

Diversitet og sesongvariasjon av marine pelagiske protister i Indre Oslofjord, undersøkt med metastrekkoding.

Seasonal variation and diversity of marine pelagic protist in the inner Oslofjord, revealed by metabarcoding.

Karoline Saubrekka



Masteroppgave ved Institutt for Biovitenskap

UNIVERSITETET I OSLO

Februar 2019

Diversitet og sesongvariasjon av marine pelagiske protister i Indre Oslofjord, undersøkt med metastrekkoding.

Seasonal variation and diversity of marine pelagic protist in the inner Oslofjord,
revealed by metabarcoding.

Karoline Saubrekka
karoline_saubrekka@hotmail.com
+47 911 99 537

Veiledere

Professor Bente Edvardsen
bente.edvardsen@ibv.uio.no

Doktor Anders Kristian Krabberød
a.k.krabberod@ibv.uio.no

«Det er på grunn av temperaturen». *Bjørn Saubrekka*. Hele mitt liv.

Seksjon for akvatisk biologi og toksikologi

Institutt for biovitenskap

Universitet i Oslo

Februar 2019

© Karoline Saubrekka

År: 2019

Tittel: Diversitet og sesongvariasjon av marine pelagiske protister i Indre Oslofjord, undersøkt med metastrekkoding.

Forfatter: Karoline Saubrekka

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Protister finnes i de fleste habitater og spenner over et stort spekter av størrelser og former, fra mikroskopiske encellede arter til store velkjente tang- og tare-arter. Protister har en rekke avgjørende roller i det marine økosystem, som primærprodusenter, predatorer, nedbrytere og parasitter. I denne oppgaven fokuseres det på marine pelagiske protister.

I overvåkingen er det viktig med en metode som gir resultater som kan sammenlignes uavhengig av tid og sted. Ved å studere artssammensetning og mengde kan man få kunnskap om tilstanden i det marine økosystemet. Man kan registrere forekomst nye invaderende arter og arter som produserer gifter som potensielt kan være skadelig for oss mennesker. Ved å studere langtidsserier kan man oppdage trender og endringer i samfunnet som er forårsaket av eutrofiering eller klimaendringer. Ut ifra dette kan det iverksettes forbedringstiltak, for å ivareta marine høstbare ressurser, det biologiske mangfoldet og kyst- og vannmiljøet som rekreasjons og friluftsområdet.

Tradisjonelt i marin miljøovervåking blir fyttoplankton i vannprøver identifisert og telt i lysmikroskop. Denne metoden er avhengig av dyktige taksonomer og begrenses til celler som er store nok til å bli sett og godt nok preservert til å gjenkjenne i lysmikroskop. I denne studien ble det undersøkt muligheten for å bruke metastrekkoding i overvåkingen av marine pelagiske protister. Metastrekkoding er en molekylærbiologisk metode som overkommer begrensinger ved bruk av lysmikroskop. Det ble tatt vannprøver fra to dyp gjennom ca. en toårsperiode ved en av overvåkningsstasjonene i Indre Oslofjord. Vi oppformerte V4-regionen i 18S rRNA-genet fra protistene i vannprøvene og sekvenserte med Illumina MiSeq teknologien.

Dette er den først langtids-metastrekkodingstudien som er gjort i Indre Oslofjord. Det ble påvist protister fra 22 forskjellige protistgrupper, som tilhørte 17 rekker og 8 supergrupper. Protistsamfunnets sammensetning viste sesongvariasjon hvor diversiteten var størst om vinteren og lavest om sommeren. Ved bruk av V4-regionen i 18S rRNA-genet som markør var prøvene dominert av dinoflagellater, etterfulgt av diatoméer. Den mest abundante OTUen (Operational Taxonomic Unit) i datasettet ble klassifisert til ordenen Gymnodiniales. Diatoméene dominerte våroppblomstringen, og måneden etter så vi en økning i den relative abundansen til den heterotrofe stramenopile MAST-kladen. Sommer- og høstprøvene var ofte dominert av en dinoflagellat-OTU, som varierte fra måned til måned. Våre data viste ingen signifikant forskjell i protistsammensetning eller hydrografiske data ved de to dypene 0-2 m og

5 m. Denne studien påviste taksa som ikke ble påvist med lysmikroskop, spesielt en stor mengde dinoflagellater innen den parasittiske ordenen Syndiniales. Av dinoflagellat OTUene som ble klassifisert til art påviste metastrekkoding nakne- og cyste-dannende arter som ikke ble observert i lysmikroskop. Markøren V4-regionen i 18S rRNA-genet er tilstrekkelig til å skille de store protistgruppene og enkelte arter. Men var uegnet til å skille enkelte nært beslektede arter innen slekten *Dinophysis*. Dette kan forklares av mangel på referansesekvenser i den brukte referansedatabasen PR² eller at artene har for liten interspesifikk variasjon i V4-regionen.

Derfor foreslår jeg de to metodene lysmikroskop og metastrekkoding per i dag kan ansees for å være komplementære, og metastrekkoding vil være et godt supplement til den marine miljøovervåkingen. Resultatene våre kan bidra til større forståelse av protistsamfunnets sammensetning og sesongvariasjon i kystnære områder og av protistergruppers globale utbredelse.

Abstract

Protists are present in most habitats and span a wide range of sizes and shapes, from microscopic one-celled species to well-known seaweeds. Marine protists serve several key functions in marine eco-systems – as primary producers, predators, decomposers and parasites. This thesis has been concerned with marine pelagic protists.

When monitoring marine eco-systems it is important to employ a method that yields data which is reproducible and can be compared to other studies. By studying the diversity and abundance of species in samples of seawater, knowledge about the water quality of marine eco-system can be gathered. In the short term, invasive and disruptive species or species that produce toxins harmful to humans can be detected. In the long term, changes and trends caused by eutrophication or shifts in climate can be discovered. This in turn can allow us to protect, safeguard or even improve harvestable marine resources and the biological diversity, and to maintain the coastal area as a place of recreation.

Traditionally in marine environmental monitoring, phytoplankton species composition and abundance in water samples have been studied using optical microscopes. This method has its drawbacks: it requires skilled taxonomists and is limited to cells that are large enough to be seen and recognized under the light microscope. In this study, we investigated whether it is possible to use metabarcoding as a tool in the monitoring of marine pelagic protists. Metabarcoding is a molecular method, which can overcome the limitations connected with the use of optical microscopes. We took samples from two depths over a nearly two years period at one of the monitoring stations in the inner Oslofjord. We amplified the V4-region of the 18S rRNA gene from the protists in our samples and sequenced this marker gene using the Illumina MiSeq technology.

This is the first long-term metabarcoding study of the plankton in the inner Oslofjord. Protists from 22 different protist groups, belonged to 17 phyla and eight supergroups were detected. The composition of the protist community showed a seasonal variation where the diversity peaked during winter and had a minimum during summer. Using the V4-region in the 18S rRNA as marker, the samples were dominated by dinoflagellates, followed by diatoms. The most abundant OUT (Operational Taxonomic Unit) in our dataset was assigned to the order Gymnodiniales. The diatoms dominated during the spring bloom, followed by a high relative abundance of heterotrophic Strameopile MAST-clade. The summer and fall samples were often

dominated by a single dinoflagellate OTU that differed between the months. Our data did not show any significant difference in protist composition or hydrographic data between samples taken from the two depths, 0-2 and 5 meters. We detected taxa by metabarcoding that were not detected using optical microscopy, in particular a large amount of dinoflagellates in the parasitic order Syndiniales. Of the dinoflagellate OTUs that were classified to species, the metabarcoding detected naked and cyst-forming species that were not detected using optical microscopy. The V4-region is sufficient as a marker to differentiate between the large groups of protists and some species. But could not differentiate between some closely related species, such as some species of the dinoflagellate genus *Dinophysis*. This may be explained either by the lack of reference sequences in the reference database PR² or by the species having too low interspecific variation in the V4-region.

I therefore suggest that optical microscopy and metabarcoding can be considered as being complementary. Metabarcoding can be an efficient supplement to the monitoring of marine environments. Our results can contribute to a greater understanding of the composition and seasonality in protist communities in coastal areas and the global distribution of protists groups.

Forord

Dette prosjektet har vært knyttet til det NFR-støttede forskningsprosjektet “PhD Plankton - Change in phytoplankton community structure in the inner Oslofjord during a century of sampling”, prosjektnummer 262960. Elisabeth Lundsør har vært ansatt som nærings-PhD-student på prosjektet og samlet inn noen av prøvene fra DK1 (tabell S.1). Takk til Fagrådet for vann- og avløpsteknisk samarbeid i indre Oslofjord for toktid og tilgjengeliggjøring av CTD-data. Alle molekylære analyser ble utført på Institutt for biovitenskap ved Universitetet i Oslo (UiO). Sekvensering av DNA-prøver ble gjort av Norsk sekvenseringscenter (NSC), UiO på Ullevål.

Å skrive denne masteroppgaven har vært en læringsprosess fra start til slutt. Jeg har etterstrebet å få mest mulig ut av den tiden jeg har hatt til rådighet og jeg føler selv at læringsutbyttet har vært enormt. Jeg er lang ifra utlært og har kun berørt en liten del av overflaten i dette enorme faget, som jeg gleder meg til fremtidig fordyping i. Jeg håper denne oppgaven gjenspeiler hvor gøy og givende jeg syntes dette har vært.

Det som har gjort dette spesielt er alle menneskene, uten dere hadde jeg ikke lært noen ting. Jeg vil først takke mine veiledere Bente Edvardsen og Anders K. Krabberød. Da jeg skulle velge oppgave, valgte jeg ikke oppgave men veileder, Bente. Jeg beundrer deg, din lidenskap for mikroalger og din ubegrenset kunnskap. Anders, jeg har aldri før møtt en med dine pedagogiske ferdigheter, du har lært meg å elske for-løkker, koding i R og "dobbeltnutter". Videre vil jeg takke Stein Fredriksen som anbefalte meg å skrive på norsk. Det har nok vært noe av det mest utfordrende, men i gjengjeld har det gitt meg ferdigheten til å uttrykke og forstå faget på eget språk. En som har vært uunnværlig har vært Tom Andresen, den klokeste mannen jeg kjenner, som har lært meg at jeg ikke skal sparke inn åpne dører og som selv alltid har hatt døren åpen for å svare på mine utallige spørsmål. Takk til Bushra Muzammil og Luka Šupraha for opplæring på lab, til Josefin Titelman for å si at jeg har dårlig tid, til Hilde Tanstad for å dele sine ordinasjonsferdigheter, til Sandra Gran-Stadniczeňko for å dele sine R-skript, til Rikke Runde for å lære meg illustrator, til Kai Håkon Christiansen for et privat lynkurs i oseanografi, til Wenche Eikrem for å hjelpe meg å forstå den marine overvåkingen i Norge, til Bård Sjøberg og Kjell Magnus Sjøberg for korrekturlesing og kommentarer, til Marit F. M. Bjorbækmo for gjennomgang av litteraturlisten.

Takk til kaptein Sindre og mannskapet, Jan og Tom på Trygve Braarud for fine tokt og opplæring i bruk av utstyr på båten. Takk til Norsk sekvenseringscenteret for å svare på mine henvendelser om hvordan sekvenseringen fungerer. Takk til AKVA, med sitt gode miljø og sine fine, kunnskapsrike og hjelpsomme folk.

En som har gjort hverdagen i denne oppgaven til et samarbeid og vennskap er Hulda K. Bjørneklett. Jeg er så glad for og i deg, og ikke minst stolt av alt vi har fått til.

Til slutt, med hele mitt hjerte vil jeg takke Anders Sjøberg, min mann, for å få endene til å møtes i vår hverdag som småbarnsforeldre. Til Helmer, vår sønn, som hver dag har fått meg til å være i nuet.

Til mamma og pappa; nå gleder jeg meg til å dra på hytta med dere!

Anerkjennelser

Flere av CTD-datene brukt i denne oppgaven er mottatt fra Fagrådet for vann- og avløpsteknisk samarbeid i indre Oslofjord. For noen av prøvetakningsdatoene mangler det CTD-data og det ble ikke alltid tatt prøver fra dypet 0-2 m (listet i tabell S.1).

I denne oppgaven har jeg valgt å ikke følge norskstandard med komma som desimalseparator, men heller benyttet engelskstandard med punktum. Dette er fordi programvarepakke R bruker punktum. For et helhetlig og ryddig inntrykk ønsket jeg at dette skulle samsvare gjennom oppgaven.

I figur 3.1, 3.5, 3.9-3.11 er fargekoder identiske. Den samme fargen representerer den samme gruppen/rekken protister i alle figurene. Fargekodene beskrives best i figur 3.1, der supergruppene er oppgitt. Senere i oppgaven vil ikke supergruppen bli tydeliggjort i like stor grad.

I figurer der dypene 0-2 m og 5 m sammenlignes er 0-2 m i grønt og 5 m i blått, figur 3.17 – 3.23.

Supplementære figurer (figur S.1 – S.7) og tabeller (tabell S.1 – S.11) tilhører stortsett hovedteksten, men er plassert bak hovedteksten (fra s.125) før appendiks.

Det taksonomiske klassifiseringssystemet brukt i denne oppgaven er i hovedsak systemet brukt i referansedatabasen PR² (Guillou mfl., 2013). I noen tilfeller har klassifisering blitt korrigert for feil mot ALGAEBASE (algaebase.org), som i stor grad bruker det samme referansesystemet som PR².

Årstidene i denne oppgaven er definert til:

Vår	15 februar - mai
Sommer	Juni - august
Høst	September - oktober
Vinter	November -14 februar

Det er tre datasett som er brukt til analyser i denne studien; det opprinnelige datasettet, protist datasettet og subsamplet protistdatasett. Det blir gjennom oppgaven påpekt hvilket datasett som blir omtalt/fremstilt.

Hele OTU-tabellen er tilgjengelig på figshare

DOI: 10.6084/m9.figshare.7637036.v2

Den blir gjennom oppgaven referert til som "OTU-tabell i figshare" og ligger i litteraturliste under Saubrekka & Bjørneklett, 2019.

Ordforklaring

Alignmet er DNA-sekvenser som ligger under hverandre slik at man kan sammenligne baserekkefølgen i sekvensene.

Abundans antall sekvenser.

Diversitet et annet ord for mangfold

Cyste en hvilespore.

Eutrofiering i havet er økt fytoplankton- og planteproduksjon på grunn av økt næringstilførsel.

Indikator art indikerer spesielle miljøforhold.

Jevnhet forteller om hvor like arter er i antall individer (her sekvenser) i et samfunn.

Kimær en hybrid sekvens fra mer enn en organisme.

Klustre å gruppere/slå sammen sekvenser etter likhet. Et ledd i dannelsen av en OTU.

Kontig overlappende korte deler av en sekvens som tilsammen utgjør hele sekvensen.

Opportunistiske art er en konkurransesvak art som kan forkomme i høyt antall om forstyrrelser har fjernet de konkurransesterke artene.

OTU står for "Operational Taxonomic Unit" og er en pragmatisk erstatning for art.

Pielou indeks (J') er en jevnhetsindeks som bygger på Shannon indeks (H'), gitt med

$$\text{Pielou indeks } J' = \frac{H'}{H'_{max}}$$

der H'_{max} maksimumsverdi for H' .

Pipeline er en kjede av prosesser der data blir behandlet.

Relativ abundans prosentandel sekvenser.

Rikhet antall OTUer.

Shannon indeks er en diversitetsindeks som øker med antall arter (her OTUer) i samfunnet og er en vanlig brukt indeks for å karakterisere diversitet i et samfunn.

$$\text{Shannon indeks } (H') = - \sum_{i=1}^s p_i \ln p_i$$

der p = relative forekomst (n/N) av antallet sekvenser til en spesifikk OTU (n) delt på det totale antallet sekvenser (N). \ln er den naturlige logaritmen, \sum er summen til beregningen og s er antall OTUer.

Sprangsjikt er det vannlaget der tetten raskt endres fra lav til høy. Vannet sjiktes når vann med lavere tetthet legger seg over tyngere lag med vann.

Subsample ta ut en prøve av en større prøve.

Øvre blandingslag Vannmassene over sprangsjiktet danner det øvre blandingslaget og er definert til å være homogent i tetthet, temperatur og saltholdighet.

Innholdsfortegnelse

1	Innledning.....	1
1.1	Protister.....	1
1.2	Indre Oslofjord	7
1.3	Marin miljøovervåkning.....	10
1.3.1	Overvåking av Indre Oslofjord	11
1.4	Metastrekkoding	13
1.5	Formål med oppgaven	18
2	Materialer og metoder	19
2.1	Lokalitet.....	19
2.2	Innsamling	20
2.3	Molekylære metoder.....	21
2.3.1	DNA isolering og kvantifisering.....	21
2.3.2	PCR	23
2.3.3	Forberedelse til Illumina sekvensering	25
2.3.4	Illuminasekvensering	27
2.4	Analyse av sekvensdata	30
2.4.1	Bioinformatikk	30
2.4.2	Statistiske analyser og fremstilling av data.....	33
3	Resultater.....	36
3.1	Metastrekkoding-data	36
3.2	Variasjon gjennom året.....	45
3.2.1	Sesongvariasjon i miljøvariabler.....	45
3.2.2	Protistsamfunnets sammensetning gjennom årstider	48
3.3	Dinoflagellat-taksa.....	58
3.3.1	De mest abundant og frekvente OTUene i Dinoflagellata	58
3.3.2	Dinoflagellat-taksa sammenlignet med lysmikroskopfunn.....	64
3.4	Forskjellen mellom dypene	67
4	Diskusjon.....	77
4.1	Protistdiversitet undersøkt med metastrekkoding.....	77
4.2	Variasjon gjennom året.....	84
4.3	Lysmikroskop vs. metastrekkoding.....	88

4.4	Ingen signifikant forskjell mellom dypene	92
4.5	Metodekritikk	96
4.6	Fremtidige studier og konklusjon	99
4.6.1	Hadde jeg hatt mer tid med datasettet ville jeg	99
4.6.2	Hvordan drive overvåkning i framtiden	101
4.6.3	Konklusjon	102
	Litteraturliste	103
	Supplementære figurer	125
	Supplementære tabeller	131
	Appendiks.....	149

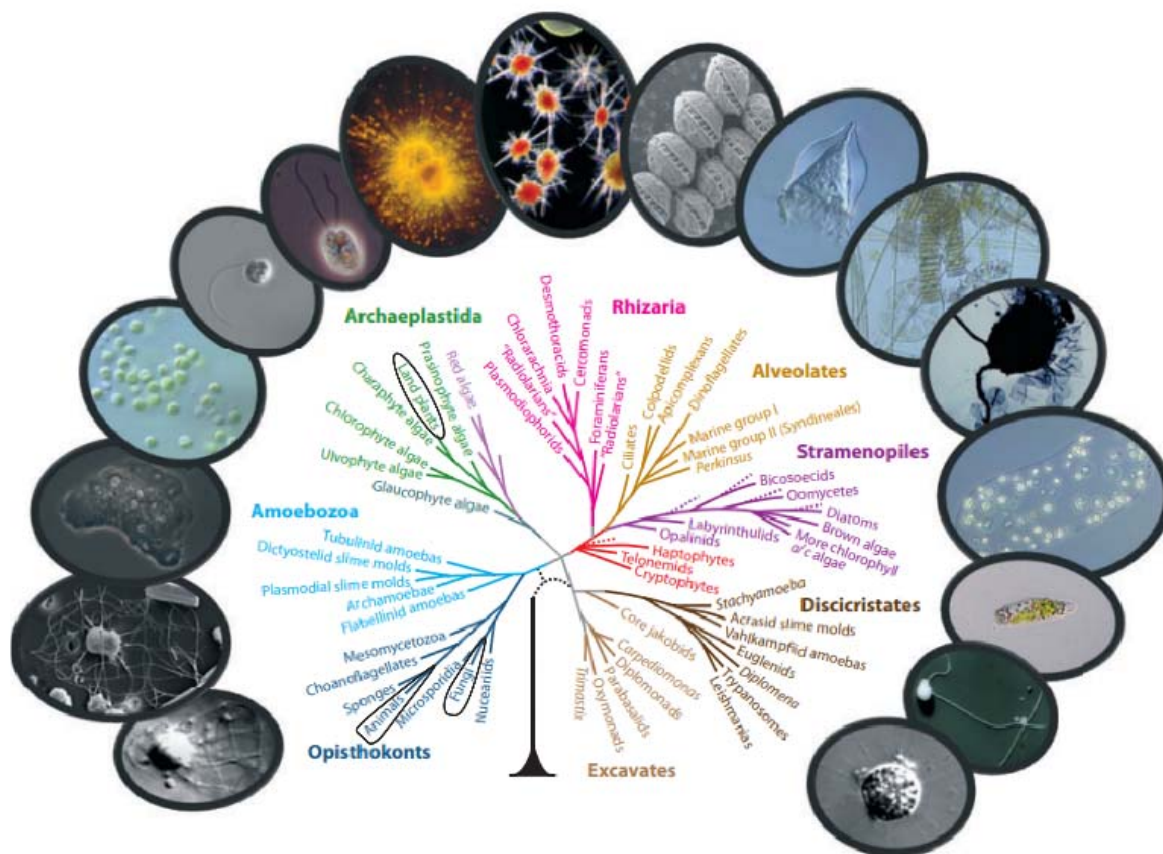
1 Innledning

1.1 Protister

Protister er en mangfoldig gruppe av eukaryote organismer, som hovedsakelig er encellede eller forekommer som enkle kolonier. Rike Protista ble først introdusert av Haeckel i 1866, og bestod av det han kalte primitive eukaryote organismer som ikke kunne plasseres i plante- eller dyreriket. Rike Protista har blitt revidert mange ganger (Adl mfl., 2005, 2007; Corliss, 1984; Scamardella, 1999; Whittaker, 1969, 1972) og fylogenetiske analyser har vist at denne gruppen er parafyletisk. I moderne taksonomi klassifiseres ikke lenger Protista som et rike, men ordet Protister har blitt et samlebegrep som omfatter eukaryote organismer som ikke er planter (Streptophyta), sopp (Fungi) eller dyr (Metazoa), og mange er langt ifra primitive (Cavalier-Smith, 1993). Protister er spredt utover i det eukaryote slektskapstreet og utgjør majoriteten av de eukaryote linjene (figur 1.1). Å studere protister har derfor vært av stor interesse for å forstå den evolusjonære utviklingen og fylogenetiske sammensetningen i livets tre (Burki, 2014; Keeling mfl., 2005). Protister er til stede i de fleste habitater og spenner over et stort spekter av størrelser, former og fylogenetisk tilhørighet (Graham mfl., 2009). I denne oppgaven fokuseres det på marine pelagiske protister, som hovedsakelig befinner seg innenfor størrelsesordenen pico- (0.2-2 μm), nano- (2-20 μm) og mikrop plankton (20-200 μm) og de fototrofe gruppene vektlegges.

Marine pelagiske protister er plankton, det vil si at de lever i de frie vannmassene og har så liten horisontal egenbevegelse at de driver med havstrømmene. Det er antatt at vi har kjennskap til mindre enn 10% av de marine protistene som finnes, og antagelig enda færre når det kommer til nano- og picoplankton ettersom det stadig oppdages nye organismegrupper (López-García mfl., 2001; Massana mfl., 2004; Not mfl., 2007).

Protister har en rekke avgjørende funksjoner i det marine økosystem, som primærprodusenter, predatorer, nedbrytere og parasitter (Tabell 1.1). Fytoplanktonet danner fundamentet til den marine næringskjeden. Ved å omdanne karbondioksid og vann til sukkerforbindelser og oksygen ved hjelp av sollys, gjør de denne energien tilgjengelig for høyere trofiske nivå ved å selv bli spist.



Figur 1.1. Fylogenetisk tre av eukaryote supergrupper (uthevet skrift) og rekker basert på publiserte molekylær- fylogenetiske og morfologiske data. Utringet i sort er linjer som hovedsakelig ikke inneholder protister. Figuren er modifisert fra Caron mfl., 2012 og basert på Baldauf, 2008.

I tillegg til å fiksure karbon i fotosynten, har mange et høyt innhold av fettsyrer (f.eks. omega-3-fettsyrene DHA og EPA), som akkumuleres oppover i næringskjeden og er essensielle for bl.a oss mennesker (Doughman mfl., 2007). Det høye innholdet av disse fettsyrene som vi finner i fisk er på grunn av direkte eller indirekte opptak av fytoplankton som har produsert disse. Selv om fytoplanktonet (inkludert prokaryote cyanobakterier) kun utgjør 0.2% av jordas fotosyntetiserende biomasse (Falkowski mfl., 1998), er de alene ansvarlig for opp mot 40-50% av den globale årlige primærproduksjonen (Field mfl., 1998; Longhurst mfl., 1995). Det vil si at omtrent halvparten av oksygenet vi puster inn er produsert av fytoplankton i havet.

Fytoplanktonet har rask vekst og kan om forholdene er gode mer enn doble antallet celler i løpet av en dag (Eppley, 1972). Gode forhold er tilstrekkelig tilgang på sollys, optimal temperatur, få predatorer og ubegrenset tilgang på næringsalter. Tre potensielt begrensende næringsalter er fosfor, nitrogen (Howarth, 1988; Paasche & Erga, 1987; Ryther & Dunstan, 1971) og silikat

Tabell 1.1. Oversikt over marine protistgrupper og deres ernæringsmåte, samt eventuell giftighet for menneske. Hovedernæring = x og tilleggsernæring = (x), med tilhørende kildehenvisning. F = Fototrof, H = Heterotrof, P = Parasittisk og G = Giftig.

SUPERGRUPPE	REKKE/GRUPPE	F	H	P	G	KILDER
ALVEOLATA	Apicomplexa		x	x		- Ciancio mfl., 2001 - Leander, 2006; 2007 - Salomon mfl., 2003
	Ciliophora		x	x		- Lynn, 2012 - Sherr & Sherr, 1994 - Skovgaard & Daugbjerg, 2008
	Dinoflagellata	x	x	x	x	- Edvardsen mfl., 2003 - Guillou mfl., 2008 - Jeong mfl., 2010 - Not mfl., 2012 - Sherr & Sherr, 1994
	Perkinsea		x	x		- Azevedo, 1989 - Burreson & Ragone Calvo, 1996 - Park & Choi, 2001
AMOEBOZOA	Amoebozoa		x	x		- Cavalier mfl., 2004 - Pawlowski & Burki, 2009
APUSOZOA	Apusozoa		x			- Cavalier-Smith & Chao, 2003b
ARCHAEPLASTIDA	Chlorophyta	x				- Eikrem & Throndsen, 1990 - Graham mfl., 2009, kap. 16-20 - Not mfl., 2004
	Rhodophyta	x				- Graham mfl., 2009 kap. 15 - Throndsen mfl., 2007
EXCAVATA	Euglenozoa	x	x			- Heyden mfl., 2004
HACROBIA	Centroheliozoa		x			- Burki mfl., 2009
	Cryptophyta	x	(x)			- Marin mfl., 1998
	Haptophyta	x	(x)		x	- Edvardsen & Paasche, 1992 - Not mfl., 2012
	Katablepharidophyta		x			- Okamoto & Inouye, 2005
	Picozoa		x			- Seenivasan, 2013
	Telonemia			x		- Klaveness mfl., 2005
OPISTHOKONTA	Mesomycetozoa		x	x		- Medina mfl., 2001
	Choanoflagellata		x			- Espeland & Throndsen, 1986; - Gold mfl., 1970
RHIZARIA	Cercozoa		x			- Cavalier-Smith & Chao, 2003a, 2003b
	Foraminifera		x	x		- Pawlowski mfl., 2003
STRAMENOPILES	Radiolaria		x			- Nikolaev mfl., 2004
	Heterotrofe Stramenopiles		x			- Massana mfl., 2004
	Ochrophyta	x	(x)		x	- Bates mfl., 1998 - Graham mfl. 2009, kap. 10, 12-14 - Not mfl., 2012

(Conley mfl., 1993). Disse finnes naturlig i ganske små mengder i sjøvann i forhold til behovet hos fytoplanktonet, men kan forekomme i store mengder i kystnære områder. Eutrofiering i havet er økt fytoplankton og planteproduksjon pga. økt tilførsel av næringsalter. Eutrofiering har blitt et globalt problem i kystnære områder der befolkningsveksten har vært stor. Næringsalter i avløpsvann fra befolkning og industri, sammen med økt avrenning fra land som bringer med seg næringsalter fra gjødsel brukt i jordbruk og organiske stoffer, ender opp i havet og endrer økosystemet der tilførselen er stor (Diaz & Rosenberg, 2008). Enkelte, også såkalte opportunistiske arter, kan raskt utnytte det økte næringsstilskuddet og få en oppblomstring på bekostning av andre arter. Fytoplanktonvekst i vannmassene og unormalt stor tilførsel av organisk materiale fører til en akkumulering av dødt materialet i bunnområdene og en økt biologisk nedbrytning som krever oksygen. Oksygenmangel kan igjen føre til tap av bunndyr og lite tilgang på sollys kan føre til tap av bunnflora. I tillegg kan store oppblomstringer av giftige alger føre til både fiskedød og gift i matnyttige skjell som blåskjell.

I planktonet er tre av de viktigste marine primærprodusenter diatoméer (i rekken Ochrophyta), dinoflagellater (rekken Dinoflagellata) og haptofytter (rekken Haptophyta) (Not mfl., 2012).

Diatoméer (kiselalger) er den viktigste algegruppen i våre farvann (Skjoldal mfl., 2004) og er fordelt på de tre klassene Bacillariophyceae, Coscinodiscophyceae og Mediophyceae i rekken Ochrophyta. Diatoméene er encellet og flere arter kan danne lange kjede-formede kolonier som kan bestå av over 20 celler (Smayda & Boleyn, 1966). Majoriteten av marine diatoméer er fototrofe og de er den mest diverse fytoplankton gruppen av alle (Armbrust, 2009). De kan forekomme i stort antall i tempererte havområder (Thronsen mfl., 2007) og i tropiske områder med oppstrømning av næringsrikt dypvann (Not mfl., 2012). Noen arter kan til og med ha store oppblomstringer under isen i Arktis (Arrigo mfl., 2012) og Antarktis (Ichinomiya mfl., 2007). Det er estimert at diatoméene er ansvarlig for 20% av den årlige globale karbonfikseringen (Armbrust mfl., 2004; Graham mfl., 2009, s. 223).

Diatoméene har et ytre skall som består av kisel (silisium). I det dominerende stadiet av livssyklusen har diatoméene i pelagialen ingen betydelig egenbevegelse. Tilsynelatende har de andre strategier som kan hjelpe dem å regulere deres plassering i vannsøylen. Hos noen arter er det vist at det å danne kjeder kan bidra til å holde dem flytende (Smayda & Boleyn, 1966), når næringsstilgangen er god (Takabayashi mfl., 2006). Men hos andre virker det som om det er en strategi for å synke (Smayda & Boleyn, 1966a) til f.eks. mer næringsrikt vann. Noen diatoméer regulerer ione-sammensetningen i saftvakuolen for å bli lettere eller tyngere enn vannet som

omgir cellen (Boyd & Gradmann, 2002). Hos slekten *Chaetoceros* har cellene lange utvekster kalt "børster", som kan bidra til senke synkehastigheten og øke næringsopptak (Bienfang mfl., 1982). Disse børstene kan under store oppblomstringer danne problemer for fiskeoppdrett, ved at børstene påfører mekanisk skade på gjellene til fisken (Yang & Albright, 1992). Noen arter innen slekten *Pseudo-nitzschia* danner gifter som kan gi nerveskader hos oss mennesker (Bates mfl., 1998). Diatoméer dominerer ofte vår oppblomstringen, men tilgangen på silisium kan være en begrensende faktor (Paasche & Østergren, 1980).

Dinoflagellater (fureflagellater) hører til rekken Dinoflagellata, i supergruppen Alveolata. Dinoflagellatene har egenbevegelse ved hjelp av to flageller og de finnes sjeldent i kolonier (Thronsen mfl., 2007). Noen dinoflagellater kalles nakne (athekate), det vil si at de mangler cellevegg med cellulose (Thronsen mfl., 2007) og gjelder hele ordenen Gymnodiniales (de Salas mfl., 2003). Omtrent halvparten av de kjente dinoflagellatene kan drive fotosyntese (Taylor mfl., 2009). Av disse er mange miksotrofe, som er en blanding av foto- og heterotrof ernæring (Jeong mfl., 2010; Not mfl., 2012). De heterotrofe dinoflagellatene har flere forskjellige strategier for næringsopptak som fagocytose, bruk av pallium og peduncel (Hansen, 1991). Dinoflagellatene er diverse og tallrike, og har vid utbredelse i både tropiske og polare områder (Taylor mfl., 2009). Dinoflagellatene er kjent for å danne store oppblomstringer og noen arter er ansvarlige for naturfenomener som morild og oppblomstringer som farger vannet rødt, kalt "red tide" på engelsk (Huang & Qi, 1997; Millie mfl., 1997; Schmitter, 1971). I tillegg kan mange arter være svært giftig for mennesker og vannlevende dyr (Anderson mfl., 1990; Edvardsen mfl., 2003; Jellett mfl., 1995; Yasumoto mfl., 1985). Dinoflagellat-artene i Norske farvann som er mest giftig for mennesker og som overvåkes i Norge i dag, finner vi i slektene *Alexandrium*, *Azadinium* og *Dinophysis*.

Haptofytter (svepeflagellater) hører til rekken Haptophyta og deles inn i klassene Prymnesiophyceae og Pavlovophyceae. Disse er vanligvis mindre enn diatoméene og dinoflagellatene, med en cellediameter på 3 – 70 µm (Thronsen mfl., 2007). De er hovedsakelig fototrofe. Cellene er stort sett dekket av organiske skjell. Hos coccolithophorene er cellene i tillegg omgitt av kalkplater (coccolith) (Thronsen mfl., 2007). Coccolithophorene, f.eks *Emiliana huxleyi*, kan forekomme i store mengder og gir da sjøen et blakket grønnlig utseende (Holligan mfl., 1993). Noen arter, som *Emiliana huxleyi* og *Phaeocystis* spp., kan produsere store mengder DMSP (dimetylsulfidpropionat) som spaltes til DMS i vannet, som i sin tur frigjøres til atmosfæren under oppblomstringer og kan påvirke

klimaet (Graham mfl., 2009, s. 35–37). I atmosfæren omdannes DMS til svoveldioksid og sulfat. Sulfat avkjøler klima ved å spre solstråler i atmosfæren og genere skyer som reflekterer sollyset tilbake ut i universet og dermed redusere innstrålingen som treffer jorda. Svoveldioksid kan gi et bidrag til svovelinnholdet i sur nedbør. Noen svepeflagellater kan være giftige, som *Prymnesium polylepis* (tidligere *Chrysochromulina polylepis* (Edwardsen mfl., 2011)). I 1988 hadde *P. polylepis* en massoppblomstring over 60 000 km², i Kattegat og Skagerrak, som forårsaket stor skade på fisk og bunnfaunaen (Skjoldal og Dundas, 1991).

Det er vist at små celler generelt har høyere vekstrate enn store (Graham mfl., 2009, s. 488–491). En mulig forklaring er at celler med et større forhold mellom overflate og volum, raskere kan ta opp næringsalter over celleoverflaten enn store arter med stort volum i forhold til overflate. I enkelte marine økosystem er pico- og nanoplanktonet ansvarlig for over 60% av primærproduksjonen (Uitz mfl., 2010). I tillegg til å være primærprodusenter er fytoplanktonet en viktig komponent i den globale biogeokjemiske sirkulasjonen. På grunn av økt fokus på klimaendringer, har det de siste par tiårene vært stor interesse i å forstå protisters rolle i karbonsyklusen (Riebesell, 2004; Uitz mfl., 2010). Den "korte-karbonsyklusen" kan beskrives slik: CO₂ i gassform blir gjennom fotosyntese omdannet til organiske karbonforbindelser i fytoplanktonet og næringskjeden. I organismers celleånding går det organiske karbonet tilbake til CO₂ i gassform. Ved at noe av planktonet synker når det dør kan det bidra til å fjerne karbondioksid fra det "korte-karbonkretsløpet", ved vertikal transport av karbon til havbunnen. Det som ikke spises av bunnlevende organismer eller brytes ned av bakterier vil over tid bli fossilt karbon, som f.eks. kalkstein. Det er estimert at kalsifiserende coccolithophorer alene bidrar til 25% av den årlig dypvannsakkumuleringen av karbon (Rost & Riebesell, 2004).

Selv om fytoplanktonet deler av året dominerer i biomasse i det øvre, godt opplyste lag (den eufotiske sonen) er det flere heterotrofe protistgrupper i planktonet som kan utgjøre en stor del av diversiteten (de Vargas mfl., 2015) (Tabell 1.1). De heterotrofe protistene kan være bakterivore, herbivore eller spise andre heterotrofe protister. Samlet er de et essensielt ledd fra det mikrobielle næringsnett til den tradisjonelle næringskjeden vi mennesker benytter oss av. Artssammensetningen og størrelsen på det heterotrofe samfunnet bestemmer hvor stor del av primærproduksjonen som bindes opp i det mikrobielle næringsnett og hvor mye som blir overført til høyere trofiske nivå. Det er estimert at heterotrofe herbivore consumerer 50% av den daglige primærproduksjonen (Calbet & Landry, 2004). Omtrent halvparten av de kjente frittlevende dinoflagellatene er heterotrofe og utgjør en viktig del av økosystemet ved å beite

på en rekke forskjellige fyttoplankton og bakterier. I tillegg produserer dinoflagellatene en stor mengde organiske materialer som bidrar til "marin snø", som i tillegg består av fekale peller fra dyreplankton og ofte tomme diatomé-skall (Alldredge & Silver, 1988). Disse aggregatene synker nedover gjennom vannsøylen og er viktige komponenter i næringskjeden i de dypere vannmassene. Rekken Ciliophora (flimmerdyr eller ciliater) blir ansett som topppredatorer i det mikrobielle-næringsnett, hvor noen arter også er parasitter. Apicomplexa og Perkinsea inneholder en mengde arter som er parasitter på og i marine evertebrater og vertebrater.

