*Prochlorococcus marinus* responses to light and oxygen

**Mireille Savoie**, Aurora Mattison, Julie Nadeau, Laurel Genge, Sylwia Śliwińska-Wilczewska, Maximilian Berthold, Naaman Omar, Ondřej Prášil, Amanda M. Cockshutt, and Douglas A. Campbell

2024-02-16

Table of Contents

[Abstract 2](#_Toc158981612)

[Introduction 3](#_Toc158981613)

[0.1 *Prochlorococcus* evolutionary and functional diversity 3](#_Toc158981614)

[0.2 *Prochlorococcus* responses to changing niches 6](#_Toc158981615)

[0.3 Photosystem II maintenance as a limitation on *Prochlorococcus* growth 11](#_Toc158981616)

[Materials and methods 13](#_Toc158981617)

[MetaProteomics 13](#_Toc158981618)

[*Prochlorococcus* culturing and experimental design 14](#_Toc158981619)

[MetaProteomics bioinformatic analyses and limitations 18](#_Toc158981620)

[Growth data management and analysis 20](#_Toc158981621)

[Generalized additive model of growth responses 20](#_Toc158981622)

[0.3.1 Estimation of Photosynthetically Usable Radiation 21](#_Toc158981623)

[Results and Discussion 22](#_Toc158981624)

[Supplemental 34](#_Toc158981625)

[1 Acknowledgements 34](#_Toc158981626)

[2 Funding 34](#_Toc158981627)

[References 34](#_Toc158981628)

1 Department of Biology, Mount Allison University, 53 York St., Sackville NB, Canada, E4L 1C9.  
2 2  
3 Department of Biology, University of British Columbia  
4 3  
5 Umweltbundesamtm, Dessau-Roßlau, Germany  
6 4  
7 Algatech, Czech Academy of Science  
8 5  
9 Department of Chemistry, St. Frances Xavier University

# Abstract

# Introduction

## 0.1 *Prochlorococcus* evolutionary and functional diversity

*Prochlorococcus marinus*, a genus of the phylum Cyanobacteria, is the smallest known photosynthetic prokaryote, with cell diameters ranging from 0.5 to 0.7 µm [**chisholmProchlorococcusMarinusNov1992?**]. Despite small cell size, *P. marinus* contribute 13 to 48% of net primary production in oligotrophic oceans, corresponding to about 30% of global oxygen production [**partenskyProchlorococcusMarinePhotosynthetic1999?**]. *P. marinus* growth is currently constrained to between latitudes of 40°N to 40°S in open ocean waters, from surface to 300 m depth, thus spanning 3 orders of magnitudes of light (Partensky et al., 1999) (Chisholm et al., 1992) [**partenskyProchlorococcusMarinePhotosynthetic1999?**].

*Prochlorococcus marinus* is functionally differentiated from other cyanobacteria primarily by the absence of phycobilisomes [**chisholmProchlorococcusMarinusNov1992?**]. Instead, *Prochlorococcus* utilizes intra-membrane Prochlorophyte Chlorophyll Binding (Pcb) proteins bound to divinylchlorophyll a (Chl a2), and α-carotene, as the major light harvesting complex [**goerickePigmentsProchlorococcusMarinus1992?**]; [**larocheIndependentEvolutionProchlorophyte1996?**], with diversity in the number of genes encoding Pcb, expression of Pcb under changing light [**garczarekMultiplicationAntennaGenes2000?**] [**rocapGenomeDivergenceTwo2003?**] [**garczarekMultiplicationAntennaGenes2000?**], and in thylakoid organization [**partenskyProchlorococcusMarinePhotosynthetic1999?**]. The genes encoding most of the rest of the photosynthetic apparatus is highly conserved in *Prochlorococcus* [**kettlerPatternsImplicationsGene2007?**], though the photosynthetic efficiency may vary among clades [**hessPhotosyntheticApparatusProchlorococcus2001?**,**moorePhotophysiologyMarineCyanobacterium1999?**].

*Prochlorococcus marinus* comprises many strains, organized into clades, defined by 16S-23S intergenic transcribed ribosomal sequence signatures (Rocap et al., 2003). The clades inhabit distinct ecological niches [**moorePhysiologyMolecularPhylogeny1998?**], originally defined by Low Light (LL) or High Light (HL). Of at least 12 known *Prochlorococcus* clades only 5 to date have cultured representatives; HLI, HLII, LLI, LLII/III and LIV [**billerProchlorococcusStructureFunction2015?**].

XXXXMIREILLE Check strain descriptions for accuracyXXXX Low-Light clades thrive in deeper ocean waters, extending beyond 200 m in depth [**partenskyProchlorococcusMarinePhotosynthetic1999?**], where only ~1% of the surface irradiance penetrates, primarily in the blue (450 nm) to green (520 nm) spectral ranges [**holtropVibrationalModesWater2021?**]. Clade LLIV, including cultured strain MIT9313, falls near the base of the *Prochlorococcus* radiation, and is characterized by preference for low light, at depths from 120 m and 200 m (Partensky et al., 1999). LLIV members are, as yet, the only cultured strains found in Oxygen Minimum Zones (OMZ). Some as yet uncultured *P. marinus* strains in clades LLV and LLVI thrive in OMZ of the subtropical Atlantic and Pacific Oceans, where dissolved oxygen concentrations [O2] can be less than 20 µM [**ulloaCyanobacteriumProchlorococcusHas2021?**] (Lavin et al., 2010) ADDITIONAL REFS XXXX. (Johnson et al., 1999). [**goerickeNovelNicheProchlorococcus2000?**,**garcia-robledoCrypticOxygenCycling2017a?**]. *P. marinus* LL ecotypes may indeed dominate the phytoplankton wihthin OMZ [**lavinNovelLineagesProchlorococcus2010?**,**ulloaCyanobacteriumProchlorococcusHas2021?**] [**partenskyProchlorococcusAdvantagesLimits2010?**].

Clades LLII and LLIII are often grouped together as the second oldest phylogenetic lineages in the *Prochlorococcus* radiation, and share an affinity for low light; LLIII cultured strain SS120, was isolated from 135 m depth in the Sargasso Sea.

Clade LLI includes cultured strains NATL1A and NATL2A, which prefer moderate irradiances found between 30 and 100 m depth.

HL clades are the most recently evolved lineages, with reduced genome sizes in comparison to LL ecotypes. High-Light clades are typically dominant picophytoplankton in near-surface, oligotrophic waters, characterized by high light levels. Clade HLI, including cultured strain MED4,is adapted to higher iron, and lower temperatures, and originated from the Mediterranean Sea, just 5 m below the surface (Partensky et al., 1999). Clade HLII adapted to higher iron, and higher temperatures, is the most abundant group in the North Atlantic and North Pacific Oceans, often constituting over 90% of the total population (Partensky et al., 1999), and are most numerous around 50 m depth (Partensky et al., 1999). Clade HLIII/IV is adapted to lower iron [**kentParallelPhylogeographyProchlorococcus2019?**,**johnsonNichePartitioningProchlorococcus2006?**,**zinserInfluenceLightTemperature2007?**].

*P. marinus* clades are found in environments beyond their optimal habitats. HL clades inhabit depths overlapping with LL ecotypes [**westNichePartitioningProchlorococcusPopulationsStratified1999?**,**zinserInfluenceLightTemperature2007?**,**delmontLinkingPangenomesMetagenomes2018?**] [**moorePhotophysiologyMarineCyanobacterium1999?**]. [**lavinNovelLineagesProchlorococcus2010?**] and [**mattisonPhotosyntheticStrategiesWild2020?**] show that LL ecotypes can occupy regions in the OMZ at depths above 40 m, reaching ambient light levels above what LL clades were thought to tolerate.

## 0.2 *Prochlorococcus* responses to changing niches

Our changing climate, characterized by warming ocean waters, is rapidly altering conditions for these highly specialized strains of marine picophytoplankton. Predictions indicate a net global increase of *P. marinus* cell abundances of 29%, along with poleward latitudinal shifts of at least 10° in marine phytoplankton niches by the end of this century [**flombaumPresentFutureGlobal2013?**] in response to warming waters, with increases in *Prochlorococcus* of approximately 50% in the more poleward regions. As the solar incidence angles are lower and photoperiod seasonality is more pronounced at higher latitudes, this shift will carry these phytoplankters into new combinations of photoperiod, seasonal regimes and changing spectral regimes [**bartonAnthropogenicClimateChange2016?**]. A 10° poleward shift of *P. marinus*, coupled with a dramatic increase in cell abundance between 40°N and 50°N, could have a large impact on ocean ecosystems and global biogeochemical cycles [**flombaumPresentFutureGlobal2013?**], and thus will undoubtedly impact the ecosystem balance off our Canadian waters.

Global warming is also rapidly changing the chemistry of our oceans. By the end of this century, surface ocean pH is projected to decline by 0.1 to 0.4 due to projected increases in carbon dioxide concentrations [**garcia-sotoOverviewOceanClimate2021?**]. Moreover, substantial changes in the global water cycle, leading to extensive changes in worldwide precipitation patterns, are affecting ocean salinity levels on a global scale, and ice melts due to rising temperatures are impacting salinity levels in the Arctic and Northwest Atlantic oceans [**leeRecognizingSalinityThreats2022?**]. Increasing sea temperatures are also causing a decrease in [O2] across our global oceans [**matearLongtermChangesDissolved2003?**], with significant declines in [O2] observed toward the poles in both hemispheres [**helmObservedDecreasesOxygen2011?**]. Warmer ocean waters increase stratification, and decrease oxygen solubility at the surface, which in turn decreases oxygen mixing downwards by ocean currents [**garcia-sotoOverviewOceanClimate2021?**]. Models predict that the OMZ in the Pacific and Indian Oceans are expanding [**garcia-sotoOverviewOceanClimate2021?**,**buseckeDivergingFatesPacific2022?**]. In light of these expanding OMZ, a recent study by [**buseckeDivergingFatesPacific2022?**] highlights an intriguing finding; the cores of the OMZ, where the oxygen levels are lowest, is projected to contract. Given the predicted decrease in overall oceanic [O2],projected expansion of OMZs, and possible decline in the OMZ core volumes, we set forth to examine the influence of [O2] on *P. marinus* growth, as this area of research has been relatively understudied.

