CHARACTERISATION OF THE GENES ENCODING MANNURONAN-C5-EPIMERASE IN THE BROWN ALGA Lessonia variegata



Lessonia variegata, Moa Point, Island Bay, Wellington; April 2012. Photo by Bongkot Wichachucherd

SAYANI GHOSH

A thesis submitted to the Victoria University of Wellington in partial fulfillment of the requirements for the degree of Masters of Science

Victoria University of Wellington

2012

Abstract

Alginate is known to be a commercially valuable polysaccharide, of great importance in industries such as food, cosmetics, medicine and pharmaceuticals. It is obtained commercially by harvesting brown algae. The final step in the alginate biochemical pathway involves the epimerization of D-mannuronic residues into L-guluronic residues, catalyzed by the enzyme mannuronan-C5-epimerase. This final step has been found to be responsible for controlling the physicochemical properties of the produced alginate. This study is the first to characterize the genes encoding for the enzyme mannuronan-C5epimerase in the Northern, Southern and Wellington lineages of the brown alga Lessonia variegata (Phaeophyceae). The gene of interest was amplified by standard PCR and cloning. Cloning PCR results revealed the presence of two distinct copies of the gene in Lessonia variegata. The coding region of the copies was found to be very conserved with very little sequence variation. The Lessonia variegata sequences were compared with those of Laminaria digitata and Saccharina japonica, which indicated that at least one gene duplication event has occurred in *Lessonia variegata*, leading to the formation of two gene duplicates. The possible mechanisms by which the gene paralogs may control the structure and function of the produced alginate have been discussed.

Acknowledgments

I would like to thank the following people who have helped me through my MSc thesis: Joe Zuccarello, my supervisor, without whose guidance, unwavering support and endless patience, this project would not have been possible. Thank you for all the help with the molecular work, for the many discussions and for editing and proofreading my thesis, all of which have helped me to learn and develop my laboratory, writing and critical thinking skills.

Christian Boedeker for the help with molecular work, the sequence analysis and editing assistance.

Peter Martin for letting me use his DNA samples, without which this project would have been severely delayed.

A big thank you to the rest of the Zuccarello research group - Laura Sanchez, Narongrit Muangmai, Michael Gemmell, Maren Preuss and Bongkot Wichachucherd, and also to Nicky Fitzgibbon and Kerry Charles in the School of Biological Sciences for all the help, the discussions and the general banter. This project would not have been half as much fun without you guys.

In the outside world, a special thank you to my family and Thomas Ramsay for their constant support and encouragement. Also, thank you to Tony Ramsay for proofreading my thesis and for the helpful suggestions.

Table of Contents

Abstract	ii
Acknowledgments	
List of Figures	vi
List of Tables	vii

Chapter One

1.1 Introduction	1
1.2 Taxonomy	2
1.3 Characteristics of Phaeophyceae	2
1.4 Morphology and Life History of Laminariales	4
1.5 Distribution and life history of Lessonia variegata J. Agardh	6
1.6 Carbohydrates produced by brown algae and their uses	8
1.6.1 Properties of alginate	10
1.6.2 Biosynthesis of alginate	13
1.6.3 Mannuronan-C5-Epimerase genes involved in encoding the last enzyme in the alginate biosynthetic pathway	14
1.6.4 Alginate production in <i>Lessonia variegata</i>	16
1.7 Outline and scope of this study	18

Chapter two: Materials and methods

2.1 Sampling	20
2.2 DNA Extraction	20
2.3 PCR with Primer Set I	20
2.3.1 Primer Set I designed from Genbank data	20
2.3.2 DNA Amplification	21
2.3.3 DNA extraction and ligation	21
2.3.4 Making LB-agar-ampicillin plates	22

2.3.5 Preparation and transformation of chemically competent E. coli	23
2.3.6 Transformation of competent DH5 α E. coli cells	23
2.3.7 Miniprepping of plasmids	23
2.3.8 Restriction digestion of plasmid	24
2.4 PCR with Primer Set II	25
2.4.1 Primer Set II designed for the specific amplification of Lessonia variegata sequences	25
2.4.2 Ligation, PCR clean up and transformation	25
2.4.3 Colony PCR	27
2.5 Analysis of obtained DNA sequences	28
2.6 RNA extraction	30

Chapter three

3.1 RESULTS	35
3.1.1 Amplification with Primer Set I	35
3.1.2 Sequences obtained from PCR with Primer Set II	36
3.2 DISCUSSION	45
3.2.1 Gene duplication: multiple copies of Man-C5-E gene in Lessonia variegata	45
3.2.2 Allelic variations in a diploid organism	47
3.2.3 Introns and chimeric sequences in the dataset: possibilities	49
3.2.4 Mannuronan C5 Epimerase gene duplicates may control alginate structure	51
3.2.5 RNA extraction	53
3.2.6 Conclusion	54

4.0 References	55
5.0 Appendices	65
Appendix 1	65
Appendix 2	66
Appendix 3	67

List of Figures

Figure		Page
Fig. 1.1:	Alternation of generation in the life cycle of members of the Laminariales	4
Fig. 1.2:	The Structure of the Laminariales	5
Fig. 1.3:	Map of New Zealand showing the approximate distribution of the different lineages of <i>Lessonia variegata</i> .	7
Fig. 1.4	Alginate structure	12
Fig. 1.5	Biosynthesis pathway of alginate in Phaeophyceae	13
Fig. 1.6	Histogram of monthly yield of alginate within Wellington lineage of Lessonia variegata	17
Fig. 3.1	Unrooted neighbour joining topology of Man-C5-E exon sequences	39
Fig. 3.2	Haplotype network showing nucleotide substitutions between exon sequences	40
Fig. 3.3	Unrooted neighbour-joining tree of Man-C5-E intron sequences	42
Fig. 3.4	Bayesian, Maximum-likelihood and Neighbour-joining analysis of sequences	44
Fig. 3.5	Diagram illustrating a possible gene duplication in Lessonia variegata	47

List of Tables

Table	Page
Table 2.1: Location of samples used	32
Table 2.2: Brown algal sequences used	33
Table 2.3: Primers designed	34
Table 3.1: Observed frequency of nucleotide changes (bp) between three <i>Lessonia variegata</i> sequences from the Wellington lineage C568, obtained from primer set I.	36
Table 3.2: <i>Lessonia variegata</i> sequences (variable positions) obtained from primer set II	37
Table 3.3: Observed range of nucleotide changes within and between exon groups	41
Table 3.4: List of possible alleles in each exon group	41
Table 3.5: Observed range of nucleotide changes within and between Man-C5-E DNA sequences of <i>Laminaria digitata</i> groups	45

CHAPTER ONE

1.1 Introduction

Algae constitute a large and diverse group of unicellular or multicellular organisms, which are generally autotrophic. The most commonly known algal classes, which includes both micro- and macroalgae, are diatoms (Class Bacillariophyceae), green algae (Class Chlorophyceae), red algae (Class Rhodophyceae), yellow green algae (Class Xanthophyceae), golden algae (Class Chrysophycae) and brown algae (Class Phaeophyceae) (Chen et al., 2009). The Phaeophyceae, Rhodophyceae and the Chlorophyceae are generally referred to as macroalgae or seaweeds. Algae provide one of the main sources of food in benthic marine coastal systems and are thus one of the most important groups of primary producers (Phillips et al., 2008a). They are also of the most important ecosystem engineers, helping to fix carbon dioxide (CO_2) and thus oxygenating the water and making it possible for other organisms in the community to survive and thrive (Tsai *et al.*, 2012). Both micro- and macroalgae are used for various commercial purposes. Research has shown that in addition to being important primary producers, brown algae (Phaeophyceae) are also used for a number of commercial purposes, such as, biosorption of heavy metals (Davis et al., 2003) and various medical and pharmaceutical uses (Athukorala et al., 2007; Rupérez et al., 2002; Kelecom, 2001). Some of the most commercially important chemicals produced by brown algae are alginates. There have been several studies on the enzymes involved in alginate production, the chemical structure and biotechnological applications of alginate. However, studies on the genes controlling the biochemical pathway of alginate synthesis are lacking in number. There is a need to understand how these genes function, for both academic and commercial

reasons, as knowledge about the genetic mechanisms will lead to understanding the genes encoding the enzymes in the biochemical pathway, how the genes are expressed and variation of the expression levels spatially, seasonally and between different tissues of the same alga, thus providing information about the regulation of alginate production. The following study attempts to shed some light on the genes encoding mannuronan C5 epimerase, the last enzyme in the biochemical pathway of alginate in *Lessonia variegata* J. Agardh.

1.2 Taxonomy

Brown algae belong to the class Phaeophyceae within the division Heterokontophyta (Phillips *et al.*, 2008b). They are almost exclusively marine multicellular algae found in both tropical and temperate ecosystems (Phillips *et al.*, 2008b). The range of distribution is wide, extending from the marine ecosystems in the subpolar regions to the equator (De Reviers, Rousseau, & Draisma, 2007). They act as foundation species and structure marine ecosystems by providing food, habitat, protection and also serving as recruitment sites for a variety of marine organisms (Vásquez, 2007).

1.3 Characteristics of Phaeophyceae

Brown algae range in size from microscopic filaments to giant kelps. The life history of all brown algae, except the members of the order Fucales, exhibit alternation of generation between a haploid gametophyte and a diploid sporophyte phase (Fig. 1.1) (Graham and Wilcox, 2000). Brown algae have three types of flagellate reproductive cells – meiospores, asexual zoospores and gametes (Graham and Wilcox, 2000). These cells have typical heterokont flagella, consisting of two long, unequal flagella, one anterior and one posterior, which beat differently (Daugbjerg and Anderson, 1997). Phaeophycean motile cells are distinctive as their flagella usually emerge laterally rather than apically or sub-apically (Graham and Wilcox, 2000). In some brown algae the posterior flagellum is shorter than the anterior flagellum and some even have only a single anterior flagellum, while flagella are completely absent from the egg cells of the most oogamous species (Graham and Wilcox, 2000). Some brown algae are isogamous, that is, the flagellated gametes, while motile, cannot be morphologically distinguished from each other. However, one gamete typically settles at the bottom, while the other actively swims for a longer period (Graham and Wilcox, 2000). Anisogamous brown algae have one gamete larger than the other, while some species exhibit oogamy, where there is one non-motile egg cell and a smaller flagellate sperm cell (Graham and Wilcox, 2000). Gametes are produced by multicellular gametophytes within specialized plurilocular gametangia (Graham and Wilcox, 2000). Plurilocular gametangia are produced by mitosis and are internally subdivided into chambers, each of which contains a gamete (Graham and Wilcox, 2000). Sporophytes sometimes produce a singlecompartment reproductive structure called a unilocular sporangia, which is absent from the thalli of the gametophyte generation (Graham and Wilcox, 2000). Unilocular sporangia undergo meiosis leading to the formation of haploid zoospores. In some cases, the sporophyte produces a plurilicolar sporangia, which produces asexual zoospores that develop into multicellular thalli, which is morphologically similar to the parent (Graham and Wilcox, 2000).



Fig. 1.1: Diagram showing the alternation of generation in the life cycle of members of the Laminariales. *Source: Drawn by Sayani Ghosh, modified from Graham and Wilcox (2000).*

1.4 Morphology and Life History of Laminariales

The order Laminariales, commonly called kelp, belongs to the class Phaeophyceae. They include the largest algae in the world, reaching up to a length of around 50m (Van den Hoek *et al.*, 1995). Kelps are found in cold temperate regions dominating hard subtidal substrates. These aggregations are called kelp beds if the algae do not form a surface canopy and kelp forests if they do form a floating surface canopy (Nybakken and Bertness, 2005). The size and biomass of kelps creates an unique habitat for various species (Steneck *et al.*, 2002) by providing a source of food, substrate and protection from predators. Although the morphology varies among the members of the Laminariales, all kelp are attached to the substrate by a holdfast of branched haptera, rhizoids or disc (Lane *et al.*, 2006). A stipe arises centrally from the holdfast. The stipe

has intercalary meristematic tissue at its upper end, at the transition zone between the stipe and the terminal blade (Fig. 1.2) (Lane *et al.*, 2006). Extension of the stipe and the blade occurs due to cell divisions in this transition zone and the resulting increase in girth produced by a superficial meristem is called the meristoderm (Bold and Wynne, 1985). Kelp are the only seaweeds that have specialized cells called trumpet hyphae, present on the blades, that are used for transport of nutrients, which are subsequently stored as laminaran and mannitol (Lüning, 1990).



Fig. 1.2: The Structure of the Laminariales. Source: Diagram by Erasmo Macaya Horta. Used with permission.

The Laminariales exhibit an alternation of generations of a macroscopic sporophytes and microscopic gametophytes (Kai *et al.*, 2006). Reproduction is by unilocular sporangia, which are born in sori on general or localized portions of the sporophyte. The sporangium undergoes meiosis, leading to the formation of haploid zoospores, which then settle on a substrate and form male and female gametophytes (Dayton, 1985). The reproduction by

sporophytes is of particular importance as it helps to expand the distribution via zoospore dispersal (Kai *et al.*, 2006). The rocky, subtidal habitats of the Northern Hemisphere are dominated by members of the Laminariales, which are less abundant in the Southern Hemisphere. The Northern Atlantic only has a fraction of the species diversity found in the North Pacific, where 40 species are currently recognized along the coast of North America and 41 from Asia (Lane *et al.*, 2006). Only eight species of kelp occur in the North Atlantic. Kelp species in the Southern Hemisphere are limited in number. They are restricted to *Eisenia galapagensis* W.R. Taylor, *Laminaria abyssalis* A.B. Joly & E.C. Oliveira, *Laminaria brasiliensis* A.B. Joly & E.C. Oliveira, three species of *Macrocystis pyrifera* (Linnaeus) C. Agardh and *Macrocystis integrifolia* Bory de Saint-Vincent, the introduced species *Undaria pinnatifida* (Harvey) Suringer, and the largely Southern Hemisphere genera *Ecklonia* and *Lessonia*. Lessoniaceae is exclusively southern in distribution (Lane *et al.*, 2006).

