conc (ng/ul)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Muscle

<table>
<thead>
<tr>
<th></th>
<th>XL-2-110-male muscle</th>
<th>XM-3-809 male muscle</th>
<th>F1-B-410 male muscle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gDNA</td>
<td>HpaII</td>
<td>MspI</td>
</tr>
<tr>
<td>From (s)</td>
<td>88.93</td>
<td>80.31</td>
<td>75.53</td>
</tr>
<tr>
<td>To (s)</td>
<td>96.81</td>
<td>96.53</td>
<td>100.36</td>
</tr>
<tr>
<td>Area</td>
<td>32.9</td>
<td>6.9</td>
<td>5.6</td>
</tr>
<tr>
<td>% of Total</td>
<td>62</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>Average Size (bp)</td>
<td>14,179</td>
<td>10,719</td>
<td>10,463</td>
</tr>
<tr>
<td>Size distribution</td>
<td>14.4</td>
<td>27.3</td>
<td>43.3</td>
</tr>
<tr>
<td>Conc (ng/ul)</td>
<td>1.96</td>
<td>2.57</td>
<td>4.78</td>
</tr>
</tbody>
</table>
Figure 3.6 Gel Electrophoresis for a) gDNA and b) Restriction digest products
Figure 3.7 Electrophenogram for male muscle tissues (X. laevis). There is smearing around the 10000 bp (upper marker) for the gDNA and the Msp1 and HpaII digest. Since there is overlap, it was considered not reliable.
Figure 3.8 Electrophenogram for male Muscle Tissues (*X. muelleri*). The smearing of *Msp1* overlaps that of the gDNA, hence it is a probably a carry-over so it was not considered significant.
Figure 3.9 Electrophenogram for male Muscle Tissues (F1 Hybrid).
F1 male hybrid muscle. No smearing observed with the digest from the restriction enzymes. It was therefore considered as highly methylated.
Figure 3.10 Electrophenogram for male Liver Tissues (*X. laevis*)
No smearing observed for the gDNA or the digests from the restriction enzymes
Figure 3.11 Electrophenogram for male Liver Tissues (*X. muelleri*)
No smearing observed with the digests from the two restriction enzymes.
Figure 3.12 Electrophenogram for male Liver Tissues (F1 Hybrid)
There was smearing in both digests from HpaII and MspI.
Discussion

Comparison of genome sizes between parental species and their hybrids

Genome size presents a parameter by which a genome may be characterized (Thiebaud & Fischberg 1977). Even though X. laevis and X. muelleri are both tetraploid with 36 chromosomes, they have different genome lengths (Thiebaud & Fischberg 1977). There is evidence for genome downsizing across major clades of flowering plants with increase in ploidy levels (Leitch and Bennett 2004). This pattern suggests that a genuine loss of DNA accompanies genome duplication in many organisms (Kron et al. 2007). The results are comparable to estimates for genome size using erythrocytes (X. laevis and X. muelleri). This study also points to the fact that the genome size of X. muelleri is slightly larger than that of X. laevis. However, it is interesting that the genome size of the hybrid is slightly larger than that of X. laevis but closer to that of X. muelleri. However due to the small sample size, one has to treat this information with caution.

Genome Methylation

De novo chromosome changes such as genome-wide undermethylation, retro-viral element amplification and chromosome remodeling is very common in other interspecies hybrids (O’Neill et al. 1998). O’Neill et al. (1998) in their study of several interspecific mammalian hybrid (Macropus eugenii x Wallabia bicolor) found that the hybrids are undermethylated. This was also the finding in this study with respect to liver tissues. The F1 hybrids that did survive in the crosses conducted by the Michalak lab and used in this study are phenotypically normal with a few exceptions that show some form of
deformities. However, all hybrid males are sterile. Some F1 hybrid females were used in backcrosses and they produced viable offspring.

