IMPROVED PROCEDURE FOR EXTRACTION AND PURIFICATION OF HIGH MOLECULAR WEIGHT DNA FROM *LILIUM PARDALINUM*

A University Thesis Presented to the Faculty of California State University, East Bay

In Partial Fulfillment of the Requirements for the Degree Master of Science in Biological Sciences

By,

Aruna Arcot

June, 2015
Abstract

Previous studies indicate that genome size is not an indicator of organismal complexity. Moreover, non-coding portions of the genome, including transposable element (TE) repeats contribute to genome obesity. It is not clear, however, if TE content and therefore genome obesity is advantageous in evolution. The family Liliaceae is ideally suited for this investigation, since its family members show a range of genome sizes as well as significantly obese genomes. Sequencing the genomes is the first step in addressing this question in Liliaceae. Typically shotgun sequencing is the method of choice for whole genome sequencing. However, this approach is not suitable for complex genomes with sequence redundancy, such as those in Liliaceae plants. In these cases, BAC library based sequencing rather than shotgun sequencing is the preferred. In order to obtain high molecular weight DNA required for BAC library preparation, whole nuclei extraction for DNA purification was used in previous attempts at creating BAC library for Fritillaria species, a plant with one of the largest known genomes, by researchers Gangavarapu, Jayakar, Mishra and Patel. Their efforts yielded mixed results at best and it was seen that the DNA purified from the procedure was insufficient in quality and quantity.
This thesis research outlines two specific aims:

(1) To develop an alternate protocol to extract and purify high molecular weight (HMW) DNA suited for ligation with BAC vector, from the leaf tissue of *Lilium pardalinum*, a species representative of the Liliaceae family.

(2) To demonstrate through a flow cytometry assay that the whole nuclei based DNA extraction is not suitable for this purpose.

Once the genome is successfully extracted and purified, the high molecular weight (HMW) DNA can then be used to create BAC clones for whole genome sequencing of *L.pardalinum*, as well as other downstream applications.
IMPROVED PROCEDURE FOR EXTRACTION AND PURIFICATION OF
HIGH MOLECULAR WEIGHT DNA FROM LILIUM PARDALINUM

By
Aruna Arcot

Approved:
Dr. Christoph Baysdorfer
Dr. Claudia Fordstone
Dr. Maria Gallegos
Dr. Kenneth Curr

Date:
5/29/15
2015 Jun 3
6/3/2015
June 5, 2015
Acknowledgements

This thesis research effort is dedicated to my son, Vishal Arcot and to my husband, Krishna Arcot.

Many thanks to Dr. Claudia-Uhde Stone, Dr. Maria Gallegos and Dr. Kenneth Curr, for being my thesis committee members, for your encouragement and all that I have learned from you.

My sincere appreciation and thanks to my friend and lab partner Vincent Jieh-Yeng Choo, for your help in taking this project across the finish line.

My heartfelt thanks to my friends Soma’n’Sarvesh Surumpudi, for your support and for always being there.

Above all, my deep gratitude goes to Dr. Chris Baysdorfer for giving me the opportunity to work on this project, for your kindness and your confidence in me and for the chance to become a part of the Biocore family.
## Table of Contents

Abstract.................................................................................................................................ii

Acknowledgments..................................................................................................................v

List of Figures..........................................................................................................................ix

Background...............................................................................................................................1

  Previous Attempts at Creation of Fritillaria BAC Library .................................................7

Methods and Materials ..........................................................................................................10

  Effect of centrifugation on nuclear membrane integrity during whole nuclei isolation from lily plants using Flow Cytometry .........................................................10

  Extraction of high molecular weight DNA using Qiagen’s Genomic-tip 20/G and Genomic DNA Buffer set .........................................................................................12

  DNA quantification using picogreen assay for Nucleic Acids on Nanodrop 3300 Spectrometer ..............................................................................................................14

  Accessing DNA quality using BioRad’s Pulsed Field Gel Electrophoresis ........................15

  DNA end repair using Ion Torrent’s DNA End Repair Kit .................................................16

  HindIII adapter ligation to end repaired DNA using FastLink ligation Kit .......................17

  3% agarose Gel Electrophoresis to purify extract/insert of un-ligated adapters ...............18

  Insert-vector cohesive ended ligation using FastLink ligation kit ...................................19

  GELase treatment to digest agarose plugs and desalting protocol .................................19
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td>21</td>
</tr>
<tr>
<td>Isolation of whole nuclei, In-plug extraction and HindIII digestion of DNA from Asiatic lily leaf tissue</td>
<td>21</td>
</tr>
<tr>
<td>First Extraction of genomic DNA from leaf tissue of <em>Lilium pardinum</em> using Qiagen’s Genomic Buffer Kit and Genomic Tip 20/G</td>
<td>23</td>
</tr>
<tr>
<td>Second DNA Extraction from <em>L.</em> <em>pardalinum</em> Leaf Tissue</td>
<td>26</td>
</tr>
<tr>
<td>Hourly PFGE for DNA purification</td>
<td>27</td>
</tr>
<tr>
<td>PFGE on <em>L.</em> <em>pardalinum</em> DNA to remove fragments smaller than 50 KB (plug purification)</td>
<td>28</td>
</tr>
<tr>
<td>Third DNA Extraction from <em>L.</em> <em>pardalinum</em> Leaf Tissue</td>
<td>31</td>
</tr>
<tr>
<td>PFGE on <em>L.</em> <em>pardalinum</em> DNA to remove fragments smaller than 50KB</td>
<td>32</td>
</tr>
<tr>
<td>PFGE to Assess DNA Quality Under Storage</td>
<td>33</td>
</tr>
<tr>
<td>Fourth DNA Extraction from <em>L.</em> <em>pardalinum</em> Leaf Tissue</td>
<td>35</td>
</tr>
<tr>
<td>Fifth DNA Extraction from <em>L.</em> <em>pardalinum</em> Leaf Tissue</td>
<td>37</td>
</tr>
<tr>
<td>Sixth DNA Extraction from <em>L.</em> <em>pardalinum</em> Leaf Tissue</td>
<td>44</td>
</tr>
<tr>
<td>Seventh DNA Extraction from <em>L.</em> <em>pardalinum</em> Leaf Tissue</td>
<td>47</td>
</tr>
<tr>
<td>FACS Assay Results</td>
<td>52</td>
</tr>
<tr>
<td>Test flow cytometry assay to assess the effect of centrifugation on nuclear membrane integrity</td>
<td>52</td>
</tr>
<tr>
<td>Optimization assays to capture fluorescence from test species nuclei</td>
<td>56</td>
</tr>
<tr>
<td>Flow assay to assess the effect of centrifugation on <em>L.</em> <em>pardalinum</em> nuclei</td>
<td>63</td>
</tr>
</tbody>
</table>
Flow assay to determine effect of incubation time on nuclei integrity ........71

Discussion ........................................................................................................................................74

References ..........................................................................................................................................78

Appendix............................................................................................................................................83
# List of Figures

1: Post-enzyme digest electrophoresis of Asiatic Lily DNA plugs .................................. 22
2: PFGE of Asiatic Lily plugs purified of smaller fragments .............................................. 22
3: *L. pardalinum* DNA (Extract 1) on a 1% TAE/Agarose gel ........................................ 24
4: PFGE for *L. pardalinum* DNA (Extract 1) for eluate 1 ............................................. 25
5: PFGE for *L. pardalinum* DNA Second extraction (Extract 2) .................................... 26
6: Hourly PFGE of Pulse Marker 0.1 to 200KB Ladder .................................................. 28
7: *L. pardalinum* Extract 2 plug DNA purification for 2.5hrs on PFG ............................. 29
8: PFGE of *L. pardalinum* Extract 2, purified of smaller fragments ............................... 30
9: PFGE for *L. pardalinum* Extract 3 ........................................................................... 31
10: PFGE for plug purification of Extract 3 ...................................................................... 32
11: PFGE to check DNA fragment size quality for all samples Extract 1 to 3 ............... 34
12: PFGE of *L. pardalinum* Extract 4 ........................................................................... 36
13: 2% TAE Gel Electrophoresis for plug purification ....................................................... 36
14: PFGE on Extract 5, plug purified Extracts 2 &3, and 4 ........................................... 37
15: 10 min gel electrophoresis for plug purification of Extracts 5 and 4 ....................... 39
16: 2% Gel electrophoresis to assess fragment quality of adapter ligated Extract 5 .............. 40
17: Plug purification of Extracts 5 & 4, after insert-vector ligation .................................. 41
18: Post GELase treatment gel picture of Extract 5 & 4 ligation products .................. 42
19: HindIII digestion and Test ligation of pGem 3Zf+ vector...........................................43

20: A. Gel electrophoresis on end repaired Extract 6 to remove small DNA fragments. B. PFGE on Extract 6 plugs, post TAE gel electrophoresis (A).........45

21: PFGE on Pre-ligation and Post-vector ligation Extract 6 ...........................................46

22: PFGE on Extract 7 and post plug purified Extract 7.......................................................48

23: PFGE to excise adapter ligated Extract 7 for electro-elution. Post electro-elution gel picture .................................................................................................................................50

24: PFGE of electro-eluted Extract 7, and Post vector ligation of electro-eluted Extract 7 .................................................................................................................................52

25: Flow assay to determine effect of centrifugation on P.smithii nuclei ..........................53

26: Flow assay to test the effect of centrifugation on P.smithii and L.pardalinum whole nuclei.........................................................................................................................55

27: Flow assay to optimize the capture of P.smithii fluorescence in specific channels .................................................................................................................................56

28: Flow assay to optimize the capture of L.pardalinum fluorescence in specific Channels .................................................................................................................................57

29: Flow assay to optimize the capture of P.smithii and L.pardalinum fluorescence in specific channels .................................................................................................................................58

30: Flow assay to assess fluorescence capture due to LB01 and/or PI............................59

31: Flow assay to optimize the capture of L.pardalinum fluorescence in specific channels .................................................................................................................................60

32: Flow assay to visualize the capture of P.smithii fluorescence in specific channels .................................................................................................................................61

33: Flow assay to visualize L.pardalinum fluorescence in specific channels ..........62

34: Flow assay to visualize P.smithii and L.pardalinum fluorescence
35: Fluorescence captured from P. smithii nuclei prior to centrifugation

36: Fluorescence captured from P. smithii nuclei after centrifugation

37: Fluorescence from L. pardalinum nuclei prior to centrifugation

38: Fluorescence from L. pardalinum nuclei after centrifugation

39: Fluorescence from L. pardalinum nuclei prior to centrifugation

40: Fluorescence from L. pardalinum nuclei pellet after centrifugation at 100rpm

41: Fluorescence from pelleted nuclei after centrifugation at 3200rpm

42: Fluorescence from supernatant after centrifugation at 3200rpm

43: Fluorescence from L. pardalinum nuclei after 15min incubation with PI

44: Fluorescence from L. pardalinum nuclei after 1hr incubation with PI

45: Fluorescence from L. pardalinum nuclei after 2hr incubation with PI
Background

The C-value (Swift, 1950) is the total amount of DNA within a haploid nucleus of an organism and is the genome size of that organism. The term C-value paradox (Thomas, 1971) is based on the observation that genome size does not necessarily correlate with the organism’s biological complexity and the number of coding genes within its genome. Genome sizes (C-values) therefore seem to differ primarily due to the variation within the non-coding part of the genome. That it is unclear what overall effects the non-coding DNA has on the organism is referred to as the C-value enigma (Gregory, 2001).

