

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

Copyright © by Alie Patrick Koroma 2010

All Rights Reserved

## 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 assistance 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 Betrán, 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 assistance 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 helped 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% ( $75.0 \pm 4$ ; mean  $\pm$  SD) and 67.0% to 73.8% ( $69.1 \pm 3.2$ ; mean  $\pm$  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
<i>Xenopus</i> .....	6
2. METHYLATION PATTERNS AMONG HYBRIDS AND PARENTS ( <i>Xenopus laevis</i> and <i>Xenopus muelleri</i> ) 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 ( <i>X. l. laevis</i> and <i>X. muelleri</i> ) 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

Figure	Page
2.1	Diagrammatic representation of the AFLP Technique.....17
2.2	Summary graph showing the results of methylated band (alleles) patterns across populations [a) liver and b) muscle tissues].....22
2.3	Principal Coordinates Analysis (PCoA) plot of MSAFLP for all individual pair-wise genetic distances [liver tissue] - PCoA via Covariance with data standardization.....25
2.4	Principal Coordinates Analysis (PCoA) plot of MSAFLP for all individual pair-wise genetic distances [muscle tissue] - PCoA via Covariance with data standardization.....26
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].....29
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 tissues].....33
2.7	Liver MSAP fragments observed in individuals from <i>X. laevis</i> , <i>X. muelleri</i> and F1 hybrid males and females.....39
2.8	Muscle MSAP fragments observed in individuals from <i>X. laevis</i> , <i>X. muelleri</i> and F1 hybrid males and females.....40
3.1	Histogram of the mean PI-A measurements .....57
3.2	Flow cytometric histograms of fluorescence intensity for a) <i>X. laevis</i> and b) <i>X. muelleri</i> .....58
3.3	Flow cytometric histograms of fluorescence intensity for a) F1 hybrid and b) <i>X. laevis/muelleri</i> .....59

3.4	Flow cytometric histograms of fluorescence intensity for a) <i>X. laevis</i> /F1 hybrid and b) <i>X. muelleri</i> /F1 Hybrid.....	60
3.5	Flow cytometric histograms of fluorescence intensity for <i>X. laevis</i> / <i>muelleri</i> / F1 hybrid.....	61
3.6	Gel electrophoresis for a) gDNA and b) restriction digest products.....	64
3.7	Electrophenogram for male muscle tissues ( <i>X. laevis</i> ).....	65
3.8	Electrophenogram for male muscle tissues ( <i>X. muelleri</i> ).....	66
3.9	Electrophenogram for male muscle tissues (F1 hybrid).....	67
3.10	Electrophenogram for male liver tissues ( <i>X. laevis</i> ).....	68
3.11	Electrophenogram for male liver tissues ( <i>X. muelleri</i> ).....	69
3.12	Electrophenogram for male liver tissues (F1 hybrid).....	70

## LIST OF TABLES

Table		Page
2.1	Descriptive statistics of AFLP loci from liver and muscle tissues [band frequencies, estimated allele frequencies and estimated heterozygosity by population for binary (diploid) data].....	21
2.2	Summary statistics for Principal Coordinate Analysis (PCoA) for all individuals [liver and muscle tissues].....	24
2.3	Results of AMOVA performed on all populations identified as clusters [liver and muscle tissues].....	28
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).....	31
2.5	Results of pair-wise population unbiased Nei genetic distance and unbiased Nei genetic identity values as [liver and muscle tissues].....	32
2.6	Primers used in MSAP.....	36
2.7	Methylated MSAP fragments observed in parents ( <i>X. laevis</i> and <i>X. muelleri</i> ) and hybrids ( <i>X. laevis</i> x <i>X. muelleri</i> ) generated by <i>EcoRI-HpaII</i> and <i>EcoR1-Msp1</i> digestion (one selective primer combination)- [liver].....	37
2.8	Methylated MSAP fragments observed in parents ( <i>X. laevis</i> and <i>X. muelleri</i> ) and hybrids ( <i>X. laevis</i> x <i>X. muelleri</i> ) generated by <i>EcoRI-HpaII</i> and <i>EcoR1-Msp1</i> digestion (one selective primer combination - [muscle Tissues] .....	41
2.9	Percentage of Methylated Fragment Liver and Muscle.....	41

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-Fernàndez 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 allopolyploidization and range from diploid (n=20) to dodecaploid (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.

Polyploid 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).

AFLP is 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 were 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 bred 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 cytosine methylation. *HpaII* will not cleave if internal cytosine is methylated and *MspI* is only sensitive to methylation of the external cytosine on both strands. If either the external cytosine 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

cytosine 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 cytosine.

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 were 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 homoplasmy 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).

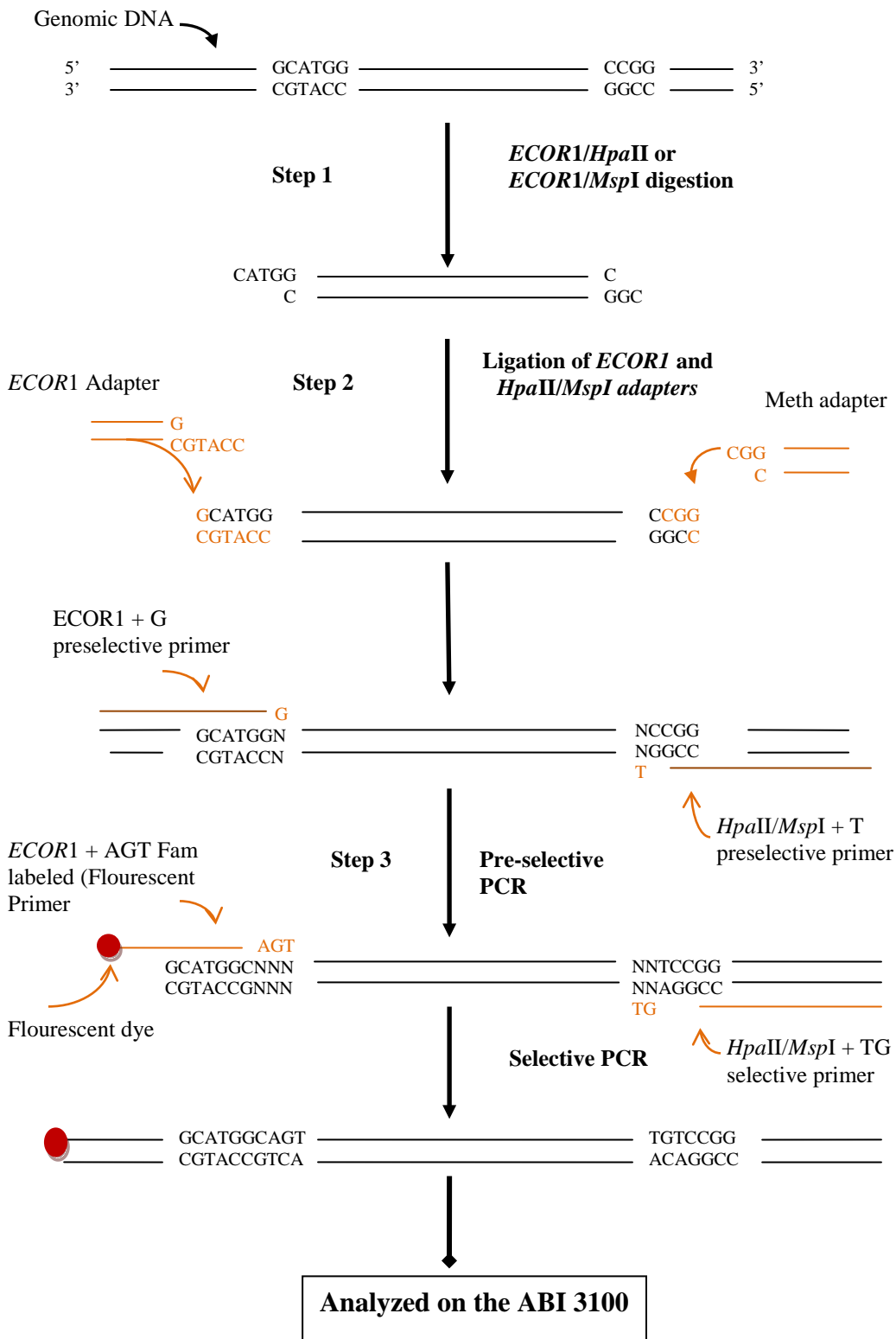


Figure 2.1 Diagrammatic representation of the AFLP Technique

### Data quality/genotyping error

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 group 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.

### Cluster Analysis/Population Structure

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 markers 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 2.1 Descriptive Statistics of AFLP Loci from Liver and Muscle Tissues [Band Frequencies, Estimated Allele Frequencies and Estimated Heterozygosity by Pop for Binary (Diploid) Data]

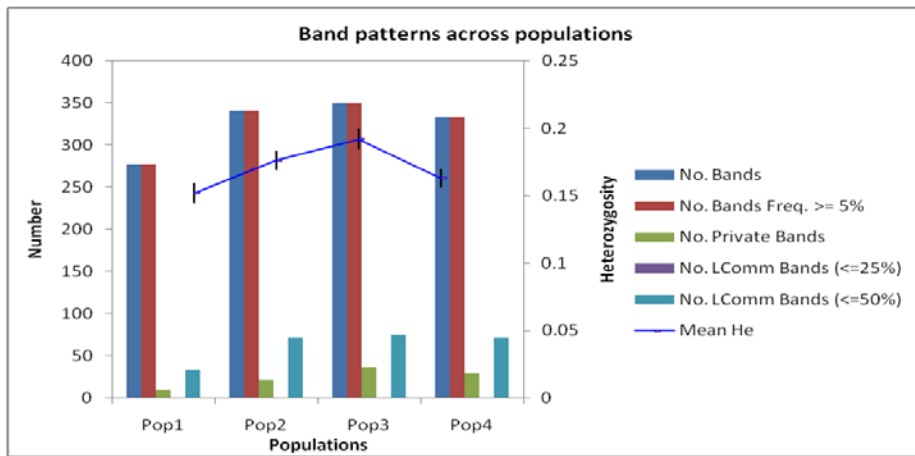
Liver

Pop	n	No. of Loci	Na (SE)	Ne (SE)	I (SE)	He (SE)	Uhe (SE)	Percent Polymorphic Loci
F1-female	6.000	277	1.091 0.044	1.234 0.013	0.241 0.011	0.152 0.007	0.166 0.008	54.17%
F1-male	8.000	341	1.343 0.042	1.270 0.013	0.282 0.011	0.176 0.007	0.188 0.008	66.67%
<i>X. laevis</i>	8.000	350	1.381 0.041	1.303 0.014	0.302 0.011	0.192 0.008	0.205 0.008	68.65%
<i>X. muelleri</i>	8.000	333	1.321 0.042	1.242 0.012	0.265 0.010	0.163 0.007	0.174 0.007	66.07%
TOTAL (SE)	7.500 0.019	504	1.284 0.021	1.262 0.006	0.272 0.005	0.171 0.004	0.183 0.004	63.89% 3.29%

Muscle

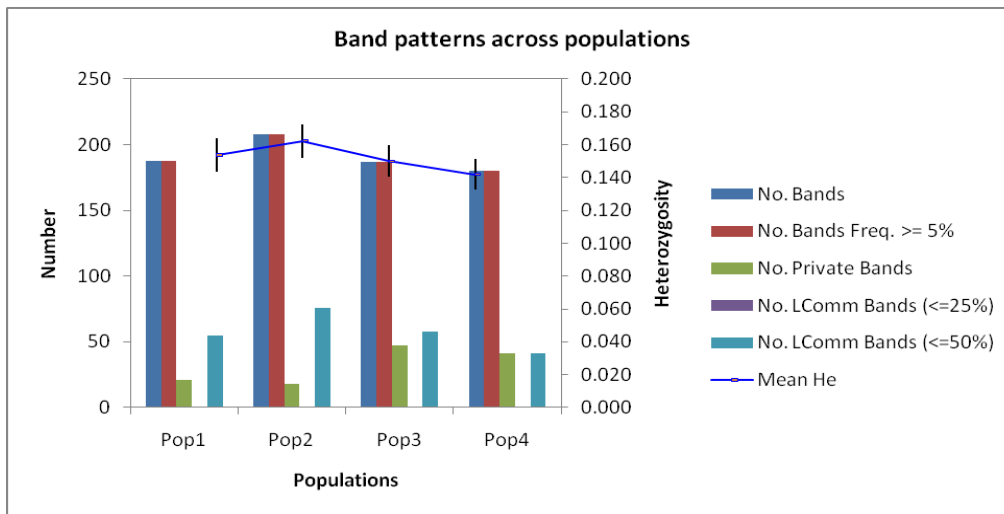
Pop	n	No. of Loci	Na (SE)	Ne (SE)	I (SE)	He (SE)	Uhe (SE)	Percent Polymorphic Loci
F1-female	4.000	188	0.956 0.050	1.259 0.018	0.232 0.014	0.154 0.010	0.176 0.011	43.96%
F1-male	4.000	208	1.030 0.049	1.274 0.019	0.244 0.015	0.162 0.010	0.185 0.012	45.88%
<i>X. laevis</i>	4.000	187	0.953 0.050	1.249 0.018	0.228 0.014	0.150 0.010	0.171 0.011	43.96%
<i>X. muelleri</i>	4.000	180	0.926 0.051	1.230 0.017	0.217 0.014	0.142 0.009	0.162 0.011	43.13%
TOTAL (SE)	4.000 0.000	364	0.966 0.025	1.253 0.009	0.230 0.007	0.152 0.005	0.174 0.006	44.23% 0.58%

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,  $q = (1 - \text{Band Freq.})^{0.5}$  and  $p = 1 - q$ .