1.5 Distribution and life history of Lessonia variegata J. Agardh

The Lessoniaceae are conspicuous and the dominant species of algae along many rocky coastlines of the Pacific and Southern Oceans (Chin *et al.*, 1991). including the rocky shores and reefs of New Zealand. The sporophytes are usually found subtidally (Nelson and Schwartz, 2005), while the microscopic gametophytes are presumed to be present in the mid to low intertidal zone (Fox and Swanson, 2007).

Lessonia variegata J. Agardh is a species which is endemic to New Zealand. There are four cryptic species of *Lessonia variegata* – the Northern lineage, extending from outer

South Head (Auckland) to East Cape; the Wellington lineage, which exends from East Cape to White Bay (Marlborough Sounds); the Kaikoura lineage, extending from Cape Cambell to Goose Bay; and the Southern lineage, extending from Tumbledown bay (Canterbury) to Stewart Island (Fig. 1.3) (Martin, 2011; Abbott, 2011). The sporophytes are distributed along some of the most exposed shorelines of New Zealand, from the upper subtidal zone to 15-20m depth (Nelson, 2005). *Lessonia spp.* exhibits the pattern of growth typical of Laminariales (Hawes *et al.*, 2004).



Fig. 1.3: Map of New Zealand showing the approximate distribution of the different lineages of *Lessonia* variegata. *Source: Martin (2011). Used with permission.*

Schwartz *et al.* (2006) conducted field experiments on *Lessonia variegata*, collected from the three sites along the Wellington south coast (Red Rocks, Island Bay and Quarry), and found that *Lessonia variegata* was fertile from April to September (autumn to spring),

based on the presence of dark brown sori on the blades, with the fertility reaching its peak in August and September. Sori were absent from the blades by November and this coincided with the beginning of summer with warmer temperatures and longer days. The elongation rate of the blades of *Lessonia variegata*, and thus the growth of the algae, was found to be maximal over the summer months from November to February. Nelson and Schwartz (2005) conducted culture experiments on male and female gametophytes of Lessonia variegata by varying the temperature, daylength and light intensity. They found that more gametes were formed under longer days (15 hours of light) and warmer (15°C) conditions, rather than shorter days (9 hours of light) and colder conditions (10° C). It was found that under low light conditions, the gametophytes grew as filamentous clumps instead of forming sperm and eggs. When the conditions were altered from winter (low light) to summer (high light) conditions, the gametophytes responded rapidly and formed gametes that underwent fertilization and developed into zygotes and ultimately formed sporophytes. Thus the gametophytes appear to serve as 'seed banks' for kelp under unsuitable conditions and produce gametes under suitable conditions (Nelson and Schwartz, 2005). In many parts of the world, kelps are harvested from wild populations and some species are cultivated in aquaculture operations to use as human food, animal feed, fertilisers and soil conditioners, as well as a source of many useful extracts such as gelling agents and pharmaceutical compounds (Graham and Wilcox, 2000).

1.6 Carbohydrates produced by brown algae and their uses

Brown algae produce various carbohydrates that are economically important and have uses in different industries ranging from food, the cosmetic and the pharmaceutical

industries to microbiology and biotechnology. Cellulose, laminaran, alginate, fucoidan and mannitol are the major polysaccharides produced by brown algae (Graham and Wilcox, 2000). All of these polysaccharides have various functions in the Phaeophyceae. One of the major photosynthetic end products, mannitol, along with laminarin (a β -1,3glucan), acts as the main source of carbon storage in the Phaeophyceae (Rousvoal et al., 2011). Fucoidans play an important role in the algal cell wall organization, and are involved in the morphogenesis of algal embryos (Bisgrove and Kropf, 2001). Fucoidans may also be involved in cross-linking alginate and cellulose (Mabeau *et al.*, 1990). Cellulose comprises about 1-10% of the thallus dry weight, alginate may comprise up to 35% of the thallus dry weight, mannitol accounts for about 20-30% (Graham and Wilcox, 2000), fucoidan may vary between 5-20% (Davis et al., 2003) while laminaran percentage may vary between 0.2-6.5% (Rioux et al., 2009). Alginates, responsible for the flexibility of different parts of the algae, are mainly present in the intercellular matrix. They play an important role in preventing dessication and also help in ion exchange (Graham and Wilcox, 2000). The chemical structures and the content of some of these carbohydrates like laminaran, mannitol and fucoidan may vary according to the season, age of population and the geographic location (Rioux et al., 2009; Rioux et al., 2010; Zvyagintseva et al., 2003).

Laminarin and fucoidans are commercially important for their various biological activities, while alginate is commercially used mainly in the food and cosmetic industry (Rioux *et al.*, 2007). Hayashi *et al.* (2008) found that sulfated fucoidans isolated from the edible brown alga *Undaria pinnatifida* inhibits the *in vivo* replication of herpes simplex

virus type 1 (HSV-1) in mice by increasing the activities of the natural killer (NK) and cytotoxic T lymphocyte (CTL) cells. They have also been found to have anti-coagulant and anti-thrombotic properties, interfere with the proliferation of cells, and may interfere with certain fertilization-related mechanisms (Berteau and Mulloy, 2003). Laminarin has been found to suppress apoptotic cell death (Kim *et al.*, 2006), stimulate humoral immunity (Awad and Osman, 2003) and show anticoagulant properties (Hawkins and O'Neill, 1955). Experiments using laminarin to treat hydrofluoric acid induced corneal burns in rabbits have shown positive results (Hatipoglu *et al.*, 2008).

1.6.1 Properties of alginate

Alginate, a cell wall polysaccharide, is believed to act as a structure forming component, having similar physiological properties as cellulose in terrestrial plants (Draget *et al.*, 2005). The intercellular alginate gel matrix is responsible for the mechanical strength and flexibility of kelps (Anderson *et al.*, 1977). Alginate is a linear polysaccharide made up of 1,4-linked- β -D-Mannuronic acid (M) and its C-5 epimer α -L-Guluronic (G) acid arranged in heteropolymeric and homopolymeric blocks (Figure 1.4) (Campos *et al.*, 1996; Nyvall *et al.*, 2003). A G block-rich alginate provides an increased rigidity to the stipes and holdfasts of the algae. On the other hand, alginates rich in M or MG blocks which are more flexible polymers, are found in blades exposed to wave action (Kloareg and Quatrano, 1988). The stipe and holdfast of *Laminaria hyperborea* (Gunnerus) Foslie, an alga that grows in very exposed coastal areas, have very high contents of guluronic acid that gives it mechanical rigidity, while the blades have much lower guluronic acid that allow them to be flexible (Draget *et al.*, 2005). The composition and content of alginate have been found to vary according to species, geographical locations and tissues. In 1984, Cragie *et al.* found that the polyguluronate composition (G- block rich alginate) was the lowest (~3%) in the blades of *Laminaria longicruris* Bachelot de la Pylaie, collected from the site with strong tidal currents, which were the most flexible, compared to blades from other sites (~14%), which were more rigid. On comparison of the alginate content in the blade and stipe of *Laminaria digitata* (Hudson) J.V. Lamouroux to that of *Laminaria longicruris*, it was found that the G-block content was much higher and the MG content lower in *Laminaria digitata* compared to *Laminaria longicruris*. Moreover, comparison of alginate composition between *Laminaria digitata* collected from Sandy Bay, Nova Scotia and that from east coast of Britain (Stockton *et al.*, 1980) showed significant differences both in the blade and stipe, thus proving that alginate composition varies intraspecifically (Cragie *et al.*, 1984).





Fig. 1.4: Diagram showing a) chemical structures of alginate monomers; b) an example of block distribution of alginate.

Source: Draget et al.(2005)

Commercially, alginate is a very valuable chemical, used in a number of industries. Alginate is able to form heat-resistant gels in the presence of Ca^{+2} (Campos *et al.*, 1996), a property which makes it useful as a gel-former, and viscosifier (South and Whittick, 1987; Fenoradosoa *et al.*, 2010) in a wide range of industrial applications (Tenhaken *et al.*, 2011), such as to immobilize enzymes (Tanriseven and Doğan, 2001), as a thickening agent in food and cosmetics (Podkorytova *et al.*, 2007; Gómez-Díaz and Navaza, 2003) or to cover organs during transplantation as a barrier between the transplant and the host immune system (Kühtreiber *et al.*, 1999). It is also used in cell encapsulation technology (Ertesvag and Valla, 1998).

b)

1.6.2 Biosynthesis of alginate

The biosynthetic pathway of alginate production in brown algae is shown in Fig. 1.5. Studies on the enzymes in the biosynthetic pathway of alginate in brown algae are very few. The final step of the pathway is the epimerization of D-mannuronic residues into L-guluronic residues in the polymer chain. This reaction is catalyzed by the enzyme mannuronan C-5 epimerase (Nyvall *et al.*, 2003). In addition to stimulating the reaction, the presence of low concentration of Ca^{2+} favours introduction of neighbouring G units, whereas at high concentration it has been found to favour the introduction of single G units, thus creating a polymer with a more alternating structure (Rødde and Larsen, 1997).



Fig. 1.5: Biosynthesis pathway of alginate in Phaeophyceae.

Source: Nyvall et al., (2003)

The GDP- mannose dehydrogenase (GMD) enzyme, which converts GDP- mannose to GDP-D-mannuronic acid via a four electron oxidation, has been studied in *Ectocarpus siliculosus* (Dillwyn) Lyngbye (Tenhaken *et al.*, 2011). The GMD enzyme in *Ectocarpus siliculosus* is significantly different from related bacterial enzymes due to the absence of a ~120 amino acid sequence at its C-terminal (Tenhaken *et al.*, 2011). The GMD from brown algae is highly active at alkaline pH (8.5-9.0) and is inactive below pH 7. It contains a catalytic Cys-residue which is highly sensitive to heavy metals. It is temperature sensitive, with the activity steadily increasing from 0°-30°C and rapidly getting deactivated at temperatures above 30°C (Tenhaken *et al.*, 2011). The enzymatic activity was found to increase in the presence of salts, the optimum being around 500mM for Na₂SO₄ and NaCl and slightly lower for KCl.

1.6.3 Mannuronan-C5-Epimerase genes involved in encoding the last enzyme in the alginate biosynthetic pathway

Studies conducted on the genes involved in the biosynthetic pathway of alginate in members of the phaeophyceae are limited in number. One of the two studies done on members of the Laminariales was conducted by Nyvall *et al.* (2003). They conducted an expressed sequence tag (EST) analysis on the kelp *Laminaria digitata* and established a total of 905 ESTs from sporophytes and gametophytes, which led to the characterization of 600 transcripts. The corresponding cDNAs of two ESTs were found to exhibit a significant similarity with the coat protein GP1 of the virus (EsV-1) infecting *Ectocarpus siliculosus*. The sequenced cDNAs, named Man-C5-E1 and Man-C5-E2, revealed a significant similarity with the mannuronan C5 epimerase (Man-C5-E) genes from the

alginate producing bacteria Azotobacter vinelandii and Pseudomonas aeruginosa. A northern blot on colonies containing cDNAs of L. digitata sporophytes, using the coding region of Man-C5-E1 as a probe, revealed four new cDNAs (Man-C5-E3-E6). The analysis of the amino acid sequences of the six cDNAs indicated that Man-C5-E1 and Man-C5-E6 were the only two full length cDNAs, encoding proteins of 55 kD. All the other cDNAs were partial cDNAs. The six cDNAs had large, variable 3'-UTRs, varying in length between 1318-1825 base pairs (bp) and shared a more or less conserved section (802 bp or 267 amino acid long) of the coding region (54% nucleotide identity and 69% amino acid identity). The 5' end was also variable between the copies. This indicated the presence of multiple copies of the Man-C5-E gene. The expression levels of the Man-C5-E genes were examined by northern blot analysis and reverse transcriptase PCR (RT-PCR) in sporophytes collected at regular intervals over the whole year. A single 3.5 kb transcript was observed, based on the hybridization with the cDNA probe (Man C5-E1). However, cloning and sequencing of the RT-PCR products indicated the presence of 16 different epimerase genes in the sporophytes. Both northern blot and RT-PCR showed that the transcript expression level was the highest in winter and early spring, which was also when the seawater nutrient levels and the active growth and alginate content of *Laminaria digitata* were maximal in Brittany, France.

This was the first study to discover the presence of multiple copies of the Man-C5-E gene and the fluctuations in the expression levels of the copies, in a member of the Laminariales, thus throwing some light on the genetical mechanisms behind the last enzyme in the alginate biochemical pathway. Inoue *et al.* (2007) conducted a study on Man-C5-E genes in *Saccharina japonica* (Areschoug) C.E. Lane, C. Mayes, Druehl and G.W. Saunders. Although the study was not published, the sequence of the gene is available from GenBank. In the Ectocarpales, the sequenced genome of *Ectocarpus siliculosus* (Cock *et al.*, 2010) revealed the presence of 20 copies of the gene and 25 isoforms of the Man-C5-E enzyme.

1.6.4 Alginate production in Lessonia variegata

Lessonia variegata, like all brown algal species, produces the cell wall building chemical alginate. Alginate is the major matrix component of the brown algal cell wall, accounting for 10.8% to 30.9% of the dry weight in the Wellington lineage (Abbott, 2011). Abbott (2011) measured the amount of alginate in *Lessonia variegata* from March 2010 to February 2011 and her findings indicated that the alginate levels fluctuate throughout the year, showing the highest levels in May, November and December (Fig. 1.6).



Fig 1.6: Histogram showing the monthly mean yield (\pm standard deviation, n= 9-10) for alginate within the Wellington lineage of *Lessonia variegata* for the period of March 2010 – February 2011, (expressed as percent dry weight of the *Lessonia variegata* milled samples). Columns with the same letters are not significantly different at P < 0.05 *Source: Abbott (2011). Used with permission.*

Skriptsova *et al.* (2004) found that in *Undaria pinnatifida* (Laminariales, Phaeophyceae) the alginate content was highest just before plant maturation and sporophyll development, which occurs during winter. This was hypothesized to be related to the development of reproductive structures. A fluctuation of alginate content is also seen in Abbott's (2011) findings. The highest level of alginate production is seen in November, which coincides with the maximum growth rate and rate of blade elongation in *Lessonia variegata* (Schwartz *et al.*, 2006). Since alginate is a cell wall polysaccharide responsible for maintaining the structure of the cell wall, the production of alginate is expected to increase during the growth of the algae.

