THE GENETIC BACKGROUND OF RAPID GROWTH IN THE GREEN SEAWEED *ULVA* (SEA LETTUCE):

UNRAVELING THE GREEN TIDE PROBLEM AND SUPPORTING THE SUSTAINABLE FUTURE OF BLUE ECONOMY

Abdul Hawil Abas

Student number: 02202642

Promotor(s): Prof. dr. Olivier de Clerk & dr. Jonas Blomme
Tutor(s): Emma Coulembier Vandelannoote, M.Sc.

Copyright

Gent, 7 juni 2024

The first promotor,   The second promotor,

Prof. Olivier de Clerck   *Dr. Jonas Blomme*

The tutor,   The author,

Emma Coulembier Vandelannoote M.Sc.   *Abdul Hawil Abas*

Preface

As I present this thesis, I am deeply grateful to those who have supported and guided me throughout this journey. I would like to express my heartfelt thanks to my promotors, Prof. Olivier de Clerck and Prof. Jonas Blomme, whose expertise, patience, and encouragement were invaluable in shaping this work. My sincere gratitude also goes to my tutor, Emma Vandelannoote M.Sc., for her insightful feedback and unwavering support.

I extend my thanks to Prof. Thomas Jacobs for allowing me to use the facilities in his lab, "Plant Genome Edit," and to all the lab members, especially Mattias Vermeesch, who assisted with the analysis. To my family, your constant love and belief in me provided the foundation I needed to persevere. Lastly, to my friends, thank you for your understanding, companionship, and the moments of joy that kept me grounded. This accomplishment is as much yours as it is mine.

Table of Contents

List of abbreviations

AFLP: Amplified Fragment Length Polymorphism
ANOVA: Analysis of Variance
CNY: Chinese Yuan
CO$_2$: Carbon Dioxide
COIII: Cytochrome c Oxidase Subunit III
CRISPR-Cas9: Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated Protein 9
CTAB: Cetyltrimethylammonium Bromide
DMS: Dimethyl Sulfide
DMSP: Dimethyl Sulfoniopropionate
DNA: Deoxyribonucleic Acid
EUR: Euro
F1: First Filial Generation
FACS: Fluorescence-Activated Cell Sorting
FAO: Food and Agriculture Organization
GWAS: Genome-Wide Association Studies
HPLC: High-Performance Liquid Chromatography
IGIS: In Vitro Growth Imaging System
IMTA: Integrated Multitrophic Aquaculture
ITS: Internal Transcribed Spacer
MAS: Marker-Assisted Selection
NIL: Near-Isogenic Line
NIR: Near Infrared
PCR: Polymerase Chain Reaction
PVE: Phenotypic Variance Explained
QTL: Quantitative Trait Loci
RAD-Seq: Restriction Site-Associated DNA Sequencing
RB: Retinoblastoma
RIL: Recombinant Inbred Line
RNA: Ribonucleic Acid
ROH: Runs of Homozygosity
RubisCO: Ribulose-1,5-bisphosphate Carboxylase-Oxygenase
SBF: S phase-promoting Complex
SLAF-seq: Specific Length Amplified Fragment Sequencing
SNP: Single Nucleotide Polymorphism
SPX: Simplex
SSR: Simple Sequence Repeat
UCM: Ulva Culture Medium
VL: Linear Growth Rate
WGS: Whole Genome Sequencing
kbp: Kilobase pairs
kt: Kilotonne

Abstract

Green seaweeds from the genus *Ulva* are known for their rapid, expansive growth, often leading to green tide events that significantly impact marine ecosystems. These algae blooms pose ecological risks such as hypoxia and obstructed sunlight penetration, accelerating decay and fostering harmful microorganisms. This thesis investigates the genetic factors behind *Ulva*'s rapid growth by creating the first mapping population of green seaweed through crossing a fast-growing (*slender*) strain and a slow-growing (wild-type) strain of *Ulva*. The population was genotyped and phenotyped, with mating types determined using PCR, and 22 single nucleotide polymorphisms (SNPs) analyzed via Multiplex PCR and Oxford Nanopore sequencing. Growth rates were measured using the IGIS platform, and additional phenotypic traits were assessed through direct observation. Results revealed successful crossbreeding, yielding 166 isolated segregating lines with a nearly even distribution of mating types (43% mt+ and 52% mt-, with 5% displaying both types). Phenotyping indicated structural variations, including tubular thalli (74), blades (52) and hard to determine (12), with rhizoid presence significantly linked to blade structure. Growth projected area analysis using IGIS showed that segregating lines have variable growth (from 25.86 to 227.8 mm^2 at end time point). Multiplex PCR and Nanopore sequencing confirmed the presence of SNPs, though expected differences between *slender* and wild-type strains were not validated. This research enhances our understanding of *Ulva*'s genetic growth factors and sets the stage for future studies to optimize aquaculture productivity and manage green tide impacts.

Keywords: green seaweed, *Ulva,* green tide, mapping population, growth rate

I. INTRODUCTION

Green tides are blooms formed by green algae, usually *Ulva* spp. The rapid growth of *Ulva* can cause both ecological and economic problems. For instance, in the Yellow Sea of China, green tides of *Ulva prolifera* significantly impact marine ecosystems. It starts with their rapid growth that leads to increased seawater pH due to efficient carbon assimilation. This shift can alter phytoplankton communities and disrupt ecosystem functions. *Ulva* also outcompetes other seaweeds by producing allelochemicals (Gao et al., 2014).

As the green tide wanes, the decomposition of *Ulva* fronds consumes oxygen, creating hypoxic conditions detrimental to marine life and aquaculture. Decomposition also releases nutrients that degrade water quality and contribute to coastal acidification (Zhang et al., 2019). Green tides are happening worldwide and there are more than 40 regions where green tides happen regularly (Ye et al., 2011). Decomposition of *Ulva* often produces toxic substances such as hydrogen sulfide gas which is toxic to humans, causing breathing problems. This has already caused the death of two people in 2015. The decomposition of *Ulva* can also be toxic to other marine organisms like abalone (C. Wang et al., 2011), oyster (Green-Gavrielidis et al., 2018), and fish (Q. Fan et al., 2022). Economically, the *U. prolifera* green tide in 2021 resulted in 1.44 million tons biomass and the government of Qingdao spent 350 million CNY (~45 million EUR), consumed 46.7 kt oil-Eq of fossil fuels, and released 331.0 kt CO_2-Eq of greenhouse emissions to clean it up (Chen et al., 2022).

The major cause of these macroalgae blooms has been linked to eutrophication, a high input of nutrients usually from anthropogenic activities like agriculture (J. Cai et al., 2023). Therefore, preventing *Ulva* blooms by limiting the nutrient flow from the agriculture to the coastal environment has been the major preventing measures (Xia et al., 2022). However, it has been demonstrated that the influx of nutrients from anthropogenic activities alone is not the sole cause of green tides (Keesing et al., 2016). A green tide in Yatsu tidal flat, Japan even happened during remediation efforts (Yabe et al., 2009). Moreover, it has been demonstrated that there are differences in growth rate and metabolism between strains that can or cannot form green tides. These differences can occur due to different selective pressures. It has been suggested that the green tide strains can be utilized in aquaculture, due to for example, their faster growth (Fort et al., 2020). This hints that there is a genetic component to green tides occurrence. Substantial efforts have been made to study green tides from an environmental perspective. However, the genetic background of this rapid growth of *Ulva* has not been studied clearly (Blomme et al., 2023). Therefore, it is clear that there is a research gap to find out the genetics that contribute to the rapid growth of *Ulva* green tide forming strains. One way to study this is to create a mapping population.

Mapping populations are essential for genetic research. It provides a basis for identifying genetic markers linked to phenotypic traits through the crossing of parents with distinct characteristics. It is used to analyze trait segregation and genetic underpinnings in the progeny. Mapping populations are vital for uncovering genes associated with distinct traits like growth rate, metabolic activity or specific nutrient utilization. The mapping population in this study is created from crossing two *Ulva* lab strains, the faster growing *slender* and the slower growing wild-type. Concurrently, phenotyping is the precise measurement of physical and biochemical traits in various environments. It plays a crucial role in deciphering the genotype's expression and its interaction with the environment, enhanced by high-throughput technologies for efficient trait measurement. In this study, the IGIS (in vitro growth imaging system) platform was utilized for phenotyping the

mapping population. *Ulva* are cultivated on this platform in petri dishes that are rotated on a disk, and pictures of each plate are taken every hour in order to measure the growth rate over time (Dhondt et al., 2014). These measurements are complemented with direct observations of general morphology and smaller structures such as the rhizoid area with a binocular microscope. Genotyping complements this by determining *Ulva* genetic composition through DNA analysis, using techniques such as PCR, Single Nucleotide Polymorphisms (SNP) analysis, and whole-genome sequencing (WGS) to associate genetic variations with phenotypic traits. This synergy between mapping populations, phenotyping, and genotyping is instrumental in this research.

Building on this foundation, Genome-Wide Association Studies (GWAS) and Quantitative Trait Loci (QTL) mapping will be central in associating genes with *Ulva* rapid growth traits. GWAS scans genomes to associate genetic markers linked to specific traits, providing insights into complex traits influenced by multiple genes and environmental factors. This method can identify novel genes and genetic pathways by analyzing genetic variants across a broad spectrum (Santure & Garant, 2018). QTL mapping complements GWAS by focusing on regions of the genome associated with quantitative traits like growth rate, pinpointing specific chromosomal regions that significantly influence these traits (Kumar et al., 2017). SNPs are variations at a single position in the DNA sequence among segregating lines of a species (Fadason et al., 2022). SNPs serve as crucial genetic markers that can illuminate genetic diversity, population structure, and adaptive traits (Kinnby et al., 2020; X. Wang et al., 2018). The significance of SNPs in seaweed research is particularly profound as they may help in understanding the genetic basis of important traits such as growth rates, stress tolerance, and reproductive strategies.

This thesis aims to investigate the genetic determinants of growth rate differences between the *slender* and wild-type strains of *Ulva* by creating a mapping population and utilizing phenotypic and genotypic analyses. The objectives are; 1.) cross *slender* with wild-type to generate a mapping population. 2.) phenotype the morphology and growth rates of the mapping population. 3.) genotype the mapping population by determining mating types through PCR of specific sex-determining regions and 4.) establish SNP analysis using multiplex PCR. These steps will lay the foundation for future studies employing GWAS and QTL analyses to associate specific genes with fast growth and possibly green tide formation.

II. LITERATURE REVIEW

2.1. Introduction to *Ulva* and seaweed biology, variability and applications

Seaweeds or macroalgae encompass a diverse group of photosynthetic algae thriving in marine environments. These organisms have adapted to various depths and ecological niches in polar, temperate and tropical seas (Anburaj et al., 2024). Generally, macroalgae are divided into three groups: Green (Chlorophyta), red (Rhodophyta), and brown (Phaeophyceae) seaweeds (Vieira et al., 2021). The distinction between the three groups is described in Table 1 and Figure 1. These organisms exhibit a wide range of morphologies, from simple unicellular forms to complex multicellular structures that can reach up to several meters in length. Macroalgae can have significant morphological differences both within and between species (Gaspar et al., 2017; Torres et al., 2015). Seaweeds play a crucial role in marine ecosystems as primary producers, contributing significantly to oceanic carbon capture and forming the basis of marine food webs (Cotas et al., 2023). They possess unique life cycles that can include both asexual and sexual reproduction methods, often featuring alternation of generations (X. Liu et al., 2017). Their adaptability allows them to inhabit a broad spectrum of habitats and makes them a vital component of marine biodiversity, but also an essential resource for coastal communities (Cotas et al., 2023).
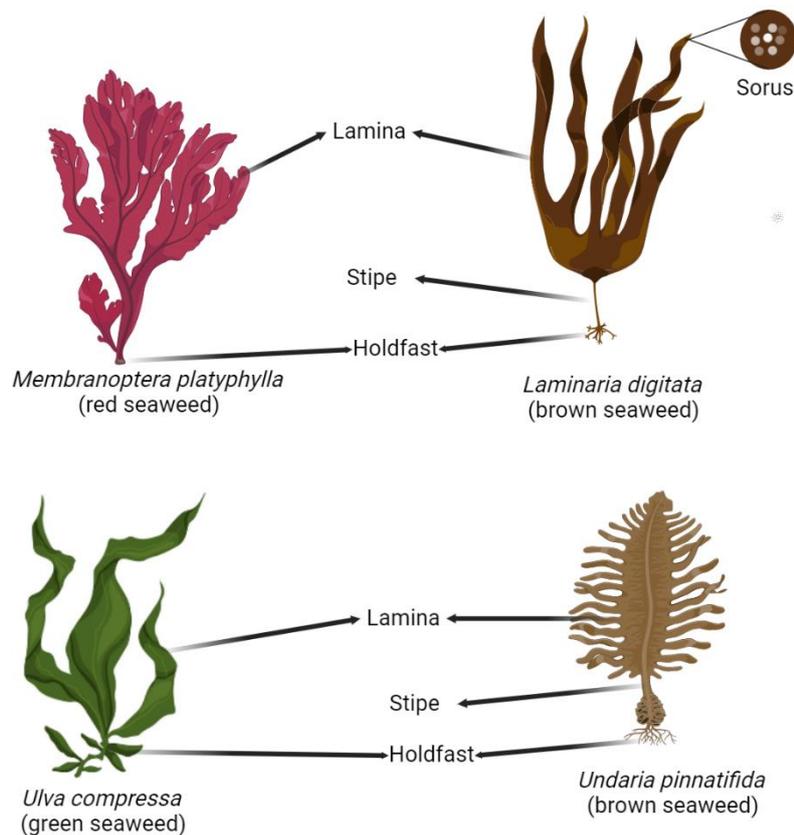

Figure 1. Seaweed general structure. Not drawn to scale. Figure made by author using Biorender.

Table 1. The distinct features of three macroalgae groups.

	Red algae	Brown algae	Green algae
Color	Red due to phycoerythrin masks chlorophyll (De Clerck et al., 2012)	Olive green to brown due to fucoxanthin dominate chlorophyll (Din et al., 2022)	Green due to chlorophyll a and b (De Clerck et al., 2012)
Habitat	Mostly marine, some species in deep water to efficiently absorb green, blue and violet light (De Clerck et al., 2005)	Mostly cold marine waters. Intertidal zone forming underwater forest (Bringloe et al., 2020)	Variety of environments, incl. freshwater, saltwater, moist land (Tuya et al., 2014)
Cell structure	Mostly multicellular with complex structures. Cell walls contain agar or carrageenan (J. L. Jiang et al., 2021)	Primarily multicellular and grow very large. Cell walls composed of cellulose and alginic acid (Bogolitsyn et al., 2020)	Unicellular, multicellular, colonial, filamentous. Cell wall mainly consists of cellulose (De Clerck et al., 2012)
Photosynthetic pigments	Contains chlorophyll a, d and phycoerythrin along with phycocyanin in some species (De Clerck et al., 2012)	Chlorophyll a, c and fucoxanthin (Din et al., 2022)	Mainly chlorophyll a and b, small amounts of carotenoids and xanthophyll (De Clerck et al., 2012)
Examples	*Corallina*, *Porphyra* (nori) and *Gracilaria*	*Fucus*, *Sargassum*, giant kelps (*Macrocystis* and *Laminaria*)	*Spirogyra*, *Chlamydomonas*, *Ulva*

Ulva, commonly called sea lettuce or green laver, is a genus of green algae in the *Ulvaceae* family, class *Ulvophyceae*, phylum *Chlorophyta*. It exemplifies green algae with its chlorophyll-a and chlorophyll-b photosynthesis. Noted for its diversity, *Ulva* includes many species that share a typical cellular structure characteristic of green algae (Tran et al., 2022). *Ulva* are ubiquitous, meaning they can occur almost everywhere in the world and grow in salt, brackish and freshwater habitats (Rybak, 2021; Sebök & Hanelt, 2023; Wu et al., 2018). *Ulva* morphology can be very diverse, morphologies are often variable and can overlap even within species. This makes it difficult to determine the species based on morphology alone (Ismail & Mohamed, 2017; Kazi et al., 2016). One example of intraspecific variation is found in the lab strains of *Ulva mutabilis/compressa*. The tubular lab strain *slender* is usually floating, unbranched and does not have secondary rhizoids. Meanwhile, the wild-type strain forms flat blades and has a better developed rhizoid area (Spoerner et al., 2012). Some *Ulva* form a thin thallus that grows long while others grow more evenly to make a wide, blade-like thallus (Figure 2).
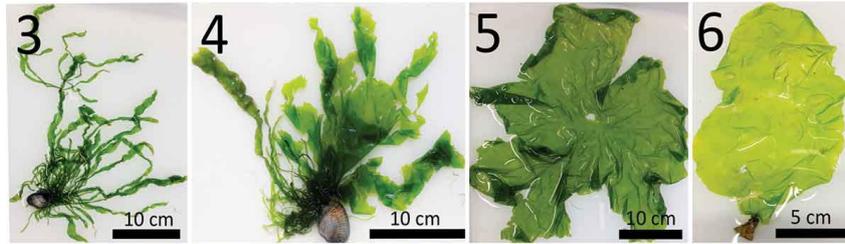
Figure 2. *Ulva* different morphologies within the same species (*Ulva compressa*). (3) Tubular, branching form of *Ulva* with a central rhizoid part. (4). Similar structure with figure 3 but shows a mix of tubular and blade like structure. (5). *Ulva* without central rhizoid and blade like structure. (6). Blade like structure of *Ulva* with holdfast/stipe structure that helps to attach to a concrete. Figure by Steinhagen, Weinberger, et al. (2019).

Ulva exhibits a haplodiplontic life cycle, which is characterized by the alternation of generations between a haploid gametophyte and a diploid sporophyte. In this life cycle, both the gametophyte and sporophyte stages are morphologically similar, a phenomenon known as isomorphism. The gametophyte produces gametes through mitosis, which then fuse to form a zygote, developing into the sporophyte. The sporophyte, in turn, undergoes meiosis to produce haploid spores (zoids), which germinate to give rise to new gametophytes. This cycle ensures genetic diversity and allows *Ulva* to adapt to various environmental conditions, contributing to its widespread presence in marine ecosystems around the world (De Clerck et al., 2018; Wichard et al., 2015). The whole *Ulva* life cycle is described in Figure 3.
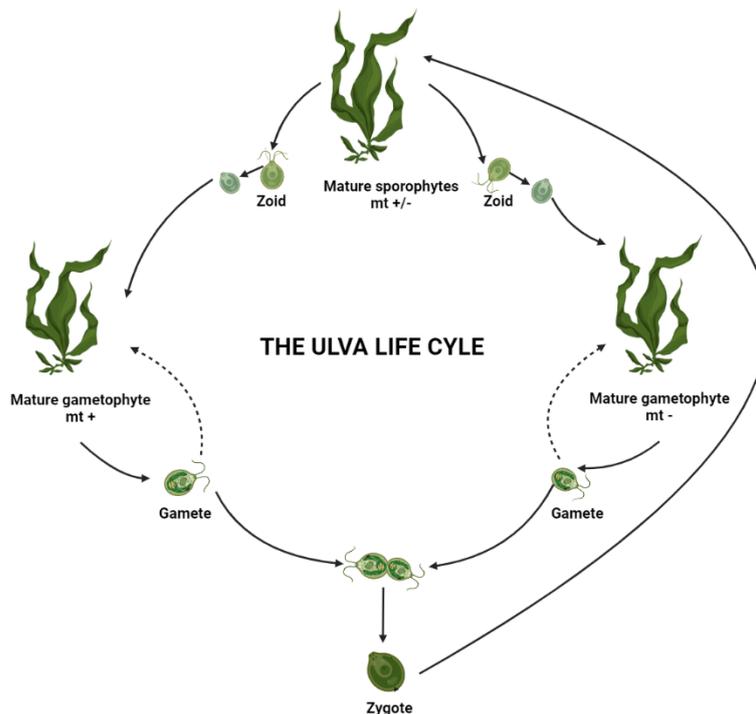
Figure 3. *Ulva* spp. life cycle. Not drawn to scale. Figure made by author using Biorender.

Ulva has found applications across various fields, demonstrating its versatility. In the culinary world, *Ulva* species have been embraced as a nutritious food source, rich in vitamins, minerals, and dietary fibers (Debbarma et al., 2016). They are utilized in a variety of dishes, from

salads to soups, offering a subtle flavor and a boost of green color. Beyond nutrition, *Ulva's* high protein content make it an appealing ingredient for health-conscious consumers. Its use in the culinary field extends to the emerging sector of plant-based alternatives, where *Ulva* is being explored as a sustainable ingredient for vegan products, showcasing its potential to contribute to food security and dietary diversity (E et al., 2023; Magnusson et al., 2019; Paiva et al., 2017).

In environmental and bioengineering applications, *Ulva*'s role is equally significant. Its ability to absorb nutrients and heavy metals from water bodies makes it a valuable agent in bioremediation efforts, helping to improve water quality and combat eutrophication in aquatic ecosystems (Bews et al., 2021; Oliveira et al., 2024). Furthermore, researchers are exploring *Ulva* as a feedstock for biofuels, leveraging its rapid growth rate and high biomass yield. This application presents a promising avenue for renewable energy production, minimizing reliance on fossil fuels and reducing greenhouse gas emissions (Bikker et al., 2016). Additionally, *Ulva* is used in the production of bio-based materials, such as bioplastics and packaging materials, offering an eco-friendly alternative to petroleum-based products (Sonchaeng et al., 2023). These diverse utilizations underscore *Ulva*'s potential to contribute to sustainable practices across multiple industries, from food and health to environmental management and renewable energy.

2.2. Understanding the biology of *Ulva*: -omics studies

2.2.1. Genome

The genome represents a focal point of interest within phycological research due to its implications for understanding both evolutionary biology and the specific adaptive mechanisms of seaweeds. Genomic studies in the world of macroalgae has been done more in the red and brown algae compared to green algae. In red algae, the genome of *Pyropia yezoensis* or susabi-nori (Nakamura et al., 2013), *Porphyra umbilicalis* (Brawley et al., 2017), and *Chondrus crispus* (Brawley et al., 2017) has been sequenced and studied. These studies give insights about the metabolites production, evolutionary biology and growth strategies. The brown seaweed *Ectocarpus siliculosus* genome reveals capabilities to live in the tidal environment by having a complete set of the pigment biosynthesis and unique metabolic processes (Cock et al., 2010). Moreover, the *Saccharina* genome study focused on the biosynthesis of polysaccharide (Ye et al., 2015) and the *Cladosiphon okamuranus* genome study resulted in an identification of the genes associated with the biosynthetic pathway for sulfated fulcans and alginate (Nishitsuji et al., 2016).

The *Ulva mutabilis/compressa* genome exhibits distinct variances when compared to unicellular organisms, underscoring its unique evolutionary path. Notably, the genome contains a smaller number of transcriptional regulators compared to other green algae families. Additionally, *Ulva* appears to have developed multicellularity through a different mechanism, as it lacks the critical cell division genes of the retinoblastoma (RB)/E2F pathway along with D-type cyclins. Similarly, genes such as Cln 2/3, SBF, and Whi5, which are involved in the G1-S transition in yeast, were absent in *Ulva's* genome. This suggests a unique regulatory mechanism for multicellularity development (De Clerck et al., 2018). Another study about *U. prolifera* proposed that the cause of the green tide or rapid growth is associated with the genes of cell cycle, phosphorylation and stress resistance (He et al., 2021). When compared to the genome reported by De Clerck et al., a recent draft genome research of *U. compressa* (a Chilean isolate) indicates numerous noteworthy differences (Osorio et al., 2022). The Chilean genome is smaller (80,8 Mb) compared to the 98.5

Mb reported for *U. mutabilis/compressa*. Nevertheless, this smaller genome has a greater gene density (238 genes per Mb). This higher gene density suggests a more compact and possibly more functionally diverse genome, which could have implications for the organism's adaptability and biological processes.

In other green seaweeds, genome study has provided more information towards unique cell biology features. The high regeneration ability of the coenocytic feather green algae *Byopsis* was influenced by the genes encoding BPL-1/Bryohealin which were majorly duplicated and absent in other green seaweeds. There is 30+ genes encoding kinesin but only one myosin gene. This is suggested to the main cause of the coenocytic structure because myosin is more important in cytokinesis (Ochiai et al., 2024). Another study of a siphonous green algae *Caulerpa lentilifera* showed that the expression of expanded transport regulators varies regionally, implying that the structural patterning approach of a multinucleate cell is dependent on nuclear pore protein diversification (Arimoto et al., 2019).

Besides the nuclear genome, organelle DNA content has been studied in *Ulva*, mostly in the field of phylogeny and taxonomy. Numerous rearrangement events have resulted in a very flexible structure of *Ulva* chloroplast genomes (F. Liu & Melton, 2021). However, this is not applicable to all *Ulva* species. *Ulva flexuosa* chloroplast genomes showed that there was significant rearrangement outside of *Ulva* spp. and great conservation of the chloroplast genomes between them (C. Cai et al., 2017). These insights highlight the importance of genetic approaches in understanding the diversity and evolutionary history of *Ulva* species.

DNA-based delimitation is a method used to identify and classify species based on genetic information rather than morphological characteristics. Traditional morphological methods have often proven insufficient due to *Ulva*'s high phenotypic plasticity and morphological convergence. Utilizing DNA barcoding, particularly through the sequencing of genetic markers such as the 18S rDNA, ITS regions, and *tufA* gene, has provided more precise and reliable species identification. This molecular approach has revealed cryptic species, which were previously indistinguishable morphologically, and has clarified phylogenetic relationships within the genus. Consequently, DNA-based delimitation not only enhances the accuracy of species classification but also aids in the conservation and management of these ecologically significant algae. Taxa from *Ulva* and their evolutionary connections according to the *tuf*A gene is depicted in Figure 4.
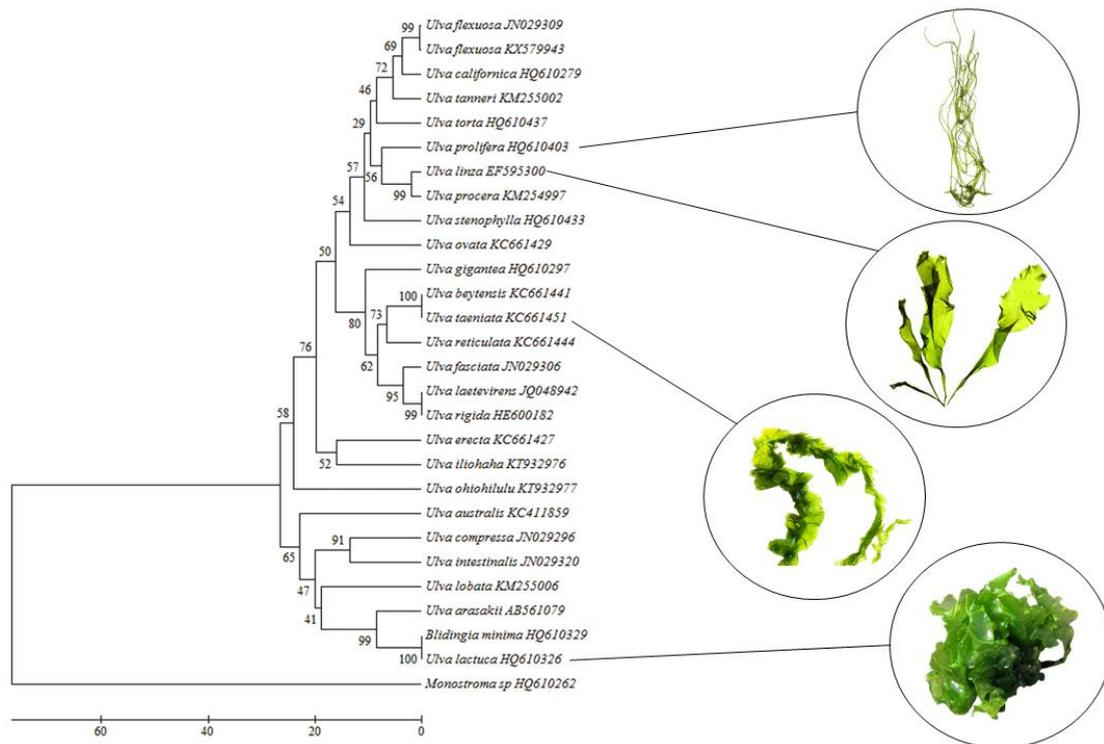
Figure 4. *Ulva* evolutionary relationships based on *tuf*A gene. Figure by C. Cai et al. (2017).