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*

a



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*

b

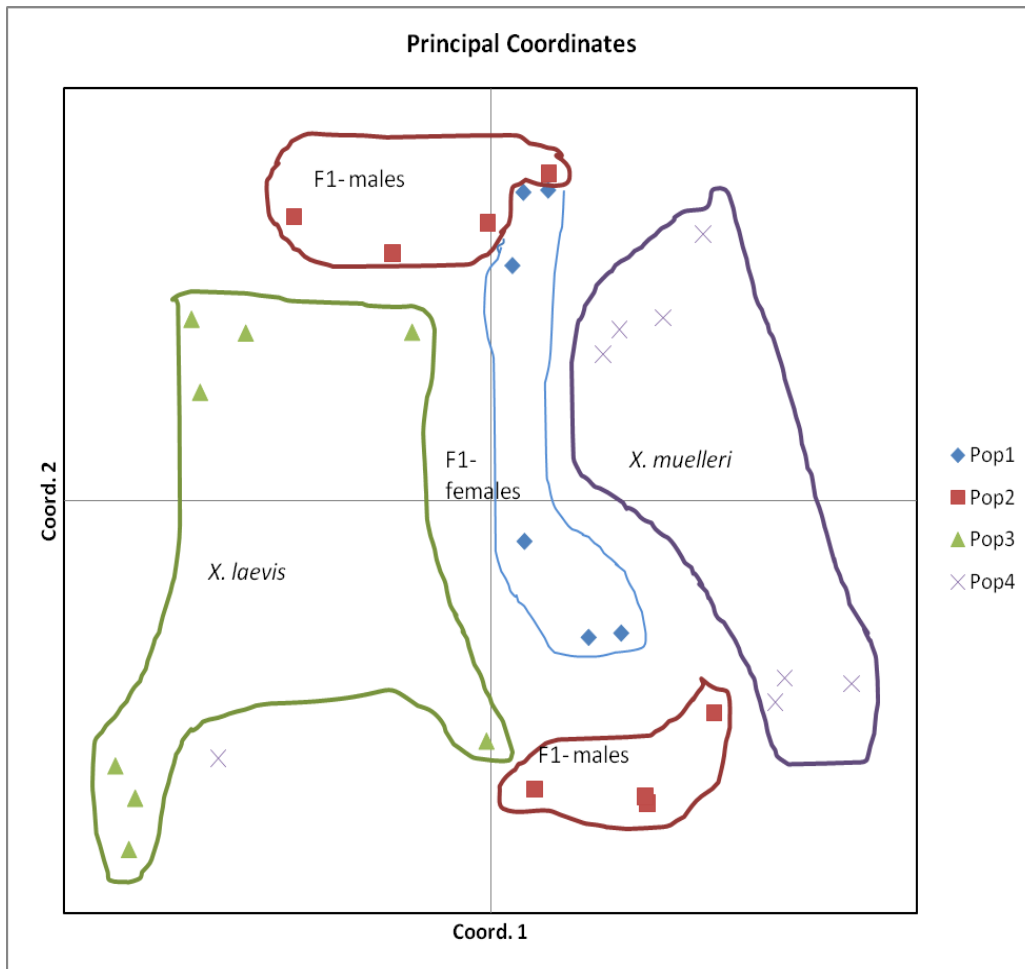
Figure 2.2 Summary graph showing the results of methylated band (alleles) patterns across populations [a) Liver and b) Muscle Tissues]

Cluster Analysis/Population Structure.

The results of the Principal coordinate analysis (PCoA) analysis of the individual pairwise 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]

	1st Eigenvalue (% Variance)	1st Eigenvalue (% Variance)	1st Eigenvalue (% Variance)	Cummulative % variance
Liver	44.643 (25.68)	39.714 (22.84)	23.341 (15.59)	64.11
Muscle	48.601 (31.14)	35.823 (22.95)	22.643 (14.51)	68.59



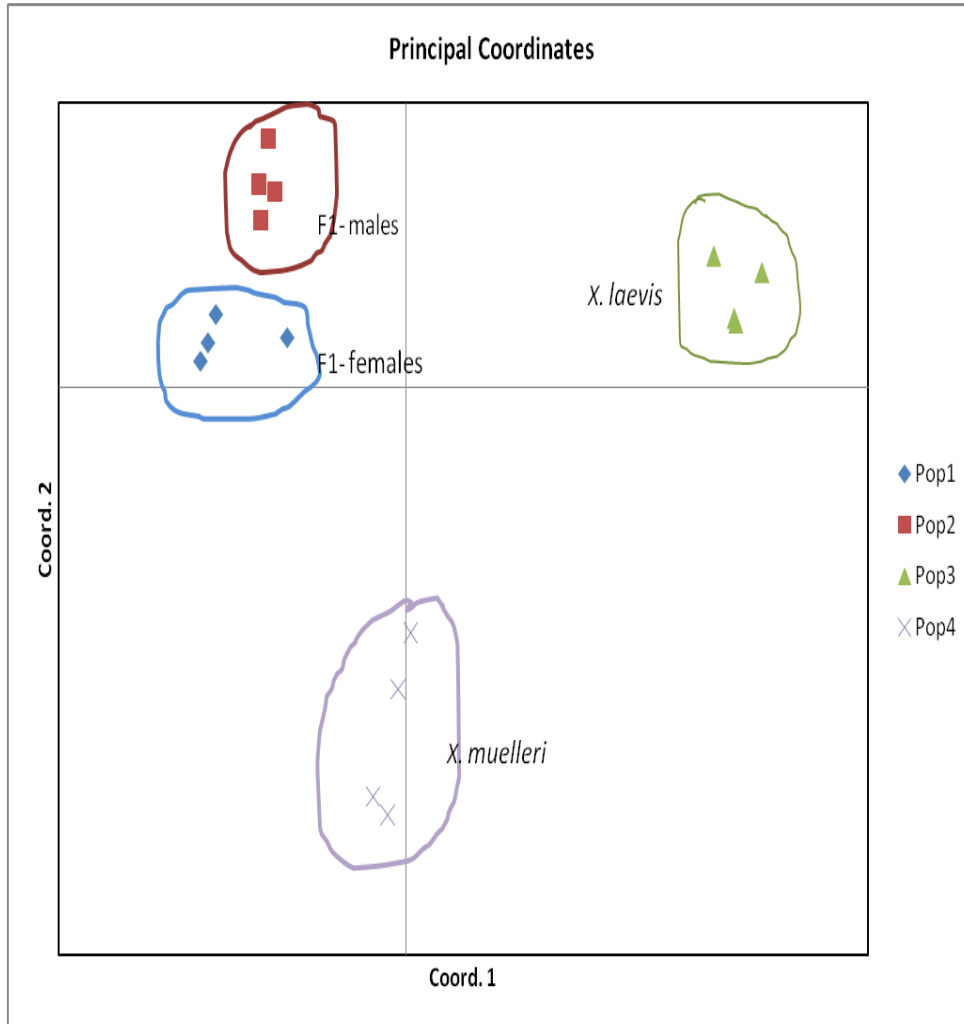
Pop 1 = F1-females (*X. laevis* x *X. muelleri*)

Pop 2 = F1-males (*X. laevis* x *X. muelleri*)

Pop 3 = *X. laevis*

Pop 4 = *X. muelleri*

Figure 2.3 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



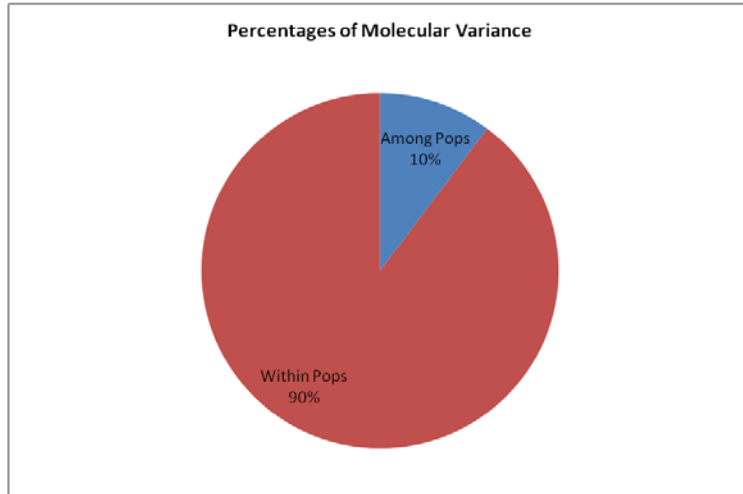
### Analysis of Molecular Variance

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.

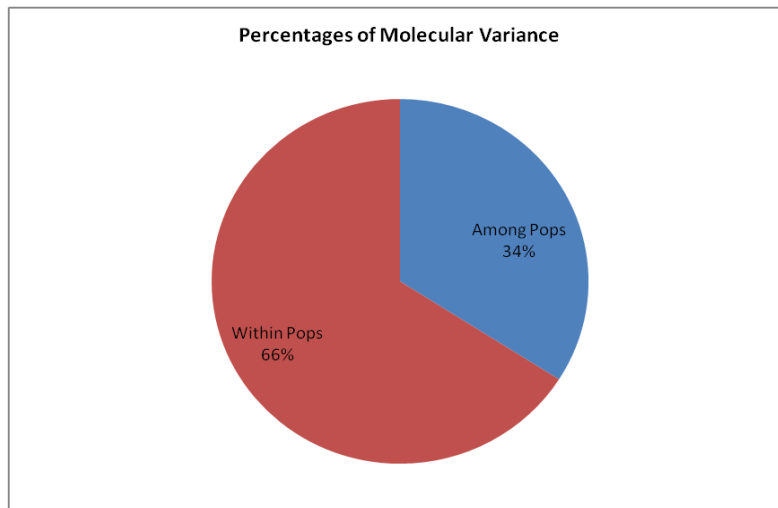
Table 2.3 Results of AMOVA performed on all populations identified as clusters [Liver and muscle Tissues]

Tissue	Source of variation	df	Variance component	% variance	$\phi$ -Statistic	P-Value
Liver	Among	3	7.715	10	0.102	<0.0001
	Within	26	67.755	90		
	Total	29	75.470	1		
Muscle	Among	3	22.781	34	0.34	<0.0001
	Within	12	44.167	66		
	Total	15	66.948	100		

Probability, P (rand<=data), for PhiPT is based on 9999 permutation across the full data set



a



b

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  $0.166 \pm 0.008$ ,  $0.188 \pm 0.008$ ,  $0.205 \pm 0.008$  and  $0.174 \pm 0.007$  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  $\geq 5\%$  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  $0.176 \pm 0.011$ ,  $0.185 \pm 0.012$ ,  $0.171 \pm 0.011$  and  $0.162 \pm 0.011$  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  $\geq 5\%$  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