Near the equator, photoperiod remains relatively constant at the ocean surface, approximately 12 hours (h) of daylight and 12 h of darkness throughout the entire year. The effective length of the photoperiod does, however attentuate with depth as dawn and dusk light drops below the level needed for biological processes. As *P. marinus* potentially expands its temperature-permissive niches poleward into temperate regions [**flombaumPresentFutureGlobal2013?**,**bartonAnthropogenicClimateChange2016?**], it will encounter more pronounced seasonal variations in photoperiod regimes at surface and at depth. With the influence of light attenuation with depth, temperate regions exhibit more combinations of photoperiod, compared to equatorial regions, with potentially complex effects upon viable growth niches [**liDiatomGrowthResponses2017?**] [**prezelinDielPeriodicityPhytoplankton1992?**]. [**vaulotGrowthProchlorococcusPhotosynthetic1995?**] showed that *Prochlorococcus* replication of DNA occurs in the afternoon, while cell division occurs at night.

We therefore analyzed the growth and physiological responses of representative strains from three clades under a matrix of [O2], light levels, spectral waveband ranges, and photoperiods, to approximate eco-physiological conditions representative of current and hypothetical future ocean zones. *P. marinus* MED4, a clade HLI strain (Figure ??), was isolated near the ocean surface (5m depth) of the Mediterranean Sea where [O2] is near saturation, light levels are high and spectral bias from full solar irradiance is minimal. *P. marinus* SS120, a clade LLIII (Figure ??), was isolated from the Sargasso Sea at a depth of 120 m, while *P. marinus* MIT9313, a clade LLIV (Figure ??), was isolated from the North Atlantic Gulf Stream at a depth of 135 m. At these depths, light attenuation and spectral shifts occur, resulting in low blue and green light levels, while [O2] varies from near-surface saturation levels to decreased concentrations, but does not necessarily decrease systematically with depth [**billheimerOxygenSeasonalityUtilization2021?**].

We then analyzed growth rates in terms of cumulative diel Photosynthetically Usable Radiation (PUR). Photosynthetic organisms rely on absorbing light energy within the PAR range, 350 to 700 nm, for photosynthesis [**morelAvailableUsableStored1978?**]. PUR represents the fraction of PAR that is absorbed by the pigments of photosynthetic organisms [**morelAvailableUsableStored1978?**]. PUR, therefore, takes into account the pigment composition of photosynthetic organisms and the specific spectral wavebands these pigments absorb. *Prochlorococcus marinus* light-harvesting complexes, [**chisholmProchlorococcusMarinusNov1992?**,**goerickePigmentsProchlorococcusMarinus1992?**], show an absorption maxima of 442 nm for divinyl chlorophyll a and 478 nm for divinyl chlorophyll b [**goerickePigmentsProchlorococcusMarinus1992?**] allowing *P. marinus* to efficiently harvestblue light in the 400 nm to 500 nm range [**morelAvailableUsableStored1978?**] prevailing in deep ocean habitats, where only blue spectral wavelengths prevail [**holtropVibrationalModesWater2021?**]. In *Prochlorococcus* small cell diameters, from 0.5 to 0.7 µm [**chisholmProchlorococcusMarinusNov1992?**], and simple cell structures, minimize the complication of pigment package effect or intracellular self-shading [**morelProchlorococcusSynechococcusComparative1993?**]. Given the different spectral light regimes typical of the niches of different ecotypes, expressing growth rates in terms of cumulative diel absorbed PUR might simplify different photoperiods, spectral bands, and PAR levels into a common parameter, making growth response comparisons across strains and different oxygen levels more accessible. We also aimed to detect whether growth responses are driven simply by cumulative diel absorbed PUR, or whether specific photoperiods, spectral bands or PAR levels have independent, albeit interacting, effects on growth.

## 0.3 Photosystem II maintenance as a limitation on *Prochlorococcus* growth

*Prochlorococcus* remain challenging to culture, as their reduced genomes – the smallest of any known oxyphototroph – render them partially dependent upon mutualistic heterotrophic bacteria to detoxify reactive oxygen species (Moore et al., 2007; Morris et al., 2008). MED4, SS120 and MIT9313 have been successfully cultured in laboratories, and used to show that ecotypic classifications correspond to biochemical differences among strains. CITATONS OUR LAB and OTHERS.

Under full atmospheric [O2], LL clades of *Prochlorococcus* are restricted to growth under low light, in part because they suffer photoinhibition of Photosystem II (PSII) through several paths, including direct absorbance of UV or blue light, in parallel with generation of Reactive Oxygen Species (ROS) if the electron flow is slowed, [**aroPhotoinhibitionPhotosystemII1993?**], producing damaging singlet oxygen (1O2). [**aroPhotoinhibitionPhotosystemII1993?**,**soitamoPhotoinhibitionMarinePicocyanobacteria2017?**] [[**murphyPhotoinactivationPhotosystemII2017?**]][**hakalaEvidenceRoleOxygenevolving2005?**]. Repair of inhibited PSII process relies on the removal of damaged PsbA (Mann et al., 2000; Adam et al., 2005), followed by reassembly with newly synthesized PsbA [**nixonRecentAdvancesUnderstanding2010?**]. Thus, degradation of PsbA is a rate-limiting step in recovery from photoinhibition (Kanervo et al., 1993), mediated largely by a heterohexamer of (FtsH12)3, a membrane-bound (Sacharz et al., 2015)(Zak et al., 2001) metalloprotease (Chiba et al., 2002) (Yoshioka-Nishimura and Yamamoto, 2014). (Adam et al., 2005; Boehm et al., 2012) (Nixon et al. (2010)) (Komenda et al., 2007). (Pisareva et al., 2007; Sacharz et al., 2015), (Mann et al., 2000), (AMANDA REF from ALGATECH).

*Prochlorococcus* genomes encode 4 FtsH proteins, henceforth referred to as FtsH1-4, homologous to the characterized FtsH isoforms of the model cyanobacterium *Synechocystis* FtsH, with presumably parallel functions (Table 1). Upon a shift to higher light HLIII MED4 upregulated expression of FtsH1 and FtsH2 (Bonisteel et al. (2018)), homologs to the *Synechocystis* slr0228 and slr1604, implicated in PSII repair. In contrast representative LLIV strain MIT9313 showed less overall expression of the FtsH proteases, and thus has fewer FtsH serving each photosystem. Furthermore MIT9313 expressed primarily FtsH3, homologous to Sll1463 involved in PSI biogenesis, and FtsH expression did not increase in response to light stress. Through adaptation to steady low light, LLIV *Prochlorococcus* instead allocates resources processes other than dynamic regulation of PSII repair.

## # A tibble: 3 x 5  
## Organism FtsH1 FtsH2 FtsH3 FtsH4   
## <chr> <chr> <chr> <chr> <chr>   
## 1 Prochlorococcus marinus FtsH1 FtsH2 FtsH3 FtsH4   
## 2 Synechocystis sp. PCC6803 SlrO22 Slr1604 Slr1463 Slr1390   
## 3 Function PSII Repair PSII Repair PSI biogenesis Cell viabili~

# Materials and methods

## MetaProteomics

The OceansMap Protein Portal (OPP; <https://proteinportal.whoi.edu/>) is an open access online data repository (Woods Hole Oceanographic Institute, WHOI) of mass spectroscopy data on marine microbial peptides, sampled from various depths and locations worldwide. We screened a subset of the OPP for proteins annotated as from *Prochlorococcus* strains, to identify differential strategies employed by strains living at varying depths and oxygen levels within the marine water column. We focussed on proteins mediating photosynthesis and protein metabolism. The samples for metaproteomic analyses were collected from 12 locations in the tropical North Pacific ocean along 150 W from 18 N of the equator between October 1, 2011 and October 25, 2011 during the voyage of the R/V Kilo Moana MetZyme cruise KM1128 (<https://www.rvdata.us/search/cruise/KM1128>; original datasets in the Biological and Chemical Oceanography Data Management Office repository; <https://www.bco-dmo.org/project/2236>).

Collection and treatment of protein samples were performed following Saito et al., (2014, 2015) XXXXADD to ZoteroXXXX . Briefly, samples of seawater from depths of 20, 40, 50, 60, 70, 80, 90, 120, 150, 200, 250, 300, 380, 400, 500, 550, 600, and 800 m below the ocean surface were pumped through a 0.2 µm filter, preserved in RNAlater and frozen at -80°C until extraction. Proteins were extracted from the filter in an SDS-based detergent, embedded in tube gel, alkylated and reduced prior to in-gel trypsin digestion. Peptide spectra were generated using a Q-Exactive Orbitrap Mass Spectrometer, searched in the SEQUEST CITATIONXXXX and labelled with the most likely protein and species annotation from Uniprot. Oxygen levels at the location of sampling were recorded.

## *Prochlorococcus* culturing and experimental design

Three xenic cultures of *P. marinus* were obtained from Bigelow Labs, NCMA Maine, US (REFERENCE). MED4 (CCMP1986) is from High Light-adapted (HLI) clade; SS120 (CCMP1375) is from Low Light-adapted (LLIII) clade; and MIT9313 (CCMP2773) is from Low Light-adapted (LLIV) clade. Cultures were maintained in incubators set to 22°C with a light/dark cycle of 12 h, with illumination from XXXXX. The PAR level for maintenance cultures reflected Photosynthetically Active Radiation (PAR) in the source niche of the ecotype; MED4, of 160 µmol photons m-2 s-1; SS120 and MIT9313 at 30 µmol photons m-2 s-1. To maintain active growth all strains were transferred weekly with XX to YY dilution with Pro99 media [1] prepared with autoclaved artificial seawater (National Center for Marine Algae and Microbiota (NCMA) protocol by combining salt solution I and salt solution II using the enriched artificial seawater (ESAW) recipe. REPLACe with CITATION OF NCMA protocol)XXXX

Controlled growth experiments were performed using PSI Multicultivators (Figure 0.1; MCMIX-OD orMC1000-OD, Brno, Czech Republic ). Each multicultivator individually controls 8 tubes at a common temperature of 22°C. Each tube containing 70 mL of Pro99 media was inoculated with 10 mL of growing maintenance culture. In a factorial matrix design, each tube was then subject to an individual combination of sinusoidal photoperiod (4, 8, 12, 16 h); reaching a peak PAR (30, 90, 180 µmol photons m-2 s-1), with defined spectral bandwidth (AA, BB, CC, DD nm). [O2] levels (2.5 µM, 25 µM, 250 µM) were imposed by bubbling tubes with varying ratios of air and Nitrogen (N2), with consistent 0.05% of Carbon Dioxide (CO2) gas, delivered through a 0.2 μm sterile microfilter via a G400 gas mixing system XXXXGIVE SOURCEXXXX. [O2] in situ was verified using oxygen optodes (PyroScience, Germany) inserted into tubes for real-time measurements, with a temperature probe in the aquarium of the bioreactor to correct [O2] measures for temperature fluctuations. In addition, the Pyroscience software corrected [O2] based on the salinity of the media (32 ppt). The flow rate of the gas mixture was controlled, but variations in bubbling speed, PAR and culture density affected the [O2] achieved in each tube. A low [O2] of 0.5 µM - 5 µM (reported as 2.5 µM hereafter), was achieved by sparging with a gas mixture containing 99.95% N2 and 0.05% CO2. An intermediate [O2] of 10 - 25 µM (reported hereafter as 25 µM) was achieved by sparging with a gas mixture containing 98.95% N2, 0.05% CO2 and 1% O2. A high O2 of 200 µM - 280 µM (reported hereafter as 250 µM) was achieved by sparging with lab air (78% N2, 21% O2, 1% Ar and 0.05% CO2).