There is only one study on the alginate production of *Lessonia spp.*, conducted on *Lessonia variegata* (Abbott, 2011). The studies which have been conducted all focus on the chemistry of the alginate and its biological and commercial uses. There is no study on the biochemical pathway of alginate synthesis or the enzymes and genes involved in the pathway in *Lessonia spp.* It is assumed that the biochemical pathway of alginate synthesis is similar to that of other members of the Laminariales. However, as the genome of this species has not yet been sequenced, there is no information about the genes encoding the enzymes in the pathway. Due to the lack of data about the genetic sequences, the pattern of expression of the genes is not yet known.

1.7 Outline and scope of this study

This study aims to develop a better understanding of the genetic mechanisms leading to the production of alginate in the endemic New Zealand species *Lessonia variegata*. To start this, I have attempted to characterize the genes encoding mannuronan C5 epimerase on the four lineages of *Lessonia variegata*. I have compared the genetic sequences of the four lineages to check for the presence of multiple gene copies, thus gathering more information about the phenomenon of gene duplications, multigene families and genes under selection. I have also compared the Man-C5-E gene sequences of *Lessonia variegata* with those of *Laminaria digitata* and *Saccharina japonica* to examine if the genes share a monophylectic origin.

This study is important in increasing our knowledge about *Lessonia variegata*, both in an academic and commercial sense, as this species could be a very important source of alginate production and extraction for a lot of industries. Understanding how the genes

behind the biosynthetic pathway of alginate production work could potentially lead to further studies on discovering more copies of the gene, the expression levels of the gene paralogs and fluctuations in the expression levels seasonally, spatially and also between different tissues of the same algae. This knowledge would be valuable in providing information about manipulating the genes to increase alginate production.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Sampling

Previous samples collected and prepared by a PhD student, Peter Martin, were used for this study. For this study, six samples of the Wellington lineage, one sample from the Northern lineage, one from the Kaikoura lineage and one from the Southern lineage were used (Table 2.1).

2.2 DNA Extraction

Peter Martin extracted the DNA using a modified CTAB protocol (Zuccarello and Lokhorst, 2005).

2.3 PCR with Primer Set I

2.3.1 Primer Set I designed from Genbank data

Primer set I was designed by downloading and aligning the cDNA sequences of *Laminaria digitata* and *Saccharina japonica* from GenBank (Table 2.2), using the bioinformatics software Geneious (Drummond *et al.*, 2012). The most conserved region was found to be 568 bp long, between nucleotide positions 1056 and 1623 (64.3% nucleotide identity), corresponding to *Laminaria digitata* Man-C5-E6 sequence. The aligned sequences in this region were thus used to design two primers, called C5-1056F and C5-1623R (Table 2.3), to amplify ~500bp of the coding region of the gene of interest in *Lessonia variegata*.

2.3.2 DNA Amplification

All PCR amplifications were carried out in a PTC – 100 Thermal Cycler (MJ Research). Each 30 µl reaction volume contained 1 µl of diluted DNA sample, 1X ThermoPol Reaction buffer (New England Biolabs, Ipswich, MA, USA), 0.83 mM dNTP, 0.5 mM Mg²⁺ (New England Biolabs), 0.25% BSA (New England Biolabs), 0.75 nM of each primer, and 1U of *Taq* polymerase (New England Biolabs). A touchdown PCR was carried out on the DNA sample C568. An initial denaturation step at 94°C for 4 minutes was followed by 9 cycles, where the initial annealing temperature of 55°C decreases by 1°C each cycle until it reaches 45°C. This was followed by 29 cycles at an annealing temperature of 45°C. An initial denaturation step at 94°C for 4 minutes was followed by 37 cycles (1 minute at 94°C, 1 minute at 45°C and 1 minute at 72°C), followed by a final extension step at 72 °C for 10 minutes.

2.3.3 DNA extraction and ligation

PCR products were checked for yield on 1% agarose gel. The bands were excised using a clean scalpel and extracted according to the specifications of a Qiagen Gel extraction kit. The excised gel slice was weighed and an equal volume of binding buffer was added to it. The mixture was then incubated at 58° C till the gel was completely melted and then the mixture was vortexed for 3 minutes. 600-700 μ l of the DNA/agarose solution was placed in a Hibind[®] mini column and centrifuged at 10,000x g for 1 minute. The flow through was discarded and 300 μ l of binding buffer was added to the same Hibind[®] mini column

and centrifuged at 10,000x g for 1 minute at room temperature to wash the column. The flow through was discarded and the column was washed by adding 700 μ l of SPW wash buffer diluted with absolute ethanol. It was then centrifuged at 10,000x g for 1 minute and the step was repeated. The column was then placed into a clean microcentrifuge tube and 30 μ l of elution buffer was added directly to the column matrix, incubated at room temperature for 1min and then centrifuged for 1minute at maximum speed to elute the DNA.

Extracted DNA was then ligated into pGEM® T-Easy vector (Promega, Madison, USA) by adding 1µl of vector, 1X ligation buffer (Promega, Madison, USA), 3U of T4 DNA ligase (Promega, Madison, USA) to 7 µl of purified PCR product to make a final volume of 10 µl. This mixture was then put in a water bath and incubated at 4°C overnight.

2.3.4 Making LB-agar-ampicillin plates

LB-Agar Plates were made by adding 1% agar to 500ml of Luria Broth (LB) (Sigma-Aldrich, New Zealand), which was then autoclaved. 1ml of Ampicillin (50mg/ml) was added to the mixture after cooling it down to 55°C, followed by pouring the mixture into petri dishes. 20µl each of X-gal (20mg/ml) and IPTG (0.1M) were then added to the LB-Agar plates. The plates were then left partially open for 20 minutes for the X-gal and IPTG to dry out.

2.3.5 Preparation and transformation of chemically competent E. coli

The DH5α *Escherichia coli* cells were made according to the protocol adapted by Jeremy Owen (2010) based on the method described in Hanahan (1983).

2.3.6 Transformation of competent DH5α E. coli cells

Competent DH5 α bacterial cells were transformed by adding 10µl of ligation to 100µl of competent cells, followed by leaving the cells on ice for 20 minutes and heat shocking at 42°C for 90 seconds. 900µl of LB was then added and the cells were allowed to recover for 1 hour at 37°C. Thereafter, 250µl of transformed DH5 α cells containing the ligated DNA were added to each plate and colonies were grown overnight at 37°C.

White colonies were picked with sterile toothpicks and transferred to another LB-agarampicillin plate (without X-gal and IPTG), along with one blue colony, and grown overnight at 37°C.

2.3.7 Miniprepping of plasmids

Colonies grown on the master plates were then picked with sterile toothpicks and inoculated in 3ml of LB and 6µl of Ampicillin (50mg/ml) and grown overnight at 37°C. The bacteria from the culture tube was then transferred to a 1.5ml eppendorf tube and centrifuged for 5 minutes at 13,000 rpm at 4°C. The bacterial pellets were then resuspended in 100µl of ice cold solution I (1M glucose, 1M Tris, pH 8.0, 0.5 EDTA, pH 8.0) by vortexing until all the pellets are completely dissolved. 200µl of freshly prepared solution II (0.2N NaOH, 1% sodium dodecyl sulphate) was then added and the contents mixed by inverting the tubes five times. 150µl of ice cold solution III (5M potassium acetate, 96% glacial acetic acid) was then added and the contents mixed by inverting five times. This mixture was then centrifuged at maximum speed for 20 minutes at 4°C, followed by transferring the supernatant to a fresh tube. 450µl of 1:1 phenol:chloroform was then added to each tube. The mixture was shaken well and centrifuged at maximum speed for 5 minutes. 400µl of the upper aqueous layer was then transferred to a fresh tube and an equal volume of 24:1 chloroform-isoamyl alcohol mixture was added to each tube. This was mixed and centrifuged at maximum speed for 2 minutes. 350µl of the upper aqueous layer was transferred to a fresh tube and 870 µl of cold 99% ethanol was added to each tube. The mixture was centrifuged at maximum speed for 15 minutes at 4°C. The supernatant was then removed and 200µl of 70% ethanol was run down the sides of each tube and the mixture centrifuged for 5 minutes at maximum speed. The supernatant was removed and the pellet air dried. The nucleic acid pellet was then redissolved in 40 µl of sterile water.

2.3.8 Restriction digestion of plasmid

Restriction digestion of the miniprepped plasmid was carried out by adding 3µl of the plasmid, containing the insert, to 1X *Eco*RI buffer (Roche Applied Science, New Zealand), 10U of the restriction enzyme *Eco*RI (Roche Applied Science, New Zealand), 1µl of 2mg/ml RNAse A (Invitrogen, Life technologies), and 4µl of sterile water, to make up a total volume of 10µl. This was incubated at 37°C for 3 hours. The resulting digested plasmid was then run on a 1% agarose gel to check the length of the inserts, which were then commercially sequenced (Macrogen Inc., Seoul, South Korea).

2.4 PCR with Primer Set II

2.4.1 Primer Set II designed for the specific amplification of Lessonia variegata sequences

The three gene sequences obtained were aligned, using Geneious (Drummond *et al.*, 2012), with the Man-C5-E gene sequences of *Laminaria digitata* and *Saccharina Japonica* to design primer set II, Lv.E-F49 and Lv.E-R473 (Table 2.3), specific for the Wellington lineage of *Lessonia variegata*. This was done to minimize chances of non-specific amplification of DNA sequences.

PCR amplifications were carried out at an annealing temperature of 55°C with the new primers on all 4 lineages. An initial denaturation step at 94°C for 4 minutes was followed by 37 cycles (1 minute at 94°C, 1 minute at 55°C and 1 minute at 72°C), followed by a final extension step at 72 °C for 10 minutes.

2.4.2 Ligation, PCR clean up and transformation

As the PCR amplification by the primers Lv.E-F49 and Lv.E-R473 (Table 2.3) showed multiple bands. 5% DMSO was added to increase specificity. The PCR products were then purified according to the specifications in the High Pure PCR product purification kit (Roche Applied Science, New Zealand) to get more concentrated DNA for ligation. 25µl of PCR product was mixed with 125µl of binding buffer. The mixture was then transferred to the upper reservoir of a High Pure Filter Tube, inserted into a collection tube, and centrifuged for 1 minute at 13,000x g at room temperature. The flow through was discarded and 125µl of wash buffer was added to the upper reservoir, which was

then centrifuged for 1 minute at 13,000x g at room temperature. The resulting flow through was discarded and 50µl of wash buffer was added to the filter tube and centrifuged as before. The collection tube with the flow through was discarded and the filter tube was inserted into a 1.5ml microcentrifuge tube. 10µl of elution buffer was added to the filter tube, and centrifuged for 1 minute at maximum speed. The resulting flow through contained purified PCR product.

The PCR products were then directly ligated into pGEM® T-Easy vector (Promega, Madison, USA) by adding 1µl of vector, 1X ligation buffer (Promega, Madison, USA), 3U of T4 DNA ligase (Promega, Madison, USA) to 7µl of PCR product to make a final volume of 10 µl. This mixture was then put in a water bath and incubated at 4°C overnight.

LB-agar plates were made following the procedure described previously. Chemically competent JM109 High Efficiency Competent Cells (Promega, Madison, USA) were transformed with the ligated DNA by following the transformation procedure provided by the manufacturer. 2µl of each ligation was added to a 1.5ml microcentrifuge tube set on ice. Frozen JM109 cells were placed in an ice bath until just thawed. The competent cells were then mixed by gently flicking the tube and 50µl of competent cells were transferred to each of the tubes containing ligations. The tubes were gently flicked to mix the contents and placed on ice for 20 minutes. The cells were then heat-shocked for 45-50 seconds in a water bath at exactly 42°C. The tubes were then immediately returned to the ice for 2 minutes. Then 950µl of room temperature super optimal broth with catabolite

repression (SOC) (2% bactro-tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl₂, 20mM glucose) was added to the tubes and the mixture was incubated for 1.5 hours at 37°C with shaking. 100 μ l of each transformation was then plated onto each LB/ampicillin/X-gal/IPTG plates. The plates were then incubated overnight (16-24 hours) at 37°C to facilitate the growth of bacterial colonies.

White colonies were picked with sterile toothpicks and transferred to another LB/agar/ampicillin plate (without X-gal and IPTG), along with one blue colony, and grown overnight at 37°C.

2.4.3 Colony PCR

PCRs were carried out on the bacterial colonies growing on the LB/agar/ampicillin plate, in order to determine if the plasmids in the competent cells contained the desired insert (~500bp). All PCR amplifications were carried out in a PTC – 100 Thermal Cycler (MJ Research). Each 30 μ l reaction volume contained 1X ThermoPol Reaction buffer (New England Biolabs, Ipswich, MA, USA), 0.83 mM dNTP, 0.5 mM Mg²⁺ (New England Biolabs), 0.25% BSA (New England Biolabs), 0.75 nM of each primer (Primer Set II; Table 2.3), and 1U of *Taq* polymerase (New England Biolabs). Sterile toothpicks were used to pick up each colony and then dipped into the 1.5ml microcentrifuge tube containing the PCR mixture and swirled around to mix well. PCR amplifications were carried out at an annealing temperature of 60°C for all samples. An initial denaturation step at 94°C for 4 minutes was followed by 37 cycles (1 minute at 94°C, 1 minute at 60°C and 1 minute at 72°C), followed by a final extension step at 72 °C for 10 minutes. PCR products were then run on 1% agarose gel to check for the presence of the desired insert. The samples containing the insert were then purified with ExoSAP-IT (USB, Cleveland, Ohio, USA) and sent away to be commercially sequenced (Macrogen Inc., Seoul, South Korea).