2.2.2. Epigenome and transcriptome

The study of the *Ulva* epigenome opens fascinating avenues into understanding the intricate layers of regulation that overlay its genomic blueprint. Epigenetic modifications, including DNA methylation, histone modifications, and RNA-mediated processes, serve as dynamic regulators of gene expression without altering the underlying DNA sequence. The *Ulva* epigenome has not yet been studied clearly. Research on this topic must begin to unravel how epigenetic landscapes in *Ulva* are remodeled in response to developmental cues, external stress factors such as salinity shifts, nutrient availability, and light intensity, facilitating rapid acclimatization. For example, cytosine methylation levels can cause protoplasts of *Ulva reticulata* to regenerate into diverse morphotypes (Gupta et al., 2012). The exploration of the *Ulva* epigenome will not only enrich our understanding of algal biology but also offers potential insights into the evolution of epigenetic control mechanisms across the tree of life.

The transcriptome analysis of *Ulva* species provides profound insights into the dynamic expression of genes across different developmental stages, environmental conditions, and stress responses, reflecting the organism's remarkable adaptability and metabolic versatility. Transcriptomic studies can unveil the complex gene expression networks that underpin *Ulva*'s life cycle, including the switch between haploid and diploid phases, and its ability to undergo morphogenesis in response to environmental stimuli. For instance, the transcriptome dynamics of *U. mutabilis* gametogenesis revealed that 45% genes were differentially expressed during gametogenesis. This study identified a conserved RWP-RK transcription factor in the activation of sexual reproduction in *Ulva* (X. Liu et al., 2022).

2.2.3. Proteome and metabolome

By cataloging protein expression under various conditions, proteome studies in *Ulva* illuminate the molecular machinery driving its growth, photosynthesis, and stress responses. Fan et al. (2018) presents a comparative proteomic analysis on how *U. prolifera* reacts to high-temperature stress. 1,223 differentially expressed proteins in response to heat stress were identified, revealing significant upregulation in proteins linked to stress response, oxidative phosphorylation, and signal transduction pathways, among others. Downregulation was observed in proteins associated with photosynthesis and carbon fixation. This investigation enhances understanding of the molecular adaptations of *U. prolifera* to high-temperature stress, marking a significant step towards exploring the genetic and protein-level responses of algae to environmental stressors (M. Fan et al., 2018).

The exploration of the *Ulva* metabolome, encompassing the complete set of metabolites present within the organism, has opened up new vistas in understanding its biochemical diversity and ecological interactions. Metabolomic studies have detailed the array of primary and secondary metabolites in *Ulva*, including amino acids, fatty acids, vitamins, and unique bioactive compounds, which play pivotal roles in nutrition, metabolism, and defense mechanisms. For instance, the identification of sulfur-containing compounds like dimethyl sulfoniopropionate (DMSP) has highlighted their role in osmotic regulation and as antioxidants, besides their contribution to the global sulfur cycle through the release of dimethyl sulfide (DMS) (Kessler et al., 2018). Additionally, the presence of a variety of phenolic compounds and terpenoids underscores *Ulva*'s chemical defense strategies against predators, pathogens, and biofouling organisms (Alagan et al., 2017; Anjali et al., 2019; Vargas Cárdenas et al., 2023). Metabolomic profiles have also provided insights into the adaptability of *Ulva* to varying environmental conditions, revealing the metabolic adjustments it undergoes to optimize energy production and stress resilience (Alsufyani et al., 2017; Ghaderiardakani et al., 2022; Gupta & Kushwaha, 2017; He et al., 2018, 2019; Sanchez-Arcos et al., 2022).

2.2.4. Microbiome;

The studies of the *Ulva* microbiome have unveiled the complex and dynamic interactions between this green alga and its associated microbial communities (Figure 5), highlighting a fascinating aspect of marine ecology and symbiosis. The microbial consortia associate with *Ulva* tissues and comprise bacteria, fungi, and microalgae, playing integral roles in nutrient cycling, defense against pathogens, and the overall health and growth of the alga. When cultivated without the presence of microorganisms, *Ulva* forms a loose, callus-like clump of cells with deformed cell walls. Complete morphogenesis is only seen upon exposure to specific bacterial strains. In lab research, axenic *Ulva* can be exposed to a particular strain of *Maribacter* and a certain *Roseovarius* in order to establish morphogenesis. Research has demonstrated that specific bacterial communities are essential for inducing *Ulva*'s normal development, particularly in its early life stages, where certain bacteria produce signaling molecules that trigger the settlement and morphogenesis of *Ulva* spores (van der Loos et al., 2021).

Moreover, the *Ulva* microbiome is involved in the biogeochemical cycling of elements, such as nitrogen and sulfur, enhancing nutrient availability through processes like nitrogen fixation and the breakdown of complex organic compounds. This symbiotic relationship not only supports *Ulva*'s growth and resilience but also influences its chemical environment, impacting broader marine

ecosystems (Aires et al., 2019; Kessler et al., 2018; Kong et al., 2023). Additionally, studies on the *Ulva* microbiome have shed light on the alga's capacity to resist colonization by pathogenic microbes, likely mediated through the production of antimicrobial compounds by both *Ulva* and its beneficial microbial partners (Busetti et al., 2017). Understanding the intricacies of the *Ulva* microbiome offers valuable insights into microbial ecology, symbiotic relationships, and the potential for harnessing these interactions in biotechnology, aquaculture, and environmental restoration projects.
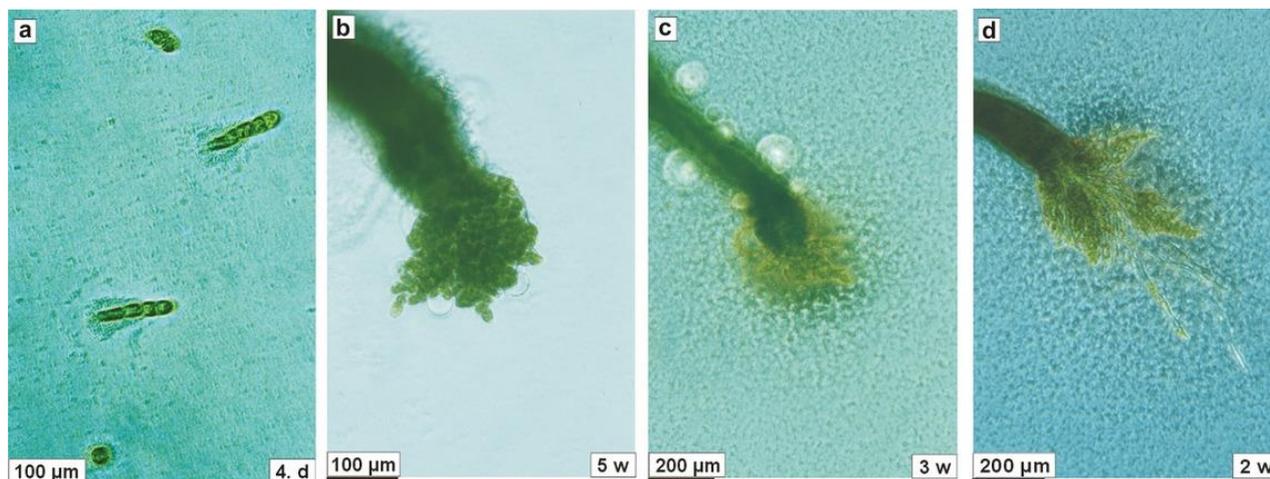
Figure 5. The figure displays the interaction between the symbiotic bacteria and the wild-type *U. mutabilis*: (a) axenic germlings of the *U. mutabilis* mutant affixed to cover glasses were put in petri dishes containing *Roseobacter* MS2 in UCM. (b) Holdfast of a wild-type germling developed in the presence of *Roseobacter* MS2 (separated from the substrate without the bacterial biofilm). (c) Similar to (b), but still connected to the substrate (containing the bacterial biofilm). (d) Similar to (c), but grown in the presence of *Roseobacter* MS2 and *Cytophaga* MS6. Figure by (Spoerner et al., 2012b).

2.2.5. Phenomics

Phenotyping refers to the process of accurately and precisely measuring the phenomes, the physical and biochemical traits of individuals as they interact with the environment. In land plants, this includes traits such as height, leaf size, flowering time, and yield. In seaweed, the common traits are thalli size, metabolite production, blade length/width and wet weight. Phenotyping is a critical step in understanding the expression of genetic information (genotype) in varying environmental contexts, thereby allowing for the elucidation of the complex interactions between genes and the environment. In land plants, advances in technology have led to high-throughput phenotyping, which uses (semi-)automated and non-invasive methods to measure a wide range of plant traits more efficiently (Figure 6). This is crucial in identifying traits of interest for crop improvement and in understanding the genetic basis of these traits. However, high-throughput automated methods for seaweed are not developed yet. In microalgae, the automated process exists but only to quantify biomass as the measurement of growth (Calmes et al., 2020).

Efforts to make an automated phenotyping platform for seaweed are ongoing and some have interesting result. A machine learning algorithm has been developed to use NIR spectroscopy on seaweed. However, this is only to measure the protein content of the seaweed and lacking other important phenotypic traits (Tadmor Shalev et al., 2022). Similar effort is done to quantify the xylose, rhamnose, glucuronic acid and glucose contents (Shefer et al., 2017). Other than that, phenotyping in seaweed has been done manually. For example, an exploration of *Phyllospora* in

Australia measures the thallus circumference, frond width, primary length, wet weight, number of vesicles and other traits by manual measuring method (Wood et al., 2022). Even in a breeding program, seaweed morphological measurements are done manually and after that the data are analyzed by statistical analyses like principal component analyses (Camus et al., 2018). This approach is satisfactory for small sample size phenotyping. In mapping population, segregating line number can reach 100 to 300 or even more. In this sample size, automated approach is more advisable.

Figure 6. Overview of High-throughput Land Plant Phenotyping Platforms: (A) Small-scale: At Palacký University, Czechia, the XYZ PlantScreen™ chamber uses top-view RGB imaging for screening biostimulants in *Arabidopsis*. (B) Medium-scale: The Modular System includes advanced imaging technologies for detailed phenotyping and cultivation of mid to large plants. (C) Mobile unit: The James Hutton Institute's Phenomobile, equipped with hyperspectral imagers, specializes in fruit trees and bushes. (D) Large-scale: The autonomous PlantScreen™ Field System, mounted on a robotic arm, features comprehensive sensors for extensive field phenotyping. Figure by Rouphael et al., 2018.

A semi-automated or computer assisted phenotyping of seaweed measurements and morphological has been done. In the past, it is limited to image analysis with a software help. The researcher still manually takes the picture and manually operates a computer to use a software to generate the data (Monro et al., 2007). High-throughput automated system have been described recently, where *Ulva* discs are cut from the thalli and imaged over time to measure growth rates (Fort et al., 2019). The system is similar with the system developed in this study. In this experiment, an automated high-throughput phenotyping platform originally developed for *Arabidopsis* is used (Dhondt et al., 2014). Some adjustments are made in order to generate the data (see the methodology section). This is limited to measure the growth rate or growth speed (thalli size over time), while other traits are still documented manually with a camera and microscope.

2.3. Genomic analysis techniques in seaweed

2.3.1. Mapping populations

Mapping populations are foundational to the study of genetic traits and their inheritance patterns. These populations are specifically created and used by researchers to identify genetic markers associated with phenotypic traits. The creation of a mapping population involves crossing two parents with distinct phenotypes to produce progeny, which are then analyzed to study how traits segregate and to identify the underlying genetic bases of these traits. There are several types of mapping populations, including F2 populations, backcrosses, recombinant inbred lines (RILs), and near-isogenic lines (NILs), each serving different purposes in genetic research. Their differences are described in Table 2. Mapping populations are critical in plant and seaweed breeding and genetics because they enable the identification of genetic loci associated with desirable traits, such as yield, disease resistance, and drought tolerance.

Table 2. Different types of mapping population.

Mapping Population	Origin	Generations	Genetic Composition*	Typical Use
F2 Population	Derived from F1 hybrids between two distinct parents.	Second generation after the initial cross.	Mix of homozygous and heterozygous genotypes at each locus.	Initial genetic mapping and analysis of trait segregation.
Recombinant Inbred Lines (RILs)	Developed by selfing or sibling mating from F2 generation onwards.	Many generations (often 6-10) to achieve homozygosity.	Nearly homozygous at all genetic loci; different lines capture different recombination of parental alleles.	Detailed genetic mapping, especially for quantitative traits; reusable for multiple studies.
Backcross Population	Involves backcrossing an F1 hybrid to one of the original parents (or a parent-like genotype).	Typically, one or two generations of backcrossing.	Primarily the genetic background of the recurrent parent with some introgression from the other parent.	Mapping specific traits and introgression of particular alleles into a desired genetic background.
Doubled Haploids	Produced either through chemical treatment of microspores or by culture of anthers or unfertilized ovules.	Single generation as haploids are immediately doubled.	Completely homozygous at all loci.	Precise and rapid genetic analysis; useful in plant breeding for creating uniform lines.
Near Isogenic Lines (NILs)	Typically developed by repeated backcrossing to a recurrent parent, selecting for a specific trait from the donor parent.	Multiple generations of backcrossing with selection.	Mostly the genetic background of one parent with one or few traits/genes from the other parent.	Fine mapping of specific traits; studying gene function and interaction.

* For general diploid / polyploid organisms. *Ulva* mapping population is gametophytes (haploid).

2.3.2. Genotyping

Genotyping is the technique of examining an organism's DNA sequence to ascertain its genetic composition. In the context of plant genetics, genotyping is used to identify variations in the

genetic code that are responsible for the diversity in phenotypic traits observed within and between plant species. This is achieved through various molecular biology techniques, including PCR (polymerase chain reaction), SNP analysis, and WGS. Genotyping allows researchers to link specific genetic variations to phenotypic traits, facilitating the identification of genes that control important characteristics. This information is invaluable in breeding programs, where genotypic data can be used to predict phenotypic performance under different environmental conditions, thus enabling the selection of superior genotypes for improvement.

In seaweed, intra- and interspecific variability is very well observed. However, for the isomorphic haplodiplontic *Ulva* both mating types of the gametophyte and the sporophyte have the same morphologies. Genotyping is the key to differentiate species, strains and life stages. Different genotypes may have different phenotypic characteristics that aren't easily observed. For instance, research identified 11 genotypes of *wakame* (*Undaria pinnatifida)* from Japan, China and South Korea with PCR of the COIII and RubisCO genes. They found that only 2 genotypes of the 11 are found in 31 of commercial *wakame* products. While the other 9 were only found in 1 to 4 products. It suggested that genotyping can be important to maintain commercial value (Yoshinaga et al., 2014) because it ensures the consistency and quality of the products by identifying and selecting strains with desirable phenotypic traits. This can help in producing seaweed with optimal growth rates, nutritional content, and other commercially valuable characteristics, thereby enhancing marketability and consumer trust. A study found that in the giant kelp, significant genetic underpinnings for phenotypic differences are indicated by the significant correlation between morphological and genetic data (Camus et al., 2018).

Studying genotypes has ecological relevance. *Eucheuma* and *Kappaphycus* are the most cultured seaweed in Zanzibar. However, many of the cultures have an Asian origin due to an introduction in 1989. This genotype starts to grow in coral reefs and drifts toward seagrass meadows (Halling et al., 2013). In the wild, one area can have very different genotypes of seaweed while another might have little variability. A microsatellite genotyping population of *Undaria pinnatifida* in Nagasaki, Japan (native region), Dunedin, New Zealand (introduced by shipping activities) and Brittany, France (introduced for aquaculture) showed large differences between the populations. Nagasaki exhibited the lowest variability with 58% polymorphic loci. The highest diversity is observed in Dunedin where 90% of the loci were polymorphic. However, this high diversity was not observed in Brittany with only 60% of polymorphic loci, despite both are introduced. The research suggests that the introduction process, the spreading and the environmental condition may play a role in the genotype variabilities. Whole comparison of all genotyping techniques is explained in Table 3.

Table 3. Genotyping methodologies differentiation.

Method	Features	Advantages	Limitations
Microsatellites or simple sequence repeats (SSRs)	Short repeating DNA sequences	High polymorphism; good for fine-scale studies	Labor-intensive; requires prior sequence knowledge
Single Nucleotide Polymorphisms (SNPs)	Single base pair variations	High-throughput; accurate; scalable	Requires some genome knowledge; higher cost
Amplified Fragment Length Polymorphism (AFLP)	Digestion of DNA followed by selective PCR amplification	Generates many markers; no prior sequence knowledge needed	Labor-intensive; not easily comparable between studies

Restriction site-associated DNA Sequencing (RAD-Seq)	DNA fragments adjacent to restriction enzyme sites are sequenced	Efficient for discovering and genotyping SNPs	Requires careful experimental design; data management can be complex
Whole-genome sequencing	Sequencing of the entire DNA	Most comprehensive; facilitates detailed studies	Expensive; requires significant computational resources
Targeted Sequence Capture	Enrichment of specific genomic regions before sequencing	Focuses on relevant regions; efficient for specific genes	Setup cost for initial probe design; limited to targeted regions

These genetic insights gained from genotyping are crucial for further applications in QTL mapping. QTL mapping is a method used to identify regions of the genome associated with specific phenotypic traits. By linking genotypic data to phenotypic traits, researchers can pinpoint the genetic bases of important characteristics, such as growth rates, stress resistance, and nutrient content in seaweeds. This bridging of genotypic and phenotypic data through QTL mapping is essential for advancing our understanding of complex traits and enhancing breeding programs to develop superior seaweed varieties. It is important to delve into the methodologies and significance of QTL mapping in seaweed research.

2.3.3. QTL

Mapping populations in green algae have not been reported yet. Meanwhile, several studies exist in red and brown macroalgae. For example, Shan et al. (2015) developed a high-density genetic map for the brown alga *Undaria pinnatifida* using the specific length amplified fragment sequencing (SLAF-seq) technology. This research aimed at addressing the scarcity of genetic and genomic information for this economically significant species in East Asia. A key finding was the identification of a major sex-associated QTL on linkage group 22, pinpointed to a region flanked by 68 SLAF markers. This region, along with the SSR marker and five other SLAF markers, was closely linked to the sex phenotype, suggesting it as the sex-determining region. This study marks the first construction of a high-density genetic linkage map for *Undaria pinnatifida*, offering a foundational tool for further research in this species and other kelp species (Shan et al., 2015).

In a 2018, QTL mapping and candidate gene screening associated with blade length and width was reported in *Saccharina japonica*. Using SLAF-seq, the team developed a high-density genetic linkage map with 7,627 SNP markers distributed across 31 linkage groups. This map facilitated the identification of 12 QTLs for blade length and 10 QTLs for blade width. Notably, some QTL intervals were relevant to both traits, indicating a genetic correlation. Following the QTL mapping, 14 Tic20 genes and three peptidase genes were associated with these yield-related traits. This study provides valuable markers for breeding efforts aimed at improving yield traits in this economically important seaweed (X. Wang et al., 2018). A follow-up study aimed to enhance the efficiency of cultivar development through marker-assisted selection (MAS). Utilizing 167 SSR markers and a mapping population of 125 segregating lines, they constructed a genetic linkage map with 27 linkage groups. This map was then integrated with *S. japonica*'s genome sequences to align with 23 of 31 pseudo-chromosomes. A significant QTL, closely associated with the SSR marker M1895, was identified. This QTL harbored the SjPT gene, coding for a high-affinity

phosphate transporter, potentially critical for kelp blade development. Transcriptional analysis confirmed the gene's involvement in blade size, offering a valuable marker for MAS in *S. japonica* breeding programs. This research underscores the utility of genetic mapping and QTL analysis in understanding and manipulating economically relevant traits in seaweeds (X. Wang et al., 2023).

Xu et al. (2015) constructed the first high-density genetic linkage map *for Pyropia haitanensis*, a marine macroalgae of significant economic importance in China. This high-density map was employed to identify QTLs associated with six economically important traits: frond width, length, growth rates, fresh weight, and thickness of both frond length and fresh weight. Fifteen QTLs were identified, with phenotypic variance explained (PVE) by segregating line QTLs ranging from 9.59% to 16.61%. This detailed mapping provides a robust platform for further genetic studies, offering insights into the genetic basis of key economic traits and facilitating marker-assisted breeding in *P. haitanensis* (Xu et al., 2015). To further elucidate the genetic architecture of important traits in seaweed, GWAS have been increasingly employed.

2.3.4. GWAS

Understanding the genetic underpinnings of traits such as growth rate, temperature tolerance, and nutrient content is vital for the sustainable exploitation and conservation of seaweeds. GWAS present a robust approach to uncover genetic variants linked to such traits, providing insights that are pivotal for selective breeding and conservation efforts. GWAS is a research approach used to identify genetic variants associated with specific traits or diseases. It involves scanning the genomes of many individuals to find genetic markers that occur more frequently in those with a particular trait or condition compared to those without it.

GWAS in green seaweed has not been done much compared to red and brown seaweed. This is maybe due to red and brown seaweed having more species that are commercially important and have been used a lot in many industries. For instance, the absence of consistent methods for measuring traits and collecting phenotypic data, combined with limited knowledge of genes influencing economically important traits, has hindered the breeding and variety improvement of the industrial algae *Gracilariopsis lemaneiformis*. In response, a study created a dependable method to assess the linear growth rate (V_L) and branching by using tip culture. Using this method, they conducted a GWAS on 174 haploid strains, identifying 483,670 SNPs. Notably, 68 SNPs were linked to V_L and 12 to branching. Additionally, within 20 kb of these SNPs, they identified 79 genes associated to V_L and 16 to branching. Validation of SNP loci Chr3–969,142 and Chr13–2,948,547 confirmed their involvement in trait regulation. Further analysis showed a link between V_L and genes related to the cell cycle and ubiquitin pathways, suggesting their roles in trait development. This study lays a foundation for future genetic research and breeding programs for *G. lemaneiformis* (Feng et al., 2023).

Similar to genotyping one or more marker genes, GWAS can also be used to study how the seaweed introduction, dispersal and cultivation efforts influence it genetics. Graf et al. (2021) did a GWAS to study the effect of human introduction of *Undaria pinnatifida* and they found that the natural, cultivated and introduced population are genomically different. They also provide some insights about the effect of introduction and cultivated on genomic scale. Introduced populations are identified from France and New Zealand. Despite both populations being introduced, they found

that the population in France had lower genetic diversity compared to the population in New Zealand (Figure 7). The population in New Zealand has high genetic diversity due to multiple introductions from the shipping industry. Meanwhile, the introductions in France have occurred less frequently since it was originally meant to support the oyster aquaculture.
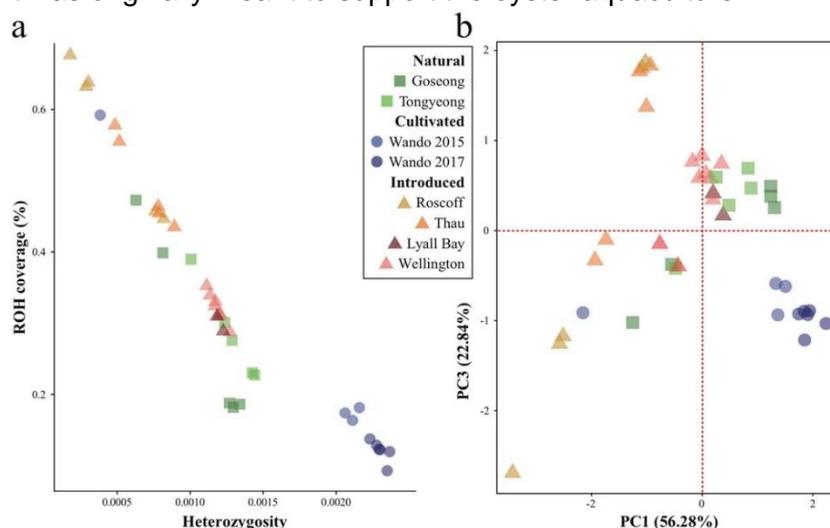

Figure 7. (a) The average genome-wide heterozygosity as a function of the run of homozygosity (ROH) coverage for *Undaria pinnatifida* populations. Natural populations are represented by squares (Goseong, Tongyeong), cultivated populations by circles (Wando 2015, Wando 2017), and introduced populations by triangles (Roscoff, Thau, Lyall Bay, Wellington). (b) Principal Component Analysis (PCA) based on the total length, average length, and number of ROH for the same populations. The PCA differentiates natural, cultivated, and introduced populations, illustrating genomic distinctions influenced by human introduction and cultivation practices. Figure by Graf et al. (2021).

The successes of GWAS studies in red and brown seaweeds can be replicated in green seaweeds. GWAS study using fast-growing versus slow-growing *Ulva* strains can offer insights into the genetic determinants of growth rates in seaweeds. The potential identification of genetic markers that are associated with growth could pave the way for breeding programs aimed at enhancing the productivity and sustainability of *Ulva* cultivation. This study is particularly compelling as it addresses a fundamental biological trait, growth rate, whose variation could be crucial for monitoring green tide formation with growth-associated markers. This approach provides a model for similar research in other less-studied seaweed species, influencing both ecological dynamics and the economic viability of seaweed farming.

In lab settings, all the seaweed, especially *Ulva*, are grown in the same medium, same place, and under identical conditions, effectively removing environmental variability. Despite this controlled environment, the seaweed still exhibits a wide range of phenotypes. By studying the genotype and linking it to the phenotype in the lab, we can more directly attribute variations in phenotype to genetic causes. This provides a strong foundation for understanding the genotype-phenotype link before studying it in natural populations, where environmental effects play a role. These foundational insights pave the way for advances in culturing *Ulva*, enabling more controlled and efficient cultivation practices that can be optimized based on genetic information.

2.4. Advances in culturing *Ulva*

Recent advancements in culturing and genetic techniques have significantly enhanced the potential for utilizing *Ulva* in various industries, including bioenergy (Bruhn et al., 2011), bioplastics (Helmes et al., 2018), and functional foods (Fort et al., 2024). Culturing *Ulva* requires specialized media that replicate their natural seawater environment to optimize growth rates and biochemical composition. Adjustments in the concentrations of minerals, vitamins, and other essential nutrients are crucial for promoting healthier and more robust *Ulva* strains. For instance, the *Ulva* Culture Medium (UCM) is fully synthetic and optimal for culturing most Ulvophyceae (Stratmann et al., 1996).

While optimized culture media like UCM are ideal, many *Ulva* species and other seaweeds can also grow in filtered or autoclaved seawater (Bruhn et al., 2011). This can be particularly useful for research where standardized culture media are not yet available or when large volumes are needed to reduce costs. Studies have shown that the culture media significantly impact seaweed growth and performance, emphasizing the need for standardized media in other seaweed research. For instance, it has been shown that the decreased pH (7.8) and the increased pH (9.0) in seawater inhibited the nitrogen uptake and nitrogen reductase in *Pyropia haitanensis*, highlighting the importance of maintaining the pH level in *P. haitanensis* research (H. Jiang et al., 2018).

Culture techniques for commercial purposes have been well developed and documented, resulting in several handbooks and FAO manuals (Foscarini & Prakash, 1990; Redmond et al., 2014). However, these resources are often limited to commercially important seaweed species that are relatively easy to culture. Fortunately, *Ulva* falls into this category. In coastal systems, *Ulva* is often grown in open sea settings using raft systems or ropes. These setups are also used in more offshore conditions where the seaweed can benefit from the nutrient-rich waters (Castelar et al., 2014). Additionally, *Ulva* is cultivated in land-based systems, such as ponds and raceways (Neori et al., 2020), which provide controlled environments that can optimize growth. Moreover, *Ulva* is an integral part of Integrated Multitrophic Aquaculture (IMTA) systems, where it is grown alongside other aquatic organisms. This method not only enhances the overall productivity of the aquaculture system but also contributes to environmental sustainability by utilizing the waste products from other species as nutrients for *Ulva* (Ben-Ari et al., 2014).

Other emerging important seaweed species, may face constraints in commercial production and scalability. For instance, the red seaweed *Asparagopsis* has become a trending topic due to the production of a metabolite that reduces the methane emission in cows (Roque et al., 2019). However, the process to domesticate and scale is challenging due to the size and distinct form of the algae. It is well recognized that the genus has a heteromorphic triphasic alternation of generations life cycle consisting of a filamentous diploid tetrasporophyte, an attached diploid carposporophyte, and a macroscopic haploid gametophyte (Dishon et al., 2023). While other seaweed has problem to culture, other type of seaweed culture techniques is keep being developed.