The full crossing of all factor levels would yield 4 x 3 x 4 x 3 = 144 treatments, x 3 strains for 432 possible combinations. Consistent absence of growth of some strains under some levels of photoperiod, PAR, or [O2] meant we completed 291 growth factor treatment combinations.

In situ XXXXItalicsXXX measurements of Optical Density (OD) 680 nm, a proxy for cell suspension density, cell size dependent scatter and cell chlorophyll content; and OD 720 nm, a proxy for cell suspension density and cell size dependent scatter, were recorded every 5 minutes over least 8 to 14 days, depending on the duration of the lag phase, if any. All data from the Multicultivator were saved as comma separated values file (CITATION of FILES LOCATION).

PAR of 180, 90 or 30 µmol photons m-2 s-1, and spectral wavebands (white LED full spectrum, 660 nm, 530 nm, and 450 nm) were chosen to approximate light levels and spectral colors spanning the vertical ocean water column, from near-surface to the lower euphotic zone depths. Photoperiods were chosen to approximate diel cycles characteristic of current and hypothetical future niches of *P. marinus*; 16 h represents temperate (45°N) summer at the ocean surface; 12 h for equatorial (0°N) ocean surface or temperate (45°N) spring and fall ocean surface or temperate (45°N) summer at deeper ocean depths; 8 h for temperate (45°N) winter at the surface or at temperate (45°N) spring and fall at depth and equatorial (0°N) deep ocean depths; and 4 h for temperate (45°N) winter or deep ocean depths during temperate (45°N) spring and fall.

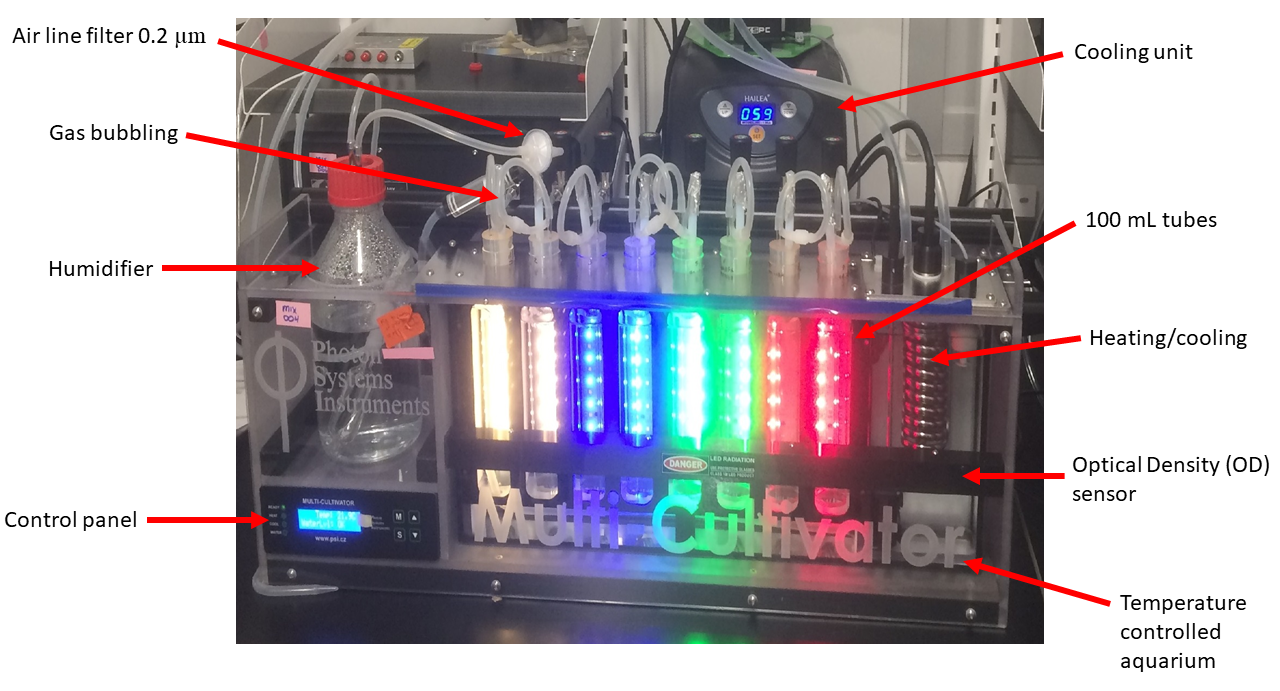


Figure 0.1: **PSI MCMIX-OD Multicultivator.** Spectral wavebands and light levels are individually controlled for each culture tube. Real time Optical Density (OD) measurements eliminate intrusive subsampling of sterile cultures. The temperature of culture tubes are collectively controlled via heating or cooling of the aquarium water. Gas with specific oxygen concentrations is bubbled through a humidifier and passed through a 0.2 um filter.

## MetaProteomics bioinformatic analyses and limitations

Metaproteomic datasets were obtained from the KM1128 entry in the BCO-DMO database (<https://www.bco-dmo.org/deployment/59053>) accessed via the OPP in June 2019. Datasets contained: i) Peptide sequences and sample identification (ID) number; ii) Sample ID number, station, depth in meters below the surface the sample was collected at, best-hit BLASTP protein and species annotation and the corresponding Uniprot Entry number for the identified proteins; iii) Sample station depth and [O~2].  
The depth and [O~2] were joined to peptide sequence and BLASTP annotations by ID number, depth and station using the functions in the Tidyverse package (<https://www.tidyverse.org/>) in RStudio v1.2.5019 (<https://rstudio.com/>) (LOCATION OF CODE ON GITHUBO. The resulting merged dataset was filtered for *Prochlorococcus* peptides, detected from 0 to 300 m below the surface, annotated as a subunit of *Prochlorococcus* chlorophyll binding proteins (Pcb); Photosystem II (PSII); Cytochrome b6f (Cytb6f); Photosystem I (PSI); NADPH Dehydrogenase (NDH); Terminal Oxidase (PTOX); plastocyanin (PC); ferredoxin (Fd); Ribulose-1,5-bisphosphate oxygenase (RUBISCO); ATP Synthase; FtsH proteases (FtsH) or ribosomes. Detected peptides were re-annotated for consistency and labelled according to strain, clade, subunit and protein complex (where possible). Full protein sequences corresponding to detected proteins were obtained from UniProt (<https://www.uniprot.org/>) and analyzed in Molecular Evolution and Genetic Analyses X (MEGAX) software (<https://www.megasoftware.net/>). Sequences for each protein and/or subunits for each of the thirteen Prochlorococcus strains identified in the dataset were aligned with MUSCLE using UPGMA cluster method and a lambda of 24 with a -2.9 gap open penalty and 1.20 hydrophobicity multiplier. Overall mean pairwise distance between protein sequences was determined using bootstrap variance estimation methods. Maximum likelihood phylogenetic trees were assembled using 1000 bootstrap replications with a 95% site coverage cut off. *Prochlorococcus* FtsH isoform identities were specified by phylogenetic comparison to the characterized FtsH proteases of Synechocystis sp. PCC6803 (CITATONS). Data for each strain was plotted against depth and [O~2] and sampling station.

Graphs depict the detection of mass-spectrally detected peptides at varying depths and [O~2] in the water column for *Prochlorococcus* strains and clades identified in the dataset. When assessing the presence of a particular protein complex, all peptides belonging to all subunits of the complex were included to give the greatest number of data points. Although not a direct measurement, it is assumed that light intensity is inversely correlated with depth. As this data was acquired on a discovery mission rather than through targeted peptide approaches, it is difficult to discern accuracies of strain annotations, particularly as the proteins of interest in this study are highly conserved in sequence across strain. We are confident in clade and ecotypic classification for each protein examined, although exact strain annotation may be ambiguous. Another caveat to interpretation of this data is the “peptide detection bias” inherent to mass spectrometry: certain peptides are more easily detected and thus dominate the spectra over less easily detected peptides. The data is also limited by the number and nature of protein spectra in the SEQUEST database: a peptide sequence cannot be determined unless there is already a known spectrum for that peptide in the SEQUEST database, hence some peptides of interest may not be identifiable. Therefore, the absence of a particular does not necessarily rule out its presence overall. Furthermore, a peptide must be detected above a certain threshold abundance in order to be considered an accurate ‘hit’. Therefore, the ‘absence’ of a particular protein structure is here interpreted as low in quantity, while its presence is interpreted as higher in quantity.

