1997

Establishment of a model genomic library for endangered insect species

Shizhi Yu
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Yu, Shizhi, "Establishment of a model genomic library for endangered insect species" (1997). *Master's Theses*. 1486.

DOI: <https://doi.org/10.31979/etd.py49-33hz>

https://scholarworks.sjsu.edu/etd_theses/1486

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

ESTABLISHMENT OF A MODEL
GENOMIC LIBRARY
FOR ENDANGERED INSECT SPECIES

A Thesis

Presented to
The Faculty of the Department of
Biological Sciences
San Jose State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

By

Miao Yu

May 1997

UMI Number: 1384735

**Copyright 1997 by
Yu, Miao**

All rights reserved.

**UMI Microform 1384735
Copyright 1997, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

© 1997

Miao Yu

ALL RIGHTS RESERVED

Approved for the Department of Biological
Sciences

Dr. Steve White
Associate Professor
Thesis Advisor

Dr. Robert. G. Fowler
Professor

Dr. Ruthann Kibler
Associate Professor

APPROVED FOR THE UNIVERSITY

ABSTRACT

ESTABLISHMENT OF A MODEL GENOMIC LIBRARY FOR ENDANGERED INSECT SPECIES

by Miao Yu

Large areas of forests are being destroyed. As a result, many insect species are endangered or have become extinct due to habitats lost. This thesis presents procedures to save genomic information of endangered insect species in genomic libraries.

A simple method was developed for preservation of high molecular weight (HMW) DNA for an extended period by storing homogenized tissue in diluted Dodecyltrimethylammonium Bromide (DTAB) solution at ambient temperature. After DNA was isolated using the DTAB/Cetyltrimethylammonium Bromide (CTAB) method, the HMW DNA was modified and ligated to lambda-GEM 12 vector to obtain two genomic libraries (mouse standard and insect libraries).

The percentage capture of libraries was estimated and it was lower than expected. There are many reasons that could explain the low percentage capture of a library. Some of them will be discussed in detail in this thesis. Some suggestions will be given to improve percentage capture in future work.

Acknowledgments

I would like to thank Dr. Steven White, my major advisor, who gave me the opportunity to work on this project, who originally suggested the nature of this work, and who has provided inspiration, support, and direction throughout the course of the project with his patience. Without him, this project could not have been completed. I would also like to express my sincere gratitude to my committee for their help and encouragement.

I thank all those who assisted me during the project, including Dr. Ronald Stecker and Dr. Gordon Edward for their valuable advice and Mr. Martin Sorich for his technical assistant during the work. This work would not have been completed without their help.

Thanks are also given to my parents and my brother for their patience and help when I need them.

The special thanks are extended to my lovely husband for his love, patience and support financially and spiritually.

Table of Contents

list of Figures	viii
List of tables	ix
Chapter I Introduction	1
Chapter II Materials and Methods	36
Chapter III Results	50
Chapter IV Discussion	67
Literature Cited	96
Appendix	101

List of Figures

- Figure 1 5
Steps involved in the construction of cDNA and genomic DNA libraries (from Current Protocols in Molecular Biology, 1989)
- Figure 2 33
Schematic model for partial fill-in reaction.
- Figure 3 87
Agarose gel electrophoretic (AgGE) fractionation of DNA isolated from lady bugs fixed intact (ie. "whole") in various solutions and stored for 21 days at room temperature prior to DNA extraction.
- Figure 4 88
AgGE fractionation of DNA isolated from lady bugs that were minced and stored in various solutions for 7 days (Panel A) or 27 days (Panel B) at room temperature prior to DNA extraction.
- Figure 5 89-90
AgGE fractionation of DNA isolated from lady bugs that were immersed and immediately minced and stored in DTAB solution for varying lengths of time (periods ranging from a few minutes to 1.5 years) prior to isolation of the DNA using the DTAB/CTAB procedure.
- Figure 6 91
Effect of differing NaCl concentrations on the yield of genomic DNA during the DTAB/CTAB DNA isolation procedure.
- Figure 7 92
Test for the presence of residual DNase activity in isolated genomic insect DNA and demonstration of isolated insect genomic DNA as a good substrate for restriction enzyme digestion.
- Figure 8 93-94
Determination of the restriction enzyme (RE) concentration required to generate genomic DNA fragments of the optimal size range.
- Figure 9 95
Large-scale digestion of insect and mouse genomic DNAs using Sau 3A.

List of Tables

Table 1 Restriction titration for mouse DNA	59
Table 2 Restriction titration for insect DNA	59
Table 3 Ligation reaction	63
Table 4 Results of ligation and packaging reaction	66

Chapter I

Introduction

While rapid industrialization, modern technology and expanding urbanization have in many ways dramatically improved the quality of human life, these same "developments" are also swiftly depleting the natural resource base of the planet. The world's forests, for example, represent a complex resource that has undergone significant destruction within the last 100 years. In many parts of the world (especially in poorer, underdeveloped countries exhibiting a high birth rate and expanding population) the forests are being destroyed simply to supply basic survival needs... clearing land to grow food and feed livestock, harvesting hardwoods to construct simple shelters. The "slash-and-burn" agriculture typically used in poor countries is often then further exacerbated by larger-scale industrial exploitation... the destruction of forest lands as a result of laying down logging and mining roads, the over-harvesting of hardwood lumber as a "cash crop", the stream and river pollution associated with mining operations, the flooding of forest lands as a result of hydroelectric power production, etc. Under this continued

assault large areas of world's forests have simply disappeared.

Tropical forests, for example, are disappearing at the rate of 20.4 million hectares each year (Warlker, 1991), roughly equivalent to an area the size of Nebraska. The continually diminishing size of the forests does not, however, reflect a reduced importance. Mature tropical forests cover just 7% of earth's land surface but are believed to still hold over 50% of the planet's species (Warlker, 1991). Since 90% of world's insect species are believed to be found in tropical forests (Warlker, 1991), many of these insect species are now endangered (or at the brink of extinction) due to habitat loss. This profound environmental destruction and projected loss of insect biodiversity are well known, but little success has been met in halting or reversing this trend. It is predicted that as a result of the expanding human population in poor countries and the high demand for forest products in industrialized ones, it will be increasingly difficult to halt the cycle of habitat destruction and species extinction.

This situation has recently driven biochemically-trained conservationist to look for and/or develop methods that might be used to at least preserve the genomic information contained within endangered species. The development of DNA cloning techniques provides an

opportunity to preserve the genetic information (as archival genomic libraries) of endangered species whose habitats seem doomed to certain destruction. A variety of zoos, for example, have committed themselves to the construction of genomic libraries from endangered mammals, and various pharmaceutical firms have begun cloning plant species that may possess commercial value. Little attention, however, has been given to insects. Towards this end, we sought to develop a practical method for stabilizing DNA *in the field* (ie. on location in the rain forest), for isolating this DNA in high molecular weight form, and to demonstrate the functionality of this DNA by creating a genomic library.

General review of gene libraries

Commonly used gene libraries in biological studies fall into two categories, cDNA libraries and genomic libraries. A cDNA library is a representation of genes active at the moment of mRNA isolation. Those cDNA libraries are thus extremely useful in the study of gene expression and in efforts directed towards the identification of functional genes. For the goal of this research, however, there are two problems with a cDNA library. First, a cDNA library lacks all the flanking and intron sequences that are present in the corresponding genomic DNA library. Those sequences often serve regulatory functions that are essential for the

control of gene expression. Second, cDNA libraries represent only transcriptionally active genes. Genes which are active only in certain developmental stages or are expressed in tissues different from those that served as the source for RNA isolation may not appear in cDNA libraries. A cDNA library, therefore, will never represent the complete genome of a given species and will thus not achieve the archival purpose envisioned by this project.

In contrast, a genomic library is expected to contain all the DNA sequences of the genome from which it has been derived. To construct a library with as complete a coverage as possible and with as few clones as possible, the cloned DNA fragments should randomly represent the genome in the library. Under those conditions the probability (P) that a unique DNA sequence is represented in N plaques is given by

$$N = \ln(1-P)/\ln(1-f)$$

where f is the size of the fragment expressed as a fraction of the genome (Clark & Carbon, 1976; Seed et al., 1982).

For preservation of genomic information, complete genomic libraries are required (ie., P should be at least 0.95, and ideally would be 0.99).

Construction of genomic libraries

The fundamental processes underlying construction of genomic libraries are shown in Figure 1: (i) isolation of

high molecular weight (HMW) DNA from tissues, (ii) generating appropriate-sized DNA fragments, (iii) joining foreign DNA to a vector, (iv) introducing the recombinant molecule into a host cell in which it can replicate, and (v) screening for an interested clone of recipient cells that has the required recombinant or amplifying for storage.

Figure 1. Steps involved in the construction of cDNA and genomic DNA libraries (from Current Protocols in Molecular Biology, 1989)

TISSUE PRESERVATION

Isolation of purified HMW DNA is an important part of the construction of genomic libraries. Many DNA extraction procedures are available, but the quality of DNA obtained varies significantly between methods. These differences can be important, since the quality of DNA that is appropriate for some techniques (e.g., polymerase chain reaction amplification, which can successfully use small amounts of highly sheared DNA) may not be suitable for others (e.g. restriction enzyme mapping and DNA fingerprinting, which typically require larger amounts of intact HMW DNA). Construction of lambda-based genomic libraries, for example, requires that the following conditions be met: (i) the DNA should be high molecular weight (size fragments at least 100 Kilo base pairs (Kbp) after isolation), (ii) the DNA should be rigorously freed of nucleases and other proteins, chaotropic salts, heavy metals, precipitating resins, RNA, and agarose fragments, and (iii) the high molecular weight DNA fragments should be intact (ie., the fragments should not have suffered cross-linking, alkylation, or depurination). Such "high quality" DNA is required if the nucleic acid is to be efficiently digested, modified (ie. phosphatase-treated) or ligated. Clearly, any tissue fixatives which compromise the quality of the DNA must be avoided.

Tissue from a living organism is the the most commonly used source for isolating HMW insect DNA. Unfortunately, there are many problems associated with bringing live insects from the field back to a laboratory, not the least of which involve how to keep the insects alive during the three to nine week long transportation period (ie. problems of proper thermal control, obtaining a proper food source, and meeting water requirements while making the cage "escape-proof" during weeks in transit) and how to legally satisfy the medical and agricultural restrictions associated with live insect importation. Conversely, while the procedures most commonly employed during a DNA extraction typically require a well-equipped laboratory, it is simply not realistic to set up such a laboratory in the field (ie. deep in the Amazon rain forest). These restrictions thus require that all insect specimens first be killed in the field, the DNA stabilized (to prevent degradation), the material rendered non-infectious (many insects act as vectors for disease-producing microbial organisms), and the properly packaged material then sent to a laboratory for DNA extraction. Immediate inactivation of nuclease activity is critical if HMW DNA is to be preserved, since under normal conditions cellular necrosis is coincident with nuclease activation and DNA degradation. A rapid and effective fixation method (ie. one that immediately and completely

inactivates nuclease activity while not damaging the DNA) is thus required.

Investigators have developed a variety of methods to preserve tissue for DNA isolation. The most effective method is freezing the tissue at or below -80°C . Often small tissue samples or cell pellets are simply plunged into liquid nitrogen (about -176°C) and allowed to "snap freeze". At these "ultra-low" temperatures the samples can be stored for an almost unlimited time without detectable DNA degradation. Freezing insects in liquid nitrogen, however, has obvious drawbacks in the field. The logistics of transporting massive steel cylinders of pressurized liquid nitrogen and/or large, heavy Dewar's flasks through the tropical rain forest are daunting. The high volatility of liquid nitrogen further limits its field use since tissue samples frozen in liquid nitrogen must be maintained at that temperature for the remainder of the study (ie. weeks or months), and it is highly unlikely that stocks of liquid nitrogen could be easily and rapidly replenished.

One simple alternative method for preservation of HMW DNA involves rapid air drying of whole blood (or a similar protein-rich cell suspension). Aggarwal et al. (1992), for example, have shown that HMW DNA could be isolated from small samples of alligator or crocodile blood which was spotted onto clean slides and air dried at room temperature.

After drying, the blood samples on slides can be transported or stored at room temperature for two-three months and still serve as good sources of HMW DNA. Unfortunately, while this method may be adequate for animals with closed circulatory systems containing blood rich in nucleated cells, it has little practical application in our research effort since insects possess only an open circulatory system with few circulating blood cells (Fox & Fox, 1964).

Another approach is to chemically and physically stabilize the DNA via the use of one or more "fixatives". For use under field conditions, an ideal fixative should possess the following characteristics.

- 1) *The fixative itself must be stable under field conditions.* Since the entire fixation process must be performed under field conditions, the components of a fixative should be thermally stable (ie. to at least 45°C over prolonged periods), should be highly resistant to oxidative damage and UV attack, should be of low volatility, and if it is a "reactive fixative" (ie. as are many aldehyde-based fixatives) should not be significantly depleted (even when employed at high tissue:fixative mass ratios).

- 2) *The fixative must penetrate tissue quickly and act rapidly.* While slow tissue fixation protocols are often

acceptable in routine histological procedures, those would not be possible here. Preservation of DNA as high molecular weight material requires the immediate and complete inactivation of DNase activity. The use of a slowly penetrating fixative would provide time for DNase to hydrolyze the DNA into fragments too small for use in genomic library construction.

3) *The fixative must preserve the DNA as HMW material under field conditions (ie. elevated temperatures, physical agitation etc) for prolonged periods (ie. many months).*

Fixatives should have the ability to preserve HMW DNA for a relatively long time so that there is enough time for shipping the samples from field location to a laboratory.

4) *The tissue preparation (if any) and fixation procedure should be convenient.* The fixation process should be easy for field biologists to use, employing a minimum of equipment and tissue preparation steps (ie. a simple homogenization step or heating step may be acceptable while a centrifugation step would not).

5) *The fixative should be relatively safe and should not present a transportation problem when crossing international borders.* The ideal fixative should be relatively safe for humans (ie. exhibit low volatility, low

flammability, low toxicity, and low carcinogenic or teratogenic potential) during the operation and transportation/mailling processes. Ideally, the fixative should also require no elaborate (especially large or heavy) transportation packages.