<b>Pop1</b>	<b>Pop2</b>	<b>PhiPT</b>	<b>LinPhiPT</b>	<b>#Pop1</b>	<b>#Pop2</b>	<b>P (rand &gt;= data)</b>	<b>No. PW Pm</b>
F1-female	F1-male	0.041	0.042	6	8	0.097	9999
F1-female	<i>X. laevis</i>	0.119	0.135	6	8	0.001	9999
F1-male	<i>X. laevis</i>	0.108	0.121	8	8	0.002	9999
F1-female	<i>X. muelleri</i>	0.071	0.077	6	8	0.022	9999
F1-male	<i>X. muelleri</i>	0.090	0.099	8	8	0.005	9999
<i>X. laevis</i>	<i>X. muelleri</i>	0.155	0.183	8	8	0.001	9999

Muscle

<b>Pop1</b>	<b>Pop2</b>	<b>PhiPT</b>	<b>LinPhiPT</b>	<b>#Pop1</b>	<b>#Pop2</b>	<b>P(rand &gt;= data)</b>	<b>No. PW Pm</b>
F1-female	F1-male	0.193	0.238	4	4	0.027	9999
F1-female	<i>X. laevis</i>	0.412	0.700	4	4	0.030	9999
F1-male	<i>X. laevis</i>	0.384	0.624	4	4	0.029	9999
F1-female	<i>X. muelleri</i>	0.294	0.417	4	4	0.029	9999
F1-male	<i>X. muelleri</i>	0.335	0.503	4	4	0.029	9999
<i>X. laevis</i>	<i>X. muelleri</i>	0.382	0.618	4	4	0.030	9999

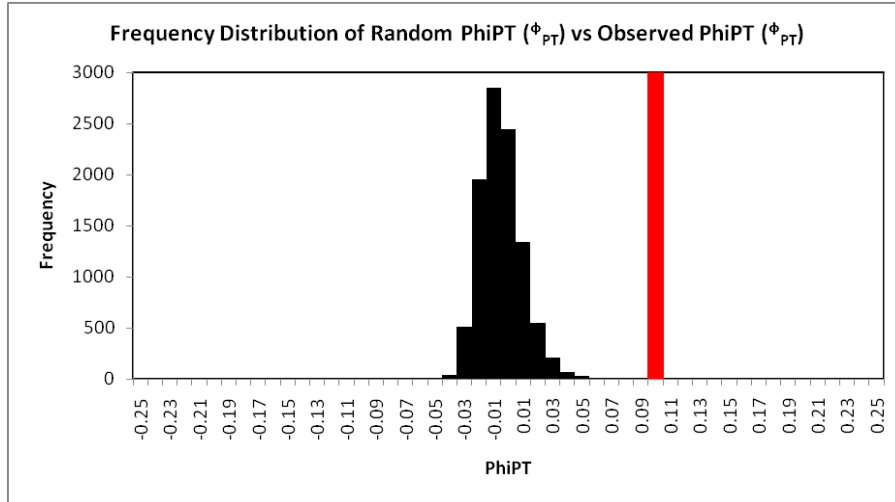
Table 2.5 Results of Pair wise Population Unbiased Nei Genetic Distance and Unbiased Nei Genetic Identity Values as Table [liver and muscle tissues]

Liver

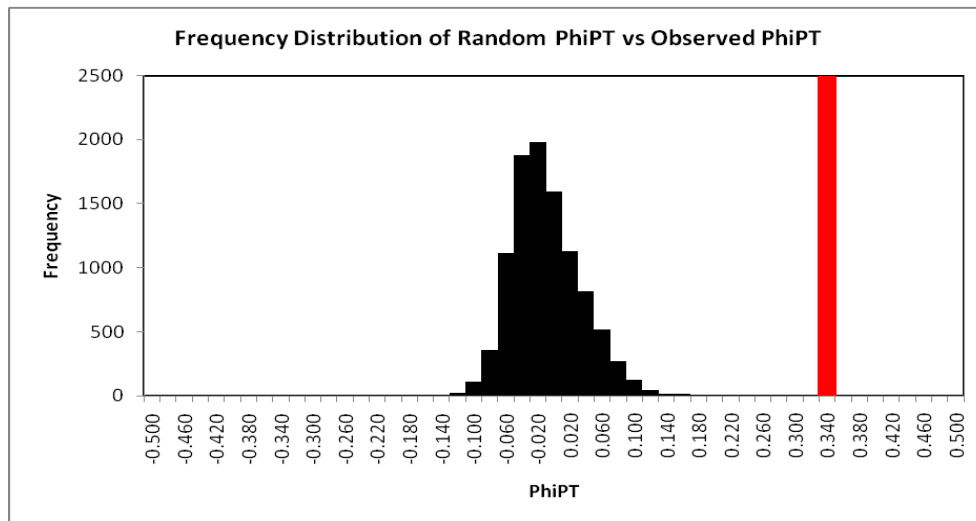
<b>Pop1</b>	<b>Pop2</b>	<b>UNei GD</b>	<b>UNei ID</b>	<b>#Pop1</b>	<b>#Pop2</b>
F1-female	F1-male	0.017	0.983	6	8
F1-female	<i>X. laevis</i>	0.028	0.972	6	8
F1-male	<i>X. laevis</i>	0.029	0.972	8	8
F1-female	<i>X. muelleri</i>	0.016	0.984	6	8
F1-male	<i>X. muelleri</i>	0.021	0.979	8	8
<i>X. laevis</i>	<i>X. muelleri</i>	0.033	0.968	8	8

Muscle

<b>Pop1</b>	<b>Pop2</b>	<b>UNei GD</b>	<b>UNei ID</b>	<b>#Pop1</b>	<b>#Pop2</b>
F1-female	F1-male	0.017	0.983	6	8
F1-female	<i>X. laevis</i>	0.028	0.972	6	8
F1-male	<i>X. laevis</i>	0.029	0.972	8	8
F1-female	<i>X. muelleri</i>	0.016	0.984	6	8
F1-male	<i>X. muelleri</i>	0.021	0.979	8	8
<i>X. laevis</i>	<i>X. muelleri</i>	0.033	0.968	8	8



a



b

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 2.6 Primers used in MSAP

	Sequences
<b>Adaptors</b>	
<i>ECOR1</i>	5 CTC GTA GAC TGC GTA CC – 3 5 AAT TGG TAC GCA GTC TAC -3
<i>HpaII</i> <i>MspI</i> (METHADPR)	5 GAC GAT GAG TCT AGA A -3 5 CGT TCT AGA CTC ATC -3
<b>Preselective Primers</b>	
<i>ECORI</i> preG	5 GAC TGC GTA CCA ATT CG -3
<i>HpaII</i> <i>MspI</i> pres-T	5 GAT GAG TCT AGA ACG GT -3
<b>Selective Primers</b>	
<i>ECORI</i> (Fluorescent primer) – FAM-labeled. Have 2 extra bases at the end of the Preselect.	<i>ECORI</i> -AGT  5 GAT GAG TCT AGA ACG GTT G -3
<i>HpaII</i> <i>MspI</i> sel-TG	

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-MspI* digestion (one selective primer combination) - [Liver]

	Parental fragments	Parental fragments in both F1 sexes	Parental fragments missing in both F1 sexes	Parental fragments missing in only one F1 sex		Fragments shared only by Hybrids
				Male	Female	
Common parental fragments	247	158 (64%)	34 (13.8%)	20 (8.1%)	35 (14.2%)	
Diagnostic fragments						
From <i>X. laevis</i>	115	36 (31.3%)	36 (31.3%)	27 (23.5%)	16 (13.9%)	
From <i>X. muelleri</i>	100	31 (31%)	29 (29%)	15 (15%)	25 (25%)	
Total number of diagnostic fragments between parents	215	67 (31.2%)	99 (46%)	62 (28.8%)	76 (35.3%)	
Private/Unique fragments	-			21	10	11

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 *EcoRI-MspI* digestion (one selective primer combination) - [Muscle Tissues]

		Parental fragments	Parental fragments in both F1 sexes	Parental fragments missing in both F1 sexes	Parental fragments missing in only one F1 sex		Fragments shared only by Hybrids
					Male	Female	
Common parental fragments		78	40 (51.3%)	20 (25.6%)	9 (11.5%)	9 (11.5%)	
Diagnostic fragments	From <i>X. laevis</i>	110	24 (21.8%)	47 (42.7%)	12 (10.9%)	27 (24.5%)	
	From <i>X. muelleri</i>	101	39 (38.6%)	41 (40.6%)	9 (8.9%)	12 (11.9%)	
Total number of diagnostic fragments between parents		211	63 (29.9%)	108 (51.2%)	30 (14.2%)	48 (22.7%)	
Private/Unique fragments					21	18	36

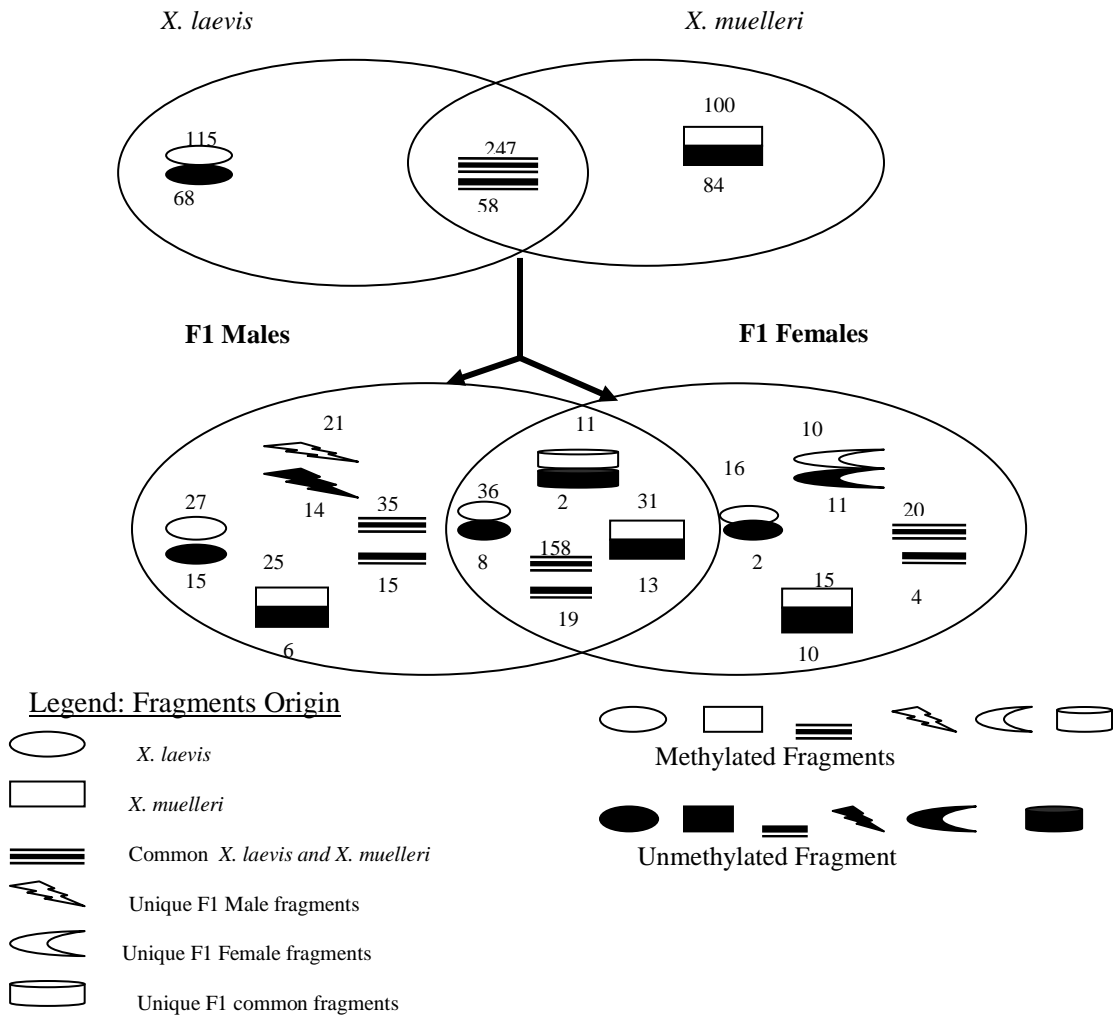


Figure 2.7 Liver MSAP fragments observed in individuals from *X. laevis*, *X. muelleri* and F1 hybrid males and females