2.5 Analysis of obtained DNA sequences

An unrooted neighbour-joining tree consisting of all the cloned Man-C5-E DNA sequences (full 429bp dataset) was constructed using Jukes-Cantor (Jukes and Cantor, 1969) distances. This was done in Geneious (Drummond *et al.*, 2012). The resulting topology showed gene clusters that indicated the presence of multiple identical sequences per sample (Appendix 1). Each cloned Man-C5-E gene sequence was thus manually checked and identical sequences as well as sequences differing from each other by only one or two base pairs, removed. The clones were also screened for singletons and chimeric PCR products (Wang and Wang, 1997). 'Singletons' or base pairs which differ in position in one cloned sequence when compared with other clones in the same sample, could be errors which normally arise due to the non-proof reading ability of *Taq* DNA polymerase (Zuccarello et al., 2009). These point mutations, which are not restricted to third codon positions, and can occur randomly in first and second codon positions as well, were checked manually and removed. The clones were also checked for chimeric PCR products (PCR-produced errors in the form of mixed template sequences) using the Bellerophon server (Huber et al., 2004). A BLAST search (blastx) of the samples was done to confirm the correct orientation (5'-3') and to check for homology to existing algal

Man-C5 epimerases in the GenBank database. Each alignment was edited and formatted in Se-Al (Rombaut, 1996) prior to constructing trees.

The final *Lessonia variegata* Man-C5-E gene dataset consisted of 20 sequences, belonging to the three lineages (Northern, Southern and Wellington lineage) and a maximum of 429 characters per sequence. Two unrooted neighbour joining (NJ) trees, using Jukes-Cantor (Jukes and Cantor, 1969) distances, were constructed, one with the first 127 characters (the exon) and one with the remaining 302 characters (the intron) (refer to section 3.1.2). For bootstrap analysis (BP), 1000 replicates were generated from resampled data. The NJ trees were constructed in Geneious (Drummond *et al.*, 2012). A haplotype data file of the first 127 characters of 20 Man-C5-E DNA sequences was generated using DnaSP v5 (Librado and Rozas, 2009). A haplotype network was calculated and drawn in Network 4.6.1.0 (Fluxus Technology Ltd.) using median-joining (Bandelt *et al.*, 1999).

A Man-C5-E DNA dataset consisting of 29 sequences, belonging to 4 species (*Lessonia variegata, Laminaria digitata, Saccharina japonica* and *Ectocarpus siliculosus*) and the 127 characters of the exon (28 parsimony-informative positions, 9 parsimony uninformative positions and 90 constant characters), was analyzed using three algorithms: a neighbour-joining procedure was implemented in PAUP*4.0b10 (Swofford, 2002). General Time Reversible (GTR) (Tavaré, 1986) distances were used to determine bootstrap percentages (1000 replicates), which was used to construct a NJ tree; a maximum likelihood (ML) analysis was implemented in PAUP*, using GTR+ proportion
invariant + gamma (GTR+I+G) (Tavaré, 1986) distances, with 1 random sequence addition, tree-bisection-reconnection (TBR) branch swapping, unordered and unweighted characters, and gaps treated as missing data. For bootstrap analysis (BP), 100 bootstraps were generated from resampled data (10 random sequence additions). The models for both the analyses were executed in ModelTest 3.7 (Posada and Crandall, 1998), using the Akaike Information Criterion (AIC) (Akaike, 1974). A Bayesian tree was inferred using Mr Bayes 3.1.2 (Huelsenbeck and Ronquist, 2001) using two parallel runs of each three heated chains and a cold one, and 3 X 10^6 generations with sampling every 100 generations. Tracer v1.4 (Rambaut and Drummond, 2007) was used to check the convergence between the runs and the stationarity of scores. A burn-in of 3000 trees was removed from each run separately, and the posterior probability estimated from the remaining 27000 trees. Transition model + Proportion invariant + gamma (TIM+I+G) distances for Mr Bayes analysis were estimated in ModelTest using AIC.

2.6 RNA extraction

Lessonia variegata samples were collected from Moa Point, Island Bay, Wellington for RNA extraction. Tissues $(1-2 \text{ cm}^2)$ were collected from the base (youngest part), middle and tip (oldest part) of the blades, which were then patted dry, sealed in bags, frozen immediately in liquid nitrogen and stored at -80°C.

The TRIzol[®] method (Invitrogen, Life technologies) was used to extract RNA. A sterilized mortar and pestle was used for homogenization of the tissues. Liquid nitrogen was added to the mortar and pestle and the tissue was ground up until powdery. The

tissue samples were then homogenized in 1ml of TRIzol[®] Reagent per 50-100 mg of tissue. The homogenate was then centrifuged at 12,000x g for 10 minutes at $2^{\circ}-8^{\circ}C$ to remove excess polysaccharides. The supernatant containing RNA was then transferred to a new 1.5ml microcentrifuge tube. The tubes were then incubated at $15^{\circ}-30^{\circ}$ C for 5 minutes to allow the nucleoprotein complex to dissociate completely. Then 0.2ml of chloroform per 1ml of TRIzol[®] Reagent was added, followed by vigorous shaking by hand for 15 seconds and incubating the tubes at 15°-30°C for 2-3 minutes. The samples were then centrifuged at 12,000x g for 15 minutes at 2°-8°C. Following centrifugation, the mixture separated into a lower red phenol-chloroform phase, an interphase, and a colourless upper aqueous phase containing RNA. The aqueous phase was then transferred to a fresh tube and the RNA was precipitated by adding 0.5ml of isopropyl alcohol per 1ml of TRIzol[®] Reagent added for homogenization. The samples were then incubated at 15°-30°C for 10 minutes and then centrifuged at 12,000x g for 10 minutes at 2°-8°C. A gel-like pellet was seen on the bottom of the tube. The supernatant was then removed and the pellet was washed by adding 1ml of 75% ethanol (in DEPC-treated water) per 1 ml of TRIzol[®] Reagent added for homogenization. The sample was then centrifuged at 7,500x g for 5 minutes at 2° -8°C. The RNA pellet was then briefly air-dried for 5 minutes, taking care not to let it dry out completely. The pellet was then dissolved in 40μ l of RNase-free water (0.01% diethylpyrocarbonate or DEPC added to sterile water, left overnight and autoclaved) by passing the solution a few times through a pipette tip and incubating at 55°-60°C for 10 minutes.

The samples were then run on 1% agarose gel to check for bands of RNA.

Table 2.1 Location of samples used in this study

The following table lists the *Lessonia variegata* samples used in this study, the region from where it was collected, the geographical coordinates of the location, the name of the collector and the date it was collected.

Sample no.	Species	Phytogeographic region	Sample site	Latitude N-S	Longitude W-E	Collection	Date
A103	L.variegata	Kaikoura	Paia Point	42°29'22.43"S	173°31'27.05"E	Peter Martin	06.02.10
B070	L.variegata	Northeastern	Cape Wiwiki	35°09'22.56"S	174°07'19.62"E	Peter Martin	07.02.15
C132	L.variegata	Fiordland	Doubtful Sound at Causet Cove	45°17'47.63"S	166°54'33.01"E	Wendy Nelson	08.01.25
C318	L.variegata	Cook	Turakirae Head 500	41°26'18.63"S	174°55'10.87"E	Peter Martin	08.12.01
C549	L.variegata	Cook	Cape Palliser 0	41°36'44.84"S	175°17'49.87"E	Peter Martin	09.01.29
C550	L.variegata	Cook	Cape Palliser 0	41°36'44.84"S	175°17'49.87"E	Peter Martin	09.01.29
C568	L.variegata	Cook	Cape Palliser 300	41°36'43.24"S	175°17'58.33"E	Peter Martin	09.01.29

Table 2.2 Brown algal sequences used in the study

The following table lists the brown algal sequences downloaded from GenBank, which were used in this study: the name of the species and the gene, the GenBank accession numbers and the source article.

Species	GenBank Accession Numbers	Source
<i>Laminaria digitata</i> Man- C5-E6	AJ496454	Nyvall <i>et al.</i> , 2003
<i>Laminaria digitata</i> Man- C5-E5	AJ496453	Nyvall <i>et al.</i> , 2003
<i>Laminaria digitata</i> Man- C5-E4	AJ496452	Nyvall et al., 2003
<i>Laminaria digitata</i> Man- C5-E3	AJ496451	Nyvall et al., 2003
<i>Laminaria digitata</i> Man- C5-E2	AJ496450	Nyvall et al., 2003
<i>Laminaria digitata</i> Man- C5-E1	AJ496449	Nyvall <i>et al.</i> , 2003
Saccharina japonica	AB299380	Inoue <i>et al.</i> , 2007
Ectocarpus siliculosus	FN648516	Cock et al., 2010
Ectocarpus siliculosus	FN648702	Cock et al., 2010

Table 2.3 Primers designed for this study

The table lists the primers that were designed for this study. C51056 and C51623 are the primers, obtained by aligning *Laminaria digitata* and *Saccharina japonica* sequences, used in the original PCR, while the rest were designed specifically for *Lessonia variegata*, to be used in the later PCR. The names given to each primer, the DNA sequences, melting temperature (Tm) and percentage of GC content of each primer are noted below; F= Forward primer, R = Reverse primer.

Primers	Sequence (5'-3')	Primer Set	Melting Temperature	%GC
C5-1056 (F)	CAC RTA CGG YCA CSA GMA GG		(1 m) 59 3°C	62.5
(nucleotide			57.5 C	02.5
1056-1076)		T		
C5-1623 (R)	SMG TCB TCG AAG CGG ATC GT	I	59.9°C	60.8
(nucleotide				
1603-1623)				
Lv.E-F49 (F)	TGG AGG AGG AAC AAG ATG CAC		60.9°C	54.1
(nucleotide	GAC			
49-73)				
Lv.E-	TAC GGC CAC GAG CAG GCG G	II	65.5°C	65.5
R473(R)				
(nucleotide				
454-473)				

CHAPTER THREE

3.1 RESULTS

3.1.1 Amplification with Primer Set I

The restriction digestion and sequencing of the cloned products from the original PCR, using Primer Set I (Table 2.3) produced three distinct *Lessonia variegata* DNA sequences, differing by 8-14bp (Table 3.1). Each of the obtained sequences were 474 bp long and shared 97.3% nucleotide identity between them. When aligned with the *Laminaria digitata* and *Saccharina japonica* epimerase sequences, the sequences were found to be aligned between nucleotide positions 1057 and 1540 of *Laminaria digitata* Man-C5-E6 sequence (Genbank Accession number: AJ496454; Table 2.2). However, the first 156bp of the *Lessonia variegata* sequences showed the highest nucleotide identity (79.5%) with the *L. digitata* and *S. japonica* sequences (between nucleotide positions 1057 - 1212; Table 2.2), while the remaining 318bp of the amplified sequences could not be aligned sensibly, due to the presence of a high amount of gaps. Stop codons in all reading frames were observed in all the *L. variegata* sequences between positions 181-494 (corresponding to *Lessonia variegata* sequence alignment).

Table 3.1: List of the observed number of nucleotide changes (bp) between three *Lessonia variegata* sequences from the Wellington lineage C568, obtained from primer set I.

Number of nucleotide changes (bp)	C568.1	C568.2	C568.3
C568.1			
C568.2	10		
C568.3	8	14	

3.1.2 Sequences obtained from PCR with Primer Set II

The cloning PCR using Primer Set II (Table 2.3; see Materials and Methods) produced multiple Man-C5-E gene sequences for the Wellington, Northern and Southern lineages, the longest amplicon size being 429bp. Each of the sequences was assigned a number. The sequences were compared manually to identify identical copies and singletons (see Materials and Methods), which were then deleted. This produced a total of 20 sequences for *Lessonia variegata* (Table 3.2). PCR amplifications in *Lessonia variegata* were unsuccessful in amplifying from the Kaikoura lineage (Table 2.3).

When the full 429bp *Lessonia variegata* dataset was aligned, the first 127bp of the alignment was found to be very conserved, with base pair variations mostly seen at the third codon positions. The remaining 302bp alignment showed more variations in the alignment, with the presence of gaps and stop codons after position 180 in the alignment. Hence, the alignment was divided up into two parts - the first 127bp (deduced to be

exons) and the remaining 302bp (deduced to be introns). Each part was analysed

separately.

Table 3.2: List of *Lessonia variegata* sequences (only variable positions shown) obtained from primer set II

Each sequence per sample was assigned an unique number. The table lists the DNA isolation number-*Lessonia variegata* lineage-clone number of each sample and the point mutations observed at each position in the first 127bp (exon) of the alignment of the *Lessonia variegata* clones. Non-synonymous mutations are highlighted; bold:1st codon transition, Italics and underlined: 2nd codon transition.

					Poir	nt Mu	itatio	ons (I	Positi	ons i	n the	aligi	nment	:)	
Samples	Lineage	9	18	21	30	36	39	43	50	73	75	78	87	111	120
B70-N-1	Northern	С	Т	С	С	Т	Т	С	Т	Α	С	Α	С	Т	С
B70-N-2	Northern	С	Т	С	С	Т	Т	С	Т	А	С	G	С	С	Т
B70-N-12	Northern	C	Т	С	С	С	С	Т	Т	А	С	А	Т	С	С
C132-S-1	Southern	Т	Т	Т	Т	С	С	Т	Т	А	С	G	С	С	Т
C132-S-2	Southern	С	Т	С	С	Т	Т	С	Т	А	С	G	С	С	Т
C132-S-5	Southern	Т	Т	Т	С	С	С	Т	<u>C</u>	А	С	А	С	Т	C
C132-S-7	Southern	Т	Т	Т	Т	С	С	Т	Т	А	С	А	С	Т	С
C132-S-22	Southern	Т	Т	Т	С	Т	Т	С	Т	А	С	А	С	С	Т
C318-W-1	Wellington	С	Т	С	С	Т	Т	С	Т	А	Т	G	С	С	Т
C318-W-2	Wellington	Т	С	С	С	Т	Т	С	Т	А	С	G	С	Т	С
C549-W-2	Wellington	С	Т	Т	С	С	С	Т	Т	G	С	А	С	Т	С
C549-W-4	Wellington	С	Т	С	С	Т	Т	С	Т	А	Т	G	С	С	Т
C549-W-10	Wellington	С	Т	С	Т	Т	Т	С	Т	А	Т	G	С	С	Т
C549-W-17	Wellington	Т	С	С	С	Т	Т	С	Т	G	С	G	С	Т	С
C550-W-1	Wellington	С	Т	Т	С	С	С	Т	<u>C</u>	А	С	А	С	Т	С
C550-W-2	Wellington	С	Т	Т	С	С	С	Т	Т	А	С	А	С	Т	С
C550-W-5	Wellington	С	Т	С	С	Т	Т	С	Т	А	Т	G	С	С	Т
C550-W-6	Wellington	Т	С	Т	С	С	С	Т	Т	А	С	А	С	Т	С
C550-W-12	Wellington	Т	С	С	С	Т	Т	С	Т	А	С	G	С	Т	С
C550-W-27	Wellington	Т	С	С	С	Т	Т	С	Т	А	С	G	С	С	Т

An unrooted neighbour joining tree (Jukes-Cantor distances) of the first 127 bp (exon) of *Lessonia variegata* lineages produced two distinct groups (Fig. 3.1). A weakly supported relationship (<50%) of the groups clearly showed the presence of at least two distinct copies of the Man-C5-E gene in *Lessonia variegata*. The groups differed mainly due to base pair changes in the third codon positions (10 variable sites), all of them being

synonymous mutations. Non-synonymous mutations were only observed in five sequences at positions 50 and 73 in the alignment, due to transitions at the second and first codon positions respectively (Table 3.2).