The idea that DNA methylation arose as a result to defend the host genome against the deleterious effects of transposable elements (TE) is put to the test in these scenarios. However O’Neill (1998) suggested that there may be other alternative mechanisms of genomic methylation, which resulted in developmentally important genes in the interspecies marsupial they studied to be unaffected despite the deficient methylation of the genome and activation of retroelements, which usually lead to gross changes in genome structure.
CHAPTER 4

DISCUSSION

DNA methylation has evolved not only as a means of controlling of transcription of genes involved in development (Kakutani et al. 2004) but also to protect the genome from the deleterious effects of transposons (Yoder et al. 1997; Matzke et al. 1999; Selker et al. 2003). Accordingly, undermethylation might pose a serious problem for F1 hybrids because of the plausible eruption of transposons as a result of hybridization. In the absence of any mechanism for controlling this, there can be a disruption of the mechanism involved in gene transcription. In Arabidopsis ddm1 (a gene necessary for DNA methylation), DNA hypomethylation mutation results in a variety of developmental abnormalities inducing heritable changes in other loci. Some of the hybrids produced in our lab had noticeable phenotypic deformities and the survival from tadpoles to adult was very difficult (Michalak pers com.). This might be due to undermethylation in the hybrids. The loss of DNA methylation may have induced developmental abnormalities through transposon mobilization and perturbation of transcription as was observed in Arabidopsis (Kakutani et al. 2004).

The methylation pattern of the hybrids is unique and completely different from that of the parental species. Methylation profile is not inherited from the germ line (Kafri et al. 1992). The parental methylation pattern is erased before implantation and the
embryo has de-novo methylases that can modify the entire genome (Okano et al 1999). In *Xenopus laevis* the paternal genome is not subjected to active demethylation of 5-methyl cytosine immediately after fertilization (Stancheva et al. 2002). This study however looked at the embryos from parents of the same species. It did not consider what happens within the embryos of interspecies hybrids when differing genomes are brought together causing enormous disruption. This can lead to disruption in chromosome pairing during meiosis leading to defective gametes.

In *Xenopus* the males from interspecies F1 hybrids are evolutionary dead-ends because they are sterile both in natural pairings and when their testis were used for artificial insemination. Because the differentiation of the eggs precedes meiosis the viability of the female gametes remains intact (Kobel 1981). The females can therefore produce fertile eggs and can interbreed with the parents. Production of interspecies *Xenopus* hybrids even though can be achieved in laboratory settings is very uncommon in the wild. It has been proposed that these species are ecologically separated. They can live sympatrically but are found in different ecological environments (Fischer et al 2000).

The results presented here indicate that methylation changes may occur very early in interspecies hybridization. The parental species genome having evolved over a very long period of time, have been able to deal with the internal chaos arising from allopolyplodization. One way they may have done this is through DNA methylation.

Also, the results indicate that there is a substantial proportion of methylated fragments within the *Xenopus* genome. In general, polyploids produce a large number of AFLP fragments (Fay et al 2005; Guo et al 2006; Kardolus et al 1998). The error rate
related to MSAP fragments in this study was much higher than 2-5% that has traditionally been reported for AFLP. This is also what Holland et al (2008) found in their study. They obtained error rates ranging from 9% to 18% (*Ipomoea*) and 6% to 13% (*Ourisia*). They suggested that one has to be cautious in making inter-study error rates comparisons. The divergence of individuals, the study sample size, errors resulting from the raw AFLP profiles (e.g. PCR errors) and from the scoring process (type of procedure and software employed) can all affect the error rates. This study indicated inter-tissue differences in methylation patterns. These differences are however very small. How significant these differences are needs to be further investigated. DNA methylation patterns appear to be consistent across different kinds of tissue types for a given species (Dolinoy et al. 2006; Eckhardt et al. 2006). This variation may suggest that the usefulness of MSAP fragments in answering different molecular questions will depend on the tissue, age and the sex of the individuals. It is however noteworthy that age-dependent variation in methylation at the MSAP loci seems to be rare (Verhoeven et al. 2010).