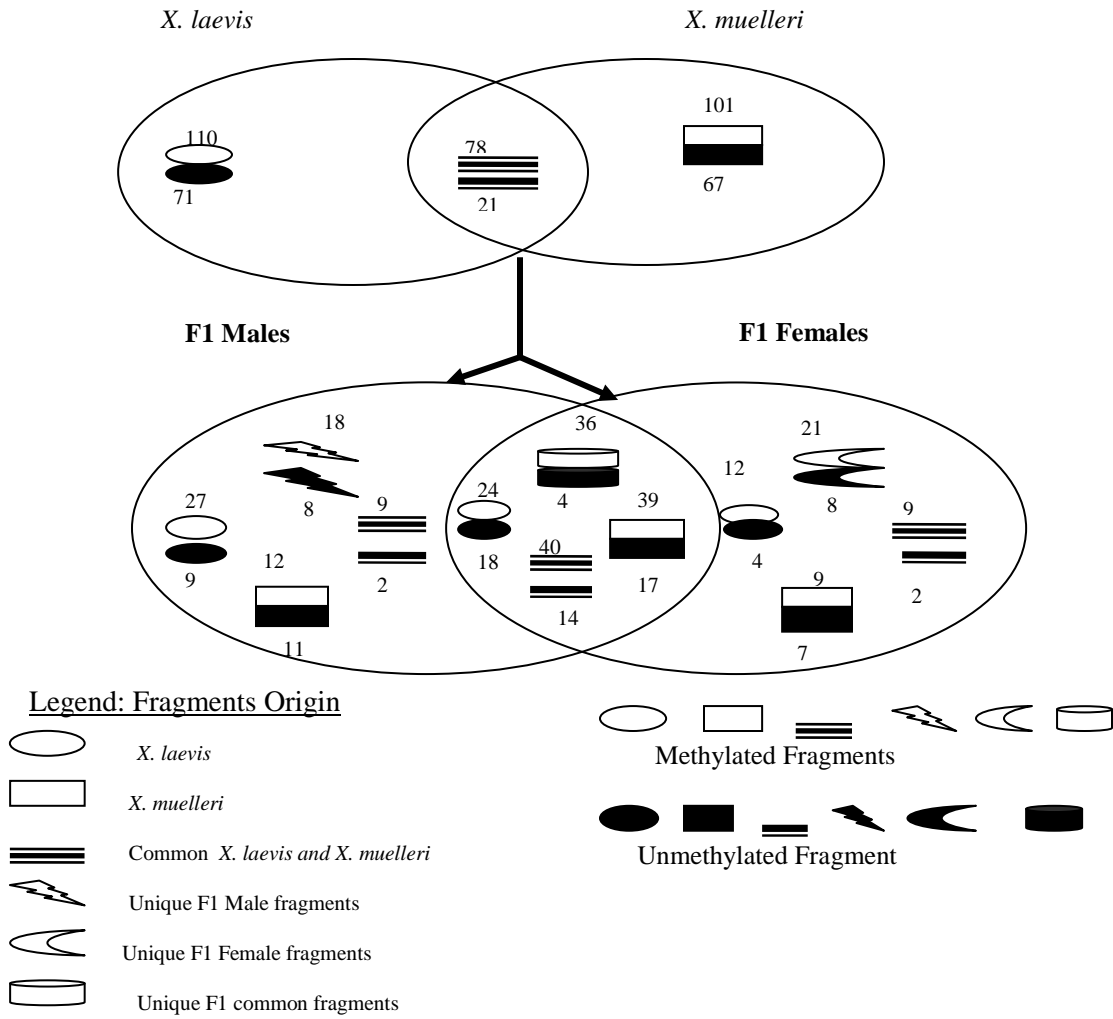


Figure 2.8 Muscle MSAP fragments observed in individuals from *X. laevis*, *X. muelleri* and F1 hybrid males and females

Table 2.9 Percentage of Methylated Fragment for Liver and Muscle

Liver

Cluster/Population	MSAP Fragments			% Methylated fragments
	Nonmethylated	Methylated	Total	
F1 female	79	277	356	77.8
F1 male	92	341	433	78.8
<i>X. laevis</i>	126	350	476	73.5
<i>X. muelleri</i>	142	333	475	70.1

Muscle

Cluster/Population	MSAP Fragments			% Methylated fragments
	Nonmethylated	Methylated	Total	
F1 female	86	188	274	68.6
F1 male	74	208	282	73.8
<i>X. laevis</i>	92	187	279	67.0
<i>X. muelleri</i>	88	180	268	67.2

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

	Parental fragments in both F1 sexes	Parental fragments missing in both F1 sexes	P Value	Parental fragments missing in male F1	Parental fragments missing in female F1	P Value
<i>X. laevis</i>	36	36	0.86303	27	16	0.02846
<i>X. muelleri</i>	31	29		15	25	

Muscle

	Parental fragments in both F1 sexes	Parental fragments missing in both F1 sexes	P Value	Parental fragments missing in male F1	Parental fragments missing in female F1	P Value
<i>X. laevis</i>	24	47	0.07074	12	27	0.40245
<i>X. muelleri</i>	39	41		9	12	



## 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  $44.23\% \pm 0.58\%$  and  $63.89\% \pm 3.29\%$  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* ( $0.205 \pm 0.008$ ) the values for the liver which range from  $0.166 \pm 0.008$  to  $0.188 \pm 0.008$  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  $75.0\% \pm 4$  and  $69.1\% \pm 3.2$  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 polyploidy 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.

## CHAPTER 3

### COMPARISON OF GENOME SIZE OF HYBRIDS TO PARENTAL SPECIES (*X. l. laevis* and *X. muelleri*) USING FLOW CYTOMETRY

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; Kozłowski 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 Strelman 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 polyploidy. 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; C<sub>x</sub>-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 C<sub>x</sub>-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 mm<sup>3</sup> 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 *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 *X. laevis* (n = 2) is 70445, *X. muelleri* (n=2) is 799332.17 and for hybrids (n = 2) is 79115.5. From the results the peak of *X. l. laevis* is located at a lower fluorescence when compared to *X. muelleri*. The peak of the hybrid appears to fall between the peaks of *Xenopus l. laevis* and *X. muelleri*. However it appears to be closer to *X. muelleri* since the peaks of the hybrid and *X muelleri* merges and only show one peak when the two samples are combined. The genome size of *X. muelleri* is 113.5% of *X. laevis* whereas that of the F1 hybrid is 112.3%.

Table 3.1 Mean Fluorescence (PI-A) for three different measurements

Samples	Mean Fluorescence (PI-A)			
	A	B	C	Mean
<i>X. laevis</i> -3 (female)	71599	73112	73218	72643
<i>X. laevis</i> -4 (female)	68084	68396	68282	68254
<i>X. muelleri</i> -3 (female)	80121	79823	80192	80045.3
<i>X. muelleri</i> -4 (male)	78879	80248	80330	79819
Hybrid-3 ( <i>X. laevis</i> x <i>X. muelleri</i> ) -female	80433	80994	80993	80806.7
Hybrid-4 ( <i>X. laevis</i> x <i>X. muelleri</i> ) -male	77183	77320	77770	77424.3

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

Specimen	No. of specimens	DNA content (Mean arbitrary units)	DNA content (% of <i>X. laevis</i> )
<i>X. Laevis</i>	2	70448.5	100
<i>X. muelleri</i>	2	79932.17	113.5
<i>F1</i> ( <i>X. laevis</i> x <i>X. muelleri</i> )	2	79115.5	112.3

Table 3.3 Fluorescence (PI-A) for the mixed samples

	Peak 1	Peak 2	Peak 1	Peak 2	Peak 1	Peak 2	Peak 1	Peak 2
<i>X. laevis</i> - <i>X. muelleri</i>	68603	81039	68369	80796	68390	80628	68454	80821
<i>X. laevis</i> - Hybrid	65694	75655	66737	76466	67014	76567	66481.67	76229.3
<i>X. muelleri</i> - Hybrid	0	79774	0	79878	0	80027	0	79893
<i>X. laevis</i> - <i>X. muelleri</i> - Hybrid	67813	79503	68289	79964	68197	79984	68099.67	79817

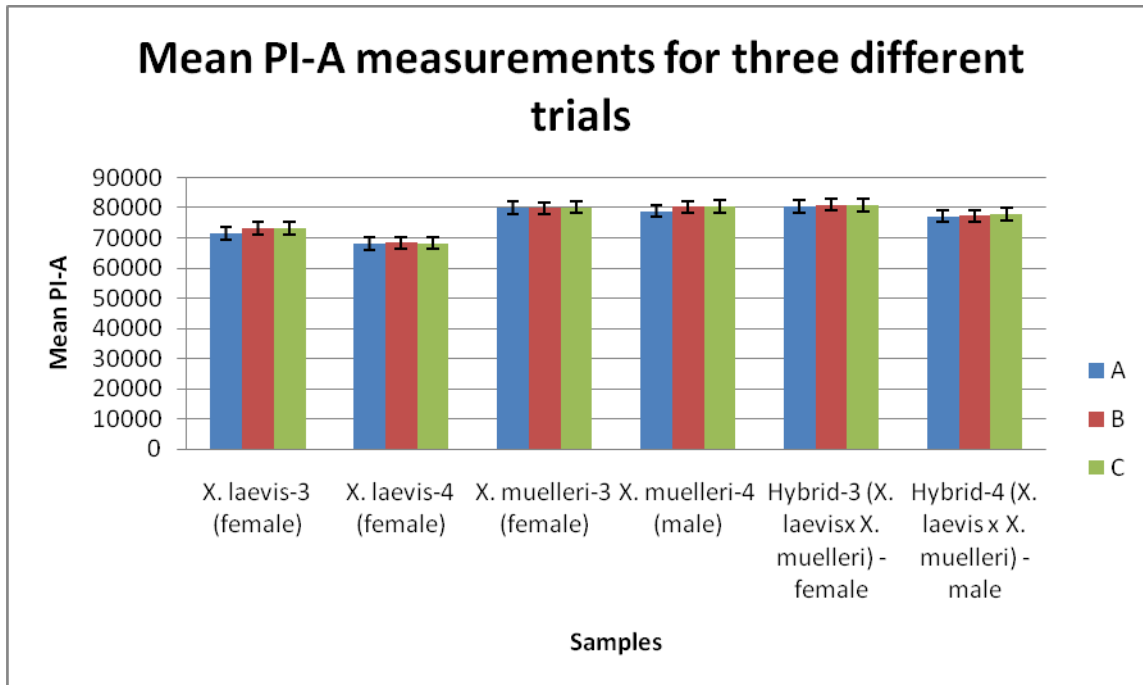
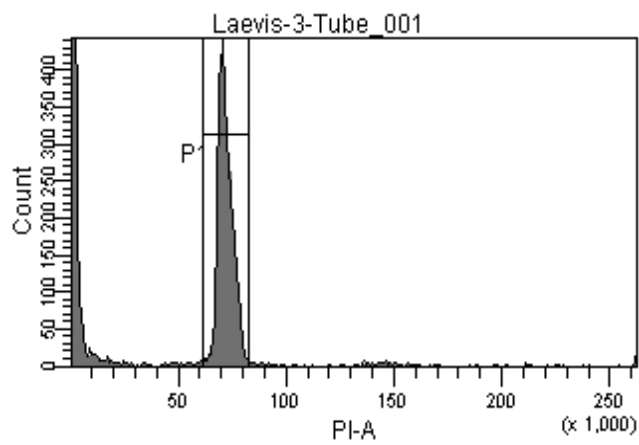


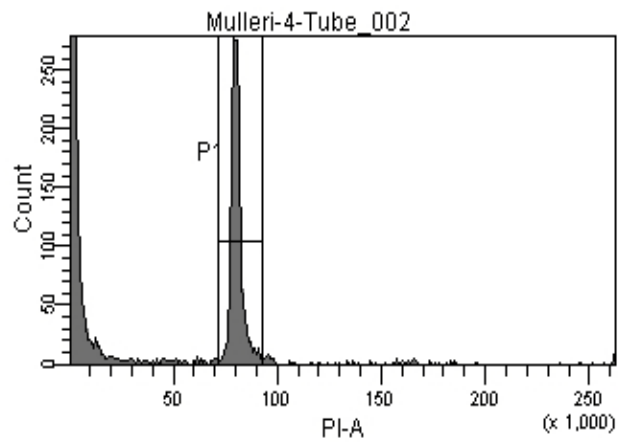
Figure 3.1 Histogram of the mean PI-A measurements

