Identification and Characterization of a Novel LYR/LVR Gene Highly Expressed During Embryogenesis in Douglas-fir

by

UMESH RAMACHANDRAN
M.Sc, UNIVERSITY OF MADRAS, 2002

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF BIOCHEMISTRY

in the Department of Biochemistry and Microbiology

© UMESH RAMACHANDAN, 2006
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

SUPERVISORY COMMITTEE

Identification and Characterization of a Novel LYR/LVR Gene Highly Expressed During Embryogenesis in Douglas-fir

by

Umesh Ramachandran
M.Sc, University of Madras, 2002

Supervisory Committee

Dr. Santosh Misra (Department of Biochemistry & Microbiology)

Supervisor

Dr. Chris Upton (Department of Biochemistry & Microbiology)

Departmental Member

Dr. Peter Constabel (Department of Biology)

Outside Member

SUPERVISORY COMMITTEE

Supervisor: Dr. Santosh Misra (Biochemistry & Microbiology)

Departmental Member: Dr. Chris Upton (Biochemistry & Microbiology)

Outside Member: Dr. Peter Constabel (Department of Biology)

Abstract

In order to elucidate the molecular and biochemical events occurring in embryogenesis in Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco), an essential gene expressed highly during early embryogenesis was identified, cloned and further characterized in this study. Douglas-fir LYR/LVR cDNA was obtained using RT-PCR with specific primers, followed by cloning and sequencing. Northern blot analysis showed higher amounts of LYR/LVR transcripts in early-cotyledonary embryonic stages and megagametophytes when compared with mid- and late-cotyledonary embryos. LYR/LVR transcript levels declined in seeds (mature embryos) and seedlings. Differential regulation of LYR/LVR gene expression with response to brassinosteroid treatment of Douglas-fir seeds was studied. LYR/LVR mRNA showed higher accumulation in seeds treated with different concentrations of brassinosteroids. Bioinformatic analysis showed that Douglas-fir LYR/LVR protein may be an essential inner mitochondrial protein, NADH oxidoreductase necessary for energy production. The phylogenetic tree analysis was used to investigate the evolutionary relationship of the newly identified Douglas-fir LYR/LVR

protein with closely related proteins (LYR family) in different organisms. InterPro, UniProt and Pfam results showed the sequence similarity of Douglas-fir LYR/LVR protein with other related members in *Arabidopsis thaliana* and *Oryza sativa*, indicating that the LYR complex contains short stretches of closely related proteins that are essential for energy production. Amino acids 19-90 in the LYR/LVR protein were highly conserved and is likely the functional LYR motif necessary for oxidoreductase activity.

Supervisor: Dr. Santosh Misra, (Department of Biochemistry and Microbiology)

Table of Contents

Supervisory Committee....ii
Abstract....iii
Table of Contents....v
List of Tables....viii
List of Figures....ix
List of Abbreviations....x
Acknowledgements....xii
Dedication....xiv

CHAPTER 1. INTRODUCTION....1

Zygotic embryogenesis....2
Somatic embryogenesis....4
Embryogenesis-specific genes in conifers....7
Molecular analysis of SE....9
Hormones and conifer embryogenesis....13
Importance of brassinosteroids....14
Main Objectives....16

CHAPTER 2. MATERIALS AND METHODS....17

Plant materials....17
Seed germination....19
Hormone treatment....19
Reagents and media....20
RNA experiments....20
Total RNA extraction....20
Trizol extraction method....20
Lithium chloride extraction....21
Northern Hybridization....21
DNA experiments....23

Douglas-fir genomic DNA isolation....................23
Plasmid DNA isolation....................24
DNA fragment purification....................25
Labelling DNA fragment....................26
Random primer labelling....................26
Southern hybridization....................27
Prehybridization and hybridization....................28
PCR amplification of Douglas-fir cDNA probe used for hybridization....................28
RT-PCR amplification of early embryogenesis gene from Douglas-fir RNA....................29
First strand cDNA synthesis using superscript II RT....................29
PCR amplification....................29
Molecular cloning....................30
Preparation of *E.coli* competent cells....................30
Transformation of *E.coli*....................31
Bioinformatic Analysis....................33
Base-By-Base....................33
InterPro....................34
CLUSTALW....................35

CHAPTER 3. RESULTS....................37

Cloning of an early embryogenesis gene of Douglas-fir....................37
Isolation of Douglas-fir LYR/LVR cDNA using RT-PCR amplification....................37
Cloning of Douglas-fir LYR/LVR gene....................40
Sequence analysis of Douglas-fir LYR/LVR gene....................40
Bioinformatic analysis of LYR/LVR gene....................42
Analysis of LYR/LVR protein....................47
Phylogenetic analysis....................51
Characterization of LYR/LVR gene....................53
Genomic Southern blot analysis....................53
Northern Blot Analysis....................53
Effect of brassinosteroid on LYR/LVR gene expression in mature seeds....................56

CHAPTER 4. DISCUSSION....................58

Genes involved in Douglas-fir embryogenesis....................58
Identification and cloning of Douglas-fir LYR/LVR cDNA....................59
Bioinformatics findings....................59
LYR complex....................61
Differential expression of Douglas-fir LYR/LVR during embryogenesis....................64
Effect of brassinosteroids on Douglas-fir LYR/LVR gene expression....................64

CHAPTER 5. CONCLUSIONS AND FUTURE STUDIES....................67

Conclusions..67
Future studies..68

BIBLIOGRAPHY...**69**

APPENDIX 1. NCBI BLASTP (protein-protein blast) showing the alignment of translated amino acid sequences of Douglas-fir LYR/LVR protein..........................78

APPENDIX 2 A. Primers used for the synthesis of LYR/LVR probe for Southern and northern hybridizations..80

APPENDIX 2 B. Primers used for the isolation of Douglas-fir LYR/LVR cDNA by RT-PCR amplification...80

List of Tables

Table 1. Different stages of Douglas-fir zygotic embryos and their collection dates......18

Table 2. Percentage amino acid sequence identity of the Douglas-fir LYR/LVR protein with LYR/LVR proteins of other related organisms.......................45

Table 3. Percentage identity scores between LYR/LVR amino acid sequences of Douglas-fir, *Arabidopsis thaliana* and *Oryza sativa*...........................46

List of Figures

Figure 1. Different stages of Douglas-fir somatic and zygotic embryos....................5
Figure 2. RT-PCR using total RNA isolated from different stages of zygotic embryos with a set of primers (Appendix 2) based on *Arabidopsis thaliana* LYR gene 38
Figure 3. 1% agarose gel electrophoresis of restriction digested plasmids from positive white colonies..39
Figure 4. Douglas-fir LYR/LVR cDNA sequence and Translated +3 frame of the LYR/LVR protein...41
Figure 5. CLUSTALW alignment of the Douglas-fir LYR/LVR amino acid sequence with *Arabidopsis thaliana* and *Oryza sativa* LYR proteins...........43
Figure 6. CLUSTALW alignment of Douglas-fir LYR amino acid sequence with LYR/LVR protein of other related organisms.............................44
Figure 7. Taxonomic coverage of Douglas-fir LYR/LVR protein matches with other closely related known proteins of the same LYR family (complex 1) in different organisms..48
Figure 8. InterPro scanning showing different LYR proteins of the LYR complex family in *Arabidopsis thaliana*....................................49
Figure 9. InterPro scanning to determine different LYR proteins in *Oryza sativa* constituting the LYR complex......................................50
Figure 10. Phylogenetic tree analysis of LYR/LVR proteins...............................52
Figure 11. Genomic Southern analysis of Douglas-fir LYR/LVR..........................54
Figure 12. Northern blot analysis of LYR/LVR gene expression in different stages of Douglas-fir zygotic embryos and other tissues..........................55
Figure 13. Northern blot analysis of Douglas-fir seeds treated with different concentration of brassinosteroids at different time intervals....................57

List of Abbreviations

ABA	Abscisic acid
Amino acid	A : alanine; C : cysteine; D : aspartic acid; E : glutamic acid; F : phenylalanine; G : glycine; H : histidine; I : isoleucine; K : lysine; L : leucine; M : methionine; N : asparagine; P : proline; Q : glutamine; R : arginine; S : serine; T : threonine; V : valine; W : tryptophan; Y : tyrosine
ATP	Adenosine 5´ triphosphate
Nitrogenous bases	A : Adenine C : Cytosine G : Guanine T : Thymine
bp	Base pairs
BSA	Bovine serum albumin
BR	Brassinosteroids
cpm	Counts per minute
CTAB	Hexacetyltrimethylammonium bromide
DEPC	Diethyl pyrocarbonate
deoxynucleotides	dATP: deoxyadenosine-5´-triphosphate dCTP: deoxycytidine-5´-triphosphate dGTP: deoxyguanosine-5´-triphosphate dTTP: deoxythymidine-5´-triphosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
h	Hour(s)
IPTG	Isopropylthiogalactoside
kb	kilobase pairs
kDa	kilodaltons
LB	Luria-Bertani medium (1% NaCl, 1% Bacto-tryptone, 0.5% yeast extract, 10 mM MgSO4, 0.2% maltose)
M	Molar
μg	microgram
μl	microliter

ml	millilitres
mg	milligram
min	minute(s)
mM	millimolar
MOPS	3-[*N*-morpholino]propanesulfonic acid
MW	Molecular weight
PEG	Polyethylene glycol
RNA	Ribonucleic acid
mRNA	Messenger ribonucleic acid
rRNA	Ribosomal ribonulceic acid
RNAase	Ribonuclease
rpm	revolutions per minute
S	Svedberg sedimentation unit
SM	Suspension medium (0.1 M NaCl, 8 mM MgSO4.7H2O, 50 mM Tris-HCl pH 7.5, 0.01% gelatin)
SSC	Sodium chloride-sodium citrate buffer (1X SSC : 0.15 M sodium chloride, 0.015 M sodium citrate)
TAE	Tris-acetate EDTA buffer (1X TAE : 0.04 M Tris-acetate, 0.002 M disodium-EDTA, pH8.0)
TBE	Tris-borate EDTA buffer (1X TBE : 0.089 M Tris, 0.089 M boric acid, 0.008 M disodium-EDTA, pH 8.0)
TE	Tris-EDTA buffer (1X TE : 10 mM Tris-HCl pH 8.0, 1 mM EDTA)
Tris	Tris-[hydroxymethyl]-aminomethane
X-gal	5-bromo-4-chloro-3-indolyl-β-D galactopyranoside

Acknowledgements

I kindly thank my supervisor Dr. Santosh Misra for her valuable support, guidance and encouragement throughout the progress of this study. I would like to thank my committee members, Dr. Peter Constabel and Dr. Chris Upton for their support, time, advice and guidance for completing this study. I would like to thank Dr. Chris Upton for assisting me with the Bioinformatics work involved in this study and Dr. Peter Constabel for guiding me during the sabbatical leave of my supervisor early this year. I would like to thank Dr. Barbara Hawkins for her kind acceptance and valuable suggestions as an external examiner. I express my gratitude to the Center for Forest Biology, University of Victoria, especially to Brad Binges for his help in collecting Douglas-fir cones and providing other necessary facilities. I kindly acknowledge the Pacific Forestry Centre for providing the Douglas-fir seeds used for this study. I would also like to thank Scott Scholz, Albert Labossiere and Stephen Horak for their technical assistance and support.

Thanks to all my colleagues in Dr. Misra's laboratory Dr. Ravinder Kumar, Dr. Dmytro Yevtushenko, Teresa Francescutti, Apurv Bharghava, Dr. Milan Osusky and Mariana Vetrici. Special thanks to Melissa Da Silva and Aijazuddin Syed for helping me with the bioinformatics part of my study. Special thanks to Dr. Ravinder Kumar and Dr. Dmytro Yevtushenko for their help and suggestions during the course of my research work. Thanks to departmental secretaries Melinda Powell, Deb Penner and administrative officer John Hall for their assistance. Special thanks to Janice Keliher, Dr. Barbara

Hawkins, Dr. Peter Constabel, Dr. David Levin for their continuous encouragement and moral support. Also, I would like to thank all my friends Sudarsanji, Venkat, Shiva, Deepak, Rambabu, Sanal, Anna, Prabhakar, Mike and many others for their friendship and support. I give my sincere appreciation to my loving brothers Mahesh, Anand, Anoop and all my family members including my aunt Leela who visited me in Victoria and supported me during hard times.

Most importantly, my sincere wishes to my mother, father and my grandmother for their consistent encouragement, love, support and confidence in everything I do.

Dedication

I would like to dedicate this thesis to my mother, Macheth Saraswathi, my father Ramachandran Rhaghavan and to my grandmother who always supported and encouraged me in everything I pursued.

Chapter 1

INTRODUCTION

Douglas-fir (*Pseudotsuga menziesii*) is a member of the Pinaceae, the largest family of conifers. Apart from its wide distribution in the Pacific North-West, it is also commonly found in New Zealand, Europe and Australia. Its high quality lumber, efficiency for extreme fiber stress in bending, resistance to warping and resilience against strong winds, storms and earthquakes make it the most economically important conifer species. Consequently, Douglas-fir is the major component of reforestation programs (Abo El-Nil, 1987). It is also an important species ecologically as its seeds are an important feed source for birds and other mammals.

This creates a continuous demand for Douglas-fir seeds used in reforestation. The achievements using conventional breeding and forestry practices have been slow as the reproductive cycle of Douglas-fir extends over a period of 17 months (Allen and Owens, 1972). Also the seed crops can be unpredictable due to failure during reproductive processes, including lack of cone bud development, lack of pollination, seed immaturity and pre-embryo abortion (Owens *et al.*, 1991). Hence, environmental control and the need for genetically superior trees with high-value lumber are the primary objectives of Douglas-fir reforestation programs.

Elite trees are generally selected using various breeding processes of recurrent selection: however, due to the slow growth habit of conifers, it may take many years before a tree is ready for assessment of phenotype, and even longer before cone production (Timmis *et al.*, 1987). In contrast, clonal propagation is more expeditious means of propagating genotypes exhibiting desired traits (Timmis *et al.*, 1987). Micropropagation has been suggested as an alternative method of clonal propagation and includes three main types- shoot elongation, organogenesis (Thorpe *et al.*, 1982), and embryogenesis. Embryogenesis is the most preferred method of mass micropropagation because a) greater numbers of propagules are produced, b) embryogenesis is faster, there is a lower probability of abnormalities arising in plantlets produced by embryogenesis than organogenesis, c) embryoids produced *in vitro* may be used to prepare artificial seed, and it acts as the basis for large volume tissue culture using cell suspensions (Goldberg, 1998).

Zygotic Embryogenesis

Zygotic embryogenesis is a highly organized and complex process that plays a vital role in the life cycle of higher plants. Following fusion of haploid male and female gametes resulting in the formation of a diploid zygote, embryo development and concomitant cellular differentiation commence. The term fertilization in plants may involve more than the fusion of a male gamete with an egg. In gymnosperms, fertilization may involve other and unusual events from the time the pollen tube reaches the neck of the archegonium until the first division of the zygote (Allen and Owens, 1972). In Douglas-fir, the fate of all nuclei and cells that pass through the pollen tube must be considered because they all

do not disappear after the fusion of the male gamete and the egg. Fertilization in Douglas-fir generally occurs between June 1 and June 20 at lower elevations in the Pacific Northwest and British Columbia. In any one year, there is a variation of several days between ovules depending on an individual tree (Allens and Owens, 1972). Subsequent early events in embryogenesis give rise to an embryo proper and a separate suspensor region (Schwartz *et al.*, 1997, Owens *et al.*, 1993). In angiosperms, mature embryos or seeds act as products of the process of double fertilization, where one of the two sperm nuclei fertilizes the haploid egg cell producing a diploid zygote, and the remaining sperm nucleus fertilizes the polar nuclei in the central cell resulting in a triploid endosperm (Reiser and Fischer, 1993, Mordhost *et al.*, 1997). As the endosperm and embryo develop, the ovule enlarges and forms a seed.

In contrast to the double fertilization event and triploid endosperm characteristic of angiosperms, gymnosperm embryogenesis involves a single fertilization of the female oocyte. Hence, embryos develop in a haploid female tissue, the megagametophyte (Misra S, 1994, 1995). In angiosperms, the first zygotic division determines the basal cell, which results in the formation of the suspensor, and a terminal cell, which gives rise to the embryo proper (Yeung *et al.*, West *et al.*, 1993). Gymnosperms undergo a free-nuclear phase followed by several nuclear divisions before the formation of cell wall. Another round of division results in the formation of a four-tiered 16-celled proembryo. The four cells in the distal tier give rise to the embryo proper with respect to the mycropylar end of the seed, and the next tier form the suspensor (Schultz *et al.*, 1969, Krasowski *et al.*, 1993). Also in gymnosperms, a phenomenon called cleavage polyembryony is common,

where each embryo proper can divide into four individual embryos, each containing its own suspensor (Spurr, 1949).

During the process of embryogenesis, the zygote undergoes various complex morphological and cellular changes aiding in the formation of a mature embryo consisting of an embryo axis with shoot and root apices and cotyledons (Misra and Green, 1991) (**Figure 1**). Embryogenesis usually involves three main steps: (1) initial morphogenetic phase (early embryogenesis), characterized by cell division followed by cell differentiation (2) maturation phase (mid-cotyledonary stage) involving accumulation of major storage products and preparation for seed desiccation, dormancy and germination (3) late embryogenesis, characterized by water loss and the embryo becomes metabolically quiescent, desiccated and developmentally inert (Kermode, 1990). Much less is known about conifer embryogenesis probably due to the long reproductive cycle.

Somatic Embryogenesis

Somatic embryogenesis (SE) is the process by which somatic cells are induced to develop into differential plants through characteristic embryo development. SE is the most efficient method to produce high quality trees with desirable characteristic features as somatic cells from elite trees produce superior quality somatic embryos with improved germination rates (Gupta *et al.*, 1995). Embryogenesis can be initiated either from somatic (diploid) or gametic (haploid) tissues or cells. In addition, haploid embryogenesis

A

Late-stage proembryos

Precotyledonary somatic embryo

Cotyledonary somatic embryo

Mature somatic embryo

B

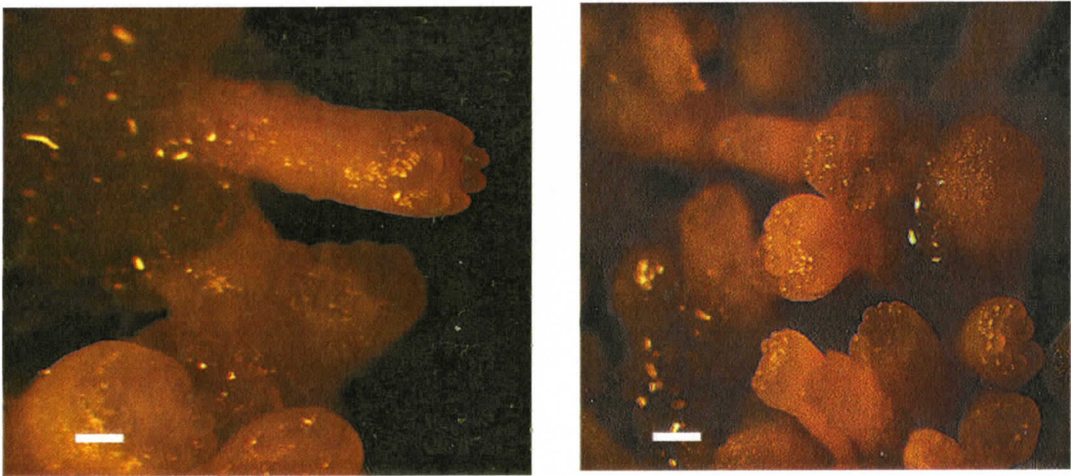


Figure 1 (A-B). Different stages of Douglas-fir somatic and zygotic embryos. **A)** Four important stages of Douglas-fir somatic embryos are shown (Chatthai *et al.*, 1998). **B)** Douglas-fir mature zygotic embryos. Bars equal 0.5 mm.

can be further subdivided into cultures derived from male tissues (androgenesis) or female tissues (gynogenesis). *In vitro* embryogenesis was first observed in angiosperms (*Daucus carota*) in 1958 (Thorpe, 1998); however, it was not achieved in conifers until 1985, when it was shown in both *Picea abies* (Hackman *et al.*, 1990, 1993, Gupta *et al.*, 1986, Boulay *et al.*, 1988) and *Larix decidua* (Nagmani and Bonga, 1985). Subsequently, the focus of tissue culture research is to develop reliable methods of somatic embryogenesis that is applicable to a broad range of species and genotypes. Generally, *in vitro* embryogenesis consists of three important stages: induction, proliferation, and development (de Jong *et al.*, 1998, Zimmerman, 1993). During the first stage, the explant material is induced to produce embryogenic tissue. The callus tissue- undifferentiated parenchyma cells forming an explant and then may give rise to somatic embryos (Durzan, 1988, Dunstan *et al.*, 1995). In conifers, this tissue is an organized mass of compact embryonal cells associated with long suspensors, comparable to the early embryo of conventional zygotic development (Durzan *et al.*, 1988, Nomura *et al.*, 1986, Jokinen *et al.*, 1994, Dodeman *et al.*, 1997).

