

Durham E-Theses

The nature of growth in the biofuel feedstock and bloom-forming green macroalga Ulva

YOON, KEVIN,JONG,COOK

How to cite:

YOON, KEVIN,JONG,COOK (2020) *The nature of growth in the biofuel feedstock and bloom-forming green macroalga Ulva*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/13489/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

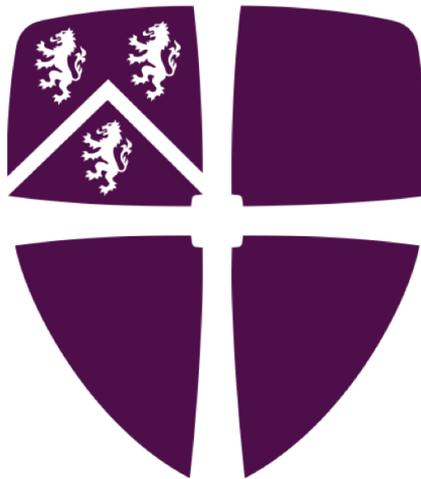
The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

The nature of growth in the biofuel
feedstock and bloom-forming green
macroalga *Ulva*

Kevin Jong Cook Yoon

A thesis presented for the degree of
Doctor of Philosophy



Department of Biosciences

Durham University

United Kingdom

2019

Acknowledgements

In no particular order, I would like to express my utmost gratitude with love always to:

Henry, who has loved and supported me indefatigably, and always forgave me, especially when I struggled to forgive myself.

My mother, who has always nurtured my inquisitive nature and taught me all that I know about love, dedication, resilience and patience.

My father and stepmother, who have shown pertinacious support for my pursuit of knowledge, truth and especially happiness, comprehension notwithstanding.

My sister, who has taught me so much about loyalty and human nature.

Lyn, Jeff and Jack, who have been my home away from home, and my anchor in an often-turbulent journey.

John, who has been the best mentor I could have ever asked for in my studies and in my life.

The multitude of amazing people I have met along this journey that have had varying contributions from advice to collaborations, and have shaped who I am as a researcher and as a person.

Declaration

I hereby declare that the work in this thesis is entirely my own and original, unless specifically mentioned and acknowledged to be the outcome of work done in collaboration. The contents of this thesis have never been submitted in whole or in part for consideration for any other degree or qualification, at Durham University or any other university or institution.

Copyright © 2019 by Kevin Jong Cook Yoon.

“The copyright of this thesis rests with the author. No quotation from it should be published without the author’s prior written consent and information derived from it should be acknowledged”.

**The nature of growth in the biofuel feedstock and bloom-forming
green macroalga *Ulva*
Kevin Jong Cook Yoon**

Abstract

Ulva is a genus of multicellular green algae that is phylogenetically similar to unicellular green algae such as *Chlamydomonas* and *Ostreococcus*. *Ulva* is present in much of the coastal benthic zones worldwide, and is of great interest for three main reasons. Firstly, *Ulva* is an important feedstock for biofuels. Secondly, many *Ulva* species are massively proliferating organisms that cause Harmful Algal Blooms, which are ecologically devastating. Finally, *Ulva* is an important model organism that could elucidate the evolution of multicellularity. This thesis investigates the physiology of growth in *Ulva* in four sequential results chapters. The first establishes a statistical proof for the goodness of fit of gene family occupancy data to a discrete power law model. This was an assumption used in the only *Ulva* genome study, which found no genomic signature for multicellularity. This establishes the baseline for the investigation of bottom-up morphogenesis in *Ulva*. The second is the investigation of differential growth, by identifying cell tessellation patterns in different morphologies of *Ulva* thalli, namely the “ribbon” and “leaf” morphotypes, with mathematical models using Voronoi tessellations. The third expands investigates differential growth in the ribbon and leaf morphotypes with a focus on identifying potential mechanisms with further mathematical models using Centroidal Voronoi Tessellations. The fourth aims to develop experimental techniques to confirm the hypotheses arising from the mathematical modelling in the second and third chapters. The first part involves the use of EdU cellular proliferation assays. The remainder of the chapter will investigate the development of a live-imaging biomass monitoring system that aims to improve the accuracy, reliability and temporal resolution of aquatic biomass measurements. It can be concluded that *Ulva* does not show a genomic signature for multicellularity, and bottom-up mechanisms likely explain its morphogenesis and morphologies.

Contents

Declaration	ii
List of Figures	vii
List of Tables	xv
1 General Introduction	1
1.1 The Global Importance of Marine Macroalgae	1
1.2 A Sea of Threats: Anthropogenic Climate Change	6
1.3 A Sea of Opportunities: Biofuels	8
1.4 A Sea of Threats: Green Tides	12
1.5 A Sea of Opportunities: Understanding Fundamental Biology	14
1.6 Ulvophyceae	19
1.7 Plant Morphogenesis and Early Development of Tissue	20
1.8 Morphogenesis and Early Development of <i>Ulva</i>	22
1.8.1 Morphology of <i>Ulva</i>	22
1.8.2 Life History of <i>Ulva</i>	25
1.8.3 Symbiosis in <i>Ulva</i>	28
1.9 The Hologenome Theory of Evolution: Applications to <i>Ulva</i>	30
1.10 Summary	31
2 General Methods	33
2.1 Challenges in experimental work	33
2.2 Sample collection	34

2.3	Sample culturing	35
2.4	Fluorescence Microscopy	36
2.4.1	History of Fluorescence Microscopy	36
2.4.2	The Principles of Fluorescence Microscopy	37
2.4.3	History of Confocal Microscopy	39
2.4.4	Principles of Confocal Laser Scanning Microscopy	40
2.4.5	Stain Selection	40
2.4.6	Scanning Parameters	47
3	Modelling Gene Family Occupancy	48
3.1	Introduction	48
3.2	Methods	55
3.3	Results	58
3.4	Discussion	62
4	Bottom-up proliferation of <i>Ulva</i> growth patterns and cellular architecture	66
4.1	Introduction	66
4.1.1	Filling Flat Space in Nature	67
4.1.2	Fundamental Drivers of <i>Ulva</i> proliferation patterns and Voronoi Tessellation	72
4.2	Methods	76
4.2.1	Stain Selection, Image Collection and Analysis	76
4.2.2	Pattern Modelling	79
4.2.3	Statistical Analysis	83
4.3	Results	84
4.3.1	Voronoi Tessellation Fit	84
4.3.2	First Phase Model Statistics	85
4.3.3	Second Phase Model Statistics	85
4.3.4	Leaf vs Ribbon Morphotype Comparisons	87
4.4	Discussion	100
5	Fundamental mathematical and physical mechanisms underlying optimal distribution in <i>Ulva</i> cell growth and proliferation patterns	106

5.1	Introduction	106
5.2	Methods	113
5.2.1	Image Collection and Analysis	113
5.3	Results	115
5.4	Discussion	131
6	Investigating the spatial and temporal patterns of growth in <i>Ulva</i>	142
6.1	Introduction	142
6.2	Optical Imaging System	143
6.2.1	Introduction	143
6.2.2	Methods and Results	146
6.3	EdU Cell Proliferation Assays	159
6.3.1	Introduction	159
6.3.2	Methods and Results	160
6.4	Discussion, Limitations and Further Studies	168
7	General Discussion	170
7.1	General Summary	170
7.2	Future direction of <i>Ulva</i> Research	173
	Bibliography	176
	Appendix A Mathematica Code - CVT	214
	Appendix B Mathematica Code - KummerU Function	216
	Appendix C R Code - Voronoi Data Extraction	218
	Appendix D Cumulative Distribution Functions	219
	Appendix E Cumulative Distribution Function Confidence Intervals	222

List of Figures

- 1.1 Four examples of the various challenges and opportunities posed by macroalgae around the world are highlighted here using four photographs with their imaging location marked on a world map. The colours of the lines indicate the group of algae, with yellow indicating Phaeophyceae, red indicating Rhodophyta and green indicating Chlorophyta. (A) a golden tide, or *Sargassum* bloom in Belize (Hunt, 2019). (B) an *Undaria* farm in the south-east of Korea (Kookje News, 2012). (C) a green tide, or *Ulva* bloom on a beach in Monterey Bay, California (Mazur, 2014). (D) an *Eucheuma* farm in the Philippines (Keats, 2010). 5
- 1.2 *Ulva* spp. bloom in Qingdao, China. Blooms like these can cover an area of several thousand square kilometres and cost millions to clean up (Jing, 2013). 12
- 1.3 Phylogenetic tree, visualised using iTOL v4.4.2. (Letunic and Bork, 2019), showing the three clades of marine macroalgae and land plants (Embryophyta) with example genera. The colours of the branches represent the colour of the algae, with Phaeophyceae shown in brown, Rhodophyta shown in red and Chlorophyta shown in green. Chlorophyta and Embryophyta are both members of the Viridiplantae clade and as such are both shown in the same colour. 15
- 1.4 A typical *Ulva* thallus displaying the rosette morphology (Novák, 2018) 23
- 1.5 Longitudinal section of an *Ulva* thallus (Fernández in Arora and Sahoo, 2015) 24
- 1.6 *Ulva* alternate between gametophytic and sporophytic generations. Red arrows represent the parthenogenetic development of unfused gametes into gametophytes (Modified from Morrissey et al., 2016) 25

1.7	Light Microscopy images of the differentiation of vegetative cells into reproductive cells in <i>Ulva</i> . (A) Vegetative Cells, (B) Formation of gametes after 26 hours, (C) Reproductive cells with fully formed gametes after 44 hours, (D) Reproductive cells with fully formed and released gametes after 46 hours (Carl et al., 2014)	26
1.8	Molecular structure of thallusin (Matsuo et al., 2005). Thallusin induces rhizoid formation and normal cell wall synthesis without the unusual callus-like colourless protrusions observed in axenic cultures (Spoerner et al., 2012; Wichard, 2015).	29
2.1	Typical samples of the (A) leaf and (B) ribbon morphotypes of <i>Ulva</i> used throughout this project	34
2.2	Molecular structure of Calcofluor White M2R (Fluorescent Brightener 28) (Sigma-Aldrich, 2019)	43
2.3	The chemical structure of Direct Yellow 96 (Ismael et al., 2013)	44
3.1	The Cumulative Distribution Functions (CDFs) for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for (A) <i>Volvox carteri</i> , (B) <i>Ostreococcus tauri</i> , (C) <i>Chlamydomonas reinhardtii</i> and (D) <i>Ulva mutabilis</i> . For a CDF, $F(x)$ is a probability between 0 and 1. The legend is shown below.	59
3.2	The Cumulative Distribution Function (CDF) plots of the confidence intervals for the discrete power law model calculated by 100 simulations for (A) <i>Volvox carteri</i> , (B) <i>Ostreococcus tauri</i> , (C) <i>Chlamydomonas reinhardtii</i> and (D) <i>Ulva mutabilis</i> . The grey data points indicate the estimated points, and the black data points indicate the observed gene family occupancy data, as above in figure 3.1. For a CDF, $F(x)$ is a probability between 0 and 1.	61
4.1	Two views (xy and zy) of the densest possible configuration of sphere packing: (A) face-centred cubic and (B) hexagonal close packing arrangements. Note that in (A), there are three repeating layers that form a unit, whilst in (B) there are only two.	68
4.2	Hexagonal rock formations at Giant’s Causeway in Northern Island (Reddit, 2017).	70

4.3	Descartes's figure of convex regions in space. S,E, ε ,A are the centres (generators). All matter within the area FGGF belongs to the cell generated from S, and likewise for HGGH and the generator ε (Gaukroger, 1998).	73
4.4	Confocal microscope image of <i>Ulva</i> spp stained with Calcofluor White.	77
4.5	Confocal microscope image of <i>Ulva</i> spp stained with Direct Yellow 96.	78
4.6	Voronoi Diagram obtained from a leaf morphotype sample, coloured to visualise data. Red represents the Edge Cells, which were excluded from various analyses. Green represents the Valid Cells included in various analyses. The diagram is generated in Wolfram Mathematica using the VoronoiMesh tool (Wolfram, 2015). See Appendix A for code.	79
4.7	Representative patterns generated by of models 1 to 4 used in the analysis . . .	80
4.8	Representative patterns generated by models 5, 6 and 7 used in the analysis. .	82
4.9	Representative confocal image of the (A) leaf and (B) ribbon morphotypes, both stained with Direct Yellow 96. The blue dots indicate the centre points of the cells, which were then used to generate Voronoi tessellation diagrams that are overlaid on the confocal images. Note that the Voronoi edges align almost perfectly with the cell walls.	84
4.10	The polygon distribution of leaf and ribbon morphotypes and model 7 (50% jitter).	88
4.11	Side by side comparison of the Voronoi tessellation of (A) a 50% Jitter model and (B) a ribbon morphotype sample.	89
4.12	Side by side comparison of the Voronoi tessellation of (A) a 50% Jitter model and (B) a ribbon morphotype sample. (B) has been illustrated with red lines to denote some of the centre lines of division that can be observed in clusters of cells that have likely divided from the same parent cells.	90
4.13	Q-Q plots of (A) two iterations of a 50% jitter model, (B) A 50% jitter model vs a 20% jitter model, (C) A 50% jitter model vs a ribbon morphotype sample, (D) 50% jitter model vs a leaf morphotype sample and (E) a leaf morphotype sample vs a ribbon morphotype sample. The blue line is the line $y=x$	93

4.14	The relative areas and perimeters of all leaf and ribbon morphotype samples plotted against each other. Green denotes the leaf morphotype, and the black denotes the ribbon morphotype. The blue line in each case is the expected curve given a perfectly regular polygon.	96
4.15	Voronoi tessellation diagrams of a (A) leaf and (B) ribbon morphotype samples. Different polygons are represented by different colours: quadrilaterals = gold , pentagons = green , hexagons = orange , heptagons = purple and octagons = brown . Diagrams generated in Wolfram Mathematica. See Appendix A for code.	97
4.16	The profiles of Voronoi tessellation diagrams generated from (A) leaf and (B) ribbon morphotype samples. Higher peaks indicate higher pixel density.	98
4.17	Profile plots of the Voronoi tessellation diagrams of the four quarters of a ribbon morphotype sample: (A) top left quarter, (B) top right quarter, (C) bottom left quarter and (D) bottom right quarter. Higher peaks indicate higher pixel density.	99
5.1	Example of (A) a Voronoi tessellation diagram and (B) a Centroidal Voronoi Tessellation diagram (right). In (A), the dots represent the Voronoi generators and the circles represent the centroids. Note that the (B) does not require this distinction as the two are exactly the same (Du et al., 1999).	112
5.2	Voronoi Diagram obtained from a leaf morphotype sample, coloured to visualise data. Red represents the Edge Cells, which were excluded from various analyses. Green represents the Valid Cells included in various analyses. Diagram generated in Wolfram Mathematica. See Appendix A for code.	114
5.3	Visual comparison of the Voronoi tessellation diagrams (A) before and (B) applying Lloyd's algorithm to a leaf morphotype sample. Diagrams generated in Wolfram Mathematica. See Appendix A for code.	116
5.4	Change of proportion of each polygon out of total sum of polygons tracked throughout Lloyd's iterations in a leaf morphotype sample. The numbers represent the number of edges in the polygon, i.e. 3 = triangles, 4 = quadrilaterals and so forth.	117
5.5	Visual comparison of the Voronoi tessellation diagrams (A) before and (B) applying Lloyd's algorithm to a ribbon morphotype sample. Diagrams generated in Wolfram Mathematica. See Appendix A for code.	118

5.6	Change of proportion of each polygon out of total sum of polygons tracked throughout Lloyd’s iterations in a ribbon morphotype sample. The numbers represent the number of edges in the polygon, i.e. 3 = triangles, 4 = quadrilaterals and so forth.	119
5.7	Confocal image of <i>Ulva</i> dyed with the cell wall stain Direct Yellow 96 showing (A) cells, (B) patches, (C) overlay of patches onto cells at lower opacity for visualisation, (D) overlay of patches onto cells at 100% opacity on one half of the image.	121
5.8	Two TEM images of <i>Ulva</i> thalli. The darker the shade, the higher the electron density, indicating a higher density of matter present.	124
5.9	Coloured Voronoi tessellation diagrams of a leaf morphotype sample (A) before and (B) after and a ribbon morphotype sample (C) before and (D) after applying Lloyd’s iteration to reach Centroidal Voronoi Tessellation. Different polygons are represented by different colours: quadrilaterals = gold , pentagons = green , hexagons = orange , heptagons = purple and octagons = brown . Diagrams generated in Wolfram Mathematica. See Appendix A for code.	125
5.10	Profile plots of the Voronoi tessellation diagrams of a leaf morphotype sample (A) before and (B) after, and a ribbon morphotype sample (C) before and (D) after applying Lloyd’s algorithm to reach Centroidal Voronoi Tessellation. Higher peaks indicate higher pixel density.	126
5.11	Profile plots of the Voronoi tessellation diagrams of a leaf morphotype sample (A) before and (B) after, and a ribbon morphotype sample (C) before and (D) after applying Lloyd’s algorithm to reach Centroidal Voronoi Tessellation, without the edge cells data. Higher peaks indicate higher pixel density.	127
5.12	Profile plots of the Voronoi tessellation diagrams of the four quarters of a ribbon morphotype sample: (A) top left quarter, (B) top right quarter, (C) bottom left quarter and (D) bottom right quarter. Higher peaks indicate higher pixel density.	128
5.13	Centroidal Voronoi Tessellation diagrams of the (A) leaf morphotype and (B) ribbon morphotype marked with the outlines of their respective patches in cyan. Different polygons are represented by different colours: pentagons = green , hexagons = orange , heptagons = purple . Diagrams generated in Wolfram Mathematica. See Appendix A for code.	129

5.14	Voronoi Tessellation diagrams of the (A) leaf morphotype and (B) ribbon morphotype after 50 Lloyd's iterations, marked with the outlines of their respective patches. Different polygons are represented by different colours: pentagons = green , hexagons = orange , heptagons = purple . Diagrams generated in Wolfram Mathematica. See Appendix A for code.	130
6.1	Algem bioreactor by Algenuity seen on the left, with a touchscreen control panel. To the right, the dedicated cooling system can be seen, and the flow control unit in the middle, with all of the pipes containing cooling fluid covered with an insulating foam layer.	146
6.2	A standard wide-mouth 1 litre beaker fitted with the specialised bung supplied by Algenuity for the Algem. The various nozzles can be configured to allow nutrient and gas supply to the culture.	147
6.3	Visualisation of the camera and lens configuration initially used in the preliminary experiments.	150
6.4	Approximate light path of the camera and lens configuration used in the preliminary experiments.	151
6.5	The final camera and varifocal lens configuration used for the duration of the present experiment. The camera is mounted directly onto the growth chamber lid.	152
6.6	First prototype design for the fixation of <i>Ulva</i> thalli within a 1 litre beaker. . .	153
6.7	Second prototype design for the fixation of <i>Ulva</i> thalli within a 1 litre beaker. .	154
6.8	Final prototype design for the fixation of <i>Ulva</i> thalli within a 1 litre beaker. . .	156
6.9	Regression curve of known mass in grams versus mass estimated by image analysis in grams.	158
6.10	Confocal image of a section of an <i>Ulva</i> thallus, (A) shows the single channel of 543 nm laser excited emissions (EdU and autofluorescence) and (B) shows the dual channel of 458 nm and 543 nm laser excited emissions (EdU, autofluorescence and Direct Yellow 96).	163
6.11	<i>Ulva</i> thalli (A) before and (B) after incubation in ethanol and acetic acid (3:1)	164

6.12	Confocal microscope image of an <i>Ulva</i> thalli stained with DY96 and EdU Alexa Fluor 555 azide. This image is a stitched 10 x 10 image of 1024 x 1024-pixel images, resulting in a total resolution of 10240 x 10240 pixels, or an area of 2460.3 x 2460.3 μm	166
6.13	Cropped view of a confocal microscope image of an <i>Ulva</i> thalli stained with DY96 and EdU Alexa Fluor 555 azide. This image is a stitched 3 x 10 image of 1024 x 1024-pixel images, resulting in a total resolution of 3072 x 10240 pixels, or an area of 738.1 x 2460.3 μm	167
D.1	The Cumulative Distribution Function (CDF) plots for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for <i>Picochlorum</i> sp. SENEW3 (SE3). For a CDF, F(x) is a probability between 0 and 1. The legend is shown below.	219
D.2	The Cumulative Distribution Function (CDF) plots for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for (A) <i>Bathycoccus prasinus</i> , (B) <i>Chlorella</i> sp. NC64A, (C) <i>Coccomyxa subellipsoidea</i> C-169 and (D) <i>Gonium pectorale</i> . For a CDF, F(x) is a probability between 0 and 1. For a CDF, F(x) is a probability between 0 and 1. Please refer to the legend above.	220
D.3	The Cumulative Distribution Function (CDF) plots for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for (A) <i>Micromonas pusilla</i> strain CCMP1545, (B) <i>Micromonas</i> sp. RCC299, (C) <i>Ostreococcus lucimarinus</i> and (D) <i>Ostreococcus</i> sp. RCC809. For a CDF, F(x) is a probability between 0 and 1. For a CDF, F(x) is a probability between 0 and 1. Please refer to the legend above.	221
E.1	The Cumulative Distribution Function (CDF) plot of the confidence intervals for the discrete power law model calculated by 100 simulations for <i>Picochlorum</i> sp. SENEW3 (SE3). The grey data points indicate the simulated points, and the black data points indicate the observed gene family occupancy data. For a CDF, F(x) is a probability between 0 and 1.	222

E.2	The Cumulative Distribution Function (CDF) plots of the confidence intervals for the discrete power law model calculated by 100 simulations for (A) <i>Bathycoccus prasinos</i> , (B) <i>Chlorella</i> sp. NC64A, (C) <i>Coccomyxa subellipsoidea</i> C-169 and (D) <i>Gonium pectorale</i> . The grey data points indicate the simulated points, and the black data points indicate the observed gene family occupancy data. For a CDF, $F(x)$ is a probability between 0 and 1.	223
E.3	The Cumulative Distribution Function (CDF) plots of the confidence intervals for the discrete power law model calculated by 100 simulations for (A) <i>Micromonas pusilla</i> strain CCMP1545, (B) <i>Micromonas</i> sp. RCC299, (C) <i>Ostreococcus lucimarinus</i> and (D) <i>Ostreococcus</i> sp. RCC809. The grey data points indicate the simulated points, and the black data points indicate the observed gene family occupancy data. For a CDF, $F(x)$ is a probability between 0 and 1. . . .	224

List of Tables

2.1	The growth parameters set for <i>Ulva</i> in the incubators and the justifications for their use throughout this project.	35
2.2	The mean wet and dry mass, percentage change and water content of <i>Ulva</i> over five days in the laboratory culture conditions outlined above.	36
3.1	The species name, genome size, habitat and body organisation of the species investigated in this chapter. The results for the first four are directly shown in this chapter.	58
4.1	Wilcoxon Signed Rank Test p-values for Model 1 (Regular Hexagon Honeycomb), Model 2 (Squashed Hexagon Honeycomb), Model 3 (Random) and Model 4 (Square Grid).	85
4.2	Wilcoxon Signed Rank Test p-values for Model 1 (Regular Hexagon Honeycomb) and associated jitter models.	86
4.3	Wilcoxon Signed Rank Test p-values for Model 2 (Squashed Hexagon Honeycomb) and associated jitter models.	86
4.4	Wilcoxon Signed Rank Test p-values for models of equal jitter (regular vs squashed honeycomb).	87
4.5	Wilcoxon Signed Rank Test p-values for leaf vs ribbon morphotypes. For each of the 12 samples of each morphotype, the samples were compared with their corresponding sample number of the other morphotype (i.e. leaf sample 1 vs ribbon sample 1). Sample 2 has been highlighted as an anomalous result, as addressed in the discussion.	87

5.1	Iterations required to reach Centroidal Voronoi Tessellation (CVT) in both leaf and ribbon morphotypes. The ribbon morphotype sample 2 has been highlighted as an anomalous result, as addressed in the discussion of chapter 4, as well as again in the discussion of this chapter.	115
5.2	The mean start, end and total change in percentage in polygon distribution, between the first and last iteration in the leaf morphotype.	116
5.3	Mean percentage change of each polygon between the given ranges of iterations for the leaf morphotype.	117
5.4	The mean start, end and total change in percentage in polygon distribution, between the first and last iteration in the ribbon morphotype.	118
5.5	Mean percentage change of each polygon between the given ranges of iterations for the ribbon morphotype.	119
5.6	The mean, standard deviation and coefficient of variation for the start and end percentage, and percentage change of hexagons in the leaf and ribbon morphotypes.	120
5.7	The percentage of total positional changes occurring in the iteration ranges given, for both morphotypes.	120

General Introduction

1.1 The Global Importance of Marine Macroalgae

Around half of the World's primary production occurs in its waters (Falkowski and Raven, 2007; Field et al., 1998; Vitousek et al., 1986). The bulk of this carbon fixation is carried out by the unicellular phytoplankton of the ocean's surface, but an important part is generated by the macroalgae, or seaweeds, found in its coastal regions. This macroalgal primary production comprises a small, but disproportionately important, fraction of the total: marine macroalgae are the dominant primary producers in coastal benthic zones (Charpy-Roubaud and Sournia, 1990; Middelburg et al., 2005) and a significant intertidal carbon sink (Field et al., 1998; Krause-Jensen and Duarte, 2016; Smith, 1981). Furthermore, marine macroalgae serve a number of critical ecological services. Marine macroalgae, like seagrasses and other coastal vegetation, provide a major ecosystem and economic service by wave attenuation and sea-bed stabilization (Asano, 2006; Kobayashi et al., 1993). This ecological facilitation aids recruitment of other marine macroalgae and plants (Bennett and Wernberg, 2014), allowing for a diverse assemblage that is stable and resilient (Stachowicz et al., 2008). Marine macroalgae also act as habitats and essential nursery grounds for fish and invertebrates, and as important food sources (Harley et al., 2012; Santelices et al., 2009).

This rich primary productivity is likely enabled by the enormous physiological diversity of the macroalgae, which allows for the realisation of more ecological niches (Helaouët and Beaugrand, 2009), and by their corresponding taxonomic diversity: the macroalgae have over 11,000 known species from across 1,400 genera (Guiry and Guiry, 2019; Lee, 2008;

van den Hoek et al., 1995).

If we are to understand marine macroalgae, therefore, we need to consider this broad range of physiologies and life histories, looking at species that are benthic, pelagic and holopelagic. The three major areas of interest to humans when studying marine macroalgae are ecological, environmental and the economic benefits they can provide.

First, the marine macroalgae display distinct and unique ecological adaptations and patterns. For example, marine macroalgae show a diverse biogeography, with many species that are near fully cosmopolitan and those that are localised, sometimes even to a single site (Garbary, 2001). The biogeography of marine macroalgae shows an unusual characteristic, in that their global diversity gradients are the opposite of the classical latitudinal patterns seen in most other clades (Bolton, 1994; Keith et al., 2014; Kerswell, 2006). One of classical ecology's oldest and most fundamental recognised patterns is the decrease of species richness from tropical latitudes towards the polar latitudes (Brown et al., 1998). This classical latitudinal richness pattern, often termed the Latitudinal Diversity Gradient (LDG), is seen regardless of taxa, longitudinal geography, and time (Willig et al., 2003). That is to say that, from animals to plants, from terrestrial organisms to aquatic organisms, from millions of years ago to this very day, the same pattern has held true. Curiously, there are no definite explanations regarding the mechanisms underlying this pattern, with at least 30 different hypotheses having been proposed without any significant consensus emerging (Kerswell, 2006; Rosen, 1988; Willig et al., 2003). The fact that marine macroalgae as a whole are one of the few main exceptions, alongside gymnosperms (Fragnière et al., 2015), to the LDG provides an opportunity to help inform a model that may elucidate the underlying driving mechanisms of this global ecological pattern.

Second, while the biogeography of the marine macroalgae may open up new ecological vistas, a sad fact they share with many other species is that their distributions are increasingly affected by anthropogenic climate change and habitat destruction. As with other species affected by climate change, the type and extent of the effects are yet to be appreciated. Therefore, a particularly important question is the response of marine macroalgae to anthropogenic climate change and ocean acidification. These are topics that are poorly understood, with no shortage of critical gaps in the literature. The biogeography of marine macroalgae is expected to change as dissolved inorganic carbon rises. A literature review of more than 100 species of macroalgae found that more than 85 % of macroalgae utilise

C₃ photosynthesis, and only around 15 % utilise C₄ photosynthesis (Koch et al., 2013). The rate of photosynthesis in species that utilise C₄ photosynthesis is not limited by CO₂ concentration due to the presence of a CO₂ concentrating mechanism, whereas the rate of photosynthesis is limited by CO₂ concentration in species that utilise C₃ photosynthesis (Bowes et al., 2002). This means that as atmospheric CO₂ concentrations increase, the rate of photosynthesis in many marine macroalgal species will increase, which should in general lead to increased growth rates similar to those expected of C₃ land plants (Koch et al., 2013).

However, calcifying macroalgae, which are a diverse range of organisms encompassing several genera in all major marine macroalgal taxonomical divisions, will be detrimentally affected by the changes in carbonate biogeochemistry caused by ocean acidification due to decreases in cell wall rigidity and consequent disturbances of cell development, morphogenesis and functioning (Guinotte and Fabry, 2008; Hofmann and Bischof, 2014; Koch et al., 2013). Increased atmospheric CO₂ causes ocean acidification and consequent changes in carbonate biogeochemistry, which lessen coral competitiveness whilst enhancing macroalgal competitiveness, making anthropogenic climate change a significant driver of phase shifts, alongside complex ecological shifts of grazers, from coral dominated ecosystems to macroalgae dominated ecosystems (Diaz-Pulido et al., 2011; Hoegh-Guldberg et al., 2007). There are a handful of well documented examples of range shifts in macroalgae in response to anthropogenic climate change over relatively long time periods, of at least 10 years. This includes range expansions and contractions due to both local cooling and warming, for example the westward shift of *Fucus serratus*, at a rate of 1.7 km yr⁻¹, and *Himantalia elongata*, at 4.4 km yr⁻¹, in northern Spain due to the warming of the Bay of Biscay (Straub et al., 2016). All of this combined will likely lead to several significant distribution range and ecosystem composition shifts in a variety of marine habitats, though substantial projections of global effects and patterns are still non-existent due to the paucity of data (Koch et al., 2013; Porzio et al., 2011).

Third, the macroalgae have long been of great economic importance, especially in Asia, where many farmers are dependent on aquaculture of marine macroalgae for their livelihoods. In addition to those whose livelihoods depend directly upon macroalgae aquaculture, the uncertainty regarding macroalgal responses to global anthropogenic climate change crisis will undoubtedly be a cause for much concern to the billions of people around the

world utilising macroalgae for a plethora of uses including as food, feeds, fertilisers, cosmetics, medicines and biofuels (Buschmann et al., 2017). In 2016, the total production of marine macroalgae from aquaculture was 30.1 million tonnes, representing an economic value of approximately USD 11.7 billion (FAO, 2018). Most of the world's aquatic plant aquaculture occurs in Asia, with China and Indonesia being the dominant producers. The key species cultivated include *Eucheuma* and *Kappaphycus* for carrageenan extraction in Indonesia, *Laminaria japonica*, *Porphyra* spp., *Undaria pinnatifida* and *Sargassum fusiforme* for human food in China, Japan and Korea (FAO, 2018). It is expected that by 2050, 50 to 70 % more food will need to be produced to meet current consumption trends (Bjerregaard et al., 2016). Marine macroalgae could represent a critically important food source in the coming years, especially in a world with increasingly precarious global food security due to the rising production demands in increasingly adverse conditions. Marine macroalgae are generally low in fat (>5 %), rich in carbohydrates (44 – 52 %) and protein (7 – 44 %), as well as high in a variety of minerals, vitamins and other micronutrients (Kim, 2011; Pereira and Yarish, 2008; Radulovich et al., 2014). Furthermore, marine macroalgae could represent a significant economic resource for many developing nations. Increasing global macroalgal agriculture by 14% yr⁻¹ would result in a global production rate of 500 million dry tonnes by 2050, representing a value of approximately USD 500 billion (Bjerregaard et al., 2016). Figure 1.1 below illustrates the challenges and opportunities that can arise from macroalgae, and their widespread geography.



Figure 1.1: Four examples of the various challenges and opportunities posed by macroalgae around the world are highlighted here using four photographs with their imaging location marked on a world map. The colours of the lines indicate the group of algae, with yellow indicating Phaeophyceae, red indicating Rhodophyta and green indicating Chlorophyta. (A) a golden tide, or *Sargassum* bloom in Belize (Hunt, 2019). (B) an *Undaria* farm in the south-east of Korea (Kookje News, 2012). (C) a green tide, or *Ulva* bloom on a beach in Monterey Bay, California (Mazur, 2014). (D) an *Eucheuma* farm in the Philippines (Keats, 2010).

Finally, *Ulva* is a potentially important feedstock for third generation biofuels, which offers an attractive solution to both the climate change crisis that we face, as well as the current and impending fuel crises. Growth optimisation of *Ulva* will also benefit certain cultural usage, as *Ulva*, or “aonori” is interestingly an important food crop in Japan and Korea (Carl et al., 2014; McHugh, 2002). Given the favourable nutritional profile of low fat, low sodium, high protein (up to 20%) and high vitamin A and B content, in addition to the

rapid growth rate and the fact that no arable land is required for growth, *Ulva* may become a more globally important food crop in a world that has 83 million new mouths to feed each year, and is projected to reach a population of 11.2 billion people in 2100 (McHugh, 2002; United Nations, 2017).

1.2 A Sea of Threats: Anthropogenic Climate Change

Anthropogenic climate change is arguably the greatest threat humanity has ever faced. The primary cause of climate forcing is anthropogenic burning of fossil fuels (IPCC, 2013). As a result of decades of recklessly burning fossil fuels, on 9th May 2013, Global atmospheric carbon dioxide (CO₂) concentrations surpassed the 400 ppm mark for the first time since accurate continuous CO₂ measurements started (NOAA, 2016; UNEP, 2013). This marks the highest concentration of CO₂ since the Pliocene Epoch, 5.3 to 2.6 million years ago, when the global temperature was on average 2-3 °C above pre-industrial temperatures and the sea levels were up to 30 m higher than present (Dwyer and Chandler, 2009; IPCC, 2013; Robinson et al., 2008). The environmental shifts that are expected as a result of recent warming will be catastrophic to modern society.

The most common measure of impact to society is the economic measure; from politicians to insurance companies, everyone seems to value money as the best measurement of devastation. The economic loss caused by climate change already exceeds the cost of transition to a low-carbon economy (Edenhofer et al., 2010). The cost of inaction against climate change will be approximately £52 trillion GBP; this figure has been calculated from the costs of just 13 aspects of climate change, including sea level rise, ecosystem damage and extreme weather amongst others (Watkiss et al., 2005). A particularly costly risk for the least developed countries is sea level rise, as they are generally the least prepared and most vulnerable to it (McKinnon and Hickey, 2009). However, major coastal cities around the world are predicted to suffer losses of more than £660 billion GBP, regardless of how developed they are (Hallegatte et al., 2013). The global sea level has risen 195 mm from January 1870 to December 2004, and the rate of rise is accelerating significantly (Church and White, 2006). Some claim the upper limit of sea level rise is expected to be approximately 1.8 m by the end of the 21st Century (Jevrejeva et al., 2014). However there have been doubts cast on the figure, for example due to palaeoclimatological and palaeoceanographic

graphic studies confirming a 6 m rise in sea level due to the collapse of just the Greenland Ice Sheet during the late-Quaternary interglaciations, suggesting similar catastrophic rises are possible (Reyes et al., 2014). A sea level rise of just 1 m would cause a minimum £619 billion GBP loss from GWP, without taking into consideration the damage control and management costs (Anthoff et al., 2006).

However, it is not just the economic costs that should be considered. The direct effects of climate change, such as sea level rise, extreme weather events and desertification, already account for 400,000 deaths a year, with an additional 4.5 million deaths annually attributed to the effects of the carbon-intensive industries (DARA and Climate Vulnerable Forum, 2012). Additionally, some 310 million people are already facing the direct pressures of the various effects of climate change, with a further 600 million people at risk in the near future from sea level rise alone (DARA and Climate Vulnerable Forum, 2012; McGranahan et al., 2007). Furthermore, plants and animals are also facing great pressures with the rate of habitat loss being the highest ever (Laurance, 2010). The pressure is evident, with the rate of loss of biodiversity being so outrageously high, at up to 140,000 species year⁻¹, that current times have been termed the Holocene Extinction, or the Sixth Mass Extinction (Kolbert, 2014; McCallum, 2015; Pimm et al., 1995).

Ocean acidification is expected to be responsible for much of the mass extinction of marine organisms, and is often dubbed the “other CO₂ problem” (Doney et al., 2009; Orr et al., 2005). Ocean acidification is the decrease in pH caused by CO₂ dissolving in sea water to form carbonic acid. The IPCC AR5 Representative Concentration Pathways models estimate a likely decrease (high confidence) between 0.14 and 0.32 by 2100 (IPCC, 2013). This affects the carbonate chemistry as increased [H⁺] will lead to a decrease in [CO₃⁻], making it difficult for organisms to form biogenic CaCO₃ and decrease calcification rates (Orr et al., 2005; Zondervan et al., 2001). This decreases cell and tissue rigidity and stability as well as development, severely affecting key calcifying organisms such as coral and certain plankton (Orr et al., 2005). In a business as usual scenario, much of the world’s oceans could suffer from undersaturation of calcium carbonate by 2100 (Orr et al., 2005).

There is no doubt that climate change and ocean acidification are gargantuan problems, but perhaps more relevant to our consumption-obsessed society, mankind is simply running out of fuel. As of 2020, it is estimated that there are approximately 34 and 36 years left until reserve depletion for oil and gas respectively at our current rate of consumption (Owen

et al., 2010; Shafiee and Topal, 2009). The age of cheap and readily available fossil fuels is over; there is no choice but to find alternative sources of energy.

1.3 A Sea of Opportunities: Biofuels

For all of the aforementioned reasons and more, fossil fuel alternatives are sought after. It is a common idea to the majority of the population in developed countries that a biofuel is a low-emission, renewable and therefore an ideal, environmentally-friendly form of energy (Renewable Fuels Agency, 2008). It is understandable then, that biofuels are highly regarded as being a major component of climate change mitigation.

However, this is an overly simplistic view of the realities of biofuels, omitting crucial details that critically affect the viability of various biofuels (Sims et al., 2010). Traditionally, the carbon costs of the production of biofuels have often been neglected when discussing the use of biofuels, whilst the carbon benefits are usually exaggerated excessively and unrealistically (Fargione et al., 2008; Searchinger et al., 2008). These carbon costs include major problems such as the carbon debt accumulated during land conversion for feedstock growth, which is a concept that is addressed in more detail below, the societal and moral problems that are caused by redistribution of land and resources, as well as relatively more minor carbon costs such as the greenhouse gas emission from the processing and transporting of crops and biofuels (Fargione et al., 2008; Lawrence et al., 2010; Nuffield Council on Bioethics, 2011; Renewable Fuels Agency, 2008; Searchinger et al., 2008).

There are many political complications regarding the production of biofuels. Due to targets set out by various organisations, such as the EU target of 10% biofuel in all transport fuel by 2020, many countries must redistribute land use or crop use in order to meet these targets. One of the most scrutinised examples is that of the United States, in which 69.1 billion litres, approximately 11.3%, of total motor fuel consumption must be from biofuels in 2012, according to the Energy Independence and Security Act of 2007 (US Government, 2007). This figure was to rise to 75.24 billion litres, approximately 12.4%, in 2013 and continue to grow and reach 36 billion litres, approximately 26.9%, by 2022 (EPA, 2012). This is an extremely problematic legislation, as there are no other legislations that limit the use of food crop to produce biofuels. The US must use approximately 44.4% of its total corn production for producing biofuels in the forthcoming years, which will result

inevitably in a food shortage both domestically and internationally (USDA, 2016). This is especially reckless, as the US has recently suffered the worst drought years since the 1950s, and is bracing itself for the potential of the worst drought in 1,000 years (Cook et al., 2015; NOAA, 2012). The total corn production used in the production of bioethanol in 2016 was 5.124 billion bushels, which is enough to feed 640 million people for one whole year (Runge and Senauer, 2007; USDA, 2016).

It is tremendously difficult to argue that meeting the biofuel target is of greater importance than immediately relieving famine and reducing poverty, and yet the US remain adamant on increasing the corn bioethanol production despite the food shortage it creates, arguing that the long term benefits of meeting these targets are far greater (US Government, 2007). Furthermore, the US argue that it is stated in the 1992 Rio Declaration on Environment and Development that although preservation of the environment must be an essential part of development, states have the ultimate right to exploit their own natural resources as long as negative impacts on the environment of other states are not caused (UN, 1992). Therefore, the US has a valid international policy argument for claiming that it is acceptable for the US to redistribute the use of crops as it sees fit. This has attracted widespread criticism, including by Jean Ziegler, the independent expert on right to food at the UN, who heavily criticised the US policies claiming that using corn to produce biofuels is “a crime against humanity”, causing millions around the world to starve (Ziegler, 2013). Jean Ziegler is not alone in his concerns, with countless other experts and organisations voicing similar concerns about food shortage and biofuel ethics (Doornbosch and Steenblik, 2008; Lawrence et al., 2010; Nuffield Council on Bioethics, 2011; Renewable Fuels Agency, 2008). It is obvious that not only is the production of biofuels riddled with socio-economic and moral dilemmas, but they also have a detrimental impact on greenhouse gas emissions and carbon debt throughout the world.

The concept of carbon debt is one that is increasingly becoming the focus of the scientific community. This refers to the devastating amount of damage in carbon release that is done when biofuels are produced by converting land such as rainforests, peat lands, savannahs and grasslands into crop fields (Fargione et al., 2008; Searchinger et al., 2008). The two largest organic terrestrial sinks of carbon are soil and plant biomass, which, combined, store 274% more carbon than the atmosphere (Schlesinger and Bernhardt, 2013). Therefore, releasing this carbon store is far more detrimental to carbon emissions than increased

greenhouse gas emissions. During the land conversion process, carbon is released in the short term when the land is cleared rapidly by burning the existing vegetation and the finer roots and plant matter such as leaves decompose. This is followed by a prolonged period of carbon release due to the decomposition of remaining plant matter, primarily lignocellulosic biomass such as roots, branches and wood (Fargione et al., 2008). The combined short-term and long-term (50 years) release of carbon is termed the carbon debt. In order to repay the carbon debt incurred by the land conversion processes in the production of biofuels, the US will need at least 93 years, Brazil will need some 319 years and Southeast Asia will require a staggering 423 years (Fargione et al., 2008).

The complications that have just been discussed apply primarily to first generation biofuels, which are crop plants repurposed for bioenergy, such as those derived from corn, soybeans and sugar cane. These complications are less prohibitive for more advanced generations of biofuel, so this is where research effort, and our attention, is now turning.

Second-generation biofuels are biofuels derived from non-food crops, as well as feedstocks including forestry and agricultural waste, and also including various grasses such as silvergrass (*Miscanthus*) and short rotation forestry crops such as poplar (*Populus*) (Lee and Lavoie, 2013; Naik et al., 2010; Saladini et al., 2016).

Third-generation biofuels are biofuels derived from algal, both microalgal and macroalgal, biomass.

Fourth-generation biofuels are a broader category of biofuels, within which many different organisations and researchers classify different types of biofuels, ranging from metabolically engineered algae to photobiological solar biofuels and electrofuels derived from synthetic biological approaches (Aro, 2016; Cameron et al., 2014; Dutta et al., 2014).

Macroalgal biofuels are very attractive to many investors, as they do not require arable land (in aquatic culture systems they do not require any land for growth), they do not compete with food, and they grow biomass far more efficiently and rapidly than first generation feedstock (Sims et al., 2010; Singh et al., 2011).

The principle is similar to first-generation biofuels, such as sugar cane, with the final products being ethanol and methane. Traditional fermenters are based on microorganisms that convert cellulose into simple sugars, and therefore are not as effective at breaking down the diverse polysaccharides obtained from macroalgal feedstock, such as laminarin

and mannitol in brown algae, agar and carrageenan in red algae, and mannan and ulvan in green algae (Jung et al., 2013). Macroalgal feedstock have an important advantage over traditional lignocellulosic feedstock, as they lack lignin (Ross et al., 2008; van den Hoek et al., 1995), therefore they do not necessitate expensive and harsh pre-treatment such as mechanical comminution, steam explosion, acid hydrolysis, alkali swelling and exposure to supercritical fluids (Agbor et al., 2011; McMillan, 1994).

In general, therefore, there are three major areas on which research effort is currently focused:

First, is the improvement of algal strains, focusing on macroalgal breeding through Next Generation Sequencing marker-assisted selection to create new biofuel stocks that are not genetically modified organisms, a critical consideration for any feedstock that must be grown in the open sea.

Second, the identification, characterisation and improvement of biomass processing enzymes, screening both known and novel degradative environments (grazer intestinal flora, rotting or diseased biomass) for culturable degradative microbes and metagenomic clones from unculturable microbes that can contribute to macroalgal pre-processing (biomass to polymers) and processing (polymers to oligomers).

Third, systems and synthetic biology approaches, combining experimental manipulation and theoretical modelling of known and discovered metabolic capabilities to develop microbial platforms capable of advanced biofuel generation from macroalgal biomass (oligomers to longer chain alcohols).

With the advancement of biofuel technology in these areas, including those using macroalgal biomass, is expected to yield desirable hydrocarbons for use as direct replacements for automotive fuels (Regalbuto, 2009). The current bottleneck is the lack of technology. However, the technology for these processes is expected to improve quickly and significantly in the coming years with stimulation from the generous tax breaks, subsidies and investments, including almost £ 15 billion GBP of tax breaks and subsidies worldwide and more than £3.5 billion GBP of new global investments in 2014 into research and development (FS-UNEP Collaborating Centre, 2015; Gerasimchuk et al., 2012).

1.4 A Sea of Threats: Green Tides

Over the past decade, the ability of chlorophyte macroalgae to grow rapidly has been increasingly evident in the “Green tides”, which is the common name for the Harmful Green Algal Blooms that now recurrently plague intertidal flats and sublittoral zones along much of the world’s coasts (Ansell et al., 1998). The most famous recent example was similar to that illustrated in figure 1.2 below.



Figure 1.2: *Ulva* spp. bloom in Qingdao, China. Blooms like these can cover an area of several thousand square kilometres and cost millions to clean up (Jing, 2013).

Figure 1.2 shows the annual bloom of *Ulva Spp.* in Qingdao’s coasts in the Yellow Sea. Smaller blooms occur worldwide, from the beaches of Brittany to the coasts of Australia, with many being so large that remote sensing techniques must be used to monitor them (Hu et al., 2010; Shen et al., 2012; Ye et al., 2011).

Green tides cause significant disruptions to the local ecosystems by outcompeting other species for nutrients and often having allelopathic effects, causing biochemical inhibition of growth, survival or reproduction, on other algae and invertebrates (Castaldelli et al., 2003; Nelson et al., 2003; Valiela et al., 1997). The exaggerated increase in primary productivity by opportunistic macroalgae such as *Ulva* can cause the creation of hypoxic and anoxic zones, in which dissolved oxygen is temporarily but drastically reduced, causing a dystrophic crisis. These eutrophication events cause mass mortality of fauna and flora

(Castaldelli et al., 2003; Herbert, 1999). *Ulva* can further this problem by creating a physical barrier within water columns due to the large mean size of its thalli, which means that a lower water column can be starved of oxygen and kept in a hypoxic or anoxic state regardless of whether the *Ulva* is dead or alive (Viaroli et al., 1996). Even after the collapse and death of these blooms, the decomposition can also lead to large scale dystrophic crises (Castaldelli et al., 2003).

Green tides, like other algal blooms, are thought to be caused by a range of conditions and combinations of eutrophication and specific local conditions such as the topography, thermocline, pycnocline and current systems (Fletcher, 1996; Lapointe, 1999; Ye et al., 2011). For example, in Qingdao the cause is thought to be the removal of unwanted *Ulva* from edible seaweed nori (*Porphyra yezoensis*) aquaculture rafts nearby (Hu et al., 2010; Liu et al., 2010). Anthropogenic eutrophication of coastal waters in particular is thought to have a strong link with Harmful Algal Blooms, and due to the increasing amounts of anthropogenic input, the issue will only get worse, as shown by the case of the gargantuan recurrent green tides along the coasts of Qingdao that has been a major issue since at least 2007 (Paerl and Whitall, 1999; Ye et al., 2011).

In particular, anthropogenic nitrogen loading is of major concern in the formation of Harmful Algal Blooms, as well as in the restructuring of the indigenous floral and faunal communities by opportunistic macroalgae as nitrogen availability is widely considered to be one of the limiting factors for primary productivity in coastal benthic zones (Herbert, 1999). Anthropogenic nitrogen loading can be from a wide array of sources including, but not limited to, fertiliser run-off, sewage discharges, aquaculture and atmospheric deposition from dry and wet fall (Balls, 1994; Paerl and Whitall, 1999; Postma, 1984; Soulsby et al., 1985).

Ulva responds rapidly to increased nitrogen availability in both lab and field settings, and takes up dissolved inorganic nitrogen in both nitrate or ammonium form, though it responds even better to ammonium, which is the dominant form of nitrogen in anthropogenic fertilisers (Erisman et al., 2007; Smil, 2001; Teichberg et al., 2008). In California, *Ulva* has been observed to enhance growth rates in other *Ulva* species such as *Ulva intestinalis* by a mechanism of enriching the availability of dissolved organic nitrogen (Fong et al., 1996).

Eutrophication is an extremely urgent issue, like climate change, that also has its roots in

the exponential increase in the consumption of goods by humans, and can severely impact terrestrial, freshwater and marine ecosystems, also much like climate change (Smith et al., 1999). Alarmingly, in addition to the consequences of nitrogen loading, anthropogenic eutrophication has been shown to have even more severe impacts on coastal zone carbonate chemistry than ocean acidification, which in itself is cause for extreme concern (Borgesa and Gypensb, 2010).

The clean-up of the more than one million tons of *Ulva* from Qingdao's coast in 2008 cost more than \$100 million USD, involving a clean-up crew of over 10,000 people, 1,000 vessels and 24 km of containment booms amongst other measures (Hu et al., 2010; Hu and He, 2008). This presents a serious economic issue, in addition to the ecological issue, the gravity of which will only continue to increase in parallel with the growth of the scale of green tides. It is important then, that these green tides are studied, and the causes identified, in order that they may be prevented or at least the damage minimised.

1.5 A Sea of Opportunities: Understanding Fundamental Biology

Marine macroalgae have evolved a variety of life cycles to survive and thrive in the ocean. There are three broad categories of macroalgal life cycles: benthic, pelagic and holopelagic. The life cycles of benthic macroalgae involve attachment to substrates in the benthic zone ranging from gravel and sand to boulders and stones, with a holdfast for the entirety of their lives after their initial attachment. *Macrocystis*, *Nereocystis*, *Lessonia*, *Ecklonia* and *Laminaria* are all examples of benthic marine macroalgae; specifically marine macroalgae that form vast kelp forests, up to 8 km long and 1 km wide along the Pacific Coast of North America (Gaines and Roughgarden, 1987), in the benthic zones of the coast of all five continents (Raffaelli and Hawkins, 1999; Steneck et al., 2002).

Pelagic marine macroalgae can be attached or free floating, in various species this depends on the life cycle stage as well as the environmental pressures, such as turbulence. *Ulva* and *Sargassum* both show this type of life cycle, being attached to the benthic zone for growth, but being able to survive by forming rafts and continuously undergoing vegetative reproduction in the open ocean (Lee, 2008).

Holopelagic marine macroalgae are exclusively pelagic, meaning that every stage in the entire life cycle occurs in the open ocean and no stage involves attachment to the benthic zone.

Currently, only two species are considered holopelagic: *Sargassum natans* and *Sargassum fluitans* (Smetacek and Zingone, 2013).

It should be noted that it is often useful to discuss marine macroalgae by behaviour as many fill similar niches in both ecosystems and in human exploitation. However, it is important to acknowledge that there are distinct divisions of marine macroalgae, with significant phylogenetic, physiological and biochemical differences. Marine macroalgae are largely classified into three ecologically converged, but evolutionarily distinct phyla: the Rhodophyta (red), Phaeophyceae (brown) and Chlorophyta (green) as shown below in figure 1.3 (Keith et al., 2014).

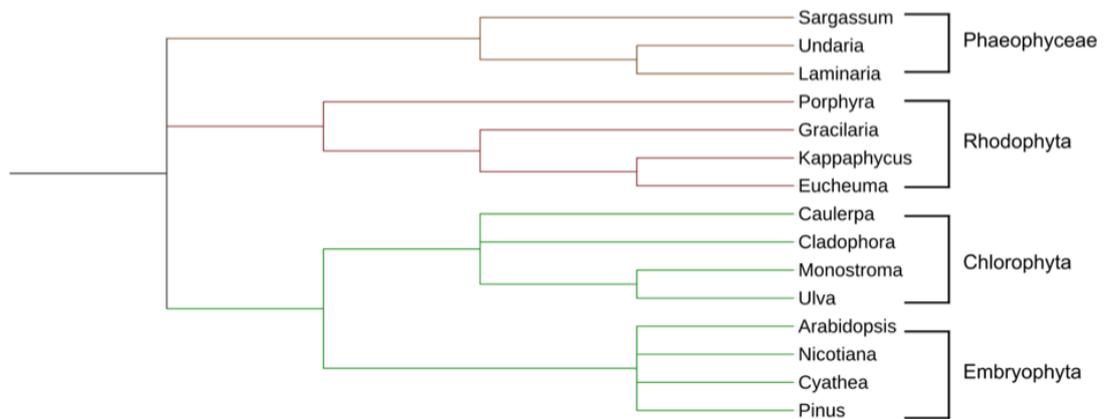


Figure 1.3: Phylogenetic tree, visualised using iTOL v4.4.2. (Letunic and Bork, 2019), showing the three clades of marine macroalgae and land plants (Embryophyta) with example genera. The colours of the branches represent the colour of the algae, with Phaeophyceae shown in brown, Rhodophyta shown in red and Chlorophyta shown in green. Chlorophyta and Embryophyta are both members of the Viridiplantae clade and as such are both shown in the same colour.

The most familiar macroalgae along the shores of the UK are probably the Phaeophyceae, or brown algae, which includes kelps and wracks, are a class of exclusively multicellular eukaryotic algae in the phylum Ochrophyta, composed almost entirely of marine algae, with the exception of four freshwater genera (Lee, 2018; van den Hoek et al., 1995). The characteristic colour of the Phaeophyceae is due to the presence of the carotenoid fucoxanthin (Lee, 2018). Unlike members of the Rhodophyta and Chlorophyta, members of the Phaeophyceae generally do not undergo vegetative reproduction by fragmentation due to the nature of their highly specialised tissue structures, with the exception of the pelagic genera *Sargassum* (Kim, 2011).

The Phaeophyceae life cycle involves alternating haploid gametophyte and diploid sporo-

phyte phases (Bell, 1997; van den Hoek et al., 1995). Meiosis occurs in the unilocular sporangia in the sporophyte, and the gametes are released to form gametophytes with plurilocular sporangia (Bell, 1997; Lee, 2018). There have been more Phaeophyceae sexual hormones identified than for the Rhodophyta or Chlorophyta, and curiously these are quite non-specific in nature, with the same pheromones having been identified in different genera (Frenkel et al., 2014).

The Phaeophyceae are also critically important for human exploitation, accounting for approximately 35 % of all aquaculture production of aquatic plants, almost exclusively for direct human consumption of *Laminaria japonica* (Korean: dasima; Japanese: makombu), *Undaria pinnatifida* (Korean: miyeok; Japanese: wakame) and *Sargassum fusiforme* (Korean: tot; Japanese: hijiki) (FAO, 2018).

It is important to note that whilst Phaeophyceae is often confused and discussed as a phylum “Phaeophyta”, Phaeophyceae is now considered a class within the phylum Ochrophyta (Heterokonta/Stramenopiles), which encompasses at least 18 classes, including the Xanthophyceae (yellow-green algae), Chrysophyceae (golden-brown algae) as well as classes of diatoms and siphonous fungi (Blackwell, 2009; Garbary, 2001; Guiry, 2012; Lee, 2018; van den Hoek et al., 1995). The class Phaeophyceae is often discussed as an equal to the phyla Rhodophyta and Chlorophyta, as the class Phaeophyceae encompasses all multicellular marine brown algae of interest, such as *Laminariales* (kelp) and *Fucales* (*Fucus*) (Lee, 2018; van den Hoek et al., 1995).

The Rhodophyta, or red algae, are a monophyletic group consisting solely of the eukaryotic algae class Rhodophyceae within the newly proposed Archaeplastida clade that also includes, the Glaucophyta, green algae, including the Chlorophyta, and the Plantae (Lee, 2018; Yoon et al., 2016). Their characteristic red colour is a result of the phycobilioproteins including R-phycocyanin, allophycocyanin, and B, R and C-phycoerythrins (Lee, 2018). Rhodophyceae lack many key traits found in other algae, including flagella, light-sensing phytochromes and the glycosylphosphatidylinositol (GPI)-anchor biosynthesis and macroautophagy pathways (Yoon et al., 2016). The Rhodophyceae utilise floridean starch grains, distinct from green algal starches, in the cytoplasm as a storage product (Lee, 2018; Yoon et al., 2010). The Rhodophyceae life cycle involves one haploid phase, wherein the monosporangia and parasporangia produce asexual spores that vegetatively develop into a genetic clone of the parent thallus. There are also two diploid phases, wherein, upon fer-

tilisation of the carpogonium by spermatium from the male gametophyte by means of currents, the female gametophyte produces gonimoblast filaments that form carposporangia, that form diploid carpospores that are released. The diploid carpospores then develop into diploid tetrasporophytes and finally haploid tetraspores (Lee, 2018; Yoon et al., 2010).

The Rhodophyceae encompasses at least ten orders of diverse organisms, both unicellular and multicellular, with some showing unique traits such as calcification in *Corallinales*, which are vital organisms to coral reef ecology as they are significant contributors to calcification of the reefs, as well as inducing larval settlement of many benthic organisms (Fabricius and De'ath, 2001; Lee, 2018).

The Rhodophyceae are also of significant commercial importance, accounting for approximately 61 % of all aquaculture production of aquatic plants, with two important hydrocolloids, agar and carrageenan, being primarily obtained from *Gelidium* and *Gracilaria*, and *Kappaphycus* and *Eucheuma*, respectively (FAO, 2018; Lee, 2018). Genera such as *Porphyra* are important farmed red marine macroalgae for direct human consumption (FAO, 2018).

Of all the major algal clades, marine macroalgal members of the Chlorophyta remain poorly understood in comparison to Rhodophyta and Phaeophyceae, with only one whole genome having been sequenced (De Clerck et al., 2018) and no significant commercial use as of yet (FAO, 2018). The Chlorophyta is an exciting clade with much potential for commercial applications including biofuels, food, feed, biofiltration, bioremediation, bioindicators, medicines, as well as being an important model organism to study the rise of multicellularity.

The Chlorophyta, or green algae, is a phylum of primarily freshwater algae, with around 10 % of its total species being marine (Lee, 2018). The Chlorophyta are one of two clades within the Viridiplantae, the other being the streptophyta, which encompasses all embryophytes (land plants) in addition to a number of other organisms such as the Charophyceae (Smith et al., 2009). Chlorophyte taxonomy is a dynamic field, with new information regularly changing classifications. Several reviews of chlorophyte taxonomy over the past ten years have proposed changes in classifications, including the creation and removal of classes and inter-phyta changes (Becker and Marin, 2009; Lee, 2018; Leliaert et al., 2011; Lewis and McCourt, 2004; Marin, 2012; van den Hoek et al., 1995). However, the three relatively

undebated core classes within the Chlorophyta are considered to be the Chlorophyceae, Trebouxiophyceae and Ulvophyceae (Cocquyt et al., 2010; Leliaert et al., 2011; Marin, 2012; Wichard, 2015).

All of the chlorophytes derive their characteristic vibrant green colour from the presence of chlorophylls a and b as their primary photosynthetic pigments, with no accessory pigments masking them (Lee, 2018; van den Hoek et al., 1995). Chlorophyta store reserve polysaccharides in the form of starch granules around the pyrenoids and in the chloroplast stroma (Lee, 2018; van den Hoek et al., 1995).

The Chlorophyta show significant diversity in morphology, sometimes even within species, and life histories, ranging from unicellular species to colonial unicellular and multicellular species. They can be uninucleate or multinucleate (Lee, 2018), with a well-known example demonstrating the unique range of Chlorophyta morphology being the *Caulerpa* genus, which contains the largest unicellular organisms in the world, regularly growing several feet long (Jacobs, 1994; Lee, 2018). Despite being unicellular, *Caulerpa* can display astonishingly complex differentiation of structure, developing leaves, stems and roots (Jacobs, 1994). Many species of the Chlorophyta reproduce vegetatively by fragmentation and can also form zoospores to reproduce asexually (Lee, 2018). The Chlorophyta sexual reproduction can be isogamous, anisogamous or oogamous, and gametes in isogamous species are not formed in specialised cells, while in oogamous species they are formed in specialised gametangia (Lee, 2018).

The Chlorophyta have been less exploited for direct human use than the Rhodophyta and Phaeophyceae, accounting for less than 4 % of all aquaculture production of aquatic plants. This is an upper figure and the actual figure is likely to be significantly lower as this is an estimate for all the species unaccounted for by the “State of World Fisheries and Aquaculture” report published by the FAO, with no chlorophyte genera or species being directly named (FAO, 2018). However, this reflects the lack of exploitation of the Chlorophyta, not any lack of potential. With advances in aquaculture management, many commercially important uses for the chlorophytes have been proposed, including as part of an integrated aquaculture system for bioremediation or biofiltering, bioindication and as feed (Amosu et al., 2015; Bolton et al., 2009; Yokoyama and Ishihi, 2010). *Monostroma latissimum* is the most commercially cultivated genera of the Chlorophyta, mostly in eastern Asia (Korea, Taiwan, the Philippines, China and Japan), although it is also cultivated in South America

(Pereira, 2016). *Monostroma latissimum*, like most members of the Chlorophyta, are still only exploited for direct human consumption. *Ulva intestinalis* and *Ulva lactuca* are also often bundled together with *Monostroma latissimum* when discussing culinary uses, and local names often refer to all of these species (Korean: *parae*; Japanese: *aonori*) (Pereira, 2016).

It is worth noting that macroalgal taxonomy, and algal taxonomy in general, is an extremely dynamic field, especially with the advent of molecular data aided classification. As such, even two well regarded textbooks “Algae: an introduction to phycology” by the Dutch phycologist Christiaan van den Hoek, and “Phycology” by the American phycologist Robert Edward Lee, show various difference in classifications despite their relative closeness in publication dates (first editions published in 1978 and 1980, respectively). This is further evidenced by various debates regarding how to classify clades, such as the heterokont (Cavalier-Smith, 1986) or stramenopiles (Patterson, 1989) clades which are both used interchangeably (Derelle et al., 2016) and as different levels of classification (Brown and Sorhannus, 2010). Given the relative infancy of many modern taxonomic methodologies, and the consequent volatility of the field, taxonomy of marine macroalgae should be of contextual and supporting information, rather than defining in most cases.

1.6 *Ulvophyceae*

Within the Chlorophyta, the seaweeds of the *Ulvophyceae* are of immediate ecological, environmental and economic interest. *Ulvophyceae* encompass all commercially cultivated green marine macroalgae, but we will follow the genus *Ulva* as an example. *Ulva* presents significant potential for exploitation as well as a significant problem. *Ulva* has a cosmopolitan range, having been observed in every ocean and estuary in the world, though the highest concentrations and species diversity are found mainly in temperate waters (Lee, 2018; Messyasz and Rybak, 2011; OBIS, 2019; Tanner and Guiry, 2019; van den Hoek et al., 1995). *Ulva* is able to rapidly respond to enhanced nutrient supply, for example following upwelling, but this characteristic is becoming more problematic as anthropogenic eutrophication increases (Lee, 2018). This also means that the potential for Green tides is also cosmopolitan, and are now widespread in the intertidal zone along much of the world’s coastlines, especially in temperate, nutrient-enriched waters (Raffaelli et al., 1998;

Schramm and Nienhuis, 2012; Teichberg et al., 2010; Valiela et al., 1997). The tendency of *Ulva* to grow rapidly in a range of environmental conditions, and rapidly proliferate in response to anthropogenic nutrient supplementation likely underlies their increasing appearance in Harmful Algal Blooms, or Green tides. However, the main advantage of this characteristic is that it also demonstrates the capacity of *Ulva* to be rapidly grown for various commercial uses. This gives *Ulva* a unique status, as both a problem and a solution.