The family Liliaceae, belonging to the order Liliales, is comprised mostly of diploid species and has several genera of flowering plants remarkable in that they show a wide range of C-values, including species with giant genome sizes (Leitch et al., 2007). Genome sizes among the Liliaceae vary, with the smallest being 3pg for Prosartes hookeri and the largest, 127.4pg (triploid) for Fritillaria assyriaca (RBG Kew DNA C-values database query; Bennett and Smith, 1976). The family members of Liliaceae therefore provide for a good example of the C-value paradox. Although polyploidy accounts for some of the C-value variation in other families, most of the Liliaceae plants owe their big genome size to the presence and prevalence of transposable elements (Fay et al., 2005).
Transposable elements are repetitive genetic DNA sequences which can proliferate within a genome either by a cut-paste mechanism (DNA transposons) or copy-paste (RNA mediated retrotransposons). Both DNA transposons as well as retrotransposons have been discovered and catalogued within the genomes of the Lily family (Leeton and Smyth, 1993, Ambrozova et al., 2011). In higher plants, transposable elements form a significant portion of the non-coding portion of the genome and have been known in general to interfere with gene function either by inserting themselves into the coding part of a gene sequence (Bennetzen, 2000), in the promoter region (Kidwell, 2002) or create gene duplications (copy and paste mechanism of transposons) (Xiao et al., 2008). It has been suggested that they can alter transcriptional activity by providing promoter, enhancer or transcription factor binding sites (Warnefors et al., 2010). They lead to an increase in genome size (Hartl, 2000, Gregory, 2001) and cause stable mutations that accumulate over time (Kunze et al., 1997). They form a key driving force for genome evolution (Vicient et al., 1999, Xiao et al., 2008). It has also been suggested that genome size may affect annual versus perennial lifestyle in plant species (Bennett, 1987). The prevalence of repetitive DNA possibly puts greater demands on the plant life cycle and propagation (Palevitz, 2000); due to the increased demand on resources and therefore the duration it might take for a
larger genome to replicate. For example, *Fritillaria assyriaca* (124GB) takes about six weeks to undergo meiosis (Bennett, 1997), while the entire lifecycle in *Arabidopsis* (135MB) lasts about six weeks. Evidence therefore points to potential impact of the transposable elements on plant genomes at molecular level, as well as on the species morphology and adaptation.

An interesting question to address would be whether genome size and the transposable elements that contribute to genome obesity confer any evolutionary advantage or disadvantage to the survival of the plant and if they do what the mechanisms are. The plants of the Liliaceae family are uniquely suited for this study, due to the fact that they comprise of a wide range of haploid genome sizes as well as giant genome sizes. The plant of choice for this thesis research was *Lilium pardalinum*.

*Lilium pardalinum* is an herbaceous perennial plant, native to Northern America. It is found along the coast of California and Oregon, and in the Sierra Nevada (USDA/NRCS Plant Guide). It is commonly known as the leopard lily and has alternate names ‘tiger lily’ or ‘panther lily’. The plant is listed as an endangered species, according to the USDA website. It has five sub species: pardalinum, pitkinense (rare), shastense, vollmeri, and wigginsii. The genome size of *Lilium pardalinum* has been estimated by this researcher as 41-42pg.
Compared to the human genome size of 3pg, this represents a significantly obese genome and therefore, *L. pardinum* is a good candidate for investigation into the C-value paradox. This plant species was used as the primary sample specimen during this project.

Whole genome sequencing is increasingly becoming the initial step in genomic research. As helpful as sequencing the genic part of a genome is, recent focus on the non-coding DNA outlines the potential importance of the intergenic regions as well. Quantitative analysis of the repetitive DNA sequences, which are a key driving force of genome expansion, is better facilitated by whole genome sequencing. The presence of the non-coding DNA and the question of its importance on the organism’s survival and evolution being at the heart of the concept of C-value enigma, genome sequencing is an important tool to address the question of genome size and it’s relation to organismal complexity. With the advances in sequencing techniques leading to high throughput sequencing capability, it has become easier to sequence the genomes of a variety of organisms from the plant and the animal kingdom (NCBI Complete Genomes List). Traditional DNA sequencing by chain termination and the shotgun sequencing technique are effective at sequencing shorter fragment lengths, starting from about 700bp up to 8-10KB long. During whole genome shotgun
sequencing, an amplified genome is randomly shredded by physical means into small sized fragments, and the fragments are sequenced and arranged into contigs by software programs. When the sequence reads are shorter, the programs map them to available reference genome sequences. For larger genomes that are riddled with sequence redundancy due to presence of high transposable element content, as might be the case with Liliaceae plants where closely related reference genomes are not commonly available, shotgun sequencing of the genome might fall short. This is due to the fact that it would be difficult to map the random sheared repetitive fragments to specific locations on the chromosomes, especially, since repeat sequences (as in a transposable element containing sequence) can map incorrectly or map to more than one location. In such cases, hierarchical sequencing involving BAC to BAC sequencing could prove to be a better solution.

An invaluable tool for cloning large DNA fragments and creating libraries, a BAC vector is a synthetic DNA construct that has its origins in the fertility factor (F factor) of E.coli. The fertility factor (Lederberg et al., 1952) is an extra chromosomal circular DNA contained within the bacterium and can independently replicate. F factor exchange can occur during bacterial conjugation and this transfer is known to carry genes coding for virulence from the host cell
to the recipient. This ability of the F factor plasmid to transfer between bacterial cells and to incorporate within and carry additional DNA segments was utilized by Shizuya et al. in 1992 to create the first Bacterial Artificial Chromosome (BAC). A typical F factor based BAC vector codes for genes that are required for its own replication as well as to maintain its copy number to 1-2 copies per cell. The oriS and repE regulatory genes code for replication, while the parA and parB regulatory genes are responsible for maintaining the plasmid copy number (Shizuya et al., 1992) in daughter cells. BAC vectors are stable, can carry DNA inserts up to 300KB long (Shizuya et al., 1992) and successfully and stably replicate within the host bacterium. BACs are therefore well suited for cloning high molecular weight DNA fragments for sequencing as well as other downstream applications. BAC sequences can be used for end sequencing, and physical mapping of the sequences to the genome. Restriction analysis of BACs and selection of a minimal tiling path provides for creating contigs for assembly of uninterrupted sequences of a large part of chromosomes generated by the clones (Marek and Shoemaker, 1997, Marra et al., 1997), as well as reduce the complexity associated with sequencing a genome riddled with repeat elements. BAC cloning and sequencing, although cumbersome, is therefore preferable to shotgun sequencing for complex genomes as those of the Liliaceae species.
YAC or Yeast Artificial Chromosome is a vector that can stably replicate within the single celled eukaryote, yeast and is used to clone DNA fragments up to 1MB long. Initially cloned by Burke et al. (1987), a YAC is a linear vector consisting of two arms of chromosome, a telomere, origin of replication, centromere, a cloning site and markers for selection (Zhang H, 2000). Although BAC vectors incorporate cloned fragments of much smaller sizes, they are considered more suitable as vectors for long fragment length DNA cloning, since a BAC is more stable than a YAC system (Shizuya et al., 1992, Boysen et al., 1997). It is harder to purify YAC DNA from yeast cells, and the formation of chimera and chromosomal rearrangement within YACs (Green et al., 1991, Nagaraja et al., 1994) are potential drawbacks of using a YAC for cloning. Another fact to be taken under consideration is that the YAC is maintained as a linear molecule which will make it prone to mechanical shearing when compared with the circular or supercoiled BAC.

**Previous Attempts at Creation of a Fritillaria BAC library:**

Development of a *Fritillaria* (a genus in the family Liliaceae) BAC library was attempted by previous researchers Jayakar (2005), Gangavarapu (2006), Mishra (2008) and Patel (pers comm). For library creation, these researchers used the whole nuclei isolation method developed by Zhang et al. (2000). Whole
nuclei were extracted from leaf tissue by either chopping the fresh leaf tissue in nuclei extraction buffer or by grinding frozen leaf tissue in homogenization buffer. The extract was centrifuged at low speeds (between 500-2000 g), the nuclei containing pellet re-suspended into appropriate buffer and poured into agarose plugs. The plugs were used in order to immobilize the nuclei and prevent mechanical shearing of the DNA within. Nuclear membrane lysis and subsequent buffer washes to remove contaminants was conducted in the plugs. The DNA fragments were then digested in the plug with HindIII, size selected for high molecular weight fragments, extracted from the plug and ligated into a linearized BAC vector (pBeloBAC11) with HindIII compatible ends. Jayakar (2005) was able to obtain sufficient amount of *Fritillaria* DNA fragments of the range 50-200kb, but the second size fractionation and subsequent GELase treatment reduced the DNA concentration to 10ng/ul. Theoretically, this would still have been sufficient amount of DNA for transformation. However, the insert-vector ligation results were unclear and the electroporation did not yield any colonies. When Gangavarapu attempted the nuclei isolation using the Zhang protocol, it was noticed that there were on average 40-45 nuclei obtained from each extraction (Gangavarapu 2006), while it is estimated that concentration of 5 – 10 X 10^7 nuclei/ml for a 1GB genome size is the preferred number for extraction
and purification of high molecular weight DNA (Zhang H, 2000). So it was
evident that the amount of starting material for DNA extraction was highly
insufficient and accordingly the high molecular weight DNA obtained from the
first size selection was very dilute. DNA concentration obtained by electro-
elution of DNA from the size selection procedure was only 1pg/ul. The electro-
eluted DNA was used in ligation and resulted in very few colonies with inserts
of about 50KB in size. One of the reasons for the very low transformation
efficiency was thought to be that the vector molecules re-circularized without the
uptake of insert DNA and therefore produce hundreds of blue colonies,
containing empty vectors (Gangavarapu, 2006). Mishra continued the work to
attempt the preparation of a BAC clone from high molecular weight DNA of
*Fritillaria* agrestis. It seemed initially that increasing the amount of leaf tissue
used for DNA extraction 5 fold, produced a good concentration of DNA,
observed as bright fluorescence of DNA on the diagnostic gel. Only when the
restriction digestion was conducted on the DNA in plug was it seen that hardly
any DNA was present in the plug. Therefore, the DNA was being lost in the
initial extraction step itself, and not enough was being used for further
purification (Mishra, 2008). The nuclei isolation protocol was fine-tuned by
testing on different plant species, both from Lily and non-Lily family by Patel.
Flow Cytometry based quantification of whole nuclei was conducted, and it was seen that *Fritillaria* agrestis leaf tissue yielded low numbers of nuclei. The in-plug DNA extraction and agarose gel based size selection yielded did not show the desired results of high molecular weight DNA (pers comm). The results obtained by these researchers suggest that the initial extraction of DNA was either insufficient in quality and quantity or that the DNA used for vector ligation was of poor quality in terms of fragment length which undermines their primary objective of creating BAC libraries with high molecular weight insert DNA. This necessitated the development of an alternate procedure to obtain suitable concentration and quality of HMW DNA from the Lily species for creation of BAC clones and library.

**Methods and Materials**

**Effect of centrifugation on nuclear membrane integrity during whole nuclei isolation from lily plants using Flow Cytometry.**

Fluorescence based genome content (2n) estimation from whole nuclei of different Liliaceae species was conducted earlier in this lab by Noel Canio (2009), using Coulter EPICS-XL Flow Cytometer. The protocol used by this researcher was followed for the Flow Cytometry assay.

The following buffers were prepared for whole nuclei extraction:
Buffer LB01 – Nuclei isolation buffer (Dolezel et al., 1992), stored in -20°C: 15 mM Tris pH 7.4, 2 mM Na₂EDTA, 0.5 mM Spermine tetrahydrochloride, 80 mM KCl, 20 mM NaCl, and 0.1% (v/v) Triton X-100. Before use, 15 mM β-mercaptoethanol and 1mg/ml RNase were added to the LB01 buffer.