In conifers, somatic embryogenesis is induced from primary explants such as immature and mature zygotic embryos that are cultured on medium containing auxin, cytokinin and kinetin. Two types of cells are produced: large, vacuolated non-embryogenic cells and smaller, compact embryogenic cells (Lelu and Bornman, 1990). Only the latter type develops into a fully mature somatic embryo. The process of induction of embryos to undergo cleavage and form numerous identical embryos is called cleavage polyembryony and occurs even during seed development of some species (Dodeman *et al.*, 1997).

Mature embryos can be obtained by transferring singulated embryos into a medium containing abscisic acid (ABA), gibberellic acid (GA), high osmotic polyethylene glycol (PEG) and activated charcoal (Gupta *et al.*, 1991).

SE technology involving the development of somatic embryos has been reported in several coniferous species (Tautorous *et al.*, 1991) including Douglas-fir (Durzan and Gupta, 1987). In addition, several methods have been successful for inducing somatic embryogenesis from mature tissues through stress or hormonal treatments (Boutilier *et al.*, 2002, Ikeda-Iwai *et al.*, 2003). Douglas-fir somatic embryo cultures show discontinued growth after six months due to unknown reasons and hence molecular markers are needed for proper understanding of the embryogenic process at the molecular level. Also, there is no evidence that SE may be induced from vegetative cells such as needles or bark.

Embryogenesis-specific genes in conifers

Rapid advancement in molecular biology techniques and their application to plant breeding has yielded several genetic gains in agricultural crops, some of which have already entered the market (Walter *et al.*, 1998). Molecular markers are generally used to identify elite genotypes in both forestry and agriculture and detect specific genomic regions necessary for important traits, resulting in the potential for characteristic improvements with respect to clone identification and detection of individuals with superior performance (Wilcox *et al.*, 1996).

Genetic engineering, in comparison with agricultural biotechnology applications, is lagging behind in conifers, partly because of difficult regeneration, tissue culture protocols and long generation times for the species involved (Aitken-Christie, 1991, 1992). Large scale identification of genes and expression requires modern tools of genomics. Conifer functional genomics aims to define the roles played by most of the genes in a given species. Furthermore, it seems that the conifer functional genomics will lead to new insights on manipulating tree genomes for practical benefits such as increasing yield or altering wood quality. However, the increasing demand for paper, pulp and timber products apart from the growing awareness of the high ecological and social value of indigenous forests, have stimulated several molecular tree improvement programs. DNA markers aid to the increased genetic gains from tree improvement programmes by the identification of DNA markers associated with traits of economic value and revealing differences in the DNA sequence among individual trees (Silen, 1978). Marker aided selection with DNA markers is very promising for use in conifers (Walter *et al.*, 1998). One of the reasons for the slow progress of genetic engineering of forest trees when compared to agricultural crops is their long rotation times. Genetic engineering is considered a useful tool for the development of commercial forests because a range of desired traits that are not available in the breeding population could potentially be transferred into conifers from other sources. Rapid production of large quantities of genetically improved and uniform tree seedlings for restocking harvested land is an essential factor for sustainable forest management programs. Somatic embryogenesis acts as an efficient testing system for introducing genes for the investigation of wood quality and genes necessary for growth and energy production in

conifers. There are cogent reasons for replacing natural seed with somatic or synthetic seed, especially in coniferophyta where the natural seed can survive in the desiccated state (Bornman, 2002). The nucleus of the artificial seed constitutes the somatic embryo and the cost of this type of seed will be related to the viability and efficiency of somatic embryo production. Gymnosperms are outcrossing, and the most important method of mass propagation is via clonal seed-orchard production. This provides a platform for genetic engineering to test genes that may be involved in wood quality and growth. These trees could then be used to detect the effects of specific genes, data that is very important in transgenic applications, for marker-assisted selection and combinations.

Molecular analysis of SE

Vegetative propagation through somatic embryogenesis has the potential to meet this need in conifers and provide additional benefit of ensuring consistent seedling quality. In commercial production, conifer mass production through somatic embryogenesis is relatively new and there are various unknown biological reasons during this complex developmental pathway (Taniguchi *et al.*, 2004). In order to provide more insight into the embryo developmental processes, two-dimensional electrophoresis was employed to quantitatively assess the expression level of proteins in different stages of development of *Picea glauca* embryos (0, 7, 21 and 35 days post ABA treatment) (Bommineni *et al.*, 1998). Forty-nine differentially expressed proteins were identified including those that showed significant changes in abundance and could be detected as early as day 7 of embryo development. These proteins are involved in various cellular processes and most of them were not previously been associated with embryo development (Lippert *et al.*,

2005). The availability of sufficient expressed sequence tags (EST) resource developed for white, sitka and interior spruce aided in completion of the study conducted by Lippert and his colleagues. The combined use of spruce (ESTs) with respect to GenBank accessions for other plants improved the rate of protein identification from 38 % to 62 % when compared with GenBank alone, using automated and highly sophisticated techniques. This specifies the importance of the utility of EST resources in a proteomic study of any species for which a genome sequence is unavailable (Lippert *et al.*, 2005).

To understand the molecular events involved in conifer embryogenesis, extensive work is required to explore stage-specific gene expression during zygotic and somatic embryogenesis. Pines and other gymnosperms, in contrast to angiosperms, form well-developed suspensors in somatic embryogenic cultures (Cairney *et al.*, 1999, 2000). The study of gene expression during the early stages of loblolly pine embryogenesis resulted in the identification of PtNIP1:1 transcript in immature loblolly pine (*Pinus taeda*) zygotic and somatic embryos. PtNIP1:1 constitutes the nodulin-like members of the major intrinsic protein superfamily branch of the aquaporins (Ciavatta *et al* ., 2001, 2002). However, this protein is undetectable in later-stage embryos, megagametophytes, roots, stems, one year-old seedlings and was shown to act as an active aquaglyceroporin (glycerol permease).

Advancement towards identifying individual genetic elements that regulate embryo development in angiosperms has been significantly aided by cloning of important regulatory genes such as LEC1, ABI3, FUSCA3 and VP1 (Kagaya *et al.*, 2005). Recent

studies have shown that ectopic expression of LEC1 is sufficient to induce embryogenesis pathway in transgenic *Arabidopsis thaliana* vegetative tissues (Lotan *et al.*, 1998). *LEAFY COTYLEDON 1* is a CCAAT box-binding transcription factor and acts as a critical regulator of early embryogenesis. Recently, LEC1 gene was isolated in Douglas-fir by using PCR with a set of degenerate primers based on the conserved domain of Arabidopsis LEC1 (Vetrici and Misra, 1998). Studies are in progress to identify Douglas-fir LEC1 (PmLEC1), examine its tissue-specific expression and induction in mature tissues by hormones and other stress chemicals. Molecular cloning of PmLEC1 raises the possibility of modulating its expression and function for forestry benefits.

In order to better understand the SE process for Douglas-fir and white spruce, immunoscreening and differential screening of cDNA libraries from mid-maturation stage embryos were used for the isolation of cDNAs corresponding to the genes expressed during maturation of zygotic and somatic embryos. cDNAs encoding 11S and 2S seed storage protein genes (Chatthai *et al.*, 1998) and a MT-like gene were identified and selected for further characterization (Chatthai *et al.*, 1997). Also, antibody and cDNA probes were used to analyze temporal changes in mRNA, protein profiles in developing zygotic and somatic embryos of different genotypes and to examine the effect of specific cultural characteristics on the quality of embryos with respect to the expression of these genes. Several cDNAs were also isolated from germination-specific cDNA libraries of Douglas-fir. The cDNAs encoding NADPH-P450 reductase (Tranbarger *et al.*, 2000), thiol-protease (Tranbarger *et al.*, 1996) and BiP, a molecular chaperone (Forward *et al.*,

2001) were identified and characterized in Douglas-fir. These were used to monitor conversion of somatic embryos into viable plants. A genomic library of Douglas-fir was constructed to elucidate the regulation of the genes discussed above and the genomic sequences of corresponding cDNAs were isolated. The various constructs (PmBiPPro-1) using GFP/GUS chimeric gene have been stably transformed in *Arabidopsis* and or tobacco (Forward *et al.*, 2001). These studies have provided basic information on molecular biology of embryo development in conifers.

Characterization has been more difficult for genes that encode proteins of low abundance, or with transitory, ill-defined, or completely unknown activities. For example, genes involved in signal transduction- the process of coordinating growth, development, and environmental responsiveness have not been well studied in conifers. Recently work has been done to identify and elucidate these genes with the help of genomics (Lippert *et al.*, 2005).

Some of the embryo-specific genes encode mitochondrial membrane proteins, knowing more about these genes will contribute to our understanding of different developmental pathways during embryogenesis and energy production. Global gene expression assays such as differential display and DNA based arrays both in zygotic and somatic embryos of *Picea glauca* indicated that approximately 5% of genes were actively regulated during embryogenesis (Lippert *et al.*, 2005). Approximately, one quarter among the actively regulated genes were functionally similar in both somatic and zygotic embryos. Further, several embryogenesis related genes of pine were homologous to *Arabidopsis thaliana*.

Despite differences in certain aspects of angiosperm and gymnosperm embryogenesis, certain molecular pathways are shared.

Hormones and Conifer Embryogenesis

Somatic embryogenesis in higher plants is induced by exogenous hormones and culture-related stress (Dong *et al*., 2000). In tissue culture, ethylene may act as either a promoter or inhibitor of regeneration depending upon specific plant species (Minocha *et al*., 2000). Work has been done to elucidate the regulation of specific proteins/factors in the megagametophyte (MG) and its relation with induction and early stages of embryogeny. Recent studies indicate that metabolites such as carbohydrates act as essential signaling molecules to regulate embryo development. Moreover, there are evidences to show the interaction between sugar signaling and signals mediated by phytohormones. Sugars are known to play an important role in induction and maturation of conifer somatic embryos (Minocha *et al*., 2000, Brocard-Gifford *et al*., 1991, 2003). In Norway spruce, induction of SE was increased in the media containing lower levels of sucrose (1.5-2 %) along with the presence of 2,4-D, whereas in loblolly pine and white pine, the concentration of sucrose (6 %) in combination with ABA promoted maturation of SE (Black, 1991, Schneider *et al*., 1994).

In order to understand the role of megagametophyte and the molecular basis of early embryogeny, studies on changes in metabolites and hormones in the MG and embryo with respect to the concentration of carbohydrates (hexoses and starch) and phytohormones (ABA, auxin and gibberellic acid) have been analyzed. Sugars, ABA and

PEG stimulate embryo development and maturation, while auxins and cytokinins play important roles in regeneration (Minocha *et al.*, 2000). More work needs to be done to examine the effect of these treatments on the expression patterns of embryo-specific genes and how their activity is modulated. Ikeda-Iwai and co-workers (2003) performed experiments where they used certain chemical stressors to induce somatic embryogenesis and demonstrated the increased expression of embryogenesis-specific genes with response to the stimuli. In *Arabidopsis thaliana*, explants from shoot apex were placed on phytohormone-free medium containing stress inducers such as sucrose, sodium chloride, cadmium chloride, sorbitol, mannitol or dehydration. After few days, the explants were transferred to phytohormone-free medium without the stressors and somatic embryos were formed on the surface of the shoot-apical tip explants. Expression of embryo-specific genes such as FUS3 and ABI3 was observed in the somatic embryos but not in callus (Kagaya *et al.*, 2005). These experiments showed that certain chemical stressors induce SE and this could be confirmed by identifying somatic embryos and expression of embryo-specific genes.

Importance of Brassinosteroids

Brassinosteroids (BR) are a new group of natural plant growth regulators (Steroidal Lactones - including brassinolide and its analogs) found in many plant species. Angiosperm seeds have been shown to have diverse, species-specific and tissue-specific responses to BR including the stimulation of cell elongation, cell division and differentiation, reproductive biology, senescence, ethylene production and increased

resistance to abiotic stress (Brosa 1999, Clouse and Sasse 1998, Clouse 2001, Bishop *et al.*, 2002).

In most conifer species, only a few cells become embryogenic and not all genotypes are embryogenic. BRs are known to induce embryogenesis at very low concentrations in conifers (Pullman *et al.*, 2002a, 2002b). Pullman and colleagues (2003) studied the effect of brassinolides on somatic embryogenesis in conifers and rice. Brassinolides promoted conifer embryonic tissue formation and growth at very low concentration and was most stimulatory at a concentration of 0.1 μM. The initiation percentage for loblolly pine, Douglas-fir increased relatively from 15 to 30.1%, to 16.1% to 36.3% respectively.

Neff and co-workers (1999) identified BAS1 – a gene regulating brassinosteroids levels and light responsiveness in *Arabidopsis*. Based on mutant screening and genetic analysis, it was determined that the *Arabidopsis* bas1-D mutation suppresses the long hypocotyl phenotype caused by photoreceptor phytochrome B (phy B). Russinova *et al.*, (2004), showed the heterodimerization of Brassionsteroid (BR) receptor (BRR) with *Arabidopsis thaliana* somatic embryogenic receptor-like kinase 3 (AtSERK3). Limited information is available about the effect of BR with respect to embryo-specific genes in plants. In conifers, more work is needed on the differential expression pattern of the early- and late-embryogenesis genes with respect to the brassinosteroid treatment of different developing tissues and genotypes.

To understand the molecular basis of early embyogeny and effect of BRs, degenerate primers (Appendix 2) were designed based on known sequences from *Arabidopsis thaliana* COP1 (Constitutively Photomorphogenic 1). Several PCR products were obtained, one of them showed identity to a mitochondrial membrane protein (LYR/LVR). LYR/LVR protein is related to energy production, but not much is known about its expression and regulation in embryogenesis. The primary focus of this thesis was to clone LYR/LVR and examine its expression and regulation in developing embryos and seedlings.

Main objectives were:

- ❖ Identification of the early embryogenesis expressed genes in Douglas-fir.
- ❖ Characterization of an LYR/LVR cDNA and examine its temporal and tissue-specific expression.
- ❖ Study the effect of brassinosteroid treatment on LYR/LVR gene expression.
- ❖ Analyse Douglas-fir LYR/LVR cDNA using bioinformatics.

Chapter 2
MATERIALS AND METHODS

PLANT MATERIALS

Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) immature cones were obtained from trees near the University of Victoria, Victoria, BC, between May 15 and August 23, 2004- 2005. The cones were collected in paper bags and then stored at 4 °C until they were dissected within two days. The seeds were collected on May 15 and 30, June 8, 18 and 29, July 14 and 29, and August 10 and 25. Developing seeds were isolated from the cones and their seed coats removed. Both cones and seeds varied significantly in size as the trees in British Columbia contain smaller seeds (about 132,000/kg or 60,000/lb) when compared to the larger seed (about 51,000/kg or 23,000/lb) from trees in California.

The seeds collected on specified dates were dissected to obtain the megagametophytes and the early embryos. After the separation process, the tissues were immediately frozen and stored at –80 °C. Dry Douglas-fir seeds were obtained from The BC Ministry of Forests. The morphological features of different developmental stages of Douglas-fir embryos were compared with those criteria established by Allen and Owens (1972). The morphological features associated with different stages of zygotic embryos and their collection dates are summarized in Table 1.

Stage	Collection Dates	Developmental stages
1	May 15	Pre-fertilization Ovule consists of a multicellular haploid megagametophyte
2	May 30, June 8	Embryonal mass Zygote contains 2-4 nuclei close to the center of the archegonium
3	June 18	Early-embryo Megagametophyte enlarges, formation of proembryo from the zygote and the proembryo consists of 10-12 cells
4	June 29	Club-shaped embryo Proembryo club-shaped along with suspensor cells
5	July 14	Early-cotyledonary embryo Club-shaped proembryo enlarges and differentiates forming dome-shaped initial cotyledons
6	July 29	Mid-cotyledonary embryo Embryo differentiates further with polar growth of the cotyledon and epicotyl
7	August 10	Late cotyledonary embryo Embryo enlarges containing fully developed cotyledons
8	August 25	Mature embryo Formation of seeds

Table 1. Different stages of Douglas-fir zygotic embryos and their collection dates. The different developmental stages of zygotic embryos were examined and the embryos were dissected and frozen in liquid nitrogen.

SEED GERMINATION

Douglas-fir seeds (*Pseudotsuga menziesii* [Mirb.] Franco) were imbibed overnight in sterile distilled water at 4 °C. After blotting dry, the seeds were placed in a plastic bag for 3 weeks at 4 °C for stratification, a treatment that facilitates the germination of several conifer species including Douglas-fir (Tranbarger *et al*., 2000, Allen, 1941, El-Kassaby *et al*., 1992). Stratified seeds were then placed in small germination chambers containing a single layer of sterile Kimpak and two layers of sterile filter paper (Rose Scientific Ltd., Edmonton, AB). These seeds were germinated in a controlled incubator or growth chamber maintaining day and night temperatures of 30 °C and 20 °C, respectively, with 16 hours light and 8 hours dark photoperiod.

HORMONE TREATMENT

The germination chambers were wrapped with parafilm to avoid contamination and maintain humidity before placing them in a photo incubator. Analyses of the effects of plant steroidal lactones such as brassinosteroids on the early embryogenesis-specific gene expression were performed on the stratified seeds. The stratified seeds were treated with different concentrations of brassinosteroids; 0.1 μM, 0.5 μM, 1 μM, 5 μM and 10 μM by incubation in sterile Petri dishes containing Whatman 1MM filter papers soaked with the hormone solutions. The Petri dishes containing the treated seeds were wrapped with aluminium foil and incubated in the dark at specific time intervals; 6 hours, 1 day and 1 week, after which the samples were collected and frozen in liquid nitrogen, and stored at –80 °C.

REAGENTS AND MEDIA

Culture media, buffers, antibiotics and other solutions were prepared based on procedures described in Sambrook *et al.* (1989) with appropriate storage conditions. All the laboratory chemicals were obtained from BDH Chemicals, Fisher Scientific, Gibco BRL or Sigma. Restriction enzymes were obtained from New England Biolabs and used per their specifications. Ingredients necessary for cloning such as tryptone, sodium chloride, and yeast extract were purchased from EMD. Antibiotics, isopropyl-β-D-thiogalactoside (IPTG), and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) were purchased from Sigma. For RNA experiments, all the glassware was first baked at 180 °C in a hot air oven for 15-17 hours to inactivate RNases, and the solutions were treated with 0.05% diethylpyrocarbonate (DEPC; Sigma) overnight and then autoclaved.

RNA EXPERIMENTS

Total RNA Extraction

Total RNA was extracted from Douglas-fir zygotic embryos, megagametophytes, seeds and seedlings using Trizol (Invitrogen) extraction method and protocol described by Verwoerd and co-workers (1989) with some modifications.