## Growth data management and analysis

Data files saved form from the Multicultivator software were imported into R-Studio (CITATION) for data management [2], growth rate calculations CITATION, comparisons of model fits CITATION, and generation of figuresggplot2 CITATION. The chlorophyll proxy optical density (OD680 - OD720; ΔOD) was used to determine the chlorophyll-specific exponential growth rate (µ, h-1) for each treatment combination. We first we used a rolling mean from the RStudio zoo package [3] to calculate the average ΔOD data over a 1-hour window, to lower influence of outlier points XXXXother trimmingXXXXX using a Levenberg-Marquardt algorithm [4] modification of the non-linear least squares, using the R package minpack.lm [5], to fit a logistic equation (0.1) where ΔODmax is maximum ΔOD, ΔODmin is minimum ΔOD, t is time duration over the growth trajectory. OTHER MOREXXXXXX

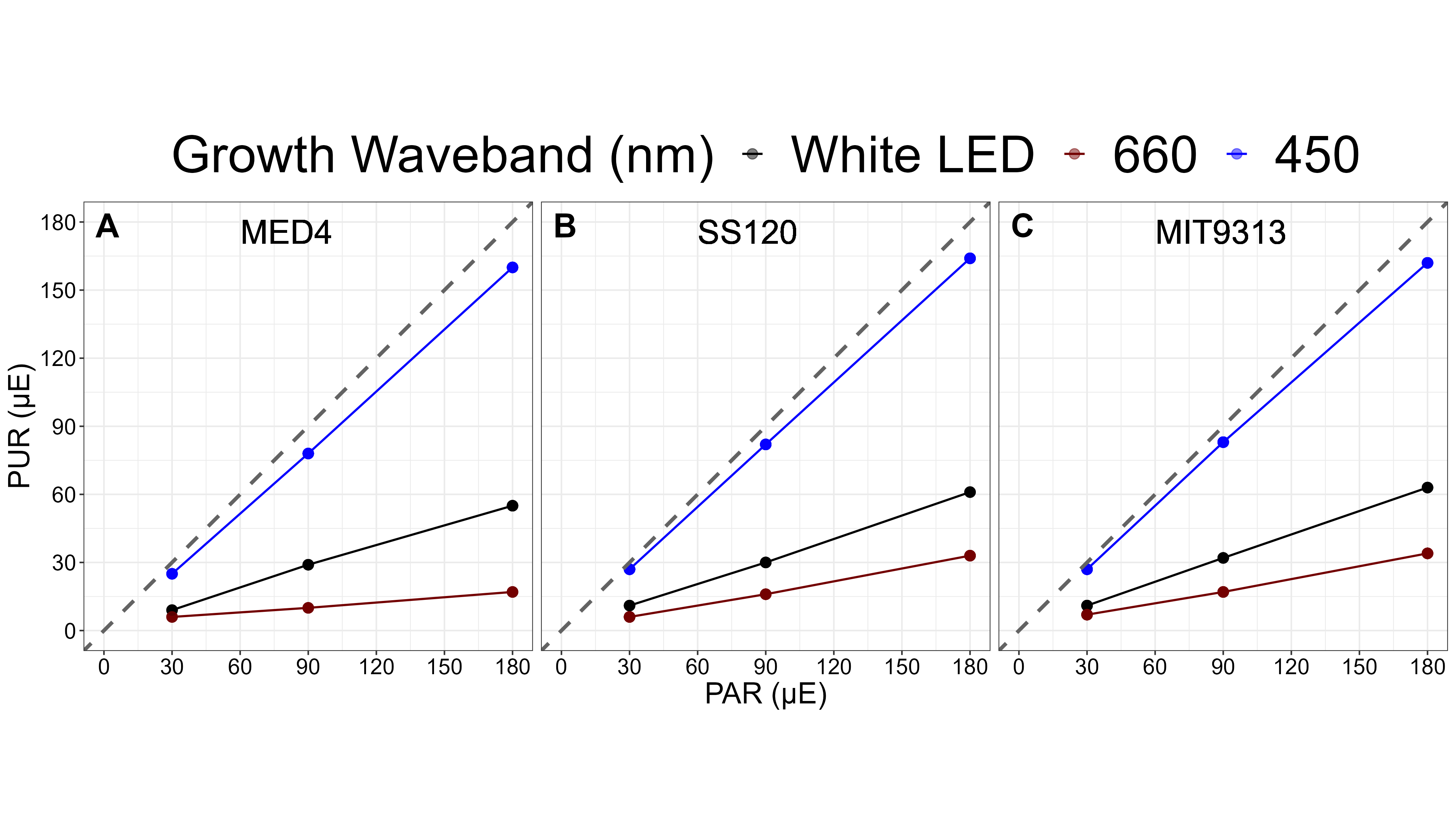
CITAtIONS of github locations of code etc. XXXXX

### Generalized additive model of growth responses

%%%%%% Taken from my thesis word for word, need to reword and shorten %%%%

A Generalized Additive Model (GAM) was applied to examine the relationship between chlorophyll-specific µ, h-1, for each [O2] level, across the blue and red spectral wavebands, photoperiods and PAR levels for each *P.marinus* strain in this study. The gam function from the R package mgcv [6] was used to model the growth rate with smoothing terms to indicate the 90, 50 and 10% quantiles for growth rate across the levels of factors. Only data XXXXXWHAT DATAXXXX BE SPECIFICXXXX below a standard error tolerance of 30% of the fit was used in the model. Our priority was studying the effect of blue light on growth trends, since blue light is the most ecologically relevant spectral waveband for deep ocean niches. We included analyses of responses of red light, which is not ecophysiologically relevant but might prove mechanistically informative (CITE MURPHY ETC)

### 0.3.1 Estimation of Photosynthetically Usable Radiation

XXXXTEXT NEEDEDXXXX 

# Results and Discussion

%%%% OceanProteinPortal data show complexes in environment %%%

Furthermore, Evidence of proteins derived from HL ecotypes of *P. marinus* inOMZ at depths up to 200 meters, with O2 of 15 µM. The extent to which [O2] defines the niches occupied by different *P. marinus* ecotypes, as compared to potentially covarying environmental variables like photoperiod, light spectrum, and light level, is poorly described.

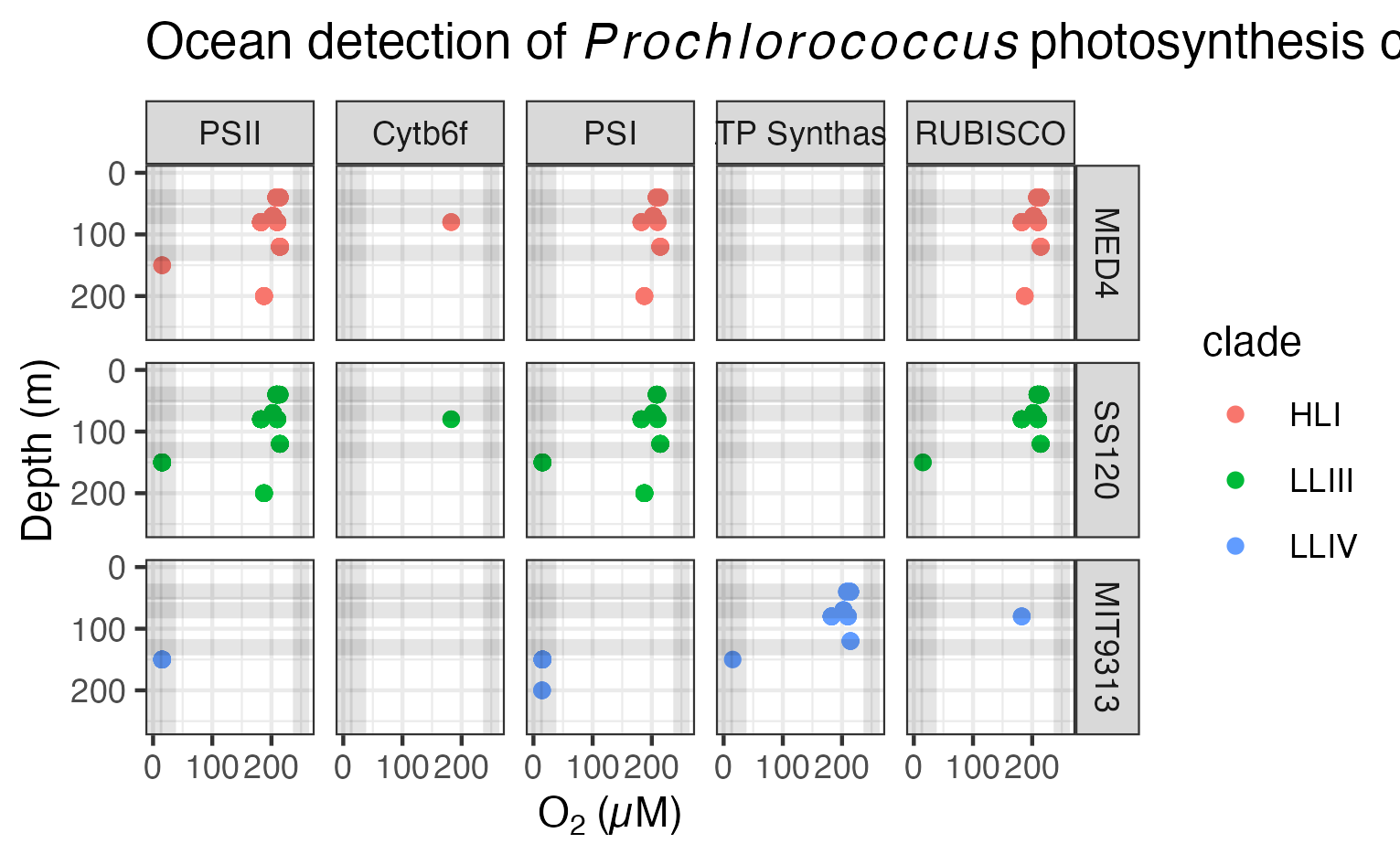


Figure 0.3: **Ocean detection of Prochlorococcus marinus photosynthesis complexes.** Protein detections are plotted vs. O2 (µM) (X axis) and depth (m) (Y axis) at sample origin. Rows separate data annotated as from Prochlorococcus marinus strains MED4, SS120 and MIT9313. Columns show detections of proteins annotated as Photosystem II (PSII), Cytochromeb6f complex (Cytb6f), Photosystem I (PSI), ATP Synthase or Ribulose-1,5-bisphosphate oxygenase carboxylase (RUBISCO). Culture growth experimental conditions indicated by horizontal grey lines for depths approximating Photosynthetically Active Radiation (µmol photons m-2 s-1) and vertical grey lines for O2 (µM). Point colours identify Prochlorococcus clade assignments Data from OceanProteinPortal (<https://www.oceanproteinportal.org/>).