Ved å studere artssammensetning og mengde kan man få viktig kunnskap om tilstanden i det marine økosystemet. Man kan oppdage nye invaderende arter, og ved å studere langtidsserier kan man oppdage trender og endringer i protistsamfunnet over tid. Endringer kan være forårsaket av menneskelig påvirkning som eutrofiering eller klimaendringer. Utfra dette kan det iverksettes tiltak for å ivareta marine høstbare ressurser, bevare biologisk mangfold, kyst- og vann-miljøet som rekreasjons- og friluftsområder. I det norske statsbudsjett settes årlig av flere millioner kroner til forskning på havets helse (Nærings- og fiskeridepartementet, 2018) som bl.a. innebærer å identifisere og kvantifisere marine protister.

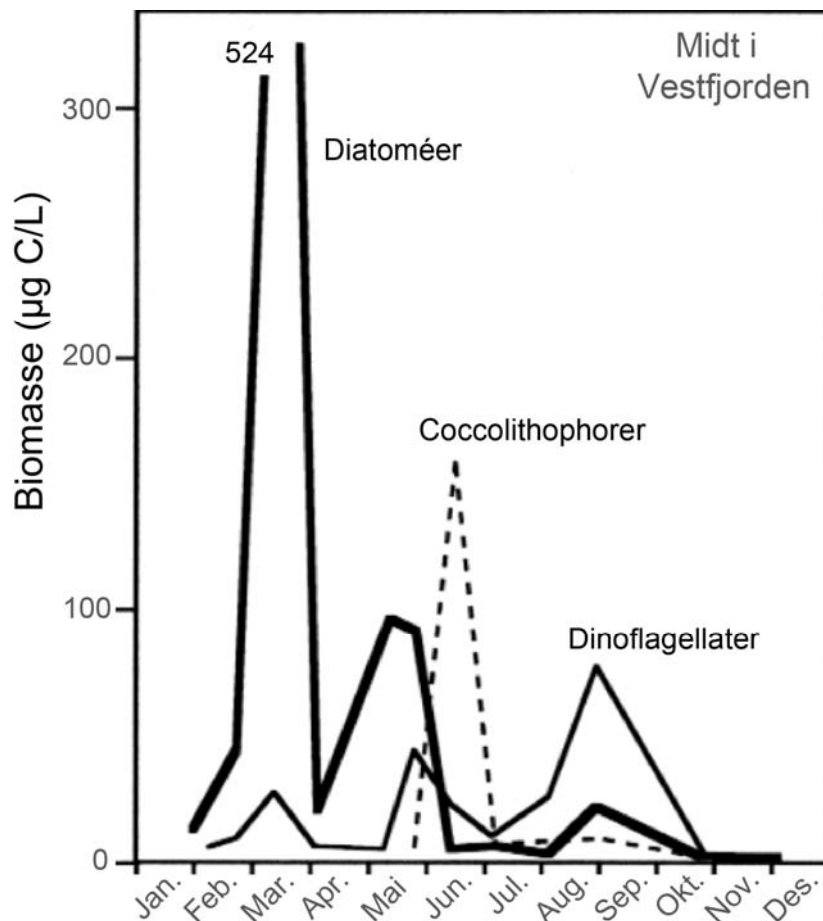
1.2 Indre Oslofjord

Vannet i Indre Oslofjord kommer via Ytre Oslofjord fra Skagerrak og fra elver. I Skagerrak blandes vannet fra tre havstrømmer: Atlanterhavsstrømmen (Golfstrømmen), Jylland-strømmen og den Baltiske strøm (figur 1.2). Atlanterhavsstrømmen er forbundet med verdenshavene. En gren fra denne strømmen renner inn i Nordsjøen og videre til Skagerrak. Jylland-strømmen bringer vann fra Den engelske kanal og Tyskebukta. Den Baltiske strøm bringer brakkvann fra Østersjøen nordover langs svenskekysten via Kattegat (Baalsrud & Magnusson, 2002).



Figur 1.2. Havstrømmene som fører vann til Skagerrak.

Vannfornyelse i Indre Oslofjord skjer som regel på vinteren, når nordavinden er vedvarende og sterk. Den fører de øvre vannmassene ut av fjorden og det dannes en motstrøm av dypere saltvann som strømmer inn over den grunne Drøbaksterskelen (20 m dyp og 1 km bred). Saltholdigheten i de øvre vannmassene varierer i stor grad gjennom året. Hvor lav saltholdigheten kan bli i dette laget, avhenger mye av hvor stor ferskvannstilførselen er i fra Drammenelva og Glomma, i Ytre Oslofjord. Dette gjør Indre Oslofjord til et omvendt estuarint system. Det vil si at i motsetning til de fleste fjorder på Vestlandet, kommer ikke den største ferskvannspåvirkningen fra elver som munner ut innerst i fjorden, den kommer utenfra og inn. I tider da nedbørmengden er stor kan småelvene og vann fra nedbørsfelt i Indre Oslofjord være med på å påvirke saltholdigheten nevneverdig.



Figur 1.3. Fytoplanktonets års-syklus i 1976 i midten av Vestfjorden, som ligger midt i Indre Oslofjord. Figuren er trolig basert på upubliserte data fra Paasche og Østergren og blir brukt i Norconsult (2017, 2018) og Paasche (2005) for å illustrere en typisk fytoplankton års-syklus i Indre Oslofjord. Figuren er modifisert.

Tidevannet i Indre Oslofjord føres inn og ut av fjorden to ganger i døgnet og vannmassene som påvirkes begrenses av den grunne Drøbaksterskelen. Det er i stor grad de samme vannmassene fra de øverste 20 meterne som fraktes inn og ut over terskelen. Vannstandsvariasjonen mellom flo og fjære er i snitt 28 cm og det er store vannmasser som skal fraktes ut og inn av fjorden daglig (Baalsrud & Magnusson, 2002).

Det dynamiske blandingslaget er det øverste vertikale vannlaget hvor turbulent omrøring av vannmassene er forårsaket av fysiske krefter (f.eks. vind, bølger, båttrafikk). Omrøringen i dette laget kan skje i løpet av kort tid, sekunder til minutter eller timer (Brainerd & Gregg, 1995). Det øvre blandingslaget er svært viktig for fytoplanktonvekst. Et sterkt sprangsjikt begrenser tilgangen på næringssalter fra de dypere lagene og de øvre vannmassene kan bli næringsfattig i perioder med stor fytoplanktonvekst (Paasche & Erga, 1987).

Det kan skilles mellom to typer primærproduksjon: ny og regenerert. Ny produksjon er vekst av fytoplankton basert på tilførsel av næringssalter til det øvre blandingslaget fra land og de dypere vannlagene (Rinde mfl., 1998, s. 59). Regenerert produksjon baseres på næringssalter som frigjøres fra bakterier, fototroft-, heterotroft- og dyreplankton på grunn av nedbryting, ekskrementer, urin, utskillelse fra cellene og "sloppy feeding".

Etter vinteren har dypt og næringsrikt vann kommet til overflaten og er tilgjengelig for fytoplanktonet. Når dagene blir lenger og sollyset tilstrekkelig vil våroppblomstringen kunne starte. Samtidig forekommer det en sjikning av vannmassene som er forårsaket av snøsmelting og oppvarming. I Indre Oslofjord er våroppblomstringen normalt i februar-mars (Paasche, 1990) og domineres vanligvis av diatoméer (figur 1.3). Diatoméene kan vokse eksponentielt i antall så lenge det er nok med næringssalter og mengden beiting er liten. Diatoméenes vekst begrenses enten av tilgangen på silisium, som de bruker til å bygge skallet sitt, eller nitrogen når lystilgangen er tilstrekkelig (Kristiansen mfl., 2001). Fytoplanktonbiomassen kan avta i samspill med at mengden heterotrofe protister og dyreplankton øker, siden de beiter ned diatoméene eller fytoplanktonet synker til bunn på grunn av næringsmangel (Kristiansen mfl., 2001). Senere om våren bringer elvene med seg smeltevann med et høyt innhold av næringssalter og en ny oppblomstring kan oppstå på forsommeren, vanligvis mai-juni i Indre Oslofjord (Paasche, 1990). Mot sommeren blir kystvannet delt av et skarpt sprangsjikt som starter på rundt 5-10 m (Paasche & Erga, 1987). I det øvreblandingslaget blir tilgangen på næringssalter lav og algebiomassen avtar. Det er vanlig å se dinoflagellatene dominere på sommeren. Dinoflagellatene kan holde seg i det øvre blandingslaget om dagen og drive

fotosyntese, for så å vertikalt migrere ned til 10 m dyp etter solnedgang (Hasle, 1950). Trolig for å hente næringssalter eller unngå predatorer. Det finnes også mange heterotrofe dinoflagellater og mikstotrofe, som begynner å beite på bakterier når fotosyntesen ikke gir tilstrekkelig næringsutbytte. Om høsten begynner sprangsjiktet å brytes ned og man kan se en høst oppblomstring, men når og i hvor stor grad avhenger av lokale klimatiske og hydrografiske forhold.

1.3 Marin miljøovervåkning

Forvaltning av norsk vannmiljø regulert gjennom Vannforskriften (Vannforskriften, 2007) og har som hovedmål at alle vannforekomster skal oppnå minst "god tilstand". Vannets tilstand måles ut ifra et sett med godkjente grenseverdier for ulike kvalitetskriterier. I regi av Miljødirektoratet gjennomføres en rekke overvåkningsprogrammer langs Norges kyst. De innhenter kunnskap om utvalgte komponenter i det marine økosystem og har som mål å fange opp endringer på et tidlig stadium. Spesielt fokuserer de på uønsket endring av næringssaltnivåer som fører til eutrofiering og klimaendringers påvirkning på det marine økosystem.

Marine økosystemtjenester, som bl.a. matproduksjon, råmaterialer, rekreasjon, resirkulering av næringsstoffer og gassregulering, har en svært høy årlig verdi som globalt ble estimert til 21 billioner (eng. trillions) amerikanske dollar (US\$) i 1997 (Costanza mfl., 1997), som per i dag mest sannsynlig er langt høyere. Fjord og kyst er globalt noen av de mest brukte økosystemer og nå mest truet av alle jordas natur-systemer (Halpern mfl., 2008; Lotze mfl., 2006; Worm mfl., 2006).

I Naturmangfoldloven §1 står det "Lovens formål er at naturen med dens biologiske, landskapsmessige og geologiske mangfold og økologiske prosesser tas vare på ved bærekraftig bruk og vern, også slik at den gir grunnlag for menneskenes virksomhet, kultur, helse og trivsel, nå og i fremtiden, også som grunnlag for samisk kultur."

Tap av artsmangfold og habitat truer havet som ressurs og vil i stor grad påvirke mennesket som benytter det. For å håndtere og bevare biologisk mangfold må man vite hva som er der, hva som går tapt, hvor og hvorfor.

1.3.1 Overvåking av Indre Oslofjord

Fagrådet for vann- og avløpsteknings samarbeid i indre Oslofjord er ansvarlig for overvåkingen i Indre Oslofjord. Overvåkningsprogrammet utføres av næringsbedrifter, som frem til 2015 var NIVA (Norsk institutt for vannforskning) og fra 2015-2018 var Norconsult. NIVA har siden 2015 fortsatt å gjennomføre årlige observasjoner med FerryBox-systemet av overflatevannet i Oslofjorden, som tidligere var en del av hovedprogrammet i overvåkingen av Indre Oslofjorden (Norli & Eikrem, 2018).

Historikk

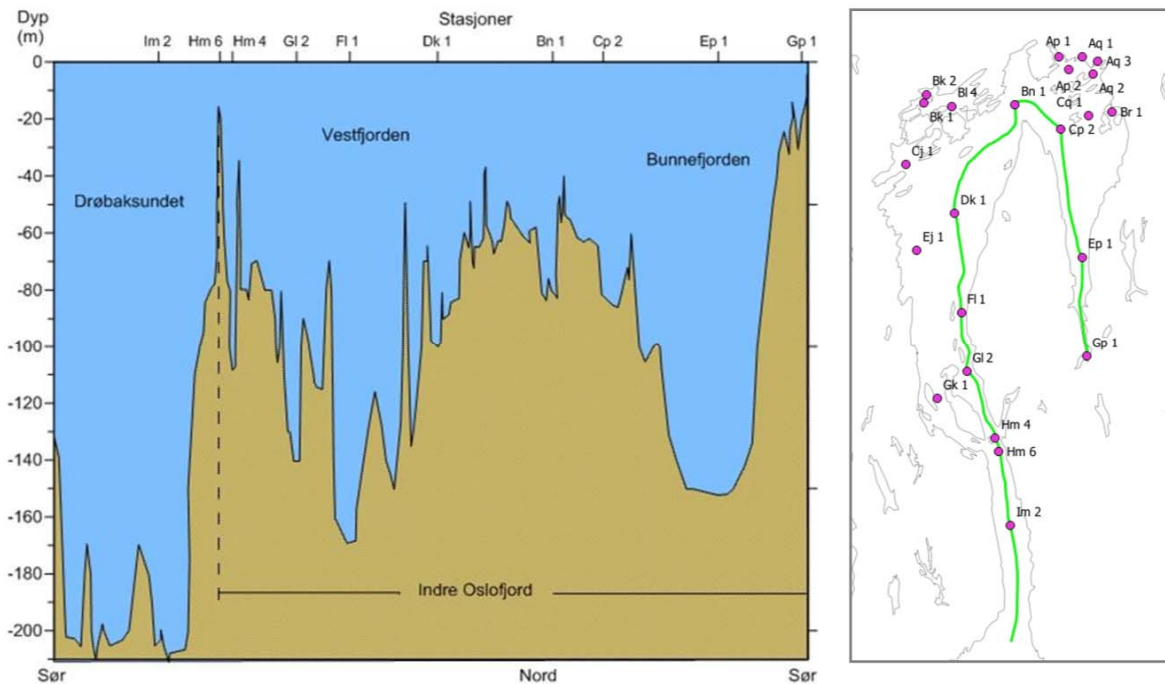
Siden starten av 1900-tallet har overgjødning vært et av hovedproblemene i Indre Oslofjord. Før 1970-årene ble avløpsvann fra Oslos befolkning ført rett ut i fjorden (Miljødirektoratet, 2017). I 1970-årene var tilførselen av næringssalter og organiske stoffer på sitt høyeste på grunn av stadig økende befolkning og industri. Overvåkningsprogrammet i Indre Oslofjord ble startet opp i 1973 og det ble iverksatt en rekke tiltak for å forbedre vannkvaliteten. Utbygging av ledningsnett og tunneller for oppsamling og transport av avløpsvann til effektive renseanlegg - VEAS og Bekkelaget - har ført til betydelig bedring av vannkvaliteten de siste ti-årene (Thaulow & Faafeng, 2013).

Utførelse

I overvåkningsprogrammet gjennomført av Norconsult i 2017 ble vannprøver samlet fra 25 vannprøvetakningsstasjoner (figur 1.4). Overvåking av fytoplanktonet skjedde i hovedsak ved å måle mengde klorofyll a og næringssalter i vannprøvene fra 10 av stasjonene og bestemme artssammensetning og biomasse ved å telle arter i lysmikroskop i vannprøver fra stasjonene Dk1 og Ep 1.

Klorofyll a er et mål på fytoplankton biomasse og inngår i som et biologisk kvalitetselement i Vannforskriften og utføres etter Veileder 02:2013 (Norconsult, 2018).

Artsammensetning og mengde analyseres fra 25 mL vannprøve som konserveres i 1% Lugols løsning og observeres i omvendt lysmikroskop i henhold til NS-EN 15972:2006 (Veiledning for kvantitative og kvalitative undersøkelser av marine planktonalger) og NS-EN 15204:2006 (Utermöhls metode om omvendtmikroskop). Som kvalitativanalyse inkluderes vertikalt hovtrekk, med maskevidde 10 µm, fra 30 – 0 meters dyp (Norconsult, 2018).



Figur 1.4. Bunntopografi i Indre Oslofjord (venstre) langs transsekt markert i grønt (høyre). Oversikt over vannprøvetakningsstasjoner i overvåkningsprogrammet i Indre Oslofjord (høyre og venstre) (Norconsult, 2017).

Utermöhls metode har globalt blitt standard metode for kvantitativ analyse av fytoplankton (Lund mfl., 1958; Paxinos & Mitchell, 2000). Metoden baserer seg på at fytoplanktonet sedimenterer i et tellekammer før identifisering og telling av celler.

Fordeler og begrensninger med metoden lysmikroskop

Fordeler med å bruke lysmikroskop som metode i overvåkingen er at dette er en relativt billig metode, så lenge man har lysmikroskop og taksonomisk kompetanse tilgjengelig. Et viktig moment er at lysmikroskopi er en metode som er brukt i lang tid og i mange land, slik at nye celledata kan settes inn i eksisterende langtidsserier som er viktig for å oppdage trender i protistsamfunnet. En mengde bestemmelseslitteratur finnes der arter er beskrevet basert på morfologiske karakterer for observasjon i lysmikroskop, også for arter observert i norske farvann (Thronsen mfl., 2007; Thronsen og Eikrem, 2001). Å identifisere arter i lysmikroskop krever tid samt vesentlig kyndighet og taksonomisk kompetanse. Genetisk variasjon i kryptiske arter og fenotypisk plastisitet er to store utfordringer. Selv for veldig synlig grupper, som *Pseudo-nitzschia*, er det eksempler på at individer som er morfologisk like i et lysmikroskop er genetisk forskjellige arter (Amato mfl., 2007). Det finnes også eksempler på

morfologisk distinkte individer av samme art, som til og med kan endre utseende etter tid på døgnet (Pizay mfl., 2009). En tredje utfordring er at bestemmelseslitteratur ofte beskriver kun et bestemt livsstadie til en art, slik at arter som befinner seg utenfor dette livsstadiet potensielt ikke blir identifisert (Hebert mfl., 2003).

Nøyaktig identifikasjon av arter kan altså ikke alltid gjøres basert på cellemorfologi observert i lysmikroskop alene og er mer kritisk for protister mindre enn 20 μm som mangler iøynefallende former (Massana, 2011). Størrelsen på arter som kan identifiseres i lysmikroskop er vanligvis begrenset til celler over ca. 15-20 μm og er avhengig av at cellene er godt preservert. Preservering i Lugols kan være fordelaktig da cellene sedimenterer raskere i tellekammeret og det farger cellene brune, som gjør de synligere og iblant lettere å telle. Lugols kan på den annen side også forårsake at cellene bli deformert og ugjenkjennelige. Dette gjelder spesielt "nakne" arter som mangler eller har tynne organisk skall (Gómez mfl., 2015; Hansen mfl., 1991). Det er også et lite vannvolum som telles, noe som kan ekskludere arter som har lav abundans.

For å overkomme begrensningene man møter på med bruk av lysmikroskop, har økologer lenge vært på utkikk etter en metode som er repliserbar og lett å gjennomføre for ikke-taksonomer. Metastrekkoding er foreslått som en løsning.

1.4 Metastrekkoding

DNA-strekkoding er et diagnostisk verktøy som benytter DNA-sekvenser til å identifisere arter (Gall & Saunders, 2010; Hebert mfl., 2003; Lahaye mfl., 2008; Lin mfl., 2009). En DNA-strekkode er en artsspesifikk DNA-sekvens i en valgt region av et gen (markør) i artens genom. Dette kan ses som analogt med strekkoden på en vare i en butikk. En DNA-sekvens på 15 baser kan ha 15^4 - over en milliard - forskjellige kombinasjoner av DNAets fire baser (Adenin, Cytosin, Guanin og Thymin). Det er mer enn 100 ganger det antallet kombinasjoner man trenger hvis hver enkelt levende art har sin unike strekkode med 15 baser (Hebert mfl., 2003). DNA-sekvenser som ikke blir konservert gjennom evolusjon er utsatt for mutasjoner, og artspesifikke sekvenser kan oppstå i de variable regionene av genomet.

En god markør (DNA-region) bør være lett å oppformere med PCR (Polymerase Chain Reaction), finnes hos et stort antall eukaryote-linjer og inneholde høy interspesifikk- men lav intraspesifikk variasjon. De beste markørene har vist seg å ha konserverte regioner som gjør det

mulig å designe primere samt en variabel region som er taksa-spesifikk: artens strekkode (Lin mfl., 2009).

Den tradisjonelt mest brukte markøren for å undersøke protister har vært hele eller deler av genet som koder for ribosomalt RNA i den lille ribosomale subenheten (SSU), **18S rRNA-genet** (Brown mfl., 2016; de Vargas mfl., 2015; Edwardsen mfl., 2000; Egge mfl., 2015a; Gran-Stadniczeñko mfl., 2017; Logares mfl., 2012; Massana mfl., 2015; Zimmermann mfl., 2011). Det er mange fordeler med å bruke 18S rRNA-genet (18S) som markør; 18S finnes i alle eukaryote celler, det finnes vanligvis mange kopier av 18S i genomet og primærstrukturen til 18S består av en mengde konserverte og variable regioner (figur 1.5). V4-regionen er en hyper-variabel region i 18S og har av CBOL-protist arbeidsgruppen (ProWG) blitt utnevnt til den beste markøren for å skille de store gruppene protister (Pawlowski mfl. 2012). En rekke andre markører for protister har også blitt foreslått som alternativer til 18S:

Det mitokondrielle genet for cytochrome c oxidase (**COI**, også kalt *cox1*) har blitt en veletablert markør for deteksjon av dyr (Hebert mfl., 2003; Vences mfl., 2005) og har vist seg å fungere godt på enkelte dinoflagellater (Lin mfl., 2009), diatoméer (Evans mfl., 2007), rødalger (Saunders, 2005) og brunalger (McDevit & Saunders, 2009).

Kloroplast-genet som koder for RUBISCO (**rbcL**) er en veletablert markør for landplanter som også kan brukes til å skille flere diatoméer til art (MacGillivray & Kaczmarska, 2011).

Ribosomalt **ITS** (internal transcribed spacer) som er markøren for sopp (Schoch mfl. 2012) kan fungere for noen dinoflagellater (Litaker mfl., 2007) og diatoméer (Moniz & Kaczmarska, 2010).

28S rRNA-genet som koder for ribosomalt RNA i den store ribosomale subenheten (LSU) fungerer godt som markør for noen ciliater (Gentekaki & Lynn, 2009), haptofytter (Liu mfl., 2009) og potensielt noen diatoméer (Trobajo mfl., 2010).

I dag finnes det flere DNA-databaser hvor DNA-sekvenser er lagret og tilegnet taksonomisk tilhørighet. NCBI (National Center for Biotechnology Information) (Edgar mfl, 2002) og EBI (European Bioinformatics Institute) er de største databasene, de kalles også genbanker og inneholder alle typer gener og genomer.

Databasene PR² (Guillou mfl., 2013), Silva (Quast mfl., 2013) og Dinoref (Mordret mfl., 2018) inneholder DNA-sekvenser av 18S rRNA-genet. Disse DNA-databasene gjør det mulig å taksonomisk klassifisere sekvenser av 18S rRNA-genet mer nøyaktig enn NCBI og EBI.

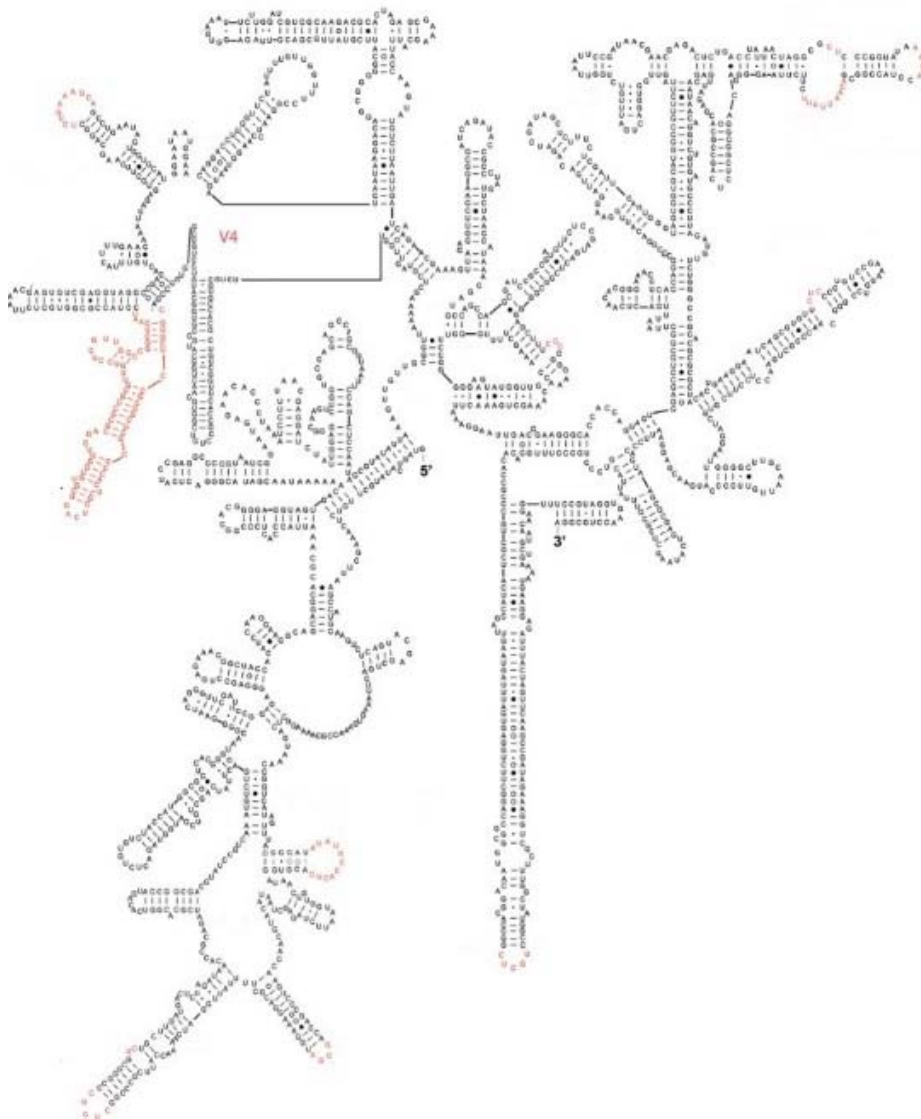
I økologiske diversitets studier har det den siste tiden blitt mer og mer populær å benytte seg av metastrekkoding (Brown mfl., 2016; de Vargas mfl., 2015; Massana mfl., 2015). Metastrekkoding er en metode som kombinerer DNA-strekkoding og " high-throughput" sekvensering (HTS) (figur 1.6). Fra en miljøprøve blir den valgte markøren oppformert fra DNAet til alle organismer som er tilstede i prøven, med PCR. Utviklingen av HTS-platformer har gjort det mulig å raskt sekvensere millioner av korte DNA-sekvenser samtidig (Reuter mfl., 2015). For øyeblikket er det Illumina sekvenserings-platformen (Illumina Inc., USA) som er den ledende aktøren på markedet og gir den beste sekvenseringsdybden per sekvensering (Reuter mfl., 2015). Med god sekvenseringsdybde menes at en stor del av sekvensene som er tilstede i prøven blir sekvensert (Sims mfl., 2014). Den gode sekvenseringsdybden gjør det mulig å få et detaljert bilde av protistsamfunnets komposisjon og i tillegg detektere sjeldne taksa. Illumina MiSeq sekvensering (forklart i seksjon 2.3.4) resulterer i sekvenser lange nok til å dekke markører som er vanlig å bruke i deteksjon av protister (Pawłowski mfl., 2012). Flere undersøker nå muligheten for å benytte metastrekkoding i miljøovervåkning (Ji mfl., 2013; Pawłowski mfl., 2016; Valentini mfl., 2016; Zaiko mfl., 2015).

Fordeler og begrensninger med metoden metastrekkoding

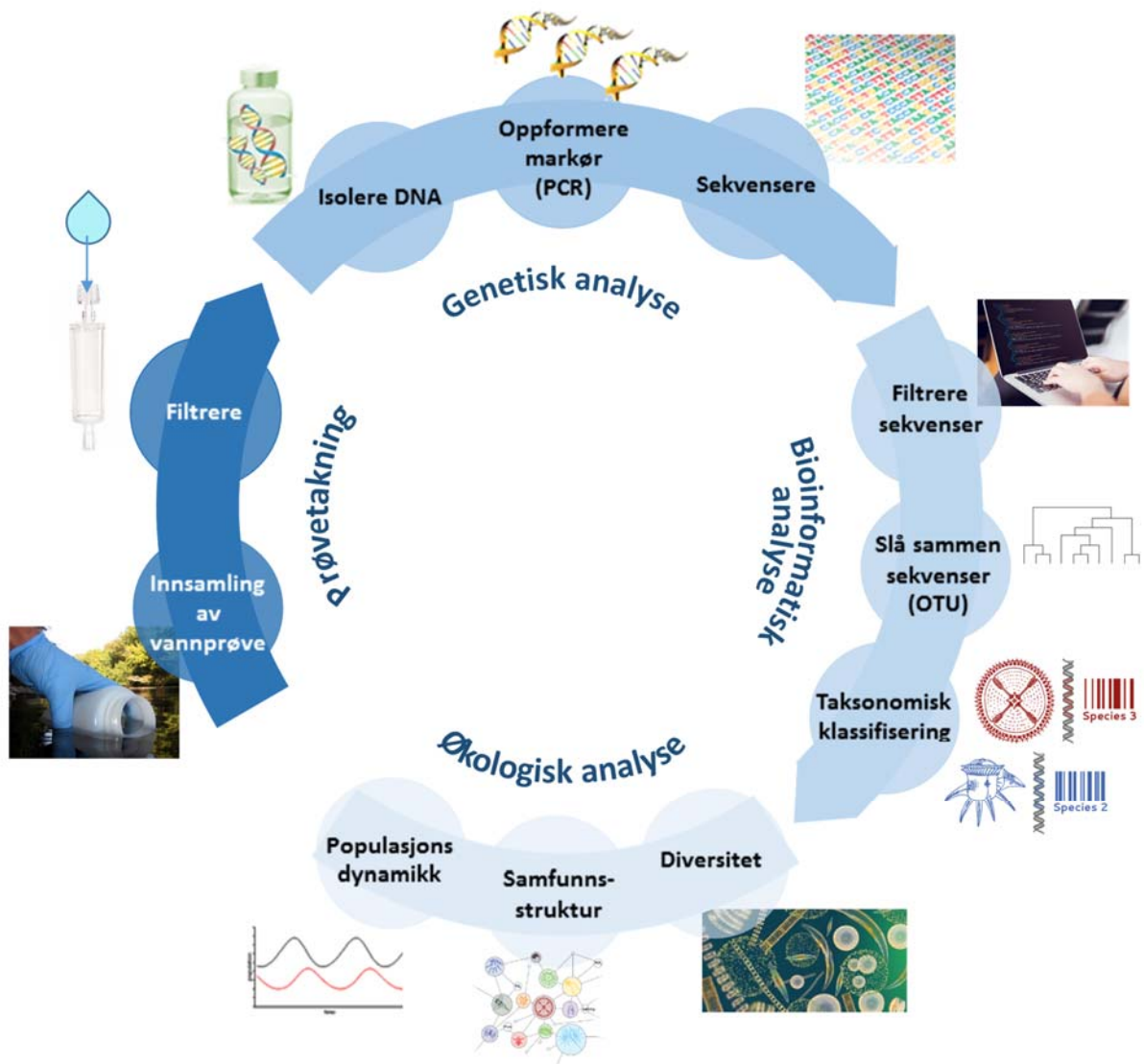
Mange av fordelene med metastrekkoding utfyller begrensningene møtt med den tradisjonelle metoden lysmikroskop. Metastrekkoding gir muligheten til en automatisert håndtering av miljøprøver som kan bli analysert med en standardisert protokoll. Dette øker muligheten til å sammenligne datasett fra forskjellige studier og laboratorier (Zimmermann mfl., 2015). Ved hjelp av denne metoden kan man identifisere taksa uavhengig av livsstadiet og størrelse samt uten å ha taksonomisk og morfologisk kunnskap om arten.

Den største begrensningen med metastrekkoding ligger i tilgangen på referansesekvenser. Metoden er avhengig av gode og oppdaterte referansedatabaser for identifisering av markøren. En annen begrensning er valg av markør og hvilken primer som blir brukt for å amplifisere den ønskede markøren. Å lage et primerpar som er universelt nok til å fange opp diversiteten i prøven uten å ha høyere affinitet for enkelt taksa (Taberlet mfl., 2012) er ønskelig.

I metastrekkoding oppformes markøren med PCR. Metoden er derfor kun semi-kvantitativ: Mengden sekvenser kan ikke overføres til antall celler som er i prøven. Forholdet mellom sekvenser i en prøve gir likevel et inntrykk av hva det er mye av i prøven og hva som er sjeldent.



Figur 1.5. Skjematisert fremstilling av sekundærstrukturen til 18S rRNA-genet, dvs. den lille subenheten av ribosomenes RNA. Den hyper-variable V4-regionen er markert øverst til venstre. Modifisert figur fra Yu, mfl., 2013.



Figur 1.6. Arbeidsflyten i metoden metastrekkoding. Modifisert fra Siano & Arnaud-Haond, 2016.

1.5 Formål med oppgaven

I denne studien ble protistplanktonsamfunnet i Indre Oslofjord ved stasjonen DK1 undersøkt gjennom en toårsperiode ved bruk av metastrekkoding. Da Norconsult i 2015 overtok overvåkingen av Indre Oslofjord, ble det opprettet et ekstra overvåkningsdyp, 0-2 meter, i tillegg til det tradisjonelle dypet på 5 meter for kvantifisering av fytoplankton. Formålet med denne oppgaven var å beskrive diversitet og sesongvariasjon ved overflaten (0-2 m) og 5 m dyp, utforske nytten av å bruke metastrekkoding i miljøovervåkingen og hvorvidt protistsamfunnet er forskjellig i de to dypene.

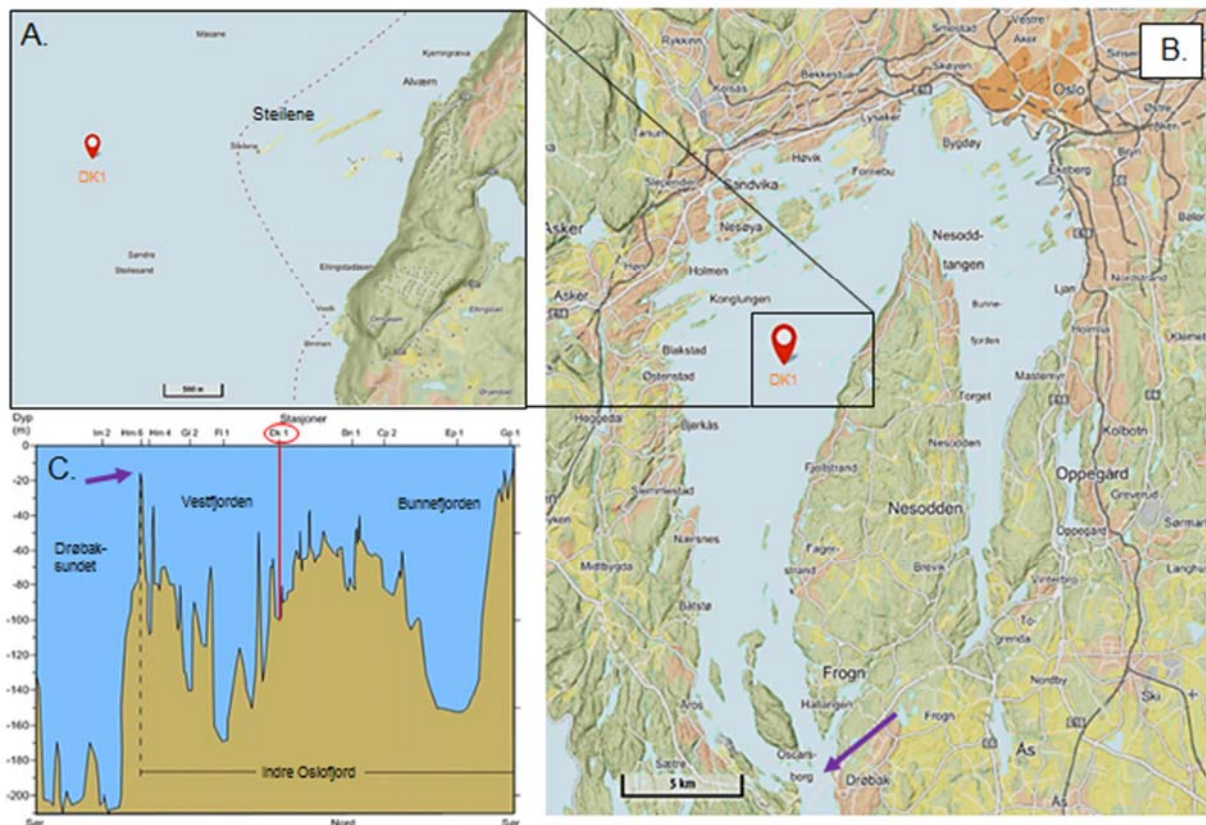
Spørsmål som ble stilt var:

1. Hvilke grupper protister i planktonet kan påvises i Indre Oslofjord ved bruk av metastrekkoding?
2. Hvordan varierer metastrekkoding-prøvene gjennom året?
3. Påvises det forskjellige dinoflagellat-taksa med metastrekkoding enn det kan observeres med lysmikroskop?
4. Er det signifikant forskjell i protistsamfunnet mellom de to dypene 0-2m og 5m og hvorfor?