The ability of commonly used cell and tissue fixatives (ie., those routinely used in light microscopy) to simultaneously fix tissue and stabilize/preserve HMW DNA has been recently reexamined. Formaldehyde, for example, has been used routinely in a wide variety of cytological studies since before the turn of the century (Humason, 1979), but only recently has its ability to preserve intact high molecular weight DNA been investigated (Koshiba et al., 1993). Koshiba et al., (1993) found that formaldehyde-based fixatives preserved HMW DNA when fixation and storage were carried out at 4°C but, interestingly, was unable to do so when material was fixed and stored at room temperature (Koshiba et al., 1993). Degradation occurred in formalin when the pH or salt concentration was low, or the formic acid level was high. Relatively intact DNA (ie., about 50 Kbp) was only obtained from tissues fixed and stored in 10% phosphate-buffered formalin (pH 7.4, 173 mM Na salt) at 4°C (Koshiba et al., 1993). The low temperature requirement and inherent toxicity of formaldehyde, however, rendered this

fixative essentially useless for the field applications we envisioned.

Reviewing the literature thus suggested that there was still much to be learned about the ability of simple fixatives (especially low molecular weight alcohols) to simultaneously fix tissue (thus preserving tissue morphology) and preserve/stabilize HMW DNA. The use of these alcohols seemed particularly attractive, since it was expected that these fixatives would be fairly easily removed and/or compatible with downstream DNA extraction procedures. Therefore studies were conducted on the ability of methanol, ethanol and isopropanol to both fix tissue and preserve and stabilize HMW DNA. In the interests of comparison, the ability of a commonly used histological fixative, Carnoy's solution (a mixture of methanol and glacial acetic acid), to fix tissue and preserve HMW DNA was also investigated.

DNA ISOLATION

Classical approaches in molecular biology essentially trade off tissue preservation for rapid protein denaturation (especially nuclease inactivation) and DNA stabilization and solubilization. Perhaps the best known and most widely used method to effect simultaneous protein denaturation and DNA solubilization involves protease K digestion followed by phenol-chloroform extraction (Eickbush et al., 1978; Adeli

et al., 1990). This method, however, has a number of limitations. It is difficult to remove residual phenol from the extracted DNA, and this phenol in turn may affect the function of enzymes used in subsequent procedures. The repeated phenol extractions, and multiple chloroform extractions generally employed in removing residual phenol, also often cause significant loss of DNA (Singer, et al., 1989), generating unacceptably low yields when starting from limited amounts of tissue (a situation routinely expected, given that the endangered insect specimens will most likely be extremely rare). More importantly, carrying out the enzymatically-based protein destruction step using protease K would not really be practical under field conditions, since long-term storage of this enzyme at -20°C in the rain forest would be "difficult" at best. Finally, phenol is extremely sensitive to oxidative and UV attack, especially under conditions of elevated temperature (Clark, 1964), as would be likely in rain forest environments.

The use of detergent and guanidium-based salt solutions to effect rapid tissue disruption, protein denaturation and DNA solubilization was considered. Standard formulations of guanidium salt solutions, however, are also considered to be UV and heat sensitive (personal communication, Sigma technical representative) and are classified as hazardous

and toxic (making mail shipment across international borders more cumbersome), and were thus not used in this study. Most of searches for an optimal method thus focused on the initial use of powerful denaturing detergents, since such reagents are expected to exhibit high thermal stability, relatively low toxicity, and good compatibility with other known DNA extraction protocols.

A DNA isolation method involving the detergent Cethyltrimethylammonium Bromide (CTAB), was initially used in bacteria (Jones, 1963) and later modified to obtain DNA from plant tissues (Murray et al., 1980; Doyle et al., 1990; Roger et al., 1985). CTAB forms an insoluble complex with nucleic acids when the initial NaCl concentration is low, about 0.5 M (Roger et al., 1985). Since the nucleic acid-CTAB complex is only soluble in high salt, the detergent is removed by raising the NaCl concentration to solubilize the DNA. This is followed by ethanol precipitation and centrifugation to separate the pelleted nucleic acids from the detergent-rich supernatant. The CTAB DNA isolation method was modified to extract DNA from insect tissues. Boyce et al. (1989), used a protocol modified from one developed by Doyle et al. (1987) to isolate total genomic DNA from Bark Weevills.

The detergent Dodecyltrimethylammonium Bromide (DTAB) represents a more recently developed thermally stable detergent exhibiting extremely powerful denaturing and tissue solubilization properties. It was the recognition of DTAB's ability to rapidly denature proteins (Gustincich et al., 1991) that led to the development of a linked DTAB/CTAB DNA extraction protocol for use with freshly isolated human blood. In this protocol, DTAB is used to both rapidly lyse and solubilize tissue and to immediately denature all proteins in the lysate. The denatured proteins are then removed and DNA partitioned into the aqueous phase via a chloroform extraction. CTAB is then added to the aqueous phase to selectively precipitate DNA as described above for the original CTAB procedure. We have developed various modifications of this original DTAB/CTAB DNA isolation protocol, and the one we have found to be the best suited for isolation of DNA from small amounts of insect tissue will be presented.

DNA PARTIAL DIGESTION

Construction of a genomic library involves the fragmentation of HMW genomic DNA, since extremely long fragments are not suitable for insertion into typical cloning vectors. Some methods do not involve restriction enzymes. One of them is random breakage of the DNA molecule

by simple mechanical shearing, which produces non-specific breakage of long double stranded DNA. The thread-like duplex DNA molecule is very rigid and can easily be broken by shearing force in solution. The use of intense sonication with ultrasound can further reduce the length of DNA molecules, down to an average size of about 300 base pairs (bp). Both methods result in non-specific breakage so that the ends of a DNA fragment are irregular and undefined, a condition which increases the difficulty of subsequent ligation reactions. The use of restriction endonucleases (RE) is much more common, however, since such enzymes produce well-defined DNA fragment ends after cutting.

Partial digestion is the most commonly used method to prepare DNA for genomic library construction. There are two basic methods for generating relative large size fragments (≈ 20 Kbp). One is restricting the DNA with a mixture of two REs (Maniatis et al., 1978). Those enzymes have tetranucleotide recognition sites which occur frequently in target DNA. The restriction reaction is carried out to only partial digestion so that fragments with sizes between 15 to 20 Kb can be produced. DNA fragments within this size range are suitable for insertion into a phage vector. A more common and convenient method involves using a single RE to carry out a partial digestion. One of often used enzymes in this application is Sau3AI, which cuts the DNA frequently.

Since Sau3AI generates fragments compatible with several convenient lambda vectors, a great advantage of using it is that Sau3AI-digested DNA fragments can be readily inserted into high-capacity phage vectors which are digested by other restriction enzymes. This partial digestion method coupled with an *in vitro* packaging reaction has been widely used for creating genomic libraries (Firschauf, 1978).

PREVENTION OF CLONING ARTIFACTS

Genomic libraries are constructed using partially digested DNA. These DNA fragments should have ends rendering them capable of joining covalently (ligating) to vector DNA in the presence of DNA ligase. DNA fragments created by digestion with a single restriction enzyme will, however, also be able to ligate to each other. When insert DNA is ligated into the phage vector, it is very important to prevent ligation of genomic fragments that are not contiguous in the genome. Three methods described here are commonly used to prevent construction of such "artificial" sequences.

Size fractionation in sucrose gradients and gels.

Partially digested DNA can be size fractionated by centrifugation through sucrose gradients. Digested DNA is laid on to the top of a 10-40% (w/v) sucrose gradient and the gradient is subjected to ultracentrifugation at 150,000

x g for 16-24 hours, during which time the DNA fragments will migrate according to their size (mass); the smallest fragments will have migrated farthest down the tube (usually into the 40% sucrose zone) while the longest fragments will have migrated the shortest distance. The resulting gradient is then fractionated (typically 20-40 fractions are collected), and small aliquotes of each fraction analyzed by agarose gel electrophoresis (together with size marker fragments) to determine the size distribution of the fragments in each fraction. The DNA in fractions of the correct size can then be precipitated. The disadvantages of the method are that it is very time consuming, requires a costly ultracentrifuge, rotor(s) and tubes, and typically requires that large amounts of DNA be processed.

Size fractionation also can be carried out by gel electrophoresis. The challenges here are loading sufficient amounts of DNA on a gel and then purifying the sized fractionated DNA fragments from the isolated agarose plugs following electrophoresis. One method for extraction of DNA from agarose gel plugs is to add agarase to destroy glycosidic bonds in the agarose allowing the agarose strands to revert to the solution state; the DNA fragments thus released are then purified away from the agarose strands either by filtration or selective precipitation. Unfortunately, this method does not work efficiently for

large DNA fragments, and it is also very costly in terms of time and materials. Since this research project deals with endangered insects, there typically will not be large amount of starting materials to work with. It is thus highly advantageous to use as small an amount of DNA as possible to construct a complete genomic library. As a result, size fractionation either by sucrose gradients or agarose gels was not deemed suitable for this study.

Dephosphorylation. Treatment of DNA with alkaline or calf intestinal phosphatase leads to removal of the 5'-phosphate from the DNA strands and prevents ligation of any two dephosphorylated DNA molecules; this method thus demands that only one population of DNA (ie. either vector or insert) be modified. Since it is important to prevent ligation of the genomic DNA fragments to each other, dephosphorylation of insert DNAs is typically performed. The lambda arms, which must remain fully phosphorylated, can therefore still self-ligate. Often times a large fraction of the right and left phage arm populations become consumed in this non-productive (non-insert carrying, non-packagable) ligation process. Large amounts of lambda arms are thus needed using this protocol, and when only a limited amount of lambda arms is available this self-ligation will result in decreased packaging efficiency.

Partial fill-in reaction. The use of a partial fill-in (partial "back fill") reaction prevents the random association and covalent linkage of genomic DNA fragments during cloning. In essence, the Klenow fragment of DNA polymerase I and a reaction mixture containing only two of four dNTPs are used to partially fill in the sticky ends of both insert and vector DNAs. There are several significant advantages to the use of a fill-in reaction. First, by converting the ends of the genomic fragments into non-hybridizable termini, the fill-in reaction prevents the ligation of genomic DNA fragments and thus the production of randomly ligated "hybrid artifact" inserts. As a consequence of this conversion, the fill-in reaction also prevents the formation of over-sized "hybrid artifact" inserts (ie., hybrid inserts larger than 23 Kbp). These over-sized inserts, when ligated to phage arms, would exceed the maximum packagable size. Second, subjecting the phage arms to a partial fill-in reaction will prevent the association and ligation of right and left arms in the absence of inserts. This increases the efficiency of the *in vitro* package reaction and decreases the number of arms used for a given amount of insert DNA.

LIGATION

Currently, three methods are available for joining the DNA fragments together. The first depends on the ability of DNA ligase to covalently join the annealed cohesive ends produced by restriction enzymes. The second method uses DNA ligase (at very high unit activity ratios) from phage T4 infected *E. coli* to catalyze the formation of phosphodiester bonds between blunt-ended fragments. The third method utilizes the enzyme terminal deoxynucleotidyltransferase to make complementary homopolymeric 3'-single-stranded tails at the ends of insert and vector DNA fragments (ie., adding either dATP/dTTP or dCTP/dGTP at the ends). Among the three, the first method is usually used in genomic library construction.

Joining target DNA fragments and lambda arms was carried out by T4 DNA ligase. T4 DNA ligase in the presence of ATP and dNTPs has the ability to seal single-stranded nicks between the adjacent nucleotides in a duplex DNA, forming intact strands (Olivera et al., 1968). The ideal temperature for optimum functioning of DNA ligase is 37°C, but at this temperature the hydrogen-bonded joint between the cohesive ends is unstable. Commonly used restriction endonucleases produce ends with 4 bp overhanging, and at 37°C the hydrogen bonds are unable to hold two termini together. If the two fragments disassociate, ligase will

not work. Since higher temperature favors greater enzyme activity, and lower temperature favors association of cohesive ends, there is a balance between the activity of enzyme and the degree of annealing of the cohesive ends. Therefore, the optimum temperature for the ligation reaction is typically in the range of 10-15°C (Dugaicyzk, et al., 1975; Ferritti & Syaramella, 1981). In the procedures presented in this thesis, however, there are only two bases overhanging at the ends after a partial fill-in reaction. It is thus extremely important to keep the reaction temperature as low as possible in order to minimize strand disassociation. Since 4°C is not the optimum temperature for the DNA ligase function, prolonged incubation times (12-24 hours) were used in an attempt to maximize ligation efficiency.

IN VITRO PACKAGING

Originally, the conversion of recombinant DNA to viable phage particles was carried out by transfection of Ca^{+2} -treated, permeabilized cells with DNA. The subsequent gene expression and multiplication of the recombinant DNA produced functional phage (Lederbery & Cohen, 1974). The efficiency of transfection, however, is very low. Newly developed technologies allow encapsulation of the

recombinant lambda DNA *in vitro*, using exogenously supplied phage coat proteins.

The principle of packaging *in vitro* is to supply the ligated recombinant DNA with a high concentration of phage head precursors, packaging proteins, and phage tails. *In vitro* packaging of recombinant DNA molecules into mature phage particles results in two to five logs greater transformation efficiency than do CaCl_2 transfection procedures (Mandel & Higa, 1970; Hohn & Murray, 1977). Since well-developed protocols and materials are available for phage packaging *in vitro*, commercially prepared packaging extracts ("Packagene" extracts from Promega, Inc.) were chosen for the construction of the genomic libraries.

VECTORS

In the most general sense, vectors are autonomously replicating DNA molecules into which DNA fragments are inserted for amplification ("cloning"), expression and/or operative manipulations (DNA sequencing, for example). DNA fragments taken up by these vectors are replicated as the vectors replicate to yield numbers suitable for subsequent molecular studies. There are many different vectors used in cloning studies, the most common being: plasmids, cosmids and phagemids, bacterial viruses (ie. phage like Lambda,

M13, Mu, P2), eukaryotic cell viruses (represented most heavily by certain adenoviruses, herpes viruses and retroviruses), yeast artificial chromosomes (YACs) and the recently developed bacterial artificial chromosomes (BACs). A brief description of the more important vectors and our rationale for selecting lambda phage is presented below.