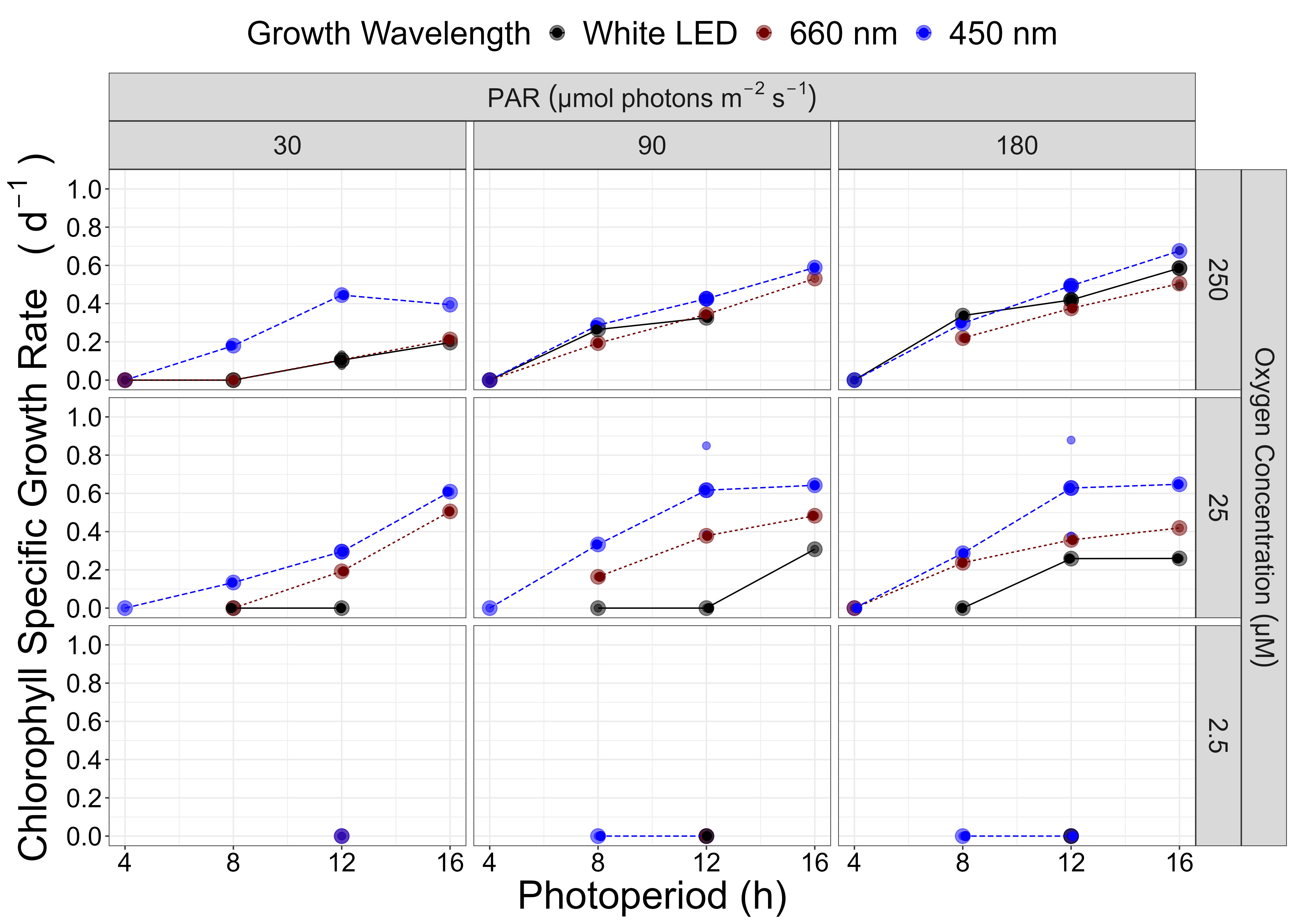


Figure 0.4: **Chlorophyll specific growth rate (d-1) for Prochlorococcus marinus MED4 (High Light (HLI) near surface clade) vs. photoperiod (h).**  3 levels of growth Photosynthetically Active Radiation (PAR) (µmol photons m-2 s-1) are in columns; 3 levels of imposed growth dissolved O2 concentrations (µM) are in rows. Colors represent the actinic spectral waveband (nm). Large circles show mean or single determinations of growth rate from logistic curve fits; small circles show values for replicate determinations, if any: replicates often fall with larger circles.

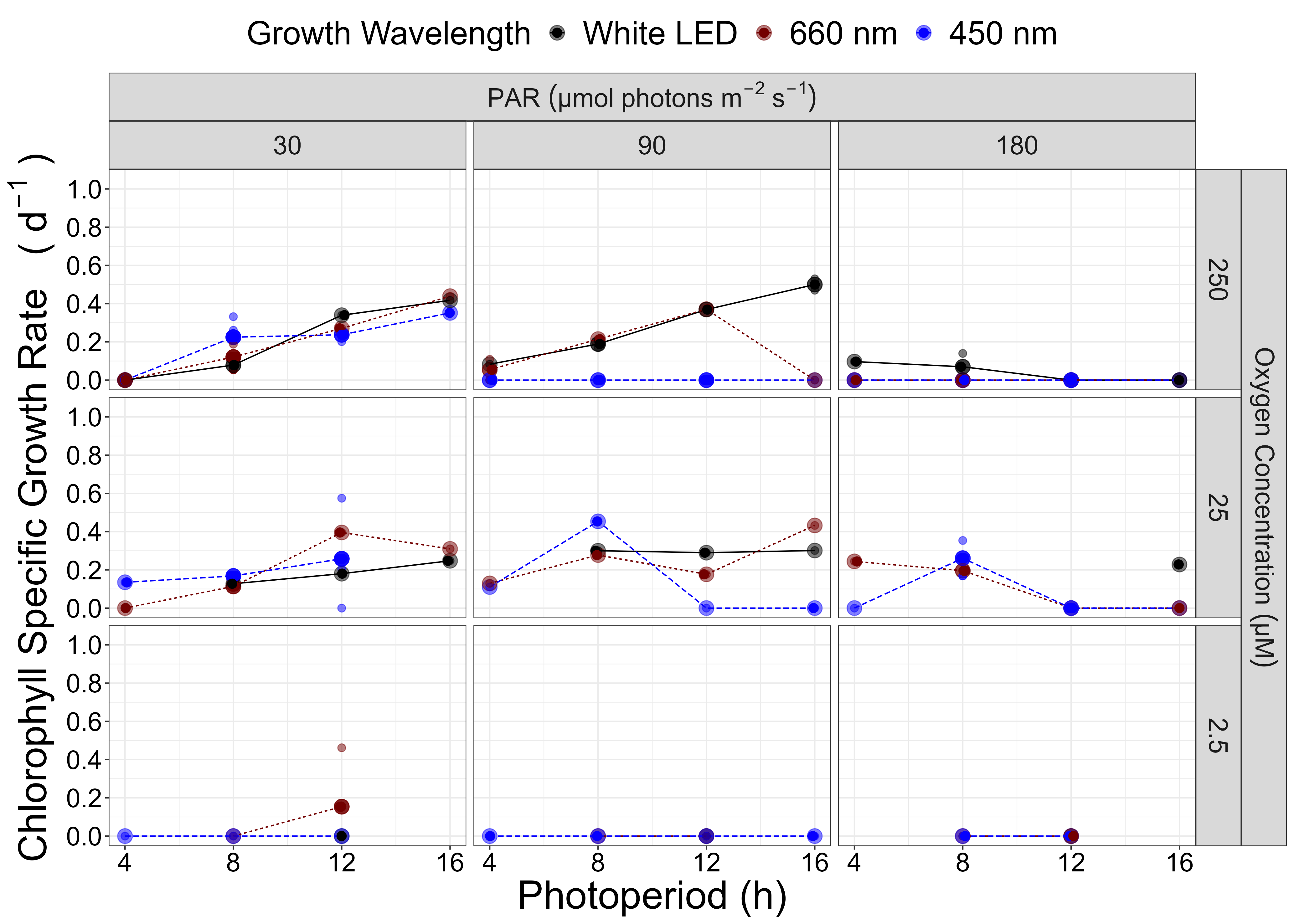


Figure 0.5: **Chlorophyll specific growth rate (d-1) for Prochlorococcus marinus SS120 (Low Light (LLIII) deep ocean clade) vs. photoperiod (h).**  3 levels of growth Photosynthetically Active Radiation (PAR) (µmol photons m-2 s-1) are in columns; 3 levels of imposed growth dissolved O2 concentrations (µM) are in rows. Colors represent the actinic spectral waveband (nm). Large circles show mean or single determinations of growth rate from logistic curve fits; small circles show values for replicate determinations, if any: replicates often fall with larger circles.



Figure 0.6: **Chlorophyll specific growth rate (d-1) for Prochlorococcus marinus MIT9313 (Low Light (LLIV) deep ocean clade) vs. photoperiod (h).**  3 levels of growth Photosynthetically Active Radiation (PAR) (µmol photons m-2 s-1) are in columns; 3 levels of imposed growth dissolved O2 concentrations (µM) are in rows. Colors represent the actinic spectral waveband (nm). Large circles show mean or single determinations of growth rate from logistic curve fits; small circles show values for replicate determinations, if any: replicates often fall with larger circles.