Trizol Extraction method:

For the Trizol extraction method, 100-200 mg of frozen tissues were ground in a mortar-pestle using liquid nitrogen and transferred to sterile 1.5ml microcentrifuge tubes. The extraction procedure was followed as per the protocol mentioned in the Trizol (Invitogen) RNA extraction catalogue.

Lithium chloride extraction method:

Approximately, 50-100 mg of frozen tissues were homogenized using liquid nitrogen and transferred to sterile 1.5 ml microcentrifuge tubes. The homogenized powder was first mixed with 0.5 ml of hot extraction buffer (100 Mm LiCl, 100 mM Tris-HCl pH 8.0, 1 % SDS 10 mM EDTA, in phenol) at 80 °C, and 0.25 ml of chloroform: isoamylalcohol (24:1, v/v) was added and thoroughly mixed by vortexing. The samples were centrifuged at 14,000 rpm in a microcentrifuge (Eppendorf) for 5 minutes, after which, the aqueous phase was transferred to a sterile microcentrifuge tube and 1 volume of 4 M LiCl (Fisher Scientific) was added. The samples containing total RNA were precipitated at –20 °C and collected by centrifugation. The RNA pellets were dissolved in 100-200 μl of DEPC water, and then mixed with 0.1 volume (25 μl) of 3M sodium acetate, pH 5.2. Total RNA was precipitated with 2 volumes of cold 100 % ethanol (500 μl) for 1 hour at –20 °C. The pellets were further washed with 70 % ethanol. RNA was resuspended in DEPC treated water and the concentration was estimated based on the absorbance at 260 and 280 nm using a Beckman DU-65 spectrophotometer. RNA samples extracted using Trizol (Invitrogen) extraction method were preferred as this method provided better yield when compared to the Lithium chloride extraction method.

Northern Hybridization

Total RNA (25 μg/lane) was mixed with 5X RNA loading buffer (720 μl of 37% formaldehyde, 80 μl of 500 mM EDTA (pH 8.0), 3 ml of 100% formamide, 2 ml 100% glycerol), and 4 ml of 10X FA (formaldehyde-agarose) gel buffer (200 mM MOPS, 50 mM sodium acetate and 10 mM EDTA, pH 7.0). RNA samples were separated on a 1.2

% agarose gel (100 ml) containing 10 ml of 10X FA gel buffer and 1.8 ml of 37% (12.3M) formaldehyde. In this method (Qiagen Rneasy Mini Handbook), higher concentration of RNA loading buffer was used allowing a larger volume of RNA sample to be loaded onto the gel compared to conventional protocols (Sambrook *et al.*, 1989). The final concentration of formaldehyde in the 5X loading buffer was 2 % v/v and the concentration in the agarose gel was 0.6 % v/v. The gel was equilibrated in 1X FA gel running buffer for at least 30 minutes and the samples were electrophoretically separated at 60 V (5-7 V per cm) and transferred to Biodyne nitrocellulose membrane (Pall) for 18-24 hours using 20 X SSC as transfer buffer. After transfer, the membranes were dried at 80 °C for 40 minutes under vacuum using a Bio-Rad gel dryer. The blots were then probed with a PCR amplified genomic fragment representing the Douglas-fir LYR gene labelled by random priming (Invitrogen) with (α - ^{32}P dCTP (Perkin Elmer). Pre-hybridization was done using Perfect Hyb plus hybridization buffer (Sigma) for 2-3 hours at 65 °C and then hybridized overnight in the same buffer at 65 °C. The probe was purified by Probequant GT purification columns from Amersham and then added to the hybridization buffer. After hybridization, the membranes were washed with 2X SSC, 0.1% SDS first at 40 °C for 30 min, then at 45 °C for 30 min, and then at 50 °C for 20 min, followed by two final washes, 1X SSC, 0.1% SDS at 50 °C for 15 min and then 0.5X SSC, 0.1% SDS at 50 °C for 15 min at the end. After washing, the membranes were exposed to Kodak BioMax MR film for 24-72 hours depending on the clarity of the blots.

DNA EXPERIMENTS

Douglas-fir genomic DNA isolation:

Forty to fifty grams of spring-flush needles were used for Douglas-fir genomic DNA extraction (Wagner *et al.*, 1987; Aagard, 1997). Using liquid nitrogen, the needles were crushed and made into a fine powder. The powdered mixture was homogenised in 100 ml of ice-cold extraction buffer in sterile centrifuge tubes for less than 10 seconds. The CTAB extraction buffer contains 50 mM of Tris-HCl (pH 8.0), 5 mM EDTA (pH 8.0), 0.35 M sorbitol, 0.1 % bovine serum albumin (stock: 200mg/ml), 0.1 % β-mercaptoethanol (added just before using), 10 % polyethylene glycol 4000, 0.1 % spermidine and 0.1 % spermine (stock: 200 mg/ml) for 1 litre of distilled water. The homogenate was filtered through 3-4 layers of cheesecloth and 2 layers of Mira cloth. The tubes were centrifuged at 9000 rpm for 15 min at 4 °C in a Beckman's GSA rotor. The pellets were resuspended in 5 ml of CTAB wash buffer and incubated at room temperature. The CTAB wash buffer contains 50 mM Tris-HCl (pH 8.0), 5 mM EDTA (pH 8.0), 0.35 M sorbitol, 0.1% β- mercaptoethanol for 1 litre of distilled water. Then, 1/5 volume of 5 % sarkosyl was added to the tubes and incubated for 15 minutes at room temperature. One seventh volume of 5 M NaCl was added and gently mixed by inverting the tubes, followed by the addition of 1/10 volume of 8.6 % CTAB in 0.7 M NaCl. After 15 min incubation at 60 °C, 10 ml of chloroform/octonol or isoamyl alcohol (24:1) was added and gently mixed 20 times until an emulsion was formed in the tubes. After centrifugation at 5000 x g for 10 min at 4 °C, the upper aqueous phase was transferred to another tube and 2/3 volume of isopropanol or ethanol was added and gently mixed by inversion to precipitate the DNA and again centrifuged at 3000 x g to lightly compress

the precipitate. Twenty millilitres of 76 % ethanol/10 mM ammonium acetate was added and incubated for 20 min, followed by centrifugation at 3000 x g for 10 min. The precipitate was transferred to a 1.5 ml sterile eppendorf tube containing 250 μl of sterile water or TE buffer. In order to dissolve the DNA pellets, the tube was kept for 30 min at room temperature and then mixed by gentle inversion. If not completely dissolved, the DNA was incubated at 65 °C for 3-5 min before storing at 4 °C.

Plasmid DNA Isolation

Three millilitres of overnight grown bacterial culture was taken into two sterile 2 ml eppendorf tubes, centrifuged at 13,000 rpm in a microcentrifuge (Eppendorf) for 15 min and the bacterial pellets were resuspended in 340 μl of GET. Fifty microlitres of lysozyme solution (10 mg/ml) was added, mixed quickly by inversion and incubated at room temperature for 10 min, followed by the addition of 400 μl of alkaline lysis solution and after gentle mixing placed on ice for 5 min. Subsequently, 400 μl of neutralizing solution was added and placed on ice for 5 min after gentle mixing. Centrifugation was done for 5 min at 4 °C in a microcentrifuge (Eppendorf) and the supernatant obtained was divided into two sterile eppendorf tubes (approximately 600 μl each), followed by the addition of 450 μl of room temperature isopropanol solution into each tube with gentle mixing. The mix was incubated for 2 min at room temperature. After centrifugation for 5 min at 4 °C, the supernatant was discarded and the remaining liquid removed after spinning briefly. The pellet obtained was dissolved in 98 μl of sterile distilled water and after combining the solutions; 4 μl of DNAse-free RNAase (10 mg/ml) was added and incubated at 37 °C for 15-20 min. Then, 120 μl of alkaline lysis solution and neutralizing

solution was added, gently mixed, placed on ice for 5 min and centrifuged for 5 min at 4 °C. The supernatant collected was mixed with 310 µl of room temperature isopropanol and after mixing was incubated for 5 min at room temperature. The supernatant was discarded after centrifugation for 5 min at 4 °C and the pellet was dissolved in 200 µl of TE buffer (pH 8.0), to which 140 µl of 20 % PEG-8000, 2.5 M NaCl, 10 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0) were added. This mix was kept frozen at –80 °C for 15 min, or preferably overnight. After thawing pellets at room temperature, the tubes were centrifuged immediately at 4 °C for 15 min. The pellets were first rinsed with 250 µl of ice cold 70 % ethanol, dried under vacuum for 8-10 min and, finally, the plasmid DNA was suspended in 10-20 µl of TE buffer (pH 8.0) and stored at –20 °C.

DNA Fragment Purification

PCR amplified DNA fragments and restriction endonuclease digested DNA fragments were electrophoretically separated on 1% agarose gel in 1X TBE and the required DNA fragments were visualized and excised using a sterile blade on a long-wavelength UV light box. Extraction from the gel was done using a Qiagen gel extraction kit. The excised agarose fragments were melted at 65 °C for 10 min and 3 volumes of Qiagen extraction buffer was added to 1 volume of gel with further incubation at 50 °C for 10 min until the gel slice had completely dissolved. Then, 1 gel volume isopropanol was added and mixed by inverting the tube. The sample was transferred to a QIAquick column and centrifuged for 1 min in a microcentrifuge (Eppendorf), followed by addition of 0.5 ml of gel extraction buffer to remove traces of agarose and centrifuged for 1 min. The column was finally washed with 0.75 ml of buffer PE containing ethanol and centrifuged for 1 min.

After an additional centrifugation step, the column was placed into a sterile 1.5 ml eppendorf tube, followed by elution of the DNA using 25-30 μl of elution buffer containing 10 mM of Tris Cl, pH 8.5. Centrifugation was done for 1 min after allowing the column stand for 1 min and the tube was stored at –20 °C.

Labelling DNA Fragment

Random Primer Labelling

Douglas-fir cDNA probe was radio-labelled using a random primer DNA labelling kit (Invitrogen) and subsequently used as probe for Southern and northern hybridizations. Twenty-five nanograms of DNA dissolved in 5-20 μl of sterile distilled water was denatured by heating for 5 min in a boiling water bath and immediately cooled on ice. Two microlitres of dATP, dGTP, dTTP, 15 μl of random primer buffer mixture and 5 μl of α - ^{32}P dCTP (approximately 50 μCi) were added to the denatured probe and made up to 49 μl with distilled water. The tube was gently mixed by tapping followed by the addition of 1 μl of Klenow DNA polymerase enzyme and incubated at 25 °C for 3-4 hours inside a lead container. The radio-labelled probe was purified using Probequant G-50 Micro columns (Amersham Biosciences). The probequant G-50 micro column was first prepared by resuspending the resin in the column by vortexing and then placing in a sterile 1.5 ml eppendorf tube. The column was pre-spun for 1 min at 3000 rpm in a microcentrifuge (Eppendorf). The column was placed in a new sterile 1.5 ml tube and 50 μl of the sample was applied to the top-center of the resin, followed by centrifugation at 2000 rpm for 2 min and cpm/μl of labelled probe was determined. Two microlitres of purified probe was added to 98 μl of STE (0.438 g of NaCl in 40 ml of TE to prepare 50

ml of STE) buffer in a eppendorf tube. Fifty microlitre aliquots of the diluted sample were first dispensed into duplicate microcentrifuge tubes, then transferred to scintillation vials and finally counted using Beckman's scintillation counter. The average cpm values were obtained for the duplicates followed by cpm/μl value of the purified, labelled DNA. Finally, total cpm incorporated into the purified sample was calculated by multiplying cpm/μl by the volume of the sample recovered from the column.

$$\% \text{ Label incorporation} = \frac{\text{average cpm for column or purified sample} \times 100}{\text{average cpm for total sample}}$$

For each reaction, more than 55% incorporation was achieved after labelling the probe.

Southern Hybridization

Restriction digested genomic DNA samples were electrophoresed on a 1 % agarose gel in 1X TBE at 25 V (5-7 V per cm) for 16-18 hours. After electrophoresis, depurination of the DNA was done by incubating the gel briefly for 15-20 min in 2 volumes of 0.25 M HCl solution. The gel was rinsed with distilled water to remove excess HCl then incubated in 2 volumes of high salt denaturation solution until the bromophenol blue dye reverted back to blue. The gel was removed from the high salt denaturation solution and placed on a supported wick of paper with the ends of the wick placed in alkaline transfer buffer (0.4 N NaOH). A dry piece of Biodyne membrane was placed directly on the gel and three sheets of blotting paper (Whatman) and a stack of blotting pads (Sigma) were

placed on the top of the membrane. A glass plate was placed over the pads (Sigma) with approximately 300 g of weight placed on the cover for even transfer.

Prehybridization and Hybridization:

After 16-17 hours of transfer, the membrane was baked at 80 °C for 40 min using a Bio-Rad gel dryer and rinsed in 5X SSC just before prehybridization. Denatured salmon sperm DNA (10 mg/ml) was added to the hybridization tube containing membrane and the hybridization buffer (Sigma) and prehybridization was carried out at 65 °C for 3-4 hours. The DNA probe was labelled using random primer labelling (Invitrogen) and purified using Probequant G-50 microcolumns (Amersham Biosciences) as described earlier. The labelled probe was denatured by boiling for 5 min, followed by snap cooling for 1 min and then added to the prehybridization solution. Hybridization was conducted at 65 °C for 16-17 hours. The membrane was washed first with 2X SSC, 0.1 % SDS at 45 °C for 40 min, then at 50 °C for 25 min, then twice with 1X SSC, 0.1% SDS at 50 °C for 25 min, followed by 0.2X SSC, 0.1 % SDS at 55 °C for 30 min. The membrane was exposed to a Kodak BioMax MR X-ray film with an intensifying screen at – 80 °C for 48-72 hours.

PCR Amplification of Douglas-fir cDNA Probe used for Hybridization

Douglas-fir cDNA probe was synthesized using PCR amplification of Douglas-fir cDNA with synthetic primers (Appendix 2) based on the newly obtained cDNA sequence resulting in a product size of 280 bp. The primers were ordered from Invitrogen and the *Taq* polymerase obtained from Qiagen. The following conditions were used for touch

down PCR: 95 °C for 5 min; 94 °C for 30 s, annealing temperature of 51 °C depending upon the melting temperature of the primers, followed by 72 °C for 1 min for 35 cycles; extension time of 72 °C for 15 min. The PCR products were separated on a 1 % agarose gel and compared with DNA markers from Invitrogen. The obtained probe (280 bp) was further used for both Southern and northern Hybridization procedures.

RT-PCR Amplification of Early Embryogenesis Gene from Douglas-fir RNA

First-strand cDNA synthesis using superscript II RT

Total RNA was isolated from different stages of zygotic embryos using Trizol extraction method (Invitrogen). One to two micrograms of total RNA was added to a nuclease-free microcentrifuge tube containing 1 μl Oligo (dT), 1 μl dNTP mix and sterile distilled water to 12 μl. The mixture was heated at 65 °C for 5 min and quickly chilled on ice, followed by brief centrifugation to collect the contents of the tube. Four microlitres of 5X first-strand buffer, 2 μl of 0.1M DTT and 1 μl of RNaseOUT were added to the mixture and incubated at 42 °C for 2 min, followed by the addition of 1 μl of SuperScript II RT (Reverse Transcriptase). The mixture was gently mixed by pipeting up and down and incubated at 42 °C for 50 min. After incubation, the reaction was inactivated by heating at 70 °C for 15 min.

PCR Amplification

The first strand cDNA synthesized was subjected to touch down PCR amplification. PCR amplifications were done using *Taq* polymerase (Qiagen) together with synthetic primers ordered from Invitrogen. All the reagents used for RT-PCR amplification were obtained

from Invitrogen. The following conditions were used for touch down PCR: 95 °C for 5 min; 94 °C for 30 s, annealing temperature starting from 5 °C over the optimal melting temperature of the corresponding primers for 1 min per cycle and reducing it by one degree per cycle to the optimum temperature, followed by 72 °C for 1 min, then 35 cycles at 51 °C (optimum temperature) with same elongation/extension conditions; extension time of 72 °C for 15 min. Touchdown PCR was preferred as this method of PCR amplification involved higher annealing temperatures first and gradually reduced to optimum melting temperatures of the corresponding primers used, so that more time is given for the heterologous primers to bind with the template. The PCR products were separated on a 1 % agarose gel and compared with DNA markers from Invitrogen.

Molecular Cloning

Preparation of *E.coli* competent cells

Escherichia coli DH5-α competent cells were prepared using the following protocol described by Sambrook and colleagues (1989). The cells prepared using this protocol were efficient and useful for long-term storage at –80 °C. Fifty millilitres of LB medium was inoculated with 500 μl of overnight culture of *E.coli* DH5-α cells and incubated at 37 °C overnight with continuous shaking until A_{600} = 0.3-0.5 and then chilled on ice for 10 min. The cells were transferred into sterile centrifuge tubes and centrifuged at 3,000 x g for 10 min at 4 °C followed by resuspension in 25 ml of sterile ice-cold buffer containing 0.1 M magnesium chloride, 10 mM Tris-HCl pH 8.0 and incubated on ice for 10 min. The cells were centrifuged at 3,000 x g for 10 min at 4 °C and the pelleted cells were resuspended in 17 ml of sterile ice-cold buffer containing 0.1M calcium chloride, 10

mM Tris-HCl, pH 8.0 followed by incubation on ice for 30 min. The cells were collected by centrifugation at 3,000 x g for 10 min at 4 °C and resuspended in 17 ml of ice-cold sterile buffer containing 0.1 M calcium chloride, 10 mM Tris-HCl pH 8.0, followed by slow addition of 1 ml of 15 % glycerol with gentle swirling. The cells were dispensed into 100 μl fractions in sterile eppendorf tubes, frozen in liquid nitrogen and stored at –80 °C.

Transformation of *E.coli*

For cloning, competent cells prepared using protocol described by Sambrook and colleagues (1989) and TOP10F' cells (Invitrogen TOPO cloning kit) were used. Cloning was done using conventional protocol (Sambrook *et al*., 1989) and TOPO cloning method (Invitrogen). Five microlitres of the plasmid was added to 50 μl *E.coli* DH5-α competent cells and incubated on ice for 45 min. Cells were heat shocked by placing them first at 42 °C for 90 s and immediately cooled on ice for 2 min. One millilitre of LB medium was added and the cells were incubated at 37 °C for 1 hour with gentle shaking. The cells were collected by brief centrifugation at room temperature, resuspended in 200 μl of LB broth and plated on selective plates containing 50 μg/ml kanamycin. Forty microlitres of 20 mg/ml X-gal dissolved in dimethylformamide was spread on different selective LB plates and incubated at 37 °C for at least 30 minutes before plating cells. The positive clones were identified using a blue-white screening system. All the white colonies obtained were inoculated in 2 ml of LB broth and incubated overnight at 37 °C for plasmid DNA preparation. Alternatively, cloning was also done using a TOPO cloning kit (Invitrogen) using chemically competent One Shot TOP10F' *E.coli* cells. TOPO

cloning method was preferred over the conventional method as better results were obtained using this method. The cDNA fragments obtained from RT-PCR amplification were cloned using TOPO cloning kit and the positive colonies were subjected to further investigation. The TOPO cloning reaction was prepared by mixing 4 μl of fresh PCR product, 1 μl of salt solution, 1 μl of TOPO vector and incubated for 10 min at room temperature. The reaction was placed on ice until proceeding with transformation of One Shot competent cells. The TOPO cloning reaction could be stored at –20 °C overnight. Two microlitres of TOPO cloning reaction was added into a vial containing One Shot chemically competent *E.coli* cells and mixed gently. After incubation on ice for 15 min, the cells were heat-shocked for 30 s at 42 °C without shaking. The tubes were transferred immediately to ice and 250 μl of room temperature SOC medium was added and incubated horizontally (200 rpm) at 37 °C for 1 hour. The selective LB plates were warmed at 37 °C for 30 min and 40 μl of 40 mg/ml X-gal was spread on each LB plate, followed by incubation at 37 °C before spreading. Ten to 120 μl from each transformation was spread on a prewarmed selective plate and incubated overnight at 37 °C. Approximately, 20-30 white colonies were picked and inoculated in LB medium containing 100 μg/ml ampicillin or 50 μg/ml kanamycin overnight at 37 °C in a shaking incubator. The plasmid DNA was isolated as per the procedure discussed earlier. The plasmids were analysed by restriction digestion with *Eco*RI, and positive clones were identified visually using agarose gel electrophoresis and confirmed by the appropriate insert size.