Plasmids

Plasmids are autonomously replicating double stranded circular DNA molecules widely used in cloning, gene expression and sequencing efforts. An ideal plasmid cloning vehicle is one that has low molecular weight, has the ability to confer readily selectable phenotypes on the host, and has single sites for a large number of restriction endonucleases. Other important features might include prokaryotic and eukaryotic promoter cassettes, enhancer sequences, splicing and polyadenylation sequences, depending upon the particular applications intended. There are two general classes of plasmid vectors: the "naturally" occurred plasmids (e.g., Col E1 and pSC101), and artificially constructed plasmids (e.g., pBR322). Both types of plasmid cloning vectors typically have a common disadvantage: most are capable of accepting only small "insert" fragments. If the size of inserted DNA is too large, these plasmids (and thus the inserts they carry) are unstable in host cells.

This size-dependent instability *generally* makes plasmid vectors a poor choice for constructing genomic libraries where inserts are typically quite large.

YACs

Yeast artificial chromosomes (YACs) are typically constructed when it becomes necessary to clone very large genomic DNA fragments. The yeast artificial chromosomes are constructed by ligating the foreign DNA between two linearized plasmid components. One plasmid contains at least one prokaryotic replication origin (ori), one yeast autonomously replicating sequence (ARS), some type of prokaryotic selection marker (often one conferring antibiotic resistance), some type of eukaryotic selection marker (often a LEU2⁺ marker), and insertion site (ideally a multiple cloning site), and a telomere sequence (either from the protozoan *Tetrahymena*, or from yeast itself). The other plasmid component should contain all of the above sequence elements plus a yeast centromeric (CEN) sequence. The presence of the prokaryotic selection marker (ie. Amp^r) will allow propagation and amplification of the plasmid vectors in bacteria, while the LEU2⁺ marker can be used for selection of transformants in yeast (Murray & Szostak, 1983). The YACs can clone DNA fragments from about 50 Kbp to up to 1000 Kbp in length and replicate in yeast cells

very efficiently, but with smaller fragments YACs do not work well. Besides the fact that the stability of very small YACs may actually be lower than that of circular plasmids, there are other problems that make working with YACs as cloning vectors difficult. Transformation of yeast cells, for example, is neither particularly convenient nor highly efficient; this makes transformation of recombinant DNA molecules tedious and difficult. Isolation of intact YACs is also a laborious process.

BACs

The bacterial artificial chromosome (BAC) is based on *Escherichia coli* and its single-copy plasmid F factor. The pBAC108L vector, for example, includes the genes that are essential to control replication (ie., *oriS*, *repE*) and copy number (ie., *parA* and *parB*) of the F plasmid as well as a chloramphenicol resistant marker and a cloning segment. The cloning segment involves lambda *cosN* and P1 *loxP* sites for specific cleavage by RE, two cloning sites (*HindIII* and *BamHI*) for insertion of foreign DNA, G-C rich sites for rare-cutter restriction enzymes (*NotI*, *EagI*, *SmaI*, *BglI* and *SfiI*) for potential excision of inserts, and a sequence containing T7/SP6 promoters for generating RNA probes or DNA sequencing (Shizuya et al., 1992). This circular plasmid is

capable of maintaining human genomic DNA fragments up to 300 Kbp with a high degree of structural stability in the host, even after 100 generations of serial growth. Equally important, the large plasmid can be transfected to *E. coli* by electroporation with 10-100 times higher efficiency than yeast spheroplast transfection. The benefit of a low copy number of F plasmid is that it reduces the potential for recombination between DNA fragments carried by the plasmid. The other advantage of BACs over YACs is that BACs allow easier manipulation of the large DNA molecules in solution with less of a shearing problem than is typically seen with YACs, since the BACs exist as supercoiled circular plasmids in *E. coli* (Shizuya et al., 1992) and thus represent a more hydrodynamically stable structure. Isolation of plasmids in the 70-100 kbp range, however, is still fairly difficult, and typically results in very low yields. Screening BAC libraries is also difficult, due in large part to the fact that colony lift assays must be performed using clones that typically possess only a few copies of each BAC per cell (thus creating a problem of low target abundance during the screening effort).

Cosmids

A cosmid is the combination of a plasmid and a fragment of DNA including the *cos* site (Collins & Bruning, 1978;

Collins & Hohn, 1979). The cosmid can be used as the cloning vector in conjunction with the *in vitro* packaging system. Because a cosmid is a small molecule (5 Kbp) it can clone DNA fragments as large as 32-40 Kbp, much larger than a lambda phage vector can accommodate. Cosmids, packaged *in vitro*, can then infect suitable hosts. The recombinant cosmid infects the host and circularizes as phage DNA, but then replicates as a normal plasmid without the expression of any phage functions. Cosmids provide an efficient means of cloning "relatively" large pieces of foreign DNA. Because of their capacity to accept large fragments of DNA and the high transformation efficiency afforded by *in vitro* packaging, cosmids are often used for constructing eukaryotic genomic libraries. Like the BAC system, however, screening still involves a colony lift procedure, greatly increasing the time required to screen the entire library.

Bacteriophages

While cosmids are rapidly gaining popularity, lambda phage remains the most commonly used vector for rapidly constructing genomic libraries. The DNA of the phage is a double-stranded linear molecule of about 48.5 Kbp in size. The center section, about 25% of the genome, is not essential for lytic functions of the phage. This

"unessential DNA" is an advantage, since it can be removed and replaced with a foreign DNA fragment without impairing the ability of the phage to infect and reproduce in host cells. Many derivatives of the wild-type phage have been produced to act as cloning vectors. The newly developed phage vectors are improved in several respects, such as an increased capacity for foreign DNA fragments (Murray, 1983), alterations allowing for efficient positive selection of recombinants, and the development of phage expression vectors. The most recent generation of vectors combines a large capacity for foreign DNA (maximum 23 Kbp) together with features that allow "relatively" simple and efficient library construction. Addition of an *in vitro* phage packaging extract to these recombinant viral genomes then supplies the phage coat proteins and enzymes required for the formation of functional phage. Packaging *in vitro* increased the likelihood that an appropriately large number of independent recombinants can be recovered, which theoretically would give an almost completely representative library.

As stated above, however, while both phage and cosmids can be used to clone relatively large fragments of DNA, lambda phage libraries offer certain practical advantages. First, screening libraries of phage recombinants by plaque hybridization is quicker (phage plaques are screened instead

of bacterial colonies) and the signal typically cleaner (less background hybridization is usually encountered when screening phage plaques). Second, phage libraries are more efficiently stored and amplified. With phage, the initial recombinant DNAs are packaged and plated out. The plate containing recombinant plaques can be washed to give an amplified library of recombinant phage, and these recombinant phages can be stored almost indefinitely (since phages have a long shelf-life) and/or replated and screened with different probes at later times. While bacterial colonies containing recombinant cosmids can also be stored and amplified (Hanahan & Meselson, 1980), bacterial cells tend to suffer a greater degree of viability loss during the freeze-down and thawing procedures. Third, if an expression vector is used, it is entirely possible that some proteins may be produced that are toxic to *E. coli*. The use of phage, which rapidly takes over the host cell machinery in the production of progeny, would be more likely to maintain such sequences in the library (ie. even if a DNA sequence encoding a toxic protein were expressed, at least some progeny phage would probably be produced during the rapid lytic cycle). In contrast, cosmids containing DNA sequences which code for such "toxic proteins" would quite rapidly be deleted from the library. These reasons, when combined with the vast body of literature on phage and phage libraries and

the many commercially produced reagents for phage library construction now available, makes lambda phage a very suitable vector for this project.

THE EXPERIMENTAL SYSTEM.

Choice of host cells.

Recombinant molecules need to be introduced into suitable host cells for screening or amplification. The most commonly used host for gene libraries is *E. coli*. The genomic libraries in the phage vectors are expected to contain most of the sequences of the genome from a given species. Some sequences, however, may not be found in a library of an eukaryotic genome because deletions can occur during cloning of eukaryotic DNA. The evidence shows that inverted repeat sequences can not be cloned in phage unless special host strains are used (Leach & Stahl, 1983). It is very important to choose a suitable host since some deletions are preventable by growing recombinant phage on mutant strains. *E. coli* LE392 strain [F^- , *hsdR514*, *supE44*, *SupF58*, *lacY1* or $(lacIZY)_6$, *galK2*, *galT22*, *metB1*, *trpR55* $^-$] was chosen as host cells for this study. There are a number of advantages to the use of LE392: (i) This strain is restriction-deficient and modification-deficient; thus it does not digest or modify exogenous DNA (i.e., phage and/or

insert DNA). (ii) LE392 does not contain an F plasmid and therefore lacks the ability to transfer its genetic material to other cells. (iii) This strain contains *supE44* and *supF58* that will suppress the mutations in phage and host cells.

The Lambda GEM-12 replacement vector.

Many types of lambda vectors are commercially available, most of which exhibit reasonably high efficiencies in ligation and packaging reactions. The GEM-12 vector (available from Promega Corp.) was chosen for this study since it was designed for use in a Klenow back-fill reaction (Figure 2). The lambda DNA is digested with XhoI and then partially filled at the sticky ends of the XhoI sites to leave dinucleotide 5'overhangs. These arms can then form complementary base pairs with the insert DNA fragments which have themselves been partially filled-in at Sau3A sites. The real value of this vector rests in the fact that the possibility of insert-to-insert ligations (and thus the production of "artifactual sequences") is minimized. The added burden of the Klenow back-fill reaction (and subsequent repurification) was worth the benefit of minimizing the production of artifacts in the genomic libraries.

Figure 2. Schematic model for partial fill-in reaction. XhoI sites cannot self-ligate but can accommodate partially fill-in, non-self-ligatable Sau3A ends (Old & Primrose, Principles of Gene Manipulation, 1994)

Objective of the study

Creation of genomic "archival" libraries for endangered insect species whose habitats are being destroyed is the goal of this research project. In this particular study, a model of the genomic archival library was established. The research primarily concentrated on testing various DNA stabilization reagents and protocols for preservation of HMW DNA that could be initiated in the field and finished in the research laboratory, followed by construction of a genomic library using this DNA. Manipulation of the library (ie. screening using homologous probes from other species, estimation of the fraction of unique vs moderately repetitive vs highly repetitive sequences via Cot analysis) was not intended to be a part of this study, although the library is now available for such analysis.

Significance of the study

Molecular genetic techniques provide powerful tools to both analyze and preserve genomic information. This work represents our first step in indentifying "field workable" protocols that will allow us to establish genomic libraries from endangered ("soon-to-be-extinct") insect species. As such, this work also represents an archival record, as much as it does a "methods development" thesis. Ultimately, it is also hoped that these archival genomic libraries will be

of benefit to those studying insect systematics, population biology, biogeography, evolutionary relationships between species and, perhaps, a reminder of the frailty of life on a planet under human siege.

Chapter II

Materials and Methods

This chapter covers specimens, detailed experimental procedures, and techniques used in this research.

REAGENTS AND EQUIPMENT.

Live specimens of *Hippidamia converglms*, a type of beetle commonly referred to as a "lady bug"), were purchased from a local nursery store. The individual beetles averaged about 6 mm in length and 8.5 mg in weight. Mouse tissue was obtained from freshly killed Balb/c mice and stored at -80°C until ready for use. Lambda GEM-12 arms, Packagene Extracts, restriction enzymes PstI and Sau3A, Klenow DNA polymerase and T4 DNA ligase were purchased from Promega Corporation. All salts, stains and agaroses were obtained from Sigma Chemicals Inc. Power sources and gel electrophoresis units were from Hoefer and BioRad, respectively. Photographs of gels were obtained using a Polaroid "instant" camera (Fisher Biotech Electrophoresis systems) and Type 667 high speed black and white film. The concentrations of DNA solutions were measured using either a

Fluorescence Spectrophotometer (Perkin-Elmer 204) or an UV Spectrophotometer (Shimadzu UV 160U).

EXPERIMENTAL PROCEDURES

This section describes all the test procedures for all protocols conducted during this study. The best protocols were selected later (Chapter III) to establish the complete process for the genomic library construction.

TISSUE PRESERVATION PROTOCOLS

Two basic fixation protocols were utilized in this study.

Fixation protocol 1: Insects (ie. lady bugs) were fixed in either 95% ethanol, 100% methanol, 100% isopropyl alcohol or methanol: glacial acetic acid (3:1, vol:vol) and stored at room temperature for varying periods of time prior to DNA isolation. The lady bug were either fixed "intact" (whole) or immersed in fixative and immediately minced (using sterile scissors). The amount of the lady bug tissue added per ml of fixative was varied from about 3.5 mg tissue/ml fixative up to 34 mg tissue/ml fixative.

Fixation protocol 2: Lady bugs were either fixed "intact" or immediately minced in a dilute DTAB solution (2 volumes stock DTAB solution: 1 volume reagent grade deionized water) in microfuge tubes at room temperature.

The DTAB stock solution is 8% DTAB in 1.5 M NaCl, 100 mM Tris Cl, 50 mM EDTA (pH 8.6). The amount of lady bug tissue added per ml of diluted DTAB solution varied from about 18.9 mg/ml to 38.8 mg/ml.

DNA ISOLATION

Stored samples were subjected to different preparative steps prior to DNA extraction because they were preserved using different methods. Specimens fixed "intact" in either methanol, ethanol, isopropanol or methanol: glacial acetic acid (3:1) were removed from the fixative, blotted briefly (to remove excess fixative), and ground in liquid nitrogen to a fine powder in a pre-cooled mortar. The powdered material was then transferred to a microfuge tube containing 900 μ l of diluted DTAB solution (ie. 2 volumes DTAB stock solution : 1 volume DI water), and the tube incubated in a 65°C water bath for 6 minutes. Specimens minced in methanol, ethanol, isopropanol or methanol: glacial acetic acid (3:1) were first centrifuged at about 12,000 x g to sediment the tissue fragments, the supernatants discarded, the pellet resuspended in 900 μ l of diluted DTAB solution (ie. 2 volumes DTAB stock solution : 1 volume DI water), and the tube incubated in a 65°C water bath for 6 minutes. Specimens stored "intact" in diluted DTAB were minced immediately prior to the 65°C incubation step, while

specimens stored as "minced tissue" were incubated at 65°C for 6 minutes without any further processing.