5ml of the Nuclei Isolation buffer was maintained on ice just before the extraction procedure. Approximately 200mg of washed L.pardalinum leaf tissue was finely chopped with a sterile blade, in the Nuclei Isolation buffer. Supernatant from the extract was filtered through Spectra nylon mesh into a FACS tube, with a wide bore barrier pipette tip and incubated on ice with 100ul of Propidium Iodide (1mg/ml) for 15min or more. Alternately, to test the effect of centrifugation, the supernatant from the tissue extract was centrifuged at varied speeds for 3 minutes or greater at 4°C. The supernatant was removed and the pellet re-suspended in Nuclei Isolation Buffer, and incubated for at least 15 minutes with 100ul Propidium Iodide (1mg/ml). The prepared extracts were assayed on the flow cytometer to determine the number of nuclei measured by the fluorescence channel FL3. Nuclei Isolation Buffer with 100ul Propidium Iodide (at 1mg/ml) was used to blank the instrument. Histograms were recorded for number of events (nuclei) versus linear as well as log FL3 value. The number of events recorded by the cytometer for the nuclei extract without centrifugation
was compared to the number of events recorded for the extracts that were centrifuged to estimate if centrifugation had an effect on the *L.*pardalinum nuclei. Significant difference between the numbers recorded with and without centrifugation indicated that the lily nuclei were probably not staying intact during centrifugation, with the nuclear membrane lysing and the DNA being lost in the supernatant. That would mean not enough DNA was available for purification and cloning. Since the centrifugation was the first required step in the nuclei isolation protocol (Zhang H, 2000), as well as multiple centrifugation and washing steps are required for obtained a nuclei pellet, this indicated that the amount of starting tissue material might still not make a significant difference for nuclei extraction procedure, and invariably result in loss of intact nuclei.

**Extraction of high molecular weight DNA using Qiagen’s Genomic-tip 20/G and Genomic DNA Buffer set.**

Qiagen’s Genomic DNA buffer set of lysis buffer G2, equilibration buffer QBT, wash buffer QC, elution buffer QF and Genomic-tip 20/G was used for the column based DNA purification.

Approximately 100mg of leaf tissue from -80°C refrigerator was homogenized in 2ml buffer G2 and 4ul RNaseA. 100ul of ProteinaseK was added and the homogenate incubated at 50°C for 2hrs or more till it was clear. The
homogenate was centrifuged for 15 min at 4°C at 14000 RPM. Supernatant from the centrifugate was diluted with an equal volume of buffer QBT and loaded onto a QBT pre-equilibrated Genomic-tip 20/G. The homogenate was allowed to drip through the column under gravity. Multiple washes of the column with buffer QC were done, each time 1ml of the buffer used to drip through, removing cellular contaminants. DNA was eluted from the gravity column using 1ml of buffer QF. Since QF is a high salt buffer, the eluate was subjected to alcohol washes. 700ul (0.7 times the volume of buffer QF) room temperature Isopropanol was added and sample centrifuged for 15 min at 4°C at 14000 RPM. Supernatant was removed, taking care to not to disturb the DNA pellet. 1ml of 70% ice-cold ethanol was added to the pellet and centrifuged at 14000 RPM for 10 minutes at 4°C. Supernatant was removed and the pellet briefly air dried and re-suspended in 100ul of low TE buffer at ph 8, by incubating at 55°C for 2 hrs. This protocol is described in pages 30-31, 44-46 of Qiagen Genomic DNA handbook (Qiagen, 2001).
DNA quantification using picogreen assay for Nucleic Acids on Nanodrop 3300 Spectrometer.

Invitrogen’s Quant iT™ Picogreen dsDNA Assay kit was used for the \textit{L.pardalinum} DNA quantification. Kit components include 20X TE buffer, 200X Picogreen dye and dsDNA of known concentration.

Standard Curve generation: The dsDNA of known concentration from the kit was serially diluted over a range of decreasing concentrations in 1X TE (1:20 dilution). dsDNA dilutions were incubated with equal volumes of 1X picogreen (1:200 dilution) for at least 5 minutes at room temperature, covered with a foil. A few ul of 1X TE incubated with equal volume of 1X picogreen was used as negative control. Nanodrop 3300 software was used, with the picogreen protocol for nucleic acids. 1X TE was used as blank measurement. 2ul of each dilution (numbered as a distinct standard on the assay program) was used for the reference standard tab on the program. 3-5 measurements were collected for each dilution. The standard curve generated by the program was checked for a good \( R^2 \) value (\( = \) or > 0.99), for an acceptable standard curve.

\textit{L.pardalinum} DNA quantification: The standard curve generated by using the known dsDNA standards was loaded on the program. 1 ul of DNA extracted by the Genomic tip/20 column based purification was serially diluted in 1X TE
buffer, for a 1:10, 1:100, 1:1000 and a 1:10000 dilutions of the original DNA. Each dilution was incubated with equal volume of 1X picogreen. 2ul of each dilution mix was placed on the pedestal of the Nanodrop, and the concentration measured against the known standards. Approximately a 10 fold gain in fluorescence (RFU) as well as DNA concentration as the dilution of DNA decreases 10 fold in the measured sample, was an indicator of the correctness of the quantification procedure.

**Accessing DNA quality using BioRad’s Pulsed Field Gel Electrophoresis.**

Pulsed Field Gel Electrophoresis procedure was conducted to assess the fragment lengths of DNA in the extract to determine if the extracted DNA had high molecular weight DNA, greater than 50KB, suitable for cloning. Reagents used for this procedure were 10X TBE (Tris-Borate EDTA) buffer, Pulsed Field Certified™ Agarose (Bio-Rad, CA), Incert® Agarose (Lonza, ME), distilled water, Promega Pulsemarker or NEB Low Range PFG Marker, 1KB ladder (Invitrogen or NEB or Promega).

**Preparation of Incert Agarose – L.pardalinum gel plug:** A 1% Incert agarose gel solution was prepared in 1X TE buffer, the mix microwaved for few seconds a time until the agarose completely dissolved in the buffer and the solution cooled to around 50°C. Equal volumes of this pre-warmed incert agarose and
L. pardalinum DNA were mixed with wide bore pipette tips (up to a maximum of 80ul), loaded into plug mold and set on ice for at least 30 min, to make an agarose/DNA plug.

PFG Electrophoresis: 125ml of 10X TBE buffer (108g Tris Base, 55g Boric acid, 40 ml O.5M EDTA, pH 8.0) was mixed with 2375 ml of distilled water for 2.5 liters of 0.5X TBE. 2 liters of the buffer was poured into the gel apparatus, the cooling chamber’s temperature set to 14°C. 1% agarose gel was prepared by boiling 1 gram of Bio-Rad’s PFGE Certified agarose in 100ml of the 0.5X TBE, and poured into the gel cassette. A half DNA/agarose plug was placed into each well of the gel, and sealed with 1% LMP agarose. The gel was subjected to a Pulsed Field Gel Electrophoresis procedure using the following conditions: Initial and Final switch time of 10sec, voltage of 6V and a runtime of 16hrs. Low range PFG marker and 1KB ladder were used. After the electrophoresis the gel was stained in Ethidium Bromide, de-stained and the gel picture was used to verify that fragments larger than 45KB were present in the sample, so it could be used for further processing.

**DNA end repair using Ion Torrent’s DNA End Repair Kit.**

Extracted L. pardalinum DNA was subjected to end repair to obtain uniform, blunt ended DNA fragments for ligation with HindIII adapters. Ion
Torrent’s Ion Fragment Library Kit and Agencourt AMPure XP kit were used for the procedure. Reagents: 5X End Repair Buffer, End repair enzyme, AMPure beads, 70% ethanol and low (1X) TE buffer. The DNA end repair protocol from Ion Fragment Library kit User guide, page 11 (Ion Torrent, CA) was used and set up as follows: 40ul of 5X end repair buffer and 2ul of end repair enzyme were added to L.pardalinum DNA and the reaction mix brought up to 200ul volume with nuclease free water, to be incubated at room temperature for 20 minutes. 360ul of room temperature Agencourt AMPure beads solution was added and the sample mixed on a rotator at 8-10rpm, at room temperature for 10 minutes. The solution was placed on Dynamag magnetic rack, and the supernatant discarded. Two washes of the DNA beads were made with 500ul each of 70% ethanol. Finally air-dried DNA on the beads was eluted with 50ul Low TE.

**HindIII adapter ligation to end repaired DNA using FastLink ligation kit.**

Reagents from FastLink ligation kit were 10X Fast-Link ligation buffer, 100mM ATP and Fast-Link Ligase. In order to create cohesive ends on the L.pardalinum insert DNA suitable for ligation with the CopyControl HindIII vector, HindIII/EcoRI adapters from Genelink (Hawthorne, NY) were used to conduct a blunt ended ligation. End repaired L.pardalinum DNA was incubated with 1-2ul of Genelink HindIII/EcoRI adapters and DNA grade water for a
reaction volume of 87ul, at 55°C for 10 minutes and then cooled to room temperature for 15 minutes. To this solution, 10ul of 10X Ligation buffer, 2ul of Fast Link ligase and 1ul of 100mM ATP were added and the reaction mix incubated at 16°C overnight. Ligase action was deactivated at 65°C for 15 min.

3% agarose Gel Electrophoresis to purify extract/insert of un-ligated adapters.

During the blunt ended ligation of adapters with the L.pardalinum DNA, due to the low ligation efficiency, or due to abundance of adapter molecules in the ligation mix, there could be excess adapters left after the ligation procedure. These could ligate to CopyControl vector molecules, to create circularized vectors lacking the L.pardalinum insert. A 3% agarose gel electrophoresis was therefore conducted after adapter ligation, for 10 min at 60V, in order to run off the un-ligated, excess adapters in the ligation mix. First, the adapter ligated DNA was combined with equal volume of 1% incert agarose (in low TE) and poured into plug molds, and set on ice for 1 hour. The plug was placed into the well of a 3% agarose gel (made in 1X TAE buffer), and the gel run conducted at 80V for 15 minutes. After the run, the gel was cooled in 4°C and the plug excised from the gel. A gel picture was taken with Chemidoc to confirm the removal of small sized fragments from the DNA-adapter plug.
Insert-vector cohesive ended ligation using FastLink ligation kit.

The pCC1BAC (Copy Control) vector from EpiCentre (Madison, WI) is dephosphorylated and linearized at specific restriction enzyme recognition sites - for HindIII, EcoRI and BamHI, by the manufacturer. The vector has HindIII recognition sites at the end of the linearized vector sequence, which could be ligated to the HindIII adapter ends of the *L. pardalinum* DNA, for a cohesive ended ligation. Adapter ligated *L. pardalinum* plug was incubated with 1-3ul of the BAC vector at 55°C for 10 minutes, along with sufficient amount of DNA grade water for an approximate final volume of 87ul. The mix was cooled at room temperature for 15 minutes. FastLink ligation kit (EpiCentre, Madison, WI) was used, 10ul of 10X ligase buffer, 2ul of FastLink ligase and 1ul of 100mM of ATP were added and the reaction mix of 100ul was incubated at 16°C overnight. Ligation was then stopped by incubating the reaction mix at 65°C for 15 minutes. The plugs containing putative insert-vector (recombinant) plasmids could then be used for the electroporation procedure.

GELase treatment to digest agarose plugs and desalting protocol.