Advancements in controlled environment systems, such as photobioreactors, have revolutionized *Ulva* cultivation. These systems allow precise control over light intensity, temperature, CO2 supply, and mixing rates, leading to higher biomass yields and reproducibility. A pilot-scale flat panel photobioreactor has been developed for growing *Ulva lactuca* as an alternative

to open-sea cultivation. Using poultry litter extract as a substitute nitrogen source, the highest productivity achieved was 303 g m−2 d−1 (fresh weight), which could scale up to 910 tons ha−1 yr−1 when the biomass composition matched the control medium (Mhatre et al., 2018). Additionally, another *Ulva* photobioreactor has been developed for indoor, urban-style farming, demonstrating the potential for urban sea farming (Figure 8) (Chemodanov et al., 2017).
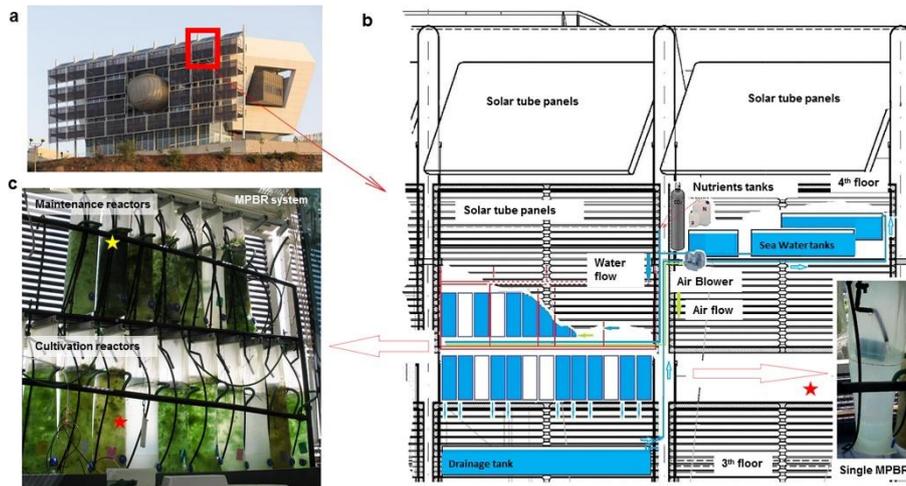
.

Figure 8. Flat panel diagram of building attached urban-style *Ulva* photobioreactor by Chemodanov et al. (2017).

III. MATERIALS AND METHODS

3.1. *Ulva* strains source and culturing method

The two lab strains used in this study are *U. mutabilis slender* and wild-type. Both are descendants from individuals isolated in Portugal in the 1950s (Steinhagen, Barco, et al., 2019) and have been maintained in labs ever since. *slender* is described as the fast-growing strain with a hollow tubular structure and having no secondary rhizoid. Meanwhile, the wild-type is slow-growing with blade like structure and well-developed rhizoids (Spoerner et al., 2012).

The lab strains and the mapping population are grown in UCM (Stratmann et al., 1996) that contains several salts, minerals and vitamins that support both the growth of the *Ulva* and the microbiome that surrounds it. UCM consist of 5 different solutions. When UCM is needed, UCM solution 1 is prepared in batches for 5-10 L and mixed with magnetic stirrers (IKA C-MAG MS 7, Staufen, Germany), then divided into 1 L glass bottles that are autoclaved at 121°C for 15 minutes. 10 mL of the UCM solution 2, 3 and 4 and 2 mL of solution 5 needs to be added to 1L of solution 1. The exact compositions of UCM solution 1-5 are described in appendix A. The UCM solution 2-5 are filter-sterilized and stored in cold room (4°C) as stocks. Leftover mixed UCM are also stored at the cold room.

50mL UCM is poured into 150 x 20 mm crystal-grade polystyrene, E.O gas sterilized petri dish plate (SPL life sciences, Pocheon, South Korea). The plates are sealed with micropore tape (3M, Minnesota, United States) and placed in the 21°C tissue culture room under light intensity of 100 µmol photons m^{-2} s^{-1} (Spectralux Plus NL-T8 36 W/840/ G13 fluorescent lamp) under a day-night rhythm of 16:8 hours.

3.2. Gamete, spores release induction and crossing method

The induction of *Ulva* gametes started by selecting plates containing mature individuals. It takes about 1-2 months for *slender* and 2-3 months for the wild-type to grow to this stage. The amount, size and density of the individuals inside the plate is also taken into account. The selected parent strains are then cut manually with a single edge blade (GEM scientific, Bentley, United Kingdom) to fragments of a few mm^2. After incubating at least 30 minutes in fresh UCM, the tissue fragments are filtered via funnel and Miracloth (475855-1R, pore size 22-25 µm, Merck milipore, Darmstadt, Germany) and 50 mL of fresh UCM is added. The culture is then placed back in the tissue culture room for 72 hours. After that, the UCM is refreshed to get rid of the swarming inhibitors. Then, the plate is partly covered with an aluminum. This will trigger the gametes natural behavior to swim towards the light. The plates are left unbothered under the light for ~1 hour. For *slender*, the accumulation of gametes towards the light can be seen with the naked eye. These gametes are then harvested simply by pipette aspiration. The accumulation of gametes for wild-type is not so clear for the naked eye. Therefore, the media is filtered again and centrifuged (5 minutes at 4000 rpm) to pellet the gametes. The gametes harvested are then counted with the hemocytometer (Fuchs Rosenthal, Labor Optik, Bad-Homburg, Germany) to know the approximate density by dilutions and observation.

Equal amounts of gametes are then pipetted on a plate and mixed. The mixed gametes are put in the tissue culture room for 1 hour, then 50 mL of UCM is added. The culture is then grown for 14 days for both strains. Then, individuals from each plate are isolated into 12 well plate (VWR tissue culture plates, Radnor, USA) to test for ploidy by genotyping methods. After identifying the

sporophytes (2n), the sporophyte is induced. The induction of spore (n) release in sporophytes is similar to the gamete induction mentioned in this section. After the spore release, a dense culture was left in the culture room to develop for about 3 weeks. Then segregating lines are manually isolated one by one to create the mapping population. The cultures of the spores are diluted by 10x to ease the isolating process. The overall process of the gamete induction is depicted in Figure 9.
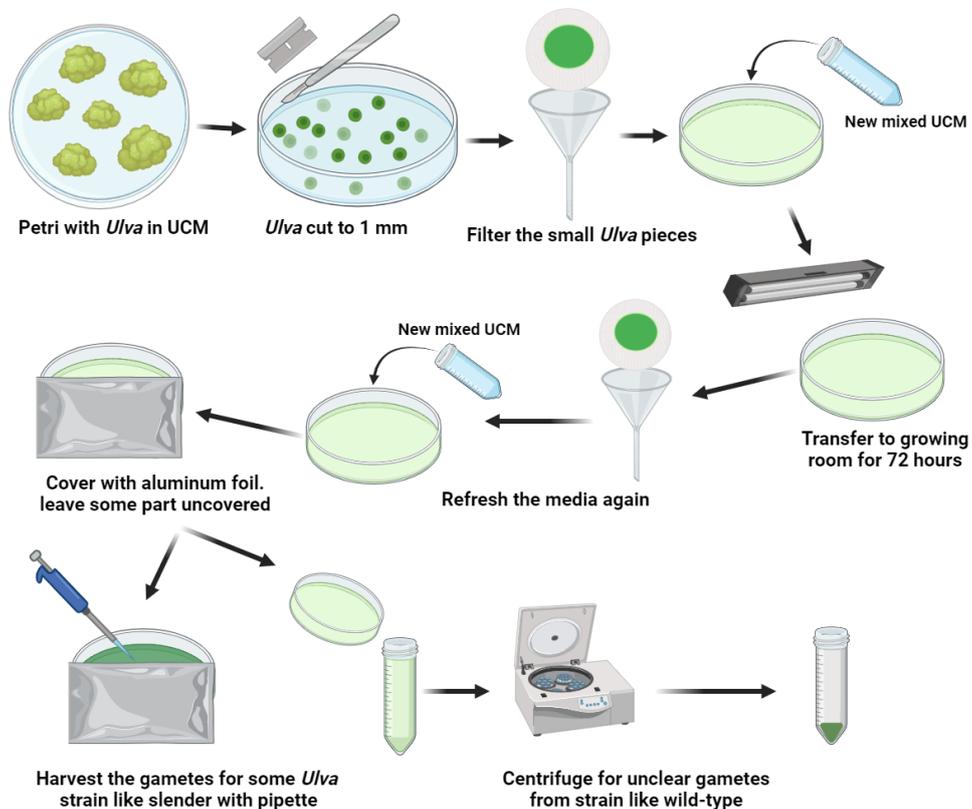
Figure 9. Workflow to induce and harvest *Ulva* gametes. Figure by author using Biorender..

3.3. Genotyping methods

Genotyping in this study was done to identify a sporophyte from *slender* and wild-type cross attempt and to acquire information about the mating type of the mapping population after the induction of the sporophytes. DNA was extracted by the CTAB method according to the adopted protocol by Van den Daele Hilde in September 2005 for the Department of Plant Systems Biology in VIB (Appendix B). The acquired DNA is in 2 mL tubes (Eppendorf, safe lock tubes, Hamburg, Germany) and stored at -20°C. For the PCR reaction, the DNA sample is mixed with the ALLin^{TM} Red Taq Mastermix (highQu, Kraichtal, Germany), forward and reverse primers and water at per sample composition below:

- Redtaq mix   : 5 µL
- DNA sample   : 2 µL
- Primer F   : 0.2 µL
- Primer R   : 0.2 µL
- MiliQ water   : 2.6 µL

Four primer pairs are to distinguish both gametophyte mating types (mt+ and mt-; Table 4). The detailed protocols for the PCR are described in the appendices (Appendix B). After PCR, the sample is loaded on an agarose gel (1,5%) containing SYBR safe (Invitrogen SYBR® Safe, 10,000X in DMSO; 1.5%, Carlsbad, USA). A Benchtop 1 kbp DNA ladder was used (Promega, Madison, USA) as the reference. The gel electrophoresis run for 25 minutes at 100 volts. For every PCR reaction, three controls are included; *slender* DNA, wild-type DNA and a water control.

Beside the method described above, segregating lines were also genotyped using a Phire Tissue Direct PCR Master Mix (F170S, ThermoFisher Scientific, Waltham, Massachusetts, United States). This method is used for segregating lines from the mapping population that could not be genotyped by using the RedTaq-protocol. This methodology started by cutting the small part of *Ulva* tissue (0.5 mm). Then, PCR wells are filled with the primers, the Phire Tissue Direct PCR Master Mix and H_2O. The exact composition used are as follows:

- 2x Phire PCR Master Mix   : 25 µL
- Primer F   : 2.5 µL
- Primer R   : 2.5 µL
- MiliQ water   : 20 µL

After the *Ulva* tissue is added (`0.5 mm^2), the reaction can go directly in to the PCR thermal cycler. Gel electrophoresis is performed as described above. This method eliminates the need for a DNA-extraction prior to PCR and therefore reduces the probability of contamination issues during DNA extraction (Figure 10).
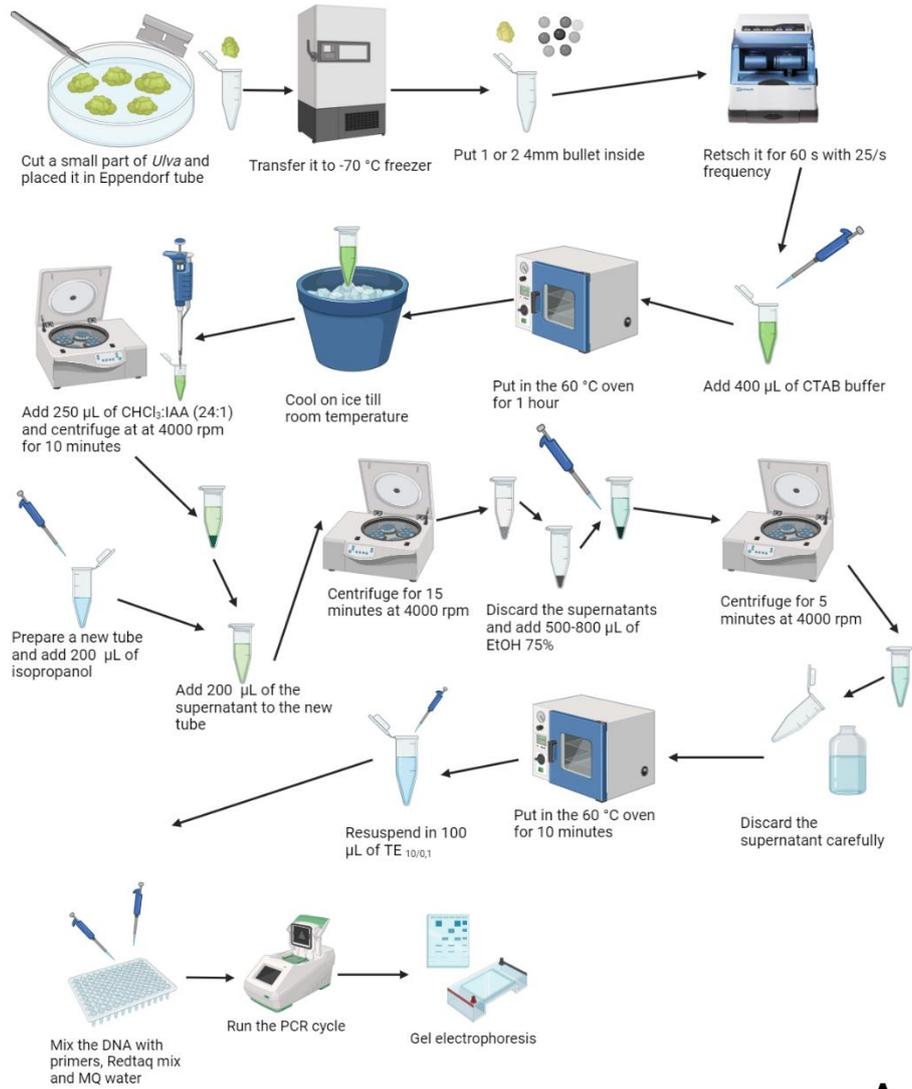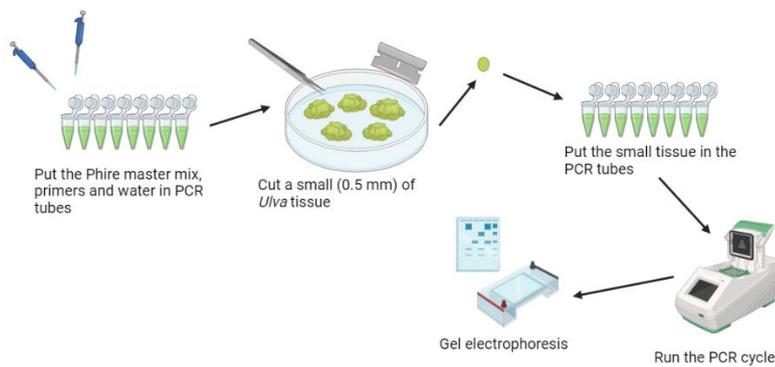
Figure 10. 2 different workflows of the genotyping done in this study. (A) CTAB DNA extraction method. (B) Phire direct plant master mix method. Figure by author using Biorender..

Table 4. 4 primers to identify mating type in this study.

Name	Annotation	5'-3'
SL019_0090-1F	Mt+ #1	GACARATGYATWTACAGCCA
SL019_0090-2R		GCYTCTGCTTGYTSYTGATC
UMSL025_0049F	Mt+ #2	ACAAATTCTCGGTCTAACG
UMSL025_0049R		GAACCCACTGAACTCCTCT
WT_29-1F	Mt- #1	GCCWCCAAGATCCAAATG
WT_29-1R		CATGTCTTTCTCACACTCATG
UMWT029_0065F	Mt- #2	CATTTCGCCAGCTTACAC
UMWT029_0065R		CACTCCATCTATCAGGGTCA

3.4. Phenotyping

Phenotyping is important to measure the different phenotypes of the mapping population that can be linked to the genetic background. A binocular stereomicroscope (Leica MZ16, Heerbrugg, Switzerland), facilitated by mount camera (Toupcam UCMOS14000KPA, ToupTek photonics, Hanzhou, China) with illuminator (KL 1500 lcd, Schott, Mainz, Germany) and direct observation is used to characterize the morphology of the mapping population such as the presence of rhizoids and overall structure. Furthermore, the IGIS platform was used to phenotype the growth rate of the mapping population (Dhondt et al., 2014). The IGIS platform consists of cameras taking one picture of the *Ulva* mapping population every hour. The exact setup of the platform is described in Figure 11. The culture media for the mapping population phenotyping trial under IGIS is modified form the typical mixed UCM solution. The media also contains 0.2% of plant tissue culture agar (Lab M limited, Manchester, UK). This is to ensure that the *Ulva* is not moving during the phenotyping process.

Figure 11. IGIS platform general setup. PLC: Programmable logic controller. LED: Light emitting diodes. NIR: near infrared. Figure by author using Biorender.

The selection of segregating lines to be put under the IGIS platform starts by selecting the mapping population that contains a very dense small culture growing in the base of the plate. This is because this dense new population is usually uniform and are in the relatively same age (Figure

12). This dense population is scooped with a sterilized tweezer (Dumont Swissmade tweezer, Montignez, Switzerland) then transferred into a new Petri dish with UCM. After 1 week, the segregating lines have grown sufficiently and are transferred to the agar containing UCM plate. The organization of these new segregating lines follows a uniform based placement (Figure 13). The segregating line is placed on the media and not inside of it. The plate is then placed on the IGIS platform and runs for 4 weeks. The same subculture also used for rhizoid observation, meanwhile thalli observation used the original mapping population culture which is quite dense and old. This is done because not all segregating lines form this dense new population for subculture. Moreover, almost all structure while they're young forms a tube.
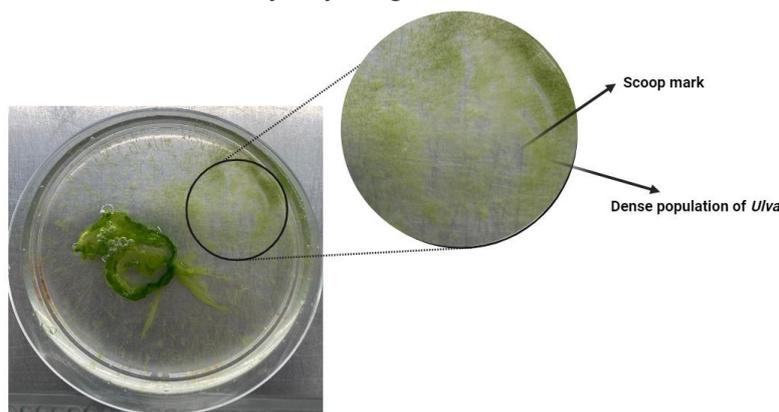
Figure 12. Example of dense population used for phenotyping with IGIS platform. (*Ulva* in the picture is from mapping population, segregating line number 64). Figures by author using Biorender.
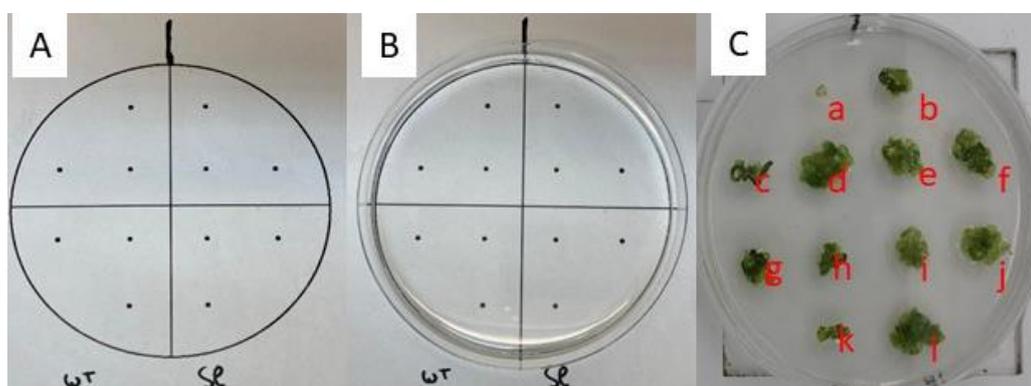
Figure 13. Template of *Ulva* placement in the plate for IGIS phenotyping. A single individuals is placed on each black dot. 6 individuals used for each segregating line (A) The template on top of a paper alone. (B) The template with plate containing agar UCM placed above. (C) *Ulva* segregating line plant labeling and placement on the agar UCM.

The pictures from the IGIS platform are analyzed with ImageJ software (Abràmoff et al., 2004)s. The IGIS platform was developed for *Arabidopsis* imaging, where the data analysis of the size actually relies on a NIR (Near InfraRed) filter. The pictures generated by NIR show a very nice contrast between the plate and the *Arabidopsis* tissue. However, this is not the case for *Ulva* since under NIR, *Ulva* appears darker and lacks contrast with the plate, making analysis difficult. The IGIS setup was adjusted to abandon the NIR filter in order to work with *Ulva*. This makes the data acquisition during the night impossible. The exact duration skipped is 8 hours and the last day picture was connected to the next one as 8 hours apart pictures. Normal pictures are used and to be able to analyzed the size, the pictures need to be turned into black and white. This black and

white picture shows a nice contrast between the plate and the *Ulva* tissue. However, turning a picture into black and white creates a new problem. Sometimes, not the entire tissue of the *Ulva* actually turns into white. Making an unconnected tissue that disrupt the analysis. Vice versa, areas of the agar near the *Ulva grown* can also turn yellowish that in some cases also turn into white and detected by the analysis. The process of the *Ulva* analysis under IGIS platform and problem with analysis are described in Figure 14. The way to deal with this unconnected *Ulva* picture was by simply removing the pictures.
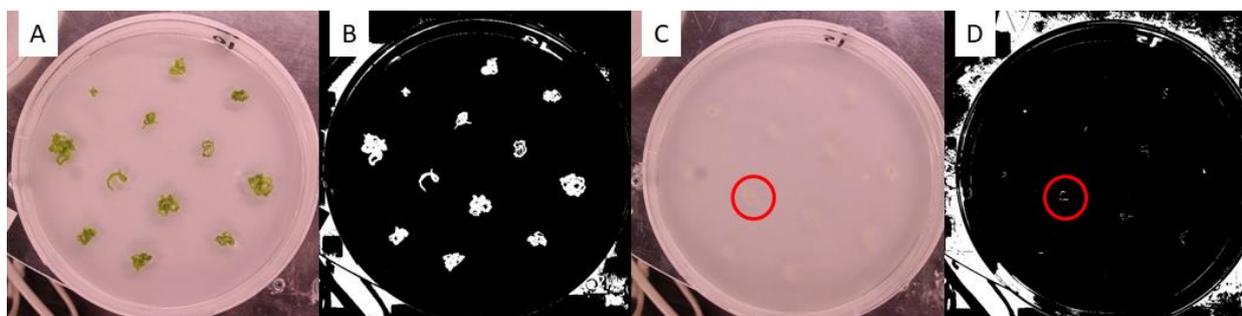
Figure 14. A. picture taken from IGIS platform. B. The same picture turned into black and white for analysis with ImageJ. C and D. Example of an unconnected black and white picture of one segregating line of the *Ulva* caused problem with analysis.

The data generated from this pipeline is in pixels initially. It is then converted to mm^2 following a ratio of 21.83 pixels / mm for IGIS platform one and 17.21 pixels / mm for IGIS platform two. They have different ratio due to them having different camera as well. The ratio obtained from the pixels of the camera compared to the width of the petri which is 150 mm. The conversion follows a simple calculation follows:

1. Calculate the conversion factor squared:

$$Conversion\ factor^2 = \left(\frac{1\ mm}{pixel\ to\ mm\ ratio}\right)^2$$

2. Convert the area from pixels to mm^2:

$$Area\ in\ mm^2 = Area\ in\ pixels \times \left(\frac{1\ mm}{pixel\ to\ mm\ ratio}\right)^2$$

For example, we will use the 21.83 pixels/mm ratio and 86591 pixels. Plug in the values:

$$\left(\frac{1}{21.83}\right)^2 = (0.0458)^2 = 0.0021\ mm^2/\ pixel^2$$
$$Area\ in\ mm^2 = 86591\ pixels \times 0.0021\ mm^2/\ pixel^2 = 182\ mm^2$$

3.5. Simplex and Multiplex PCR

Since multiplex PCR on *Ulva* DNA samples is a new technique in the lab, we designed a proof-of-concept experiment in this thesis. The primers for the amplicons were designed by SMAP-design (Develtere et al., 2023) and visualized using the Geneious Prime (v. 2024.0) software. For this test experiment, 22 amplicons were selected that should contain few SNPs between wild-type

and *slender*. The 22 amplicons were selected within gene sequences in different contigs. The 44 primers sequence and the SNPs detail are listed in appendix C. First, a simplex PCR test with wild-type and *slender* DNA was done. The simplex amplicons were visualized with gel electrophoresis. PCR products were purified by HighPrep PCR protocol (Magbio Genomics Inc., Gaithersburg, USA; appendix D) and analysed using Sanger sequencing (Mix2Seq; Eurofins Genomics, Ebersberg, Germany). The multiplex PCR utilized the Platinum™ Multiplex PCR Master Mix (ThermoFisher Scientific, Waltham, Massachusetts, United States). The detailed procedure is listed in appendix E and PCR set up is as follows:

- Platinum Multiplex PCR Master Mix : 25 uL
- Primer mix   : 5 µL
- GC enhancer   : 10 µL
- Sample DNA   : 5 µL (8 µL for WT, to ensure 100 ng/µL conc.)
- Water   : 5 µL (2 µL for wild-type)

After multiplex PCR, clean-up was done with the DNA Clean & Concentrator-5 kit (D4004, Zymo Reseaerch, California, USA). The procedure is listed in appendix F. We used Oxford Technology Nanopore sequencing (Eurofins Genomics, Ebersberg, Germany) for amplicon sequencing. All analyses were performed using Geneious Prime (v. 2024.0) software.

IV. RESULTS

4.1. Culturing and crossing outcomes

Ulva crossing started by inducing the gametophytes of the lab strains *slender* and wild-type. Results of the gamete induction of these two strains are different. On average, *slender* produced 13 million cells/ml. For the wild-type, after concentrating the gametes by centrifugation, we obtained a solution of 4,3 million cells/ml (Table 5).

Table 5. Cell counts for *slender* and wild-type gamete induction.

slender		
Cells / quadrant (16 squares)	Cells / mL (1/100)*	Cells / mL (actual)
27	135000	13500000
23	115000	11500000
28	140000	14000000
	Average	13000000
	Standard deviation	±1322875.66
wild-type		
83	415000	4150000
98	490000	4900000
85	425000	4250000
	Average	4433333.33
	Standard deviation	±407226.39

*1 small square corresponds to 0.0125 ul per quadrant = 16*0.0125=0.2 ul. Therefore, if we count 100 cells per quadrant, this means that this corresponds to 500 cells per µl (x5) or therefore to 500,000 cells per ml (x5000).

From the table above, we calculated how many wild-type and *slender* gametes we need in order to cross approximately the equal amount. Approximately three volumes of wild-type gamete suspension is equivalent to one volume of *slender* gamete suspension. We made three cross attempts. We mixed 5, 10 or 20 µL of *slender* gametes (roughly 65,000 cells/5 µL) with 15, 30 or 60 µL of wild-type gametes (roughly 66,500 cells/5 µL). This resulted in many segregating lines that were fully grown over the plate (Figure 15). Since selecting sporophytes visually is impossible in the dense cultures, segregating lines are scooped and diluted in a new plate. This resulted in segregating lines that grew less dense and bigger (Figure 16).
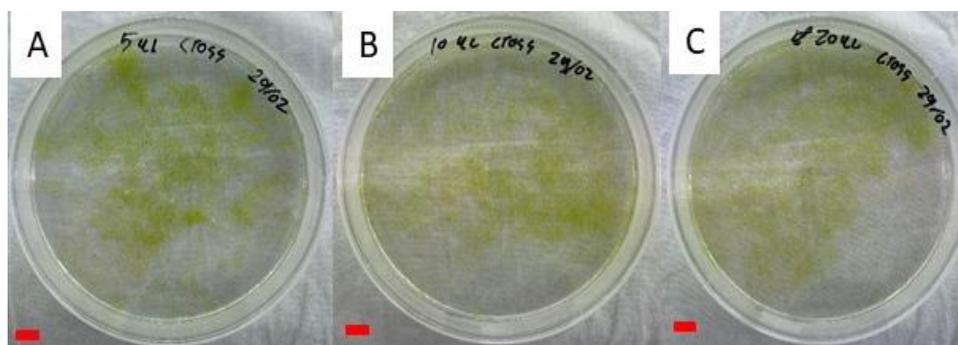

Figure 15. Original plates of three independent *Ulva* cross attempts. (A) 5 µL slender and wild-type gametes solution. (B) and (C) are 10 µL and 20 µL respectively. Red bar = 1 cm.
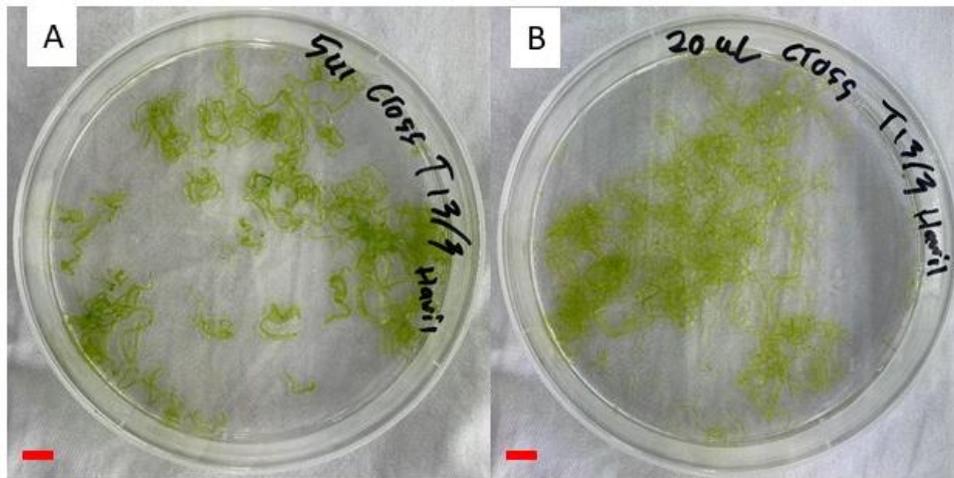
Figure 16. Selected segregating lines from original cross culture were grown in separate petri dishes. (A) from 5 µL cross between *slender* and wild-type gametes. (B) 20 µL of each gamete solutions. Red bar = 1 cm.