2 Materialer og metoder

2.1 Lokalitet

Prøvetakningsstasjonen DK1 (59 48.900 N, 10 34.163 Ø) ligger omtrent midt i Indre Oslofjord, rett ut for Steilene i den delen av fjorden som heter Vestfjorden (figur 2.1.A og B). Allerede i 1916 var DK1 etablert som prøvetakningsstasjon, og den gang hadde DK1 nettopp navnet Steilene (Braarud & Nygaard, 1967). I dag er DK1 en av 25 vannprøvetakningsstasjoner i Fagrådets overvåkingsprogram av Indre Oslofjord (Norconsult, 2018). DK1 er omtrent 100 m dypt og omgitt av grunnere poller/terskler på hver side som skaper et undervannsbasseng (figur 2.1.C). Vannmassene i dette bassenget kommer hovedsakelig inn Oslofjorden fra Skagerak og må krysse den grunne Drøbakerskelen. I det øverste vannlaget finnes det gjennom året en varierende grad av ferskvannstilførsler fra Glomma og Drammenselva, sammen med nedbør og avrenning fra land.



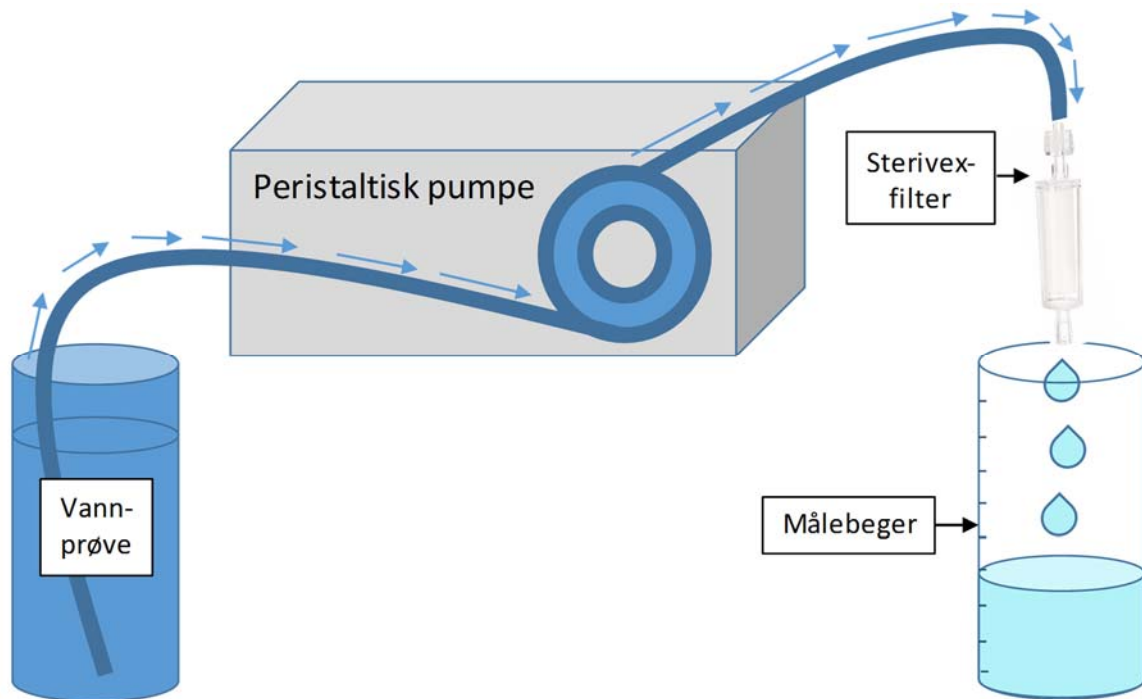
◀Figur 2.1. Prøvetakningsstasjonen DK1. (A) Stasjonen DK1 er markert i nærbilde, fremhevet fra (B) kart over Indre Oslofjord. Begge kartene er laget på Norgeskart.no. (C) Bunntopografien i Indre Oslofjord, dyp er angitt på y-aksen og øverste x-akse er navnene og plasseringen til prøvetakningsstasjoner i overvåkingen oppført. DK1 er uthevet med rødt. Bilde er hentet fra Norconsult (2018) og modifisert. (B og C) Drøbakterskelen er markert med lilla pil.

2.2 Innsamling

Det ble samlet inn 30 vannprøver fra to dyp, 0-2m og 5m, over ca. en toårsperiode (tabell S.1). Det ble tatt vannprøver tilnærmet annenhver måned i 2016. I 2017 hver måned med unntak av januar og juli. Hver innsamling ble gjennomført fra forskningsfartøyet F/F Trygve Braarud, samtidig som vannets ledningsevne (C, "conductivity", som er et mål på saltholdighet), temperatur (T), trykk (D, dybde), oksygenmetning og fluorescens ble målt med en CTD-måler (SBE 9plus CTD, Sea-Bird Scientific, USA).

Vannprøvene fra de to dypene ble samlet inn på hver sin måte. Overflatevannet (0-2 m) var en blandeprøve som ble tatt med en 2 m integrert vannhenter (KC Denmark A/S, Danmark). Vannhenteren er et rør som penetrerer vannsøylen ned til ønsket dybde. Ved at røret lukkes i bunn omslutter det vannsøylen og gjør det mulig å ta ut et 2 L utsnitt som kan hentes om bord. Vannet ble overført til en ren 10 L plastdunk, som etter gjentatte vannhenter ble full. Vannet fra 5 m ble samlet i Niskinflasker (2x5 L) festet til en rosett på CTD-måleren. CTD-målingene ble registrert fortløpende da måleren ble senket ned til ~100m dyp. På vei opp ble måleren stanset på 5m og vannprøvene ble samlet.

Vannet fra de to dypene ble filtrert på hvert sitt Sterivex-filer (Sterivex GV 0.22 µm, EMD Milipore Corporation), som fanger opp plankton større enn 0.22 µm. Dette ble gjort ved at vannet ble pumpet med en hastighet på 600-800 rtm, vha. en peristaltisk pumpe (ISMATEC, Tyskland), gjennom filteret, til et målebeger (figur 2.2). Målet var at det skulle filtreres 2 L vann på ett filter. De gangene det ikke ble filtrert 2 L var dette fordi filteret var tett (tabell S.1), som i vår-sommermånedene da det var mye plankton i vannet eller fordi det ikke ble registrert. Etter filtreringen ble filteret forseglet med heftemasse (Tack-it 50g, Faber-castell), merket og lagt i kjøleskap inntil det ble fraktet i kjølebagg til -80 °C fryser, hvor det ble oppbevart frem til DNA-ekstraksjon. Det ble laget to filter per prøve, et som ble brukt og et i reserve.



Figur 2.2. Skjematisk fremstilling av hvordan vannprøven pumpes ved hjelp av en peristaltisk pumpe gjennom et Sterivex-filter til et målebeger.

2.3 Molekylære metoder

2.3.1 DNA isolering og kvantifisering

DNA ble ekstrahert fra Sterivex-filteret og isolert ved bruk av DNeasy PowerWater Sterivex Kit (50) (Qiagen, Tyskland), etter protokoll fra produsent med noen modifikasjoner (appendiks 1). I Sterivex-filteret ble det først sprøytet inn ST1B-løsning som får cellene til å løsne fra filtermembranen. Filteret ble ristet i en Vortex-maskin (Vortex Genie 2, Scientific Industries, USA). Deretter ble det sprøytet inn bufferen MBL som ødelegger celleveggen og får cellene til å gå i oppløsning, cytolyse, ved inkubering på høy varme. Filteret ble igjen ristet i Vortex-maskinen. Etterpå ble den homogene løsningen trukket ut av filteret med en sprøyte og overført til et rør med små glasskuler. Siden celleveggen var ødelagt og cellemembranen allerede var punktert kunne de små kulene slå ut og frigjøre cellenes og kjernenes innhold av DNA ved at røret ble ristet i Vortex-maskinen. Etter sentrifugering ble kun supernatanten overført med pippette til et rent rør. For å rense supernatanten ble det tilsatt IRS. IRS er en løsning som binder seg til uorganiske- og organiske stoffer som ikke er DNA, slik som proteiner og humus-stoffer. Etter røret ble ristet, inkubert i kjøleskap og sentrifugert ble IRS-bunnfallet med avfallsstoffene

unngått og supernatanten med DNAet overført til et rør med en silika(silisiumdioksid)-membran. DNA binder seg til silika ved høye konsentrasjoner av salt. Saltløsningen MR ble tilsatt og DNAet i supernatanten bandt seg til silika-membranen. Cellerester, som membraner og proteiner, samt annet uorganisk materiale ble vasket bort med først etanol, som sentrifugeres gjennom membranen. Deretter sentrifugeres gjennom en PW-løsning som også er etanol-basert og som fjerner saltrester. Til slutt en siste sentrifugering og vask med etanol. Det renvaskede DNAet ble frigjort fra silika-membranen med EB-buffer (Elution Buffer 250 ml, Qiagen, Tyskland) og isolert. Ekstraksjon ble utført på 4-6 filtre av gangen, dette ble gjort for å unngå at DNA skulle degraderes siden hvert steg i protokollen tar lenger tid med økende antall filter.

Som en forberedelse til PCR ble konsentrasjonen (ng/μL) av det isolerte DNAet fra hver enkelt prøve kvantifisert med Qubit® 3.0 fluorometer (Thermo Fisher Scientific, USA). Arbeidsløsning ble blandet med Qubit® dsDNA HS Assay Kit, 500 assays (Invitrogen, Thermo Fisher Scientific, USA) etter protokoll fra produsent. I denne løsningen festes fluorokrom til dobbelttrådet DNA fra prøven. Fluorokrom som er bundet til DNAet vil emitte lys når det blir belyst med ultrafiolett lys (UV). Det emitterte lyset blir registrert og gir et mål på mengden DNA i prøven. Dersom resultatet av målingen viste "Too Low", altså ingen målbar DNA-konsentrasjon i prøven (<10 pg/μL), ble prøven kastet og DNA-ekstraksjon fra reservefilteret ble utført ved et senere tidspunkt. Viste derimot målingen en høyere DNA-konsentrasjon enn 5 ng/μL, ble prøven fortynnet med PCR-vann (Water, Mol Bio grade DNase-, RNase-, and protease-free, Tyskland) til konsentrasjonen 5 ng/μL.

Fortynningen ble utført etter likningen:

$$c1 * V1 = c2 * V2$$

Der **c1** = DNA konsentrasjon i prøven (ng/μL), **c2** = ønsket DNA konsentrasjon (5 ng/μL), **V1** = x (ukjent volum av prøven som skal tilsettes totalvolum) og **V2** = totalvolum (100 μL).

Som gir

$$V1 = \frac{5 \text{ ng}/\mu\text{L} * 100 \mu\text{L}}{c1 \text{ ng}/\mu\text{L}} = \frac{500}{c1} \mu\text{L}$$

2.3.2 PCR

I PCR-reaksjonsmiksen (tabell 2.1) ble det blandet KAPA, fremover- og revers-primer etter oppskriften under, der n er antall prøver:

Tabell 2.1. PCR-reaksjonsmikse. Tabellen viser navn (og produsent) på de tre ingrediensene i reaksjonsmiksen, hva de inneholder og etter hvilket forhold de skal blandes (basert på n antall prøver).

	Inneholder:	Mengde:
KAPA (KAPA HiFi HotStart Sekvensy Mix (X2), Kapabiosystems, USA)	DNA-polymerase, dNTP, MgCl ₂ og stabilisatorer.	12.5 µL * n
Fremover-primer: V4F_illumina	Oligonukleotider: (5' – TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA GCA SCY GCG GTA ATT CC– 3')	5 µL * n
Revers-primer: V4R_AZig_illumina	Oligonukleotider: (5' –GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GAC TTT CGT TCT TGA TYR ATG A– 3')	5 µL * n

KAPA inneholder enzymet DNA-polymerase, deoksiribose nukleotidene A, T, C og G (dNTP), magnesiumklorid og stabilisatorer. Primere velges etter hvilken del av DNAet man ønsker å oppformere i PCR-reaksjonen. Primerne som her ble valgt er universale for eukaryoter og binder seg til flankene av V4 regionen i den lille ribosomale subenheten (18S) (Piredda mfl., 2017). Primeren har festet på seg et overheng (figur 2.4) som senere inngår i å feste indekser til sekvensene (Seksjon 2.3.3).

Regionen V4 er i gjennomsnitt 380-410 basepar (bp) lang (Zimmermann mfl., 2011) og vanligvis artsspesifikk (Zhan mfl., 2013). DNA-polymerasen er avhengig av at primeren fester seg først til DNA-templatet, slik at den kan bruke den ledige 3'-enden på primeren til å starte syntesen.

Det ble pipettert 22.5 µL reaksjonsmikse, deretter 2.5 µL DNA-templat (fortynnet til 5 ng/µL), til hvert av de åtte rørene i PCR-stripsen, utenom det siste røret. Den var blank, kun 22.5 µL reaksjonsmikse og 2.5 µL PCR-vann, som kontroll på at ikke mastermiksen eller PCR-vannet var kontaminert. PCR-maskinen (Eppendorf AG, Tyskland) ble stilt inn på PCR-programmet som ble benyttet: først 3 min denaturering ved 95°C, fulgt av 30 sykluser med denaturering ved 98°C i 20 s, hybridisering av primer ved 62°C i 30 s og DNA-syntese ved 72°C i 1 min og 30 s. Avsluttende forlengelse ble gjort ved 72°C i 10 min. Før hybridiseringstemperaturen ble valgt,

ble det kjørt en test-PCR for å undersøke hvilken hybridiseringstemperatur som ga best resultat. Temperaturgradienten som ble testet var 62°C, 65.2°C og 67.9°C, hvor 62°C ga best resultat (appendiks 2).

Etter PCR ble agarosegelelektroforese benyttet for å undersøke at PCR-produktet inneholdt DNA-fragmenter med ønsket lengde (~380 bp) og at den blanke kontrollen var uten DNA. Det ble laget en 0.8% agarosegelé ved å blande 0.8 g agarosepulver (Lonza, USA) med 100 mL TAE x1-buffer (appendiks 3). Løsningen ble varmet i mikrobølgeovn og etter kort avkjøling ble det tilsatt 4 µL GelRed (Biotium, USA). GelRed binder seg til DNA og fluorescerer i UV-lys og derfor gjør det mulig å visualisere DNA-fragmentenes plassering på gelen ved å belyse den med UV. Den homogene løsningen ble overført til et elektroforesekammer med kam, som lagde brønnene i gelen. Da gelen hadde stivnet ble 1 µL "loadingbuffer" (Termo Scientific, USA) blandet med 5 µL PCR-produkt applisert i hver sin brønn. Første og siste brønn var forbeholdt 6 µL størrelsesmarkør (FastRuler™ Low Range DNA Ladder, Termo Scientific, USA), som gjorde det mulig å kontrollere at DNA-fragmentene hadde den ønskede lengden. Gelen ble kjørt i 40 min, ved 80V og 400mA (Electrophoresis Power Supply – EPS 301, Bio-Sciences AB, Sweden). Gelen ble observert og fotografert i et UV-kabinett (Gene Genius, Bio Imagine System, Syngene, England).

Kimærer

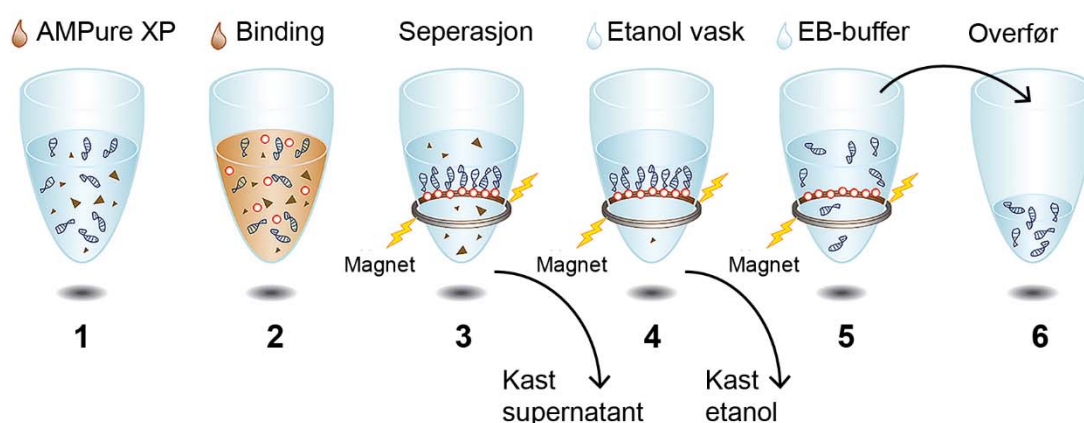
Kimærer oppstår når to enkelttrådet DNA fra hver sin organisme ved en feil/tilfeldighet smelter sammen under PCR. Dette kan skje ved at konserverte regioner i rDNAet fungerer som primer; dermed kan det dannes et hybrid av tråder fra to forskjellige organismer som bruker hverandre som primer og templat (Berney mfl., 2004; Hugenholtz, 2003). For å redusere amplifiseringsfeil ble hver prøve oppformert i tre separate PCR-reaksjoner (Acinas mfl., 2005; Kanagawa, 2003). De tre PCR-produktene ble etter gel-kjøring og kontroll blandet til ett PCR-produkt, som heretter omtales som et amplikon.

2.3.3 Forberedelse til Illumina sekvensering

Amplikonet ble forberedt til Illumina Miseq-sekvensering ifølge produsentens protokoll "16S Metagenomic Sequencing Library Preparation Preparing *16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System*".

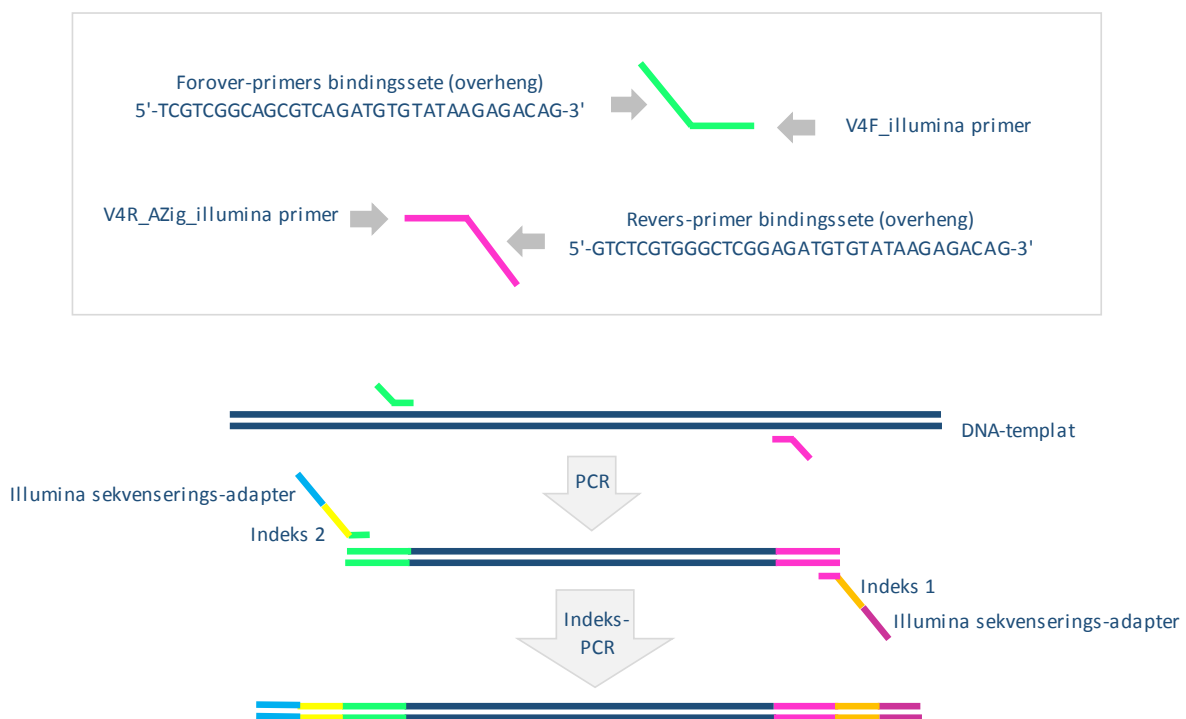
https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf

Denne protokollen har flere trinn, sidetall blir oppgitt. Etter PCR ble amplikonet vasket (s.8-9). Dette ble gjort med AMPure-kuler (Agencourt® AMPure® XP, Beckman Coulter Inc., USA) og 80% etanol (figur 2.3). AMPure-kulene har den egenskapen at de er paramagnetiske, det vil si at de kun er magnetiske i et magnetiskfelt. Kulene vil ikke tiltrekke hverandre og danne klumper som blir bunnfall i løsningen, men isteden binde det negativt ladete DNAet. Overflaten på kulen er dekket med karboksyl-molekyler som reversibelt binder DNAet ved tilstedeværelsen av PEG (polyetylen glykol) og salt (natriumklorid) (DeAngelis mfl., 1995). Etter at DNAet bandt seg til AMPure-kulene, ble rørene plassert i et magnetstativ (DynaMag™-96 Side Skirted, Life Technologies AS Oslo, England) som aktiverte magnetismen i kulene, som da sammen med DNAet ble trukket mot vegg i røret der magneten stod. Det ble derfor mulig å fjerne den DNA-fri supernatanten og vaske med etanol. Det ble tilsatt EB-buffer som reverserte bindingen av DNAet til AMPure-kulene, og supernatanten som da inneholder DNAet ble overført til et rent rør.



Figur 2.3. Skjematisk oversikt over stegene i vasking av PCR-produkt med AMPure-kuler. Figuren er hentet fra Beckman Coulter, Inc. SPRIselect User Guide © 2012 og modifisert.

Neste trinn i protokollen er Indeks-PCR (s.10-12). Det ble ved hjelp av PCR festet tre oligonukleotider til DNA-templatet i ampliconet (figur 2.4). Oligonukleotidene var primer-bindingssteder, indekser og Illumina Sekvenserings-adaptere (Nextera XT Index Kit, Illumina, USA). Indeksene er spesifikke for hver enkelt prøve og gjør det mulig å gjenkjenne hvilke sekvenser som tilhører hvilken prøve etter sekvenseringen. Illumina Sekvenserings-adapterne er nødvendig for at sekvensene skal feste seg til "flow"-cellen under sekvenseringen, som vil bli forklart i seksjon 2.3.4. Illuminasekvensering. Alle ampliconene samlet omtales heretter som et amplicon-bibliotek.



Figur 2.4. Figuren illustrer hvordan primere med overheng gjør det mulig å feste indekser (indeks 1: orange, indeks 2: gul) og Illumina sekvenserings-adaptere (i hver ende, blå og lilla) vha. PCR til DNA-templatet i ampliconet.

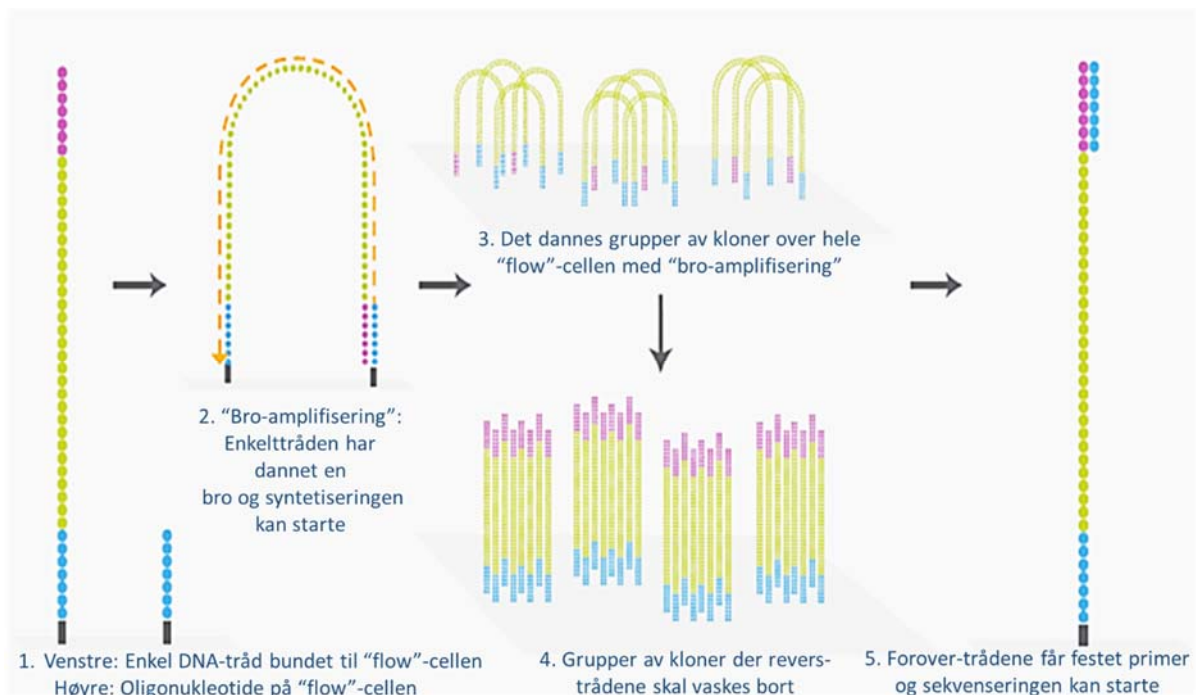
Etter Indeks-PCR ble amplicon-biblioteket igjen vasket med AMPure-kuler og etanol (s.13-14), deretter ble DNA fluorometrisk kvantifisert med Qubit. Arbeidsløsning ble laget med Qubit® dsDNA LS Assay Kit, 500 assays (Invitrogen, Thermo Fisher Scientific, USA) etter protokoll fra produsent.

Som en siste sjekk på at DNA-fragmentene i biblioteket er like lange (~600bp=380 bp + primer-bindingssete + indeks + adapter), ble det kjørt på en 0.8% agarosegel. Alle prøvene i biblioteket ble fortynnet med EB-buffert til lik konsentrasjon (10 nM), etter beskrivelse på s.16 i protokollen

og til slutt blandet til én blandeprøve ved at det ble pipettert 5µL av hver prøve over til et nytt rent rør. Blandeprøven gjennomgikk en siste kvalitetssjekk der prøvens renhet og mengde DNA (ng/µL) ble målt med Nanodrop ND-1000 spektrometer (Thermo Scientific, USA). Renheten til prøven ble gitt etter forholdet mellom absorpsjon av bølgelengdene 260nm og 280nm. DNA absorberer lys ved bølgelengden 260nm. En 260/280 ratio på 1.8-2.1 bli ansett som en ren prøve, men er ratioen derimot lavere kan det indikere at prøven inneholder proteiner, fenoler eller andre forurensninger som absorberer lys ved bølgelengden 280nm. Til slutt ble prøven kjørt på 0.8% agarosegel for å kontrollere at alle DNA-fragmentene var like lange. Disse resultatene ble sendt med prøven til Norsk sekvenseringscenter ved Universitetet i Oslo.

2.3.4 Illuminasekvensering

Oppformering av kloner med "bro-amplifisering"

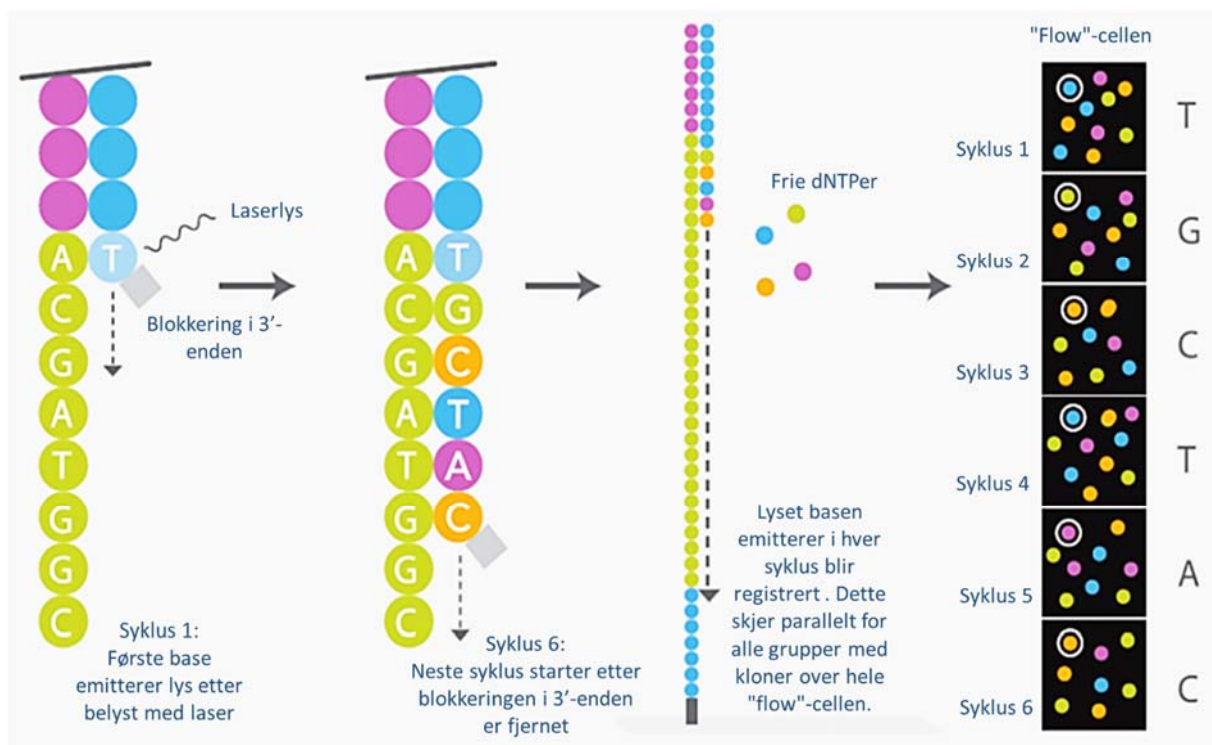


Figur 2.5. Figuren viser trinnvis hvordan det blir dannet kloner med "bro-amplifisering". Figuren er modifisert © Illumina, Inc.

Illumina MiSeq System sekvenserer ved syntese. Det vil si at baserekkefølgen blir registrert etterhvert som DNA-tråden syntetiseres. Først blir prøven denaturert med natriumklorid (NaOH). NaOH bryter ikke bare hydrogenbindingene i DNAets dobbelhelix, men forhindrer også enkeltrådene i å danne sekundærstrukturer (Wang mfl., 2014). Prøven blir overført til "flow"-cellen som er en glassplate behåret med to ulike typer oligonukleotider. Disse er

komplementære til de to Illumina sekvenserings-adapterne som tidligere har blitt festet til hver ende på amplicon-prøvene. Hybridisering skjer ved at de denaturerte fragmentene i prøven fester seg til den komplementære oligonukleotiden. Et polymerase-enzym danner en komplementær DNA-tråd til det hybridiserte fragmentet, det dobbeltrådede molekylet blir deretter denaturert og det originale templatet vaskes bort. Det blir dannet kloner av tråden ved "bro-amplifisering" (figur 2.5). I denne prosessen bøyer den fastsittende tråden seg over og binder seg til det andre oligonukleotiden, slik at tråden blir stående i en broform. En polymerase bruker tråden som templat slik at broen blir dobbeltrådet. Broen blir denaturert, som resulterer i to enkeltrådet kopier av molekylet som sitter fast i hver sin oligonukleotidet i "flow"-cellen. Dette blir gjentatt igjen og igjen, og skjer samtidig for millioner av DNA-tråder over hele "flow"-cellen, som resulterer i grupper av tilnærmet 1000 kloner for alle sekvensfragmentene i prøven. Etter klone-amplifisering blir revers-trådene vasket bort og kun forover-trådene står igjen på "flow-cellen". 3'-endene blir blokkert for å forhindre uønsket priming og endelig kan selve sekvenseringen starte.

Sekvensering



Figur 2.6. Figuren viser sekvensering ved syntese, med syntetiseringen (venstre) og hvordan dette ser ut på "flow"-cellen (helt til høyre). Figuren er modifisert © Illumina, Inc.

Sekvenseringen starter med at forover-primer blir bundet til sitt bindingssetet på forover-tråden (figur 2.6). Teknologien "Sekvensering ved syntese" bruker dNTPer som er merket med hvert sitt fluorokrom samt reversibelt blokkert i 3'-enden. I hver syklus konkurrerer nukleotider om den ledige plassen i den voksende tråden. Nukleotidene blir belyst med en grønn- (for G og T) og rød- (for A og C) laser som får fluorokromet til å emittere lys. Lysets bølgelengde og styrke blir gjennom flere filtre tolket og det nukleotidet resultatet tilsvarer blir registrert (figur 2.6). Hvert registrert nukleotid i sekvensrekkefølgen får en tilhørende kvalitetsverdi, som sier noe om hvor korrekt denne avlesningen er kalkulert til å være. Overflødige nukleotider, fluorokromet samt blokkeringen i 3'-enden blir kjemisk fjernet og neste syklus kan starte. Sekvenseringen stanser etter et bestemt antall sykluser, som da tilsvarer lengden på den første sekvenserte tråden. Den sekvenserte tråden blir vasket bort, indeks 1 blir lest av, forover-tråden danner igjen en bro og indeks 2 blir lest av. Som ved "bro-amplifiseringen" blir revers-tråden syntetisert og alle forover-trådene vaskes bort. På samme måte som forover-tråden blir revers-tråden sekvensert. Dette kalles "paired-end"-sekvensering og resulterer i to overlappende sekvenser.

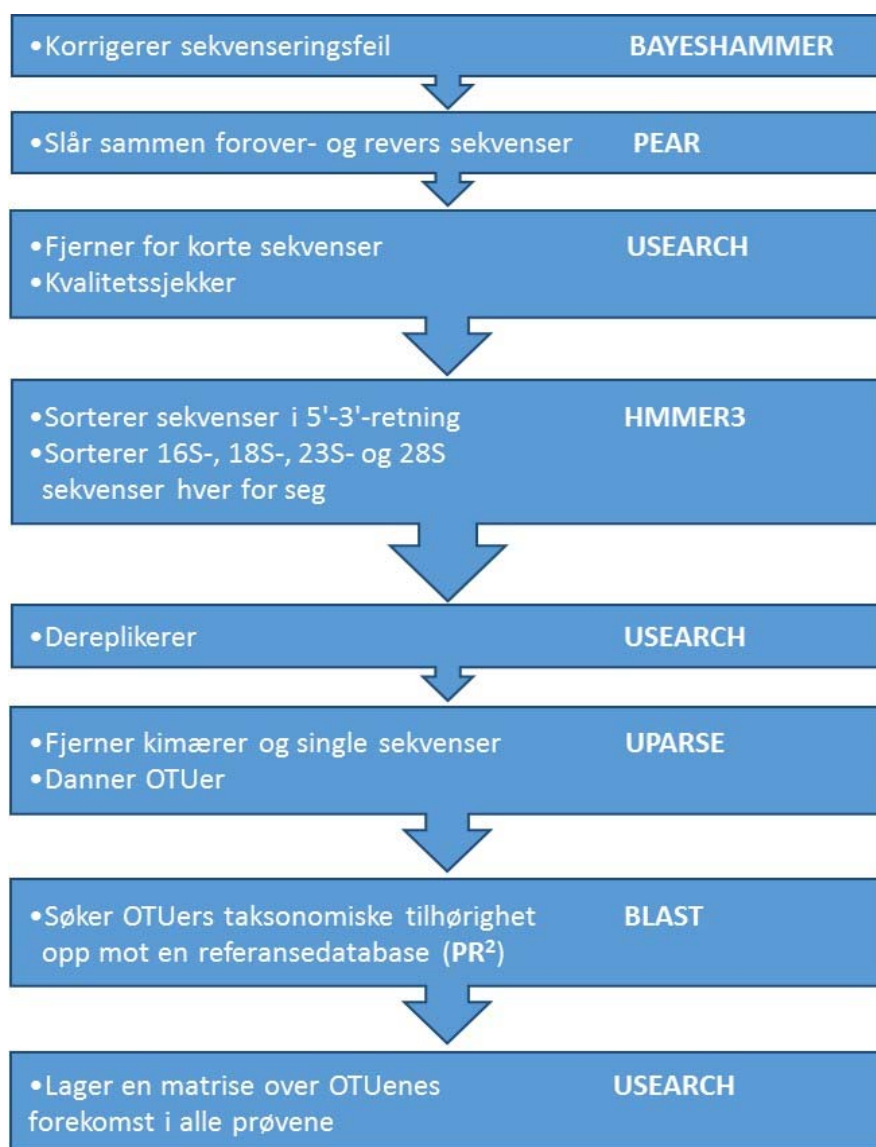
Sekvenseringsfeil

Den vanligste feilen i Illumina sekvensering er substitusjonliknende registreringer av nukleotidet. Dette oppstår på grunn av sterk korrelasjon mellom emisjonsspekteret til A og C, og T og G, og begrensninger i filtrene som brukes til å tolke/skille signalet. Et annet problem oppstår når ikke alle klonene i en gruppe emitterer lys i lik bølgelengde, kjent som faseforskyvning og pre-faseforskyvning. Faseforskyvning skjer når blokkeringen i 3'-enden på det sist syntetiserte nukleotidet ikke er blitt tilstrekkelig fjernet. Dette kan få noen av klonene i en gruppe til henge etter de andre i syntetiseringen. Pre-faseforsyning på den annen side skjer når ikke alle nukleotidene er tilstrekkelig vasket bort i slutten av en syklus og 3'-enden på disse blir fjernet. Dette gir noen kloner muligheten til å ligge litt foran de andre i syntetiseringen fordi de syntetiserer flere nukleotider i en syklus enn hva som blir registrert. Begge faseforskyvningene resulterer i følgefeilforstyrrelser i lyssignalet sendt ut fra den gitte gruppen kloner (Kircher mfl., 2009; Schirmer mfl., 2015).