Agarose is a polysaccharide polymer, made up of repeating units of D-galactose and 3,6-anhydro-L-galactopyranose, linked by alpha- and beta-glycosidic bonds. Agarose polymer chains are helical in structure and the rigidity
of its polymer structure depends on the concentration of agarose (Stephen and Phillips, 2006). Solid agarose has a mesh-like structure held together by hydrogen bonds and formed by the polymer chains. The pores within the mesh structure enable movement of biomolecules such as nucleic acids, proteins, etc. In the case of the vector ligated *L. pardalinum*, the recombinant DNA is trapped within the incert (low melting point) agarose gel plugs, and not accessible for electroporation. Therefore, the agarose plug was enzymatically digested using the enzyme GELase, to break down agarose into its component oligonucleotides. For the plug digestion, the High Activity protocol (page 5, handbook) was used. The gel plug were weighed and then incubated for 1 hr with 1X GELase buffer (diluted from 50X, by combining 1:50 ratio of GELase stock buffer and nuclease free water), such that the volume of buffer added was 3X the weight of the gel plug. The plug was thoroughly melted at 70°C, and then equilibrated to 45°C. 1.2U of the enzyme was added and the reaction mix incubated at 45°C for 30 minutes or more, until the plug became completely molten. The reaction product was then desalted 1 hr on ice, in agarose cones prepared with 50% glucose and agarose mix in 1X TAE. The glucose-agarose in TAE was heated till dissolved, and cooled to around 55°C. It was poured into a 2ml microcentrifuge tube, and a smaller 0.5ml PCR tube was placed into the mix. The agarose mix was allowed to
solidify around the inner PCR tube, which was removed to create an agarose cone (CopyControl BAC Cloning Kit, Epicentre). The desalting procedure was conducted to ensure that the salt content within the GELase reaction mix did not interfere with electroporation, by causing arcing.

**Results**

**Isolation of whole nuclei, In-plug extraction and HindIII digestion of DNA from Asiatic lily leaf tissue:**

Approximately 2g of fresh Asiatic Lily leaf tissue was chopped in Nuclei Isolation Buffer, as per the whole nuclei isolation, DNA extraction and purification in plug protocols (Appendix, pages 90-93). A half of a treated plug was incubated with 1ul of HindIII, 10ul of NEB buffer 2 at 37°C for 15 min and the reaction was stopped by incubating the mix at 80°C for 20 min. The mix was placed on ice, and the re-formed gel plug was used to run on a 2% Agarose/TAE gel for 20min to remove small molecular weight DNA fragments from the plug. The plug was excised from the well, and the gel was stained in Ethidium Bromide for 10 min, and de-stained for 45min. The gel picture was taken using Chemidoc. Based on the gel picture (Fig. 1), it was deduced that small DNA fragments approximately in the size range of 0.5Kb to 1KB were removed from
the Asiatic Lily DNA plugs. A Pulsed Field Gel Electrophoresis (PFGE) was conducted on the plugs as per the protocol listed on page 15-16.

Fig. 1: Post-enzyme digest electrophoresis of Asiatic Lily DNA plugs.

Fig. 2: PFGE of Asiatic Lily plugs purified of smaller fragments
As seen from the gel picture in Fig. 2, only lane 5 showed very slight traces of DNA in a smear ranging between the sizes 23.1KB and about 145KB. Although the molecular weight of the purified DNA was in the acceptable range, the amount of DNA that would be available for BAC library prep was minimal and insufficient. There seemed to be insufficient DNA remaining after the purification process conducted as described in the Whole nuclei isolation protocol (Zhang, H, 2000).

First Extraction of genomic DNA from leaf tissue of *Lilium pardalinum* using Qiagen’s Genomic Buffer Kit and Genomic Tip 20/G.

Approximately 100mg of *L.* pardalinum frozen leaf tissue from -80°C was ground inside a tissue grinder/homogenizer, in 2ml of Buffer G2 and 4ul of RNaseA. The homogenate was incubated and loaded on pre-equilibrated Genomic-tip 20/G, and the DNA was extracted as per protocol listed on pages 12-13 of this document, into 1ml buffer QF, twice. 2ul of the Extract 1 in TE buffer was retained for picogreen quantification and 5ul each of eluate 1 and 2 were loaded in wells of a 0.4% Agarose/1X TAE gel for electrophoretic analysis (Fig. 3), compared against a 1KB ladder.
Fig. 3: *L.pardalinum* DNA (Extract 1) on a 1% TAE/Agarose gel.

It is seen from the gel picture in Fig. 3, that the lane 2 with eluate 1 shows presence of more DNA than eluate 2. Both eluates seem to be of DNA greater than 10KB in size. 2ul of the eluate 1 (Extract 1) DNA was used to conduct Picogreen quantification assay, according to the protocol listed on pages 13-15 of this document. An average reading of 3975ng/ml was recorded, indicating that approximately 0.4ug of *L.pardalinum* DNA was present in the 100ul of TE. For further analysis of fragment size, 80ul of the Extract 1, eluate 1 was poured into plugs and subjected to PFGE, using Pulsemarker DNA ladder.
It was observed that the extracted DNA was a bright smear of size 145KB or smaller, as compared to the ladder (Fig. 4). The region corresponding to 48.5 KB size of the ladder seemed especially bright, indicating that the DNA was concentrated around that fragment size. Although the quality and quantity of extracted DNA was satisfactory, more extractions were conducted to ensure the reproducibility of the results. The remaining Extract 1 (20ul) was poured into a plug and stored at 4°C.

Fig. 4: PFGE for *L. pardalinum* DNA (Extract 1) for eluate 1.
Second Extraction for DNA from *L.pardalinum* Leaf Tissue:

DNA extraction from frozen leaf tissue of *L.pardalinum* was conducted and 20ul of extracted DNA (Extract 2) was combined with equal quantity of 1% incert agarose and poured into plugs. The plug was used to conduct pulsed field gel electrophoresis, with a low range PFG ladder as well as a 1KB+ ladder for reference. 2ul of the extracted DNA was used in a picogreen assay for an average recorded DNA concentration of 7590ng/ml or approximately 0.76ug in the 100ul of TE buffer.

![PFG E for L.pardalinum DNA Second extraction attempt (Extract 2).](image)

Fig. 5: PFGE for *L.pardalinum* DNA Second extraction attempt (Extract 2).
As seen in Fig. 5, there was sufficient quantity of DNA in the size range of 45-145KB extracted that can be used for further purification. Extract 2 was saved at -20°C for further purification procedures.

**Hourly PFGE for DNA purification.**

In order to determine the time required to purify the extracted DNA of fragments lesser than 50KB size, I conducted a PFGE with 1 hour increments. A small sliver of low range marker was added to each well of the prepared 1% ultra-pure agarose gel, and the gel run was conducted at 6V, and initial and final switch time of 10 sec. After 1hr, the run was stopped, and the gel slice including one well was longitudinally cut and removed. The gel run was continued for a second hour, after which another strip of gel with the second well was cut out. The gel run was conducted for seven hours. At the end of seventh hour, the gel slices were stained in Ethidium Bromide and de-stained. The gel picture was taken using Chemidoc.
As seen from Fig. 6, the low range Marker showed the presence of a bright band at the end of 3 hours, indicating that the 48.5KB band of the marker had migrated out of the well and into the gel. Therefore the extracted *L. pardalinum* DNA was to be subjected to 2.5hrs of PFGE, to remove fragments smaller than 50KB sizes.

**PFGE on *L.pardalinum* DNA to remove fragments smaller than 50KB (plug purification):**

*L.pardalinum* DNA Extract 2 was subjected to end repair procedure, described on page 17, poured into agarose plugs, and a PFGE was conducted for 2.5hours. At the end of the run, the top portion of the gel, containing the DNA plugs was cut off and the DNA plug was excised from the well and stored in -
20°C. The rest of the gel was stained in Ethidium bromide, de-stained and the gel picture was taken.

**Fig. 7: L.pardalinum Extract 2 plug DNA purification for 2.5hrs on PFG.**

It was seen that fragment sizes 48.5KB or lower only had moved into the gel after 2.5 hours of Pulsed Field Gel Electrophoresis (Fig. 7). It is therefore assumed that the DNA plug excised from the gel would possibly contain fragment sizes upwards of 50KB. In order to verify that the plugs now contained the desirable size range of DNA, another PFGE was conducted with part of plug purified Extract 2.
Fig. 8: PFGE of *L. pardalinum* Extract 2, purified of smaller fragments.

As seen from Fig. 8 Extract 2 lane, there was only a small amount of DNA left in the plug. This could have been due to improper excision of the plug in the previous plug purification step, or because of loss of DNA during the 2.5hr PFGE. The remaining DNA was mostly present in region between 23.1KB and 145KB. Since the size range was acceptable for cloning, the remaining Extract 2 was retained for adapter ligation and insert-vector procedures. Another DNA extraction was attempted to try to retain more DNA after plug purification.

**Additional DNA extraction attempts:** An additional DNA extraction procedure was conducted and the eluted DNA was saved at -20°C.
Third Extraction of DNA from \textit{L. pardalinum} Leaf Tissue:

DNA extraction from frozen leaf tissue of \textit{L. pardalinum} was conducted as described on pages 12-13, except that a Spectra Mesh Nylon Filter was used to filter the supernatant before loading on the Qiagen column. DNA was eluted in 800ul of pre-warmed buffer QF. 20ul of Extract 3 was combined with equal quantity of 1% incert agarose and poured into a plug. The plug was used to conduct PFGE, along with a low range PFG ladder.

\textbf{Fig. 9: PFGE for \textit{L. pardalinum} Extract 3.}

Sufficient amount of DNA was extracted, as seen in Fig. 9. The brightness of the band centered corresponding to the 48.5KB size of the ladder indicates that majority of fragment sizes were of 50KB or less size. However, since there
seemed to be enough DNA of higher fragment sizes present in the sample, it was
used for further purification steps. 60ul of extracted DNA was end repaired
using the Ion Torrent AMpure beads and Dynamag kit. 2ul of Extract 3 was used
for a picogreen quantification, for an average reading of 8343 ng/ml or ~0.67ug in
the 80ul of extracted DNA.

**PFGE on L.pardalinum DNA to remove fragments smaller than 50KB:**

End repaired Extract 3 was mixed with equal volume of 1% incert agarose
to make plugs and a PFGE was conducted for 2.5hours. The plugs were excised
from the gel and stored at 4°C. The rest of the gel was stained in Ethidium
Bromide and de-stained. The gel picture was taken using Chemidoc.

Fig. 10: PFGE for plug purification of Extract 3.

As seen in Fig. 10, the 48.5KB band of the PulseMarker had moved into
the gel and so it can be deduced that the bands that moved into the agarose gel
after 2.5hrs of PFGE were 50KB as smaller. Unlike the previous attempt, however, ladder sizes that were greater than 48.5KB also had moved into the gel, but a corresponding smear of DNA was not present clearly in the sample lanes. It was unclear if fragment sizes greater than 50KB were still present in the excised DNA plug or not. Nanodrop based quantification was not a suitable procedure for DNA in plug. The plug purified Extracts 2 and 3 were combined (Extract 2/3) to ligate with HindIII adapters followed by ligation with Copy control vector (protocols on pages 18-19), and the plug was saved in -20°C for assessment.

**PFGE to assess DNA quality under storage:**

A PFGE procedure was conducted to check the stability of extracted *L. pardalinum* DNA stored in 4°C and 20°C. Extract 1 plug saved at 4°C, Extract 2 and 3, and the additional extract saved at -20°C, were used for the PFGE. The vector ligated combined extract 2/3 was also run on the gel, and the subjected to a PFGE and the gel was stained in Ethidium Bromide, de-stained and the picture was analyzed (Fig. 11). There was detectable, and useable amount of DNA, as was seen in Fig. 11 in lanes 2, 3, 4 and 6 indicating that the extracted DNA stayed mostly stable at -20°C for up to 4 months. DNA stored in plug at 4°C was lost, as seen from lane 1. It was also seen that there was a very smear present in the lane corresponding to the ligation product, apart from a bright band of vector
molecule of 8KB (lane 5). It was possible that the ligation was successful, but there was low concentration of insert DNA. It could be that there was loss of insert DNA prior to ligation step itself and that there was not sufficient DNA present during insert-vector ligation.