There can however be no doubt that there is some sexual dimorphism in DNA methylation. Pogribny et al. (2004) reported sex- and tissue-specific radiation induced methylation. Mouse liver tissues were hypomethylated in the females and not in males. However the spleen in both sexes was hypomethylated but this was more pronounced in females. DNA hypomethylation was observed in males F1 hybrids. It will be very important to investigate the same phenomenon in F1 females and compare the amount of methylation between the sexes. According to Adams et al. (2003) subfunctionalization is
a consequence of allopolyploid formation, because of the differential expression patterns between homeologous genes in different tissues.

The methylation effects of hybridization and allopolyploidization (in which two diverged genomes are brought together) are usually more pronounced than the effects of genome doubling per se (Salmon et al. 2005; Wang et al. 2006). This study detected no genome doubling in the F1 hybrids. Instead the genome size of the hybrids is intermediate between the parental species. This is in agreement with the findings of Fischer et al. (2000).

The Principal Coordinate Analysis plot also revealed the same information with hybrids forming clusters between the parents indicating that they carry the properties from both parents.

There is loss of methylated fragments from the parental species in the F1 hybrids. This loss is directed at both parent species. It is biased to both the maternal and paternal genome depending on the sex of the F1 in both the liver and the muscle tissues. The bringing together of two diverse genomes results in a phenomenon that McClintock (1984) refer to as "genome shock", she defined as the result of an extensive restructuring that take place due to a preprogrammed response to an unusual challenge. Madlung et al. (2002) further postulated that the "unusual challenge" may involve epigenetic gene silencing. The silencing can be as a result of homologous DNA-DNA or DNA-RNA interactions. This marriage of two genomes in the F1 can lead to genome instability (Jones and Hegarty 2009). This instability can lead to genomic rearrangement (Madlung et al. 2002). Ozkan et al. (2001) and Shaked et al. (2001) both found rapid and
widespread loss of DNA sequences and changes in DNA methylation in synthetic hybrids of wheat (*Triticum aestivum*).

In 1876 Charles Darwin systematically characterized the phenomenon known today as hybrid vigor or heterosis. This has been defined as heterozygotes have a higher fitness in a population than the homozygotes (Chen 2010). According to Pikaard (2001) hybridization, establishment of duplicate-gene cytotypes, can result in permanent heterozygosity especially if the duplicate gene copies confer similar function. The increased genetic variability in polyploids or hybrids can lead to exploitation of new habitats wherein certain combinations of alleles may be more favorable. There is very little evidence for natural hybridization of *Xenopus* (Fischer et al. 2000), despite the phenomenon of heterosis. Poynton and Broadley (1985) postulated that in even in the zones of sympatry, *Xenopus* frogs populations are not necessarily in actual contact.

The results presented in this study collectively point to the high degree of genetic similarity in the methylated markers between the F1 hybrids and the parental species. This high degree of similarity may be as result of retention of highly conserved methylation markers in the F1 hybrid from the parental species. This study also revealed unique MSAP markers in the F1 hybrid. The occurrence of methylation repatterning after merging two diverse genomes is well established (Wendel 2005). Therefore the appearance of unique/de novo MSAP markers in the F1 hybrids continue to add to the knowledge that hybridization can trigger *de novo* methylation patterns in the newly formed hybrids.
This study left several unanswered questions that needed to be addressed in the near future. The following are some of these questions: What is the global methylation of the F1 and parental species. Is the heterozygosity seen in this study significant across different tissues such as the gonads, brain, liver and muscle? It may also be interesting to investigate any association between the methylation level in a tissue and phenotypic outcomes. Also it would be very interesting to clone certain methylated bands and see if there is any association between them and genes or phenotype.
APPENDIX A

CHECKING MSAP FRAGMENT DATA QUALITY
<table>
<thead>
<tr>
<th>Samples</th>
<th>100.4</th>
<th>100.9</th>
<th>101.4</th>
<th>102</th>
<th>103.1</th>
<th>103.6</th>
<th>104.2</th>
<th>104.7</th>
<th>105.8</th>
<th>106.9</th>
<th>107.9</th>
<th>108.7</th>
<th>111</th>
<th># of Frag.</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-A-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>F1-A-2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>F1-B-1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>F1-B-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>F1-C-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>F1-C-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