2.4 Analyse av sekvensdata

Primer-bindingssteder, indekser og adaptore ble fjernet fra råsekvensene av Norsk sekvenseringscenter og råsekvensenes lengde var ~250 baser.

2.4.1 Bioinformatikk

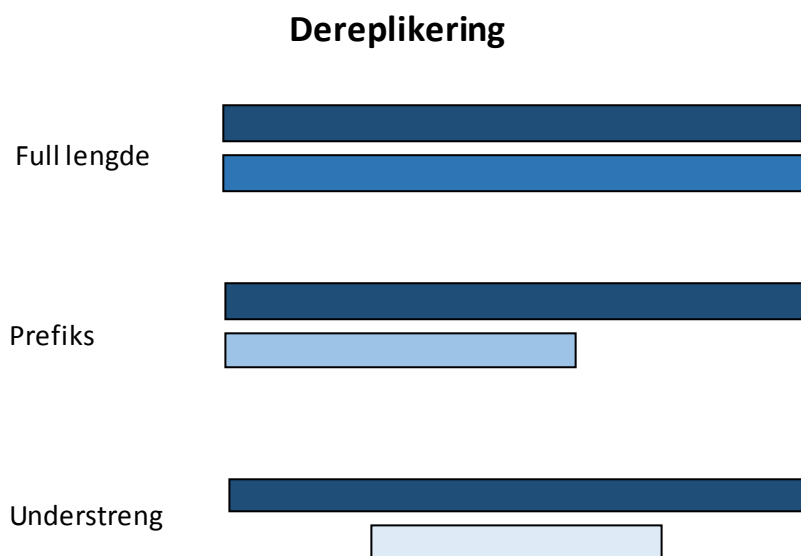


Figur 2.7. Arbeidsflyten i den bioinformatiske-pipelinen som ble brukt til å behandle råsekvensene og beskriver jobben de forskjellige programmer/algoritmer har utført.

På datasettet ble det brukt en pipeline designet for filtrering og klassifisering av Illumina MiSeq råsekvenser (Logares, 2017) (appendiks 6, figur 2.7). Alle analyser i pipelinen ble gjort på superdatamaskinen Abel ved Universitet i Oslo (<https://www.uio.no/english/services/it/research/hpc/abel/>). I korte trekk ble fastq råsekvensene

først korrigerert for feil med BAYESHAMMER (Nikolenko mfl., 2013). BAYESHAMMER bruker i førsteomgang grupperingsalgoritmen i HAMMER til å gruppere de tilnærmet identiske råsekvensene, for deretter å redefinere gruppene ved å sub-gruppere kontiger (overlappende fragmenter av sekvensen som til sammen representerer hele sekvensen) etter kvalitetsverdiene gitt av Illumina. For hver tilordnede sub-gruppe gis "Bayesianske-straffeverdier" (BIC) basert på Bayesiansk statistikk (dette blir ikke forklart nærmere siden det går utover rammene for denne oppgaven). Til sammen inngår kvalitetsverdiene og straffeverdiene i en ny algoritme som korrigerer de antatte sekvenseringsfeilene i kontigene. BAYESHAMMER setter sammen kontigene til hele sekvenser og gir nye kvalitetsverdier til basene i sekvensene. De sekvensene som da har for dårlige verdier fjernes, siden det indikerer mange sekvenseringsfeil. "Pair-end"-sekvenseringen genererte to sekvenser for hvert DNA-templat, en for hver av primerne (forover og revers). Det komplementære overlappet i disse sekvensene ble parett og sekvensene slått sammen med PEAR (Zhang mfl., 2014). Hvis det var baser i overlappet som ikke var komplementære, vurderte PEAR de nye kvalitetsverdiene til baseparet opp mot hverandre og beholdt den basen med best verdi. Basen som ble beholdt fikk korrigerert ned sin kvalitetsverdi. Med USEARCH ble sekvenser som etter sammenslåingen var kortere enn 100 baser fjernet. Grensen ble satt til 100 for å unngå type 1 feil, altså at den fjerner reelle sekvenser, men den blir ikke satt lavere da dette vil øke sjansen for type 2 feil, at falske sekvenser beholdes. De gjenværende sekvensene gjennomgikk en kvalitetskjekk, med en maksimumsforventningsfeil = 1, der sekvenser med en kvalitetsverdi som tilsvarer over 1% feil-rate ble filtrert bort (Edgar & Flyvbjerg, 2015).

Med metoden vi har brukt, er det mer enn 18S sekvenser som kan bli amplifisert og sekvensert. Det er derfor viktig å identifisere 18S sekvensene fra alt annet. For å kunne sortere ut 18S sekvenser ble verktøyet HMMER3 brukt. HMMER3 er et verktøy som søker i sekvensdatabaser eller forhåndlagede hmmer-profiler etter homologe sekvenser (Eddy, 1998; <http://hmmer.org/>). I pipelinen er det forhåndsdefinerte profiler for den lille- og den store ribosomale subenheten for både prokaryoter (henholdsvis 16S og 23S) og eukaryoter (18S og 28S). De sekvenser som ikke var 18S ble fjernet, og de resterende sekvensene ble sortert i samme retning (5'-3'). Sekvensene ble deretter dereplisert med USEARCH (Edgar, 2010). Dereplisering betyr fjerning av duplikater, altså sekvenser som er like. Dette gjøres for å minimere størrelsen på dataene man jobber med i pipelinen og derfor sparer mye tid på de jobbene som kommer etter.



Figur 2.8. Illustrasjon av de tre typene duplikater som kan ha oppstått til en sekvens. "Full lengde" er et duplikat som er identisk sekvenseni lengde. "Prefiks" er et duplikat i starten (eller slutten) av sekvensen. "Understreng" er et duplikat innad i sekvensen.

Det er tre typer duplikater (figur 2.8): "Full lengde" duplikatet er det ønskede produktet etter PCR og er identiske sekvenser, også i antall baser, som navnet tilsier. "Prefiks" er et duplikat i starten (eller slutten) av sekvensen, men har en kortere lengde. "Understreng" er et duplikat innad i sekvensen, som er kortere enn sekvensen fordi starten og slutten av sekvensen

mangler. Den representative OTU-sekvensen man står igjen med etter dereplikering inneholder informasjon om hvor mange duplikater den representerer og til hvilken prøve duplikatet tilhørte.

Som tidligere nevnt er mange sekvenser tilnærmet like. Disse ønsket vi å slå sammen til det som kalles en OTU. OTU står for "Operational Taxonomic Unit" og er en pragmatisk erstatning for art, som her ble definert til sekvenser med en 98% likhet (Ribeiro mfl., 2018). Dette ble gjort med UPARSE V8.1.1756 som parallelt fjernet kimærer og OTUer som bare bestod av en enkelt sekvens (Edgar, 2013). En OTU som består bare av en sekvens, også kalt en singel sekvens, er en sekvens som kun har forkommet i én enkelt prøve én gang. Etter å ha tatt prøver over nesten to år som hver har gjennomgått tre PCRer og vi har fått 9 863 086 sekvenser, er sannsynligheten liten for at vi vil stå igjen med én singel sekvens fra en organisme. Disse ble derfor vurdert som PCR- eller sekvenseringsfeil og fjernet. OTUenes taksonomiske tilhørighet ble bestemt ved å bruke programmet BLAST (Altschul mfl., 1997). BLAST tar den representative sekvensen for hver OTU og generer et søk opp mot en valgt referansedatabase, her PR² (Guillou mfl., 2013). PR² (The Protist Ribosomal Reference database) er en kuratert database for RNA og DNA til 18S hos eukaryote. Avsluttende ble det med USEARCH generert en matrise over OTUenes forekomst i hver prøve. Sekvensene blir brukt som estimat for abundans, OTUenes abundans estimeres for hver enkelt prøve ved å telle hvor mange sekvenser

(reads på engelsk) som finnes i Illumina-biblioteket som korresponderer til den representative OTU-sekvensen i den gitte prøven.

2.4.2 Statistiske analyser og fremstilling av data

Alle analyser, databehandling og fremstilling ble enten utført i Excel eller i R versjon 3.5.0 (R Core Team, 2018) ved bruk av pakkene ggplot2 (Wickham mfl., 2018) og Vegan versjon 2.5-3 (Oksanen mfl., 2018). R-skript til alle diagrammer og analyser er i appendiks 7. Protistsamfunnets totale sammensetning ved stasjonen DK1 i toårsperioden, basert på totalt antall OTUer og sekvenser, ble fremstilt med treemap-pakken (Tennekes & Ellis, 2017). For å kunne sammenligne samfunnet i prøvene ble datasettet ble subsamplet med funksjonen "rrarefy" etter den prøven med lavest antall sekvenser (10 530). Dette gjør prøvene like store og sammenlignbare.

Miljøvariabler

For å fremstille miljøvariablene ved 0-2 m og 5 m, ble målinger hentet ut fra CTD rå-datafilene ved hjelp av R (appendiks 7).

For å undersøke om hver enkelt miljøvariabel samvarierer mellom de to dypene ble det benyttet lineære regresjon. Viser miljøvariabelen en lineær sammenheng mellom de to dypene, kan man ved å undersøke likningen til regresjonslinjen si noe om ulikheter og avvik fra samvariasjon. En regresjonslinje beskrives med likningen $y = ax + b$, der a er stigningstallet til linjen og b er skjæringspunktet med x-aksen. Overflatemålingene (0-2 meter) ble plottet på y-aksen og 5 meter ble plottet x-aksen. Er stigningstallet a , tilnærmet lik 1 er det stor samvariasjon mellom de to dypene.

Det øvre blandingslaget ble bestemt ved å fremstille tetthets-profiler med Oce-pakken (Kelley & Richards, 2018) og lese av dypet der sprangsjiktet startet.

Fylogenetisk analyse

For å verifisere den taksonomiske klassifiseringen tildelt av PR² og for å undersøke plasseringen til de mest abundante dinoflagellat-OTUene, ble det konstruert et fylogenetisk tre. Det ble valgt ut 17 OTUer i gruppen Dinoflagellata. I disse 17 OTUene inngår de mest abundante og frekvente OTUene fra både protistdatasettet og det subsamplede datasettet. For hver OTU ble det manuelt gjort to Blast-søk til genbanken NCBI

(<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), ett som inkluderte miljøsekvenser og ett som ekskluderte miljøsekvenser. Ett til to av de beste treffene ble plukket ut fra hvert søk. Sammen ble disse brukt som en del av referansesekvensene i det fylogenetiske treet (tabell S.5). Resten av referansesekvensene tilhørte typiske arter innen de forskjellige ordener i Dinoflagellata, samt seks sekvenser som utgjorde utgruppen Apicomplexa, alle hentet fra PR². Sekvensene ble alignet med L-ins-algoritmen i MAFFT v7.300 (Kato & Standley, 2013) og deretter trimmet med trimAL etter standardinstillinger (Capella-Gutiérrez mfl., 2009). Det ferdige alignmentet inneholdt 196 taksa og 1 908 karakterer. Maksimum-likhets-analyse ble utført med programmet RAxML v8.0.26, ved å bruke fortløpende bootstrapping og ML-søk med GTRGAMMA-modellen (Stamatakis, 2014). Antallet bootstrapper ble automatisk satt basert på MRE bootstopp kriterier, og stoppet etter 352 replikater.

Alignment ble visuelt kontrollert i MAFFT (<https://mafft.cbrc.jp/alignment/software/>).

Ordinasjon

For å undersøke hvordan prøvene fra de to dypene var relatert til hverandre ble det benyttet ordinasjon. Som foreslått av van Son og Halvorsen (2014) ble det fortatt en multippel parallell ordinasjon (MPO), som inkluderte DCA (detrending correspondance analysis) (Hill & Gauch, 1980) og GNMDS (global-non-metric multidimensional scaling) (Kruskal, 1964). For å sammenligne ordinasjonsaksene fra DCA og GNMDS ble det brukt Kendall's tau korrelasjonskoeffesient, τ . DCA ble kjørt med standardinnstillinger for decorana-funksjon (Hill, 1979), 4 reskaleringssykluser med 26 segmenter i hver syklus. DCA-aksene er skalert i S.D enheter. En to-dimensjonal GNMDS løsning ble laget, der Bray Curtis-ulikheter ble sammenlignet med distansematrisen. Der avstanden var så stor at den ikke var til å stole på (>0.8) ble nye Geodesic avstander beregnet. GNMDS-aksene er skalert i "Half Change"-enheter. Miljøvariabler ble skalert med "env.fit"-funksjonen og fremstilt sammen med ordinasjonen i et biplot.

For å teste om prøvenes plassering på ordinasjonsaksene var lik for de to dypene, ble det benyttet Welch tosidig t-test med funksjonen "t.test" og likhetsanalyse (Analysis of Similarity – ANOSIM) for Bray-Curtis verdier ble utført med funksjonen "ANOSIM" i Vegan-pakken.

Rikhet, diversitet og jevnhet

Indekser for rikhet (antallet OTUer i hver prøver), Shannon diversitet, H' (Shannon, 1948) og jevnhet, J' (Pielou, 1966) ble bestemt med "diversity"-funksjon i Vegan-pakken, og ble brukt til å undersøke sesongvariasjon i protistsamfunnet ved de to dypene. Funksjonen "gam" i mgcv-pakken (Wood, 2018) tilpasset en ikke-parametrisk generalisert additiv modell (GAM) til indeksdataene, slik at trender gjennom prøvetakningsperioden ble fremhevet. Det ble benyttet Welch tosidig t-test for å teste om indeksene var forskjellig ved de to dypene.

Sammenligning med lysmikroskop data

For å undersøke om metoden metastrekkoding påviser de samme dinoflagellat-taksa observert med den tradisjonelle metoden lysmikroskop ble det sammenlignet tilstedeværelsen av dinoflagellat-taksa registrert med de to metodene. Metastrekkoding-dataene er resultatene i denne studien fra de 11 prøvene som ble samlet på 5 m dyp i 2017. Lysmikroskopi-dataene er basert på 18 prøver samlet med FerryBox systemet gjennom 2017. Lysmikroskop-prøvene er tatt ved samme prøvetakningsstasjon (DK1) og dyp som metastrekkoding-dataene. Prøvene ble tatt i lysmikroskop av Wenche Eikrem på NIVA (Norli & Eikrem, 2018). Étt taksa ble definert til navnet gitt som nærmeste taksonomiske klassifisering i de to metodene. I metastrekkoding-dataene ble OTUene taksonomisk klassifisert mot referansedatabasen PR² og duplikate taksa-navn ble fjernet. Ubestemte dinoflagellater innen en størrelsesorden (μm) eller innen en klade regnet som étt taksa.

Senere ble det kun sammenlignet taksa som ble klassifisert til art med de to metodene. Det ble antatt at alle arter identifisert i lysmikroskop er 100% korrekt klassifisert. Alle taksa som ikke ble klassifisert til art med stor sikkerhet (<99% identitet med referansesekvens i PR²) ble fjernet. Tilslutt ble taksalistene for de to metodene kontrollert for synonymer ved å søke opp hver enkelt art i ALGAEBASE (<http://algaebase.org>).

Vind

På Landbruksmeteorologisk tjeneste sine nettsider (<http://lmt.bioforsk.no/>) ble det hentet ut data på gjennomsnittlig vindhastigheten, for hver dag fra februar 2016 til desember 2017 (prøvetakningsperioden) ved meteorologiskmålestasjonen på ÅS. Den meteorologiske målestasjonen på Ås er den nærmeste målestasjonen til vår prøvetakningslokalitet (DK1) i Indre Oslofjord.

3 Resultater

3.1 Metastrekkoding-data

For hver enkelt av prøvene ble det kontrollert at den inneholdt DNA-fragmenter med ønsket lengde, ~600 bp (appendiks 4), deretter ble de blandet til én blandeprøve. Før blandeprøven ble sendt til Norsk sekvenseringssenter for Illumina MiSeq sekvensering ble konsentrasjon og renhet sjekket (appendiks 5). Blandeprøvens renhet ble undersøkt med Nanodrop hvor 260/280 nm ratioen var 1.97, og dermed innenfor kravet på 1.8-2.1. Mengden DNA ble målt til 2.7 ng/ μ L, som tilsvarer 6.82 nM, noe lavere enn den ønskede mengden på 10 nM.

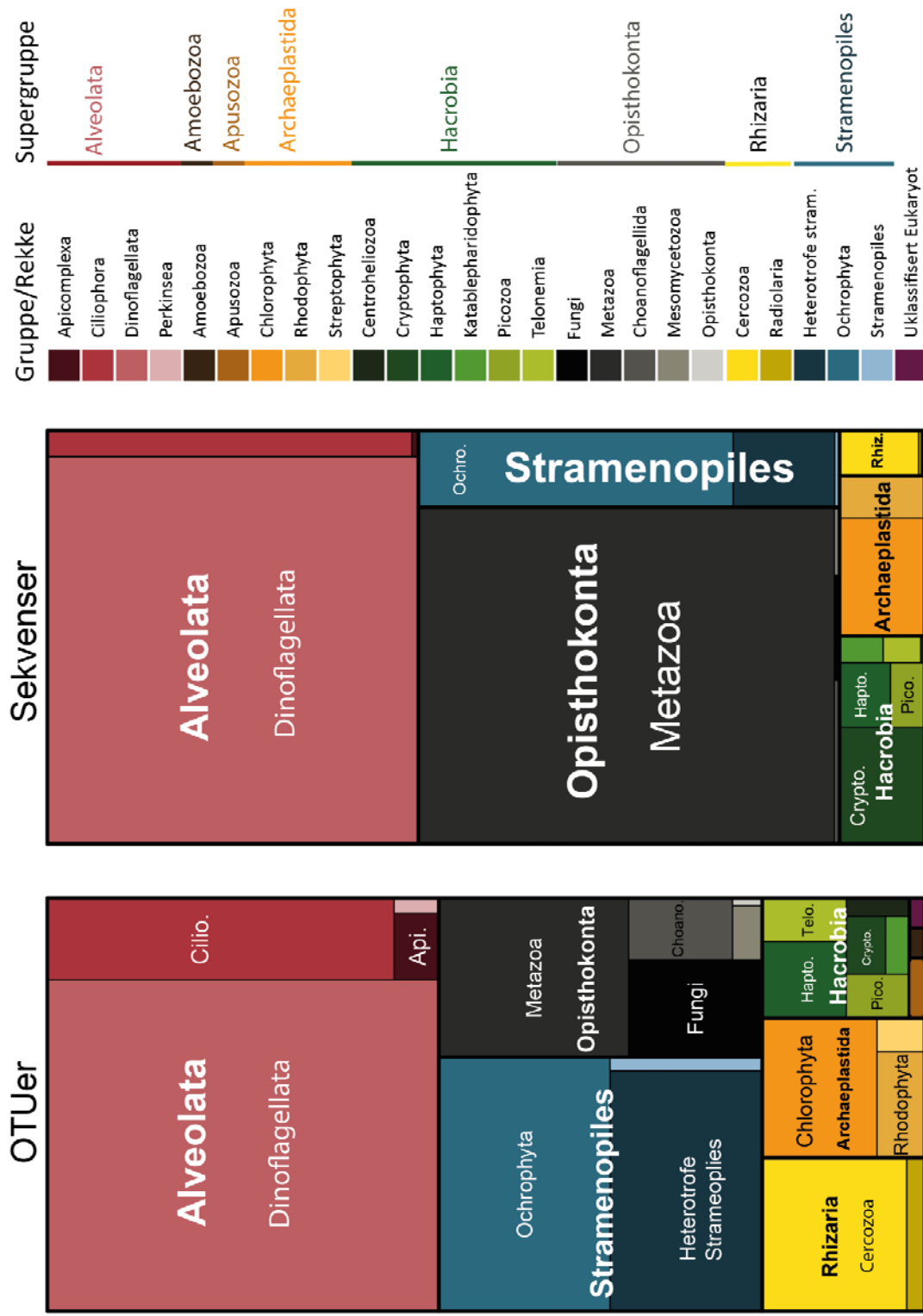
Fra Norsk Sekvenseringssenter mottok vi 60 (30×2) filer i formatet fastq.gz, det var 2 filer per prøve som inneholdt et likt antall sekvenser (én fil for forover- og én for revers sekvenser). I filene tilsammen var det totalt 19 726 172 ($9\,863\,086 \times 2$) sekvenser. Gjennom pipelinen ble over 45% av sekvensene filtrert bort (tabell S.2.). I tillegg ble det fjernet 10 408 kimærer og single sekvenser som til sammen utgjorde 285 323 sekvenser ("Rådata" i tabell S.3). Videre ble det manuelt fjernet OTUer som totalt i datasettet hadde under 10 sekvenser og OTUer som forkom i kun én prøve (tabell 3.1).

Tabell 3.1 Antallet OTUer og tilhørende sekvenser som manuelt er fjernet fra datasettet da de kun forkom i én prøve eller til sammen hadde under 10 sekvenser.

	Antall OTUer	Antall sekvenser
Rådata før manuell filtrering	-	5 057 831
Forkommer i én prøve	95	4 093
<10 sekvenser	1 044	4 892
Totalt fjernet	1 139	8 985
Gjenværende i datasettet	1 584	5 048 846

Det opprinnelige datasettet

De gjenværende sekvenser (tabell 3.1) tilhører 1 584 unike OTUer som har 5 048 846 representative sekvenser med en gjennomsnittslengde på 417 bp. Disse utgjør det "opprinnelige datasettet" (tabell S.3) og ble taksonomisk klassifisert etter det taksonomiskesystemet brukt i referansedatabasen PR². OTUene tilhørte 26 grupper som i hovedsak er taksonomiske rekker, tilhørende 8 supergrupper og uklassifiserte eukaryoter (figur 3.1).



Figur 3.1. Treemap som inkluderer Fungi, Metazoa og Streptophyta. Antallet OTUer (venstre) og deres relative tetthet i sekvens (høyre) er proporsjonalt inndelt etter tilhørende taksonomiske gruppe/rekke og supergruppe (uthevet), basert på det opprinnelige datasettet.

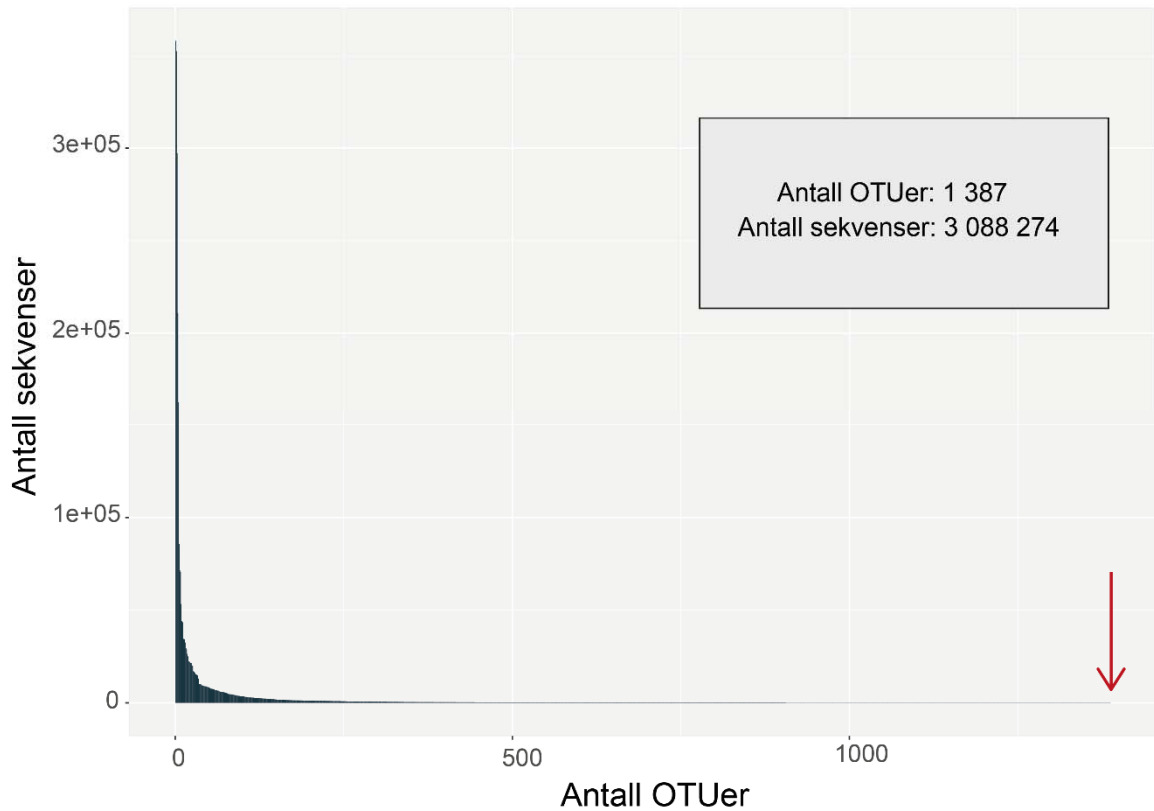
I tabell S.4 kan man finne antallet OTUer og sekvenser for hver av de taksonomiske gruppene. Protist-gruppene blir beskrevet i større detalj i seksjon 3.2.2. Figur 3.1 illustrerer hvilke grupper i planktonsamfunnet som ble påvist ved hjelp av metastrekkoding. Gruppene som ikke består av protister (Fungi, Metazoa, Opisthokonta (uklassifiserte) og Streptophyta) vil kun raskt nevnes under, siden denne oppgaven i hovedsak dreier seg om protister.

Ikke-protist-gruppene utgjorde 197 OTUer med 39% av antallet sekvenser, hvor Metazoa (dyr) dominerte med 132 OTUer og nesten 2 millioner sekvenser (38.6%) (tabell S.4). De OTUene med flest sekvenser i datasettet var OTU 1 og OTU 2, som til sammen hadde over 1 million sekvenser (21.1%) (OTU-tabell i figshare). Begge OTUene ble klassifisert til familien Maxillopoda i klassen Athropoda (leddyr), og OTU 1 ble bestemt til hoppekreps-arten *Paracalanus parvus*. Det var kun 39 av OTUene i Metazoa som ikke ble klassifisert til art. Av de 57 OTUene som utgjorde gruppen Fungi (sopp) ble alle klassifisert minst til taksonomisk orden og 10 av disse bestemt til art. Det var kun én OTU fra uklassifiserte opisthokonter og 7 OTUer er Streptophyta (høyere planter).

Protister

Alle OTUer og tilhørende sekvenser som ikke ble klassifisert som protister (markert i gult i supplementær tabell S.4.) ble fjernet fra det originale datasettet, som resulterte i "Protistdatasettet". Dette datasettet bestod av 3 088 274 sekvenser fordelt på 1 387 OTUer (figur 3.2). OTUene ble klassifisert til 22 grupper/rekker som tilhørte 8 forskjellige supergrupper (figur 3.5). Det var en ujevn fordeling av sekvenser per OTU. Hele 1226 OTUer (88%) ble definert som sjeldne OTUer (<0.05% sekvenser per OTU) (OTU-tabell i figshare) og kun 15 OTUer som vanlige (>1% sekvenser per OTU) (tabell 3.2). Protistdatasettet ble subsamplet til et likt antall sekvenser (10 530), etter den prøven med færrest sekvenser. For å teste om subsamplingen påvirket den relative OTU-sammensetningen mellom prøver ble det gjort gjentatte subsamplinger som ble sammenlignet. De alle viste en korrelasjon på >0.999 med p-verdi <0.001. Etter subsamplingen bestod det subsamplede datasettet av 1315 OTUer og ca. 1/10 av sekvensene fra protistdatasettet (tabell S.4). Det var 72 OTUer som gikk tapt under subsamplingen og antallet sekvenser som tilhørte disse utgjorde 0.04% av protistdatasettet. Antallet OTUer og sekvenser fordelte seg fortsatt svært likt på de taksonomiske gruppene om man sammenligner protistdatasettet før og etter det ble subsamplet (figur 3.5). De 1 315 OTUene ble taksonomisk klassifisert nærmest til 1 095 referansesekvenser i referansedatabasen PR², 220 OTUer ble altså tildelt samme referansesekvensen som en eller flere andre OTUer.

Som eksempel ble 6 forskjellige OTUer taksonomisk klassifisert nærmest FJ402945 *Triplos candelabrus*, med varierende prosent identitet til referansesekvensen (OTU-tabell i figshare). Det subsamplede datasettet fordeles etter dyp i Seksjon 3.3.



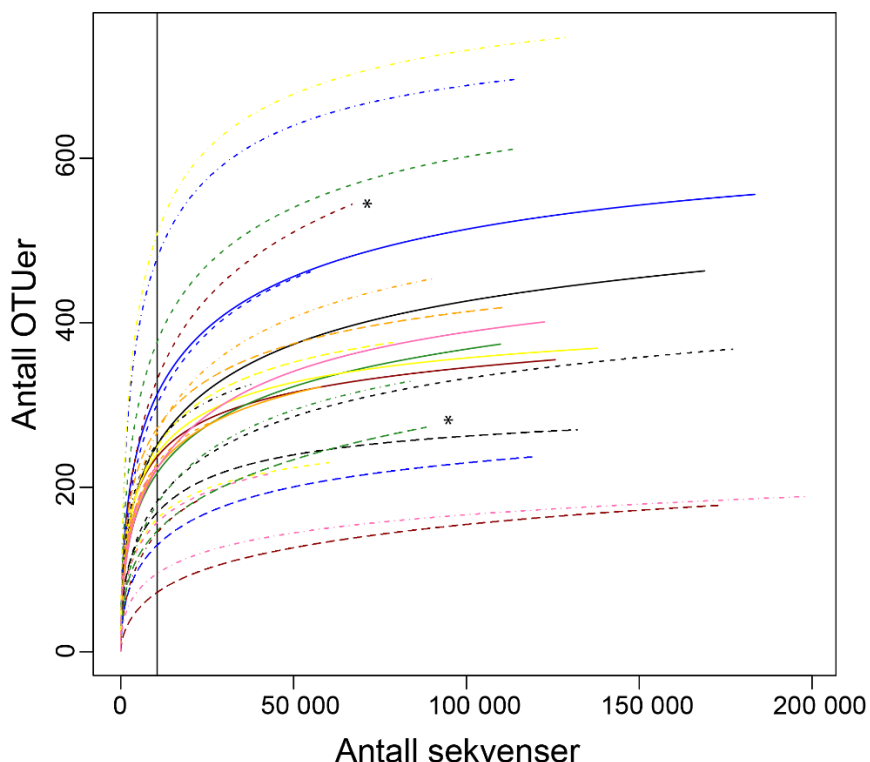
Figur 3.2. Viser hvordan de totalt 3 088 274 sekvenser i protistdatasettet er fordelt på de 1 387 OTUene, i rekkefølge fra mest til minst abundant OTU (OTU nr. 1 387 har 10 sekvenser, markert med rød pil). Legg merke til at de aller fleste OTUene har få sekvenser.

Tabell 3.2. Tabellen viser de 15 mest abundante (>1% sekvenser) OTUene i protistdatasettet samt deres taksonomiske klassifisering tildelt av Blast-søk mot referansedatabasen PR². Til tre OTUer ble taksonomiske klassifiseringer (kasse, orden og familie) korrigert mot algaebase.com og er markert med *.

OTU NR.	SUPERGRUPPE	REKKE	KLASSE	ORDEN	FAMILIE	ART
OTU 22	Alveolata	Dinoflagellata	Dinophyceae	Dinophysiales	Dinophysiaceae	<i>Dinophysis acuminata</i>
OTU 3				Gonyaulacales	Ceratiaceae	<i>Tripos candelabrus</i>
OTU 103						<i>Tripos fusus</i>
OTU 8					Goniodomataceae	<i>Alexandrium hiranoi</i>
OTU 4				Peridiniales	Heterocapsaceae	<i>Heterocapsa</i> sp.
OTU 10				Suessiales	Ukjent	Ukjent
OTU 29				Ukjent	Ukjent	Ukjent
OTU 1033						Ukjent
OTU 14						Ukjent
OTU 16	Archaeplastida	Chlorophyta	Mamiellophyceae	Mamiellales	Bathycoccaceae	<i>Bathycoccus prasinos</i>
OTU 19			Chlorodendrophyceae	Chlorodendrales	Chlorodendraceae	Ukjent
OTU 11	Hacrobia	Cryptophyta	Cryptophyceae	Pyrenomonadales	Geminigeraceae	<i>Teleaulax amphioxeia</i> *
OTU 15	Stramenopiles	Ochromytha	Mediophyceae	Thalassiosirales	Thalassiosiraceae	<i>Thalassiosira</i> sp. *
OTU 12				Stephanodiscales	Stephanodiscaeae	<i>Cyclotella</i> sp. *
OTU 13				Chaetocerotales	Chaetocerotaceae	<i>Chaetoceros</i> sp. *

Rarefaction

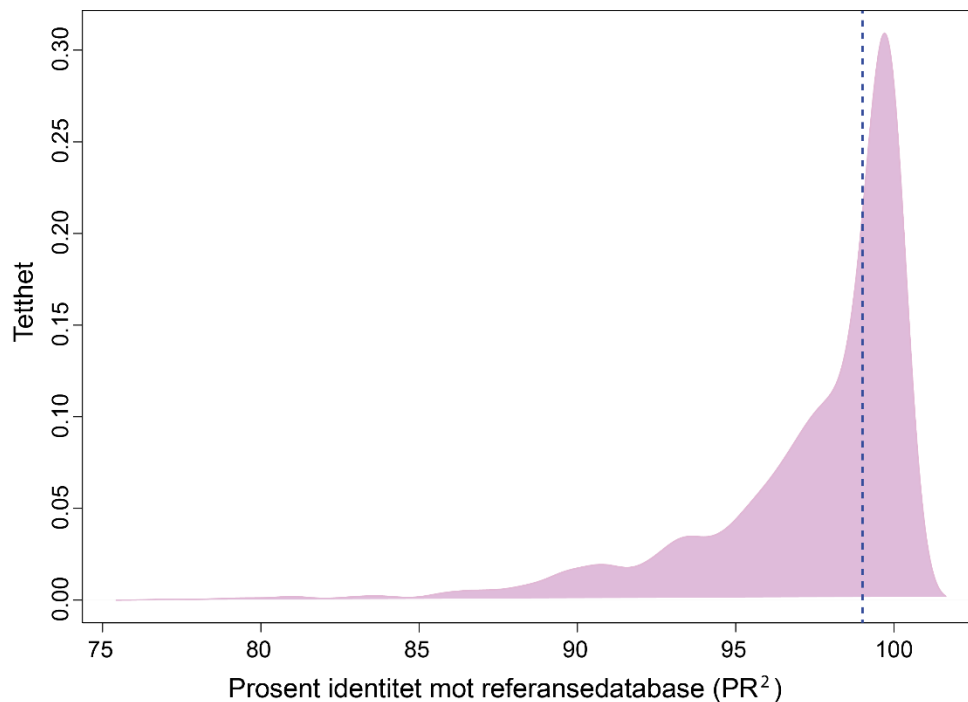
For å undersøke om sekvenseringsdybden var tilstrekkelig til å fange opp protistsamfunnets struktur og sammensetning ble det laget et rarefaction-plot basert på protistdatasettet (figur 3.3). Rarefactionkurver blir dannet ved å ta ut et gitt antall tilfeldige sekvenser fra hver enkelt prøve og fremstille hvor mange OTUer disse tilhører. Dette blir gjort gjentatte ganger og antallet sekvenser man tar ut øker for hver gang, inntil man har tatt ut det totale antallet sekvenser i prøven. Kurven viser OTU rikheten for et gitt antall sekvenser og gir et estimat på antallet OTUer som hypotetisk kan bli oppdaget. Hvis sekvenseringsdybden er tilstrekkelig vil kurven for prøven flate ut. Dette betyr at det ikke blir plukket ut sekvenser som tilhører nye OTUer. Alle prøvene i protistdatasettet flater ut, utenom to (markert med * i figur 3.3.), som tilsier at sekvenseringsdybden stort sett var tilstrekkelig til å fange opp OTU-rikheten som potensielt kunne vært i hver enkelt av disse prøvene. Kurven til prøve nr.10 (tatt den 25.11.16, fra 0-2 m) og prøve nr.26 (tatt den 18.10.17, fra 5 m) flater ikke ut. Dette kan tolkes som at det fantes OTUer i disse prøvene som ikke ble sekvensert.



Figur 3.3. Rarefaction-diagram basert på protistdatasettet for alle 30 prøver. Hver linje i diagrammet representerer én prøve. Vertikal svart linje i diagrammet markerer antallet sekvenser protistdatasettet ble subsamlet på (10 530).