![PFGE gel image](image)

**Fig. 11: PFGE to check DNA fragment size quality for all samples Extract 1 to 3**

The presence of bright bands in lanes 2 and 3 indicate that the loss of DNA in the vector ligated product (Lane 5) was not due to lack of starting material due
to storage, but rather due to loss of DNA in the purification steps. All results from the PFGE point to the conclusion that, although good quantity of HMW DNA was being extracted, it was being lost at one or more consecutive purification or ligation steps.

**Fourth Extraction of *L. pardalinum* DNA:**

DNA extraction from frozen leaf tissue of *L. pardalinum* was conducted as before, except that 2.5ml of buffer G2 and 150ul of Proteinase K were used and Extract 4 was eluted twice in 1ml of pre-warmed buffer QF, and combined into Extract 4, in 100ul TE. 40ul of Extract 4 was combined with equal quantity of 1% incert agarose and poured into a plug. The plug was used to conduct PFGE, along with a low range PFG ladder. Picogreen assay value was an average of 4982ng/ml or ~0.5ug. DNA obtained from the fourth extraction was mostly concentrated between the fragment sizes of 23.1KB and 97KB, as seen in Fig. 12. The remaining DNA extract was poured into plugs, which were run at 60V for 10 minutes on a 2% TAE agarose gel in TAE buffer. This was tested as an alternative to the 2.5hrs of PFGE to remove smaller fragments.
It was not clear from the gel picture (Fig. 13) after purifying the plugs, whether smaller fragments were removed since there was no discernable DNA on the gel, and if bigger fragments were retained or not. It seemed that at least a portion of the DNA extract was not excised from the gel wells.
Fifth DNA extraction from *L. pardalinum* leaf tissue:

*L. pardalinum* DNA was extracted, eluted into 100ul of TE buffer and saved as Extract 5. 20ul of Extract 5 was mixed with incert agarose to make a plug and subjected to PFGE (Fig. 14). A portion of Extract 4 plug after purification was also run on the same gel (lane 4) to assess the fragment quality after plug purification.

![PFGE gel image](image)

**Fig. 14:** PFGE on Extract 5, plug purified Extracts 2 & 3, and 4.

It was seen that the purification procedure conducted on Extract 4, to remove fragments smaller than 50KB had only partially worked, with a smear extending from 5KB upwards. There also seem to be loss of DNA, since the
smear was much fainter than the original DNA extract. It was possible that the DNA was degraded during end repair procedure itself, prior to the plug purification or during the plug purification. Since the end repair of the extracted DNA was critical in order to ligate adapters to insert DNA for vector ligation, it was necessary to confirm the effect of the end repair protocol on the extracted *L.pardalinum* HMW DNA. It was also important to determine and demonstrate if there was loss of extracted DNA during consecutive purification and ligation procedures to determine if the current protocol was not suited for *L.pardalinum* HMW purification.

The entire 80ul of Extract 5 was subjected to end repair procedure and collected in TE. It was poured into plugs and run on 2% TAE gel electrophoresis at 60V for 8 minutes, to remove small sized fragments. Adapter ligation was conducted on the end repair Extract 4 plugs, and they were also run on the gel to purify the plug DNA of the adapters. After the gel run, all plugs were excised, and the gel picture was taken.
Fig. 15: 10 min gel electrophoresis for plug purification of Extracts 5 and 4.

It was seen (Fig. 15) after 8 min, DNA 3KB or lesser had definitely moved into the gel. There was still un-excised plug left behind for Extract 5, which had particularly bright band compared to Extract 4. The gel picture also indicated possible loss of DNA during adapter ligation, based on the band of DNA higher than 3KB in the Adapter ligated Extract 4 lane. No band was detected for un-ligated adapters on the gel picture.

The excised plug for Extract 5 was subjected to adapter ligation and the plug was purified to remove the un-ligated adapters, again by running the plug on a 2% TAE gel at 60V for 8min. The plug was excised, and a portion was then run on a 2% TAE gel, using a 1KB+ ladder as positive control, to assess fragment quality.
Fig. 16: 2% Gel electrophoresis to assess fragment quality of adapter ligated Extract 5.

A very faint band was seen (Fig. 16) higher than 12KB on the gel, indicating that the DNA in the plug was of higher fragment size than 12KB. The entire remaining adapter ligated Extracts 4 and 5 plugs were used for an insert-vector ligation procedure, by warming the plugs in 50°C water bath, and then following the ligation protocol listed on page 20 of this document. The plugs were re-formed from the ligation mix, and both the plugs were once again subjected to a 2% TAE gel electrophoresis for 10 min, to remove any un-ligated vectors from the plugs.
Fig. 17: Plug purification of Extracts 5 & 4, after insert-vector ligation.

The gel picture (Fig. 17) clearly showed a vector band for both extracts, indicating that excess vector was in fact removed from the ligated product. The wells indicated the presence of left over plug DNA, but this picture was not sufficient to indicate if the ligation had taken place or not. The vector ligated plugs were then subjected to GELase treatment/high activity protocol (pages 19-20) to digest the agarose, and recover the plug DNA in the GELase buffer (91ul of buffer was used). After the treatment, 15ul of the samples were each used to run on a 2% gel, with a 1KB ladder.
The gel picture (Fig. 18) showed no presence of DNA, ligated or unligated, indicating that no DNA was recovered from the GELase treatment, or that there was not sufficient DNA prior to the treatment itself.

In order to test if the ligases used for the ligation procedures were functionally active, and to ensure that an inactive ligase were responsible for any potential failure of ligation procedures, the following test was conducted. 2ul of pGem 3Zf+ (1.3ug/20ul) vector was incubated with 2ul of 10X Buffer 2, 15.5ul of DNA grade water and 0.5ul of HindIII enzyme for 2hrs to ensure complete digestion of the pGem vector. Ligation of the linearized vector was conducted using T4 ligase as well as the Fastlink ligase used in the L.pardalinum ligation
procedures. The reaction mix was set up as –2ul of digested vector mix, 2ul of 10X ligase buffer, 15ul of DNA grade water and 1ul of T4 ligase or Fastlink ligase. The tubes were labeled and incubated at 16°C for 5 hours each. At the end of the incubation, a sample of HindIII digested vector, T4 ligase ligated vector, Fastlink ligase ligated vector, and uncut pGEM vector were all subjected to electrophoresis on a 2% TAE gel.

Fig. 19: HindIII digestion and Test ligation of pGem 3Zf+ vector.

The Chemidoc picture (Fig. 19) showed that the vector was linearized indicating proper digestion. The ligated vectors showed very faint bands corresponding to the linearized vector, indicating incomplete ligation of at least a portion of the DNA, as well as presence of nicked DNA, as a band running higher than the uncut plasmid. However, a band corresponding to uncut plasmid was seen for both ligation products, indicating that the ligases were
effective. Combined with the results seen in Fig. 15-18, the failure to purify a vector ligated HMW *L.pardalinum* DNA seems mainly due to loss of DNA during the purification steps.

**Sixth DNA Extraction from *L.pardalinum* Leaf Tissue:**

Due to the potential loss of DNA during different purification steps, I decided to extract higher amounts of DNA to begin with. Another DNA extraction procedure was conducted using approximately three times the amount of starting leaf tissue, 6ml of buffer G2, 12ul of Rnase A and 300ul of Proteinase K. The purified DNA was eluted into 100ul X 3 buffer QF, collected into three tubes. 2ul of each sample was subjected to picogreen assay to obtain 27340ng/ml, 170 ng/ml and 10850 ng/ml or approximately 2.7ug and 1.08ug in the 100ul extracted sample. The second eluate was not used further.

An end repair procedure was conducted on the DNA combined from both eluates of Extract 6. The end repaired DNA was poured into plugs, and run on a 2% TAE agarose gel for 8 minutes. The plugs were excised and the gel picture was taken. A portion of each plug was used to conduct a PFGE, to determine the fragment quality after the TAE electrophoresis. Bright bands higher than 3KB were observed on the gel picture taken after end repair, indicating that smaller fragments were removed from the extracted DNA (Fig. 20 A). Based on the PFGE
picture (Fig. 20 B), there seemed to be enough usable DNA upwards of 23KB, indicating that the plug purification was successful and higher insert sizes were potentially available for ligation with the vector.

Fig. 20: A. 2% TAE gel electrophoresis on end repaired Extract 6 to remove small DNA fragments.  
B. PFGE on Extract 6 plugs, post TAE gel electrophoresis (A).

The entire plug of the remaining end repaired DNA was combined and subjected to an adapter ligation and subsequently a ligation with the vector. A portion of the plug after each procedure was retained and used to run on a PFG, to assess the quality of the fragment sizes after adapter ligation and after insert-vector ligation. The PFGE picture (Fig. 21) for the adapter ligated and vector ligated products showed a degradation of the quality of the end repaired DNA in
terms of concentration as well as fragment size, when compared to the end repaired DNA (Fig. 20 B). Therefore, it seems that the initial DNA extract was being lost during the end repair, and more during the subsequent adapter ligation and insert-vector ligation steps, as was seen with the previous attempt with Extract 5. It is possible that the magnetic beads (AMPure beads) based DNA purification during end repair and adapter ligation was detrimental to HMW DNA and was possibly causing the DNA to shear.

**Fig. 21: PFGE on Pre-ligation and Post-vector ligation Extract 6**
Seventh DNA Extraction from L.pardalinum Leaf Tissue:

In order to ensure reproducibility of the extraction procedure to obtain higher amounts of DNA to begin with (as in the sixth extraction attempt), another DNA extraction procedure was conducted using approximately three times the amount of starting leaf tissue, 6ml of buffer G2, 12ul of RNase A and 300ul of Proteinase K. The purified DNA was eluted into 100ul X 3 buffer QF, collected into three tubes. Plugs were made with the extracted DNA from the three tubes, while 2ul of each eluate was used for picogreen quantification assay, for an average reading of 295.4ng/ml or ~0.03ug for tube 1, and 1437ng/ml or 0.14ug for tube 2 and 34043 ng/ml or 3.4ug. A portion of each plug was retained for a PFGE. The remaining DNA plugs were subjected to TAE gel electrophoresis for 8min., at 60V to remove small fragments. The plugs were excised, and along with the original plugs for Extract 7, were all subjected to PFGE.
Since the picogreen assay had indicated that there was minimum amount of DNA in eluate 1, only one plug was run for verification. The PFG picture (Fig. 22) showed that *L.* pardalinum DNA was mostly concentrated between 48.5KB to approximately 180KB fragment sizes, which was the desired outcome of the extraction procedure. However there seemed to be a significant loss of
DNA in the next step itself, purifying the extract of small DNA fragments. Also, DNA less than 50KB was still retained, so the purification procedure was not fully successful. The plugs purified of smaller fragments were combined by incubating in water bath at 50°C for 15 min., and used for an adapter ligation procedure (page 17).

To check if electro-elution could be used to recover the DNA as an alternative to the GELase treatment (which required incubating the DNA at 70°C for 30 min or more), the adapter ligated DNA was subjected to another PFGE, with the bulk of the DNA plug was run in two lanes – lane 2 and 5 (Fig. 23), and three other lanes 1, 3, 4 (Fig. 23) were run with a small amount of DNA, to be used as marker lanes. After the gel run, the entire strip of gel for lanes 2 and 5 were sliced and removed. The rest of the gel was stained in Ethidium Bromide, de-stained and the gel picture was taken using Chemidoc. The unstained DNA slices were laid back with the rest of the gel, and using the gel picture as reference, the region between 48.5KB and 145.5KB were excised.
Fig. 23: PFGE to excise adapter ligated Extract 7 for electro-elution. Post electro-elution gel picture.