Bioinformatic Analysis

Base-By-Base:

The protein alignment for the identified Douglas-fir gene with closely related species was done using Base-By-Base program (Brodie *et al.*, 2004) developed in Dr. Chris Upton's lab. The Base-By-Base (BBB) program has different tools for aligning protein sequences such as CLUSTALW (Thompson *et al.*, 1994), T–coffee (Notredame *et al.*, 2000) and Muscle. The protein sequences from closely related organisms were converted in FASTA format and subjected to CLUSTALW alignment. Both FASTA and CLUSTAW (.aln) formatted text files can be loaded into the program along with multiple sequences, placed in a single file and additional sequences can be added at any time. The alignments may be exported in either of these formats and the native BBB file format stores the sequence alignment, gene features and other specific annotations for the sequences. Base-By-Base is a pairwise and multiple alignment editor program. The program elucidates the differences between pairs of alignments and it allows the user to navigate large alignments of similar sequences. This software also provides visualization tools to enable researchers to identify and rectify alignment errors in large, multiple genome alignments, and provides graphical and tabular output differences between the genomes at the nucleotide level (Brodie *et al.*, 2004). By utilising the annotation information about the aligned genomes, this software can list each predicted gene with nucleotide differences and elucidates whether differences occur within promoter regions or coding regions and whether these changes result in any amino acid substitutions. This software is available via the Internet using Java Web Start and runs on Macintosh, PC and Linux operating systems with the Java 1.4 virtual machine. BBB provides access to both CLUSTALW

(Thompson *et al.*, 1994 and 1997) and T-coffee (Notredame *et al.*, 2000) software. Entire genome sequences or specific regions can be selected and aligned within the program.

InterPro:

InterPro is a database containing protein domains and functional sites that act as vital resources for the prediction of protein functions (Mulder *et al.*, 2005). Earlier, several signature-recognition methods have evolved to address different sequence analysis problems, resulting in rather different and, in most cases, independent databases. Diagnostically, these resources have different areas of optimum application corresponding to the different strengths and weakness of their underlying analysis methods. In order to obtain best results, a search strategy should involve all of them. InterPro (Mulder *et al.*, 2005) is a collaborative scanning system aimed at providing an integrated layer involving the most commonly used signature databases by creating a unique, non-redundant characterization of a given protein family, domain or functional site. The InterPro database integrates PROSITE, PRINTS, UniProt, Pfam, ProDom, SMART, TIGRFAMs, PIR superfamily, SUPERFAMILY, Gene 3D and PANTHER (Quevillon *et al.*, 2005). In this study, Pfam and UniProt (Wu *et al.*, 2006) databases of protein domain families were used. Pfam contains multiple sequence alignments for each family and corresponding hidden Markov models (HMMs). HMMs are statistical models of the primary structure consensus of a sequence family (Finn *et al.*, 2006). The protein sequences obtained were converted into FASTA format and used for scanning using the Pfam database. Examples of sequences that are depicted in FASTA format are:

>LVR clone gene_1|GeneMark.hmm|141_aa

MASPLRQVGTGANSQSLDEARKRVFHFFREACRSIPQIMETYNLHEVITPSQLRS
VVAAQFRKQAHVTNPKVIDMLIIKGDEELRNCLDHSKQRHHIVGQYVIGQEGLIP
SNVGAVSSGGSDFLRKFYDTLNVRRDFGDKE

>Arabidopsis thaliana catalytic/ oxidoreductase, acting on NADH or NADPH

MAAPFALRKIGVPPNSANLTEARRRVFDFFRAACRSIPTIMDIYNLQDVVAPSQL
RYAISAQIRNNAHITDPKVIDLLIFKGMEELTDIVDHAKQRHHIIGQYVVGEGLVQ
NTGNKDQGKTDFLKNFYTSNYF

CLUSTALW:

One of the breakthrough of modern bioinformatics is the comparison or alignment of both nucleotide and protein sequences. Using multiple sequence alignment programs, researchers can study the sequence pattern conserved through evolution and relationships between different organisms. The corresponding sequences could be aligned only in certain regions (local alignment) or can be aligned across the entire length (global alignment) (Chenna *et al.*, 2003). The CLUSTAL series of programs are also used for the preparation of phylogenetic trees. The new features include FASTA and NEXUS format output, printing the range numbers and faster tree calculation.

CLUSTALW improves the sensitivity of the commonly used multiple sequence alignment by assigning individual weights to each sequence in a partial alignment in order to up-weight the most divergent sequences and down-weight close-duplicate

sequences. Using the file menu containing the option "add sequences to alignment from your local file in FASTA format" in the Base-By-Base program, the Douglas-fir LYR protein sequence could be added along with other protein sequences from different organisms of interest. The amino acid substitutions are varied at different alignment stages with respect to the divergence of the sequences to be aligned (Thompson *et al.*, 1994). The basic alignment method involves the following important stages: a) each pair of sequences are aligned separately so that a distance matrix providing the divergence of each pair of sequences could be calculated; b) a phylogenetic tree is calculated from the distance matrix; c) the sequences were aligned based on the branching order depicted in the guide tree.

Chapter 3

RESULTS

CLONING OF AN EARLY EMBRYOGENESIS GENE OF DOUGLAS-FIR

Isolation of Douglas-fir LYR/LVR cDNA using RT-PCR amplification

Characterization of early embryogeny genes is necessary for understanding the molecular and biochemical processes in conifer embryogenesis. Various PCR products were obtained from RT-PCR amplification of total RNA obtained from early embryonic stages of Douglas-fir using degenerate primers based on *Arabidopsis thaliana* COP1 (Constitutively Photomorphogenic 1) cDNA. One of the product was found to be a LYR/LVR cDNA based on (Appendix 2) similarity with *Arabidopsis thaliana* LYR/LVR gene. Further work was done to obtain the cDNA fragment. Two microlitres of the first strand cDNA synthesized from total RNA obtained from early-stage embryos (collected in May and June) was PCR amplified using touchdown RT-PCR (Materials and Methods) (**Figure 2A**). The two cDNA fragments approximately of 679 bp fragment and an 870 bp were cloned and sequenced. The positive controls used for the reaction included, a 680 bp fragment obtained using PCR amplification of *Arabidopsis thaliana* seedling RNA with same set of primers (**Figure 2B**).

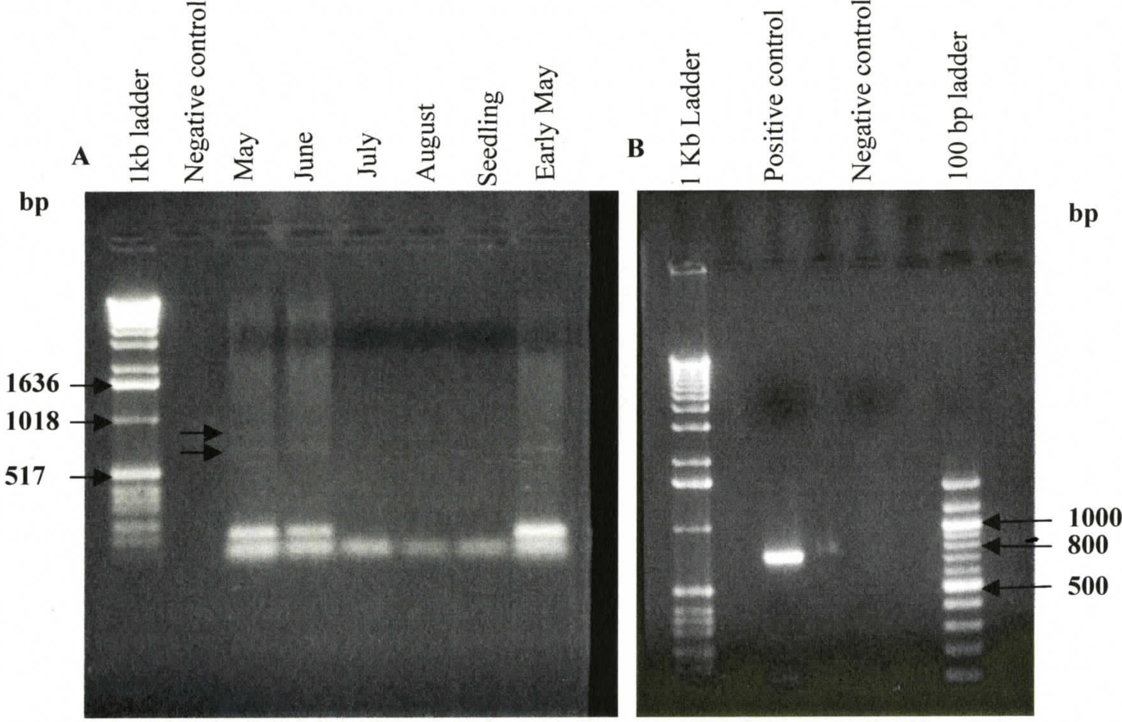


Figure 2 (A-B). RT-PCR using total RNA isolated from different stages of zygotic embryos with a set of primers (Appendix 2) based on *Arabidopsis thaliana* LYR gene. **A.** 1 % agarose gel showing cDNAs obtained from total RNA isolated from various stages of seed development. Early May is mRNA obtained from megagametophyte and embryos. Two fragments of approximately 679 bp and 870 bp from RNA isolated from May, June and early May are shown (arrow heads shown in lane 3). **B.** 1 % agarose gel showing both positive and negative controls used for this reaction.

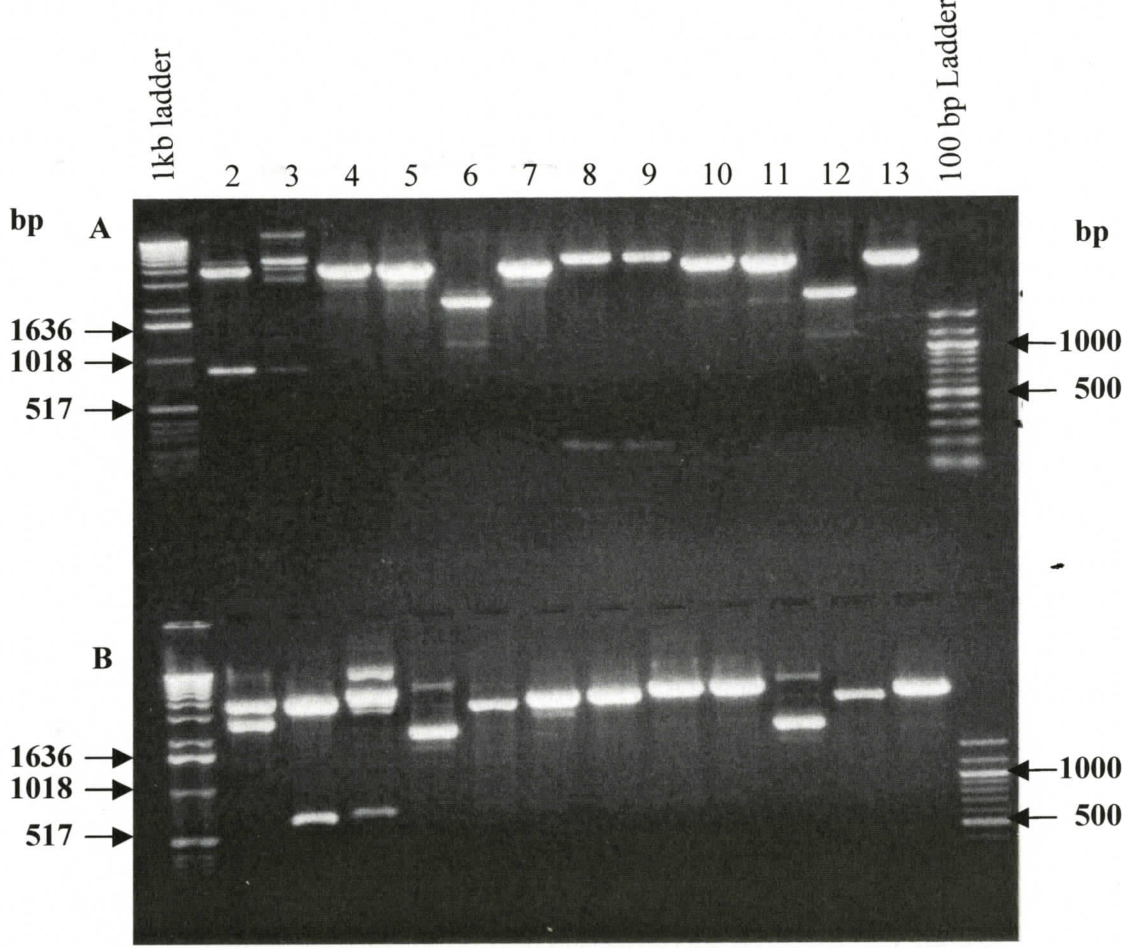


Figure 3 (A-B). 1% agarose gel electrophoresis of restriction digested plasmids from positive white colonies. The cDNA clones were restriction digested with *Eco*RI. Lanes 2-13 correspond to the restriction digestion of plasmids isolated from different positive white colonies. **A.** Lanes 2-3, ~ 870 bp. **B.** Lanes 3-4, ~ 679 bp insert.

The negative control included PCR master mix and other reagents without any template RNA (**Figure 2B**).

Cloning of Douglas-fir LYR/LVR gene

The 679 bp Douglas-fir LYR/LVR cDNA and the 870 bp cDNA fragments were cloned using the TOPO cloning vector (see Materials and Methods). The white colonies observed on plates with different dilutions were screened for the presence of the 679 bp and 870 bp inserts (**Figure 3 A-B**). The selected white colonies were subcultured in 5 ml LB broth containing 100 μg/ml ampicillin and plasmids were isolated from the corresponding subcultures. Plasmid isolation was followed by restriction digestion with *EcoR*I and different samples were separated electrophoretically on 1% agarose gel in 1 X TBE. One of the four positive colonies containing the 679 bp insert was selected and the plasmid DNA was isolated and sequenced. The selected colony was subcultured in LB broth containing 100 μg/ml ampicillin and incubated overnight at 37 °C. After incubation, five millilitres of the culture was subjected to glycerol stock preparation and stored at -80 °C.

Sequence Analysis of Douglas-fir LVR/LYR gene

The 679 bp cDNA fragment was identified as Douglas-fir LYR/LVR cDNA and the 870 bp fragment was identified as an unrelated protein. Hence, the 679 bp cDNA fragment was subjected to further analysis. The positive clone was sequenced at the DNA Sequencing Centre, Centre for Biomedical Research, University of Victoria (**Figure 4**). The deduced Douglas-fir LYR/LVR DNA sequence was subjected to NCBI BLASTN

Figure 4. Douglas-fir LYR/LVR cDNA sequence and Translated +3 frame of the LYR/LVR protein.

A. Nucleotide Sequence

CCGATTTGTATGCAGATTATTTGCCACCGAAGAGGAGAGAGAAGAGGCGGAGCTCGCAG
TTGGCCATGGCGTCCCCGTTGAGACAGGTGGGAACAGGAGCAAATTCACAGTCACTGGA
TGAGGCGAGAAAGCGGGTTTTTCACTTCTTTAGGGAAGCCTGCCGTTCCATCCCACAGA
TCATGGAGACCTACAACCTCCATGAAGTAATTACCCCTTCTCAGTTGCGTTCTGTTGTA
GCTGCCCAGTTTCGAAAGCAGGCCCATGTTACCAATCCAAAGGTTATTGACATGCTTAT
TATCAAAGGGGATGAAGAACTACGGAACTGTCTTGATCATTCAAAGCAGCGGCATCACA
TTGTAGGCCAATATGTTATTGGTCAAGAAGGGTTAATACCATCAAACGTAGGTGCTGTC
AGCAGTGGAGGTTCAGATTTTCTGCGGAAGTTCTATGATAGTAATAACTTCTGATTTAG
GCAACAACATTAGAGATGTCATGTTCTGTCAAAACTTTTTATGCAGATACTATTGGGAT
GTCTGATTAGGCTGTTAATCTAGAAGTGCATTTAAAGTATACAGCAAGATCTCTCACTG
TGTTTGTTGTTTATGCTTGCTTTTTTATTGATCAGCACTGAATGTTAGAAGGATTTTG
GAGACAAAGAATAATCTGCATACAAATCGG

B. Amino acid Sequence Translated in the +3 Frame

MASPLRQVGTGANSQSLDEARKRVFHFFREACRSIPQIMETYNLHEVITPSQLRSVVAA
QFRKQAHVTNPKVIDMLIIKGDEELRNCLDHSKQRHHIVGQYVIGQEGLIPSNVGAVSS
GGSDFLRKFYDSNNFFRQQHRCHVLSKLFMQILLGCLIRLLIKCISIQQDLSLCLLFML
AFLLISTECKGFWRQRIICIQI

(Altschul *et al.*, 1997) analysis and then translated using Six Frame Translation tool. The frame +3 of the LYR protein was used and subjected to protein-protein blast (BLASTP), which showed approximately 59% amino acid sequence identity over 134 residues with *Arabidopsis thaliana* LVR family protein encoding NADH oxidoreductase (NCBI and TIGR locus tag-AT3G12260, Gene ID- 820406) (ref. NP566416.1/ Unigene At48649).