OTU-rikhet og sekvenser fordelt på taksonomiske grupper

De aller fleste OTUene ble klassifisert til en taksonomisk gruppe, utenom to OTUer som omtales som uklassifiserte eukaryote. De taksonomiske gruppene er 17 rekker og 5 grupper. De heterotrofe stramenopilene er en gruppe. Gruppen Stramenopiles er andre stramenopiler som ikke med sikkerhet kan klassifiseres som heterotrofe eller som fototrofe ochrofytter. I PR² blir Amoebozoa og Apusozoa omtalt som supegrupper (Guillou mfl., 2013) og de få OTUene påvist i disse supergruppene, ble samlet i grupper med samme navn. Den siste gruppen er uklassifiserte eukaryote. Omtrent halvparten (48%) av OTUene viste >99% likhet med en sekvens i referansedatabasen PR² (figur 3.4). Det var 287 OTUer som hadde 100% identitet med en referansesekvens i PR² og 54 av disse ble klassifisert til art (tabell S.8).

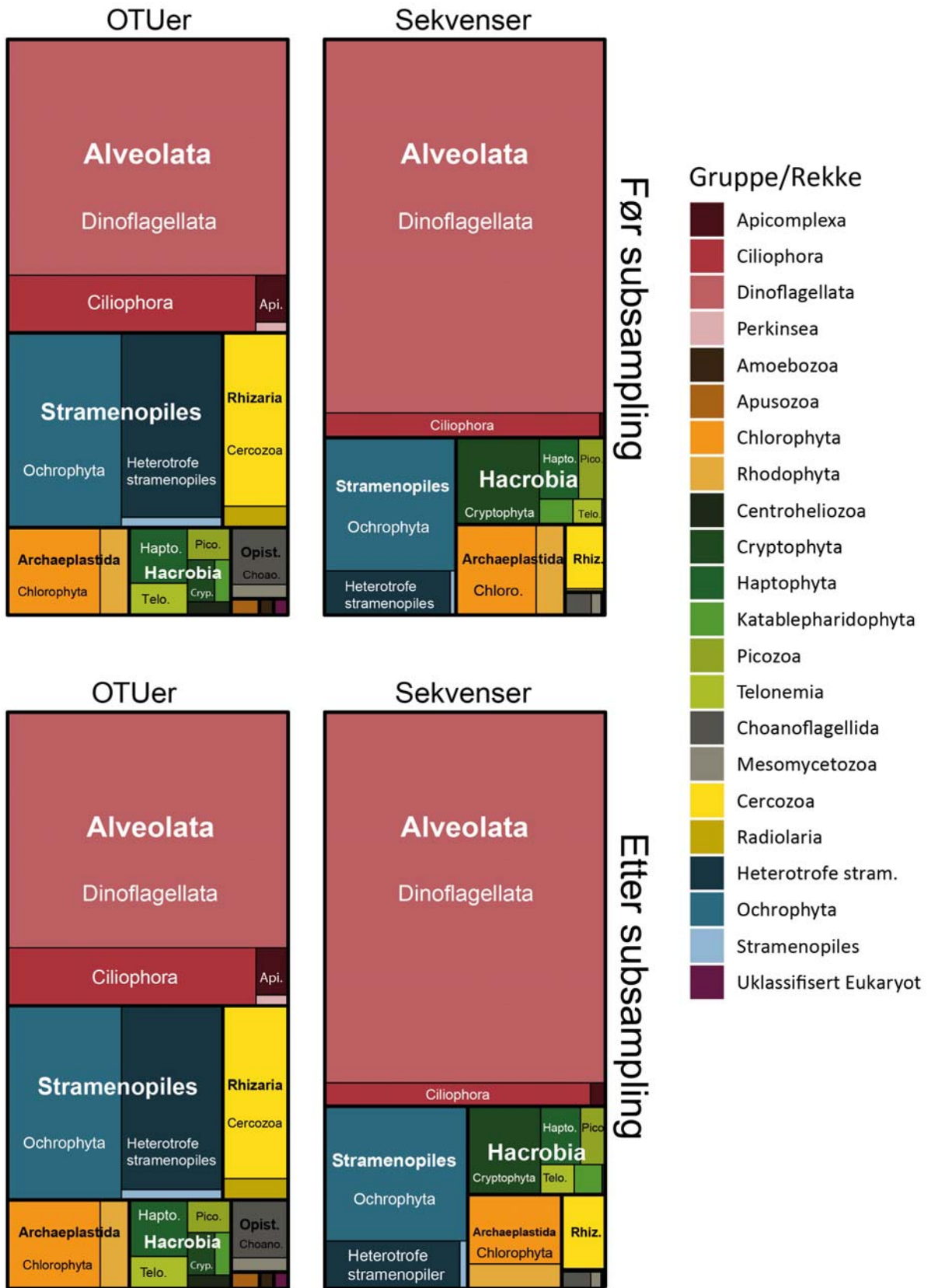


Figur 3.4. Tetthetsdiagram (glatt histogram) som viser fordelingen av OTUenes identitetsprosent mot en referansesekvens i den kuraterte referansedatabasen PR². Striplet linje markerer 99%-likhet og vi ser at omtrent halvparten (48%) av OTUene har mer enn 99% likhet med referansesekvens. Basert på protistdatasettet.

Tabell 3.3. Andel OTUer og sekvenser regnet i prosent (%) for hver supergruppe i protist- og det subsamplede datasettet.

	PROTISTDATASETT		SUBSAMPLET DATASETT	
	OTUer (%)	Sekvenser (%)	OTUer (%)	Sekvenser (%)
ALVEOLATA	50.97	69.16	50.65	68.36
AMOEBOSOA	0.14	<0.005	0.08	<0.005
APUSOZOA	0.29	<0.005	0.30	<0.005
ARCHAEPLASTIDA	6.63	6.16	6.77	5.44
HACROBIA	5.55	7.97	5.70	7.56
OPISTHOKONTA	2.52	0.53	2.66	0.44
RHIZARIA	7.86	1.68	7.91	2.04
STRAMENOPILES	25.88	14.47	25.78	16.11
UKL. EUKARYOTA	0.14	0.03	0.15	0.03

Forholdet mellom andelen OTUer og andelen sekvenser i datasettet viste liten korrelasjon og fordelte seg ujevnt på de forskjellige supergruppene (tabell 3.3). Alveolatene dominerte tydelig datasettet med flest OTUer og sekvenser, etterfulgt av stramenopilene. Supergruppen Rhizaria og Ochrophyta viste lav andel sekvenser i forhold til OTUer, og det motsatte gjelder for Hacrobia. Archaeplastida var supergruppen som viste størst likhet mellom andel sekvenser og OTUer i datasettet. Amoebozoa, Apusozoa og uklassifiserte eukaryoter utgjorde så liten del av datasettet at de ikke vil bli diskutert videre.



Figur 3.5. Treemap av protistgrupper. Antallet OTUer (venstre) og deres relative tetthet i sekvenser (høyre) er proporsjonalt inndelt etter tilhørende taksonomiske gruppe/rekke og supergruppe (uthevet), basert på protistdatasettet (topp) og det subsamlet datasett (bunn).

3.2 Variasjon gjennom året

3.2.1 Sesongvariasjon i miljøvariabler

De hydrografiske forholdene og klorofyll-a-fluorecens ved DK1 varierte gjennom prøvetakningsperioden (figur 3.6). Prøvetakningsdatoene for målingene er listet i tabell S.1. Temperaturmålingene for overflaten (0-2 m) og 5 m var identiske ved de fleste målingene og viste en sterk sesongvariasjon, med varmeste registrering på 19°C i august og kaldeste på 1.3°C i februar (figur 3.6.A).

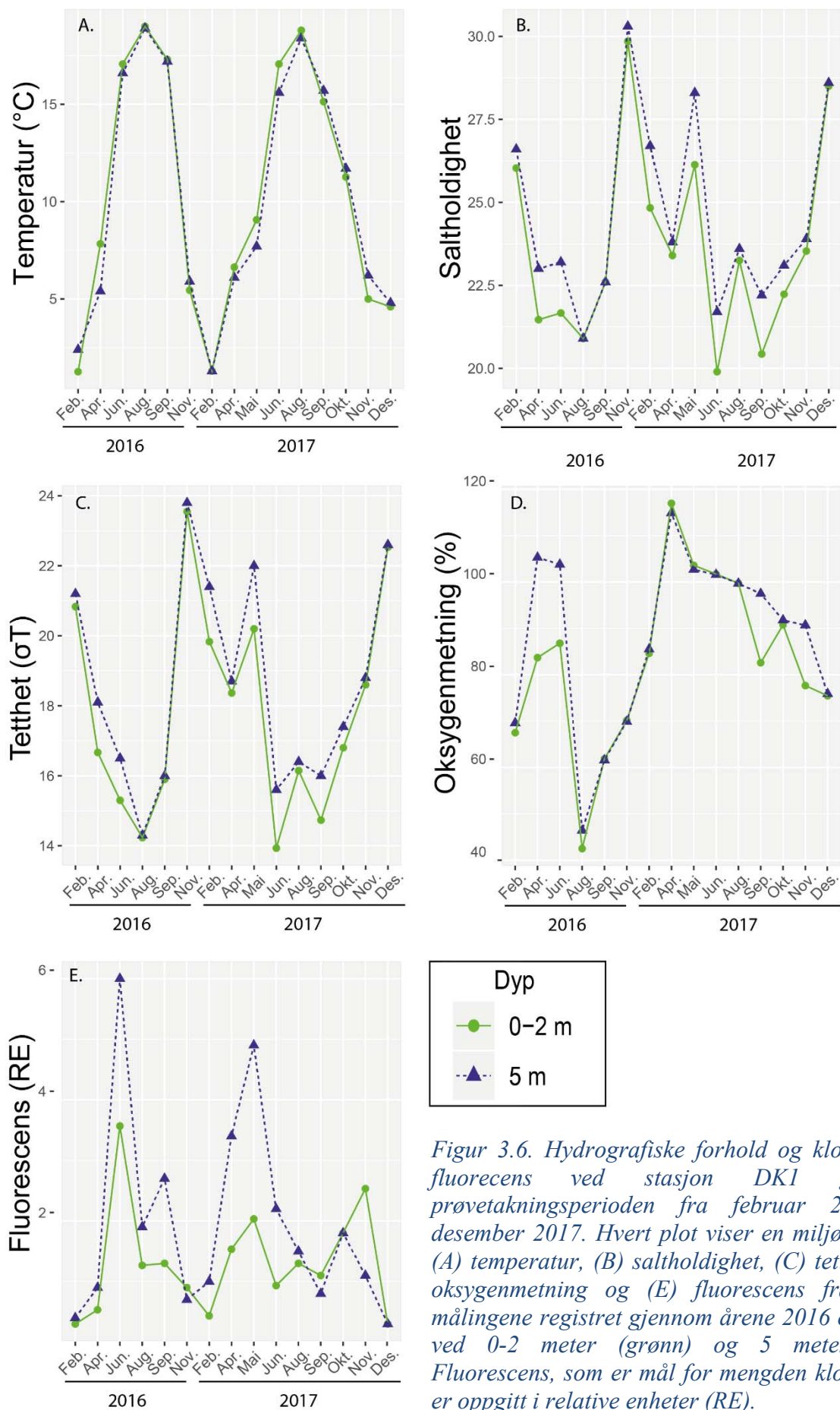
Saltholdigheten varierte mellom 19.9 og 30.3 og tettheten (σ_t) fra 14.2 til 23.8 (figur 3.6.B-C). Saltholdighet og tetthet var sterkt korrelert for begge dyp (0-2 m: $R^2 = 0.9397$, p-verdi >0.001 ; 5m: $R^2 = 0.9377$, p-verdi >0.001). Overflatevannet viste generelt noe lavere verdier for saltholdighet og tetthet enn ved 5 m, spesielt i månedene april-juni (2016), samt juni og september (2017).

Dybden på det øvre blandingslaget varierte mye gjennom de to årene (figur 3.7 og figur S.3). Det dypeste var på 12 m i april 2016 og grunnest var på 2 m i mai 2017. I mai 2017 var det en markant økning i saltholdighet og tetthet (figur 3.6.B-C).

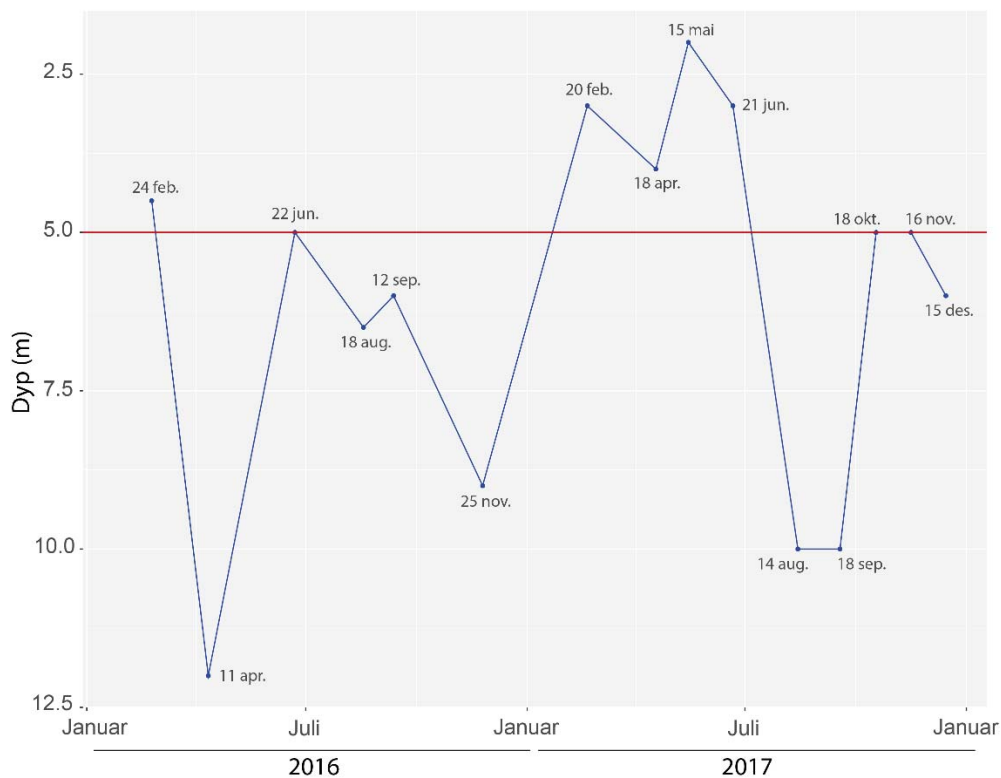
Oksygenmetningen var lavere i overflatevannet i vårmånedene (april-juni) i 2016 og september og november (2017) enn ved 5 m (figur 3.6.D). Metningen ved 5 m overskred 100% i noen av vårmånedene begge år, og det samme gjaldt for overflatevannet i 2017. Det kan være fordi primærproduksjonen denne perioden er så høy at havet kan gi fra seg oksygen til luften. For at dette skal være mulig forventes det en stor algebiomasse.

Dette ble bekreftet av forhøyede fluorescens-målinger ved spesielt 5 m samme periode (figur 3.6.E). Fluorescens er et mål på mengden klorofyll-a som igjen er et estimat på algebiomasse, her oppgitt i relative enheter. Mengden fluorescens viste store svingninger gjennom året for begge dyp og har ved 5 m to tydelige toppe i vår-sommermånedene. Det var en svak økning i fluorescens i høstmånedene (sep-nov) begge år.

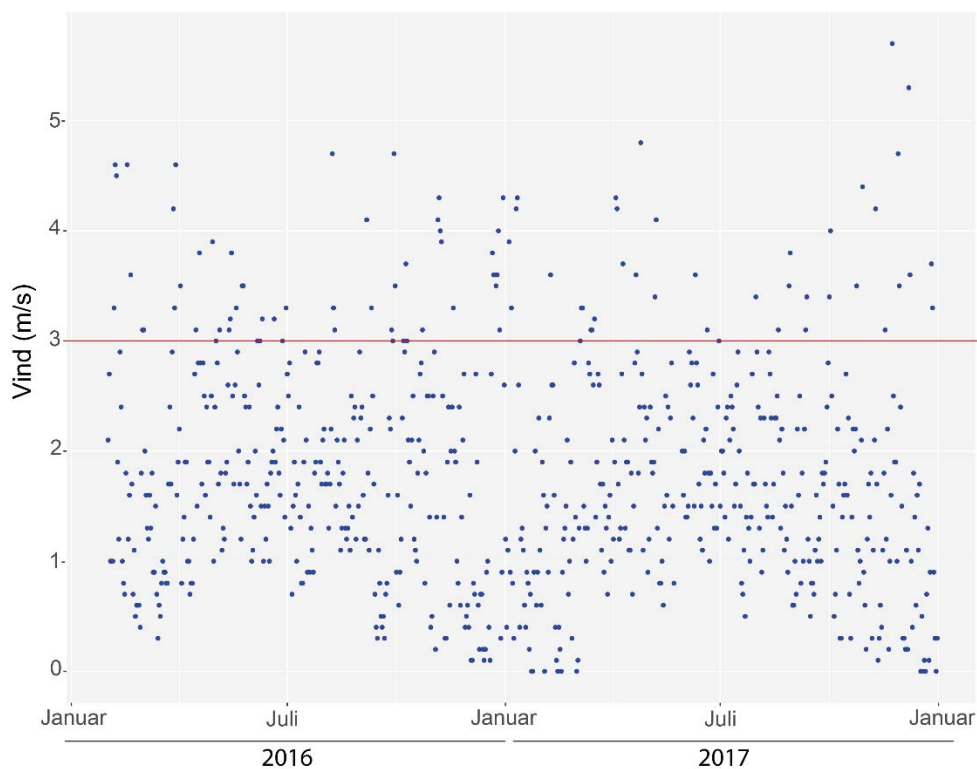
Vindhastigheten varierte tydelig med sesong (figur 3.8). Det var størst spredning i vindhastighet i vintermånedene, hvor det var dager uten vind og dager med vind helt opp til 6 m/s. I sommermånedene blåste det hver dag, men relativt sjeldent over 3 m/s. Vindhastigheten i prøveperioden (feb. 2016 – des. 2017) viste en gjennomsnittlig vindhastighet på 1.77 m/s per dag; 93 av dagene var vindhastigheten over 3 m/s (markert med rød strek i figur 3.8). Alle månedene, utenom juli 2016, hadde minst en dag (i snitt 4.2 dager) der vindstyrken var over 3 m/s.



Figur 3.6. Hydrografiske forhold og klorofyll-a-fluorescens ved stasjon DK1 gjennom prøvetaksperioden fra februar 2016 til desember 2017. Hvert plot viser en miljøvariabel (A) temperatur, (B) saltholdighet, (C) tetthet, (D) oksygenmetning og (E) fluorescens fra CTD-målingene registret gjennom årene 2016 og 2017, ved 0-2 meter (grønn) og 5 meter (blå). Fluorescens, som er mål for mengden klorofyll a, er oppgitt i relative enheter (RE).



Figur 3.7. Dypet til det øvre blandingslaget gjennom prøvetakningsperioden, ved DK1. Datoer for prøvetaking er oppgitt. Den røde horisontallinjen markerer 5 meters dyp.



Figur 3.8. Gjennomsnittlig vindstyrke (meter per sekund) for hver dag, i perioden februar 2016 til desember 2017, ved metrologiske målestasjonen på Ås (nærmeste metrologiske målestasjon til prøvetakningsstasjonen DK1). Den røde horisontallinjen markerer vindhastighet 3 m/s.

3.2.2 Protistsamfunnets sammensetning gjennom årstider

Sesongvariasjonen til protistsamfunnet ble undersøkt ved å se på antallet OTUer (figur 3.9), prosentandel andel sekvenser fordelt på gruppe/rekke (figur 3.10) og prosentandel sekvenser fordelt på klasse (figur 3.11) for hver enkelt prøve tatt ved de to dypene (0-2 m og 5 m). Prøven representerer måneden den ble tatt (foruten om 2 og 20 februar 2017), eksakte datoer står listet i tabell S.1.

Hver supergruppe vil nå bli presentert hver for seg. Beskrives antall OTUer refereres det til figur 3.9 og OTU-tabell i figshare (Saubrekka og Bjørneklett, 2019). Antall OTUer er et mål på diversitet. En gruppe med et stort antall OTUer har høy diversitet, altså et stort mangfold og rikhet. Prosentandelen sekvenser i prøven forteller om tettheten til gruppene i forhold til hverandre; deres relative abundans. Det refereres til figur 3.10 for gruppenes prosentandel sekvenser og figur 3.11 for prosentandel sekvenser fordelt på taksonomisk klasse. Disse tre figurene burde for et helhetlig bilde sees i sammenheng og derfor presenteres hver supergruppe isolert.

Alveolata

Supergruppen Alveolata dominerte datasettet og viste størst diversitet i vintermånedene november 2016 og desember 2017 og lavest diversiteten i vår-sommer månedene (mars-juni). Diversiteten ved begge dyp var gjennomgående lik gjennom året, foruten synlig færre OTUer i overflaten (0-2 m) i mai og november 2017 (figur 3.9).

Det kan være noen få OTUer som dominerer samfunnet med en stor andel sekvenser, altså høy relativ abundans. Dette så vi i prøven fra mai 2017 (0-2m), der én dinoflagellat-OTU (OTU 29) utgjorde 93% av alle sekvensene i prøven (figur 3.10 og tabell S.7). Dinoflagellatene viste ikke en tydelig sesongvariasjon i andelen sekvenser, med totalt 567 unike OTUer i protistdatasettet som varierte i relativ abundansen gjennom året. Det var spesielt månedene juni og september (2016), mai, juni og oktober (2017) som ble dominert av noen få dinoflagellat-OTUer i klassen Dinophyceae (figur 3.11 og tabell S.7). Denne klassen utgjorde hele 60% av alle sekvenser i datasettet og var klart dominant da vi sammenlignet alle 79 klassene protister som ble påvist. Hele 9 av de 15 vanligste OTUene i protistdatasettet tilhører klassen Dinophyceae (tabell 3.2), som plasseres i et fylogenetisk tre i seksjon 3.3. OTU 3, som ble klassifisert til *Tripos candelabrus* i PR², dominerte ofte prøvene ved begge dyp gjennom hele prøvetakningsperioden (tabell S.7).

Syndinophyceae var den nest mest abundante dinoflagellat-klassen, med 297 OTUer som til sammen hadde over 5% av sekvenser i datasettet, som alle tilhørte Dino-gruppe I eller Dino-gruppe II i ordenen Syndiniales (Figur 3.11 og OTU-tabell i figshare).

Ciliatene (Ciliophora) var tilstede gjennom hele året ved begge dyp og hadde til sammen et høyt antall OTUer (124) (figur 3.9), men lav abundans (figur 3.10). OTU 31, klassifisert til ordenen Strombidiida, var mest frekvent av ciliatene, med høy abundans spesielt i februar og november 2016 og februar 2017 (OTU-tabell i figshare).

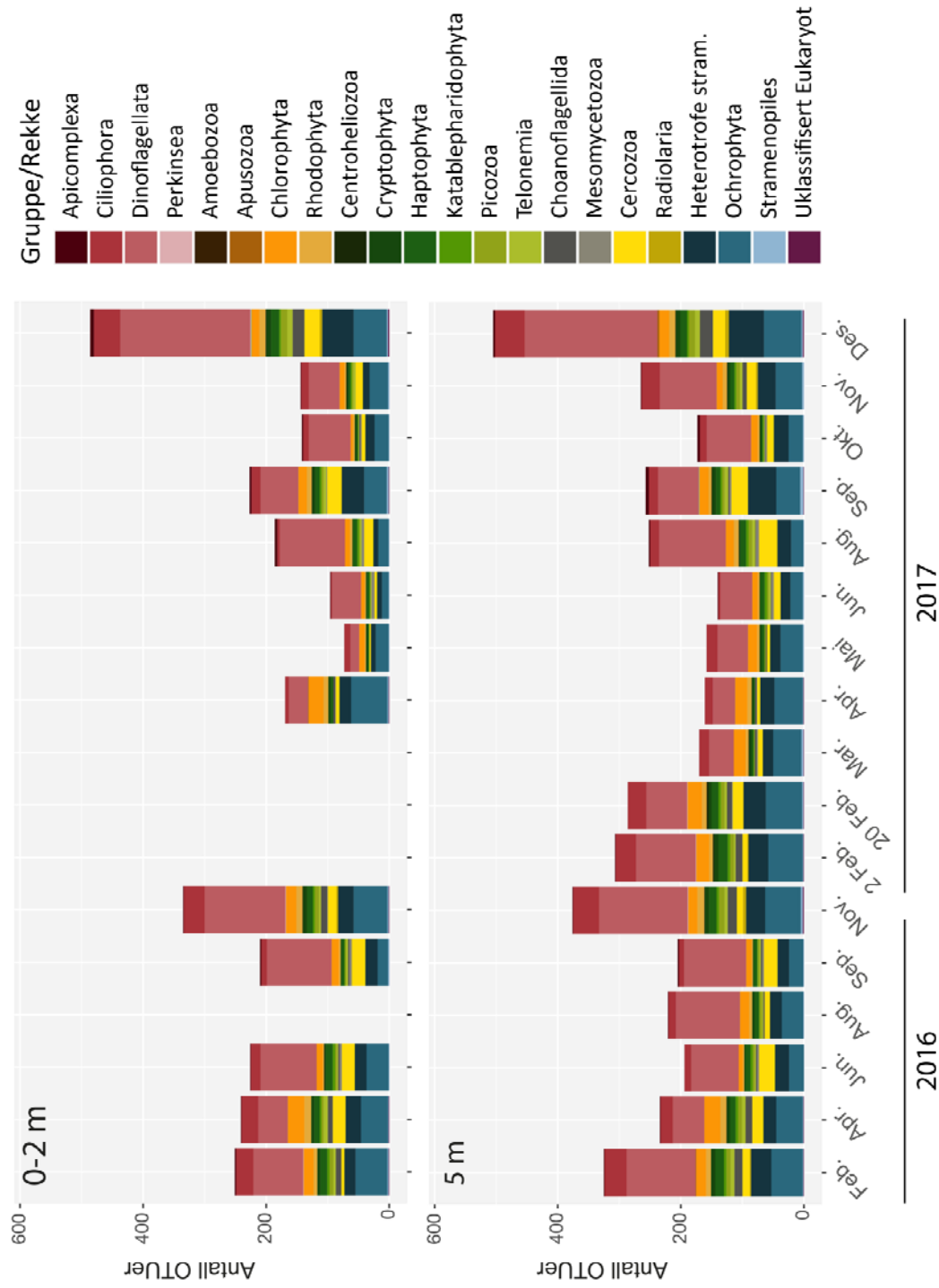
Alveolatene hadde lav andel sekvenser i mars ved 5 meter og det var tydelig den stramenopile supergruppen (Stramenopiles) som dominerte (figur 3.10).

Stramenopiles

Stramenopilene viste en ganske lik sesongvariasjon som Alveolatene, med lavest diversitet i vår-sommermånedene (april - juli) og høyest i vintermånedene (nov. - des.) (figur.3.9). Den relative abundansen til disse to supergruppene så ut til å fluktuere parallelt: der den ene avtar øker den andre og motsatt (figur 3.10). Abundansen av stramenopiler var høy ved overflaten i juni 2016 og april 2017. Ved 5 m var abundansen høy i august 2016 og februar-april 2017.

Det var rekken Ochrophyta som innen stramenopilene hadde størst diversitet og høyest relativ abundans (figur 3.9 og figur 3.10). Klassen Mediophyceae (senriske diatoméer) dominerte prøven fra 5 m i mars 2017 (figur 3.11) og spesielt OTU 15 hadde høy relativ abundans (tabell S.7). OTU 15 ble klassifisert til *Thalassiosira* sp. Andre diatomé-OTUer som var realtvt abundante (ca. 5% sekvenser) under våroppblomstringen i mars tilhørte slekten *Rhizosolenia* (OTU-tabell i figshare). OTU 15 var sammen med OTU 12 og OTU 13, som ble klassifisert til henholdsvis *Cyclotella* sp. og *Cheatoceros* sp., blant de 15 vanligste OTUene i protistdatasettet (tabell 3.2). Klassen Mediophyceae viste også høy relativ abundans på 2 m i juni og 5 m i august 2016, samt november 2017 ved begge dyp (figur 3.11). Innen klassen Bacillariophyceae tilhørte majoriteten av sekvensene slektene *Pseudo-nitzschia* og *Thalassionema*.

De heterotrofe stramenopilene, spesielt den abundante MAST-kladen viste en høy abundans april 2017 ved 5 m dyp (figur 3.11).



Figur 3.9. Antall OTUer. For hver prøve tatt gjennom 2016–2017, ved dypene 0–2 meter (øverst) og 5 meter (nederst) fremstilles totalt antall OTUer og hvordan de fordeler seg på de taksonomiske gruppene/rekkene. Fargene i hver søyle har lik rekkefølge som tegnforklaringen til høyre. Over de datoene det mangler en søyle, er det ikke tatt prøver ved dette dypet (0–2 meter) den datoen. Diagrammet er basert på subsamplet protidsdatasett fordelt etter dyp.

Hacrobia

Hacrobia, som bl.a. inneholder rekken Haptophyta (svepeflagellater), hadde som de andre supergruppene størst diversitet i vintermånedene og lavest i vår-sommer månedene. Den viste verken et tydelig sesongmønster eller trender i antall OTUer eller andel sekvenser gjennom de to årene (figur 3.9 og figur 3.10).

Klassen Cryptophyceae (svelgflagellater) var den mest abundante og eneste klassen i rekken Cryptophyta (figur 3.11). Innen denne klassen fant vi én av de 15 vanligste OTUene i datasettet, som ble klassifisert til arten *Teleaulax amphioxeia* (tabell 3.2).

Av haptofyttene (Haptophyta) viste klassen Prymnesiophyceae en plutselig stor økning i relative abundans i august 2017 ved 5 meters dyp (figur 3.11). I prøven tatt i august 2017 var det to OTUer som dominerte; den mest dominante tilhørte dinoflagellaten *Heterocapsa* sp. og den andre til slekten *Emiliana* (tabell S.7). Slekten *Chrysochromulina* var den slekten med høyest diversitet innen klassen Prymnesiophyceae, med seks OTUer som hadde høy relativ abundans i feb.-apr. 2016 og sep. 2017. Kun én av *Chrysochromulina*-OTUene ble klassifisert til art, *Chrysochromulina leadbeateri* (OTU-tabell i figshare).

Rekken Telonemia hadde høy relativ abundans i sep. 2017, ved 5 m. De heterotrofe rekkene Centroheliozoa, Katablepharidophyta, Picozoa og Telonemia viste en lav abundans gjennom gjennom hele prøvetakningsperioden, foruten Picozoa som hadde en høy abundans i desember 2017 (figur 3.10). Innen disse rekkene var det kun OTUene som tilhørte Katablepharidophyceae som ble bestemt til klasse (figur 3.11). Hacrobia var supergruppen der antallet OTUer var positivt korrelert med relative abundans i protistsamfunnet ved flere måneder.

Rhizaria

Supergruppen Rhizaria, hvor vi finner de heterotrofe rekkene Cercozoa og Radiolaria, viste en økning i relative abundans i september 2017, en måned etter Haptophyta (figur 3.10)

Rekken Cercozoa hadde en høy diversitet med 94 OTUer i subsamlet protistdatasett (tabell S.4). Av disse var OTUen som ble klassifisert til *Cryothecomonas aestivalis* den mest frekvente og abundante (OTU-tabell i figshare). Ellers hadde Rhizaria relativt få sekvenser gjennom de to årene, som fordelte seg på et varierende antall OTUer.

Archaeplastida

Supergruppen Archaeplastida (planter) viste høy OTU-rikhet ved begge dyp fra desember til april (figur 3.9). Antallet OTUer korrelerte ikke med andel sekvenser, som varierte mellom dypene (figur 3.10). Dette så vi spesielt i februar og november 2016, august og september 2017, hvor antallet OTUer var relativt likt for begge dyp. Andelen sekvenser var synlig større i et av dypene; ved 0-2 m i feb. 2016 og aug 2017 og ved 5 m i nov. 2016 og sep. 2017. Rekken Chlorophyta (grønnalger) var mest tallrik i antall OTUer og sekvenser i denne supergruppen. En av de vanligste OTUene i protistdatasettet tilhørte *Bathycoccus prasinus* i klassen Mamiellophyceae (tabell 3.2.). Slekten *Micromonas*, med fire OTUer, hadde tilnærmet like mange sekvenser til sammen som *Bathycoccus prasinus* (OTU-tabell i figshare). Klassen Mamiellophyceae hadde høy relativ abundans ved begge dyp i februar-april i 2016 og spesielt ved 5 m i november samme år (figur 3.11). Klassen Chlorodendrophyceae synliggjorde seg merkbart i augustprøven fra 2017, dominert av bl.a. OTU 19 klassifisert til ordenen Chlorodendrales (figur 3.11 og tabell S.7).

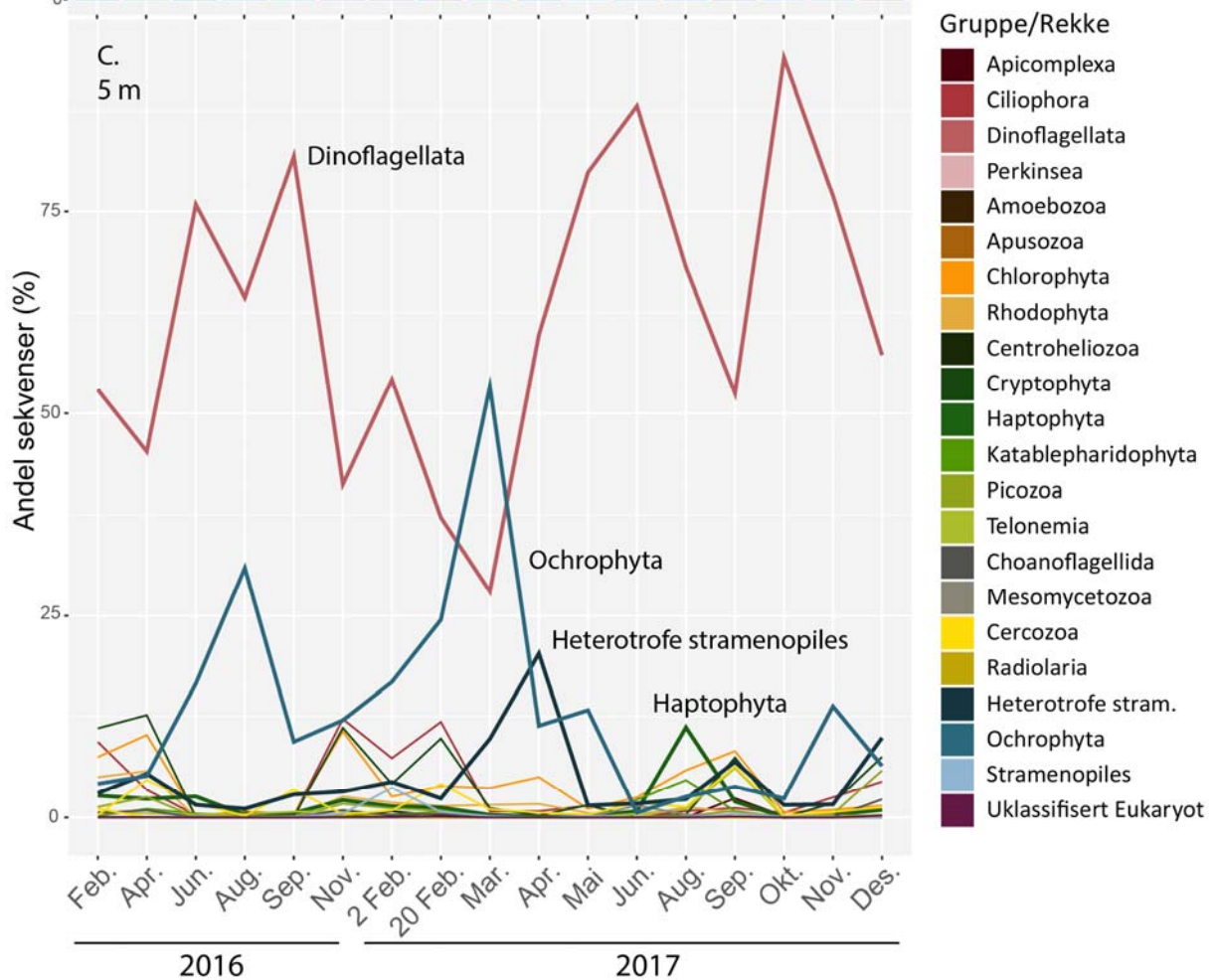
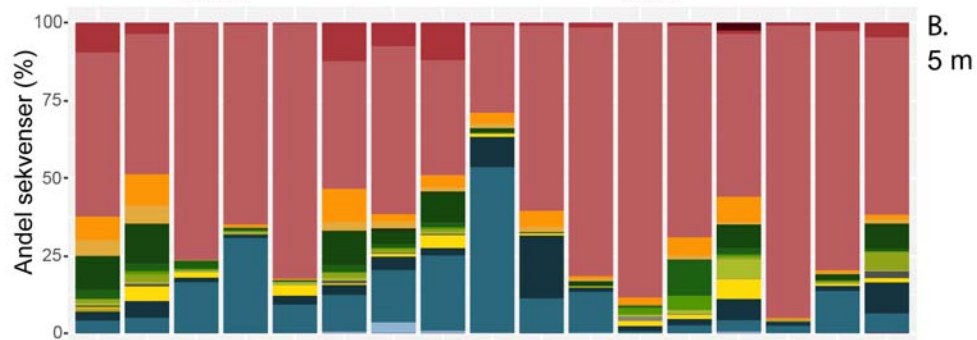
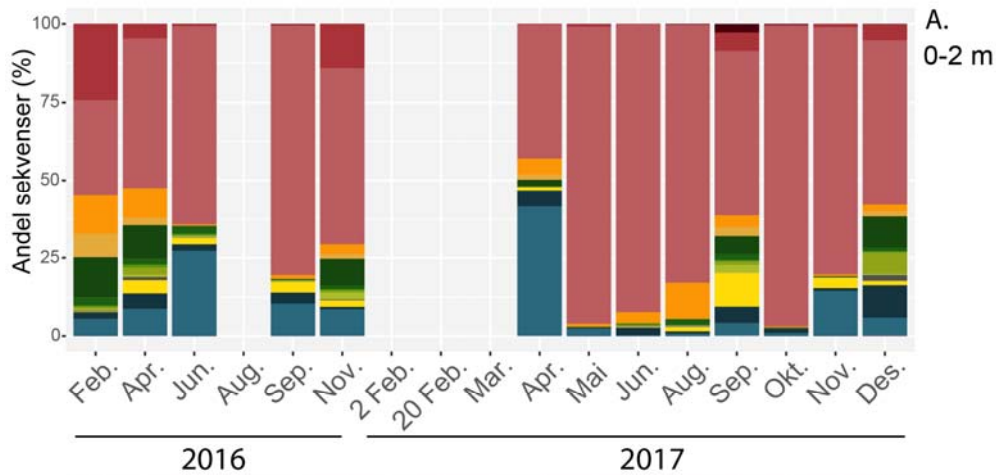
I rekken Rhodophyta (rødalger) ble det påvist 22 OTUer, hvor over halvparten viste lav prosent identitet (~90%) med referansesekvensene i PR² (OTU-tabell i figshare). Dette gjaldt alle OTUer klassifisert til klassene Compsopogonophyceae og Stylonematophyceae. De mindre abundante OTUene var mest til stede i april og august 2017, hvor fem ble med 100% identitet klassifisert til art innen klassene Florideophyceae (tabell S.8).