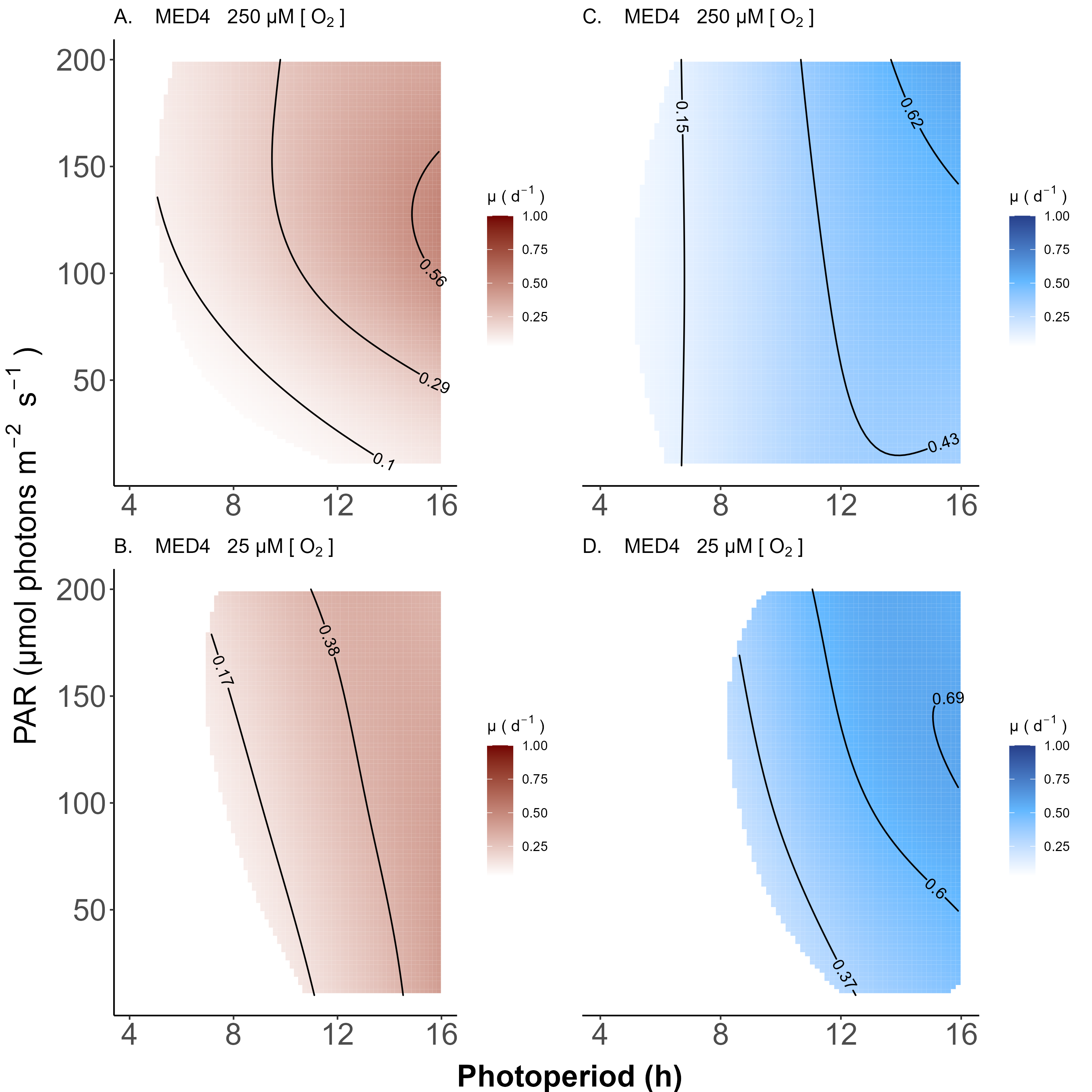


Figure 0.7: **A contour plot of a Generalized Additive Model (GAM) representing the chlorophyll specific growth rate (d-1) for Prochlorococcus marinus MED4 grown under 660 nm (red) and 450 nm (blue) light.** X-axis is photoperiod (h). Y-axis is actinic Photosynethetic Active Radiation (PAR, µmol photons m-2 s-1). **A.** represents the model under 250 µM of O2 and red light. **B.** represents the model under 25 µM of O2 and red light. **C.** represents the model under 250 µM of O2 and blue light. **D.** represents the model under 25 µM of O2 and blue light. Legends represent a colour gradient of growth rate from no growth (white) to 1.00 d-1 (dark red or dark blue). Labeled contour lines indicate the 90%, 50%, and 10% quantiles for achieved growth rate.

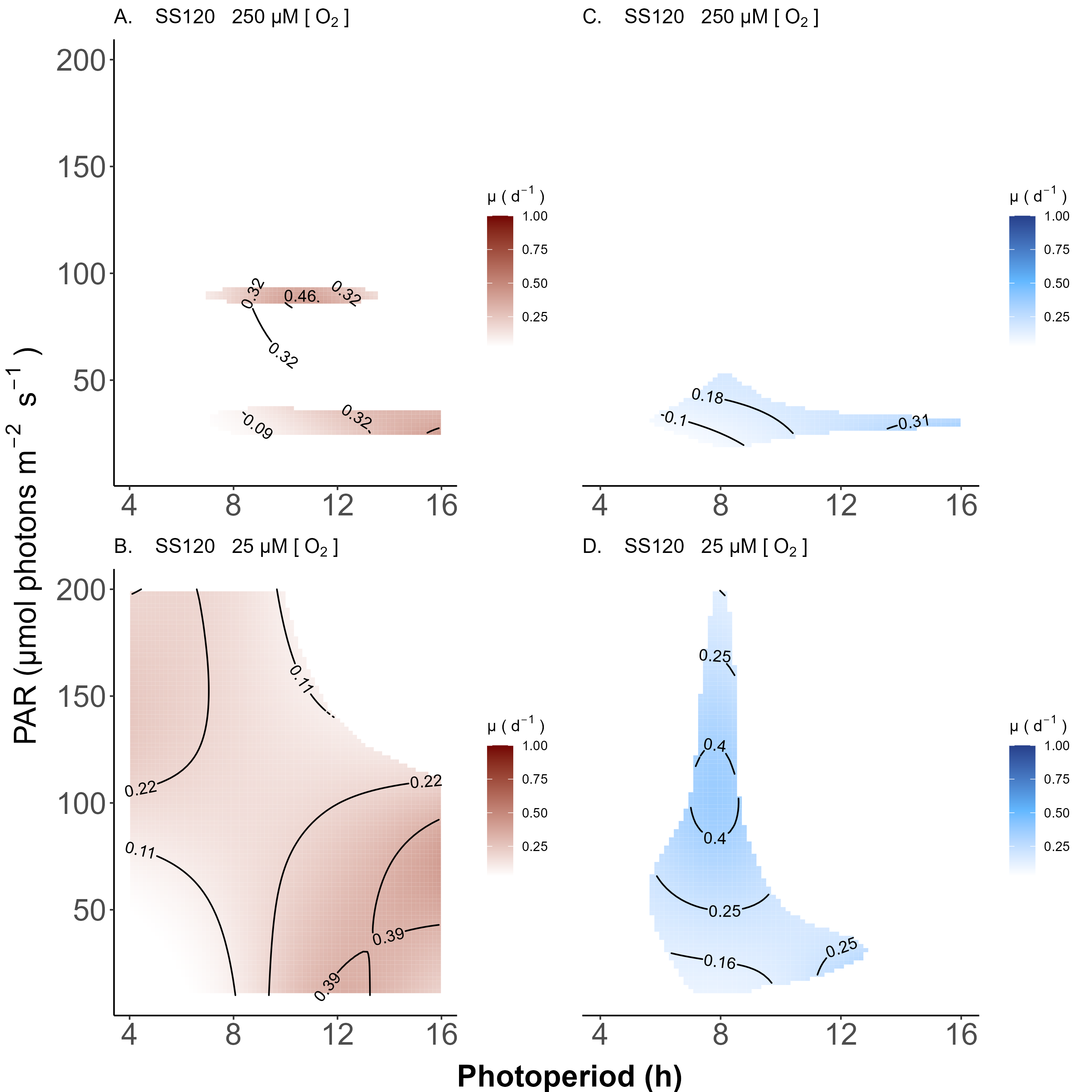


Figure 0.8: **Contour plot of a Generalized Additive Model (GAM) representing the chlorophyll specific growth rate (d-1) for Prochlorococcus marinus SS120 grown under 660 nm (red) and 450 nm (blue) light.** X-axis is photoperiod (h). Y-axis is actinic Photosynethetic Active Radiation (PAR, µmol photons m-2 s-1). **A.** represents the model under 250 µM of O2 and red light. **B.** represents the model under 25 µM of O2 and red light. **C.** represents the model under 250 µM of O2 and blue light. **D.** represents the model under 25 µM of O2 and blue light. Legends represent a colour gradient of growth rate from no growth (white) to 1.00 d-1 (dark red or dark blue). Labeled contour lines indicate the 90%, 50%, and 10% quantiles for achieved growth rate.

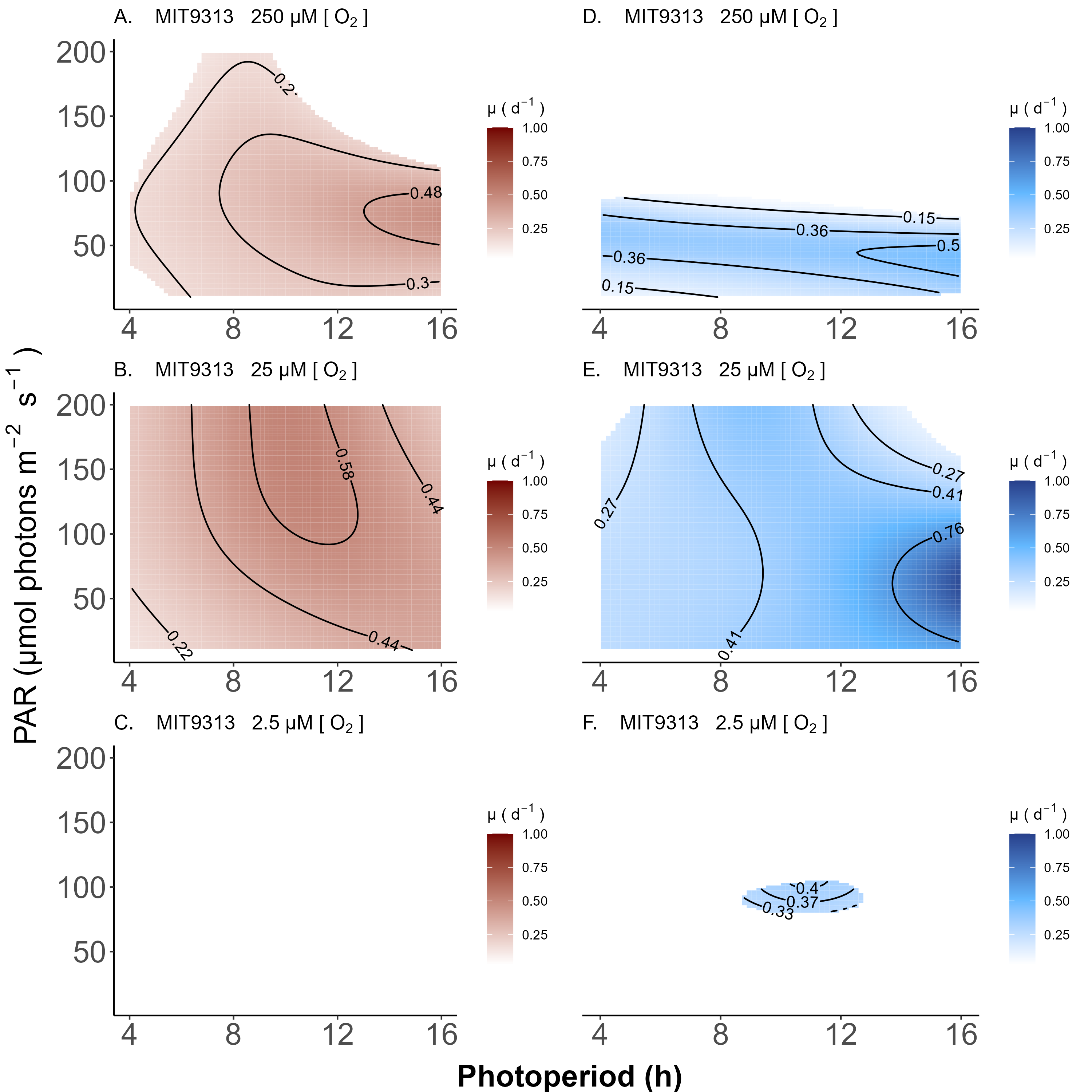


Figure 0.9: **Contour plot of a Generalized Additive Model (GAM) representing the chlorophyll specific growth rate (d-1) for Prochlorococcus marinus MIT9313 grown under 660 nm (red) and 450 nm (blue) light.** X-axis is photoperiod (h). Y-axis is actinic Photosynthetically Active Radiation (PAR, µmol photons m-2 s-1). **A.** represents the model under 250 µM of O2 and red light. **B.** represents the model under 25 µM of O2 and red light. **C.** represents the model under 2.5 µM of O2 and red light. **D.** represents the model under 250 µM of O2 and blue light. **E.** represents the model under 25 µM of O2 and blue light. **F.** represents the model under 2.5 µM of O2 and blue light. Legends represent a colour gradient of growth rate from no growth (white) to 1.00 d-1 (dark red or dark blue). Labeled contour lines indicate the 90%, 50%, and 10% quantiles for achieved growth rate.