BIOINFORMATIC ANALYSIS OF LYR/LVR GENE

The Base-By-Base program developed in Dr. Chris Upton's lab, University of Victoria was used to perform multiple sequence alignment of the Douglas-fir LYR/LVR protein sequence. The protein sequences of Douglas-fir LYR and other closely related species were converted to FASTA format using text edit tool. Using, Base-By-Base program file menu, the protein sequences were added and after complete selection, multiple sequence alignment was done using the CLUSTALW tool available in the program. The protein alignment for Douglas-fir LYR with respect to *Arabidopsis thaliana* (gi/18399551) and *Oryza sativa* (gi/50510168) LYR/LVR is shown in **Figure 5**. The protein sequence of Douglas-fir LYR was also compared with LYR protein sequences from other groups of living organisms (**Figure 6**) apart from plants such as; *Ustilago maydis* (gi/46098242), *Ixodes ricinus* (gi/33521688), *Homo sapiens* (gi/49456947) and *Rattus norvegicus* (gi/27663138). The multiple sequence alignment indicates that amino acids from 19-90 in the LYR/LVR protein are highly conserved (**Figure 5**). The sequence identity table generated using the Base-By-Base program (**Table 2 and 3**) provided information regarding the similarity between the amino acids encoding the Douglas-fir LYR protein

CONSERVED REGION

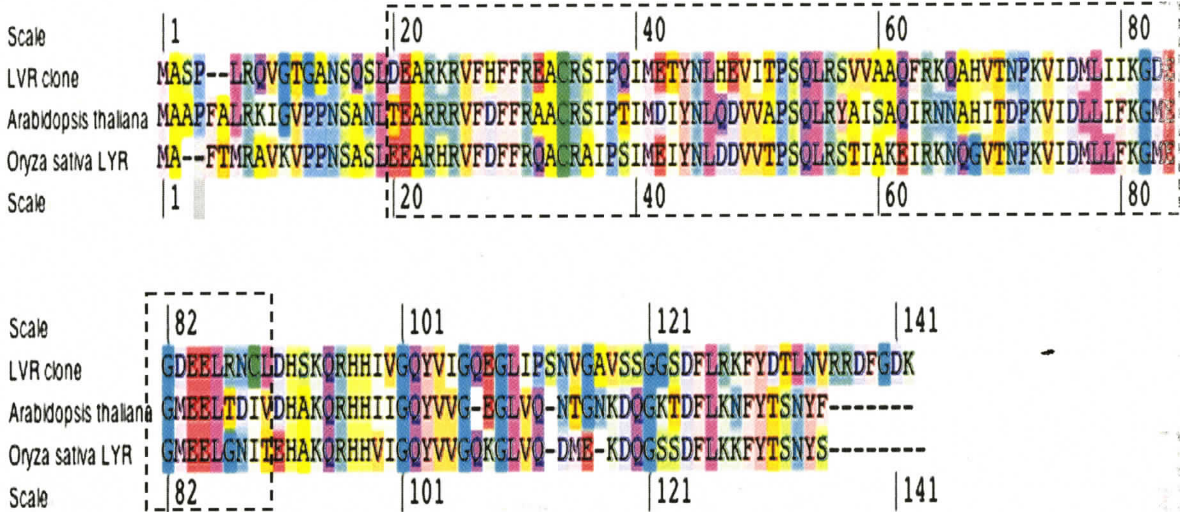


Figure 5. CLUSTALW alignment of the Douglas-fir LYR/LVR amino acid sequence with *Arabidopsis thaliana* and *Oryza sativa* LYR proteins. The amino acid sequences of the compared species were converted to FASTA format, followed by CLUSTALW alignment using Base-By-Base program. The alignment was confirmed by Interpro scanning using the Pfam database. Amino acids glycine (G) and asparagines (N) are represented in cyan, amino acids proline (P) and aspartic acid (D) are represented in blue, amino acids lysine (K), arginine (R), serine (S) and cysteine (C) are represented in green, amino acids glutamic acid (E), valine (V), threonine (T) and phenylalanine (F) are represented in orange, amino acids isoleucine (I), alanine (A) and histidine (H) are represented in yellow, amino acids leucine (L), methionine (M) and tyrosine (Y) are represented in pink, amino acid glutamine (Q) is represented in violet. The dotted lines indicate amino acids 19-90 that are highly conserved in the LYR proteins in Douglas-fir, *Arabidopsis* and rice.

LYR Repeat

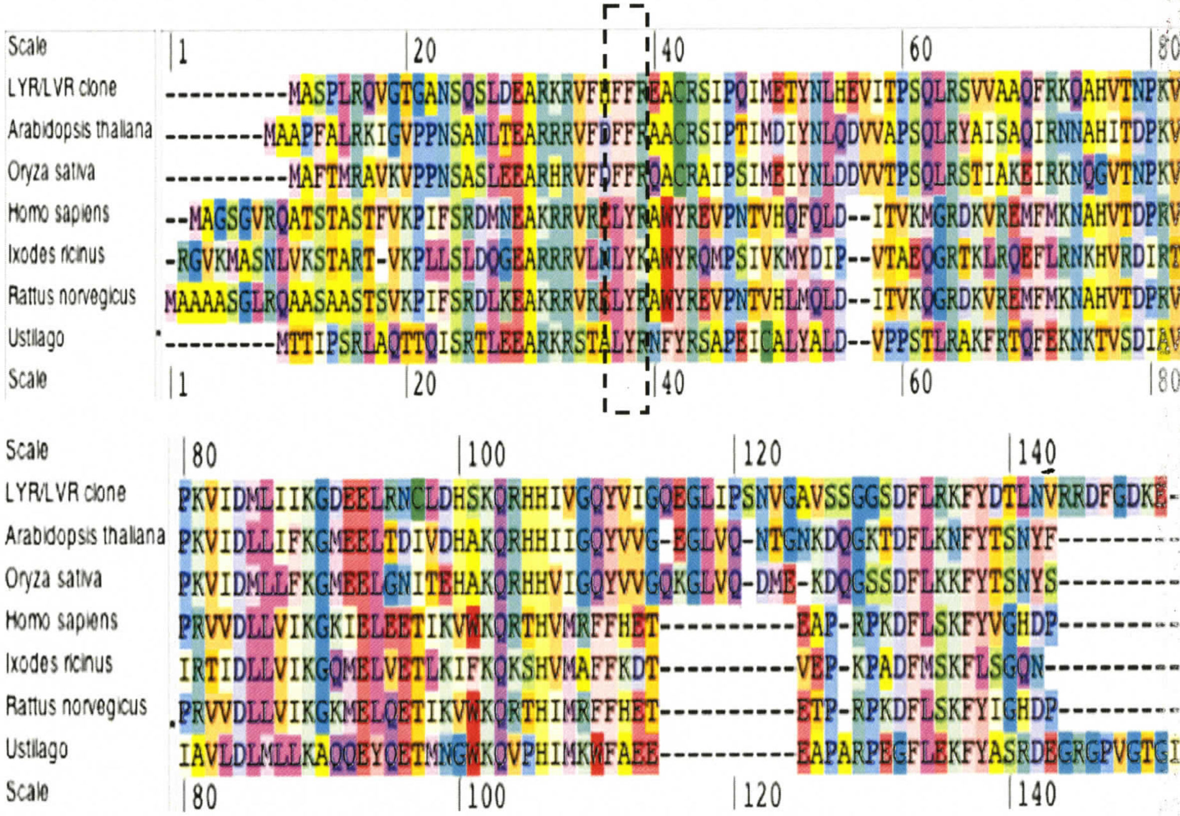


Figure 6. CLUSTALW alignment of Douglas-fir LYR/LVR amino acid sequence with LYR/LVR protein of other related organisms. CLUSTALW tool in Base-By-Base program was used to align the amino acid sequences of LYR protein in FASTA format from different organisms. This data was used to construct a phylogenetic tree. Amino acids glycine (G) and asparagines (N) are represented in cyan, amino acids proline (P) and aspartic acid (D) are represented in blue, amino acids lysine (K), arginine (R), serine (S) and cysteine (C) are represented in green, amino acids glutamic acid (E), valine (V), threonine (T) and phenylalanine (F) are represented in orange, amino acids isoleucine (I), alanine (A) and histidine (H) are represented in yellow, amino acids leucine (L), methionine (M) and tyrosine (Y) are represented in pink, amino acid glutamine (Q) is represented in violet. The dotted lines indicate amino acids (37-49) L, Y, R and F, F, R (plants).

PERCENT IDENTITY SCORES							
Sequences	1	2	3	4	5	6	7
1	100%	56.49%	56.78%	32.5%	29.66%	35.83%	30.23%
2	56.49%	100%	48.46%	30.33%	28.33%	33.61%	29.51%
3	56.78%	68.46%	100%	31.09%	30.77%	31.93%	30.83%
4	32.5%	30.33%	31.09%	100%	48.41%	87.5%	38.84%
5	29.66%	28.33%	30.77%	48.41%	100%	48.03%	33.61%
6	35.83%	33.61%	31.93%	87.5%	48.03%	100%	40.5%
7	30.23%	29.51%	30.83%	38.84%	33.61%	40.5%	100%

Table 2. Percentage amino acid sequence identity of the Douglas-fir LYR/LVR protein with LYR/LVR proteins of other related organisms. This data was generated by performing amino acid alignment CLUSTALW and the percentage identity tool in the Base-By-Base program. Numbers 1-7 denote the amino acid sequences from different organisms, 1- Douglas-fir LYR/LVR protein, 2- *Arabidopsis thaliana*-NADH oxidoreductase (LYR family), 3- *Oryza sativa* LYR protein, 4- *Homo sapiens* NADH-ubiquinone oxidoreductase, 5- *Ixodes ricinus* NADH-ubiquinone oxidoreductase, 6- *Rattus Norvegicus* NADH- ubiquinone-1 subcomplex, 7- *Ustilago maydis* Hypothetical complex1 (LYR).

SEQUENCES	*Arabidopsis thaliana*-NADH oxidoreductase (LYR family)	Douglas-fir LYR/LVR protein	*Oryza sativa* LYR protein
Arabidopsis thaliana-NADH oxidoreductase (LYR family)	100%	58.76%	70%
Douglas-fir LYR/LVR protein	58.76%	100%	59.69%
`*Oryza sativa* LYR protein	70%	59.69%	100%

Table 3. Percentage identity scores between LYR/LVR amino acid sequences of Douglas-fir, *Arabidopsis thaliana* and *Oryza sativa*. The amino acid sequences of Douglas-fir LYR, *Arabidopsis* and rice were aligned using CLUSTALW tool in Base-By-Base program. The percentage identity scores of the different protein sequences of the plants compared were determined using Base-By-Base program. Douglas-fir LYR/LVR protein shows approximately 59 % identity with *Arabidopsis thaliana* and 60 % identity with *Oryza sativa* LYR proteins, respectively.

when compared to other organisms. The Douglas-fir LYR protein showed approximately 59 % amino acid similarity to *Arabidopsis thaliana* LYR protein and 60 % identity to *Oryza sativa* protein sequence. The percent identity scores of LYR/LVR protein in plant group (**Table 3)** and other organisms (**Table 2**) were calculated to compare the percent identity scores among the different groups of organisms as this software does not count the gaps observed in the multiple sequence alignment (**Figures 5 and 6**). There is no significant difference in the identity scores as the Douglas-fir LYR/LVR percent identity is increased by approximately 1 % (**Table 3**) when compared with LYR/LVR protein in other group of organisms (**Table 2**) apart from plant group.

Analysis of the LYR/LVR protein

Characterization of LYR protein was done using InterPro scanning system which is an integrated system using different databases providing information about the protein. The amino acid sequence of the Douglas-fir LYR/LVR protein was converted to FASTA format and scanned using Pfam database (Finn *et al.*, 2006) and compared with amino acids from closely related species. The Pfam database showed possible similarities of Douglas-fir LYR with catalytic oxidoreductase enzyme constituting the LYR complex in *Arabidopsis thaliana* and other closely related species. Comparison of sequences using UniGene suggested that the Douglas-fir LYR/LVR protein may be the catalytic oxidoreductase acting on NADH or NADPH. The InterPro scanning system (IPR008011) showed possible similarities and functions of Douglas-fir LYR protein with 159 closely related proteins in LYR complex I in different organisms. The InterPro scanning (Quevillon *et al.*, 2005) also resulted in determining the taxonomic coverage of Douglas-

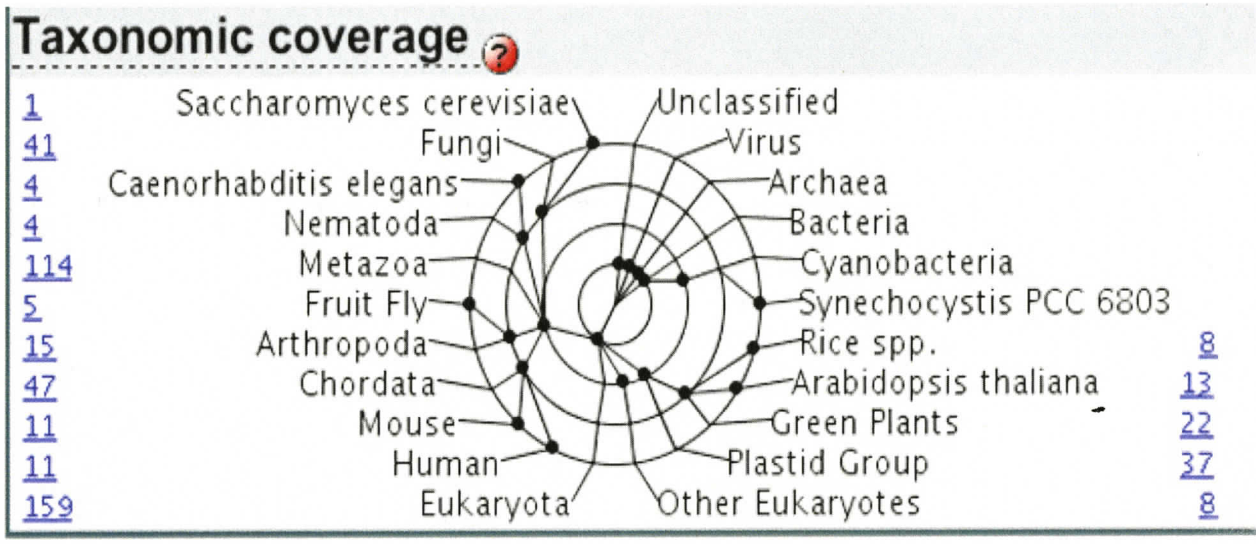
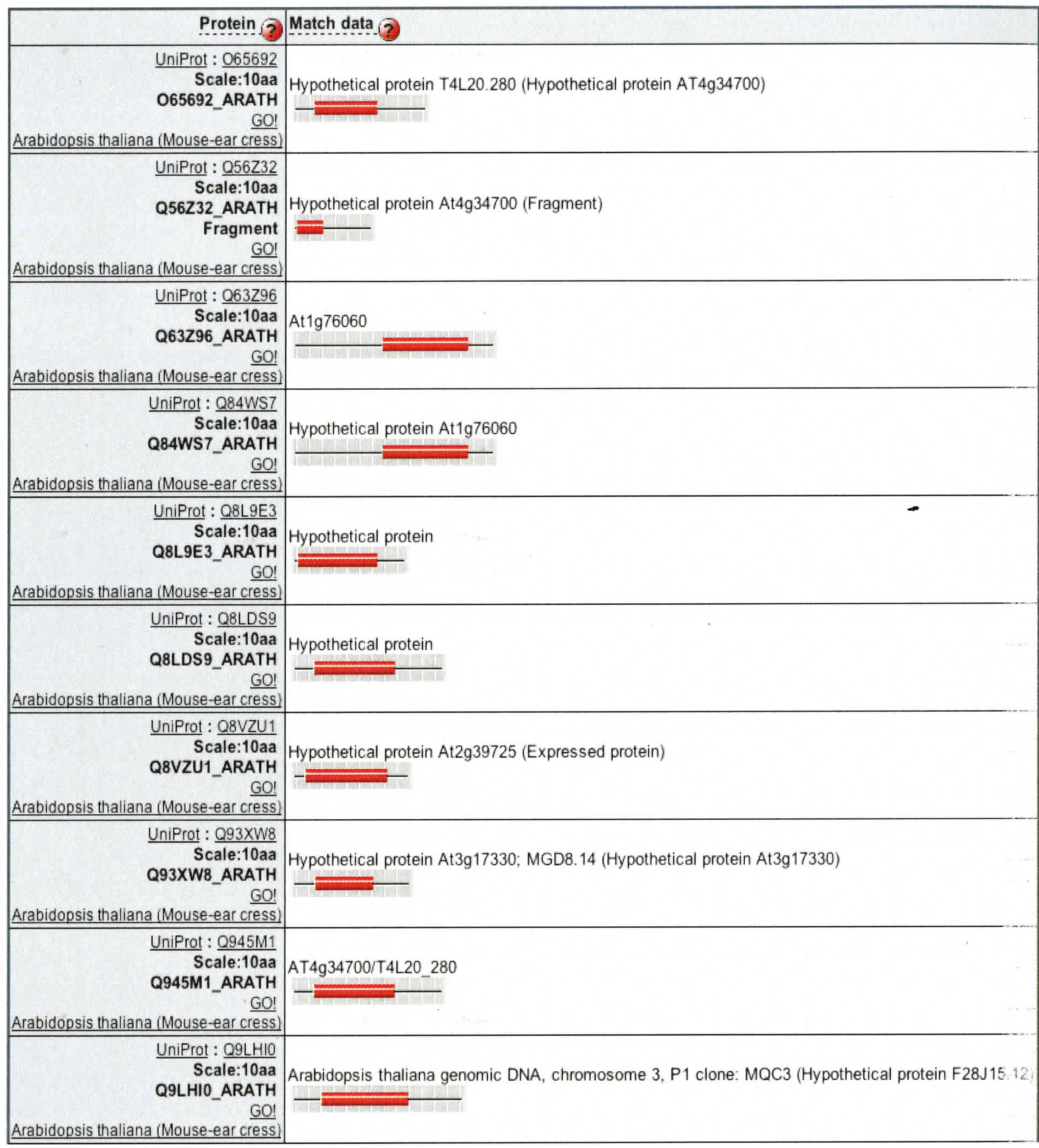
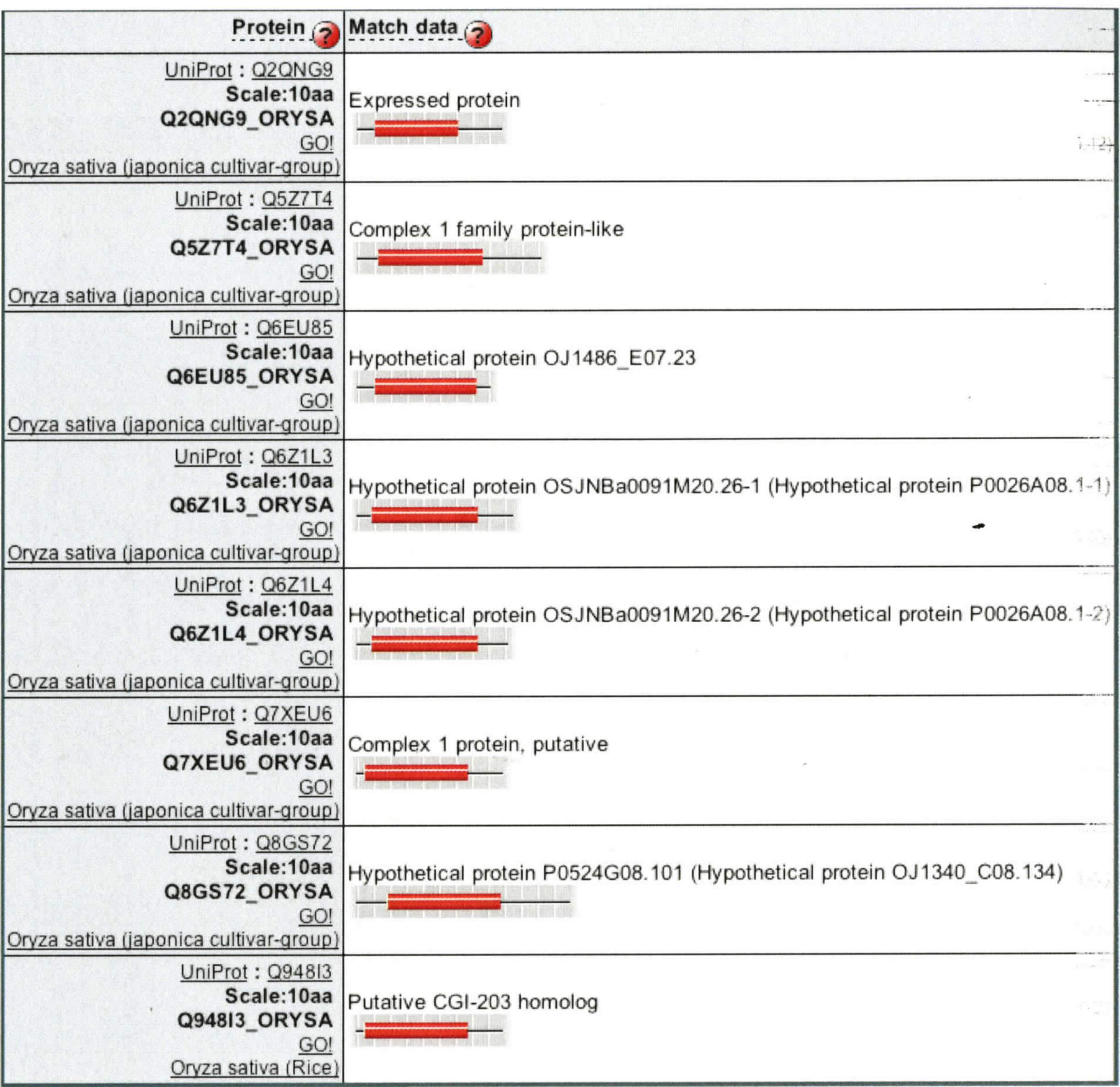


Figure 7. Taxonomic coverage of Douglas-fir LYR/LVR protein matches with other closely related known proteins of the same LYR family (grouped under respiratory complex 1) in different organisms. The numbers indicate total number of closely associated LYR/LVR proteins in individual organism constituting the LYR complex (Pfam ID: is PF05347). The total number of proteins identified in *Arabidopsis thaliana* and *Oryza sativa* are 13 and 8 respectively.