Opisthokonta

Opisthokonta var den supergruppen med gjennomgående lavest OTU-rikhet og andel sekvenser, ved begge dyp (figur 3.9 og figur 3.10). Innen rekken Choanoflagellida (krageflagellater) var det klassen Choanoflagellata som hadde flest sekvenser (figur 3.11). Denne klassen var tilstede gjennom hele prøvetakningsperioden ved 5 meters dyp, men uteble i overflaten i aug.-okt. 2017.

De mindre tallrike og abundante gruppene fulgte hverandre gjennom året. De så ut til å øke i andel sekvenser da Alveolatene avtok og minke da stramenopilene økte i abundans (figur 3.10).

► *Figur 3.10. Søylediagram og linjediagram over andel sekvenser (%). For hver prøve tatt gjennom 2016-2017 ved de to dypene (A) 0-2 meter og (B og C) 5 meter fremstilles andelen sekvenser i prosent fordelt på taksonomiske grupper/rekker. Fargene i hver søyle har lik rekkefølge som tegnforklaringen. Der det mangler søyle (A), er det ikke tatt prøver ved dette dypet (0-2 meter) for den gjeldende datoen. Basert på subsamplet protistdatasett fordelt på dyp.*



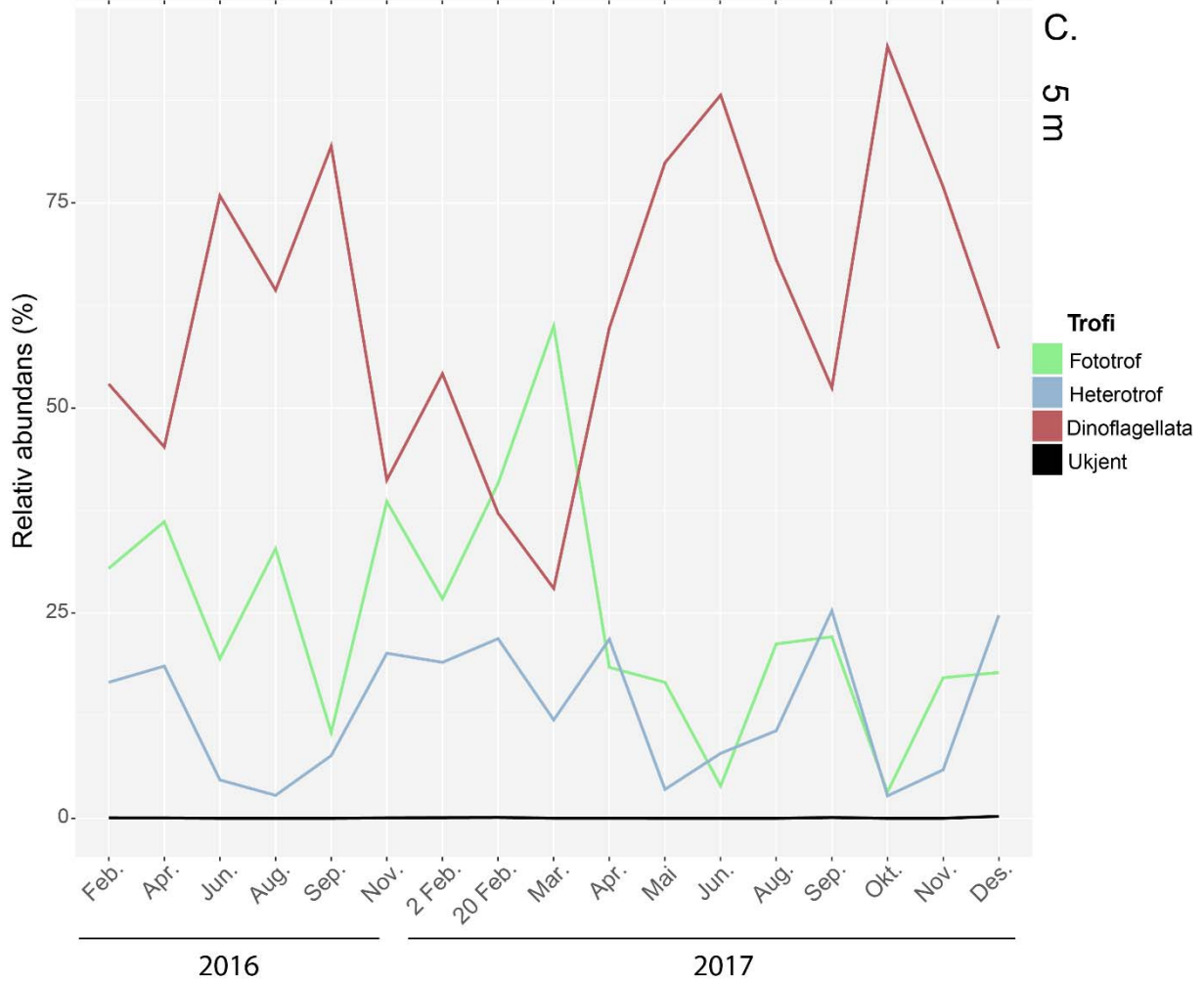
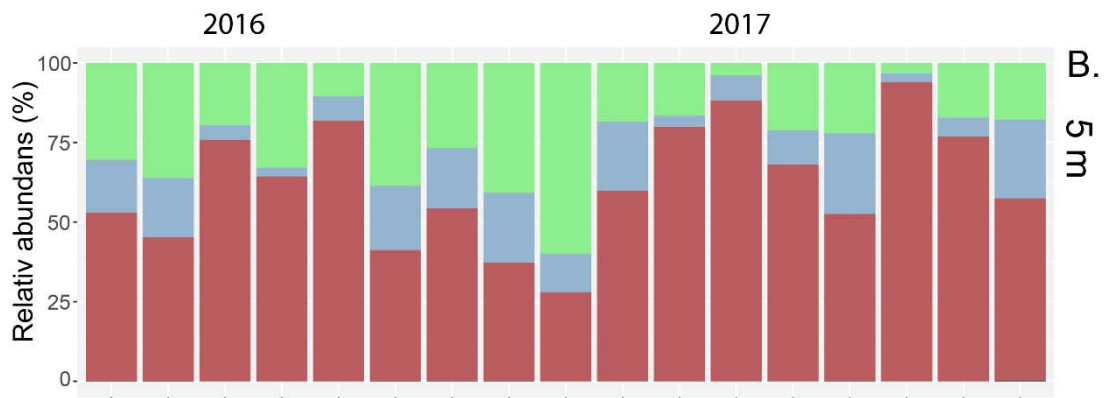
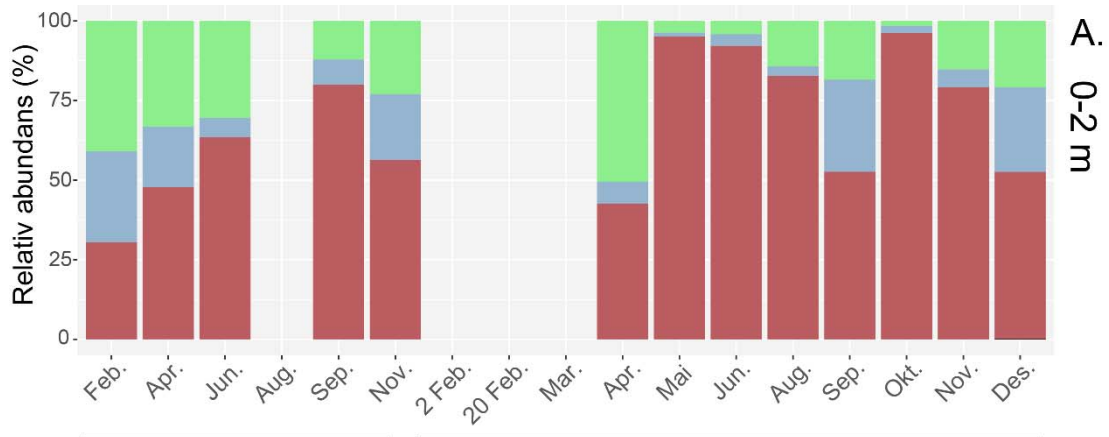
◀ *Figur 3.11. Hver rad i varmekartet representerer en taksonomisk klasse (tilegnet av PR² og korrigeret mot ALGAEBASE). Klassene er rangert i supergruppe, utenom uklassifiserte eukaryoter og klassene innen Apusozoa og Amoebozoa. Alle klassene som inngår i samme gruppe har lik farge på klasse-navnet, for noen er gruppenavn oppført. Første celle i hver rad (i A og i B) er klassens prosentandel i datasettet ved det dypet. Prosentverdier er oppgitt i hver celle, med to desimaler. For prosentverdier under 0.005 står dette oppført som 0. Mørkere farge i cellen indikerer høyere prosentverdi og i blanke celler er klassen fraværende, prosentenverdi 0. Basert på subsamplet protistdatasett fordelt på dyp.*

Sesongdynamikk hos funksjonelle grupper

OTUene ble klassifisert etter trofi basert på deres taksonomiske tilhørighet (figur 3.12). De heterotrofe består av Apicomplexa, klassen Syndiniophyceae i rekken Dinoflagellata, Amoebozoa, Apusozoa, Centroheliozoa, Katablepharidophyta, Picozoa, Telonemia, Choanoflagellida, Mesomycetozoa, Cercozoa, Radiolaria og de heterotrofe ikke-ochrophyte stramenopilene. De fototrofe er medlemmer av Chlorophyta, Rhodophyta, Chryptophyta, Haptophyta, Ochrophyta. Dinoflagellatene blir gruppert for seg siden de består av fototrofe, heterotrofe og miksotrofe taksa uten om den heterotrofe klassen Syndiniophyceae. Det ble ikke tatt hensyn til at det finnes unntak innen noen fototrofe grupper der medlemmer har mistet sin evne til å drive fotosyntese gjennom evolusjon eller er miksotrofe, siden dette kan variere på slekts-nivå og lite er kjent om omfanget.

Dinoflagellatene dominerte med høyest relativ abundans hele prøvetakningsperioden, ved begge dyp, for uten om i overflaten i feb. 2016 og april 2017 samt ved 5 m i mars 2017 (figur 3.12). De fototrofe gruppene hadde høy tetthet i vår-månedene og dominerte i overflaten i februar 2016 (figur 3.12.A) og vår-månedene i 2017 ved 5m (figur 3.12.B og C). Andelen fototrofe var relativt lik ved hvert dyp i prøvetakningsperioden og hadde høyest relativ tetthet i månedene hvor dinoflagellatene avtok samtidig som den heterotrofe andelen var lav. Økning av heterotrofe viste i 2017 et sterkt samspill med andelen fototrofe, de øker i mengden relative abundans i perioden etter det har vært en høy andel fototrofe (figur 3.12.C). Andelen heterotrofe er lav når andelen dinoflagellater er høy.

▶ *Figur 3.12. Søylediagram over relative abundans fordelt på trofisk nivå (i prosent %). For hver prøve tatt gjennom 2016-2017 ved de to dypene (A) 0-2 meter og (B og C) 5 meter fremstilles den relative abundansen (%) for to trofinivåer, fototrof og heterotrof, samt dinoflagellater og ukjente. Over de datoene de mangler en søyle, er det ved denne datoen ikke tatt prøver ved dette dypet (0-2 meter). Basert på subsamplet protistdatasett fordelt på dyp.*



3.3 Dinoflagellat-taksa

3.3.1 De mest abundant og frekvente OTUene i Dinoflagellata

I det fylogenetiske treet fremstilles de mest abundante og frekvente dinoflagellat-OTUene i protistdatasettet (figur 3.13. For fullstendig tre som inkluderer utgruppen se figur S.2). I treet ble 12 av 17 OTUer plassert sammen med kjente grupper og de resterende 5 OTUene ble plassert i hver sin monofyletiske gruppe med miljøsekvenser, som ble tildelt et gruppenavn. Resultatene under burde leses sammen med tabell S.5.

Av figur 3.13.A ser vi at **OTU 638** plasseres sammen med to andre miljøsekvenser, samlet har de blitt gitt navnet Gruppe 1. En av sekvensen funnet i NCBI var identiske regeransesekvensen tildelt OTU 638 av PR² (tabell S.5). PR² ga litt mer opplysninger enn NCBI, da den klassifiserte sekvensen som en sekvens innen klassen Dinophyceae, ikke kun som miljøsekvens. Ser vi videre på resultatet fra Blast-analyse mot NCBI-databasen der miljøsekvenser ble utelukket, var en sekvens av *Prorocentrum* nærmeste treff. OTU 638 er ikke en *Prorocentrum* ifølge plasseringen i treet, men det kan heller se ut til at Gruppe 1 er en basal søstergruppe til resten av dinoflagellatene. Siden grenens støtteverdi er <65, er Gruppe 1 sin plassering i treet usikker.

OTU 733 ble passert i Gruppe 2 sammen med to andre miljøsekvenser. Denne gruppen ser ut til å være en søstergruppe til orden Peridinales, men kan også være en del av Peridinales. Den ene sekvensen med best Blast-treff i NCBI da miljøsekvenser ble ekskludert, tilhørte *Podolampas spinifera* som er innen ordenen Peridinales (tabell S.5). Den andre sekvensen derimot, som hadde en like stor likhet med OTU 733, var en sekvens innen orden Gymnodinales. PR² viste at OTU 733 hørte til klassen Dinophyceae. Utifra det overnevnte kan man ikke med sikkerhet klassifisere OTUen i en taksonomisk orden.

Vi kan med sikkerhet si at **OTU 4** tilhører *Heterocapsa rotundata*, med god støtte av Blast-analyse mot PR² og NCBI, samt plassering i treet.

OTU 20 ble plassert i Gruppe 3 sammen med to andre miljøsekvenser. OTU 20 er identisk med en av miljøsekvensene som ble nærmeste treff med Blast-søk mot NCBI. Det er ikke inkludert noen referansesekvenser fra søket der miljøsekvenser ble utelatt, da treffene viste <90% identitet. Gruppe 3 står som søstergruppe til Gonyaulacales, med lave støtteverdier. Grenens plassering i treet kan derfor være tilfeldig. Gruppe 3 har en lang gren, som kan tyde på at den er svært forskjellig fra andre dinoflagellater inkludert i analysen.

OTU 38 ble plassert i Gruppe 4, sammen med en annen miljøsekvens, som var identisk referansesekvensen tildelt av PR². Det nærmeste Blast-treffet i NCBI ble ved en feil ikke inkludert i referansesekvensene til det fylogenetiske treet. OTU 38 sin nærmeste slektning i treet ble *Spiniferites ramosus*, med 100% støtte.

OTU 103 og **OTU 3** ble begge klassifisert som *Tripos*-arter. OTU 103 var mest sannsynlig fra *Tripos fusus*, som ble støttet av PR² og NCBI. Ved visuell inspeksjon av et MAFFT-alignment med OTU 3 og to *Tripos*-sekvenser, var OTU 3 100% lik *T. muelleri* og skilte seg fra *T. candelabrus* med >40 bp. Dette indikerer at OTU 3 var fra *Tripos muelleri*, ikke fra *Tripos candelabrus* som ble nærmeste Blast-treff mot PR².

OTU 8 ble plassert til *Alexandrium hiranoi* i treet, som ble støttet av Blast-treff mot PR² og NCBI. Siden dette er en art som ikke tidligere er påvist i norskefarvann, ble det laget et MAFFT-alignment med sekvensen til *Alexandrium andersonii* og sekvensen til *Alexandrium pseudogonyaulax*, som er vanlig arter i Norge. OTU 8, *A. hiranoi* og *A. pseudogonyaulax* var 100 % identiske i V4-regionen og *A. andersonii* skilte seg fra disse med 28 bp.

OTU 37 ble plassert som *Alexandrium margaelefii*, som støttes av Blast-treff mot PR² og NCBI. Ved visuell inspeksjon av MAFFT-alignmentet med disse to sekvensene, skilte OTU 37 seg fra *A. margaelefii* med 2 bp.

OTU 22 ble klassifisert som *Dinophysis acuminata* i PR², og ble i treet plassert midt i ordenen Dinophysiales (figur 3.13.B). Et MAFFT-alignment viste at de tre *Dinophysis*-artene *D. acuta*, *D. acuminata* og *D. norvegica* er 100% identiske og OTU 22 skilte seg fra disse med ett bp helt i starten av sekvensen. På grunnlag av det overnevnte kan OTU 22 med god sikkerhet klassifiseres i slekten *Dinophysis*, men siden disse tre artene er identiske i V4-regionen kan OTU 22 ikke klassifiseres til art.

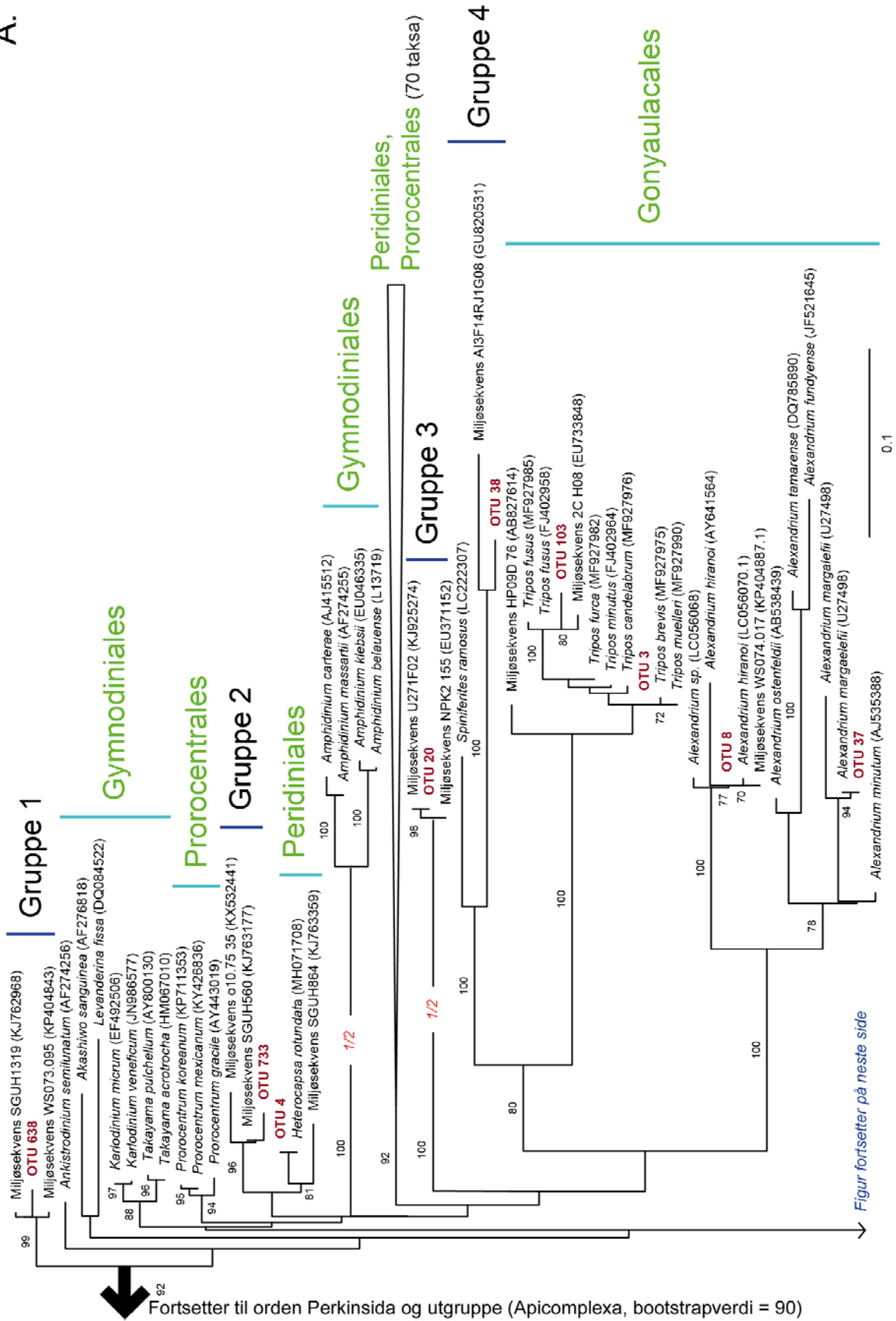
OTU 14, **OTU 639** og **OTU 10** ble plassert i ordenen Gymnodiniales. OTU 14 passerer sammen med en annen miljøsekvens, hvor nærmeste slektning ble en *Warnowia*-art. OTU 639 ble plassert i monofyletisk gruppe med to andre miljøsekvenser og ble av PR² klassifisert som en *Gymnodinium*-art, som i treet er nærmeste søstergren. OTU 10 tilhører med god støtte *Gymnodinium beii* (H. J. Spero) synonym til *Pelagodinium bei*.

OTU 51 ble plassert i Gruppe 5, med den lengste grenen i treet (figur 3.13.C). Av den grunn ble det mistenkt at sekvensen til OTU 51 var en kimær. Den ble visuelt inspisert i et MAFFT-

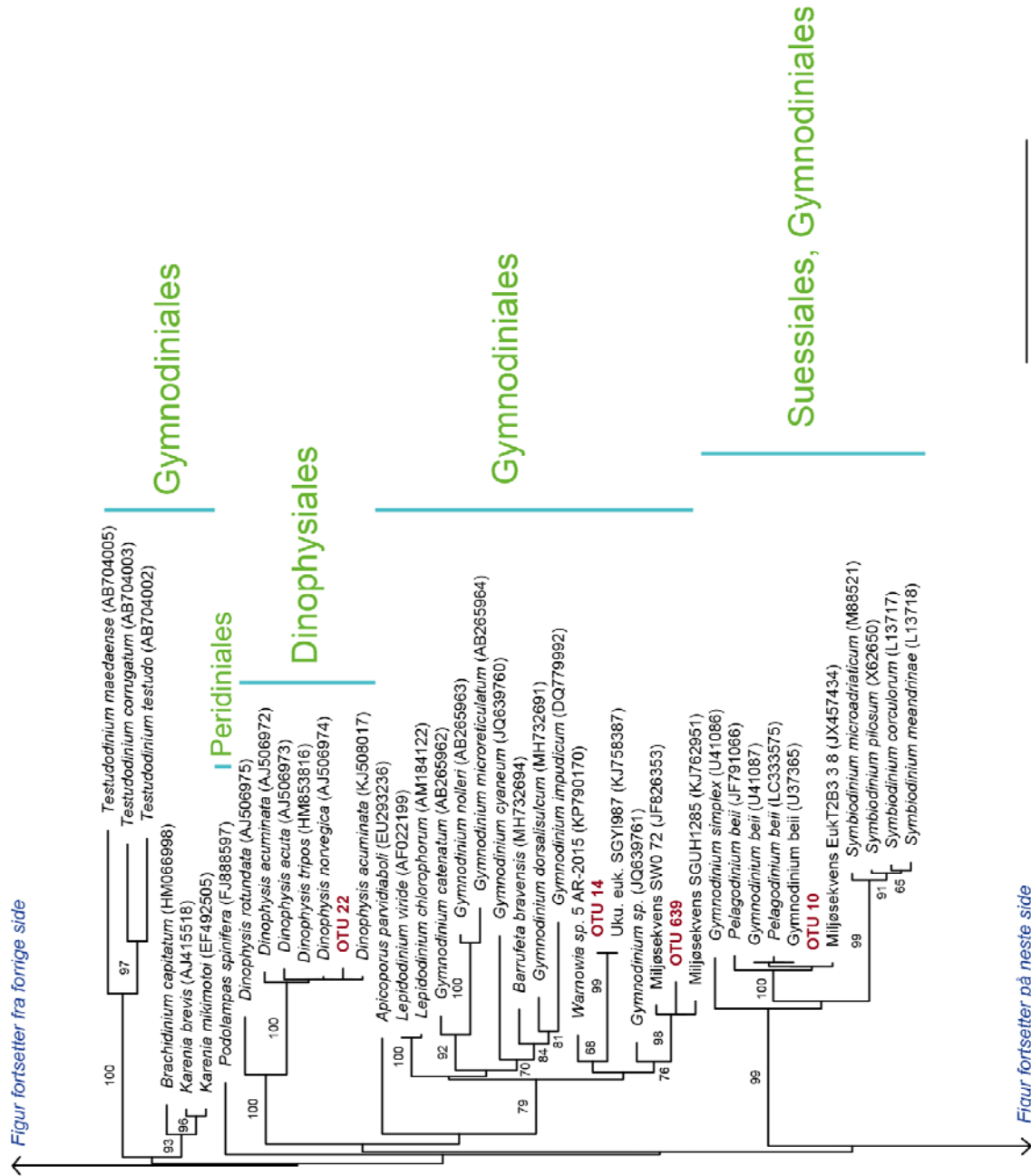
alignment mot andre dinoflagellat-arter og deler av sekvensen ble hver for seg søkt opp mot NCBI. Det var ingenting som tydet på at OUT 51 var en kimær. Dette gjør denne gruppen veldig spennende, da det kan være en gren innen dinoflagellatene som ikke er genetisk karakterisert enda.

OTU 30, OTU 29 og **OTU 1033** ble alle klassifisert til *Gyrodinium*-arter. OTU 30 ble plassert som *G. gutrula*. OTU 29 var den OTUen i protistdatasettet med flest sekvenser, og interessant nok ble plassert i treet med ferskvanns-arten *Gyrodinium helveticum* som nærmeste slektning. Et MAFFT-alignment viste at OTU 8 var 6 bp forskjellig fra de to *G. helveticum*-sekvensene som var nærmeste Blast-treff i NCBI (tabell S.5). OTU 1033 ble plassert sammen med en annen miljøsevens, med *G. fusiforme* som nærmeste slektning.

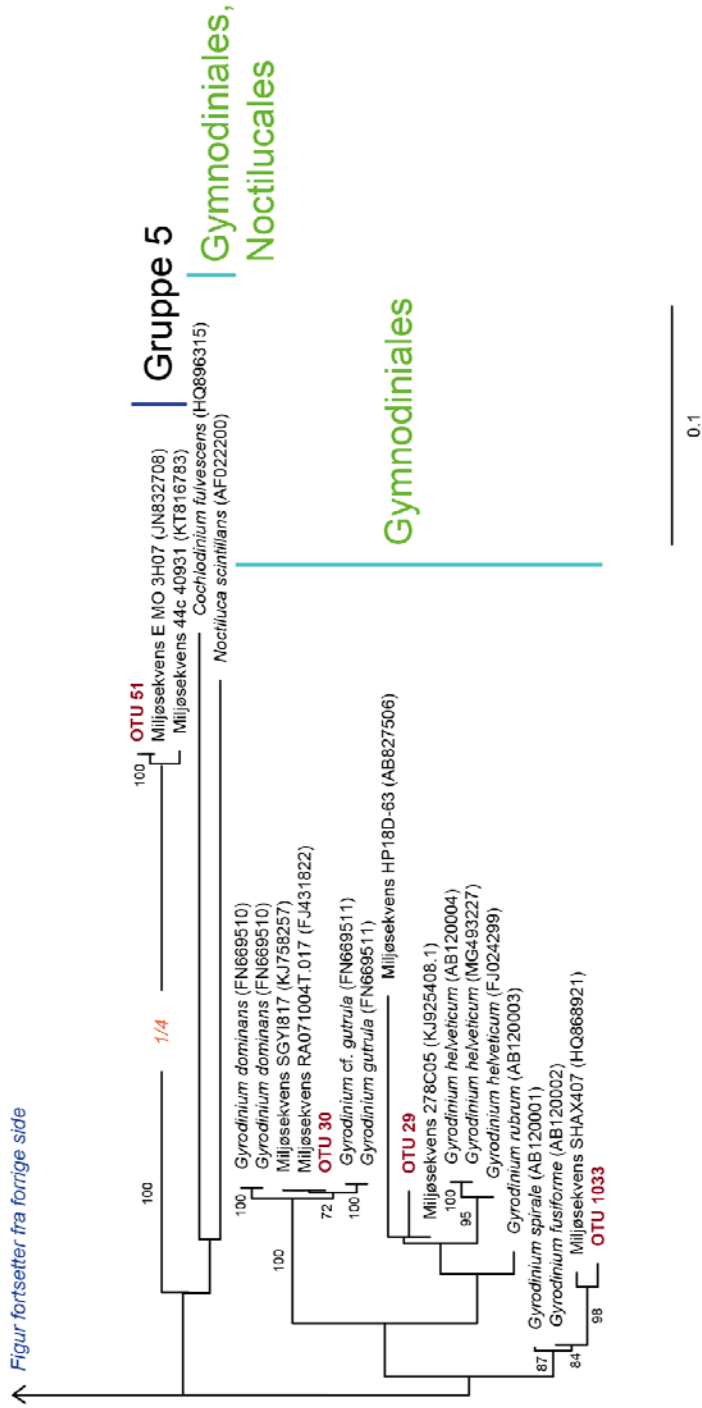
A.



Fortsetter til orden Perkinsida og utgruppe (Apicomplexa, bootstrapverdi = 90)



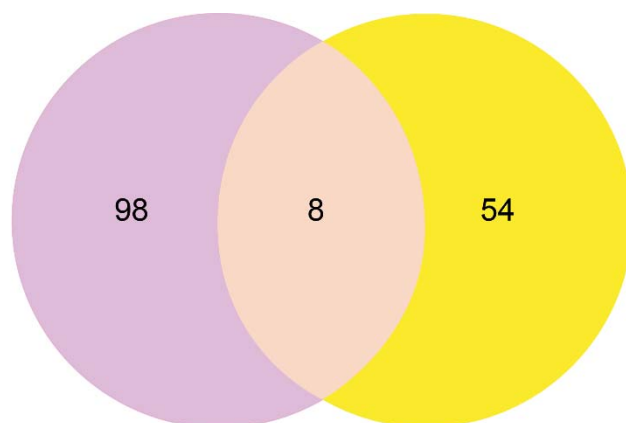
C.



Figur 3.13. Fylogenetisk tre basert på 18S rDNA over de mest abundante og frekvente *Dionoflagellat* OTUer i Protist- og subsamplet datasett. Treet er konstruert med RAxML i superdatamaskinen Albel. Tallene forbundet med nodene er bootstrappverdier for RAxML. OTUene er uthevet mørkrøde, bootstrappverdier under 65% vises ikke. Hvilken orden de monofyletiske grenene tilhører er markert i grønn skrift. Legg merke til sort tykk pil øverst i treet som viser til den ekskluderte søsterggruppen *Perkinsida* (bootstrappverdi = 92) og utgruppen *Apicomplexa* (bootstrappverdi = 90). Grener som er forkortet er markert med hvilket forhold de er forkortet (brøk i rødt). Oppmålt avstand (0.1) representerer genetisk avstand på 0.1 basepar.

3.3.2 Dinoflagellat-taksa sammenlignet med lysmikroskopfunn

De to metodene påviste svært forskjellige dinoflagellat-taksa (figur 3.14 og tabell S.6). I protistdatasettet ble det påvist 567 unike dinoflagellat-OTUer (tabell S.4). Av disse var 401 OTUer tilstedet i prøvene tatt på 5 m i 2017. Disse ble taksonomisk klassifisert til 106 unike dinoflagellat-taksa (OTU-tabell i figshare).



Unike dinoflagellat-taksa: Totalt = 160
■ Metastrekkoding = 106 ■ Lysmikroskop = 62 ■ Felles = 8

Tilnærmet dobbelt så mange taksa ble detektert med metastrekkoding som med lysmikroskopi (figur 3.14 og tabell S.6). Det er få taksa som er blitt registrert med begge metoder,

kun åtte arter (*Dinophysis acuminata*, *Tripos furca*, *Tripos fusus*, *Akashiwo sanguinea*, *Gyrodinium fusiforme*, *Protoperdinium depressum*, *Prorocentrum triestinum* og "ubestemt dinoflagellat").

Med lysmikroskop var det 9 taksa (tekate- og atekate dinoflagellater) som ble registrert som ubestemte etter størrelsesorden (μm) og 41 taksa som med metastekking ble registrert som ukjente innen klader i dino-gruppe 1 og 2 i ordenen Syndiniales. Metastrekkoding fanget opp 32 slekter som ikke ble registrert med lysmikroskop. Av disse var 9 av slektene innen ordenen Gymnodiniales. Det ble registrert 7 slekter med lysmikroskop som ikke ble registrert med med metastrekkoding. Det ble påvist 6 *Dinophysis*-arter med lysmikroskop, hvor kun én av disse ble påvist i metastrekkoding-OTUene klassifisert med PR² (tabell S.6).

Dinoflagellat-arter

Ved å fjerne alle taksa som ikke ble taksonomisk klassifisert til art fra metastrekkoding- og lysmikroskop-taksalisten (tabell S.6) halveres det totale antallet unike taksa (figur 3.15). Taksa defineres heretter som en art. Mange av metastrekkoding-OTUene har blitt klassifisert med lav

Figur 3.14. Venndiagram over Dinoflagellat-taksa detektert med metastrekkoding og lysmikroskop i 2017, innsamlet ved stasjon DK1 fra 5 meter. Metastrekkodingtall er basert på tall fra protistdatasett og lysmikroskoptall er publisert i rapporten *Overvåkning med FerryBox* (Norli & Eikrem, 2018).

prosent identitet mot referansesekvenser i referansedatabasen PR² (<99%). Det ble fjernet 30 taksa som var påvist med metastrekkoding på grunn av usikker klassifisering.

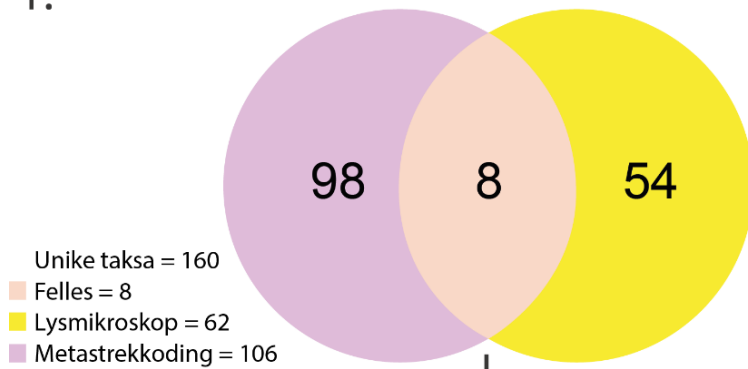
Fordelingen av taksa identifisert med de to metodene er nå jevnere og det er registrert 3 flere taksa med lysmikroskop (punkt 3 i figur 3.15). Ett taksa som var felles i den opprinnelige sammenligningen ble ikke påvist med metastrekkoding med høy nok identitet mot referansesekvensen. Derfor ser vi en økning i ett taksa kun påvist med lysmikroskop mellom punkt 2 og 3 i Figur 3.15. En siste gjennomgang av taksalistene innebar sammenligning av vitenskapelig artsnavn og synonymer. Dette var for å se om det er flere taksa som er felles for begge metoder, men skjult i bruk av ulike artsnavn. Artsnavn i taksalistene ble søkt opp ALGAEBASE hvor synonymer er listet (tabell S.10). Dette resulterte i to ekstra felles taksa siden *Lingulodinium polyedrum* er synonym med *Lingulodinium polyedra* og *Phalacroma rotundatum* er synonym med *Dinophysis rotundata*. *Tripos candelabrus* som påvist med metastrekkoding ble i et alignment-sjekk omklassifisert til *Tripos muelleri*, som resulterte i totalt 8 felles taksa for de to metodene.