In Group 1(92.1% nucleotide identity), gene clusters consisting of samples from all three lineages were observed. A sample from the Southern lineage (C132-S-5) and one from the Wellington lineage (C550-W-1) were found to share a common nonsynonymous transition at a second codon position (Table 3.2), causing an amino acid substitution (isoleucine \rightarrow threonine). A non-synonymous transition leading to the amino acid change (asparagine \rightarrow aspartic acid) was seen in C549-W-2 (Table 3.2).

In Group 2 (92.9% nucleotide identity), the Wellington lineage samples C549-W-4, C318-W-1 and C550-W-5 were found to have identical sequences, sharing point mutations at positions 21, 36, 39, 43, 75, 78, 111 and 120 (Table 3.2). The Northern and Southern lineage samples were found to cluster together, rather than with the Wellington lineage samples, with B70-N-2 and C132-S-2 having identical sequences (Fig. 3.1). Interestingly, the non- synonymous transversion in the first codon position leading to an amino acid change (asparagine \rightarrow aspartic acid) in C549-W-2 (Group 1) was also observed in C549-W-17 (Group 2) (Table 3.2).



Fig. 3.1: Unrooted neighbour joining topology of Man-C5-E exon sequences (127 characters, Jukes-Cantor distances) in Northern, Southern and Wellington lineages of *Lessonia variegata*. Lineages are color-coded. B70-N Northern lineage (green); C132-S Southern lineage (red); C318-W, C549-W, C550-W Wellington lineage (blue)

The two gene clusters (Fig. 3.1) are defined as two groups (exon group 1 and 2) based on the number of base pair differences between the groups, ranging from 5-11bp (Fig. 3.2; Table 3.3). The variations within each exon group, was found to be less than that between the two groups (Table 3.3). The number of base pair differences within exon group 1 was found to be within 1-6bp, while that within exon group 2 was found to range from 0-7bp (Table 3.3). The haplotype network (Fig. 3.2) shows the base pair positions (in the *Lessonia variegata* alignment) where the substitutions have taken place, and thus indicates the number of base pair changes between sequences. The number of possible alleles in each sample is shown in Table 3.4.



Fig. 3.2: Haplotype network showing nucleotide substitutions between sequences. Numbers on branches indicate positions where the transversions have taken place in the *Lessonia variegata* lineages, also listed in Table 3.1. B70-N: Northern lineage (green); C132-S: Southern lineage (red); C318-W, C549-W, C550-W: Wellington lineage (blue).

Table 3.3: List of observed range of nucleotide changes within and between exon groups. The nucleotide change has been calculated per lineage: W = Wellington lineage consisting of samples C550, C549 and C318; N= Northern lineage consisting of sample B70; S= Southern lineage consisting of sample C132

Number		Exo	n Group 1		Exon Group 2			
of				ä			~	
nucleotide		W	N	S	W	Ν	S	
changes	(bp)							
Exon	W	1-3						
Group								
1	N	3-5	N/A					
	S	1-6	5-6	3-5				
Exon Group	W	5-11			0-7			
2	N	7-9	5-6		1-5	3		
	S	6-11	6-9	6-9	1-7	0-3	2	

Table 3.4: A list of possible alleles in each exon group, as seen in fig 3.1

Lineages	Samples	Exon Group 1	Exon Group 2
		Alleles	Alleles
Northern	B70-N	1	2
Southern	C132-S	2	2
Wellington	C318-W	-	2
	C549-W	1	3
	C550-W	3	3

A very low nucleotide identity (52.3%) was observed when the remaining 302bp of the *L. variegata* sequences were aligned with each other. All three lineages were found to have stop codons after position 180, in all three reading frames. When aligned with the Man-C5-E gene copies of *Laminaria digitata* and *Saccharina japonica* (Table 2.2), the alignment was found to be very poor (15.6% nucleotide identity) with a high amount of gaps after the first 127bp. This poor alignment, combined with the presence of stop codons and the high variability between sequences, indicates the presence of an intron in the gene, after the first 127bp of the sequence. This is further supported by the presence of the GT sequences at position 127-128 (corresponding to all *Lessonia variegata* sequences), as per the GT-AG intron splice junction rule (Mount, 1982). However, the putative ending of the intron, signalled by the presence of the AG sequence was not found.



Fig. 3.3: Unrooted neighbour-joining topology of the intronic sequences of Man-C5-E gene (Jukes-Cantor distances) of the Northern, Southern and Wellington lineages of *Lessonia variegata*. Lineages are colour-coded. B70-N Northern lineage (green); C132-S Southern lineage (red); C318-W, C549-W, C550-W Wellington lineage (blue)

When an unrooted neighbour-joining tree (Jukes Cantor distances) was constructed with the intron sequences of the Man-C5-E gene (Fig. 3.3), the sequences were found to form different clusters than those observed previously (Fig. 3.1). A BLAST search to identify the presence of transposons in the introns yielded no result.

The search for chimeric sequences in the dataset indicated the presence of one such sequence in the Wellington lineage, C549-W-17. Bellerophon (see Materials and Methods) indicated that the parent sequences of the putative chimera, were C549-W-1 and C549-W-2. The chimeric breakpoint was observed to be present within the intronic region (nucleotide position 200 in *Lessonia variegata* sequences). Manual inspection revealed a high percentage (>95%) of nucleotide identity of the putative chimera with the parent sequences, on both sides of the breakpoint and 100% nucleotide identity with C549-W-17 within the intronic region, after the chimeric breakpoint.

A Bayesian tree constructed (GTR+G distances; exon only 127 characters) with the epimerase sequences of *Laminaria digitata*, *Saccharina japonica* and *Lessonia variegata*, with *Ectocarpus siliculosus* epimerase sequences as outgroups (Fig. 3.4), revealed two distinct groups of *L. variegata* exon sequences, as observed previously (Fig. 3.1). Assuming the topology is correct, the phylogenetic tree reveals at least one gene duplication event took place in *Lessonia variegata*, leading to the formation of two gene duplicates. The topology also shows one gene duplication event in *Laminaria digitata*. The base pair substitution frequency within and between the two *Laminaria digitata*

groups (Table 3.5), was found to be greater than that in the *Lessonia variegata* groups (Table 3.3).



0.01 substitutions/site

Fig. 3.4: Neighbour-Joining topology of Man-C5-E gene DNA sequences (127 characters, GTR distances) of select taxa in the Laminariales. *Ectocarpus siliculosus* chosen as outgroups. PP/ML-BP/NJ-BP values shown if \geq 50% BP/0.7 PP. BP- bootstrap analysis; PP - posterior probability; NJ- neighbour joining; ML - maximum likelihood. '-' indicates <50% support in a particular analysis.

Table 3.5: List of observed range of nucleotide changes within and between Man-C5-E DNA sequences of *Laminaria digitata* groups.

Frequency of nucleotide changes (bp)	L. digitata Group 1	L. digitata Group 2
<i>L. digitata</i> Group 1	10-12	
<i>L. digitata</i> Group 2	13-20	2-10

3.2 DISCUSSION

3.2.1 Gene duplication: multiple copies of Man-C5-E gene in Lessonia variegata

The high nucleotide similarity (Appendix 2) of the first 127 bp (exon) is an indication of the exon sequences being 'under selection' (Bergthorsson *et al.*, 2007). If the production of a certain protein is essential to the organism, the DNA sequences are subjected to 'selection' or functional restraints to minimize divergence. As alginate is an important cell wall polysaccharide, its production is very important for the alga. This makes it beneficial to select for mutations that contribute to alginate production. Thus the rate of divergence is less to maintain specific functions, as well as to minimize the rate of acquiring deleterious mutations.

The occurrence of gene duplication in *Lessonia variegata* was hypothesized based on the discovery of multiple Man-C5-E gene copies in *Laminaria digitata* (Nyvall *et al.*, 2003) and *Ectocarpus siliculosus* (Cock *et al.*, 2010).

In addition to the bootstrap support (>50%) in the neighbour-joining analysis (Fig. 3.1), the base pair variations between exon groups 1 and 2, being greater than that within the

groups, indicated the possible presence of at least two duplicates of the gene. This was further supported by the Bayesian analysis of the *Lessonia variegata* exons (Fig. 3.4), which supports one gene duplication event that occurred on the branch leading to the two groups, thus forming two gene paralogs in *Lessonia variegata*. The Bayesian analysis also supports at least one gene duplication event in *Laminaria digitata*. This brings up the question - when did the gene duplication event in *Lessonia variegata* take place? It is hard to ascertain that fact due to the lack of support for the branches (Fig. 3.4). However, there are two possible scenarios.

Since the NJ topology (Fig. 3.3) supports a common origin of the epimerase gene in the Laminariales, the first scenario is that the gene duplication could have occurred simultaneously with the diversification of Lessoniaceae from Laminariaceae. According to Silberfeld et al. (2010), the Laminariales started diversifying around 84.4 million years ago, during the upper Cretaceous era. The ontogenetic split between Laminariaceae and Lessoniaceae took place sometime in the late Paleogene (Cenozoic) era, which was approximately 20 million years ago (Silberfeld et al., 2010). Thus, if the gene duplication event occurred on the branch leading to these two families, then we would expect two clades, each clade containing one Man-C5-E gene copy from Laminaria digitata and one from Lessonia variegata. However, since the topology also indicates that the Lessonia *variegata* exon groups are closer to each other than to the *Laminaria digitata* groups, the more likely explanation, the second scenario, is that the gene duplication event occurred after the ontogenetic split of Lessoniaceae from Laminariaceae (Fig. 3.5). Our data would indicate an independent gene duplication event in Lessonia variegata, separate from that in *Laminaria digitata*. This theory provides an explanation for the high percentage of

nucleotide identity and low number of base pair variations (Table 3.3) between the *Lessonia variegata* sequences, compared to the *Laminaria digitata* sequences (Table 3.5), as a recent duplication would mean an insufficient amount of time for the gene to acquire sufficient point mutations to promote differences (Hurley, 2004; Zhang, 2003). Further sequencing and analysis are necessary to ascertain this theory.



Fig. 3.5: Diagram illustrating a possible gene duplication that may have occurred in *Lessonia variegata* after the ontogenetic split between Laminariaceae and Lessoniaceae.

3.2.2 Allelic variations in a diploid organism

The term 'allele' refers to one of the different forms of a gene that is present at a particular locus in the genome (Feero *et al.*, 2010; Lodish *et al.*, 2000). Diploid organisms, like *Lessonia variegata*, have one allele on each chromosome. If both alleles are the same, then the individual is referred to as homozygous, whereas, if the alleles are different, it is referred to as heterozygous (Lodish *et al.*, 2000). The exon sequences seen in the NJ tree

(Fig 3.1) were deduced to be alleles of the gene duplicates (Table 3.4). Samples with identical, or near identical sequences (Fig 3.2) were assumed to represent a single allele. However, it is not possible to ascertain homozygous and heterozygous individuals from the limited dataset. The sporophyte of *Lessonia variegata* being diploid, each sample of Lessonia variegata would exhibit, at most, only two alleles of the gene. Thus one gene duplication event would lead to the formation of four alleles in each sample per lineage. The presence of more than four alleles in the Wellington lineage sample C550-W and the Southern lineage sample C132-S may be an indicator still of the presence of PCR artifacts in the dataset. In spite of the removal of singletons and identical sequences from the dataset, it is possible that some variations in the sequences were the result of point mutations introduced by *Taq* polymerase during PCR which were amplified during subsequent PCR cycles and ultimately during cloning. Taq polymerase is known to have a non-proof reading ability (Zuccarello *et al.*, 2009) and an error rate of 10^{-5} to 10^{-4} error per nucleotide synthesized (Eckert and Kunkel, 1991; Zhou et al., 1991). The presence of only three alleles in the Northern lineage and the Wellington lineage sample C318-W can be attributed to insufficient amount of sequencing or that these individuals are homozygous at one of the two loci. Moreover, further sequencing and cloning of the three lineages may help in distinguishing true alleles from sequences which are simply PCR artifacts.

The topology of the NJ tree of exon sequences (Fig. 3.1) indicates that the Wellington lineage samples C549-W-17, C318-W-2, C550-W-12 and C550-W-27 may be very divergent alleles of the Man-C5-E gene, differing from the other Wellington lineage

samples in exon group 2, by 6-10bp (Fig. 3.2). If that hypothesis is correct, then it could indicate that the Wellington lineage allele diversified from the Northern and Southern lineages after the gene duplication occurred, thus forming the exon group 1 alleles (Fig. 3.1).The presence of the divergent Wellington lineage allele can also be explained by the fact that all three lineages are cryptic species of *Lessonia variegata* (Martin, 2011). However, further cloning and analysis of the Northern and Southern lineages is necessary to rule out the presence of a similar allele in these two lineages.