The use of *Ulva* as a solution is increasingly more viable and may even become essential. As anthropogenic climate change alters the global climate, the biogeography of organisms limited by climate also changes (Molinos et al., 2017). Climate change and anthropogenic eutrophication are thought to be the primary drivers behind the increase in the frequency and scale of Harmful Algal Blooms (IPCC, 2013). It is only since the early 2000s that there have been reports of unprecedented and repeated *Sargassum* blooms along the coasts of the West Indies, the Caribbean, Brazil and West Africa (Djakouré et al., 2017; Sissini et al., 2017). Green tides have only been a significant recurring problem in the South China Sea and Yellow Sea since 2008 (Gao et al., 2017; Zhou et al., 2015). *Eucheuma* and *Kappaphycus* infestation with ice-ice disease and *Porphyra* infestation by suminori disease have been associated with increased temperatures, and taken with current temperature projections and its potential to significantly damage coastal economies, this is a major source of concern for the world's aquaculture (Arasamuthu and Edward, 2018; IPCC, 2013; Kusuda et al., 1992; Largo et al., 1995; Largo, 2002). It is undeniable that climate change threatens devastating damage to the world's aquaculture, as well as to coastal ecosystems and could significantly affect the potential of the marine biotic carbon sink, which represents roughly 3 Gt C of the 1,020 Gt C stored in the surface ocean and continues to sequester 0.173 Gt C yr⁻¹ (Siegenthaler and Sarmiento, 1993).

1.7 Plant Morphogenesis and Early Development of Tissue

Current thinking holds that multicellularity may arise in either of two ways: clonal, in which a multicellular organism arises from the division of a single cell, and aggregative, in which multiple cells of genetically distinct lineages aggregate to form an organism (Grosberg and Strathmann, 1998, 2007). Genetic diversity arises from somatic mutation, intergenotypic

fusion or pathogen infection in clonal multicellular organisms and from the association of genetically distinct lineages (Grosberg and Strathmann, 2007; Michod and Roze, 2001). Clonal multicellularity often involves a unicellular phase of a sexually produced zygote, which is hypothesised to eliminate deleterious mutations and control parasitic replicators (Grosberg and Strathmann, 1998).

All multicellular organisms have evolved from unicellular organisms, and the competitive advantages that multicellularity offers may explain why it has independently evolved multiple times, throughout the kingdoms animalia, plantae and fungi, and multiple times within some phyla (Claessen et al., 2014; Kaiser, 2001; King, 2004; Rokas, 2008; Whitaker, 1969). Though simple multicellularity has evolved multiple times in various lineages, complex multicellularity is only seen in six clades, namely metazoans, embryophytes, florideophytes, laminarialean brown algae, ascomycetes and basidiomycetes (Knoll, 2011). Complex multicellularity here refers to the presence of cell to cell adhesion, intercellular communication and tissue differentiation by top-down controls.

It was once thought that transitioning from unicellularity to multicellularity required a significant change in the genome. Prochnik et al. (2010) found in their genome study of *Volvox carteri*, a multicellular green volvocine alga of the Chlorophyta clade, that the genome is remarkably similar to *Chlamydomonas*, a unicellular volvocine alga, thereby demonstrating that it is possible for multicellularity to arise without significant changes in genome content or protein domains (Merchant et al., 2007; Prochnik et al., 2010). This demonstrated that simple multicellularity can arise with relatively little difficulty. However, the difference in genome size between simple and complex multicellular organisms is vast, with *Physcomitrella* and *Arabidopsis* genomes containing three and four times more protein domains respectively, than *Volvox* or *Chlamydomonas* (Prochnik et al., 2010). This, then, suggests that the evolutionary challenge lies in the transition from simple to complex multicellularity, rather than from unicellularity to multicellularity.

Chlorophyte algae provide a unique opportunity to study the rise of multicellularity, as they encompass a variety of morphologies and life histories. Of the Chlorophyta, Ulvophyceae, and the genus *Ulva* in particular, are especially intriguing as they display unparalleled morphological and cytological diversity (Cocquyt et al., 2010). The four main cytomorphological types of Ulvophyceae are uninucleate unicellular, uninucleate multicellular, such as Ulvales, multinucleate unicellular, such as Bryopsidales, which includes the aforemen-

tioned *Caulerpa* and multinucleate multicellular, such as Cladophorales (Cocquyt et al., 2010; Lee, 2018). *Ulva* also demonstrates an early form of clonal multicellularity, without differentiation of cells into tissues. Knowing the example of *Volvox*, one could also expect that *Ulva* will not show significant changes in genome content or protein domain when compared with *Chlamydomonas*. This was shown to be true by De Clerck et al. (2018).

Plant growth coordination usually involves both top-down and bottom-up controls. For example, the auxin gradient that result in root shoot differentiation is a top-down control, and phosphorus dependent root hair density is a bottom-up control (Ma et al., 2001; Overvoorde et al., 2010). The genomic differences begs the question: how is multicellular growth coordinated in organisms that lack top-down control? It is likely that the answer will encompass a number of different bottom-up controls, the specific controls depending on the unique characteristics of each organism. However, there is a strong case to be made that in simple multicellular organisms, physical constraints may be a strong driver of cellular organisation as there is little specialisation of tissue.

1.8 Morphogenesis and Early Development of *Ulva*

1.8.1 Morphology of *Ulva*

Of all the Chlorophyta, the best known and most studied macroalgae are those in the *Ulva* genus. *Ulva* is a genus of multicellular green algae present in many coastal benthic zones worldwide; it is phylogenetically close to unicellular green algae such as *Chlamydomonas* and *Ostreococcus* (Ansell et al., 1998; Xu et al., 2012; Zhang et al., 2012). The best known *Ulva* species in the UK is probably *Ulva lactuca*, commonly referred to as “sea lettuce” due to its similarity in appearance to lettuce leaves, as shown in figure 1.4 below.



Figure 1.4: A typical *Ulva* thallus displaying the rosette morphology (Novák, 2018)

Ulva was originally one single genus classified by Linnaeus in 1753, but in 1952 was split into distinct genera, namely *Ulva* and *Enteromorpha*, based on morphological characteristics, and the genus *Chloropelta* was later separated from *Ulva* by Tanner (1980), containing a single species *Chloropelta caespitosa* Tanner (Hayden et al., 2003; Linnaeus, 1753; Silva, 1952; Tanner, 1980).

Ulva species identification was once commonly done using morphological characteristics, however there was frequent and steady doubt cast on the reliability of these morphological attributes in *Ulva* taxonomy due to the extreme intraspecific morphological variation (Malta et al., 1999; Mshigeni and Kajumulo, 1979). *Ulva spp.* morphology is extremely plastic and even within species, thalli can display different morphologies; this process is thought to be environmentally determined (Hoeksema and van den Hoek, 1983; Malta et al., 1999; Mshigeni and Kajumulo, 1979; Wichard, 2015).

Morphological variations in *Ulva* can be so extreme that three genera were recently reclassified by a detailed phylogenetic study analysing the nuclear ribosomal internal transcribed spacer DNA (ITS nrDNA), the chloroplast-encoded Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RuBisCO) (*rbcL*) gene, and a combined *rbcL* and ITS nrDNA data set (Hayden et al., 2003). *Ulva*, *Enteromorpha* and *Chloropelta* were all found to be of the genus *Ulva*. Thus, *Enteromorpha* and *Ulva* were reclassified as the same genus *Ulva*,

and *Chloropelta caespitosa* Tanner was also reclassified as *Ulva*, with the new name *Ulva tanneri* H.S.Hayden & J.R.Waaland. Malta et al. (1999) even found that three *Ulva* species found in South West Netherlands previously recognised by taxonomic identification of morphology are in fact one polymorphic species, as well as finding that two samples from different collection locations previously thought to be of the same species were actually different species. These findings remind us that *Ulva* can show extreme variations in morphology and is even more difficult to identify when proliferating as free-floating thalli, as much taxonomic identification emphasizes variation in basal regions of the thallus (Hoeksema and van den Hoek, 1983; Koeman and van den Hoek, 1981).

Ulva only has three cell types (blade, rhizoid and stem) arranged into two opposing one-cell thick sheets, without differentiation into tissues and as such, there are no active mechanisms of cell signalling or overarching cellular architecture coordination within the plant itself (Wichard, 2015).

Ulva thalli are often referred to as “blades” and these blades can reach more than one metre in length (Sahoo and Seckbach, 2015). Cells in *Ulva* thalli are uninucleate, contain a single cup shaped chloroplast at the surface edge and a single pyrenoid, isodiametric in shape, and are distromatic (arranged into a bilayer), as shown in figure 1.5 below (Arora and Sahoo, 2015; Wichard, 2015).

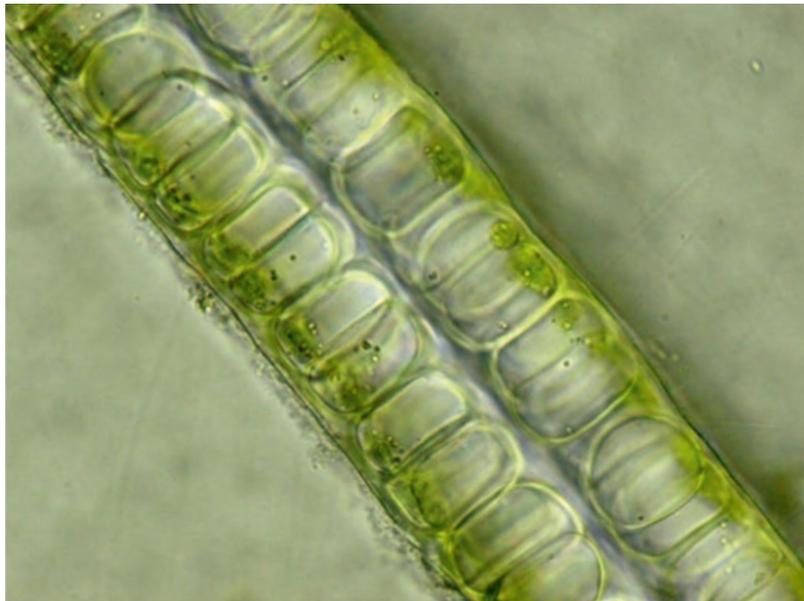


Figure 1.5: Longitudinal section of an *Ulva* thallus (Fernández in Arora and Sahoo, 2015)

The extreme intra-specific variations in morphological features as well as the isomorphism of distinct *Ulva* species and the lack of genomic architecture that gives rise to these morphologies indicate that there is a more basic and universal mechanism that determines morphology in the genus *Ulva*.

1.8.2 Life History of *Ulva*

All species of *Ulva* undergo alternation of generations between haploid gametophytes and diploid sporophytes, as shown below in figure 1.6 (Morrissey et al., 2016; Wichard, 2015). Interestingly, *Ulva* gametophytes and sporophytes show extremely similar sizes and morphologies; the reasons behind this isomorphic alternation remain a matter of debate. Sporophytes of a generation will produce zoospores that go on to form gametophytes of the next generation. These haploid gametophytes will go on to produce haploid gametes that fuse to form a diploid gamete and develop into a sporophyte, closing the cycle. *Ulva* zoospores have four flagella (quadriflagellate), and haploid gametes have two (biflagellate), allowing for some motility (Morrissey et al., 2016). Unfused gametes can undergo parthenogenetic development into gametophytes and the cycle continues, indicated with red arrows below in figure 1.6.

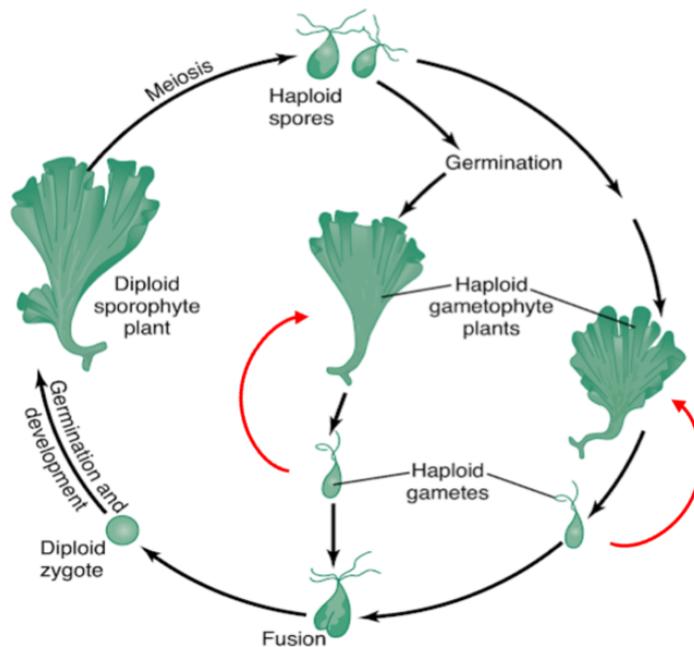


Figure 1.6: *Ulva* alternate between gametophytic and sporophytic generations. Red arrows represent the parthenogenetic development of unfused gametes into gametophytes (Modified from Morrissey et al., 2016)

Major disturbances, such as fragmentation, induce the formation of sporangia and the release of spores in *Ulva*, peaking approximately two to three days after the disturbance event both in lab and field settings (Carl et al., 2014; Dan et al., 2002; Gao et al., 2010, 2017; Hiraoka and Oka, 2008). This reproductive response to fragmentation and Sporulation Inhibitors may be key to *Ulva* bloom formation and expansion, as well as for the future of *Ulva* cultivation for maximum rates of biomass production.

The rate of *Ulva prolifera* zygote and meispore germination is approximately three times greater than gamete germination through parthenogenesis, indicating that asexual reproduction through fusion and sexual reproduction through meiosis is the key driver of *Ulva* proliferation, not parthenogenesis (Cui et al., 2018).

Vegetative cells in an *Ulva* thallus can directly differentiate into gametangia or sporangia as shown in figure 1.7 below (Carl et al., 2014; Dan et al., 2002).

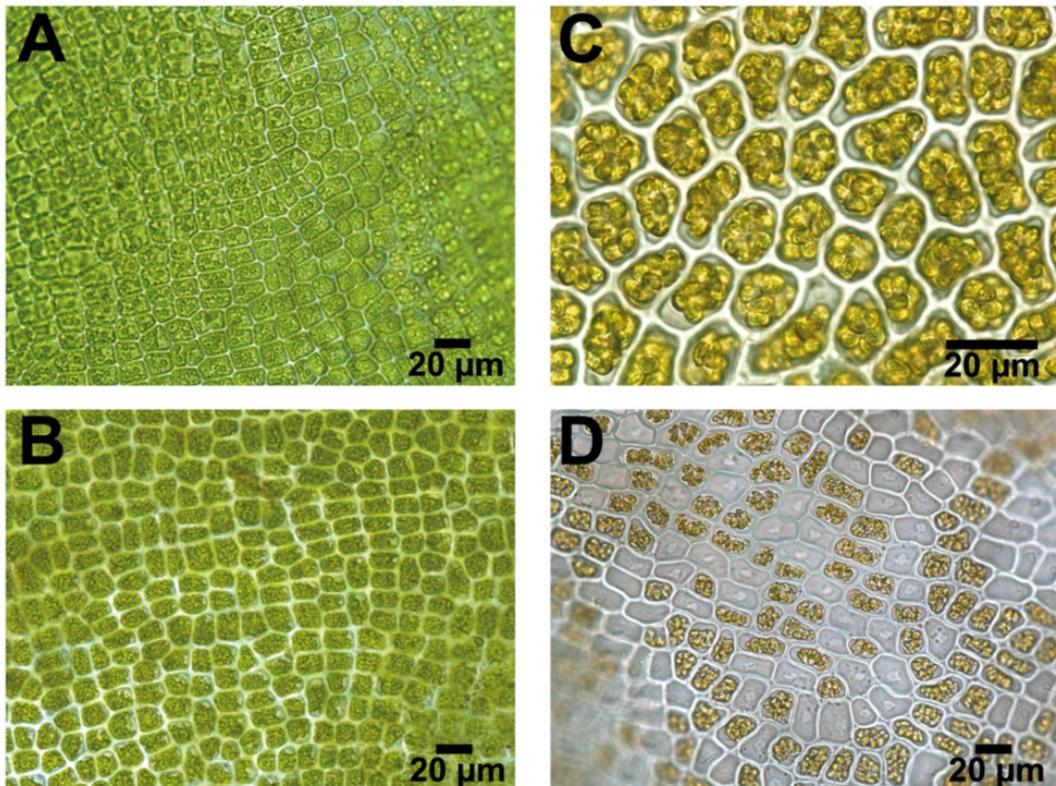


Figure 1.7: Light Microscopy images of the differentiation of vegetative cells into reproductive cells in *Ulva*. (A) Vegetative Cells, (B) Formation of gametes after 26 hours, (C) Reproductive cells with fully formed gametes after 44 hours, (D) Reproductive cells with fully formed and released gametes after 46 hours (Carl et al., 2014)

Ulva can grow anchored or free floating, and massive proliferation events, known as green tides as previously discussed in section 1.4, occur when the environmental conditions allow for the development of free-floating *Ulva*. *Ulva* is an opportunistic macroalga that quickly dominates the marine floral composition following disturbance events as well as nutrient enrichment events such as eutrophication, with or without disturbance (Kraufvelin, 2007). This success of *Ulva* as a primary succession macroalga can be credited to a number of factors, including the extreme tolerance and rapid adaptability to temperature, light intensity and UV exposure (Cui et al., 2015; Rautenberger and Bischof, 2006) and salinity variations (Shimada et al., 2008), and the ability to maintain extremely high growth rates of up to 54.5% per day (Hiraoka and Oka, 2008), all of which contribute to the cosmopolitanism of *Ulva* (Kirkendale et al., 2013; Shimada et al., 2008).

Ulva is known to produce and utilise at least two types of sporulation inhibitors: a cross-linked glycoprotein, Sporulation Inhibitor-1 (SI-1), and a class of non-protein sporulation inhibitors (SI2), of very low molecular mass and as yet unknown molecular structure (Stratmann et al., 1996). SI-1 levels are found to be elevated in cell walls and are also detected in significant amounts in the extracellular environment. SI-1 inhibits the differentiation of vegetative cells into gametangia until the thallus is fertile. The production of SI-1 in *Ulva* gradually decreases as the thallus matures, until the concentration of SI-1 falls below inhibitory levels, at which point gametogenesis is induced. The complete artificial removal of SI-1 from the extracellular environment also induces gametogenesis (Wichard and Oertel, 2010).

The decrease in the concentration of a Swarming Inhibitor (SWI) is required for the gamete release of *Ulva mutabilis*, though light is the only prerequisite for gamete release in *Ulva lactuca* (Stratmann et al., 1996; Wichard and Oertel, 2010).

Ulva is a symbiotic organism, and *Ulva* zoospores settle preferentially at sites where N-(3-oxododecanoyl)-homoserine lactone (AHL), a bacterial quorum sensing signal molecule, is detected and the absence of AHL abolishes the *Ulva* zoospores attraction to the bacterial biofilms (Joint et al., 2002, 2007; Tait et al., 2005; Wheeler et al., 2006).

1.8.3 Symbiosis in *Ulva*

When we consider the conditions required for the normal development of *Ulva*, an increasingly important area of study is on the various symbioses that are characteristic of this genus. For example, as aforementioned, N-(3-oxododecanoyl)-homoserine lactone (AHL) is a bacterial quorum sensing signal molecule that induces *Ulva* zoospores to settle preferentially, acting as a strong chemoattractant and the absence of which abolishes the *Ulva* zoospores attraction to the bacterial biofilms (Joint et al., 2002, 2007; Tait et al., 2005; Wheeler et al., 2006).

Furthermore, dimethyl sulfoniopropionate (DMSP) is a known osmolyte in *Ulva*, and along with its cleavage products, is already known to have multiple functions, acting as an anti-herbivore defense, a reactive oxygen species (ROS) scavenger and a cryoprotectant (Karsten et al., 1992; Nishiguchi and Somero, 1992; Van Alstyne et al., 2001; Van Alstyne and Puglisi, 2007). Recent evidence suggests that in *Ulva*, DMSP also has another function as a bacterial chemoattractant, which acts to signal bacteria of the presence of a food source, in this case glycerol as a carbon source (Kessler et al., 2017).

These relationships with bacteria alone, strongly suggest that *Ulva* is a symbiotic organism. The symbiotic nature of *Ulva* was found to be irrefutable as normal morphogenesis in *Ulva* requires cross-kingdom signalling between itself and bacteria. In the absence of associated bacteria, *Ulva* will grow as undifferentiated lumps that display atypical morphology (Burke et al., 2009; Joint et al., 2002, 2007). More specifically, growth becomes slow and callus-like, with colourless exterior cell wall protrusions (Spoerner et al., 2012; Wichard, 2015). Spoerner et al. (2012) set about identifying the exact bacteria involved in this symbiotic relationship by isolating the bacterial strains associated with non-axenic *Ulva* cultures. From the isolates, three strains (MS1, MS2, MS3) were found to induce growth. However, individual strains on their own could not induce normal morphogenesis; the thalli grew as they would in axenic cultures, but with increased rates of growth and a larger overall size. Revealingly, when *Ulva* was cultured with any of MS1, MS2 or MS3 as well as the seemingly inactive strain MS6, normal morphogenesis was restored, even though cultures grown only with strain MS6 did not show visible thallus growth. Cultures containing strains MS1 and MS6 eventually saw a drop in pH, growth suppression and bleaching of the thalli, but this was not the case with the MS2/MS6 or MS3/MS6 combinations. On closer inspection,

strain MS2 was found to be concentrated around the base of the thallus in the growth medium, and not diffuse like MS3, so it was hypothesised that strain MS2 forms a closer symbiotic relationship as this location specificity may indicate further specific signalling.

The associated bacterial epiphytes were originally classified as *Roseobacter* sp. Strain MS2 and *Cytophaga* sp. Strain MS6. A spatial separation experiment demonstrated that *Halomonas* MS1, *Roseobacter* sp. MS2 and *Sulfitobacter* MS3 excrete specific regulatory factors into the medium that induce enhanced cell division and formation of an *Ulva* thallus, much like cytokinin in higher plants. *Cytophaga* sp. MS6 excretes a factor that induces enlargement and stretching of newly divided algal cells, much like auxin in higher plants, which resulted in blade cell expansion and rhizoid growth, as well as eliminating the cell wall abnormalities observed in its absence (Spoerner et al., 2012).

The morphogen excreted by *Maribacter* MS6 is thallusin, the structure of which, illustrated below in figure 1.8, was first elucidated by Matsuo et al. (2005), having isolated it from the epiphytic marine bacterium strain YM2-23 (deposition no. MBIC 04683) obtained from *Monostroma oxyspermum* (Weiss et al., 2017).

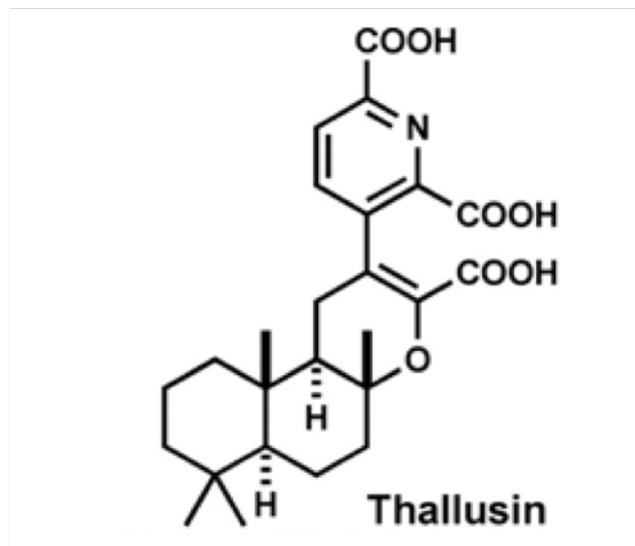


Figure 1.8: Molecular structure of thallusin (Matsuo et al., 2005). Thallusin induces rhizoid formation and normal cell wall synthesis without the unusual callus-like colourless protrusions observed in axenic cultures (Spoerner et al., 2012; Wichard, 2015).

Following further phylogenetic assessment, strains *Roseobacter* MS2 and *Cytophaga* MS6 have now been reclassified as *Roseovarius* sp. MS2 and *Maribacter* sp. MS6 respectively (Grueneberg et al., 2016). A reliable, reproducible model system of a tripartite symbiotic community has therefore been identified (Grueneberg et al., 2016; Weiss et al., 2017;

Wichard, 2015). This *Ulva* holobiont requires a chemosphere that includes morphogenetic compounds, bacterial chemo-attractants, carbon sources, quorum sensing signal molecules and metallophores (Theis et al., 2016; Wichard, 2015).

1.9 The Hologenome Theory of Evolution: Applications to *Ulva*

The conclusion of the previous section referred to the symbiotic community of *Ulva* and its associated bacteria as the *Ulva* holobiont. The term ‘holobiont’ refers to the organisms, the host and all its symbionts, involved in a symbiotic relationship wherein the host and symbiont have undergone, and are undergoing natural selection as a single unit (Rosenberg and Zilber-Rosenberg, 2018). The symbionts in this case encompasses the entirety of the endocellular and extracellular microbiome. First proposed by Rosenberg et al. (2007) in regards to the coral holobiont, the hologenome theory of evolution aimed to address the lack of a nuanced framework to discuss the complex and interconnected nature of host-symbiont interactions, development and evolution.

The hologenome theory recognises a holobiont and its hologenome as a distinct organisational level, upon which evolutionary selection operates, aiding the clarity of lexicon when discussing host-symbiont evolution and co-evolution (Theis et al., 2016). The hologenome theory most critically addresses all plants and animals as holobionts, as they all are inextricably associated with substantial and diverse microbiota (Rosenberg and Zilber-Rosenberg, 2018). There has been much discussion regarding the merits of the concept of the hologenome, with more than 106 papers having been published in the past 12 years. There have been both support for and legitimate criticisms of the theory that have shaped the development of the concepts, leading to the current framework for understanding (Rosenberg and Zilber-Rosenberg, 2018).

It is known that *Ulva* has multiple symbionts that are essential to normal morphogenesis. Given the relative simplicity of the *Ulva* genome, it is reasonable to speculate that there are more as yet unknown symbionts that are essential to *Ulva* growth and reproduction. This makes it useful to apply the concept of the hologenome in order to investigate the nature of their relationships.

1.10 Summary

The chlorophyte seaweeds are of great economic and environmental importance, and *Ulva* are strong organisms with which to study this group. Specifically, the investigation of their growth, including their genetics, genomics and physiology underpins a more complete understanding of their large-scale spread in both controlled (farmed) and uncontrolled (green tide) conditions. There are a number of key gaps in the literature regarding the nature of growth in *Ulva*.

First, the genomic architecture underpinning multicellular growth in *Ulva* had not been addressed at the commencement of this project. With the publication of the study by De Clerck et al. (2018), the first *Ulva* genome is now available. Amongst the number of interesting findings from this investigation, the lack of a genomic signature for multicellularity is of the greatest interest for this thesis. However, the analysis in this publication was based on a number of assumptions that, whilst highly unlikely to be incorrect, were not technically supported by a robust statistical framework. This lack of statistical framework is the first issue to be addressed.

Second, given the lack of a genomic signature, the most plausible explanation for *Ulva* morphogenesis is a physical bottom-up mechanism. This is a gap in the literature that has not yet been addressed in any capacity, and this thesis sets out addressing not only the bottom-up mechanisms underpinning *Ulva* proliferation patterns, but also a framework for studying this pattern in other tissue types that show this type of growth, for example, metazoan and plant epithelial tissues.

Finally, the investigation of growth in *Ulva* could help elucidate the evolution of multicellularity and the transition from simple to complex multicellularity in the green lineage. There are surprising gaps in the fundamental biology underpinning macroalgal development. An area of particular concern is the lack of knowledge regarding the growth patterns and rhythms of macroalgae, both spatially and temporally. Spatially, the growth patterns, including the areas of growth and cell proliferation are as yet unknown. Temporally, the circadian growth patterns have not been able to be studied in *Ulva*, due to the lack of resolution in measurement techniques available. Both examples are fascinating in the context of the research of a potential model organism, that could provide useful information on the

evolution of multicellularity as well as the evolution of cell differentiation (Bonner, 2009; Hori et al., 1985; Wichard, 2015).

General Methods

2.1 Challenges in experimental work

Ulva is a difficult organism to research in a laboratory setting and presents a number of challenges when designing experiments, as the literature regarding established techniques and protocols is barren, when compared to more well studied organisms such as *Arabidopsis* and *Nicotiana*. Among the many challenges, three are most substantial and persistent in relation to the types of experiments conducted in this project.

The first major challenge is the culturing method. Unlike land plants for which there are incubators and growth rooms at every major research facility, *Ulva* requires specialist growth equipment. There is only one model system for *Ulva* growth, which is the aforementioned model tripartite community of *Ulva mutabilis*, *Roseovarius MS2* and *Maribacter MS6*. Even this system is only a few years old, with the underlying mechanisms having only been partially elucidated following decades of research.

The second major challenge is the monitoring of growth. Currently, the only method of growth monitoring is by physical removal of the tissue and weighing on scales. This presents a number of issues. Firstly, data obtained from different experiments or even the same experiment in different facilities is incomparable due to the lack of a standardised protocol and the inevitable inconsistencies in methods. The inconsistencies will inevitably give inaccurate measurements in some cases given that *Ulva* is an aquatic organism that has significant fluctuations in water content, and therefore mass, over relatively short periods of time.

The third major challenge is the lack of a standard fluorescent staining method. Again,

unlike well-studied land plants, there is no standard method of fluorescent staining and so protocols used for other species have to be extrapolated and adapted to suit *Ulva* specifically.

The following subsections in this general methods chapter will attempt to broadly address these major problems and their resolutions.

2.2 Sample collection

The samples used in this project were collected from beaches of Seaham, located in the North East of the United Kingdom (coordinates: 54°51'02.1"N 1°20'31.9"W). Figure 2.1 below shows representative samples collected for use in this project.

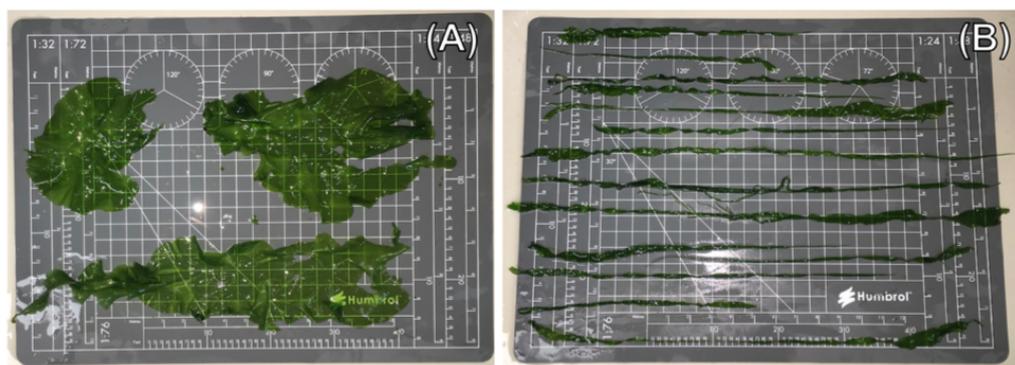


Figure 2.1: Typical samples of the (A) leaf and (B) ribbon morphotypes of *Ulva* used throughout this project

Both morphotypes have been sequenced, and the leaf and ribbon morphotypes have been identified as *Ulva lactuca* and *Ulva mutabilis*, respectively. However, for the purposes of the investigation of physical proliferation patterns, the morphology is of greater importance and interest, therefore this thesis will refer to each sample by their morphotype rather than their species name.

Sample collection was only conducted during low tide of less than 1 m, as higher tides would not allow for access to the samples. Samples are typically most healthy during the winter months when the sea surface temperature is relatively low. Samples collected during the summer months, especially towards the end of summer when the sea surface temperature is at its height, were typically not viable.

2.3 Sample culturing

Ulva tends to thrive in a variety of conditions, which is one of the reasons for its cosmopolitan range and proliferative growth that can result in bloom formation. For the duration of this experiment, constant culture conditions were maintained as shown in table 2.1 below.

Table 2.1: The growth parameters set for *Ulva* in the incubators and the justifications for their use throughout this project.

Parameter	Value	Justification
Temperature	10 °C	Mean sea surface temperature in the North Sea (Morris et al., 2018; Rayner et al., 2003)
Light	12:12 light/dark cycle	Standard light/dark cycle for plant growth experiments
Turbidity	100 RPM	Algem bioreactor environment simulation value for North Sea

Ulva is a symbiotic organism, and is known to require at least two symbionts (*Roseovarius* sp. MS2 and *Maribacter* sp. MS6) for normal morphogenesis and growth rates when grown in culture (Wichard, 2015; Grueneberg et al., 2016). However, the associated bacteria are present in collected wild samples, and the effects of the absence of these bacteria is immediately obvious in two ways: the growth of uncharacteristic lumps of undifferentiated cells that display atypical morphology and decreased rate of growth (Wichard, 2015; Spoerner et al., 2012). Further, the experiments that elucidated the presence of these bacteria involved sterile growth conditions, monitored by PCR tests, of carefully prepared axenic *Ulva* gametes (Spoerner et al., 2012). Our experiments only consisted of culture growth of adult wild samples and neither of the signs of bacterial absence were observed at any point. No difference in growth rate was observed between samples that were freshly collected and samples that had been growing in the lab for multiple weeks, with both sets of samples maintaining growth rates of approximately 40 % week⁻¹. Therefore, it was concluded that it can be safely assumed that the associated bacteria were present in culture.

The medium used throughout this experiment to culture the *Ulva* was artificial seawater. Preliminary experiments were conducted to test the suitability of filtered seawater, however the inconsistencies in results using filtered seawater led to the use of a standard artificial seawater mix. 1,000 ml of the artificial seawater used throughout was prepared with 1,000 ml milli-Q purified water, 35 g Sigma sea salt, 70 mg NaNO₃ and 10 mg β-Glycerophosphate disodium salt hydrate.

In order to characterise the culturing conditions and *Ulva* wet and dry mass measurements under these conditions, a simple preliminary growth trial was performed for *Ulva* by growing one hundred samples in culture and taking wet mass measurements of ten random samples (9 repeat measurements) and dry mass measurements of ten sample every day of the trial. The results are summarised below in table 2.2 below.

Table 2.2: The mean wet and dry mass, percentage change and water content of *Ulva* over five days in the laboratory culture conditions outlined above.

Day	Mean Wet Mass/g	Percentage Change	Mean Dry Mass/g	Percentage Change	Water Percentage
1	0.7225	-	0.1297	-	82.0
2	0.8444	16.9	0.1327	2.3	84.3
3	0.9011	6.7	0.1452	9.4	83.9
4	0.9322	3.5	0.1542	6.2	83.5
5	0.9989	7.2	0.2255	46.2	77.4
Total Change	0.2764	38.3	0.0958	73.9	-5.6

This showed that *Ulva* grows well in the culture conditions, showing a mean 38.3% increase in wet mass and a mean 73.9% increase in dry mass over the five days, equating to a daily growth rate of 8.5% in wet mass and 16.0% for dry mass. This is in line with previously reported growth rates in laboratory culture conditions (Mantri et al., 2011; Trivedi et al., 2013), but not as high as some have reported under specialist conditions utilising a technique called "germling cluster" (Hiraoka and Oka, 2008). However, this method cannot be replicated without access to equipment that is not available to us.

2.4 Fluorescence Microscopy

2.4.1 History of Fluorescence Microscopy

Fluorescence was first scientifically documented in 1565 by the Spanish physician Nicolás Bautista Monardes in his three-part publication "*Historia medicinal de las cosas que se traen de nuestras Indias Occidentales*" (Acuña and Amat-Guerri, 2007). Here, he describes the spectacular blue opalescence of a water infusion made from the wood of a tree that came to be known as *Lignum nephriticum*, which is now known to be *Eysenhardtia polystachya* or Mexican kidneywood (Acuña, 2007; Acuña and Amat-Guerri, 2007; Marcu et al., 2014; Valeur and Berberan-Santos, 2011). This blue fluorescence is reportedly

already well known by Aztec doctors as noted by 16th Century Spanish missionaries such as Bernardino de Sahagun and Francisco Hernandez (Acuña and Amat-Guerri, 2007). The fluorescent compound in this infusion was only recently identified as matlaline, named after the Nahuatl word “matlali” for blue (Acuña et al., 2009).

In the following centuries, the optical properties *Lignum nephriticum* became the subject of great interest. Many of the greatest minds of the 16th and 17th Centuries such as Francesco Maria Grimaldi, Athanasius Kircher, Sir Isaac Newton and Robert Boyle, studied these properties, paving the path for subsequent physicists such as to arrive at more complex, comprehensive and accurate descriptions of the physical properties of light and fluorescence (Acuña and Amat-Guerri, 2007; Marcu et al., 2014; Valeur and Berberan-Santos, 2011).

The renowned British physicist and the Lucasian Professor of Mathematics at the University of Cambridge Sir George Gabriel Stokes made the first substantial, accurate and notable contribution to the understanding of the physics of photoluminescence in his 1852 paper entitled “On the Refrangibility of Light” (Acuña and Amat-Guerri, 2007; Valeur and Berberan-Santos, 2011). Sir George Gabriel Stokes is credited with the introduction of the term “fluorescence” and describing “Stokes’ Shift”, which is the wavelength difference between the maxima of the excitation and emission spectra (Marcu et al., 2014; Valeur and Berberan-Santos, 2011). However, the French physicist Edmond Becquerel had already published a paper in 1842 noting that the emission wavelength is longer than the excitation wavelength (Valeur and Berberan-Santos, 2011), though his contributions are not credited to him, satisfying Stigler’s law of eponymy.

In the ensuing century and a half, fluorescence microscopy has emerged as an essential tool in various fields, and especially in life sciences due to the significant benefits over other microscopy techniques. The four main benefits of fluorescence microscopy offers are extremely high contrast, due to the absence of background, high specificity, due to the selective binding of stains, the quantitative nature of the fluorescence signals and the possibility of live imaging (Lichtman and Conchello, 2005; Muller, 2006).

2.4.2 The Principles of Fluorescence Microscopy

Fluorescence, simply, is the emission of, typically longer wavelength, light from an object following illumination by an incident light (Lichtman and Conchello, 2005). The difference

itself is due to the loss of energy from the incident light (Lakowicz, 2007), and is termed the Stokes' shift, as aforementioned. This means that any object that displays fluorescence, whether by autofluorescence or fluorescent labelling, can be excited by incident light of a certain wavelength and the emission of a different wavelength can be observed using a fluorescence microscope. The both the excitation and emission light are measured in nanometres (nm), and the relative intensity is plotted to give a spectrum of each. The wavelength of highest relative excitation and emission are termed the excitation and emission maxima. The observation of the emission spectrum is the principle of fluorescence microscopy in its simplest form.

However, many problems present themselves in the practical realisation of this principle. The first of which, is the overlap of the excitation and emission spectra. In practice, this is overcome by using a series of filters. In a traditional fluorescence microscope, these filters are typically the excitation filter, the barrier filter and the dichroic (dichromatic) mirror (Lichtman and Conchello, 2005). The purpose of these filters is best illustrated by following the light path from the incident source.

The incident light travels from the source, typically an arc lamp, first meeting the excitation filter, which filters the incident light to only allow light of the excitation wavelength to pass through. This filtered light then hits the dichroic mirror, placed at a 45° angle, which splits the light and reflects light of excitation wavelength downwards to the sample, whilst allowing the rest to pass through. The light then excites the sample and the emission light from the sample goes back up towards the dichroic mirror, which reflects any light of the excitation wavelength, and allows just the emission light to reach the barrier filter, which only allows the longer wavelength emission light through. The resulting emission light can then be collected by an image sensor to be viewed on a computer or observed directly through an eyepiece.

All of the filters used are highly wavelength-specific interference filters, which are composed of multiple stacks of thin alternating layers of different refractive index. The physics of this entire process is dependent on small differences in the wavelengths of light, meaning that specific filter combinations must be used according to the specific excitation and emission spectra of the fluorophore being used. Modern epifluorescence microscopes have filter sets or filter cubes, which allow for the convenient and rapid selection of prefabricated arrangements of filters for specific excitation-emission wavelengths.

Another principle limitation of fluorescence microscopy is that all of the emission light in the frame is recorded at once. This means that some of the emission light will be out of focus, limiting the resolution that can be achieved. This is a problem that is addressed by confocal microscopy.

2.4.3 History of Confocal Microscopy

In 1955 the renowned cognitive scientist and artificial intelligence expert Marvin Minsky, who was then a young postdoctoral fellow at Harvard University, invented a microscope that utilised a stage-scanning confocal optical system (Minsky, 1988). This microscope utilised sequential single-point illumination of an object to construct an image, with the incident light being focussed with a pinhole. The emission light is ensured to be from the focal point by utilising a second pinhole, thereby removing any out of focus light, and is collected by a photomultiplier to eventually construct the image on a long-persistence screen (Inoué, 2010; Semwogerere and Weeks, 2008). This was the genesis of confocal microscopy and sparked feverish excitement in the biological research community in the decades that followed. Many different approaches to the confocal microscope were made, such as the Tandem Scanning Microscope made by Petráň et al. (1968), which formed the basis of the modern disk-scanning confocal microscope (Toomre and Pawley, 2010). This project exclusively utilised Confocal Laser Scanning Microscopy, which is the modern incarnation of Minsky’s confocal microscope, and any mentions of “Confocal Microscopy” will be referring to Confocal Laser Scanning Microscopy. Despite the many approaches that researchers took towards confocal microscopy in the mid 20th Century, the word ‘confocal’ itself was not used until 1979 when Brakenhoff et al. (1979) built the Confocal Scanning Light Microscope, which is the predecessor of the Confocal Laser Scanning Microscope. The term itself refers to both, ‘con’, the focus point, ‘focal’, of the objective lens and the pinhole being the same (Peterson, 2010).

The advances in technology in the late 20th Century, including but not limited to the availability of high-powered lasers and powerful computers to analyse the vast sums of data generated, fuelled the development of confocal microscopy as one of the most important contemporary development in biological research. The international efforts led to the development of the Confocal Laser Scanning Microscope as we know it today at the MRC

laboratory of molecular biology at Cambridge University (Amos and White, 2003).

2.4.4 Principles of Confocal Laser Scanning Microscopy

A Confocal Laser Scanning Microscope builds on the same principles as previous incarnations of the confocal microscope, such as the Confocal Scanning Light Microscope, in that it also builds an image by illuminating and focusing on one point at a time. The Confocal Laser Scanning Microscope uses a high-powered laser, in place of an arc lamp, as it allows high-intensity light to be focused on one point at a time, rather than the whole field, and the emission signal is also amplified using a photomultiplier tube, which allows clear resolution of extremely faint signals. This combination allows for an extremely high contrast and high resolution, more specific detection and protects the sample from extended exposure to high intensity light (Inoué, 2010).

An extremely important development allowed by Confocal Laser Scanning Microscope is the construction of high-resolution z-stacks, as only one point is illuminated at any given time. This means that high-resolution true 3D images can be collected, which is an incredible advantage that is impossible to achieve with traditional fluorescence microscopy.

2.4.5 Stain Selection

The ratio of the number of photons emitted to the number of photons absorbed is termed the quantum yield (Φ) (Sommer et al., 2015). Though the term quantum efficiency is sometimes used in fluorescence microscopy to describe fluorophores (Laverdant et al., 2011), caution is advised as there are multiple definitions in multiple fields. For example, in physics it is associated with photosensitive devices such as solar panels (Hagfeldt et al., 2012) and image sensors (Fowler et al., 1998), in chemistry to describe photocatalytic systems (Serpone and Salinaro, 1999). Therefore, quantum efficiency would be a more suitable term in relation to the image sensor, rather than the fluorophore, in the case of fluorescence microscopy. Here, the only the term quantum yield will be used throughout, for clarity and consistency.

The quantum yield (Φ) is one of the two most important characteristic of a fluorophore when imaging. The higher the quantum yield, the brighter the emission of the fluorophore. The quantum yield is measured as a ratio; therefore, the highest value is 1, which is termed

as unity. For example, fluorophores such as rhodamines, which have some of the highest quantum yields, are said to be approaching unity (Lakowicz, 2007).

The second characteristic is the fluorescence lifetime of the fluorophore, measured in nanoseconds, which is simply the mean time taken for the fluorophore to return to the ground state from its excited state (Lakowicz, 2007). The differences in fluorescence lifetime of a fluorophore can be exploited for a specialised type of microscopy called Fluorescence Lifetime Imaging Microscopy (FLIM). In FLIM, images are created using fluorescence lifetimes rather than emission intensity (Bastiaens and Squire, 1999). Fluorophores with the same emission wavelength can be differentiated using their fluorescence lifetimes, meaning that different molecular environments within cells can be identified using FLIM (Bastiaens and Squire, 1999; Duncan, 2006; Suhling et al., 2015). These two characteristics of fluorophores are vital to note when selecting the type of microscopy. However, other considerations must be made regarding the characteristics of the fluorophore, that will affect the suitability of a given fluorophore in its intended use.

One of the great advantages of fluorescence microscopy is the ability to image with extremely high specificity due to the wide range of fluorophores available. For example, the *In Vitro* Molecular Probes Handbook published by Life Technologies Corporation, now in its 11th edition, includes over 3,000 fluorescent probes that can stain almost every component of a biological system (Johnson and Spence, 2010). There often are multiple probes of differing spectral properties available to label the same component, allowing for a tailored labelling solution for virtually any experiment that can range from simple single component labelling to multi-channel labelling of multiple components.

With the almost overwhelmingly large selection of fluorophores available, selecting a suitable fluorophore for a specific experiment can be difficult. Many dyes can be used for the same purpose, but that does not always mean that they are all equally suitable, and the suitability depends on the experimental aims. Different fluorophores have specific optical properties and bind specifically to different targets. However, even fluorophores that bind to the same target sites have unique properties to consider when assessing the suitability for use in a particular experiment. For example, DAPI (4', 6-diamidino-2-phenylindole) is a widely used DNA-specific fluorophore that binds to the minor grooves of A-T rich repeat sequences in DNA, and has also been shown to bind well to A-U rich repeated sequences in RNA, but by intercalation unlike in DNA (Kapuscinski, 1995; Lachner et al., 2001;

Tanious et al., 1992). The Hoechst family of dyes are also minor groove-binding dyes, with Hoechst 33342 and 33258 being the most commonly used (Telford et al., 1992). DAPI, Hoechst 33342 and 33258 dyes are all excited by blue light, and they all have an emission maximum at around 461 nm, and so they can be used almost interchangeably on the whole (Bestvater et al., 2002; Bucevičius et al., 2018; Portugal and Waring, 1988). However, Hoechst dyes and DAPI react differently under similar conditions, due to a one base pair difference in binding site size, which is an important consideration when investigating the effects of drug-DNA (Portugal and Waring, 1988).

Similarly, different fluorophore that stains the same molecule can have different mechanisms of doing so. For example, ethidium bromide and propidium iodide are further examples of DNA staining fluorophores, however these dyes both bind by intercalation, and therefore are mutagenic (Sudharsan Raj and Heddle, 1980; Ueckert et al., 1995).

Even within the same family of dyes, there can be certain differences that make one fluorophore favourable over another. The Hoechst family of dyes is a good example of this. Hoechst 33342 and 33258 binds the same molecule at the same site and have the same excitation emission maxima. However, Hoechst 33342 permeates cells significantly better than Hoechst 33258 due to a lipophilic ethyl group, and therefore is better suited for vital staining (Bucevičius et al., 2018). It is clear that selecting the right fluorophore for a certain experiment is crucial for the experimental results.

In order to investigate cell patterns in, a reliable cell wall staining protocol is required. However, there is no standard *Ulva* cell wall staining protocol, so I set out to establish one. The first dye to be explored was Calcofluor White M2R (Fluorescent Brightener 28), which is a very commonly used cell wall stain in biological research. Calcofluor White is a water-soluble, non-specific fluorochrome with a high affinity for β 1-3 and β 1-4 polysaccharides like those in cellulose and chitin, so has been steadily used in biological research for decades, though the specific component that it binds to has not yet been demonstrated with certainty (Harrington and Hageage, 2003; Harrington and Raper, 1968; Herth, 1980; Herth and Schnepf, 1980; Hoch et al., 2005). The structure of Calcofluor White is given below in figure 2.2.

2010; Hoch et al., 2005; Liesche et al., 2013; J. Thomas et al., 2013). Direct Yellow 96 can also be used as an alternative (Hoch et al., 2005; Ursache et al., 2018), however there is extremely limited information on the dye, especially in biological research. Pontamine Fast Scarlet 4B and Direct Yellow 96 both show three valuable traits, and two are clear advantages over Calcofluor White M2R. Both of the dyes, like Calcofluor White M2R, are apparently non-toxic to live cultures of various fungi (Hoch et al., 2005). Furthermore, Pontamine Fast Scarlet 4B and Direct Yellow 96 show higher intensity and longer duration of fluorescence when compared to Calcofluor White M2R (Hoch et al., 2005).

Pontamine Fast Scarlet 4B was the first candidate for staining trials and was prepared using deionised water and 1X PBS (0.01 M Phosphate buffer, 0.154 M NaCl, pH7.4), at concentrations ranging from 0.01% to 1%. Various permeabilisation and fixing steps were used to facilitate the dye to penetrate the tissue, however no combination of steps allowed for an even stain, reliably observable with a confocal microscope.

Direct Yellow 96 (CAS No. 61725-08-4), also known as C.I. Direct Yellow 96, Diphenyl Brilliant Flavine 7GFF and Solophenyl Flavine 7GFE 500, is a dye with many unresolved questions (Ismael et al., 2013; Sigma-Aldrich, 2012; Ursache et al., 2018). The molecular structure is illustrated in figure 2.3 below:

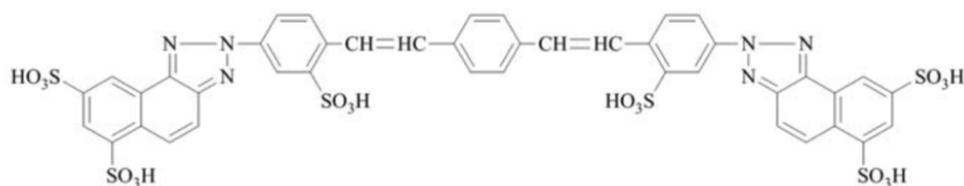


Figure 2.3: The chemical structure of Direct Yellow 96 (Ismael et al., 2013)

Direct Yellow 96 has been used as a dye for cotton and viscose fibres, and has been shown to dye cellulose and chitin so can be intuitively hypothesised that it binds to β 1-3 and β 1-4 polysaccharides, much like Calcofluor White M2R (Hoch et al., 2005; Ismael et al., 2013; Was-Gubala and Machnowski, 2014).

There currently is no firm consensus on the excitation and emission spectra of Direct Yellow 96. A number of resources, of which the only source that could potentially be considered reliable is the Fluorochrome Data Table by Olympus Life Sciences, indicate the same excitation of emission of Diphenyl Brilliant Flavine 7GFF to be 430 nm and 520 nm respectively (Abramowitz and Davidson, 2019). In his report “The QTRACER2 Program

for Tracer-Breakthrough Curve Analysis for Tracer Tests in Karstic Aquifers and Other Hydrologic Systems” prepared for the United States Environmental Protection Agency, Field, (2002) lists the maximum excitation and emission of Dipheyl Brilliant Flavine 7GFF to be 415 nm and 489 nm, respectively. Finally, Ursache et al., (2018) reported using a 488 nm laser to excite samples of *Arabidopsis* stained with Direct Yellow 96, achieving a maximum emission of 519 nm.

I performed a lambda scan on *Ulva* thalli stained with Direct Yellow 96 at the maximum resolution of 3 nm on a Zeiss 880 with Airyscan confocal microscope. This showed that Direct Yellow 96 bound to the cell walls of *Ulva* thalli is excited best by the 458 nm laser, which showed higher relative intensity than the 405 nm and 488 nm lasers. Though the excitation spectrum could not be tested due to the nature of the fixed wavelength lasers available, the maximum emission was found to be between 504 nm and 507 nm, regardless of the excitation laser (405, 458, 488 nm). The maximum emission was statistically closer to 504 nm than 507 nm, however this is simply an observation as there is a limit to inferring the exact wavelength without empirical data.

Regarding the toxicity and ecotoxicity of the dye, there is extremely limited information. The manufacturer Sigma-Aldrich lists hazard statements H315, H319 and H335, which refer to causing skin irritation, serious eye irritation and respiratory irritation respectively (Sigma-Aldrich, 2012). However, the vast majority of physical and chemical properties, stability, reactivity and toxicity and ecotoxicity information regarding Direct Yellow 96 has not yet been thoroughly investigated. Though there is extremely limited information regarding any of the properties, Direct Yellow 96 has been previously used as a tracer dye in some hydrological studies for groundwater tracing (Bayless et al., 1994; Thrailkill, 1985). Furthermore, Direct Yellow 96 was found to have no acute toxicity at or substantially above the usage concentration in groundwater tracing studies (Field et al., 1995).

After initial trials, the final staining protocol that I used throughout this thesis was as follows:

1. Solution Preparation

- 500 ml 1X PBS
- 150 ml 0.1% Triton X-100 diluted with 1X PBS

- 50 ml Ethanol : Acetic Acid (3 : 1)
- 50 ml FAA (Formaldehyde Alcohol Acetic Acid: 10% : 50% : 5% + 35% dH₂O)
- 2 ml Stock solution of 5 mg/ml Direct Yellow 96 made with 10X PBS (Store at 4 °C, protected from light).

2. Sample preparation

- Take desired number of *Ulva* thalli and wash in appropriate amount of 1X PBS (3 washes, 5 minutes each).

3. Permeabilisation

- Incubate in 0.1% Triton X-100.

4. Pigment Removal

- Incubate in a solution of Ethanol : Acetic Acid (3 : 1) for 1 hour, or until the leaves are colourless.

5. Fixation

- Fix in FAA overnight in centrifuge tubes or appropriate glassware with a lid.

6. Wash

- Wash in 1X PBS (3 washes, 5 minutes each)

7. Staining

- Prepare a working solution of 0.1 mg/ml Direct Yellow 96
- Incubate for 1 hour at 4 °C, protected from light

8. Wash

- Wash in 1X PBS (3 washes, 5 minutes each)

9. Image

- Keep samples protected from light unless actively imaging
- Samples can be stored in 1X PBS at 4 °C, protected from light for further imaging.

2.4.6 Scanning Parameters

Confocal image collection was performed primarily on a Leica SP5 Laser Scanning Confocal Microscope. After dozens of initial experimentation and optimisation trials, a standard set of parameters was chosen as a best compromise between image quality and scanning time. The scan area was chosen as 1024 x 1024 pixels, or 246.03 x 246.03 μm , with the images taken with a 63 X magnification water immersion lens, through a 111.4 μm (1 airy) pinhole.

For excitation of Direct Yellow 96, the 458 nm Argon laser was most effective, so this laser was used at 21% power. For detection, the HyD detector was used at 463 nm to 589 nm. The lower bound of 463 nm allowed detection of the Direct Yellow 96 signal without laser bleed, and the upper bound of 589 nm allowed images that effectively remove the signal from chlorophyll *a* and *b* fluorescence, which would have been excited by the 458 nm laser as they are excited by light around 400 to 700 nm, and emits mostly around 600 – 800 nm (Hägele et al., 1978; Kalaji et al., 2017, 2014; Pfarrherr et al., 1991). Furthermore, a lambda scan of *Ulva* autofluorescence showed that the vast majority of emission, of the presumed chlorophyll autofluorescence, was between 634 and 720 nm, with peak emission at 684 nm. A further lambda scan showed that the signal of Direct Yellow 96 is most intense from 474 to 522 nm with a peak at 504 and 507 nm. Therefore, the finalised detector settings allowed for the measurement of the vast majority of the emission signal from Direct Yellow 96, without laser bleed from the incidence laser or chlorophyll fluorescence.

The frame average was set to 8, which allowed for high-resolution images without the long scanning times of 16 averages, and the scanner was set to uni-directional scanning, to allow for higher resolution images, at a speed of 100 Hz. The scanning speed of 100 Hz produced high-resolution images without the significantly longer scanning times required by 10 Hz, which is the next step interval in scanning speed.

Modelling Gene Family Occupancy

3.1 Introduction

Ulva is a cosmopolitan genus of green alga that uniquely displays multicellularity in the Chlorophyte clade, and is found in both seawater and brackish waters (Lee, 2018; Messyasz and Rybak, 2011; OBIS, 2019; Tanner and Guiry, 2019; van den Hoek et al., 1995). Yet despite the cosmopolitan range, its potential as a biofuel feedstock and its nuisance nature as a bloom-forming alga, the literature on the *Ulva* genome, and chlorophytes in general, is extremely limited. Genomic studies into the Ulvophytes lag behind those in other macroalgae. Though there has been interest for decades, the earliest genomic studies of *Ulva* was published by Stanley et al., (2005). Since then, there have been a handful of *Ulva* transcriptome studies, including the investigation of C₃ and C₄ photosynthetic pathways (Xu et al., 2012), stress tolerance in intertidal zones (Zhang et al., 2012) and high temperature stress (Fan et al., 2017), with the majority of studies having been conducted in Asia because *Ulva*'s proliferative green tide growth comes at significant economic cost to the region (Jing, 2013b; Trainer and Yoshida, 2014; Yang et al., 2018).

The first full organellar genomes (chloroplast and mitochondrial) for *Ulva* sp. UNA00071828 were published by Melton et al., (2015) and the full *Ulva mutabilis* genome was published by De Clerck et al. (2018). One of the most striking discoveries by De Clerck et al. (2018) was that *Ulva* lacks 10 families of transcription factors and two families of transcriptional regulators present in other green algae. The most notable gene families that have been lost in the evolution of *Ulva* are the genes of the retinoblastoma / E2F pathway and associated D-type cyclins. Comparative genomic analyses of chlorophytes have shown that the

retinoblastoma pathway is a key step towards multicellularity in this lineage (Featherston et al., 2018; Hanschen et al., 2016; Prochnik et al., 2010), so the lack of the retinoblastoma pathway in *Ulva* is most surprising. Simply put, *Ulva* shows no genomic signature for multicellularity.

One of the assumptions made during the genome annotation and curation of *Ulva mutabilis* was that gene family occupancy follows a heavy-tailed distribution. This was due to the fact that there needs to be a known distribution in order to estimate whether particular gene families are expanded or contracted more than would be expected. Heavy-tailed distributions refer to distributions that have a heavier tail than an exponential distribution, so that they tend to zero more slowly than an exponential distribution. Common examples include the Pareto and Cauchy distributions, although the term “heavy-tailed” is often used more or less loosely depending on the subject area, with some fields regularly including distributions such as the log-normal and Weibull distributions in the same category as other heavy-tailed distributions (Bryson, 1974; Foss et al., 2013; Nair et al., 2013).

We can see, then, that systems that exhibit heavy tails can deliver periodic shocks of large values, or sometimes just a single extravagantly outlying value, while at least one of the four moments (mean, variance, skew and kurtosis), usually either the mean and variance, is infinite or does not exist (Adler et al., 1998; Nair et al., 2013). It should be stressed that this is a mathematical definition, as the mathematical criterion for a distribution to be heavy-tailed is that the moments of the distribution are not all finite, meaning that distributions such as the lognormal and Weibull, whose moments are all finite, are not considered heavy-tailed in mathematics (Foss et al., 2013). The conditions set out by the mathematical definitions of heavy tailed distributions affect the extremes of the distributions, which is rarely of concern in practice. The mathematical definition is not universally adopted by all academic fields, and many discuss the lognormal and Weibull distributions as heavy-tailed distributions (Adler et al., 1998; Foss et al., 2013; Nair et al., 2013). In practice, there is no authoritative definition of heavy-tailed distributions, and the general use definition is often regarding the absence of exponential moments.

The heavy tail means that the probability of finding a value many times higher than the central one is relatively high; this is often seen in headline-grabbing “wild statistics”, such as the fact that the richest individual in the world is worth more than 36 million times as much as the World’s median individual (Forbes, 2019; Shorrocks et al., 2018). This

stands in contrast to the “mild statistics” that arise from more normal distributions with, for example, the tallest known individual in history being only around 1.5 times taller than the mean adult male height in the United States of America (Fryar et al., 2016; Guinness World Records, 2019; Resnick, 2007).

Heavy-tailed distributions are often seen in biological “big data”. Most obviously, genomic data generally deals with thousands to hundreds of thousands of gene count numbers; high enough to be able to draw distributions from them, but not so high that they may be approximated as infinite. We will, therefore, use the relaxed definition of heavy-tailed distributions as there is little to be gained in labouring over the minute details of the definition of a term that whose definition varies between fields. More importantly, we are less interested in an exact description of heavy-tailed distributions and more interested in how we can use their general properties to improve our understanding of gene family analysis.

Specifically, heavy-tailed distributions may be generated by a number of different processes. We would like to be able to identify whether any particular class of process can generate the heavy-tailed curves that describe gene family occupancy in *Ulva* and in the green algae more generally. If we could do this, we may be able to more accurately investigate patterns of gene family and genome evolution in this important clade.

My first step in this chapter is, therefore, to consider the most common candidates for generating biological heavy-tailed distributions, which include the Poisson, exponential, log normal and power law distributions. I will consider each of these in turn.

Named after the French mathematician Siméon Denis Poisson, the Poisson distribution was first proposed in 1837, and is arguably the most important and commonly used discrete probability distribution (Box et al., 2005; Stigler, 1982). The Poisson distribution expresses the probability of a given number of events in a fixed period of time, given a known average rate and independence from preceding events, and is applicable to systems in which there are a large number of outcomes and each outcome has a small probability of “success” (Box et al., 2005; Stigler, 1982).

The Poisson distribution has the probability mass function shown in equation 3.1 below.

$$P(X = c) = \frac{\lambda^c e^{-\lambda}}{c!} \quad (3.1)$$

where $c \in \{0, 1, 2, \dots\}$

and $\lambda > 0$

Although the Poisson distribution was initially used by Poisson himself to analyse data on convictions in France (Hacking, 1990; Maltz, 1996), it is now widely used throughout a plethora of academic fields including nuclear physics (Cannizzaro et al., 1978), ecology (Young and Young, 2013), aerospace medicine (Gildea et al., 2018) and microbiology (Koyama et al., 2016). Given that genome data is discrete count data, the Poisson distribution is an obvious and immediate candidate for analysis.

The second distribution of interest is the exponential, which is another of the most widely used distributions in statistics. Indeed, as previously mentioned, the exponential distribution is the baseline against which the weight of a distribution's tail is measured in order to determine heavy-tailedness or light-tailedness. The exponential distribution is closely related to the Poisson distribution: the Poisson distribution describes the probability of the number of events occurring in fixed time, whilst the exponential distribution describes the time elapsed between such events. As such, while this is a useful model, it is irrelevant to the particular line of inquiry in this investigation and will only be included for completeness of comparison.

The third candidate distribution is the lognormal, developed by the British statistician Francis Galton and physician Donald McAlister as a method of accounting for the limits of a normal distribution, while ignoring random variability and its effects (Hart, 1990; Limpert et al., 2001). The lognormal distribution has a few general characteristics, first and foremost of which is that, given a random variable x , its logarithm $\log(x)$ is normally distributed, hence the etymology of the term (Limpert et al., 2001).

Accordingly, the lognormal distribution has the probability density function show in equation 3.2:

$$p(x) = \frac{1}{x \cdot \sigma \sqrt{2\pi}} e^{-\frac{(\log(x) - \mu)^2}{2\sigma^2}} \quad (3.2)$$

given $x > 0$

The lognormal distribution is often discussed as an alternative to a power law distribution, and scholars from disciplines varying from biology (Hoffmann et al., 2007; Jain and Ramakumar, 1999) and geophysics (Nakajima and Higurashi, 1998) to economics (Mandelbrot, 1997) and hydrology (Bowers et al., 2012), have had many a discussion as to which distribution is a better fit for a particular set of empirical data, often compromising between the two without a definite solution (Mitzenmacher, 2004).

Finally, power law distributions are a class of distributions that use the basic concept of power law relationships in which, given two variables, the value of one variable is generated as a power of the value of the other variable. Power law distributions are pervasive throughout the natural world, as well as being often seen in the social sciences and humanities. For example, power law distributions are often used to describe economic situations in which a small number of individuals hold the most value, for example a ‘vital few’ earn vastly more than the ‘trivial many’ (Reed, 2001). The classic examples used to demonstrate power law distributions include wealth distribution (Klass et al., 2006), the size distribution of cities (Rosen and Resnick, 1980) and website traffic (Kumar et al., 2009) amongst others.

A power law relationship is mathematically described as shown in equation 3.8 below.

$$P(X \geq x) = cx^{-\alpha} \quad (3.3)$$

where $c > 0$

and $\alpha > 0$

Given the basic property of logarithms that

$$\log(ab) = \log(a) + \log(b) \quad (3.4)$$

We can take the log values of the generalised power law relationship and derive:

$$\log(y) = -\alpha \log(x) + c \quad (3.5)$$

Simply, if a dataset follows a power law distribution, plotting that dataset on a log-log plot gives a straight line.

The most commonly known and used power law distribution is the Pareto distribution which is often inaccurately called “the” power law distribution. The Pareto distribution is “a” power law distribution that occurs under specific conditions, and so it should be clearly noted that whilst all Pareto distributions are power law distributions, not all power law distributions are Pareto distributions. The concept of the ‘vital few and trivial many’ was introduced earlier but was intentionally left brief in order to first make this clear distinction, as the Pareto distribution is inextricably tied with the classical examples of power law distributions.

Named after the first proponent of the principle, Vilfredo Pareto, the Pareto distribution was originally published in 1906 and described the property ownership distribution in Italy, where 20 % of the population owned 80% of the properties (Pareto, 2014). This was the inception of the Pareto principle, also referred to as a multitude of other names including the principle of the vital many and trivial few, the 80/20 rule, Pareto’s law, Pareto’s rule and the law of factor sparsity.