Det var 22 unike dinoflagellat-taksa kun observert i lysmikroskop, hvor flesteparten enten var *Dinophysis*-arter, *Tripos*-arter eller *Protoperdinium* (tabell S.11).

Med metastrekkoding ble det påvist OTUer som med stor sikkerhet ble klassifisert til 18 taksa fordelt på 16 forskjellige slekter, som ikke ble observert i lysmikroskop (tabell S.10). Ved en gjennomgang av artene påvist med metastrekkoding ble det vist at omtrent halvparten av artene er nakne dinoflagellater, noen kan danne cyster/hvilesporere, en var parasitt og en var mindre enn 20 µm (tabell S.10).

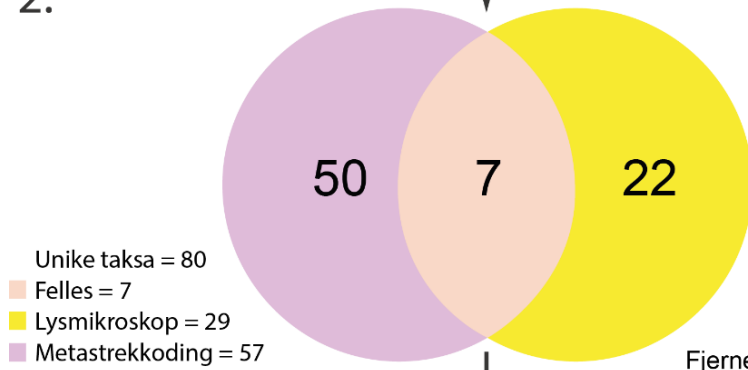
► *Figur 3.15. Sammenligning av dinoflagellat-taksa. Metastrekkoding-taksaliste og lysmikroskop-taksaliste ble modifisert fra (1) opprinnelig sammenligning av alle taksa ved å (2) fjerne taksa som ikke ble klassifisert til art, (3) beholde taksa som ble klassifisert med stor sikkerhet og (4) kontrollere artsnavn for synonymer mot algaebase.org og korrigere listene mot hverandre. Venndiagram (1,2,3, og 4) fremstiller fordelingen av taksa i hver liste i hvert trinn av sammenligningen.*

1.



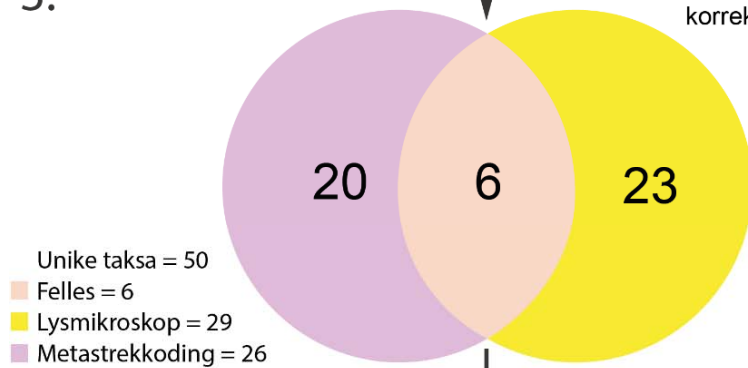
Fjerner alle taksa som ikke ble klassifisert til art.

2.



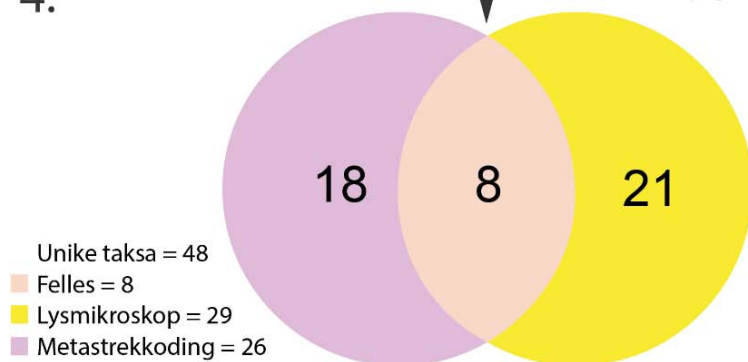
Fjerner alle metastrekkoding-taksa som hadde <99% identitet med referansesekvens i PR².
Antar at alle lysmikroskop-taksa er 100% korrekt klassifisert.

3.



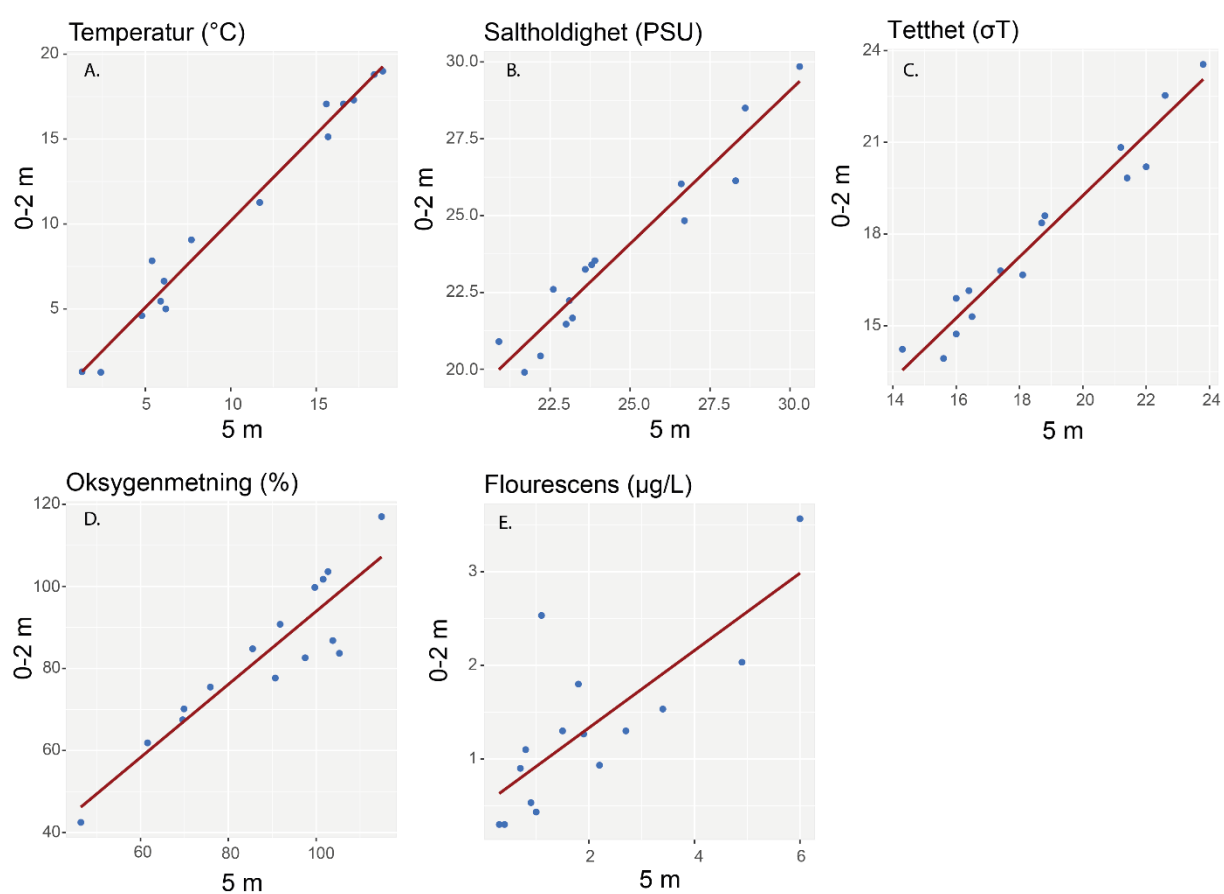
Kontrollerer begge taksa-lister for synonymer (algaebase.org) og korrigerer mot hverandre.

4.



3.4 Forskjellen mellom dypene

Alle miljøvariablene viste en signifikant lineær sammenheng mellom de to dypene (p-verdi <0.0001) (figur 3.16 og tabell 3.4). Temperaturen ved de to dypene var tilnærmet perfekt lineær ($y = x$). Utfra likningene til saltholdighet og tetthet (likning B og C i tabell 3.4), ser man at stigningstallet er <1, som vil si at forskjellen mellom dypene blir større når verdiene for målingene øker. Oksygenmetningen var gjennomgående lavere ved 0-2 m enn ved 5 m. Det samme gjelder for fluorescens.



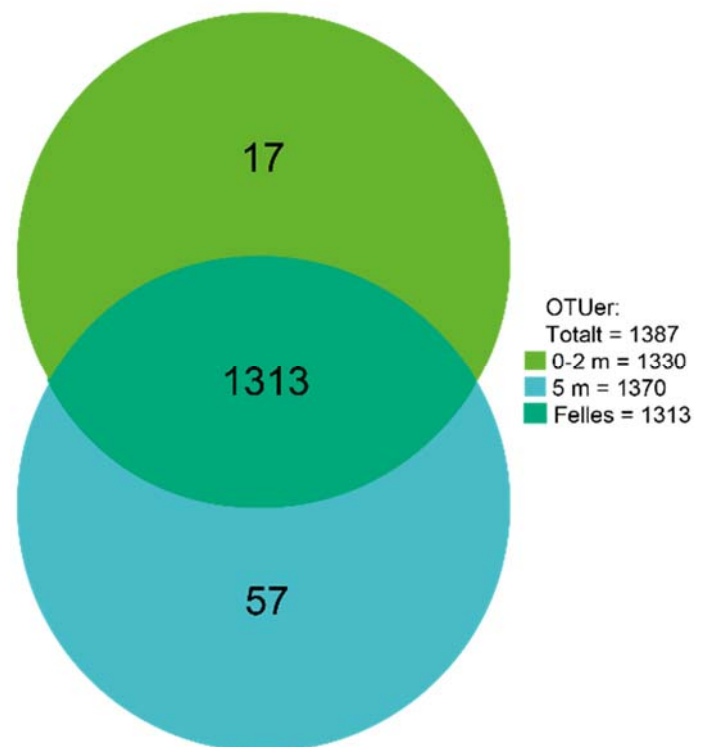
Figur 3.16. Lineær regresjonsanalyse. Miljøvariablene (A) temperatur, (B) saltholdighet, (C) tetthet, (D) oksygenmetning og (E) fluorescens, målt gjennom prøvetakingsperioden ved stasjon DK1 ved de to dypene 0-2 meter og 5 meter.

Tabell 3.4. Målingene for hver av miljøvariablene ((A) temperatur, (B) saltholdighet, (C) tetthet, (D) oksygenmetning og (E) fluorescens) fra dypene 0-2 meter og 5 meter ble sammenlignet med linærregresjonsanalyse. Alle viser en signifikant lineær sammenheng (p -verdi >0.0001).

Miljøvariabel	R ²	p-verdi	Likning
A. Temperatur	0.9771	4.849e-12	$y = 1.02x - 0.02$
B. Saltholdighet	0.9303	6.772e-09	$y = 0.99x - 0.87$
C. Tetthet	0.9525	5.514e-10	$y = x - 0.76$
D. Oksygenmetning	0.8407	1.522e-06	$y = 0.89x + 4.93$
E. Fluorescens	0.5923	0.0007926	$y = 0.41x + 0.51$

OTUene

Det var 1387 unike OTUer totalt i protistdatasettet. Av disse var 1313 OTUer (95%) felles for de to dypene, 0-2 m og 5 m (figur 3.17). Det var 17 OTUer som var særegne i overflaten, hvor 8 av disse ble klassifisert til art av PR² (tabell 3.5.A). Ved 5 m var det 57 OTUer som ikke var tilstede i overflaten, av disse ble 6 bestemt til art (tabell 3.5.B). Av de særegne OTUene bestemt til art, er det kun 2 av artene der klassifiseringen har relativt god støtte ($>99\%$ identitet med referansesekvens), ciliaten *Pseudocohnilembus persalinus* og dinoflagellaten *Oxyrrhis marina*, begge var kun i overflaten (tabell 3.5).



Figur 3.17. Venndiagram som viser hvordan OTUene i protistdatasettet fordeler seg på de to dypene 0-2 meter (grønn) og 5 meter (blå) og hvor mange OTUer som er felles for begge dyp (turkis).

Tabell 3.5. OTUer klassifisert til art (av PR²) som var unike ved (A) 0-2 meters dyp og (B) 5 meters dyp ved DK1, gjennom prøvetakningsperioden fra feb. 2016 til des. 2017. IDENT. = prosent likhet mellom OTU-sekvens og referansesekvens i PR². "QUERY COVER" = prosent av OTU-sekvens søkt opp mot referansesekvens. Arter markert i gult er ikke unike arter ved et av dydene siden flere OTUer i protistdatasettet også ble tilegnet samme art av PR², og disse var tilstedet ved begge dyp.

A.

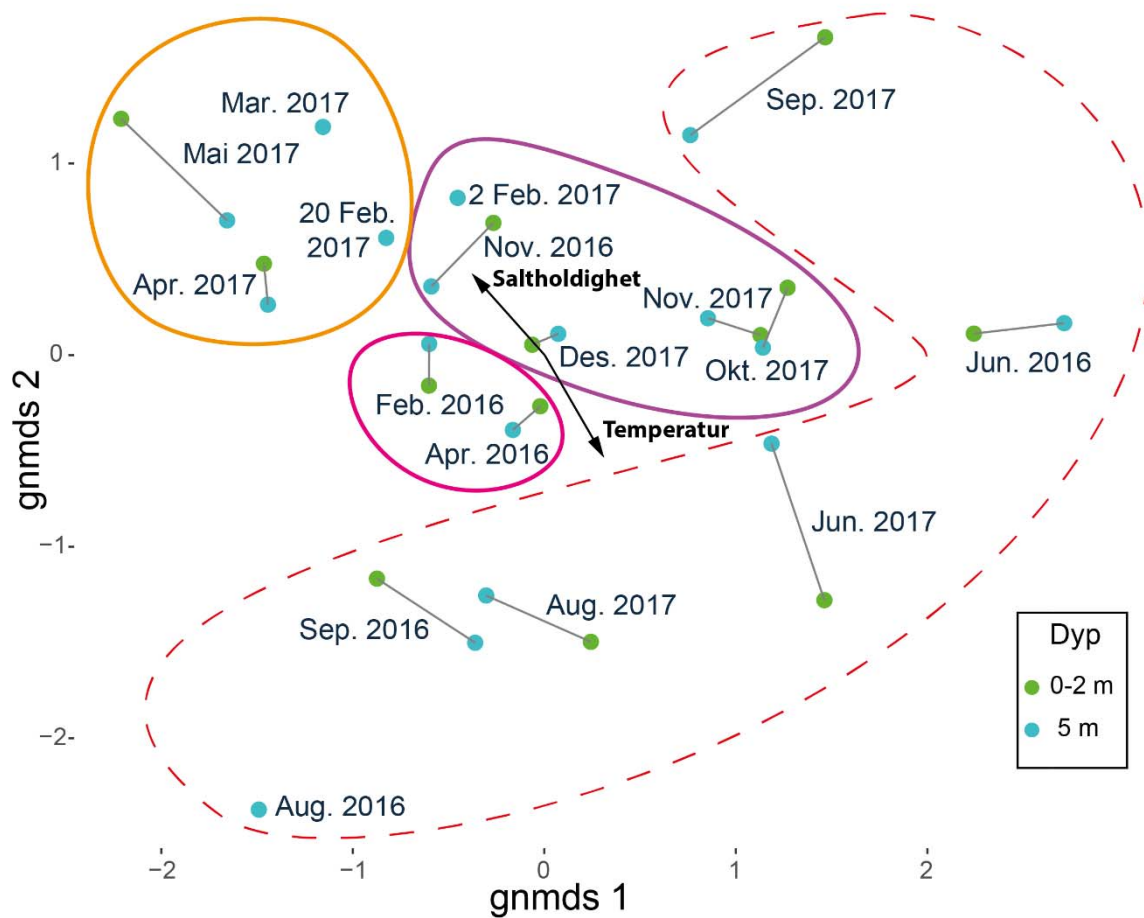
OTU	SUPERGRUPPE	GRUPPE	ART	IDENT.	BASEPAR AVVIK	"QUERY COVER"
1528	Alveolata	Apicomplexa	<i>Colpodella tetrahymenae</i>	96.1 %	10	100 %
290		Ciliophora	<i>Pseudocohnilembus persalinus</i>	99.5 %	2	100 %
2036		Dinoflagellata	<i>Gymnodinium dorsalisulcum</i>	98.1 %	8	100 %
2031			<i>Tripes candelabrus</i>	95.2 %	20	100 %
918			<i>Oxyrrhis marina</i>	99.8 %	1	100 %
1016	Archaeplastida	Chlorophyta	<i>Hamakko caudatus</i>	98.8 %	5	100 %
1296	Stramenopiles	Ochrophyta	<i>Haslea crucigera</i>	98.6 %	6	100 %
1658			<i>Hydrurus foetidus</i>	98.4 %	7	100 %

B.

OTU	SUPERGRUPPE	GRUPPE	ART	IDENT.	BASEPAR AVVIK	"QUERY COVER"
651	Alveolata	Dinoflagellata	<i>Amphidiniopsis dragescoi</i>	96.2 %	16	100 %
1802			<i>Blastodinium mangini</i>	96.9 %	9	99.5 %
1089	Archaeplastida	Chlorophyta	<i>Tetracystis pampae</i>	87.7 %	27	53.4 %
890	Rhizaria	Cercozoa	<i>Reckertia filosa</i>	98.4 %	7	100 %
618	Stramenopiles	Ochrophyta	<i>Chaetoceros calcitrans</i>	90.8 %	25	98.8 %
2554		Labyrinthomorpha	<i>Labyrinthuloides haliotidis</i>	98.8 %	0	98.7 %

Ordinasjon

Ordinasjonen viser at prøver tatt samme dato ved de to dypene, er relativt like (figur 3.18). Dette ble underøkt ved å benytte ordinasjon av typen MPO. Første akse i DCA og GNMDS viste sterk korrelasjon, τ -verdi: 0.64, p-verdi <0.0001. Andre aksene viste en svakere korrelasjon, τ -verdi: 0.35, p-verdi = 0.006. Siden DCA-plottet viste tydelig tunge-effekt (figur S.1), ble det valgt å kun analysere og diskutere GNMDS-ordinasjonen.

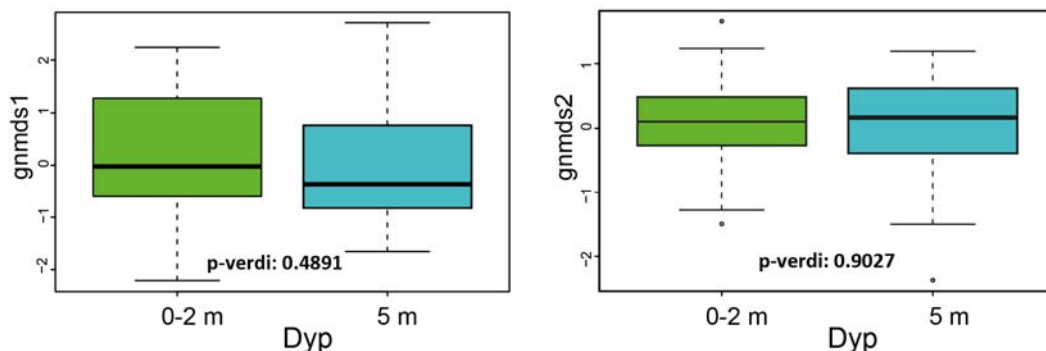


Figur 3.18. Ordinasjonsplot (gnm1 og gnm2) basert på subsamplet protistdatasettet som viser prøvenes plassering langs ordinasjonsaksene (gnm1 og gnm2). Prøvenes farge forteller ved hvilket dyp de er tatt, 0-2 meter (grønn) og 5 meter (blå). Prøvetakningsdatoen for hver prøve er oppgitt og prøver tatt på samme dato er knyttet sammen med en grå linje. Der prøven ikke er knyttet til en annen, er det kun tatt prøver ved ett dyp på den gitte dato. Miljøvariablene er fremstilt med sorte piler. Gruppering etter sesong: (Rosa) vårprøver 2016, (orange) vårprøver 2017, (rød striplet) sommer-høstprøver og (lilla) vinterprøver, inkludert oktober 2017.

Shepards-diagram (figur S.4) viser liten spredning om linjen som indikerer at den opprinnelige ulikheten i datasettet blir godt preservert når datasettets dimensjoner reduseres til to i ordinasjonen.

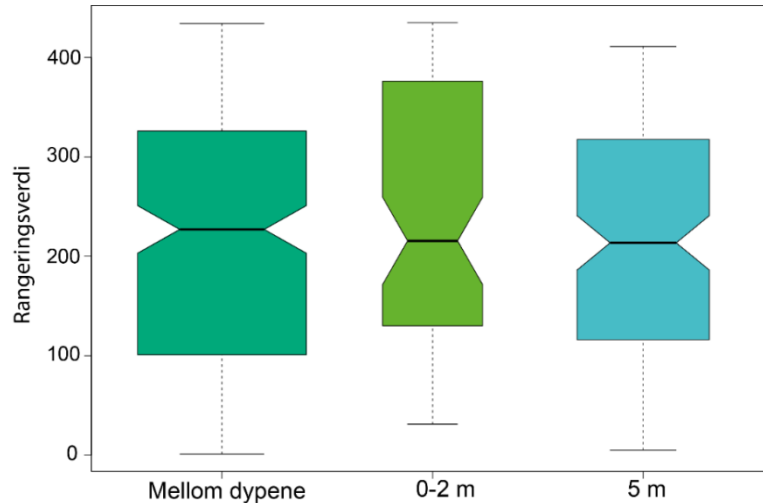
Prøver tatt samme dato ved hvert sitt dyp ble plassert i nærhet til hverandre i ordinasjonsplottet (figur 3.18). Dette indikerer at disse prøvene har protistsamfunn med en lik sammensetning av OTU-diversitet og abundans. Prøver tatt på samme årstid dannet distinkte grupper. Vårprøvene (feb.-apr.) i 2016, ble plassert mellom -1 og 0 på ordinasjonsakse 1 og 2. Vårprøvene (20. feb-mai) i 2017 ble plassert med lavest verdi på ordinasjonsakse 1 og høyest verdi på ordinasjonsakse 2. Høst-vinterprøvene (okt. til feb.) fikk spredning rundt midtverdiene på ordinasjonsakse 1 og 2. Sommer-høstprøvene begge årene viste en stor spredning langs begge aksene, rundt de andre sesong-grupperingene.

Miljøvariablene temperatur og saltholdighet forklarte prøvenes plassering i like stor grad, men i motsatt retning. Økt temperatur forklarte prøvene plassert med høy verdi på ordinasjonsakse 1 og lav verdi på ordinasjonsakse 2. Økt salinitet forklarte prøver plassert med lav verdi på ordinasjonsakse 1 og høy verdi på ordinasjonsakse 2.



Figur 3.19. Boksdiagram representerer prøvenes spredning langs ordinasjonsakse 1 (venstre) og ordinasjonsakse 2 (høyre) for hvert dyp, 0-2 meter og 5 meter. Boksen går fra nedre første kvartil (25.persentil) til øvre tredje kvartil (75.persentil) og svart horisontal strek indikerer medianen (50.persentil). Striplet linje over og under boksen strekker seg til nedre og øvre verdi som ligger innenfor halvannen interkvartilavstand. Punkter utenfor dette kalles uteliggere og representerer ekstremverdier. p-verdi fra Welch tosidig t-test er oppgitt i hvert diagram, som viser ingen signifikant forskjell mellom de to dypene. Basert på subsamplet protistdatasett.

Welch tosidig t-test viste ingen signifikant forskjell mellom prøvenes plassering langs ordinasjonsaksene (gnm1 og gnm2), som sier at det ikke er detekterbar variasjonen i prøvene mellom de to dypene (figur 3.19 og tabell S.9). Dette ble bekreftet i likehetsanalysen (ANOSIM) hvor $R = -0.013$ med p-verdi >0.5 (figur 3.20).



Figur 3.20. Boksdiagram for likhetsanalyse (ANOSIM), som viser ulikheter mellom og innad i hvert dypene 0-2 meter og 5 meter. Basert på rangeringsverdier tilegnet Bray-Curtis verdier i ulikhetsmatrisen. Boksen går fra nedre første kvartil (25.persentil) til øvre tredje kvartil (75.persentil) og svart horisontal strek indikerer medianen (50.persentil). Striplet linje over og under boksen strekker seg til nedre og øvre verdi som ligger innenfor halvannen interkvartilavstand. Teststastikken ($R = -0.013$ med p -verdi >0.5) viser ingen signifikant forskjell i prøvene fra hvert dyp.

Ordinasjonsavstand og sprangsjikt

Ved å rangere ordinasjonsavstanden i figur 3.18 fra størst til minst (figur S.7) og sammenligne med dypet til det øvre blandingslaget så vi ingen korrelasjon (tabell 3.6).

Tabell 3.6. Sammenligning av ordinasjonsavstand og dybden på det øvre blandingslaget for hver prøve tatt i prøvetakningsperioden fra februar 2016 til desember 2017. Ordinasjonsavstandene er rangert fra (1) størst til (12) minst, der avstand er synonymt med forskjell i protistsamfunnets diversitet og abundans mellom prøven tatt på 0-2meter og 5 meters dyp.

DATO	ORDINASJONSAVSTAND	ØVRE BLANDINGSLAG (METER)
SEP. 2017	1	10
JUN. 2017	2	3
MAI 2017	3	2
SEP. 2016	4	6
AUG. 2017	5	10
JUN. 2016	6	5
NOV. 2016	7	9
OKT. 2017	8	5
NOV. 2017	9	5
FEB. 2016	10	4.5
APR. 2017	11	4
APR. 2016	12	12
DES. 2017	13	6

Rikhet, diversitet og jevnhet

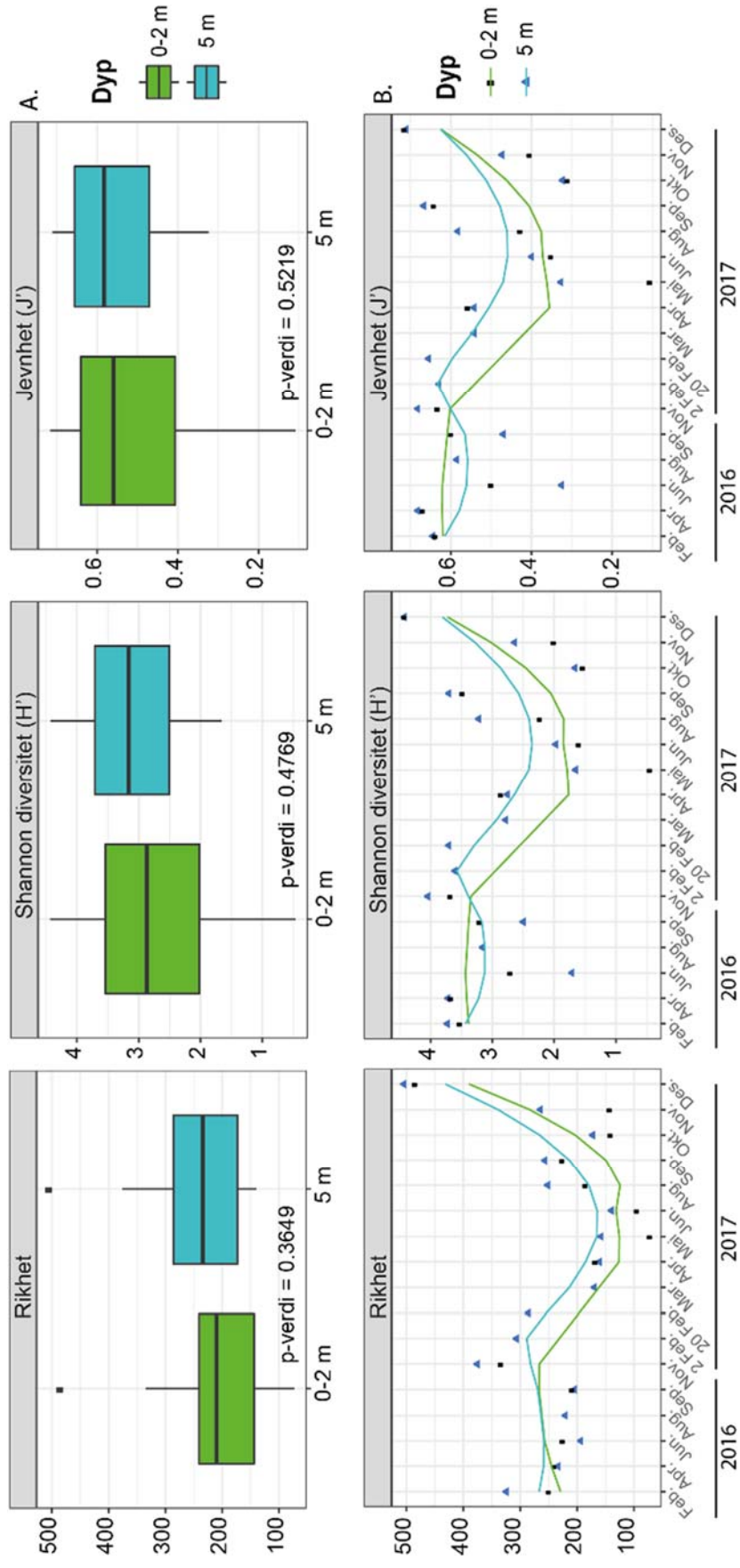
Indekser for rikhet, diversitet og jevnhet viser ingen signifikant forskjell mellom prøvene fra de to dypene (figur 3.21), siden indeksene mellom dypene er sterkt korrelert (tabell 3.7). Prøvene viser tydelig sesongvariasjon, spesielt hvis man ser isolert på 2017 (figur 3.21). Lavest rikhet, diversitet og jevnhet er i vår-sommerprøvene (april-august), ved begge dyp. Rikheten er synonymt med antallet OTUer, når det er lav rikhet er få OTUer tilstede i prøven. Om diversiteten i tillegg er lav, betyr det at noen OTUer har høy relativ abundans. Dette bekreftes av den lave jevnheten, som viser at mengden sekvenser er fordelt skjevt på de OTUene som er tilstede. Utover høsten øker verdiene for de tre indeksene og viser høyest verdi i vinterprøvene (desember-februar). Rikheten viser at et høyt antall forskjellig OTUer tilstede i vinterprøvene. Diversiteten i vinterprøvene er høy, da andelen sekvenser er fordelt jevnt på flere OTUer, som bekreftes av høy jevnhet. Dette gjelder for prøver tatt ved begge dyp. Rikhet, diversitet og jevnhet er korrelert med hverandrandre ved begge dyp (tabell 3.8).

Tabell 3.7. Korrelasjonstest mellom dypene. Indeksene rikhet, diversitet (Shannon, H') og jevnhet (Pielou, J'), for dypene 0-2 meter og 5 meter, ble testet mot hverandre med Kendalls korrelasjonstest. For hver test p-verdi og Kendalls tau (τ) oppgitt i tabellen.

		0-2 m					
		Rikhet		Diversitet (H')		Jevnhet (J')	
		p-verdi	τ	p-verdi	τ	p-verdi	τ
5 m	Rikhet	<0.001	0.67				
	Diversitet (H')			<0.001	0.79		
	Jevnhet (J')					<0.001	0.69

Tabell 3.8. Korrelasjonstest for hver av dypene. For hvert dyp, 0-2 meter og 5 meter, er de tre indeksene rikhet, diversitet (Shannon, H') og jevnhet (Pielou, J') testet mot hverandre med Kendalls korrelasjonstest. For hver test er p-verdier og Kendalls tau (τ) oppgitt i tabellen.

Dyp:	Indekser testet mot hverandre:		
		p-verdi	τ
0-2 m	Rikhet – Divesitet (H')	<0.001	0.88
	Rikhet – Jevnhet (J')	<0.001	0.77
	Diversitet (H') – Jevnhet (J')	<0.001	0.90
5 m	Rikhet – Divesitet (H')	<0.001	0.69
	Rikhet – Jevnhet (J')	<0.001	0.57
	Diversitet (H') – Jevnhet (J')	<0.001	0.88



Figur 3.21. Indekser for OTU-rikhet, diversitet (Shannon, H') og jevnhet (J') ved begge dyp. Øverst er boksdiagram med tilhørende p-verdier fra Welch tosidig t-test, nederst fremstilles indeks verdiene for hver enkelt prøve tatt gjennom 2016 og 2017, med tilpasset Generalisert Additiv Model (GAM) for å fremheve trender (heltrukket linje).

Indekser mot miljøvariabler

Det ble også undersøkt hvordan rikhet, diversitet og jevnhet forklares av miljøvariablene. Dette ble gjort ved å plote hver av indeksene mot en miljøvariabel av gangen. Det var kun temperatur som viste et interessant mønster, en parabel (figur S.6). Tross en tilsynelatende tydelig trend i diagrammet, fikk denne trenden ingen støtte da modellen ble testet mot en lineær- og en nullmodell (p-verdi >0.05). Dette gjaldt for alle indekser og miljøvariabler da de ble testet mot hverandre. Disse resultatene er ikke inkludert i oppgaven.

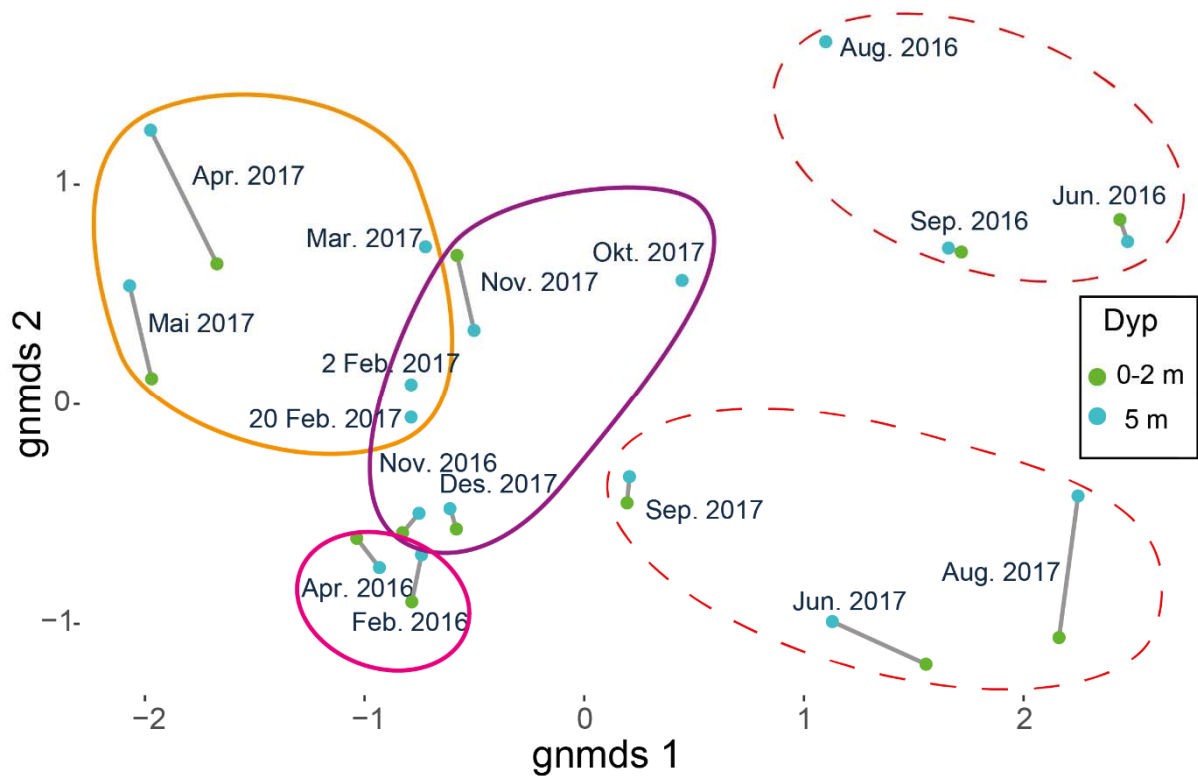
Ekskludering av dinoflagellater

Som et siste forsøk på å teste om det er forskjell i protistsamfunnet ved de to dypene ble det laget et nytt datasett der dinoflagellatene ble fjernet. Siden dinoflagellatene er så abundante kan det tenkes at de overskygger andre trender i samfunnet. Dette ble gjort ved å manuelt fjerne alle OTUer som tilhørte Dinoflagellata fra det originale datasettet, og deretter subsample dette etter den prøven med lavest antall sekvenser. Prøve nr. 25 (18.10.17, 0-2 m) ble utelukket, da denne hadde kun 824, som er 1/6 av antallet sekvenser det ble subsamplet på.