The putative sporophytes were subsequently placed in 12 well plates. These putative sporophytes are left to grow 1-2 weeks before genotyping test was done using two mt+ and two mt- specific amplicons. The success rate from this attempt is low. From 72 putative sporophytes (24 for each cross attempt), only 5 were identified as a sporophyte, with one individual where four markers were amplified, two individuals with three and two individuals with two amplicons of different mating type markers (Figure 17). All putative sporophytes were induced to release their spores. After the induction, UCM containing the spores is placed in the tissue culture room. After 3-4 weeks the cultures of all the induced sporophytes have grown to about 0.5 to 1 cm.

In this step, we used the same genotyping method to determine if the progeny showed different mating types, this confirms that the parent was a sporophyte and we have a mapping population. However, if they show all of mainly the same mating type, then it is from a gametophyte. In this experiment, out of 5 putative sporophytes, only one (the one with 4 mt amplicons) showed a roughly equal amount of mating types from 6 samples taken with three confirmed mt+, two mt- and one unclear mt with amplicons in all primers. Meanwhile the other 4 showed all one mating type.

While the process to obtaining mapping population were being done, genotyping and phenotyping were done on the previously generated mapping population in the lab. In total, there are 166 segregating lines. Therefore, all of the analysis done were done to this pre-existing mapping population.

In summary, the *Ulva* crossing experiments demonstrated significant differences in gamete production between the *slender* and wild-type strains, with the *slender* strain producing notably more gametes. Genotyping revealed a low success rate, with only one out of five putative sporophytes confirmed through progeny analysis. This sporophyte led to the development of a mapping population consisting of at least ~100 segregating lines. This new mapping population will support the established already in the lab mapping population of 166 segregating lines, providing a strong foundation for further genetic studies.
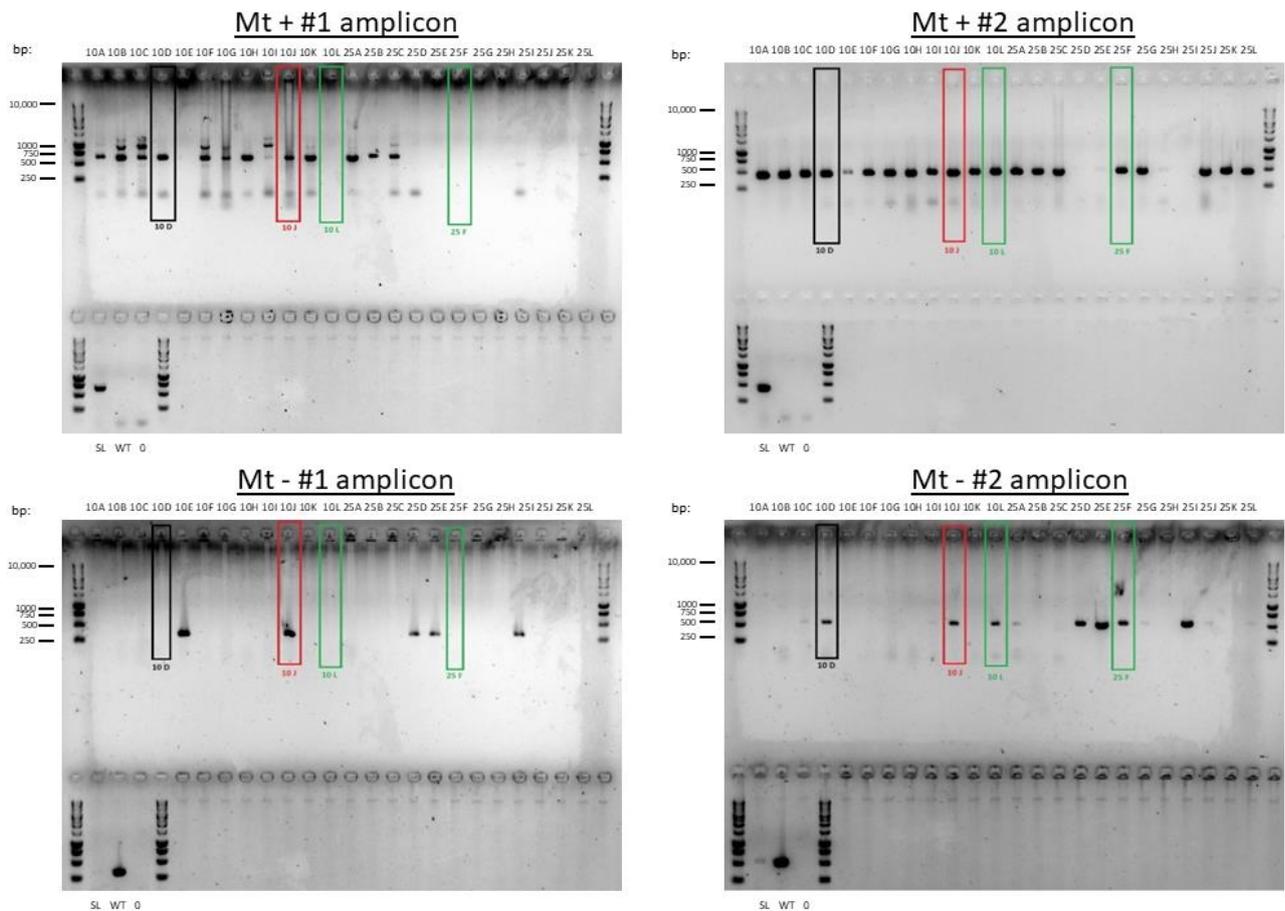
Figure 17. Gel electrophoresis result of sporophyte determination for four mating type specific amplicons. Black: 3 bands, red: 4 bands, green: 2 bands. SL contains *slender* DNA (mt+), WT contains wild-type DNA (mt-) and 0 is water control.

4.2. Genotyping results of the mapping population

A genotype refers to the genetic make-up of an organism. Here, we determined the mating type of all segregating lines of a pre-existing putative mapping population of 166 segregating lines. We successfully identified 70 segregating lines as mt+ and 85 segregating lines as mt-. Originally, unclear results were obtained for 26 segregating lines. We repeated genotyping using the Phire tissue direct PCR method and narrowed this down to 11 segregating lines where markers of two different mating types were amplified (Figure 18).

Figure 18. (Top) Pie chart of mating type distribution in the *Ulva* mapping population. (bottom) Gel electrophoresis results of segregating line number 145-166 of the mapping population with 4 different mating type primers. Red box: segregating line number 153 showing 2 bands for mt+ primers and no bands for mt- primers, indicating a + mating type. Blue box: segregating line number 155 showing 2 bands for mt- primers and no bands for mt+ indicating a – mating type. Black box: unclear mating type showing bands in all primers (marked "?"). Each number above the lane indicate the mapping population segregating lines. SL contains *slender* (mt+) DNA, WT contains wild-type (mt-) DNA and 0 is the water control.

To determine if the observed distribution of 85 mt+ and 70 mt- *Ulva* significantly differs from the expected 50:50 distribution, we performed a chi-square test for goodness of fit. The chi-square statistic is approximately 1.45. Since the p-value (0.228) is greater than the significance level (α=0.05), we fail to reject the null hypothesis. This means there is no significant difference between the observed distribution of 85 mt+ and 70 mt- *Ulva* and the expected 50:50 distribution. In summary, our genotyping approach shows that the two different mating types segregate as expected in the mapping population.

4.3. Phenotyping results

4.3.1. Morphological observations

The parental strains of the mapping population display different phenotypes. *Slender* segregating lines have a tube like structure and no fully developed rhizoid area. Thalli of wild-type

segregating lines have a blade structure and develop a rhizoid. We expect these phenotypic traits to segregate as well in our mapping population. The IGIS platform was used to assess growth rate, and the morphology was studied using a binocular microscope. The different phenotypes of the parent strains and the mapping population are shown in Figure 19.
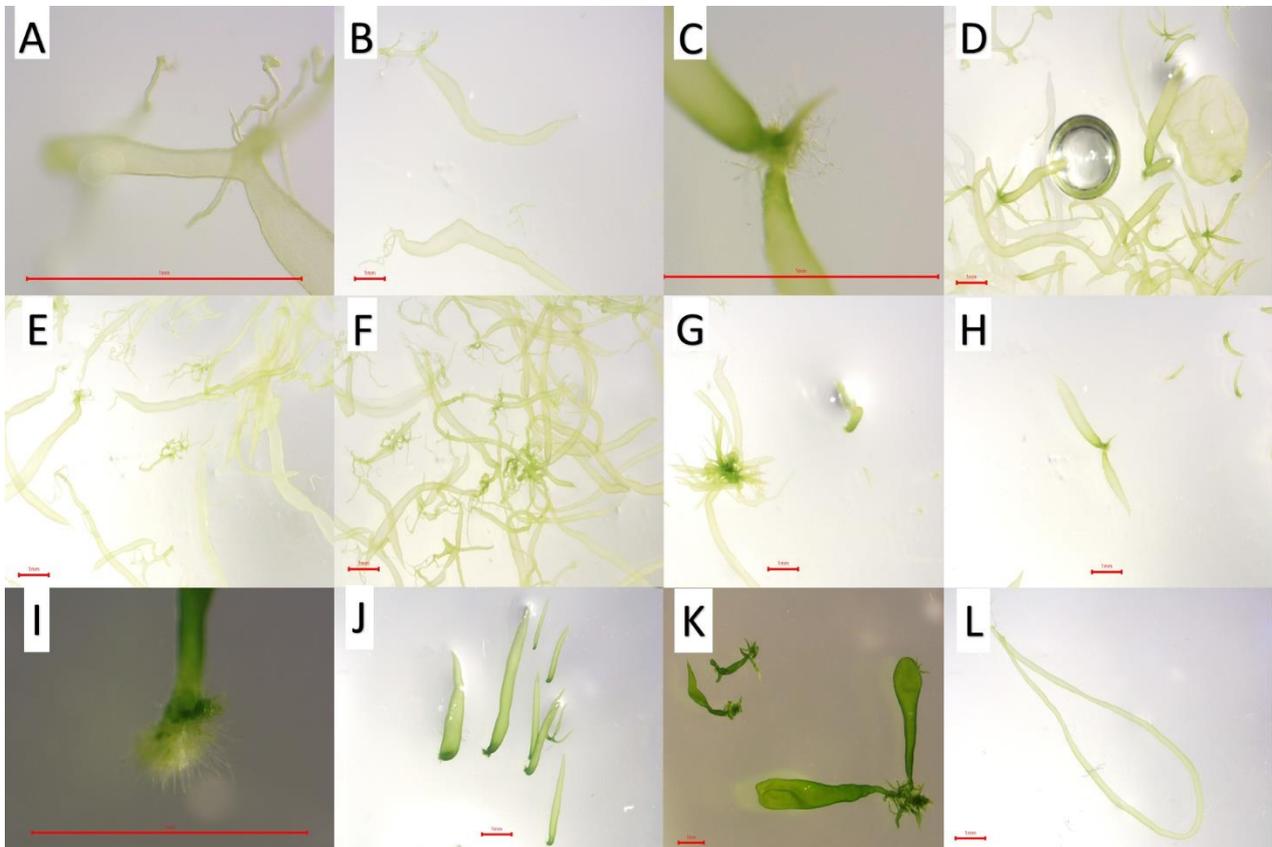

Figure 19. *Slender* and wild-type structural differences in parent strains (A-H) and in mapping population (I-L). (A-B) *slender* segregating lines. (C) wild-type fully developed rhizoid. (D) Wild-type general structure showing blade formation. E-F *slender* general morphology. (G-H) Wild-type segregating lines. (I) fully developed rhizoid of segregating line in the mapping population (no. 8). (J-K) rhizoid developed with tube-like thallus (segregating line no. 36 (J) and no. 11 (K)). (L) tubular thallus without rhizoid (segregating line no.2). All red bars = 1 mm.

We anticipate that the phenotypes in the mapping population segregate independently from the mating type if the trait is not sex-linked. We observed 72 segregating lines having a fully developed rhizoid: 31 in mt+, 36 in mt- and 5 in mt"?". 82 Segregating lines has no clear rhizoid area: 34 in mt+, 44 in mt- and 4 in mt"?". We could not determine the rhizoid phenotype in the 12 remaining segregating lines: 5 in mt+, 5 in mt- and 2 in mt"?". A chi-square test of independence was performed to determine if there was a significant association between mating type and the presence of rhizoids. The results showed a chi-square statistic of 2.569 with a p-value of 0.632, indicating no statistically significant association.

For the thallus morphology of the 166 segregating lines of the mapping population, 74 (23 mt+, 45 mt-, and 6 mt"?") displayed a tubular structure, 52 (28 mt+, 21 mt-, and 3 mt"?") were blade-like and 40 (19 mt+, 19 mt-, and 2 mt"?") displays an unclear thallus morphology (e.g., having a

flattened tube not yet wide enough as a blade or many new thalli emerging from mother thallus). Similar with the rhizoid, a chi-square test of independence was performed to determine if there was a significant association between mating type and structure type. The results showed a chi-square statistic of 7.201 with a p-value of 0.126, indicating no statistically significant association.

The link between rhizoid and the thalli structure are also analyzed. The segregating lines without fully developed rhizoid were subdivided in 53 segregating lines with tubular thallus, 23 hard to determine (ND) and 6 with blades. Meanwhile, the segregating lines with rhizoid consists of 15 segregating lines with a tubular thallus, 14 ND, and 43 with a blade structure. For the segregating lines where rhizoid is not determined, 6 have a tubular thallus, 3 a ND and another 3 have a blade-like structure. The link between rhizoid presence and the thalli structure was analyzed. A chi-square test was conducted to assess the relationship between rhizoid presence and thalli structure. The results showed a chi-square statistic of 51.543 with 4 degrees of freedom, and a p-value of 1.719×10^{-10}. Since the p-value is significantly less than 0.05, we reject the null hypothesis, indicating a statistically significant association between the presence of rhizoid and the type of thalli structure. Specifically, the data suggests that the presence of rhizoid is strongly linked to the blade structure, while its absence is predominantly associated with a tubular thallus (Figure 20).
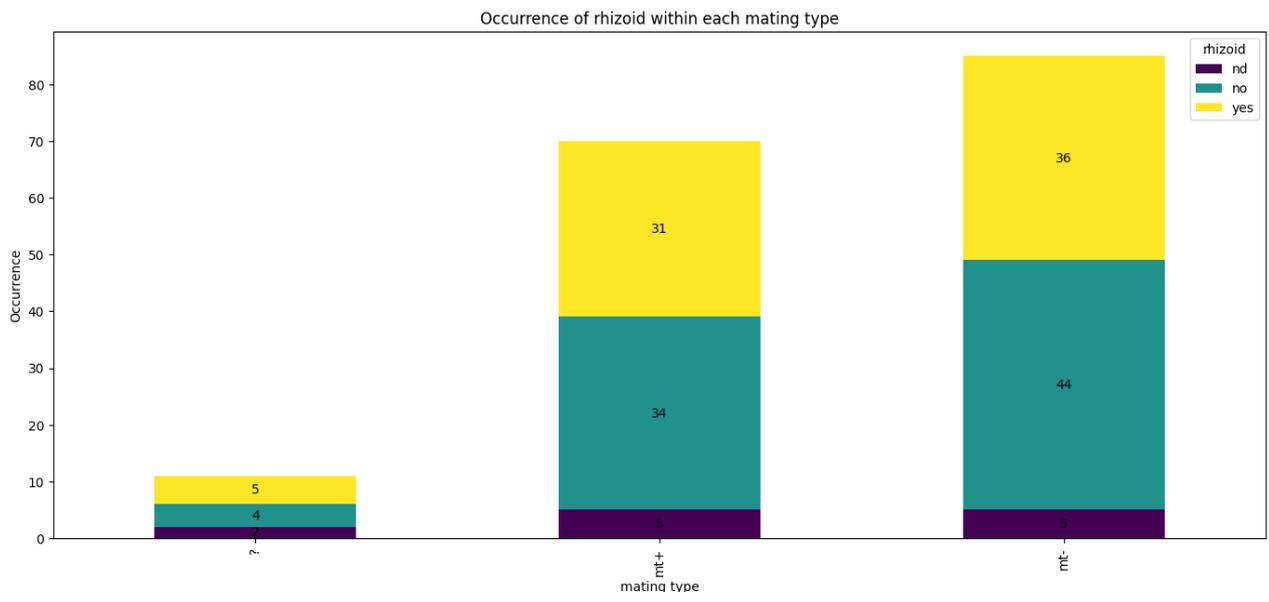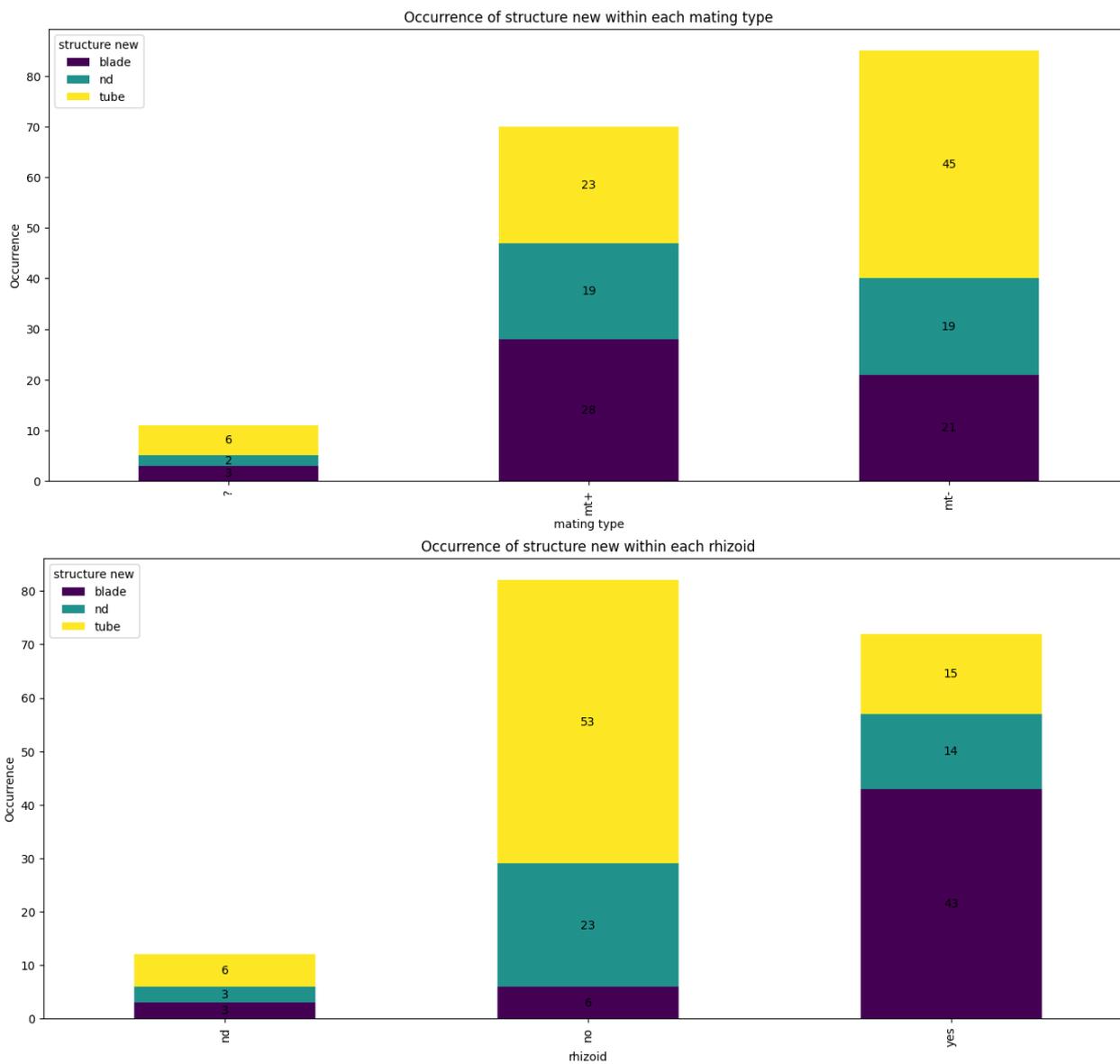
Figure 20. Occurrence of mapping population's presence of rhizoid and structure of each mating type and structure of each rhizoid condition. nd= not determined.

4.3.2. IGIS results

We used the semi-automated IGIS platform to quantify the growth rate of the mapping population. The data was analyzed with ImageJ in order to get the projected area of growth over time. 36 segregating lines from the mapping population were assessed with this pipeline with *slender* and wild-type included as controls in each run. We have imaged 36 additional segregating lines in a second imaging run, but the data was not extracted at the time of submitting this thesis. We successfully documented differences in growth of the mapping population. Unfortunately, our fast-growing *slender* control did not grow in this experiment, likely because the segregating lines were too small upon transfer to agar-containing Petri dishes.

When averaging all of the projected area of the plants per segregating line, different segregating line takes the lowest and highest spot. The lowest unexpectedly taken by our *slender* control at platform one (IGIS 1) with only 3022 pixels projected area at the end time point. This is because it was observed that the *slender* individuals were failed to grow at all. Our wild-type control grew and reached a projected area of 48801 pixels in IGIS 1 and 22484 pixels in IGIS 2. Different segregating lines have a final projected area different from the wild-type control. Segregating line number 70 was the largest with average 86591 pixels projected area at the end time points and the lowest was segregating line number 8 with 12314 pixels (Figure 21).
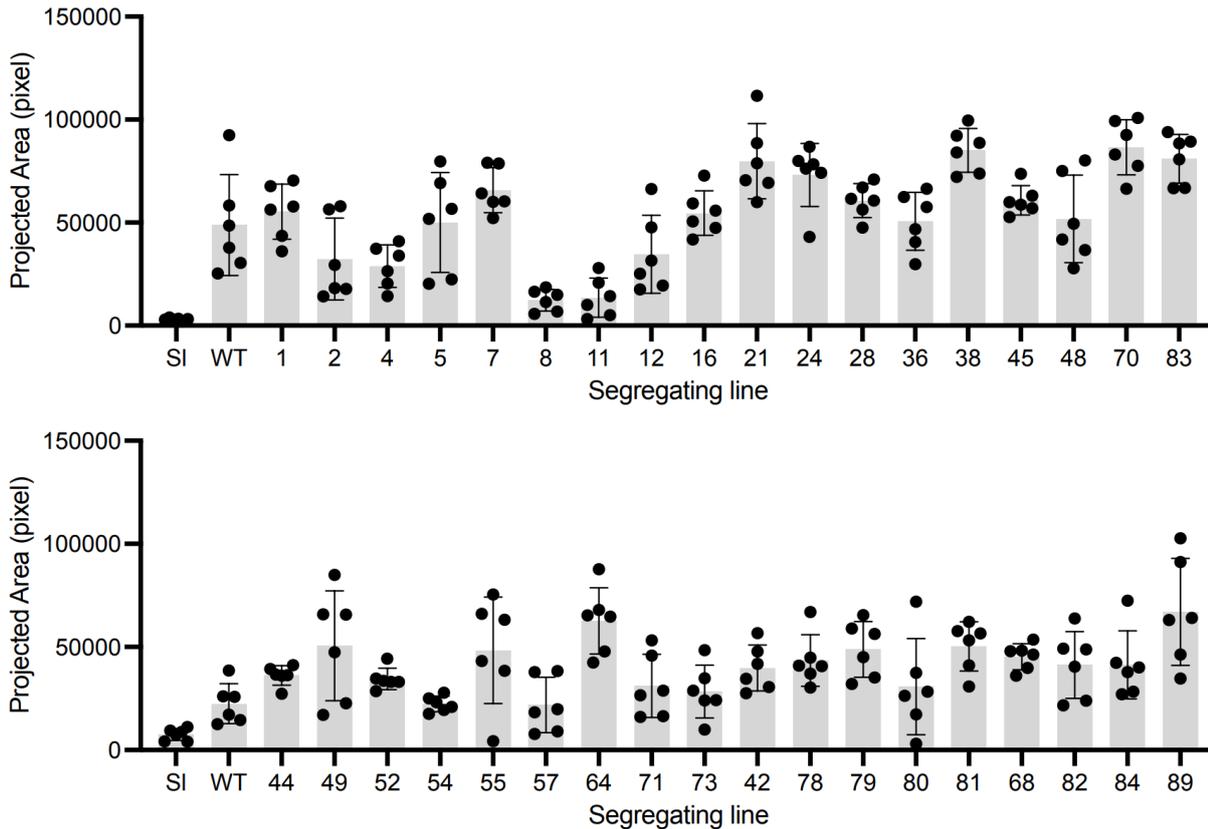
Figure 21. IGIS result showing projected area at the end time point (676 hours) of segregating lines from mapping population. Grey bar showing the average of all plants per segregating line. Black dot is the projected area of each plant. Error bar showing standard deviation (n=6).

The projected area over time also illustrates differences in growth behavior of our mapping population (Figure 22). Some segregating lines display a consistent large or small area over time increase meanwhile other seems to change the area over time increase. Segregating line number 70 displays superior area over time compared to other segregating lines. Similar consistency is shown in segregating line number 8 and 11 with consistent being the slowest area over time increase. Meanwhile, segregating line number 21 shows a relatively higher area over time at the end period and catching up to segregating line number 83 surpassing segregating line number 24. In contrast, segregating line number 45 shows a slower area over time increase at the end period allowing segregating line number 7 to surpass.
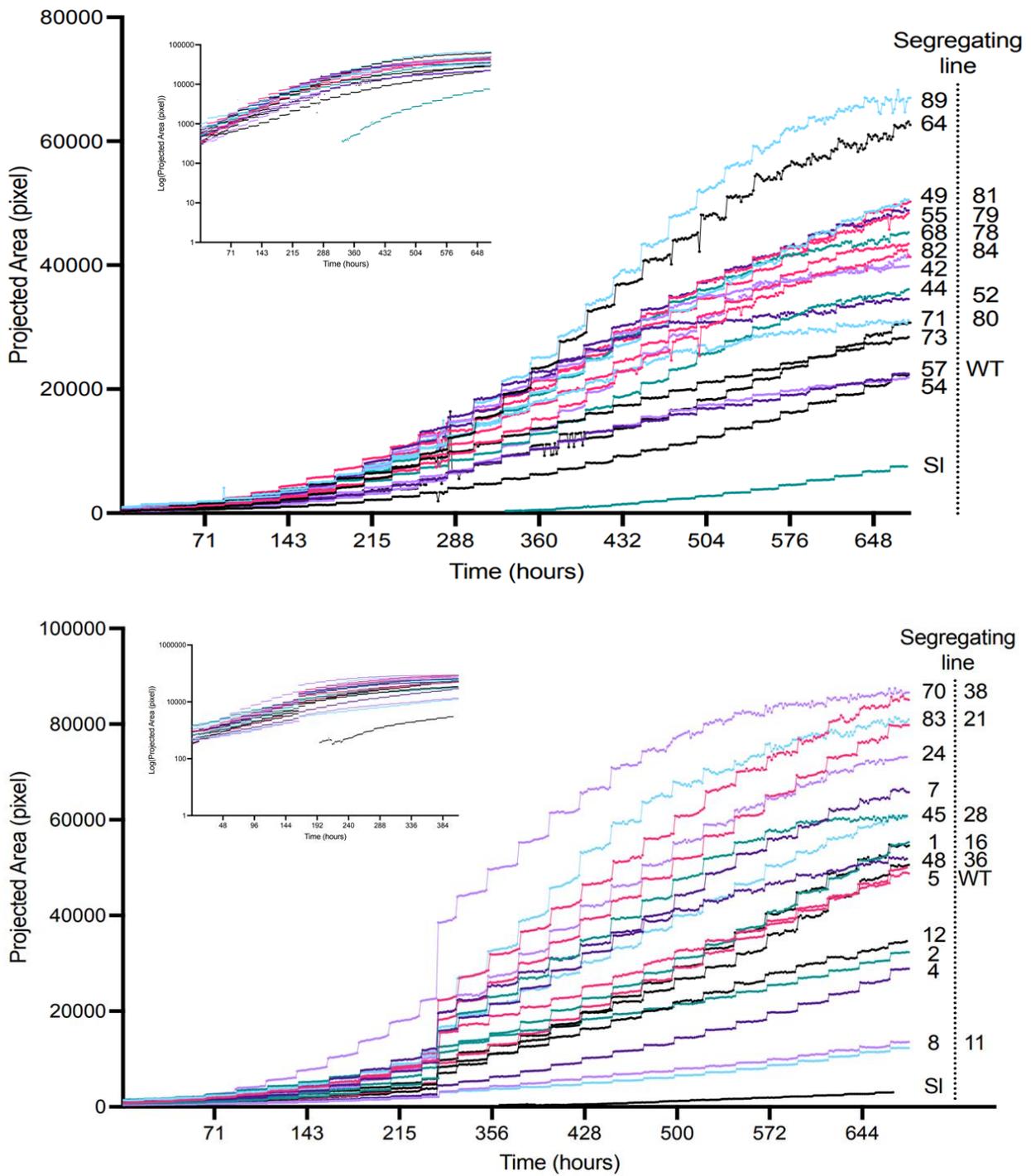
Figure 22. The projected area growth over time of each segregating line. (top) IGIS platform 1. (bottom) IGIS platform 2. Every datapoint is the average of six segregating lines, connected with a line. Insert: logarithmic transformation of the data.