In recent years many popular writers have, wildly incorrectly, taken the Pareto principle to supposedly represent a means, or “life hack”, capable of massively improving productivity, revolutionising marketing, radicalising management, changing one’s life, or even predicting the future (!). Its alter ego the “80/20 rule” is often mentioned in these cases, although 80/20 is a simple ratio restricted to cases in which the Pareto exponent gives rise to that ratio; this is common, but not universal. The pervasive nature of the Pareto pattern, especially in so many aspects of human culture and the lack of an all-encompassing fundamental explanation, often gives it an air of mysticism and laymen could easily be convinced that there is a simple underlying phenomenon, or secret magic formula, that generates all of these observed patterns (Foss et al., 2013). However, there is no great magic: the Pareto distribution is simply a power law distribution with the following conditions:

$$P(X \geq x) = \left(\frac{x}{c}\right)^{-\alpha} \quad (3.6)$$

where $c > 0$

and $\alpha > 0$

and $X \geq c$

The Pareto distribution has the probability density function shown in equation 3.7 below.

$$p(x) = \frac{\alpha c^\alpha}{x^{\alpha+1}} \quad (3.7)$$

where $c > 0$

and $\alpha > 0$

The Pareto distribution is a continuous distribution; however, the Pareto principle can be applied to discrete data through its close relative, the Zipf distribution. Zipf's law is a discrete power law distribution named after the American linguist George Kingsley Zipf, who wrote extensively regarding his eponymous distribution in the early 20th century in order to identify and provide an explanation for the most commonly used words in natural languages (Powers, 1998; Rousseau, 2002). Zipf's law states that the frequency, x , of word occurrence is inversely proportional to its rank, α , such that:

$$P(X = x) = cx^{-\alpha} \quad (3.8)$$

where $c > 0$

and $\alpha > 0$

Zipf's law has the probability mass function shown in equation 3.9 below.

$$p(x) = \frac{1/c^\alpha}{H_{N,\alpha}} \quad (3.9)$$

where $c \in \{1, 2, \dots\}$

and $\alpha > 0$

and $H_{N,\alpha}$ is the N^{th} generalised harmonic number

Zipf's law is a discrete power law distribution and is therefore one of the probability distributions of most interest in this investigation, alongside the Poisson and discrete lognormal distributions.

The Poisson, exponential, lognormal and power law distributions are the most used distributions when modelling skewed distributions that may be heavy-tailed. Fitting observed gene family occupancy data to these distribution models will allow for the comparison, selection and analysis of the most accurate statistical model.

3.2 Methods

Gene family data were collected from the picoPlaza (<https://bioinformatics.psb.ugent.be/plaza/versions/pico-plaza/>) database created and maintained by our collaborators at Ghent University (Vandepoele et al., 2013).

The models used in this investigation were 7 different types of heavy-tailed distribution models. Using the `powerLaw` package on R, Discrete Power Law, Poisson, Discrete Log Normal, Discrete Exponential, Continuous Power Law, Continuous Log Normal and Continuous Exponential models were fitted against the observed gene family occupancy data. The code used for this is as follows:

```
gfs<-read.csv("~/root/data.csv",header=T)
olu<-gfs[,3]
orc<-gfs[,4]
ota<-gfs[,5]
bpr<-gfs[,6]
mpu<-gfs[,7]
mrc<-gfs[,8]
vca<-gfs[,9]
gop<-gfs[,10]
cre<-gfs[,11]
ulm<-gfs[,12]
cvu<-gfs[,13]
pse<-gfs[,14]
cnc<-gfs[,15]
```

```
olu0<-olu[olu>0]
orc0<-orc[orc>0]
ota0<-ota[ota>0]
bpr0<-bpr[bpr>0]
mpu0<-mpu[mpu>0]
mrc0<-mrc[mrc>0]
vca0<-vca[vca>0]
gop0<-gop[gop>0]
cre0<-cre[cre>0]
ulm0<-ulm[ulm>0]
cvu0<-cvu[cvu>0]
pse0<-pse[pse>0]
cnc0<-cnc[cnc>0]
test.data<-sp0
library(powerLaw)
sp1<-displ$new(test.data)
sp1$setPars(estimate_pars(sp1))
m2 = dispois$new(test.data)
m2$setPars(estimate_pars(m2))
m3 = dislnorm$new(test.data)
m3$setPars(estimate_pars(m3))
m4 = disexp$new(test.data)
m4$setPars(estimate_pars(m4))
m5 = conpl$new(test.data)
m5$setPars(estimate_pars(m5))
m6 = conlnorm$new(test.data)
m6$setPars(estimate_pars(m6))
m7= conexp$new(test.data)
m7$setPars(estimate_pars(m7))
plot(sp1)
lines(sp1, lty=1, lwd=3)
lines(m2, col="red", lty=1, lwd=3)
```

```
lines(m3, col="blue", lty=1, lwd=3)
lines(m4, col="green", lty=1, lwd=3)
lines(m5, col="pink", lty=1, lwd=3)
lines(m6, col="cyan", lty=1, lwd=3)
lines(m7, col="yellow", lty=1, lwd=3)
```

The confidence intervals for the Discrete Power Law distributions were simulated by using 100 iterations of the model. This was done by using built-in functions within the `powerLaw` package. The code used to do this is given below:

```
test.plot<-sp0
library(powerLaw)
m1<-displ$new(test.plot)
m1$setPars(estimate_pars(m1))
test.length<-length(test.plot)
total.val<-list()
total.cdf<-list()
fit.coef<-m1$pars
for (i in 1:100){
  test<-rpldis(test.length,1,fit.coef)
  test.val<-sort(unique(test))
  test.cdf<-sapply(1:length(test.val), function(x)
length(test[test>=test.val[x]])/length(test))
  total.cdf[[i]]<-test.cdf
  total.val[[i]]<-test.val
}
test.val.plot<-sort(unique(test.plot))
test.cdf.plot<-sapply(1:length(test.val.plot), function(x)
length(test.plot[test.plot>=test.val.plot[x]])/length(test.plot))
max.x<-max(test.plot)
min.y<-min(test.cdf.plot)
test.val<-total.val[[1]]
test.cdf<-total.cdf[[1]]
plot(test.val,test.cdf,log='xy', pch=19,xlim=c(1,max.x),ylim=c(min.y,1))
```

```

for (i in 2:100) {
test.val<-total.val[[i]]
test.cdf<-total.cdf[[i]]
points(test.val,test.cdf, pch=19,col='lightgrey')
}
points(test.val.plot,test.cdf.plot,pch=21)

```

3.3 Results

The species used in this chapter are shown below in table 3.1.

Table 3.1: The species name, genome size, habitat and body organisation of the species investigated in this chapter. The results for the first four are directly shown in this chapter.

Species	Total genes in genome	Habitat	Body organisation
<i>Chlamydomonas reinhardtii</i>	17,741	Freshwater	Unicellular
<i>Volvox carteri</i>	14,971	Freshwater	Colonial Multicellular
<i>Ulva mutabilis</i>	14,273	Marine	Colonial Multicellular
<i>Ostreococcus tauri</i>	7,668	Marine	Unicellular
<i>Gonium pectorale</i>	16,290	Freshwater	Colonial Unicellular
<i>Micromonas</i> sp. RCC299	10,197	Marine	Unicellular
<i>Micomonas pusilla strain</i> CCMP1545	9,937	Marine	Unicellular
<i>Chlorella</i> sp. NC64A	9,791	Freshwater	Unicellular
<i>Coccomyxa subellipsoidea</i> C-169	9,629	Freshwater	Unicellular
<i>Bathycoccus prasinus</i>	7,847	Marine	Unicellular
<i>Ostreococcus lucimarinus</i>	7,805	Marine	Unicellular
<i>Ostreococcus</i> sp. RCC809	7,492	Marine	Unicellular
<i>Picochlorum</i> sp. SENEW3 (SE3)	6,998	Marine	Unicellular

The Cumulative Distribution Function (CDF) plots showing the seven models and their fits against the observed data for the first four species in table 3.1 are shown below in figure 3.1. The CDF plots for the remaining species can be found in Appendix D.

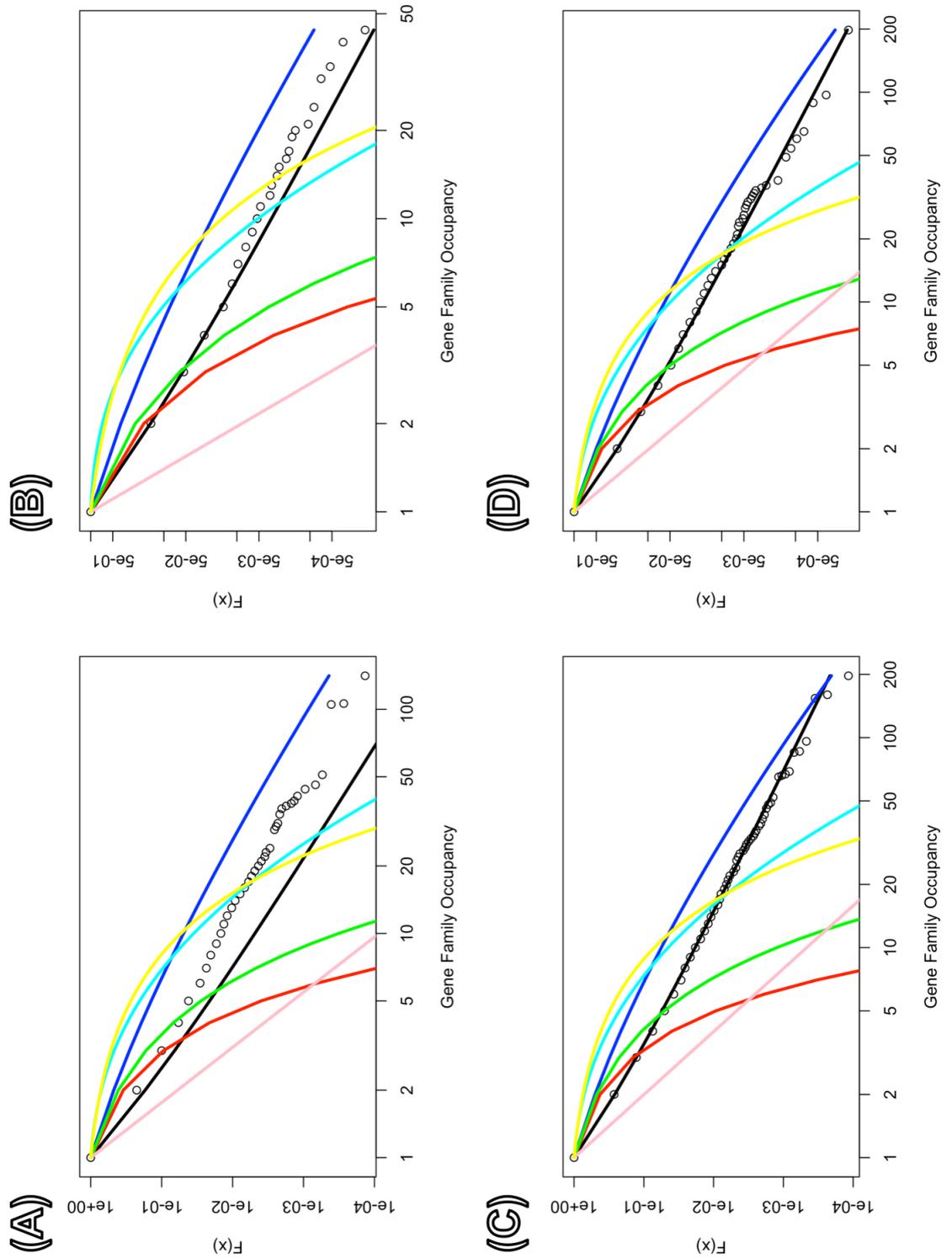


Figure 3.1: The Cumulative Distribution Functions (CDFs) for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for (A) *Volvox carteri*, (B) *Ostreococcus tauri*, (C) *Chlamydomonas reinhardtii* and (D) *Ulva mutabilis*. For a CDF, $F(x)$ is a probability between 0 and 1. The legend is shown below.

Legend	
Circles	Observed gene family sizes
Black ---	Discrete Power Law
Red ---	Poisson
Blue ---	Discrete Log Normal
Green ---	Discrete Exponential
Pink ---	Continuous Power Law
Cyan ---	Continuous Log Normal
Yellow ---	Continuous Exponential

These results suggest that a discrete power law model is the best single-parameter model to describe the observed gene family occupancy data than any other. A Poisson model is the second-best fit; however, it still shows a poor fit, with consistent and significant deviation of the observed data from the model.

To study the goodness of fit, the confidence intervals were calculated for the discrete power law by simulating the distribution 100 times for each species. The plot of confidence intervals for the four species shown above in figure 3.1 are shown in figure 3.2 below. The plot of confidence intervals for the remaining species can be found in Appendix E.

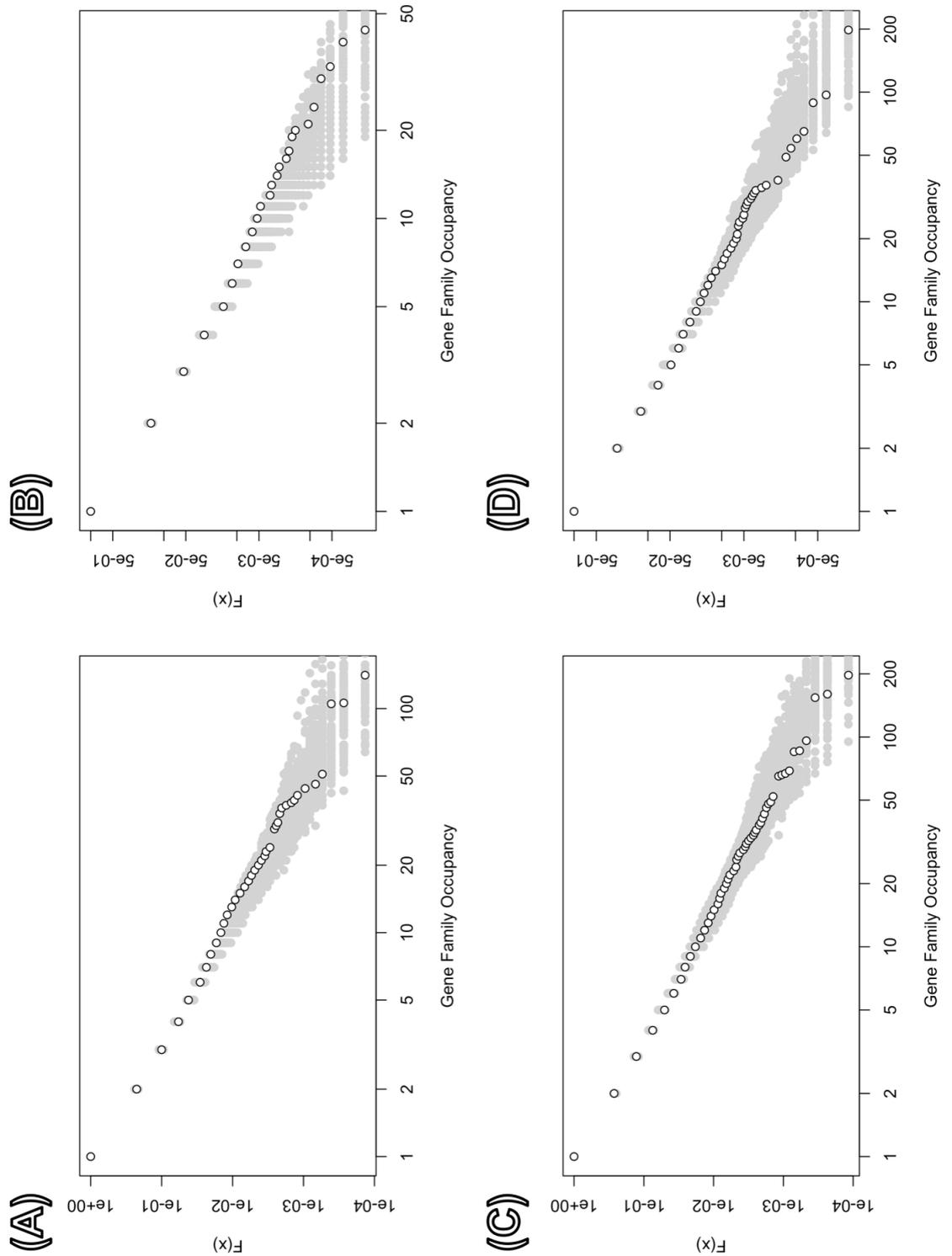


Figure 3.2: The Cumulative Distribution Function (CDF) plots of the confidence intervals for the discrete power law model calculated by 100 simulations for (A) *Volvox carteri*, (B) *Ostreococcus tauri*, (C) *Chlamydomonas reinhardtii* and (D) *Ulva mutabilis*. The grey data points indicate the estimated points, and the black data points indicate the observed gene family occupancy data, as above in figure 3.1. For a CDF, $F(x)$ is a probability between 0 and 1.

These results show that the observed gene family occupancy data generally lie within the bounds of the expected values generated by a discrete power law model.

3.4 Discussion

My results show that a discrete power law model is the best single-parameter model for all genomes. The fit is close enough that using a discrete power law model to detect significant expansions of certain gene families, for example gene families associated with multicellularity, should be possible with high confidence.

The confidence intervals for the discrete power law model were simulated for each species, and from the results it can be concluded that the deviation of the observed data is not beyond what is expected from random deviation. Taken individually, each of the models would suggest that the discrete power law model is able to explain the observations and is therefore a good fit. However, though the discrete power law model is a good fit, it still is not a perfect fit. There are similar patterns of deviation from the discrete power law model within broad classifications of the genomes, such as marine genomes and freshwater genomes. Marine unicellular species show the same type of deviation from the discrete power law model. From an occupancy of 10 onward, the deviation is entirely above the model with the exception of *Picochlorum* sp SENEW3 (SE3) which shows some deviation below the model. Freshwater unicellular species also show a consistent deviation below the discrete power law model from an occupancy of 50 onward. Furthermore, larger genomes show closer adherence to the model than smaller ones. The key here being that though each individual model may be a good fit, the analysis of all of the models show consistent patterns of deviation. Overall, the discrete power law model consistently underestimates the occupancy, indicating biological noise in the data that cannot be fully accounted for by the model.

The species used for these statistical models are entirely restricted to the chlorophyte clade, primarily in order to minimise phylogenetic variation in the analyses that followed. However, this limitation underlines a weakness in this approach: the lack of any species that show a different type of multicellularity. Only three species (*Volvox carterii*, *Ulva mutabilis* and *Gonium pectorale*) in the analysis shows multicellularity and they all show colonial multicellularity. Future studies could improve upon this by including species that display

complex multicellularity.

“All models are wrong, but some are useful.” is a common statistical aphorism attributed to the British statistician George Box, who often discussed the principles of good modeling and their limits (Box, 1976, 1979). In order to make the most accurate model, the underlying mechanisms that generate the observed distribution must be considered and reduced to their core components. The more parameters there are in a model, the more assumptions have to be made, and the less transparent each parameter becomes, and so the more difficult it becomes to infer processes from patterns.

The basis of a good model is first and foremost, selecting the best probability distribution. As aforementioned, the Poisson distribution is used for discrete count data. This characteristic means that the Poisson distribution is useful and has been often used for analysing genome data (Hsieh et al., 2003; Lander and Kruglyak, 1995; Munsky et al., 2012; Pushkarev et al., 2009; Yanai et al., 2000). However, the Poisson distribution has three strict conditions: first, the data is in a fixed period of time, second, given a known average rate and, third, independence from preceding events. The last assumption is not always a realistic assumption and does not wholly reflect the biological mechanisms underpinning genomic data. The birth and death of genes are not always independent of preceding events as there are selection pressures acting on them. Simply, the Poisson model is a fully stochastic model while genome data is both stochastic and deterministic. Furthermore, the Poisson model is a single-parameter model in which biological noise and technical artefacts cannot be fully accounted for. Given the restrictions, it is expected that the Poisson distribution is not a perfect fit for gene family occupancy data. The results of this chapter have shown that the Poisson distribution is not a good fit, and that a discrete power law model is a better fit for all genomes investigated. Larger genomes show closer adherence to the model as expected from the Law of Large Numbers, which was incidentally named as such by Poisson (Hacking, 1990). The next closest distribution is the discrete lognormal model, which is expected, as they are intrinsically connected distributions (Mitzenmacher, 2004). There is no doubt that the gene family occupancy data for the 13 species tested are best predicted, and well predicted, by a discrete power law model. The next logical line of inquiry would be to determine how this particular power law distribution is generated.

A commonly used model to generate power law distributions is the Yule process, which was

developed first in 1925 to model stochastic speciation (Yule, 1925). The Yule process is a Poisson process, meaning that the birth of new genera and species in a given time occur at a given rate, and independently of each other (Kumamoto and Kamihigashi, 2018). The Yule process also neglects death events, which makes little biological sense when considering genome data.

In 1955 the Yule-Simon process was developed from the Yule process, adding a death process (Simon, 1955). For large numbers, the tail of the Yule-Simon distribution is equivalent to the Zipf distribution (Garcia, 2011). This addition of a death process makes more biological sense as a model for genome data. However, the Yule-Simon process is still a stochastic process and adding a level of determinism to the process would be more representative of the biological mechanisms underpinning genome data. This led to a recent approach of using Birth-Death-Innovation/Birth-Death-Mutation/Duplication-Loss-Change models that aim to characterise the balance between stochasticity and determinism by accounting for biological mechanisms (Gabetta and Regazzini, 2010; Karev et al., 2002; Koonin et al., 2002; Maruvka et al., 2011).

The probability mass function for a particular Birth-Death-Mutation (BDM) model is described by the Kummer function $U(a, b, c)$ (Maruvka et al., 2011). So, the probability n_m to pick at random a gene family occupancy of m , is described by equation 3.10 below.

$$n(m) = \frac{R_c \Gamma(1 + \nu)}{m} U \left(1 + \nu, 0, \frac{R_c m}{N_0} \right) e^{-\frac{R_c m}{N_0}} \quad (3.10)$$

where $\mu > \gamma$

$$R_c \equiv \frac{2N_0|\gamma - \mu|}{\sigma^2}$$

$$\sigma^2 = \text{Var}(n)$$

$$\nu \equiv \frac{\gamma}{\mu - \gamma}$$

Here, μ denotes the mutation rate and γ denotes the growth rate, which is a function of the birth and death rates, and these are the two input parameters. I have coded for this BDM model (appendix B), however the input parameters still need to be refined, with sound biological and mathematical justification. The work presented in this thesis chapter is currently being taken forward within our group as an ongoing project.

Nevertheless, the results and conclusions from this chapter statistically confirms the model used by De Clerck et al. (2018) to reach their conclusions, and raises an extremely intriguing question: how is *Ulva* morphogenesis coordinated in the absence of an overarching genomic architecture, or top-down controls?

Plant growth coordination usually involves both top-down and bottom-up controls. For example, the auxin gradient that result in root shoot differentiation is a top-down control, and phosphorus dependent root hair density is a bottom-up control (Ma et al., 2001; Overvoorde et al., 2010). The focus of investigation in the next chapter is now to look more deeply at which broad category of mechanisms can generate the observed *Ulva* morphologies: top-down or bottom-up?

Bottom-up proliferation of *Ulva* growth patterns and cellular architecture

4.1 Introduction

In the general introduction to this thesis, I discussed that the Ulvophytes are the only Chlorophyte clade that contains large multicellular species. These species adopt characteristic patterns that are still used in their identification. However, I also discussed how *Ulva* morphology is increasingly realised to be so plastic that its use in taxonomy can be suspect (Hayden et al., 2003; Malta et al., 1999). In my previous chapter, I helped to show how a lack of multicellular growth coordination genes may underline this plasticity.

Plant growth coordination usually involves both top-down and bottom-up controls. For example, the auxin gradient that result in root shoot differentiation is a top-down control, and phosphorus dependent root hair density is a bottom-up control (Ma et al., 2001; Overvoorde et al., 2010). The focus of investigation in this chapter is now to look more deeply at which broad category of mechanisms can generate the observed *Ulva* morphologies: top-down or bottom-up?

It has been previously suggested that many cellular arrangements in various metazoan tissue types, ranging from *Drosophila* eyes to human biceps follow space filling patterns (Gibson et al., 2006; Sánchez-Gutiérrez et al., 2016). In comparison to the differentiated tissue types of metazoans, *Ulva*, is a much simpler organism with only three cell types:

stem, rhizoid and blade cells (Wichard, 2015). The *Ulva* genome lacks land plant hormonal receptor orthologs such as for auxin, gibberellins and IAA (De Clerck et al., 2018). *Ulva* growth is also known to be symbiotic, as it requires associated bacteria for normal morphogenesis, as introduced previously (Burke et al., 2009; Joint et al., 2002, 2007). Given the lack of hormonal top-down controls and the unlikelihood of cross-kingdom signalling to be entirely responsible for growth coordination, there is a strong case to be made for bottom-up proliferation. This all suggests that simple mathematical and physical constraints may be the key driver of *Ulva* cell proliferation patterns and gives rise to the first hypothesis of this chapter: different morphologies in *Ulva* are results of different bottom-up cellular proliferation.

4.1.1 Filling Flat Space in Nature

I start with the observation that the thalli of many *Ulva* species form flat sheets in which cells are arranged to fill all the available space. This is a very simple pattern of cell distribution and we know that more complex patterns in flowering plants may be explained using simple generative methods. For example, many geometric patterns of morphogenesis have been observed and geometric rules have been established in *Bellis perennis* (common daisy), *Brassica oleracea* “Romanesco” (Romanesco broccoli), *Pinus sp.* (pine cone), *Helianthus annuus* (Common sunflower), all of which demonstrate phyllotaxis (Douady and Couder, 1992; Hernández, 1995; Hotton et al., 2006). Phyllotaxis is an iterative self-organising process of morphogenesis, with variations of the same geometric phenomenon (Fibonacci, Lucas, bijugate and multijugate sequences) occurring depending on the changes in the position of the initial primordia (Atela, 2011; Douady and Couder, 1992; Hotton et al., 2006). Phyllotactic morphogenesis is strikingly similar to patterns of self-organisation found in crystals, and they both share a set of geometric rules: the first is that new units added in succession are at least near equivalent, and the second is that the interactions with previous units determine the position of the new units (Hotton et al., 2006).

An obvious place to start with the flat thalli of *Ulva* is with tessellation across a plane. This is a long-standing area of enquiry. Tessellations are found in areas that are seemingly completely unrelated, from rock formations and insect eyes to bubbles and honeycombs.

The singular most space filling, surface area minimising and stable three-dimensional shape

is a sphere; they allow for the maximum volume with the least surface area. Therefore, soap bubbles, gas bubbles in liquids and water in zero gravity all take the form of spheres. Likewise, the most space-filling, surface area minimising and stable two-dimensional shape is a circle.

To explain how regular global patterns can arise from local interactions, let us imagine a collection of rigid spheres, such as cannonballs, packed as closely as possible, as in figure 4.1 below, which shows the two densest packing of spherical packing possible, known as the face-centred cubic and the hexagonal close-packing methods.

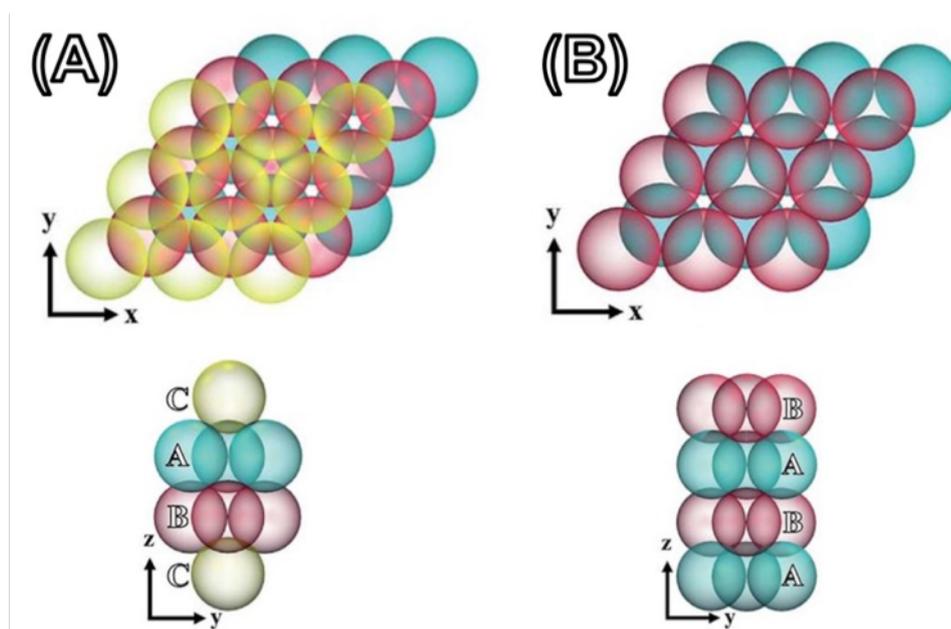


Figure 4.1: Two views (xy and zy) of the densest possible configuration of sphere packing: (A) face-centred cubic and (B) hexagonal close packing arrangements. Note that in (A), there are three repeating layers that form a unit, whilst in (B) there are only two.

Examining these configurations of rigid sphere packing, it is obvious that there are still plenty of gaps. Indeed, Kepler conjectured in 1611 that these two methods are the densest possible arrangements at $\frac{\pi}{\sqrt{18}}$ or 74.05 % density, but proof was only provided in 1998 by the American mathematicians Thomas C. Hales and Samuel P. Ferguson and confirmed in 2017 (Hales et al., 2017; Hales and Ferguson, 2011).

If the spheres were flexible, for example as it is in the case in soap bubbles, then the spheres will deform to eliminate the wasted area and form the most optimal space-filling arrangement possible, that minimises the surface area. Mathematically, it has been established

there are only three possible regular polygons can yield perfect tessellations: triangles, squares and hexagons. An extremely simple mathematical calculation can suggest that hexagons minimise surface area for a given internal area as follows:

Given an equilateral polygon of Area = a

Side length = s

The perimeter of each of the following polygons will be:

Triangles:

$$a = \frac{1}{2}s \cdot \frac{s^2\sqrt{3}}{2} = \frac{s^2\sqrt{3}}{4}$$

$$s^2\sqrt{3} = 4a$$

$$s = \sqrt{\frac{4a}{\sqrt{3}}}$$

$$Perimeter = 3\sqrt{\frac{4}{\sqrt{3}}}\sqrt{a}$$

$$\approx 4.56\sqrt{a}$$

Squares:

$$a = s^2$$

$$s = \sqrt{a}$$

$$Perimeter = 4\sqrt{a}$$

Hexagons:

$$a = s^2 \frac{3\sqrt{3}}{2}$$

$$s = \sqrt{\frac{2a}{3\sqrt{3}}}$$

$$Perimeter = 6\sqrt{\frac{2}{3\sqrt{3}}}\sqrt{a}$$

$$\approx 3.72\sqrt{a}$$

These simple demonstrations for the three regular polygon tessellations are the logical first steps to introducing the millennia-old “Honeycomb Conjecture”, which states: “any partition of the plane into regions of equal area has perimeter at least that of the regular

hexagonal honeycomb tiling” or more simply that a regular hexagonal tessellation is the least perimeter method of enclosing infinitely many unit areas. Though assumed to be true for centuries the mathematical proof for this conjecture has only recently been given (Hales, 2001).

Thus, it can be shown that a tessellation of hexagons abides by mathematical and physical laws to be in the most efficient and stable state. Each side of the bubbles will meet in triplets, to form the most mechanically stable junctions of $\cos^{-1}\left(-\frac{1}{2}\right) = 120^\circ$, known as a Plateau border, and these Plateau borders will meet in quadruplets at a vertex to form tetrahedrons, at $\cos^{-1}\left(-\frac{1}{3}\right) = 109.47^\circ$. This means that in sheets of bubbles, they will tessellate into hexagons and have tetrahedral vertices in a three-dimensional array.

The same principles apply to the spectacular rock formations in Giant’s Causeway in Ireland, shown in figure 4.2, Devils Postpile in California and Jusangjeolli Cliff in South Korea. As lava slowly cools, cracks form to release tension and angles of 120° abides by the principle of maximum strain energy release rate, allowing the most amount of energy release per crack, resulting in the hexagonal columns (Hofmann et al., 2015).



Figure 4.2: Hexagonal rock formations at Giant’s Causeway in Northern Island (Reddit, 2017).

Having considered the theoretical mathematics underpinning the geometry of hexagonal tessellation, actual bee honeycombs are even more logical, in that the perimeter is beeswax, which is a biologically expensive resource, requiring 8.4 kg of honey per 1 kg of beeswax (Winston, 1991). Described by Charles Darwin as “absolutely perfect in economising labour and wax”, honeycombs are now synonymous with perfect hexagonal tessellation, being the namesake. The honeybees have had the reputation of being nature’s best architects. However, as aforementioned, honeycombs, just like bubbles, rock formations and insect eyes, are simply the product of mathematics. The cells of honeycomb are initially produced as circles, which in turn become hexagons upon the application of heat (Karihaloo et al., 2013). It is intuitive then, that a biological organism, would allow mathematical and physical laws to govern cellular tessellation patterns rather than endlessly use energy to work against them. Thus, hexagonal tessellations are commonplace, in order to achieve the most mechanically stable, space filling and resource efficient configuration, unless there are other mechanisms, such as cell and tissue differentiation, that warrant the use of extra energy.

The prevalence of hexagons in cell proliferation patterns has also been long reported. One of the earliest accounts of this kind was by Grew (1682), who wrote of “pentangular, sex-angular, and septangular” parenchymal cells. Many early observations of cell distributions and morphology include a reference to the “bubble-like” nature of the arrangement of said cells. Hooke (1667) noted “a kind of solid or hardened froth, or a congeries of very small bubbles” and Grew (1682) noted “much the same thing, as to its conformation, which the froth of beer or egg is”.

Throughout the first half of the 20th Century, Frederic Thomas Lewis was among the pioneering scientists in the field of mathematical biology, specifically morphology and the underlying mathematics of cellular polygon distribution patterns. He wrote extensively regarding the cellular polygon distribution patterns in both animals and plants, and his work on epithelial mosaics (Lewis, 1928) are especially interesting in a modern context, given the advances in technology and the current capabilities of practically furthering the theoretical framework set out by the likes of Lewis. Lewis (1928), found that much like parenchymal cells of other species, cucumber epithelia largely consisted of three polygons, and this distribution was dominated by hexagons, and with more pentagons than heptagons.

Cell arrangements are not usually as simple as honeycomb, since cell differentiation and tissue formation are required for the functioning of complex organisms such as embryophytes. However, initial analyses have shown that *Ulva* cell tessellation patterns are dominated by hexagons, much like other types of epithelial tissues that have been previously studied. This in conjunction with the lack of multicellular genomic architecture makes *Ulva* an excellent candidate for the exploration of the underlying proliferation mechanism.

4.1.2 Fundamental Drivers of Ulva proliferation patterns and Voronoi Tessellation

Mathematical approaches to space-filling arrangements centres around the generation of global patterns from local interactions. However, biologists typically take the opposite approach to the problem of space-filling arrangements, attempting to determine the underlying mechanism given a pattern.

There are numerous mathematical frameworks and tools developed to address the investigation of various natural patterns, from fractals formed by phyllotaxis and honeycomb tessellation to bubble formation and quasicrystal geometry. Space filling patterns are often investigated using a method called Voronoi tessellation.

The earliest record of the use of Voronoi diagrams was in posthumous publication of a Latin translation in 1644, of the works by the renowned French philosopher René Descartes in his book “The World” (also known as Treatise on the Light; French: *Traité du monde et de la lumière*), written between 1629 and 1633 (Gaukroger, 1998; Okabe et al., 2009). Descartes used the idea of vortices of matter revolving around fixed points in convex regions of space, as shown in figure 4.3 below.

archaeology to astronomy, art and electrical engineering (Du et al., 2006; Okabe et al., 2009).

A Voronoi tessellation diagram is a method of investigating the most optimally space filling arrangement of polygons, given a set of generating seeds, or generators.

Voronoi tessellation is defined as follows by Du et al. (1999):

“Given an open set $\Omega \subseteq \mathbb{R}^N$, the set $\{V_i\}_{i=1}^k$ is called a *tessellation* of Ω if $V_i \cap V_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^k \bar{V}_i = \bar{\Omega}$. Let $|\cdot|$ denote the Euclidean norm on \mathbb{R}^N . Given a set of points $\{z_i\}_{i=1}^k$ belonging to $\bar{\Omega}$, the *Voronoi region* \hat{V}_i corresponding to the point z_i is defined by

$$\hat{V}_i = \{x \in \Omega \mid |x - z_i| < |x - z_j| \text{ for } j = 1, \dots, k, j \neq i\}.$$

The points $\{z_i\}_{i=1}^k$ are called *generators*. The set $\{V_i\}_{i=1}^k$ is a *Voronoi tessellation* or *Voronoi diagram* of Ω , and each \hat{V}_i is referred to as the *Voronoi region* corresponding to z_i .”

The definition given by Du et al. (1999) of Voronoi tessellation is not the exclusive definition, with Okabe et al. (2009) having defined Voronoi tessellations as closed sets, for example. I found that the manner in which Du et al. (1999) defined Voronoi regions before Voronoi generators can be a cause of confusion, and that an alternative approach using a different perspective can clarify the definitions and understanding of the concepts. Therefore, I have written a more explicit definition of Voronoi edges and the derived definition of Voronoi regions and tessellation to aid with the understanding of Voronoi tessellations.

Given an open set $\Omega \subseteq \mathbb{R}^N$, a Voronoi edge is defined as:

$$\tilde{V}_{i,j} = \{x \in \Omega \mid |x - z_i| = |x - z_j| \quad \forall j = 1, \dots, k, i \neq j\}$$

$$\hat{V}_i = \bigcup_{j=1}^k \tilde{V}_{i,j}$$

$$\hat{V}_i = \left\{ \hat{x} \in \bar{\Omega} \mid \hat{x} < \tilde{V}_{i,j} \quad \forall j = 1, \dots, k, i \neq j \right\}$$

The definitions of some useful terminology can be clarified for some, using written explanations and definitions, as follows. Voronoi tessellations are constructed from a given set of centre points or Voronoi generators and refer to the entire arrangement of Voronoi regions. A Voronoi region is constructed by calculating the distance from a generator (A) to its nearest neighbouring generator (B), and drawing a line, termed a Voronoi edge, to bisect that distance equally, so that the line is equidistant from generator A and B. This process is repeated for all neighbouring generators, until Voronoi edges encapsulate the generators, forming polygons, that are termed Voronoi regions. A Voronoi region, then, is the convex area in Euclidean space that encapsulates every point that is closer to one Voronoi generator than any other. The terms Voronoi generator, Voronoi region, Voronoi edge and Voronoi tessellation from henceforth will be referring to the above definitions.

As is intuitively apparent, calculating Voronoi cells is extremely labour-intensive, so Voronoi tessellation diagrams did not see any use in disciplines such as biological sciences where thousands of calculations have to be completed for a single sample, until technological advances made their use feasible. Honda (1978) is widely credited for the first use of Voronoi tessellation in a biological context, more than a hundred years after Voronoi established the method Bock et al. (2010).

Voronoi tessellation diagrams in biological sciences have been used largely to model morphogenesis and cancerous proliferation (Barrio et al., 2013; Bock et al., 2010; Kansal et al., 2000). Most of the work in the literature focus on metazoan tissue and complex multicellular plant tissues; there is currently no literature concerning the Voronoi tessellation analysis of macroalgae, observed or modelled. Voronoi tessellations give an objective measurement of space-filling. When one wishes to study cell patterns, the Voronoi tessellation gives the objective pattern from the centre points of said cells, which can be compared to the cells and models. However, there is no “goodness of fit” measure yet defined, therefore descriptive methods are used here.

This all gives rise to two hypotheses that are the subject of investigation in this chapter:

It is possible for local information to result in global patterns, as observed in the likes of root hair cells developing using DELLA transcription factors switched on by positional cues (Lee and Schiefelbein, 2002). This is the basis of the first hypothesis of this chapter: different morphologies of *Ulva* can arise from different bottom-up cellular proliferation

patterns.

It is intuitive that a biological organism, would tend towards hexagonal tessellations in order to achieve the most mechanically stable, space filling and resource efficient configuration. This is the basis of the first set of computer models that will lay the foundation for further modelling, and the second hypothesis of this chapter: cell tessellation patterns in simple multicellular organisms such as *Ulva* can be replicated by applying random deviations, or “jitter” in the Voronoi generating centre points of a regular hexagonal honeycomb pattern.

4.2 Methods

4.2.1 Stain Selection, Image Collection and Analysis

In order to investigate cell patterns, one must first image the cell walls ideally by themselves. Light microscopy is not suitable for this purpose, meaning that confocal microscopy is the ideal method. To perform confocal imaging, a reliable cell wall staining protocol is required. However, there is no standard *Ulva spp.* cell wall staining protocol, so I set out to establish one. The first tested were Calcofluor White and Pontamine Fast Scarlet 4B. Calcofluor White is a non-specific fluorochrome with a high affinity for the β 1-3 and β 1-4 polysaccharides like those in cellulose and chitin, so has been steadily used in biological research for decades, with excitation-emission maxima at 365 nm and 435 nm (Harrington and Hageage, 2003; Harrington and Raper, 1968; Herth, 1980; Herth and Schnepf, 1980; Sigma-Aldrich, 2019a). Pontamine Fast Scarlet 4B is a much newer alternative to Calcofluor White that has higher specificity, with excitation-emission maxima at 507 nm and 615 nm (Anderson et al., 2010; Hoch et al., 2005; Liesche et al., 2013; Sigma-Aldrich, 2019; Thomas et al., 2013).

Due to the advantages that dyes higher specificity provide, Pontamine Fast Scarlet 4B was the first candidate for staining trials and was prepared using deionised water and 1X PBS (0.01 M Phosphate buffer, 0.154 M NaCl, pH7.4), at concentrations ranging from 0.01% to 1%. Various permeabilisation and fixing steps were used to facilitate the dye to penetrate the tissue, however no combination of steps allowed for an even stain that was reliably observable with a confocal microscope.

Calcofluor White was the next candidate for experimentation. At 0.1% in 1X PBS, Calcofluor White gave a fairly even stain, reliably detectable when excited with a 405 nm laser on a confocal microscope. Figure 4.4 below shows a representative confocal image taken of an *Ulva* sample stained with Calcofluor White, on a Leica SP5 confocal microscope.

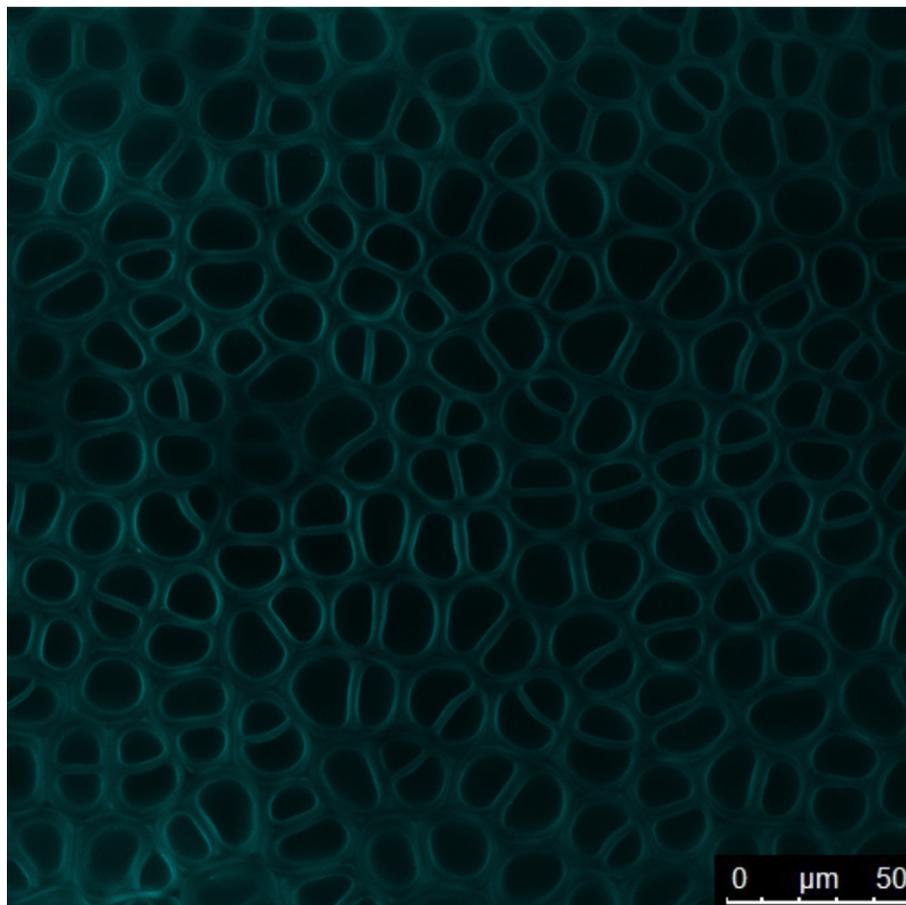


Figure 4.4: Confocal microscope image of *Ulva* spp stained with Calcofluor White.

Unfortunately, the images showed weak signal, low resolution and noise from autofluorescence. The final dye tested was a little-known one that has seen only limited use in biological research, called Direct Yellow 96 (Diphenyl Brilliant Flavine 7 GFF). Figure 4.5 below shows a representative confocal image taken of an *Ulva* sample stained with Direct Yellow 96, on a Leica SP5 microscope.

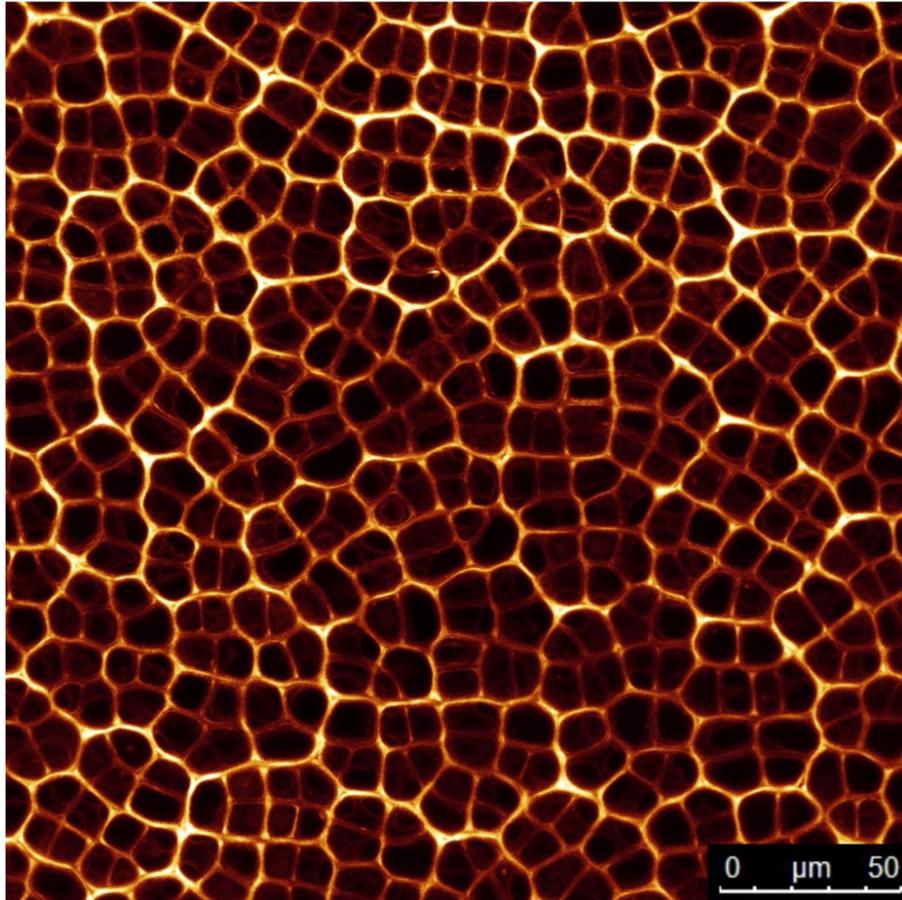


Figure 4.5: Confocal microscope image of *Ulva* spp stained with Direct Yellow 96.

Microscope images capture a representative field of cells, but some cells at the edge of the field of view are not entirely captured. Therefore, for the analysis of certain attributes such as polygon distribution and change, a simple exclusionary definition of cells to be analysed, Valid Cells, was used.

Valid Cells were defined as:

$$\text{Valid Cells} = \text{All Cells} - \text{Edge Cells},$$

Where Edge Cells = Any cell in contact with the edge of the frame.

This is visually represented in figure 4.6 with Valid Cells being represented in green and Edge Cells being represented in red.

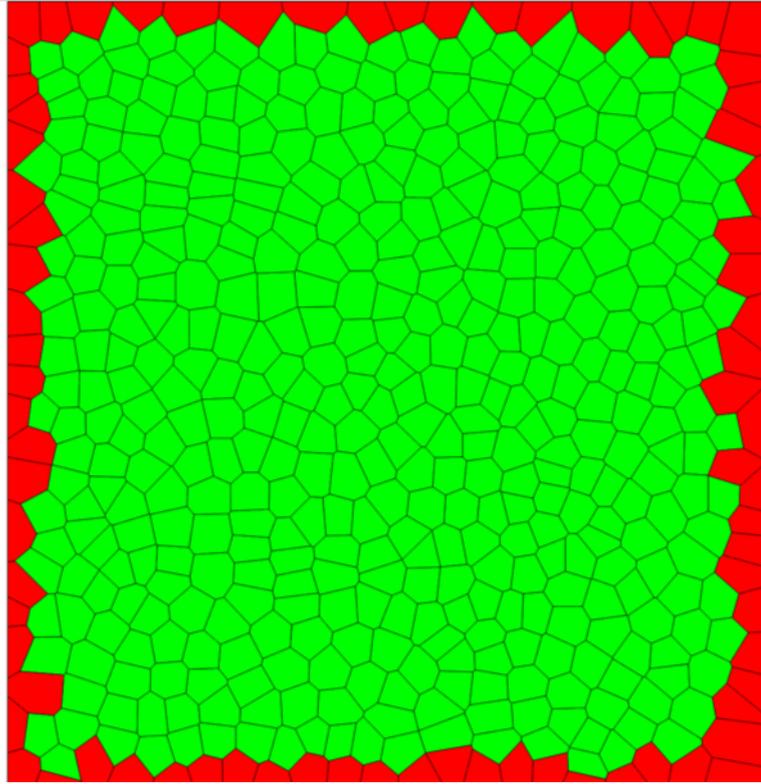


Figure 4.6: Voronoi Diagram obtained from a leaf morphotype sample, coloured to visualise data. Red represents the Edge Cells, which were excluded from various analyses. Green represents the Valid Cells included in various analyses. The diagram is generated in Wolfram Mathematica using the VoronoiMesh tool (Wolfram, 2015). See Appendix A for code.

4.2.2 Pattern Modelling

Initial simple computer modelling was conducted in two phases to establish a framework by which to analyse the real collected data, and then to investigate the second hypothesis that *Ulva* cellular proliferation patterns could be achieved using random variations in centre points of a regular honeycomb pattern.

The first set of models were devised to establish the comparative framework:

Model 1 was a completely regular hexagonal “honeycomb” sheet.

Model 2 was a sheet of irregular hexagons, with each hexagon measuring half the height and double the width of model 1.

Model 3 was a Voronoi diagram from completely randomly generated centre points, and it was introduced as a reference point, as all statistics should show a significant difference

with a completely random model. There were multiple iterations of model 3, in order to identify the frequency of false positives.

Model 4 was a regular square grid, introduced as another control. Again, all statistical analyses should show significant differences.

These models are visually represented below in figure 4.7.

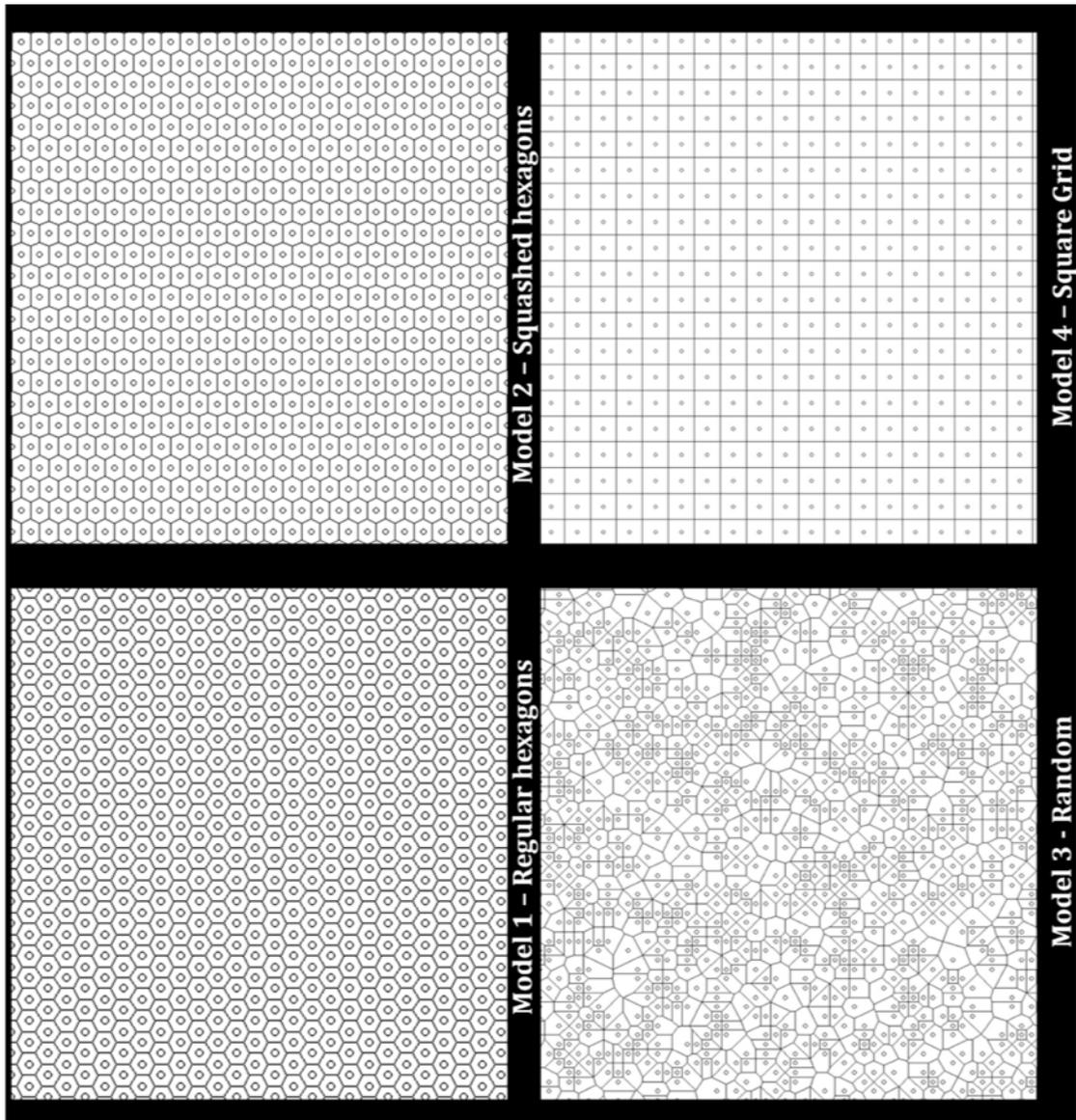


Figure 4.7: Representative patterns generated by of models 1 to 4 used in the analysis

Each of these models were then given “jitter”: randomly deviations of the centre points, to test the robustness of the second hypothesis of the chapter. This gave the following models:

Model 5: Regular hexagons + 10% random jitter in centre points

Model 6: Regular hexagons + 20% random jitter in centre points

Model 7: Regular hexagons + 50% random jitter in centre points

Model 8: Squashed hexagons + 10% random jitter in centre points

Model 9: Squashed hexagons + 20% random jitter in centre points

Model 10: Squashed hexagons + 50% random jitter in centre points

Models 5, 6 and 7 are visually represented below in figure 4.8.

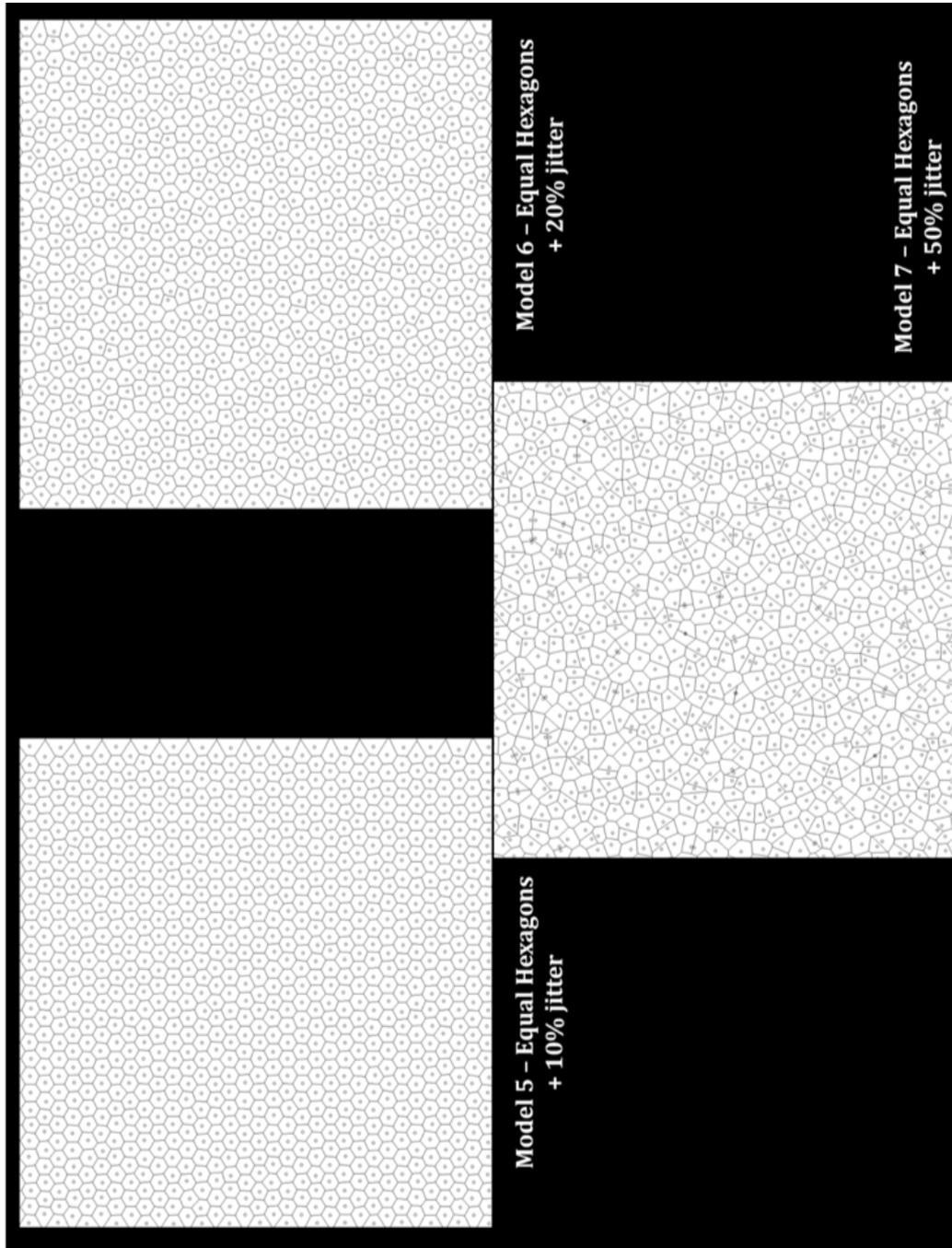


Figure 4.8: Representative patterns generated by models 5, 6 and 7 used in the analysis.

Given that the jitter is by definition random deviations, they will be different every time. In order to reliably conduct statistical analyses, three iterations of each of the jitter models were generated. Models 5 to 10 all had three iterations, and each individual iteration was used in the statistical analyses.

The analysis of the models examined three main variables:

1. Edge length (the length of each side of a polygon)
2. Area of each polygon
3. Edge Count (the number of edges of each polygon)

The main aim of the latter part of this investigation is to better characterise the natural patterns of cellular proliferation in *Ulva*. Twelve samples of the leaf morphotype (figure 2.1 **A**, section 2.2) and twelve samples of the ribbon morphotype (figure 2.1 **B**, section 2.2) were dyed with Direct Yellow 96. The resulting confocal image was then used to gather information on the centre points of each cell, which was in turn used to generate Voronoi tessellation diagrams for analysis. The analysis of these natural patterns focused on the same three main variables as the models.

4.2.3 Statistical Analysis

The Wilcoxon signed-rank test was utilised to test the difference between the leaf and ribbon morphotypes, as they are related samples. In order to apply the same principle of related comparisons, all models were tested against their corresponding pair. For example, all of the jitter models were compared against their non-jittered hexagonal honeycomb counterparts from which they were generated. The same principle was applied to the comparisons of leaf and ribbon morphotypes, with the twelve samples of each morphotype being labelled arbitrarily from 1 to 12 and compared to the corresponding sample of the other morphotype. All statistical analyses were conducted using native functions in the core stats package (v3.6.2) on R (R-Project, 2019).

4.3 Results

4.3.1 Voronoi Tessellation Fit

The first notable finding is that cell patterns in *Ulva* can be described using Voronoi tessellations, as shown in figure 4.9 below.

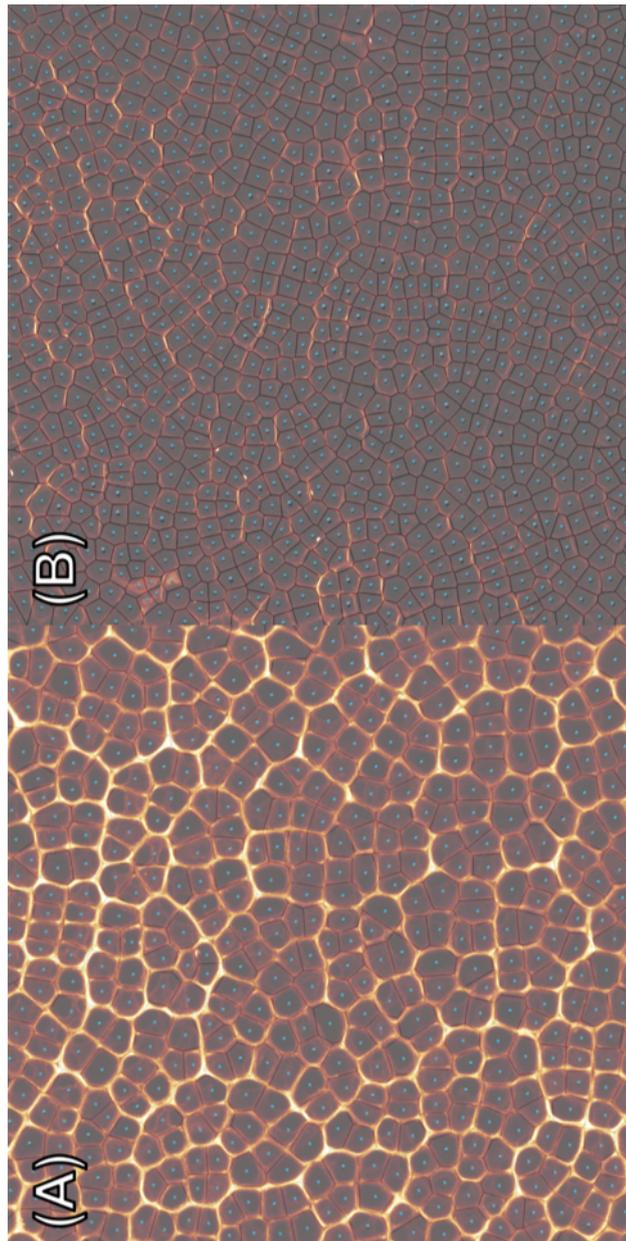


Figure 4.9: Representative confocal image of the (A) leaf and (B) ribbon morphotypes, both stained with Direct Yellow 96. The blue dots indicate the centre points of the cells, which were then used to generate Voronoi tessellation diagrams that are overlaid on the confocal images. Note that the Voronoi edges align almost perfectly with the cell walls.

All 24 (12 leaf, 12 ribbon) of the Voronoi tessellation diagrams show close adherence to the cell walls of the samples they were generated from. There are no deviations observed, making a compelling case for the morphogenesis of *Ulva* thalli being a space-filling bottom-up proliferation mechanism.

4.3.2 First Phase Model Statistics

The statistical analyses of the models showed that there is a significant difference ($p < 0.05$) between edge lengths when comparing the regular and squashed hexagonal models (Model 1 and 2). There were significant differences between all of the variables when comparing hexagonal models to the random model (Model 1 and 2 with 3). All variables showed significant differences when comparing hexagonal models to the square grid model (Model 1 and 2 with 4).

There is no significant difference between edge count and area when comparing the regular and squashed hexagonal models. This means that even if the areas and the edge counts show no significant difference, there can be a significant difference in edge length. Uniform distribution models differ significantly from random models, and uniform models of other shapes. These results are summarised below in table 4.1.

Table 4.1: Wilcoxon Signed Rank Test p-values for Model 1 (Regular Hexagon Honeycomb), Model 2 (Squashed Hexagon Honeycomb), Model 3 (Random) and Model 4 (Square Grid).