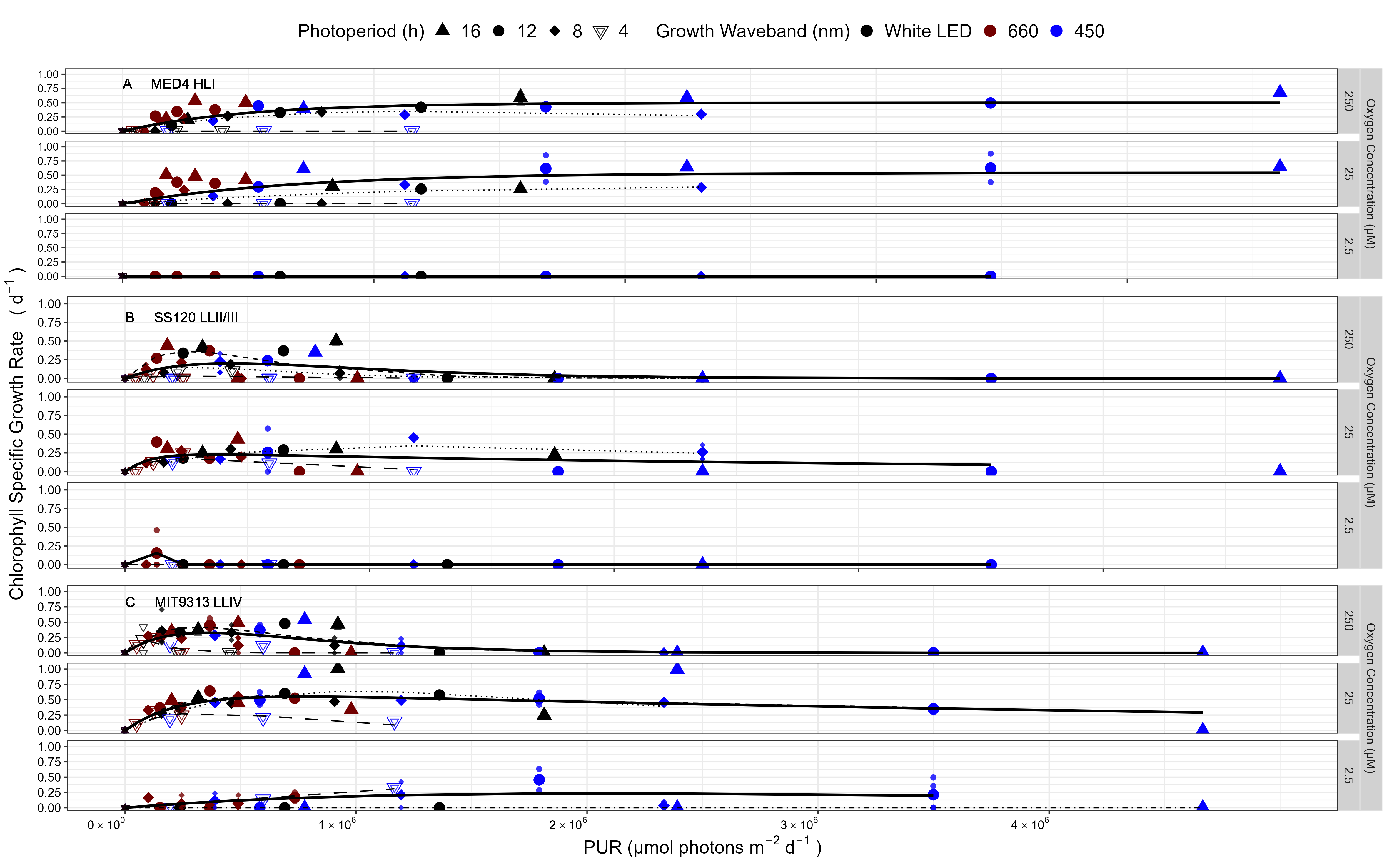


Figure 0.10: **Chlorophyll specific growth rate (d-1) vs. cumulative diel Photosynthetic Usable Radiation (PUR, µmol photons m-2 d-1).** **A.** MED4 High Light I ecotype under 250, 25, or 2.5 µM of O2. **B.** SS120 Low Light II/III (LLII/III) ecotype under 250, 25, or 2.5 µM of O2. **C.** MIT9313 Low Light IV (LLIV) ecotype under 250, 25, or 2.5 µM of O2. 3 levels of imposed dissolved O2 concentrations (µM) are in rows for each strain. Shapes show the imposed photoperiod (h); 4 h (hollow inverted triangle), 8 h (solid diamond), 12 h (solid circle), 16 h (solid upright triangle). Symbol colours show the spectral waveband for growth; white LED (black symbols), 660 nm (red symbols), and 450 nm (blue symbols). Large symbols show mean of growth rate from logistic curve fits; small symbols show values for replicate determinations, if any. Platt (add Platt citation here) 4 parameter model fit to data pooled for each combination of strain and dissolved oxygen shown with solid lines. Seperate models fit to photoperiod data and shown if significantly different from the pooled model; 4 h (long dashed line); 8 h (dotted line); 12 h (dashed line); and 16 h (dot dashed line).

%%%% OceanFtsHRibosomeProteins figure from OceanPortal repository %%%

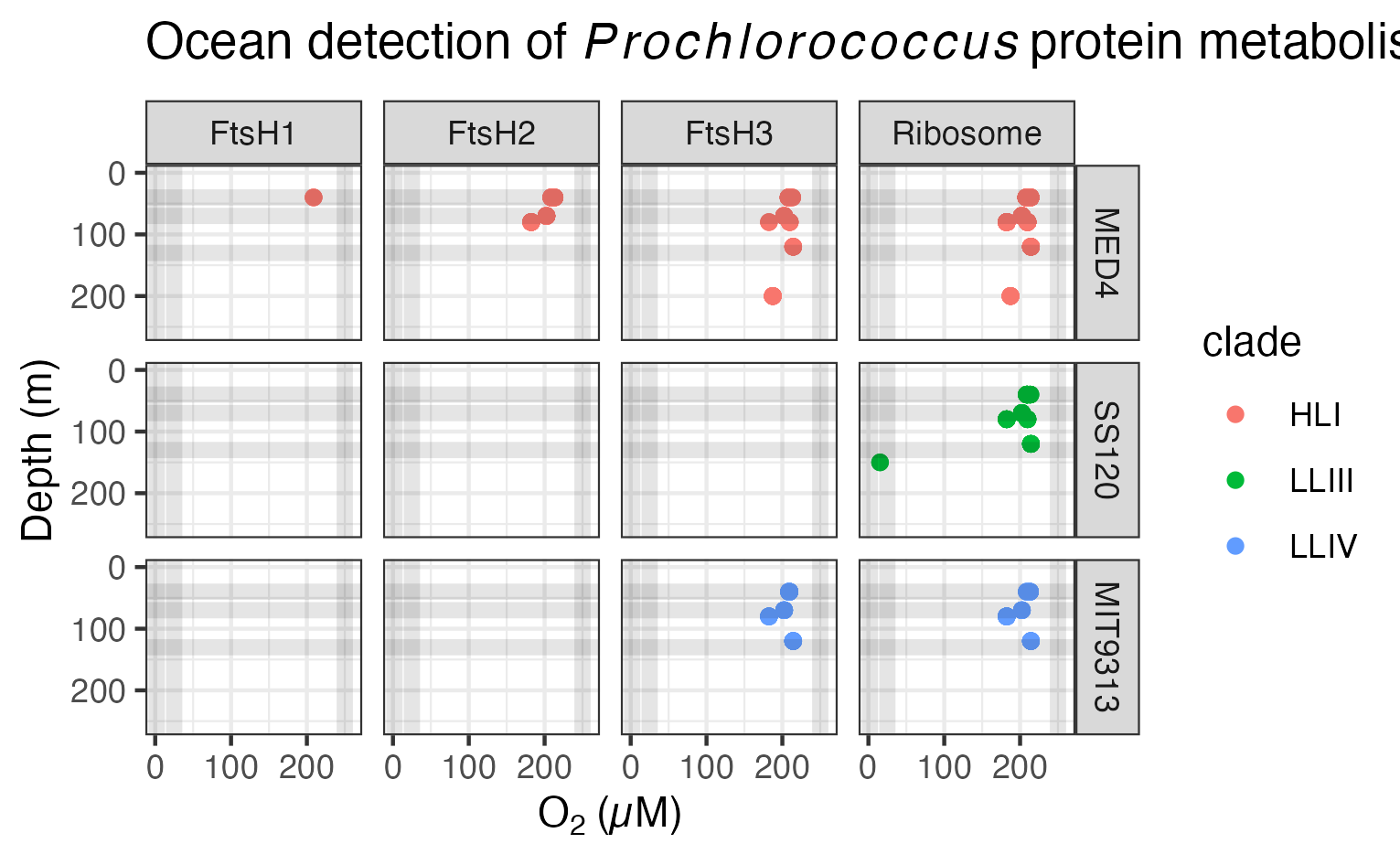


Figure 0.11: **Ocean detection of Prochlorococcus marinus protein metabolism complexes.** Protein detections are plotted vs. O2 (µM) (X axis) and depth (m) (Y axis) at sample origin. Rows separate data annotated as from Prochlorococcus marinus strains MED4, SS120 and MIT9313. Columns show detections of proteins annotated as FtsH Protease Complexes (FtsH1, FtsH2, FtsH3) or the Ribosome). Culture growth experimental conditions indicated by horizontal grey lines for depths approximating Photosynthetically Active Radiation (µmol photons m-2 s-1) and vertical grey lines for O2 (µM). Point colours identify Prochlorococcus clade assignments Data from OceanProteinPortal (<https://www.oceanproteinportal.org/>).