For further analysis, the pixel area is converted to mm². The conversion follows the ratio of 21.83 pixels / mm for IGIS platform one and 17.21 pixels / mm for IGIS platform two. This is obtained

from the image pixel data compared to the actual width of the petri dish (150 mm). The difference of the ratio is due to them having different camera. Since the pixel/mm of the two are quite different, it is interesting to convert the data to mm^2. After converting, the highest area or size at the end time point in mm^2 changed to segregation line 89 with 227.8 mm^2. Meanwhile, the lowest was still segregation line 8 with 25.86 mm^2 (Figure 23).

As the results showed that the segregating lines have big differences at their growth projected area, it is beneficial to categorize them. In order to do that, we do K-means clustering analysis to find clear cutoff points. The results of the analysis give small-medium cutoff at 84.82 mm^2 and medium-large cutoff at 146.33 mm^2. With these cutoff points, we were able to categorize the end time point data of the IGIS results with 10 segregating lines categorized as small, 17 as medium and 13 as Large. The clustering is depicted in Figure 23.
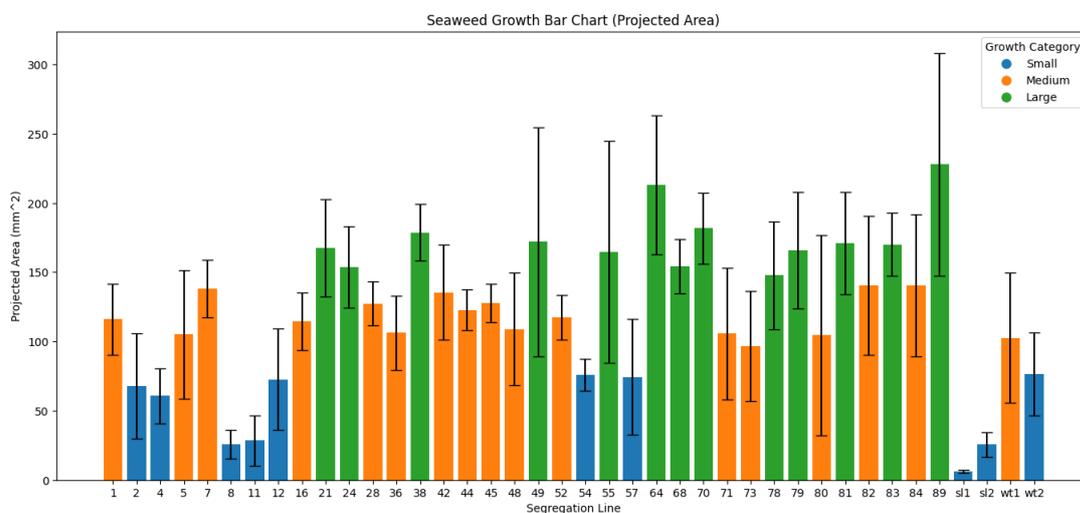
Figure 23. Clustering of the IGIS result using the end time point data. Error bars = standard deviation (n=6).

This difference in growth by projected area is interesting to see if it associated with another phenotypic data present in this study. However, important to note that not all segregation line has been putted under the IGIS platform and no replicates are being done. We did an ANOVA test for the mating type, rhizoid and thalli structure tested with the projected area data (Figure 24.). There are no statistically significant differences in the projected area across the different categories for mating type (p=0.39), rhizoid (p=0.23), and structure (p=0.56).
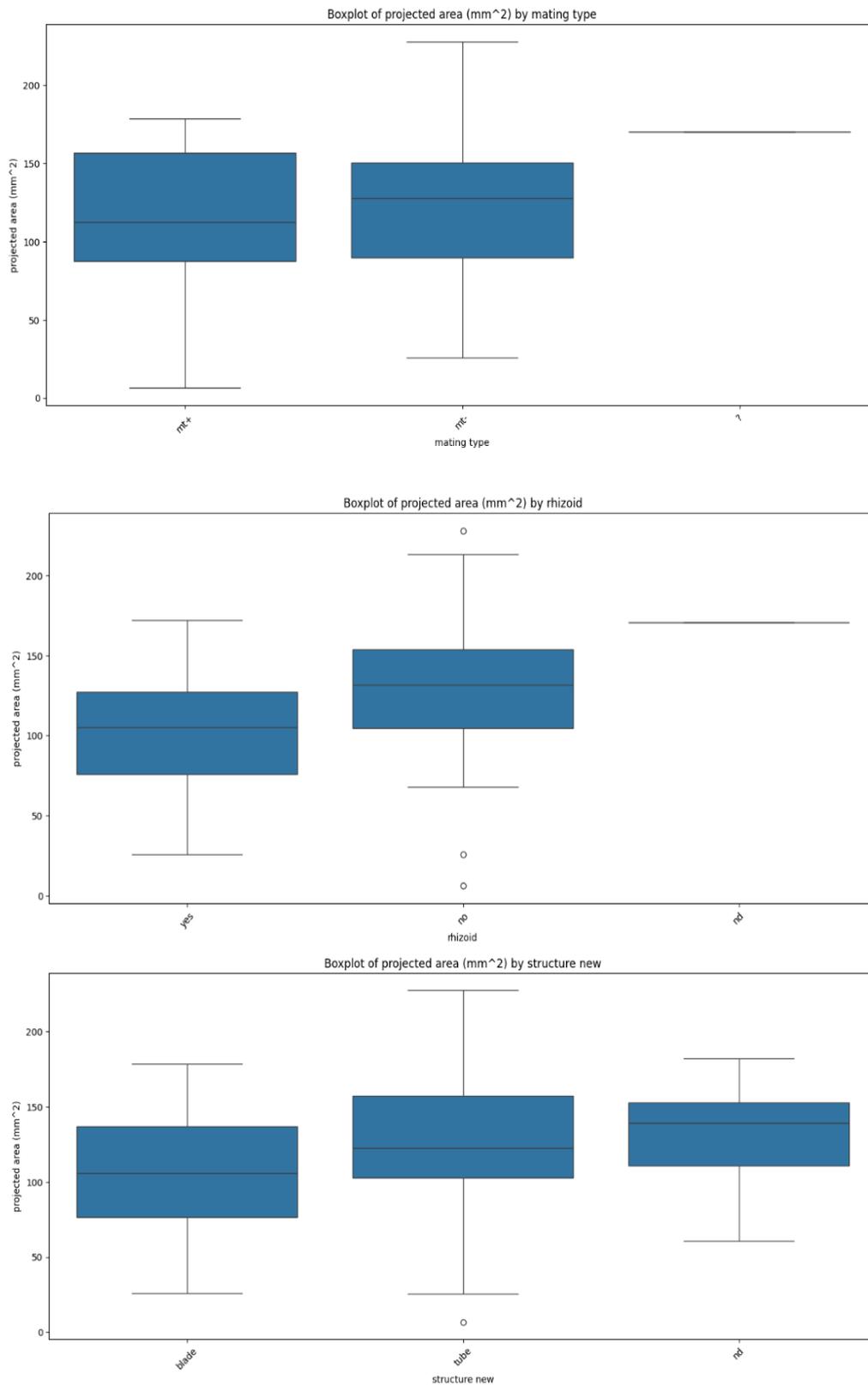
Figure 24. Box plot visualizes the relationship between the mating type, rhizoid and structure categories and the "projected area (mm^2)" in the dataset.

In summary, the IGIS platform effectively quantified the growth rates of a mapping population, revealing significant differences among segregating lines. By analyzing the projected area of growth over time, we identified segregating line number 89 as having the highest average projected area and a *slender* control as having the lowest. Growth patterns varied, with some segregating lines maintaining a steady growth rate while others fluctuated. To further understand these differences, K-means clustering was employed, categorizing segregating lines into small, medium, and large growth groups. We do not observe statistically significant differences in projected area across these phenotypic categories. These findings highlight the variability in growth behaviors within the population and provide a basis for further genetic and phenotypic analysis.

4.4. Results of the genetic analysis

In this study the genetic makeup of the mapping population will not be detected by WGS due to lack of time. Instead, we explored if the genetic makeup can be efficiently analyzed by using multiplex PCR analysis. As a test-case, we designed an experiment to confirm SNPs identified in a genome comparison of *slender* (unpublished) and wild-type (De Clerck et al., 2018). Opting for multiplex PCR analysis instead of WGS can offer several advantages, particularly in terms of efficiency, cost, and specificity. Multiplex PCR allows simultaneous amplification of multiple specific DNA regions, making it an ideal method for targeted genotyping. By using this method, we can efficiently generate SNP data pertinent to phenotypic observations without the need to resequencing the entire genome. A preliminary simplex test was done to test 22 primer pairs individually on either *slender* or wild-type DNA. Based on gel electrophoresis analysis (Figure 25), we managed to amplify 18/22 amplicons in *slender* and 16/22 in wild-type.
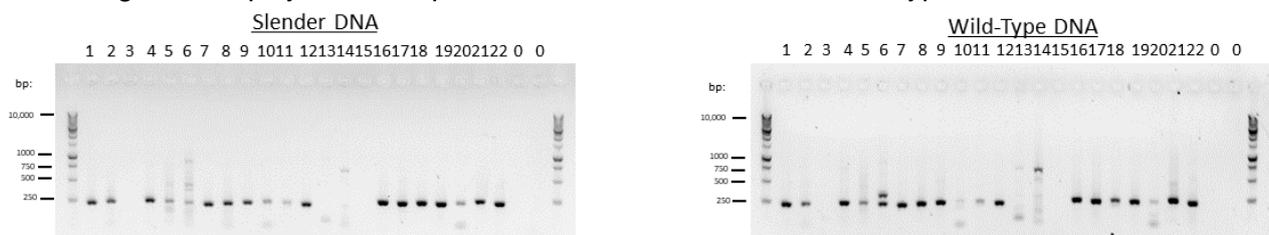
Figure 25. Simplex result of each primer tested on *slender* and wild-type DNA. The numbers represent the primers (MP-) and 0 represent blank (no primer).

To confirm if the anticipated genomic region was amplified, the PCR products of the simplex test were purified then sent for Sanger sequencing. The sequence was mapped to the *slender* and wild-type genome. Compared to the wild-type genome, a total of 10 SNPs was detected, 5 was not detected and 3 are not possible due to bad sequence read or technical limitations, since Sanger sequencing did not accurately sequence SNPs in close proximity to the sequencing primer. However, in all SNPs identified, surprisingly, we did not observe a difference between wild-type and *slender*. We conducted a follow-up genotyping experiment with the mating-type specific primers and confirmed that our input DNA for this experiment was not contaminated. The data from this limited set of amplicons suggests that the genome assembly of wild-type and *slender* still contain several sequencing errors that have been interpreted as SNPs (Figure 26).
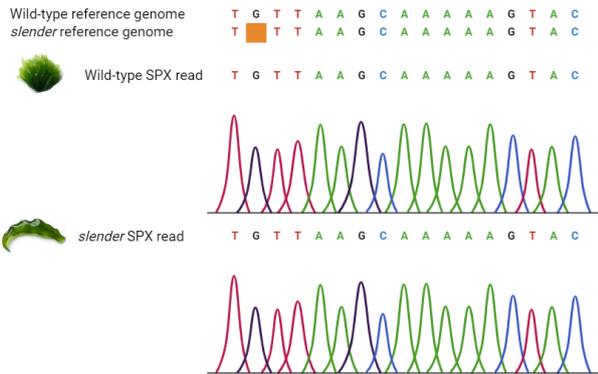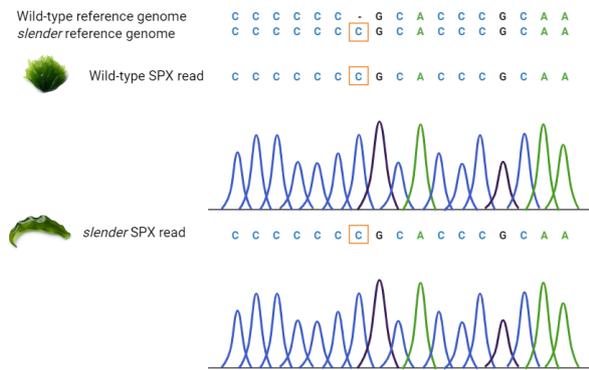
Figure 26. Example of an amplicon simplex DNA sequence result mapped to the *Ulva slender* and wild-type genome. (A) part of MP1 amplicon sequence resulted in no SNPs detected. (B) part of MP2 primer sequence resulted in SNP detected but not as predicted (both wild-type and *slender* DNA shows no difference). Orange box is the SNP place. SPX=simplex

A similar result was detected in the multiplex PCR. Both the *slender* DNA and wild-type was successfully amplified by the primer mix visualized by gel electrophoresis (Figure 27). The PCR products sent to Oxford nanopore sequencing give 621 reads for *slender* and 351 reads for wild-type. The reads were trimmed for high quality then mapped to the reference amplicons. In total 444 reads from *slender* were mapped to the anticipated amplicons (41.69%), while only 55 were mapped for the wild-type (13.55%). Reads for both strains mapped to the same 13 out of 22 amplicons, with one additional amplicon for *slender*. The highest number of reads were observed by amplicon 7 (56 reads) for *slender* and for wild-type it was the amplicon 16 (41 reads). Meanwhile, the lowest if not zero is amplicon 4 (2 reads) and amplicon 13 (6 reads) for *slender* and wild-type, respectively.
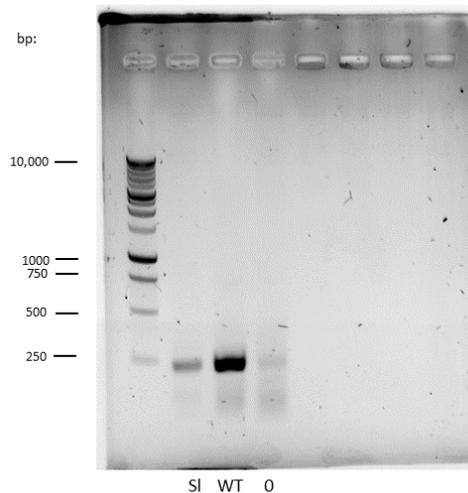
Figure 27. The Multiplex PCR results visualized in gel electrophoresis. Sl contains *slender* DNA. WT contains wild-type DNA. 0 is water control.

Although Nanopore sequencing is error-prone and not ideal to perform a SNP analysis, we evaluated the anticipated SNPs based on the reads we obtained. Real SNPs could be detected since they were consistently called in multiple reads as opposed to sequencing errors that were limited to one or a few reads. For both *slender* and wild-type, we obtained the same consensus sequence for the analyzed amplicons. Out of the 14 amplicons, six are similar to the wild-type genome, four are similar to the *slender* genome, three have the both a SNP similar to wild-type and *slender* reference and one amplicon could not be evaluated due to a low quality read (Figure 28 and Appendix F).
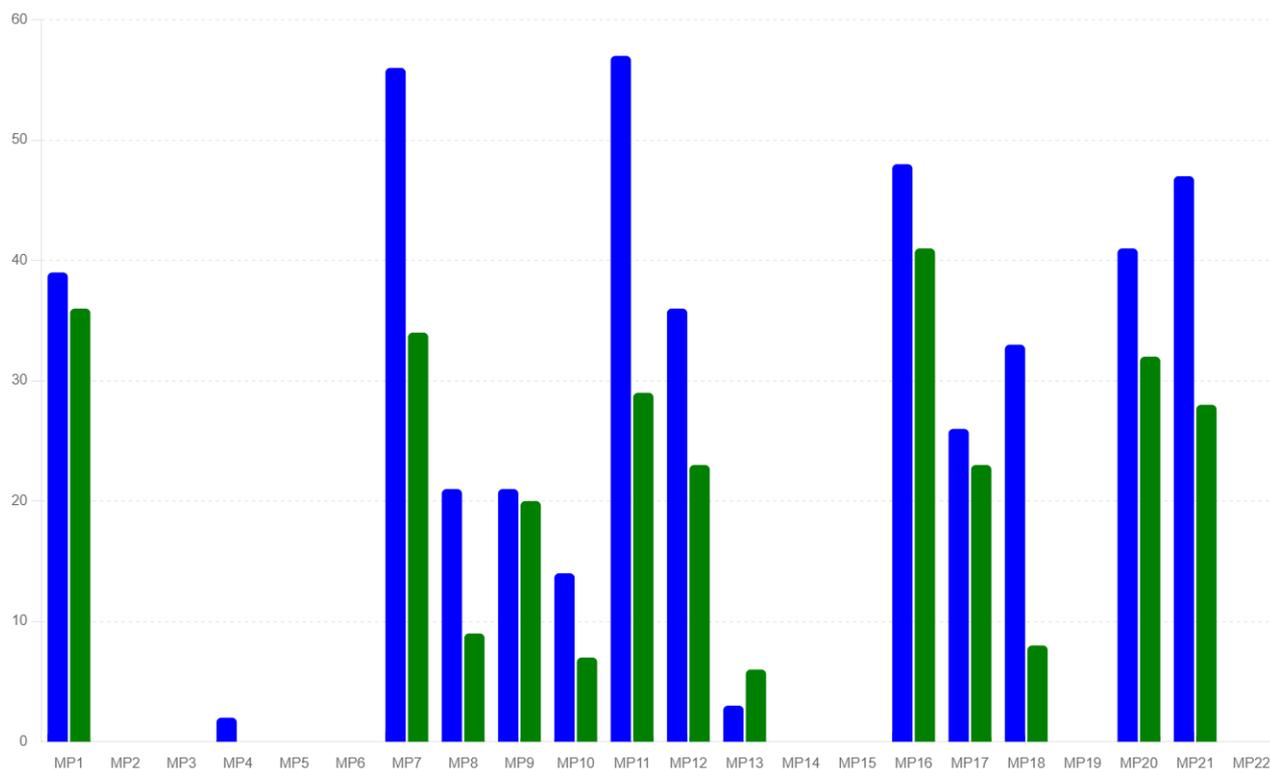
Figure 28. Multiplex reads mapped distribution of each amplicon (MP1 – MP22). Blue bars = *slender* DNA. Green bars = wild-type DNA

In summary, the multiplex PCR analysis provided an efficient and cost-effective method for genotyping and detecting SNPs in the slender and wild-type genomes. Despite some technical limitations, this approach successfully identified genetic variations and highlighted the need for further refinement in the genome assemblies of both strains.

V. DISCUSSION

5.1. Insights from Culturing and Crossing of Ulva

The *Ulva* crossing experiments revealed notable differences in gamete production between *slender* and wild-type strains, with the *slender* strain producing significantly more gametes compared to the wild-type strain. This discrepancy required adjusting the volumes of gametes used in crossing experiments to achieve approximately equal amounts of gametes from both strains. The differences in gamete production between the two strains suggest intrinsic genetic factors influencing reproductive success. This disparity in gamete production highlights a potential area for further investigation, particularly in understanding the genetic regulation of reproductive traits in *Ulva*. Or simply, the induction protocols for wild-type might needs to be optimized. Another medium change, or longer gametangia time to develop may needed for wild-type strains.

In comparison, a study on the brown algae *Fucus vesiculosus* natural populations showed a lower gamete concentration, with the highest egg concentration at 1500/L and the highest sperm concentration at 8000/L. Despite the lower gamete concentrations, the fertilization rate was nearly 100% (Berndt et al., 2002). This indicates that factors other than gamete concentration, such as gamete viability and compatibility, might play crucial roles in fertilization success.

Additionally, another study on the red algae *Gracilaria gracilis* showed that females on average produce 19.5 cystocarps per dm of thallus. This study suggested that fertilization success is influenced not only by the distance between male and female gametes but also by male/male competition and female choice (Engel et al., 1999). This highlights the importance of gamete interactions and selection mechanisms in reproductive success.

The success rate of *Ulva* crossings was relatively low, with only one out of five putative sporophytes confirmed through progeny analysis. This low success rate may be attributed to several factors, including the challenges in visually selecting sporophytes from dense cultures and potential differences in gamete viability or compatibility between the strains. Similar to the findings in *Fucus vesiculosus* and *Gracilaria gracilis*, the reproductive success in *Ulva* may depend on more complex interactions beyond gamete production, such as gamete viability, compatibility, and selection mechanisms.

Overall, while *Ulva* and the studied brown and red algae differ in their gamete production and reproductive strategies, the importance of gamete viability, compatibility, and selection mechanisms appears to be a common factor influencing fertilization success across these species. Understanding these factors in *Ulva* could provide insights into the genetic regulation of reproductive traits and improve crossing success rates in future experiments involving creating mapping population.

The successful generation of a mapping population from the confirmed sporophyte is a significant achievement, providing a valuable resource for further genetic studies. The new mapping population will complement the 166 segregating lines from the existing mapping population. This number of segregating lines is higher compared to other mapping population of seaweed. For example, the mapping of sex-linked locus for the brown alga *Undaria pinnatifida* only accounts for 50 female and 51 male gametophytes (Shan et al., 2015). Another SNP-based QTL

mapping of *S. japonica* worked with 178 sporophytes with SLAF-seq method (X. Wang et al., 2018), and other worked with 125 segregating lines with SSR method (X. Wang et al., 2023).

The low success rate of identifying sporophytes highlights the need for improved methods for selecting and confirming sporophytes in future experiments. However, this molecular techniques using the mating type primers are new and more objective compared to previous techniques which only relies on morphology.Techniques such as fluorescence-activated cell sorting (FACS) or advanced molecular markers could be developed to do sporophyte selection and increase the success rate of crossing experiments. For instance, to utilize FACS for selecting sporophytes in creating an *Ulva* mapping population, we can start by identifying sporophyte-specific biomarkers through transcriptomic or proteomic analysis. Then, develop fluorescent tags, such as antibodies or dyes, for these biomarkers. Isolate and stain *Ulva* cells with these fluorescent tags, then calibrate the FACS machine to detect the specific fluorescence signal. Sort the cell populations, and verify the presence of sporophytes using targeted sex amplicon PCR. This method combines molecular biology and advanced cell sorting to achieve precise sporophyte selection despite the isomorphism with gametophytes. Additionally, FACS have been sufficient to screen microalgae in a study by Lin et al. (2020). FACS also have been done to differentiate *Fucus* sperm and egg (Berndt et al., 2002).

5.2. Implications from the Genotyping Results of the Mapping Population

The genotyping of the pre-existing mapping population identified 70 segregating lines as mt+ and 85 segregating lines as mt-, with 11 segregating lines showing unclear results with amplicon in opposite mating types targeted PCR. The chi-square test for goodness of fit indicated no significant difference between the observed distribution and the expected 50:50 ratio. This finding is consistent with the expected behavior of a simple Mendelian trait and provides a solid foundation for further genetic analyses of the mapping population.

The presence of segregating lines with unclear genotyping results underscores the importance of using robust and reliable genotyping methods. The use of the Phire method successfully reduced the number of unclear segregating lines, demonstrating its effectiveness. However, further optimization and validation of genotyping protocols are necessary to minimize ambiguous results such as amplicon in opposite mating types and ensure accurate determination of mating types.

Other genotyping methodologies could also be explored in genotyping *Ulva*. Since a study showed that while findings from Bayesian clustering analysis showed a general congruence between the two marker classes, six SNP loci produced considerably different patterns of intrapopulation genetic diversity compared to those generated using seven moderately polymorphic microsatellite sites (Provan et al., 2012). Other method of DNA extraction could also be explored, for example using different extraction buffer or some cleanup before extraction. For instance, a diversity study of *Asparagopsis taxiformis* used Roche buffer and remove the carpogonia before DNA extraction (Andreakis et al., 2009).

5.3. Understanding of the Phenotyping Results

5.3.1. Morphological Observation Comprehensions

The phenotypic analysis of the mapping population revealed a significant correlation between the presence of rhizoids and the type of thalli structure. Segregating lines with fully developed rhizoids predominantly exhibited blade-like structures. However, we still do see segregation of this trait, meaning that they can still be inherited independently and do not always co-occur in every instance. Meanwhile those lacking rhizoids were mostly tubular. This association suggests that the genetic factors governing rhizoid development may also influence overall thalli morphology. The absence of a significant association between mating type and these morphological traits further supports the hypothesis that these traits are controlled by distinct genetic loci.

The phenotypic analysis of the mapping population revealed distinct differences compared to the parental strains. The *slender* (mt+) strain exhibited a tube-like structure and lacked fully developed rhizoids, while the wild-type (mt-) strain displayed a blade-like structure with well-developed rhizoids. These phenotypic traits segregated in the mapping population, with a variety of structural forms observed, including tube and blade structures in both mating types.

5.3.2. IGIS Results proof differences in growth

The use of the IGIS platform for high-throughput phenotyping provided valuable insights into the growth dynamics of the mapping population. The observed variation in growth by projected area among different segregating lines underscores the genetic diversity within the population. The clustering analysis categorized segregating lines into small, medium, and large growth groups, highlighting the potential for selecting fast-growing segregating lines for further analysis. Interestingly, it was observed that the *slender* phenotype, which was expected to grow rapidly, actually grew slower compared to the wild-type. This unexpected result may be attributed to the fact that the slender segregating lines simply did not grow as anticipated; they might have been too small when moved to the IGIS platform and could have experienced stress. This underscores the importance of using replicates in experiments to account for variability and ensure that observed trends are robust and not artifacts of segregating line sample responses.

The absence of significant associations between growth rate and mating type, rhizoid presence, or thalli structure indicates that growth rate is likely influenced by other genetic factors not directly linked to these phenotypic traits. This means the growth rate segregates independently from these traits. Future replicates are needed to validate these findings and further explore the underlying genetic mechanisms affecting growth rate. A similar result was observed in a study assessing the ploidy distribution of *Ulva* in the USA. The study found that there is no significant difference in growth rate between gametophyte and sporophyte despite sporophytes having significantly larger cells, suggesting that growth rate did not affected from the phenotype of cell size (Potter et al., 2016).

5.4. Subsequent Research following Success of the Genetic Analysis

The genetic analysis using multiplex PCR provided an efficient and cost-effective method for genotyping and SNP detection in the *slender* and wild-type genomes. Although technical limitations and sequencing errors were encountered, the approach successfully identified SNPs

between the strains. However, it is different than the expected results with both strain reads are always similar whether showing or not showing the SNPs. The discrepancies observed between the predicted and actual SNPs highlight the need for further refinement of the genome assemblies for both strains.

The preliminary success of the multiplex PCR method demonstrates its potential for targeted genotyping in future studies of *Ulva*. However, the results also underscore the importance of validating predicted SNPs through additional sequencing methods or using larger datasets to ensure accuracy. The continued development and refinement of genetic tools will be crucial for advancing our understanding of the genetic basis of phenotypic traits in *Ulva*.

5.5. Relevance to the Global South

This thesis significantly contributes to the global south by advancing the genetic understanding and cultivation practices of *Ulva*, a green seaweed with substantial ecological and economic importance. The global south, particularly countries in Asia, Africa, and South America, are major producers of seaweed, with vast coastal regions suitable for aquaculture. For instance, the country where author from, Indonesia, Is the 2nd biggest producer of seaweed in the world (García-Poza et al., 2020). Seaweeds like *Ulva* are critical for local economies as sources of food, biofuels, and bioactive compounds, and play a role in environmental management through bioremediation. However, *Ulva* and many other green seaweeds are underutilized despite having a fast growth, robust, reliable, ubiquitous and highly adaptive. Therefore, by elucidating the genetic factors behind *Ulva*'s rapid growth, this study provides valuable insights that can help optimize *Ulva* (and other seaweed) production, making it more sustainable and economically viable.