3.2.3 Introns and chimeric sequences in the dataset: possibilities

The results indicated the presence of an intron, at least 300bp long, in the DNA sequences of *Lessonia variegata*, even though primer set I (Table 2.3) was designed from the cDNAs of *Laminaria digitata* and *Saccharina japonica* (see Materials and Methods). A probable explanation is that the primers attached to the template DNA near an intron/exon boundary, as a result of which the intron was amplified during subsequent PCR cycles. However, due to the absence of the putative AG sequence signalling the end of an intron, it is likely that only a portion of a full length intron was amplified. It is probable that the primer attached to the edge of the intron and the putative end of the intron was cut off during sequence trimming of the obtained PCR product. Introns were also found in the Man-C5-E gene of *Laminaria digitata*. Nyvall *et al.* (2003) discovered six introns, ranging in size from 295-523bp in the reading frame of a genomic clone (EpiG), on comparison of the clone with the six cDNAs in *Laminaria digitata*. As the EpiG sequence data was not available in Genbank, it is not possible to determine if the location of the introns is the same as that observed in *Lessonia variegata*.

The NJ tree constructed with only the intronic sequences of the Lessonia variegata samples exhibited a different gene clustering pattern than that exhibited by the corresponding exon sequences. Introns, the non-coding regions of a gene, are presumed to have a higher rate of change than the corresponding protein-coding exon regions, as they are subjected to fewer functional constraints (Li, 1997; Villablanca et al., 2002). This is exhibited in the dataset by the high number of variable sites within the intronic sequences (Appendix 3). Intronic recombination events also provide a possible explanation for the intronic gene clusters. It has been proposed that in eukaryotes, a high rate of genetic recombination occurs within the eukaryotic introns or 'spliceosomal introns', which promotes events like 'exon shuffling' to increase the protein diversity within an organism (Roy and Gilbert, 2006; Kolkman and Stemmer, 2001). Intronic recombination can also occur via transposons or 'jumping genes' (McClintock, 1950). However, that is an improbable event in this case, as the search for transposons present in the intronic sequence was found to be negative. An event like 'jumping PCR' (Pääbo et al., 1989) generally occurs when the DNA template is somewhat degraded. If the degraded DNA sample is amplified, the lengths of some of the degraded sequences may not be sufficient to span the entire distance between the two primers. During such an event, the first primer is only extended until the end of that short fragment in the first extension cycle by *Taq* polymerase. In the subsequent cycle, that primer will 'jump' to another complimentary DNA fragment and the extension will continue until the original DNA sequence has been reformed. Eventually, a double stranded DNA will be formed, which will serve as a template for subsequent rounds of amplification, thus forming a

recombinant DNA sequence (Howe, 1995; Pääbo *et al.*, 1989). This is the leading cause of formation of chimeric sequences during PCR amplification. It is possible that some of the sequences contained a recombinant intron. The formation of some recombinant introns in the dataset may provide an explanation for different intronic gene clusters, but it is unlikely that this event occurred in all 20 sequences. Although such theories may elude to possibilities explaining the formation of intronic gene clusters that are different from the corresponding exon clusters, further research is necessary to ascertain this phenomenon. Thus this remains a curiously unresolved part of our analysis.

It is not clear whether there is an actual chimeric sequence in the dataset, or if the analysis indicated the presence of a chimera due to a higher percentage of variations in the intron region (52.8% nucleotide identity), compared to the exon region (89% nucleotide identity). The high percentage of nucleotide identity between the putative chimera and the parent sequences in the exon region, that is, before the chimeric breakpoint, can be explained by the exon region being under selection, and thus highly conserved in all sequences of *Lessonia variegata*, as discussed previously.

3.2.4 Mannuronan C5 Epimerase gene duplicates may control alginate structure

In *Lessonia variegata*, the gene duplicates may encode Man-C5-E enzymes that control the content of the M and G blocks of the alginate in different parts of the plant, thus giving some parts more flexibility than others. This was hypothesized by Nyvall *et al.* (2003) due to the discovery of 16 different gene copies in sporophytes of *Laminaria digitata*. They hypothesized that in addition to epimerization of the D-mannuronic residues into L-guluronic units, the different epimerases may also be involved in the

tailoring of the relative contents and distributions of G blocks, M blocks and MG blocks in alginate chains, depending upon the season, age of plant and tissue type. Venegas et al. (1993) conducted a study on Lessonia trabeculata Villouta & Santalices and found that the blades of plants collected from exposed areas, produced alginates with a significantly lower M/G ratio and had a higher percentage of G blocks than those from protected sites. The blades of sporophytes collected from the exposed sites were also significantly narrower and thicker. They also found that when the blades from exposed sites were transplanted to a protected bay, there was a significant decrease of M/G ratio and increase of MM blocks, while blades transplanted from the protected bay to the exposed beach showed a significant increase of M/G ratio and increase in content of GG blocks. This showed that algae acclimatized to changes in the hydrodynamic activity of the site through modifications in alginate biosynthesis. It is possible that in *Lessonia variegata*, the multiple gene copies play a similar function in modifying the chemical structure of alginate in response to environmental conditions or to different requirements in different parts of the thallus. It was also shown that the expression of Man-C5-E genes in *Laminaria digitata* may be stress and defence induced (Roeder *et al.*, 2005).

Alginates with different proportions of M and G blocks have been found to have different biotechnological applications. For example, alginate with high M block content (over 50%) has been found to have a wide range of immunological applications. These include applications, such as, cell transplantations, due to the resulting alginate gels being less viscous than alginates with more guluronic acid content. This makes the gels more stable, have less diffusion permeability and thus be more efficient in protecting immobilized cells against the host immune system (Klöck *et al.*, 1997). Since the exon sequences are highly conserved with very few amino acid substitutions, the paralogs may encode for enzymes that may synthesize or modify the alginates in a different way. The presence of the intron indicates that the Man-C5-E gene copies are modular, and thus possibly subject to alternative splicing (Black, 2003). Alternative splicing has been found to be responsible for the formation of multiple protein isoforms (Graveley, 2001). Su *et al.* (2006) suggest that alternative splicing and gene duplication may exhibit co-evolution and that in the early stages after gene duplication, the newly evolved duplicates may take over some of the protein functional diversity which, prior to gene duplication, was carried out by alternative splicing of the ancestral gene.

The identification and linking of such mechanisms to alginate structure and composition would be of particular importance in the biotechnological industry.

3.2.5 RNA extraction

RNA extraction from the blades of *Lessonia variegata* after several attempts with the TRIzol[®] method was unsuccessful. A possible explanation for the failure of the TRIzol[®] method may be due to the presence of high amounts of polysaccharides in the algal cell walls which can bind or co-precipitate with the RNA, thus causing very low or no yield of RNA (Falcão *et al.*, 2008; Salzman *et al.*, 1999). Alternatively, RNA degradation could have occurred during the extraction procedure, despite taking the requisite precautions. Some studies have developed alternate methods of extracting RNA, such as a modification of the TRIzol[®] method (Falcão *et al.*, 2008) or using soluble polyvinylpyrrolidone (PVP) and ethanol (ETOH) precipitation (Salzman *et al.*, 1999), which should be explored in future studies.

3.2.6 Conclusion

This is the first study on the genetic aspect of the biochemical pathway of alginate in the endemic New Zealand species *Lessonia variegata*. With the exception of a few studies on the genetics of alginate production in the Laminariales (Nyvall *et al.*, 2003; Inoue *et al.*, 2007), most of the studies on alginate have focused on the enzymes involved in the pathway and the chemical aspects of alginate. However, with the increasing importance of alginate in various industries, such as the food, cosmetics and medicine, it is of both an academic and commercial importance to characterize and gather more information about the genes involved in the production of alginate, which would help in the regulation of alginate production.

At least one gene duplication event was found to have occurred in *Lessonia variegata*, leading to the formation of two Man-C5-E genes. Putative alleles were discovered in the Northern, Southern and Wellington lineages. Further analyses are required to separate the true alleles from PCR artifacts, as well as to ascertain the presence of a diverging allele in the Wellington lineage. It would be of great interest to see if there are more alleles of the gene in *Lessonia variegata*, as that could be an indication of more than one gene duplication event having taken place in the brown alga. Future studies are also required to determine the number of gene paralogs in the Kaikoura lineage. It is also of academic and commercial interest to determine whether all the gene paralogs are expressed and how the expression level differs seasonally and between different tissues. While this study raises the possibility of the occurrence of alternative splicing, further studies should be done to determine the existence of the actual phenomenon in alginate production and the genetic mechanisms behind the production of alginates with different block structures.

4.0 References

Abbott, L.J. (2011). Polysaccharide content and growth rate of *Lessonia variegata* J. Agardh: investigating its potential as a commercial species. Victoria University of Wellington, MSc thesis.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions* on Automatic Control, 19: 716-723.

Anderson, I., Skipnes, O., Smidsrod, O., Ostgaard, K., & Hemmer, P. (1977). Some biological functions of matrix components in benthic algae in relation to their chemistry and the composition of seawater. *ACS Symposium Series*, 48: 361-381.

Athukorala, Y., Lee, K.W., Kim, S.K., & Jeon, Y.J. (2007). Anticoagulant activity of marine green and brown algae and collected from Jeju island in Korea. *Bioresource Technology*, 98: 1711-1716.

Awad, E., & Osman, O. (2003). Laminarin enhanced immunological disorders of septicimeric albino rats infected with *Aeromonas hydrophila*. *Egyptian Journal of Immunology*, 10: 49-56.

Bandelt, H-J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16: 37-48.

Bergthorsson, U., Andersson, D.I., & Roth, J.R. (2007). Ohno's dilemma: evolution of new genes under continuous selection. *Proceedings of the National Academy of Sciences*, 104: 17004-17009.

Berteau, O., & Mulloy, B. (2003). Sulfated fucans, fresh perspectives: structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. *Glycobiology*, 13: 29-40.

Bisgrove, S.R., & Kropf, D.L. (2001).Cell wall deposition during morphogenesis in fucoid algae. *Planta*, 212: 648-658.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Reviews of Biochemistry*, 72: 1-64.

Bold, H., & Wynne, M. (1985). *Introduction to the Algae*. Englewood Cliffs: Prentice-Hall Inc.

Campos, M., Martinez-Salazar, J., Lloret, L., Moreno, S., Nunez, C., Espin, G., *et al.* (1996). Characterization of the gene coding for GDP-Mannose Dehydrogenase (algD) from *Azotobacter vinelandii*. *Journal of Bacteriology*, 178: 1793-1799. Chen, P., Min, M., Chen, Y., Wang, L., Li, Y., Chen, Q., Wang, C., Wan, Y., Wang, X., Chen, Y., Deng, S., Hennessy, K., Ling, X., Liu, Y., Wang, Y., Martinez, B., & Ruan, R. (2009). Review of the biological and engineering aspects of algae to fuels approach. *International Journal of Agricultural and Biological Engineering*, 2: 1-30.

Chin, N.K., Brown, M.T., & Heads, M.J. (1991). The biogeography of Lessoniaceae, with special reference to *Macrocystis* C. Agardh (Phaeophyta: Laminariales). *Hydrobiologia*, 215: 1-11.

Cock, J.M., Sterck, L., Rouze, P., Scornet, D., Allen, A.E., Amoutzias, G., Anthouard, V., Artiguenave, F., Aury, J.M., Badger, J.H., Beszteri, B., Billiau, K., Bonnet, E., Bothwell, J.H., Bowler, C., Boyen, C., Brownlee, C., Carrano, C.J., Charrier, B., Cho, G.Y., Coelho, S.M., Collen, J., Corre, E., Da Silva, C., Delage, L., Delaroque, N., Dittami, S.M., Doulbeau, S., Elias, M., Farnham, G., *et al.* (2010). The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature*, 465: 617-621.

Cragie, J., Morris, E., Rees, D., & Thom, D. (1984). Alginate block structure in Phaeophyceae from Nova Scotia: variation with species, environment and tissue-type. *Carbohydrate Polymers*, 4: 237-252.

Daugbjerg, N., & Anderson, R.A. (1997). A molecular phylogeny of the heterokont algae based on analyses of chloroplast-encoded *rbcL* sequence data. *Journal of Phycology*, 33: 1031-1041.

Davis, T., Volesky, B., & Mucci, A. (2003). A review of the biochemistry of heavy metal biosorption by brown algae. *Water Research*, 37: 4311-4330.

Dayton, P.K. (1985). Ecology of Kelp Communities. *Annual review of ecology and systematic*, 16: 215-245.

De Reviers, B., Rousseau, F., & Draisma, S. (2007). Classification of the Phaeophyceae from past to present and current challenges. In J. Brodie, & J. Lewis (Eds.), *Unravelling the Algae - the past, present and future of algal systematics*. London: CRC Press: Taylor and Francis Group.

Draget, K.I., Smidsrod, O., & Skjak-Braek, G. (2005). Alginates from algae. In A. Steinbüchel, & S.K. Rhee (Eds.), *Polysaccharides and Polyamides in the Food Industry: Properties, Production and Patents*. Trondheim: Wiley-VCH Verlag Gmbh & Co.

Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T., & Wilson, A. (2012). *Geneious v5.6*, available from http://www.geneious.com

Eckert, K.A., & Kunkel, T.A. (1991). DNA polymerase fidelity and the polymerase chain reaction. *Genome Research*, 1: 17-24.

Ertesvag, H., & Valla, S. (1998). Biosynthesis and Applications of Alginates. *Polymer Degradation and Stability*, 59: 85-91.

Falcão, V.D.R., Tonon, A.P., Oliveira, M.C., & Colepicolo, P. (2008). RNA isolation method for polysaccharide rich algae: agar producing *Gracilaria tenuistipitata* (Rhodophyta). *Journal of Applied Phycology*, 20: 9-12.

Feero, W.G., Guttmacher, A.E., & Collins, F.S. (2010). Genomic Medicine - an updated primer. *The New England Journal of Medicine*, 362: 2001-2011.

Fenoradosoa, T.A., Delattre, G.A.C., Laroche, C., Petit, E., Wadouachi, A., & Michaud, P. (2010). Extraction and characterization of an alginate from the brown seaweed *Sargassum turbinarioides* Grunow. *Journal of Applied Phycology*, 22: 131-137.

Fluxus Technology Ltd. (2011). Network 4.6.1.0. Available at www.fluxusengineering.com

Fox, C.H., & Swanson, A.K. (2007). Nested PCR detection of microscopic life stages of laminarian macroalgae and comparison with adult forms along intertidal height gradients. *Marine Ecology Progress Series*, 332: 1-10.