After incubation at 65°C for 6 minutes in diluted DTAB, all samples were centrifuged for 10 seconds at top speed in a microfuge (approx. 12,000 x g) at room temperature to pellet any unsolubilized debris. The pellet was discarded and the supernatant transferred into a new tube (large-bore pipettor tips were used for transfer as this minimizes mechanical shearing of long DNA strands) and mixed with an equal volume of chloroform (typically 900 µl). The sample was then centrifuged for 2 minutes (12,000 x g) at room temperature and the upper aqueous phase was transferred into a new tube. The transferred solution was mixed with 1/9th of a volume of (ie. 100 µl) CTAB stock solution (5% CTAB in 0.4 M NaCl;) and 900 µl of deionized H₂O, and the samples then mixed well by 30-40 gentle inversions to selectively precipitate DNA. The samples were then spun for 5 minutes (12,000 xg) at room temperature to pellet the DNA. The supernatants were removed and discarded, and the internal wall of each centrifuge tube then carefully wiped with a sterile cotton swab to remove as much of the supernatant as possible.

Each DNA/CTAB pellet was next resuspended in 300 µl of 1.2 M NaCl by flicking the tube until the pellet completely dissolved. To reprecipitate the DNA, 2.5 volumes (ie. 750 µ

1) of 100% ethanol was then added to each tube and the contents mixed by 30-40 gentle inversions. The tubes were then centrifuged for 15 minutes at about 12,000 x g at room temperature. The supernatants were poured off and 1 ml of 75% ethanol was added to wash the DNA pellets. After centrifuging for 10 minutes (12,000 x g at room temperature), the supernatants were discarded and excess liquid was removed by wiping the internal wall of each tube with a sterile cotton swab or twisted Kimwipe. The pellets were then dried completely (about one hour) under vacuum and each pellet gently resuspended in 300 µl of TE buffer (10 mM Tris Cl, 0.1 mM EDTA, pH 7.4). A small amount of DNA sample (about 15 µl) was then taken to examine on a 0.8% agarose gel. The whole procedure took less than three hours. This protocol has been used successfully to isolated genomic DNA from a number of different insect species and mouse tissues.

SALT STUDY

In the DNA isolation protocol used throughout this work a salt solution was used to displace the detergents in the DNA precipitate. To examine the effect of salt concentration on the efficiency of DNA extraction, different concentrations of NaCl were used to resuspend the CTAB/DNA precipitate. The concentrations of NaCl solution were 0.5 M, 1.2 M, 1.4 M, 1.6 M, 1.8 M, and 2.0 M. Isolated and

resuspended DNA samples were then subjected to agarose gel electrophoresis and the yield in each case estimated relative to lambda Hind III size/mass standards.

HORIZONTAL AGAROSE GEL ELECTROPHORESIS

Agarose gels (ranging from 1% to 0.5 % wt/vol) were prepared with regular agarose (Sigma, high gelling temp.). The samples were run in 1X TAE buffer (40X TAE buffer is 1.6 M Tris Cl, 0.8 M sodium acetate, 40 mM EDTA, pH 7.2) at 1 volt per centimeter of gel length for 16-18 hours (for "high resolution" gels greater than or equal to 14 cm in length) or about 5 volts/cm gel length for 2.5-3 hours (for 10 cm long mini-gels). HindIII digested lambda DNA fragments were typically used as size ("molecular weight") standards. Before loading, these lambda fragments were incubated in a 65°C water bath for 5 minutes and then immediately plunged into ice and kept there until loading. The gels were then stained in 0.5 µg/ml ethidium bromide solution for at least one hour (overnight staining was common) and examined using an ultraviolet (UV) transilluminator.

EXAMINATION OF QUALITIES OF ISOLATED DNA

Evaluation of the size of the isolated genomic DNA.

Estimations of the size of DNA fragments were carried out by agarose gel electrophoresis, as described above.

Assay for the presence of residual DNase activity.

Aliquots of extracted genomic DNA were incubated in low salt TE buffer at 37°C for an hour. The samples were then subjected to agarose gel electrophoresis and evaluated for evidence of DNase activity (ie. conversion of HMW DNA to smaller sized fragments).

Suitability of the DNA preps for restriction enzyme digestion. Isolated lady bug DNA (typically 0.5-2 µg) was incubated at 37°C for one hour in the presence or absence of restriction endonuclease (typically 2-10 units of Pst1 or Sau3A) and the appropriate reaction buffer. Restrictions were terminated at the end of the incubation period by immediately placing the tubes on ice and adding 1/8th of a volume of 0.1 M EDTA stock solution. The DNA samples were then examined on 0.8% agarose gels to check for DNA degradation.

DNA PARTIAL DIGESTION

Restriction titration assay. A stock solution of Sau3A restriction enzyme (10 units/ μ l in a stablization buffer containing 50% glycerol, Promega Corp.) was used to prepare an enzyme dilution series using 1X Sau3A restriction enzyme buffer as dilutent (10x RE buffer: 100 mM TrisCl pH 7.5, 100 mM $MgCl_2$, 10 mM DTT, 1 mg/ml BSA or gelatin). All enzyme dilutions were prepared on ice immediately prior to use. An equal amount of lady bug genomic DNA then was added into microfuge tubes (typically 0.8 μ g DNA per tube) and these assay tubes were placed on ice. Each assay tube then received an equal volume of a dilute enzyme solution (or an equal volume of 1X RE buffer for the undigested control), the contents of each tube were mixed well, and all tubes briefly spun for 3-5 seconds to concentrate materials to the bottom of each tube. The total reaction volume in all tubes was 20 μ l/tube. All tubes were then incubated in a 37°C water bath for 30 minutes. The reactions were stopped by the addition of 0.1 M EDTA (as described above) and immediately plunged into ice. After the addition of 1/9th a volume of 10X DNA sample loading buffer (20% Ficoll 400, 0.1 Na_2EDTA , 1.0% SDS, 0.25% (w/v) bromophenol blue and 0.25% xylene cyanol) to each tube, the samples were then examined by agarose gel electrophoresis as described previously.

Large-scale partial digestion. The optimal ratio of "μg of DNA:units of restriction enzyme" was determined from the titration assay described above. Appropriate amounts of Sau3A enzyme and genomic DNA (typically about 50 μg of either mouse DNA or lady bug DNA) and RE buffer were then added to microfuge tubes, the contents of each tube carefully mixed, and the tubes incubated in a 37°C water bath for 30 minutes. The reactions were stopped using 0.1 M EDTA as described above. A small aliquot of each sample was then examined by agarose gel electrophoresis to determine the size distribution of the digestion products. Once the appropriate degree of digestion had been confirmed, an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to simultaneously destroy residual endonuclease activity and strip proteins from the aqueous phase. Each sample was then centrifuged at top speed (approx. 12,000 x g) for 10 minutes at room temperature to help break up the phases, the organic (lower) phase discarded, and the aqueous (upper) phase transferred to a new tube. This extraction step was repeated once more. The aqueous phase was then mixed with an equal volume of chloroform:isoamyl alcohol (24:1) to extract residual phenol. After centrifuging at 12,000 x g for 10 minutes at room temperature, the upper aqueous phase was transferred into another new tube. One half volume of freshly made 7.5 M ammonium acetate was then

added to the recovered aqueous phase, followed by two volumes of 100% ethanol to precipitate the DNA. The samples were then held at -80°C overnight and spun for 20 minutes at room temperature the next morning. The pelleted DNA was then rinsed with 70% ethanol, dried under vacuum about one hour, and finally resuspended in TE buffer. The concentrations of the DNA samples were subsequently determined by spectrophotometry.

PARTIAL FILL-IN REACTION

A partial fill-in reaction was performed on the Sau3A digested genomic DNA fragments in order to prevent the random ligation of these DNA fragments to each other during subsequent steps. The reaction tubes were prepared as follows:

Partially digested genomic DNA: 20 μ l*

Fill-in 10X Reaction Buffer(with dATP and dGTP): 5 μ l**

Nuclease-free water (DEPC-treated): 23 μ l

Klenow DNA Pol. enz. (5 units/ μ l; added last): 2 μ l

* A total of 6.2 μ g mouse DNA or 10.6 μ g lady bug DNA was used in the reaction.

** The 10X reaction buffer is 60mM Tris, pH7.4, 60 mM NaCl, 60 mM MgCl₂, 0.5% gelatin and 10mM dithiothreitol. The final concentrations of the dATP and dGTP are 0.166 mM.

The contents of the tubes were carefully mixed, the tubes spun for 3-5 seconds to concentrate all materials to the bottom of the tubes, and the tubes incubated at 37°C for 45 minutes. After incubation, the phenol:chloroform:isoamyl alcohol extraction procedure was done as described above in the partial digestion step. Finally, the pellets were resuspended in 20 µl of sterile water. The concentrations of DNAs were determined using spectrophotometry on a diluted aliquot.

LIGATION AND *IN VITRO* PACKAGING REACTIONS

Partially filled-in DNA fragments (1.0 µg and 0.5 µg for mouse DNA and 1.05 µg for lady bug DNA) and 2 µl of positive control inserts (16 Kb, 0.5 µg/µl available from Promega Corp.) were mixed with 2 µl of lambda-GEM12 arms (0.5 µg/µl, Promega Corp.) and 1 µl of T4 ligase (10-15 Weiss units/µl, Promega, Corp.) to a total volume of 10 µl. After incubation at 4°C overnight, 5 µl of the ligation products were mixed with 25 µl of packagene extract (Promega, Corp.). The tubes were gently tapped several times and incubated at room temperature (22°C) for 3 hours. Following this incubation, 223 µl of phage buffer (20 mM Tris-Cl, pH 7.4, 100 mM NaCl, 10 mM MgSO₄) and 12.5 µl of chloroform were added into the tubes. The tubes were gently

mixed and the chloroform was allowed to settle down to the bottom of the tubes. (For lady bug DNA, all materials used in the *in vitro* packaging reaction were doubled.) The samples were then stored at 4°C until used in the phage titration (plaque) assay.

INTRODUCTION OF RECOMBINANT PHAGE INTO *E. COLI* CELLS

A single colony of *E. coli* LE 392 cells [genotype is F^- , *hsdR154*(r_k^- , m_k^-), *supE44*, *supF58*, *lacY1* or (*lacIZY*)₆, *galK2*, *galT22*, *metB1*, *trpR55*] were inoculated into 10 ml of lambda broth (10 g tryptone, 2.5 g NaCl in 1.0 l), supplemented with 0.2% maltose and 10 mM MgSO₄ in a 50 ml flask and grown at 37°C with shaking (200-250 rpm) overnight.

Ten-fold serial dilutions of recombinant phage stocks were prepared in suspension medium (SM: 5.8 g NaCl, 2 g MgSO₄·7H₂O, 50ml TrisCl pH 7.5 and 5ml 2% gelatin in 1.0 l). In each tube, 0.1 ml of diluted phage and 0.1 ml of prepared *E. coli* cells were mixed and incubated in a 37°C water bath for 30 minutes (without shaking) to allow the phage particles to absorb. Three ml of top agar (10 g tryptone, 2.5 g NaCl, 7 g agar in 1.0 l 45°C) were then added to the mixture in each tube, vortexed gently, and immediately poured onto 37°C prewarmed plates containing hardened bottom

agar. After incubation at 37°C overnight, the numbers of plaques in each plate were counted and the plate with the most plaques was used to harvest the phages. Five ml of SM were added onto the plate, and the plate then stored at 4°C for 3 hours with intermitten, gentle shaking 5 to 6 times. After the incubation, as much as possible of SM was transferred into a 10 ml centrifuge tube using a pasteur pipette. Another 1 ml of fresh SM was added into the plate which was then stored for 15 minutes in a tilted position to allow all the fluid to drain into one area. The SM was then again removed and combined with the first harvest. After 0.1 ml of chloroform was added into the tube, the poured SM was centrifuged at 4000 x g for 10 minutes at 4°C. The supernatant was recovered and stored at 4°C after 0.3% of chloroform was added into the harvested phages (Maniatis, et al., 1982).

SPECTROFLUORIMETRY

Standard DNA solutions were prepared in TEN buffer (10 mM Tris-Cl, pH 7.4, 1mM EDTA, 100 mM NaCl) containing 0.1 μ g/ml Hoechst 33258 dye, referred to as Hoechst "working solution". The DNA concentrations were typically prepared in this "working solution" at 800 ng/ml, 400 ng/ml, 200 ng/ml, 100 ng/ml and 50 ng/ml. Aliquouts of genomic DNA (typically 2-5 μ l) were also diluted in the Hoechst "working solution". The mixtures were allowed to stand in the dark for about two hours, after which time the emission intensity of the solutions (standards and unknowns) was determined using a spectrofluorimeter. The genomic DNA concentrations were read off the standard curve.

Chapter III

Results

Development of a suitable "tissue fixative/DNA stabilization" solution.

Our initial objective was to identify (or develop) a suitable fixation procedure, one which could preserve HMW DNA in insect tissues fixed under field conditions. Several different procedures were tested in this study. After fixation, all tissues were subjected to the "DTAB/CTAB" DNA extraction protocol and the isolated DNA then analyzed by agarose gel electrophoresis. Gels containing the fractionated nucleic acids were then stained in ethidium bromide solution and examined via UV transillumination.

Neither 100% methanol, 95% ethanol, 100% isopropanol, methanol:glacial acetic acid (3:1) nor 5% citric acid : DTAB (2:1) preserved high-molecular-weight DNA.