Furthermore, the creation of the first mapping population of green seaweed and the subsequent phenotypic and genotypic analyses set a precedent for future studies aimed at improving seaweed cultivation. This research highlights the potential for selective breeding programs to enhance growth rates and resilience in *Ulva* strains, which is crucial for the global south where seaweed farming is a key livelihood. Improved seaweed strains can lead to higher yields, better quality products, and more efficient bioremediation practices, addressing both economic and environmental challenges. This study paves the way for genome-wide association studies (GWAS) and quantitative trait loci (QTL) mapping, which can further refine our understanding of genetic determinants in seaweed. In best scenario, the identified genetic factor for rapid growth can be implemented or upregulated in other seaweed, ultimately leading to innovations in aquaculture that benefit the global south's economies and ecosystems.

5.6. Broader Implications, Future Directions and Significance in general

This study contributes significantly to our understanding of the genetic factors that influence the rapid growth of *Ulva*; a green seaweed implicated in green tide events. After identifying specific SNPs associated with growth rates and other phenotypes, we show that SNPs analysis with multiplex PCR and amplicon sequencing is possible in *Ulva*. This could lead to the development of targeted breeding programs aimed at producing *Ulva* strains optimized for aquaculture. Such strains could enhance productivity and sustainability within the blue economy, offering a viable alternative for biofuel production, bioremediation, and as a nutritious food source.

Moreover, this genetic understanding can be pivotal in managing and mitigating the ecological impacts of green tides. By identifying the genetic markers linked to rapid growth, it might be possible to predict and control the proliferation of green tide-forming *Ulva* strains. This could lead to more effective environmental management practices and policies that balance ecological health with economic benefits.

Future research should focus on expanding the genetic and phenotypic analyses of *Ulva* by incorporating a broader range of strains and environmental conditions. This would help to better understand the interaction between genetic factors and environmental variables in influencing growth rates and other phenotypic traits. GWAS and QTL mapping should be conducted on a larger scale to pinpoint more precise genetic determinants of desirable traits. Additionally, the development of more refined molecular tools and techniques for *Ulva*, such as CRISPR-Cas9 gene editing, could facilitate functional studies to confirm the roles of identified SNPs and other genetic elements. Long-term field studies should also be undertaken to observe the performance of genetically characterized strains in natural and semi-natural environments, providing insights into their ecological impacts and commercial viability.

One of the primary limitations of this study is the reliance on laboratory conditions means that the findings might not be directly transferable to natural settings, where environmental factors play a significant role. The genotyping techniques used, while innovative, also have limitations in their resolution and the completeness of the genomic coverage they provide. Furthermore, the study primarily focuses on growth rate as the main phenotypic trait, potentially overlooking other important traits such as stress tolerance, nutrient uptake efficiency, and reproductive strategies. These traits could also significantly impact the ecological and economic outcomes of *Ulva* cultivation and green tide management.

This research is significant as it lays the groundwork for the genetic improvement of *Ulva*, supporting its role in the sustainable blue economy. By identifying key genetic factors that influence growth, the study opens up new avenues for the selective breeding of *Ulva* strains that are not only more productive but also potentially less likely to contribute to harmful green tides. The findings also have broader ecological implications, offering new strategies for managing green tide events that are less reliant on chemical interventions and more focused on ecological balance. This could lead to more sustainable practices in coastal management and the aquaculture industry.

VI. CONCLUSION

The findings of this research underscore the intricate relationship between genetic factors and phenotypic traits in *Ulva*, particularly its rapid growth which contributes to the formation of green tides. The significant genetic diversity within the mapping population, evidenced by the phenotypic variability observed in growth rates, rhizoid development, and thallus structure, points to the complex genetic architecture underlying these traits. The use of the IGIS platform for high-throughput phenotyping and the implementation of multiplex PCR for efficient genotyping have proven valuable in identifying and analyzing genetic variations. Although some technical challenges were encountered, the methodologies employed provide a robust foundation for future genetic studies on *Ulva*, paving the way for advancements in selective breeding and sustainable aquaculture practices.

Furthermore, the study highlights the potential of integrating genetic insights into environmental management strategies to mitigate the ecological impacts of green tides. By identifying genetic markers associated with rapid growth and other critical traits, this research opens up possibilities for predicting and controlling the proliferation of green tide-forming *Ulva* strains. This approach not only enhances our understanding of the genetic determinants of *Ulva*'s rapid growth but also offers practical solutions for balancing ecological health with economic benefits. Future research should focus on expanding the genetic and phenotypic analyses to include a broader range of strains and environmental conditions, leveraging advanced molecular tools and techniques to further elucidate the genetic basis of *Ulva*'s growth and development. This comprehensive understanding will support the development of more sustainable practices in coastal management and the aquaculture industry, ultimately contributing to the sustainable blue economy.

Reference list

Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International*, *11*(7), 36–42.

Aires, T., Muyzer, G., Serrão, E. A., & Engelen, A. H. (2019). Seaweed loads cause stronger bacterial community shifts in coastal lagoon sediments than nutrient loads. *Frontiers in Microbiology*, *10*(JAN). https://doi.org/10.3389/fmicb.2018.03283

Alagan, V., Valsala, R., & Rajesh, K. (2017). Bioactive Chemical Constituent Analysis, in vitro Antioxidant and Antimicrobial Activity of Whole Plant Methanol Extracts of Ulva lactuca Linn. *British Journal of Pharmaceutical Research*, *15*(1), 1–14. https://doi.org/10.9734/bjpr/2017/31818

Alsufyani, T., Weiss, A., & Wichard, T. (2017). Time course exo-metabolomic profiling in the green marine macroalga Ulva (Chlorophyta) for identification of growth phase-dependent biomarkers. *Marine Drugs*, *15*(1). https://doi.org/10.3390/md15010014

Anburaj, R., Kathiresan, K., & PRASANNAKUMAR, C. (2024). Exploring the depths: A comprehensive review of marine algae. *RECENT TRENDS IN*, 176.

Andreakis, N., Kooistra, W. H. C. F., & Procaccini, G. (2009). High genetic diversity and connectivity in the polyploid invasive seaweed Asparagopsis taxiformis (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Molecular Ecology*, *18*(2), 212–226.

Anjali, K. P., Sangeetha, B. M., Devi, G., Raghunathan, R., & Dutta, S. (2019). Bioprospecting of seaweeds (Ulva lactuca and Stoechospermum marginatum): The compound characterization and functional applications in medicine-a comparative study. *Journal of Photochemistry and Photobiology B: Biology*, *200*. https://doi.org/10.1016/j.jphotobiol.2019.111622

Arimoto, A., Nishitsuji, K., Higa, Y., Arakaki, N., Hisata, K., Shinzato, C., Satoh, N., & Shoguchi, E. (2019). A siphonous macroalgal genome suggests convergent functions of homeobox genes in algae and land plants. *DNA Research*, *26*(2), 183–192. https://doi.org/10.1093/dnares/dsz002

Ben-Ari, T., Neori, A., Ben-Ezra, D., Shauli, L., Odintsov, V., & Shpigel, M. (2014). Management of Ulva lactuca as a biofilter of mariculture effluents in IMTA system. *Aquaculture*, *434*, 493–498. https://doi.org/10.1016/j.aquaculture.2014.08.034

Berndt, M.-L., Callow, J. A., & Brawley, S. H. (2002). Gamete concentrations and timing and success of fertilization in a rocky shore seaweed. *Marine Ecology Progress Series*, *226*, 273–285.

Bews, E., Booher, L., Polizzi, T., Long, C., Kim, J. H., & Edwards, M. S. (2021). Effects of salinity and nutrients on metabolism and growth of Ulva lactuca: Implications for bioremediation of coastal watersheds. *Marine Pollution Bulletin*, *166*. https://doi.org/10.1016/j.marpolbul.2021.112199

Bikker, P., van Krimpen, M. M., van Wikselaar, P., Houweling-Tan, B., Scaccia, N., van Hal, J. W., Huijgen, W. J. J., Cone, J. W., & López-Contreras, A. M. (2016). Biorefinery of the green seaweed Ulva lactuca to produce animal feed, chemicals and biofuels. *Journal of Applied Phycology*, *28*(6), 3511–3525. https://doi.org/10.1007/s10811-016-0842-3

Blomme, J., Wichard, T., Jacobs, T. B., & De Clerck, O. (2023). Ulva: An emerging green seaweed model for systems biology. *Journal of Phycology*, *59*(3), 433–440. https://doi.org/10.1111/jpy.13341

Bogolitsyn, K., Parshina, A., & Aleshina, L. (2020). Structural features of brown algae cellulose. *Cellulose*, *27*(17), 9787–9800. https://doi.org/10.1007/s10570-020-03485-z

Brawley, S. H., Blouin, N. A., Ficko-Blean, E., Wheeler, G. L., Lohr, M., Goodson, H. V., Jenkins, J. W., Blaby-Haas, C. E., Helliwell, K. E., Chan, C. X., Marriage, T. N., Bhattacharya, D., Klein, A. S., Badis, Y., Brodie, J., Cao, Y., Collén, J., Dittami, S. M., Gachon, C. M. M., … Prochnik, S. E. (2017). Insights into the red algae and eukaryotic evolution from the genome of Porphyra umbilicalis (Bangiophyceae, Rhodophyta). *Proceedings of the National Academy of Sciences of the United States of America*, *114*(31), E6361–E6370. https://doi.org/10.1073/pnas.1703088114

Bringloe, T. T., Starko, S., Wade, R. M., Vieira, C., Kawai, H., De Clerck, O., Cock, J. M., Coelho, S. M., Destombe, C., Valero, M., Neiva, J., Pearson, G. A., Faugeron, S., Serrão, E. A., & Verbruggen, H. (2020). Phylogeny and Evolution of the Brown Algae. *Critical Reviews in Plant Sciences*, *39*(4), 281–321. https://doi.org/10.1080/07352689.2020.1787679

Bruhn, A., Dahl, J., Nielsen, H. B., Nikolaisen, L., Rasmussen, M. B., Markager, S., Olesen, B., Arias, C., & Jensen, P. D. (2011). Bioenergy potential of Ulva lactuca: Biomass yield, methane production and combustion. *Bioresource Technology*, *102*(3), 2595–2604. https://doi.org/10.1016/j.biortech.2010.10.010

Busetti, A., Maggs, C. A., & Gilmore, B. F. (2017). Marine macroalgae and their associated microbiomes as a source of antimicrobial chemical diversity. *European Journal of Phycology*, *52*(4), 452–465. https://doi.org/10.1080/09670262.2017.1376709

Cai, C., Wang, L., Zhou, L., He, P., & Jiao, B. (2017). Complete chloroplast genome of green tide algae Ulva flexuosa (Ulvophyceae, Chlorophyta) with comparative analysis. *PLoS ONE*, *12*(9). https://doi.org/10.1371/journal.pone.0184196

Cai, J., Ni, J., Chen, Z., Wu, S., Wu, R., He, C., Wang, J., Liu, Y., Zhou, W., & Xu, J. (2023). Effects of ocean acidification and eutrophication on the growth and photosynthetic performances of a green tide alga Ulva prolifera. *Frontiers in Marine Science*, *10*. https://doi.org/10.3389/fmars.2023.1145048

Calmes, B., Strittmatter, M., Jacquemin, B., Perrineau, M. M., Rousseau, C., Badis, Y., Cock, J. M., Destombe, C., Valero, M., & Gachon, C. M. M. (2020). Parallelisable non-invasive biomass, fitness and growth measurement of macroalgae and other protists with nephelometry. *Algal Research*, *46*. https://doi.org/10.1016/j.algal.2019.101762

Camus, C., Faugeron, S., & Buschmann, A. H. (2018). Assessment of genetic and phenotypic diversity of the giant kelp, Macrocystis pyrifera, to support breeding programs. *Algal Research*, *30*, 101–112. https://doi.org/10.1016/j.algal.2018.01.004

Castelar, B., Reis, R. P., & dos Santos Calheiros, A. C. (2014). Ulva lactuca and U. flexuosa (Chlorophyta, Ulvophyceae) cultivation in Brazilian tropical waters: Recruitment, growth, and ulvan yield. *Journal of Applied Phycology*, *26*(5), 1989–1999. https://doi.org/10.1007/s10811-014-0329-z

Chen, Z., Liu, M., Yang, Y., Bi, M., Li, M., & Liu, W. (2022). Environmental and Economic Impacts of Different Disposal Options for Ulva prolifera Green Tide in the Yellow Sea, China. *ACS Sustainable Chemistry and Engineering*, *10*(35), 11483–11492. https://doi.org/10.1021/acssuschemeng.2c02638

Cock, J. M., Sterck, L., Rouzé, P., Scornet, D., Allen, A. E., Amoutzias, G., Anthouard, V., Artiguenave, F., Aury, J. M., Badger, J. H., Beszteri, B., Billiau, K., Bonnet, E., Bothwell, J. H., Bowler, C., Boyen, C., Brownlee, C., Carrano, C. J., Charrier, B., … Wincker, P. (2010). The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature*, *465*(7298), 617–621. https://doi.org/10.1038/nature09016

Cotas, J., Gomes, L., Pacheco, D., & Pereira, L. (2023). Ecosystem Services Provided by Seaweeds. *Hydrobiology*, *2*(1), 75–96. https://doi.org/10.3390/hydrobiology2010006

De Clerck, O., Bogaert, K. A., & Leliaert, F. (2012). Chapter Two - Diversity and Evolution of Algae: Primary Endosymbiosis. In G. Piganeau (Ed.), *Advances in Botanical Research* (Vol. 64, pp. 55–86). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-391499-6.00002-5

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., Del Cortona, A., D'Hondt, S., Fernandez-Pozo, N., Gachon, C. M., Hanikenne, M., Lattermann, L., … Bothwell, J. H. (2018). Insights into the Evolution of Multicellularity from the Sea Lettuce Genome. *Current Biology*, *28*(18), 2921-2933.e5. https://doi.org/10.1016/j.cub.2018.08.015

De Clerck, O., Tronchin, E. M., & Schils, T. (2005). Red algae. *Rhodophyceae. Guide to the Seaweeds of KwaZulu-Natal. Scripta Bot. Belgica*, *33*, 131–267.

Debbarma, J., Madhusudana Rao, B., Narasimha Murthy, L., Mathew, S., Venkateshwarlu, G., & Ravishankar, C. N. (2016). Nutritional profiling of the edible seaweeds Gracilaria edulis, Ulva lactuca and Sargassum sp. *Indian Journal of Fisheries*, *63*(3), 81–87. https://doi.org/10.21077/ijf.2016.63.3.60073-11

Develtere, W., Waegneer, E., Debray, K., De Saeger, J., Van Glabeke, S., Maere, S., Ruttink, T., & Jacobs, T. B. (2023). SMAP design: a multiplex PCR amplicon and gRNA design tool to screen for natural and

CRISPR-induced genetic variation. *Nucleic Acids Research*, *51*(7), E37–E37. https://doi.org/10.1093/nar/gkad036

Dhondt, S., Gonzalez, N., Blomme, J., De Milde, L., Van Daele, T., Van Akoleyen, D., Storme, V., Coppens, F., Beemster, G. T. S., & Inzé, D. (2014). High-resolution time-resolved imaging of in vitro Arabidopsis rosette growth. *Plant Journal*, *80*(1), 172–184. https://doi.org/10.1111/tpj.12610

Din, N. A. S., Mohd Alayudin, 'Ain Sajda, Sofian-Seng, N. S., Rahman, H. A., Mohd Razali, N. S., Lim, S. J., & Wan Mustapha, W. A. (2022). Brown Algae as Functional Food Source of Fucoxanthin: A Review. In *Foods* (Vol. 11, Issue 15). MDPI. https://doi.org/10.3390/foods11152235

Dishon, G., Resetarits, H. M., Tsai, B., Black, K., Grossmann, J., & Smith, J. E. (2023). Image-based analysis and quantification of biofouling in cultures of the red alga Asparagopsis taxiformis. *Journal of Applied Phycology*, *35*(1), 209–218. https://doi.org/10.1007/s10811-022-02884-y

E, H. M., S, M., & Baskaran, R. (2023). The sea lettuce Ulva sensu lato: Future food with health-promoting bioactives. In *Algal Research* (Vol. 71). Elsevier B.V. https://doi.org/10.1016/j.algal.2023.103069

Engel, C. R., Wattierf, R., Destombe, C., & Valero, M. (1999). Performance of non–motile male gametes in the sea: analysis of paternity and fertilization success in a natural population of a red seaweed, Gracilaria gracilis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1431), 1879–1886.

Fadason, T., Farrow, S., Gokuladhas, S., Golovina, E., Nyaga, D., O'Sullivan, J. M., & Schierding, W. (2022). Assigning function to SNPs: Considerations when interpreting genetic variation. *Seminars in Cell and Developmental Biology*, *121*, 135–142. https://doi.org/10.1016/j.semcdb.2021.08.008

Fan, M., Sun, X., Liao, Z., Wang, J., Li, Y., & Xu, N. (2018). Comparative proteomic analysis of Ulva prolifera response to high temperature stress. *Proteome Science*, *16*(1). https://doi.org/10.1186/s12953-018-0145-5

Fan, Q., Shi, K., Zhan, M., Xu, Q., Liu, X., Li, Z., Liu, H., Xia, Y., Chen, Y., Shi, X., & Sha, Z. (2022). Acute damage from the degradation of Ulva prolifera on the environmental microbiota, intestinal microbiota and transcriptome of Japanese flounder Paralichthys olivaceus. *Environmental Pollution*, *302*. https://doi.org/10.1016/j.envpol.2022.119022

Feng, X., Xiao, B., Jiang, M., Li, P., Wu, Q., Dong, Y., Wang, J., & Sui, Z. (2023). Identification of candidate genes related to two economic traits using GWAS in Gracilariopsis lemaneiformis (Rhodophyta). *Algal Research*, *76*. https://doi.org/10.1016/j.algal.2023.103309

Fort, A., Lebrault, M., Allaire, M., Esteves-Ferreira, A. A., McHale, M., Lopez, F., Fariñas-Franco, J. M., Alseekh, S., Fernie, A. R., & Sulpice, R. (2019). Extensive variations in diurnal growth patterns and metabolism among ulva spp. Strains. *Plant Physiology*, *180*(1), 109–123. https://doi.org/10.1104/pp.18.01513

Fort, A., Mannion, C., Fariñas-Franco, J. M., & Sulpice, R. (2020). Green tides select for fast expanding Ulva strains. *Science of the Total Environment*, *698*. https://doi.org/10.1016/j.scitotenv.2019.134337

Fort, A., Monteiro, J. P., Simon, C., Rosário Domingues, M., & Sulpice, R. (2024). Short term decreases in salinity, combined with the right choice of species, can allow for a more nutritious sea lettuce lipid profile. *Food Chemistry*, *437*. https://doi.org/10.1016/j.foodchem.2023.137865

Gao, Z., Xu, D., Meng, C., Zhang, X., Wang, Y., Li, D., Zou, J., Zhuang, Z., & Ye, N. (2014). The green tide-forming macroalga Ulva linza outcompetes the red macroalga Gracilaria lemaneiformis via allelopathy and fast nutrients uptake. *Aquatic Ecology*, *48*(1), 53–62. https://doi.org/10.1007/s10452-013-9465-9

García-Poza, S., Leandro, A., Cotas, C., Cotas, J., Marques, J. C., Pereira, L., & Gonçalves, A. M. M. (2020). The evolution road of seaweed aquaculture: Cultivation technologies and the industry 4.0. In *International Journal of Environmental Research and Public Health* (Vol. 17, Issue 18, pp. 1–42). MDPI AG. https://doi.org/10.3390/ijerph17186528

Gaspar, R., Pereira, L., & Neto, J. M. (2017). Intertidal zonation and latitudinal gradients on macroalgal assemblages: Species, functional groups and thallus morphology approaches. *Ecological Indicators*, *81*, 90–103. https://doi.org/10.1016/j.ecolind.2017.05.060

Ghaderiardakani, F., Langhans, L., Kurbel, V. B., Fenizia, S., & Wichard, T. (2022). Metabolite profiling reveals insights into the species-dependent cold stress response of the green seaweed holobiont Ulva (Chlorophyta). *Environmental and Experimental Botany*, *200*. https://doi.org/10.1016/j.envexpbot.2022.104913

Graf, L., Shin, Y., Yang, J. H., Choi, J. W., Hwang, I. K., Nelson, W., Bhattacharya, D., Viard, F., & Yoon, H. S. (2021). A genome-wide investigation of the effect of farming and human-mediated introduction on the ubiquitous seaweed Undaria pinnatifida. *Nature Ecology and Evolution*, *5*(3), 360–368. https://doi.org/10.1038/s41559-020-01378-9

Green-Gavrielidis, L. A., MacKechnie, F., Thornber, C. S., & Gomez-Chiarri, M. (2018). Bloom-forming macroalgae (Ulva spp.) inhibit the growth of co-occurring macroalgae and decrease eastern oyster larval survival. *Marine Ecology Progress Series*, *595*, 27–37. https://doi.org/10.3354/meps12556

Gupta, V., Bijo, A. J., Kumar, M., Reddy, C. R. K., & Jha, B. (2012). Detection of Epigenetic Variations in the Protoplast-Derived Germlings of Ulva reticulata Using Methylation Sensitive Amplification Polymorphism (MSAP). *Marine Biotechnology*, *14*(6), 692–700. https://doi.org/10.1007/s10126-012-9434-7

Gupta, V., & Kushwaha, H. R. (2017). Metabolic regulatory oscillations in intertidal green seaweed Ulva lactuca against tidal cycles. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-15994-2

Halling, C., Wikström, S. A., Lilliesköld-Sjöö, G., Mörk, E., Lundsør, E., & Zuccarello, G. C. (2013). Introduction of Asian strains and low genetic variation in farmed seaweeds: Indications for new management practices. *Journal of Applied Phycology*, *25*(1), 89–95. https://doi.org/10.1007/s10811-012-9842-0

He, Y., Hu, C., Wang, Y., Cui, D., Sun, X., Li, Y., & Xu, N. (2018). The metabolic survival strategy of marine macroalga Ulva prolifera under temperature stress. *Journal of Applied Phycology*, *30*(6), 3611–3621. https://doi.org/10.1007/s10811-018-1493-3

He, Y., Shen, S., Yu, D., Wang, Y., Yin, J., Wang, Z., & Ye, Y. (2021). The Ulva prolifera genome reveals the mechanism of green tides. *Journal of Oceanology and Limnology*, *39*(4), 1458–1470. https://doi.org/10.1007/s00343-020-0212-5

He, Y., Wang, Y., Hu, C., Sun, X., Li, Y., & Xu, N. (2019). Dynamic metabolic profiles of the marine macroalga Ulva prolifera during fragmentation-induced proliferation. *PLoS ONE*, *14*(5). https://doi.org/10.1371/journal.pone.0214491

Helmes, R. J. K., López-Contreras, A. M., Benoit, M., Abreu, H., Maguire, J., Moejes, F., & van den Burg, S. W. K. (2018). Environmental impacts of experimental production of lactic acid for bioplastics from Ulva spp. *Sustainability (Switzerland)*, *10*(7). https://doi.org/10.3390/su10072462

Ismail, M. M., & Mohamed, S. E. (2017). Differentiation between some Ulva Spp. by morphological, genetic and biochemical analyses. *Vavilovskii Zhurnal Genetiki i Selektsii*, *21*(3), 360–367. https://doi.org/10.18699/VJ17.253

Jiang, H., Zou, D., Lou, W., Deng, Y., & Zeng, X. (2018). Effects of seawater acidification and alkalization on the farmed seaweed, Pyropia haitanensis (Bangiales, Rhodophyta), grown under different irradiance conditions. *Algal Research*, *31*, 413–420. https://doi.org/10.1016/j.algal.2018.02.033

Jiang, J. L., Zhang, W. Z., Ni, W. X., & Shao, J. W. (2021). Insight on structure-property relationships of carrageenan from marine red algal: A review. In *Carbohydrate Polymers* (Vol. 257). Elsevier Ltd. https://doi.org/10.1016/j.carbpol.2021.117642

Kazi, M. A., Kavale, M. G., & Singh, V. V. (2016). Morphological and molecular characterization of Ulva chaugulii sp. Nov., U. Lactuca and U. Ohnoi (Ulvophyceae, Chlorophyta) from India. *Phycologia*, *55*(1), 45–54. https://doi.org/10.2216/15-11.1

Keesing, J. K., Liu, D., Shi, Y., & Wang, Y. (2016). Abiotic factors influencing biomass accumulation of green tide causing Ulva spp. on Pyropia culture rafts in the Yellow Sea, China. *Marine Pollution Bulletin*, *105*(1), 88–97. https://doi.org/10.1016/j.marpolbul.2016.02.051

Kessler, R. W., Weiss, A., Kuegler, S., Hermes, C., & Wichard, T. (2018). Macroalgal–bacterial interactions: Role of dimethylsulfoniopropionate in microbial gardening by Ulva (Chlorophyta). *Molecular Ecology*, *27*(8), 1808–1819. https://doi.org/10.1111/mec.14472

Kinnby, A., Jonsson, P. R., Ortega-Martinez, O., Töpel, M., Pavia, H., Pereyra, R. T., & Johannesson, K. (2020). Combining an Ecological Experiment and a Genome Scan Show Idiosyncratic Responses to Salinity Stress in Local Populations of a Seaweed. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00470

Kong, S., Chen, Z., Zhao, F., Liu, Q., & Li, J. (2023). *Cultivation of Ulva lactuca changes bacterial community structure and abundances of nitrogen cycling functional genes in an IMTA system*. https://doi.org/10.21203/rs.3.rs-2801471/v1

Kumar, J., Gupta, D. Sen, Gupta, S., Dubey, S., Gupta, P., & Kumar, S. (2017). Quantitative trait loci from identification to exploitation for crop improvement. In *Plant Cell Reports* (Vol. 36, Issue 8, pp. 1187–1213). Springer Verlag. https://doi.org/10.1007/s00299-017-2127-y

Lin, J., Chen, Y., Yan, H., Nahidian, B., Hu, Q., & Han, D. (2020). High-throughput fluorescence-activated cell sorting for cell wall-deficient microalgal mutants screening. *Algal Research*, *50*. https://doi.org/10.1016/j.algal.2020.102011

Liu, F., & Melton, J. T. (2021). Chloroplast Genomes of the Green-Tide Forming Alga Ulva compressa: Comparative Chloroplast Genomics in the Genus Ulva (Ulvophyceae, Chlorophyta). *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.668542

Liu, X., Blomme, J., Bogaert, K. A., D'hondt, S., Wichard, T., Deforce, D., Van Nieuwerburgh, F., & De Clerck, O. (2022). Transcriptional dynamics of gametogenesis in the green seaweed Ulva mutabilis identifies an RWP-RK transcription factor linked to reproduction. *BMC Plant Biology*, *22*(1). https://doi.org/10.1186/s12870-021-03361-3

Liu, X., Bogaert, K., Engelen, A. H., Leliaert, F., Roleda, M. Y., & De Clerck, O. (2017). Seaweed reproductive biology: Environmental and genetic controls. In *Botanica Marina* (Vol. 60, Issue 2, pp. 89–108). Walter de Gruyter GmbH. https://doi.org/10.1515/bot-2016-0091

Magnusson, M., Glasson, C. R. K., Vucko, M. J., Angell, A., Neoh, T. L., & de Nys, R. (2019). Enrichment processes for the production of high-protein feed from the green seaweed Ulva ohnoi. *Algal Research*, *41*. https://doi.org/10.1016/j.algal.2019.101555

Monro, K., Poore, A. G. B., & Brooks, R. (2007). Multivariate selection shapes environment-dependent variation in the clonal morphology of a red seaweed. *Evolutionary Ecology*, *21*(6), 765–782. https://doi.org/10.1007/s10682-006-9150-8

Nakamura, Y., Sasaki, N., Kobayashi, M., Ojima, N., Yasuike, M., Shigenobu, Y., Satomi, M., Fukuma, Y., Shiwaku, K., Tsujimoto, A., Kobayashi, T., Nakayama, I., Ito, F., Nakajima, K., Sano, M., Wada, T., Kuhara, S., Inouye, K., Gojobori, T., & Ikeo, K. (2013). The First Symbiont-Free Genome Sequence of Marine Red Alga, Susabi-nori (Pyropia yezoensis). *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0057122

Neori, A., Bronfman, Y., Van Rijn, J., Guttman, L., Krupnik, N., Shpigel, M., Samocha, T. M., Davis, D. A., Qiu, X., & Abelin, P. (2020). The suitability of Ulva fasciata, Ulva compressa, and Hypnea musciformis for production in an outdoor spray cultivation system, with respect to biomass yield and protein content. *Journal of Applied Phycology*, *32*, 3183–3197.