Model	Edge Length	p-value	
		Area	Edge Count
1 vs 2	$< 2.20 \times 10^{-16}$	0.8449	1.0000
1 vs 3	1.87×10^{-9}	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$
2 vs 3	1.98×10^{-5}	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$
1 vs 4	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$
2 vs 4	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$

4.3.3 Second Phase Model Statistics

When comparing the regular and squashed hexagonal models with 10 and 20% jitter, there are significant differences in edge length but not edge count, whilst the comparison between the two hexagonal models with 50% jitter gave significant differences in both edge length

and edge count. This means that edge length shows the most sensitivity to variables introduced, with the 10% jitter models showing significant difference. Area is the least sensitive, with no significant differences even at 50% jitter. These results are summarised in table 4.2 and 4.3 below.

Table 4.2: Wilcoxon Signed Rank Test p-values for Model 1 (Regular Hexagon Honeycomb) and associated jitter models.

Model	Edge Length	p-value	
		Area	Edge Count
3 (Random)	1.87×10^{-9}	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$
4 (Square)	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$
5 (10% Jitter)	1.81×10^{-7}	0.1948	0.6438
6 (20% Jitter)	$< 2.20 \times 10^{-16}$	0.2154	0.4131
7 (50% Jitter)	$< 2.20 \times 10^{-16}$	0.3255	0.0206

Table 4.3: Wilcoxon Signed Rank Test p-values for Model 2 (Squashed Hexagon Honeycomb) and associated jitter models.

Model	Edge Length	p-value	
		Area	Edge Count
3 (Random)	1.98×10^{-5}	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$
4 (Square)	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	2.20×10^{-16}
8 (10% Jitter)	0.0019	0.6927	0.3652
9 (20% Jitter)	1.04×10^{-9}	0.2425	0.2255
10 (50% Jitter)	$< 2.20 \times 10^{-16}$	0.0509	0.0167

Comparing models of equal jitter, so regular hexagons with 10% jitter to squashed hexagons with 10% jitter (Model 5 with 8), 20% to 20% (Model 6 with 9) and 50% to 50% (Model 7 with 10), there was no significant difference in area and edge length, whilst edge count showed a significant difference only in the first two comparisons. These results are summarised below in table 4.4.

Table 4.4: Wilcoxon Signed Rank Test p-values for models of equal jitter (regular vs squashed honeycomb).

Model	Edge Length	p-value	
		Area	Edge Count
5 vs 8 (10% Jitter)	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.9279
6 vs 9 (20% Jitter)	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.3975
7 vs 10 (50% Jitter)	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.0045

4.3.4 Leaf vs Ribbon Morphotype Comparisons

The set of ten models provided a robust framework for comparative analyses of the primary data generated from confocal images of ribbon and leaf morphotypes, the results of which are summarised below in table 4.5.

Table 4.5: Wilcoxon Signed Rank Test p-values for leaf vs ribbon morphotypes. For each of the 12 samples of each morphotype, the samples were compared with their corresponding sample number of the other morphotype (i.e. leaf sample 1 vs ribbon sample 1). Sample 2 has been highlighted as an anomalous result, as addressed in the discussion.

Sample	Edge Length	p-value	
		Area	Edge Count
1	5.04×10^{-5}	8.75×10^{-16}	0.5458
2	0.6513	0.1770	0.9918
3	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.5033
4	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.3992
5	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.9419
6	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.9206
7	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.5902
8	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.6145
9	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.4615
10	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.9431
11	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.5476
12	2.95×10^{-14}	$< 2.20 \times 10^{-16}$	0.7488

As can be seen above, no significant differences were found between edge counts of the two morphotypes, whilst edge lengths and areas always showed significant differences. The edge counts, or polygon distribution is summarised below in figure 4.10, alongside Model 7 (50% Jitter).

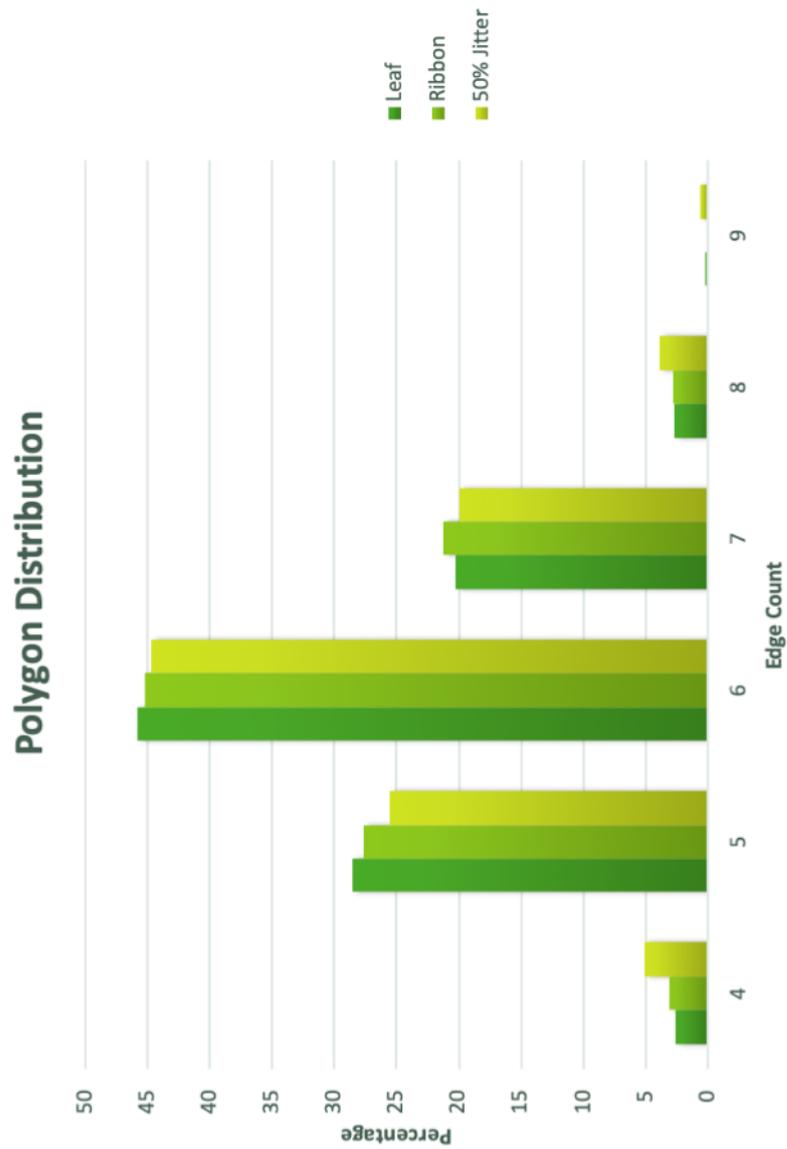


Figure 4.10: The polygon distribution of leaf and ribbon morphotypes and model 7 (50% jitter).

The p-values from the Wilcoxon signed rank tests suggest that 50% jitter models are statistically not significantly different to real data. However, visual comparison of the models show clear differences in the patterns, as shown in figure 4.11 below.

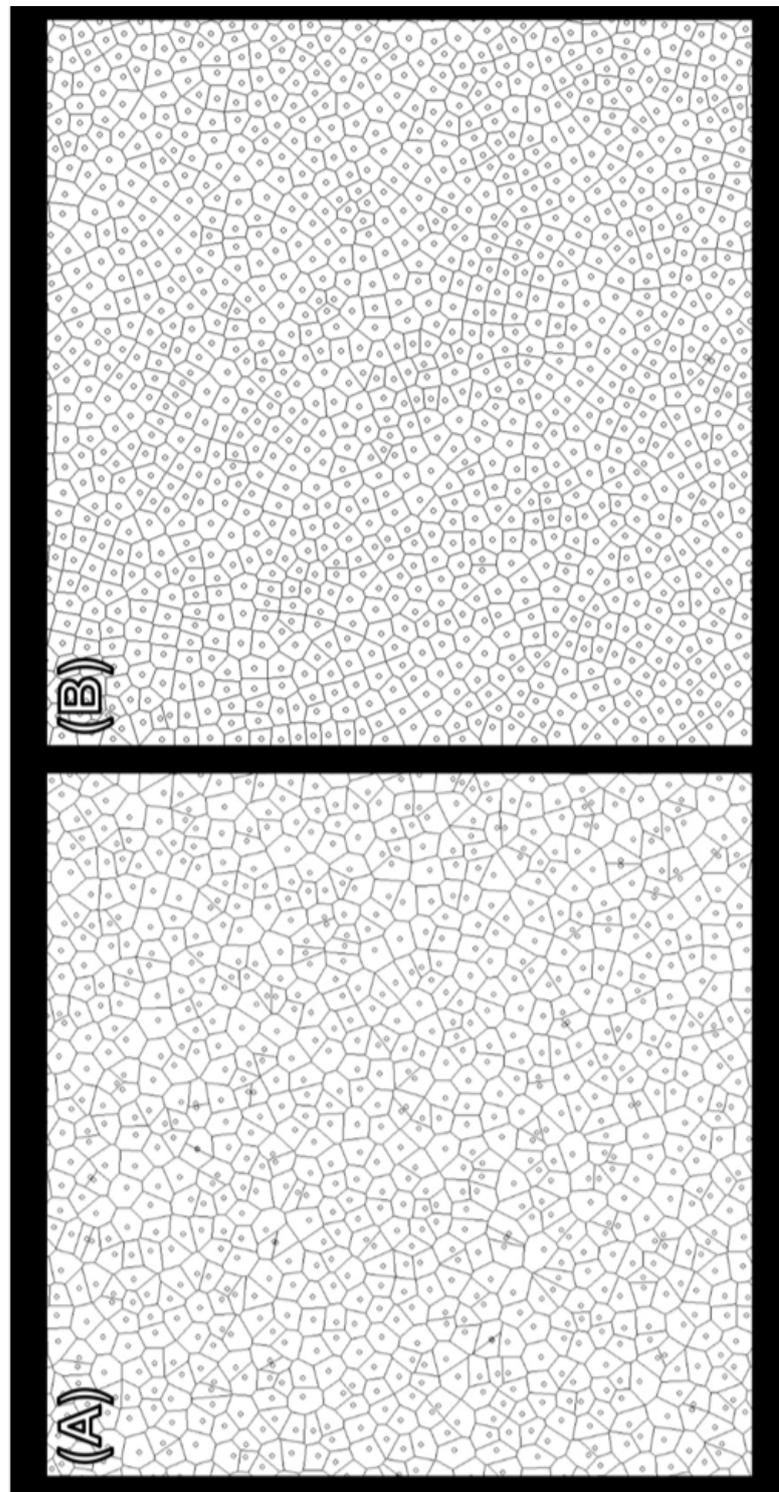


Figure 4.11: Side by side comparison of the Voronoi tessellation of (A) a 50% Jitter model and (B) a ribbon morphotype sample.

To clarify the difference, there seems to be distinct lines of cells in real samples, presumably arising from cell division. These lines are illustrated below in figure 4.12.

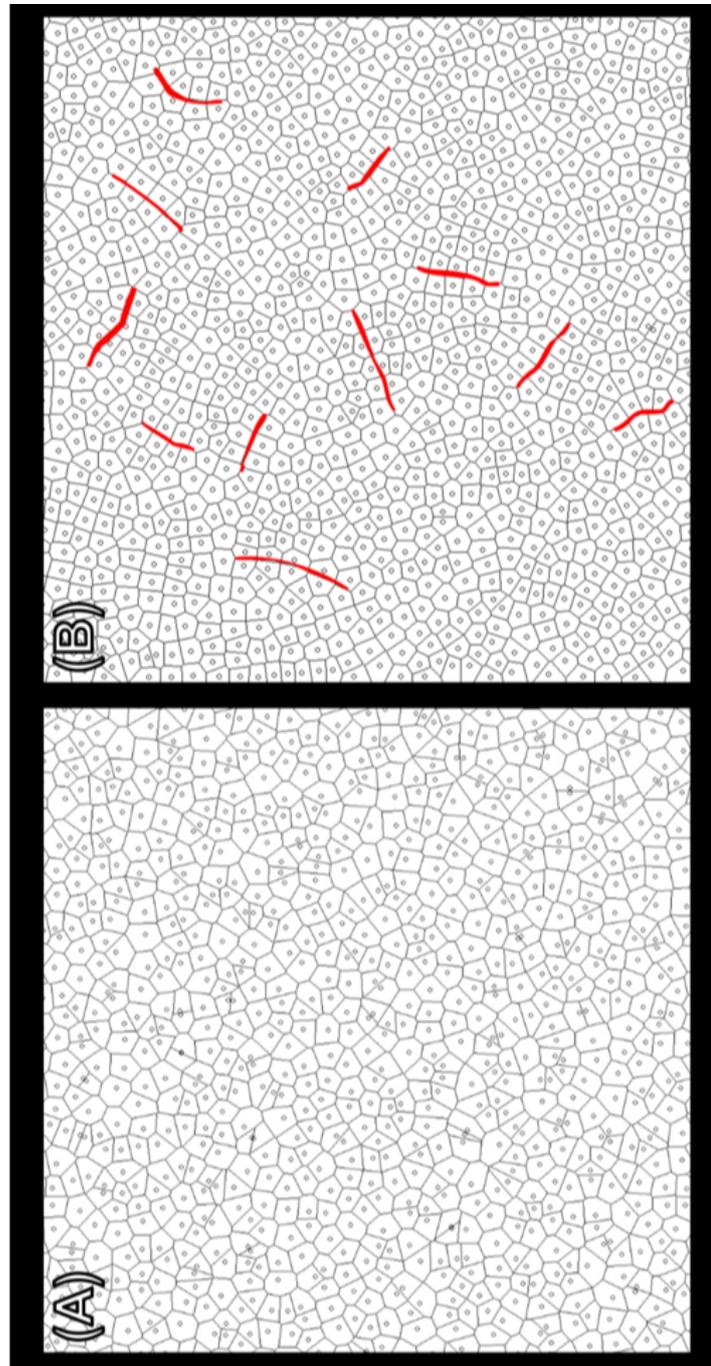
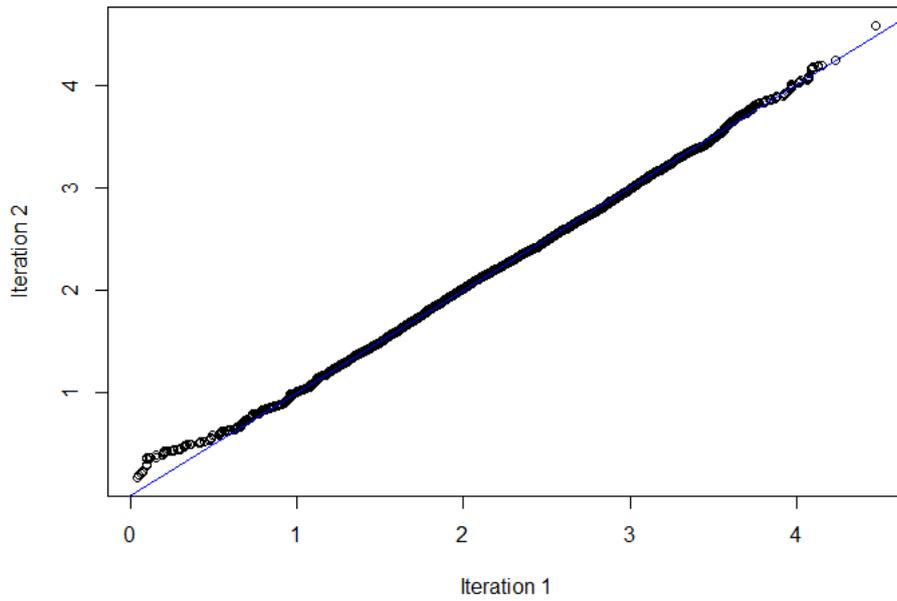


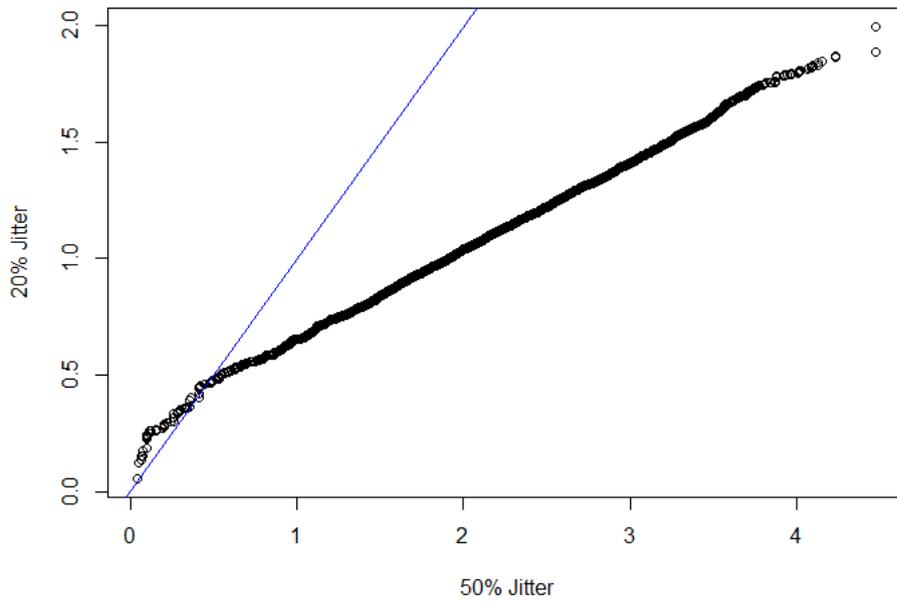
Figure 4.12: Side by side comparison of the Voronoi tessellation of (A) a 50% Jitter model and (B) a ribbon morphotype sample. (B) has been illustrated with red lines to denote some of the centre lines of division that can be observed in clusters of cells that have likely divided from the same parent cells.

Q-Q plots of various models and real data were made, and show that the ribbon and leaf morphotypes seem to be generated by the same mechanism, whilst the models and the real data are not, as shown in figure 4.13 below.

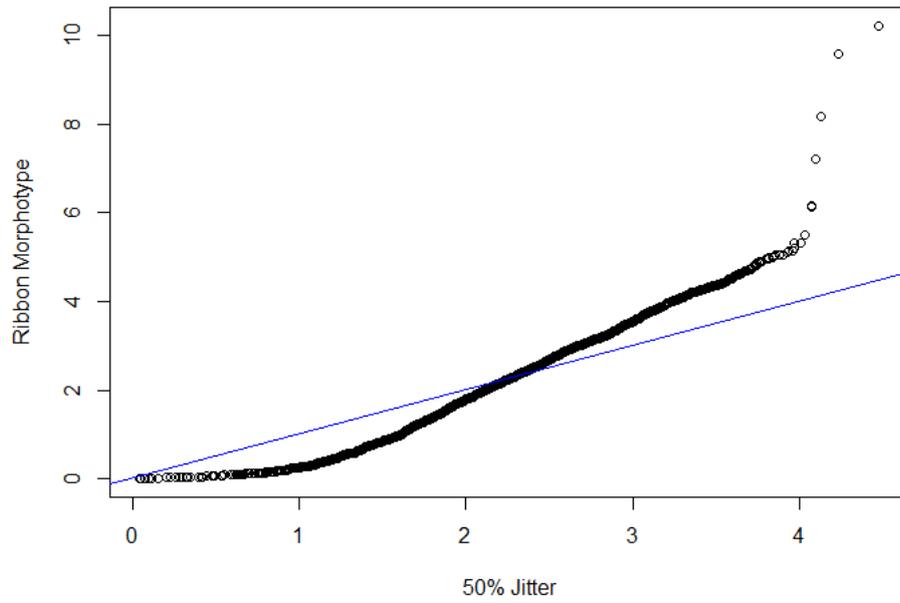
(A)



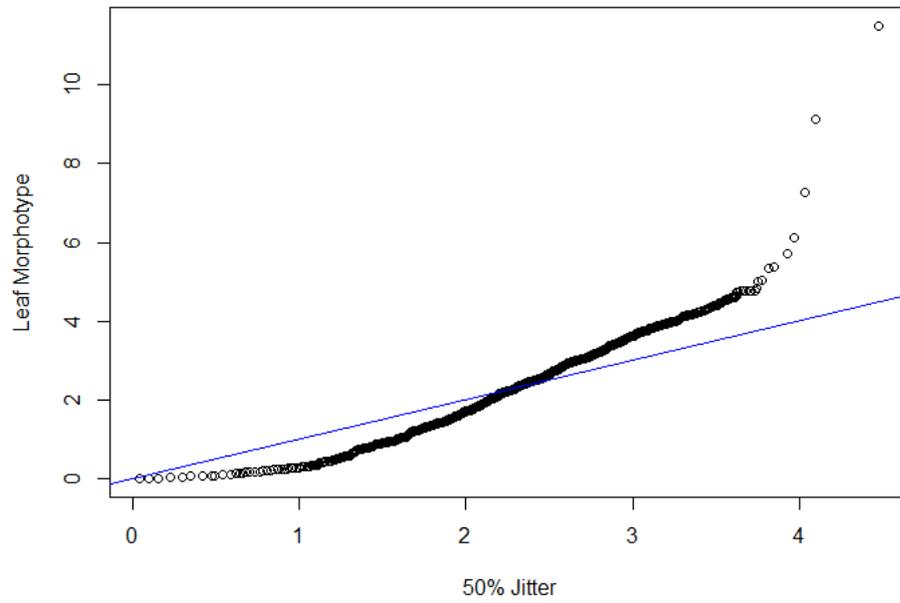
(B)



(C)



(D)



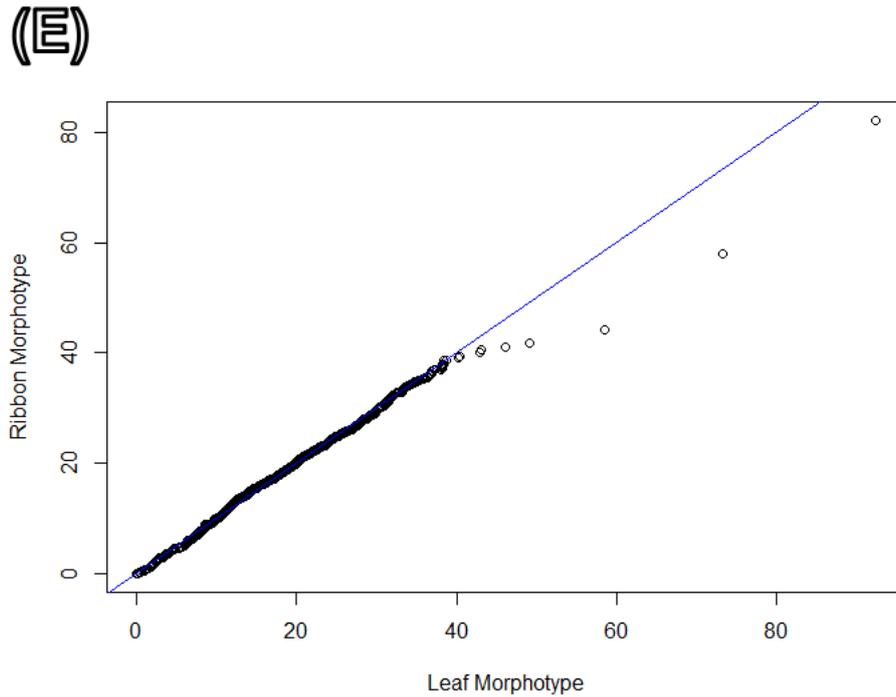
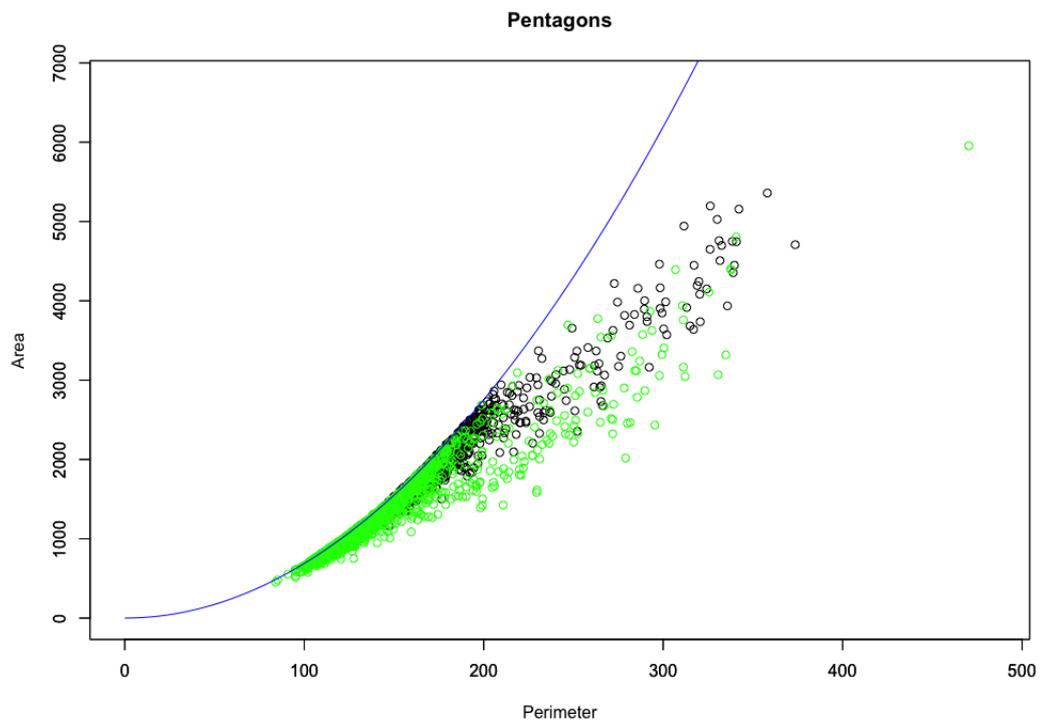
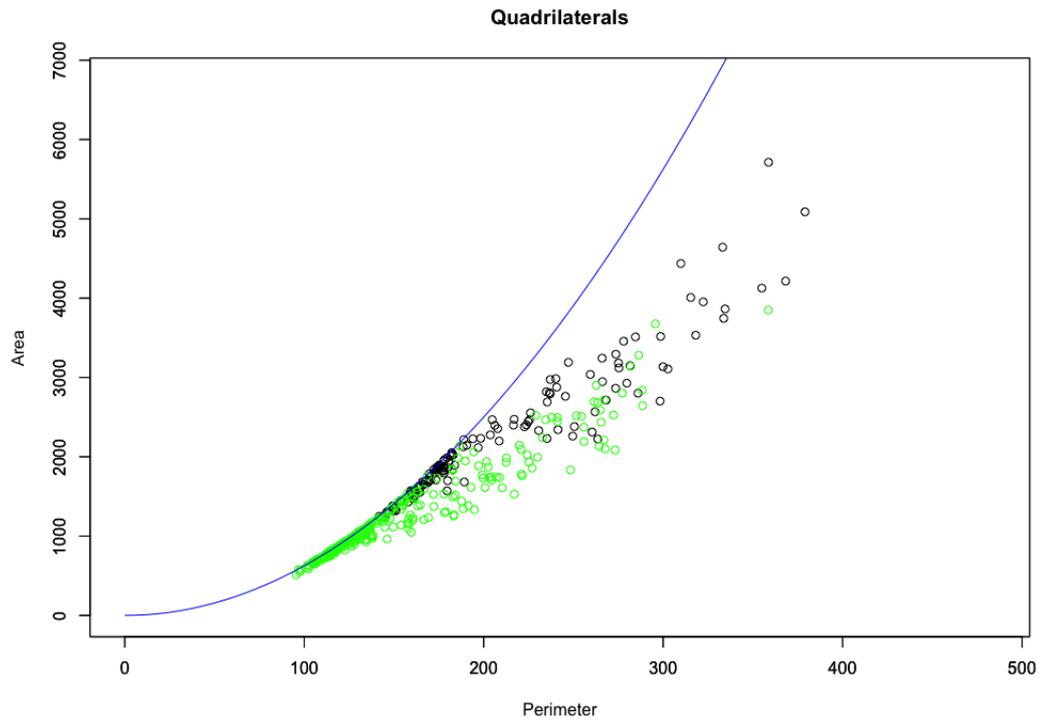
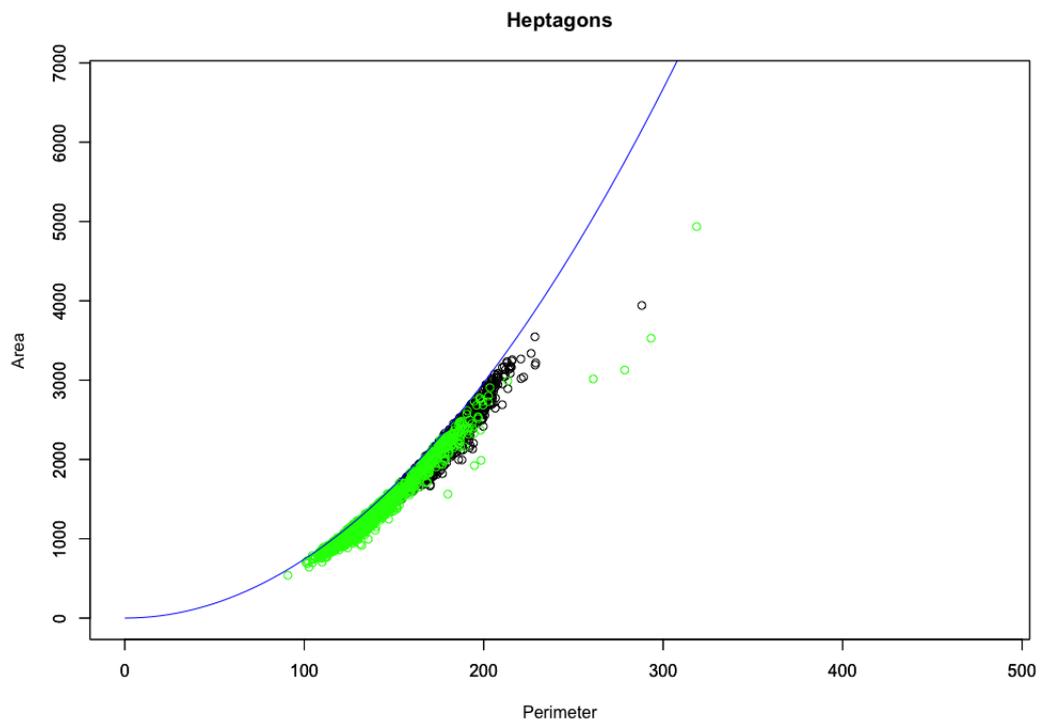
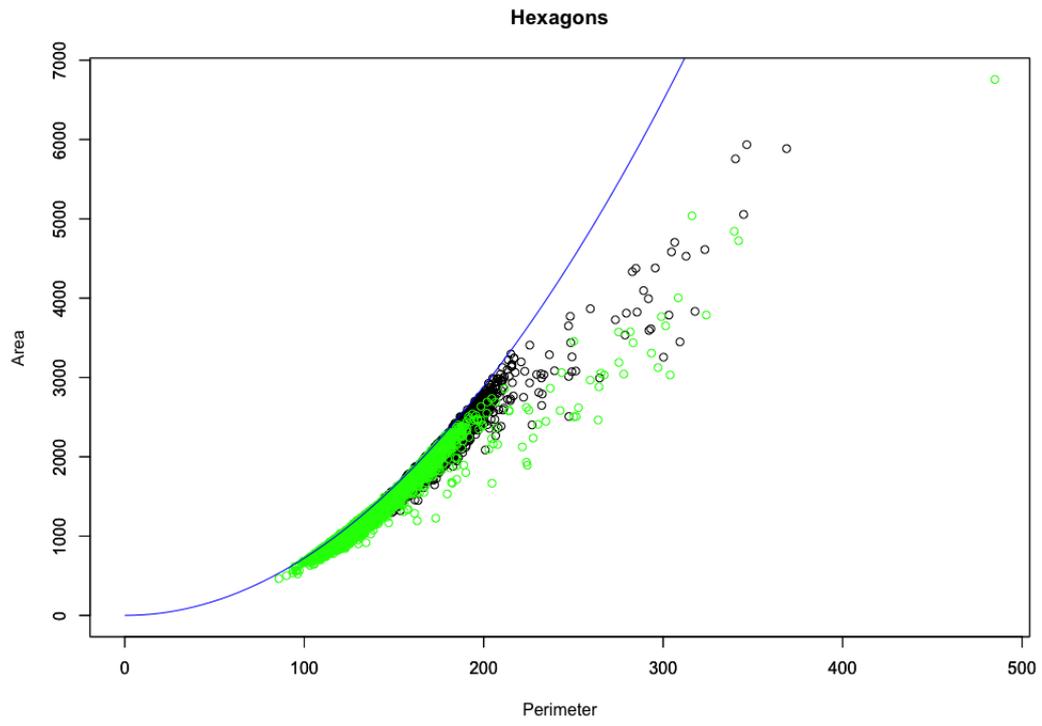


Figure 4.13: Q-Q plots of (A) two iterations of a 50% jitter model, (B) A 50% jitter model vs a 20% jitter model, (C) A 50% jitter model vs a ribbon morphotype sample, (D) 50% jitter model vs a leaf morphotype sample and (E) a leaf morphotype sample vs a ribbon morphotype sample. The blue line is the line $y=x$.

The perimeter and area of the leaf and ribbon morphotypes were plotted against each other to examine regularity. There are few differences in regularity, relative area and perimeter of quadrilaterals, pentagons and hexagons between the leaf and ribbon morphotypes, with the majority of the data points overlapping. Furthermore, the quadrilaterals, pentagons and hexagons show more irregularity in shape than the heptagons, octagons and nonagons for both morphotypes. Finally, the ribbon morphotype shows heptagons, octagons and nonagons of larger relative area and perimeter.





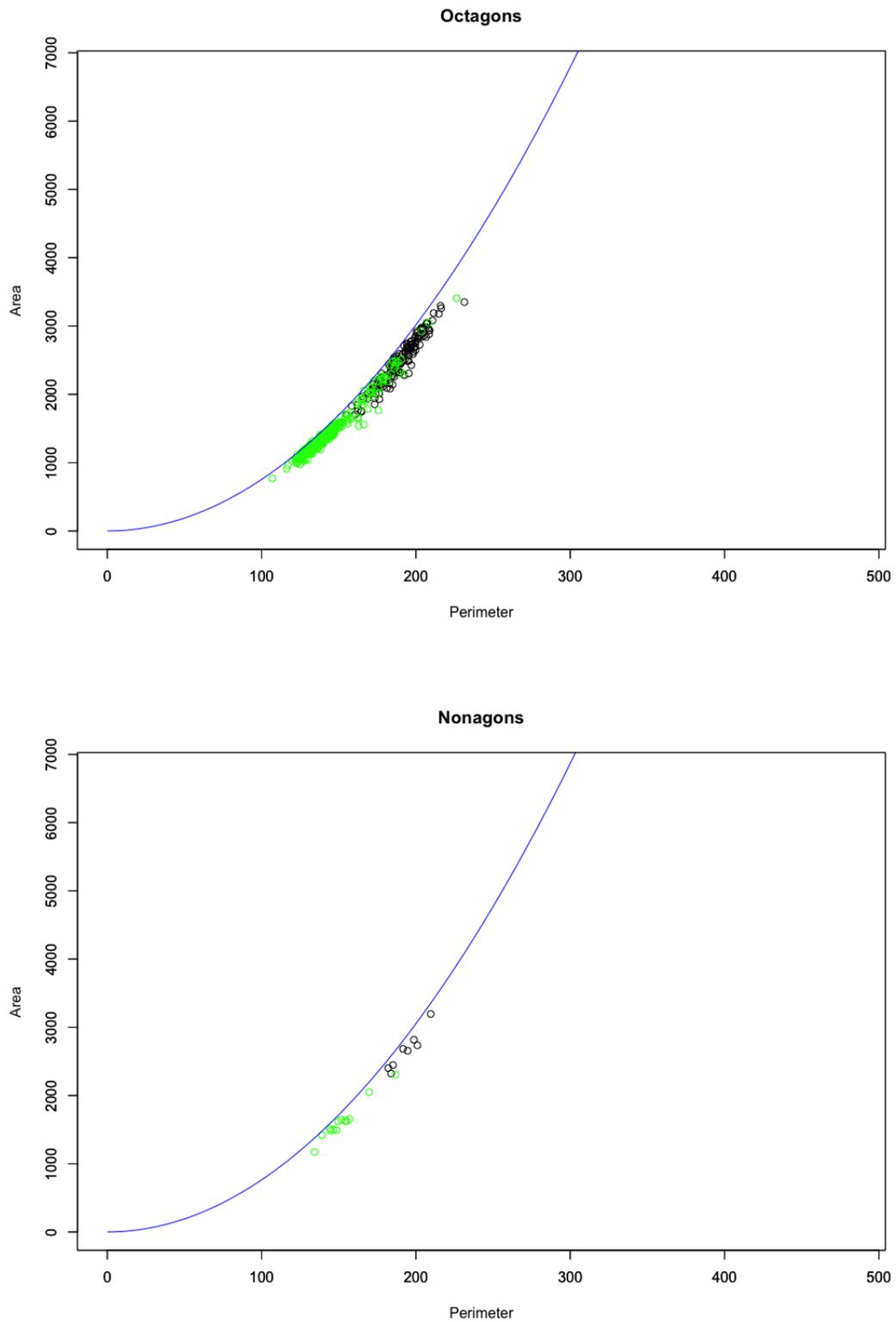


Figure 4.14: The relative areas and perimeters of all leaf and ribbon morphotype samples plotted against each other. Green denotes the leaf morphotype, and the black denotes the ribbon morphotype. The blue line in each case is the expected curve given a perfectly regular polygon.

In order to assist with visualisation, figure 4.15 shows the different polygons coloured with different colours. Upon examination, it seems as though there are lines of "irregularity" that cut through patches of regularity (hexagons).

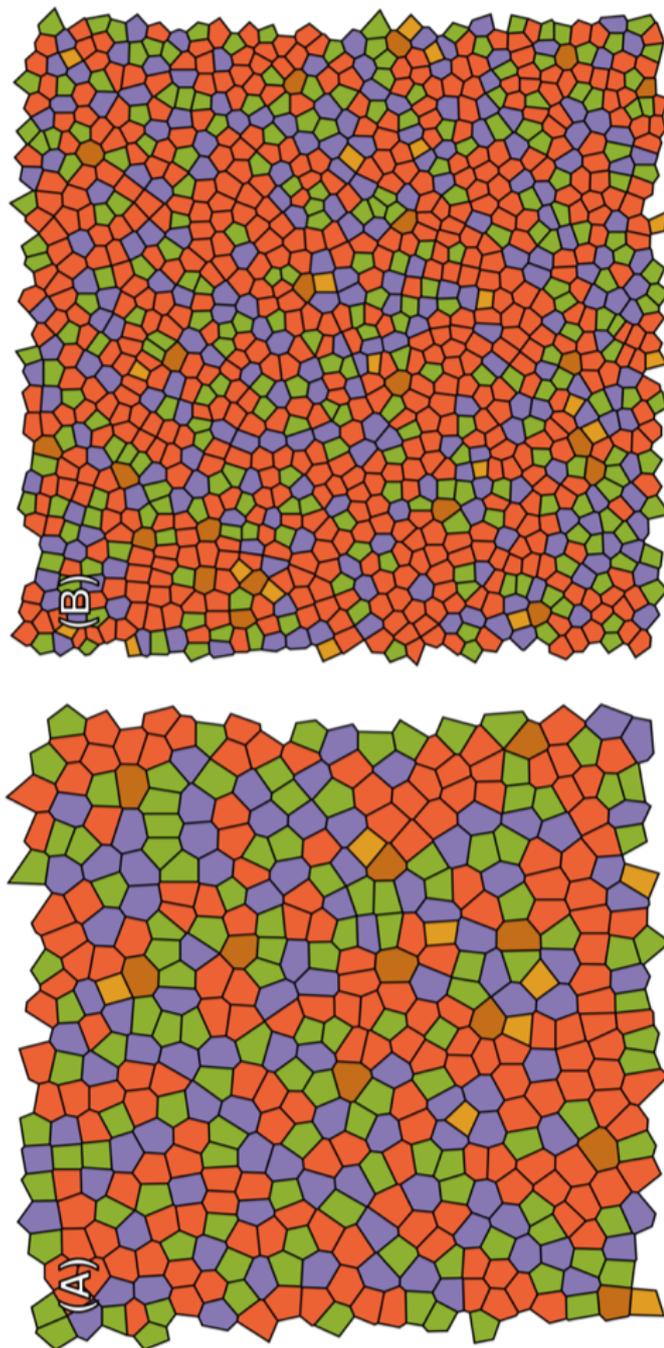


Figure 4.15: Voronoi tessellation diagrams of a (A) leaf and (B) ribbon morphotype samples. Different polygons are represented by different colours: quadrilaterals = gold, pentagons = green, hexagons = orange, heptagons = purple and octagons = brown. Diagrams generated in Wolfram Mathematica. See Appendix A for code.

Using the plot profile tool (Ferreira and Rasband, 2012) on Fiji image analysis software (Schindelin et al., 2012), the vertical pixel intensity across the horizontal plane of an image can be plotted. This was done for both leaf and ribbon morphotypes and shown below in figure 4.16.

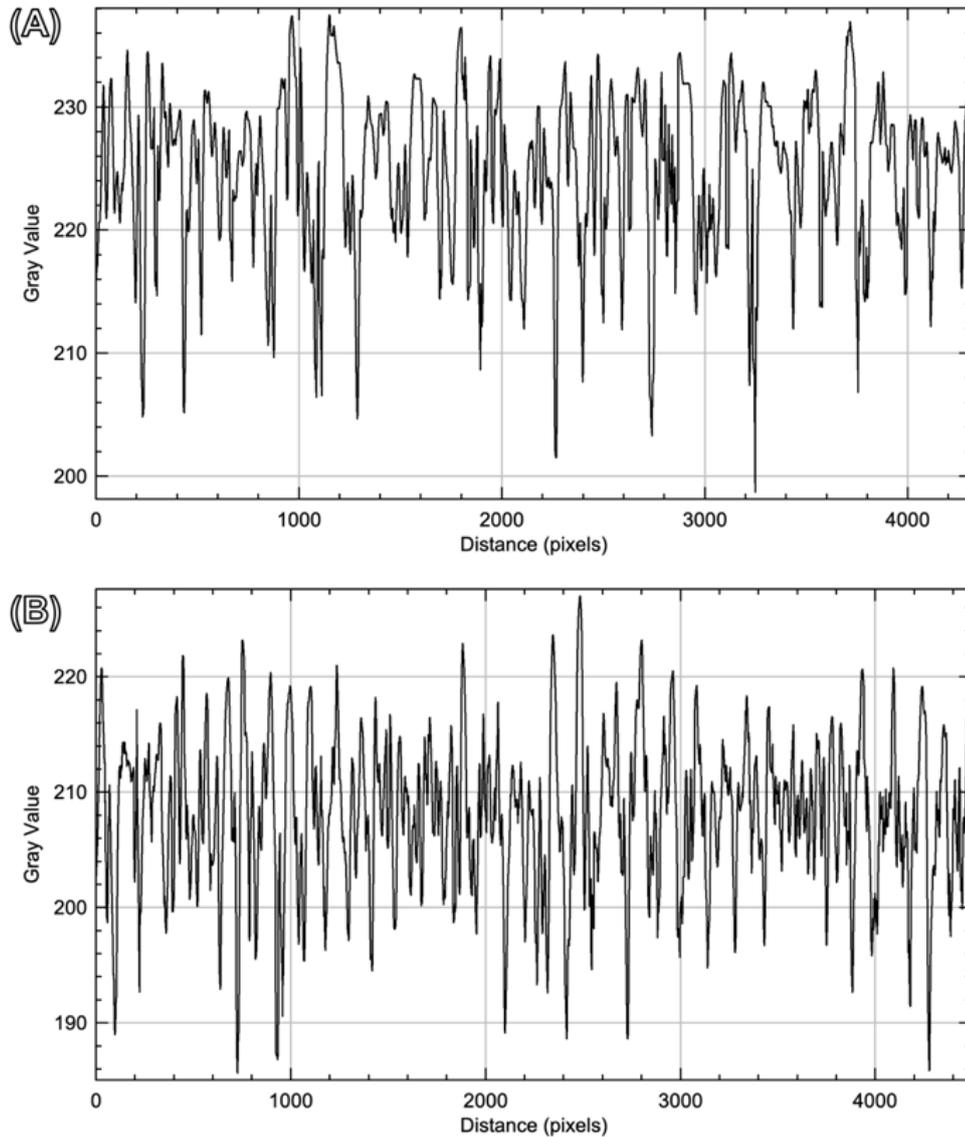


Figure 4.16: The profiles of Voronoi tessellation diagrams generated from (A) leaf and (B) ribbon morphotype samples. Higher peaks indicate higher pixel density.

Profiles of the four quarters of the ribbon morphotype diagrams were also plotted, which encompasses a roughly equivalent number of cells as seen in a whole leaf morphotype sample. This is shown below in figure 4.17 by plotting the ribbon morphotype in four quarters.

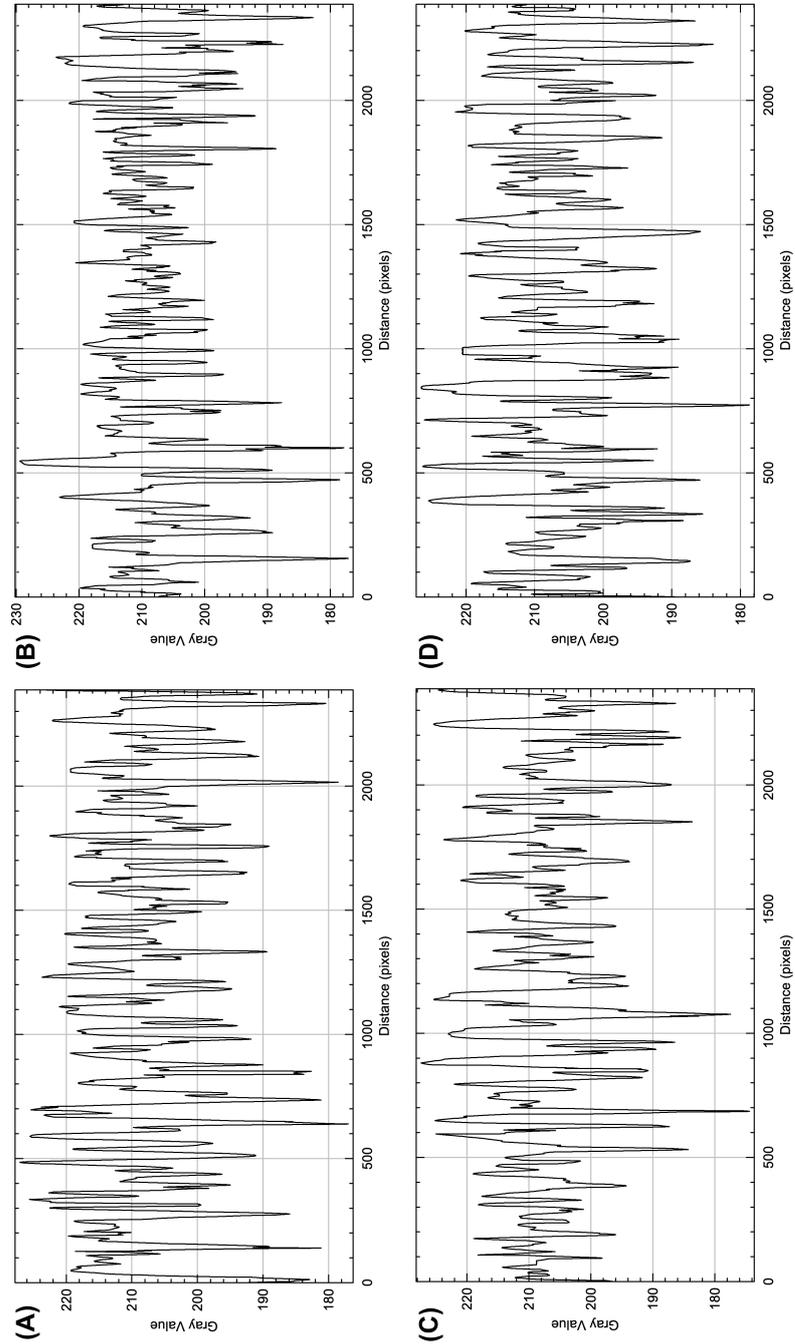


Figure 4.17: Profile plots of the Voronoi tessellation diagrams of the four quarters of a ribbon morphotype sample: (A) top left quarter, (B) top right quarter, (C) bottom left quarter and (D) bottom right quarter. Higher peaks indicate higher pixel density.

4.4 Discussion

As aforementioned in chapter 1, *Ulva* and the Chlorophyta in general show diverse intra- and inter-specific morphologies. The fundamental subject of investigation here is how different *Ulva* morphologies arise. More specifically, is it possible for different bottom-up morphogenesis to drive the generation of observed morphotypes (in this case the leaf and ribbon morphotypes)? In order to answer this question, a set of models were made to compare with observed data, as explained in the methods section of this chapter.

The statistical analysis of the first phase of modelling (Models 1 to 4) confirmed that they function as intended and set the baseline to what can be expected from the comparative analyses of the second phase of modelling (Models 5 to 7) and primary data.

50% jitter was initially chosen based on visual observations, in that 50% jitter was the starting point from which similarities to actual measured data could be seen, as shown in figure 4.11 above. The second phase of modelling showed that beyond 50% jitter, significant differences in edge counts are found when comparing with a hexagonal honeycomb tessellation pattern (Model 1 and 2). The decision to keep 50% as the threshold model was upheld having seen similar results in preliminary statistical analyses.

Models using 1%, 5%, 15%, 25% and 30% were also investigated, however their use in the analysis was excluded due to their lack of novel contribution. Models using 1%, 5%, 10% and 15% jitter largely showed the same results, and the robustness of the framework for comparative analysis was not improved by adding these models. The only notable data obtained from this experimentation of varying jitter is that models using 1% and 5% jitter showed no significant difference in edge lengths when comparatively analysed against regular and squashed hexagon models (Model 1 and 2). A jitter of 10% is the cut-off point below which no significant differences in edge lengths are found. Models using 25% and 30% jitter predictably overlapped with models using 20% jitter and added no more useful information.

Although they are useful as tools for establishing basic background hypotheses, there are many limitations to manipulation models such as those used in this chapter. The primary drawback is that manipulation models can only suggest the confirmation of hypotheses regarding observed patterns, not give any information on the reasons and mechanisms

underlying the generation of such patterns. This also ironically limits what we can manipulate. That is to say that apart from confirming observational hypotheses, manipulation of conditions for example, cannot be done. The most notable factor to be investigated as a result of this preliminary manipulation modelling, is the presence of a clear biological artefact from cell division. Due to the fact that the sheets of cells that we see have a biological mechanism for generation, i.e. mitosis, there will always be an artefact of this generation. This means that certain pairs or groups of pairs of cells will show signs of equal division, resulting in a clear centre division boundary. Figure 4.12 above is an annotated copy of figure 4.11 above, to illustrate some of the centre lines.

The initial analyses alone suggest that the polygon distributions are similar in the 50% jitter model and the ribbon and leaf morphotype samples. However, there are clear visual differences in the Voronoi tessellation diagrams due to this biological artefact. This absence of the biological mechanism for generation in the models used in this chapter can be addressed using a different type of modelling. Though manipulation models such as those used in this chapter are interesting, useful and informative, the limitations of this type of modelling are clear. The ideal model for bottom-up proliferation is a generative model, wherein given a set of parameters a single cell divides, and the daughter cells divide and continue to divide in such a way that would result in typical morphogenesis. This type of model will allow for complex modelling by incorporating known information regarding morphogenesis from experimental data collection. For example, the manipulation of conditions can be accounted for, and using primary experimental data, generative models can model growth with variations in temperature, light, mixing and nutrient concentrations, amongst others. Generative models, however, are notoriously difficult to create primarily due to the fact that there is a myriad of factors that affect growth and generation that needs to be taken into account. Due to the specificity of such factors, generative models are often limited to isolated two-dimensional scenarios in the early stages of growth in a single or groups of closely related organisms with similar underlying morphogenetic mechanisms.

The difficulties regarding the coding of a generative model means that such models have only been developed extremely recently with the advancement of computer sciences and the emergence of interdisciplinary approaches to biology. Consequently, frameworks only exist for the most well studied organisms, which are always land plants. Models such as VirtualLeaf, VPTissue, Cellzilla and CellModeller are frameworks developed by well-

funded groups of expert coders over multiple years, and are not, by design, intended to model morphogenesis in organisms such as *Ulva*. Therefore, novel adaptation of an existing generative model framework, not the development of a new framework, for *Ulva* morphogenesis would be the only feasible method to attempt in the absence of equivalent time, expertise and funding, though it is still ambitious and predictably difficult.

However, with the given data, an effective method of investigating differences like these is by using Quantile-Quantile (Q-Q) plots. The edge lengths of the jitter models, leaf and ribbon morphotype samples were plotted as shown in figure 4.13 above.

A Q-Q plot can indicate whether or not the distribution is generated by the same mechanism, as shown in figure 4.13 (A), where two iterations of a 50% jitter model can be plotted perfectly against each other, lying perfectly on the line $y=x$. A Q-Q plot is also a sensitive enough method to distinguish between different jitter models, such as in figure 4.13 (B), where a 20% jitter model is plotted against a 50% jitter model and the distribution clearly strays from the line $y=x$. Figure 4.13 (C) and (D) show that in the observed data, there is a consistent deviation from the line $y=x$, suggesting that the leaf and ribbon morphotypes are generated by the same mechanism and differs in the same manner from the mechanism that generated the 50% jitter model. This is confirmed by figure 4.13 (E), where it can be seen that the leaf and ribbon morphotypes can be plotted to lie directly on the line $y=x$.

This is further to the initial results that suggested that the leaf and ribbon morphotypes are generated by the same underlying mechanism. The edge length and area showed significant differences, expected by the mean cell density of the ribbon morphotype being 73.9% higher than the leaf morphotype. The edge count or polygon distribution, however, remained remarkably similar, and showed no statistically significant differences. The polygon distribution of the regular honeycomb 50% jitter model (Model 7) is also strikingly similar to the observed data, and it also showed no significant differences to neither the leaf nor the ribbon morphotype.

The comparison of the leaf and ribbon morphotype samples number 2 was an exception, however, showing no significant differences in any characteristics. The most likely explanation of this result is a sample anomaly. The raw sample data showed striking visual similarities in sizes and shapes of cells, unlike most of the other samples in which a clear visual difference could be noted. The initial analysis confirms this, showing 40.72% fewer

cells than the mean sum of cells in ribbon morphotype samples, with the cell density of sample 2 being more in line with leaf morphotype samples than the ribbon morphotype samples.

The polygon distribution of both leaf and ribbon morphotypes showed a heavy preference for hexagons in *Ulva*. Most importantly, the distribution itself was almost identical as can be seen above in figure 23, indicating that the pattern was generated by the same underlying mechanism, furthering the hypothesis that morphogenesis in *Ulva* is driven by bottom-up proliferation determined by physical and mathematical laws. This result is intuitive and expected based on early observations of the dominance of hexagons in polygon distribution patterns of plants. This same pattern has been observed in tissues such as plant parenchyma, *Drosophila* eyes and now *Ulva*, suggesting the same underlying mechanism of cell proliferation. That is to say, that cell proliferation patterns in *Ulva*, are also a result of a bottom up mechanism following fundamental physical and mathematical laws.

It is logical that these local interactions can generate global patterns, but each of these tissue types show a different overall morphology. How could this be explained if the underlying mechanism of the local interactions is the same? Other tissue types have top-down control, with complicated multicellular genetic architecture dictating the location, but not necessarily the proliferative mechanism of bottom-up generated tissues, with certain exceptions for specialist tissue types such as neuroepithelia (Sánchez-Gutiérrez et al., 2017). The extent of bottom-up proliferation, in morphology determination is a key question that can be studied in *Ulva*, given its lack of top-down control.

The basic principle for the next line of inquiry was that the overall morphology can be changed by changing the distribution and the sizes of the constituent polygons in the cell tessellation pattern. As a simple example, given a cell tessellation pattern, adding more heptagons and increasing their average size will give a different morphology than adding more pentagons and increasing their average size.

The first part of this inquiry has already been answered, with the leaf and ribbon morphotypes showing very similar polygon distributions. Therefore, the logical focus is on the sizes of the cells, or more specifically: are there classes of polygons with smaller or larger average areas, indicating a different pattern of relative area distribution?

This line of inquiry can be explored by plotting the relative area and perimeters of different polygons in both the leaf and ribbon morphotypes against each other. Figure 4.14 above shows this for quadrilaterals, pentagons, hexagons, heptagons, octagons and nonagons of all leaf and ribbon morphotype samples.

The first observation of note is that there are few differences in regularity, relative area and perimeter of quadrilaterals, pentagons and hexagons between the leaf and ribbon morphotypes, with the majority of the data points overlapping. Furthermore, the quadrilaterals, pentagons and hexagons show more irregularity in shape than the heptagons, octagons and nonagons for both morphotypes. Finally, the ribbon morphotype shows heptagons, octagons and nonagons of larger relative area and perimeter.

A Kolmogorov-Smirnov test on the edge lengths of all the hexagons in the leaf and ribbon morphotypes shows no significant difference ($p = 0.5847$) between the morphotypes, suggesting that the hexagons in both morphotypes are similarly regular. This is unsurprising considering the tendency towards hexagons for regularity. Repeating this for the pentagons, heptagons and octagons, there is a significant difference ($p = 0.0122$), suggesting that the regularity of these polygons is significantly different in the two morphotypes. This is also unsurprising given the results shown in figure 4.14 above. Based on this, a new hypothesis was constructed: the *Ulva* thallus consists of a sheet of similarly regular hexagons, with “impurities” of other polygons that gives rise to its morphology. Figure 4.15 above shows the Voronoi tessellation diagram generated from a leaf morphotype sample and a ribbon morphotype sample, with different polygons being coloured differently.

As aforementioned, there are distinct lines, artefacts of cell division, visible in both samples. These lines potentially correspond to areas of regularity and impurities; there are clusters of hexagons, and lines of the impurities. This idea of impurities inspired the next test, wherein the profiles of the Voronoi tessellation diagrams of the two morphotypes were analysed using Fiji (Schindelin et al., 2012). The profile plot of an image, by default, shows the vertical pixel intensity across the whole horizontal selection. This means that the more regular an image is, the more similar the peak heights will be, and the more evenly spaced said peaks will be. Figure 4.16 above shows the profiles for leaf and ribbon morphotype samples.

The leaf morphotype shows a more regular profile overall, with mostly similar sized and

spaced peaks throughout the profile, though there are relatively small and relatively large peaks spread throughout with no particularly discernible pattern. The ribbon morphotype, however, shows a clearer pattern of three clusters of large peaks with a cluster of small peaks in between each. This suggests that the leaf morphotype shows higher regularity. To ensure that the results were not an artefact of higher cell density in the ribbon morphotype, the results were checked by separately plotting the profile of the four quarters of the ribbon morphotype diagram which showed similar patterns of clusters of large and small peaks, as shown in figure 4.17 above. However, it should be noted that due to the biological noise from cell division, the profiles are very difficult to interpret and analyse. Nevertheless, these results are promising, showing that there are differences in regularity between the two morphotypes.

The first key hypothesis of this chapter was that different morphologies of *Ulva* can arise from different bottom-up cellular proliferation patterns. This hypothesis can safely be accepted as it is strongly supported by the evidence given here of fundamental physical distribution patterns, though the generative mechanism of the morphogenesis is yet to be elucidated.

The second key hypothesis was that cell tessellation patterns in simple multicellular organisms such as *Ulva* can be replicated by applying random deviations, or “jitter” in the Voronoi generators of a regular hexagonal honeycomb pattern. This must be rejected due to a lack of evidence supporting the hypothesis, as well as preliminary image analysis evidence presented here against it. However, the evidence against this hypothesis gives rise to an even more interesting question: how can we study the irregularities in the cell proliferation patterns?

Then, the logical next step in inquiry is: how do we study the regularity of the cell proliferation patterns in *Ulva*? Given the results of this chapter, two new hypotheses arise. The first is that *Ulva* cellular proliferation patterns are space-filling arrangements generated from bottom-up physical proliferation and therefore will follow fundamental physical and mathematical laws to be as close to an optimally space-filling arrangement as possible. The second is that *Ulva* cellular proliferation patterns can be described as areas of regularity permeated by lines of irregularity where cell division occurred. These will be the two hypotheses to be investigated in the next chapter.

Fundamental mathematical and physical mechanisms underlying optimal distribution in *Ulva* cell growth and proliferation patterns

5.1 Introduction

Ulva is a cosmopolitan genus of green algae that uniquely displays multicellularity in the Chlorophyte clade, and is found in both seawater and brackish waters (Lee, 2018; Messyasz and Rybak, 2011; OBIS, 2019; Tanner and Guiry, 2019; van den Hoek et al., 1995). *Ulva* displays extremely simple multicellularity with only three cell types: rhizoid, blade and stem, meaning that in the case of anchored *Ulva*, almost the entirety of a growing thallus is simply a bilayer sheet of one cell type, and in the case of free-floating *Ulva*, the entirety. Without differentiation into tissues, there are no active mechanisms of cell signalling or overarching cellular architecture coordination within *Ulva* itself.

Ulva cells are uninucleate, contain a single cup shaped chloroplast at the surface edge and a single pyrenoid, isodiametric in shape, and are distromatic (arranged into a bilayer) (Sahoo and Seckbach, 2015; Wichard et al., 2015). A sheet of *Ulva* cells, displays a pattern that appears strikingly geometric and mathematically generated, as shown previously in figure 4.5 above, much like many other epithelia and epidermises.

There is a plethora of mathematical patterns that can be observed throughout nature, including the geometric patterns found in epithelia and epidermises. From the hexagonal proliferation patterns in *Drosophila* eyes, honeycomb construction and rock formations, to the phyllotaxis in sunflowers, pinecones and Romanesco broccoli, geometric patterns following fundamental mathematical and physical laws are pervasive throughout nature.

Chapter 4 discussed fundamental physical and mathematical laws determining the space filling patterns observed in *Ulva*. One of the key findings of chapter 4 was that there seem to be areas of regularity and irregularity, potentially signalling areas of differential cell division. Therefore, the focus of this chapter is on the mathematical and physical laws that determine optimal distribution, and *Ulva*'s deviation from such laws.

There are a number of mathematical laws that determine an optimal distribution. Traditionally, three laws have been used for biological systems as indicators for optimal distribution of cells, though there are a number of other laws and principles that would also suggest the distribution in question is optimal. The first these laws is Euler's Law, which states that in a tissue with a large number of cells, the average number of neighbours of any given cell will be six (Brookes, 2018; Sánchez-Gutiérrez et al., 2017). The second is Lewis' Law, which states that the area (A_n) of each polygon increases linearly with the number of sides (n) (Chiu, 1995; Lewis, 1931). The third is Aboav-Weaire Law, which states that polygons with more sides will have fewer-sided neighbours and *vice versa* (Chiu, 1995).

In order to contextualise these laws, as well as the additional laws and principles of optimal distribution, it is helpful to use the example of soap bubbles. Soap bubbles constantly undergo minute changes in shape in order to counteract the changes in external forces, in the process of balancing internal and external forces. However, in an ideal situation with zero disturbance, a singular soap bubble will take the form of a perfect sphere, exhibiting the simple principle of area minimalization.

The first known scientific investigation of the mathematical principles governing the geometry of soap bubbles and soap films were conducted more than a century ago by Belgian physicist Joseph Plateau, who conducted a number of physical experiments and formulated a theoretical framework to address his observations (Almgren and Taylor, 1976). Following his observations, Plateau proposed the following laws that soap bubbles will abide by:

1. The mean curvature at any point of a single soap film is constant.

2. Each side of the bubbles will meet in triplets, to form the most mechanically stable junctions of $\cos^{-1}\left(-\frac{1}{2}\right) = 120^\circ$ (now referred to as a Plateau border).
3. Plateau borders will meet in quadruplets at a vertex to form tetrahedrons, at $\cos^{-1}\left(-\frac{1}{3}\right) = 109.47^\circ$. This means that in an optimally distributed sheets of bubbles, they will tessellate into hexagons and have tetrahedral vertices in a three-dimensional array.

If we isolate this set of rules and restrict it to a two-dimensional setting, it simply leads to an optimal distribution of a sheet of hexagons. More than a century after Plateau published his findings, Taylor (1976) proved Plateau's laws using geometric measure theory. Unsurprisingly, the millennia-old Honeycomb Conjecture states that a tessellation of regular hexagons is the least perimeter method of enclosing infinitely many unit areas, as proven by Hales (2001). The combination of the two mathematical principles leads to the conclusion that a regular hexagonal tessellation indeed is the most mathematically and biologically optimal cellular tessellation pattern, allowing for the most resource-efficient, space filling and mechanically stable configuration.

The late 19th Century saw a variety of hypotheses proposed regarding the position and direction of cell wall synthesis in plants. Inspired by the work of Plateau, Berthold (1886) and Errera (1888) proposed the Principle of Minimal Area, though Berthold and Errera differed on a key issue (Besson and Dumais, 2014; Sahlin and Jönsson, 2010).