*X. laevis*



a

*X. muelleri*

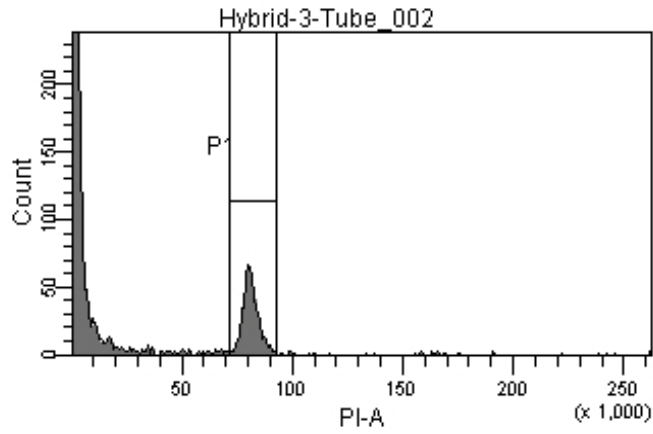


b

Figure 3.2 Flow cytometric histograms of fluorescence intensity for a) *X. laevis* and b) *X. muelleri*

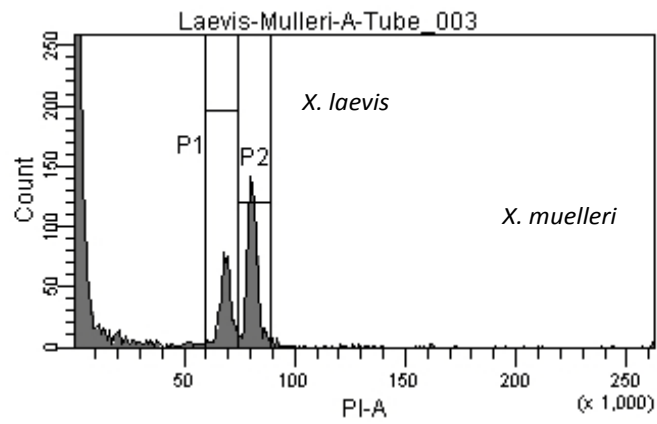


F1 Hybrid



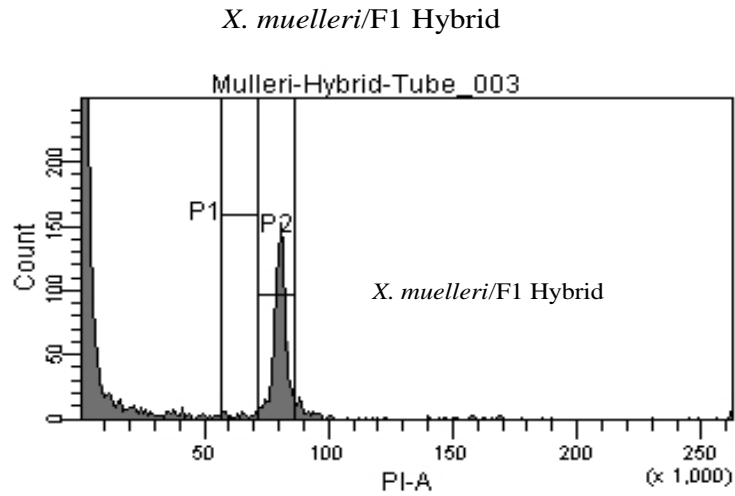
a

*X. laevis* and *X. muelleri*

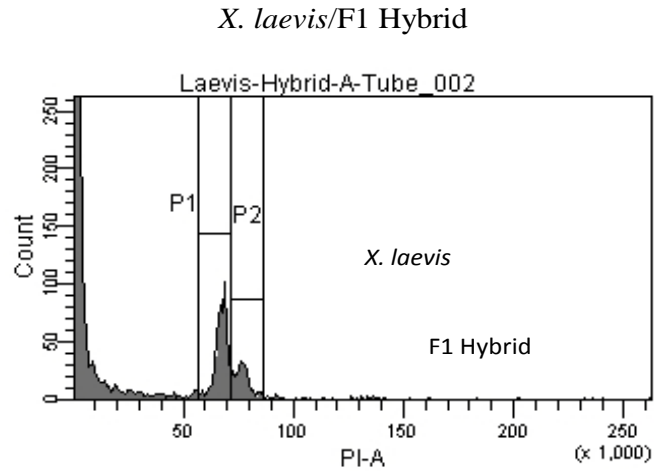


b

Figure 3.3 Flow cytometric histograms of fluorescence intensity for a) F1 hybrid and b) *X. laevis/muelleri*



a



b

Figure 3.4 Flow cytometric histograms of fluorescence intensity for a) *X. laevis*/F1 hybrid and b) *X. muelleri*/F1 Hybrid

*X. laevis*/*X. muelleri*/Hybrid

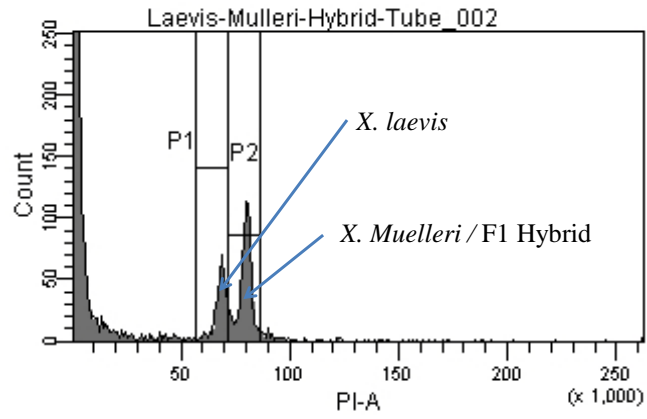


Figure 3.5 Flow cytometric histograms of fluorescence intensity for *X. laevis*/*muelleri*/ F1 hybrid

### 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 *HpaII* digestion and *MspI* digestion were compared to the gDNA. With the exception of the *X. laevis* muscle none of the *HpaII/MspI* digestion for the *X. muelleri* and the F1 hybrid has any smear pattern. The smear pattern obtained for the *X. laevis* was ruled out because there was smear in the electrophenogram of the gDNA which overlap with that of the *HpaII* and *MspI* digestion. This may therefore indicate degraded DNA around the 10,000 bp. For the liver tissue there was no smearing for the *X. laevis* and *X. muelleri* but with the F1 liver there was smearing in both the *HpaII* and the *MspI* digestion. However the smearing for the *HpaII* was excluded because it overlapped with the smear of the gDNA. The smearing in *MspI* may be an indication of undermethylation.

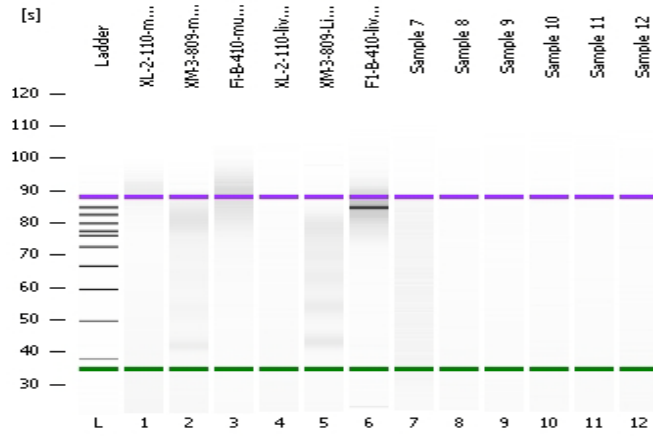
Table 3.4 Summary of Regions identified as DNA smear for Liver and muscle

Liver

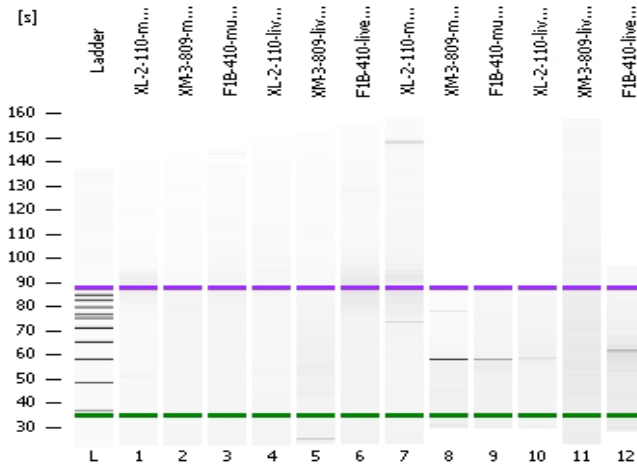
	XL-2-110-male liver			XM-3-809 male liver			F1-B-410 male liver		
	gDNA	HpaII	MspI	gDNA	HpaII	MspI	gDNA	HpaII	MspI
From (s)	-	-	-	-	-	-	73.43	78.61	54.6
To (s)	-	-	-	-	-	-	96.43	98.67	66.52
Area	-	-	-	-	-	-	380.7	5.5	5.2
% of Total	-	-	-	-	-	-	93	67	59
Average Size (bp)	-	-	-	-	-	-	7,346	10,605	571
Size distribution	-	-	-	-	-	-	48.6	35.7	11.5
Conc (ng/ul)	-	-	-	-	-	-	43.73	4.28	139.3

Muscle

	XL-2-110-male muscle			XM-3-809 male muscle			F1-B-410 male muscle		
	gDNA	HpaII	MspI	gDNA	HpaII	MspI	gDNA	HpaII	MspI
From (s)	88.93	80.31	75.53	74.61	-	-	75.54	-	-
To (s)	96.81	96.53	100.36	86.73	-	-	99.96	-	-
Area	32.9	6.9	5.6	92.3	-	-	163.8	-	-
% of Total	62	59	55	51	-	-	86	-	-
Average Size (bp)	14,179	10,719	10,463	4,172	-	-	10,268	-	-
Size distribution	14.4	27.3	43.3	49.3	-	-	43	-	-
Conc (ng/ul)	1.96	2.57	4.78	4.47	-	-	6.59	-	-



a



b

Figure 3.6 Gel Electrophoresis for a) gDNA and b) Restriction digest products

## XL2-110-male muscle

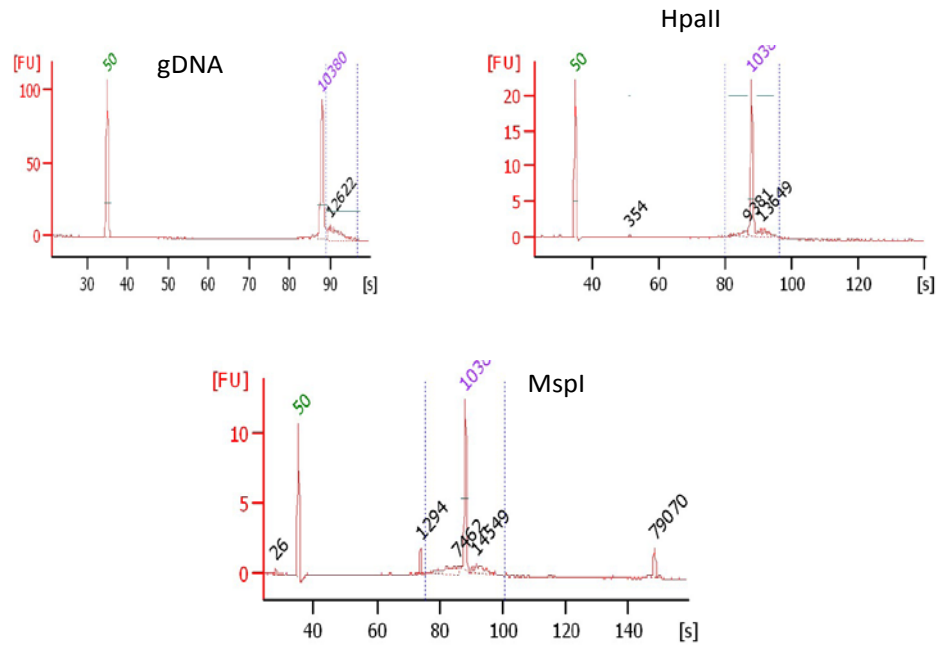


Figure 3.7 Electropherogram for male muscle tissues (*X. laevis*). There is smearing around the 10000 bp (upper marker) for the gDNA and the *MspI* and *HpaII* digest. Since there is overlap, it was considered not reliable