A strip of dialysis tubing (Spectra dialysis membrane, Spectrum) was tied on one end with a clean string. The gel slice from the PFG gel was equilibrated with 1X TAE for 1hr at 4°C, after which it was gently slipped into the tubing. 00ul of 1X TAE was pipetted into the dialysis tubing so that the other end of the tubing just above the slice, was tied with a string to keep the gel slice and buffer secure. This bag was lowered into an electrophoresis chamber filled with 1X TAE buffer, such that the bag was vertically placed, facing away from the wells. The electrophoresis was conducted at 100V for 2hrs, and then the current was reversed exactly for 1min. The power was switched off, and the tubing removed from the chamber. The bag was cut open from the top, and a wide bore pipette
tip was used to remove the buffer from inside. The gel slice was stained in Ethidium Bromide, de-stained and the gel picture was taken (Fig. 23).

75ul of the electro-eluted DNA was used for vector ligation procedure, set for overnight at 16°C. After the procedure, a portion of the ligation mix was poured into a plug, as was a portion of the electro-eluted DNA. A PFGE was run with the samples to gauge the success of both the insert-vector ligation as well as the electro-elution procedures. In lane 2 (Fig. 24), for the electro-eluted DNA, there was a very faint smear of DNA around the 9.42KB ladder band, indicating that any DNA recovered from the electro-elution procedure was less than 10KB size. Lanes 3 & 4 showed very faint bands possibly corresponding with the vector band size. If there was a smear of DNA above or below it, it was not discernable. Even it were, it was not the desired fragment size. Based on the gel picture (Fig. 24), it was deduced that the electro-elution procedure was not appropriate to purify the L. pardalinum DNA from a gel. It also seems that the purification procedures have been detrimental to the fragment quality and concentration for recovering HMW insert DNA required for BAC clone preparation.
FACS Assay Results:

Test Flow Cytometry assay to assess the effect of centrifugation on nuclear membrane integrity.

In order to learn and test the flow cytometry protocol that I created, whole nuclei extracts were prepared from 2-3mg of Prosartes smithii, and assay was conducted as described on page 10 of this document. The fluorescence emitted during the assay was captured in the channel FL3 of the instrument setting. The assay was set to run for 1600 sec.
There were two peaks detected on the FL3/count histogram (Fig. 25). It is expected that the first peak flanking the y-axis represents fluorescence other than from the P.smithii nuclei background, since there is a second distinct peak present on the histogram. In order to determine if centrifugation would alter the presence of the peak or change the number of events recorded, another extraction procedure was conducted on P.smithii leaf tissue. The extract was centrifuged at 3200rpm for 10 minutes, the supernatant was removed and the pellet was re-suspended in 5ml of LB01 buffer. 100ul of Propidium Iodide was added to the mix and incubated for 10 minutes. The mix was then used for the flow assay according to the same parameters used for the un-centrifuged nuclei extract.
It was seen from the histogram (Fig. 26) that the second peak appeared in the same FL3 channel indicating that the fluorescence could be from *P.smithii* DNA. On the Y-axis, for the un-centrifuged sample, there were about 3117 events present in each of the 10 segments on Y-axis. For the centrifuged sample, it would be about 263 events per segment. At this point, there was not much that could be concluded from the results, since it was not clear exactly how many of the events are related to the *P.smithii* nuclei.

In order to test for fluorescence detected for *L.pardalinum*, leaf tissue from *Lilium pardalinum* and *Prosartes smithii* were weighed at equal measures for approximately 300mg of final weight, and nuclei extraction procedure was conducted. Half of the extract was removed and brought up to a 5ml volume with additional LB01 buffer, incubated with 100ul of PI and assayed per the parameters discussed above. The other half of the extract was centrifuged at 100rpm for 5 minutes, the supernatant removed and the pellet re-suspended in 5ml of LB01 buffer. The supernatant was brought up to 5ml volume as well and assayed as well after being incubated with PI. A manual stop time of 503 and 506 seconds was used to stop the assay and compare the results.
Fig. 26: Flow assay to test the effect of centrifugation on *P*.smithii and *L*.pardalinum whole nuclei.

Although the number of events recorded for the un-centrifuged sample was much higher than the supernatant or the pellet values, and the events recorded for sample pellet is much smaller than that of the supernatant, it is not clear how much of the detected fluorescence could be attributed to the nuclei alone. Also, it was seen from the Count vs FL3 histogram that there is a peak showing close to the 1024 channel on the X-axis (Fig. 26). Placement of the second peak was indicative of fluorescence from nuclei of larger genome size compared to *Prosartes*. However, I could not confirm yet that the peak on Count vs FL3 histogram at the lower fluorescence channel was from the *Prosartes* nuclei.
Optimization Assays to Capture Fluorescence from Test Species Nuclei:

In order to optimize the assay and record only the relevant fluorescence from the sample whole nuclei, I set up several linear gates to independently capture the events falling within the limits of those gates. About 300mg of *P*. smithii leaf was used for a nuclei isolation procedure, and the extract incubated with PI was assayed.

Fig. 27: Flow assay to optimize the capture of *P*.smithii fluorescence in specific channels.

By using gates ‘D’ and ‘E’, I tried to capture the specific fluorescence of the second peak in the FL3/Count histogram (Fig. 27) to see if there would be an overlap that could account for the peak. However, there was count for the gated region ‘D’, but none for ‘E’. The count 10574 was not valid, given the broadness
of the gate ‘D’, but helped eliminate the counts recorded in the lower channel numbers. Next, 5ml of *L.pardalinum* nuclear extract was used for the flow assay. A peak flanking the Y-axis in the 1024 channel was noticed (Fig. 28) as in the assays where *L.pardalinum* tissue was used. It was necessary to adjust the cytometer settings in order to center the peak for proper resolution.

![Flow assay to optimize the capture of *L.pardalinum* fluorescence in specific channels.](image)

Also as seen in Fig. 28, there was a count of 4323 recorded for the gated region ‘A’, and no fluorescence recorded for gate ‘E’ and only 1108 for gate ‘D’, unlike in Fig 28, where the second peak was closer to the channel 0 on X-axis. It was deduced that the peak line close to the 1024 channel on the Y-axis must therefore correspond to the florescence from the *L.pardalinum* nuclei. In order to
confirm captures of fluorescence in two distinct channels, about 150mg each of L.pardalinum and P.smithii leaf were used for nuclei isolation procedure and assayed.

Fig. 29: Flow assay to optimize the capture of P.smithii and L.pardalinum fluorescence in specific channels.

As seen in Fig. 29, the total count in gate D seemed to be from two separate events. 464 were from the region gated ‘A’ and the rest mostly from the second peak closer to Y-axis in the region gated ‘D’. It was essential to further optimize the cytometer setting so as to be able to detect the specific signal related to Lilium nuclei in a manner where fluorescence signal and recorded events in a
Lilium/Prosartes mix was more meaningful and indicative of the actual signal capture.

To ensure that the detected signal and the recorded events were specific to the fluorescence given off from whole nuclei and not from any remnant signals or noise from PI or the LB01 buffer, the following assays were conducted. First, a cleaning/bleaching routine was conducted, and then a regular assay was conducted using the Clenz liquid. The signal from the clenz run was recorded as the following histogram and event data. Gate ‘E’ was set up to distinguish the signal from other potential signals generated. A second assay was conducted, using with 20ul of PI in 1ml of LB01 only.

![Flow assay to assess fluorescence capture due to LB01 and/or PI.](image)

**Fig. 30:** Flow assay to assess fluorescence capture due to LB01 and/or PI.
The gate ‘E’ captured the signal from the assay in both cases. As seen in Fig.30, most of the fluorescence captured when Clenz or LB01 was used was in the region gated ‘E’ only, indicating that it was lower signal intensity. Secondly, the absence of a second distinct peak noticed when P.smithii nuclei were present in the buffer mix, indicated that the second peak was derived from the fluorescence from P.smithii nuclear DNA.

About 300mg of L.pardalinum leaf was used in the nuclei extraction procedure and assayed. It was seen again that the fluorescence captured was from the channel 1024 (Fig. 31), but not resolved optimally, confirming the result previously seen in Fig. 29.

Fig. 31: Flow assay to optimize the capture of L.pardalinum fluorescence in specific channels.
In order to better capture of fluorescence signals on the histogram, about 300mg of *P*.smithii leaf tissue was chopped and assayed as above. After the second distinct peak at the expected channel appeared, the run was paused several times, while the cytometer voltage for FL3 channel was adjusted. At 560V it was observed that the peak was centered in the histogram (Fig. 46). The ‘gain’ was set at 1 for FL1 through FL4 channels, although the used values were recorded for FL3 only. The protocol on the cytometer was re-saved with the adjusted settings. To confirm that the updated protocol gave the desired results, the flow assay was run on *P*.smithii and *L*.pardalinum samples again. First, 300mg of *P*.smithii was chopped as used for nuclei isolation procedure and assayed. After the run, the distinct peak appearing in the first channel was gated ‘A’.
Fig. 32: Flow assay to visualize the capture of *P.smithii* fluorescence in specific channels.

3575 events were recorded for the peak representing the fluorescence of *P.smithii* nuclei (Fig. 32). A second assay was conducted using 300mg of *L.pardalinum* leaf. The region capturing the fluorescence from *L.pardalinum* nuclei was gated ‘B’. 1379 events were recorded for that region. It was seen that unlike the sharp peak detected from *P.smithii* nuclei, the peak from *L.pardalinum* nuclei was broad and wider characteristic of large nuclei (Fig. 33).

Fig. 33: Flow assay to visualize *L.pardalinum* fluorescence in specific channels.

A third assay was conducted using the same protocol, with about 150mg each of *L.pardalinum* and *P.smithii* leaf tissue. Gate ‘F’ was set to cover the two
peaks on the FL3/Count histogram, and it showed two well resolved peaks, once in the first channel of the FL3 channel, and another in the second-to-last FL3 channel (Fig. 34). 2365 count was measured for gate ‘A’ for P.smithii, 1254 for gate ‘B’ for L.pardalinum. Compared to the L.pardalinum standalone assay, where 300mg of leaf tissue was used, this represents an increased number of nuclei.

Fig. 34: Flow assay to visualize P.smithii and L.pardalinum fluorescence in distinct channels.

This could be due to better chopping technique, or that the L.pardalinum tissue used had more number of nuclei.

Flow Assay to Assess the effect of centrifugation on L.pardalinum nuclei:
In order to determine if centrifugation has any effect on the membrane integrity of *P*.smithii and *L*.pardalinum, the following assays were conducted. 300mg of *P*.smithii leaf was chopped in 5ml of LB01 buffer, out of which 2.5ml was incubated with 50ul of PI for 15 min., and assayed.

Fig. 35: Fluorescence captured from *P*.smithii nuclei prior to centrifugation.

A distinct peak was seen in the expected FL3 channel gated ‘A’, with a count of 8403 events (Fig. 35). The remaining 2.5ml extract was centrifuged for 3min., at 3200rpm at 4°C, the supernatant was removed except for 100ul, and the pellet was re-suspended in 2.4ml of fresh LB01 buffer. This sample was incubated with 50ul PI for 15 min and assayed.
Fig. 36: Fluorescence captured from *P*.smithii nuclei after centrifugation.

It was observed that only 929 events were recorded for the region gated ‘A’ (Fig. 36). The peak position corresponded exactly with the expected *P*.smithii peak, but the count reduced by about 9 times. This assay was repeated with 300mg of fresh *L*.pardalinum leaf tissue.

Fig. 37: Fluorescence from *L*.pardalinum nuclei prior to centrifugation.
Gate ‘B’ that was set based on earlier *L. pardalinum* assays did not capture the fluorescence. Instead it was seen in a channel prior to the expected channel (Fig. 37). It was unclear why the shift in the fluorescence occurred. The count of 4391 obtained from the gate ‘D’ was therefore only an estimate, since it includes the signal captured from other FL3 channels as well. The remaining 2.5ml of extract was centrifuged at 3200rpm for 3 min at 4°C, supernatant removed except for 100ul. The volume was brought up to 2.5ml with fresh LB01 buffer, incubated with PI and assayed.