Errera, in his hypothesis, proposes that cells will behave much like soap bubbles, in that the cell walls will “assume the form which a weightless liquid film would assume under the same conditions”, meaning that the two men agree on the general principle (Besson and Dumais, 2014). However, it seems that Errera purposefully omits a principle proposed by Berthold in his work, regarding the inevitability of absolute minimisation of area when cell walls form, provided the same surface tension as equivalent soap films. This omission was later clarified by the work of one of Errera's students, de Wildeman (1893), wherein he discusses the possibility of multiple division planes for a given cellular morphology. In his work, de Wildeman states that the system will be in a stable equilibrium if the area is simply the relative minimum as clearly stated by Plateau, rather than the absolute minimum, given that “mean curvature is and it attaches itself to the surrounding wall at right angle” (Besson and Dumais, 2014).

The discussion regarding the position and direction of new cell wall formation in plants has been ongoing since the late 19th Century, and there have been two main hypotheses. Firstly, the soap bubble hypothesis, or the Principle of Minimal Area, proposed by the aforementioned works of Errera, Berthold and de Wildeman, amongst others. Secondly, it was Hofmeister (1863) who initially proposed that cells divide perpendicular to the main direction of growth, which, incidentally, correlates with the main direction of strain in many cases. Sachs (1878) indirectly supported this with his findings that new divisions occurred perpendicularly to older walls (Sahlin and Jönsson, 2010).

Plant cell division is a discrete process to animal cell division in a number of notable ways. Firstly, plant cells are non-motile, therefore no cell migration can occur. Once a cell has divided and formed new cells, there is little to no compromise in position and orientation, apart from cell elongation (Wright and Smith, 2008). Therefore, it is vital that cell division in plants is orientated and positioned correctly, and occurs at the right time, in order to ensure normal morphogenesis and proper functioning and development. In land plants, there are hormonal signalling mechanisms that allow for asymmetric cell division in addition to the proliferative symmetric cell division, wherein a mother cell divides to give two identical daughter cells (Facette et al., 2019). Asymmetric cell division has a specific and crucial role in development of new cell types and cell differentiation. Given the nature of the role, asymmetric cell division is triggered by specific transcription factors and signalling pathways (Rasmussen and Bellinger, 2018). *Ulva* seems to lack such signalling pathways, and indeed any overarching hormonal control. Therefore, *Ulva* only shows symmetric cell division for cell proliferation, and it can be argued that it is consequently even more vital for robust symmetric cell divisions to occur in *Ulva*, as there are no mechanisms for correction or compensation.

Experimental studies have previously identified that hexagon enrichment is indeed pronounced in a variety of organisms. The first notable experimental studies on biological tissues was conducted in the early 20th Century by F. T. Lewis, after whom the aforementioned Lewis' Law is named. Lewis found in his examination of cucumber epidermis that the tessellation patterns heavily favour hexagons, with pentagons and heptagons following thereafter (Lewis, 1926, 1928). This hexagon enrichment was also found in *Ulva*, in the results of chapter 4.

Much of the investigation of optimum distribution using biomathematics has addressed the

problem as a two-dimensional one, thereby assuming a sheet of polygons. This assumption works well on the whole when investigating epithelial cells and works better still when investigating an organism such as *Ulva spp.*, which is a simple bilayer of a cells. Without differentiation into tissues, there are no active mechanisms of cell signalling or overarching cellular architecture coordination within *Ulva* itself, meaning that *Ulva spp.* are free from many of the biological mechanisms that complicate the morphology and morphogenesis. This allows for an investigation more focussed on the mathematical and physical principles that underly the cellular morphology and morphogenesis. Thus, a two-dimensional examination of polygon tessellation patterns in *Ulva spp.* cells will allow for an experimental investigation of the mathematical and physical principles that govern biological cellular pattern formation that is unprecedentedly pure.

When examining patterns of a polygon tessellation, there are various mathematical tools that have been developed for the task. As introduced in chapter 4, a popular method, and the dominant method in biological sciences, is Voronoi tessellation. Voronoi tessellation is a method of investigating space filling patterns, and has been used widely in a variety of disciplines ranging from epidemiology to electrical engineering (Okabe et al., 2009). Voronoi tessellation diagrams have previously been used to model and analyse patterns of cells in a biological context, from human biceps to *Arabidopsis thaliana* roots (Barrio et al., 2013; Sánchez-Gutiérrez et al., 2017).

Voronoi tessellation diagrams are generated from a given set of centre points or Voronoi generators. A Voronoi region is generated by calculating the distance from a generator (A) to its nearest neighbouring generator (B), and drawing a line to bisect that distance equally, termed Voronoi edge, so that the line is equidistant from seed A and B. This process is repeated for all neighbouring generators, and so polygons are generated encapsulating the generators, which are termed Voronoi regions. The terms Voronoi generator, Voronoi region, Voronoi edge and Voronoi tessellation from henceforth will be referring to the above definitions. Mathematical definition of Voronoi tessellations used in this thesis are given above in chapter 4 and will be omitted here.

Due to the extremely labour-intensive nature of calculating Voronoi cells, Voronoi tessellation diagrams did not see any use in disciplines such as biological sciences where thousands of calculations have to be completed for a single sample. Recent advances in technology has made it feasible for use in such contexts, and Honda (1978) is widely credited for the

first use of Voronoi tessellation in a biological context, more than a hundred years after Voronoi established the method (Bock et al., 2010).

A key hypothesis that arose from the results of chapter 4 was that *Ulva* morphologies are a result of bottom-up physical proliferation characterised by regions of regularity (hexagons) permeated by lines of “irregularities” (quadrilaterals, pentagons, heptagons and octagons), which occur at areas of cell division. However, the analysis was difficult due to the biological noise from cell division. This chapter will build upon this using a technique called Centroidal Voronoi Tessellation.

In the context of optimum distribution, Centroidal Voronoi Tessellation is the best method to investigate the closeness of a Voronoi Tessellation pattern to its most optimum arrangement possible. Voronoi diagrams can be iteratively “relaxed” using various algorithms. Lloyd’s algorithm is the first used for this purpose and remains the most commonly used for its observed speed, simplicity and for comparative purposes. First proposed in 1957, but published in 1982 after the completion of numerical calculations, Lloyd’s algorithm was originally developed as a technique for quantization in pulse-code modulation (PCM) (Lloyd, 1982). Since its development, Lloyd’s algorithm, also known as k-means clustering, has been used widely across various fields including computer vision, astronomy, statistics and biology (Arthur and Vassilvitskii, 2006; Lu and Zhou, 2016). Though Lloyd’s algorithm was developed more than 60 years ago, it remains one of the, if not the, most widely used clustering algorithms, as well as being one of the most widely used algorithms for data mining due to its simplicity and empirical performance (Lu and Zhou, 2016; Wu et al., 2008). However, despite its widespread use, there is limited information available on meaningful theoretical bounds of Lloyd’s algorithm. It was only established relatively recently that computation of Lloyd’s algorithm can require a superpolynomial running time (Arthur and Vassilvitskii, 2006).

In the application of Lloyd’s algorithm to a biological context, Centroidal Voronoi Tessellation is the one of the most relevant. In this process, each iteration of the diagram uses the centre of mass (centroids) of the Voronoi diagram generated from the Voronoi generators of the previous iteration, as the new Voronoi generators. This process is repeated until the centroids overlap with the Voronoi generators, at which point there is no more possible relaxation step. When a Voronoi diagram reaches this stage, it is termed a Centroidal Voronoi Tessellation diagram. The difference between a standard Voronoi diagram and a

Centroidal Voronoi Tessellation diagram is illustrated in figure 5.1 below.

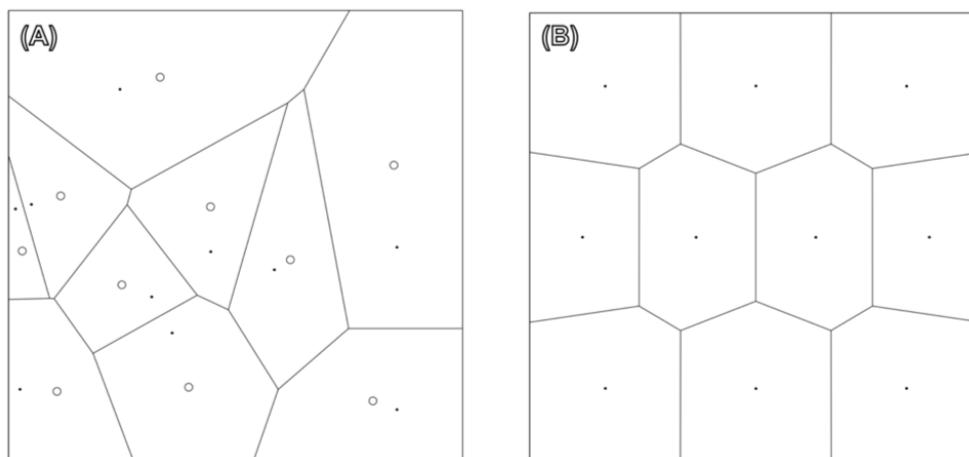


Figure 5.1: Example of (A) a Voronoi tessellation diagram and (B) a Centroidal Voronoi Tessellation diagram (right). In (A), the dots represent the Voronoi generators and the circles represent the centroids. Note that the (B) does not require this distinction as the two are exactly the same (Du et al., 1999).

As is theoretically predicted, the most favourable arrangement is that of hexagonal honeycomb tiling as this is the most ideal arrangement in space filling efficiency. Figure 5.1 illustrates this with the two centre cells, but the pattern is not immediately obvious for a key reason. The outermost cells, that is to say every cell in contact with the border of the image, will be limited by the nature of the inflexible constraints exerted by the border. Therefore, the cells in contact with the border are a matter for consideration when analysing any Voronoi Tessellation data. The exclusionary definition used in this experiment for analysis is identical to that used in Chapter 3 and is once again outlined below in the methods section. Using Centroidal Voronoi Tessellation diagrams will give a measure of how optimally distributed the cells are in *Ulva*.

Finally, the work in chapter 4 was conducted considering the thalli as a sheet of cells, each of which is assumed as an isolated unit. However, this is not a biologically accurate assumption as these patterns require a generative mechanism: cell division. This means that the majority of the cells will be inextricably linked with other cells as mother cells and daughter cells. These would essentially be patches of cells, and would be consistent with the findings of chapter 3, where patches of regularity were found. If this artefact was indeed formed due to cellular proliferation occurring within patches of cells, this would likely lead to observable physical traits, such as changes in thicknesses or density of cell walls between clusters. This was already partly observed during the initial batch of confocal

image collection, with the general pattern being noted as brighter signal, indicating more cell wall fibres for the dye to bind to, around the borders of certain clusters of cells. These clusters of cells were the final focus of this chapter and any indication for their presence was explored.

5.2 Methods

5.2.1 Image Collection and Analysis

Confocal images were collected using a Leica SP5 confocal laser scanning microscope, at 1024 x 1024-pixel resolution in order to standardise data collection and analysis. This, in combination with the standardised scanning parameters, allowed for the best compromise between image quality and a realistic timeframe for data collection. All of the scanning parameters used in this experiment are consistent with those aforementioned in detail in section 2.4.7.

Twelve leaf morphotype samples and twelve ribbon morphotype samples were stained with the cell wall dye Direct Yellow 96 using the protocol previously outlined in chapter 4, and 48 images were collected in total: 12 of the cells and 12 of the patches of both morphotypes. The centre points were isolated automatically through image analysis software and were used as the initial Voronoi generators. The resulting Voronoi diagrams were the focus of investigation in chapter 4, however the focus in this chapter was to obtain information regarding centroids.

Two controls were used for this experiment; one for the leaf morphotype and one for the ribbon morphotype. Voronoi generator coordinates were randomly generated using a pseudorandom number generator (PRNG). These coordinates were then used to generate and iteratively relax Voronoi tessellation diagrams in an identical manner to the coordinates obtained from the experimentally collected data. The number of randomly generated Voronoi generators were 459 and 828 for the leaf and ribbon morphotypes, respectively. These are the mean number of cells measured in Chapter 3, in a 1,024 x 1,024-pixel section of an *Ulva* thallus ($n = 12$).

For the analysis of certain attributes such as polygon distribution and change, a simple exclusionary definition of cells to be analysed, Valid Cells, was used, with Valid Cells being

defined as:

$$\text{Valid Cells} = \text{All Cells} - \text{Edge Cells},$$

Where Edge Cells = Any cell in contact with the edge of the frame.

This is visually represented in figure 5.2 with Valid Cells being represented in green and Edge Cells being represented in red.

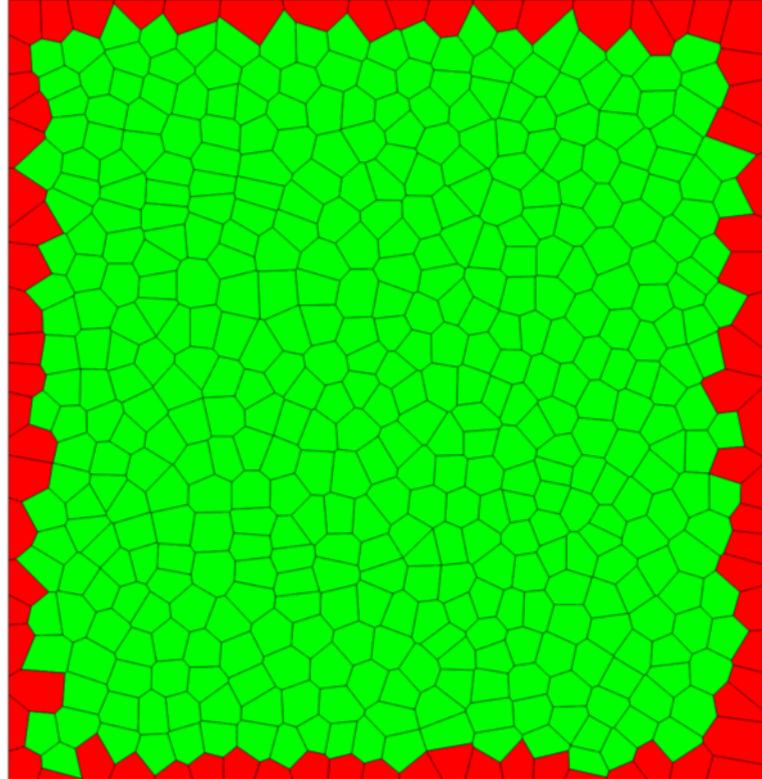


Figure 5.2: Voronoi Diagram obtained from a leaf morphotype sample, coloured to visualise data. Red represents the Edge Cells, which were excluded from various analyses. Green represents the Valid Cells included in various analyses. Diagram generated in Wolfram Mathematica. See Appendix A for code.

Lloyd's algorithm was iteratively applied to the Voronoi generators, until Centroidal Voronoi Tessellation was reached. The parameters for accepting a Centroidal Voronoi Tessellation was set to a Euclidean distance tolerance of 0.1, in the 1024 x 1024 frame. This means that if every single one of the 1,048,576 pixels in the frame has changed by less than 0.1 pixels compared to the previous iteration, the diagram is determined to have ceased changing further, and the result is accepted as the final Centroidal Voronoi Tessellation diagram. Given the computing resources available and finite time available, the distance tolerance was chosen as the best compromise.

The history of relaxation, namely the changes in the Voronoi generators coordinates, and the resulting Voronoi diagrams and centroids of each generation were tracked. The history of polygon distribution was the main focus of this analysis.

5.3 Results

The number of iterations required to reach Centroidal Voronoi Tessellation for the leaf and ribbon morphotypes are shown in table 5.1 below. The means, standard deviations and coefficients of variation were all calculated excluding the control.

Table 5.1: Iterations required to reach Centroidal Voronoi Tessellation (CVT) in both leaf and ribbon morphotypes. The ribbon morphotype sample 2 has been highlighted as an anomalous result, as addressed in the discussion of chapter 4, as well as again in the discussion of this chapter.

Sample Number	Iterations required	
	Leaf	Ribbon
1	455	306
2	514	423
3	309	327
4	433	560
5	431	321
6	350	528
7	396	395
8	442	596
9	421	426
10	499	305
11	360	331
12	412	325
Control	322	517
Mean	419	402
Standard Deviation	59.17	109.95
Coefficient of Variation	14.14	27.36

The mean number of iterations required to reach Centroidal Voronoi Tessellation was 419 for the leaf morphotype and 402 for the ribbon morphotype. The mean for the ribbon morphotype was calculated excluding sample R2, due to the anomalous nature of the sample aforementioned in chapter 4 and will be addressed again in the discussion.

Figure 5.3 below shows the comparison between the first and last iteration of applying Lloyd's algorithm, to a leaf morphotype sample.

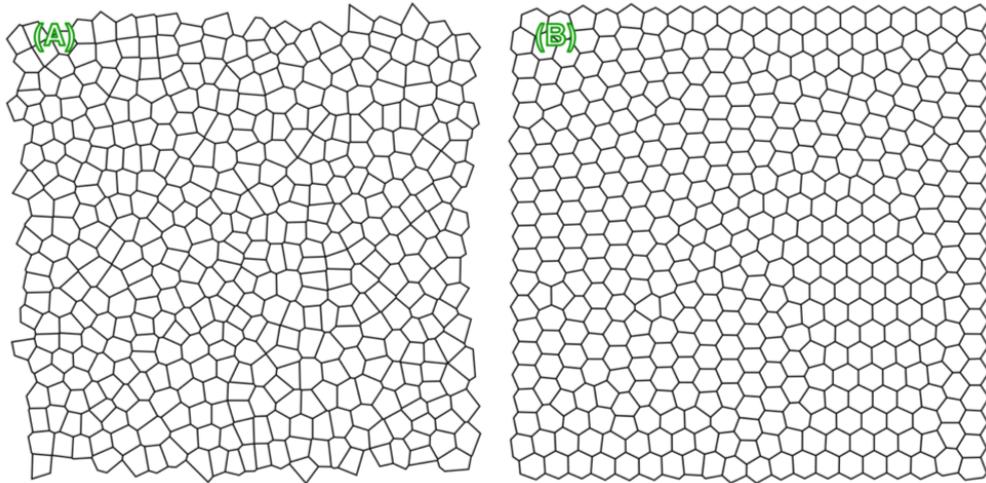


Figure 5.3: Visual comparison of the Voronoi tessellation diagrams (A) before and (B) applying Lloyd's algorithm to a leaf morphotype sample. Diagrams generated in Wolfram Mathematica. See Appendix A for code.

The mean percentage of polygon distribution of the leaf morphotype is shown in table 5.2 below.

Table 5.2: The mean start, end and total change in percentage in polygon distribution, between the first and last iteration in the leaf morphotype.

Number of Edges	Start Percentage	End Percentage	Percentage change
3	0	0	0
4	1.59	0	-100
5	27.03	8.23	-69.60
6	47.05	83.32	76.62
7	21.34	8.45	-60.50
8	2.85	0	-100
9	0.15	0	-100
10	0	0	0

As can be seen in table 5.2 above, the only polygon that shows net enrichment is the hexagon.

Figure 5.4 below is a visualisation of the polygon distribution changes throughout the iterations of a leaf morphotype sample.

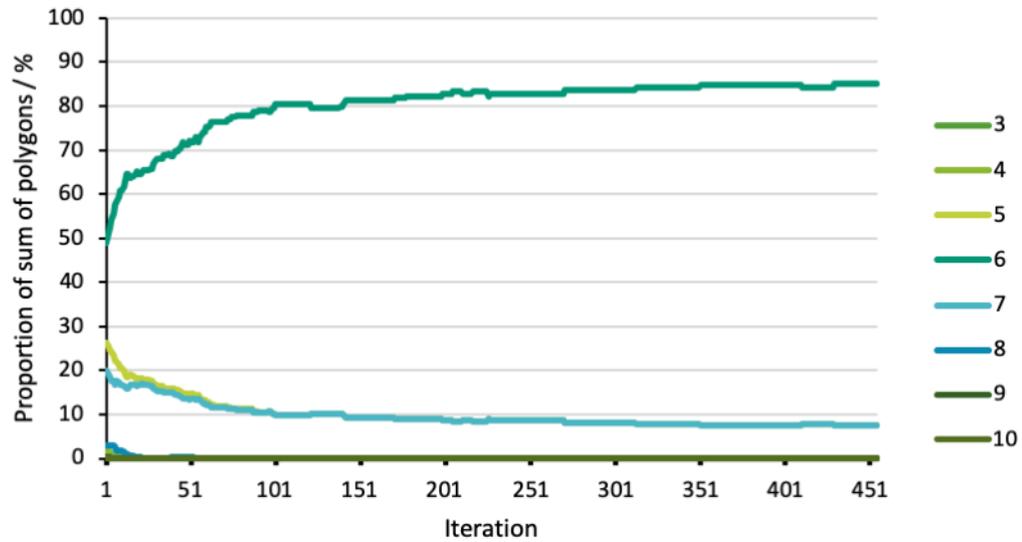


Figure 5.4: Change of proportion of each polygon out of total sum of polygons tracked throughout Lloyd's iterations in a leaf morphotype sample. The numbers represent the number of edges in the polygon, i.e. 3 = triangles, 4 = quadrilaterals and so forth.

Figure 5.4 above illustrates how rapidly the proportion of the sum each polygon occupies changes from iteration 1 to 100, after which the rate of change diminishes. The mean percentage change for the 455 iterations of this sample are shown below in table 5.3 below.

Table 5.3: Mean percentage change of each polygon between the given ranges of iterations for the leaf morphotype.

Number of Edges	Mean percentage change			
	Iteration 1 to 10	1 to 100	155 to 455	1 to 455
3	0	0	0	0
4	-9.12	-2.41	0	-0.53
5	-2.70	-0.97	-0.07	-0.27
6	2.37	0.50	0.02	0.12
7	-1.91	-0.70	-0.07	-0.21
8	-6.88	-3.05	0	-0.67
9	-10	-1	0	-0.22
10	0	0	0	0

Figure 5.5 below is a comparison of the first and last iterations of Lloyd's algorithm applied to a ribbon morphotype sample.

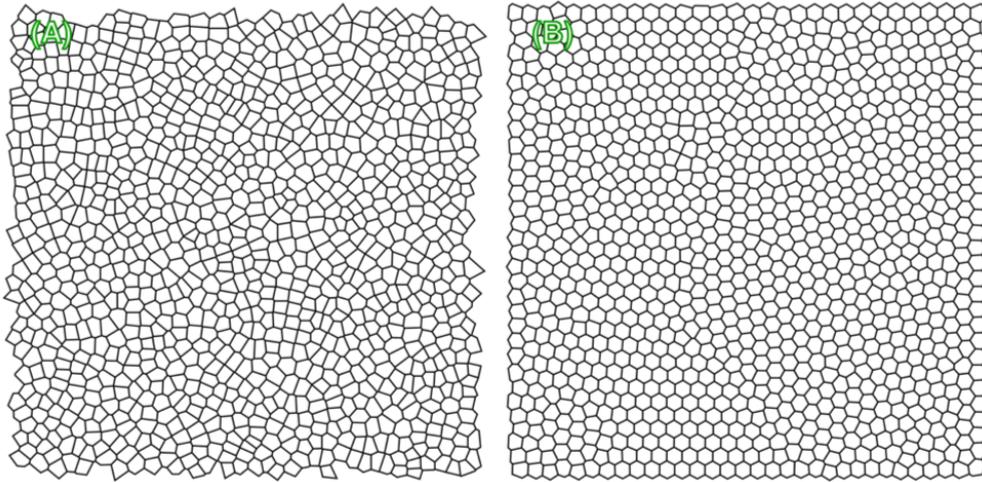


Figure 5.5: Visual comparison of the Voronoi tessellation diagrams (A) before and (B) applying Lloyd's algorithm to a ribbon morphotype sample. Diagrams generated in Wolfram Mathematica. See Appendix A for code.

As aforementioned, the mean cell density is higher in the ribbon morphotype than in the leaf morphotype, and figure 5.3 and 5.5 give a good visual representation of this observation. The mean ribbon morphotype cell density is 73.9% higher than that of the leaf morphotype. Table 5.4 below shows the mean start, end and percentage change in polygon distribution in the ribbon morphotype. The values are remarkably similar to those of the leaf morphotype shown in table 5.2 above.

Table 5.4: The mean start, end and total change in percentage in polygon distribution, between the first and last iteration in the ribbon morphotype.

Number of Edges	Start Percentage	End Percentage	Percentage change
3	0	0	0
4	2.04	0	-100
5	26.54	8.07	-69.66
6	46.16	83.72	81.20
7	22.18	8.20	-63.10
8	2.94	0.01	-99.55
9	0.15	0	-100
10	0	0	0

Figure 5.6 below shows the rate of change of the proportions occupied by each polygon

from the first to the last iteration of a ribbon morphotype sample.

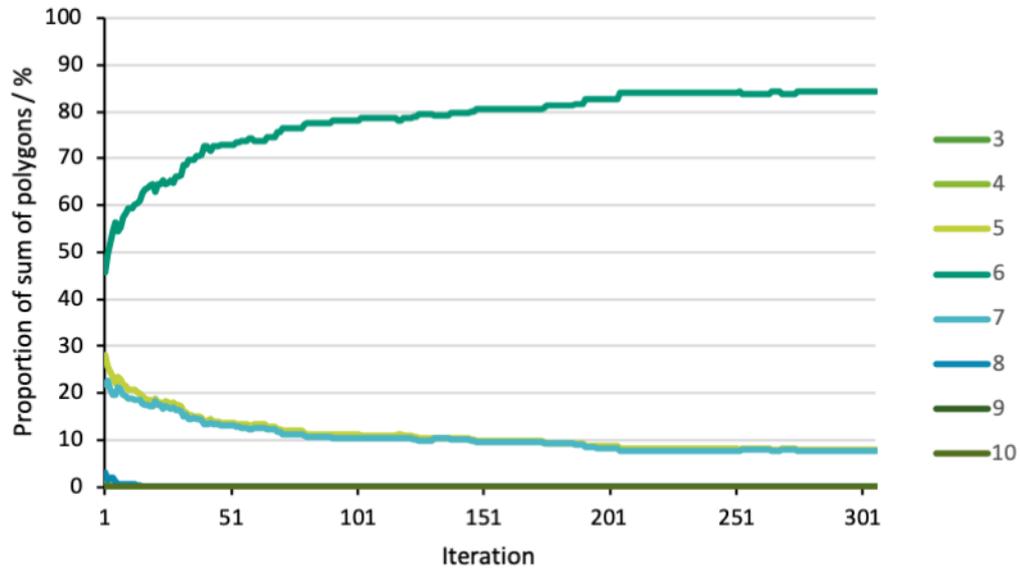


Figure 5.6: Change of proportion of each polygon out of total sum of polygons tracked throughout Lloyd's iterations in a ribbon morphotype sample. The numbers represent the number of edges in the polygon, i.e. 3 = triangles, 4 = quadrilaterals and so forth.

Table 5.5 below shows the mean percentage changes throughout the Voronoi iterations of ribbon morphotype samples.

Table 5.5: Mean percentage change of each polygon between the given ranges of iterations for the ribbon morphotype.

Number of Edges	Mean percentage change			
	Iteration 1 to 10	1 to 100	155 to 455	1 to 455
3	0	0	0	0
4	-12.54	-3.25	0	-1.07
5	-2.86	-0.90	-0.16	-0.40
6	2.71	0.55	0.04	0.20
7	-1.31	-0.70	-0.14	-0.33
8	-10.51	-2.88	-0.5	-1.27
9	-10	-1	0	-0.33
10	0	0	0	0

The analysis of the Voronoi Tessellation diagrams and their corresponding Centroidal Voronoi Tessellation diagrams showed that hexagons occupy a similar proportion of the polygon sum before and after iteratively applying Lloyd's algorithm in both leaf and ribbon morphotypes, as shown in table 5.6 below. The means, standard deviations and coefficients of variation were all calculated excluding the control.

Table 5.6: The mean, standard deviation and coefficient of variation for the start and end percentage, and percentage change of hexagons in the leaf and ribbon morphotypes.

	Start Percentage		End Percentage		Percentage Change	
	Leaf	Ribbon	Leaf	Ribbon	Leaf	Ribbon
Mean	47.05	46.16	83.32	83.72	76.62	81.20
Control	29.94	29.86	85.98	85.36	182.61	184.65
Standard Deviation	2.34	2.37	1.54	1.23	7.98	7.18
Coefficient of Variation	4.97	5.13	1.85	1.47	10.42	8.84

The rate of positional changes starts high and falls rapidly in both the leaf and ribbon morphotypes. Table 5.7 below shows the percentage of total positional changes occurring in different ranges of iterations.

Table 5.7: The percentage of total positional changes occurring in the iteration ranges given, for both morphotypes.

Morphotype	Percentage of total positional changes							
	Iteration							
	1-10	1-100	1-400	100-200	100-300	100-400	200-300	200-400
Leaf	29.4	79.7	99.3	11.1	16.7	19.8	5.8	8.8
Ribbon	30.1	79.4	98.7	12.4	18.2	19.5	5.8	7.1

Examining the uppermost surface of the *Ulva* thalli, it was immediately obvious that there are clusters of cells, henceforth referred to simply as "patches". To illustrate this, figure 5.7 below shows a collage of confocal images of *Ulva* dyed with Direct Yellow 96.

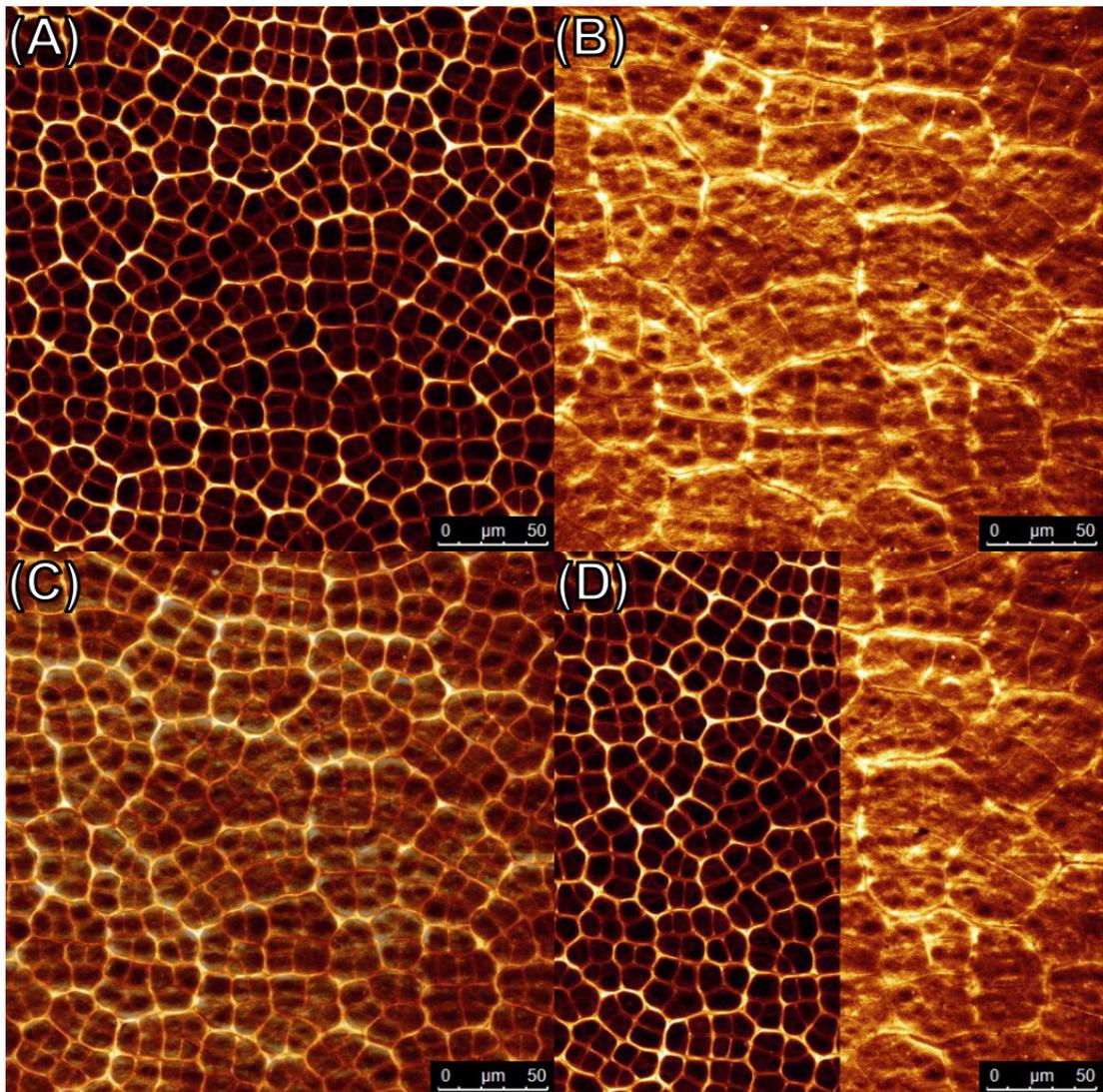


Figure 5.7: Confocal image of *Ulva* dyed with the cell wall stain Direct Yellow 96 showing (A) cells, (B) patches, (C) overlay of patches onto cells at lower opacity for visualisation, (D) overlay of patches onto cells at 100% opacity on one half of the image.

As can be seen in figure 5.7 above, when viewing a section of a confocal image of *Ulva* from a lower z depth (closer to the surface), there are distinct patches that correspond to clusters of cells. As aforementioned, these patches may be artefacts of previous cell divisions. The patches were observed in every single sample of both morphotypes.

There are two prominent, consistent and interesting observations made in every single sample. Firstly, at the surface level, the divisions between the patches of cells show weaker signal, indicating less cell wall matter being available to be stained by the dye. Secondly, the opposite is true at higher z -depths, where it is clear that the borders of patches show the strongest fluorescence, whereas the presumably more recent divisions within the patches

have weaker fluorescence; this may be a result of fewer polysaccharides available in these cells to be dyed by the stain, having had less time to grow. Taken together, this indicates that there is a variation in the amount of cell wall matter present in the borders of the patches, at different levels of the thalli, with less at the surface and more, deeper into the tissue. This suggests that the presence of the patches may have physical and mechanical implications, as they form distinctive shapes. This means that a cross-section would show a gradient of cell wall matter, and this would imply a variation in stiffness at different locations of the thalli. This gives rise to an obvious line of enquiry following this discovery: are these patches indicative of physical shapes or simply tissue compositional differences in *Ulva* thalli, and what are the structural implications, if any?

These patches could be functional, purely an evolutionarily neutral artefact that has not been selected against, a necessary artefact of cell division, or something else entirely. This inspired a query of scales, which the patches seem to resemble. By having zones of varying stiffness, a structure can be made more resistant to external stress. For example, stiff segments with flexible joints embedded on soft tissue greatly increases the stability of each segment as well as the overall structure, all the while the stiff segments providing active protection from puncture and laceration (Martini et al., 2017; Zhu et al., 2012). Combining the discovery of the patches with the known physiology of *Ulva*, namely the “stiff segments”, or patches and “soft substrate”, or the interior cavity between the bilayers, there is every potential for the patches to be mechanically functional in similar fashion to scales observed in metazoans.

The logical progression from the initial observation was to confirm this observation using a different technique. This led to the use of Transmission Electron Microscopy (TEM) in order to investigate whether it is possible to observe the aforementioned patches in a cross section. Transmission Electron Microscopy (TEM) is a type of electron microscopy that is characterised by extremely high spatial resolution, theoretically up to 0.05 Å, or 0.05 nm, which would even allow for the imaging of individual atomic planes, though there are practical limitations, such as the stains used, that restrict the resolution to approximately 10 Å, or 1 nm, when observing cellular structures (Egerton, 2005; Verkade, 2012). The principles of TEM are similar to light microscopy, in that they both utilise an illumination source, which is light in optical light microscopy or an electron source in TEM, then an array of lenses including a condenser lens, objective lens, projective lens

and an image plane (Bozzola and Russell, 1999; Verkade, 2012). TEM images are formed by transmitting electrons, accelerated using a known accelerating voltage, usually 10,000 – 120,000 V, through a thin specimen in a vacuum chamber, and mapping the transmission electron diffraction patterns on to a fluorescent screen or film (Bozzola and Russell, 1999; Egerton, 2005; Verkade, 2012).

If the hypothesis that these patches exist would hold, then one would expect to see patterns consistent with the findings from the confocal microscopy investigation. TEM uses cross sections, so then the results should show that there are sections of dense cell wall in the centre of the cell level that become less dense towards the surface, perhaps curving with the shape of the cell and leaving slight spaces or tissue of different composition. This would be expected to be mirrored in similar fashion on the other side of the thallus, or the second layer of the bilayer.

The TEM images seemed to support the hypothesis, showing clear bands of cell wall around groups of cells, as shown in figure 5.8 below.

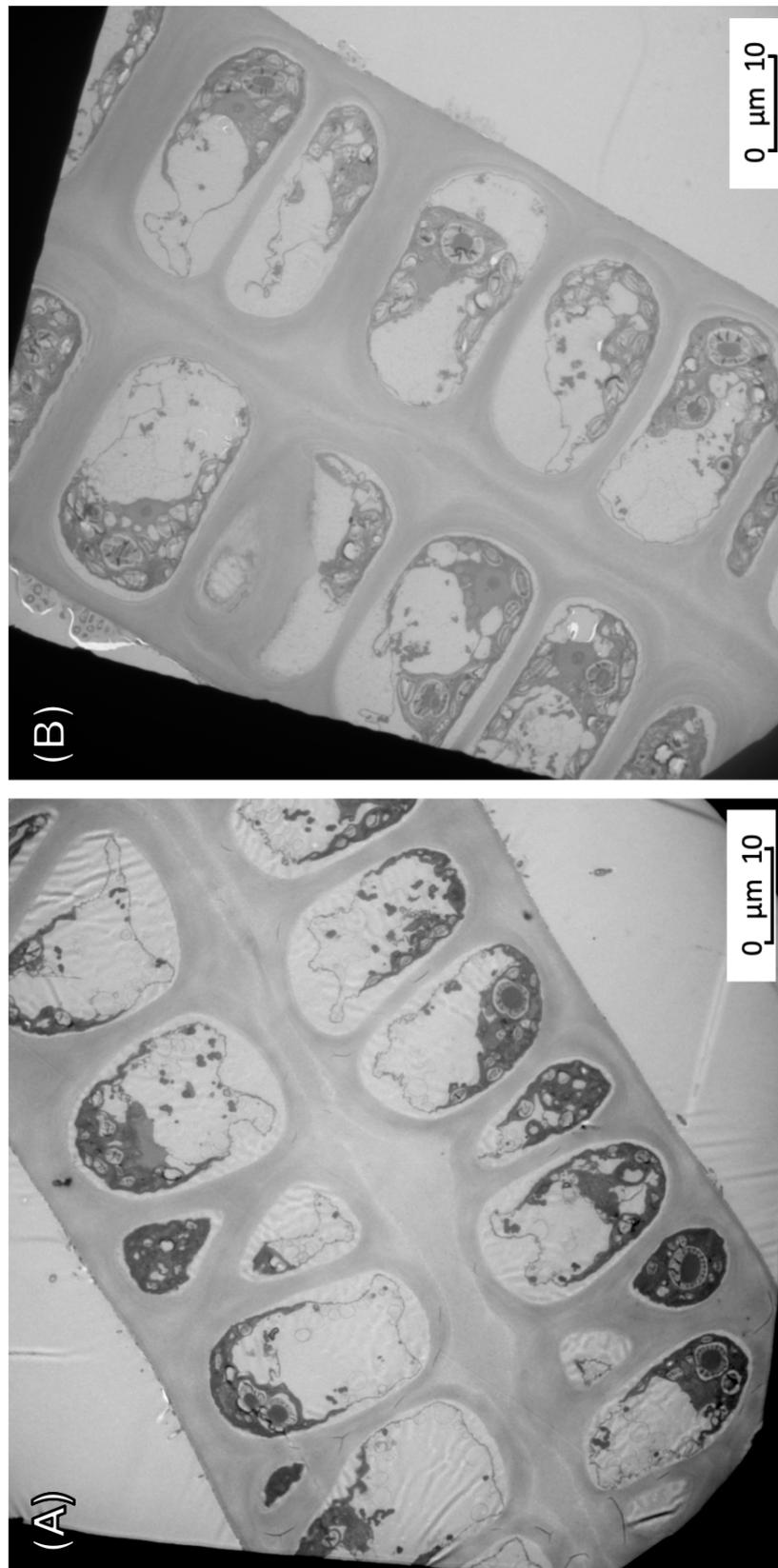


Figure 5.8: Two TEM images of *Ulva* thalli. The darker the shade, the higher the electron density, indicating a higher density of matter present.

Careful examination of the TEM images shows that there are dense bands of cell wall that follow the contours of multiple cells, and so these curves leave dimples of less dense matter near the surface, as they meet deeper into the tissue. The densest areas are the surface, the cell walls around a cluster of cells, whilst the least dense are the aforementioned “dimples” and the space between the bilayers. Interestingly, these patterns are mirrored on the other side of the cells. As the cell walls curve towards the interior of the thalli, this again leaves dimples of less dense matter. This supports the findings of the confocal microscopy, with the densest areas of the cell wall being deeper in the thalli, and the least dense areas being near the surface.

Centroidal Voronoi Tessellation was also used in this chapter as a novel method of biological noise reduction in Voronoi tessellation diagrams. This was done in order to further investigate the impurities hypothesis established in chapter 4. Figure 5.9 below illustrates the use of this technique very clearly.

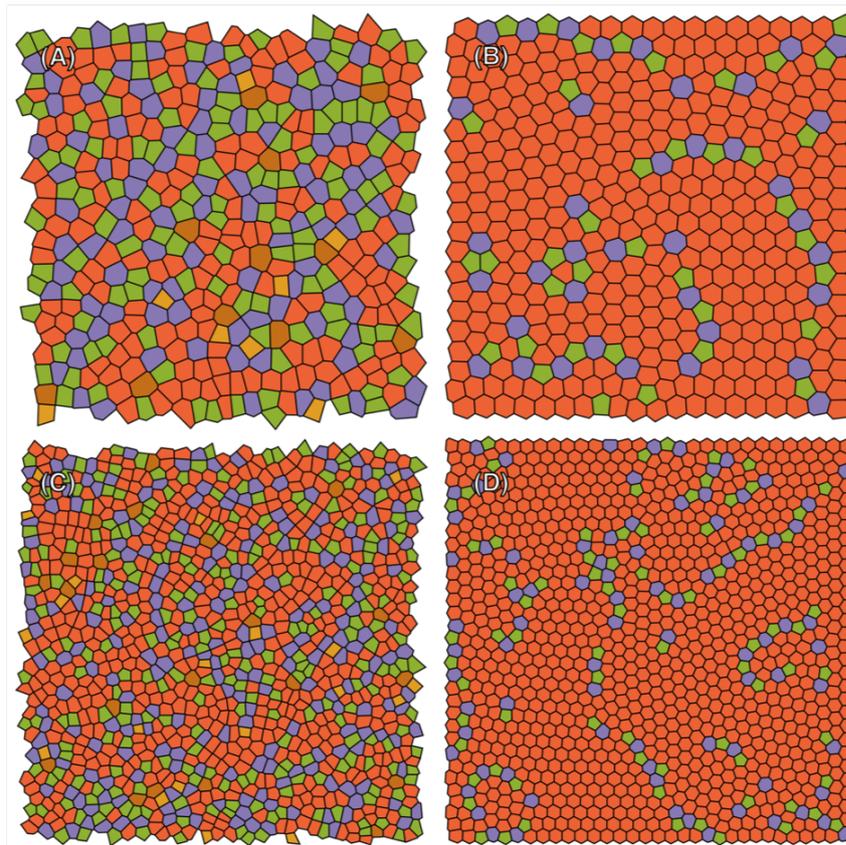


Figure 5.9: Coloured Voronoi tessellation diagrams of a leaf morphotype sample (A) before and (B) after and a ribbon morphotype sample (C) before and (D) after applying Lloyd’s iteration to reach Centroidal Voronoi Tessellation. Different polygons are represented by different colours: quadrilaterals = gold, pentagons = green, hexagons = orange, heptagons = purple and octagons = brown. Diagrams generated in Wolfram Mathematica. See Appendix A for code.

Unlike in (A) and (C), where certain patterns elude easy recognition, it can immediately be seen in (B) and (D) that there are distinct lines of impurities flowing through the sheet of regular hexagons. The profiles for all four Voronoi tessellation diagrams in figure 5.9 above were plotted as shown in figure 5.10 below.

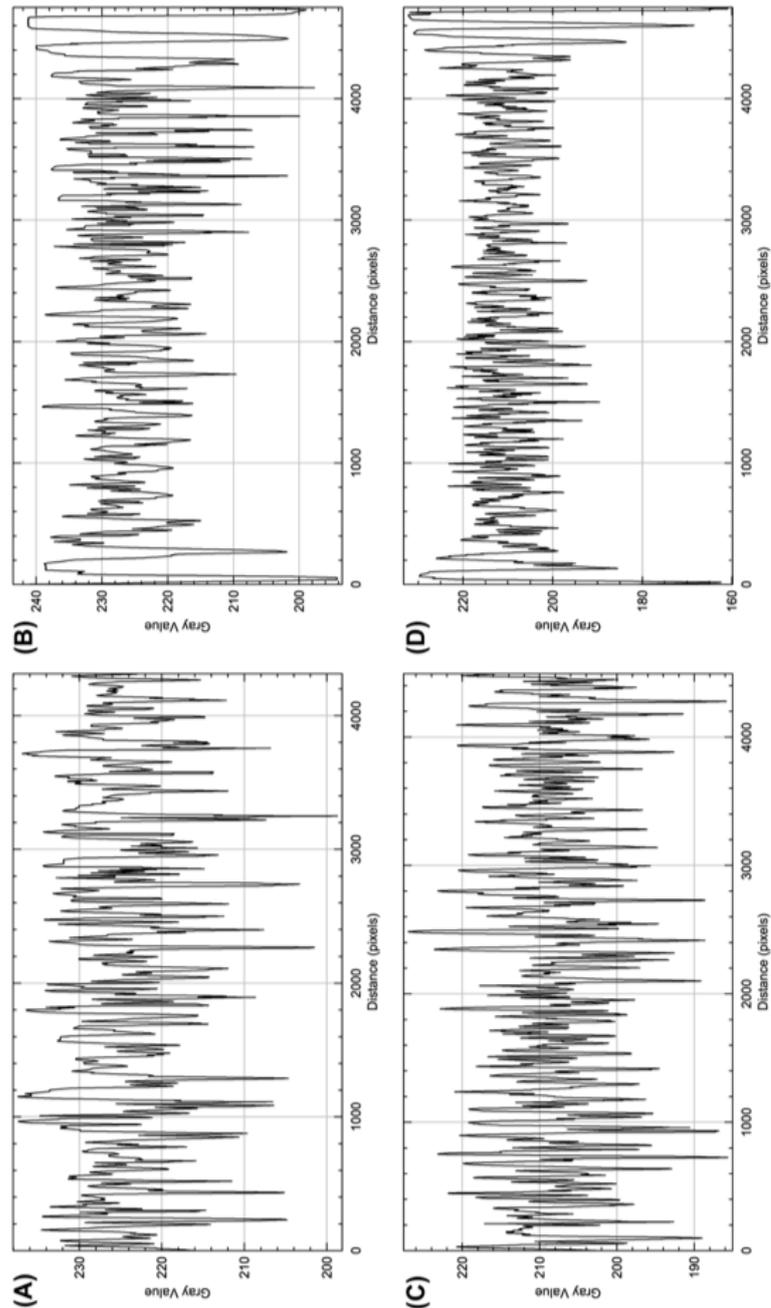


Figure 5.10: Profile plots of the Voronoi tessellation diagrams of a leaf morphotype sample (A) before and (B) after, and a ribbon morphotype sample (C) before and (D) after applying Lloyd's algorithm to reach Centroidal Voronoi Tessellation. Higher peaks indicate higher pixel density.

The reduction of the biological noise from the Voronoi tessellation diagram shows the patterns of regularity much more clearly. In order to isolate the edge effects, the profiles of the Voronoi tessellation diagrams were plotted again without the edges and shown in figure 5.11 below.

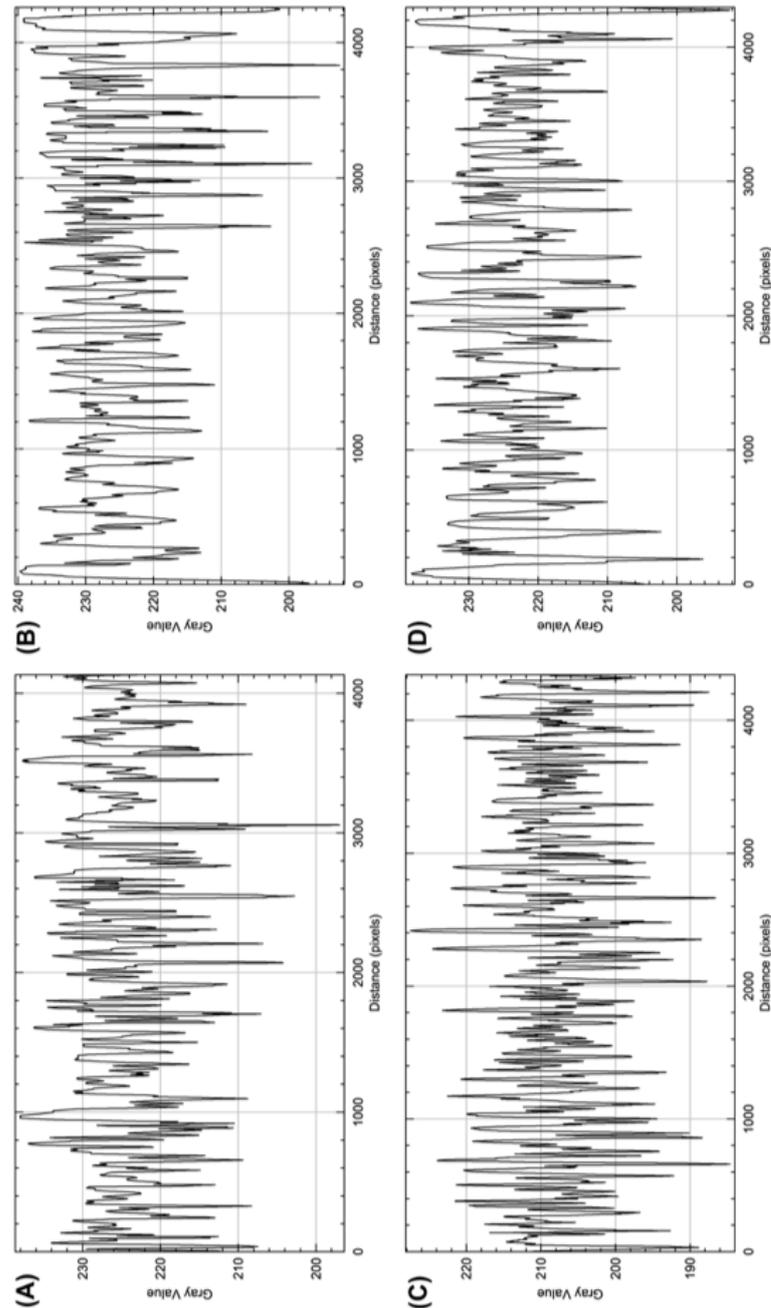


Figure 5.11: Profile plots of the Voronoi tessellation diagrams of a leaf morphotype sample (A) before and (B) after, and a ribbon morphotype sample (C) before and (D) after applying Lloyd's algorithm to reach Centroidal Voronoi Tessellation, without the edge cells data. Higher peaks indicate higher pixel density.

By plotting the profiles without the edge data, this clarifies the findings further. From this, it can be more emphatically concluded that in the leaf morphotype, there is a clear transition from an area of irregularity (small peaks) to an area of regularity (large peaks). For the ribbon morphotype, this means that there is are two transitions from an area of regularity to an area of irregularity, back to an area of regularity. The reduction in biological noise helped to see that the profiles of the leaf and ribbon morphotypes are indeed different, as shown below in figure 5.12 by plotting the ribbon morphotype in four quarters, as previously done in chapter 4.

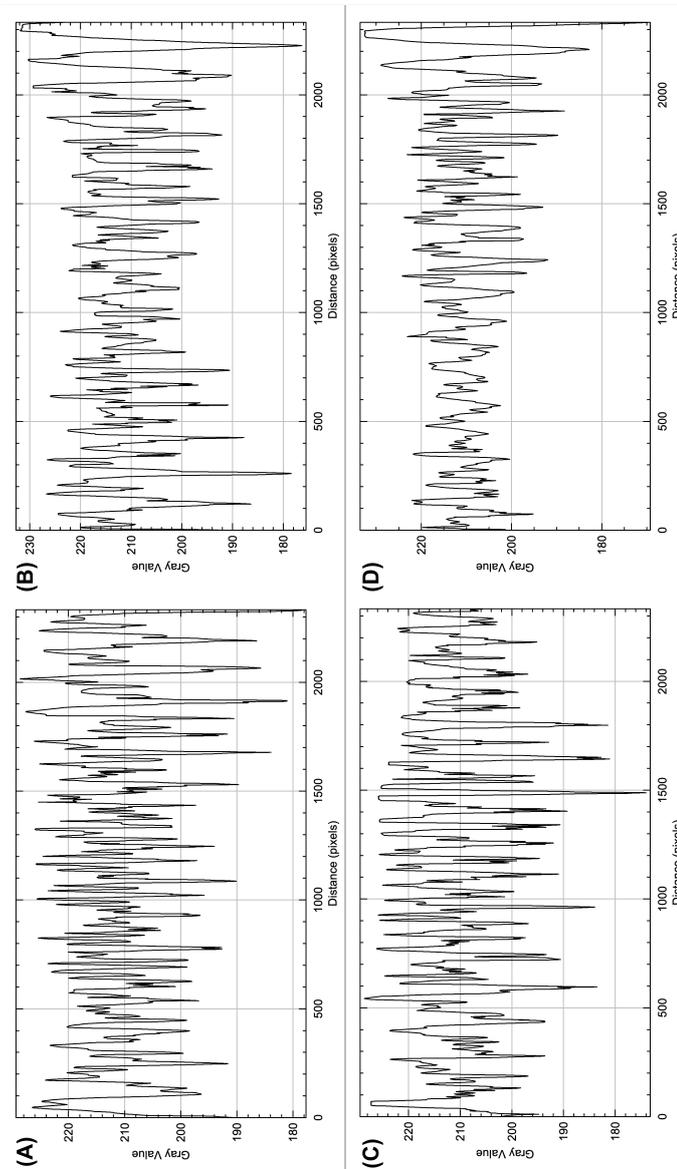


Figure 5.12: Profile plots of the Voronoi tessellation diagrams of the four quarters of a ribbon morphotype sample: (A) top left quarter, (B) top right quarter, (C) bottom left quarter and (D) bottom right quarter. Higher peaks indicate higher pixel density.

These analyses are very insightful; however, they still deal with *Ulva* thalli as a collection of individual units. What relevance are the patches? In order to investigate this, the outlines of the patches were drawn and mapped onto the Voronoi tessellation diagram, as shown in figure 5.13 below.

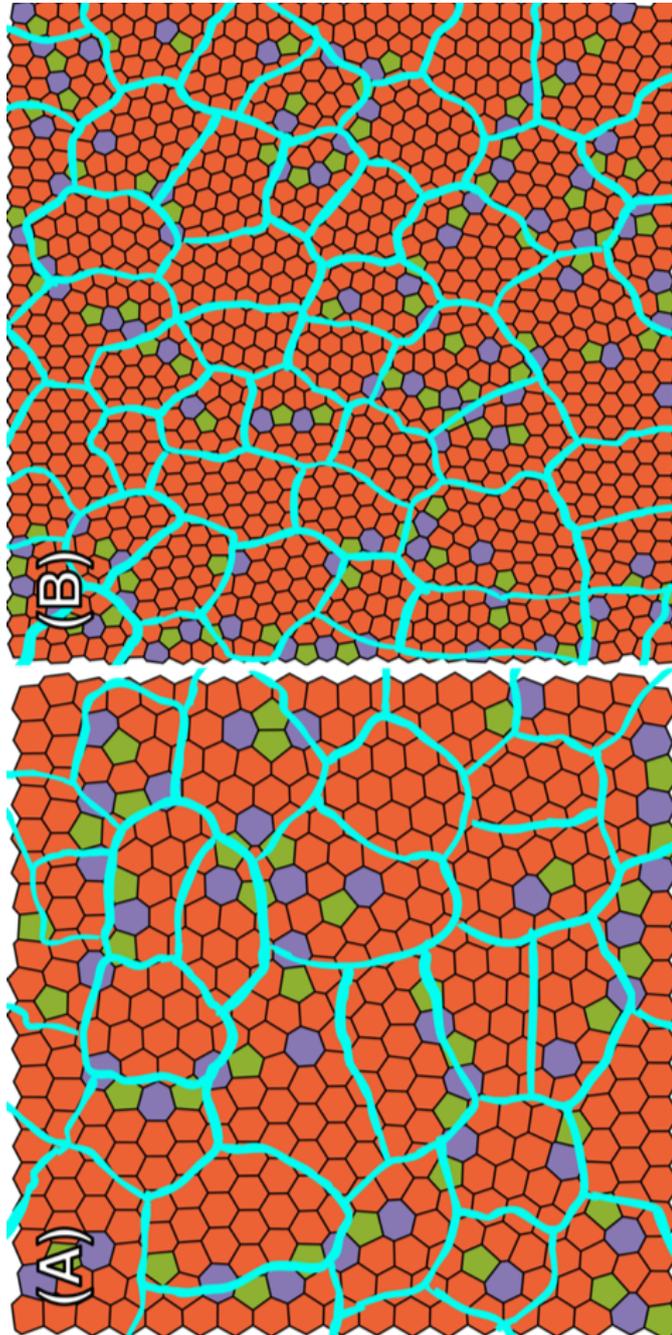


Figure 5.13: Centroidal Voronoi Tessellation diagrams of the (A) leaf morphotype and (B) ribbon morphotype marked with the outlines of their respective patches in cyan. Different polygons are represented by different colours: pentagons = green, hexagons = orange, heptagons = purple. Diagrams generated in Wolfram Mathematica. See Appendix A for code.

It can be seen that the impurities mainly lie on or near the patch outlines in both morphotypes. Using a Voronoi tessellation diagram that has undergone fewer Lloyd's iterations makes the diagram closer to the initial tessellation than the Centroidal Voronoi Tessellation diagram, and easier to interpret than the initial one, as shown in figure 5.14 below.

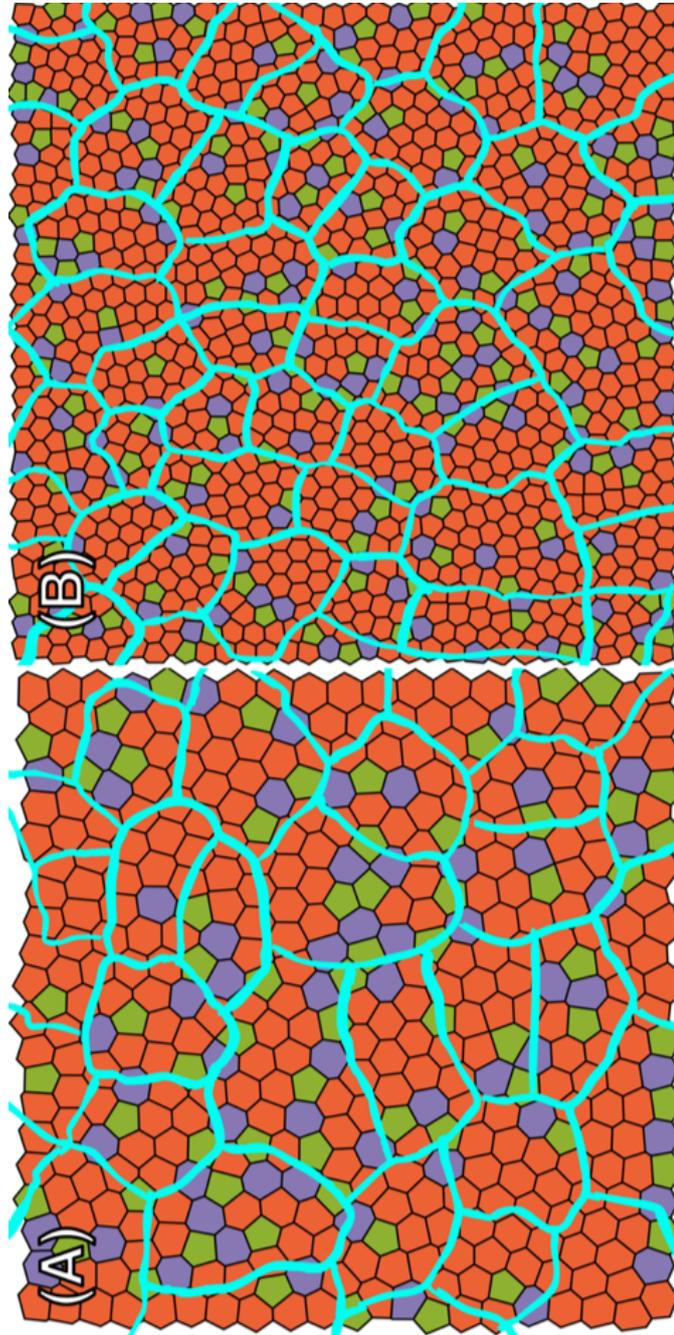


Figure 5.14: Voronoi Tessellation diagrams of the (A) leaf morphotype and (B) ribbon morphotype after 50 Lloyd's iterations, marked with the outlines of their respective patches. Different polygons are represented by different colours: pentagons = green, hexagons = orange, heptagons = purple. Diagrams generated in Wolfram Mathematica. See Appendix A for code.

5.4 Discussion

The core question of this investigation is simple: in the absence of genetic top-down controls, is the morphogenesis in both the leaf and ribbon morphotypes being driven by the same underlying physical mechanism? In order to answer this question, there are many prerequisite questions that must be answered. Taking the results from Chapter 4 and in the context that there seems to be no genomic architecture for multicellularity (De Clerck et al., 2018), one could reasonably consider the possibility that the underlying mechanisms for morphogenesis in *Ulva* are physical.

Chapter 4 discussed the possibility of accurately modelling *Ulva* cell proliferation patterns by adding random variation to the Voronoi generators of a regular hexagonal honeycomb tessellation. The random variation simulated the various physiological and environmental factors affecting development, that would limit the centre of each cell from being equidistant. Analysis of the artificially generated cell proliferation patterns, and naturally generated cell proliferation patterns showed that both patterns were extremely close to an ideal space-filling distribution, with heavy enrichments of hexagons and the entire pattern being dominated by hexagons and polygons once removed from hexagons (pentagons and heptagons), as found by Lewis (1928) in the epidermis of *Cucumis*. The hypothesis that *Ulva* thalli are sheets of regularity permeated by areas of irregularity, or hexagons with impurities of other polygons, stood and justified the original models. However, a key element missing from the random jitter models was the biological artefact from cell division that is clearly visible in all *Ulva* samples, as previously discussed in chapter 4 and shown in figure 4.12 above. This suggests that while the principle of the hypothesis is still sound, the generating mechanism of such pattern is fundamentally different in real samples versus mathematical models, as would be expected. The further analysis of this impurities hypothesis was limited by biological noise, namely the irregularity being so high, as they confounded the interpretation of the profiles.

The Centroidal Voronoi Tessellation investigated in this chapter served two purposes. First, it allowed an investigation of how close naturally generated cell proliferation patterns are to ideal, space filling distributions, as has been done previously for a number of metazoan tissue types. Second, it allowed for the reduction of biological noise when examining the impurities hypothesis, which is a novel use, meaning that the analyses had to be *ad hoc*.

Mathematically, the ideal space filling distribution is a simple sheet of regular hexagons, commonly referred to as a honeycomb pattern. The background to this, including the honeycomb conjecture solved by the mathematician Thomas C. Hales at the University of Michigan in 1999, was covered in detail in Chapter 4 , section 4.1, This also means that Centroidal Voronoi Tessellation tends towards hexagons, as the algorithm iteratively adjusts the Voronoi generators towards the most ideal space-filling and mechanically stable arrangement, as aforementioned in section 4.1.

For many years, it was believed that proliferating epithelia in metazoans have a conserved polygon distribution that arise from cell proliferating mechanisms and cellular growth rates, and it was even thought to be proven (Aegerter-Wilmsen et al., 2010; Gibson et al., 2006; Nagpal et al., 2008). Gibson et al. (2006) investigated this conserved polygon distribution using a Markov chain with empirical data gathered from time-lapse microscopy. They made a quantitative prediction that the conserved polygon distribution of approximately 28.9 % pentagons, 46.4 % hexagons and 20.8 % heptagons is the inevitable equilibrium state of proliferating epithelia, irrespective of the initial conditions. The empirical data from their experiment also seemed to support this, with the cell topology of proliferating epithelium in *Drosophila* wing discs, *Xenopus* tadpole tails and *Hydra* epidermis all showing a polygon distribution within a few percent of the predicted equilibrium.