## XM-3-809-male muscle

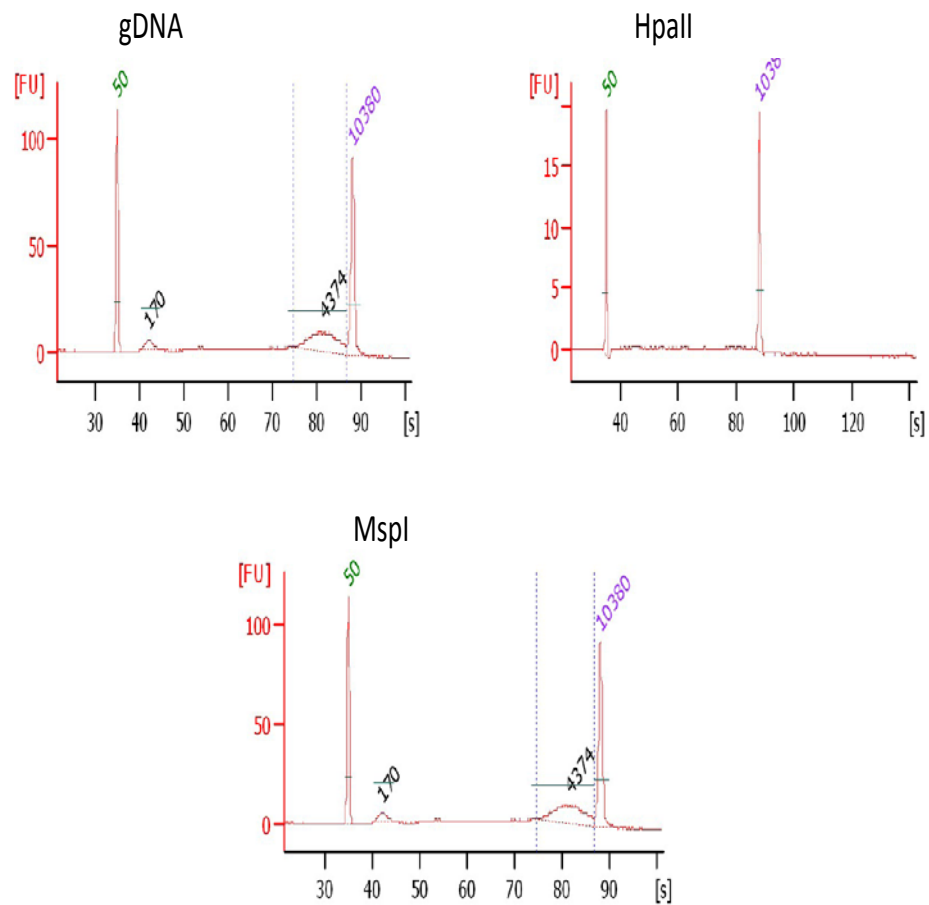


Figure 3.8 Electrophenogram for male Muscle Tissues (*X. muelleri*). The smearing of *Msp*1 overlaps that of the gDNA, hence it is probably a carry-over so it was not considered significant



## F1-410-muscle

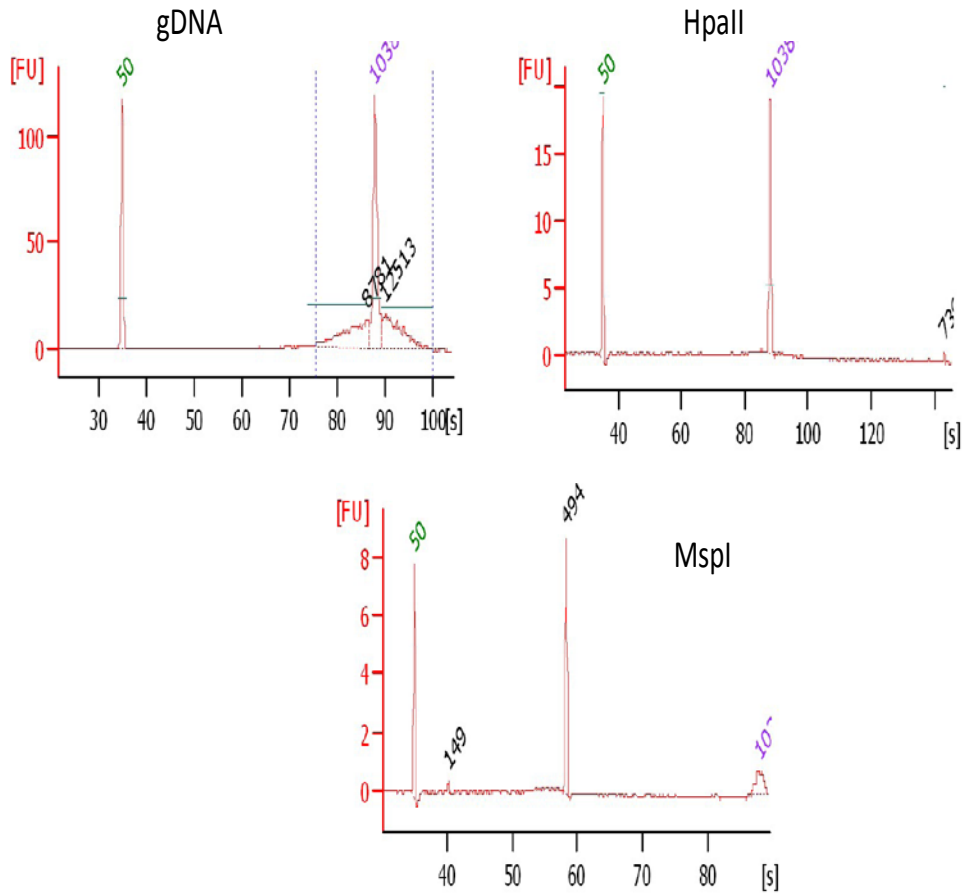


Figure 3.9 Electropherogram for male Muscle Tissues (F1 Hybrid). F1 male hybrid muscle. No smearing observed with the digest from the restriction enzymes. It was therefore considered as highly methylated.

### XL-2-110-liver

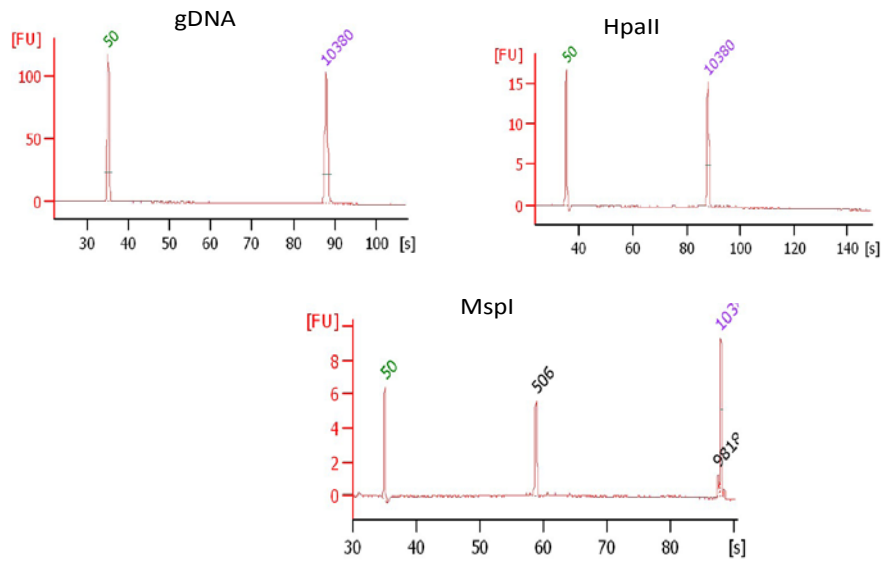


Figure 3.10 Electropherogram for male Liver Tissues (*X. laevis*)  
No smearing observed for the gDNA or the digests from the restriction enzymes

## XM-3-809-liver

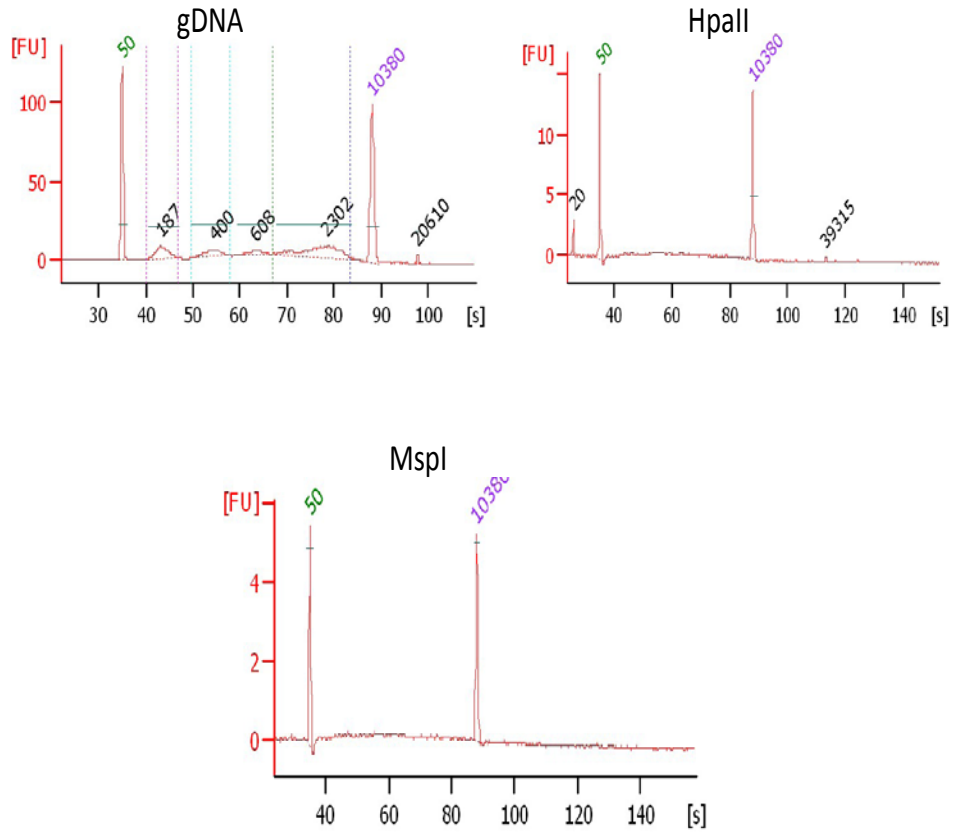


Figure 3.11 Electrophenogram for male Liver Tissues (*X. muelleri*)  
No smearing observed with the digests from the two restriction enzymes.

# F1-B-410-liver

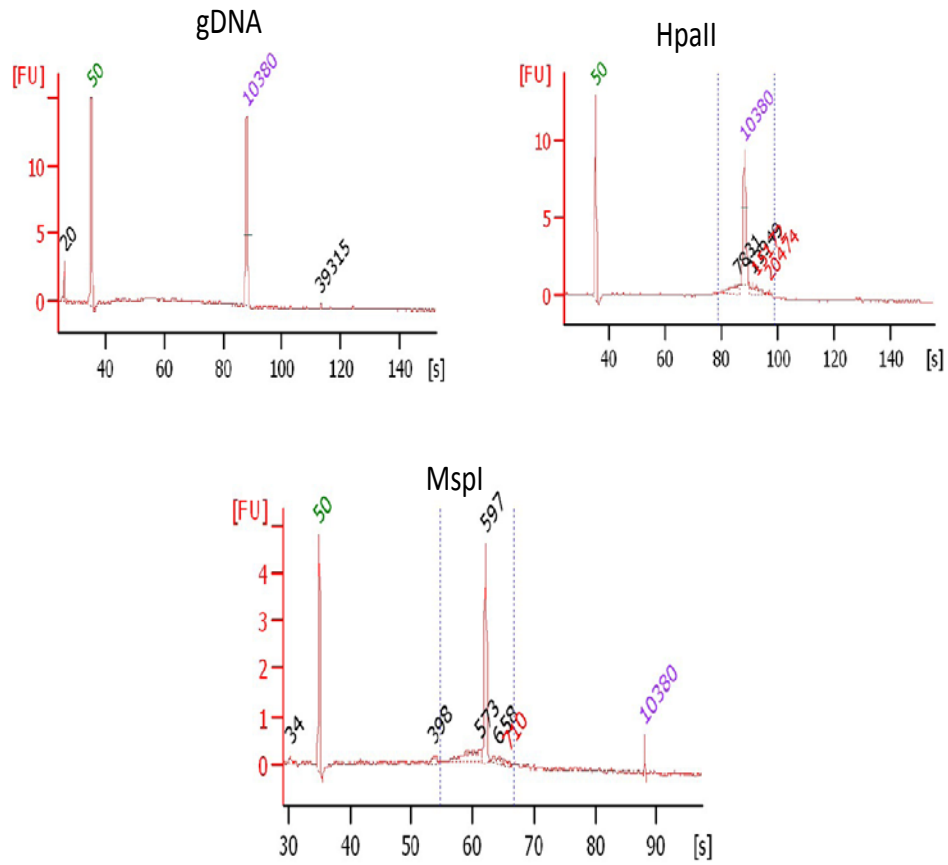


Figure 3.12 Electrophenogram for male Liver Tissues (F1 Hybrid)  
There was smearing in both digests from *HpaII* and *MspI*.