![Fluorescence from L.pardalinum nuclei after centrifugation.](image)

**Fig. 38: Fluorescence from *L.pardalinum* nuclei after centrifugation.**

There were only 238 events recorded in region gated ‘B’, but the signal was captured in the correct channel (Fig. 38). The supernatant was not assayed in this
procedure. It was possible that many nuclei had not settled in the 3min centrifugation and were removed along with the supernatant. On the other hand, it was also possible that the high centrifugation speed of 3200rpm was detrimental to the nuclei membrane integrity, but they might withstand lower speeds. To test these variables, another assay was conducted. First, 300mg of fresh *L. pardalinum* was used to conduct nuclei isolation protocol in 5ml of LB01 buffer. 2ml of the extract was incubated with 40ul of PI and assayed. 2937 events were recorded for region gated ‘D’ (Fig. 39). The gate set as ‘B’ only captured small portion of the peak.

Fig. 39: Fluorescence from *L. pardalinum* nuclei prior to centrifugation.
2ml of the remaining extract was centrifuged at 100rpm at 4°C for 5min, and supernatant except the last 100ul was removed. The pellet was re-suspended in fresh LB01 for a final volume of 2ml. 20ul of PI was added to this mix, incubated and assayed.

Fig. 40: Fluorescence from *L.pardalinum* nuclei pellet after centrifugation at 100rpm.

The count for the number of events detected was 447 (Fig. 40), indicating that the nuclei were either present in the supernatant as whole or lysed nuclei. The supernatant was centrifuged at 3200rpm for 3 min at 4°C. The supernatant from this centrifugation was removed except for the last 100ul. The pellet was re-suspended in LB01 buffer for a volume of 2ml, incubated with 40ul of PI and assayed. The number of events detected for this assay were 296 (Fig. 41),
indicating that there were a portion of nuclei left in the supernatant of the first centrifugation, but majority of the nuclei from the extract were lost. The supernatant from this second (3200rpm) centrifugation was also brought up to 2ml volume, incubated with 40ul of PI and assayed. It was seen that there was only a count of 18 detected from this centrifugation (Fig. 42). It was also noted that in spite of 3200rpm centrifugation, there were still about 300 nuclei detected in the pellet (the 100rpm centrifugation resulted in 447 counts). The time duration for which the nuclei are centrifuged might also add to the effect of centrifugation speeds.

Fig. 41: Fluorescence from pelleted nuclei after centrifugation at 3200rpm.
Fig. 42: Fluorescence from supernatant after centrifugation at 3200rpm.

This assay was repeated two more times (Appendix, pages 94-96), with the L.pardalinum nuclear extract -prior to centrifugation, the pellet after 5min of centrifugation at 100rpm re-suspended in LB01 buffer and the supernatant of the centrifugation-to ensure that the results can be reproduced. It was observed in all the three attempts that the number of nuclei significantly reduced upon centrifugation. The total count of nuclei from the pellet and supernatant fractions combined together was still much lower than the non-centrifuged sample, in each instance. Less than 3min of centrifugation at 100rpm, however seem to result in higher numbers of nuclei detected in the supernatant.
Flow Assay to Determine the Effect of Incubation Time on Nuclei Integrity:

The above procedures required for the nuclear extracts/centrifugates to be saved on ice for over an hour before the flow cytometry assay. To determine if the lower counts detected during the assay was due to the nuclei degrading on ice after extraction and not because of the centrifugation itself, the following test was conducted. 300mg of fresh *L. pardalinum* leaf tissue was used for whole nuclei extraction procedure. 3ml of the extract was divided into three flow assay tubes, 1ml into each tube, and incubated with 20ul of PI each. The first tube was assayed after 15 min of incubation, the second tube was incubated over ice for 1:15min and assayed, and the third tube was assayed with further 1hr retention on ice. FL3 vs Count was compared in all instances.

![Fluorescence graph](image)

*Fig. 43: Fluorescence from *L. pardalinum* nuclei after 15min incubation with PI.*
Fig. 44: Fluorescence from *L. pardalinum* nuclei after 1 hr incubation with PI.

Fig. 45: Fluorescence from *L. pardalinum* nuclei after 2 hr incubation with PI.
The counts of nuclei were measured for the region gated ‘G’, rather than ‘B’, since the fluorescence was detected in different channels for different incubation periods. For 15min incubation with PI, 1307 nuclei were detected, (Fig. 43) whereas for the sample assayed after one hour after that, 1788 nuclei (Fig. 44) and for 2hr incubation 1303 nuclei were detected (Fig. 45). After the 1hr incubation period, the fluorescence was detected in the channel that was expected, based on the estimated C-value of L.pardalinum (Fig. 44). After 2hrs, the number of nuclei started reducing, and the signal was also in a lower channel than expected (Fig. 45). The shift in the fluorescence was attributed to a machine error, with all experimental conditions being identical as in other flow assays. It was deduced that for the given experimental conditions, protocol and cytometer settings, up to 2hrs of incubation of L.pardalinum nuclei on ice was not detrimental to the membrane integrity of the nuclei. This assay also led to the conclusion that the reduced number of nuclei observed in the centrifuged samples might not be due to longer incubation periods, but rather due to the effects of the centrifugation itself, on the whole nuclei.
Discussion

Genome sizes of species in Liliaceae are highly varied, and huge genome sizes are highly prevalent in this family. As a comparison, the fully sequenced rice genome is of 400-430 MB and wheat is 17GB, but hexaploid. In contrast, most diploid family members of Liliaceae with large genome sizes show a range of 3GB to 127GB. However, similar to the rice and wheat genomes, the nuclear DNA of the Liliaceae plants has transposable element content which increases the sequence redundancy even as it adds to the genome size. Unlike rice and wheat, whose genomes have been sequenced with global collaborative research efforts, whole genome sequencing of the lily genomes becomes further complicated by the lack of a reference genome. Even with the current availability of high throughput sequencing technologies, and powerful sequence assembly programs, sequencing whole genomes of the Liliaceae plants is still highly cost intensive. Until such advances are made that would drastically reduce the sequencing costs, BAC clone based hierarchical sequencing still is the sequencing method of choice.

High quality, high molecular weight DNA for construction of BAC clones is necessary to provide meaningful coverage for sequencing. There are several research projects that have successfully created BAC clones with megabase
lengths of plant DNA inserts, and one such protocol developed by Dr. Hongbin Zhang was used for BAC clone creation in our lab with mixed results at best. Either due to inherent instability of the large lily nuclei or other reasons not known yet, it was seen that there was a tremendous loss of intact nuclei and high molecular weight DNA in the procedures used so far in our lab.

This thesis research aimed to develop an improved protocol for purification of sufficient quantities of high molecular weight DNA of *Lilium pardalinum* to be cloned into BAC vectors for sequencing. Qiagen’s Genomic tip-20/G - gravity flow, anion exchange column based extraction of genomic DNA of *L. pardalinum* appeared suitable to be used successfully to purify up to 200KB fragments, as shown in the multiple DNA extraction attempts by this researcher. This extraction procedure reduced the amount of starting material significantly, since it was seen that between 100-300mg of *L. pardalinum* leaf tissue was sufficient to extract between 0.5-3.4ug of DNA with proper extraction technique. It was seen that high molecular weight DNA thus extracted was stable at -20°C, for up to 3 months at least. Conversely at 4°C, it was seen that the HMW DNA was not stable even when poured into plugs. Although the extraction protocol used was sufficiently successful, the challenge was to retain the HMW DNA through the purification steps, after the extraction. The key purification steps
involved end repair of the extracted DNA to create blunt ends of fragments, ligation of the extract with HindIII adapters and ligation of the modified insert/extract with the HindIII compatible BAC vector. It was seen several times that just subjecting the HMW DNA plugs to magnetic bead based end repair and gel electrophoresis purification was sufficient to damage the fragment quality of the extracted HMW DNA. The loss of DNA at every purification step could be due to mechanical shearing of the DNA, resulting in much smaller sized fragments which ran out of the plug during electrophoresis. The remaining plug was therefore becoming devoid of usable DNA after every step. It might therefore be beneficial to conduct the end repair enzymatic step, but not the magnetic bead based clean up. The DNA extract-adapter ligation step also involves magnetic bead based cleanup, which again may have resulted in further shearing of the DNA. Size exclusion column (gel filtration) based desalting and elution of HMW DNA could potentially be a viable substitution as a purification procedure to the magnetic bead based clean up. Also, BluePippin (Sage Sciences, Beverly, MA) based size selection of DNA upwards of 50KB might be suited for purifying HMW DNA from lily leaf tissues, for construction of BAC clones for sequencing.
Another specific aim of this research was and to investigate and demonstrate the effect of centrifugation on the membrane integrity of *L. pardalinum* whole nuclei. Through the flow cytometry assay, it appeared that there was loss of up to two-thirds of the nuclei with minimal centrifugation, at 100rpm for 3 min. Multiple centrifugations and higher speeds of centrifugation could therefore affect the integrity of the lily whole nuclei membranes and possibly caused lysis of whole nuclei, as was seen in the drastically reduced number of nuclei detected in the cytometry assay. It was also seen during the cytometry assay, that a longer incubation periods (up to 2 hours) of the nuclei with the DNA binding Propidium Iodide did not adversely affect the detection of the fluorescence, and also the numbers of detected nuclei. Surprisingly, the nuclei were stable on ice for up to 2hrs, contrary to earlier expectation that the nuclei broke apart quickly after the isolation procedure. Whether due to the large size of the *L. pardalinum* nuclei or other yet unknown factors, whole nuclei did not seem to stand up well to the centrifugation.

It could be reasonably concluded that the gravity flow, ion-exchange column based DNA extraction seemed a viable alternative to the whole nuclei based HMW DNA extraction. With suitable changes made to the purification steps of the extracted DNA, it is possible to construct BAC clones from
L. pardalinum DNA which could be further used for DNA sequencing and other downstream applications.

References


Appendix

Flow Cytometry assay on nuclei extract with Propidium Iodide fluorophore

Flow Cytometry is a fluorescence based analytical technique in which microscopic particles (cells, nuclei etc.) tagged with a fluorescent dye are forced to flow one particle at a time through a stream of liquid, and as a laser of light is passed through the fluid the resulting fluorescence from the dye as well as the scattering of light due to interruption from the particles are detected and converted into electric signals. These signals are then relayed as electronic data (Shapiro, 2003). The light that is scattered off of particle’s surface is gathered as forward scatter, indicating the size, whereas the light that perpendicularly passes through a particle is gathered as side scatter, indicating the complexity of composition of the particle. The fluorescence given off from the dye bound to the particle is also measured in a specific channel of the cytometer, based on the emission maximum of the dye. For the cytometry assay, the dye used is Propidium Iodide, which is a membrane impermeable dsDNA binding fluorophore. It is an intercalating dye, showing little base preference and has shown to provide consistent nuclear genomic content estimation when compared to other evaluation methods (Dolezel et al., 1992). When bound to nucleic acids,
its fluorescence increases 20-40 fold, and its absorbance and emission maximum are 535nm and 617nm respectively. Since it binds to RNA as well, the nuclear extract needs to be treated with an RNase. Fluorescence emitted is collected in terms of recorded events, as well as one and two parameter histograms. The placement of the fluorescence peaks and number of recorded events are combined and used to confirm the genome size and the number of intact nuclei.