Protein	Match data
UniProt : O65692 Scale:10aa O65692_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein T4L20.280 (Hypothetical protein AT4g34700)
UniProt : Q56Z32 Scale:10aa Q56Z32_ARATH Fragment GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein At4g34700 (Fragment)
UniProt : Q63Z96 Scale:10aa Q63Z96_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	At1g76060
UniProt : Q84WS7 Scale:10aa Q84WS7_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein At1g76060
UniProt : Q8L9E3 Scale:10aa Q8L9E3_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein
UniProt : Q8LDS9 Scale:10aa Q8LDS9_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein
UniProt : Q8VZU1 Scale:10aa Q8VZU1_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein At2g39725 (Expressed protein)
UniProt : Q93XW8 Scale:10aa Q93XW8_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Hypothetical protein At3g17330; MGD8.14 (Hypothetical protein At3g17330)
UniProt : Q945M1 Scale:10aa Q945M1_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	AT4g34700/T4L20_280
UniProt : Q9LHI0 Scale:10aa Q9LHI0_ARATH GO! Arabidopsis thaliana (Mouse-ear cress)	Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MQC3 (Hypothetical protein F28J15.12)

Figure 8. InterPro scanning showing different LYR proteins of the LYR complex family in *Arabidopsis thaliana*. InterPro scan was used to determine the *Arabidopsis thaliana* protein subunit that showed maximum amino acid similarity with Douglas-fir LYR protein.



Protein	Match data
UniProt : Q2QNG9 Scale:10aa **Q2QNG9_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Expressed protein
UniProt : Q5Z7T4 Scale:10aa **Q5Z7T4_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Complex 1 family protein-like
UniProt : Q6EU85 Scale:10aa **Q6EU85_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Hypothetical protein OJ1486_E07.23
UniProt : Q6Z1L3 Scale:10aa **Q6Z1L3_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Hypothetical protein OSJNBa0091M20.26-1 (Hypothetical protein P0026A08.1-1)
UniProt : Q6Z1L4 Scale:10aa **Q6Z1L4_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Hypothetical protein OSJNBa0091M20.26-2 (Hypothetical protein P0026A08.1-2)
UniProt : Q7XEU6 Scale:10aa **Q7XEU6_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Complex 1 protein, putative
UniProt : Q8GS72 Scale:10aa **Q8GS72_ORYSA** GO! Oryza sativa (japonica cultivar-group)	Hypothetical protein P0524G08.101 (Hypothetical protein OJ1340_C08.134)
UniProt : Q948I3 Scale:10aa **Q948I3_ORYSA** GO! Oryza sativa (Rice)	Putative CGI-203 homolog

Figure 9. InterPro scanning to determine different LYR proteins in *Oryza sativa* constituting the LYR complex. InterPro scanning was used to identify other related LYR subunits in rice and to determine the subunit that shows highest amino acid similarity with Douglas-fir LYR/LVR protein sequence.

fir LYR protein matches with its closely related members of the LYR family in different organisms (**Figure 7**). The taxonomic diagram indicated 13 closely related proteins in *Arabidopsis thaliana* and 8 protein subunits in *Oryza sativa* constituting the LYR or respiratory complex I identified to date. In *Arabidopsis*, 10 among 13 subunits were compared with the Douglas-fir LYR/LVR protein. The different subunits of the LYR (complex I) in *Arabidopsis thaliana* and *Oryza sativa* identified were compared with the Douglas-fir LYR/LVR protein sequence (**Figure 8 and Figure 9**). This data was further confirmed using UniProt (Wu *et al*., 2006) database that showed Douglas-fir LYR/LVR protein sequence similarity with one specific protein subunit Q9LH10 and Q8GS72 among the LYR complex family in *Arabidopsis thaliana* and *Oryza sativa*, respectively. The two protein subunits in *Arabidopsis thaliana* and *Oryza sativa* showed conserved region (amino acids 19-90) relatively similar to that of Douglas-fir LYR/LVR protein.

Phylogenetic Analysis

Phylogenetic analysis of the Douglas-fir LYR/LVR clone was performed using the Neighbour joining tree tool in the Base-By-Base program. The amino acid sequences of different closely related organisms were obtained and converted to FASTA format. After addition of the amino acid sequences, multiple sequence alignment using CLUSTALW and T- Coffee was done. The alignment obtained using CLUSTALW was preferred. The entire protein sequences were selected and phylogenetic analysis was performed using the Neighbour Joining Tree tool available in "Report" menu. The Neighbour Joining Tree provided a basic phylogenetic tree comparing the newly identified Douglas-fir LYR/LVR clone with closely related sequences from other groups of organisms containing the LYR

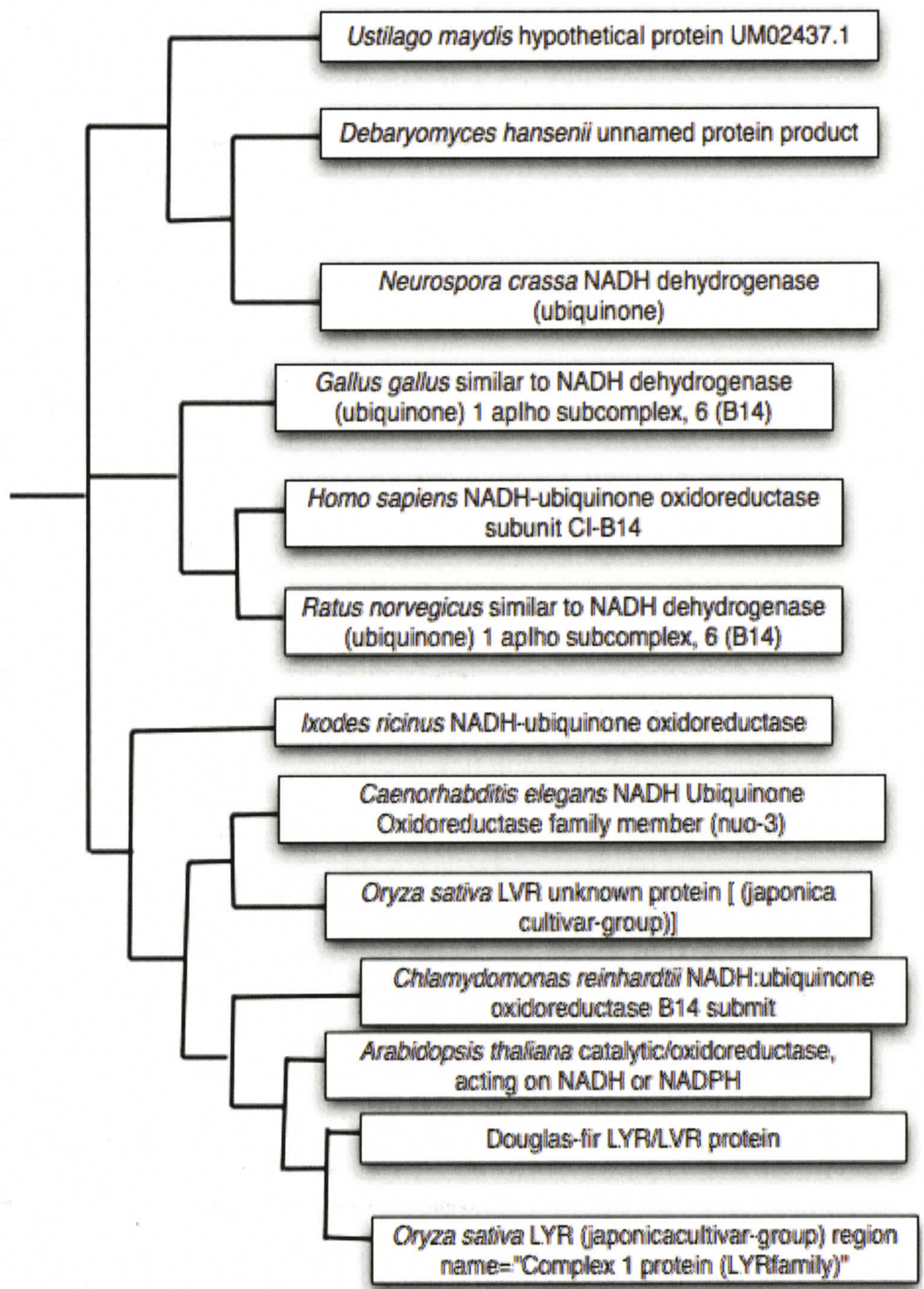


Figure 10. Phylogenetic tree analysis of LYR proteins. The phylogenetic tree was constructed using the Neighbour joining tree in Base-By-Base program. The amino acid sequences of Douglas-fir LYR/LVR with respect to other closely related species was aligned using CLUSTALW alignment in the same program.

protein (**Figure 10**). The phylogenetic analysis showed that the LYR/LVR gene has been conserved in different taxa over the course of evolution. This suggests that LYR/LVR gene encodes a catalytic enzyme that is necessary for various metabolic and cellular processes like respiration.

CHARACTERIZATION OF DOUGLAS-FIR LYR/LVR GENE

Genomic Southern Blot Analysis

To determine the presence and understand the organization of the LYR/LVR gene in Douglas-fir, genomic DNA was digested with the following restriction enzymes namely *Eco*RI, *Bam*HI, *Hind*III, *Xba*I, *Bgl*II, *Eco*RI and *Hind*III, *Bam*HI and *Hind*III and analyzed by Southern hybridization. After hybridization and washing procedures (see Materials and Methods), the blot was exposed for 4-5 days. The X-ray film showed the presence of 2 strongly reactive bands along with other weakly reactive bands per lane (**Figure 11**), indicating the presence of a gene family. Also, Bioinformatic analysis showed that the LYR proteins are grouped under respiratory complex 1 proteins and this family includes short stretch of closely related proteins. The weaker bands observed in the Southern blot may be due to lower stringency conditions used for washing the membrane.

Northern Blot Analysis

The developmental expression of LYR/LVR gene in Douglas-fir was examined by northern hybridization of total RNA isolated from different stages of zygotic embryos.

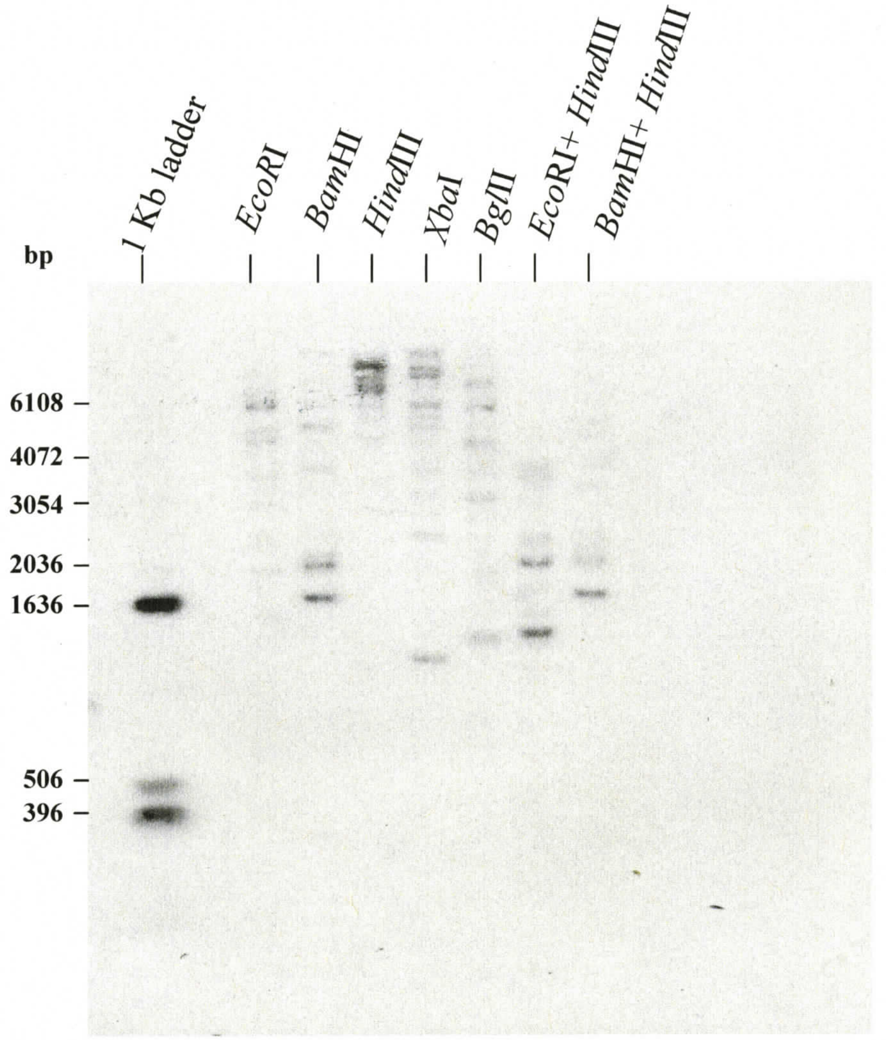


Figure 11. Genomic Southern analysis of Douglas-fir LYR/LVR. Genomic DNA (25μg/lane) was digested with restriction enzymes as indicated and separated on 1 % agarose gel. The LYR/LVR probe amplified using primers based on Douglas-fir LYR/LVR cDNA was used. Sizes are based on the migration of 1 kb DNA marker (Invitrogen) and indicated in base pairs.

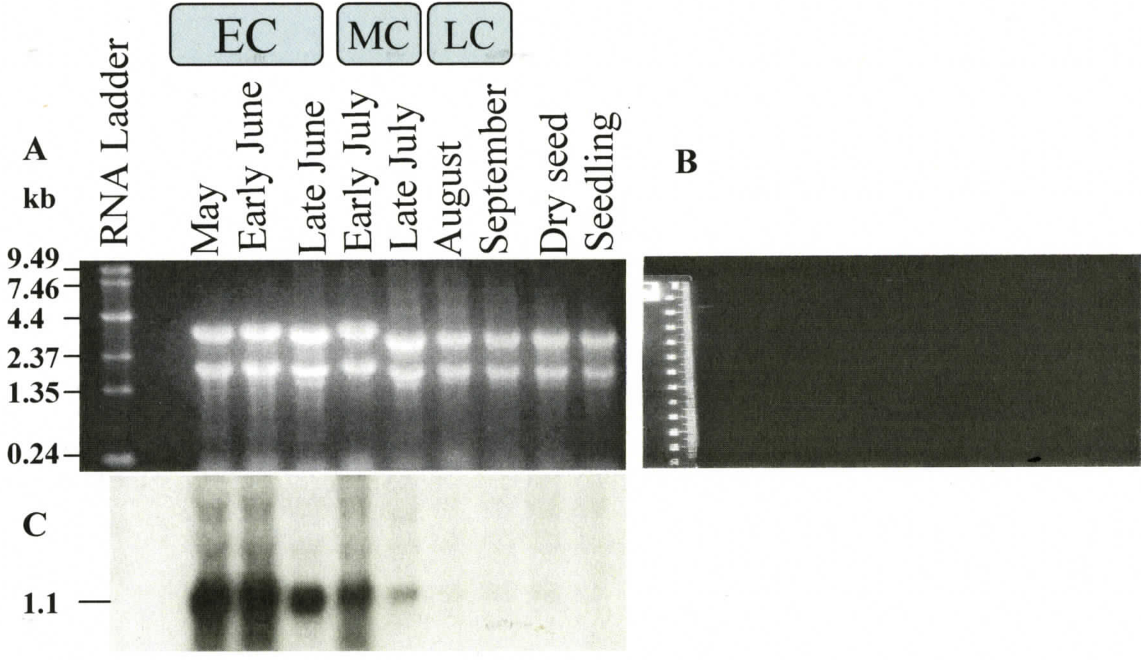


Figure 12 (A-B-C). Northern blot analysis of LYR/LVR gene at different stages of Douglas-fir zygotic embryos and other tissues. **A)** Ethidium bromide stained gel. Total RNA (25µg/lane) from different stages of zygotic embryos was separated on formaldehyde-agarose gel used, transferred to a nitrocellulose membrane and probed with ^{32}-P labeled Douglas-fir LYR/LVR fragment amplified using primers based on Douglas-fir LYR/LVR cDNA. **B)** Agarose gel after transfer and stained with ethidium to show equal transfer. **C)** Autoradiogram of Northern blot. EC- early cotyledonary embryonic stage, MC- mid cotyledonary stage and LC- late cotyledonary embryonic stage. Sizes are based on the migration of 0.24-9.5 Kb RNA marker (Invitrogen) and indicated in kilobases. The size of the LYR/LVR transcript is approximately 1-1.1 kb.

The differential expression profile was followed from before fertilization until seed maturation. In the early embryonic stage (EC), the Douglas-fir LYR/LVR transcripts were detected at a much higher level when compared to LYR/LVR transcripts in the mid cotyledonary (MC) and late cotyledonary stages (LC) (**Figure 12**). In late maturation, dry seeds and seedlings, the level of transcripts declined significantly. The size of the LYR/LVR transcripts approximately ranges between 1-1.1 kb. The expression pattern of the LYR/LVR gene suggests that tissue-specific signals play an important role in expression of LYR/LVR gene in Douglas-fir.

Effect of Brassinosteroid on LYR/LVR Gene Expression in Mature Seeds

In order to determine the effect of brassinosteroids on the expression of Douglas-fir LYR/LVR gene expression, total RNA was extracted from seeds treated with different concentrations of brassinosteroids at different time intervals. Northern hybridization showed that the LVR/LYR transcripts were present in higher amounts in Douglas-fir seeds treated with different concentrations of brassinosteroids, incubated for 1 week. The transcripts were present in lower amounts in the seeds treated for 1 day (**Figure 13**). Untreated, imbibed and stratified seeds were used as control in both experiments and the total RNA showed very low or undetectable LYR/LVR transcripts. Hence, brassinosteroids differentially regulated the expression of LVR/LYR in mature embryos or seeds.