In any entomological archival operation it is important to preserve as much of the insect's original morphology as possible. We thus first sought to fix intact specimens by simply immersing them (while still alive) in various fixative solutions. Whole lady bugs were thus immersed

intact in either 100% methanol, 95% ethanol, 100% isopropanol or a solution of methanol:glacial acetic acid (3:1), stored at room temperature for 21 days, and then subjected to the "DTAB/CTAB" DNA isolation procedure and analysis by agarose gel electrophoresis (AgGE). Unfortunately, while this procedure was very convenient it produced only minimal amounts of DNA. As shown in Figure 3, Panel A, the DNA which was extractable was of low molecular weight, as indicated by the diffuse, broad bands in the lower portions of the gel. Even if the yields could be increased, the DNA is of such small size (either as a result of shearing, nuclease attack, and/or chemical attack) that it would be useless for the construction of genomic libraries. As shown in Figure 3, Panel B, use of a mixture of citric acid and the DTAB detergent solutions was even less successful. No DNA could be recovered following the use of this fixative, and further work with citric acid/DTAB mixtures was abandoned.

I next sought to increase the access of the fixatives to the insect tissues by first crudely homogenizing (ie. mincing) the specimens immediately upon immersion. Live lady bugs were thus immersed and immediately minced in either 100% methanol, 95% ethanol, or 100% isopropanol. The minced tissues were then stored for one week at room temperature prior to DNA extraction via the "DTAB/CTAB"

procedure and analysis by AgGE. As shown in Figure 4, Panel A, this modification did produce some degree of success. Total yields of DNA increased from tissues fixed in either methanol, ethanol or isopropanol. More importantly, faint bands of HMW DNA were visible in the isopropanol-fixed tissues. Unfortunately, as is also evident from the gel, most of the DNA was still degraded and/or sheared. After four weeks of storage essentially all the HMW DNA was degraded (Figure 4, Panel B) and further work with these fixatives was abandoned.

DTAB solution was the only one found capable of preserving HMW DNA , and then only when tissue was homogenized.

Realizing that the DTAB-based DNA extraction solution employed was stable to lengthy incubation at 65°C, the possibility of using this as an initial "fixative/storage" reagent was examined. Trying the easiest procedure first, intact (whole) lady bugs were simply immersed in the DTAB detergent solution, stored for one week at room temperature, and then used for HMW DNA isolation. Despite the impervious nature of the insect exoskeleton, at least some of the detergent must have gained access to the insect body cavity, since DNA in the 10 Kbp to 40 Kbp range was observed (as shown in Figure 5, Panel B, lane 4). There was little HMW DNA (> 50 Kbp), however, which forced the use of the mincing

("homogenization") protocol originally employed for the alcohol-based fixatives.

Equal numbers of live lady bugs were thus next immersed and immediately minced in separate vials containing DTAB detergent solution, the samples stored for varying periods of time (ranging from a few minutes to 1.5 years), and the nucleic acids then isolated by carrying out the remainder of the DTAB/CTAB DNA extraction protocol. As expected, the majority of DNA isolated from freshly minced samples stored in DTAB for only 5-60 minutes (Figure 5, Panel A, lane 1) was high molecular weight material. More importantly, HMW DNA was also the predominant isolatable material in samples stored in DTAB for one, two, three, seven, nine, and 75 weeks (as shown in Figure 5.).

In summary, among all the procedures tested, the protocol in which homogenized lady bug tissue was stored in DTAB solution represents the only method capable of preserving HMW DNA for longer than eight weeks (ie., sample holding and shipping periods likely to be encountered during collecting expeditions carried out deep within tropical rain forests).

Investigation of the NaCl concentration required for optimal detergent displacement and maximal DNA yield.

In the DTAB/CTAB DNA isolation protocol, a chloroform extraction step strips away contaminating proteins (and much of the DTAB as well) and leaves the soluble DNA in the aqueous (upper) phase. After addition of the CTAB detergent solution precipitates the DNA, the DNA is spun down and the supernatant (containing the unbound detergent) discarded. A concentrated solution of NaCl is then added to displace the bound detergent (probably via simple ion exchange) and the DNA resuspended. To determine the salt concentration that optimizes this detergent displacement, NaCl solutions at 0.5 M, 1.2 M, 1.4 M, 1.6 M, 1.8 M, and 2.0 M were used in the exchange step and the effect on DNA yield analyzed by AgGE. As shown in Figure 6, HMW DNA bands were observed in each sample. Furthermore, as judged by both the brightness and the total area of the HMW bands, there was no significant difference between the various salt concentrations tested. For all ensuing DNA extraction experiments, 1.2 M NaCl solution was used in the exchange step and no further optimization was pursued.

Examination of the quality of the isolated genomic DNA.

In order to be used for constructing a genomic library, isolated DNA must meet at least three basic criteria: (i) it should be of high molecular weight (verified in the preceding figures), (ii) it should be free of residual DNase contamination (DNase may not be active in the DTAB fixative/storage solution, but might be carried through the remainder of the DNA isolation protocol and reexpress it's activity after removal of the detergent), and (iii) it should serve as a good substrate for digestion by type II restriction enzymes. Experiments were thus performed to determine if the latter two criteria were met.

Isolated genomic DNA was free of DNase contamination.

Equal amounts of lady bug genomic DNA were incubated in low salt buffer (without EDTA) at 4°C and 37°C for one hour, and then examined on a 0.8% agarose gel. The presence of DNase should be reflected in a reduction in band staining intensity of the DNA incubated at 37°C relative to the intensity of the DNA incubated at 4°C. If enough DNase was present, a brightly staining low molecular weight smear should appear in the 37°C lane. No DNase activity was detected in the genomic DNA samples (see Figure 7, Panel A).

High molecular weight DNA isolated by the DTAB/CTAB protocol served as a good substrate for restriction enzymes.

Restriction digests were performed by mixing equal masses of isolated genomic DNA with a single dilution of RE, incubating all reaction tubes at 37°C for a fixed time period, stopping further reaction in all tubes simultaneously (by the addition of 1/40th a volume of 0.5M EDTA "stop solution"), and then analyzing the size of the resulting DNA fragments via AgGE. For molecular weight determination, DNA samples were examined on a 0.8% agarose gel, using HindIII restriction fragments of lambda DNA as molecular weight (ie. size) standards. Representative results using the RE PstI are shown in Figure 7, Panel B. Similar results were obtained with other REs, confirming that the DNA was sensitive to a broad range of type II enzymes (data not shown).

DNA manipulation and library construction.

Confident that the DTAB "tissue fixation/storage" reagent and the DTAB/CTAB nucleic acid isolation procedure would provide restrictable HMW DNA, both mouse and lady bug genomic DNA were isolated for construction of two libraries and to determine the RE concentration required to produce genomic fragments of the optimal size range for ligation to

lambda arms and packaging into lambda phage heads. Partial digestion remains the most common method for producing the family of "overlapping, appropriately sized" DNA fragments required in the construction of genomic libraries. Since recombinant lambda phage genomes can only be packaged into a phage head when the genome size is 38 Kbp-52 Kbp in length (wild-type DNA has 48.5 Kbp), control of the size range of the insert DNA fragments generated during a genomic digest is critical. Exercising this control typically requires an empirical determination of the "optimal" ratio of RE units to μg of target DNA. Since the lambda vector system uses right and left arms that total about 30 Kbp in length, the Sau3A enzyme concentration (ie. units of activity) required to produce a family of RE fragments enriched in the 10-20 Kbp range was determined.

Partial digestion of isolated genomic DNA: Determination of optimal digestion conditions.

Restriction digests were performed by mixing equal masses of isolated lady bug DNA (the concentration of DNA was determined by spectrofluorimetry as described in the "Methods" section) with different dilutions of RE (see Table 1 and 2, all dilutions of RE were made in 1X RE buffer immediately before use in the assay). All reaction tubes were then incubated at 37°C for a fixed time period (30

minutes), with further reaction stopped simultaneously in all tubes by the addition of 1/40th of a volume of 0.50 M EDTA "stop solution". The size range of the resulting DNA fragments was then determined via AgGE. For molecular weight determination, DNA samples were examined on either 0.6% or 0.8% agarose gels, using uncut lambda genomic DNA and/or HindIII-restricted lambda DNA as size standards. As is shown in Figure 8, both the lady bug (Panel A) and mouse (Panel B) HMW DNAs serve as excellent substrates for the RE Sau3A. The optimal ratio of "Sau3A units:µg of DNA" is given in each panel in Figure 8.

Table 1. Restriction titration for mouse DNA

Tube#	Vol. of DNA* (μl)	Sau3A dilution	Vol. of Sau3A (μl)	Units of Sau3A (U)
1	10	1/8	1.5	0.75
2	10	1/16	1.5	0.375
3	10	1/32	1.5	0.188
4	10	1/64	1.5	0.094
5	10	1/128	1.5	0.047
6	10	1/256	1.5	0.023
7	10	1/512	1.5	0.012
8	10	1/1024	1.5	0.006

*[mouse DNA] = 0.0795 μg/μl, [Sau3A] = 4 U/μl

Table 2. Restriction titration for lady bug DNA

Tube#	Vol. of DNA* (μl)	Sau3A Dilution	Vol. of Sau3A (μl)	Units of Sau3A (U)
1	2.0	1/40	2.0	0.5
2	2.0	1/80	2.0	0.25
3	2.0	1/160	2.0	0.125
4	2.0	1/320	2.0	0.063
5	2.0	1/640	2.0	0.031
6	2.0	1/1280	2.0	0.016
7	2.0	1/2560	2.0	0.008

*[Lady bug DNA] = 0.793 μg/μl, [Sau3A] = 10 U/μl

Large-scale (preparative) digestion of genomic DNA.

Once the optimal conditions for the partial digestion had been defined, scaled-up reactions were carried out using lady bug (Figure 8, Panel A) or mouse (Figure 8, Panel B) genomic DNA in which the incubation conditions (37°C, 30 min incubation, 0.5 M EDTA as stop solution) were identical to those described above. After these large-scale partial digestions, a small aliquot of the digested material was examined on 0.8% agarose gels. As expected, the results indicated that the maximum staining intensity was in the range of 9-20 Kbp (Figure 9). After digestion the DNA fragments were purified by standard phenol:chloroform extraction and ethanol precipitation, and their concentrations determined by spectrofluorimetry. Final yields of lady bug and mouse DNA fragments after purification were 12.4 µg (in a total volume of 40 µl) and 53 µg (in a total volume of 100 µl), respectively.

Partial fill-in reaction

A partial fill-in reaction was used to prevent random association between fragments of digested genomic DNA. The RE *Sau3A* recognizes the four-nucleotide sequence 5'-GATC-3', producing complementary 5'-GATC overhanging ends upon digestion. Unfortunately, the genomic fragments can easily

become ligated together randomly (an unwanted reaction) as they can join to vector arms. This unwanted ligation will not only affect the efficiency of the packaging reaction but also will interrupt the order of the original sequence and introduce artifactual sequence stretches into the genomic library. Partial fill-in reactions were thus carried out using the Klenow fragment of *E. coli* polymerase I and selected dNTPs. In this study, reaction buffer containing only dATP and dGTP was used to convert the proximal part of each single stranded end to double stranded form. Each four base-long single stranded complementary end will thus be converted to a two base-long single stranded non-complementary end via this reaction, preventing genomic fragments from ligating to each other. Similar treatment of the lambda left and right arms ensures that only the arms and inserts will be ligatable.

A total of 6.2 µg mouse DNA and 10.6 µg of lady bug DNA were used for the partial fill-in reaction. After the reaction the DNA was recovered (once again by phenol:chloroform extraction and ethanol precipitation) and resuspended in 20 µl of sterile water. The final DNA concentrations (determined by spectrophotometry) were 0.22 mg/ml for mouse DNA and 0.263 mg/ml for lady bug DNA. The percentages of recovery were 71% for mouse DNA and 50% for lady bug DNA.

Ligation reaction and in vitro packaging reaction

The ligation reaction was performed to insert the partially back-filled DNA fragments into lambda GEM-12 arms to produce long concatemers which could be packaged into phage heads. After the partial fill-in reaction the DNA fragments had only two bases left at each overhanging end, a condition which forced a modification of the usual ligation reaction conditions. Given the fact that each complementary end could remain associated via hydrogen bonding across only two bases, even temperatures above 5 to 6°C would be expected to impart so much thermal energy (and thus molecular motion) that stable association between ends would be very unlikely. The ligation reaction temperature, therefore, was lowered to 4°C and the duration of ligation extended. In this reaction, 1.0 µg of GEM-12 arms was mixed with either 1.0 or 0.50 µg of partially digested and back-filled mouse DNA, or 1.05 µg partially digested and back-filled lady bug DNA or 1.0 µg partially digested and back-filled positive control DNA (supplied by Promega Corp.) and 10 Weiss units of T4 DNA ligase in a total volume of 10 µl (see Table 3).

Table 3. Ligation reaction

Tapes of DNA in ligation reaction	Background control	Positive control	Mouse #1	Mouse #2	Lady bug
GEM-12 λ Arms (μ g)	1.0	1.0	1.0	1.0	1.0
Positive Control DNA (μ g)	-	1.0	-	-	-
Mouse Genomic DNA (μ g)	-	-	1.0	0.5	-
Lady Bud Genomic DNA (μ g)	-	-	-	-	1.05
10X Ligase buffer (μ l)	1.0	1.0	1.0	1.0	1.0
Distilled H ₂ O (μ l)	6.0	4.0	2.6	4.3	2.0
T4 ligase (μ l)	1.0	1.0	1.0	1.0	1.0

[GEM-12 λ arms]=0.5 μ g/ μ l
 [Positive Control DNA]= 0.5 μ g/ μ l
 [T4 ligase]=10-15 Weiss units/ μ l

After incubation overnight at 4°C, the three ligation mixtures were added to separate tubes of Promega packaging ("Packagene") extract, incubated at room temperature for three hours, mixed gently with phage storage buffer and chloroform and stored at 4°C until used in the phage titration (plaque) assay, as described in Chapter 2.

For the plaque assay, dilutions of the packaged phage suspension were first prepared, and then 100 μ l of each phage dilution was mixed with 100 μ l of prepared LE 392 bacterial cells. Following adsorption, the cells were mixed with top agar and poured onto LB plates as described in "Methods". Following overnight incubation at 37°C, the viral plaques were counted and the titer determined. As shown in Table 4, no plaques were visible on the background "negative control" plates (ie. using phage packaged without insert). The titer of the positive control was 3.5×10^6 plaque-forming units per ml (pfu/ml), while the total number of plaques obtained in the positive control packaging reaction was 8.86×10^5 pfu. Given that a total of 1.0 μ g of DNA was present in the packaging reaction (0.5 μ g of vector arms and 0.5 μ g of insert DNA) the calculated packaging efficiency is 8.86×10^5 pfu/ μ g input DNA. Since only half the amount of ligation mixture was packaged, a similar number of plaques would be expected to be produced in a second packaging reaction, thus yielding a grand total of 1.77×10^6 pfu.