**Gravity Flow Column Based DNA extraction and purification, with Qiagen’s Genomic DNA kit and Genomic tip 20/G:**

DNA extraction protocols such as the Spin column protocol are efficient and suitable, when the size of the extracted DNA is not of prime consideration. However, the centrifugation causes mechanical shearing of the DNA and hampers the quality of DNA when big fragment sizes are needed. For *L. pardalinum* DNA extraction, Qiagen’s Genomic DNA Buffer Set as well as Genomic-tip 20/G was used. Genomic-tip 20/G has a gravity flow column which is composed of anion-exchange resin. Under low salt conditions provided by the lysis buffer used for leaf tissue lysis, DNA binds to the resin, while the cellular contaminants are removed with subsequent washes with medium salt buffers, with no centrifugation involved. DNA is collected using a high salt buffer, and then precipitated with alcohol (Isopropanol and ethanol) washes and stored in
Tris-EDTA buffer. Since the cellular lysate containing the DNA flows through the Genomic tip 20/G under gravity, the mechanical shearing the DNA might be subjected to is lower, when compared to the spin column based extraction. This helps in purifying higher molecular weight DNA from the cell lysate. Up to 20μg of DNA, with fragment sizes up to 150KB can potentially be extracted using the Genomic tip 20/G, with a few milligrams of starting tissue.

**Pulsed Field Gel Electrophoresis (PFGE):**

Standard agarose gel electrophoresis is insufficient to resolve and visualize high molecular weight DNA. During standard gel electrophoresis, DNA molecules greater than 20KB show similar mobility and tend to stack as a single diffused band, instead of separating into individual bands. By using the technique of Pulsed Field Gel Electrophoresis (PFGE), DNA up to 1600 kb sizes can be resolved (Ozdemir, 2000). During this electrophoresis, a small voltage is applied to the moving DNA from different directions, one direction at a time for a specific duration. This forces the DNA to change its orientation each time, under the influence of the direction of the applied voltage, causing the smaller bands to move in that direction faster than the larger molecules, which therefore tend to get separated from the smaller molecules (Joppa et al., 1992)
0.5X Tris-Borate-EDTA is the choice of buffer for PFGE due to better buffering capacity, and the relatively slower movement of DNA in TBE gel than in a TAE gel. The molecular weight ladder used was either Promega’s Pulse marker or New England Biolabs’ Low Range PFG Marker, based on availability. Both the markers have lambda phage DNA – HindIII fragments, 100 base pairs to 200KB long, embedded in 1% LMP agarose. The fragment size of 48.5KB within the marker shows as the brightest band, and can be used as a reference band size (neb.com product factsheet). A second 1KB ladder, with a highest band of 10KB size was also used for cross reference, since the PFG marker has concatemers which present as additional bands on the gel. The gel electrophoresis was conducted for up to 16hrs on Bio-Rad’s FIGE mapper cell, in which the DNA will be subjected to an 180° orientation change, causing the DNA to move backwards for certain duration.

Picogreen based DNA quantification:

Picogreen is a dsDNA binding fluorophore, preferentially binding to A-T rich regions. It has an excitation maximum of 480nm and an emission maximum of 520nm (Ahn et al., 1996). Its fluorescence increases more than a 1000 fold when bound to dsDNA and is used to quantitate dsDNA present in a given sample
(Singer et al., 1996). Unbound picogreen in a sample solution exhibits minimal fluorescence. Unlike UV absorbance based measurement, which can include single strands of DNA, picogreen is specifically bound to dsDNA and provides for a better quantitation of the DNA concentration in a sample solution. Nanodrop 3300 Fluorospectrometer from Thermo Scientific was used for the L. pardalinum DNA quantification procedure, since only 1ul of sample DNA is required for use.

**DNA End Repair for adapter ligation:**

During DNA extraction, mechanical shearing of the fragments causes the double stranded DNA to not have perfect complementarity at the fragment ends. There could be 5’ or 3’ overhangs, creating uneven strands. A DNA end repair procedure was conducted in order to prepare the extracted DNA for ligation with the BAC vector. During the end repair, the sheared ends of the DNA fragments are converted to blunt ends, using an end repair enzyme which fills in for a 5’ overhang and removes a 3’ overhang. Typically T4 DNA polymerase is the enzyme of choice due to its 5’-3’ synthesis as well as 3’-5’ exonuclease activity. It lacks 5’-3’ exonuclease activity (Sambrook et al., 2001, ThermoScientific literature for T4 Polymerase). The pCC1BAC vector used has
HindIII recognition sequences that necessitates for the inserts (extracted DNA) to have HindIII compatible ends for successful insert-vector ligation. One method to prepare the extracted DNA for this ligation would be to digest it with HindIII restriction enzyme which creates the HindIII compatible ends, but the resulting fragments would be varied in sizes. Since the desired insert fragment size is 50KB or above, blunt ending helps in two ways – it minimizes the manipulation of extracted DNA to preserve the fragment size, and prepares the DNA for blunt ended ligation.

Agencourt AMPure XP Kit from Beckman Coulter is a DNA purification system that utilizes paramagnetic beads to bind to and purify nucleic acids that are greater than 100bp long. DNA fragments that are present in the end repair reaction mix can be bound to the Agencourt AMPure beads and segregated from the rest of the reagents in the solution by magnetizing the beads on a magnetic rack. Dynamag™-2 (Life Technologies) was used for this procedure to purify the DNA. Once demagnetized, the DNA goes back into solution and can be eluted using a low salt buffer such as low Tris-EDTA.

In order to create compatible ends on the insert DNA fragments to facilitate ligation with the pCC1BAC HindIII vector, adapters from Genelink (Hawthorne, NY) were used to ligate to the blunt ended DNA fragments. Hind
III/Eco RI adapters, with the oligonucleotide sequence 5’ -AGCTTGAATTC- 3’ and 3’ -ACTTAAGp -5’ on the complementary strands were used. The adapter has a blunt end at 5’ on the 5’-3’ strand, and a phosphate on the 5’ end on the 3’-5’ strand. The adapter ligates to the DNA insert fragment at the blunt end, leaving 5’ and 3’ overhangs compatible with HindIII sites on the vector.

**Blunt ended and Cohesive ended DNA ligation:**

DNA ligation occurs when two DNA strands are covalently joined, by the formation of phosphodiester bonds on the backbone between the 5’ phosphate of one nucleotide and the 3’ hydroxyl group of an adjacent nucleotide, catalyzed by the enzyme DNA ligase. ATP is utilized for the reaction, during which the ligase transfers and AMP to the 5’ phosphate, resulting in a pyrophosphate bond, as well as a second phosphodiester bond between the 5’ phosphate and an adjacent 3’ hydroxyl group (Lehman, 1974). Although the ligase is most active at 37°C, to facilitate better annealing between the strands to be ligated, the ligation of the L.pardalinum insert and the HindIII adapters and then with the BAC vector, will be conducted at 16°C overnight. Typically ligation of cohesive ended DNA strands (as in the adapter ligated insert DNA and BAC vector with HindIII overhangs) occurs more efficiently and in the presence of less amount of ligase.
For blunt ended strand ligation (insert DNA and HindIII adapters), longer incubation of the ligation mix, as well as greater amounts of ligase would be required. Even so, the ligation efficiency is highly reduced compared to cohesive end ligation. However, a successful blunt ended ligation procedure is of critical importance for facilitating insert/adapter and therefore the insert/vector ligation and subsequent creation of BAC with high molecular weight DNA.

**Whole Nuclei Isolation Protocol for DNA Extraction and Purification:**

The following buffers are prepared for the isolation, DNA extraction and purification procedures

**Homogenization Buffer (HB) – stored at 4°C:**

10mM Tris-HCl pH 8.0 from stock 1M,

10mM EDTA pH 8.0 from stock 0.5M,

100mM KCl (FW 74.55),

4mM Spermidine (FW 145.25),

1mM Spermine (FW 348.2)

Appropriate amounts of salts were weighed or stock solutions pipetted into a conical flask. DI water was added to bring to a final volume of 1L.

**Sucrose Extraction Buffer (SEB) – stored at 4°C:**

Volume/weight needed for a final buffer volume of 300ml
Tris-HCl: 1000mM/10mM=100 dilution factor, 300ml/100=3ml

EDTA: 500mM/10mM=50, 300/50=6ml

KCl: 0.1= x gm/74.55 * 1/3 = 2.24gm

Spermine: 0.001= x gm/348.2 * 1/3L = 104.46mg

Spermidine: .004= x gm/145.25 * 1/3 = 174.3mg

Sucrose: 2M= x gm/342.3= 205.38g * 1/3 = 205.38gm for 300 ml;

To prepare a Nuclei Isolation Buffer (NEB), 4.5 ml of SEB was mixed with 1.5 ml of HB just before the extraction procedure.

Lysis buffer (stored at room temperature): -

0.5M EDTA, 1% N-Lauryl Sarcosine

1L of Lysis buffer was made using 190.1gm of EDTA, 10gm of N-Lauryl Sarcosine in distilled water. 50mg of ProteinaseK (1mg/ml) was added to 50ml of Lysis buffer before use.

TE buffer at pH 8.0 (autoclaved) – stored at 4°C:

Tris base 10mM – 121.4gm, EDTA 1mM – 2ml, Water to volume

1L of this buffer was prepared for washes.
**PMSF treatment reagent:**

50uL of PMSF from stock concentration of 100mM was added to 49.5 ml of TE buffer, for a final concentration of PMSF 0.1mM. This buffer was used for PMSF treatment.

**Procedure:**

Approximately 2gm of washed and dried Asiatic lily leaves were taken in a petri dish with 4.5ml of SEB and 1.5ml of NEB. Before use, 6uL of B-mercaptoethanol was added to the extraction buffer (1ul for 1000ul of buffer). A sterile new blade was used to finely chop the leaves in the chilled buffer. A 12ml tube with mira cloth tied on top was also pre-chilled. After chopping the leaves finely, the leaf/buffer mix was strained through the mira cloth. The solution was centrifuged twice at 100g and finally at 500g. All supernatant except the last 100uL was removed. The pellet was re-suspended using a soft brush within the buffer. Agarose plugs were prepared using the incert agarose melted at 50°C water bath. Approximately equal volume of the agarose was added to the supernatant and pellet, and gently mixed. About 80ul of the mix was added into the plug mold, and set on ice for 30 min. Lysis buffer was poured into a 50ml tube with lid, and incubated at 50°C. Prepared plugs were added into this and left in the water bath for 24hrs. The lysis buffer for the wash was changed for a
second wash, with 50ml of fresh buffer. Plugs were drained into autoclaved cheesecloth, and introduced back into the fresh buffer. This was left at 50°C again for 24hrs. Two PMSF/TE wash treatments were conducted on the plugs on rocking platform at 4°C, each wash for 1hr. This was followed by two TE washes on rocking at 4 degrees, each for 1 hr. For a third wash, the plugs were left in fresh TE buffer overnight, on a rocking platform at 4°C. A fourth TE wash was conducted in fresh buffer under the same conditions as above. These washes were required, in order to remove the PMSF from the plugs. At this point, the plugs can be stored in 0.5M EDTA at 4°C. Just before digestion of the plugs in HindIII, two washes of TE for 1hr was conducted.

FACS Supplementary Data:

Fig. 46: Voltage and Gain adjustments for optimal fluorescence visualization of *P.smithii* whole nuclei.
Fig. 47: Fluorescence from *L. pardalinum* nuclei before centrifugation.

Fig. 48: Fluorescence from *L. pardalinum* nuclei (pellet) after centrifugation.
Fig. 49: Fluorescence from *L. pardalinum* nuclei (supernatant) after centrifugation.

Fig. 50: Fluorescence from *L. pardalinum* nuclei (pellet) before centrifugation.
Fig. 51: Fluorescence from *L. pardalinum* nuclei (pellet) after centrifugation.

Fig. 52: Fluorescence captured from *L. pardalinum* nuclei (supernatant) after centrifugation.