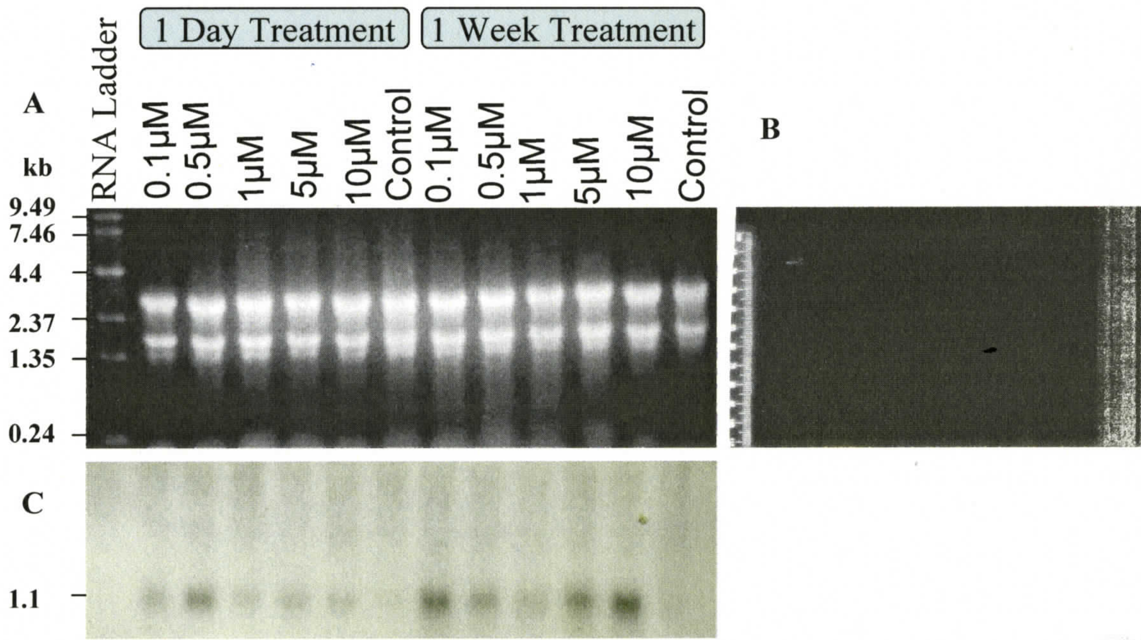


Figure 13 (A-B-C). Northern blot analysis of Douglas-fir seeds treated with different concentration of brassinosteroids at different time intervals. **A)** Ethidium bromide stained gel. Total RNA (25μg/lane) from different samples was separated in a formaldehyde agarose gel, transferred to a nitrocellulose membrane and probed with ^{32}-P labeled Douglas-fir LYR/LVR probe amplified using primers based on the Douglas-fir LYR/LVR cDNA. Total RNA from untreated seeds in Lanes 6 and 13 was used as control. **B)** Agarose gel after transfer and stained with ethidium bromide to show equal transfer. **C)** Autoradiogram of the northern blot. Sizes are based on the migration of 0.24-9.5 Kb RNA marker (Invitrogen) and indicated in kilobases. The size of the LYR/LVR transcript is approximately 1-1.1 kb.

Chapter 4
DISCUSSION

Genes involved in Douglas-fir Embryogenesis

Embryogenesis is an important process and it proceeds due to a controlled program of gene expression. Proper understanding of embryogenesis requires the identification of genes expressed during embryo development, determination of the function of the encoded proteins and the timing and location of gene expression (Thomas, 1993). Recent advancements in genomics allow regulatory genes involved in specific functions to be identified. This research focused on cloning of a LYR/LVR cDNA, and its characterization and expression in developing Douglas-fir seeds. Differential regulation in response to brassinosteroids, bioinformatic and phylogenetic analysis provided further information on the function of the LYR protein and its evolutionary relationships among different species.

In conifers, limited data is available on the identification, characterization and expression of embryogenesis-specific genes. In Douglas-fir, the available data included 11S legumin cDNA (Leal and Misra, 1993b), metallothionein-related cDNA (Chatthai *et al.*, 1997), and 2S seed storage protein cDNAs (Chatthai *et al.*, 1998, 2000) that were expressed in mid-to-late stages of somatic and zygotic embryos. Ciavatta and co-wokers (2002) identified PtNIP1:1 in loblolly pine that was expressed during early-embryogenesis. In

Picea glauca, differential screening of a cDNA library made from cotyledonary (stage 3) somatic embryos resulted in 28 different cDNAs whose temporal expression pattern during ABA-stimulated somatic embryo development were observed (Dong and Dunstan, 1996). These cDNA clones were grouped as embryo-abundant, embryo-maturation-induced, embryo development-specific and germination or non-embryogenic tissue-specific. Only a few genes expressed during early embryogenesis have been characterized in conifers up until now. Hence, this study focused on identification and characterization of a novel gene that is highly expressed during early embryogenesis.

Identification and Cloning of Douglas-fir LYR/LVR cDNA

The first strand LYR/LVR cDNA was synthesized by RT-PCR of Douglas-fir total RNA isolated using the Trizol method (see Materials and Methods). Following first strand synthesis, touchdown PCR using specific primers was done to obtain the LYR/LVR cDNA. The LYR/LVR cDNA was cloned using TOPO vector and sequenced. Sequence analysis resulted in a 679 bp LYR/LVR cDNA that was used for further analysis.

Bioinformatics Findings

The Douglas-fir LYR/LVR cDNA sequence was analysed by BLASTN and BLASTP. BLASTP (protein-protein blast) using NCBI database showed 59 % homology with *Arabidopsis thaliana* LYR expressed protein (ref:NP566416.1). The NCBI blast results identified cDNA to be one of the subunits of the LYR/LVR complex. Base-By-Base program was used to perform multiple sequence alignment of the LYR/LVR protein sequence with amino acid sequences from other closely related species (**Figure 4 and 5**).

This JAVA based program (developed in Dr. Chris Upton's lab) provides tools for protein sequence alignment, percentage identity table, phylogenetic tree analysis and visualization tools to identify and correct alignment errors. Multiple sequence alignment was performed using CLUSTALW. Douglas-fir LYR/LVR protein showed approximately 59 % amino acid identity to *Arabidopsis thaliana* LYR protein and 60 % identity to *Oryza sativa* LYR protein (**Table 3**). In order to elucidate the functional characteristics of the LYR protein, the Douglas-fir LYR protein sequence was further investigated using InterPro scanning system (Quevillon *et al.*, 2005). The results showed that Douglas-fir LYR protein encodes a NADP/NADH oxidoreductase enzyme which plays a vital role in the energy production process (InterPro ID: IPR008011). Pfam database was used to determine the functional site or conserved domain of the identified Douglas-fir LYR/LVR protein. The hidden Markov models (HMMs) provided by Pfam (Finn *et al.*, 2006) based on multiple sequence alignments for Douglas-fir LYR/LVR protein with respect to other closely related organisms predicted the primary structure consensus of the LYR protein family.

Uniprot database showed Douglas-fir LYR/LVR protein sequence similarity with Q9LH10 and Q8GS72 proteins of the LYR family in *Arabidopsis thaliana* and *Oryza sativa*, respectively (**Figures 8 and 9**). Q9LH10 and Q8GS72 belong to the LYR/LVR/Complex I family proteins and encode NADH oxidoreductase enzyme essential for energy production reactions (InterPro ID: IPR008011). In closely related *Arabidopsis thaliana* and *Oryza sativa* LYR/LVR and Douglas-fir LYR/LVR protein, the amino acid sequences between 19-90 were conserved (**Figure 5**) and may be the

functional domain (LYR motif) necessary for the oxidoreductase activity during oxidative phosphorylation in the electron transport chain (InterPro ID: IPR008011).

LYR Complex

The highly branched electron transport chain in plant mitochondria provides great flexibility for oxidation of cytosolic and mitochondrial matrix NADP. NADH:ubiquinone oxidoreductase of respiratory complex I acts as the first energy-coupling site of the respiratory chain (Grohman *et al.*, 1996). It is a large multisubunit, nuclear-encoded enzyme in the inner membrane of the mitochondria catalyzing electron transfer from NADH to ubiquinone through different redox centers such as FMN and FE-S clusters (Grohman *et al.*, 1998). The plant, mammalian and fungal counterparts contain 30-40 subunits constituting complex I whereas, the bacterial complex I contains only 14 subunits. In most species, complex I contains approximately 25-30 subunits (Heazlewood *et al.*, 2004). Based on Edman degradation and antibody cross-reaction, a small number of these subunit proteins were found to be similar to that of fungal and mammalian subunits. The sequencing of *Arabidopsis thaliana* and *Oryza sativa* nuclear genomes provided a better understanding of the nuclear-encoded components of complex I in plants (Heazlewood *et al.*, 2004, Millar *et al.*, 2001). However, in most cases, the inadequate data on the N- terminal region of the protein did not provide enough details on the different subunits investigated as only a few genes corresponding to these proteins have been sequenced from diverse plant species.

Recent studies have shown that this complex contains repeated occurrence of the tricopeptide leu-tyr-arg repeats at the N-terminal region of the LYR protein and hence the name LYR complex (Marchler-Bauer *et al.*, 2005, NCBI- Pfam 05347.5). The multiple sequence alignment of Douglas-fir LYR protein with other related LYR proteins in different organisms showed the presence of L, Y and R amino acids in the at positions 37-39 (**Figure 6**). The LYR proteins in *Arabidopsis thaliana* and *Oryza sativa* that showed similarity with the Douglas-fir LYR protein contained F, F, R amino acids in the positions 37-39 (**Figures 5 and 6**). Some of the other protein subunits identified in *Arabidopsis thaliana* that constitute complex 1 showed the presence of the tricopeptide (L, Y, R) at the same positions. Although, a few proteins contained F, F and R amino acids, they are all grouped under the same LYR family (Pfam ID: is PF05347). Complex 1 protein (LYR) family is composed of short proteins including those from the NADH-ubiquinone oxidoreductase complex 1. It is predicted that all the members of this family are components of complex 1. The family also includes the B14 subunit from bovine NADH-ubiquinone oxidoreductase and the B22 subunit from the human enzyme (Gu *et al.*, 1996). The size of *Arabidopsis thaliana* NADH-ubiquinone oxidoreductase genomic sequence is 1454 bp with a CDS length of 402 nucleotides (NCBI and TIGR locus tag-AT3G12260, Gene ID- 820406) (Haft *et al.*, 2003), encoding LYR protein containing 133 amino acids. Multiple sequence alignment and InterPro results showed the protein alignment of the Douglas-fir LYR with *Arabidopsis thaliana* and *Oryza sativa* LYR proteins and predicted functional similarity between them. Members of complex 1/LYR family are also found in yeast and in these organisms they are predicted to be required for iron-sulfur cluster biogenesis (Adam *et al.*, 2006, Wiedemann *et al.*, 2006).

From the literature cited above and available limited data, it is known that the LYR or LVR complex family consists of a short stretch of closely related proteins involved in various energy production processes. Bioinformatic analysis of the Douglas-fir LYR complex suggest that it may contain several closely related proteins apart from the protein identified in this study. This was also substantiated by the results of Southern blot analysis of Douglas-fir genomic DNA (**Figure 11**) that showed 2 strong bands in each lane, in addition to several weaker bands.

The phylogenetic tree (**Figure 10**) showed evolutionary relationship of Douglas-fir LYR/LVR protein (encoding NADH oxidoreductase) with closely related members of the same family in other groups of distant living organisms such as human, rat, chicken (data not shown), insect, and fungi such as *Neurospora*, *Debaryomyces* and *Ustilago maydis*. The phylogenetic tree showed that the Douglas-fir LYR/LVR protein was clustered with other plants LYR/LVR (NADH oxidoreductase enzyme). The Douglas-fir LYR/LVR shared a common ancestor with LYR/LVR *Arabidopsis thaliana* and *Oryza sativa.* The LYR/LVR proteins in human, rat and chicken (gi/50729240) were grouped together. The LYR proteins in the different fungi studied also shared a common ancestor. An unnamed LYR/LVR protein in *Oryza sativa* did not group with LYR/LVR gene (encoding NADH oxidoreductase) of Douglas-fir, *Arabidopsis thaliana* and *Oryza sativa*, but clustered with a hypothetical protein (NADH:ubiquinone oxidoreductase) in *Caenorhabditis elegans.* Further analysis showed that this unnamed *Oryza sativa* LYR/LVR protein did not contain the conserved LYR/LVR domain and likely was a distant member with less than 20% amino acid identity to other more closely related LYR proteins. Hence, it appears

that the LYR family genes have evolved independently and are evolutionarily conserved in different groups of organisms.

Differential expression of Douglas-fir LYR/LVR during embryogenesis

In developing seeds, the expression of Douglas-fir LYR/LVR gene was higher during early- and mid-embryogenesis. The LYR/LVR transcripts were first detected in megametophytes, prior to fertilization. In isolated zygotic embryos, high levels of transcripts accumulated during early stages, declining to low levels during late-cotyledonary stages (**Figure 11**). In mature seeds and seedlings, the LYR/LVR transcripts were barely detected.

The Douglas-fir LYR/LVR gene expression pattern was compared with the gene expression profiles of *Arabidopsis thaliana* in the NCBI Gene Expression Omnibus database (GEO), as rather limited published data is available on the LYR/LVR gene. The GEO profiles showed that *Arabidopsis* LYR/LVR gene is expressed at various developmental stages from germination to mature seeds, flower buds and roots (UniGene At.48649 GEO profile). Further work needs to be done to elucidate the different factors affecting the regulation of LYR/LVR gene expression in Douglas-fir.

Effect of brassinosteroids on Douglas-fir LYR/LVR gene expression

Recent discovery of a new class of phytohormones called brassinosteroids that act as endogenous factors affecting certain genes involved in various developmental pathways

is intriguing. In this study, the effect of BRs was examined on expression of LYR/LVR. Mature seeds treated with different concentrations of brassinosteroids at different time intervals were used as the source of RNA to determine the regulation of the LYR/LVR gene. Northern blot analysis showed differential accumulation of Douglas-fir LYR/LVR transcripts and their levels were affected by exogenous brassinosteroid treatment (**Figure 13**). LYR/LVR gene expression was induced in seeds treated with different concentrations of brassinosteroids. Prolonged exposure to brassinosteroids with concentrations from 0.1 μM to 10 μM induced the LYR/LVR gene expression in later stages of Douglas-fir embryogenesis where LYR/LVR transcripts were present at low levels.

In angiosperms, certain endogenous factors and phytohormones such as auxin, abscisic acid (ABA) are believed to play essential roles (Goldberg *et al.*, 1989). Studies have shown that certain sugars, ABA (Skriver *et al.*, 1990) and PEG are necessary for embryo development and maturation, where as auxins and cytokinins play vital roles in regeneration (Minocha *et al.*, 2000). Exogenous application of brassinosteroid at low concentrations enhanced embryonic tissue initiation in conifers such as loblolly pine and Douglas-fir (Pullman *et al.*, 2003). In conifers, limited data is available on BR-regulated gene expression during embryogenesis. Results of this study show clearly that brassinosteroids modulate the expression of the LYR/LVR gene in Douglas-fir. Since LYR/LVR gene encodes a NADH oxidoreductase enzyme (complex 1) that acts as the first energy-coupling site in the respiratory chain, its necessary to examine its role in later stages of embryogenesis where its expression is decreased. Hence, brassinosteroids act as

important inducer for LYR/LVR gene expression in mature embryos and seeds. This gene could provide a useful marker to further examine the role of BRs in conifer embryogenesis.

Chapter 5
CONCLUSION AND FUTURE STUDIES

Conclusions

The identification and characterization of genes involved in conifer embryogenesis provides necessary information to understand the biochemical and molecular events during embryo development. The isolation of molecular markers such as 11S, 2S and metallothionein-like cDNAs (Chatthai and Misra, 1997; 1998) that are expressed during mid and late- cotyledonary embryos initiated various studies to elucidate the role of these embryogenesis-specific genes in the growth and development of Douglas-fir embryos. This study provides further insight into the identification of essential molecular markers of early embryogenesis in Douglas-fir.

In this study, an early embryogenesis LYR/LVR gene was identified, cloned and characterized in Douglas-fir. The LYR/LVR gene showed higher expression levels in the early-stages of zygotic embryos and megagametophytes when compared to mid- and late-cotyledonary embryos. The transcript levels declined further in mature embryos (seeds) and seedlings. The LYR/LVR gene expression was induced in mature seeds treated with different concentrations of brassinosteroids. This gene provides a marker of early embryogenesis which will aid in the study of the effect of brassinosteroids in Douglas-fir embryogenesis. Bioinformatic analysis of Douglas-fir LYR/LVR provided information

regarding the function of LYR/LVR protein, its conserved and functional domains, and its relation with other closely related members of the LYR (complex 1) family. The Douglas-fir LYR/LVR protein is likely to be an inner mitochondrial membrane protein with NADH dependent catalytic oxidoreductase activity, necessary for reactions involved in energy production.

Future Studies

To gain further insight into transcriptional regulation and function of LYR/LVR gene in Douglas-fir, future work needs to be done as outlined below:

A) Prepare polyclonal antibodies and analyse protein levels in various tissues and developmental stages.

B) Functional analysis by overexpression of LYR/LVR full length cDNA in *Arabidopsis thaliana* and Douglas-fir.

C) Examine the expression of LYR/LVR in different Douglas-fir genotypes and other conifer species.

BIBLIOGRAPHY

Aagaard JE (1997) Genetic diversity and differentiation in Douglas-fir from RAPD markers of nuclear and mitochondrial origin. MS Thesis, Oregon State University, Corvallis, OR, USA.

Abo El-Nil MM (1987) Tissue culture of Douglas-fir and Western North American conifers. Cell and Tissue Cultue in Forestry 3, 80-100.

Adam AC, Bornhovd C, Prokisch H, Neupert W and Hell K (2006) The Nfs1 interacting protein Isd11 has an essential role in Fe/S cluster biogenesis in mitochondria. Journal of European Molecular Biology Oraganization 25, 174-83.

Aitken-Christie J (1991) "Micropropagation: Technology and application." In: Debergh PC and Zimmerman RH (eds) Kluwer Academic Publishers, pp. 363–388.

Aitken-Christie J and Connett M (1992) Micropropagation of forest trees. In Transplant Production Systems (eds) Kurata, K. & Kozai, T, pp. 163-194.

Allen G and Owens JN (1972) The Life history of Douglas fir. Environment Canada, Ottawa 139.

Allen GS (1941) A standard germination test for Douglas-fir seed. Forestry Chronicle 17, 75-78.

Altschul SF, Stephen F, Thomas L Madden, Alejandro A Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller and David J Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25, 3389-3402.

Bishop GJ and Kroncz C (2002) Brassinosteroids and plant steroid hormone signaling. The Plant Cell [Suppl], 97–110.

Black M (1991) Involvement of ABA in the physiology of developing and mature seeds. In: Davies WJ and Jones HG (eds.), Abscisic Acid Physiology and Biochemistry, Environmental Plant Biology series, Bios Scientific Publishers, pp. 99-124.

Bommineni VR, Bethune TD, Hull KL and Dunstan DI (1998) Expression of *uidA* under regulation of the ABA-inducible wheat Em promoter in transgenic white spruce somatic embryos. Plant Physiology 152, 455-462.

Bornman CH (2002) Somatic seed in conifer biotechnology — a viable alternative to natural seed? South African Journal of Botany 6, 119-126.

Boulay MP, Gupta PK, Krogstrup P and Durzan DJ (1988) Development of somatic embryos from cell suspension culture of Norway spruce (*Picea abies* Karst.) Plant Cell Reports 7, 134-137.

Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu CM, van Lammeren AA, Miki BLA, Custers JBM and van Lookeren Campagne MM (2002) Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic Growth. The Plant Cell 14, 1737–1749.

Brocard-Gifford IM, Lynch TJ and Finkelstein RR (2003) Regulatory Networks in seeds integrating developmental, abscisic acid, sugar and light signaling. Plant Physiology 131, 78-92.

Brodie R, Smith AJ, Roper RL, Tcherepanov V and Upton C (2004) Base-By-Base: Single nucleotide-level analysis of whole viral genome alignments. BMC Bioinformatics 5, 96.

Brosa D (1999) Biological effects of brassinosteroids. Critical Reviews in Biochemistry and Molecular Biology 34, 339–358.

Cairney J, Xu N, MacKay J, Pullman J (2000) Transcript profiling: a tool to assess the development conifer embryos. *In Vitro* Cellular and Developmental Plant Biology 36, 155–162.

Cairney J, Xu N, Pullman GS, Ciavatta VT, Johns B (1999) Natural and somatic embryo development in loblolly pine. Biotechnology and Applied Biochemistry 79, 5–17.

Chatthai M and Misra S (1998) Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziesii* [Mirb.] Franco. Planta 206, 138-145.

Chatthai M and Misra S (2000) Gene expression during Pseudotsuga menziesii embryogenesis: Low molecular weight proteins. In: Jain, S.M. and Minocha, S.C. (eds.) Molecular Biology of Woody Plants. Kluwer Academic Publishers, Netherlands 1, 299-313.

Chatthai M, Kaukinen KH, Tranbarger TJ, Gupta PK and Misra S (1997) The isolation of a novel metallothionein-related cDNA expressed in somatic and zygotic embryos of Douglas-fir: regulation by ABA, osmoticum and metal ions. Plant Molecular Biology 34, 243-254.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG and Thompson

JD (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research 31, 3497–3500.