## 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 allopolyploidization. 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

A: Table with Raw Data															
Sampl es	100.4	100.9	101.4	102	103.1	103.6	104.2	104.7	105.8	106.9	107.9	108.7	111	# of Frag.	Error rate
F1-A-1	1	0	0	0	0	0	1	0	1	1	0	1	1	6	
F1-A-2	1	0	0	0	0	0	1	0	0	0	0	1	1	4	
	Accept						Acce pt		Reject	Acce pt		Accept	Accept	5	
F1-B-1	0	1	0	0	1	0	0	1	0	1	0	1	0	5	
F1-B-2	0	0	0	0	0	0	1	0	0	1	0	1	1	3	
		Reject			Rejec t		Acce pt	Rejec t		Acce pt		Accept	Accept	4	
F1-C-1	0	0	0	1	0	0	0	0	0	1	0	1	1	4	
F1-C-2	0	0	0	0	0	0	0	0	0	0	0	1	1	2	
				Reje cted						Acce pted		Accept ed	Accept ed	3	
B: New Table to be used in analysis															
Sample s	100.4	100.9	101 .4	102	103.1	103 .6	104.2	104.7	105.8	106.9	107 .9	108.7	111		
F1-A	1	0	0	0	0	0	1	0	0	0	0	1	1	4	25%
F1-B	0	0	0	0	0	0	0	0	0	1	0	1	0	2	100%
F1-C	0	0	0	0	0	0	0	0	0	0	0	1	1	2	50%

C: TABLE COMPARING MSAP FRAGMENTS FROM *EcoRI/HpaII* AND *EcoRI/MspI*

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

18

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 =  $q^2 = \text{Freq. of band absence} = 1 - \text{Freq. of band presence}$

$$\text{So } q = \sqrt{(\text{Freq of band absence})}$$

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 =  $q^2 = \text{Freq. of band absence} = 1 - \text{Freq. of band presence}$

$$\text{So } q = \sqrt{(\text{Freq of band absence})}$$

$$He = 1 - \sum p^2$$

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)

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

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

- $P_{ix}$  and  $P_{iy}$  are the frequencies of the  $i$ -th allele in populations  $x$  and  $y$ .
- For multiple loci,  $J_{xy}$ ,  $J_x$  and  $J_y$  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

- Abbott RJ. 1992. Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends Ecol Evo.* 7(12):401-5.
- Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc Natl Acad Sci U. S. A.* 100(8):4649-54.
- Allis D, Jenuwein T, Reinberg D, Caparros M, editors. 2007. *Epigenetics*. Cold Spring Harbor, New York: Cold Spring Harbor Press.
- Antequera F and Bird A. 1993. DNA methylation and CPG islands. *Chromosome* :127-33.
- Bennett MD and Leitch IJ. 2005. Plant genome size research: A field in focus. *Annals of Botany* 95(1):1-6.
- Bensch S and Akesson M. 2005. Ten years of AFLP in ecology and evolution: Why so few animals? *Mol Ecol* 14(10):2899-914.
- Bensch S, Helbig AJ, Salomon M, Siebold I. 2002. Amplified fragment length polymorphism analysis identifies hybrids between two subspecies of warblers. *Mol Ecol* 11(3):473-81.
- Bestor TH, Walsh CP, Yoder JA. 1997. Does DNA methylation control transposition of selfish elements in the germline? Reply. *Trends in Genetics* 13(12):470-2.

- Bird AP. 1993. Functions for DNA methylation vertebrates. Cold Spring Harb Symp Quant Biol 58:281-5.
- Bird A. 2002. DNA methylation patterns and epigenetic memory. Genes Dev 16(1):6-21.
- Blackshear PJ, Lai WS, Thorn JM, Kennington EA, Staffa NG, Moore DT, Bouffard GG, Beckstrom-Sternberg SM, Touchman JW, Bonaldo MD, et al. 2001. The NIEHS *Xenopus* maternal EST project: Interim analysis of the first 13,879 ESTs from unfertilized eggs. Gene 267(1):71-87.
- Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. Mol Ecol 13(11):3261-73.
- Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen AC, Elven R. 2004. Polyploidy in arctic plants. Biol J Linn Soc 82(4):521-36.
- Cavalier-smith T. 1978. Nuclear volume control by nucleoskeletal DNA, selection for cell-volume and cell-growth rate, and solution of DNA C-value paradox. J Cell Sci 34(DEC):247-78.
- Cavalier-Smith T. 2005. Economy, speed and size matter: Evolutionary forces driving nuclear genome miniaturization and expansion. Annals of Botany 95(1):147-75.
- Chen, JZ. 2010. Molecular mechanisms of polyploidy and hybrid vigor. Trends in Plant Science 15(2):57-71
- Chow JC, Yen Z, Ziesche SM, Brown CJ. 2005. Silencing of the mammalian X chromosome. Annual Review of Genomics and Human Genetics 6:69-92.

- Collins AL, Levenson JM, Vilaythong AP, Richman R, Armstrong DL, Noebels JL, Sweatt JD, Zoghbi HY. 2004. Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum Mol Genet* 13(21):2679-89.
- Comai L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* 6(11):836-46.
- Comai L. 2000. Genetic and epigenetic interactions in allopolyploid plants. *Plant Mol Biol* 43(2-3):387-99.
- Comai L, Tyagi AP, Winter K, Holmes-Davis R, Reynolds SH, Stevens Y, Byers B. 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* 12(9):1551-67.
- Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol* 3(10): e314. doi:10.1371/journal.pbio.0030314
- Dolezel J and Bartos J. 2005. Plant DNA flow cytometry and estimation of nuclear genome size L093/aob/inci005. *Annals of Botany* 95(1):99-110.
- Dolezel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2(9):2233-44.
- Dolinoy DC, Weidman JR, Waterland RA, Jirtle RL. 2006. Maternal genistein alters coat color and protects A(vy) mouse offspring from obesity by modifying the fetal epigenome. *Environ Health Perspect* 114(4):567-72.
- Doolittle WF and Brown JR. 1995. Tempo, mode, the progenote, and the universal root. *Tempo and Mode in Evolution* :3-24.

- Doolittle WF and Brown JR. 1994. Tempo, mode, the progenote, and the universal root. Proc Natl Acad Sci U S A 91(15):6721-8.
- Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 38(12):1378-85.
- Evans BJ. 2008. Genome evolution and speciation genetics of clawed frogs (*Xenopus* and *Silurana*). Frontiers in Bioscience 13:4687-706.
- Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. Genetics 131(2):479-91.
- Fay MF, Conwan RS, Leitch IJ. 2005. The effects of nuclear DNA content (C-value) on the quality and utility of AFLP fingerprints. Annals of Botany 95(1):237-46.
- Feldman M, Liu B, Segal G, Abbo S, Levy AA, Vega JM. 1997. Rapid elimination of low-copy DNA sequences in polyploid wheat: A possible mechanism for differentiation of homoeologous chromosomes. Genetics 147(3):1381-7.
- Fischer WJ, Koch WA, Elepfandt A. 2000. Sympatry and hybridization between the clawed frogs *Xenopus laevis laevis* and *Xenopus muelleri* (pipidae). J Zool 252:99-107.
- Fontenot B. 2009. Natural hybridization and speciation in toads of the *Anaxyrus americanus* Group. Unpublished Ph.D. dissertation, The University of Texas at Arlington.

- Friedman R and Hughes AL. 2001. Pattern and timing of gene duplication in animal genomes. *Genome Res* 11(11):1842-7.
- Galagan JE and Selker EU. 2004. RIP: The evolutionary cost of genome defense. *Trends in Genetics* 20(9):417-23.
- Garcia-Fernandez J and Holland PW. 1996. Amphioxus hox genes: Insights into evolution and development. *Int J Dev Biol Suppl* 1:71S-2S.
- Grant V, editor. 1981. *Plant speciation*. New York: Columbia Press.
- Grant V, editor. 1971. *Plant speciation*. New York: Columbia University Press.
- Gregory TR. 2002. Genome size and developmental complexity. *Genetica* 115(1):131-46.
- Gregory TR. 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews* 76(1):65-101.
- Greilhuber J, Dolezel J, Lysak MA, Bennett MD. 2005. The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. *Annals of Botany* 95(1):255-60.
- Gu, X., Wang, Y., Gu, J. 2002. Age distribution of human gene families shows significant roles of both large- and small-scale duplications in vertebrate evolution. *Nat Genet* 31:205-9.
- Guo YP, Vogl C, Van Loo M, Ehrendorfer F. 2006. Hybrid origin and differentiation of two tetraploid *Achillea* species in East Asia: Molecular, morphological and ecogeographical evidence. *Mol Ecol* 15(1):133-44.
- Haig D. 2004. The (dual) origin of epigenetics. *Cold Spring Harb Symp Quant Biol* 69:67-70.

- He P, Friebe BR, Gill BS, Zhou JM. 2003. Allopolyploidy alters gene expression in the highly stable hexaploid wheat. *Plant Mol Biol* 52(2):401-14.
- Holland BR, Clarke AC, Meudt HM. 2008. Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Syst Biol* 57(3):347-66.
- Holland PWH and Garcia-Fernandez J. 1996. Hox genes and chordate evolution. *Dev Biol* 173(2):382-95.
- Holmgren C, Kanduri C, Dell G, Ward A, Mukhopadhyaya R, Kanduri M, Lobanenkov V, Ohlsson R. 2001. CpG methylation regulates the Igf2/H19 insulator. *Current Biology* 11(14):1128-30.
- Horner HA and Macgregor HC. 1983. C-value and cell-volume - their significance in the evolution and development of amphibians. *J Cell Sci* 63(SEP):135-46.
- Huff DR, Peakall R, Smouse PE. 1993. Rapid variation within and among natural-populations of outcrossing buffalograss [*Buchloe dactyloides* (nut) engelm]. *Theor Appl Genet* 86(8):927-34.
- Issa JP. 2003. Age-related epigenetic changes and the immune system. *Clinical Immunology* 109(1):103-8.
- Jackson JA and Tinsley RC. 2003. Parasite infectivity to hybridising host species: A link between hybrid resistance and allopolyploid speciation? *Int J Parasitol* 33(2):137-44.
- Jockusch EL. 1997. An evolutionary correlate of genome size change in plethodontid salamanders. *Proceedings of the Royal Society of London Series B-Biological Sciences* 264(1381):597-604.