Though these types of advanced analyses have yet to be done for plant tissues, the pioneering works of the American morphologist F.T. Lewis in the 1920s on the epidermis of *Cucumis* (Lewis, 1923, 1926, 1928) have been an invaluable resource for the investigation of cellular polygonal tessellation patterns since, with almost all leading literature in the field referring to Lewis' work, including the aforementioned first-order Markov model study by Gibson et al. (Gibson et al., 2006; Gibson and Gibson, 2009; Nagpal et al., 2008; Sánchez-Gutiérrez et al., 2016; Sánchez-Gutiérrez et al., 2017). Stating that Lewis' data on *Cucumis* epidermis showed remarkably similar polygon distribution patterns, Gibson et al. (2006) justified inferring that the equilibrium state of the conserved polygon distribution is not restricted to metazoans, but in most multicellular eukaryotes, and arises from cell division mechanisms alone.

Much of the work of Gibson et al. (2006) was accurate for the data they had, however their perspective that the steady state equilibrium is an inevitable mathematical consequence of cell division mechanisms may well be flawed due to the limitations of their study, which

only examined proliferating epithelial tissue. This conclusion was challenged recently by Sánchez-Gutiérrez et al. (2016), who investigated the cell topology of a non-proliferating tissue type, human bicep muscle fibres, and a different proliferating epithelia, the chicken neuroepithelium. This investigation showed that the same ‘conserved’ polygon distribution of proliferating epithelia was also present in a non-proliferating tissue type, and that proliferating epithelia don’t necessarily show this conserved polygon distribution. Sánchez-Gutiérrez et al. concluded that given a constant balance of forces, fundamental physical constraints will dictate the equilibrium state polygon distribution in epithelial tissues, without the need for any cell division mechanisms. Any changes in the balance of forces, for example due to disease, neurogenic atrophy in the bicep in this case, will alter the polygon distribution.

Investigating the leading research in the conserved polygon distribution of epithelial cellular topology, it quickly becomes evident that there is a striking gap in the literature concerning plants. Almost all research in the field refers to Lewis (1923, 1926, 1928) work, however there have been no work on multicellular plants beyond observational data, and none with any advanced modeling. *Ulva* is a fantastic organism to start to fill this gap in knowledge, due to its simplicity. With no genomic architecture to drive cellular proliferation patterns and predominantly consisting of one cell type, the entirety of *Ulva* is similar to most epithelial tissues, such as those investigated by the leading researchers in epithelial cellular topology (Aegerter-Wilmsen et al., 2010; Gibson et al., 2006; Gibson and Gibson, 2009; Nagpal et al., 2008). As shown previously in figure 4.10 in section 4.3.4, *Ulva* shows a remarkable adherence to the equilibrium state proposed by Gibson et al. (2006), and to the *Cucumis* epidermis polygon distribution patterns identified by Lewis (1928), showing only a minimum of 0.1 % and a maximum of 3.4 % difference in all polygon distributions. The polygon distribution, the simplicity of its tissue and genome make *Ulva* an ideal model system to study epithelial cell topology.

One of the primary observations of Chapter 4, aforementioned section 4.3.4, was that the leaf and ribbon morphotypes show no statistically significant difference in cell topology after accounting for cell size, with the exception of polygon regularity as discussed in section 4.4. This, coupled with the well-known tendencies of *Ulva* to show varying morphologies in seemingly due to different environmental conditions, led to the hypothesis that morphogenesis in *Ulva* arises from the same underlying physical mechanism, and is governed by

bottom-up proliferation mechanisms, rather than top-down mechanisms. The findings in this chapter further support this hypothesis in three main discoveries:

Firstly, the mean number of iterations required for the leaf and ribbon morphotypes to reach Centroidal Voronoi Tessellation were 419 and 402, respectively.

Secondly, as can be seen in table 5.7 above, the mean percentage change between iterations is highest in the first ten iterations, for both leaf and ribbon morphotypes. The leaf and ribbon morphotypes showed that 29.4 and 30.1 %, respectively, of the total positional changes occur in the first 10 iterations and 79.6 and 79.4 %, respectively, of the total positional changes occur in the first 100 iterations.

Finally, once Centroidal Voronoi Tessellation is reached, the Voronoi diagrams for all samples look remarkably like a regular hexagonal honeycomb tessellation. Having started at a strikingly similar polygon distribution to each other, and to other known epithelial cell polygon distribution patterns, *Ulva* maintains this similarity in polygon distribution between both morphotypes, with the polygon distribution being dominated by hexagons at 83.3 % and 83.7 % for leaf and ribbon morphotypes respectively. After Centroidal Voronoi Tessellation is reached, *Ulva* polygon distributions in both the leaf and ribbon morphotypes are within one percent of each other in every polygon class.

Sánchez-Gutiérrez et al. (2016) constructed a non-proliferating tissue simulation using randomly generated Voronoi generators. They found that after applying Lloyd's algorithm four times, the random distribution iteratively relaxed to give a polygon distribution that is representative as a model for healthy human adult bicep tissue. This was also remarkably similar to the steady state equilibrium that Gibson et al. (2006) proposed for proliferating epithelia. It is intriguing then, that every random control that was investigated in this experiment also reached a polygon distribution within a few percent of Gibson's steady state equilibrium, also after four iterations. This point represented a mean of 26.1 and 27.0 % of all positional changes in the leaf and ribbon morphotypes, respectively.

The mechanisms driving patterns in proliferation tissue is a new, and as such, keenly debated topic. The results of experiments from various leaders in the field such as the aforementioned studies by Aegerter-Wilmsen et al. (2010); Gibson et al. (2006, 2011); Gibson and Gibson (2009); Nagpal et al. (2008); Sánchez-Gutiérrez et al. (2016), intrinsically and inherently have significant merit as purely fundamental and mechanistic information.

However, one must question why such patterns are so universally found, especially in light of the findings such as those of the previously discussed investigation by Sánchez-Gutiérrez et al. (2016), wherein it was found that non-proliferating tissue can show the same patterns as proliferating tissues, and that proliferating tissues can show different patterns. The fundamental question that would link all of these studies, is simple: What advantage does optimal space-filling geometric packing confer to the organism? Furthermore, the simple question can be then elaborated into three more defined questions. What advantage, indeed, does optimal space-filling geometric packing confer, what would limit the tissues from being perfectly optimal (honeycomb), and what is the point at which the benefits of the optimal space-filling configuration is outweighed by the detriments?

At first sight, one could perceive the answer to the first question to be simple. Optimal space-filling arrangements reduces the total amount of biological investment in both energy and matter, as dictated by the Honeycomb conjecture that was previously discussed. Also as previously discussed, the most perimeter method of enclosing infinitely many unit areas with regular polygons of equal area, would be using triangles. Using triangular tiling and hexagonal tiling as controls, one could deem it possible to compare the Voronoi tessellation patterns of real samples with respect to how efficient or inefficient they are in comparison. The most immediately obvious comparison would be the total perimeters of each configuration, which would translate into the amount of matter used for the cell walls in this case. The three questions posed above are simply to organise the inquiry. All three questions must be considered together, as the answers are all interlinked.

The tissue of an organism does not have only the issue of conserving biological energy and matter to consider. There are other critical and paramount attributes to consider. The first to consider is the stability of the tessellation and therefore the stability of the tissue and in turn the organism. In a two-dimensional Euclidean space, hexagons are the most stable. Another consideration is the actual achievement of this configuration; cells do not just appear in a space-filling tessellation, so a mechanism must drive its path to the desirable configuration.

There exists another level of complication beyond this all, wherein the specialist functionality of the tissue warrants a deviation from a space-filling configuration. The deviation of proliferating tissue of an organism from a space-filling configuration has only been demonstrated in cases of specialist tissues, wherein the functionality of the tissue is affected by the

tissue structure and therefore requires a more complex structure than a relatively homogeneous sheet of tessellating polygons, for example in chicken neuroepithelia Sánchez-Gutiérrez et al. (2016).

All of these considerations leave one certain uncertainty: the preserved polygon distribution observed in both metazoans and plants arises from an as yet unknown, but almost definitely physical, mechanism. Gibson et al. (2006) put forward a compelling argument that this distribution is a mathematical consequence using a first-order Markov model. This was well supported with empirical data, however the empirical data used was not an exhaustive list of tissue types, as the experiment only dealt with proliferating tissue of a similar nature. Sánchez-Gutiérrez et al. (2016) cast doubt on the absolute nature of the conclusions drawn by Gibson et al. (2006), by investigating a non-proliferating tissue and artificially generated models with the same conserved polygon distribution, and a proliferating tissue that does not show the conserved polygon distribution. The results of this investigation add support to the findings of Sánchez-Gutiérrez et al. (2016), with my own artificially generated models also showing a conserved polygon distribution without the use of proliferation mechanisms. These two studies, alongside my results and the discovery by De Clerck et al. (2018) that there is no genomic signature for multicellularity in *Ulva*, make an extremely strong case for the underlying mechanism for cellular proliferation and morphogenesis in *Ulva* to be physical.

Regardless of the study, there are two key themes. The first, is the physical nature of the observed pattern, and the second is the enrichment of hexagons. The dominance of hexagons in the observed patterns is a clearly understood matter, as dictated by mathematical and physical laws to minimise perimeter to area ratios in two dimensions, and surface area to volume ratios in three dimensions. This confers clear biological advantages of conserving biological energy and matter as well as providing structurally stable tissues. The exact nature of this proliferation mechanism that produces the conserved polygon distribution, however, is a question that will continue to be debated in the years to come.

The Centroidal Voronoi Tessellation, however, was also used in this chapter as a novel method of biological noise reduction in Voronoi tessellation diagrams. This was done in order to further investigate the impurities hypothesis established in chapter 4. Figure 5.9 above illustrates the use of this technique very clearly.

Unlike in figure 5.9 (A) and (C), where certain patterns elude easy recognition, it can immediately be seen in figure 5.9 (B) and (D) that there are distinct lines of impurities flowing through the sheet of regular hexagons. The profiles for all four Voronoi tessellation diagrams in figure 5.9 above were plotted as shown in figure 5.10 above.

The reduction of the biological noise from the Voronoi tessellation diagram shows the patterns of regularity much more clearly. The ribbon morphotype showed clusters of large peaks with interspersed clusters of small peaks. The leaf morphotype shows a pattern that was not previously recognised: a transition from a cluster of regular smaller peaks to a cluster of regular larger peaks. In order to isolate the edge effects, the profiles of the Voronoi tessellation diagrams were plotted again without the edges and shown in figure 5.11 above.

By plotting the profiles without the edge data, this clarifies the findings further. From this, it can be more emphatically concluded that in the leaf morphotype, there is a clear transition from an area of irregularity (small peaks) to an area of regularity (large peaks). For the ribbon morphotype, this means that there is are two transitions from an area of regularity to an area of irregularity, back to an area of regularity.

I had mentioned in the discussion of chapter 4 that the leaf morphotype shows a more regular profile overall, with mostly similar sized and spaced peaks throughout the profile, while the ribbon morphotype shows a clearer pattern of three clusters of large peaks with a cluster of small peaks in between each. I also mentioned that the initial results suggested that there are differences in regularity between the two morphotypes. The results of this chapter allowed for a clearer analysis of the data with less biological noise. The profiles plotted in this chapter, as seen in figure 5.10 and 5.11, confirm that the differences in patterns that were observed in chapter 4 were not due to biological noise.

Just as I had done in chapter 4, I checked the results by separately plotting the profile of the four quarters of the ribbon morphotype diagram, as shown in figure 5.12 above. This showed that the profile of an area in ribbon morphotype thalli that encompasses a roughly equivalent number of cells as in leaf morphotype thalli is, again, different to that of a leaf morphotype, leading to the conclusion that the initial differences observed were not due to the difference in cell density. This means that the hypothesis formed in chapter 4, that *Ulva* morphogenesis in the leaf and ribbon morphotypes is driven by different types

of irregularities in the tissue, can be accepted and should be explored further.

These analyses are very insightful; however, they still deal with *Ulva* thalli as a collection of individual units. What relevance are the patches? In order to investigate this, the outlines of the patches were drawn and mapped onto the Voronoi tessellation diagram, as shown in figure 5.13 above.

It can be seen that the impurities mainly lie on or near the patch outlines in both morphotypes. However, it should be noted that the cell placements of the Centroidal Voronoi Tessellation diagrams are slightly different to the initial Voronoi tessellation diagram. Using the initial tessellation diagram is difficult as there is too much biological noise to meaningfully interpret the diagram. However, using a Voronoi tessellation diagram that has undergone fewer Lloyd's iterations makes the diagram both closer to the initial tessellation than the Centroidal Voronoi Tessellation diagram, and easier to interpret than the initial one, as shown in figure 5.14 above.

It can still be seen that the patch outlines generally coincide with the impurities with mostly regular hexagons in the core of the patches.

I then proceeded to pursue more mechanical information that may elucidate the functionality of the patches. There are a number of measurements available to quantify various mechanical properties of materials such as “ E ”, Young's modulus, sometimes referred to as the elastic modulus, “ G ”, shear modulus and “ B ”, bulk modulus, amongst others (Hopcroft et al., 2010). The most obvious line of inquiry in this case is the tensile strength of *Ulva* thalli under longitudinal stress, quantified by Young's modulus, measured in Pascals (Hopcroft et al., 2010; Sokolov et al., 2013). This is due to the hypothesis that variations in tensile strength across the thalli would be seen with the variation of cell wall density across the thalli in the pattern of the patches. In nanoscale biological systems, obtaining such measurements requires the use of Atomic Force Microscopy (AFM).

AFM was built on the principles developed by Scanning Tunnelling Microscopy (STM). STM is a revolutionary ultra-high resolution microscopy technique developed by Binnig et al. (1983), which won them the Nobel Prize in Physics in 1986 (Giessibl, 2003). STM uses a sharp tip, usually made of tungsten, a platinum-iridium alloy or pure iridium, that is mounted onto a three-dimensional (xyz) positioning scanner and lowered towards the surface of the sample until the distance between the tip and the sample is just a few

angstroms, allowing a tunnelling current to flow between them (Giessibl, 2003). This principle, however, limits STM to conductive materials and even then, with the exception of very few materials such as highly oriented pyrolytic graphite, they must be studied in ultra-high vacuum (Giessibl, 2003; Rugar and Hansma, 1990). In order to overcome this limitation, a number of other scanning probe microscopes were developed, with AFM being the most successful of them and becoming a laboratory staple worldwide (Giessibl, 2003; Rugar and Hansma, 1990).

Modern AFM systems utilise a microfabricated silicon, silicon oxide or silicon nitride cantilever, that is bent towards the sample by a piezoelectric motor, which is displaced due to short-range interatomic forces when in contact with the surface of the sample and this the degree of displacement is measured using laser beams being deflected off the cantilever (Giessibl, 2003; Jena et al., 2002; Rugar and Hansma, 1990). This mode of scanning is called contact or static mode, and there exists a second mode of scanning called noncontact or dynamic mode wherein the cantilever is moved 10 – 100 nm away from the sample and set to vibrate at a given frequency and amplitude, which is then able to be disturbed by long-range forces, such as electrostatic, magnetic and van der Waals forces, and the change in frequency and amplitude can be measured (Giessibl, 2003; Jena et al., 2002; Rugar and Hansma, 1990).

AFM can also be used in a number of different ways apart from simple scanning, such as for atomic manipulation (Giessibl, 2003). In biology, a particularly useful method of utilising AFM is to accurately measure the mechanical properties of cells, for example the Young's modulus (Gavara, 2017; Thomas et al., 2013). This is done by using the cantilever as a nano-indenter and measuring the deflection of the cantilever, and consequently calculating the force acting on the cantilever to obtain a force-deflection curve, which can be fitted to a model, typically the Hertz model, to calculate the Young's modulus of the sample (Gavara, 2017; Roduit et al., 2009; Thomas et al., 2013). Though there is fair criticism regarding the various assumptions of the Hertz model, including linear elasticity, the Hertz model represents the most widely solved for a variety of parameters, including for tip shapes, sample thickness and viscoelastic effects amongst others (Gavara, 2017; Sokolov et al., 2013). Young's modulus was the exact measurement we sought to investigate in *Ulva*. The results from the AFM investigation on *Ulva* showed that there are distinct patches of stiff and soft patches, approximately the size of two to three cells. However, this

is simply an exploratory result and should be taken with great caution as it was only one replicate of one sample and any interpretation of results should be delayed until a more formal investigation can be performed with more samples and replicates.

Taken together, the results of this chapter suggest that the *Ulva* thallus is comprised of areas of regularity, permeated by lines of irregularity that correspond to a physical difference in composition and structure. The lines of irregularity potentially correspond to differences in stiffness, which could indicate a mechanical functionality of the patches. However, the nature of the equipment used in this study meant that structural roles of the patches could not be directly measured and only inferred. One method to investigate the functionality of these patches would be to use a combined Confocal and Atomic Force Microscope. This would allow the observation of the patches under the confocal microscope, and the atomic force microscope would give measurements of stiffness within the imaged area, which can then be mapped onto the confocal image to see if there is indeed a pattern of variable stiffness that correlates with the patches.

Furthermore, in addition to more detailed measurements of Young's modulus for *Ulva* thalli, obtaining measurements for other elastic moduli, in particular the shear modulus, would be extremely helpful in further testing the hypothesis that patches on *Ulva* thalli serve a mechanical purpose.

It is important to note that though the iterative states may appear to show various stages of a dynamic system, Centroidal Voronoi Tessellation is a method of studying and predicting steady states and is not a dynamic model. In a dynamic cellular system, cell division will drive the patterns and so would be better represented by a generative model, like the aforementioned VirtualLeaf, VPTissue, Cellzilla and CellModeller. However, such models are frameworks developed by well-funded groups of expert coders over multiple years, and are not, by design, intended to model morphogenesis in organisms such as *Ulva*. In the absence of the time, funding and expertise, the next stage in this investigation, then, would be to adapt a generative model to *Ulva* and investigate the morphogenetic patterns, with special attention being paid to model the development of the patches.

This chapter intended to test two key hypotheses. The first is that *Ulva* cellular proliferation patterns are space-filling arrangements generated from bottom-up physical proliferation and therefore will follow fundamental physical and mathematical laws to be as close

to an optimally space-filling arrangement as possible. The second is that *Ulva* cellular proliferation patterns can be described as areas of regularity permeated by lines of irregularity where cell division occurred. Both hypotheses were well-supported by the evidence given here, and can be accepted. Having focused heavily on modelling, the next line of inquiry is to collect data on real tissue growth to characterise the nature of growth, and combining this with the hypotheses arising from modelling to inform a generative model. The next chapter will focus on this, aiming to develop methods of improving spatial and temporal resolution of growth data that will allow for such characterisation.

Investigating the spatial and temporal patterns of growth in *Ulva*

6.1 Introduction

Chapter 4 and 5 were focused on a steady final state of *Ulva*, and modelling backwards to derive said steady state, in an effort to identify the mechanisms underpinning the patterns of cell proliferation. This chapter is a direct investigation of the cell proliferation patterns, in particular addressing the hypothesis that spatial differences in growth give rise to *Ulva* morphologies, using two approaches. However, the difficulties of direct investigations of *Ulva* is the absence of established techniques. Currently, the temporal and spatial resolution of growth monitoring in *Ulva* is very low. The temporal resolution is low due to the primitive and time-consuming nature of aquatic biomass measurement, and the spatial resolution is low due to the same reasons as well as the lack of molecular toolkit protocols. This chapter aims to investigate techniques that would increase spatial resolution of biomass growth monitoring in *Ulva* and help answer the main hypothesis developed from the course of the project: *Ulva* morphology arises due to differential spatial growth patterns. In order to investigate this hypothesis, new techniques will have to be developed. The first approach using an automated optical growth monitoring system, which would increase both spatial and temporal resolution. The second approach focused on increasing the spatial resolution of *Ulva* growth monitoring using EdU cell proliferation assays.

6.2 Optical Imaging System

6.2.1 Introduction

One of the biggest and most persistent challenges that is presented in any experiment with aquatic biomass is the absence of a standardised measurement protocol, which removes the possibility of direct comparisons between multiple experiments. The water content of all plants is in constant flux, so the measurements taken from land plants can be comparable given the controlled parameters, at the least in growth proportions. However, multicellular aquatic plants present a unique problem in the residual mass of water following the removal of the tissue from its growth medium.

The current and only method of mass measurement during an ongoing experiment with multicellular aquatic plants is the measurement of the wet mass. Wet mass measurement involves the physical removal of the samples, varying forms of an intermediate drying step, and then measurement on weighing scales. A simple preliminary growth trial was performed for *Ulva* by growing one hundred samples in culture and taking wet mass measurements of ten random samples (9 repeat measurements) and dry mass measurements of ten sample every day of the trial. The results are summarised below in table 2.2 above.

This showed that under the culture conditions, the mean mass of water in wet mass measurements over the course of five days of growth in the incubation chamber is 82.2%, and this stays relatively consistent throughout the growth period. The water content regularly changes throughout the day as well as under different stresses, leading to a number of inconsistencies in the measurements obtained, so the measurements were taken at the same time of day each day. This method can yield measurements that cannot be directly comparable to those obtained in other experiments as the exact conditions that were present in one experiment cannot be assumed for another, resulting in too many variables that are uncontrollable and unaccounted for. The only acceptable comparison between experiments would be generalised relationships, obtained from the proportionate mass changes, for particular species that hold largely true between experiments. However, without this contextual information in place, any direct comparison of wet mass, even within an experiment, is unreliable and questionable. This means that experiments involving aquatic plants will require species-specific, time-consuming (potentially weeks or months of pre-

liminary growth experiments) groundwork before the commencement of any experiment or set of experiments in order to control and account for variables and establish standardised parameters and protocols. While many land plants, such as *Arabidopsis* and *Nicotiana*, are already established as model organisms, and therefore their generalised relationships are already supported by a wealth of literature, given that the only truly well-established aquatic plant model organism is *Ostreococcus*, a microalga, it is nearly impossible to extrapolate such baselines for other types of aquatic plants.

Microalgae are currently of particular commercial interest as foods, feeds, nutritional supplements, and biofuels such as biodiesel, bio-syngas, bio-oil and biohydrogen production amongst others (FAO, 2018; Li et al., 2008; Mata et al., 2010; Spolaore et al., 2006). With great interest in its potential, there is much investment into microalgal research and research techniques, leading to the development of many reliable biomass measurement techniques. There are many commonly used measurements for microalgal biomass measurements, some direct and others by proxy or inference. Some common methods include cell density measurements, which can be done by using a haemocytometer or microscope, optical density measurements of the culture using a spectrophotometer, *in vivo* chlorophyll fluorescence measurements, and dry weight measurement, often for calibration (Butterwick et al., 1982; Eriksen, 2008; Moheimani et al., 2013). There are also a multitude of other measurement techniques, and regardless of the type of measurement technique, there is a wealth of resources that allows one to choose a suitable method and have the luxury of referring to the literature.

The problem of inconsistent wet mass measurements can be wholly avoided by completely dehydrating the tissue and obtaining dry mass measurements. However, this is still a labour and time-intensive process and therefore is limited by the act of physical measurement.

Furthermore, the spatial and temporal resolution of all of the aforementioned measurements of macroalgal biomass is extremely low. The spatial resolution of mass measurements is low, as weighing scales cannot differentiate where the growth is occurring. The temporal resolution of mass measurement is limited by the number of measurements taken by the researchers themselves, meaning that it is likely that measurements are only taken daily. Assuming that there is no complex robotic configuration being utilised, which in itself would be a tremendous feat, the spatial and temporal resolution of mass measurements in multicellular aquatic plants is inevitably very poor.

Unlike in aquatic plants, obtaining standardised mass measurements in land plant is relatively easy. Furthermore, without the variable of the water interfering with the measurements, both by physical presence and movement of the biomass within it, computer vision-based growth monitoring systems have been attempted multiple times for various parts of various species, including corn (Tarbell and F. Reid, 1991), whole *Arabidopsis thaliana* plants (Bell and Dee, 2016), *Arabidopsis thaliana* roots (Monshausen et al., 2011; Moore et al., 2013), and there has even been a spectacularly complex autonomous robotic attempt for 3D growth measurements of whole *Arabidopsis thaliana* plants (Chaudhury et al., 2015). At a larger scale, automated remote sensing is fairly common for crop yield estimation and general vegetative productivity. For remote sensing, Normalised Difference Vegetation Index (NDVI) is a now commonly used ecological technique used to characterise the vegetative productivity of an area using satellite collected data (Pettorelli et al., 2005). Taking inspiration from such systems, an exploratory investigation was conducted in order to assess the feasibility of development of an optical computer vision based automated measurement system, the further development or adaptation of existing absorbance measurement systems for aquatic environments, and the potential for an optical system to aid the development of an aquatic measurement system both in laboratory and field settings. There has previously been an attempt to use photo scanning technology to quantitatively estimate *Ulva* growth by Ale et al. (2011). Their method entailed the physical removal of the *Ulva* thalli before imaging in a petri dish, whilst the thalli were secured with microscope glass slides. This method does address the issue of using wet mass as indicators of biomass growth, however it still requires the physical removal of the *Ulva* thalli, therefore not addressing the issue of time consumption, labour intensity or temporal resolution. The present experiment was designed to address both issues by using area measurements and providing a proof of concept for automating the process, thereby reducing the inconsistencies of wet mass measurements and removing the uncertainties associated with them, and at the same time reducing the amount of labour and time required to obtain growth measurements. The development of automation software for this purpose would require extensive coding, for which we do not have the funding, expertise or time available to us. Therefore, the present experiment will focus on assessing the feasibility and merit in developing such a system in the future.

6.2.2 Methods and Results

The envisaged set up was quite simple, consisting of three major components: the incubator, the camera, and the image analysis system. The incubator was the Algem bioreactor produced by Algenuity. The Algem is designed for use with microalgae, and the bioreactor allows for a reliable, precise and accurate system for the maintenance of temperature, light and turbidity. The proprietary software allows for the reproduction of average conditions in various geographic regions around the world at a given time of year. The Algem bioreactor is made of two growth chambers and a control panel, connected to a dedicated cooling system as seen below in figure 6.1.

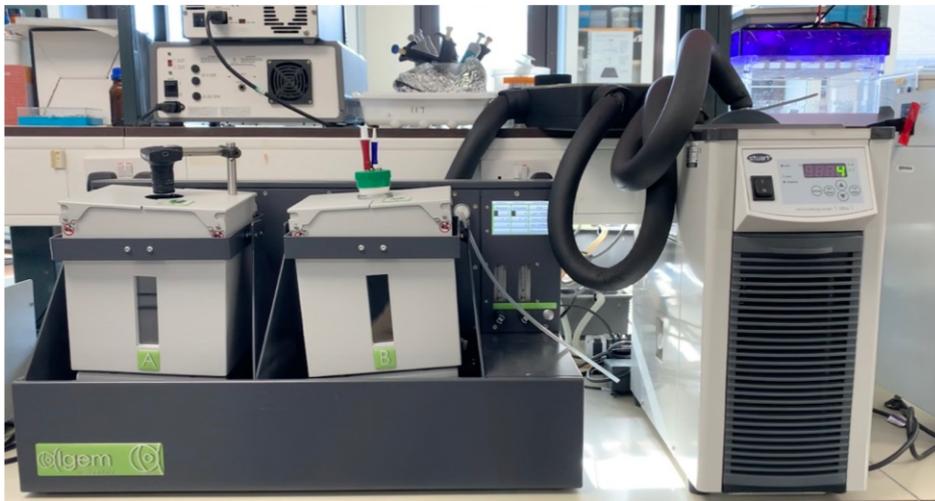


Figure 6.1: Algem bioreactor by Algenuity seen on the left, with a touchscreen control panel. To the right, the dedicated cooling system can be seen, and the flow control unit in the middle, with all of the pipes containing cooling fluid covered with an insulating foam layer.

The growth chambers are designed to use with standard 1 litre wide mouth conical flasks and are supplied with specialised bungs to allow nutrient and gas connections, as shown in figure 6.2 below.

Each growth chamber is capable of maintaining different conditions, which can be useful for comparative experiments.

The camera configuration was the first challenge. There were only three primary requirements that could not be compromised. Firstly, that the configuration must be able to capture continuous footage and still images of the experiment over the course of multiple days, at least up to and including 21 days and ideally indefinitely adjustable.



Figure 6.2: A standard wide-mouth 1 litre beaker fitted with the specialised bung supplied by Algeniuty for the Algem. The various nozzles can be configured to allow nutrient and gas supply to the culture.

Secondly, the footage and the still images captured must contain useful data, while still reducing the irrelevant background information to an absolute minimum to aid in the image processing.

Finally, the camera configuration must be adaptable enough to, at the very least, allow for adjustable focal lengths. Ideally, the configuration would be adaptable enough to allow for experimentation with various lens types, in order to investigate the ideal configuration in a variety of scenarios.

This set of requirements was considered in conjunction with the restriction in budget. The first consideration was a commercial digital camera, both compact and DSLR, that has autofocus and optical zoom capabilities. This, however, was quickly ruled out due to a number of limitations. A DSLR configuration for long-term time-lapse projects will cost thousands of pounds and are known to be difficult and temperamental amongst photographers, leading to many professional companies offering long term time-lapse photography project services for a significant cost. Compact digital cameras, then, seem like a more viable option. However, most manufacturers employ the use of various proprietary technologies to enhance and adjust images, and accounting for these variables could become problematic. These types of images would also lead to reliance on computer vision

as the effects would complicate the use of thresholding. For these reasons amongst others, the decision was made to investigate scientific camera systems.

Scientific cameras, like commercial cameras such as DSLR (digital single-lens reflex) cameras, can cost thousands of pounds for the camera body itself, for example the flagship device from industry-leading company Thorlabs costing £ 9150.26 exclusive of tax (Thorlabs, 2019). Therefore, much deliberation had to be done in order to find an acceptable compromise of cost and quality of data.

The first consideration is the sensor in the camera. Most modern commercial digital cameras in production today have a solid state array with a CCD (Charged Coupled Device), CID (Charge Injected Device) or CMOS (complementary metal-oxide semiconductor) sensor (Canon, 2019; Holst and Lomheim, 2011). CCD sensors were considered far superior to CMOS sensors for decades since the development of both types in the 1970s due to the significantly lower noise and higher light sensitivity in CCD sensors. However, the gap has narrowed significantly in recent years with the advancement of technology, and the shortcomings of CCD sensors are being exposed as engineers look to push the boundaries of both technologies.

There are four primary benefits to CMOS sensors. Firstly, CMOS sensors have faster processing speeds due to every pixel being capable of individual read-outs, as opposed to being limited to one or two read-out corners in a CCD sensor. Secondly, CMOS sensors are directly compatible with mainstream silicon chip technology and are capable of performing on-chip image processing tasks, rather than having to read out the data to a separate dedicated chip. Thirdly, CMOS sensors have much lower power consumption, which is a significant advantage for portable devices. The combination of these factors mean that the future development potential of CMOS sensors is much higher. As silicon technology advances, CMOS sensors will be able to perform even more tasks, more accurately, using less energy and in less time, all the while being significantly smaller and cheaper (Holst and Lomheim, 2011). This is reflected in the market with most devices now being manufactured with CMOS sensors, rather than CCD sensors, and manufacturers heavily preferring investing research and development resources in CMOS sensor technologies (Canon, 2019; Moynihan, 2011).

Sensor sizes then come into consideration, with 35 mm full frame (36 x 24 mm) sensors

being the reference standard. Every other size of sensor is referred to as a crop sensor, and is referred to as such, as images captured using a crop sensor shows a cropped view of an image captured using a full frame sensor, given the same lens and subject distance. Most high-end DSLR and mirrorless cameras are equipped with full frame sensors, whilst the majority of these types of cameras have a crop APS-C (22.7 x 15.1 mm) sensor (Gerlach and Gerlach, 2007). The bigger the sensor, the more expensive it is to produce, and so the price of cameras equipped with larger sensors reflect this.

The present project itself does not require extremely high picture quality, and is budget-restricted; therefore a crop frame CMOS sensor would be most logical. A monochrome sensor allows for accurate native binary thresholding and avoids many of the artefacts that are associated with demosaicing and multiband thresholding of colour images (Förstner and Wrobel, 2016).

All of these requirements led to the implementation of a camera with a 1/2-inch monochrome CMOS image sensor, with 5.2 x 5.2 μm pixels at a resolution of 1.3 Megapixels (1280 x 1024 pixels) and a standard CS-Mount for the lens. The CS-Mount requires a C-Mount adapter to use C-Mount extension tubes and lens tubes. After experimenting with various configurations, the first configuration that could yield meaningful data was the camera being attached to a lens tube (10 mm) with a lens and an iris diaphragm, then another lens tube (50 mm). The latter lens tube is inside a silicone bung, in order to insert the set up into the experimental conical flask. Figure 6.3 below summarises the initial configuration of the camera and lens.



Figure 6.3: Visualisation of the camera and lens configuration initially used in the preliminary experiments.

This would give a light path approximately shown in figure 6.4 below.

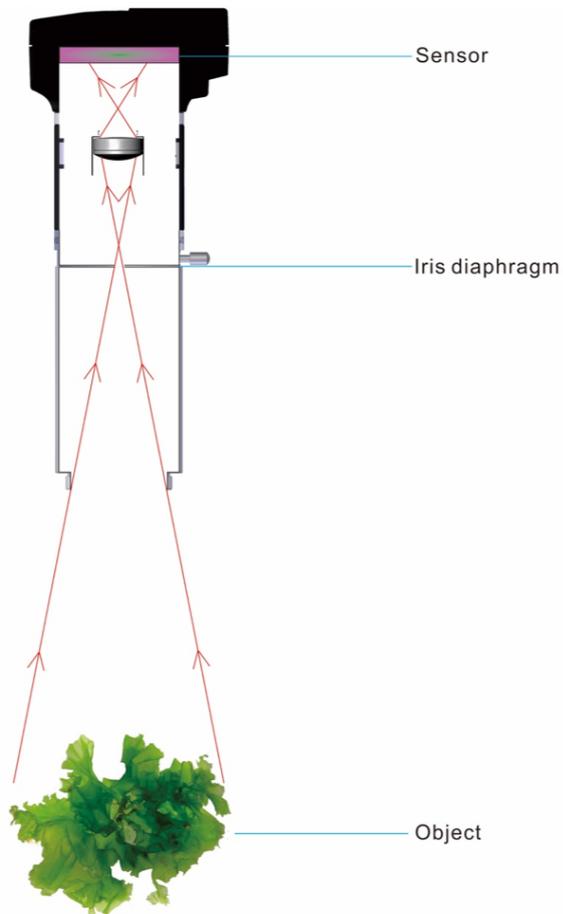


Figure 6.4: Approximate light path of the camera and lens configuration used in the preliminary experiments.

In order to make the light as diffuse as possible for accurate binary thresholding, an opal glass light diffuser was placed on top of the light source inside the growth chamber. This would give more even images that can have the threshold set more easily, allowing for higher accuracy and more margin for error in automated image processing. Having experimented extensively with a variety of configurations, it became apparent that using a system like those above would prove difficult to keep oxygenated and maintained with nutrients without regular switching of media. This was not an impossible task, with custom engineered bungs that could make the configuration work as intended. However, the crucial part was that the images obtained were too difficult to process automatically, due to physical limitations of the optics. The lens had to be placed inside the neck of the conical flask with a bung, which made achieving an ideal field of view impossible. The decision was made to abandon the conical flask and lens design, and instead of trying to compensate with a complex lens

configuration, a simpler approach was sought.

In order to achieve the simplest configuration possible, the next phase of this project involved switching the container inside the growth chamber to standard 1-litre wide-mouth beakers, instead of conical flasks. This eliminated the initial issue of a restricted field of view. In order to mount the camera, the lid of the growth chamber was machined, and a camera mount was constructed and attached. A simple varifocal lens was used in place of the complex lens configuration as shown in figure 6.5 below.

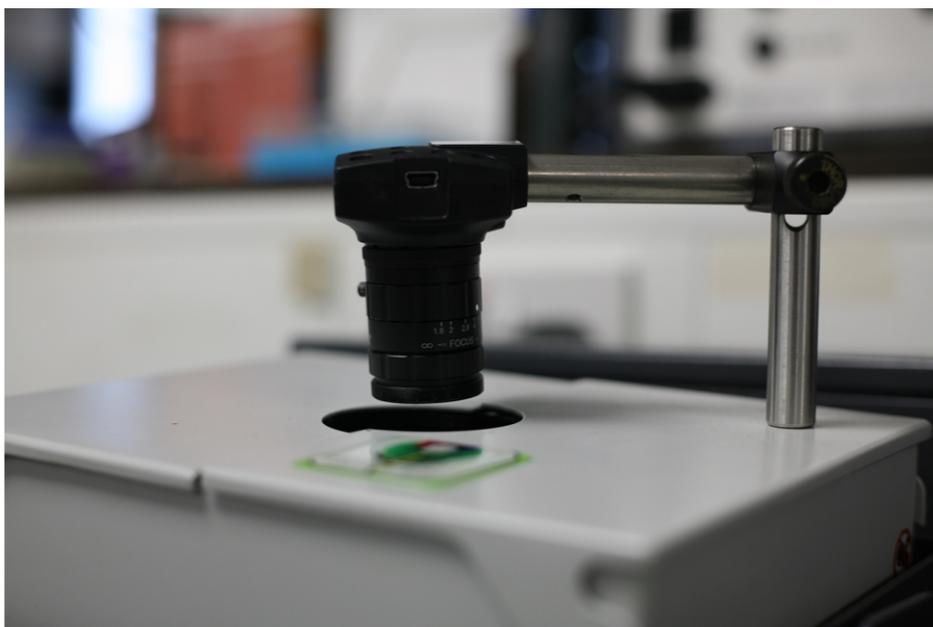


Figure 6.5: The final camera and varifocal lens configuration used for the duration of the present experiment. The camera is mounted directly onto the growth chamber lid.

This configuration allowed for easy adjustment of the focal length, as well as a more standardised system that could be replicated very easily for the other chamber on the Algem bioreactor. With the optical troubleshooting completed, a more thorough investigation of the biological system could now be done.

Ulua thalli in growth experiments of two and three days duration exhibited a slightly more problematic fluid dynamic behaviour that hadn't been previously observed. As the thalli became larger, they tended to fold within themselves at various moments in time, making the automated analysis problematic in the absence of any compensatory code. While this was being written, it occurred that it had previously been demonstrated that immobilisation of *Ulua* does not seem to affect its growth given that adequate surface removal occurs, allowing for the supply of nutrients and removal of waste products (Ale et al., 2011). With

this in mind, three methods for immobilisation of thalli were conceptualised.

The first method was to have plastic rods protruding from a plastic base, with a flat plastic piece towards the top of the rods upon which the *Ulva* thalli would rest, as shown in figure 6.6 below.

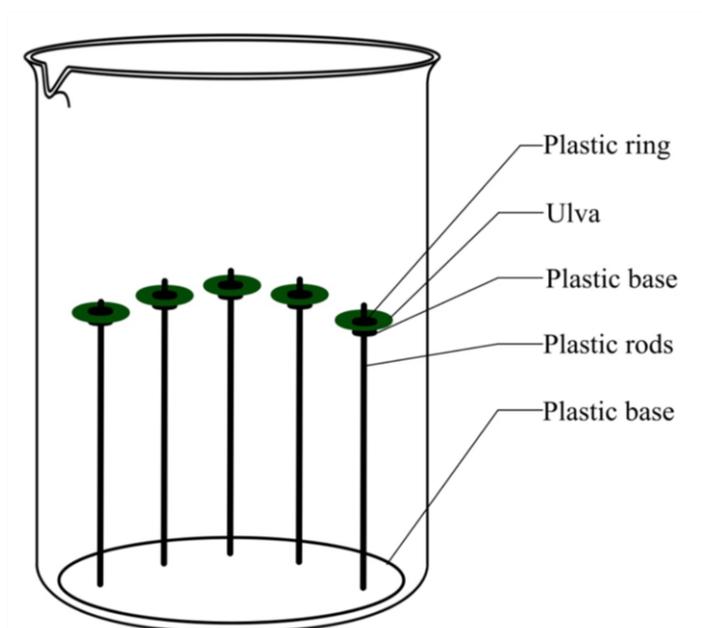


Figure 6.6: First prototype design for the fixation of *Ulva* thalli within a 1 litre beaker.

The *Ulva* thalli would effectively be skewered onto the plastic rods, rest on the plastic base and be secured in position by a plastic ring with a rubber lining on the inside of the aperture. The main material was envisaged as a biologically and chemically inert polymer, such as PVC or acrylic, in order to minimise the effects that the material itself would have on growth. The material would also have to be easily machinable and formable, so other inert options such as borosilicate glass are ruled out. Furthermore, many plastic compounds that would be considered for use are cheap and readily available, making it ideal for prototyping.

This set up will allow for surface removal, allowing for all parts of the thallus not secured by the plastic pieces to obtain nutrients and grow normally.

The obvious negatives of this apparatus is the effect physical trauma to the *Ulva* thallus would have on its growth. As previously discussed in Chapter 1, fragmentation of *Ulva* thalli is widely known to induce the formation of sporangia and subsequent release of spores, and is therefore a common technique used in laboratory cultures (Carl et al., 2014;

Dan et al., 2002; Gao et al., 2010, 2017; Hiraoka and Oka, 2008). Therefore, the growth rate of the thalli in the experimental set-up will be affected by the skewering of the thalli.

The second method, shown in figure 6.7 below, is a tangential concept building on the basic design of the first method. This again involved plastic rods protruding from a plastic base, but with a magnet fixed at the top of the rods.

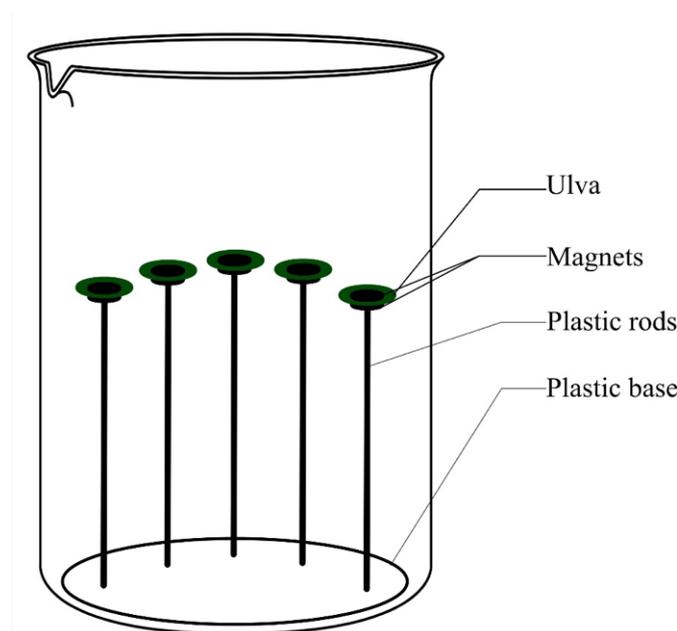


Figure 6.7: Second prototype design for the fixation of *Ulva* thalli within a 1 litre beaker.

The *Ulva* tissue would be placed on top of the magnets and secured with a magnet on top. The magnet would also be sealed in plastic, in order not to directly expose the magnet itself to the growth medium. The magnets were envisaged as neodymium-iron-boron (NdFeB) magnets, which are the strongest permanent magnets in the world. A common method of measuring magnetic strength is by measuring the maximum energy product, typically in MGOe (megagauss-oersted) or the SI unit kJ m^{-3} (kilojoule per cubic metre), where $1 \text{ MGOe} = 7.958 \text{ kJ m}^{-3}$ (Hatch and Stelter, 2001). The energy product of NdFeB magnets are approximately 56 MGOe, as opposed to approximately 1 MGOe for early steel magnets and 3 MGOe for hexagonal ferrite magnets (Gutfleisch et al., 2011). This was in order to ensure a higher energy product, which in turn would allow this set up to use much smaller magnets compared to an equivalent set up using ferrite magnets. This will allow for minimum contact area with the *Ulva* thalli and maximum surface removal, which minimises the effect the apparatus has on growth.

The most prominent argument in favour of this design is that it does not physically damage the tissue like the first method does. This design carries forward and improves on the same benefits afforded by the first design, allowing surface removal on a larger proportion of the tissue.

The clamping force exerted by the magnets could be a factor affecting the growth of the *Ulva* thalli, however the exact strength of the magnets was unable to be obtained from the manufacturer. Observationally, the magnets were strong enough to secure the thalli in place but did not leave any physical marks on the tissue. However, there was a bigger concern regarding the effects of magnets on *Ulva* growth: A crucial negative argument, and the fatal one for this method, is that there is no information available on the effects of magnetic fields on *Ulva* growth, and the limited accurate and reliable information on the effects of magnetic fields on plant growth in general. As broad examples, it has been shown that magnetic fields induce higher growth rates and affect gene expression in the model bacterium *Escherichia coli* (Potenza et al., 2004), and increases biomass production and glutathione production in the model yeast *Saccharomyces cerevisiae* (Santos et al., 2010). Since L. J. Audus, (1960), a botanist from Bedford College, published one of the first studies on the effects of magnetic fields in living systems arguing for an auxin-like effect of a phenomenon he termed “magnetotropism”, there have been a multitude of experiments examining the various aspects of plant growth and development that are affected by magnetic fields, though there is little consensus and a critical lack of any mechanistic understanding, and no shortage of fierce and, at times, scathing criticism of everything from the experimental design, to the hypotheses and analyses of the studies conducted (Barnothy, 2013; Galland and Pazur, 2005; Maffei, 2014).

However, given that the magnets themselves are a structural component of this experimental apparatus, it would be impossible to design an equivalent control experiment. Therefore, due to the absence of a suitable control experiment, this method was abandoned.

The third method was chosen after much deliberation, given that it does not suffer from either of the major drawbacks of the first and second methods that would require extensive preliminary experimentation, and that there have been experiments in the past based on similar ideas (Ale et al., 2011).

A number of designs were drafted from the concept of the third method, and a final design utilising acrylic plates was prototyped, with the design shown in figure 6.8 below.

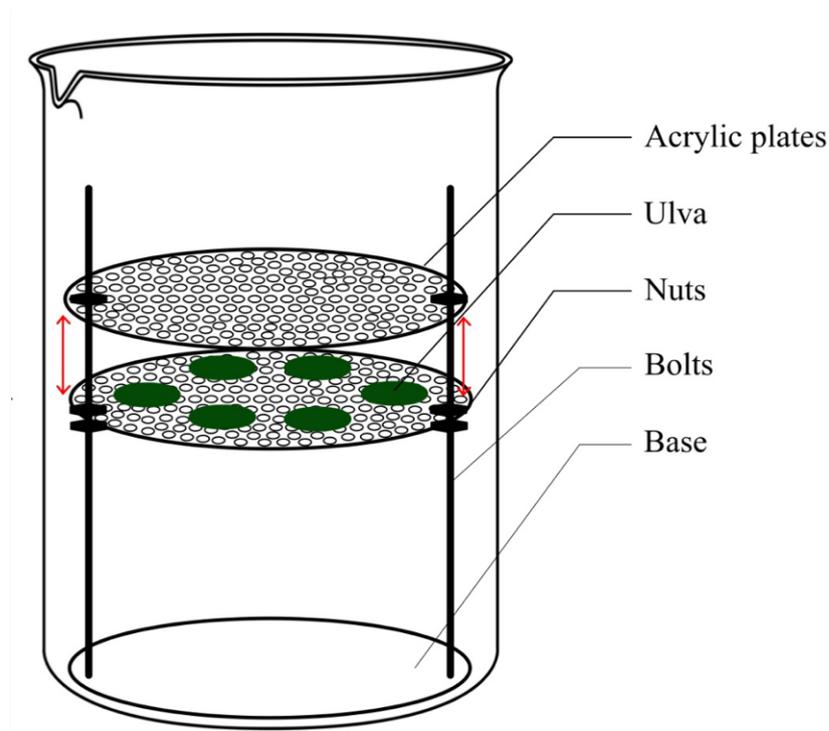


Figure 6.8: Final prototype design for the fixation of *Ulva* thalli within a 1 litre beaker.

The basic principle of this design is two bolts protruding from a plastic base, with two acrylic plates secured together. This is achieved with one nut on each bolt securing the position of the lower plate, a nut or washer on each bolt between the lower and upper plate, and a nut securing the top plate down onto the nut or washer between the two plates. As the securing is achieved by positioning nuts and washers on bolts, the position of the acrylic plates within the water can be adjusted easily. The *Ulva* thalli are secured between the two plates, and surface removal occurs through the multiple holes drilled into each of the plates.

This design was originally conceived as a low-diameter stainless-steel mesh, as it would allow for the minimisation of individual contact areas, meaning that any one part of the thalli will have minimal contact with the apparatus, and the contact will be more uniform throughout the thalli, than a design using acrylic plates. However, the difficulties and inconsistencies in machining the mesh as well as the warping of the mesh contributed towards the inconsistent background, making it more difficult to establish a baseline for

image processing. Thus, the stainless steel was replaced with acrylic.

Preliminary experimentation showed that the acrylic plates could be positioned closely enough that the thalli are essentially immobilised, in that there was not enough movement for the overall positioning of each thalli to change. However, it was possible to achieve a configuration wherein the acrylic plates were far enough apart to allow a film of water between the plate and the thalli.

The growth experiments from this design were satisfactory, with good correlation between the wet weight and image analysis measurements, after initial calibration. The calibration step involved placing square and circular pieces of paper of known area, side lengths and diameters into the beaker filled with a known volume and height of water. The pieces of paper were then imaged and analysed, then manually calibrated to set the pixel to millimetre ratio. Then, square and circular *Ulva* thalli of known area, side lengths, diameters and mass were used to double-calibrate the measurements and give a baseline for mass approximation in image analysis.

The calibration step, however, inspired another experiment using free-floating thalli instead of fixed thalli. The initial portion of the experiment involved observing the physical movement of *Ulva* within the beaker inside the growth chamber, in order to assess the fluid dynamic behaviour of free-floating *Ulva* thalli. *Ulva* seemed to behave typically of a low-density, semi-rigid solid, largely remaining on the surface of the water and rotating in the direction of the water turbulence caused by the rotation of the growth chambers. This was then extended from a short observation over the course of a working day, to a week-long observation study, with the hypothesis being that free-floating *Ulva* thalli will continue to behave in the observed manner, as a low-density, semi-rigid solid, and continue to remain largely on the surface at this size (between 1 and 5 cm).

As hypothesised, small free-floating *Ulva* thalli remain on the surface of the water at high mixing levels (over 100 RPM). In this configuration, free-floating *Ulva* thalli showed remarkably consistent movement within the growth chamber. Image analysis of free-floating thalli yielded satisfactory results, with the analysed area corresponding well to the known wet mass, and statistical analyses supported this with $p < 0.05$ and $R\text{-squared} = 0.996$, as shown in figure 6.9 below.

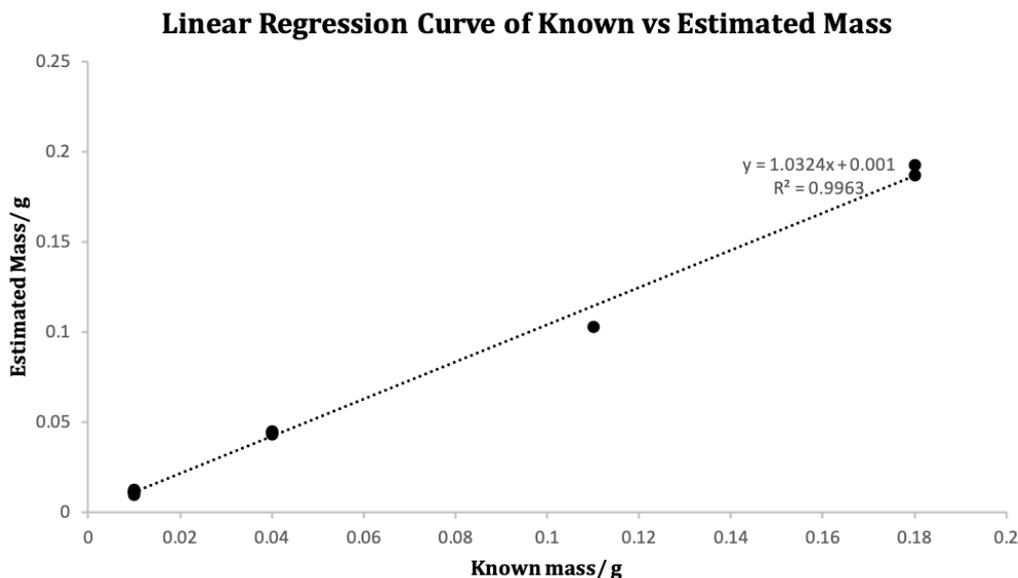


Figure 6.9: Regression curve of known mass in grams versus mass estimated by image analysis in grams.

The final prototype configuration was of free-floating *Ulva* thalli imaged by a camera with a varifocal lens mounted directly onto the growth chamber. These images were then analysed to give length and area measurements in pixels, that could be converted into millimetres from the initial calibration, and the mass could be estimated, also from the initial calibration. This analysis process could be entirely automated.

Though accurate, the given configuration is limited by the surface area of the water available for imaging, as the simplistic approach measures the top area of the thalli, with no provisions to account for overlapping within and between thalli. This is an issue that can be eventually addressed, should a larger experiment be desired, using two potential techniques. The first is by utilising multilevel thresholding, as opposed to the binary thresholding used here, for example by using established methods such as Otsu's or Kapur's methods (Baby Resma and Nair, 2018; Kotte et al., 2018). This approach can identify the increased intensity of signal caused by overlaps, and consequently account for this variation. The second method would be the addition of more cameras that could provide depth information to allow for three-dimensional or pseudo-three-dimensional imaging. Though the potential to scale up has not been investigated, controlled growth in standard 1,000 or 10,000 litre tanks does not seem impossible, though extremely difficult, with a rotating multiple camera configuration that could provide robust depth information. At the mass measurement scale,

the use of a system similar to NDVI (Nominal Difference Vegetation Index) calibrated to aquatic vegetation could potentially yield more valuable information than optical data as this would solve a bottleneck in biomass monitoring.

This experiment did not aim to develop a fully-fledged automatic biomass measurement system, but rather to explore the possibility of, and labour required in developing such a system. One of the key findings of this chapter was that the edges of the thalli experiences higher relative velocities by up to 40.3% compared to the centre. This adds to the hypothesis that *Ulva* morphology is driven by differential growth. This means that this experiment did yield useful observational data that gave confidence to undertaking the next phase of experiments.

6.3 EdU Cell Proliferation Assays

6.3.1 Introduction

EdU (5-ethynyl-2'-deoxyuridine) is a newly developed thymidine analogue, like BrdU (5-bromo-2'-deoxyuridine) which has been used in biological research for decades (Mead and Lefebvre, 2014). Thymidine analogues are incorporated into new DNA during S-phase, marking newly formed DNA and identifying newly divided cells (Flomerfelt and Gress, 2016). This technique allows for the quantification of DNA synthesizing cells, which can prove extremely useful, especially in experiments investigating proliferation patterns (Mead and Lefebvre, 2014).

The use of EdU is extremely recent, with the method only having been developed in 2008 by Salic and Mitchison (2008). While BrdU was extremely useful in studying cell cycle kinetics, DNA replication and cell proliferation, amongst other subjects, there are significant drawbacks that have always limited the capabilities of what could be investigated (Chehrehasa et al., 2009; Salic and Mitchison, 2008). The primary disadvantage is that BrdU incorporation in double-stranded DNA involves complementary base pairing, resulting in the access to BrdU subunits by anti-BrdU antibodies being blocked (Salic and Mitchison, 2008). In order to access the BrdU epitope, the sample must be subject to strong denaturing conditions, necessitating heat and strong acid treatment steps, resulting in sample degradation with much variability between experiments (Chehrehasa et al.,

2009; Salic and Mitchison, 2008). The epitope is required to be exposed as BrDU detection requires the use of an immunohistological stain, and therefore is difficult and limited by the penetration of the tissue by the antibody and the time is limited by the diffusion rate (Diermeier-Daucher et al., 2009; Salic and Mitchison, 2008).

The main advantage, and the significant improvement of EdU over BrdU is that detection is based on a Cu(I) catalysed cycloaddition of a fluorescent azide to the ethynyl group. The reagents are approximately 1/500th the size of an antibody molecule, and therefore penetrate the tissue more readily and have a higher diffusion rate (Salic and Mitchison, 2008). The nature of this “click” chemistry means that no intensive denaturing step is required for the detection of EdU opening up the possibilities for *in vivo* labelling studies with the development of cell-permeable fluorescent azides. All of these properties result in structurally well-preserved and evenly labelled samples, and detection that is not only rapid, but highly sensitive and selective, leading to low background noise (Flomerfelt and Gress, 2016; Salic and Mitchison, 2008).

EdU, then, seemed the obvious choice for investigating cell proliferation in *Ulva*. The initial concerns were the lack of pre-existing protocols and the observed tendency of *Ulva*'s resistance to many common dyes. The use of EdU will be vital in investigating the cell proliferation patterns of *Ulva*, specifically the existence and the rate of differential growth in different parts of the thalli. Therefore, to further this end, the aim for this experiment was to establish an EdU cell proliferation assay protocol for *Ulva*.

6.3.2 Methods and Results

Thermo Fisher offers EdU imaging kits with Alexa Fluor 488, 555, 594 and 647 azides, with the numbers indicating the lasers for excitation. The main aim was to allow for channel separation of double-dyed samples with Direct Yellow 96 and the EdU Alexa Fluor azide. This automatically ruled out 488, as the 488 nm laser is one of the main excitation lasers for Direct Yellow 96. The Alexa Fluor 594 and 647 azides were initially discounted as the 458 nm laser used to excite Direct Yellow 96 also excites chlorophyll a and b. The emission ranges of chlorophyll a and b is broad at around 600 to 800 nm, which would overlap with the Alexa Fluor 594 and 647 azides emission peaks of 617 nm and 668 nm respectively (Kalaji et al., 2014, 2017; Pfarrherr et al., 1991; Thermo Fisher Scientific, 2019b,c). This left the Alexa Fluor 555 azide as the prime candidate, as the

excitation laser for planned use was 543 nm, and the emission peak is 565 nm, allowing for a tightly controlled channel separation whilst also being in the range to block out much of the chlorophyll a and b fluorescence (Thermo Fisher Scientific, 2019a). Therefore, the decision was made to purchase Click-iT EdU Alexa Fluor 555 imaging kits (catalog number: C10337) from Thermo Fisher.

The initial problem, that could be identified before commencing the trials, with the standard kit provided is the Hoechst 33342 dye provided for the staining of DNA. The excitation emission peaks for Hoechst 33342 is 350 nm and 461 nm respectively, however the lowest wavelength laser available for use was 405 nm. Whilst there have been numerous reports of successful excitation of Hoechst 33342 with a 405 nm laser (Banerjee et al., 2008; Lou et al., 2013; Sabnis, 2010; Shapiro and Perlmutter, 2001; Szczurek et al., 2014), this unfortunately was not the case in these trials. There are a number of other troubleshooting steps to further identify the root cause of the problem. The first logical step is to attempt staining with Hoechst 34580, which is excited at 392 nm and is readily excited by violet lasers (Sabnis, 2010; Shapiro and Perlmutter, 2001). If DNA staining is observed with Hoechst 34580, it would confirm that the problem is not the uptake and staining, but rather the excitation and detection. However, if no staining is observed with Hoechst 34580, it would indicate that the problem could be either the uptake and staining or the excitation and detection, or both. Other DNA binding fluorescent dyes, most obviously DAPI, could then be used to attempt DNA staining in *Ulva* and continue with troubleshooting and optimisation of a protocol. However, the priority was deemed to be the actual EdU Alexa Fluor 555 azide staining, and so that was the main focus of the trials.

The next step was to proceed with the suggested protocol from the manufacturer Thermo Fisher. This protocol was attempted three times, on three samples each time. However, each attempt failed, with no detection of any fluorescence excluding autofluorescence.

Due to the tendency that *Ulva* displays in its apparent resistance to dyes, the first troubleshooting step attempted was the use of a more rigorous permeabilisation step. This involved the addition of incubation for 1 hour in 1 % Triton X-100, instead of 20 minutes in 0.5 % Triton X-100. This made no improvement to the results; therefore, advice was sought from Dr Angharad Jones, who has extensive experience with EdU in *Arabidopsis*. The protocol kindly shared by Dr Angharad Jones, was initially successful on *Arabidopsis* cotyledons that were used as a control to test the protocol. However, this protocol was unsuccessful

for detection in *Ulva*. Having experienced the difficulties of staining *Ulva* before, I made informed guesses based on observational data, to adapt the protocol for *Ulva*. The final version of the protocol used here is as follows:

1. Prepare a solution of 0.1% Triton made up with growth media containing 40 μ M EdU.
2. Incubate at 8 °C and 80 rpm in the shaking incubator overnight, protected from light.
3. Wash three times with 0.1% Triton x 100 in PBS (3 washes, 5 mins each).
4. Incubate in Ethanol:Acetic Acid (3:1) for 1 hour, or until the leaves are colourless.
5. Fix in FAA solution overnight.
6. Rinse in 3% BSA in PBS (Three washes, 5 minutes each).
7. Rinse in PBS (three washes, 5 mins each).
8. Add detection solution. e.g. for 1 ml:
 - 860 ul Click-iT reaction buffer
 - 40 ul CuSO₄
 - 2.5 ul Alexa Fluor azide
 - 100 ul reaction buffer additive
 - 0.1 mg Ascorbic Acid
9. Incubate on bench for 1-2 hours, protected from light.
10. Rinse stems with PBS (3 washes, 5 mins each).
11. Image.
12. Store in PBS at 4 °C. Protect from light.

This protocol led to successful detection of EdU in *Ulva*, and so a double stain was attempted with EdU Alexa Fluor 555 azide and Direct Yellow 96, as seen in figure 6.10 below.

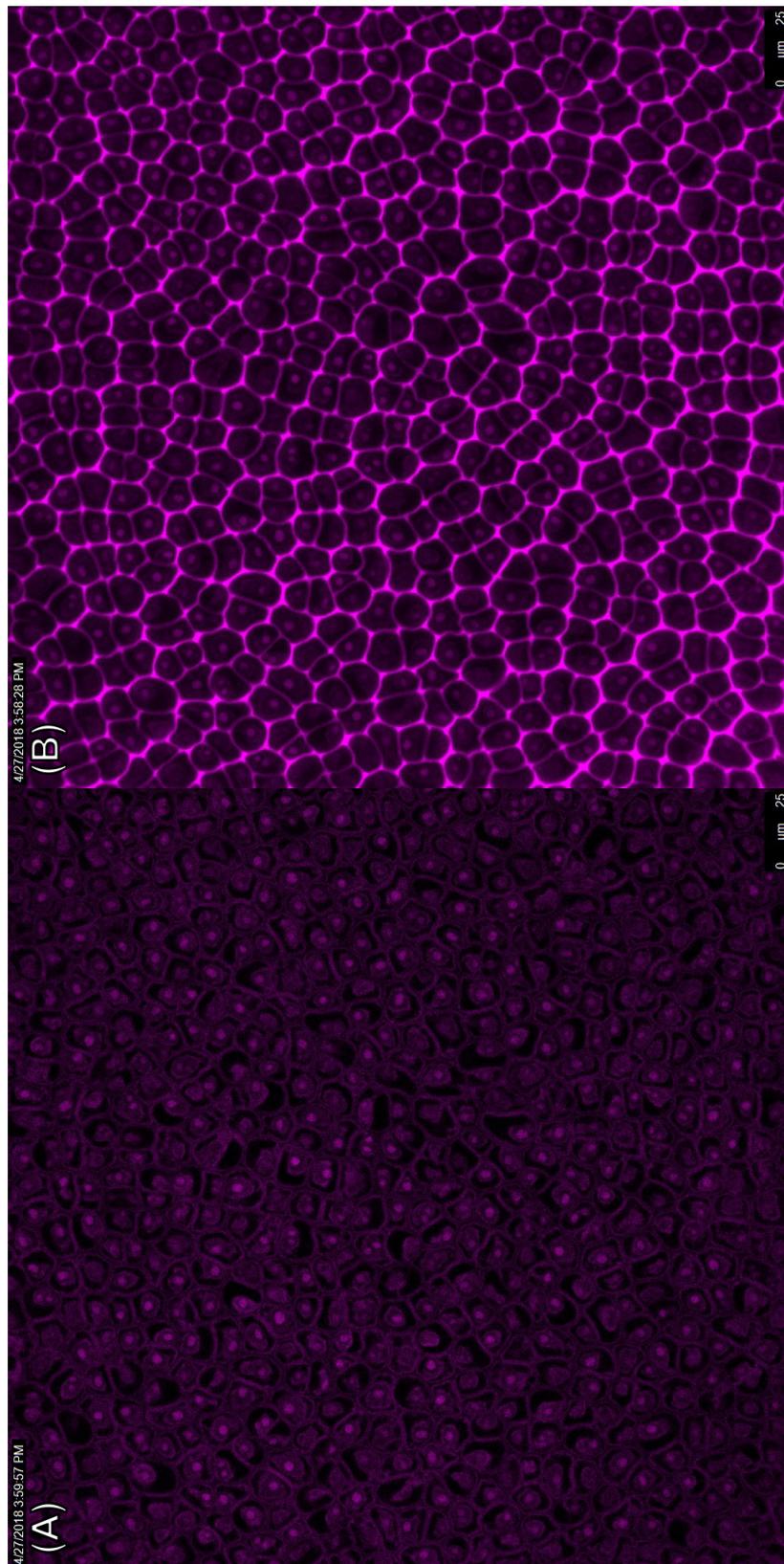


Figure 6.10: Confocal image of a section of an *Ulva* thallus, (A) shows the single channel of 543 nm laser excited emissions (EdU and autofluorescence) and (B) shows the dual channel of 458 nm and 543 nm laser excited emissions (EdU, autofluorescence and Direct Yellow 96).