A similar experimental protocol was then employed to ligate and package either lady bug or mouse restricted and partially backfilled genomic DNA fragments. The results of these experiments are also shown in Table 4. For lady bug DNA a total of 6.77×10^4 plaques were obtained per

packaging reaction (viral titer of 1.34×10^5 pfu/ml, packaging efficiency of 3.30×10^4 pfu/ug input DNA). By comparison, a representative packaging reaction using mouse DNA produced a total of about 3.06×10^4 pfu/packaging reaction (viral titer of 1.21×10^5 pfu/ml, packaging efficiency of 3.06×10^4 pfu/ug input DNA).

Table 4. Results of ligation and packaging reaction

Materials Packaged	µg λ GEM- 12 added during the packaging	µg insert DNA added during the packaging	Ratio of λ GEM-12 to insert DNA	Total vol. of diluted phage after packaging (ml)	Total pfu generated in packaging reaction	Viral titer (pfu/ml) of diluted phage after packaging	Package efficiency pfu/µg total input DNA
Background control	0.5	0	-	0.253	0	0	-
Positive control	0.5	0.5	1:1	0.253	8.86×10^5	3.50×10^6	8.86×10^5
Mouse DNA	0.5	0.25	2:1	0.253	1.85×10^4	7.30×10^4	2.47×10^4
Mouse DNA	0.5	0.5	1:1	0.253	3.06×10^4	1.21×10^5	3.06×10^4
Lady bug DNA	1	1.05	1:1.05	0.505	6.77×10^4	1.34×10^5	3.30×10^4

Chapter IV

Discussion

It is currently believed that as a result of continued habitat destruction many insect and plant species in rain forest environments have been driven to extinction. How to save those species that remain is considered an urgent problem by many biologists. While habitat protection represents the only real solution, the social, economic and political pressures operating against that effort appear unlikely to subside. Concerned scientists, however, can at least employ new cloning technologies to preserve the genomic information contained in these threatened species. With this goal of genomic preservation in mind, a model genomic library for insects was constructed using new DNA fixation/stabilizing reagents. This section discusses the capabilities and limitations of this protocol and offers some recommendations for future study.

DNA preparation

Attempts at fixation of DNA in whole insects

HMW DNA is required for construction of a genomic library, and it seemed that one way of maintaining insect

DNA as HMW material was to "fix" the insect tissue immediately upon collection. Immersing intact (ie. whole) insects in fixative seemed a reasonable means of both preserving the specimen for photography and morphological studies while simultaneously denaturing nucleases with minimal production of shear forces to the DNA. The idea was that once the specimen was fixed, the fixative could be removed at a later time, the insect transferred to and minced/homogenized in diluted DTAB solution, and the DNA isolated using the DTAB/CTAB procedure. A number of different fixatives were evaluated, including: (i) 100% methanol, (ii) 95% ethanol, (iii) 100% isopropanol, and (iv) Carnoy's fixative (3 volumes methanol: 1 volume glacial acetic acid). Unfortunately, none of these fixatives allowed the isolation of HMW DNA.

An explanation for this result remained elusive, however, since it was uncertain whether the inability to isolate HMW DNA from fixed intact specimens resulted from: (i) poor initial fixation due to poor penetration through the insect cuticle (this would allow nuclease attack on the DNA), (ii) an intrinsically poor capacity of these fixatives to rapidly inactivate all nuclease activity in insect tissue even after penetration through the cuticle, (iii) some consequence of protein precipitation around the DNA that made it particularly susceptible to shear forces during the

subsequent isolation, or (iv) direct chemical attack on the DNA (in the case of Carnoy's fixative, an acid hydrolysis). In order to investigate some of these possibilities, a series of experiments in which lady bugs were first minced/homogenized directly in fixative and then subjected to the DTAB/CTAB DNA isolation procedure were performed. Abandoning the "requirement" for preserving intact specimens also meant that a wider variety of DNA stabilization reagents... specifically, a detergent that might both rapidly penetrate and solubilize cells and simultaneously inactivate nucleases such as DTAB itself could be tried.

Attempts at isolation of HMW DNA from tissue homogenized and stabilized ("fixed") simultaneously.

We evaluated the ability of 100% methanol, 100% ethanol, 100% isopropanol, Carnoy's fixative, acidic DTAB (1 volume 5% citric acid : 2 volumes DTAB stock solution) or diluted DTAB (2 volumes DTAB stock solution : 1 volume water) to stabilize DNA in minced/homogenized tissue and allow it's subsequent isolation as high molecular weight material. Somewhat surprisingly, neither methanol, ethanol, isopropanol, Carnoy's fixative nor acidic DTAB allowed the isolation of HMW DNA. A recent study by Reiss et al. (1995)

also found that Carnoy's fixative was unable to preserve HMW DNA in insects.

The inability of acidic fixatives (ie. Carnoy's fixative or acidic DTAB) to preserve HMW DNA was not unexpected. Mild acid hydrolysis of DNA probably does occur with Carnoy's solution, but is not a problem in the histological applications in which this fixative is routinely used. Adding higher penetrative power to an acid fixative by including a detergent (ie. acidic DTAB) also appeared to be of no benefit, and suggests that penetration, per se, is not the problem. Carnoy's fixative was not added to a highly concentrated solution of purified DNA in order to further determine if acid hydrolysis occurs, but this remains an option.

None of the series of alcohols, was very effective at preserving HMW DNA for reasonable periods of time (ie., more than one week). In contrast, Reiss et al. (1995) found that homogenization of beetles in ethanol provided relatively good yields of HMW DNA up to 73 days of storage. The DNA was progressively degraded beyond 73 days in the ethanol using their protocol. There may be several reasons for obtaining such different DNA yields using two apparently similar fixation methods, but probably the two most important experimental parameters to examine are: (i) differences in the ratio of ethanol (fixative volume) to

insect tissue mass employed and/or (ii) species differences in the amount of nuclease activity per mg tissue. Both possibilities could be investigated in the future.

Since alcohol-based fixation did not adequately preserve HMW DNA, the solubilizing and denaturing detergent, DTAB was tried. As shown in Figure 5, the best procedure found for the preservation of HMW DNA from insects involves a simple homogenization/mincing step in a diluted DTAB solution at ambient temperature (ie., from about 20-45° C). This "stabilized" material can then be subjected immediately to the DTAB/CTAB-based DNA isolation protocol or may be placed under storage at ambient temperature for periods of at least one year. The stabilization procedure is easy, quick, requires essentially no equipment, employs no hazardous chemicals and provides material that can be stored for long periods of time even under field conditions.

DNA isolation using the DTAB/CTAB protocol : properties of the isolated DNA and advantages of the method.

DTAB is a positively charged denaturing detergent highly effective at inhibiting enzyme activity (Lo Piero & Petrone, 1994). In a high ionic strength buffer (1.5 M NaCl, 100 mM Tris) at elevated temperature (ie. 55°C) DTAB provides sufficient solubilization and denaturation power to both lyse cells and inactivate nucleases, thus obviating the

need for the prolonged Proteinase K digestion step found in many other DNA purification protocols. A 12-24 hour Proteinase K digestion (Adelik & Ogboma, 1990) at elevated temperature would almost certainly be impractical under the field conditions. Protein precipitation and removal of organic compounds (esp. DTAB) were then simultaneously effected by a single chloroform extraction. The advantage here is that DNA is never exposed to attack by phenol, and the use of a single extraction step decreases the loss of material that unavoidably occurs during the repetitive "extraction, centrifugation, and phase isolation" steps typical of most other protocols. The DNA is next diluted and selectively precipitated from the aqueous phase at lower ionic strength (≤ 0.4 M NaCl) in the presence of CTAB. Under these conditions, CTAB forms an insoluble salt with DNA while other substances remain in solution (Jones, 1963). Finally, the DNA is resolubilized in high salt, and reprecipitated in ethanol, and then washed in 70% ethanol to remove the salt.

The DNA obtained is typically 100 Kbp or longer can be cut with a variety of restriction enzymes, is essentially deproteinated (the OD_{260}/OD_{280} ratio obtained on this material is typically in the 1.8-2.0 range) and, more specifically, is devoid of detectable DNase activity (ie. no reduction in

the size of the DNA was observed even upon prolonged incubation at 37°C).

The presence of a low molecular weight "smear" in the 500 bp range in many of the preparations suggests that there is some amount of RNA contamination within the extracted and "purified" DNA. Since an RNase treatment was not employed in our procedure this relatively small degree of contamination was expected. Subjecting the DNA to an RNase digestion was initially considered followed by a phenol/choloform extraction series and further rounds of ethanol precipitation and washing; however, under conditions of very limited starting material the unavoidable loss of DNA incurred during those extra steps might be unacceptable. It was decided to utilize our DTAB/CTAB "purified" DNA (without supplementary RNasing and purification steps) directly in construction of genomic librabries.

Restriction, ligation, packaging and titering of recombinant phage.

In this study, partial digestions of the genomic DNA were carried out using varying amounts of RE Sau3A and the size distribution of the reaction products analyzed by agarose gel electrophoresis. One-half amount of Sau3A which gave the greatest staining intensity in the 9 to 20 Kb range on the gel was then used during the large-scale partial

digestion. These genomic fragments were then subjected to a Klenow partial "backfill" reaction (to prevent self-ligation), ligated into commercially prepared lambda arms, and then packaged *in vitro* and titered on an appropriate host strain (*E. coli* LE392). As shown in Table 4, a total of 6.77×10^4 pfu/packaging reaction was obtained in the lady bug genomic library, and 1.85 to 3.06×10^4 pfu/packaging reaction in the mouse library.

Clearly, the number of plaques obtained in the mouse library indicated that the degree of capture was very inefficient. This conclusion is based upon the Clark and Carbon equation (Clark & Carbon, 1976) as shown below (see Appendix for worked examples from the data).

Let N = the number of independent clones required to achieve a probability, p , that the library is completely representative.

$$N = \ln(1-p) / \ln[1 - (\text{average insert size} / \text{genome size})]$$
Starting, for example, from known estimates of the size of the mouse genome (ie. about 2.3×10^9 bp) and assuming an average insert size of 2×10^4 bp, one would typically need a total of about 3.44×10^5 pfu to obtain a 95% chance of having a completely represented library (ie. the probability, p , is 0.95 that any given, randomly chosen sequence within the mouse genome will be represented in the library). A considerably higher number of recombinants (ie.

plaques) is required to achieve a 99% probability (here N , the number of independent clones, must be about 5.29×10^5). Using $N = 3.06 \times 10^4$ as the total number of independent recombinants obtained following the packaging of mouse genomic fragments, then the "p" can be solved by the equation given above. Using this approach, there was only a 23.4% chance (probability= $p=0.234$) of finding all mouse genomic sequences represented in the library after one packaging reaction. If the efficiency of cloning lady bug DNA is similar to that for mouse DNA, this suggests that a similarly low probability of having a completely representative lady bug library *after only one packaging reaction*. Increasing the number of packaging reactions will, of course, increase the total number of pfu (combining the two separate mouse DNA packaging experiments, for example, yields a total of 4.91×10^4 pfu), which will then increase the percentage of the genome captured. In this case, 4.91×10^4 pfu would represent about 35% of the mouse genome. However, while this reiterative packaging approach will yield a complete library, it does not solve the fundamental problem of inefficient capture. There might be many reasons for this low capture efficiency, but some of the most important ones are examined in the following section.

Parameters affecting efficiency of the ligation and packaging reactions.

In order to show that the packaging extracts were functioning properly, a "positive packaging control" was run (by Dr. Steven White) by adding intact wild-type lambda cI857 DNA to the packaging extract and then titrating out the phage on the *E. coli* LE392 strain. The total number of plaques obtained (1.50×10^9 pfu/packaging reaction at 1 μ g input DNA/packaging reaction) indicated that the packaging extract was indeed quite functional (Dr. S. White personal communication). A background control study was performed by mixing lambda arms with T4 ligase in the absence of insert DNA fragments. After packaging the phage was allowed to infect *E. coli* LE392 and plated. Not a single plaque was visualized on this "background control" plate, suggesting either that the central stuffer fragment had been completely removed from the right and left arms by Promega Corporation (our source of vector) or the ends of the vector had been so efficiently modified that religation to the central stuffer was essentially impossible.

In this "positive packaging control" situation, however, there is no need to subject the cI857 wild-type lambda DNA to ligation; it is already full-length and supplied in the form of long multimeric concatemers as an

ideal packaging substrate. One could not obtain this packaging efficiency using insert DNA that had to undergo ligation with free right and left lambda arms.

In order to show that the ligation and packaging reactions were operating at "reasonable" efficiency, a 16 Kbp linearized, partial backfilled plasmid DNA (again supplied by Promega Corp.) was run as a "positive control insert" in both our ligation and packaging protocols. When a ligation mixture containing 0.5 µg of positive control DNA and 0.5 µg of lambda arm DNA was packaged and titered on LE392, it yielded a total of 8.86×10^5 pfu. Interestingly, the total number of plaques obtained using the "positive control insert DNA" was considerably less than that obtained when wild-type Lambda cI857 DNA was packaged. This suggested that the ligation and/or packaging reactions may not work as efficiently as expected when anything other than full length wild-type concatemeric lambda DNA is used. The genomic DNA fragments from previous experiments would probably function less efficiently than Promega's own positive control plasmid DNA insert. There are probably many reasons for this apparent inefficiency, but some of the more important ones are discussed below.