Ciavatta VT, Morillon R, Pullman GS, Chrispeels MJ and Cairney J (2001) An Aquaglyceroporin is abundantly expressed early in the development of the suspensor and the embryo proper of Loblolly Pine 1. Plant Physiology 127, 1556–1567.

Ciavatta VT, Egertsdotter U, Clapham D, Von Arnold S and Cairney J (2002) A promoter from the loblolly pine PtNIP1:1 gene directs expression in an early-embryogenesis and suspensor-specific fashion. Planta 215, 694-698.

Clouse SD (2001) Brassinosteroids. In: Somerville CR, Meyerowitz EM (eds) The Arabidopsis book. American Society of Plant Biologists, Rockville, MD. http://www.aspb.org/publications/ arabidopsis/

Clouse SD and Sasse JM (1998) Brassinosteroids: essential regulators of plant growth and development. Annual Review of Plant Physiology and Plant Molecular Biology 49, 427–451.

de Jong A, Schmidt E and De Vries S (1993) Early events in higher-plant embryogenesis. Plant Molecular Biology 22, 367-377.

Dodeman VL, Ducreux G and Kreis M (1997) Zygotic embryogenesis *versus* somatic embryogenesis. Journal of Experimental Botany 48, 1493-1509.

Dong JZ and Dunstan DI (2000) Molecular biology of somatic embryogenesis in conifers. Forestry Sciences 64, 51-88.

Dong JZ and Dunstan DI (1996) Expression of abundant mRNAs during somatic embryogenesis of white spruce [*Picea glauca* (Moench) Voss]. Planta 199, 459 466.

Dunstan DI, Tautorus TE and Thorpe TA (1995) Somatic embryogenesis in woody plants. In: Thorpe TA (ed) In vitro embryogenesis in plants. Kluwer Academic Publishers Dordrecht, Netherlands, pp. 473-539.

Durzan DJ (1988) Somatic polyembryogenesis for the multiplication of tree crops. Biotechnology and Genetic Engineering Reviews 6, 339-376.

Durzan DJ and Gupta PK (1987) Somatic embryogenesis and polyembryogenesis in Douglas-fir cell suspension cultures. Plant Science 52, 229-235.

Durzan DJ and Gupta PK (1988) Somatic embryogenesis and polyembryogenesis in conifers. Advanced Biotechnology Processes 9, 53-81.

El-Kassaby YA, Edwards DGW and Taylor DW (1992) Genetic control of germination parameters in Douglas-fir and its importance for domestication. Silvae Genetica 41: 48-54.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL and Bateman A (2006) Pfam: clans, web tools and services. Nucleic Acids Research 34, Database issue D247–D251.

Forward BS and Misra S (2001) Characterization and expression of Douglas fir luminal binding protein (PmBip). Planta 212, 41-51.

Goldberg RB (1998) Plants: novel development processes. Science 240, 1460-1467.

Goldberg RB, Barker S and Perez-Grau L (1989) Regulation of gene expression during plant embryogenesis. Cell 56, 149-160.

Grohmann L, Rasmusson AG, Heiser V, Thieck O and Brennicke A (1996) The NADH-binding subunit of respiratory chain complex I is nuclear-encoded in plants and identified only in mitochondria. The Plant Journal 10, 793-803.

Gu JZ, Lin X and Wells DE (1996) The human B22 subunit of the NADH-Ubiquinone oxidoreductase maps to the region of chromosome 8 involved in Branchio-Oto-Renal Syndrome. Genomics 35, 6–10.

Gupta PK and Durzan DJ (1986) Plantlet regeneration via somatic embryogenesis from subcultured callus of mature embryos of *Picea abies* (Norway spruce). *In Vitro* Cellular and Development Biology 22, 685-688.

Gupta PK and Grob JA (1995) Somatic embryogenesis in conifers. In: Jain S, Gupta Pkand Newton RJ (eds.) Somatic embryogenesis in Woody plants. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 81-98.

Gupta PK and Pullman GS (1991) Method for reproducing coniferous plants by somatic embryogenesis using abscisic acid and osmotic potential variation. U.S. Patent No. 5,036,007.

Haft DH, Selengut JD and White O (2003) The TIGRFAMs database of protein families. Nucleic Acids Research 31, 371-373.

Hackman I (1993) Embryology in Norway spruce (*Picea abies*). An analysis of the composition of seed storage proteins and deposition of storage reserves during seed develeopment and somatic embryogenesis. Physiologia Plantarum 87, 148–159.

Hackman I, Stabel P, Engström P and Eriksson T (1990) Storage protein accumulation during zygotic and somatic embryo development in *Picea abies* (Norway spruce). Physiolgia Plantarum 80, 441–445.

Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DD, Whelan J and Millar HA (2004) Experimental Analysis of the *Arabidopsis* mitochondrial proteome highlights

signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. The Plant Cell 16, 241–256.

Ikeda-Iwai M, Umehara M, Satoh S and Kamada H (2003) Stress-induced somatic embryogenesis in vegetative tissues of *Arabidopsis thaliana*. The Plant Journal 34, 107-114.

InterPro scan results InterPro protein ID: IPR008011

Jokinen K and Durzan DJ (1994) Properties of rescued embryonal suspensor masses of Norway spruce (*Picea abies* L. Karst.) determined by the genotype and the environment *in vitro*. Silva Fennica 28, 95-106.

Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A and Hattori T (2005) LEAFY COTYLEDON1 Controls Seed Storage Protein Genes through Its Regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. Plant and Cell Physiology 46, 399-406.

Kermode AR (1990) Regulatory mechanisms involved in the transition from seed development to germination. Critical Reviews in Plant Sciences 9, 155-195.

Krasowski MJ and Owens JN (1993) Ultrastructural and histochemical postfertilization megagametophyte and zygotic embryo development of white spruce (*Picea glauca*) emphasizing the deposition of seed storage products. Canadian Journal of Botany 71, 98–112.

Leal I and Misra S (1993b) Molecular cloning and characterization of a legumin-like storage protein cDNA of Douglas-fir seeds. Plant Molecular Biology 21, 709-715.

Lelu MA and Bornman CH (1990) Induction of somatic embryogenesis in excised cotyledons of *Picea glauca* and *Picea mariana*. Plant Physiology and Biochemistry 28, 785-791.

Lippert D, Zhuang J, Ralph S, Ellis DE, Gilbert M, Olafson R, Ritland K, Ellis B, Douglas C and Bohlmann J (2005) Proteome analysis of early somatic embryogenesis in *Picea glauca*. Molecular and Cellular Proteomics. 5, 461-473.

Lotan T, Ohto M, Yee KM, West MA, Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB and Harada JJ (1998) Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. Cell 93, 1195–1205.

Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D and Bryant SH (2005) CDD: a Conserved Domain Database for protein classification. Nucleic Acids Research 33, D192-D196.

Millar AH, Sweetlove LJ, Giege P and Leaver CJ (2001) Analysis of the *Arabidopsis* mitochondrial proteome. Plant Physiology 127, 1711–1727.

Minocha R and Jain SM (2000). Tissue culture of woody plants and its relevance to molecular biology. Forestry Sciences 64, 315-340.

Misra S (1994) Conifer zygotic embryogenesis, somatic embryogenesis, and seed germination: Biochemical and molecular advances. Seed Science Research 4, 357–384.

Misra S (1995) Molecular analysis of zygotic and somatic conifer embryos. In: Jain S, Gupta P and Newton R (eds.), Somatic Embryogenesis in Woody Plants. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 119-142.

Misra S and Green M (1991) Developmental gene expression in conifer embryogenesis and germination. II. Crystalloid protein synthesis in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss.). Plant Science 78, 61-71.

Mordhorst AP, Toonen MAJ and de Vries SC (1997) Plant embryogenesis. Critical Reviews in Plant Sciences 16: 535–576.

Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bradley P, Bork P, Bucher P, Cerutti L, Copley R, Courcelle E, Das U, Durbin R, Fleischmann W, Gough J, Haft D, Harte N, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lonsdale D, Lopez R, Letunic I, Madera M, Maslen J, McDowall J, Mitchell A, Nikolskaya AN, Orchard S, Pagni M, Ponting CP, Quevillon E, Selengut J, Sigrist CJ, Silventoinen V, Studholme DJ, Vaughan R and Wu CH (2005) InterPro, progress and status in 2005. Nucleic Acids Research 33, Database issue D201–D205.

NCBI and TIGR locus tag-AT3G12260, Gene ID- 820406, unpublished data.

NCBI- Pfam 05347.5, unpublished data.

Neff MM, Nguyen SM, Malancharuvil EJ, Fujioka S, Noguchi T, Seto H, Honda T, Yoshida S and Chory J (1999) BAS1: A gene regulating brassinosteroid levels and light responsiveness in *Arabidopsis*. Proceedings of the National Academy of Sciences 96, 15316-15323.

Nomura K and Komamine (1986) Molecular mechanisms of somatic embryogenesis. Oxford Surveys of Plant Molecular and Cell Biology 3, 456-466.

Notredame C, Higgins DG and Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of Molecular Biology 302, 205-217.

Owens J, Colangeli A and Morris S (1991) Factors affecting seed set in Douglas fir. Canadian Journal of Botany 69, 229-238.

Owens JN, Morris SJ and Misra S (1993) The ultrastructural, histochemical, and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii*. Canadian Journal of Forest Research 23, 816–827.

Pullman GS and Peter GF (2002a) Brassinolide improves embryogenic tissue initiation in conifers (abstract). In: 10th IAPTC&B Cong. Orlando, Fla., P-1138, P 67A.

Pullman GS and Peter G (2002b) Methods of initiating embryogenic cultures in plants. U.S. Patent No. 6,492,174B1. Issued December 10, 2002.

Pullman GS, Zhang Y and Phan BH (2003) Brassinolide improves embryogenic tissue initiation in conifers and rice. The Plant Cell 22, 96-104.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R and Lopez R (2005) InterProScan: protein domains identifier. Nucleic Acids Research 33, Web Server issue W116–W120.

Reiser L and Fischer RL (1993) The ovule and the embryo sac. The Plant Cell 5, 1291-1301.

Russinova E, Borst JW, Kwaaitaal M, Cano-Delgado A, Yin Y, Chory J and de Vries SC (2004) Heterodimerization and endocytosis of *Arabidopsis* brassinosteroid receptors BRI1 and AtSERK3 (BAK1). The Plant Cell 16, 3216–3229.

Spurr AR (1949) Histogenesis and organization of the embryo in *Pinus strobus* L. American Journal of Botany 36, 629-641.

Sambrook J, Fritsh EF and Maniatis T (1989) Molecular Cloning: A Laboratory Manual 2nd Edition. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.

Schäffer AA, Alejandro A, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV and Altschul SF (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic Acids Research 29, 2994-3005.

Schneider WL and Gifford DJ (1994) Loblolly pine seed dormancy. I. The relationship between protein synthesis and the loss of dormancy. Physiologia Plantarum 90, 246-252.

Schulz P and Jensen WA (1969) Capsella embryogenesis: the suspensor and the basal cell. Protoplasma 67, 139-163.

Schwartz BW, Vernon DM and Meinke DW (1997) Development of the suspensor: differentiation, communication, and programmed cell death during plant embryogenesis. In BA Larkins, IK Vasil, (eds) Cellular and Molecular Biology of Plant Seed Development. Kluwer Academic Publishers, The Netherlands, pp. 53-72.

Silen RR (1978) Genetics of Douglas-fir. USDA Forestry service research paper WO-35, 1-27.

Skriver K and Mundy J (1990) Gene expression in response to abscisic acid and osmotic stress. The Plant Cell 2, 503-512.

Taniguchi T, Kurita M, Itahana N and Kondo T (2004) Somatic embryogenesis and plant regeneration from immature zygotic embryos of Hinoki cypress (*Chamaecyparis obtusa* Sieb. et Zucc.). Plant Cell Reports 23, 26-31.

Tautorous TE, Fowke LC and Dunstan DI (1991) Somatic embryogenesis in conifers. Canadian Journal of Botany, 69, 1873-1899.

Thomas TL (1993) Gene expression during plant embryogenesis and germination: An overview. The Plant Cell 5, 1401-1410.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25, 4876-4882.

Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680.

Thorpe TA (1982) Physiological and biochemical aspects of organogenesis in vitro. In: Fujiwara A (eds.) Plant Tissue and Cell Culture Proceedings of the 5th International Congress of Plant Tissue. Japanese Association for Plant Tissue Culture, pp. 121-124.

Timmis R, Abo El-Nil MM and Stonecypher RW (1987) Potential genetic gain through tissue culture. Cell and Tissue Culture in Forestry 1, 198-215.

Tranbarger TJ and Misra S (1996) Structure and expression of a developmentally regulated cDNA encoding a cysteine protease (pseudotzain) from Douglas-fir. GENE 172, 221-226.

Tranbarger TJ, Forward BS and Misra S (2000) Regulation of NADPH-cytochrome P450 reductase expressed during Douglas fir germination and seedling development. Plant Molecular Biology 44, 141-153.

UniGene At.48649 GEO profile, *Arabidopsis thaliana* LYR protein, unpublished data.

UniGene At.48649 ref: NP_566416.1, LYR protein details, unpublished data.

Verwoerd TC, Dekker BM and Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Research 17, 2362.

Vetrici M and Misra S (2005) The Douglas-fir Leafy Cotyledon1 (PMLEC1) gene is a

transcription factor expressed during early somatic and zygotic embryogenesis and is inducible by stress and hormones. Abstract, American Society of Plant Biologists meeting, July 2005.

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP and Allard RW (1987) Chloroplast DNA polymorphisms in Lodgepole and Jack pines and their hybrids. Proceedings of National Academy of Sciences, USA 84, 2097-2100.

Walter C, Carson SD, Menzies MI, Richardson T and Carson M (1998) Review: Application of biotechnology to forestry -molecular biology of conifers. World Journal of Microbiology and Biotechnology 14, 321-330.

West MAL and Harada JJ (1993) Embryogenesis in higher plants: an overview. The Plant Cell 5, 1361-1369.

Wiedemann N, Urzica E, Guiard B, Muller H, Lohaus C, Meyer HE, Ryan MT, Meisinger C, Muhlenhoff U, Lill R and Pfanner N (2006) Essential role of Isd11 in mitochondrial iron-sulfur cluster synthesis on Isu scaffold proteins 25, 184-95.

Wilcox PH, Amerson HV, Kuhlman G, Liu GH, O'Malley DM and Sederoff RR (1996) Detection of a major gene from resistance to fusiform rust disease in loblolly pine by genome mapping. Proceedings of the National Academy of Sciences, USA 93, 3859-3864.

Wu CH, Apweiler R, Bairoch A, Natale DA, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Mazumder R, O'Donovan C, Redaschi N and Suzek B (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information. Nucleic Acids Research 34, Database issue D187–D191.

Yeung EC and Meinke DW (1993) Embryogenesis in angiosperms: development of the suspensor. The Plant Cell 5, 1371–1381.

Zimmerman J (1993) Somatic embryogenesis: A model for early development in higher plants. The Plant Cell 5, 1411-1423.

APPENDIX 1

NCBI BLASTP (protein-protein blast) showing the alignment of translated amino acid sequences of Douglas-fir LYR/LVR protein with other related LYR/LVR proteins (Complex 1 or LYR family) in different organisms (Altschul *et al*., 1997, Schaffer *et al*., 2001).

A. Ref|NP_001060420.1|Complex 1 LYR protein [*Oryza sativa* (japonica cultivar-group)]

Score = 156 bits (394), Expect = 3e-37, Identities = 77/128 (60%), Positives = 94/128 (73%), Gaps = 2/128 (1%), Query= Douglas-fir LYR/LVR protein, Length=131

```
Query 1 MASPLRQVGTGANSQSLDEARKRVFHFFREACRSIPQIMETYNLHEVITPSQLRSVVAAQ 60
        MA  +R V    NS SL+EAR RVF FFR+ACR+IP IME YNL +V+TPSQLRS +A+
Sbjct 1 MAFTMRAVKVPPNSASLEEARHRVFDFFRQACRAIPSIMEIYNLDDVVTPSQLRSTIAKE 60

Query61 FRKQAHVTNPKVIDMLIIKGDEELRNCLDHSKQRHHIVGQYVIGQEGLIPSNVGAVSSGG 120
         RK   VTNPKVIDML+ KG EEL  N +H+KQRHH++GQYV+GQ+GL+          G
Sbjct61 IRKNQGVTNPKVIDMLLFKGMEELGNITEHAKQRHHVIGQYVVGQKGLVQDM--EKDQGS 118

Query121SDFLRKFY 128
        SDFL+KFY
Sbjct119SDFLKKFY 126
```

B. Ref|NP_566416.1| catalytic/ oxidoreductase, acting on NADH or NADPH [*Arabidopsis*
thaliana] Length=133

Score = 154 bits (389), Expect = 1e-36, Identities = 77/130 (59%), Positives = 98/130 (75%), Gaps = 4/130 (3%)

```
Query  1   MASP--LRQVGTGANSQSLDEARKRVFHFFREACRSIPQIMETYNLHEVITPSQLRSVVA 58
           MA+P  LR++G   NS +L EAR+RVF FFR ACRSIP IM+ YNL +V+ PSQLR  ++
Sbjct  1   MAAPFALRKIGVPPNSANLTEARRRVFDFFRAACRSIPTIMDIYNLQDVVAPSQLRYAIS 60

Query  59  AQFRKQAHVTNPKVIDMLIIKGDEELRNCLDHSKQRHHIVGQYVIGQEGLIPSNVGAVSS 118
           AQ R  AH+T+PKVID+LI KG EEL + +DH+KQRHHI+GQYV+G EGL+  N G
Sbjct  61  AQIRNNAHITDPKVIDLLIFKGMEELTDIVDHAKQRHHIIGQYVVG-EGLV-QNTGNKDQ 118
```

C. ref|XP_425471.1| PREDICTED: similar to NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6 (B14); NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6 (14kD, B14) [*Gallus gallus*] Length=151

Score = 67.0 bits (162), Expect = 2e-10, Identities = 45/116 (38%), Positives = 61/116 (52%), Gaps = 13/116 (11%)

```
Query  14  SQSLDEARKRVFHFFREACRSIPQIMETYNLHEVITPSQLRSVVAAQFRKQAHVTNPKVI 73
           S+ L EA+ RV   +R   R +P  +  Y L   IT  Q R  V   F K AHVT+P+VI
Sbjct  45  SRDLAEAKLRVRELYRAWYREVPNAVHLYQLD--ITAKQGRDKVREMFLKNAHVTDPRVI 102

Query  74  DMLIIKGDEELRNCLDHSKQRHHIVGQYVIGQEGLIPSNVGAVSSGGSDFLRKFYD     129
           DML+IKG  +L+  ++  KQR H++ +Y    E   P           DFL KFY+
Sbjct  103 DMLVIKGKMDLQETINVWKQRTHVM-RYFHETETPRP----------KDFLSKFYE     147
```

APPENDIX 2

A. Primers used for the synthesis of LYR/LVR probe for Southern and northern hybridizations. The primers were designed based on the newly isolated Douglas-fir LYR/LVR cDNA and were obtained from Invitrogen.

LYR/LVR Forward Primer:
5' TTGAGACAGGTGGGAACAGGA 3'

LYR/LVR Reverse Primer:
5' CTACAATGTGATGCCGCTGCT 3'

B. Primers used for the isolation of Douglas-fir LYR/LVR cDNA by RT-PCR amplification. Degenerate primers based on *Arabidopsis thaliana* COP1 were used for RT-PCR amplification that yielded various cDNA products, one of them was identified as the LYR/LVR cDNA.

Forward Primer:
5' TGYATGCARATIATHAARGAYGC 3'

Reverse Primer:
5' AICTACARTAIGCAGRYGCRTCTR 3'