The separation of the channels, whilst very clear and even, was not absolutely clean due to the overlapping emission spectra of the dyes. This meant that the fluorescence of Direct Yellow 96 was detectable in the detection range of the EdU Alexa Fluor 555 azide, which was a known concern due to the previously observed broad emission spectra of Direct Yellow 96.

Regardless of a perfect channel separation, the main objective of establishing a functioning, reliable protocol for EdU detection in *Ulva* was achieved. At this stage, the inclusion of an incubation step in ethanol and acetic acid (3:1) in the final protocol meant that chlorophyll a and b fluorescence was not a major concern due to most of the pigments being removed in this step. Figure 6.11 below illustrates the visual effect of the pigment removal step on *Ulva* thalli.

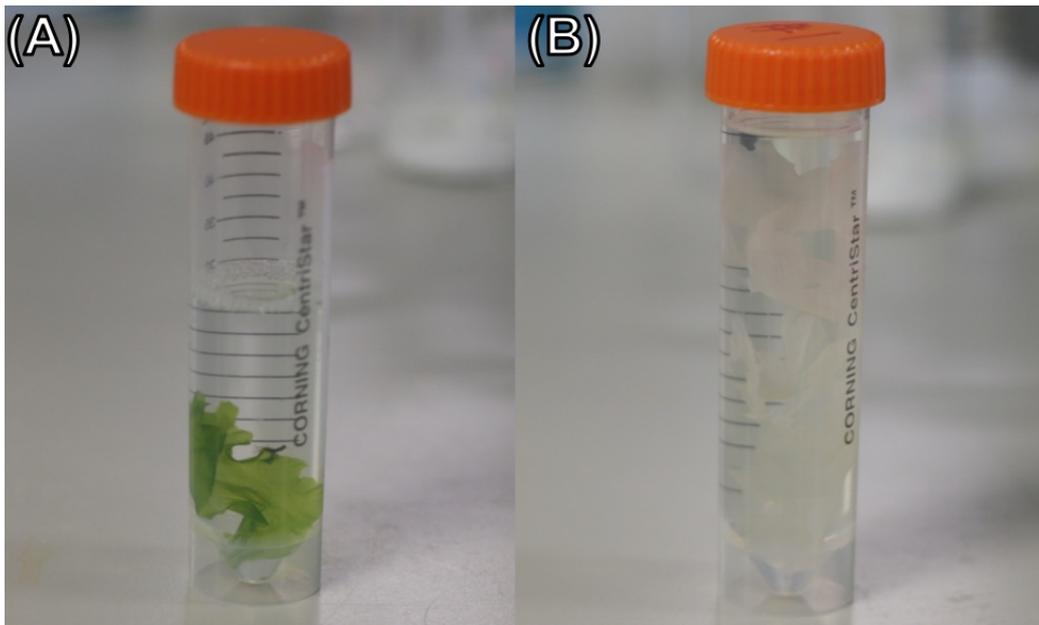


Figure 6.11: *Ulva* thalli (A) before and (B) after incubation in ethanol and acetic acid (3:1)

This means that experimentation with Alexa Fluor 594 or 647 azide could yield a cleaner separation of the Alexa Fluor azide dye and Direct Yellow 96 channels. Whilst this was not investigated due to the prohibitive cost at this point in time, it is most definitely worth consideration in future investigations.

With a working protocol having been established, a more ambitious aim was set: attain

tile scans of *Ulva* thalli, in order to investigate the differential growth in different areas of *Ulva* thalli. This is a difficult task in any tissue type, and with increasing sample size, the difficulty also increases. The nature of investigating patterns in microscopic detail over a significantly larger area leaves much room for variation, meaning that the sample preparation and detection parameters must be extremely tightly controlled in order for an investigation of this type to yield any useful data. The difficulty with *Ulva* thalli, like many other macroscopic tissues, is that the samples cannot be mounted to be completely flat, meaning that the focal depth will be different for different parts of the tissue. Leica, like most other manufacturers, offers a partial solution in the autofocus system on the SP5 Laser Scanning Confocal Microscope. However, this system is notoriously unreliable and manual focusing almost always achieves better results. Figure 6.12 below shows an example of a successful attempt.

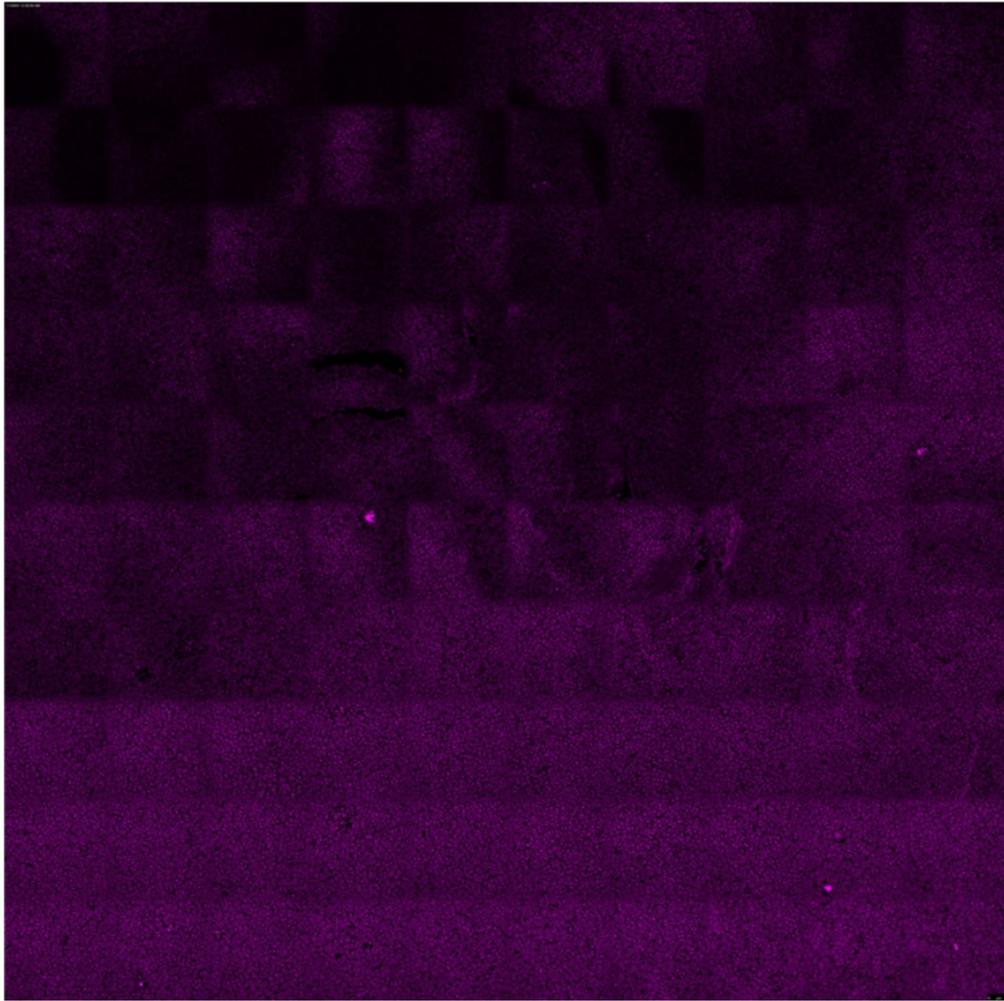


Figure 6.12: Confocal microscope image of an *Ulva* thalli stained with DY96 and EdU Alexa Fluor 555 azide. This image is a stitched 10 x 10 image of 1024 x 1024-pixel images, resulting in a total resolution of 10240 x 10240 pixels, or an area of 2460.3 x 2460.3 μm

Figure 6.12 above is a trial tile scan of a larger area, which in this instance is 2.46 x 2.46 mm. Upon close inspection, it can be observed that this image is composed of individual squares, ten squares by ten squares to be precise, resulting in one hundred images in total. Figure 6.13 below shows a cropped view of figure 6.12.

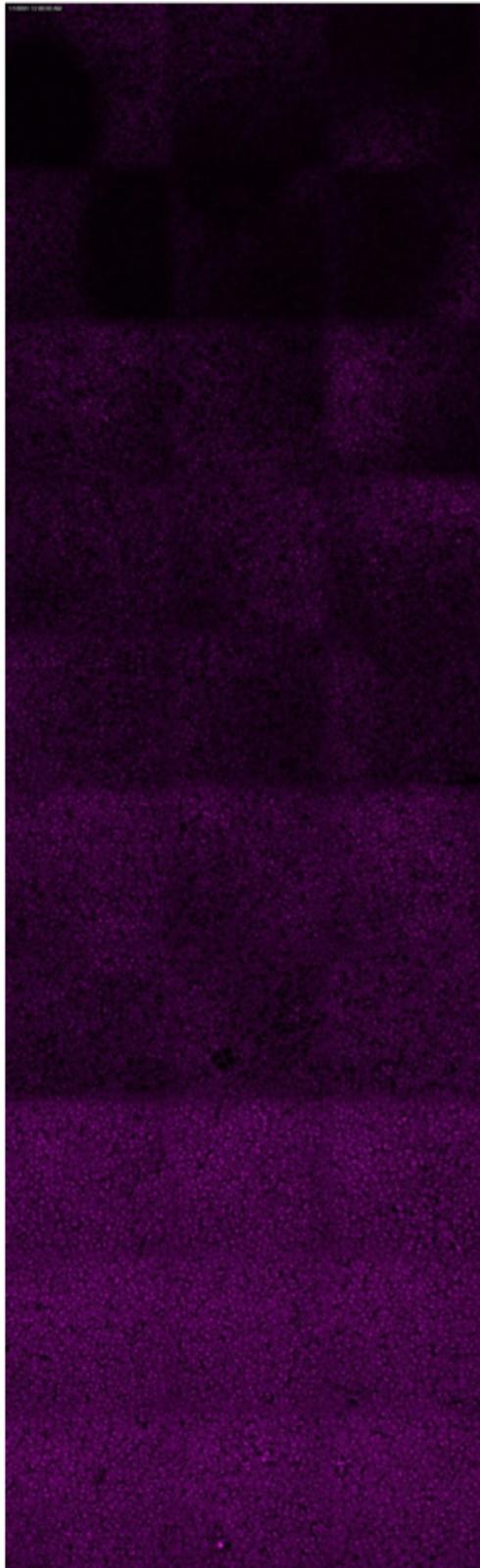


Figure 6.13: Cropped view of a confocal microscope image of an *Ulva* thalli stained with DY96 and EdU Alexa Fluor 555 azide. This image is a stitched 3 x 10 image of 1024 x 1024-pixel images, resulting in a total resolution of 3072 x 10240 pixels, or an area of 738.1 x 2460.3 μm

Figure 6.12 and 6.13 above are orientated so that the bottom of the image is closer to the edge of the thallus, and the top of the image is closer to the centre. The particular sample used for this image was a ribbon morphotype sample, chosen as it was hypothesised to illustrate a more dramatic shift in cell proliferation patterns over a shorter absolute distance. This characteristic would also lend to samples that are much more difficult to mount, due to their pronounced variation in topology. Given that the actual sample was 5.1 cm wide, this sample could capture the entirety of the variation from the centre to the edge, and a tile scan such as the one in figure 6.12 and 6.13 shows approximately 50% of the distance from the edge to the centre of the thallus. Though not illustrated in figure 6.12 or 6.13 above, the very edges of the thallus seemed to not show as much proliferation, while the same level of proliferation seen at the bottom edge of figure 6.13 is seen from a couple rows of cells into the thallus onward. This observation is potentially explained by the very edges of the thallus being more vulnerable to the exposure to the environment.

6.4 Discussion, Limitations and Further Studies

The establishment of a working protocol for *Ulva* is a pivotal step in establishing the toolkits for studies of *Ulva* cell proliferation patterns. This experiment furthers the observation that *Ulva* seems to be resistant to staining, perhaps due to the waxy cell walls. Through the use of permeabilisation, fixation and the pigment removal steps, it is possible to evenly and readily stain *Ulva* with various dyes, including Direct Yellow 96 and Alexa Fluor 555 azide, and to reliably detect these stains. Though this EdU Alexa Fluor 555 azide staining protocol for *Ulva* is functional and reliable, it would be worthwhile investigating the use of Alexa Fluor 594 or 647 azide for better channel separation from Direct Yellow 96, given that the removal of chlorophyll nullifies the initial concern of background fluorescence from chlorophyll autofluorescence.

This chapter yielded useful information about the possibilities of improving both the temporal and spatial resolution of biomass data when conducting experiments with *Ulva*. With more refinement, the optical imaging system has great potential to be an accurate and reliable method of measuring biomass growth. The potential for automation is enticing, as this will lead to high temporal resolution of data, and when the image analysis is also automated, this will prove to be a robust system for fully automatic biomass monitoring

in not only *Ulva*, but other macroalgae or aquatic organisms. The main concern with the system is the two-dimensional nature of the measurements, meaning that more complex organisms may not be best suited for this system. However, the potential to develop a three-dimensional system is also present. The spatial resolution of biomass growth monitoring can be further improved using the EdU Alexa Fluor azide protocol developed here. In conjunction, the techniques developed here enable a variety of investigations of *Ulva* cell proliferation patterns at an unprecedented spatial and temporal resolution, with the potential to improve even further.

General Discussion

7.1 General Summary

Ulva is a simple multicellular organism that is fantastic for investigating proliferative growth for a number of reasons. *Ulva* grows rapidly, being able reach growth rates of 54.5% per day (Hiraoka and Oka, 2008). *Ulva* is an extremely simple multicellular organism with only three types of cells (Wichard, 2015). The primary aim of this thesis was to investigate the nature of growth in the multicellular green alga *Ulva*. This was done through a series of experiments examining different parts of *Ulva* thalli, using varied approaches including a variety of mathematical modelling and molecular techniques.

The first set of experiments that lay the foundations that enabled the work in this project was the development of a reliable, consistent and even staining protocol. Commonly used cellulose stains such as Calcofluor White M2R and Pontamine Fast Scarlet did not yield satisfactory results in *Ulva* for reasons as yet unknown, so a trial was conducted using a variety of other dyes. Of these, a little-known dye called Direct Yellow 96 yielded excellent preliminary results, and so dozens of ensuing experiments sought to refine the protocol to achieve the desired level of reliable and even staining. This work allowed for the accurate characterisation of cellular proliferation patterns in *Ulva* and therefore allowed for the type of confocal analyses that were critical to chapter 4 and 5.

Having developed the techniques necessary for chapter 4 and 5, the next step was to substantiate the fundamental principles underpinning *Ulva* growth. Chapter 3 sought to do this by laying the foundations of the genomic background of the whole project. De Clerck et al. (2018) demonstrated that there are no expansions of transcription factors in the *Ulva*

genome in comparison to other green algae and that the retinoblastoma (RB)/E2F pathway and associated D-type cyclins gene families were absent and there is no significant expansion of transcription factors, indicating that there is no genomic signature for multicellularity in *Ulva*. The comparative genome analysis in this paper was conducted using the assumption that a discrete power law distribution is a best fit for gene family data. While entirely reasonable from initial model fitting to conclude that a discrete power law model is indeed the best fit, and indeed almost improbable that it is not, this was not supported by robust statistics. The models and robust statistics used in chapter 3 sought to mathematically and statistically confirm the assumptions made by De Clerck et al., (2018). This involved fitting an array of heavy tailed models to the gene family data for comparison, namely discrete and continuous power law, discrete and continuous lognormal, discrete and continuous exponential and Poisson distributions.

Taking the results of (De Clerck et al., 2018), chapter 4 sought to construct mathematical models to test the physical drivers of *Ulva* growth. The main method used in this chapter was Voronoi tessellation, both for the constructed model data and the collected biological data. Two phases of models were used, with four in the first phase and six in the second, giving a total of ten models. The first phase of models acted as controls and consisted of a regular hexagonal honeycomb, a “squashed” hexagonal honeycomb, a set of randomly generated diagrams and a square grid. The second phase of models were variations of the regular and “squashed” hexagonal honeycomb that introduced random deviations to the centre points that generated these models. The comparative statistics obtained from these manipulation models showed that adding 50 % random deviation models were significantly different to the preceding models. The subsequent comparison of the 50 % random deviation models to the collected biological data showed that there is no significant difference in polygon distribution between these models and the *Ulva* samples, as well as between the leaf and ribbon morphotypes.

Chapter 5 sought to build on the comparison between the leaf and ribbon morphotypes of chapter 4 and explore the generative mechanism of cellular proliferation patterns in the two morphotypes. This was done using a method called Lloyd’s algorithm to achieve a Centroidal Voronoi Tessellation (CVT) diagram, which iteratively relaxes a Voronoi diagram by using the centre of mass of a Voronoi cell as the new Voronoi generator, until the Voronoi generators are also the centre of mass. This process identifies the point at

which the diagram is the most space filling distribution possible given the set of Voronoi generators, as well as the path of progression to this distribution, or CVT path, and gives metrics of how space filling the initial distribution is. The results showed that the leaf and ribbon morphotypes are already very close to a space filling pattern, follow a near identical path towards a CVT diagram, and there are no significant differences in the distance from the initial diagram to the CVT diagram. This strongly suggests that the same physical mechanism underpins the generation of the cellular proliferation patterns seen in both morphotypes.

Chapter 6 sought to investigate the hypothesis that spatial variation in the thalli drives *Ulva* morphogenesis. In order to do this, it was necessary to develop tools that also addresses many common challenges faced in multicellular aquatic plant research as well as a problem in *Ulva* research, specifically. These are the lack of a standardised growth measurement and the lack of molecular protocols for *Ulva*, respectively. The first part of this chapter involved the development of an optical growth measurement system that would allow an investigation of live *Ulva* growth within Algem bioreactors. The results showed that optical measurements could accurately approximate mass measurements and demonstrated the differences in relative velocity between the centre and edge of the thalli. Further development of this system will allow for more reliable mass measurements at higher spatial resolutions and unprecedented temporal resolution of multicellular aquatic plants. The second part of this chapter involved the development of a working protocol to stain *Ulva* with EdU Alexa Fluor dyes. Having developed a working protocol, this was used for preliminary experiments that explored the differential growth in *Ulva* thalli, namely between the centre and edges of the thallus. The results showed that the rate of proliferation was higher in the edges of *Ulva* thalli than in the centre, supporting a long-standing hypothesis that was formed to explain the morphogenesis of the typical rosette shaped *Ulva* thalli. This experiment also led to an experiment to refine the Direct Yellow 96 staining protocol. I found that the same permeabilisation, fixation and pigment removal steps from the EdU protocol yielded a clearer, more even signal with less background noise with Direct Yellow 96 as well. This step will be indispensable in developing automated analysis using computer vision and machine learning.

Taken together, the results of the investigations in this thesis provide a strong case of fundamental physics driving this type of generative growth in *Ulva*, though it is not necessarily

limited to *Ulva*. As discussed in chapter 5, proliferating epithelia in metazoans were the first tissue type subject to the type of mathematical investigation similar to that employed in chapter 5. However, early investigations on plants inspired this field of work, and due to the physical nature of this type of proliferative growth, much of the work on one type of organism is directly relevant and applicable in other types of organisms.

7.2 Future direction of *Ulva* Research

The bottom-up growth mechanism observed in *Ulva* are a fantastic model for studying growth mechanisms of any epithelial tissue, as *Ulva* does not have any top-down control of multicellular growth. *Ulva* is the first organism without top-down controls that has been analysed in the manner presented in this thesis. Much of the literature focuses on metazoan tissues and therefore cannot categorically isolate the bottom-up growth mechanisms. These findings in *Ulva* confirms many of the fundamental concepts found in metazoan epithelial tissues (Gibson et al., 2006; Gibson and Gibson, 2009; Nagpal et al., 2008) and also the more recent developments of these concepts, especially regarding the self-organisation of tissues (Sánchez-Gutiérrez et al., 2016). Further comparative studies would benefit from the inclusion of *Ulva* for the reasons demonstrated here.

Unlike many of the organisms that have previously been studied in this field, *Ulva* will continue to be a good model organism due to its simple nature. A famous and often paraphrased quotation summarises this well: “It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.” (Einstein, 1934). Metazoan tissues have various complications regarding their top-down controls of growth, whereas *Ulva* is the simplest organism in which to study cellular proliferation patterns without losing key elements of multicellularity, such as in colonial multicellular organisms.

Much like most research in the world, the future of *Ulva* research will revolve around computational modelling, automation and Machine Learning. There are three major steps for the particular research presented in this thesis, which will ultimately lead to one goal. First, is the automation of data collection and analysis will allow for fast and accurate characterisation of polygon distributions in epithelial tissues. Second, is the automation

of a Computer Vision system that will allow for the automation of growth monitoring. Third, is the characterisation of the patches, their mechanical properties, function and their generative mechanism. These three steps will form the basis of the construction of a generative model that can accurately model growth of *Ulva* from cell to bloom. This will be one half of the model.

The other half of the model is crucial for its usefulness in applications to problems. Using, the steps above will be pivotal in the progression to manipulation experiments, wherein *Ulva* will be subject to many different growth conditions, ranging from excessive and limited nutrients to exposure to toxic substances. With enough conditions and replicates, this will allow for a database to be constructed, from which data can be extracted for simulations. Using the generative model and pairing it with the manipulation data, it will be possible to simulate ideal conditions for growth for use as biofuel feedstock, food, or other commercial uses, as well as reverse engineering the conditions that lead to green tides.

This simulation system as a whole could be applied to other organisms, given the appropriate adaptations for data collection and analysis. In particular, research in other multicellular aquatic plants, especially macroalgae, would benefit from this system. In parts, there are multiple aspects for application and the broadest application would be the Computer Vision system for growth monitoring. This would be able to be used in most multicellular aquatic plant research, given the appropriate adaptation of parameters to adjust for the unique characteristics of the organism under investigation.

On a more speculative note, a potentially important area of research that could benefit from the investigation of the proliferative mechanisms of epithelial tissues is cancer research as carcinomas account for approximately 90% of cancers in humans (Cooper, 2018). Indeed, Green Tides could be viewed as cancerous growth of macroalgae, as it is an uncontrolled proliferative growth of tissue. Understanding this mechanism of proliferation could help elucidate not only the prevention and management of Green Tides, but other types of macroalgal blooms that rely on bottom-up proliferation, as well as cancerous tissues in metazoans. However, it is important to note that there is a key difference in this comparison: cancerous cells are unlike healthy cells, whereas *Ulva* cells are all thought to be healthy even during mass proliferation events. This comparison is merely referring to the uncontrolled growth of epithelial tissues.

This thesis has broadly accomplished all of the goals it initially set out. There have been many challenges to overcome, especially given the gaps in the literature. It is inevitable to now note quite the scale of the gap in the literature, as this thesis better characterises it. Conversely, this highlights the plentiful opportunities for future research, and I believe that *Ulva* research has a bright future as it will inform key aspects of future societal demands.

Bibliography

- Acuña, A. U. (2007). More Thoughts on the Narra Tree Fluorescence. *Journal of Chemical Education*, 84(2):231.
- Acuña, A. U. and Amat-Guerri, F. (2007). Early History of Solution Fluorescence: The Lignum nephriticum of Nicolás Monardes. In Berberan-Santos, M. N., editor, *Fluorescence of Supramolecules, Polymers, and Nanosystems*, Springer Series on Fluorescence, pages 3–20. Springer, Berlin, Heidelberg.
- Acuña, A. U., Amat-Guerri, F., Morcillo, P., Liras, M., and Rodríguez, B. (2009). Structure and Formation of the Fluorescent Compound of Lignum nephriticum. *Organic Letters*, 11(14):3020–3023.
- Adler, R., Feldman, R., and Taqqu, M. (1998). *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Springer Science & Business Media. Google-Books-ID: k4ekRZs_fZkC.
- Aegerter-Wilmsen, T., Smith, A. C., Christen, A. J., Aegerter, C. M., Hafen, E., and Basler, K. (2010). Exploring the effects of mechanical feedback on epithelial topology. *Development*, 137(3):499–506.
- Agbor, V. B., Cicek, N., Sparling, R., Berlin, A., and Levin, D. B. (2011). Biomass pretreatment: Fundamentals toward application. *Biotechnology Advances*, 29(6):675–685.
- Ale, M. T., Mikkelsen, J. D., and Meyer, A. S. (2011). Differential growth response of *Ulva lactuca* to ammonium and nitrate assimilation. *Journal of Applied Phycology*, 23:345–351.

- Almgren, F. J. and Taylor, J. E. (1976). The Geometry of Soap Films and Soap Bubbles. *Scientific American*, 235(1):82–93.
- Amos, W. B. and White, J. G. (2003). How the Confocal Laser Scanning Microscope entered Biological Research. *Biology of the Cell*, 95(6):335–342.
- Amosu, A., Robertson-Andersson, D., Kean, E., Maneveldt, G., and Cyster, L. (2015). *Biofiltering and Uptake of Dissolved Nutrients by Ulva armoricana (Chlorophyta) in a Land-based Aquaculture System*, volume 18.
- Anderson, C. T., Carroll, A., Akhmetova, L., and Somerville, C. (2010). Real-Time Imaging of Cellulose Reorientation during Cell Wall Expansion in Arabidopsis Roots. *Plant Physiology*, 152(2):787–796.
- Ansell, A., Barnes, M., Gibson, R. N., and Gibson, R. N. (1998). *Oceanography And Marine Biology: An Annual Review*. CRC Press.
- Anthoff, D., Nicholls, R. J., Tol, R. S. J., and Vafeidis, A. T. (2006). Global and regional exposure to large rises in sea-level: a sensitivity analysis. This work was prepared for the Stern Review on the Economics of Climate Change 96, Tyndall Centre for Climate Change Research, Norwich.
- Arasamuthu, A. and Edward, J. K. P. (2018). Occurrence of Ice-ice disease in seaweed *Kappaphycus alvarezii* at Gulf of Mannar and Palk Bay, Southeastern India. *Indian Journal of Geo Marine Sciences*, 47(06):1208–1216.
- Aro, E.-M. (2016). From first generation biofuels to advanced solar biofuels. *Ambio*, 45(1):24–31.
- Arora, M. and Sahoo, D. (2015). Growth Forms and Life Histories in Green Algae. In Sahoo, D. and Seckbach, J., editors, *The Algae World, Cellular Origin, Life in Extreme Habitats and Astrobiology*, pages 121–175. Springer Netherlands, Dordrecht.
- Arthur, D. and Vassilvitskii, S. (2006). How Slow is the K-means Method? In *Proceedings of the Twenty-second Annual Symposium on Computational Geometry, SCG '06*, pages 144–153, New York, NY, USA. ACM.
- Asano, T. (2006). Wave attenuation and sediment deposition due to coastal vegetation. *Journal of Global Environment Engineering*, 11:29–44.

- Atela, P. (2011). The Geometric and Dynamic Essence of Phyllotaxis. *Mathematical Modelling of Natural Phenomena*, 6(2):173–186.
- Balls, P. W. (1994). Nutrient Inputs to Estuaries from Nine Scottish East Coast Rivers; Influence of Estuarine Processes on Inputs to the North Sea. *Estuarine, Coastal and Shelf Science*, 39(4):329–352.
- Banerjee, P., Lenz, D., Robinson, J. P., Rickus, J. L., and Bhunia, A. K. (2008). A novel and simple cell-based detection system with a collagen-encapsulated B-lymphocyte cell line as a biosensor for rapid detection of pathogens and toxins. *Laboratory Investigation*, 88(2):196–206.
- Barrio, R. A., Romero-Arias, J. R., Noguez, M. A., Azpeitia, E., Ortiz-Gutiérrez, E., Hernández-Hernández, V., Cortes-Poza, Y., and Álvarez Buylla, E. R. (2013). Cell Patterns Emerge from Coupled Chemical and Physical Fields with Cell Proliferation Dynamics: The Arabidopsis thaliana Root as a Study System. *PLOS Computational Biology*, 9(5):e1003026.
- Becker, B. and Marin, B. (2009). Streptophyte algae and the origin of embryophytes. *Annals of Botany*, 103(7):999–1004.
- Bell, G. (1997). The evolution of the life cycle of brown seaweeds. *Biological Journal of the Linnean Society*, 60(1):21–38.
- Bell, J. and Dee, H. M. (2016). Watching plants grow – a position paper on computer vision and Arabidopsis thaliana. *IET Computer Vision*, 11(2):113–121.
- Bennett, S. and Wernberg, T. (2014). Canopy facilitates seaweed recruitment on subtidal temperate reefs. *Journal of Ecology*, 102(6):1462–1470.
- Besson, S. and Dumais, J. (2014). Stochasticity in the symmetric division of plant cells: when the exceptions are the rule. *Frontiers in Plant Science*, 5.
- Binnig, G., Rohrer, H., Gerber, C., and Weibel, E. (1983). 7 x 7 Reconstruction on Si(111) Resolved in Real Space. *Physical Review Letters*, 50(2):120–123.
- Bjerregaard, R., Valderrama, D., Radulovich, R., Diana, J., Capron, M., Mckinnie, C. A., Cedric, M., Hopkins, K., Yarish, C., Goudey, C., and Forster, J. (2016). Seaweed aquacul-

- ture for food security, income generation and environmental health in Tropical Developing Countries. Working Paper 107147, World Bank Group, Washington, D.C.
- Blackwell, W. H. (2009). CHROMISTA REVISITED: A DILEMMA OF OVERLAPPING PUTATIVE KINGDOMS, AND THE ATTEMPTED APPLICATION OF THE BOTANICAL CODE OF NOMENCLATURE. *Phytologia*, 91(2):191–225.
- Bock, M., Tyagi, A. K., Kreft, J.-U., and Alt, W. (2010). Generalized Voronoi Tessellation as a Model of Two-dimensional Cell Tissue Dynamics. *Bulletin of Mathematical Biology*, 72(7):1696–1731.
- Bolton, J. J. (1994). Global seaweed diversity: Patterns and anomalies.
- Bolton, J. J., Robertson-Andersson, D. V., Shuuluka, D., and Kandjengo, L. (2009). Growing Ulva (Chlorophyta) in integrated systems as a commercial crop for abalone feed in South Africa: a SWOT analysis. *Journal of Applied Phycology*, 21(5):575–583.
- Bonner, J. T. (2009). *First Signals: The Evolution of Multicellular Development*. Princeton University Press.
- Borgesá, A. V. and Gypensb, N. (2010). Carbonate chemistry in the coastal zone responds more strongly to eutrophication than ocean acidification. *Limnology and Oceanography*, 55(1):346–353.
- Bowers, M. C., Tung, W. W., and Gao, J. B. (2012). On the distributions of seasonal river flows: Lognormal or power law? *Water Resources Research*, 48(5).
- Bowes, G., Rao, S. K., Estavillo, G. M., and Reiskind, J. B. (2002). C4 mechanisms in aquatic angiosperms: comparisons with terrestrial C4 systems. *Functional Plant Biology*, 29(3):379–392.
- Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356):791–799.
- Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. In Launer, R. L. and Wilkinson, G. N., editors, *Robustness in Statistics*, pages 201–236. Academic Press.
- Box, G. E. P., Hunter, J. S., and Hunter, W. G. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery, Second Edition + JMP Version 6 Software Student*

- Edition, Set.* John Wiley & Sons Inc, Hoboken, New Jersey, 2 edition. Google-Books-ID: OIudQwAACAAJ.
- Bozzola, J. J. and Russell, L. D. (1999). *Electron Microscopy: Principles and Techniques for Biologists*. Jones & Bartlett Learning. Google-Books-ID: zMkBAPACbEkC.
- Brakenhoff, G. J., Blom, P., and Barends, P. (1979). Confocal scanning light microscopy with high aperture immersion lenses. *Journal of Microscopy*, 117(2):219–232.
- Brookes, N. H. (2018). Riding the cell jamming boundary: Geometry, topology, and phase of human corneal endothelium. *Experimental Eye Research*, 172:171–180.
- Brown, J. H., Lomolino, M. V., and Brown, T. E. (1998). *Biogeography*. Oxford University Press, Incorporated. Google-Books-ID: Baa2QgAACAAJ.
- Brown, J. W. and Sorhannus, U. (2010). A Molecular Genetic Timescale for the Diversification of Autotrophic Stramenopiles (Ochrophyta): Substantive Underestimation of Putative Fossil Ages. *PLOS ONE*, 5(9):e12759.
- Bryson, M. C. (1974). Heavy-Tailed Distributions: Properties and Tests. *Technometrics*, 16(1):61–68.
- Burke, C., Kjelleberg, S., and Thomas, T. (2009). Selective Extraction of Bacterial DNA from the Surfaces of Macroalgae. *Applied and Environmental Microbiology*, 75(1):252–256.
- Butterwick, C., Heaney, S., and Talling, J. (1982). A comparison of eight methods for estimating the biomass and growth of planktonic algae. *British Phycological Journal*, 17(1):69–79.
- Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014). A brief history of synthetic biology. *Nature Reviews Microbiology*, 12(5):381–390.
- Cannizzaro, F., Greco, G., Rizzo, S., and Sinagra, E. (1978). Results of the measurements carried out in order to verify the validity of the poisson-exponential distribution in radioactive decay events. *The International Journal of Applied Radiation and Isotopes*, 29(11):649–IN1.
- Canon (2019). CCD and CMOS sensors.

- Carl, C., Nys, R. d., Lawton, R. J., and Paul, N. A. (2014). Methods for the Induction of Reproduction in a Tropical Species of Filamentous Ulva. *PLOS ONE*, 9(5):e97396.
- Castaldelli, G., Welsh, D. T., Flachi, G., Zucchini, G., Colombo, G., Rossi, R., and Fano, E. A. (2003). Decomposition dynamics of the bloom forming macroalga *Ulva rigida* C. Agardh determined using a ¹⁴C-carbon radio-tracer technique. *Aquatic Botany*, 75(2):111–122.
- Cavalier-Smith, T. (1986). The kingdom Chromista: origin and systematics. In Round, F. E. and Chapman, D. J., editors, *Progress in Phycological research*. Biopress, Bristol.
- Charpy-Roubaud, C. and Sournia, A. (1990). The comparative estimation of phytoplanktonic, microphytobenthic and macrophytobenthic primary production in the oceans. *Marine Microbial Food Webs*, 4(1):31–57.
- Chaudhury, A., Ward, C., Talasaz, A., Ivanov, A. G., Huner, N. P. A., Grodzinski, B., Patel, R. V., and Barron, J. L. (2015). Computer Vision Based Autonomous Robotic System for 3d Plant Growth Measurement. In *2015 12th Conference on Computer and Robot Vision*, pages 290–296.
- Chehrehasa, F., Meedeniya, A. C. B., Dwyer, P., Abrahamsen, G., and Mackay-Sim, A. (2009). EdU, a new thymidine analogue for labelling proliferating cells in the nervous system. *Journal of Neuroscience Methods*, 177(1):122–130.
- Chiu, S. N. (1995). Aboav-Weaire’s and Lewis’ laws—A review. *Materials Characterization*, 34(2):149–165.
- Church, J. A. and White, N. J. (2006). A 20th century acceleration in global sea-level rise. *Geophysical research letters*, 33.
- Claessen, D., Rozen, D. E., Kuipers, O. P., Sogaard-Andersen, L., and Wezel, G. P. v. (2014). Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nature Reviews Microbiology*, 12(2):115–124.
- Cocquyt, E., Verbruggen, H., Leliaert, F., and De Clerck, O. (2010). Evolution and Cytological Diversification of the Green Seaweeds (Ulvophyceae). *Molecular Biology and Evolution*, 27(9):2052–2061.

- Cook, B. I., Ault, T. R., and Smerdon, J. E. (2015). Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Science Advances*, 1:e1400082.
- Cooper, G. M. (2018). *The Cell: A Molecular Approach*. Sinauer Associates, an imprint of Oxford University Press, Oxford; New York, eighth edition. OCLC: 1039425990.
- Cui, J., Shi, J., Zhang, J., Wang, L., Fan, S., Xu, Z., Huo, Y., Zhou, Q., Lu, Y., and He, P. (2018). Rapid expansion of *Ulva* blooms in the Yellow Sea, China through sexual reproduction and vegetative growth. *Marine Pollution Bulletin*, 130:223–228.
- Cui, J., Zhang, J., Huo, Y., Zhou, L., Wu, Q., Chen, L., Yu, K., and He, P. (2015). Adaptability of free-floating green tide algae in the Yellow Sea to variable temperature and light intensity. *Marine Pollution Bulletin*, 101(2):660–666.
- Dan, A., Hiraoka, M., Ohno, M., and Critchley, A. T. (2002). Observations on the effect of salinity and photon fluence rate on the induction of sporulation and rhizoid formation in the green alga *Enteromorpha prolifera* (Müller) J. Agardh (Chlorophyta, Ulvales). *Fisheries science*, 68(6):1182–1188.
- DARA and Climate Vulnerable Forum, C. V. (2012). *Climate Vulnerability Monitor : A Guide to the Cold Calculus of a Hot Planet*. DARA and Climate Vulnerable Forum, Madrid, second edition edition.
- De Clerck, O., Kao, S.-M., Bogaert, K. A., Blomme, J., Foflonker, F., Kwantes, M., Vancaester, E., Vanderstraeten, L., Aydogdu, E., Boesger, J., Califano, G., Charrier, B., Clewes, R., Cortona, A. D., D'Hondt, S., Fernandez-Pozo, N., Gachon, C. M., Hanikenne, M., Lattermann, L., Leliaert, F., Liu, X., Maggs, C. A., Popper, Z. A., Raven, J. A., Bel, M. V., Wilhelmsson, P. K. I., Bhattacharya, D., Coates, J. C., Rensing, S. A., Straeten, D. V. D., Vardi, A., Sterck, L., Vandepoele, K., Peer, Y. V. d., Wichard, T., and Bothwell, J. H. (2018). Insights into the Evolution of Multicellularity from the Sea Lettuce Genome. *Current Biology*, 28(18):2921–2933.e5.
- Derelle, R., López-García, P., Timpano, H., and Moreira, D. (2016). A Phylogenomic Framework to Study the Diversity and Evolution of Stramenopiles (=Heterokonts). *Molecular Biology and Evolution*, 33(11):2890–2898.

- Diaz-Pulido, G., Gouezo, M., Tilbrook, B., Dove, S., and Anthony, K. R. N. (2011). High CO₂ enhances the competitive strength of seaweeds over corals. *Ecology Letters*, 14(2):156–162.
- Diermeier-Daucher, S., Clarke, S. T., Hill, D., Vollmann-Zwerenz, A., Bradford, J. A., and Brockhoff, G. (2009). Cell type specific applicability of 5-ethynyl-2'-deoxyuridine (EdU) for dynamic proliferation assessment in flow cytometry. *Cytometry Part A*, 75A(6):535–546.
- Djakouré, S., Araujo, M., Hounsou-Gbo, A., Noriega, C., and Bourlès, B. (2017). On the potential causes of the recent Pelagic Sargassum blooms events in the tropical North Atlantic Ocean. *Biogeosciences Discussions*, pages 1–20.
- Doney, S. C., Fabry, V. J., Feely, R. A., and Kleypas, J. A. (2009). Ocean Acidification: The Other CO₂ Problem. *Annual Review of Marine Science*, 1(1):169–192.
- Doornbosch, R. and Steenblik, R. (2008). Biofuels: is the cure worse than the disease? *Revista Virtual REDESMA*, 2:63.
- Douady, S. and Couder, Y. (1992). Phyllotaxis as a physical self-organized growth process. *Physical Review Letters*, 68(13):2098–2101.
- Du, Q., Emelianenko, M., and Ju, L. (2006). Convergence of the Lloyd Algorithm for Computing Centroidal Voronoi Tessellations. *SIAM Journal on Numerical Analysis*, 44(1):102–119.
- Du, Q., Faber, V., and Gunzburger, M. (1999). Centroidal Voronoi Tessellations: Applications and Algorithms. *SIAM Review*, 41(4):637–676.
- Dutta, K., Daverey, A., and Lin, J.-G. (2014). Evolution retrospective for alternative fuels: First to fourth generation. *Renewable Energy*, 69:114–122.
- Dwyer, G. S. and Chandler, M. A. (2009). Mid-Pliocene sea level and continental ice volume based on coupled benthic Mg/Ca palaeotemperatures and oxygen isotopes. *Philos Trans A Math Phys Eng Sci*, 367:157–68.
- Edenhofer, O., Knopf, B., Barker, T., Baumstark, L., Bellevrat, E., Chateau, B., Criqui, P., Isaac, M., Kitous, A., and Kypreos, S. (2010). The economics of low stabilization: model comparison of mitigation strategies and costs. *The Energy Journal*, 31:11–48.

- Egerton, R. F. (2005). *Physical Principles of Electron Microscopy: An Introduction to TEM, SEM, and AEM*. Springer US.
- Einstein, A. (1934). On the Method of Theoretical Physics. *Philosophy of Science*, 1(2):163–169.
- EPA, U. (2012). Regulation of Fuels and Fuel Additives: 2012 Renewable Fuel Standards; Final Rule. *Federal Register*, 77:1320–1358.
- Eriksen, N. T. (2008). The technology of microalgal culturing. *Biotechnology Letters*, 30(9):1525–1536.
- Erisman, J. W., Bleeker, A., Galloway, J., and Sutton, M. S. (2007). Reduced nitrogen in ecology and the environment. *Environmental Pollution*, 150(1):140–149.
- Fabricius, K. and De'ath, G. (2001). Environmental factors associated with the spatial distribution of crustose coralline algae on the Great Barrier Reef. *Coral Reefs*, 19(4):303–309.
- Facette, M. R., Rasmussen, C. G., and Van Norman, J. M. (2019). A plane choice: coordinating timing and orientation of cell division during plant development. *Current Opinion in Plant Biology*, 47:47–55.
- Falkowski, P. G. and Raven, J. A. (2007). *Aquatic Photosynthesis: Second Edition*. Princeton University Press. Google-Books-ID: l2KYDwAAQBAJ.
- Fan, M., Sun, X., Xu, N., Liao, Z., Li, Y., Wang, J., Fan, Y., Cui, D., Li, P., and Miao, Z. (2017). Integration of deep transcriptome and proteome analyses of salicylic acid regulation high temperature stress in *Ulva prolifera*. *Scientific Reports*, 7(1):11052.
- FAO (2018). The State of World Fisheries and Aquaculture 2018 - Meeting the sustainable development goals. Technical report, Food and Agriculture Organization of the United Nations, Rome.
- Fargione, J., Hill, J., Tilman, D., Polasky, S., and Hawthorne, P. (2008). Land clearing and the biofuel carbon debt. *Science*, 319:1235–8.
- Featherston, J., Arakaki, Y., Hanschen, E. R., Ferris, P. J., Michod, R. E., Olson, B. J. S. C., Nozaki, H., and Durand, P. M. (2018). The 4-Celled *Tetrabaena socialis* Nuclear

- Genome Reveals the Essential Components for Genetic Control of Cell Number at the Origin of Multicellularity in the Volvocine Lineage. *Molecular Biology and Evolution*, 35(4):855–870.
- Ferreira, T. and Rasband, W. (2012). ImageJ User Guide IJ 1.46r.
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science*, 281(5374):237–240.
- Fletcher, R. L. (1996). The Occurrence of “Green Tides”— a Review. In Schramm, D. W. and Nienhuis, P. D. P. H., editors, *Marine Benthic Vegetation*, number 123 in Ecological Studies, pages 7–43. Springer Berlin Heidelberg.
- Flomerfelt, F. A. and Gress, R. E. (2016). Analysis of Cell Proliferation and Homeostasis Using EdU Labeling. In Bosselut, R. and S. Vacchio, M., editors, *T-Cell Development: Methods and Protocols*, Methods in Molecular Biology, pages 211–220. Springer New York, New York, NY.
- Fong, P., Boyer, K. E., Desmond, J. S., and Zedler, J. B. (1996). Salinity stress, nitrogen competition, and facilitation: what controls seasonal succession of two opportunistic green macroalgae? *Journal of Experimental Marine Biology and Ecology*, 206(1):203–221.
- Forbes (2019). Jeff Bezos & family.
- Foss, S., Korshunov, D., and Zachary, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2 edition.
- Fraginière, Y., Bétrisey, S., Cardinaux, L., Stoffel, M., and Kozłowski, G. (2015). Fighting their last stand? A global analysis of the distribution and conservation status of gymnosperms. *Journal of Biogeography*, 42(5):809–820.
- Frenkel, J., Vyverman, W., and Pohnert, G. (2014). Pheromone signaling during sexual reproduction in algae. *The Plant Journal*, 79(4):632–644.

- Fryar, C., Gu, Q., Ogden, C., and Flegal, K. (2016). Anthropometric reference data for children and adults: United States, 2011–2014. Technical Report 3(39), National Center for Health Statistics, Hyattsville, Maryland.
- FS-UNEP Collaborating Centre (2015). Global Trends in Renewable Energy Investment 2015. Technical report, Frankfurt School- UNEP Collaborating Centre for Climate and Sustainable Energy Finance.
- Förstner, W. and Wrobel, B. P. (2016). *Photogrammetric Computer Vision: Statistics, Geometry, Orientation and Reconstruction*. Springer. Google-Books-ID: rMM-wDQAAQBAJ.
- Gabetta, E. and Regazzini, E. (2010). About the gene families size distribution in a recent model of genome evolution. *Mathematical Models and Methods in Applied Sciences*, 20(06):1005–1020.
- Gaines, S. D. and Roughgarden, J. (1987). Fish in Offshore Kelp Forests Affect Recruitment to Intertidal Barnacle Populations. *Science*, 235(4787):479–481.
- Gao, G., Clare, A. S., Rose, C., and Caldwell, G. S. (2017). Intrinsic and extrinsic control of reproduction in the green tide-forming alga, *Ulva rigida*. *Environmental and Experimental Botany*, 139:14–22.
- Gao, S., Chen, X., Yi, Q., Wang, G., Pan, G., Lin, A., and Peng, G. (2010). A Strategy for the Proliferation of *Ulva prolifera*, Main Causative Species of Green Tides, with Formation of Sporangia by Fragmentation. *PLOS ONE*, 5(1):e8571.
- Garbary, D. J. (2001). Biogeography of Marine Algae. In *Encyclopedia of Life Sciences*. John Wiley & Sons, Inc.
- Garcia, J. M. G. (2011). A fixed-point algorithm to estimate the Yule–Simon distribution parameter. *Applied Mathematics and Computation*, 217(21):8560–8566.
- Gaukroger, S. (1998). *Descartes: The World and Other Writings*. Cambridge University Press. Google-Books-ID: YWpErQ3g_pEC.
- Gavara, N. (2017). A beginner’s guide to atomic force microscopy probing for cell mechanics. *Microscopy Research and Technique*, 80(1):75–84.

- Gerasimchuk, I., Bridle, R., Beaton, C., and Charles, C. (2012). State of Play on Biofuel Subsidies: Are policies ready to shift? Technical report, International Institute for Sustainable Development.
- Gerlach, J. and Gerlach, B. (2007). *Digital Nature Photography: The Art and the Science*. Taylor & Francis. Google-Books-ID: iNrWzL_NpWkC.
- Gibson, M., B Patel, A., Nagpal, R., and Perrimon, N. (2006). *The emergence of geometric order in proliferating metazoan epithelia*, volume 442.
- Gibson, W. T. and Gibson, M. C. (2009). Cell topology, geometry, and morphogenesis in proliferating epithelia. *Current Topics in Developmental Biology*, 89:87–114.
- Gibson, W. T., Veldhuis, J. H., Rubinstein, B., Cartwright, H. N., Perrimon, N., Brodland, G. W., Nagpal, R., and Gibson, M. C. (2011). Control of the Mitotic Cleavage Plane by Local Epithelial Topology. *Cell*, 144(3):427–438.
- Giessibl, F. J. (2003). Advances in atomic force microscopy. *Reviews of Modern Physics*, 75(3):949–983.
- Gildea, K. M., Hileman, C. R., Rogers, P., Salazar, G. J., and Paskoff, L. N. (2018). The Use of a Poisson Regression to Evaluate Antihistamines and Fatal Aircraft Mishaps in Instrument Meteorological Conditions. *Aerospace Medicine and Human Performance*, 89(4):389–395.
- Grew, N. (1682). *The anatomy of plants : with an idea of a philosophical history of plants : and several other lectures read before the Royal Society / by Nehemiah Grew ...* [London] : Printed by W. Rawlins, for the author,.
- Grosberg, R. K. and Strathmann, R. R. (1998). One cell, two cell, red cell, blue cell: The persistence of a unicellular stage in multicellular life histories. *Trends in Ecology & Evolution*, 13(3):112–116.
- Grosberg, R. K. and Strathmann, R. R. (2007). The Evolution of Multicellularity: A Minor Major Transition? *Annual Review of Ecology, Evolution, and Systematics*, 38(1):621–654.

- Grueneberg, J., Engelen, A. H., Costa, R., and Wichard, T. (2016). Macroalgal Morphogenesis Induced by Waterborne Compounds and Bacteria in Coastal Seawater. *PLoS ONE*, 11(1).
- Guinness World Records (2019). Tallest man ever.
- Guinotte, J. M. and Fabry, V. J. (2008). Ocean Acidification and Its Potential Effects on Marine Ecosystems. *Annals of the New York Academy of Sciences*, 1134(1):320–342.
- Guiry, M. D. (2012). How Many Species of Algae Are There? *Journal of Phycology*, 48(5):1057–1063.
- Guiry, M. D. and Guiry, G. M. (2019). AlgaeBase.
- Hacking, I. (1990). *The Taming of Chance*. Cambridge University Press, Cambridge, United Kingdom.
- Hales, T. C. (2001). The Honeycomb Conjecture. *Discrete & Computational Geometry*, 25(1):1–22.
- Hallegatte, S., Green, C., Nicholls, R. J., and Corfee-Morlot, J. (2013). Future flood losses in major coastal cities. *Nature climate change*, 3:802–806.
- Hanschen, E. R., Marriage, T. N., Ferris, P. J., Hamaji, T., Toyoda, A., Fujiyama, A., Neme, R., Noguchi, H., Minakuchi, Y., Suzuki, M., Kawai-Toyooka, H., Smith, D. R., Sparks, H., Anderson, J., Bakarić, R., Luria, V., Karger, A., Kirschner, M. W., Durand, P. M., Michod, R. E., Nozaki, H., and Olson, B. J. S. C. (2016). The *Gonium pectorale* genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. *Nature Communications*, 7:11370.
- Harley, C. D. G., Anderson, K. M., Demes, K. W., Jorve, J. P., Kordas, R. L., Coyle, T. A., and Graham, M. H. (2012). Effects of Climate Change on Global Seaweed Communities. *Journal of Phycology*, 48(5):1064–1078.
- Harrington, B. J. and Hageage, G. J. (2003). Calcofluor White: A Review of its Uses and Applications in Clinical Mycology and Parasitology. *Laboratory Medicine*, 34(5):361–367.
- Hart, P. E. (1990). Lognormal Distribution. In Eatwell, J., Milgate, M., and Newman, P., editors, *Econometrics*, The New Palgrave, pages 145–147. Palgrave Macmillan UK, London.

- Hayden, H. S., Blomster, J., Maggs, C. A., Silva, P. C., Stanhope, M. J., and Waaland, J. R. (2003). Linnaeus was right all along: *Ulva* and *Enteromorpha* are not distinct genera. *European Journal of Phycology*, 38(3):277–294.
- Helaouët, P. and Beaugrand, G. (2009). Physiology, Ecological Niches and Species Distribution. *Ecosystems*, 12(8):1235–1245.
- Herbert, R. A. (1999). Nitrogen cycling in coastal marine ecosystems. *FEMS Microbiology Reviews*, 23(5):563–590.
- Hernández, L. (1995). *Pattern formation in the sunflower (Helianthus annuus L.) capitulum*. *Biophysical considerations*, volume 19.
- Hiraoka, M. and Oka, N. (2008). Tank cultivation of *Ulva prolifera* in deep seawater using a new “germling cluster” method. *Journal of Applied Phycology*, 20(1):97–102.
- Hoch, H. C., Galvani, C. D., Szarowski, D. H., and Turner, J. N. (2005). Two new fluorescent dyes applicable for visualization of fungal cell walls. *Mycologia*, 97(3):580–588.
- Hoegh-Guldberg, O., Mumby, P. J., Hooten, A. J., Steneck, R. S., Greenfield, P., Gomez, E., Harvell, C. D., Sale, P. F., Edwards, A. J., Caldeira, K., Knowlton, N., Eakin, C. M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R. H., Dubi, A., and Hatziolos, M. E. (2007). Coral Reefs Under Rapid Climate Change and Ocean Acidification. *Science*, 318(5857):1737–1742.
- Hoeksema, B. W. and van den Hoek, C. (1983). The Taxonomy of *Ulva* (Chlorophyceae) from the Coastal Region of Roscoff (Brittany, France). *Botanica Marina*, 26(2):65–86.
- Hoffmann, K. H., Rodriguez-Brito, B., Breitbart, M., Bangor, D., Angly, F., Felts, B., Nulton, J., Rohwer, F., and Salamon, P. (2007). Power law rank–abundance models for marine phage communities. *FEMS Microbiology Letters*, 273(2):224–228.
- Hofmann, L. C. and Bischof, K. (2014). Ocean acidification effects on calcifying macroalgae. *Aquatic Biology*, 22:261–279.
- Hofmann, M., Anderssohn, R., Bahr, H.-A., Weiß, H.-J., and Nellesen, J. (2015). Why Hexagonal Basalt Columns? *Physical Review Letters*, 115(15):154301.
- Holst, G. C. and Lomheim, T. S. (2011). *CMOS/CCD Sensors and Camera Systems*. JCD Publishing. Google-Books-ID: ZJxStwAACA AJ.

- Honda, H. (1978). Description of cellular patterns by Dirichlet domains: The two-dimensional case. *Journal of Theoretical Biology*, 72(3):523–543.
- Hooke, R. (1667). *Micrographia: Or, Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses. With Observations and Inquiries Thereupon*. J. Allestry, printer to the Royal Society. Google-Books-ID: qNBTLOTyhT8C.
- Hopcroft, M. A., Nix, W. D., and Kenny, T. W. (2010). What is the Young’s Modulus of Silicon? *Journal of Microelectromechanical Systems*, 19(2):229–238.
- Hori, H., Lim, B.-L., and Osawa, S. (1985). Evolution of green plants as deduced from 5s rRNA sequences. *Proceedings of the National Academy of Sciences*, 82(3):820–823.
- Hotton, S., Johnson, V., Wilbarger, J., Zwieniecki, K., Atela, P., Golé, C., and Dumais, J. (2006). The Possible and the Actual in Phyllotaxis: Bridging the Gap between Empirical Observations and Iterative Models. *Journal of Plant Growth Regulation*, 25(4):313–323.
- Hsieh, L.-C., Luo, L., Ji, F., and Lee, H. C. (2003). Minimal Model for Genome Evolution and Growth. *Physical Review Letters*, 90(1):018101.
- Hu, C. and He, M.-X. (2008). Origin and Offshore Extent of Floating Algae in Olympic Sailing Area. *Eos, Transactions American Geophysical Union*, 89(33):302–303.
- Hu, C., Li, D., Chen, C., Ge, J., Muller-Karger, F. E., Liu, J., Yu, F., and He, M.-X. (2010). On the recurrent *Ulva prolifera* blooms in the Yellow Sea and East China Sea. *Journal of Geophysical Research: Oceans*, 115(C5):C05017.
- Hunt, C. (2019). Way too much of a good thing: Sargassum and the sea.
- Inoué, S. (2010). Foundations of Confocal Scanned Imaging in Light Microscopy. In Pawley, J., editor, *Handbook of Biological Confocal Microscopy*. Springer Science & Business Media. Google-Books-ID: IKcPnaNPrhoC.
- IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Cambridge University Press, Cambridge.
- Jacobs, W. P. (1994). Caulerpa. *Scientific American*, 271(6):100–105.

- Jain, R. and Ramakumar, S. (1999). Stochastic dynamics modeling of the protein sequence length distribution in genomes: implications for microbial evolution. *Physica A: Statistical Mechanics and its Applications*, 273(3):476–485.
- Jena, B. P., Hörber, J. K. H., Wilson, L., and Matsudaira, P., editors (2002). *Atomic Force Microscopy in Cell Biology, Volume 68 - 1st Edition*, volume 68 of *Methods in Cell Biology*. Academic Press, San Diego, California.
- Jevrejeva, S., Grinsted, A., and Moore, J. C. (2014). Upper limit for sea level projections by 2100. *Environmental Research Letters*, 9:104008.
- Jing, L. (2013a). Algae bloom covers beach in Qingdao. *South China Morning Post*.
- Jing, L. (2013b). Seaweed farming linked to Qingdao’s green tide of algae. *South China Morning Post*.
- Joint, I., Tait, K., Callow, M. E., Callow, J. A., Milton, D., Williams, P., and Cámara, M. (2002). Cell-to-Cell Communication Across the Prokaryote-Eukaryote Boundary. *Science*, 298(5596):1207–1207.
- Joint, I., Tait, K., and Wheeler, G. (2007). Cross-kingdom signalling: exploitation of bacterial quorum sensing molecules by the green seaweed *Ulva*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1483):1223–1233.
- Jung, K. A., Lim, S.-R., Kim, Y., and Park, J. M. (2013). Potentials of macroalgae as feedstocks for biorefinery. *Bioresour Technol*, 135:182–90.
- Kaiser, D. (2001). Building a Multicellular Organism. *Annual Review of Genetics*, 35(1):103–123.
- Kalaji, H. M., Schansker, G., Brestic, M., Bussotti, F., Calatayud, A., Ferroni, L., Goltsev, V., Guidi, L., Jajoo, A., Li, P., Losciale, P., Mishra, V. K., Misra, A. N., Nebauer, S. G., Pancaldi, S., Penella, C., Pollastrini, M., Suresh, K., Tambussi, E., Yanniccari, M., Zivcak, M., Cetner, M. D., Samborska, I. A., Stirbet, A., Olsovska, K., Kunderlikova, K., Shelonzek, H., Rusinowski, S., and Bąba, W. (2017). Frequently asked questions about chlorophyll fluorescence, the sequel. *Photosynthesis Research*, 132(1):13–66.
- Kalaji, H. M., Schansker, G., Ladle, R. J., Goltsev, V., Bosa, K., Allakhverdiev, S. I., Brestic, M., Bussotti, F., Calatayud, A., Dąbrowski, P., Elsheery, N. I., Ferroni, L.,

- Guidi, L., Hogewoning, S. W., Jajoo, A., Misra, A. N., Nebauer, S. G., Pancaldi, S., Penella, C., Poli, D., Pollastrini, M., Romanowska-Duda, Z. B., Rutkowska, B., Serôdio, J., Suresh, K., Szulc, W., Tambussi, E., Yannicari, M., and Zivcak, M. (2014). Frequently asked questions about in vivo chlorophyll fluorescence: practical issues. *Photosynthesis Research*, 122(2):121–158.
- Kansal, A. R., Torquato, S., Harsh, G. R., Chiocca, E. A., and Deisboeck, T. S. (2000). Simulated Brain Tumor Growth Dynamics Using a Three-Dimensional Cellular Automaton. *Journal of Theoretical Biology*, 203(4):367–382.
- Karev, G. P., Wolf, Y. I., Rzhetsky, A. Y., Berezovskaya, F. S., and Koonin, E. V. (2002). Birth and death of protein domains: A simple model of evolution explains power law behavior. *BMC Evolutionary Biology*, 2(1):18.
- Karihaloo, B. L., Zhang, K., and Wang, J. (2013). Honeybee combs: how the circular cells transform into rounded hexagons. *Journal of The Royal Society Interface*, 10(86):20130299.
- Karsten, U., Kirst, G. O., and Wiencke, C. (1992). Dimethylsulphoniopropionate (DMSP) accumulation in green macroalgae from polar to temperate regions: interactive effects of light versus salinity and light versus temperature. *Polar Biology*, 12(6-7):603–607.
- Keats, D. (2010). Eucheuma farming, Philippines.
- Keith, S. A., Kerswell, A. P., and Connolly, S. R. (2014). Global diversity of marine macroalgae: environmental conditions explain less variation in the tropics. *Global Ecology and Biogeography*, 23(5):517–529.
- Kerswell, A. P. (2006). Global Biodiversity Patterns of Benthic Marine Algae. *Ecology*, 87(10):2479–2488.
- Kessler, R. W., Weiss, A., Kuegler, S., Hermes, C., and Wichard, T. (2017). Macroalgal–bacterial interactions: Role of dimethylsulfonylpropionate in microbial gardening by Ulva (Chlorophyta). *Molecular Ecology*, 27(8):1808–1819.
- Kim, S.-K. (2011). *Handbook of Marine Macroalgae: Biotechnology and Applied Phycology*. John Wiley & Sons, Hoboken, New Jersey. Google-Books-ID: zM8NoT1Gi30C.

- King, N. (2004). The Unicellular Ancestry of Animal Development. *Developmental Cell*, 7(3):313–325.
- Kirkendale, L., Saunders, G. W., and Winberg, P. (2013). A Molecular Survey of Ulva (Chlorophyta) in Temperate Australia Reveals Enhanced Levels of Cosmopolitanism. *Journal of Phycology*, 49(1):69–81.
- Klass, O. S., Biham, O., Levy, M., Malcai, O., and Solomon, S. (2006). The Forbes 400 and the Pareto wealth distribution. *Economics Letters*, 90(2):290–295.
- Knoll, A. H. (2011). The Multiple Origins of Complex Multicellularity. *Annual Review of Earth and Planetary Sciences*, 39(1):217–239. _eprint: <https://doi.org/10.1146/annurev.earth.031208.100209>.
- Kobayashi, N., Raichle, A. W., and Toshiyuki, A. (1993). Wave Attenuation by Vegetation. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 119(1).
- Koch, M., Bowes, G., Ross, C., and Zhang, X.-H. (2013). Climate change and ocean acidification effects on seagrasses and marine macroalgae. *Global Change Biology*, 19(1):103–132.
- Koeman, R. P. T. and van den Hoek, C. (1981). The taxonomy of Ulva (Chlorophyceae) in the Netherlands. *British Phycological Journal*, 16(1):9–53.
- Kolbert, E. (2014). *The sixth extinction : an unnatural history*.
- Kookje News (2012). Undaria farm in ulsan: Kookje news.
- Koonin, E. V., Wolf, Y. I., and Karev, G. P. (2002). The structure of the protein universe and genome evolution. *Nature*, 420(6912):218–223.
- Koyama, K., Hokunan, H., Hasegawa, M., Kawamura, S., and Koseki, S. (2016). Do bacterial cell numbers follow a theoretical Poisson distribution? Comparison of experimentally obtained numbers of single cells with random number generation via computer simulation. *Food Microbiology*, 60:49–53.
- Kraufvelin, P. (2007). Responses to nutrient enrichment, wave action and disturbance in rocky shore communities. *Aquatic Botany*, 87(4):262–274. WOS:000250667700002.

- Krause-Jensen, D. and Duarte, C. M. (2016). Substantial role of macroalgae in marine carbon sequestration. *Nature Geoscience*, 9(10):737–742.
- Kumar, C., Norris, J. B., and Sun, Y. (2009). Location and time do matter: A long tail study of website requests. *Decision Support Systems*, 47(4):500–507.
- Kusuda, R., Kawai, K., Salati, F., Kawamura, Y., and Yamashita, Y. (1992). Characteristics of *Flavobacterium* sp. Causing “Suminori” Disease in Cultivated *Porphyra*. *Aquaculture Science*, 40(4):457–461.
- Lakowicz, J. R. (2007). *Principles of Fluorescence Spectroscopy*. Springer Science & Business Media.
- Lander, E. and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics*, 11(3):241–247.
- Lapointe, B. E. (1999). Simultaneous top-down and bottom-up forces control macroalgal blooms on coral reefs (Reply to the comment by Hughes et al.). *Limnology and Oceanography*, 44(6):1586–1592.
- Largo, D. B. (2002). Recent developments in seaweed diseases. In *Proceedings of the National Seaweed Planning Workshop held on August 2-3, 2001, SEAFDEC Aquaculture Department, Tigbauan, Iloilo.*, pages 35–42, Tigbauan, Iloilo. Aquaculture Department, Southeast Asian Fisheries Development Center.
- Largo, D. B., Fukami, K., Nishijima, T., and Ohno, M. (1995). Laboratory-induced development of the ice-ice disease of the farmed red algae *Kappaphycus alvarezii* and *Euचेuma denticulatum* (Solieriaceae, Gigartinales, Rhodophyta). *Journal of Applied Phycology*, 7(6):539–543.
- Laurance, W. F. (2010). Habitat destruction: death by a thousand cuts. *Conservation biology for all*, 1:73–88.
- Lawrence, G., Lyons, K., and Wallington, T. (2010). *Food security, nutrition and sustainability*. Earthscan, London; Sterling, VA.
- Lee, M. M. and Schiefelbein, J. (2002). Cell Pattern in the *Arabidopsis* Root Epidermis Determined by Lateral Inhibition with Feedback. *The Plant Cell*, 14(3):611–618.