B: New Table to be used in analysis

<table>
<thead>
<tr>
<th>Samples</th>
<th>100.4</th>
<th>100.9</th>
<th>101.4</th>
<th>102</th>
<th>103.1</th>
<th>103.6</th>
<th>104.2</th>
<th>104.7</th>
<th>105.8</th>
<th>106.9</th>
<th>107.9</th>
<th>108.7</th>
<th>111</th>
<th># of Frag.</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-A</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>25%</td>
</tr>
<tr>
<td>F1-B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>F1-C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>50%</td>
<td></td>
</tr>
</tbody>
</table>
C: TABLE COMPARING MSAP FRAGMENTS FROM \textit{EcoR}/\textit{HpaII} AND \textit{EcoRI}/\textit{MspI}

<table>
<thead>
<tr>
<th>Samples</th>
<th>100.4</th>
<th>100.9</th>
<th>101.4</th>
<th>102</th>
<th>103.1</th>
<th>103.6</th>
<th>104.2</th>
<th>104.7</th>
<th>105.8</th>
<th>106.9</th>
<th>107.9</th>
<th>108.7</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1-A (HpaII)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F1-A (MspI)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>F1-A (MSAP fragments)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F1-B (HpaII)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>F1-B (MspI)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>F1-B (MSAP fragments)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F1-C (HpaII)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>F1-C (MspI)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F1-C (MSAP fragments)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

The AFLP data for two independent runs for the same individual are compared.

If a fragment is present in both runs it is automatically accepted as good and given a 1 in a new profile for this individual.

If only one run indicates the presence of the allele, it is then compared to all the runs for that group or population.

If any other run has the fragment duplicated or two-thirds of the runs contain the fragment then it will be accepted

A new profile is then generated with all the fragments that were duplicated.

The fragments that were accepted but were not duplicated are used to calculate the error rate as follows: The ratio of these fragments to the fragments that were accepted.
APPENDIX B

FORMULARS USE TO PERFORM THE STATISTICAL CALCULATION
Allele frequency (Lynch and Milligan 1994)

- Assuming random mating: Presence represents both genotypes AA and Aa
- Absence represents genotype aa
- Allele A has Freq. \( p = 1 - q \)
- Allele a has Freq. \( q = 1 - p \)
- Frequency of genotype aa = \( q^2 \) = Freq. of band absence = 1 – Freq. of band presence
  
  So \( q = \sqrt{\text{Freq of band absence}} \)

Genetic Distance (GD) (Huff et al. 1993; Maguire et al 2002)

\[
D = n \left[ 1 - \frac{2n_{xy}}{2n} \right]
\]

\( 2n_{xy} \) = the number of shared character states and \( n \) equals the total number of binary characters. When calculated across multiple loci for a given pair of samples, this is equivalent to the tally of state differences among the two DNA profiles

Expected Heterozygosity (He) of Genetic Diversity for Binary data (Maguire et al. 2002; Lynch and Milligan 1994)

- Assuming random mating:
- Presence represents both genotypes AA and Aa
- Absence represents the genotype aa
- Allele A has Freq. \( p = 1 - q \)
- Allele a has Freq. \( q = 1 - p \)
- Frequency of genotype aa = \( q^2 \) = Freq. of band absence = 1 – Freq. of band presence
  
  So \( q = \sqrt{\text{Freq of band absence}} \)

\[
He = 1 - \sum p^2
\]

Nei’s Genetic Distance (Nei D) (Hendrick 2000)

\[
Nei_D = -\ln(I)
\]
- Where \( I \) is Nei’s genetic identity
Nei’s Genetic Identity (Nei I) (Hendrick 2000)

$$\text{Nei}_I = \frac{I_{xy}}{\sqrt{(I_{xy})}}$$

$$I_{xy} = \sum_{i=1}^{k}(P_{ix}P_{iy}) ; \quad I_x = \sum_{i=1}^{k}P_{ix}^2 ; \quad I_y = \sum_{i=1}^{k}P_{iy}^2$$