Graham, L.E., & Wilcox, L.W. (2000). Algae. USA: Prentice-Hall.

Graveley, B.R. (2001). Alternative splicing: increasing diversity in the proteomic world. *Trends in Genetics*, 17: 100-107.

Gómez-Díaz, D., & Navaza, J.M. (2003). Rheology of aqueous solutions of food additives - effect of concentration, temperature and blending. *Journal of Food Engineering*, 56: 387-392.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, 166: 557-580.

Hatipoglu, F., Ogurtan, Z., Sezer, A., Uney, K., Erol, M., Ozdemir, O., Bas, A.L., & Akbuga, J. (2008). Effect of laminarin and chitosan gel formulations on the treatment of hydrofluoric acid induced corneal burns in the rabbits. *Reveu de Medecine Veterinaire*, 159: 207-214.

Hawes, I., Nelson, W., & Mercer, S. (2004). Hang on to your haptera:studies in a kelp forest. *Water & Atmosphere*, 12: 26-27.

Hawkins, W.W., & O'Neill, A.N. (1955). The anticoagulant action in blood of sulphated derivatives of laminarin. *Canadian Journal of Biochemistry and Physiology*, 33: 545-552.

Hayashi, K., Nakano, T., Hashimoto, M., Kanekiyo, K., & Hayashi, T. (2008). Defensive effects of a fucoidan from brown alga Undaria pinnatifida against herpes simplex virus infection. *International Immunopharmacology*, 8: 109-116.

Howe, C.J. (1995). *Gene cloning and manipulation*. Cambridge, UK: Cambridge University Press.

Huber, T., Faulkner, G., & Hugenholtz, P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics*, 20: 2317-2319.

Huelsenbeck, J.P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17: 754-755.

Hurles, M. (2004). Gene duplication: the genomic trade in spare parts. *Public library of Science Biology*, 2: 900-904.

Inoue, A., Sato, A., & Ojima, T. (2007). cDNA cloning of Mannuronan-C5-Epimerase from *Laminaria japonica*. Unpublished, downloaded on 15th May 2012 from http://www.ncbi.nlm.nih.gov/nuccore/AB299380

Jukes, T.H., & Cantor, C.R. (1969). Evolution of protein molecules. In H.N. Munro (Ed.), *Mammalian protein metabolism*. New York: Academic Press.

Kai, T., Nimura, K., Yasui, H., & Mizuta, H. (2006). Regulation of sorus formation by auxin in Laminariales sporophyte. *Journal of Applied Phycology*, 18: 95-101.

Kelecom, A. (2002). Secondary metabolites from marine microorganisms. *Annals of the Brazilian Academy of Sciences*, 74: 150-170.

Kim, K., Kim, Y., Kim, H., Lee, B., & Dong, S. (2006). Anti-apoptotic activity of laminarin polysaccharides and their enzymatically hydrolyzed oligosaccharides from Laminaria japonica. *Biotechnology Letters*, 28: 439-446.

Kloareg, B., & Quatrano, R.S. (1988). Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanography and Marine Biology: An Annual Review*, 26: 259–315.

Klöck, G., Pfeffermann, A., Ryser, C., Gröhn, P., Kuttler, B., Hahn, H.J., & Zimmermann, U. (1997). Biocompatibility of mannuronic acid-rich alginates. *Biomaterials*, 18: 707-713.

Kolkman, J.A., & Stemmer, P.C. (2001). Directed evolution of proteins by exon shuffling. *Nature Biotechnology*, 19: 423-428.

Kühtreiber, W.M., Lanz, R.P., & Chick, W.L. (1999). *Cell Encapsulation Technology and Therapeutics*. Boston: Birkhäuser.

Lane, C., Mayes, C., Druehl, L., & Saunders, G. (2006). A multi-gene molecular investigation of the kelp (Laminariales, Phaeophyceae) supports substantial taxonomic re-organization. *Journal of Phycology*, 42: 493-512.

Li, W.H. (1997). Molecular Evolution. Sunderland: Sinauer Associates Inc.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452.

Lobban, C.S., & Harrison, P.J. (1994). *Seaweed Ecology and Physiology*. Cambridge University Press.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). *Molecular Cell Biology* (4th edition). New York: W.H. Freeman.

Lüning, K. (1990). *Seaweeds: Their Environment, Biogeography and Ecophysiology*. New York: John Wiley & Sons Inc.

Mabeau, S., Kloareg, B., & Joseleau, J.P. (1990) Fractionation and analysis of fucans from brown algae. *Phytochemistry*, 29: 2441- 2445.

Martin, P. (2011). Phylogeny, phylogeography and population connectivity of *Lessonia* (Phaeophyceae). Victoria University of Wellington, PhD thesis.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United Stated of America*, 36: 344-355.

Mount, S.M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Research*, 10: 459-472.

Nelson, W.A. (2005). Life history and growth in culture of the endemic New Zealand kelp. *Journal of Applied Phycology*, 17: 23-28.

Nelson, W., & Schwarz, A.M. (2005). Living in the Ocean waves: kelp life histories. *Water and Atmosphere*, 13: 22-23.

Nybakken, J., & Bertness, M. (2005). *Marine Biology: An Ecological Approach* (6th edition). San Francisco: Pearson Education, Inc., publishing as Benjamin Cummings.

Nyvall, P., Corre, E., Boisset, C., Barbeyron, T., Rousvoal, S., Scornet, D., Kloareg, B., & Boyen, C. (2003). Characterization of Mannuronan C-5-Epimerase Genes from the Brown Alga Laminaria digitata. *Plant Physiology*, 133: 726-735.

Owen, J.G. (2010). Characterization, manipulation and directed evolution of nonribosomal peptide synthetase enzymes. Victoria University of Wellington, PhD thesis.

Pääbo, S., Higuchi, R.G., & Wilson, A.C. (1989). Ancient DNA and the polymerase chain reaction. *The Journal of Biological Chemistry*, 264: 9709-9712.

Phillips, N., Burrowes, R., Rousseau, F., de Reviers, B., & Saunders, G. (2008a). Resolving evolutionary relationships among the brown algae using chloroplast and nuclear genes. *Journal of Phycology*, 44: 394-405.

Phillips, N., Calhoun, S., Moustafa, A., Bhattacharya, D., & Braun, E. (2008b). Genomic insights into evolutionary relationships among heterokont lineages emphasizing the pheophyceae. *Journal of Phycology*, 44: 15-18.

Podkorytova, A.V., Vafina, L.H., Kovaleva, E.A., & Mikhailov, V.I. (2007). Production of algal gels from the brown alga *Laminaria japonica* Aresch., and their biotechnological applications. *Journal of Applied Phycology*, 19: 827-830.

Posada, D., & Crandall, K.A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics*, 14: 817-818.

Potin, P., Bouarab, K., Salaün, J.P., Pohnert, G., & Kloareg, B. (2002). Biotic interactions of marine algae. *Current Opinion in Plant Biology*, 5: 308-317.

Prescott, G. (1969). The Algae: A Review. London: Thomas Nelson and Sons Ltd.

Rombaut, A. (1996). *Se-Al, Sequence Alignment Editor (version 2.0).* Department of Zoology, University of Oxford, Oxford. Available at: http://evolve.zoo.ox.zc.uk

Rembaut, A., & Drummond, A. (2007). *Tracer v1.4*. Available at http://beast.bio.ed.ac.uk/Tracer

Rioux, L., Turgeon, S., & Beaulieu, M. (2010). Structural characterization of laminaran and galactofucan extracted from the brown seaweed *Saccharina longicruris*. *Phytochemistry*, 71: 1586-1595.

Rioux, L., Turgeon, S., & Beaulieu, M. (2009). Effect of season on the composition of bioactive polysaccharides from the brown seaweed *Saccharina longicruris*. *Phytochemistry*, 70: 1069-1075.

Rioux, L., Turgeon, S., & Beaulieu, M. (2007). Characterization of polysaccharides extracted from brown seaweeds. *Carbohydrate Polymers*, 69: 530-537.

Rødde, R.S.H., & Larsen, B. (1997). Protoplasts of *Laminaria digitata and Laminaria saccharina* (Phaeophyta) – Cultivation and biosynthesis of alginate. *Botanica Marina*, 40: 391-395.

Roeder, V., Collén, J., Rousvoal, S., Corre, E., Leblanc, C., & Boyen, C. (2005). Identification of stress gene transcripts in Laminaria digitata (Phaeophyceae) protoplast cultures by expressed sequence tag analysis. *Journal of Phycology*, 41: 1227-1235. Rousvoal, S., Groisillier, A., Dittami, S., Michel, G., Boyen, C., & Tonon, T. (2011). Mannitol-1-phosphate dehydrogenase activity in *Ectocarpus siliculosus*, a key role for mannitol synthesis in brown algae. *Planta*, 233: 261-273.

Roy, S.W., & Gilbert, W. (2006). The evolution of spliceosomal introns: patterns, puzzled and progress. *Nature Reviews Genetics*, 7: 211-221.

Rupérez, P., Ahrazem, O., & Leal, J.A. (2002). Potential antioxidant capacity of sulfated polysaccharides from the edible marine brown seaweed *Fucus vesiculosus*. *Journal of Agricultural and Food Chemistry*, 50: 840-845.

Salzman, R.A., Fujita, T., Zhu-Salzman, K., Hasegawa, P.M., & Bressan, R.A. (1999). An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. *Plant Molecular Biology Reporter*, 17: 11-17.

Schwartz, A.M., Hawes, I., Nelson, W., & Andrew, N. (2006). Growth and reproductive phenology of the kelp *Lessonia variegata* in central New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 40: 273-284.

Silberfeld, T., Leigh, J.W., Heroen, V., Cruaud, C., De Reviers, B., & Rousseau, F. (2010). A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the "brown algal crown radiation". *Molecular Phylogenetics and Evolution*, 56: 659-674.

Skriptsova, A., Khomenko, V., & Isakov, V. (2004). Seasonal changes in the growth rate, morphology and alginate content in *Undaria pinnatifida* at the northern limit in the Sea of Japan (Russia). *Journal of Applied Phycology*, 16: 17-21.

South, G.R., & Whittick, A. (1987). *Introduction to Phycology*. Oxford UK: Blackwell Scientific Publications.

Steneck, R.S., Graham, M.H., Bourque, B.J., Corbett, D., Erlandson, J.M., Estes, J.A., & Tegner, M.J. (2002). Kelp forest ecosystems: biodiversity, stability, resilience and future. *Environmental Conservation*, 29: 436–459.

Stockton, B., Evans, L.V., Morris, E.R., Powell, D.A., & Rees, D.A. (1980). Alginate block structure in *Laminaria digitata* – Implications for holdfast attachment. *Botanica Marina*, 23: 563.

Swofford, D.L. (2002). *PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sunderland, MA: Sinauer Associates.

Su, Z., Wang, J., Yu, J., Huang, X., & Gu, X. (2006). Evolution of alternative splicing after gene duplication. *Genome Research*, 16: 182-189.

Tanriseven, A., & Doğan, S. (2001). Immobilization of invertase within calcium alginate gel capsules. *Process Biochemistry*, 36: 1081-1083.

Tenhaken, R., Voglas, E., Cock, J.M., Neu, V., & Huber, C.G. (2011). Characterization of GDP-mannose dehydrogenase from the brown alga *Ectocarpus siliculosus* providing the precursor for the alginate polymer. *The Journal of Biological Chemistry*, 286: 16707-16715.

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In R.M. Miura (Ed.), *Lectures on Mathematics in the Life Sciences*. Washington, DC: American Mathematical Society.

Tsai, D.D-W., Ramaraj, R., & Chen, P.H. (2012). Growth condition study of algae function in ecosystem for CO₂ bio-fixation. *Journal of Photochemistry and Photobiology*, 107: 27-34.

Van den Hoek, C., Mann, C., & Jahns, H.M. (1995). *Algae: An Introduction to Phycology*. Cambridge University Press, Cambridge.

Vásquez, J.A. (2007). Production, use and fate of Chilean brown seaweeds: re-sources for a sustainable fishery. *Journal of Applied Phycology*, 20: 457-467.

Venegas, M., Matsuhiro, B., & Edding, M.E. (1993). Alginate composition of *Lessonia trabeculata* (Phaeophyta: Laminariales) growing in exposed and sheltered habitats. *Botanica Marina*, 36: 47-51.

Villablanca, F.X., Roderick, G.K., & Palumbi, S.R. (2002). Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology*, 7: 547-560.

Wang, G.C.Y., & Wang, Y. (1997). Frequency of formation of chimeric molecules is a consequence of PCR coamplification of 16s rRNA genes from mixed bacterial genomes. *Applied and Environmental Microbiology*, 63: 4645-4650.

Yoon, H.S., Lee, J.Y., Boo, S.M., & Bhattacharya, D. (2001). Phylogeny of Alariaceae, Laminariaceae and Lessoniaceae (Phaeophyceae) based on plastid-encoded RuBisCo spacer and nuclear-encoded ITS sequence comparisons. *Molecular Phylogenetics and Evolution*, 21: 231-243.

Zhang, J. (2003). Evolution by gene duplication: an update. *Trends in Ecology and Evolution*, 18: 292-298.

Zhou, Y., Zhang, X., & Ebright, R.H. (1991). Random mutagenesis of gene-sized DNA molecules by use of PCR with *Taq* DNA polymerase. *Nucleic Acid Research*, 19: 6052.

Zuccarello, G.C., & Lokhorst, G.M. (2005). Molecular phylogeny of the genus *Tribonema* (Xanthophyceae) using rbcL gene sequence data: monophyly of morphologically simple algal species. *Phycologia*, 44: 384–392.

Zuccarello, G.C., Oellermann, M., West, J.A., & Clerck, O.D. (2009). Complex patterns of actin molecular evolution in the red alga *Stylonema alsidii* (Stylonematophyceae, Rhodophyta). *Phycological Research*, 57: 59-65.

Zvyagintseva, T., Shevchenko, N., Chizhov, A., Krupnova, T., Sundukova, E., & Isakov, V. (2003). Water-soluble polysaccharides of some far-eastern brown seaweeds. Distribution, structure, and their dependence on the developmental conditions. *Journal of Experimental Marine Biology and Ecology*, 294: 1–13.