(i) *DNA (esp. genomic fragments) might have become degraded and/or sheared in steps prior to ligation, yielding fragments with undefined (and unligatable) ends.* Recently, it was found that just two rounds of freeze-thawing of genomic DNA produced an unacceptable degree of shearing (Mr. Tim Andriese, personal communication). Equally important was the observation that even ten single passes of HMW genomic DNA through a "P200 style" yellow pipet tip (as might occur during resuspension or mixing) was sufficient to produce an unacceptable degree of shearing. It is possible, therefore, that at least some of the 9-23 Kbp Sau3A restriction fragments might have become sheared during storage and/or manipulation. Such sheared ends would be unavailable for Klenow backfill and ligation, and would thus be unpackageable.

(ii) *DNA ends (possibly on lambda arms, but more probably on the inserts) might not have been properly backfilled during the Klenow partial fill-in reaction.* Some portion of the right and left lambda phage arms might not be available for ligation to genomic fragments because the partial fill-in reaction for the vector was not 100% efficient, allowing the unmodified right and left arms to covalently associate with each other but not with properly modified genomic fragments. Such "right : left arm ligation

products" would be still be too small to package, and would thus also be undetectable in the plaque assay. A more likely situation, however, is that inefficient Klenow partial fill-in modified only some of the insert DNA ends and/or did not properly modify those ends. Any genomic fragment without two properly modified ends would not be able to bind the required right and left vector arms and would thus not be packaged. Equally problematic is the situation in which a genomic fragment had only one properly modified end. That fragment would be able to bind up only one lambda arm and thus remain unpackageable while decreasing the number of vector arms that could otherwise be used in productive ligations.

(iii) *Contaminants might be present in the genomic DNA (esp. salts and/or residual detergent) which poison the Klenow partial filling step, the ligation step or both.* DNA ligase is fairly susceptible to inhibition by high salt. Since 1.2 M NaCl was used in the resolubilization of the CTAB-precipitated genomic DNA, inadequate washing of the DNA may have left enough residual salt to prevent efficient ligation. Recent work has since confirmed that ligation is more effective on genomic DNA that has undergone a more rigorous washing regime than was used in this work (ie. when

the DNA pellet was subjected to three 70% ethanol washes, rather than one) (Dr. Steven White, personal communication).

(iv) *Inefficient ligation might have occurred due to the selection of an inappropriate ratio of lambda arms to insert DNA during the ligation reaction.* Lambda GEM-12 arms and mouse DNA fragments were mixed together at two different mass ratios (1:1 and 2:1 respectively) and then subjected to ligation and packaging. Interestingly, no significant differences were found between the two ratios in terms of the yield of clones per microgram insert DNA (Table 4). Neither an abundance of genomic DNA, vector arms nor packaging extract was available, however, and there was no opportunity to examine the effect of different ratios on ligation efficiency.

(v) *Inefficient ligation might have occurred due to the use of partially inactivated ligase and/or ligase buffer, the use of too few units of activity for the number of picomole ends present, or both.* The partial thawing of the -70°C freezer twice during the course of this work makes it difficult to know exactly how many units of ligase activity were actually present in each DNA ligation reaction. In addition, a recent re-examination of the ligase reaction conditions found that ligation efficiency was fairly low (Dr. Steven White, personal communication). In fact,

ligation efficiency only reached acceptable levels (determined by analyzing the size of the ligation products via electrophoresis on a 0.4% agarose gel) when the number of units of ligase activity employed in each reaction was about two to four-fold higher than the amount suggested in the original protocol.

(vi) *Inefficient packaging of recombinant phage genomes might have occurred due to formation of monogenomic circular forms rather than concatemers.* The cos sites of the lambda arms could, theoretically, anneal to form circular molecules after insert DNA was ligated to the arms; such circular molecules would not be packaged into the phage head efficiently. Preincubation of the ligation mix at 42°C for one hour to allow annealing of lambda arms by their cos sites before ligation to the insert DNA (Perbal, 1984) might prove helpful here. Alternatively, increasing the concentration of DNA in the ligation reaction might also favor concatamer formation over the monomeric circle formation (Mandel & Higa, 1970).

If packaging of the recombinant genomes was inefficient, it might be possible to increase the number of functional phage generated per packaging reaction simply by increasing the amount of DNA added per packaging reaction. Promega's packaging protocol recommended the addition of

1.0-1.5 µg of total DNA per packaging reaction, and this recommended procedure was followed. While the efficiency of packaging *in vitro* is known to be independent of the DNA concentration, the total number of pfu generated in the packaging reaction, however, is not. In fact, the total number of pfu generated in any packaging reaction is linearly dependent on the amount of DNA added (Hohn, 1975). The packaging reaction was reexamined with respect to the amount of input DNA added. Increasing the amount of total input DNA from 1.5 to 3.5 µg produced about a 6-fold increase in the number of plaques obtained (to 3.12×10^5 pfu per packaging reaction) (Dr. Steven White, personal communication). This suggests that, using the new ligation reaction and packaging conditions, a complete genomic library could be constructed from an organism possessing a genome as large as a human (ie., 3×10^9 bp/haploid genome) by pooling three to four packaging reactions (generating at least 9.3×10^5 pfu total).

(vii) *Inefficient packaging of recombinant phage genomes might have occurred due to the fact that small genomic fragments could still tie up lambda arms but could not be packaged (ie. they were below the size limit for packaging into phage heads).* Titration studies were performed prior to the large large-scale digestion of

genomic DNA in order to determine the optimum ratio of enzyme to genomic DNA for partial digestion. Even when the "optimal" ratio of Sau3A units:ug of genomic DNA was identified and employed however, there was still a significant amount of DNA in the undesired range after partial digestion (Figure 8, panel B). Those fragments could still become ligated to right and/or left vector arms even though the ligation products could not undergo packaging. Even if we assume that the partial fill-in reaction was 100% efficient, many vector arms would be ligated to these poorly sized DNA fragments. As a result, the efficiency of the packaging reaction would drop.

One procedural modification that could be employed in future work is size fractionation of the restricted genomic DNA fragments using either electrophoresis through agarose gels or centrifugation through sucrose gradients. Unfortunately, both procedures demand large amounts of DNA and typically produce relatively poor yields. If adequate starting material is available, however, it may allow dramatic improvement in the amount of packagable insert DNA generated during ligations. This, in turn, would improve the packaging efficiency and the degree of sequence capture in the libraries.

Implications of this work and recommendations for future research.

Creating genomic libraries of endangered species will never take the place of species preservation. However, with an ever-increasing number of species vanishing, developing the means to preserve the genomic information accrued through hundreds of millions of years of evolution is vital. Developing a "field functional" protocol that would allow the construction of such "archival" genomic libraries is one approach.

While the use of a diluted DTAB solution clearly was found effective in stabilizing DNA as high molecular weight material, and while the DTAB/CTAB isolation protocol clearly produced HMW DNA that could be cloned into a genomic library, the sequence capture efficiency in these libraries was very low. Probably titering at least 3 or 4 packaging reactions for each species will have to be done to be "confident" ($p = 0.99$) of obtaining a fully representative archival quality library. Interestingly, while many genomic libraries have been constructed from insects (Blanchetot & Googing, 1994, 1995; Kumar & Sridhara, 1994; Kao et al., 1995; Okazaki et al., 1993), no attempts to determine sequence capture efficiency were presented in these studies. The reason for this may be that it is essentially impossible

to determine true sequence capture efficiency when the genome size is not known (the situation for most insects). Thus, in addition to defining a "field-ready" DNA stabilization reagent (diluted DTAB solution), the parallel construction of a mouse genomic library as an "external standard" to estimate sequence capture efficiency for an organism of unknown genome size represents another small conceptual and procedural advance.

Finally, this study provides an estimate of the reagent cost required to prepare an archival quality genomic library (about \$500-750 per species, given that an "external standard" mouse genomic library is also constructed in parallel). This information should prove useful in efforts directed at the procurement of needed funding to construction for archival libraries and pursue modifications in the current methodology aimed at increasing the efficiency of DNA sequence capture.

As mentioned previously, several procedures/methods discussed in this thesis need to be further modified/or investigated to increase *in vitro* packaging efficiency and to improve the efficiency of capture in a genomic library. This study not only presents a newly developed method for stabilizing HMW DNA but also provides useful information for construction of a "archival" genomic library. It is believed that an efficient procedure for construction of

"archival" genomic libraries will be of benefit to not only endangered insects but to all other species. More effort should be put into continuing these studies.

Figure 3. Agarose gel electrophoretic (AgGE) fractionation of DNA isolated from lady bugs fixed intact (ie. "whole") in various solutions and stored for 21 days at room temperature prior to DNA extraction.

Panel A. Whole insects were completely immersed in either (M) 100% methanol, (E) 95% ethanol, (I) 100% isopropanol, or (MA) methanol:glacial acetic acid (3:1). At the end of the 21 day storage (ie. fixation) period, DNA was extracted from the fixed tissues using the DTAB/CTAB protocol described in the "Methods" section. Equal volumes of extracted nucleic acid (15 μ l) were then analyzed on a 0.8% w/v agarose gel. HMW DNA could not be detected using these fixation procedures, although broad and diffuse bands of low molecular weight DNA were observed in all samples.

Panel B. In a separate experiment, insects were fixed in acidic DTAB solution (a mixture of 1 volume 5% citric acid: 2 volumes DTAB stock solution), stored for 7 days, and then subjected to the DNA extraction protocol. This fixation method yielded essentially no detectable DNA, and was immediately abandoned.

Figure 4. AgGE fractionation of DNA isolated from lady bugs that were minced and stored in various solutions for 7 days (Panel A) or 27 days (Panel B) at room temperature prior to DNA extraction. Equal numbers of insects were immersed and immediately minced in either (I) 100% isopropanol, (M) 100% methanol, or (E) 95% ethanol and stored for 7 days at room temperature, at which time the lysates were subjected to the DTAB/CTAB nucleic acid extraction protocol. Equal volumes (15 μ l) of each DNA sample were loaded and analyzed on a 0.8% gel. Similar band intensities suggest that the mass of DNA loaded in each lane was approximately equal. As shown in Panel A, at least some HMW DNA is now detectable (especially in the isopropanol-fixed tissue), but the majority of the DNA is still in the low molecular weight range. After 27 days of storage, however, essentially all of the HMW DNA has been degraded (see Panel B), indicating residual DNase activity in these fixatives.

Figure 5. AgGE fractionation of DNA isolated from lady bugs that were immersed and immediately minced and stored in DTAB solution for varying lengths of time (periods ranging from a few minutes to 1.5 years) prior to isolation of the DNA using the DTAB/CTAB procedure.

Panel A: Lane 1 is insect DNA isolated from tissue stored for less than 60 minutes in DTAB prior to DNA isolation. Lane 2 is insect DNA from tissue stored for 2 weeks in DTAB prior to DNA isolation.

Panel B: Lane 1 is uncut lambda genomic DNA. Lane 2 is insect DNA isolated from tissue stored for less than 60 minutes in DTAB prior to DNA isolation. Lane 3 is insect DNA from tissue stored for 3 weeks. Note that lane 4 is DNA isolated from insects stored intact (unminced) for 1 week in DTAB.

Panel C: Lane 1 is insect DNA isolated from tissue stored for less than 60 minutes in DTAB prior to DNA isolation. Lane 2 is insect DNA from tissue stored for 7 weeks.

Figure 5. (Cont'd)

Panel D: Lane 1 is insect DNA isolated from tissue stored for less than 60 minutes in DTAB prior to DNA isolation. Lane 2 is insect DNA from tissue stored for 9 weeks.

Panel E. Lane 1 is an aliquot of lambda DNA digested with Hind III restriction enzyme (fragment lengths in Kbp are indicated to the left). Lane 2 is insect DNA from tissue stored for 1.5 years in DTAB prior to DNA isolation.

Figure 6. Effect of differing NaCl concentrations on the yield of genomic DNA during the DTAB/CTAB DNA isolation procedure. NaCl solutions at 0.5M (lane 1), 1.0 M (lane 2), 1.2 M (lane 3), 1.4 M (lane 4), 1.6 M (lane 5), 1.8M (lane 6) and 2.0 M (lane 7) were used during the detergent displacement and DNA solubilization step (see text for explanation). Results indicate that anywhere from 0.5 to 2.0 M NaCl will allow recovery of HMW DNA, with no significant difference among the various salt concentrations.

Figure 7. Panel A: Test for the presence of residual DNase activity in isolated genomic insect DNA. Equal amounts of lady bug genomic DNA were incubated in low salt buffer at either 4°C (lanes 1 and 2) or 37°C (lanes 3 and 4) for one hour. Following this incubation the tubes were either held at room temperature (lanes 2 and 4) or boiled for 3 minutes (lanes 1 and 3), prior to analysis on a 0.8% agarose gel. The boiling step merely confirmed that the material could be denatured (and at least partially fragmented by mechanical shearing during boiling), as would be expected of DNA.

Panel B: Demonstration that isolated insect genomic DNA served as good substrate of restriction enzyme. Equal amounts of insect genomic DNA were incubated for 45 minutes at 37°C with (lane 3) or without (lane 4) PstI restriction enzyme (approximately 1.5 units). Lane 2 is uncut lambda genomic DNA, while lane 1 is Lambda DNA digested to completion with HindIII. The insect genomic DNA is clearly digested into a family of smaller fragments by PstI.

Panel A

Figure 8. Determination of the restriction enzyme (RE) concentration required to generate genomic DNA fragments of the optimal size range (ie. approximately 12-20 Kbp). Equal masses of HMW genomic DNA from lady bugs (*Panel A*) or mice (*Panel B*) were digested with decreasing concentrations of Sau 3A restriction enzyme for 30 minutes at 37°C. The samples were then analyzed on a 0.8% agarose gel.

Panel A: A total of 1.586 µg of insect genomic DNA was added to each tube (final reaction volume in all tubes was 20 ul). Bands in lane 1 are lambda/HindIII fragments. Lane 2 is uncut insect genomic DNA. Lanes 3-9 are samples of insect HMW genomic DNA digested at 0.315, 0.158, 0.079, 0.0394, 0.0197, 9.85×10^{-3} and 4.93×10^{-3} units Sau3A/µg DNA, respectively. Lane 7 was considered the optimal ratio (approximately 0.0197 units of Sau3A activity per ug insect DNA).