- \( P_{ix} \) and \( P_{iy} \) are the frequencies of the i-th allele in populations x and y.
- For multiple loci, \( I_{xy} \), \( I_x \) and \( I_y \) are calculated by summing over all loci and alleles and dividing by the number of loci. These average values are then used to calculate I

PhiPT via AMOVA without regional data (Peakall 1995)

$$\varphi_{PT} = \frac{V_{AP}}{(V_{AP} + V_{WP})}$$

\( V_{AP} \) = variance among populations
\( V_{WP} \) = variance within populations
PhiPT = correlation between individuals within a population, relative to the total.
REFERENCES


cotton is not accompanied by rapid genomic changes. Genome 44(3):321-30.

Lopez Pajares I, Delicado A, Diaz de Bustamante A, Pellicer A, Pinel I, Pardo M, Martin

Lupski JR, Deocaluna RM, Slaugenhaupt S, Pentao L, Guzzetta V, Trask BJ,
Saucedocardenas O, Barker DF, Killian JM, Garcia CA, et al. 1991. DNA duplication


A 99(9):6118-23.


Remodeling of DNA methylation and phenotypic and transcriptional changes in

mangrove species Avicennia marina (forsk.) vierh. (Avicenniaceae) detected by
AFLPs and SSRs. Theor Appl Genet 104(2-3):388-98.

Makowsky R, Chesser J, Rissler LJ. 2009. A striking lack of genetic diversity across the
wide-ranging amphibian Gastrophyryne carolinensis (anura: Microhylidae). Genetica
135(2):169-83.


Picker MD and Harrison, J.A. & Wallace, D. 1996. Natural hybridization between

*Xenopus laevis laevis* and *X. gilli* in the south-western cape province, South Africa.

17(12):675-7.


Mercury-associated DNA hypomethylation in polar bear brains via the LUmominometric methylation assay: A sensitive method to study epigenetics in wildlife. Mol Ecol


Biochemical and Biophysical Research Communications 320(4): 1253-1261.

Prokhortchouk E and Defossez P. 2008. The cell biology of DNA methylation in 
mammals. Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research
1783(11):2167-73.

Rabinovitch PS. 1994. DNA content histogram and cell-cycle analysis. Methods in Cell
Biology 41:263-96.


Tweedie S, Charlton J, Clark V, Bird A. 1997. Methylation of genomes and genes at the 

Chromosoma 44(3):335-42.

Ungerer MC, Strakosh SC, Zhen Y. 2006. Genome expansion in three hybrid sunflower 
species is associated with retrotransposon proliferation. Current Biology 16(20):R872-3.

Vandepoele K, Simillion C, Van de Peer Y. 2002. Detecting the undetectable: 

Verhoeven KJF, Van Dijk PJ, Biere A. 2010. Changes in genomic methylation patterns 


Vinogradov AE. 1997. Nucleotypic effect in homeotherms: Body-mass independent 
resting metabolic rate of passerine birds is related to genome size. Evolution 51(1):220-5.

Vinogradov AE. 1995. Nucleotypic effect in homeotherms: Body-mass-corrected basal 


BIOGRAPHICAL INFORMATION

Alie Patrick Koroma was born to Alie Paul Koroma and Janet Ballay Koroma in Lunsar Sierra Leone. He received his Bachelor of Science with honors in Zoology from Fourah Bay College, University of Sierra Leone, Sierra Leone in 1991. He then received his first Master’s degree in Tropical Forestry from Dresden University of Technology, Germany in 1997 and the second Masters from Texas Christian University Fort Worth Texas, U.S.A. in 2001. Alie received his Doctoral degree from the University of Texas at Arlington, U.S.A. in 2010. Alie has worked on a variety of projects in conservation and teaching. He served as the National Coordinate for the Important Bird Areas of Sierra Leone from 1993 to 1995. He has coauthored a book on the Important Bird Areas of Sierra Leone and articles. His current research interest is on epigenetics. Alie has served on civil organizations in Sierra Leone and in the United States. He is an active member of the City of Rowlett Community Emergency Response Team (CERT). His hobbies are reading, travel, bird watching and playing soccer. He is married with four daughters.