5.0 Appendices

Appendix 1

Neighbour-Joining topology of all cloned Man-C5-E DNA sequences of the Northern, Southern and Wellington lineages of *Lessonia variegata*, obtained from primer set II; the lineages are colour-coded. B70-N: Northern lineage (Green); C132-S: Southern lineage (Red); C318-W, C549-W, C550-W: Wellington lineage (Blue).


Appendix 2

Alignment of the exon sequences (127 characters) of the Lessonia variegata lineages. B70-N: Northern lineage; C132-S: Southern lineage; C318-W, C549-W, C550-W: Wellington lineage.

-		-					
	1	10	20	30	40	50	60
D70 N 10							
G132_S_1	AACTC	GGGCTACGGC	TTTGACCCCC	CACGACGACAG	CGACIICIIC	GACCATCCACGA	CAAC
C132-S-1	AACTO	GGGTTACGGC	TTTGATCCC	CACGACGACAG	CGACTICIIC	FACCACCACGA	CAAC
C132-S-5	AACTO	GGGTTACGGC	TTTGATCCC	CACGATGACAG	CGACIICIIC	GACCATCCACGA	CAAC
$C_{132} = 3 = 7$	AACTO	GGGIIACGGC	TTTGATCCC	CACGACGACAG	CGACIICIIC	FACCACCACGA	CAAC
C550-W-2	AACTO	GGGCTACGGC	TTTCATCCC	CACGACGACAG	CGACIICIIC	SACCATCCACGA	CAAC
C550-W-6	AACTO	GGGTTACGGC	TTCGATCCC	CACGACGACAG	CGACTTCTT	FACCATCCACGA	CAAC
C549-W-2	AACTO	GGGCTACGGC	TTTGATCCC	CACGACGACAG	CGACTTCTT	FACCATCCACGA	CAAC
B70-N-1	AACTO	GGGCTACGGC'	TTTGACCCC	CACGACGACAG	TGATTTCCT	FACCATCCACGA	CAAC
B70-N-2	AACTO	GGGCTACGGC'	TTTGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	CAAC
C132-S-2	AACTO	GGGCTACGGC'	TTTGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	CAAC
C132-S-22	AACTO	GGGCTACGGC	TTTGATCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C318-W-1	AACTO	GGGCTACGGC	TTTGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C549-W-4	AACTO	GGGCTACGGC	TTTGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C550-W-5	AACTO	GGGCTACGGC	TTTGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C549-W-10	AACTO	GGGCTACGGC	TTTGACCCC	CACGATGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C318-W-2	AACTO	GGGTTACGGC	TTCGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C549-W-17	AACTO	GGGTTACGGC	TTCGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C550-W-12	AACTO	GGGTTACGGC	TTCGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
C550-W-27	AACTC	GGGTTACGGC	TTCGACCCC	CACGACGACAG	TGATTTCCT	GACCATCCACGA	ACAAC
B70-N-12	GAGGI	CTACAACAAC	GGATACCAC	GGTATCATCGC	CTCCAAGCG	CTGCAACGGAGI	IGTCC
C132-S-1	GAGGI	CTACAACAAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C132-S-5	GAGGI	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C132-S-7	GAGGI	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C550-W-1	GAGGT	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C550-W-2	GAGGI	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C550-W-6	GAGGI	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C549-W-2	GAGGI	CTACAACGAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
B70-N-1	GAGGI	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
B70-N-2	GAGGI	CTACAACAAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C132-S-2	GAGGI	CTACAACAAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C132-S-22	GAGGI	CTACAACAAC	GGATACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C318-W-1	GAGGI	CTACAACAAT	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C549-W-4	GAGGI	CTACAACAAT	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C550-W-5	GAGGI	CTACAACAAT	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C549-W-10	GAGGI	CTACAACAAT	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
C318-W-2	GAGGI	CTACAACAAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C549-W-17	GAGGI	CTACAACGAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	GTCC
C550-W-12	GAGGI	CTACAACAAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAATGGAGI	'GTCC
C550-W-27	GAGGI	CTACAACAAC	GGGTACCAC	GGCATCATCGC	CTCCAAGCG	CTGCAACGGAGI	GTCT
B70-N-12	ATCCA	GG					
C132-S-1	ATCCA	GG					
C132-S-5	ATCCA	GG					
C132-S-7	ATCCA	GG					
C550-W-1	ATCCA	GG					
C550-W-2	ATCCA	.GG					
C550-W-6	ATCCA	GG					
C549-W-2	ATCCA	GG					
B70-N-1	ATCCA	GG					
B70-N-2	ATCCA	GG					
C132-8-23	ATCCA	GG					
$C_{132} = 3 = 22$	ATCCA	GG					
C549 - W - 4	ATCCA	GG					
C550-W-5	ATCCA	GG					
C549 - W - 10	ATCCA	GG					
C318-W-2	ATCCA	GG					
C549-W-17	ATCCA	GG					
C550-W-12	ATCCA	GG					
C550-W-27	ATCCA	GG					

Appendix 3

Alignment of the intron sequences of the *Lessonia variegata* lineages. B70-N: Northern lineage; C132-S: Southern lineage; C318-W, C549-W, C550-W: Wellington lineage.

	1	10	20	30	40	50	60
			1	1	I.		
B70-N-1	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTACCGCA	AGCAGTACC	GCATTCTGGGTAG	TCT
B70-N-2	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	AGCAGTACC	GCATTCTGGGTAG'	TCT
B70-N-12	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCGTTCTGGGTAG'	TCT
C132-S-1	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C132-S-2	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	GCAGTACC	GCATTCTGGGTAG	TCT
C132-S-5	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C132-S-7	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C132-S-22	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	GCAGTACC	GCATTCTGGGTAG	TCT
C318-W-1	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C318-W-2	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C549-W-2	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCG	GCAGTACC	GCATTCTGGGTAG'	TCT
C549-W-4	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	AGCAGTACC	GCATTCTGGATAG'	TCT
C549-W-10	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	AGCAGCACC	GCATTCTGGATAG'	TCT
C549-W-17	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C550-W-1	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGCACC	GCATTCTGGGTAG	ТСТ
C550-W-2	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C550-W-5	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C550-W-6	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C550-W-12	GTACO	GATATGTTCGGC	ACGCTTCG	CTTTGTACCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
C550-W-27	GTACO	GATGTGTTCGGC	ACGCTTCG	CTTTGTGCCGCA	GCAGTACC	GCATTCTGGGTAG'	TCT
B70-N-1	TGGTA	AGTAAGCCAAGG	GCATTTCG	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
B70-N-2	TCGTA	AGCAAGCCAAGA	GCATTTCG	CGGGTCGTCTTC	CTCAACTGG.	AGAAGTTTGTGTT	TTG
B70-N-12	TGGTA	AGTAAGCCAAGG	GCATTTCG	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C132-S-1	TCGTA	AGTAAGCCAAGA	GCATTTCG	CGGGTCGTCTTC	CTCAACTGG	AGAAATTTGTGTT	TTG
C132-S-2	TCGTA	AGTAAGCCAAGA	GCATTTCG	CGGGTCGTTTTC	CTCAACTGG	AGAAATTTGTGTT'	TTG
C132-S-5	TGGTA	AGTAAGCCAAGG	GCATTTCG	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTATT	TTG
C132-S-7	TGGTA	AGTAAGCCAAGG	GCATTTCG	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTATT	TTG
C132-S-22	TCGTA	AGTAAGCCAAGA	GCATTTCG	CGGGTCGTCTTC	CTCAGCTGG	AGAAATTTATT	
C318-W-1	TCGTA	AGTAAGCCAAGA	GCATTTCG	CG-GTCGTCTTC	CTCAACTGG.	AGAAGTTTGTGTT	TTG
C318-W-2	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C549-W-2	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C549-W-4	TCGTA	AGTAAGCCAAGA	GCATTTCG	CG-GTCGTCTTC	CTCAACTGG.	AGAAGTTTGTGTT	TTG
C549-W-10	TCGTA	AGTAAGCCAAGA	GCATTTCG	CG-GTCGTCTTC	CTCAACTGG.	AGAAGTTTGTGTT	TTG
C549-W-17	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C550-W-1	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C550-W-2	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C550-W-5	TCGTA	AGTAAGCCAAGA	GCATTTCG	CG-GTCGTCTTC	CTCAACTGG.	AGAAGTTTGTGTT	TTG
C550-W-6	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C550-W-12	TGGTA	AGTAAGCCAAGG	GCATTTTT	CGGGTCGTCTTC	CTCAGCCGG	AGAAGCTTGTT	TTG
C550-W-27	TCGTA	AGTAAGCCAAGA	GCATTTCG	CG-GTCGTCTTC	CTCAACTGG.	AGAAGTTTGTGTT	TTG
B70-N-1	GAAGI	TT-ATTCTCGAT	TGATTCCT	CAAAGAGATGAC	TATGAAAT	AACCTATTC	TCG
B70-N-2	GAAGT	T-ATTCTCGAT	TGACTCGT	CAAAGAGATGAC	CTATGAAAT	AACCTGGCAACTC	TCG
B70-N-12	GAAGT	T-ATTCTCGAT	TGATTCCT	CAAAGAGATGAC	CTATGAAAT	AACCTATTC	TCG
C132-S-1	GAAGT	T-ATTCTCGAT	TGACTCCT	CAAAGAGATGAC	CTATGAAAT	AACCTGGCAACTC'	TCG
C132-S-2	GAAGT	T-ATTCTCGAT	TGACTCCT	CAAAGAGATGA	CTATGAAAT	AACCTGGCAACTC'	TCG
C132-S-5	GAAGT	T-ATTCTCGAT	TGATTCCT	CAAAGAGATGAC	TATGAAAT	AACCTATTC	TCG
C132-S-7	GAAGI	T-ATTCTCGAT	TGATTCCT	CAAAGAGATGAC	CTATGAAAT	AACCTATTC	TCG
C132-S-22							

C318-W-1	GAAGTT-ACTCTCGATTGACTCCTCAAAGAGATGACTATGAAATAACCTATTCTCG
C318-W-2	GAAGTT-ATTCTCGATTGATTCCTCAAAGAGATGACTATGAAATAACCTATTCTCG
C549-W-2	GAAGTT-ATTCTCGATTGATTCCTCAAAGAGATGACTATGAAATAACCTATTCTCG
C.549 - W - 4	GAAGTT-ACTCTCGATTGACTCCTCAACGAGATGACTATGAAATAACCTAGCAACTCTCG
C549 - W - 10	GAAGTT-ACTCTCGATTGACTCCTCAACGAGATGACTATGAAATAACCTAGCAACTCTCG
C549 - W - 17	
C550-W-1	
C550 W 1	
CSSO-W-Z	
C550-W-5	
C550-W-6	CAAGTTAATTCTCGATTGATTCCTCAAAGAGATGACTATGAAATAACCTATTCTCG
C550-W-12	GAAGTT-ATTCTCGATTGATTCCTCAAAGAGATGACTATGAAATAACCTATTCTCG
C550-W-27	GAAGTT-ACTCTCGATTGACTCCTCAAAGAGATGACTATGAAATAACCTAGCAACTCTCG
5 50 M 1	
B/0-N-1	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
B70-N-2	CCGCAGCTATTAATGGTTCCCACGAAGGCGCCGGAGTTGGCACTGTAGATACGTGCTGCG
B70-N-12	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C132-S-1	CCGCAGCTATTAATGGTTCCCACGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C132-S-2	CCGCGGCTATTAATGGTTCCCACGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C132-S-5	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C132-S-7	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C132-S-22	TTTACGTGCTGCG
C318-W-1	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C318-W-2	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C549-W-2	
C549-W-4	
C549 - W - 4	
C549-W-10	
C549-W-17	
C550-W-1	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C550-W-2	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C550-W-5	CCGCAGCTATTAATGGTTCCCACGAAGGCGCCGGAGTTGGCGCTGTAGTTACGTGCTGCG
C550-W-6	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C550-W-12	CCGCAGCTATTAATGGTTCCCAGGAAGGCGCCGGAGTTGGCGCTGTAGATACGTGCTGCG
C550-W-27	CCGCAGCTATTAATGGTTCCCACGAAGGCGCCGGAGTTGGCGCTGTAGTTACGTGCTGCG
D70 N 1	
B/0-N-1	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
B/0-N-2	ACGTTGAACGTTATTTATTGCCCCCCCGAACCTAACACGTCCCCTCGAATATCGTCTCAT
B70-N-12	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C132-S-1	ACGTTGAACGTTATTTATCGCCCCCCGAACCTAACACGTCCCTTCGAATATCATCTCAT
C132-S-2	ACGTTGAACGTTATTTATCGCCCCCCGAACCTAACACGTCCCTTCGAATATCATCTCAT
C132-S-5	ACGCTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C132-S-7	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C132-S-22	ACGTTGAACGTTATTTATCGCCCCCCGAACCTAACACGTCCCTTCGAATATCATCTCAT
C318-W-1	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C318-W-2	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C549-W-2	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C549 - W - 4	
C549 - W - 10	
C549_W-17	
C549-W-17	
CDDU-W-I	
CSSU-W-Z	
C550-W-5	
C550-W-6	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C550-W-12	ACGTTGAACGTTGTTTATTGCCCCCCTGGACCTAACACGTCCCCTCGAATATCATCTCAT
C550-W-27	ACGTTGAACGTTATTTATTGCCCCCTCGAACCTAACACGTCCCCTCGAATATCATCTCAT
D70 M 1	moo
BIU-N-I	
B/U-N-2	TCC
B/U-N-12	TCC
C132-S-1	TCC
C132-S-2	TCC
C132-S-5	TCC

C132-S-7 TCC C132-S-22 TCC

C318-W-1	TCC
C318-W-2	TCC
C549-W-2	TCC
C549-W-4	TCC
C549-W-10	TCC
C549-W-17	TCC
C550-W-1	TCC
C550-W-2	TCC
C550-W-5	TCC
C550-W-6	TCC
C550-W-12	TCC
C550-W-27	TCC