Panel B

Figure 8.(Cont'd)

Panel B: A total of 0.795 μg mouse DNA was used in each reaction tube (final reaction volume in all tubes was 20 μl). Bands in lane 1 are lambda/HindIII fragments. Lane 2 is uncut mouse genomic DNA. Lanes 3-10 are samples of mouse genomic DNA digested at 0.943, 0.472, 0.236, 0.118, 0.059, 0.029, 0.014 and 7.37×10^{-3} units Sau3A/ μg DNA, respectively. Average of the RE unit activity in lane 7 and 8 was considered the optimal ratio (approximately 0.0440 units of Sau3A activity per μg mouse DNA).

Figure 9. Large-scale digestion of insect and mouse genomic DNAs using Sau3A.

Panel A: A total of 55.51 µg of insect genomic DNA was digested with Sau3A at the optimal "RE units:ug DNA" ratio defined in Figure 8, panel A. An aliquot was then analyzed on a 0.8% agarose gel (lane 2) along with lambda/HindIII size standards (lane 1).

Panel B: A total of 50 µg of mouse genomic DNA was digested with Sau3A at the optimal "RE units:ug DNA" ratio defined in Figure 8, panel B. An aliquot was then analyzed on a 0.8% agarose gel (lane 4) along with lambda/HindIII size standards (lane 1). Lane 2 is uncut lambda DNA and lane 3 is uncut mouse genomic DNA.

Literature Cited

Adeli, K., and Ogboma, G. (1990) Rapid purification of human DNA from whole blood for potential application in clinical chemistry laboratories. *Clinical Chemistry* 36:261-264.

Aggarwal, R. K., Lang, J. W. and Singh, L. (1992) Isolation of high-molecular-weight DNA from small samples of blood having nucleated erythrocytes, collected, transported, and stored at room temperature. *GATA* 9 (2): 54-57.

Blanchetot, A. and Gooding, R. H. (1994) Isolation and characterization of hypervariable sequences from Tsetse Fly genome. *Insect Biochem.Molec. Biol.* 24:937-941.

Blanchetot, A. and Gooding, R. H. (1995) Identification of a *mariner* element from the tsetse fl, *Glossina palpalis palpalis*. *Insect Mol. Biol.* 4: 89-96.

Boyce, T. M., Zwick, M. E. and Aquadro, C. F. (1989) Mitochondrial DNA in the Bark Weevile: size, structure and heteroplasmy. *Genetics* 123: 825-836.

Clark, L. and Carbon, J. (1976) A cloning bank containing synthetic Col EI hybrids representative of the entire *E.coli* genome. *Cell* 9: 91-99 .

Clark, N. G. (1964) Modern Organic Chemistry. An Introduction. Oxford University Press.

Collins, J. and Bruning, H. J. (1978) Plasmid usable as gene-cloning vectors in an *in vitro* packaging by *E.coli* phage lambda: 'Cosmids'. *Gene* 4: 85-107.

Collins, J. and Hohn, B. (1979) Cosmid: A type of plasmid gene cloning vector that is packagable *in vitro* in bacteriophage lambda head. *Proc. Nalt. Acad. Sci. USA.* 75: 4242-4246.

Doyle, J. J and Doyle, J. L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull* 19: 11-15.

Doyle, J. J and Doyle, J. L. (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Dugaiczyk, A., Boyer, H. W. and Goodman, H. M. (1975) Ligation of EcoRI endonuclease-generated DNA fragments into linear and circular structures. *J. Mol. Biol.* 96: 171-184.

Eickbush, E. and Moudrianakis, E. N. (1978) The compaction of DNA helicals into either continuous supercoils or folded-fibered rods and toroids. *Cell* 13: 295-306 .

Ferritti, L. and Sgaramella, V. (1981) Temperature dependence of the joining by T4 DNA ligase of termini produced by type II restriction endonucleases. *Nucleic Acids Res.* 9: 85-93.

Firschauf, A-M. (1978) *Methods in Enzymology*. In [16] Digestion of DNA: Size fractionation. 152: 184. Academy Press. Inc.

Fox & Fox, (1964) *Introduction to Comparative Entomology*. Reinhold Pub. Co.

Gustincich, S., Manfioletti, G., Del Sal, G. and Schneider C. (1991) A fast method for high-quality genomic DNA extraction from whole human blood. *BioTechnique*. 11: 298-300.

Hanahan, D. and Meselson M. (1980) Plasmid screening at high colony density. *Gene*. 10: 63-67.

Hohn, B. and Murray, K. (1977) Packaging recombinant DNA molecules into bacteriophage particles *in vitro*. *Proc. Natl. Acad. Sci. USA* 74: 3259-3263.

Humason, G., L. (1979) *Animal Tissue Techniques*. (Fourth Edition.) W. H. Freeman and Company, San Francisco.

Jones, A. S. (1963) Use of alkyltrimethylammonium bromides for the isolation of ribo-and deoxyribo-nucleic acids. *Nature* 199: 280-282.

Kao, W-Y., Hankeln, T., Schmidt, E. R. and Bergtrom, G. (1995) Sequence and evolution of the gene for the monomeric globin I and its linkage to gene coding for dimeric globin in the insect *Chironomus thummi*. *J. Mol. Evol.* 40: 354-361.

Koshiba, M., Ogawa, K., Hamazaaki, S., Sugiyama, T. and Ogawa, O. (1993) The effect of formalin fixation on DNA and the extraction of high-molecular-weight DNA from fixed and embedded tissues. *Pathology Research and Practice* 189: 448-451.

Kumar, M. N. and Sridhara, S. (1994) Characterization of four pupal wing cuticular protein genes of the silkworm *Antheraea Polyphemus*. *Insect Biochem. Molec. Biol.* 24: 291-299.

Leach, D. R. F. and Stahl (1983) Viability of lambda phage carrying a perfect palindrome in the absence of recombination nucleases. *Nature* 305: 448-451.

Lederberg, E. M. and Cohen, S. N. (1974) Transformation of *Salmonella typhimurium* by plasmid deoxyribonucleic acid. *J. Bacteriol.* 119: 1072.

Lo Piero, A. R. and Petrone, G. (1994) Effect of detergents on the chicken liver F1-ATPase: functional and structural aspects. *Biochem. Mole. Biol. International* 32: 181-189.

Mandel, W. and Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* 53: 159-162.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. and Efstradis, A. (1978) The isolation of a structural gene from libraries of eukaryotic DNA. *Cell* 15: 687-701.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning. A laboratory manual.* Cold Spring Harbor Laboratory. 65.

Murray, A. W. and Szostak, J. W. (1983) Construction of artificial chromosomes in yeast. *Nature* 305: 189-193.

Murray, M. G. and Thompson, W. F. (1980) Rapid isolation of high-molecular-weight plant DNA. *Nucl. Acids Res.* 8: 4321-4325

Murray, N. E. (1983) *Phage Lambda and Molecular Cloning.* In the bacteriophage lambda, eds Hendrix, R. W., Roberts, J. W., Stahl, F. W. and Weisberg, R. A. *Lambda II* (Monograph No. 13), Vol. 2. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York.

Murray, N. E., Brammar, W. J. and Murray, K (1977) Lambda phage that simplify the recovery of *in vitro* recombinants. *Molec. Gen. Genet.* 150: 53-61.

Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H. and Fujiwara, H (1993) Identification of a penanucleotide telomeric sequence, (TTAGG)_n, in the silkworm *Bombyx mori* and in other insects. *Mol. Cell. Biol.* 14: 1424-1432.

Olivera, B. M., Hall, Z. W. and Lahman, I. R. (1968) Enzymatic joining of polynucleotides. V. A DNA adenylate intermediate in the polynucleotide joining reaction. *Proc. Natl. Acad. Sci. USA* 87: 1451-1455.

Perbal, B. (1984) *A Practical Guide to Molecular Cloning*. New York: Wiley 1984.

Reiss, R. A., Schwert, D. P. and Ashworth, A. C. (1995) Field preservation of Coleoptera for molecular genetic analyses. *Physiological and Chemical Ecology* 24: 716-719.

Roger, S.O. and Bendich, A. J. (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5: 69-76.

Sanger, F., Coulson, A. R., Hong, G-F., Hill, D. F. and Peterson, G. B. (1982) Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.* 162: 729-773.

Seed, B., Parker, R. C. and Davidson, N. (1982) Representation of DNA segments in recombinant DNA libraries prepared by restriction enzyme partial digestion. *Gene* 19: 201-209.

Shahjahan, R. M., Hughes, K. J., Leopold, R. A. and DeVault, J. D. (1995) Lower incubation temperature increases yield of insect genomic DNA isolated by the CTAB method. *Biotechniques* 19: 332-334.

Shizuya, H. B., Birren, B., Kim, U. J., Mancino, V., Slepac, T., Tachiiri, T., and Simon, M. (1992) Cloning and stable maintenance of a 300-Kilo-base-pair fragment of human DNA in *Escherichia coli* using an F-factor based vector. *Proc. Natl. Acad. Sci. U.S.A.* 89: 8794-8797.

Singer, E., Kuenzle, C. C., Thomann, P. E., and Ulrich, H. (1989) DNA fingerprinting: Improved DNA extraction from small blood samples. *Nucleic Acid Res.* 16: 7738-7739.

Warlker, R. (1991) *Environment on File*. In " Land": The tropical forests. Ed Doyle, M. Fact on File. Inc. Ltd. Oxford.

Appendix

Total numbers of pfu in diluted phage after packaging reactions and packaging efficiency

1. Packaging reaction for positive control DNA

a. Total 5 μ l of ligation mixture (containing 1.0 μ g of input DNA) were added into 25 μ l of packagen extract and 223 μ l of phage buffer. The total volume of diluted phage was 0.253 ml.

b. A series of 10-fold dilutions were prepared from the packaged phage. From each diluted phage, 0.1 ml of a phage solution were added into 0.1 ml of maltose and Mg_2SO_4 treated *E.coli* LE 392. The phage infected bacterial was then mixed with top agar and plated.

c. Total 35 plaques were observed on the 10^{-4} plate.

- Viral titer = $35 \times 10^4 / 0.1 \text{ ml} = 3.5 \times 10^6 \text{ pfu/ml}$
- Total pfu generated in the packaging reaction = $3.5 \times 10^6 \text{ pfu/ml} \times 0.253 \text{ ml} = 8.86 \times 10^5 \text{ pfu}$.
- Packaging efficiency = $8.86 \times 10^5 \text{ pfu} / 1.0 \mu\text{g input DNA} = 8.86 \times 10^5 \text{ pfu} / \mu\text{g input DNA}$.

2. Packaging reaction for mouse DNA

- a. Two packaging reaction were performed for mouse DNA. For each reaction, 5 μ l of ligation mixture (containing 1.0 μ g of input DNA for #1 reaction and 0.75 μ g of input DNA for #2 reaction) were added into 25 μ l of package extract and 223 μ l of phage buffer. The total volume of diluted phage was 0.253 ml.
- b. A series of 10-fold dilutions were prepared from the packaged phage. From each diluted phage, 0.1 ml of a phage solution were added into 0.1 ml of maltose and Mg_2SO_4 treated *E.coli* LE 392. The phage infected bacterial was then mixed with top agar and plated.
- c. Total 121 plaques were observed on the 10^{-2} plate for reaction #1 and 73 plaques were observed on the 10^{-2} plate for the reaction #2.
- Viral titer for #1 = $121 \times 10^2 / 0.1 \text{ ml} = 1.21 \times 10^5 \text{ pfu/ml}$
Viral titer for #2 = $73 \times 1^2 / 0.1 \text{ ml} = 7.3 \times 10^4 \text{ pfu/ml}$
 - Total pfu in the packaging reaction #1 = $1.21 \times 10^5 \text{ pfu/ml} \times 0.253 \text{ ml} = 3.06 \times 10^4 \text{ pfu}$.
Total pfu in the packaging reaction #2 = $7.3 \times 10^4 \text{ pfu/ml} \times 0.253 \text{ ml} = 1.85 \times 10^4 \text{ pfu}$.
Total pfu for mouse genomic library = $(3.06 + 1.85) \times 10^4 = 4.91 \times 10^4 \text{ pfu}$.
 - Packaging efficiency for #1 reaction = $3.06 \times 10^4 \text{ pfu} / 1.0 \mu\text{g input DNA} = 3.06 \times 10^4 \text{ pfu} / \mu\text{g input DNA}$.

Packaging efficiency for #2 reaction = 1.85×10^4 pfu /
0.75 μ g input DNA = 2.47×10^4 pfu / μ g input DNA.

3. Packaging reaction for lady bug DNA

a. Total 10 μ l of ligation mixture (containing 2.05 μ g of input DNA) were added into 50 μ l of packagen extract and 455 μ l of phage buffer. The total volume of diluted phage was 0.505 ml.

b. A series of 10-fold dilutions were prepared from the packaged phage. From each diluted phage, 0.1 ml of a phage solution were added into 0.1 ml of maltose and Mg_2SO_4 treated *E.coli* LE 392. The phage infected bacterial was then mixed with top agar and plated.

c. Total 131 plaques were observed on the 10^{-2} plate.

- Viral titer = $131 \times 10^2 / 0.1 \text{ ml} = 1.31 \times 10^5$ pfu/ml
- Total pfu generated in the packaging reaction = 1.31×10^5 pfu/ml $\times 0.253 \text{ ml} = 6.62 \times 10^4$ pfu.
- Packaging efficiency = 6.62×10^4 pfu / 2.05 μ g input DNA = 3.23×10^4 pfu / μ g input DNA.

% Capture of mouse genomic library

The total numbers of pfu in the mouse genomic library, N, were 4.91×10^4 pfu. If the genome size of mouse, G, is 2.3×10^6 Kb; the size of fragments cloned is 20 Kb:

$$N = \ln(1-P) / \ln(1 - 20 \text{ kb} / 2.3 \times 10^6 \text{ Kb})$$

$$N = \ln(1-P) / \ln(1 - 8.7 \times 10^{-6})$$

$$N = \ln(1-P) / (-8.7 \times 10^{-6})$$

$$\ln(1-P) = 4.91 \times 10^4 \times (-8.7 \times 10^{-6}) = -0.43$$

$$P = 1 - (e^{-0.43}) \times 100 \% = 35\%$$