- Jones PA and Baylin SB. 2002. The fundamental role of epigenetic events in cancer. *Nature Reviews Genetics* 3(6):415-28.
- Jones RN and Hegarty M. 2009. Order out of chaos in the hybrid plant nucleus. *Cytogenetic and Genome Research* 126(4):376-89.
- Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, Mccarrey J, Cedar H, Razin A. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* 6(5):705-14.
- Kakutani T, Kato M, Kinoshita T, Miura A. 2004. Control of development and transposon movement by DNA methylation in *Arabidopsis thaliana*. *Cold Spring Harb Symp Quant Biol* 69:139-43.
- Kardolus JP, van Eck HJ, van den Berg RG. 1998. The potential of AFLPs in biosystematics: A first application in *Solanum* taxonomy (solanaceae). *Plant Syst Evol* 210(1-2):87-103.
- Kashkush K, Feldman M, Levy AA. 2003. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet* 33(1):102-6.
- Kashkush K, Feldman M, Levy AA. 2002. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* 160(4):1651-9.
- Kawara S, Takata M, Takehara K. 1999. High frequency of DNA aneuploidy detected by DNA flow cytometry in Bowen's disease. *J Dermatol Sci* 21(1):23-6.
- Kobel HR. 1996. Allopolyploid speciation. In: *The biology of Xenopus*. Tinsley, R. C. and Kobel, H. R., editor. Eds. Oxford: Clarendon Press.

- Kobel HR. 1981. Gene-mapping in *Xenopus* (anura, pipidae). *Monitore Zoologico Italiano-Italian Journal of Zoology* 15:109-18.
- Kobel HR and Dupasquier L. 1986. Genetics of polyploid *Xenopus*. *Trends in Genetics* 2(12):310-5.
- Kobel HR, Dupasquier L, Tinsley RC. 1981. Natural hybridization and gene introgression between *Xenopus gilli* and *Xenopus laevis laevis* (Anura, pipidae). *J Zool* 194(JUL):317-22.
- Kosman E and Leonard KJ. 2005. Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Mol Ecol* 14(2):415-24.
- Kozłowski J, Konarzewski M, Gawelczyk AT. 2003. Cell size as a link between noncoding DNA and metabolic rate scaling. *Proc Natl Acad Sci U S A* 100(24):14080-5.
- Kron P, Suda J, Husband BC. 2007. Applications of flow cytometry to evolutionary and population biology. *Annual Review of Ecology Evolution and Systematics* 38:847-76.
- Leitch IJ and Bennett MD. 2004. Genome downsizing in polyploid plants. *Biol J Linn Soc* 82(4):651-63.
- Leitch IJ and Bennett MD. 1997. Polyploidy in angiosperms. *Trends Plant Sci* 2(12):470-6.
- Liu B and Wendel JF. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. *Mol Phylogenet Evol* 29(3):365-79.

- Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF. 2001. Polyploid formation in 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 M. 1990. Tetraploidy in a liveborn infant. *J Med Genet* 27(12):782-3.
- 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 associated with Charcot-Marie-Tooth disease type-1a. *Cell* 66(2):219-32.
- Lynch M. 2006. The origins of eukaryotic gene structure. *Mol Biol Evol* 23(2):450-68.
- Lynch M. 2002. Intron evolution as a population-genetic process. *Proc Natl Acad Sci U S A* 99(9):6118-23.
- Lynch M and Milligan BG. 1994. Analysis of population genetic-structure with RAPD markers. *Mol Ecol* 3(2):91-9.
- Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J, Comai L. 2002. Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant Physiol* 129(2):733-46.
- Maguire TL, Peakall R, Saenger P. 2002. Comparative analysis of genetic diversity in the 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 *Gastrophryne carolinensis* (anura: Microhylidae). *Genetica* 135(2):169-83.

- Malone JH, Chrzanowski TH, Michalak P. 2007. Sterility and gene expression in hybrid males of *Xenopus laevis* and *X. muelleri*. Plos One 2(8).
- Masterson J. 1994. Stomatal size in fossil plants - evidence for polyploidy in majority of angiosperms. Science 264(5157):421-4.
- Matzke MA and Matzke AJM. 1998. Gene silencing in plants: Relevance for genome evolution and the acquisition of genomic methylation patterns. Epigenetics-Book 214:168-80.
- Matzke MA, Scheid OM, Matzke AJM. 1999. Rapid structural and epigenetic changes in polyploid and aneuploid genomes. Bioessays 21(9):761-7.
- McClintock, B. 1984. The significance of responses of the genome to challenge. Science 226: 792-801.
- McLysaght A, Hokamp K, Wolfe KH. 2002. Extensive genomic duplication during early chordate evolution. Nat Genet 31(2):200-4.
- Meudt HM and Clarke AC. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends Plant Sci 12(3):106-17.
- Mueller O, Hahnenberger K, Dittmann M, Yee H, Dubrow R, Nagle R, Ilsley D. 2000. A microfluidic system for high-speed reproducible DNA sizing and quantitation. Electrophoresis 21(1):128-34.
- Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99(3):247-57.

- O'Neill RJW, O'Neill MJ, Graves JAM. 1998. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* 393:68-72.
- Otto SP and Whitton J. 2000. Polyploid incidence and evolution. *Annu Rev Genet* 34:401-37.
- Ozkan H, Levy AA, Feldman M. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops triticum*) group. *Plant Cell* 13(8):1735-47.
- Pagel M and Johnstone RA. 1992. Variation across species in the size of the nuclear genome supports the junk-DNA explanation for the C-value paradox. *Proceedings of the Royal Society of London Series B-Biological Sciences* 249(1325):119-24.
- Panaro NJ, Yuen PK, Sakazume T, Fortina P, Kricka LJ, Wilding P. 2000. Evaluation of DNA fragment sizing and quantification by the Agilent 2100 bioanalyzer. *Clin Chem* 46(11):1851-3.
- Paun O, Fay MF, Soltis DE, Chase MW. 2007. Genetic and epigenetic alterations after hybridization and genome doubling. *Taxon* 56(3):649-56.
- Peakall R and Smouse PE. 2006. GENALEX 6: Genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6(1):288-95.
- Peakall R, Smouse PE, Huff DR. 1995. Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*. *Mol Ecol* 4(2):135-47.
- Picker MD. 1985. Hybridization and habitat selection in *Xenopus gilli* and *Xenopus laevis* in the southwestern cape province. *Copeia* (3):574-80.

- 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. .
- Pikaard CS. 2001. Genomic change and gene silencing in polyploids. Trends in Genetics 17(12):675-7.
- Pilsner JR, Lazarus AL, Nam DH, Letcher RJ, Sonne C, Dietz R, Basu N. 2010. Mercury-associated DNA hypomethylation in polar bear brains via the LUMinometric methylation assay: A sensitive method to study epigenetics in wildlife. Mol Ecol 19(2):307-14.
- Pogribny I, Raiche J, Slovack M, Kovalchuk O. 2004. Dose-dependence, sex- and tissue-specificity, and persistence of radiation-induced genomic DNA methylation changes. 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.
- Reik, W., and Walter, J. 2001. Evolution of imprinting mechanisms: The battle of the sexes begins in the zygote. Nat Genet 27:255-6.
- Rieseberg LH and Wendel J. 2004. Plant speciation - rise of the poor cousins. New Phytol 161(1):3-8.
- Rodgers L. 2006. Propidium Iodide (PI) or DAPI Staining of Unfixed Tissue Culture Cells for Flow Cytometry. Cold Spring Harb Protoc; doi:10.1101/pdb.prot4437

- Salmon A, Ainouche ML, Wendel JF. 2005. Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Mol Ecol* 14(4):1163-75.
- Scarborough PR, Hersh J, Kukulich MK, Carroll AJ, Finley SC, Hochberger R, Wilkerson S, Yen FF, Althaus BW. 1984. Tetraploidy - a report of 3 live-born infants. *Am J Med Genet* 19(1):29-37.
- Selker EU. 2004. Genome defense and DNA methylation in *Neurospora*. *Cold Spring Harb Symp Quant Biol* 69:119-24.
- Selker EU, Tountas NA, Cross SH, Margolin BS, Murphy JG, Bird AP, Freitag M. 2003. The methylated component of the *Neurospora crassa* genome. *Nature* 422(6934):893-7.
- Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13(8):1749-59.
- Shapiro H, editor. 2003. *Practical flow cytometry*. 4th ed. New York: Wiley-Liss.
- Shen, J. C., W. M. Rideout, III and P. A. Jones. 1994. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res* 22:972-6.
- Soltis DE and Soltis PS. 1999. Polyploidy: Recurrent formation and genome evolution. *Trends in Ecology & Evolution* 14(9):348-52.
- Song KM, Lu P, Tang KL, Osborn TC. 1995. Rapid genome change in synthetic polyploids of brassica and its implications for polyploid evolution. *Proc Natl Acad Sci U S A* 92(17):7719-23.

- Stancheva I, El-Maarri O, Walter J, Niveleau A, Meehan RR. 2002. DNA methylation at promoter regions regulates the timing of gene activation in *Xenopus laevis* embryos. *Dev Biol* 243(1):155-65.
- Starostova Z, Kratochvil L, Flajshans M. 2008. Cell size does not always correspond to genome size: Phylogenetic analysis in geckos questions optimal DNA theories of genome size evolution. *Zoology* 111(5):377-84.
- Stebbins GL, editor. 1950. Variation and evolution in plants. New York: Columbia University Press.
- Thiebaud CH and Fischberg M. 1977. DNA content in genus *Xenopus*. *Chromosoma* 59(3):253-7.
- Thompson JD, McNeilly T, Gray AJ. 1991a. Population variation in *Spartina anglica* C.E. Hybbard. I. Evidence from a common garden experiment. *New Phytologist* 117:115-128.
- Thompson JD, McNeilly T, Gray AJ. 1991b. Population variation in *Spartina anglica* C.E. Hybbard. II. Reciprocal transplants among three successional populations. *New Phytologist* 117:129-139.
- Thompson JD, McNeilly T, Gray AJ. 1991c. Population variation in *Spartina anglica* C.E. Hybbard. III. Response to substrate variation in a glasshouse experiment. *New Phytologist* 117:115-128.
- Tinsley, R. C. and Kobel, H. R. 2002. The biology of *Xenopus*. New York: Oxford University Press.



- Tucker KL. 2001. Methylated cytosine and the brain: A new base for neuroscience. *Neuron* 30(3):649-52.
- Tweedie S, Charlton J, Clark V, Bird A. 1997. Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol Cell Biol* 17(3):1469-75.
- Tymowska J and Fischber.M. 1973. Chromosome complements of genus *Xenopus*. *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: Uncovering duplicated segments in *Arabidopsis* by comparison with rice. *Trends in Genetics* 18(12):606-8.
- Verhoeven KJF, Van Dijk PJ, Biere A. 2010. Changes in genomic methylation patterns during the formation of triploid asexual dandelion lineages. *Mol Ecol* 19(2):315-24.
- Vermes I, Haanen C, Reutelingsperger C. 2000. Flow cytometry of apoptotic cell death. *J Immunol Methods* 243(1-2):167-90.
- 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 metabolic rate of mammals is related to genome size. *Evolution* 49(6):1249-59.

- Wang Zong-Gui, Zheng Wen-ling, Cui Dong, Ma Wen-li. 2004. The strategies of gene prolong expression in mammals. *Yichuan* 26(5):745-8.
- Wang JL, Tian L, Madlung A, Lee HS, Chen M, Lee JJ, Watson B, Kagochi T, Comai L, Chen ZJ. 2004. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* 167(4):1961-73.
- Wang JL, Tian L, Lee HS, Wei NE, Jiang HM, Watson B, Madlung A, Osborn TC, Doerge RW, Comai L, et al. 2006. Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* 172(1):507-17.
- Wang YM, Lin XY, Dong B, Wang YD, Liu B. 2004. DNA methylation polymorphism in a set of elite rice cultivars and its possible contribution to intercultivar differential gene expression. *Cell Mol Biol Lett* 9(3):543-56.
- Wendel JF. 2000. Genome evolution in polyploids. *Plant Mol Biol* 42(1):225-49.
- Wolfe KH and Shields DC. 1997. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387(6634):708-13.
- Woodward KJ, Cundall M, Sperle K, Sistermans EA, Ross M, Howell G, Gribble SM, Burford DC, Carter NP, Hobson DL, et al. 2005. Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination. *Am J Hum Genet* 77(6):966-87.
- Yager DD. 1996. Sound production and acoustic communication in *Xenopus borealis*. In: *The biology of Xenopus*. Tinsley, R. C. and Kobel, H. R., editor. Eds. Oxford: Clarendon Press. 121 p.

Yi S and Strelman JT. 2005. Genome size is negatively correlated with effective population size in ray-finned fish. *Trends in Genetics* 21(12):643-6.

Yoder JA, Walsh CP, Bestor TH. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics* 13(8):335-40.

## 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