Nishitsuji, K., Arimoto, A., Iwai, K., Sudo, Y., Hisata, K., Fujie, M., Arakaki, N., Kushiro, T., Konishi, T., Shinzato, C., Satoh, N., & Shoguchi, E. (2016). A draft genome of the brown alga, Cladosiphon okamuranus, S-strain: a platform for future studies of 'mozuku' biology. *DNA Research*, *23*(6), 561–570. https://doi.org/10.1093/dnares/dsw039

Ochiai, K. K., Hanawa, D., Ogawa, H. A., Tanaka, H., Uesaka, K., Edzuka, T., Shirae-Kurabayashi, M., Toyoda, A., Itoh, T., & Goshima, G. (2024). Genome sequence and cell biological toolbox of the highly regenerative, coenocytic green feather alga *Bryopsis. The Plant Journal*. https://doi.org/10.1111/tpj.16764

Oliveira, W. da S., Moreira, B. R., Rörig, L., Horta, P. A., Treichel, H., & Bonomi-Barufi, J. (2024). Modelling bioremediation of contaminated effluents by Ulva ohnoi. – A predictive perspective. *Environmental Pollution*, *347*. https://doi.org/10.1016/j.envpol.2024.123689

Osorio, H., Tapia-Reyes, P., Espinoza, D., Laporte, D., González, A., Castro-Nallar, E., & Moenne, A. (2022). The Genome of the Marine Alga Ulva compressa (Chlorophyta) Reveals Protein-Coding Genes with Similarity to Plants and Green Microalgae, but Also to Animal, Bacterial, and Fungal Genes. *International Journal of Molecular Sciences*, *23*(13). https://doi.org/10.3390/ijms23137279

Paiva, L., Lima, E., Neto, A. I., Marcone, M., & Baptista, J. (2017). Nutritional and Functional Bioactivity Value of Selected Azorean Macroalgae: Ulva compressa, Ulva rigida, Gelidium microdon, and Pterocladiella capillacea. *Journal of Food Science*, *82*(7), 1757–1764. https://doi.org/10.1111/1750-3841.13778

Potter, E. E., Thornber, C. S., Swanson, J.-D., & McFarland, M. (2016). Ploidy distribution of the harmful bloom forming macroalgae Ulva spp. in Narragansett Bay, Rhode Island, USA, using flow cytometry methods. *Plos One*, *11*(2), e0149182.

Provan, J., Glendinning, K., Kelly, R., & Maggs, C. A. (2012). *Levels and patterns of population genetic diversity in the red seaweed Chondrus crispus (Florideophyceae): a direct comparison of single nucleotide polymorphisms and microsatellites*. https://academic.oup.com/biolinnean/article/108/2/251/2452745

Roque, B. M., Salwen, J. K., Kinley, R., & Kebreab, E. (2019). Inclusion of Asparagopsis armata in lactating dairy cows' diet reduces enteric methane emission by over 50 percent. *Journal of Cleaner Production*, *234*, 132–138. https://doi.org/10.1016/j.jclepro.2019.06.193

Rouphael, Y., Spíchal, L., Panzarová, K., Casa, R., & Colla, G. (2018). High-throughput plant phenotyping for developing novel biostimulants: from lab to field or from field to lab? *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.01197

Rybak, A. S. (2021). Freshwater macroalga, Ulva pilifera (Ulvaceae, Chlorophyta) as an indicator of the trophic state of waters for small water bodies. *Ecological Indicators*, *121*. https://doi.org/10.1016/j.ecolind.2020.106951

Sanchez-Arcos, C., Paris, D., Mazzella, V., Mutalipassi, M., Costantini, M., Buia, M. C., von Elert, E., Cutignano, A., & Zupo, V. (2022). Responses of the Macroalga Ulva prolifera Müller to Ocean Acidification Revealed by Complementary NMR- and MS-Based Omics Approaches. *Marine Drugs*, *20*(12). https://doi.org/10.3390/md20120743

Santure, A. W., & Garant, D. (2018). Wild GWAS—association mapping in natural populations. *Molecular Ecology Resources*, *18*(4), 729–738. https://doi.org/10.1111/1755-0998.12901

Sebök, S., & Hanelt, D. (2023). Cultivation of the brackish-water macroalga Ulva lactuca in wastewater from land-based fish and shrimp aquacultures in Germany. *Aquaculture*, *571*. https://doi.org/10.1016/j.aquaculture.2023.739463

Shan, T., Pang, S., Li, J., Li, X., & Su, L. (2015). Construction of a high-density genetic map and mapping of a sex-linked locus for the brown alga Undaria pinnatifida (Phaeophyceae) based on large scale marker development by specific length amplified fragment (SLAF) sequencing. *BMC Genomics*, *16*(1). https://doi.org/10.1186/s12864-015-2184-y

Shefer, S., Israel, A., Golberg, A., & Chudnovsky, A. (2017). Carbohydrate-based phenotyping of the green macroalga Ulva fasciata using near-infrared spectrometry: Potential implications for marine biorefinery. *Botanica Marina*, *60*(2), 219–228. https://doi.org/10.1515/bot-2016-0039

Sonchaeng, U., Wongphan, P., Pan-utai, W., Paopun, Y., Kansandee, W., Satmalee, P., Tamtin, M., Kosawatpat, P., & Harnkarnsujarit, N. (2023). Preparation and Characterization of Novel Green Seaweed Films from Ulva rigida. *Polymers*, *15*(16). https://doi.org/10.3390/polym15163342

Spoerner, M., Wichard, T., Bachhuber, T., Stratmann, J., & 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. https://doi.org/10.1111/j.1529-8817.2012.01231.x

Steinhagen, S., Barco, A., Wichard, T., & Weinberger, F. (2019). Conspecificity of the model organism Ulva mutabilis and Ulva compressa (Ulvophyceae, Chlorophyta). *Journal of Phycology*, *55*(1), 25–36. https://doi.org/10.1111/jpy.12804

Steinhagen, S., Weinberger, F., & Karez, R. (2019). Molecular analysis of Ulva compressa (Chlorophyta, Ulvales) reveals its morphological plasticity, distribution and potential invasiveness on German North Sea and Baltic Sea coasts. *European Journal of Phycology*, *54*(1), 102–114. https://doi.org/10.1080/09670262.2018.1513167

Stratmann, J., Paputsoglu, G., & Oertel, W. (1996). Differentiation of Ulva mutabilis (Chlorophyta) gametangia and gamete release are controlled by extracellular inhibitors. *Journal of Phycology*, *32*(6), 1009–1021. https://doi.org/10.1111/j.0022-3646.1996.01009.x

Tadmor Shalev, N., Ghermandi, A., Tchernov, D., Shemesh, E., Israel, A., & Brook, A. (2022). NIR spectroscopy and artificial neural network for seaweed protein content assessment in-situ. *Computers and Electronics in Agriculture*, *201*. https://doi.org/10.1016/j.compag.2022.107304

Torres, A. C., Veiga, P., Rubal, M., & Sousa-Pinto, I. (2015). The role of annual macroalgal morphology in driving its epifaunal assemblages. *Journal of Experimental Marine Biology and Ecology*, *464*, 96–106. https://doi.org/10.1016/j.jembe.2014.12.016

Tran, L. A. T., Vieira, C., Steinhagen, S., Maggs, C. A., Hiraoka, M., Shimada, S., Van Nguyen, T., De Clerck, O., & Leliaert, F. (2022). An appraisal of Ulva (Ulvophyceae, Chlorophyta) taxonomy. *Journal of Applied Phycology*. https://doi.org/10.1007/s10811-022-02815-x

Tuya, F., Png-Gonzalez, L., Riera, R., Haroun, R., & Espino, F. (2014). Ecological structure and function differs between habitats dominated by seagrasses and green seaweeds. *Marine Environmental Research*, *98*, 1–13. https://doi.org/10.1016/j.marenvres.2014.03.015

van der Loos, L. M., D'hondt, S., Willems, A., & De Clerck, O. (2021). Characterizing algal microbiomes using long-read nanopore sequencing. *Algal Research*, *59*. https://doi.org/10.1016/j.algal.2021.102456

Vargas Cárdenas, J., Chávez Pérez, J., Martínez Ordinola, N., Soto Rodríguez, I., Brito, L. O., Peixoto, S. R. M., & Galvez, A. O. (2023). Phytochemical screening and antibacterial assessment of two macroalgae Ulva papenfussi and Ulva nematoidea (Chlorophyta) against the bacterium Vibrio parahaemolyticus. *Food Science and Technology International*, 10820132231165540.

Vieira, C., De Ramon N'Yeurt, A., Rasoamanenndrika, F. A., D'Hondt, S., Tran, L. A. T., Van den Spiegel, D., Kawai, H., & De Clerck, O. (2021). Marine macroalgal biodiversity of northern Madagascar: morpho-genetic systematics and implications of anthropic impacts for conservation. *Biodiversity and Conservation*, *30*(5), 1501–1546. https://doi.org/10.1007/s10531-021-02156-0

Wang, C., Yu, R., & Zhou, M. (2011). Acute toxicity of live and decomposing green alga Ulva (Enteromorpha) prolifera to abalone Haliotis discus hannai. *Chinese Journal of Oceanology and Limnology*, *29*(3), 541–546. https://doi.org/10.1007/s00343-011-0126-3

Wang, X., Chen, Z., Li, Q., Zhang, J., Liu, S., & Duan, D. (2018). High-density SNP-based QTL mapping and candidate gene screening for yield-related blade length and width in Saccharina japonica (Laminariales, Phaeophyta). *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32015-y

Wang, X., Yang, X., Yao, J., Li, Q., Lu, C., & Duan, D. (2023). Genetic linkage map construction and QTL mapping of blade length and width in Saccharina japonica using SSR and SNP markers. *Frontiers in Marine Science*, *10*. https://doi.org/10.3389/fmars.2023.1116412

Wichard, T., Charrier, B., Mineur, F., Bothwell, J. H., De Clerck, O., & Coates, J. C. (2015). The green seaweed Ulva: A model system to study morphogenesis. *Frontiers in Plant Science*, *6*(FEB). https://doi.org/10.3389/fpls.2015.00072

Wood, G., Steinberg, P. D., Campbell, A. H., Vergés, A., Coleman, M. A., & Marzinelli, E. M. (2022). Host genetics, phenotype and geography structure the microbiome of a foundational seaweed. *Molecular Ecology*, *31*(7), 2189–2206. https://doi.org/10.1111/mec.16378

Wu, H., Shin, S. K., Jang, S., Yarish, C., & Kim, J. K. (2018). Growth and nutrient bioextraction of gracilaria chorda, G. Vermiculophylla, ulva prolifera, and U. compressa under hypo-and hyper-osmotic conditions. *Algae*, *33*(4), 329–340. https://doi.org/10.4490/algae.2018.33.11.13

Xia, Z., Yuan, H., Liu, J., Sun, Y., Tong, Y., Zhao, S., Xia, J., Li, S., Hu, M., Cao, J., Zhang, J., & He, P. (2022). A review of physical, chemical, and biological green tide prevention methods in the Southern Yellow Sea. In *Marine Pollution Bulletin* (Vol. 180). Elsevier Ltd. https://doi.org/10.1016/j.marpolbul.2022.113772

Xu, Y., Huang, L., Ji, D., Chen, C., Zheng, H., & Xie, C. (2015). Construction of a dense genetic linkage map and mapping quantitative trait loci for economic traits of a doubled haploid population of Pyropia haitanensis (Bangiales, Rhodophyta). *BMC Plant Biology*, *15*(1). https://doi.org/10.1186/s12870-015-0604-4

Yabe, T., Ishii, Y., Amano, Y., Koga, T., Hayashi, S., Nohara, S., & Tatsumoto, H. (2009). Green tide formed by free-floating Ulva spp. at Yatsu tidal flat, Japan. *Limnology*, *10*(3), 239–245. https://doi.org/10.1007/s10201-009-0278-4

Ye, N. hao, Zhang, X. wen, Mao, Y. ze, Liang, C. wei, Xu, D., Zou, J., Zhuang, Z. meng, & Wang, Q. yin. (2011). "Green tides" are overwhelming the coastline of our blue planet: Taking the world's largest example. In *Ecological Research* (Vol. 26, Issue 3, pp. 477–485). https://doi.org/10.1007/s11284-011-0821-8

Ye, N., Zhang, X., Miao, M., Fan, X., Zheng, Y., Xu, D., Wang, J., Zhou, L., Wang, D., Gao, Y., Wang, Y., Shi, W., Ji, P., Li, D., Guan, Z., Shao, C., Zhuang, Z., Gao, Z., Qi, J., & Zhao, F. (2015). Saccharina genomes provide novel insight into kelp biology. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms7986

Yoshinaga, T., Nishiduka, H., & Nanba, N. (2014). Genotype analysis of commercial products of the soft seaweed Undaria pinnatifida (Laminariales, Alariaceae). *Coast Mar Sci*, *37*, 9–15.

Zhang, Y., He, P., Li, H., Li, G., Liu, J., Jiao, F., Zhang, J., Huo, Y., Shi, X., Su, R., Ye, N., Liu, D., Yu, R., Wang, Z., Zhou, M., & Jiao, N. (2019). Ulva prolifera green-tide outbreaks and their environmental impact in the Yellow Sea, China. In *National Science Review* (Vol. 6, Issue 4, pp. 825–838). Oxford University Press. https://doi.org/10.1093/nsr/nwz026

Appendices

Appendix A. *Ulva* culture media solutions ingredients.

	Name	Linear formula	g per L
Solution I	Sodium chloride	NaCl	19,14
	Sodium sulphate decahydrate	Na2SO4·10H2O	7,28
	Magnesium chloride hexahydrate	MgCl2·6H2O	8,68
	Calcium chloride hexahydrate	CaCl2·6H2O	1,24
	Sodium nitrate	NaNO3	0,085
	Ammonium sulfate	(NH4)2SO4	0,0066
Add 10ml to 1L of Solution I	Name	Linear formula	g per L
Solution II	Sodium phosphate monohydrate	NaH2PO4·H2O	0,7
	Sodium bicarbonate	NaHCO3	8,8
	Tris(hydroxymethyl)-aminomethan (trisbuffer)	Tris-OH / NH2C(CH2OH)3	10
Add 10ml to 1L of Solution I,II	Name	Linear formula	g per L
Solution III	Potassium bromide	KBr	7,84
	Potassium chloride	KCl	54,2
	Strontium chloride hexahydrate	SrCl2·6H2O	1,95
Add 10ml to 1L of Solution I-III	Name (+ stock concentration)	Linear formula	g per L
Solution IV pH8.0 (NaOH)	Tritiplex® II (Ethylenediamine tetraacetic acid)	EDTA	0,6684
	Boric acid	H3BO3	1,14

	Iron (II) sulphate heptahydrate	FeSO4·7H2O	0,199
	Copper (II) sulphate pentahydrate	CuSO4·5H2O	0,0039
	Sodium molybdate dihydrate	Na2MoO4·2H2O	0,0126
	Magnesium chloride tetrahydrate	MnCl2·4H2O	0,036
	Zinc sulphate heptahydrate	ZnSO4·7H2O	0,044
	Cobalt (II) nitrate hexahydrate	Co(NO3)2·6H2O	0,0033
	Ammonium metavanadate	NH4VO3	0,0023
	Potassium Iodide (100mM)	KI	100μL
	Sodium selenite pentahydrate (100mM)	Na2SeO3·5H2O	15μL
	Arsenic oxide (100mM)	As2O3	0,47μL
	Sodium tungstate dihydrate (100mM)	Na2WoO4·2H2O	0,2μL
	Tellurium oxide (100mM)	TeO2	0,21μL

Add 2ml to 1L of Solution I-IV	Name (+ stock concentration)	Linear formula	g/μL per L
Solution V	Vitamin B12 (136mg/10mL)	B12	3,68μL
2mL	Thiamine hydrochloride/vitamin B1	C12H18Cl2N4OS	0,2
	Niacin/nicotinic acid/vitamin B3	C6H5NO2	0,1
	Ca-pantothenate	C18H32CaN2O10	0,1
	Pyridoxine·HCl/pyridoxine hydrochloride/vit B6	C8H12ClNO3	0,04
	p-aminobenzoic acid/vitamin B10	C7H7NO2	0,01
	Biotin/vitamin H	C10H16N2O3S	0,005
	Thymine	C5H6N2O2	0,8
	Inositol	C6H12O6	1
	Orotic acid	C5H4N2O4	0,26
	Folinic acid (citrovorum) 200mg/10mL	C20H23N7O7	10μL
	Folic acid/vitamin M	C19H19N7O6	0,0025
	Putrescine dihydrochloride/putrescine·2HCl	C4H14Cl2N2	0,04
	Riboflavin/vitamin B2	C17H20N4O6	0,005
	Pyridoxamine dihydrochloride/pyridoxamine·2HCl	C8H14Cl2N2O2	0,02
	Choline chloride/choline·Cl	C5H14ClNO	0,36

Appendix B. CTAB DNA extraction procedure and PCR protocol for mating type determination

2. **METHOD:**

- Put a bullet (4 mm) in each microtube.
- Make sure that the microtubes are cooled on dry ice.
- Harvest your sample and put it immediately in the collection microtube.
- Grind the material with the RETSCH-machine (time :30 sec, frequency : 30/sec).
- Add 400 µl of CTAB-buffer and mix manually.
- Put the samples in an oven at 60 °C for 1 hour.
- Cool the samples till room-temperature (on ice).
- Add 250 µl CHCl₃/IAA.
- Mix manually for 1 min.
- Centrifuge for 10 min at 4000 rpm.
- Put 200 µl isopropanol in a new microtube. (no bullet!)
- Add 200 µl of supernatans to the microtube with isopropanol.
- Mix well.
- Centrifuge for 15 min at 4000 rpm.
- Discard the supernatans.
- Wash with 500-800 µl of EtOH 75 %.
- Centrifuge for 5 min at 4000 rpm.
- Discard the supernatans carefully.
- Dry the pellet for 10 min at 60 °C.
- Resuspend the DNA in 100 µl TE₁₀₀.₁.
- Transfer the DNA to a new microtiterplate.

- Agarose-gel :
    - 1 % agarose / 0.5 x TAE
    - 10 µl DNA + 2 µl 6x Loading-dye (Big fragments)
    - 100 V – 15 min
- Or check the DNA with the Nanodrop.

Annex, table 3: PCR standard program: used to amplify DNA with primers specific for *mt+ #2*, *mt- #1* and *mt- #2*.

	Temperature	Time	Step
1	95°C	3 min.	Initial
2	95°C	30 sec.	Denaturation
3	60°C	30 sec.	Annealing
4	72°C	1 min.	Elongation
5	40 x		Repetition from step 2
6	72°C	5 min.	Finishing
7	16°C	∞	Infinite hold

Annex, table 4: PCR program adjusted for the use of the primers specific *mt+ #1*.

	Temperature	Time	Step
1	95°C	3 min.	Initial
2	95°C	1 min.	Denaturation
3	50°C	1 min.	Annealing
4	72°C	1 min.	Elongation
5	40 x		Repetition from step 2
6	72°C	5 min.	Finishing
7	16°C	∞	Infinite hold

With Phire:

Table 2. Cycling protocol

Cycle step	2-step		3-step		Cycles
	Temp.	Time	Temp.	Time	
Initial denaturation	98 °C	5 min	98 °C	5 min	1
Denaturation	98 °C	5 s	98 °C	5 s	
Annealing (see 6.3)	-	-	X°C	5 s	
Extension (see 6.4)	72 °C	20 s ≤1 kb 20 s/kb >1 kb	72 °C	20 s ≤1 kb 20 s/kb >1 kb	40
Final Extension	72 °C +4 °C	1 min hold	72 °C +4 °C	1 min hold	1

X=55

Appendix C. 44 multiplex primers sequence

Code	Forward	Reverse	Gene ID	SNPs-indels
MP1	GGATGTTGTTTGGGCCAAGG	CCGTGTAGACATCATCGCCA	UMWT001_0240.1	GC and GA
MP2	TCCTAGCCGAAATGTCCACG	ATGGCGAAGAGGAGGAGGT	UMWT001_0248.1	TC and +C

MP3	CACTGTCCGCGCTACAAATC	CGCCGTTACCATTGTTTGCA	UMWT001_0424.1	AG, CT-TC, TG-GT, CATG-TATT, GA, GA, TTC-GTT
MP4	ACTGAGAGTTGCTTGGCGAG	CCCATCATCCAGCAGCACTT	UMWT001_0470.1	TGGCGGT-GGGGGGG, +G
MP5	GAGGTGTTTGTTGGGGTGGT	CTCCCGTGATGGTGCCTG	UMWT001_0553.1	"+G"
MP6	GTCTCCACCGTCTGCATCAT	GTCGTCTACCGCGTCCAAC	UMWT002_0093.1	"+C", +C
MP7	CCACTCTGCCCCAATCACAA	CCGGCACCTGTTTCAACAAG	UMWT002_0193.1	CG, GT
MP8	TGTACAGCTTGGACAGGCAG	TTCTCCGCGATGAGGTCATG	UMWT002_0238.1	"+G"
MP9	TGTCGACTCTGACGGTTTCG	GGAGGAGCTGATCGCAGG	UMWT002_0463.1	"+C"
MP10	GCTGCCATGCTGGAGAACT	ATTTGTTGCACACACCTGGC	UMWT002_0587.1	"+C"
MP11	GCGAGCTCTTAGTGACAGCT	TACAGATTCGGTGCGTTGCT	UMWT002_0643.1	GA, TC, GC
MP12	CTTCAGCCTGCAAGAGGTGA	GTATACTTCCGTCCCCGCAG	UMWT003_0205.1	TA, CTG-TTT
MP13	CCTTGAACGCGTGCTTCAG	CGGGAGCTGGGTGTAGATGT	UMWT003_0412.1	"+C", +C
MP14	GGCTCGCAGCAGAACATG	CAGAGGTCCTGCTCGTGATG	UMWT003_0499.2	"+G"
MP15	GGAGGAGTCGTCAGTGATGC	CTGCCTGTCTTCCCCGAAG	UMWT003_0516.1	"+G", +G
MP16	TCAACTACACATGGGCCACG	TACTGCCCCTGAACGGAAAC	UMWT004_0115.1	AG, GCTCG-ACTCC
MP17	TTGCCGCGGATGCATGTG	GCAGGCGAAGGACCTAACAG	UMWT004_0195.1	"+G"
MP18	ATTCTTCCCAGTTCTGGCCG	TTCGTACAGGCGAAAGTGCT	UMWT004_0359.1	AT
MP19	GGACGAAGAGGACGGTCAAG	TTTCGTCTGGGAACTGCAGG	UMWT005_0045.1	AGGGGGGCGGTGGGGGTGCGGA - AGGGGGGGCGGTGGGGGGTGCGGA
MP20	GAGAGGACGGCGATGAACTG	CAGCAGTACGTCGTCGACTG	UMWT005_0164.1	"+C"
MP21	CAACCAGAGCCCACAGGTG	GGAGGTAATGCATGGCTGGA	UMWT005_0311.1	"+C"
MP22	GCAGAATGAGGCCAGGTGTA	GCAAGGTGCCCGCAAATC	UMWT005_0391.1	"+G", GA-GGAG

Appendix D. DNA purification procedure by Magnetic Bead

Magnetic bead purification

We use HighPrep™ PCR (MAGBIO AC-60050) for the post PCR cleanup

- Shake thoroughly the HighPrep™ PCR reagent to fully resuspend the magnetic beads
- Add 9 µL of the HighPrep™ PCR (1,8 x) to 5 µL of the sample (remainder after gel analysis)

- Follow protocol
  [http://www.magbiogenomics.com/image/data/Literature/Protocols/HighPrep%20PCR%20Protocol.pdf]

For the sequencing, mention in the remark or to Wilson that this PCR product is purified (this sequencing is € 0,80/sample cheaper)

HighPrep™ PCR Protocol
Materials:
- PCR MAGBIO AC-60050 beads (fridge in the stockroom, same fridge as the antibiotics), normally 1 tube per group in use OR - HighPrep PCR magnetic beads
- PCR product (20-50 µl)
- 96-well magnetic rack
- Elution buffer: reagent grade water, 10 mM Tris-HCl, etc.
- 80% ethanol
- Long narrow tips

* Bring the HighPrep PCR to room temperature for at least 30 min before use.

1. Shake thoroughly the HighPrep PCR reagent (magnetic beads) to fully resuspend the magnetic beads.

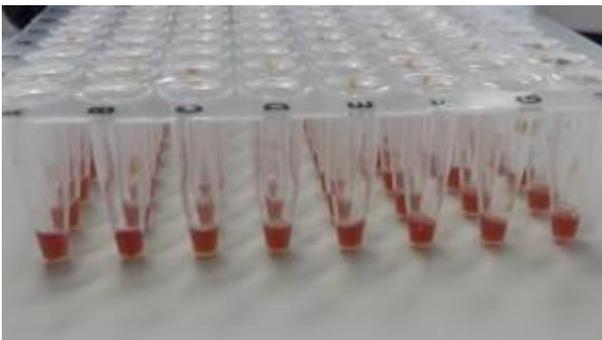2. A. Transfer PCR reaction to appropriate 96-well plate.
OR

   B. Use PCR strip tubes or plate you used for PCR

3. Add 1.8 ul of magnetic beads to 1 ul of PCR reaction.
For example: Add 9 µL of the HighPrep™ PCR (1,8 x) to 5 µL of the sample.
- Mix thoroughly the HighPrep PCR reagent and PCR sample by mix pipetting up and down 6-8 times.

- Incubate the mixture for 5 minutes at room temperature.
4. Place the sample plate on the 96 magnetic separation rack for 3 minutes or until the solution clears. Beads will stick to the side of the well towards the magnet.
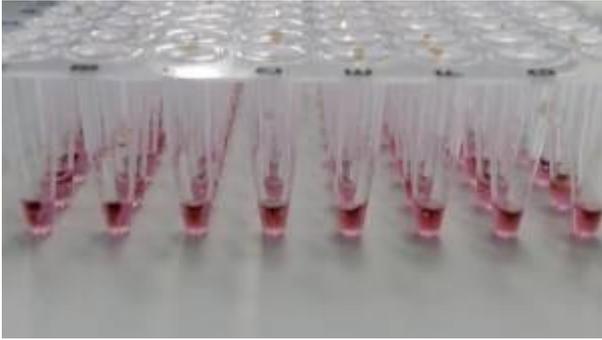
5. With the sample plate still on the magnet, remove and discard the supernatant by pipetting using long narrow tips.
Make sure not to disturb the attracted beads while aspirating the supernatant.
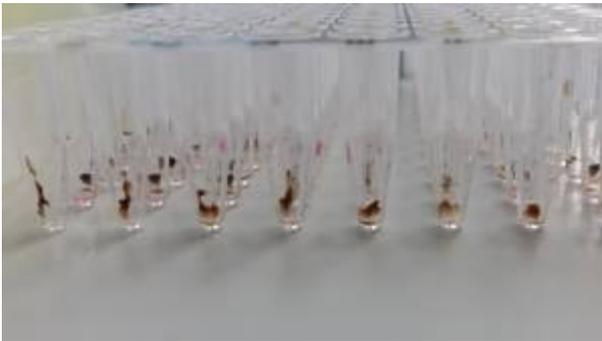
6. With the sample plate still on the magnet, immediately add 150-200µL of 80% ethanol to each well and incubate for 30 seconds at room temperature.
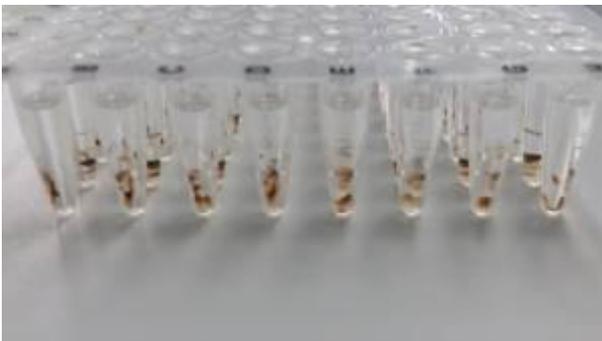
7. With the sample plate still on the magnet, remove and discard the supernatant by pipetting.