- Lee, R. A. and Lavoie, J.-M. (2013). From first- to third-generation biofuels: Challenges of producing a commodity from a biomass of increasing complexity. *Animal Frontiers*, 3(2):6–11.
- Lee, R. E. (2008). *Phycology*. Cambridge University Press, Cambridge, 4 edition.
- Lee, R. E. (2018). *Phycology*. Cambridge University Press, Cambridge, 5 edition. Google-Books-ID: rHRJDwAAQBAJ.
- Leliaert, F., Verbruggen, H., and Zechman, F. W. (2011). Into the deep: New discoveries at the base of the green plant phylogeny. *BioEssays*, 33(9):683–692.
- Letunic, I. and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, 47(W1):W256–W259.
- Lewis, F. T. (1923). The Typical Shape of Polyhedral Cells in Vegetable Parenchyma and the Restoration of That Shape following Cell Division. *Proceedings of the American Academy of Arts and Sciences*, 58(15):537–554.
- Lewis, F. T. (1926). The effect of cell division on the shape and size of hexagonal cells. *The Anatomical Record*, 33(5):331–355.
- Lewis, F. T. (1928). The correlation between cell division and the shapes and sizes of prismatic cells in the epidermis of cucumis. *The Anatomical Record*, 38(3):341–376.
- Lewis, F. T. (1931). A comparison between the mosaic of polygons in a film of artificial emulsion and the pattern of simple epithelium in surface view (cucumber epidermis and human amnion). *The Anatomical Record*, 50(3):235–265.
- Lewis, L. A. and McCourt, R. M. (2004). Green algae and the origin of land plants. *American Journal of Botany*, 91(10):1535–1556.
- Li, Y., Horsman, M., Wu, N., Lan, C. Q., and Dubois-Calero, N. (2008). Biofuels from Microalgae. *Biotechnology Progress*, 24(4):815–820.
- Lichtman, J. W. and Conchello, J.-A. (2005). Fluorescence microscopy. *Nature Methods*, 2(12):910.
- Liesche, J., Ziomkiewicz, I., and Schulz, A. (2013). Super-resolution imaging with Pontamine Fast Scarlet 4bs enables direct visualization of cellulose orientation and cell connection architecture in onion epidermis cells. *BMC Plant Biology*, 13(1):226.

- Limpert, E., Stahel, W. A., and Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question. *BioScience*, 51(5):341–352.
- Linnaeus, C. (1753). *Species plantarum, exhibentes plantas rite cognitatas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas.*, volume 2. Impensis Laurentii Salvii, Holmiae [Stockholm], Stockholm.
- Liu, D., Keesing, J. K., Dong, Z., Zhen, Y., Di, B., Shi, Y., Fearn, P., and Shi, P. (2010). Recurrence of the world’s largest green-tide in 2009 in Yellow Sea, China: *Porphyra yezoensis* aquaculture rafts confirmed as nursery for macroalgal blooms. *Marine Pollution Bulletin*, 60(9):1423–1432.
- Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Trans. Information Theory*, 28:129–136.
- Lou, Z., Li, P., Song, P., and Han, K. (2013). Ratiometric fluorescence imaging of cellular hypochlorous acid based on heptamethine cyanine dyes. *Analyst*, 138(21):6291–6295.
- Lu, Y. and Zhou, H. H. (2016). Statistical and Computational Guarantees of Lloyd’s Algorithm and its Variants. *arXiv:1612.02099 [cs, math, stat]*. arXiv: 1612.02099.
- Ma, Z., Bielenberg, D. G., Brown, K. M., and Lynch, J. P. (2001). Regulation of root hair density by phosphorus availability in *Arabidopsis thaliana*. *Plant, Cell & Environment*, 24(4):459–467.
- Malta, E.-J., Draisma, S., and Kamermans, P. (1999). Free-floating *Ulva* in the southwest Netherlands: species or morphotypes? a morphological, molecular and ecological comparison. *European Journal of Phycology*, 34(5):443–454.
- Maltz, M. D. (1996). From Poisson to the present: Applying operations research to problems of crime and justice. *Journal of Quantitative Criminology*, 12(1):3–61.
- Mandelbrot, B. B. (1997). *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk. Selecta Volume E*. Springer-Verlag, New York.

- Mantri, V. A., Singh, R. P., Bijo, A. J., Kumari, P., Reddy, C. R. K., and Jha, B. (2011). Differential response of varying salinity and temperature on zoospore induction, regeneration and daily growth rate in *Ulva fasciata* (Chlorophyta, Ulvales). *Journal of applied phycology*, 23:243–250.
- Marcu, L., French, P. M. W., and Elson, D. S. (2014). *Fluorescence Lifetime Spectroscopy and Imaging: Principles and Applications in Biomedical Diagnostics*. CRC Press. Google-Books-ID: K8uSAwAAQBAJ.
- Marin, B. (2012). Nested in the Chlorellales or Independent Class? Phylogeny and Classification of the Pedinophyceae (Viridiplantae) Revealed by Molecular Phylogenetic Analyses of Complete Nuclear and Plastid-encoded rRNA Operons. *Protist*, 163(5):778–805.
- Martini, R., Balit, Y., and Barthelat, F. (2017). A comparative study of bio-inspired protective scales using 3d printing and mechanical testing. *Acta Biomaterialia*, 55:360–372.
- Maruvka, Y. E., Kessler, D. A., and Shnerb, N. M. (2011). The Birth-Death-Mutation Process: A New Paradigm for Fat Tailed Distributions. *PLOS ONE*, 6(11):e26480.
- Mata, T. M., Martins, A. A., and Caetano, N. S. (2010). Microalgae for biodiesel production and other applications: A review. *Renewable and Sustainable Energy Reviews*, 14(1):217–232.
- Matsuo, Y., Imagawa, H., Nishizawa, M., and Shizuri, Y. (2005). Isolation of an Algal Morphogenesis Inducer from a Marine Bacterium. *Science*, 307(5715):1598–1598.
- Mazur, J. (2014). Yuck! Seaside, Monterey beaches covered in bright green algae.
- McCallum, M. L. (2015). Vertebrate biodiversity losses point to a sixth mass extinction. *Biodiversity and Conservation*, pages 1–23.
- McGranahan, G., Balk, D., and Anderson, B. (2007). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and urbanization*, 19:17–37.
- McHugh, D. J. (2002). A guide to the seaweed industry. Technical Report 441, FAO, Rome.

- McKinnon, K. and Hickey, V. (2009). Convenient Solutions to an Inconvenient Truth: Ecosystem-based approaches to climate change. *International Bank for Reconstruction and Development/The World Bank*, 2.
- McMillan, J. D. (1994). Pretreatment of Lignocellulosic Biomass. In *Enzymatic Conversion of Biomass for Fuels Production*, volume 566 of *ACS Symposium Series*, pages 292–324. American Chemical Society.
- Mead, T. J. and Lefebvre, V. (2014). Proliferation Assays (BrdU and EdU) on Skeletal Tissue Sections. In Hilton, M. J., editor, *Skeletal Development and Repair: Methods and Protocols*, Methods in Molecular Biology, pages 233–243. Humana Press, Totowa, NJ.
- Meade, M. S. and Earickson, R. (2005). *Medical Geography*. Guilford. Google-Books-ID: PCBWr5UySs8C.
- Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Maréchal-Drouard, L., Marshall, W. F., Qu, L.-H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.-L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J.-P., Riaño-Pachón, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C.-J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W. C. A., Otilar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007).

- The Chlamydomonas Genome Reveals the Evolution of Key Animal and Plant Functions. *Science*, 318(5848):245–250.
- Messyas, B. and Rybak, A. (2011). Abiotic factors affecting the development of *Ulva* sp. (Ulvophyceae; Chlorophyta) in freshwater ecosystems. *Aquatic Ecology*, 45(1):75–87.
- Michod, R. E. and Roze, D. (2001). Cooperation and conflict in the evolution of multicellularity. *Heredity*, 86(1):1–7.
- Middelburg, J. J., Duarte, C. M., and Gattuso, J.-P. (2005). Respiration in coastal benthic communities. In Giorgio, P. A. d. and Williams, P. J. l. B., editors, *Respiration in Aquatic Ecosystems*. Oxford University Press, USA. Google-Books-ID: pD5RUDW1m7IC.
- Mitzenmacher, M. (2004). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*, 1(2):226–251.
- Moheimani, N. R., Borowitzka, M. A., Isdepsky, A., and Sing, S. F. (2013). Standard Methods for Measuring Growth of Algae and Their Composition. In Borowitzka, M. A. and Moheimani, N. R., editors, *Algae for Biofuels and Energy*, Developments in Applied Phycology, pages 265–284. Springer Netherlands, Dordrecht.
- Molinos, J., Poloczanska, E., Olden, J., Lawler, J., and Burrows, M. (2017). Biogeographical Shifts and Climate Change. In *Reference Module in Earth Systems and Environmental Sciences*.
- Monshausen, G. B., Miller, N. D., Murphy, A. S., and Gilroy, S. (2011). Dynamics of auxin-dependent Ca²⁺ and pH signaling in root growth revealed by integrating high-resolution imaging with automated computer vision-based analysis. *The Plant Journal*, 65(2):309–318.
- Moore, C. R., Johnson, L. S., Kwak, I.-Y., Livny, M., Broman, K. W., and Spalding, E. P. (2013). High-Throughput Computer Vision Introduces the Time Axis to a Quantitative Trait Map of a Plant Growth Response. *Genetics*, 195(3):1077–1086.
- Morris, D. J., Pinnegar, J. K., Maxwell, D. L., Dye, S. R., Fernand, L. J., Flatman, S., Williams, O. J., and Rogers, S. I. (2018). Over 10 million seawater temperature records for the United Kingdom Continental Shelf between 1880 and 2014 from 17 Cefas (United Kingdom government) marine data systems. *Earth System Science Data*, 10(1):27–51.

- Morrissey, Sumich, J. L., and Pinkard-Meier, D. R. (2016). *Introduction to the Biology of Marine Life*. Jones & Bartlett Learning. Google-Books-ID: 6V6HDQAAQBAJ.
- Moynihan, T. (2011). CMOS Is Winning the Camera Sensor Battle, and Here's Why | TechHive.
- Mshigeni, K. E. and Kajumulo, A. A. (1979). Effects of the Environment on Polymorphism in *Ulva fasciata* Delile (Chlorophyta, Ulvaceae). *Botanica Marina*, 22(3).
- Muller, M. (2006). *Introduction to Confocal Fluorescence Microscopy*. SPIE Press. Google-Books-ID: t7N4G4x9zGcC.
- Munsky, B., Neuert, G., and Oudenaarden, A. v. (2012). Using Gene Expression Noise to Understand Gene Regulation. *Science*, 336(6078):183–187.
- Nagpal, R., Patel, A., and Gibson, M. C. (2008). Epithelial topology. *BioEssays*, 30(3):260–266.
- Naik, S. N., Goud, V. V., Rout, P. K., and Dalai, A. K. (2010). Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 14(2):578–597.
- Nair, J., Wierman, A., and Zwart, B. (2013). The Fundamentals of Heavy-tails: Properties, Emergence, and Identification. *ACM SIGMETRICS Performance Evaluation Review*, 41:387.
- Nakajima, T. and Higurashi, A. (1998). A use of two-channel radiances for an aerosol characterization from space. *Geophysical Research Letters*, 25(20):3815–3818.
- Nelson, T. A., Lee, D. J., and Smith, B. C. (2003). Are “green Tides” Harmful Algal Blooms? Toxic Properties of Water-Soluble Extracts from Two Bloom-Forming Macroalgae, *Ulva Fenestrata* and *Ulvaria Obscura* (ulvophyceae). *Journal of Phycology*, 39(5):874–879.
- Nishiguchi, M. K. and Somero, G. N. (1992). Temperature- and concentration-dependence of compatibility of the organic osmolyte β -dimethylsulfoniopropionate. *Cryobiology*, 29(1):118–124.
- NOAA (2012). Drought for June 2012. Technical report, NOAA, Online.

- NOAA (2016). Weekly mean CO₂ and historical comparisons. Technical report, NOAA.
- Nuffield Council on Bioethics (2011). *Biofuels : ethical issues*. Nuffield Council on Bioethics, London.
- OBIS (2019). Ulva Linnaeus, 1753.
- Okabe, A., Boots, B., Sugihara, K., and Chiu, S. N. (2009). *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, 2 edition. Google-Books-ID: dT7YH3mjeeIC.
- Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R. M., Lindsay, K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R. G., Plattner, G.-K., Rodgers, K. B., Sabine, C. L., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Totterdell, I. J., Weirig, M.-F., Yamanaka, Y., and Yool, A. (2005). Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*, 437(7059):681–686.
- Overvoorde, P., Fukaki, H., and Beekman, T. (2010). Auxin Control of Root Development. *Cold Spring Harbor Perspectives in Biology*, 2(6).
- Owen, N. A., Inderwildi, O. R., and King, D. A. (2010). The status of conventional world oil reserves: Hype or cause for concern? *Energy policy*, 38:4743–4749.
- Paerl, H. W. and Whitall, D. R. (1999). Anthropogenically-derived atmospheric nitrogen deposition, marine eutrophication and harmful algal bloom expansion. *Ambio*, 28(4).
- Pareto, V. (2014). *Manual of Political Economy: A Critical and Variorum Edition*. OUP Oxford. Google-Books-ID: 43NYAAQBAJ.
- Patterson, D. J. (1989). Stramenopiles: chromophytes from a protistan perspective. In Green, J. C., Leadbeater, B. S. C., and Diver, W. L., editors, *The Chromophyte algae: problems and perspectives*. Clarendon Press, Oxford.
- Pereira, L. (2016). *Edible Seaweeds of the World*. CRC Press. Google-Books-ID: sy90DgAAQBAJ.
- Pereira, R. and Yarish, C. (2008). Mass Production of Marine Macroalgae. In *Encyclopedia of Ecology*, pages 2236–2247.

- Peterson, D. A. (2010). Confocal Microscopy. In Kompoliti, K. and Verhagen Metman, L., editors, *Encyclopedia of Movement Disorders*. Academic Press. Google-Books-ID: 9OMOfuKosgC.
- Petráň, M., Hadravský, M., Egger, M. D., and Galambos, R. (1968). Tandem-Scanning Reflected-Light Microscope*. *JOSA*, 58(5):661–664.
- Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., and Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20(9):503–510.
- Pfarrherr, A., Teuchner, K., Leupold, D., and Hoffmann, P. (1991). Chlorophyll b in solution: fluorescence lifetimes, absorption and emission spectra as criteria of purity. *Journal of Photochemistry and Photobiology B: Biology*, 9(1):35–41.
- Pimm, S. L., Russell, G. J., Gittleman, J. L., and Brooks, T. M. (1995). The future of biodiversity. *Science-AAAS-Weekly Paper Edition*, 269:347–349.
- Porzio, L., Buia, M. C., and Hall-Spencer, J. M. (2011). Effects of ocean acidification on macroalgal communities. *Journal of Experimental Marine Biology and Ecology*, 400(1):278–287.
- Postma, H. (1984). Eutrophication of Dutch Coastal Waters. *Netherlands Journal of Zoology*, 35(1):348–359.
- Powers, D. M. W. (1998). Applications and Explanations of Zipf’s Law. In *Proceedings of the Joint Conferences on New Methods in Language Processing and Computational Natural Language Learning*, NeMLaP3/CoNLL ’98, pages 151–160, Stroudsburg, PA, USA. Association for Computational Linguistics. event-place: Sydney, Australia.
- Prochnik, S. E., Umen, J., Nedelcu, A. M., Hallmann, A., Miller, S. M., Nishii, I., Ferris, P., Kuo, A., Mitros, T., Fritz-Laylin, L. K., Hellsten, U., Chapman, J., Simakov, O., Rensing, S. A., Terry, A., Pangilinan, J., Kapitonov, V., Jurka, J., Salamov, A., Shapiro, H., Schmutz, J., Grimwood, J., Lindquist, E., Lucas, S., Grigoriev, I. V., Schmitt, R., Kirk, D., and Rokhsar, D. S. (2010). Genomic Analysis of Organismal Complexity in the Multicellular Green Alga *Volvox carteri*. *Science*, 329(5988):223–226.
- Pushkarev, D., Neff, N. F., and Quake, S. R. (2009). Single-molecule sequencing of an individual human genome. *Nature Biotechnology*, 27(9):847–850.

- R-Project (2019). stats package v3.6.2 | R Documentation, <https://www.rdocumentation.org/packages/stats/versions/3.6.2>.
- Radulovich, R., Umanzor, S., Cabrera, R., and Mata, R. (2014). *Tropical seaweeds for human food, their cultivation and its effect on biodiversity enrichment*, volume 436.
- Raffaelli, D. and Hawkins, S. J. (1999). *Intertidal Ecology*. Springer Netherlands.
- Raffaelli, D. G., Raven, J. A., and Poole, L. J. (1998). ECOLOGICAL IMPACT OF GREEN MACROALGAL BLOOMS. In Ansell, A., Barnes, M., and Gibson, R. N., editors, *Oceanography And Marine Biology: An Annual Review*, volume 36.
- Rasmussen, C. G. and Bellinger, M. (2018). An overview of plant division-plane orientation. *New Phytologist*, 219(2):505–512.
- Rautenberger, R. and Bischof, K. (2006). Impact of temperature on UV-susceptibility of two *Ulva* (Chlorophyta) species from Antarctic and Subantarctic regions. *Polar Biology*, 29(11):988–996.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108(D14).
- Reddit, u. (2017). Giant’s Causeway.
- Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters*, 74(1):15–19.
- Regalbuto, J. R. (2009). Cellulosic Biofuels—Got Gasoline? *Science*, 325(5942):822–824.
- Renewable Fuels Agency (2008). The Gallagher review of the indirect effects of biofuels production. Technical report, Renewable Fuels Agency London, London.
- Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York.
- Reyes, A. V., Carlson, A. E., Beard, B. L., Hatfield, R. G., Stoner, J. S., Winsor, K., Welke, B., and Ullman, D. J. (2014). South Greenland ice-sheet collapse during Marine Isotope Stage 11. *Nature*, 510:525–8.

- Robinson, M. M., Dowsett, H. J., and Chandler, M. A. (2008). Pliocene role in assessing future climate impacts. *Eos, Transactions American Geophysical Union*, 89:501–502.
- Roduit, C., Sekatski, S., Dietler, G., Catsicas, S., Lafont, F., and Kasas, S. (2009). Stiffness Tomography by Atomic Force Microscopy. *Biophysical Journal*, 97(2):674–677.
- Rokas, A. (2008). The Origins of Multicellularity and the Early History of the Genetic Toolkit For Animal Development. *Annual Review of Genetics*, 42(1):235–251.
- Rosen, B. R. (1988). Progress, problems and patterns in the biogeography of reef corals and other tropical marine organisms. *Helgoländer Meeresuntersuchungen*, 42(2):269–301.
- Rosen, K. T. and Resnick, M. (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics*, 8(2):165–186.
- Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007). The role of microorganisms in coral health, disease and evolution. *Nature Reviews Microbiology*, 5(5):355–362.
- Rosenberg, E. and Zilber-Rosenberg, I. (2018). The hologenome concept of evolution after 10 years. *Microbiome*, 6(1):78.
- Ross, A. B., Jones, J. M., Kubacki, M. L., and Bridgeman, T. (2008). Classification of macroalgae as fuel and its thermochemical behaviour. *Bioresource Technology*, 99(14):6494–6504.
- Rousseau, R. (2002). George Kingsley Zipf: life, ideas, his law and informetrics. *Glottometrics*, 3:11–18.
- Rugar, D. and Hansma, P. (1990). Atomic Force Microscopy. *Physics Today*, 43(10):23.
- Runge, C. F. and Senauer, B. (2007). How Biofuels Could Starve the Poor. Technical report.
- Sabnis, R. W. (2010). *Handbook of Biological Dyes and Stains: Synthesis and Industrial Applications*. John Wiley & Sons. Google-Books-ID: A8VvfFbNBiC.
- Sahlin, P. and Jönsson, H. (2010). A Modeling Study on How Cell Division Affects Properties of Epithelial Tissues Under Isotropic Growth. *PLOS ONE*, 5(7):e11750.

- Sahoo, D. and Seckbach, J. (2015). *The Algae World*. Springer. Google-Books-ID: nDg3CwAAQBAJ.
- Saladini, F., Patrizi, N., Pulselli, F. M., Marchettini, N., and Bastianoni, S. (2016). Guidelines for emergy evaluation of first, second and third generation biofuels. *Renewable and Sustainable Energy Reviews*, 66:221–227.
- Salic, A. and Mitchison, T. J. (2008). A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proceedings of the National Academy of Sciences*, 105(7):2415–2420.
- Santelices, B., Bolton, J. J., and Meneses, I. (2009). Marine Algal Communities. In Witman, J. D. and Roy, K., editors, *Marine Macroecology*. University of Chicago Press.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682.
- Schlesinger, W. H. and Bernhardt, E. S. (2013). *Biogeochemistry: An Analysis of Global Change*. Academic Press. Google-Books-ID: 533UOWBU3_AC.
- Schramm, W. and Nienhuis, P. N. (2012). *Marine Benthic Vegetation: Recent Changes and the Effects of Eutrophication*. Springer Science & Business Media. Google-Books-ID: leLrCAAAQBAJ.
- Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., and Yu, T.-H. (2008). Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, 319:1238–40.
- Semwogerere, D. and Weeks, E. R. (2008). Confocal Microscopy. In Wnek, G. E. and Bowlin, G. L., editors, *Encyclopedia of Biomaterials and Biomedical Engineering*. CRC Press. Google-Books-ID: 0ThZDwAAQBAJ.
- Shafiee, S. and Topal, E. (2009). When will fossil fuel reserves be diminished? *Energy policy*, 37:181–189.
- Shapiro, H. M. and Perlmutter, N. G. (2001). Violet laser diodes as light sources for cytometry. *Cytometry*, 44(2):133–136.

- Shen, L., Xu, H., and Guo, X. (2012). Satellite Remote Sensing of Harmful Algal Blooms (HABs) and a Potential Synthesized Framework. *Sensors (Basel, Switzerland)*, 12(6):7778–7803.
- Shimada, S., Yokoyama, N., Arai, S., and Hiraoka, M. (2008). Phylogeography of the genus *Ulva* (Ulvophyceae, Chlorophyta), with special reference to the Japanese freshwater and brackish taxa. *Journal of Applied Phycology*, 20(5):979–989.
- Shorrocks, A., Davies, J., and Lluberas, R. (2018). Credit Suisse Research Institute Global Wealth Report 2018. Technical Report 9, Credit Suisse Research Institute, Zurich, Switzerland.
- Siegenthaler, U. and Sarmiento, J. L. (1993). Atmospheric carbon dioxide and the ocean. *Nature*, 365(6442):119.
- Sigma-Aldrich (2019). Direct Red 23 212490.
- Silva, P. C. (1952). *A Review of Nomenclatural Conservation in the Algae from the Point of View of the Type Method*. University of California Press. Google-Books-ID: 8gScQAAACAAJ.
- Simon, H. A. (1955). ON A CLASS OF SKEW DISTRIBUTION FUNCTIONS. *Biometrika*, 42(3-4):425–440.
- Sims, R. E. H., Mabee, W., Saddler, J. N., and Taylor, M. (2010). An overview of second generation biofuel technologies. *Bioresource Technology*, 101:1570–80.
- Singh, A., Nigam, P. S., and Murphy, J. D. (2011). Renewable fuels from algae: an answer to debatable land based fuels. *Bioresour Technol*, 102:10–6.
- Sissini, M., Barros-Barreto, M. B., Széchy, M., Lucena, M., Oliveira, M., Gower, J., Liu, G., De Oliveira Bastos, E., Milstein, D., Gusmão, F., Martinelli-Filho, J., Alves-Lima, C., Colepicolo, P., Ameka, G., de Graft Johnson, K., Gouvea, L., Torrano Silva, B., Nauer, F., Nunes, J. M., and Horta, P. (2017). The floating Sargassum (Phaeophyceae) of the South Atlantic Ocean - Likely scenarios. *Phycologia*, 56(3):321–328.
- Smetacek, V. and Zingone, A. (2013). Green and golden seaweed tides on the rise. *Nature*, 504(7478):84–88.

- Smil, V. (2001). *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food Production*. MIT Press. Google-Books-ID: HvdFQgAACAAJ.
- Smith, S. A., Beaulieu, J. M., and Donoghue, M. J. (2009). Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology*, 9(1):37.
- Smith, S. V. (1981). Marine Macrophytes as a Global Carbon Sink. *Science*, 211(4484):838–840.
- Smith, V. H., Tilman, G. D., and Nekola, J. C. (1999). Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution*, 100(1–3):179–196.
- Sokolov, I., Dokukin, M. E., and Guz, N. V. (2013). Method for quantitative measurements of the elastic modulus of biological cells in AFM indentation experiments. *Methods*, 60(2):202–213.
- Soulsby, P. G., Lowthion, D., Houston, M., and Montgomery, H. A. C. (1985). The role of sewage effluent in the accumulation of macroalgal mats on intertidal mudflats in two basins in Southern England. *Netherlands Journal of Sea Research*, 19(3):257–263.
- Spoerner, M., Wichard, T., Bachhuber, T., Stratmann, J., and Oertel, W. (2012). Growth and Thallus Morphogenesis of *Ulva mutabilis* (Chlorophyta) Depends on A Combination of Two Bacterial Species Excreting Regulatory Factors. *Journal of Phycology*, 48(6):1433–1447.
- Spolaore, P., Joannis-Cassan, C., Duran, E., and Isambert, A. (2006). Commercial applications of microalgae. *Journal of Bioscience and Bioengineering*, 101(2):87–96.
- Stachowicz, J. J., Graham, M., Bracken, M. E. S., and Szoboszlai, A. I. (2008). Diversity Enhances Cover and Stability of Seaweed Assemblages: The Role of Heterogeneity and Time. *Ecology*, 89(11):3008–3019.
- Steneck, R. S., Graham, M. H., Bourque, B. J., Corbett, D., Erlandson, J. M., Estes, J. A., and Tegner, M. J. (2002). Kelp forest ecosystems: biodiversity, stability, resilience and future. *Environmental Conservation*, 29(4):436–459.

- Stigler, S. M. (1982). Poisson on the poisson distribution. *Statistics & Probability Letters*, 1(1):33–35.
- Stratmann, J., Paputsoglu, G., and Oertel, W. (1996). Differentiation of *Ulva Mutabilis* (chlorophyta) Gametangia and Gamete Release Are Controlled by Extracellular Inhibitors¹. *Journal of Phycology*, 32(6):1009–1021.
- Straub, S. C., Thomsen, M. S., and Wernberg, T. (2016). The Dynamic Biogeography of the Anthropocene: The Speed of Recent Range Shifts in Seaweeds. In Hu, Z.-M. and Fraser, C., editors, *Seaweed Phylogeography: Adaptation and Evolution of Seaweeds under Environmental Change*. Springer Netherlands.
- Szczurek, A. T., Prakash, K., Lee, H.-K., Żurek Biesiada, D. J., Best, G., Haggmann, M., Dobrucki, J. W., Cremer, C., and Birk, U. (2014). Single molecule localization microscopy of the distribution of chromatin using Hoechst and DAPI fluorescent probes. *Nucleus*, 5(4):331–340.
- Sánchez-Gutiérrez, D., Sáez, A., Gómez-Gálvez, P., Paradas, C., and Escudero, L. M. (2017). Rules of tissue packing involving different cell types: human muscle organization. *Scientific Reports*, 7:40444.
- Sánchez-Gutiérrez, D., Tozluoglu, M., Barry, J. D., Pascual, A., Mao, Y., and Escudero, L. M. (2016). Fundamental physical cellular constraints drive self-organization of tissues. *The EMBO Journal*, 35(1):77–88.
- Tait, K., Joint, I., Daykin, M., Milton, D. L., Williams, P., and Cámara, M. (2005). Disruption of quorum sensing in seawater abolishes attraction of zoospores of the green alga *Ulva* to bacterial biofilms. *Environmental Microbiology*, 7(2):229–240.
- Tanner, C. and Guiry, M. D. (2019). *Ulva* Linnaeus, 1753.
- Tanner, C. E. (1980). *Chloropelta* Gen. Nov., an Ulvaceous Green Alga with a Different Type of Development¹. *Journal of Phycology*, 16(1):128–137.
- Tarbell, K. A. and Reid, J. (1991). A COMPUTER VISION SYSTEM FOR CHARACTERIZING CORN GROWTH AND DEVELOPMENT. *Transactions of the ASAE*, 34(5):2245–2255.

- Taylor, J. E. (1976). The Structure of Singularities in Soap-Bubble-Like and Soap-Film-Like Minimal Surfaces. *Annals of Mathematics*, 103(3):489–539.
- Teichberg, M., Fox, S. E., Aguila, C., Olsen, Y. S., and Valiela, I. (2008). Macroalgal responses to experimental nutrient enrichment in shallow coastal waters: growth, internal nutrient pools, and isotopic signatures. *Marine Ecology Progress Series*, 368:117–126.
- Teichberg, M., Fox, S. E., Olsen, Y. S., Valiela, I., Martinetto, P., Iribarne, O., Muto, E. Y., Petti, M. A. V., Corbisier, T. N., Soto-Jiménez, M., Páez-Osuna, F., Castro, P., Freitas, H., Zitelli, A., Cardinaletti, M., and Tagliapietra, D. (2010). Eutrophication and macroalgal blooms in temperate and tropical coastal waters: nutrient enrichment experiments with *Ulva* spp. *Global Change Biology*, 16(9):2624–2637.
- Theis, K. R., Dheilly, N. M., Klassen, J. L., Brucker, R. M., Baines, J. F., Bosch, T. C. G., Cryan, J. F., Gilbert, S. F., Goodnight, C. J., Lloyd, E. A., Sapp, J., Vandenkoornhuise, P., Zilber-Rosenberg, I., Rosenberg, E., and Bordenstein, S. R. (2016). Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems*, 1(2):e00028–16.
- Thermo Fisher Scientific (2019a). Click-iT EdU Alexa Fluor 555 Imaging Kit - Catalog number: C10338.
- Thermo Fisher Scientific (2019b). Click-iT EdU Alexa Fluor 594 Imaging Kit - Catalog number: C10339.
- Thermo Fisher Scientific (2019c). Click-iT EdU Alexa Fluor 647 Imaging Kit - Catalog number: C10340.
- Thomas, G., Burnham, N. A., Camesano, T. A., and Wen, Q. (2013). Measuring the Mechanical Properties of Living Cells Using Atomic Force Microscopy. *Journal of Visualized Experiments : JoVE*, (76).
- Thorlabs (2019). 8 Megapixel Scientific-Grade Cameras for Microscopy.
- Toomre, D. and Pawley, J. (2010). Disk-scanning Confocal Microscopy. In *Handbook of Biological Confocal Microscopy*. Springer Science & Business Media. Google-Books-ID: IKcPnaNPrhoC.

- Trainer, V. L. and Yoshida, T. (2014). Proceedings of the Workshop on Economic Impacts of Harmful Algal Blooms on Fisheries and Aquaculture. PICES Scientific Report 47.
- Trivedi, N., Gupta, V., Reddy, C. R. K., and Jha, B. (2013). Enzymatic hydrolysis and production of bioethanol from common macrophytic green alga *Ulva fasciata* Delile. *Bioresource Technology*, 150:106–112.
- UN (1992). Rio Declaration on Environment and Development. Technical report.
- UNEP (2013). Observed Concentrations of CO₂ Cross 400 parts per million Threshold at Several Global Atmosphere Watch Stations.
- United Nations, Population Division, D. o. E. a. S. A. (2017). World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Technical Report ESA/P/WP/248, United Nations, New York.
- US Government (2007). Energy Independence and Security Act of 2007.
- USDA (2016). WASDE-552.
- Valeur, B. and Berberan-Santos, M. N. (2011). A Brief History of Fluorescence and Phosphorescence before the Emergence of Quantum Theory. *Journal of Chemical Education*, 88:731–738.
- Valiela, I., McClelland, J., Hauxwell, J., Behr, P. J., Hersh, D., and Foreman, K. (1997). Macroalgal Blooms in Shallow Estuaries: Controls and Ecophysiological and Ecosystem Consequences. *Limnology and Oceanography*, 42(5):1105–1118.
- Van Alstyne, K. L. and Puglisi, M. P. (2007). DMSP in marine macroalgae and macroinvertebrates: Distribution, function, and ecological impacts. *Aquatic Sciences*, 69(3):394–402.
- Van Alstyne, K. L., Wolfe, G. V., Freidenburg, T. L., Neill, A., and Hicken, C. (2001). Activated defense systems in marine macroalgae: evidence for an ecological role for DMSP cleavage. *Marine Ecology Progress Series*, 213:53–65.
- van den Hoek, C., Mann, D. G., and Jahns, H. M. (1995). *Algae: An Introduction to Phycology*. Cambridge University Press, Cambridge, 1 edition. Google-Books-ID: s1P855ZWc0kC.

- Vandepoele, K., Bel, M. V., Richard, G., Landeghem, S. V., Verhelst, B., Moreau, H., Peer, Y. V. d., Grimsley, N., and Piganeau, G. (2013). pico-PLAZA, a genome database of microbial photosynthetic eukaryotes. *Environmental Microbiology*, 15(8):2147–2153. [_eprint: https://sfamjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1462-2920.12174](https://sfamjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1462-2920.12174).
- Verkade, P. (2012). Electron Microscopy (TEM and SEM). In *Essential Guide to Reading Biomedical Papers*, pages 59–65. John Wiley & Sons, Ltd.
- Viaroli, P., Bartoli, M., Bondavalli, C., Christian, R. R., Giordani, G., and Naldi, M. (1996). Macrophyte communities and their impact on benthic fluxes of oxygen, sulphide and nutrients in shallow eutrophic environments. In *Coastal Lagoon Eutrophication and ANaerobic Processes (C.L.E.AN.)*, Developments in Hydrobiology, pages 105–119. Springer, Dordrecht.
- Vitousek, P. M., Ehrlich, P. R., Ehrlich, A. H., and Matson, P. A. (1986). Human Appropriation of the Products of Photosynthesis. *BioScience*, 36(6):368–373.
- Watkiss, P., Downing, T., Handley, C., and Butterfield, R. (2005). The impacts and costs of climate change. *Brussels, European Commission DG Environment*.
- Weiss, A., Costa, R., and Wichard, T. (2017). Morphogenesis of *Ulva mutabilis* (Chlorophyta) induced by *Maribacter* species (Bacteroidetes, Flavobacteriaceae). *Botanica Marina*, 60(2):197–206.
- Wheeler, G. L., Tait, K., Taylor, A., Brownlee, C., and Joint, I. (2006). Acyl-homoserine lactones modulate the settlement rate of zoospores of the marine alga *Ulva intestinalis* via a novel chemokinetic mechanism. *Plant, Cell & Environment*, 29(4):608–618.
- Whittaker, R. H. (1969). New Concepts of Kingdoms of Organisms. *Science*, 163(3863):150–160.
- Wichard, T. (2015). Exploring bacteria-induced growth and morphogenesis in the green macroalga order Ulvales (Chlorophyta). *Frontiers in Plant Science*, 6.
- Wichard, T. and Oertel, W. (2010). Gametogenesis and Gamete Release of *Ulva Mutabilis* and *Ulva Lactuca* (chlorophyta): Regulatory Effects and Chemical Characterization of the “Swarming Inhibitor”1. *Journal of Phycology*, 46(2):248–259.

- Willig, M., Kaufman, D., and Stevens, R. (2003). Latitudinal Gradients of Biodiversity: Pattern, Process, Scale, and Synthesis. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):273–309.
- Winston, M. L. (1991). *The Biology of the Honey Bee*. Harvard University Press.
- Wolfram (2015). VoronoiMesh—Wolfram Language Documentation.
- Wright, A. J. and Smith, L. G. (2008). Division Plane Orientation in Plant Cells. In Verma, D. P. S. and Hong, Z., editors, *Cell Division Control in Plants*, Plant Cell Monographs, pages 33–57. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., and Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1):1–37.
- Xu, J., Fan, X., Zhang, X., Xu, D., Mou, S., Cao, S., Zheng, Z., Miao, J., and Ye, N. (2012). Evidence of Coexistence of C3 and C4 Photosynthetic Pathways in a Green-Tide-Forming Alga, *Ulva prolifera*. *PLOS ONE*, 7(5):e37438.
- Yanai, I., Camacho, C. J., and DeLisi, C. (2000). Predictions of Gene Family Distributions in Microbial Genomes: Evolution by Gene Duplication and Modification. *Physical Review Letters*, 85(12):2641–2644.
- Yang, Y., Boncoeur, J., Liu, S., and Nyvall-Collen, P. (2018). Economic assessment and environmental management of green tides in the Chinese Yellow Sea. *Ocean & Coastal Management*, 161:20–30.
- Ye, N.-h., Zhang, X.-w., Mao, Y.-z., Liang, C.-w., Xu, D., Zou, J., Zhuang, Z.-m., and Wang, Q.-y. (2011). ‘Green tides’ are overwhelming the coastline of our blue planet: taking the world’s largest example. *Ecological Research*, 26(3):477–485.
- Yokoyama, H. and Ishihi, Y. (2010). Bioindicator and biofilter function of *Ulva* spp. (Chlorophyta) for dissolved inorganic nitrogen discharged from a coastal fish farm — potential role in integrated multi-trophic aquaculture. *Aquaculture*, 310(1):74–83.
- Yoon, H. S., Nelson, W., Lindstrom, S. C., Boo, S. M., Pueschel, C., Qiu, H., and Bhattacharya, D. (2016). Rhodophyta. In Archibald, J. M., Simpson, A. G., Slamovits, C. H.,

- Margulis, L., Melkonian, M., Chapman, D. J., and Corliss, J. O., editors, *Handbook of the Protists*, pages 1–45. Springer International Publishing, Cham.
- Yoon, H. S., Zuccarello, G., and Bhattacharya, D. (2010). Evolutionary History and Taxonomy of Red Algae. In *Red Algae in the Genomic Age*, pages 25–42.
- Young, L. J. and Young, J. (2013). *Statistical Ecology*. Springer Science & Business Media. Google-Books-ID: 8a_VBwAAQBAJ.
- Yule, G. U. (1925). II.—A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 213(402-410):21–87.
- Zhang, X., Ye, N., Liang, C., Mou, S., Fan, X., Xu, J., Xu, D., and Zhuang, Z. (2012). De novo sequencing and analysis of the *Ulva linza* transcriptome to discover putative mechanisms associated with its successful colonization of coastal ecosystems. *BMC Genomics*, 13:565.
- Zhou, Y., Tan, L., Pang, Q., Li, F., and Wang, J. (2015). Influence of nutrients pollution on the growth and organic matter output of *Ulva prolifera* in the southern Yellow Sea, China. *Mar Pollut Bull*, 95:107–14.
- Zhu, D., Ortega, C. F., Motamedi, R., Szewciw, L., Vernerey, F., and Barthelat, F. (2012). Structure and Mechanical Performance of a “Modern” Fish Scale. *Advanced Engineering Materials*, 14(4):B185–B194.
- Ziegler, J. (2013). Burning food crops to produce biofuels is a crime against humanity.
- Zondervan, I., Zeebe, R. E., Rost, B., and Riebesell, U. (2001). Decreasing marine biogenic calcification: A negative feedback on rising atmospheric pCO₂. *Global Biogeochemical Cycles*, 15(2):507–516.

Mathematica Code - CVT

```
LXCP = Import["Root\Data.csv"];
myvmesh = VoronoiMesh[LXCP, {{0, 1024}, {0, 1024}}]
Poly = MeshCells[myvmesh, 2]; )
XandYbounds = {{0, 1024}, {0, 1024}}
InputPoints = LXCP;
HistoryOfPoints =
  With[{maxit = 1000,
        tol = 0.1},
    FixedPointList[
      Function[InputPoints,
        Block[{cells},
          cells = MeshPrimitives[VoronoiMesh[InputPoints, XandYbounds],
            "Faces"];
          RegionCentroid /@
            cells[[SparseArray[
              Outer[#2@#1 &, InputPoints, RegionMember /@ cells, 1],
              Automatic, False]["NonzeroPositions"][[All, 2]]]]],
        InputPoints, maxit,
        SameTest -> (Max[MapThread[EuclideanDistance, {#1, #2}]] <
          tol &)]];
Print["Number of iterations: ", Length[HistoryOfPoints]];
```

```

myv1 = VoronoiMesh[HistoryOfPoints[[1]],
  XandYbounds]
myvmid = VoronoiMesh[HistoryOfPoints[[MID]],
  XandYbounds]
myv2 = VoronoiMesh[HistoryOfPoints[[-1]],
  XandYbounds]
Polys = MeshCells [myv1, 2];
Polys2 = MeshCells [myv2, 2];
Centroidy = MeshCellCentroid [myv1]
Show[myv1, Graphics[{Red, Point[PropertyValue[{myv1, 2}, MeshCellCentroid]}]]]
mycentroids = PropertyValue[{myv1, 2}, MeshCellCentroid];
mycentroids2 = PropertyValue[{myv2, 2}, MeshCellCentroid];
FNmyv[i_] :=
  FNmyv[i] = VoronoiMesh[HistoryOfPoints[[i]], XandYbounds]
FNcells[i_] := FNcells[i] = MeshPrimitives[FNmyv[i], 2];

FNregb[i_] := FNregb[i] = RegionBounds[FNmyv[i]];
FNinout[i_] :=
  FNinout[i] = IntersectingQ[Flatten@FNregb[i], Flatten@RegionBounds[#]] & /@
  FNcells[i];
FNin[i_] := FNin[i] = Pick[FNcells[i], FNinout[i], False];
FNout[i_] := FNout[i] = Pick[FNcells[i], FNinout[i], True];
Table[FNin[i], {i, 1, MAX}];
NumberOfPolysOfGivenEdgeCount[polys_, n_] :=
  Length[Select[polys, Length#[[1]] == n &]]
ListOfEdgeLengths [polys_, maxn_] :=
  Table[NumberOfPolysOfGivenEdgeCount[polys, i], {i, 3, maxn}]
myedgehistory = Table[ListOfEdgeLengths[FNin[i], 10], {i, 1, MAX}];
Transpose[myedgehistory];

```

Mathematica Code - KummerU Function

```
BDM Simulator beta v1.4

ulvadata = CompressedData["Data"];
NO = 9874;
ObservedGeneFamilies = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,
  16, 17, 18, 18, 19, 20, 21, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35,
  36, 38, 49, 54, 60, 65, 89, 97, 198};
ulvadataexc0s = CompressedData["Data"];
sigma2inc0 = Variance[ulvadata];
sigma2exc0 = Variance[ulvadataexc0s];

DataCount[i_] := Count[ulvadata, i];

TableUlva = Table[DataCount[i], {i, 1, Max[ulvadata]}];
FreqTableUlva = TableUlva/16047;

MuValue=RandomReal[{0.01,0.05},50000];
GammaValue=RandomReal[{0,0.01},50000];

Export["C:\\...path\\MuValue.csv", MuValue];
Export["C:\\...path\\GammaValue.csv", GammaValue];
```

```
MuValueImported = Import["C:\\...path\\MuValue.csv"];
GammaValueImported = Import["C:\\...path\\GammaValue.csv"];

RcValue[x_] := (2*N0*(Abs[(GammaValueImported[[x]] - MuValueImported[[x]])])/
  sigma2exc0;
NuValue[x_] := (GammaValueImported[[x]]/(MuValueImported[[x]] -
  GammaValueImported[[x]]);

FirstElement[x_, y_] := (RcValue[x]*Gamma[1 + NuValue[x]]/(OGFNo0[[y]])
Kummer[x_, y_] :=
  Abs[HypergeometricU[NuValue[x], 0, ((RcValue[x]*(OGFNo0[[y]]))/N0)]]
ExponentialElement[x_, y_] := (ee)^(-((RcValue[x]*(OGFNo0[[y]]))/N0))
Probabilitynm[a_,
  b_] := (FirstElement[a, b])*(Kummer[a, b])*(ExponentialElement[a, b])

SimulationN[N_] := Table[Probabilitynm[x, y], {x, 1, N}, {y, 1, 43}];

Simulation10 = SimulationN[10];

MADSim100 = Map[MedianDeviation, Simulation100];

Length[MADSim100]
```

100

R Code - Voronoi Data Extraction

```
library(deldir)
z<-deldir(x,y)
w<-tile.list(z)
plot(w, showpoints=FALSE)
tileInfo(z)
print.tileInfo(w)
plot.tile.list(z)

a<-tileInfo(z)
b<-a[1]
allEdgeCounts<-a[2]
tabEdgeCounts<-a[3]
allEdgeLengths<-a[4]
Areas<-a[5]
uniqueEdgeLengths<-a[6]
individual_tiles<-lapply(a$indivTiles, "[", 1)
max_length <- max(sapply(individual_tiles, length))
new_individual_tiles <- lapply(individual_tiles, function(v) { c(v, rep(NA,
max_length-length(v))))})
```

Cumulative Distribution Functions

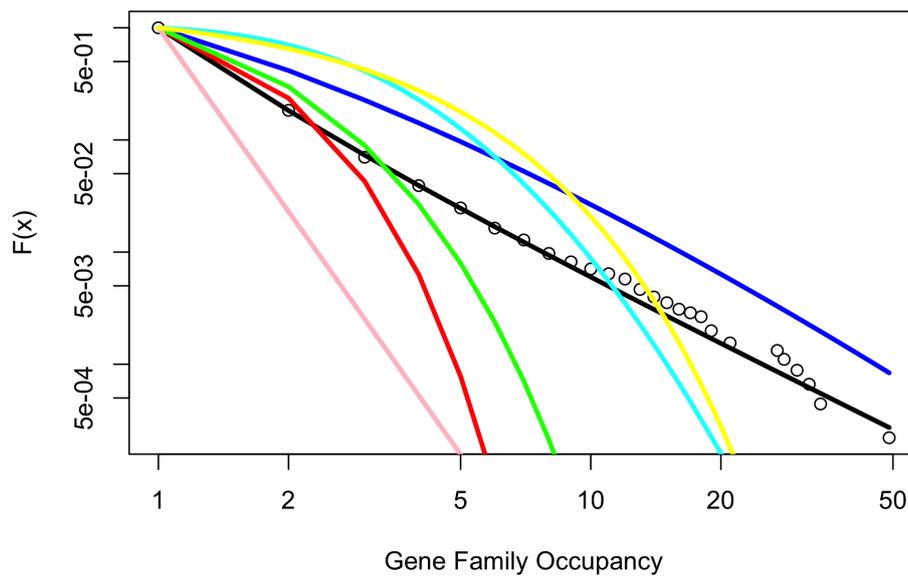


Figure D.1: The Cumulative Distribution Function (CDF) plots for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for *Picochlorum* sp. SENEW3 (SE3). For a CDF, $F(x)$ is a probability between 0 and 1. The legend is shown below.

Legend	
Circles	Observed gene family sizes
Black ---	Discrete Power Law
Red ---	Poisson
Blue ---	Discrete Log Normal
Green ---	Discrete Exponential
Pink ---	Continuous Power Law
Cyan ---	Continuous Log Normal
Yellow ---	Continuous Exponential

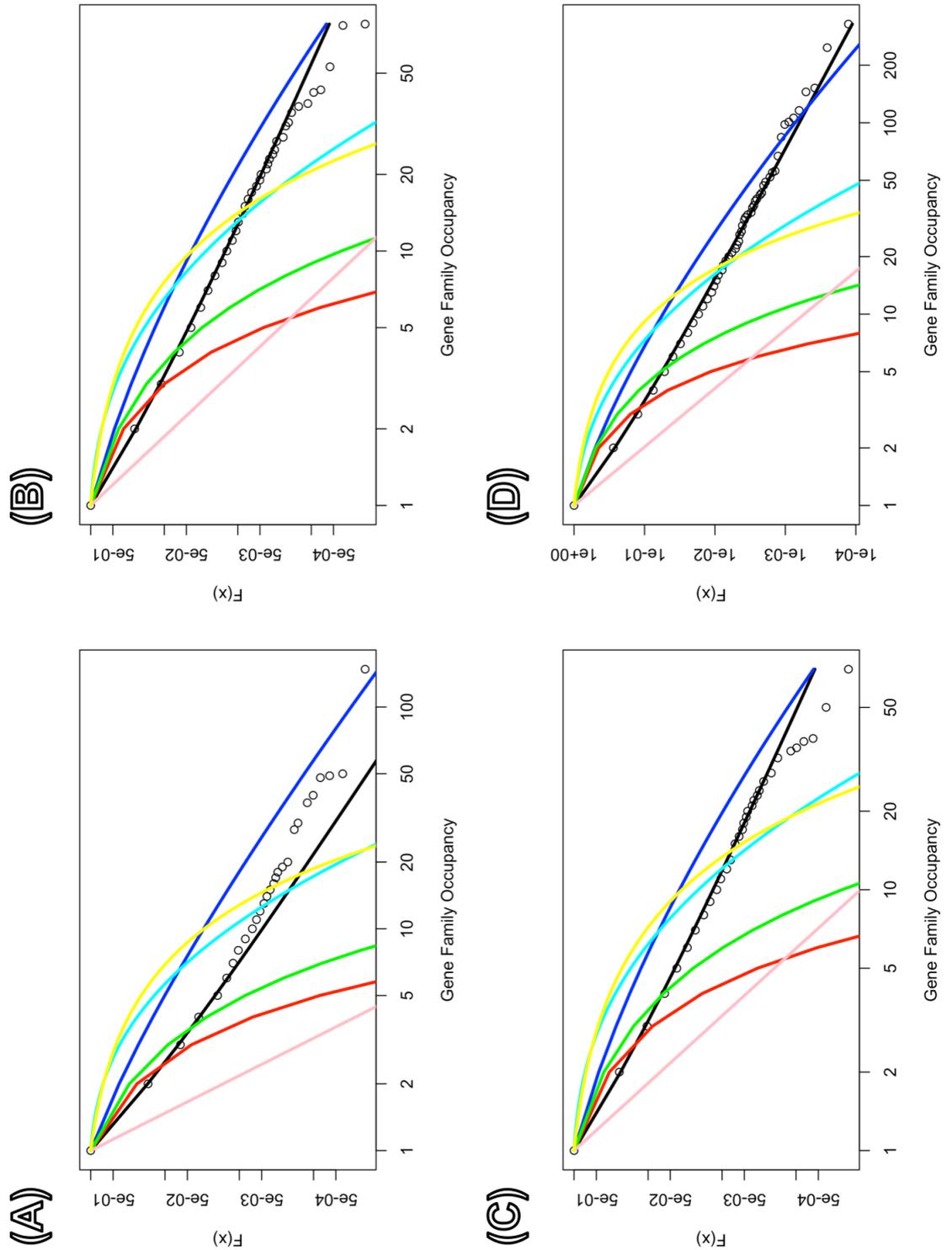


Figure D.2: The Cumulative Distribution Function (CDF) plots for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for (A) *Bathycoccus prasinos*, (B) *Chlorella* sp. NC64A, (C) *Coccomyxa subellipsoidea* C-169 and (D) *Gonium pectorale*. For a CDF, $F(x)$ is a probability between 0 and 1. For a CDF, $F(x)$ is a probability between 0 and 1. Please refer to the legend above.

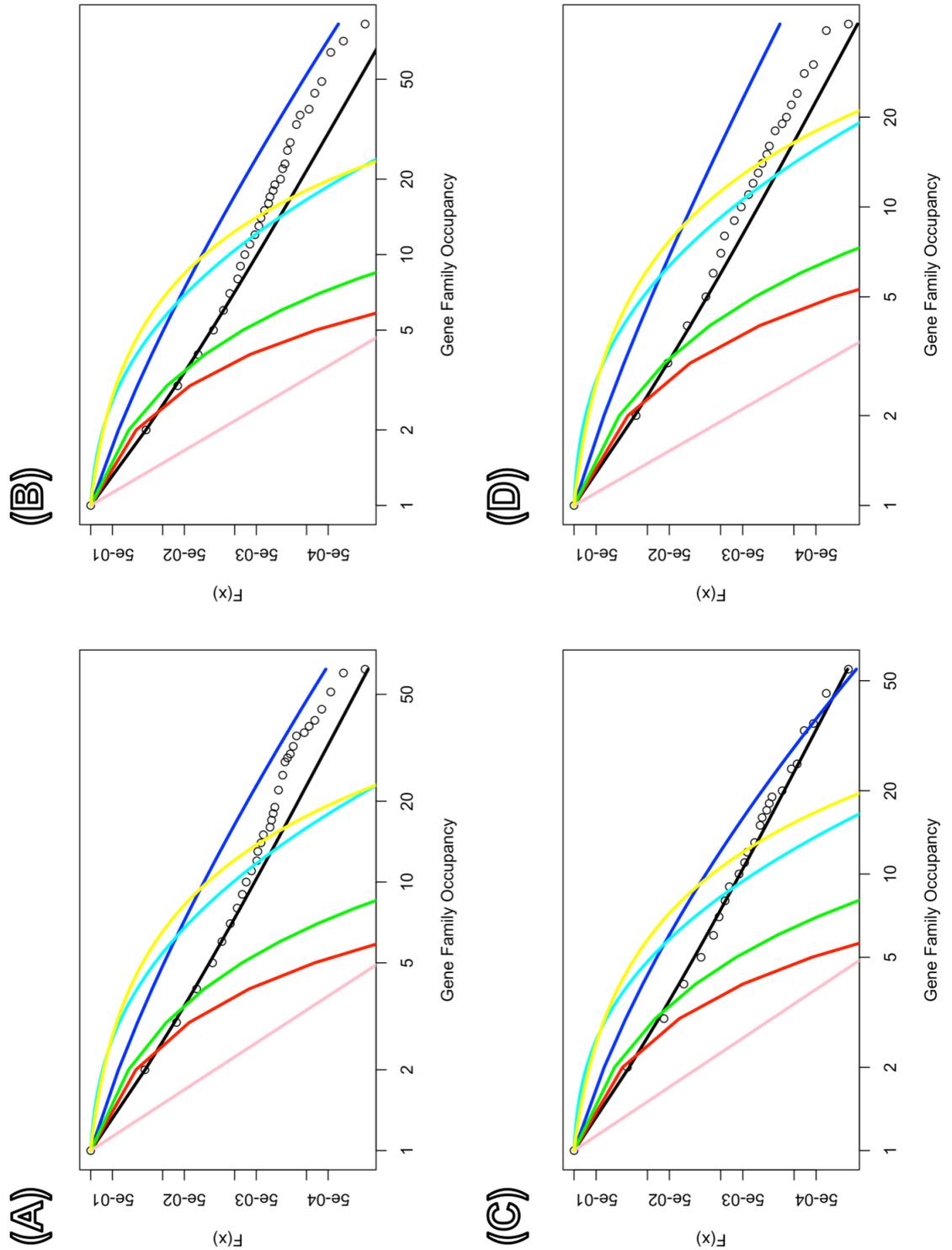


Figure D.3: The Cumulative Distribution Function (CDF) plots for gene family occupancy predicted by seven heavy tailed distribution models plotted against the observed data for (A) *Micromonas pusilla* strain CCMP1545, (B) *Micromonas* sp. RCC299, (C) *Ostreococcus lucimarinus* and (D) *Ostreococcus* sp. RCC809. For a CDF, $F(x)$ is a probability between 0 and 1. For a CDF, $F(x)$ is a probability between 0 and 1. Please refer to the legend above.

Cumulative Distribution Function Confidence Intervals

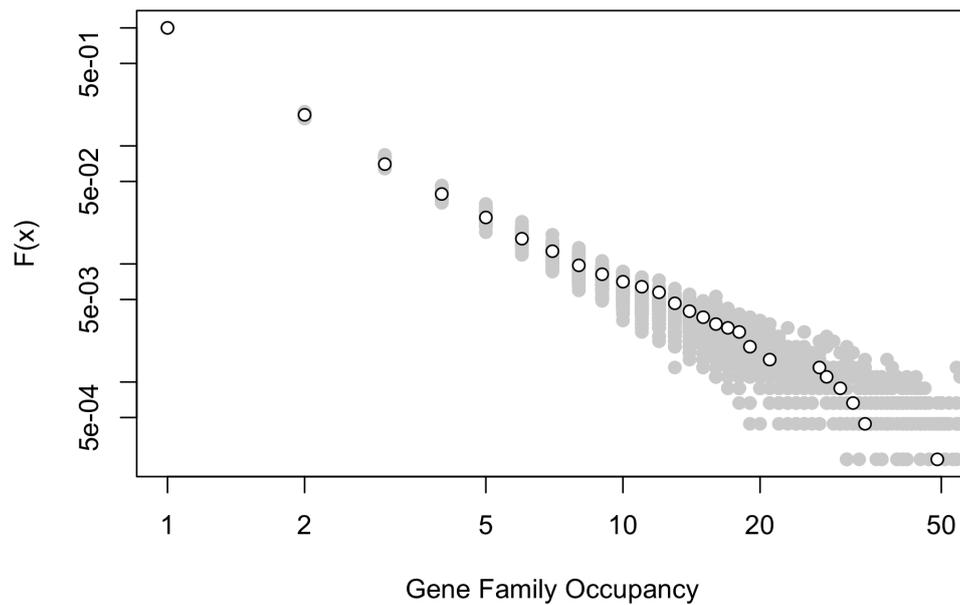


Figure E.1: The Cumulative Distribution Function (CDF) plot of the confidence intervals for the discrete power law model calculated by 100 simulations for *Picochlorum* sp. SENEW3 (SE3). The grey data points indicate the simulated points, and the black data points indicate the observed gene family occupancy data. For a CDF, $F(x)$ is a probability between 0 and 1.

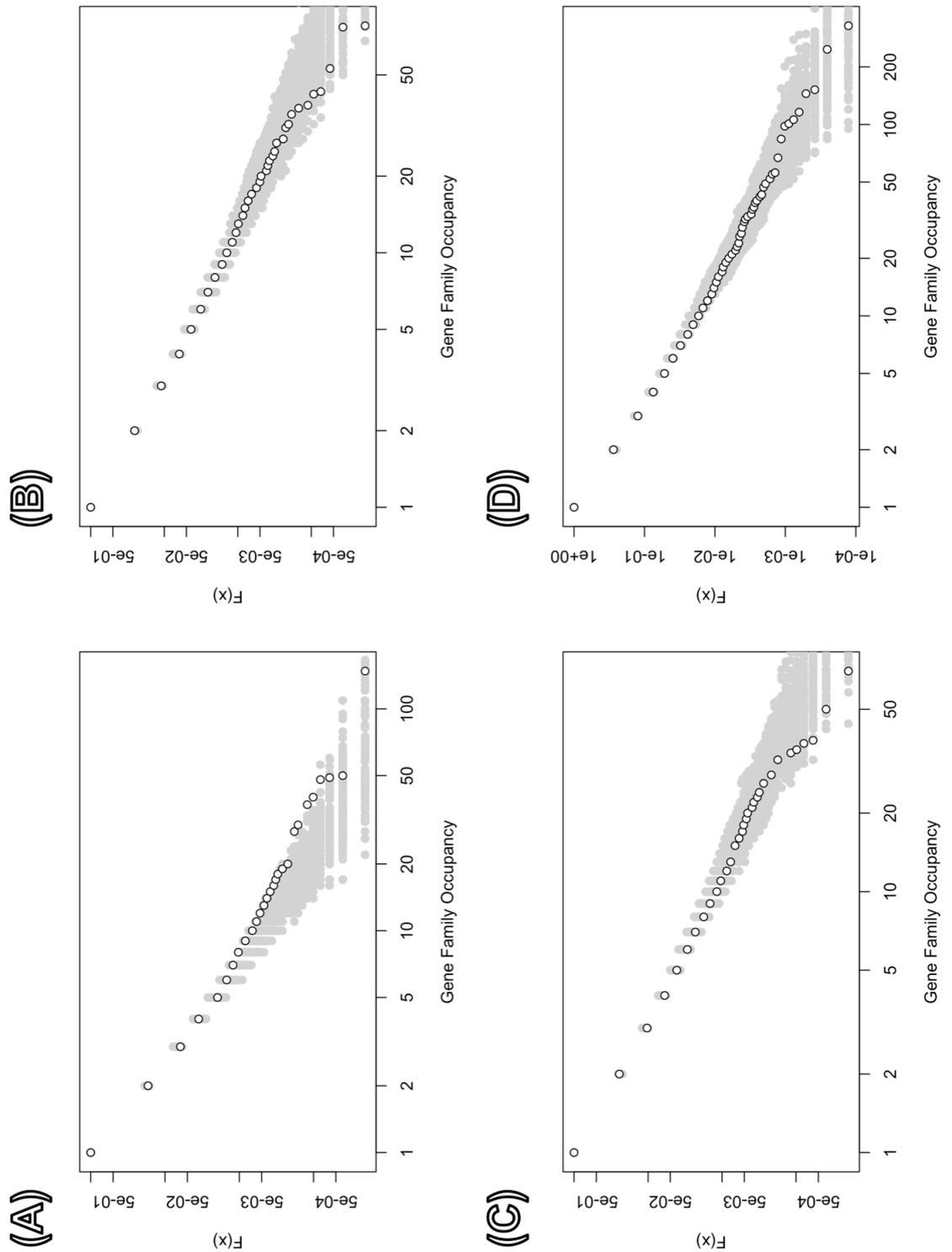


Figure E.2: The Cumulative Distribution Function (CDF) plots of the confidence intervals for the discrete power law model calculated by 100 simulations for (A) *Bathycoccus prasinos*, (B) *Chlorella* sp. NC64A, (C) *Coccomyxa subellipsoidea* C-169 and (D) *Gonium pectorale*. The grey data points indicate the simulated points, and the black data points indicate the observed gene family occupancy data. For a CDF, $F(x)$ is a probability between 0 and 1.

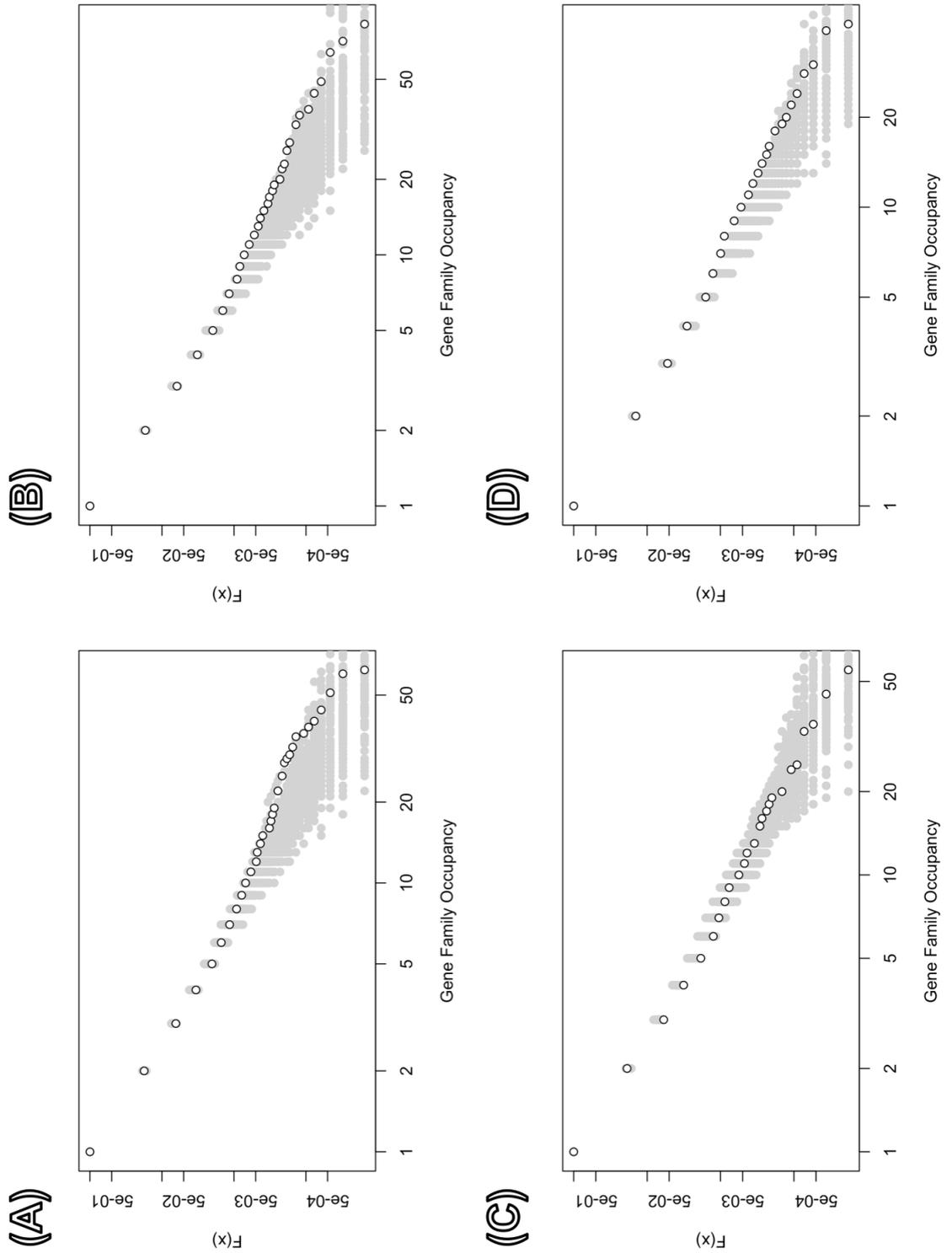


Figure E.3: The Cumulative Distribution Function (CDF) plots of the confidence intervals for the discrete power law model calculated by 100 simulations for (A) *Micromonas pusilla* strain CCMP1545, (B) *Micromonas* sp. RCC299, (C) *Ostreococcus lucimarinus* and (D) *Ostreococcus* sp. RCC809. The grey data points indicate the simulated points, and the black data points indicate the observed gene family occupancy data. For a CDF, $F(x)$ is a probability between 0 and 1.