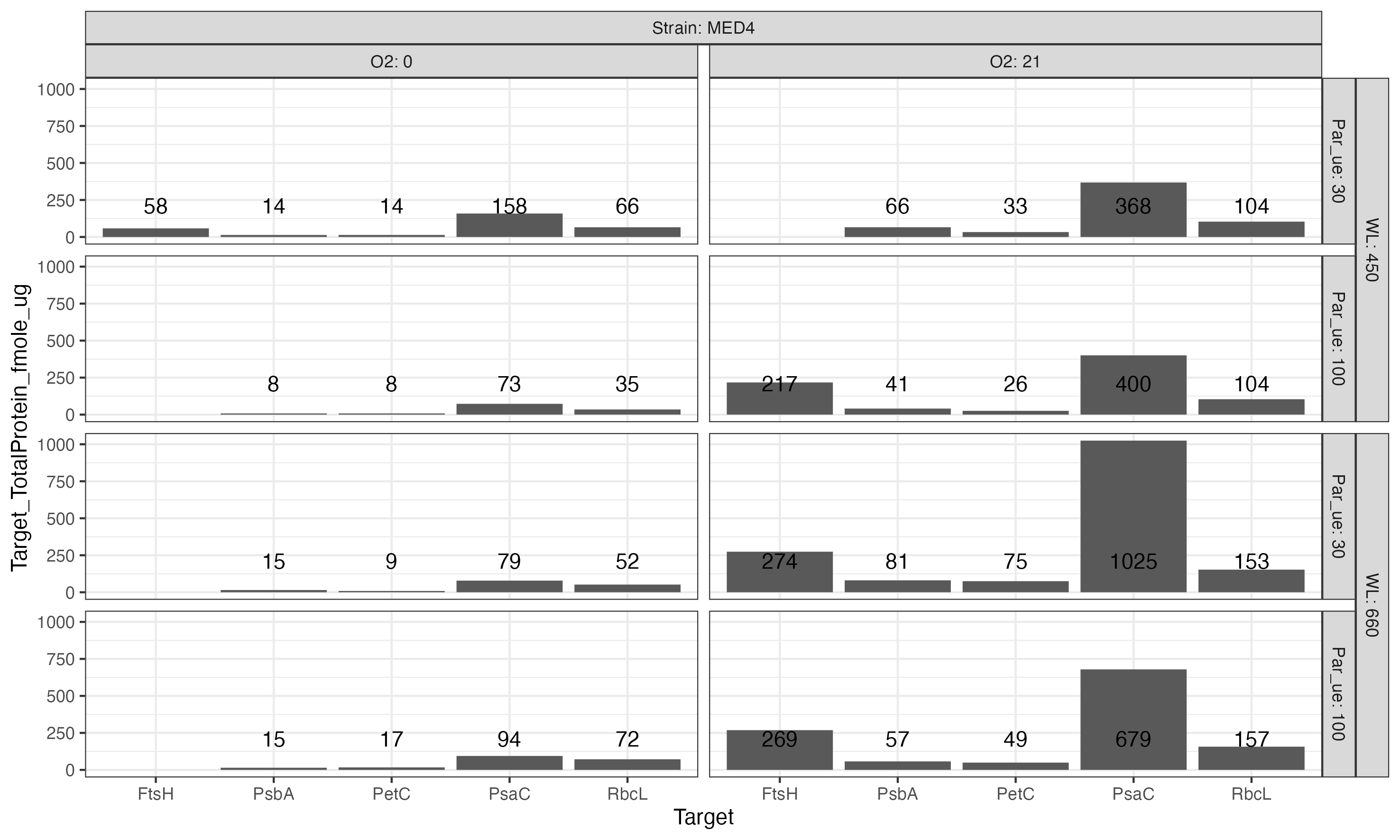


Figure 0.12: **fmole target protein per ug total protein for Prochlorococcus marinus MED4 (High Light (HLI) near surface clade).**  Growth Photosynthetically Active Radiation (PAR) (µmol photons m-2 s-1) and spectral wavelength are in rows; 2 levels of imposed growth dissolved O2 concentrations (µM) are in columns. Numbers over each bar are fmole/ug

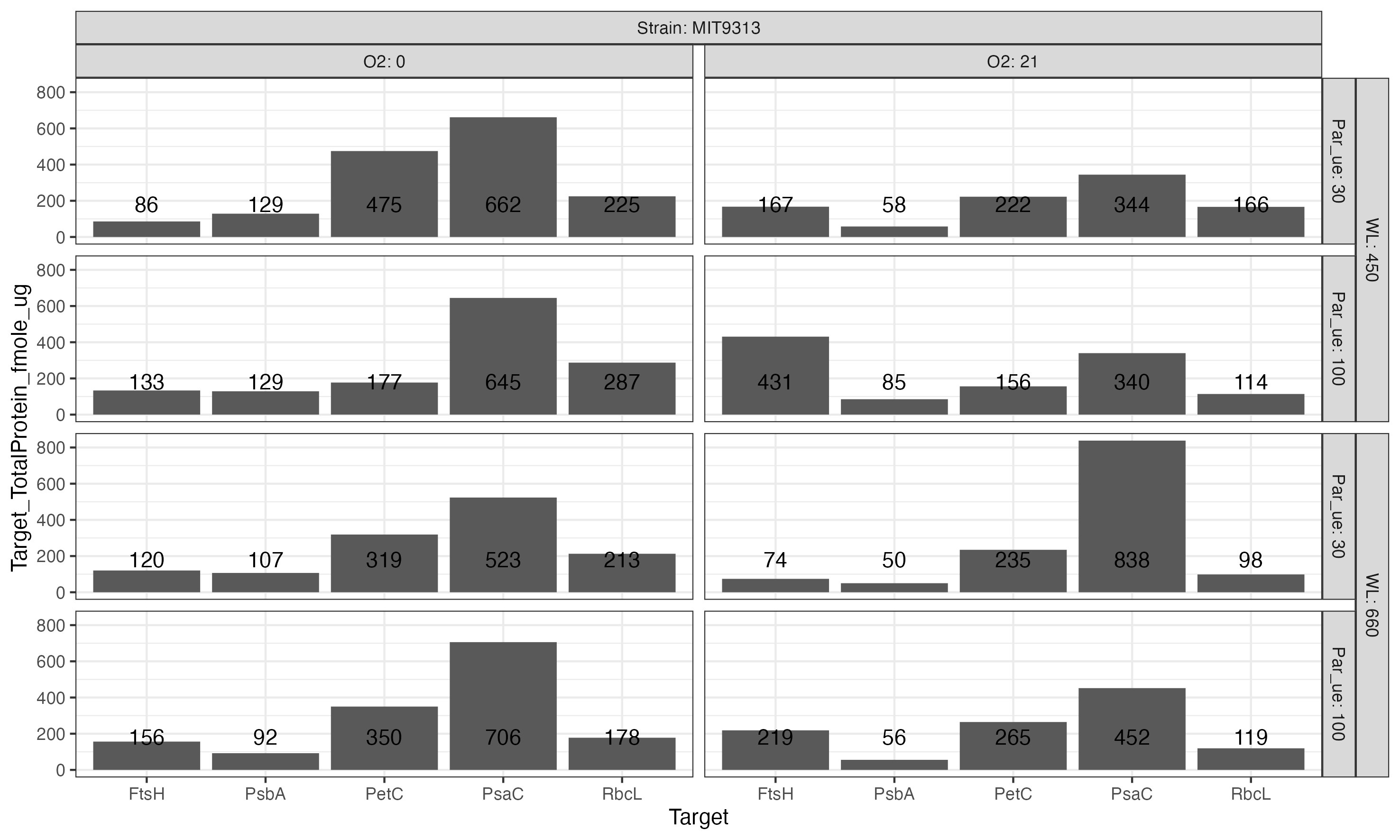


Figure 0.13: **fmole target protein per ug total protein for Prochlorococcus marinus MIT9313 (Low Light (LLIV) deep ocean clade).**  Growth Photosynthetically Active Radiation (PAR) (µmol photons m-2 s-1) and spectral wavelength are in rows; 2 levels of imposed growth dissolved O2 concentrations (µM) are in columns. Numbers over each bar are fmole/ug

# Supplemental

# 1 Acknowledgements

Jonah Sheinin assisted with growth of some cultures for protein analyses, while Carlie Barnhill assisted with code for import of multicultivator growth data files.

# 2 Funding

To Be Entered through PLoSONe system and deleted here Czech Academy of Science visiting fellowship (DAC) Canada Research Chair in Phytoplankton Ecophysiology (DAC) Natural Sciences and Engineering Research Council of Canada, ‘Latitude and Light’ (DAC) Canada Foundation for Innovation (DAC) New Brunswick Foundation for Innovation (DAC) Rice Graduate Fellowship (MS)

# References

1. Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D, et al. Culturing the marine cyanobacterium Prochlorococcus: Prochlorococcus culturing. Limnology and Oceanography: Methods. 2007;5: 353–362. doi:[10.4319/lom.2007.5.353](https://doi.org/10.4319/lom.2007.5.353)

2. Wickham H. Tidyverse: Easily install and load the tidyverse. 2023. Available: <https://CRAN.R-project.org/package=tidyverse>

3. Zeileis A, Grothendieck G, Ryan JA. Zoo: S3 infrastructure for regular and irregular time series (z’s ordered observations). 2021. Available: <https://zoo.R-Forge.R-project.org/>

4. Bellavia S, Gratton S, Riccietti E. A LevenbergMarquardt method for large nonlinear least-squares problems with dynamic accuracy in functions and gradients. Numerische Mathematik. 2018;140: 791–825. doi:[10.1007/s00211-018-0977-z](https://doi.org/10.1007/s00211-018-0977-z)

5. Elzhov TV, Mullen KM, Spiess A-N, Bolker B. Minpack.lm: R interface to the levenberg-marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. 2016. Available: <https://CRAN.R-project.org/package=minpack.lm>

6. Wood S. Mgcv: Mixed GAM computation vehicle with automatic smoothness estimation. 2022. Available: <https://CRAN.R-project.org/package=mgcv>