8. Repeat steps 6-7 for a total of two 80% ethanol washes. Try to remove as much residual ethanol as possible by pipetting (this makes it dry faster)

9. Dry the beads by incubating the plate for 10-15 minutes at room temperature with the plate still on the magnetic separation device.
It is critical to completely remove all traces of alcohol but take caution in not over-drying the beads as this will reduce the yield.
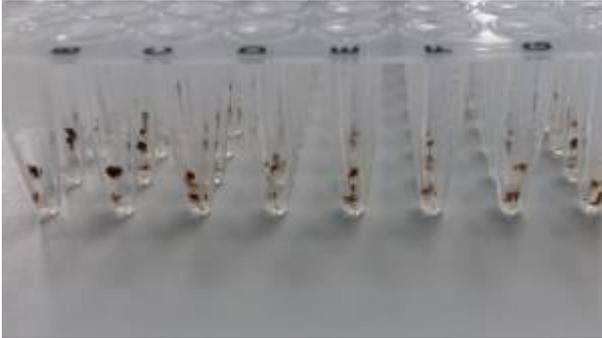
10. Remove the sample plate from the magnetic separation rack. Add 40µL of elution buffer (reagent grade water, 10 mM Tris-HCl, etc.) to each well and pipet up and down 5 times to mix. Note: Using less elution buffer will not give you a more concentrated product.
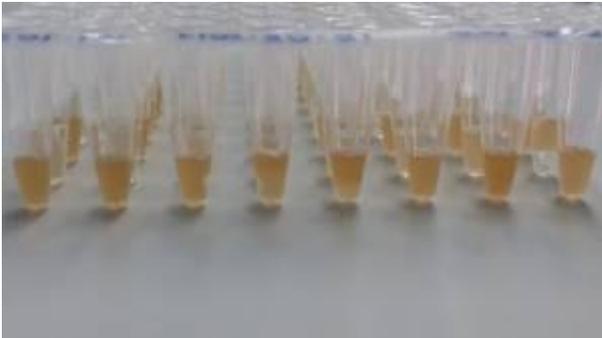

- Incubate for 2 minutes at room temperature.

11. Place the sample plate back on the magnetic separation rack and wait 1 minute or until the magnetic beads clear from the solution.
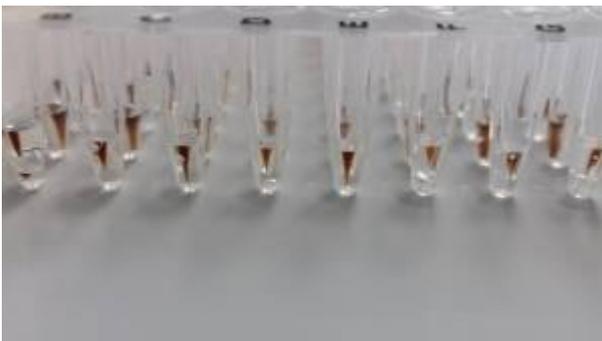
12. Transfer the eluate (clear supernatant) to a new plate (or new Eppendorf tubes) for storage or for subsequent applications.
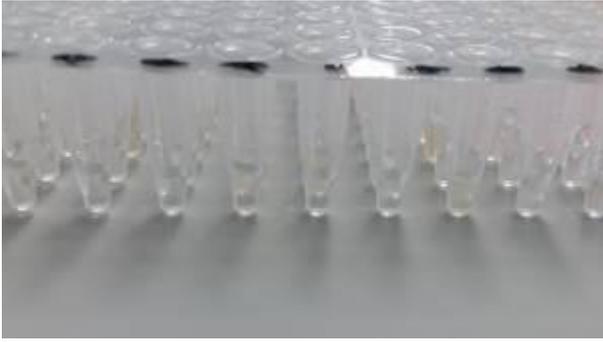
Appendix E. Platinum™ Multiplex PCR Master Mix procedures

Before starting

- Design primers for multiplex PCR. See Appendix B, "Design, verify and synthesize the primers" for details.
- Verify the primers produce a single band of the correct size using singleplex PCR. See "Guidelines for primer verification" on page 15 for details.

Prepare primer mix

1. Combine all primers for the multiplex PCR reaction in one tube, and adjust the final concentration to 0.5 μM per primer using 0.1✕ TE (10 mM Tris-HCl, 0.1 mM EDTA) buffer.

2. Dispense the primer mix into aliquots. Store at –20°C for up to one year or 4°C for up to two months.

Preparations for PCR

- Before you begin PCR, review "PCR good laboratory practices" on page 16.
- Select an instrument and reaction plate
- Calculate the number of reactions needed to perform each assay, including no-template control (NTC) reactions. Add at least one extra reaction for every 10 required reactions to account for volume loss from pipetting. For example, for a 96-well plate, prepare enough volume for approximately 110 reactions.
- Multiple PCR reactions can be run on one reaction plate, but include controls for each run on the plate.

Prepare the PCR Reaction Mix

1. Thaw the primer mixes, templates, GC Enhancer (optional), and Platinum™ Multiplex PCR Master Mix.

2. Mix the Platinum™ Multiplex PCR Master Mix by gently inverting the tube 10 times, then place on ice.

   IMPORTANT! Avoid bubble formation during mixing.

3. Mix the remaining reagents by inverting each tube a few times, then centrifuge the contents briefly. Place the tubes on ice.

4. Combine the following components for the number of reactions required, plus an additional 10% of volume.

Component	Volume per 50-μL reaction	Final concentration
Platinum™ Multiplex PCR Master Mix, 2X	25 μL	1X
Primer mix (0.5 μM each primer)	5–10 μL [1]	50–100 nM each primer
(*Optional*) GC Enhancer	0–10 μL [2]	0–20%
Nuclease-free Water	50 μL minus template DNA volume[3]	—

[1] Use 10 μL of primer mix (100 nM final concentration of each primer) when less than 0.1 μg of DNA is used.
[2] For targets with 65 to 75% GC, use 6 μL of GC Enhancer in a 50-μL reaction. For targets with >75% GC, start with 10 μL in a 50-μL reaction. See "Enhancer concentration" on page 17 for details.
[3] This protocol assumes that identical volumes of template DNA will be used for each reaction.

5. Mix the PCR Reaction Mix by inverting the tube a few times.

6. Centrifuge briefly to bring the PCR Reaction Mix to the bottom of the tube and eliminate air bubbles.

Prepare the PCR reaction plate

1. Dispense appropriate volumes (50 μL minus the template DNA volume) of the PCR Reaction Mix to the wells of a reaction plate or PCR tubes.

2. Add 0.1–0.2 μg of DNA template to the PCR Reaction Mix to bring the final volume to 50 μL.

   Note: Adjust the final reaction volume of each no-template control well or tube to 50 μL with Nuclease-free Water.

3. Seal the reaction plate with optical adhesive film, or cap the tubes.

4. Mix the contents well by inverting the plates or tubes a few times.

   Note: Inverting the plate gives more uniform mixing across the reaction plate than vortexing.

5. Centrifuge briefly to bring the contents to the bottom of the wells or tubes and eliminate air bubbles.

Set up and run the PCR instrument

See the appropriate instrument user guide for detailed instructions to program the thermal-cycling conditions or to run the plate.

1. Select the appropriate amplification protocol based on your analysis method.

   Amplification protocol for analysis by agarose gel electrophoresis or Agilent™ Bioanalyzer™ Instrument.

Step	Cycles	Temperature	Time
Activation	1	95°C	2 min
Denaturation		95°C	30 s
Annealing	30–40	60°C	90 s
Extension		72°C	60 s/kb of largest amplicon
Final extension	1	72°C	10 min
Final hold	1	4°C	Hold

   Amplification protocol for analysis by capillary gel electrophoresis.

Step	Cycles	Temperature	Time
Activation	1	95°C	2 min
Denaturation		95°C	30 s
Annealing	25–40	60°C	90 s
Extension		72°C	60 s/kb of largest amplicon
Final extension	1	60°C	30 min
Final hold	1	4°C	Hold

2. Load the reaction plate or tubes into the PCR instrument, then start the run.

Code	Forward	Reverse	Gene ID	SNPs-indels	amplicon size	simplex PCR SI	PCR WT	SNP	multiplex reads SI	reads WT	SNP
MP1	GGATGTTGTTTGGGCCAAGG	CCGTGTAGACATCATCGCCA	UMWT001_0240.1	GC and GA	228	ok	ok	no	39	36	no
MP2	TCCTAGCCGAAATGTCCACG	ATGGCGAAGAGGAGGAGGT	UMWT001_0248.1	TC and +C	229	ok	ok	no / yes	0	0	
MP3	CACTGTCCGCGCTACAAATC	CGCCGTTACCATTGTTTGCA	UMWT001_0424.1	AG, CT-TC, TG-GT, CATG-TATT, GA, GA, TTC-GTT	231	failed	failed		0	0	
MP4	ACTGAGAGTTGCTTGGCGAG	CCCATCATCCAGCAGCACTT	UMWT001_0470.1	TGGCGGT-GGGGGGG, +G	244	ok	ok	no / yes	2	0	ND
MP5	GAGGTGTTTGTTGGGGTGGT	CTCCCGTGATGGTGCCTG	UMWT001_0553.1	"+G"	238	failed	failed		0	0	
MP6	GTCTCCACCGTCTGCATCAT	GTCGTCTACCGCGTCCAAC	UMWT002_0093.1	"+C", +C	226	ok	failed	ND / yes	0	0	
MP7	CCACTCTGCCCCAATCACAA	CCGGCACCTGTTTCAACAAG	UMWT002_0193.1	CG, GT	221	ok	ok	ND / no	56	34	no
MP8	TGTACAGCTTGGACAGGCAG	TTCTCCGCGATGAGGTCATG	UMWT002_0238.1	"+G"	224	ok	ok	yes	21	9	mix
MP9	TGTCGACTCTGACGGTTTCG	GGAGGAGCTGATCGCAGG	UMWT002_0463.1	"+C"	232	ok	ok	yes	21	20	yes
MP10	GCTGCCATGCTGGAGAACT	ATTTGTTGCACACACCTGGC	UMWT002_0587.1	"+C"	245	ok	failed	no	14	7	yes
MP11	GCGAGCTCTTAGTGACAGCT	TACAGATTCGGTGCGTTGCT	UMWT002_0643.1	GA, TC, GC	235	ok	ok	no	57	29	no
MP12	CTTCAGCCTGCAAGAGGTGA	GTATACTTCCGTCCCCGCAG	UMWT003_0205.1	TA, CTG-TTT	220	ok	ok	no	36	23	no
MP13	CCTTGAACGCGTGCTTCAG	CGGGAGCTGGGTGTAGATGT	UMWT003_0412.1	"+C", +C	236	failed	failed		3	6	mix
MP14	GGCTCGCAGCAGAACATG	CAGAGGTCCTGCTCGTGATG	UMWT003_0499.2	"+G"	249	failed	failed		0	0	
MP15	GGAGGAGTCGTCAGTGATGC	CTGCCTGTCTTCCCCGAAG	UMWT003_0516.1	"+G", +G	250	failed	failed		0	0	
MP16	TCAACTACACATGGGCCACG	TACTGCCCCTGAACGGAAAC	UMWT004_0115.1	AG, GCTCG-ACTCC	243	ok	ok	no	48	41	no
MP17	TTGCCGCGGATGCATGTG	GCAGGCGAAGGACCTAACAG	UMWT004_0195.1	"+G"	225	ok	ok	yes	26	23	mix
MP18	ATTCTTCCCAGTTCTGGCCG	TTCGTACAGGCGAAAGTGCT	UMWT004_0359.1	AT	241	ok	ok	no	33	8	no
MP19	GGACGAAGAGGACGGTCAAG	TTTCGTCTGGGAACTGCAGG	UMWT005_0045.1	AGGGGGGGCGGTGGGGGTGCGGA - AGGGGGGGCGGTGGGGGGTGCGGA	225	ok	ok	yes	0	0	
MP20	GAGAGGACGGCGATGAACTG	CAGCAGTACGTCGTCGACTG	UMWT005_0164.1	"+C"	229	ok	ok	yes	41	32	yes
MP21	CAACCAGAGCCCACAGGTG	GGAGGTAATGCATGGCTGGA	UMWT005_0311.1	"+C"	250	ok	ok	yes / ND	47	28	yes
MP22	GCAGAATGAGGCCAGGTGTA	GCAAGGTGCCCGCAAATC	UMWT005_0391.1	"+G", GA-GGAG	223	ok	ok	yes	0	0	
								total reads	621	351	
								total reads mapped	444	55	

no / yes = 2 SNPs in the amplicon, one is no and the other is yes
no = both SI and WT reads are similar to WT sequence
yes = both SI and WT reads have the same SNP compared to the WT genome
mix = a mix of sequences with and without the SNP. Probably caused by a region that is difficult to sequence.
ND = not possible to determine: low quality read or too few reads to properly analyses

Appendix G. Codes

Code for the chi-square test

a.) for the mating type difference (R)
```
# Observed data
observed_counts <- c(85, 70)
# Expected data for a 50:50 distribution
total_counts <- sum(observed_counts)
expected_counts <- c(total_counts / 2, total_counts / 2)
# Perform the Chi-square test
chi_square_test <- chisq.test(observed_counts, p = expected_counts / total_counts)
# Print the test results
print(chi_square_test)
```

b.) for the mating type linked with rhizoid (Python)
```
import pandas as pd
from scipy.stats import chi2_contingency
# Creating the contingency table
data = [[31, 36, 5],
        [34, 44, 4],
        [5, 5, 2]]
# Defining row and column labels
row_labels = ['Fully Developed', 'Not Developed', 'Difficult to Determine']
col_labels = ['mt+', 'mt-', 'mt?']
# Creating a DataFrame
df = pd.DataFrame(data, columns=col_labels, index=row_labels)
# Performing the chi-square test
```

```python
chi2, p, dof, ex = chi2_contingency(df)
chi2, p, dof, ex
```

c.) for the mating type linked with structure (Python)
```python
# Creating the contingency table for structure type
data_structure = [[23, 45, 6],
                  [28, 21, 3],
                  [19, 19, 2]]
# Defining row and column labels
row_labels_structure = ['Tubular', 'Blade-like', 'Flat Tube']
col_labels_structure = ['mt+', 'mt-', 'mt?']
# Creating a DataFrame
df_structure = pd.DataFrame(data_structure, columns=col_labels_structure, index=row_labels_structure)
# Performing the chi-square test
chi2_structure, p_structure, dof_structure, ex_structure = chi2_contingency(df_structure)
chi2_structure, p_structure, dof_structure, ex_structure
```

d.) for the rhizoid linked with structure. (Python)
```python
from scipy.stats import chi2_contingency
# Creating the contingency table for chi-square test
contingency_table = df.pivot(index='Rhizoid Status', columns='Structure Type', values='Count').fillna(0)
# Performing the chi-square test
chi2, p, dof, expected = chi2_contingency(contingency_table)
chi2, p, dof, expected
```

Code for all other data analyses (Python)
```python
import subprocess
import sys
from scipy import stats
from sklearn.decomposition import PCA
from sklearn.cluster import KMeans
import pandas as pd
# Function to install packages
def install(package):
    subprocess.check_call([sys.executable, "-m", "pip", "install", package])
# List of packages to install
packages = ['pandas', 'scipy', 'statsmodels', 'openpyxl', 'scikit-learn']
# Install each package
for package in packages:
    try:
        __import__(package)
    except ImportError:
        install(package)
# Now import the required libraries
import pandas as pd
# Load the data from the Excel file
file_path = r"C:\Users\hawil\OneDrive\Desktop\ULVA THESIS\Data\DATA FOR PYTHON.xlsx"
data = pd.read_excel(file_path)
# Print the column names to verify them
```

```python
print("Column names:", data.columns)
# Clean column names (if necessary)
data.columns = data.columns.str.strip()
# Normalize the case of categorical columns
data['Mating Type'] = data['Mating Type'].str.lower()
data['Rhizoid'] = data['Rhizoid'].str.lower()
data['structure new'] = data['structure new'].str.lower()
# Fill missing values in growth columns with the mean of the column
data['Average Green Pixel Count (mm^2)'] = data['Average Green Pixel Count (mm^2)'].fillna(data['Average
Green Pixel Count (mm^2)'].mean())
data['Average Max Contour Area (mm^2)'] = data['Average Max Contour Area (mm^2)'].fillna(data['Average
Max Contour Area (mm^2)'].mean())
# Ensure all data is numeric
data['Average Green Pixel Count (mm^2)'] = pd.to_numeric(data['Average Green Pixel Count (mm^2)'],
errors='coerce')
data['Average Max Contour Area (mm^2)'] = pd.to_numeric(data['Average Max Contour Area (mm^2)'],
errors='coerce')
# Descriptive Statistics for growth rates
descriptive_stats_green_pixel = data['Average Green Pixel Count (mm^2)'].describe()
descriptive_stats_contour = data['Average Max Contour Area (mm^2)'].describe()
print("Descriptive Statistics for Green Pixel Growth Rate:")
print(descriptive_stats_green_pixel)
print("\nDescriptive Statistics for Contour Growth Rate:")
print(descriptive_stats_contour)
# T-tests: Compare the growth rates (green pixel and contour) between different mating types (mating type
+ vs. mating type -).
ttest_green_pixel = stats.ttest_ind(
    data[data['Mating Type'] == 'mt+']['Average Green Pixel Count (mm^2)'],
    data[data['Mating Type'] == 'mt-']['Average Green Pixel Count (mm^2)'],
    nan_policy='omit'
)
print(f"\nT-test for Mating Type (Green Pixel): {ttest_green_pixel}")
ttest_contour = stats.ttest_ind(
    data[data['Mating Type'] == 'mt+']['Average Max Contour Area (mm^2)'],
    data[data['Mating Type'] == 'mt-']['Average Max Contour Area (mm^2)'],
    nan_policy='omit'
)
print(f"T-test for Mating Type (Contour): {ttest_contour}")
# ANOVA: Compare the growth rates among different structures (tube, flat tube, blade).
anova_green_pixel = stats.f_oneway(
    data[data['structure new'] == 'tube']['Average Green Pixel Count (mm^2)'],
    data[data['structure new'] == 'flat tube']['Average Green Pixel Count (mm^2)'],
    data[data['structure new'] == 'blade']['Average Green Pixel Count (mm^2)']
)
print(f"\nANOVA for Structure (Green Pixel): {anova_green_pixel}")
anova_contour = stats.f_oneway(
    data[data['structure new'] == 'tube']['Average Max Contour Area (mm^2)'],
    data[data['structure new'] == 'flat tube']['Average Max Contour Area (mm^2)'],
    data[data['structure new'] == 'blade']['Average Max Contour Area (mm^2)']
)
```

```python
print(f"ANOVA for Structure (Contour): {anova_contour}")
# T-tests: Compare the growth rates (green pixel and contour) between rhizoid presence (yes vs. no).
ttest_rhizoid_green_pixel = stats.ttest_ind(
    data[data['Rhizoid'] == 'yes']['Average Green Pixel Count (mm^2)'],
    data[data['Rhizoid'] == 'no']['Average Green Pixel Count (mm^2)'],
    nan_policy='omit'
)
print(f"\nT-test for Rhizoid Presence (Green Pixel): {ttest_rhizoid_green_pixel}")
ttest_rhizoid_contour = stats.ttest_ind(
    data[data['Rhizoid'] == 'yes']['Average Max Contour Area (mm^2)'],
    data[data['Rhizoid'] == 'no']['Average Max Contour Area (mm^2)'],
    nan_policy='omit'
)
print(f"T-test for Rhizoid Presence (Contour): {ttest_rhizoid_contour}")
# ANOVA: Compare the growth rates among different rhizoid presence (yes, no, not determined).
anova_rhizoid_green_pixel = stats.f_oneway(
    data[data['Rhizoid'] == 'yes']['Average Green Pixel Count (mm^2)'],
    data[data['Rhizoid'] == 'no']['Average Green Pixel Count (mm^2)'],
    data[data['Rhizoid'] == 'nd']['Average Green Pixel Count (mm^2)']
)
print(f"\nANOVA for Rhizoid Presence (Green Pixel): {anova_rhizoid_green_pixel}")
anova_rhizoid_contour = stats.f_oneway(
    data[data['Rhizoid'] == 'yes']['Average Max Contour Area (mm^2)'],
    data[data['Rhizoid'] == 'no']['Average Max Contour Area (mm^2)'],
    data[data['Rhizoid'] == 'nd']['Average Max Contour Area (mm^2)']
)
print(f"ANOVA for Rhizoid Presence (Contour): {anova_rhizoid_contour}")
# Pearson and Spearman correlation: Assess the correlation between green pixel growth rate and contour
growth rate.
pearson_corr, pearson_p = stats.pearsonr(data['Average Green Pixel Count (mm^2)'], data['Average Max
Contour Area (mm^2)'])
spearman_corr, spearman_p = stats.spearmanr(data['Average Green Pixel Count (mm^2)'], data['Average
Max Contour Area (mm^2)'])
print(f"\nPearson Correlation between Green Pixel and Contour Growth Rates: {pearson_corr}, p-value:
{pearson_p}")
print(f"Spearman Correlation between Green Pixel and Contour Growth Rates: {spearman_corr}, p-value:
{spearman_p}")
# PCA: Reduce dimensionality and identify the most important components that explain the variability in the
growth data
features = ['Average Green Pixel Count (mm^2)', 'Average Max Contour Area (mm^2)']
x = data[features]
pca = PCA(n_components=2)
principal_components = pca.fit_transform(x)
principal_df = pd.DataFrame(data=principal_components, columns=['Principal Component 1', 'Principal
Component 2'])
print("\nPCA Results:")
print(principal_df.head())
print(f"Explained Variance Ratios: {pca.explained_variance_ratio_}")
# K-Means Clustering: Cluster segregating lines based on their growth rates
kmeans = KMeans(n_clusters=3, random_state=0)
```

```python
data['Cluster'] = kmeans.fit_predict(data[['Average Green Pixel Count (mm^2)', 'Average Max Contour Area
(mm^2)']])
print("\nK-Means Clustering Results:")
print(data[['Segregating line', 'Average Green Pixel Count (mm^2)', 'Average Max Contour Area (mm^2)',
'Cluster']].head())
# Display the number of samples in each cluster
print("\nNumber of samples in each cluster:")
print(data['Cluster'].value_counts())
# Show detailed information for each cluster
print("\nDetailed cluster information:")
for cluster in range(3):
    print(f"\nCluster: {cluster}")
    cluster_data = data[data['Cluster'] == cluster]
    print(cluster_data[['Segregating line', 'Mating Type', 'Rhizoid', 'structure new', 'Average Green Pixel
Count (mm^2)', 'Average Max Contour Area (mm^2)']])
```

Code for all the figures (Python)

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans

# Load the data from the Excel file
file_path = r"C:\Users\hawil\OneDrive\Desktop\ULVA THESIS\Data\DATA FOR PYTHON.xlsx"
data = pd.read_excel(file_path)

# Clean column names (if necessary)
data.columns = data.columns.str.strip()

# Normalize the case of categorical columns
data['Mating Type'] = data['Mating Type'].str.lower()
data['Rhizoid'] = data['Rhizoid'].str.lower()
data['structure new'] = data['structure new'].str.lower()

# Fill missing values in growth columns with the mean of the column
data['Average Green Pixel Count (mm^2)'] = data['Average Green Pixel Count (mm^2)'].fillna(data['Average
Green Pixel Count (mm^2)'].mean())
data['Average Max Contour Area (mm^2)'] = data['Average Max Contour Area (mm^2)'].fillna(data['Average
Max Contour Area (mm^2)'].mean())

# Ensure all data is numeric
data['Average Green Pixel Count (mm^2)'] = pd.to_numeric(data['Average Green Pixel Count (mm^2)'],
errors='coerce')
data['Average Max Contour Area (mm^2)'] = pd.to_numeric(data['Average Max Contour Area (mm^2)'],
errors='coerce')

# 1. Histograms: Growth Rate Distribution
plt.figure(figsize=(14, 6))
plt.subplot(1, 2, 1)
sns.histplot(data['Average Green Pixel Count (mm^2)'], kde=True)
```

```
plt.title('Growth Rate Distribution (Green Pixel)')

plt.subplot(1, 2, 2)
sns.histplot(data['Average Max Contour Area (mm^2)'], kde=True)
plt.title('Growth Rate Distribution (Contour)')
plt.tight_layout()
plt.show()

# 2. Box Plots: Growth Rates by Categories
plt.figure(figsize=(14, 12))
plt.subplot(3, 2, 1)
sns.boxplot(data=data, x='Mating Type', y='Average Green Pixel Count (mm^2)')
plt.title('Growth Rates by Mating Type (Green Pixel)')

plt.subplot(3, 2, 2)
sns.boxplot(data=data, x='Mating Type', y='Average Max Contour Area (mm^2)')
plt.title('Growth Rates by Mating Type (Contour)')

plt.subplot(3, 2, 3)
sns.boxplot(data=data, x='Rhizoid', y='Average Green Pixel Count (mm^2)')
plt.title('Growth Rates by Rhizoid Presence (Green Pixel)')

plt.subplot(3, 2, 4)
sns.boxplot(data=data, x='Rhizoid', y='Average Max Contour Area (mm^2)')
plt.title('Growth Rates by Rhizoid Presence (Contour)')

plt.subplot(3, 2, 5)
sns.boxplot(data=data, x='structure new', y='Average Green Pixel Count (mm^2)')
plt.title('Growth Rates by Structure (Green Pixel)')

plt.subplot(3, 2, 6)
sns.boxplot(data=data, x='structure new', y='Average Max Contour Area (mm^2)')
plt.title('Growth Rates by Structure (Contour)')
plt.tight_layout()
plt.show()

# 3. Bar Charts: Frequency of Categories
plt.figure(figsize=(14, 6))
plt.subplot(1, 3, 1)
sns.countplot(data=data, x='Mating Type')
plt.title('Frequency of Mating Types')

plt.subplot(1, 3, 2)
sns.countplot(data=data, x='Rhizoid')
plt.title('Frequency of Rhizoid Presence')

plt.subplot(1, 3, 3)
sns.countplot(data=data, x='structure new')
plt.title('Frequency of Structures')
plt.tight_layout()
```

```
plt.show()

# 4. Enhanced Box Plots for Contour Growth Rate
plt.figure(figsize=(14, 6))
plt.subplot(1, 2, 1)
sns.boxenplot(data=data, x='Mating Type', y='Average Max Contour Area (mm^2)')
plt.title('Enhanced Box Plot for Contour Growth Rate by Mating Type')

plt.subplot(1, 2, 2)
sns.boxenplot(data=data, x='Rhizoid', y='Average Max Contour Area (mm^2)')
plt.title('Enhanced Box Plot for Contour Growth Rate by Rhizoid Presence')
plt.tight_layout()
plt.show()

# Enhanced Box Plots for Structures
plt.figure(figsize=(14, 6))
sns.boxenplot(data=data, x='structure new', y='Average Max Contour Area (mm^2)')
plt.title('Enhanced Box Plot for Contour Growth Rate by Structure')
plt.tight_layout()
plt.show()

# 5. Pie Charts: Proportions of Categories
plt.figure(figsize=(18, 6))
plt.subplot(1, 3, 1)
data['Mating Type'].value_counts().plot.pie(autopct='%1.1f%%', startangle=90)
plt.title('Proportions of Mating Types')

plt.subplot(1, 3, 2)
data['Rhizoid'].value_counts().plot.pie(autopct='%1.1f%%', startangle=90)
plt.title('Proportions of Rhizoid Presence')

plt.subplot(1, 3, 3)
data['structure new'].value_counts().plot.pie(autopct='%1.1f%%', startangle=90)
plt.title('Proportions of Structures')
plt.tight_layout()
plt.show()

# 6. Heatmaps: Correlation Matrix
plt.figure(figsize=(10, 8))
corr_matrix = data[['Average Green Pixel Count (mm^2)', 'Average Max Contour Area (mm^2)']].corr()
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()

# 7. Pair Plots: Multivariate Relationships
sns.pairplot(data, vars=['Average Green Pixel Count (mm^2)', 'Average Max Contour Area (mm^2)',
'Average Total Contour Area (mm^2)'], hue='structure new')
plt.suptitle('Pair Plots by Structure', y=1.02)
plt.show()
```

```
sns.pairplot(data, vars=['Average Green Pixel Count (mm^2)', 'Average Max Contour Area (mm^2)',
'Average Total Contour Area (mm^2)'], hue='Rhizoid')
plt.suptitle('Pair Plots by Rhizoid', y=1.02)
plt.show()


sns.pairplot(data, vars=['Average Green Pixel Count (mm^2)', 'Average Max Contour Area (mm^2)',
'Average Total Contour Area (mm^2)'], hue='Mating Type')
plt.suptitle('Pair Plots by Mating Type', y=1.02)
plt.show()


# 8. Cluster Plots: K-Means Clustering Results
kmeans = KMeans(n_clusters=3, random_state=0)
data['Cluster'] = kmeans.fit_predict(data[['Average Green Pixel Count (mm^2)', 'Average Max Contour Area
(mm^2)']])


# Map cluster labels to meaningful names
cluster_labels = {0: 'slow', 1: 'medium', 2: 'fast'}
data['Cluster'] = data['Cluster'].map(cluster_labels)


plt.figure(figsize=(10, 8))
sns.scatterplot(data=data, x='Average Green Pixel Count (mm^2)', y='Average Max Contour Area (mm^2)',
hue='Cluster', palette='viridis')
plt.title('K-Means Clustering Results')
plt.legend(title='Cluster', labels=['Slow', 'Medium', 'Fast'])
plt.show()
```