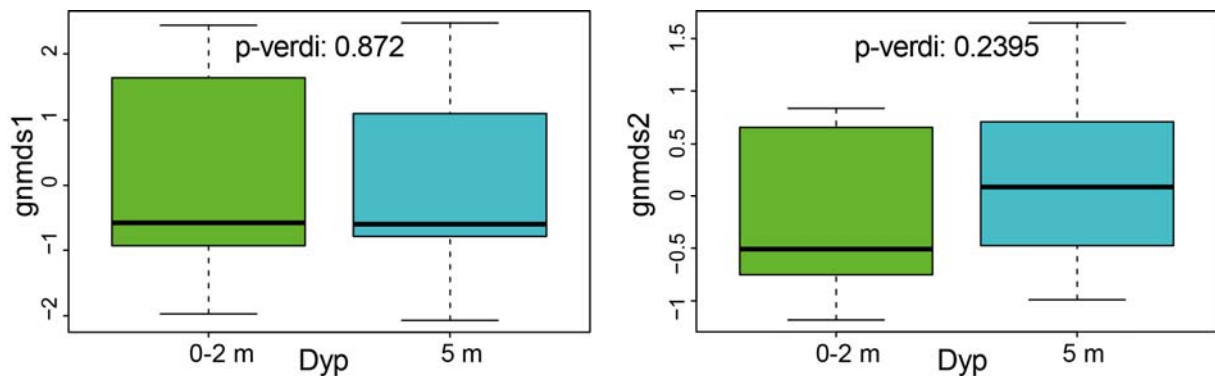
Det ble utført gnmds-ordinasjon på det subsamplede datasettet hvor dinoflagellater var ekskludert (figur 3.22), som viste at prøvene tatt samme dato ved hvert sitt dyp er svært like. Ordinasjonen viser de samme trendene som vi så i ordinasjonen som inkluderer dinoflagellatene (figur 3.18). Prøver tatt hver samme dato plasseres nær hverandre og grupperes etter sesong. Sommer-høstprøvene grupperes i figur 3.22 i tillegg etter hvilket år de er tatt, og viser ikke en stor spredning som i ordinasjonen med dinoflagellatene.

Shepards-diagram (figur S.5) viser liten spredning om linjen som indikerer at den opprinnelige ulikheten i datasettet blir godt preservert når datasettets dimensjoner reduseres til to i ordinasjonen.

Prøvenes plassering langs ordinasjonsaksene ble testet for hvert dyp og viste ingen signifikant forskjell mellom de to dypene (figur 3.23).



Figur 3.22. Ordinasjonsplot (gnmnds) basert på subsamlet datasettet der dinoflagellater er fjernet. Viser prøvenes plassering langs ordinasjonsaksene (gnmnds1 og gnmnds2). Prøvenes farge forteller ved hvilket dyp de er tatt, 0-2 meter (grønn) og 5 meter (mørk blå). Prøvetakningsdatoen for hver prøve er oppgitt og prøver tatt på samme dato er knyttet sammen med en grå linje. Der prøven ikke er knyttet til en annen, er det kun tatt prøver ved ett dyp på den datoen. Prøver plassert nær hverandre har liten variasjon. Gruppering etter sesong: (Rosa) vårprøver 2016, (orange) vårprøver 2017, (rød striplet) sommer-høstprøver og (lilla) vinterprøver, inkludert oktober 2017.



Figur 3.23. Boksdigram representerer prøvenes spredning langs ordinasjonsakse 1 (venstre) og ordinasjonsakse 2 (høyre) for hvert dyp, 0-2 meter og 5 meter. Boksen går fra nedre første kvartil (25.persentil) til øvre tredje kvartil (75.persentil) og svart horisontal strek indikerer medianen (50.persentil). Striplet linje over og under boksen strekker seg til nedre og øvre verdi som ligger innenfor halvannen interkvartilavstand. p-verdi fra Welch tosidig t-test er oppgitt i hvert diagram, som viser ingen signifikant forskjell mellom de to dypene. Basert på subsamlet protistdatasett der dinoflagellater er ekskludert.

4 Diskusjon

4.1 Protistdiversitet undersøkt med metastrekkoding

Dette er den første langtidsstudien som har benyttet metastrekkoding for undersøke protistsamfunnet i pelagialen i Indre Oslofjord. Vi påviste 22 protistgrupper (17 rekker, 4 grupper og uklassifiserte eukayote) som tilhørte 8 supergrupper. Sekvenser klassifisert til supergruppene Alveolata og Stramenopiles var de mest abundante. Innen disse supergruppene er det henholdsvis dinoflagellater og diatoméer som dominerer i eutrofe kystnære områder (Not mfl., 2012). Majoriteten (88%) av de 1 387 OTUene i protistdatasettet var sjeldne (<0.05% sekvenser per OTU) og de vil i liten grad bli diskutert her.

Alveolata

Dinoflagellatene hadde høyest diversitet og relative abundans av sekvenser gjennom stort sett hele prøvetakningsperioden. Dette er i samsvar med flere tidligere metastrekkoding-studier av marint protistplankton (f.eks. (Gran-Stadniczeňko mfl., 2018; Kiliyas mfl., 2014; Piredda mfl., 2017)).

Mange dinoflagellat-arter, som medlemmer av slekten *Triplos*, er relativt store (>100 µm lange (Thronsen mfl., 2007)). To av de mest abundante og vanligste dinoflagellat-OTUene i vårt datasett (>1% sekvenser per OTU), ble fylogenetisk klassifisert som nettopp *Triplos*-arter. Det er tidligere vist positiv korrelasjon mellom cellestørrelse og antall rRNA-gen kopier i genomet hos eukaryoter (Godhe mfl., 2008; Prokopowich mfl., 2003; Zhu mfl., 2005). Dinoflagellater har en stor cellekjerne kalt dinokaryon (Thronsen mfl., 2007) og en stor andel av biomassen til dinoflagellater går til arveanlegget (Tang, 1996). Overrepresentasjonen av dinoflagellat-sekvenser som vi ser kan være relatert til den store genom-størrelsen i denne gruppen (Piredda mfl., 2017) og forholdet mellom genom-størrelsen og antallet rRNA-gen kopier i genomet (Prokopowich mfl., 2003).

OTU 3 ble i PR² klassifisert til *Triplos candelabrus* (vitenskapelignavn *Triplos candelabrum*). I et MAFFT-alignment med *T. muelleri*, *T. candelabrus*, *T. brevis* og OTU 3 viste OTU 3 seg å mest sannsynlig tilhøre *Triplos muelleri*. OTU 3 og *Triplos muelleri* (tidligere *Ceratium triplos*,

O. F. Müller) var 100% like i V4-regionen og OTU 3 skilte seg fra *T. candelabrus* med >40 bp. Hvorfor ikke OTU 3 ble klassifisert til *T. muelleri* kan forklares med at denne sekvensen ikke finnes i den versjonen av PR² som her ble brukt. *Tripes muelleri* ble registrert i høyt antall i lysmikroskopi-celletelldata tatt på samme lokalitet som våre prøver, men det ble ikke talt noen *T. candelabrus/candelabrum* (Norli & Eikrem, 2018). *Tripes candelabrum* er distinkt stor varmtvanns-art (100-200 µm), der cellekroppen er bredere enn lang og er lett gjenkjennelig i mikroskop (Dogde, 1982). *Tripes muelleri* er en stor art som kan være inntil 300 µm lang, den er en kosmopolitt og er vanlig å se i norske kystvann spesielt om sommeren og høsten (Thronsen mfl., 2007). På grunn av det overnevnte antar vi at OTU 3 tilhører *Tripes muelleri*. OTU 3 dominerte på flere tidspunkt gjennom hele preøvetakningsperioden og viste ingen preferanser for hverken dyp eller årstid.

Den OTUen med flest sekvenser i datasettet var OTU 29, som ble fylogenetisk plassert nærmest ferskvanns-arten *Gyrodinium helveticum* (synonym *Gymnodinium helveticum*, Penard). Dette stemmer godt med metastrekkoding-studie fra Ytre Oslofjord, der en av de mest abundante OTUene ble fylogenetisk plassert som *G. helveticum* (Gran-Stadniczeñko mfl., 2018). I Artsdatabanken.no (2018) er det kun registrert ett funn av *Gyrodinium helveticum* i Norge, i Tyrifjorden i Ringerike. Tyrifjorden har utløp i Vikersund og danner starten på Drammenselva som munner ut ca. midt i Ytre Oslofjord. I løpet av mai 2017 hadde Drammenselva økende vannføring til over 850 m³/s (Naustvoll mfl., 2018) og i vår mai-prøven fra 2017 dominerte OTU 29. Det er fristende å spekulere at OTU 29 er ferskvannsarten *G. helveticum*, som ble fraktet med Drammenselva fra Tyrifjorden i en periode hvor vannføringen var stor nok til å føre arten inn i Indre Oslofjord. Men dette forklarer ikke den høye tilstedeværelsen av OTU 29 resten av året. En ferskvanns-art som overføres til saltvann vil oppleve et enormt osmoregulatorisk stress og vil mest sannsynlig dø innen kort tid (Bouvier & Giorgio, 2002; Ning mfl., 2002). Våre CTD-målinger fra mai 2017 viste i tillegg en *økning* i saltholdighet sammenlignet med april, som tyder på at vannføring fra Drammenselva trekker overflatevannet ut av fjorden, ikke skyver det inn. Logares og medarbeidere (2007) har vist at *G. helveticum* har nært slektskap til marine grupper ved at arten ikke har nevneverdig forskjell i 18S-sekvensen sammenlignet med andre marine arter. *Gyrodinium helveticum* ble fylogenetisk plassert sammen med den marine arten *Gyrodinium rubrum* i flere studier (Gran-Stadniczeñko mfl., 2018; Hansen & Daugbjerg, 2004; Logares mfl., 2007). Logares og medarbeider (2007) omtaler denne kladen GYRO og foreslår at den kan representere en linje som har kolonisert ferskvann. *Gyrodinium*-arter er stort sett heterotrofe, få finnes i kultur og mange beskrevne arter er ennå ikke sekvensert (Hansen &

Daugbjerg, 2004). OTU 29 skilte seg fra *G. helveticum* med 6 bp, som kan indikere at dette er en nær marin slektning av *G. helveticum* som ennå ikke er genetisk karakterisert.

Dinoflagellatene er, etter diatoméene, ansett å for være de viktigste marine eukaryote primærprodusentene i kystvannet (Not mfl., 2012). Arter innen denne rekken kan danne store oppblomstringer som potensielt kan være giftige. Blant de mest abundante dinoflagellat-OTUene i protistdatasettet, ble tre plassert fylogenetisk i to slekter med giftige alger, *Alexandrium* og *Dinophysis*. *Alexandrium*-arter produserer nervegiften saxitoksin som forårsaker PSP (Paralytic Shellfish Poisoning) som potensielt kan være dødelig (Anderson mfl., 1990; Jellett mfl., 1995). *Dinophysis*-arter produserer diaré-giftene okadasyre og dinophysistoksiner, som forårsaker gastrointestinal DSP (Diarrhetic Shellfish Poisoning) (Yasumoto mfl., 1985). Mennesker kan bli utsatt for disse forgiftningene ved å spise skjell der giftene har blitt oppkonsentrert. I Norge blir forekomsten av giftige algeoppblomstringer og mengden gift i skjell nøye overvåket (Duinker mfl., 2016; Havforskningsinstituttet, 2007; <http://www.matportalen.no/verktoy/blaskjellvarsel/>; <http://algeinfo.imr.no/>). Ved å studere algeinformasjon for i Indre- og Ytre Oslofjord gjennom 2017 (se <http://algeinfo.imr.no/>), er artene *D. acuminata*, *D. acuta*, *D. norvegica*, *A. tamarensis* og *A. ostenfeldii* registrert. Dette stemmer godt med våre funn. I langtidsstudier fra Flødevigen og Skagerrak ble det årlig sett oppblomstringer av arter innen begge slektene (Havforskningsinstituttet, 2007). OTU 8 ble klassifisert i PR² og fylogenetisk plassert som *Alexandrium hiranoi*. Inspeksjon av alignment mellom OTU 8, *A. hiranoi* og *A. pseudogonyaulax* viste at disse tre var identiske i V4-regionen. *A. hiranoi* er ikke tidligere påvist i Norge og vi kan anta at OTU 8 tilhører *A. pseudogonyaulax*. *Alexandrium pseudogonyaulax* er en kosmopolitt, den finnes verden over og trives best i brakkvann i kystområder (Triki mfl., 2016) og er vanlig i Norge (Thronsen mfl., 2007). Den er detektert i celletelldata fra hele Ytre Oslofjord i juni og juli 2016 (Waldy mfl., 2017), noe som støtter vår tolkning. Den er tidligere registrert i Indre Oslofjord og blir omtalt som et årviss fenomen i Oslofjorden (Berge mfl., 2015).

Alle OTUene i klassen Syndiniophyceae tilhørte Syndiniales som er en utelukkende parasittisk orden (Guillou mfl., 2008). Syndiniales-arter innen Dino-gruppe II, som endoparasitten *Amoebophrya* sp. er kjent for å infisere bl.a. *Alexandrium*- og *Dinophysis*-arter. *Amoebophrya* ble påvist i *Dinophysis norvegica* da den i 2001 hadde stor oppblomstring i Østersjøen og Nordsjøen (Salomon mfl., 2003) og i *Alexandrium* sp. som hadde oppblomstring utenfor den franske kyst samme år (Chambouvet mfl., 2011). Syndiniales-arter kan også infisere mange

andre dinoflagellat-arter, arter innen Radiolaria og ciliater (Bråte mfl., 2012; Coats mfl., 1996). Den store diversiteten av Syndiniales-OTUer vi ser i dette studiet, kan indikere at endoparasitter er vanlig i protistsamfunnet i Indre Oslofjord som ellers i havet. Tara Ocean ekspedisjonen påviste en stor mengde parasitter (MALV – marine alveolates) i sine data (de Vargas mfl., 2015), som bekrefter våre antagelser.

OTUer som er klassifisert til ciliater ble registrert gjennom hele prøvetakningsperioden, med relativt få sekvenser utenom i november og februar. Dette stemmer dårlig med tidligere celletellestudier fra Kielbukta som viser at ciliatene her har høyest abundans om våren og på høsten og lavest gjennom vinteren (Smetacek, 1981). På den annen side er det i Oslofjorden tidligere blitt påvist få ciliater under våroppblomstringen (Kristiansen mfl., 2001), hvor antall ciliater øker fra sensommeren og ut mot høsten (Paasche & Kristiansen, 1982). Ciliater beiter på både nano- og picoplankton, diatoméer, dinoflagellater og bakterier og blir selv beitet på av zooplankton (Pierce & T. Turner, 1992) og blir derfor ansett som et viktig ledd i det marine næringsnett.

Stramenopiles

Rekken Ochrophyta var etter dinoflagellatene den rekken med størst diversitet og relativt abundans av sekvenser i protistdatasettet. Disse funnene stemmer dårlig sammenlignet med lysmikroskop-celletelledata fra Indre Oslofjord, hvor diatoméene dominerte over dinoflagellatene (Norconsult, 2017, 2018; Norli & Eikrem, 2018). Kisel-skallet til diatoméene er vanskelig å knuse, som kan forhindre isolering av DNAet og muligens forklare at diatoméene er underrepresentert i våre data (Medinger mfl., 2010). På den annen siden kan våre data inkludere mange dinoflagellater som ikke lar seg observere i lysmikroskop, som parasitter innen ordenen Syndiniales, og kan derfor muligens fremstille et mer omfattende bilde av det totale protistsamfunnet enn resultater fra mikroskopi celletelledata.

I klassen Mediophyceae (sentrisk diatomeer) så vi tre av de vanligste OTUene i protistdatasettet, som tilhørte slektene *Chaetoceros*, *Cyclotella* og *Thalassiosira*. Arter innen disse tre slektene kan danne lange kjeder. *Thalassiosira* er i norske farvann en av de best undersøkte og mest artsrike marine plankton slektene, med mer enn 100 kjente arter (Thronsdén mfl., 2007). Den første marine diatomeen som fikk hele sitt genom sekvensert var *Thalassiosira pseudonana* (Armbrust mfl., 2004). Dette er en vanlig art å se i Indre- og Ytre Oslofjord (Thronsdén mfl., 2007) og man tror at *T. pseudonana* er nedstammet fra en stamfar i

slekten *Cyclotella* (Alverson mfl., 2011). De fleste *Cyclotella*-artene hører hjemme i fersk- og brakkvann og er vanlig å se i norske fjorder (Thronsen mfl., 2007). *Cyclotella choctawhatcheeana* har blitt foreslått som en potensiell indikatorart på eutrofiering i kystvann i Østersjøen (Jaanus mfl., 2009).

Det ble registrert et stort antall *Chaetoceros* celler i celletelldata fra overvåkingen (Norli & Eikrem, 2018), som stemmer godt med funnene vi har gjort. Blant de heterotrofe stramenopilene i vårt studie dominerte den relativt nyoppdagede MAST-kladen (Massana mfl., 2004). MAST-kladen består av hovedsakelig bakterivore picoplankton, men også herbivore arter. De finnes i alle fem verdenshav (Massana mfl., 2002; 2004; Piwosz & Pernthaler, 2010).

Hacrobia

Innen supergruppen Hacrobia var cryptofyten *Teleaulax amphioxeia* blant de vanligste OTUene i protistdatasettet. Dette er en art som det er normalt å finne i brakkvannsområder i Skagerrak og Kattegat (Thronsen mfl., 2007). Dette var også en av de mest abundante OTUene i artikkelen til Gran-Stadiczeňko og medarbeidere (2018) i Ytre Oslofjord.

Innen rekken Haptophyta tilhørte flest sekvenser OTUer i klassen Prymnesiophyceae. I tidligere metastrekkodingstudier fra Ytre Oslofjord hvor det ble brukt haptofytt-spesifikke primere, var ordenen Prymnesiales den mest abundante, frekvente og hadde høyest diversitet (Egge mfl., 2015; 2015a). Artene *Emiliana huxleyi* og *Chrysochromulina simplex* ble av både Egge og medarbeidere (2015) og Gran-Stadiczeňko og medarbeidere (2018) påvist med høy abundans. I våre data hadde slekten *Chrysochromulina* relativt høy diversitet, med seks OTUer. En ble klassifisert til arten *Chrysochromulina leadbeateri*. Dette er en vanlig art i kystvannet (Thronsen mfl., 2007). *Chrysochromulina leadbeateri* er i Norge spesielt kjent etter en giftig masseoppblomstring i Vestfjorden (Nordland) i 1991, hvor flere hundre tonn oppdrettslaks døde (Edwardsen & Paasche, 1992; Eikrem & Thronsen, 1998). Studier har siden dette vist at *C. leadbeateri* vanligvis er lite til ikke giftig (Edwardsen & Paasche, 1992). Innen slekten *Emiliana* ble det kun påvist én OTU som ikke ble klassifisert til art. Denne sammen med én annen OTU dominerte vannprøven fra 5 m i august 2017. Sannsynligheten er stor for at dette er *Emiliana huxleyi*, siden lysmikroskop celletelldata viste et høyt antall *Emiliana huxleyi* i prøven tatt i august 2017 fra 5 m på samme lokalitet som våre prøver (Norli & Eikrem, 2018).

OTUer påvist i de heterotrofe rekkene Centrohelioczoa, Katablepharidophyta, Picozoa og Telonemia ble sjeldent klassifisert nærmere enn til klasse. Alle rekkene ble påvist med lav abundans, foruten Picozoa i desember 2017. Picozoa er en ganske nylig beskrevet rekke, som inkluderer grupper picoeukayoter som tidligere hadde navnet «picobiliphytes» (Seenivasan, 2013). Picozoa blir regnet som heterotrofe, siden det ikke er funnet DNA fra disse som tilhører kloroplaster (Yoon mfl., 2011). Innen denne gruppen er det foreløpig kun beskrevet en art, *Picomonas judraskeda* (Seenivasan mfl., 2013).

Archaeplastida

Innen rekken Chlorophyta var picoflagellaten *Bathycoccus prasinus* som tilhører klassen Mamiellophyceae blant de 15 vanligste OTUene. *Bathycoccus prasinus* ble for første gang beskrevet av Eikrem og Throndsen (1990) fra vannprøver tatt i Napolibukten utenfor Italia og ble i samme arbeid bekreftet i vannprøver tatt fra Austevoll i Hordaland og Flødevigen i Aust-Agder. Utbredelsen til denne arten i Norge er lite kjent, mye fordi den er så liten (1-2 µm) at den er vanskelig å identifisere i lysmikroskop og lett kan forveksles med *Micromonas pusilla* (Eikrem & Throndsen, 1998; Throndsen mfl., 2007). *Bathycoccus prasinus* er tidligere påvist som en av de mest abundante grønnalgene blant picoplanktonet i Den engelske kanal (Not mfl., 2004) og kystnære områder i Arktis (Not mfl., 2005).

OTUene med høy abundans i rekken Rhodophyta hadde lav identitet med referansesekvenser i PR² (<90%). I rekken Rhodophyta er majoriteten av de kjente artene flercellede bentiske alger og det finnes kun noen få encellede arter i planktonet som alle tilhører klassen Bangiophyceae (Throndsen mfl., 2007). Ingen av OTUene i våre data tilhørte klassen Bangiophyceae. Blant de mindre abundante rødalge-OTUene i protistdatasettet ble fem 100% klassifisert til art, *Ceramium rubrum* (vanlig rekeklo), *Dasya baillouviana* (strømgarn), *Leptosiphonia schousboei*, *Polysiphonia fucoides* (svartdokka) og *Rhodomela confervoides* (teinebusk). Arter innen slektene *Ceramium* og *Polysiphonia* er blant de aller vanligste rødalgene i norsk flora og begge slektene er representert i alle verdens hav med et høyt antall arter (Rueness, 2006). Teinebusk, vanlig rekeklo og svartdokka er vanlige arter å se i Oslofjorden (Artsdatabanken.no, 2018).

Siden vi har sett på DNA og ikke RNA er det vanskelig å si om sekvensene vi observerer stammer fra levende eller døde celler. DNA brytes langsommere ned i vann enn RNA. Rødalge OTUene kan stamme fra levende celler som sporer og vegetative-spredningsenheter eller dødt

og delvis nedbrutt materialet. Rødalgenes sporer mangler flageller og er avhengig av strømmer i vannet for spredning (Searles, 1980). Strømgarn ble først observert i Norge i 1966 og ble ansett som en fremmedart i Skandinavia. Den har nå blitt vanlig å se langs sør-kysten og inn i Skagerrak (Lein, 1999), som kan indikere god spredningsevne. Mange arter slipper sporene sine i løpet av våren og sommeren. Dette samsvarer med tiden de ble observert i våre prøver. Etter hardt vær er det vanlig å se alger drivende i vannoverflaten og over tid vil disse brytes ned til mindre fragmenter. Indre Oslofjord er hverken vær- eller bølgeeksponert, men man kan spekulere i om økt menneskelig aktivitet i sommerhalvåret langs kystlinjen og antallet fritidsbåter kan bidra til økt spredning av både sporer, vegetative celler og løse plante-rester.

Rhizaria

Innen supergruppen Rhizaria hadde sekvenser tilordnet rekken Cercozoa høy diversitet, men lav relativ abundans. Cercozoa var tilstede gjennom hele prøvetakningsperioden og viste en plutselig økning i september 2017, som var forårsaket av spesielt en OTU som ble klassifisert til *Cryothecomonas aestivalis*. *Cryothecomonas*-arter er vanlig å se i Arktis og Barentshavet (Lovejoy mfl., 2006) og ble først beskrevet fra prøver tatt i den danske Isefjorden som er knyttet til Kattegat (Thomsen mfl., 1991). I Nordsjøen er det vanlig å se *Cryothecomonas aestivalis* om sommeren og høsten, hvor den beiter på diatomeen *Guinardia delicatula*, ved å penetrere cellen og fagocyttere cellens cytoplasma ved hjelp av et pseudopodium (Drebes mfl., 1996). I våre data var OTUen som ble klassifisert til *Guinardia delicatula* ikke tilstede i hverken august- eller september-prøvene fra 2017, som kan indikere at *Cryothecomonas aestivalis* også beiter på andre diatoméer.

Opisthokonta

I supergruppen Opisthokonta var det krageflagellatene som dominerte av protistene. Krageflagellatene er evolusjonært viktig, da de er den nærmeste kjente slektningen til metazoa (King mfl., 2008). Krageflagellatene har en global utbredelse, både marint og i ferskvann (Buck & Garrison, 1988; Thomsen & Larsen, 1992).

Sammensetningen av protistgruppene vi så i Indre Oslofjord hadde en høy diversitet – av både brakkvann og saltvannsarter – som stemmer godt med hvor vannmassene i Indre Oslofjord kommer fra. Planktonet har liten horisontal egenbevegelse og blir fraktet med havstrømmene

som forflytter seg over store avstander. Vi fant arter som er typiske i Nordsjøen, Den engelske kanal, Skagerak, Kattegat, Østersjøen og de fem verdenshav. Videre fant vi en god blanding av arter fra forskjellige nivå i næringskjeden fytoplankton, heterotrofe beitere, predatorer, bakterivore, samt parasitter.

4.2 Variasjon gjennom året

Sesongvariasjon i fytoplanktonet

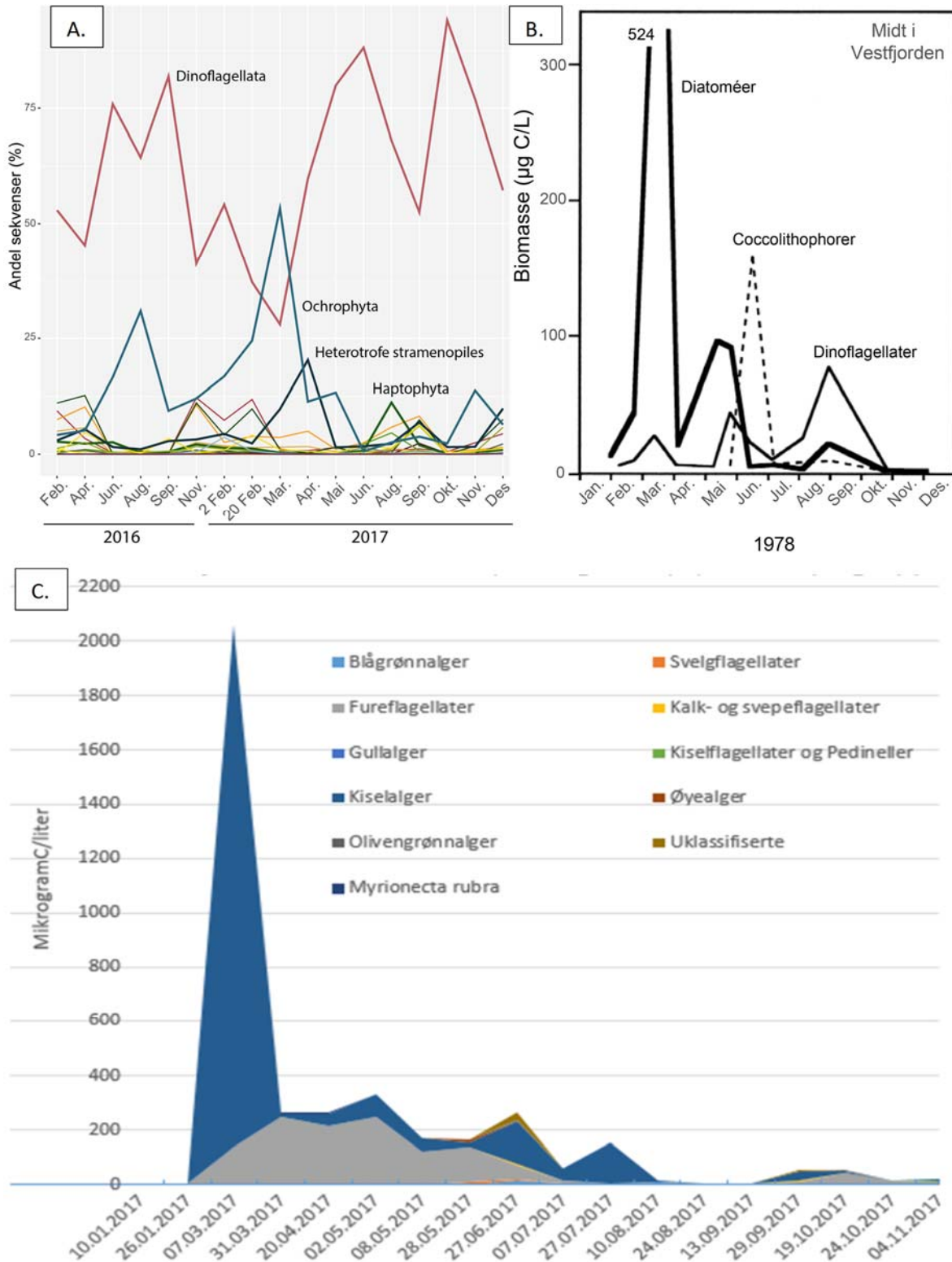
Figur 4.1 fremstiller hvordan noen grupper fytoplankton varierer gjennom året ved prøvetakningsstasjonen DK1.

Det skal være nevnt at det er mange feilkilder i det å sammenligne de tre diagrammene som fremstilles i figur 4.1.

1. Metodene er forskjellige. Vi vet ikke om andel sekvenser er sammenlignbart med biomasse, men det er et håp om at det i fremtiden vil være mulig å regne ut biomasse for ulike grupper basert på sekvenser.
2. Metoden for innsamling forskjellig for alle tre diagrammene.
3. Vi vet ikke fra hvilket dyp cellenelledataene i diagram B er samlet fra.
4. Oppblomstringer av fytoplanktonet varierer fra år til år i mengde og tidspunkt.

Det er allikevel interessant å se hvor like trender de tre diagrammene fremstiller. Det er tydelig at diatoméene/Ochrophyta/Kiselalger dominerer under en våroppblomstring som tilsynelatende finner sted i mars, etterfulgt av en mindre oppblomstring i mai. Haptofyttene/coccolithophorer/kalk- og svepeflagellatene har vanligvis en oppblomstring på sommeren. I diagram C er denne vanskelig å se, men man ser en tynn gul linje mellom grå og blå over prøven tatt 27.06.2017. Dinoflagellatene/fureflagellatene varierer med flere oppblomstringer gjennom året. Det forekommer en høstoppblomstring som består av diatoméene/ochrophyta/kiselalger og dinoflagellatene/fureflagellatene.

Et poeng med denne sammenligningen er at metoden metastrekkoding har potensialet til å avsløre like trender som er avdekket med tradisjonelt lysmikroskopi. Det er relativt vanlig å se at fytoplankton sammensetningen i havet skifter fra diatoméer til dinoflagellater og visa versa (Tang, 1996; Taylor mfl., 1993), som kommer tydelig fram i våre data.



Figur 4.1. Sesongvariasjon i grupper fytoplankton i (A) 2016-2017 (våre data), (B) i 1978 (Norconsult, 2017, 2018; Paasche, 2005) og (C) 2017 (Norli & Eikrem, 2018) ved prøvetakningsstasjonen DK1, som ligger midt i Vestfjorden i Indre Oslofjord. Mengde fytoplankton er (A) andel sekvenser fra metastrekkoding studie og (B og C) biomasse (mikrogram karbon/ liter) basert på å telle celler i lysmikroskop.

Sesongvariasjon i prøvene

De påviste protistgruppene varierte gjennom året, både i diversitet og relativ abundans. Indekser for rikhet og Shannon diversitet (H') og Pielou jevnhet (J') viste sesongvariasjoner ved begge dyp, med høyest indeksverdier i vintermånedene og lavest i vår-sommermånedene. Temperatur viste tydelig sesongvariasjon med lavest måling i februar og høyest i august. Temperatur alene kunne ikke forklare sammensetningen av protister i prøvene.

Vinter

I ordinasjonen grupperte vinterprøvene seg tydelig sammen, som indikerer at OTU sammensetningen i disse prøvene var ganske lik. Indeksverdiene for rikhet, diversitet (H') og jevnhet (J') var signifikant korrelert med hverandre og viste høyest verdier i vintermånedene. Den prøven med høyest OTU-rikhet ble tatt i desember 2017, med >500 OTUer. Den høye diversiteten på vinteren kan forklares ved at høy jevnheten (J') muligens tillater deteksjon av et større antall OTUer (Caporaso mfl., 2012).

Det er spesielt på vinteren hvor man ser en vannfornyelse i Indre Oslofjord. Pågående sterk nordavind vil frakte det øvre vannlaget ut av fjorden og som kompensasjon dannes en dypere motstrøm av vann inn i fjorden (Baalsrud & Magnusson, 2002). Det nye vannet fra Skagerrak kan frakte med seg arter som vanligvis ikke er å se ellers i året. I våre data så vi en økt saltholdighet i vintermånedene og et svakere sprangsjikt, noe som kan bekrefte at en vannfornyelse og omveltning i vannmassene har funnet sted. Et svakere sprangsjikt og forflytting av vannmassene oppover tillater arter som vanligvis befinner seg i de dypere vannlagene å komme til overflaten. I tillegg vil dypvannet bringe med seg næringssalter til overflaten som tillater større diversitet ettersom det er mindre konkurranse om næringssaltene. Om vinteren er temperaturen i vannet på det laveste og daglengden på det korteste, som gir lav vekstrate hos de fototrofe protistene (Lawrence & Menden-Deuer, 2012). Allikevel så vi en markant økning av diatoméer i november 2017, som bekreftes av fluorescens-målingene. Vi kan anta at en økning i næringssalter er årsaken til denne høstoppløstringen. Saltholdigheten er ikke spesielt høy i november 2017 noe som indikerer at næringssaltene ikke kommer fra en vannfornyelse, men muligens på grunn av økt nedbør som fører til avrenning fra land.

Vår

I ordinasjonen grupperes også vårprøvene seg tydelig sammen. Når våren kommer blir dagene lenger og økt sol-innstråling gjør vannet varmere. Protistsamfunnet har etter vinteren lav biomasse og tilgangen på næringsalter er god. Det fytoplanktonet som har evnen til å raskest utnytte de overnevnte forholdene og samtidig ha høy maksimal vekstrate vil ha en mulighet til å oppnå en høy biomasse på kort tid. OTUene som tilhørte fototrofe rekker og grupper hadde høy abundans i slutten av februar og i mars 2017. På våren er det vanlig å se en algeoppblomstring dominert av diatoméer i kystnære områder (Bruno mfl., 1983; Peperzak mfl., 1998) og i Indre Oslofjord (Norconsult, 2018; Norli & Eikrem, 2018; Paasche & Østergren, 1980). Dette stemmer godt med våre resultater. Diatoméene kan ha stor overflate i forhold til sitt volum, som kan tillate raskt næringsopptak. De har også en høy maksimal vekstrate i forhold til andre fytoplanktongrupper (Chan, 1980). Dette er fordi diatoméene har et lavt forhold mellom karbon og klorofyll a, som tillater høy fotosyntese-hastighet per enhet karbon.

At vi ikke ser en tidlig våroppblomstring i våre fluorescensdata kan forklares med at det mangler CTD-data for prøvene tatt i mars. I prøvene våre ser vi tydelig at de sentriske diatoméene dominerer i mars 2017, spesielt OTU 15 som ble klassifisert til *Thalassiosira* sp. Mengden bakterier øker vanligvis under våroppblomstringen og det blir høy tetthet av bakterier som er spesialisert i å bryte detritus fra fytoplanktonet (Riemann mfl., 2000). Dette forklarer godt økningen av den bakterivore MAST-kladen som vi ser i april 2017. Mengden diatoméer begrenses hovedsakelig av tilgangen på silisium (Harrison mfl., 1977; Paasche & Østergren, 1980). Hos noen av diatoméene øker synkehastigheten ettersom silisium-mangelen begrenser lengden på setae ("børstene") i *Chaetoceros*-arter og stavene ("processes") som holder *Skeletonema*-artene i lange kjeder og mulig flytende (Smayda & Boleyn, 1966; Takabayashi mfl., 2006). Diatomésamfunnet kan bl.a. derfor gå fra høy abundans til lav på relativt kort tid. Mengden heterotrofe grupper og dinoflagellater øker markant i april 2017 etter våroppblomstringen, som også kan ha bidratt med å beite ned diatoméene. I mai 2017 ser vi en ny økning av diatoméer som kan forklares av moderat vannfornyelse og dermed næringsalter. I de hydrografiske målingene ser vi en økning i saltholdighet i mai 2017, som bekrefter en vannfornyelse. Det er vanlig i Indre Oslofjord å se i slik moderat vannfornyelse i mai (Baalsrud & Magnusson, 2002). Om våren vil også smeltevann fra innlandet bringe med seg næringsalter til fjorden, som kan bidra til økt fytoplanktonvekst.

Sommer - høst

I sommerprøvene var indeksverdiene for rikhet, diversitet (H') og jevnhet (J') på det laveste og prøvene var dominert av dinoflagellater. Sommerprøvene viste en stor spredning i ordinasjonsdiagrammet som hovedsakelig kan indikere en større variasjon mellom prøvene sammenlignet med resten av året. Den store variasjonen forklares av at hver prøve ofte var dominert av én OTU som var forskjellig for hver prøve. I juni 2016 dominerte OTU 8 som mest sannsynlig tilhørte *Alexandrium pseudogonyaulax*. Juniprøvene fra 2017 ble dominert av spesielt en OTU innen ordenen Suessiales, som ble fylogenetisk plassert til arten *Pelagodinium bei*. *Pelagodinium bei* er endosymbiont i f.eks. den pelagiske foraminiferen *Orbulina universa* (Spero, 1987). I september 2016 og august 2017 dominerte OTU 4 som ble fylogenetisk plassert til *Heterocapsa rotundata*. *Heterocapsa rotundata* er en miksotrof dinoflagellat som beiter på både bakterier og fytoplankton (Millette mfl., 2017). OTU 3 ble identifisert til *Tripos muelleri* og dominerte prøvene fra oktober i 2017. Dinoflagellatene har god egenbevegelse og påvirkes derfor ikke av lite næringsalter i det øvre blandingslaget om sommeren. I tillegg er flere fototrofe dinoflagellater kjent som miksotrofe og kan beite på bakterier og andre protister (Millette mfl., 2017) om fotosyntesen ikke kan dekke næringsbehovet.

4.3 Lysmikroskop vs. metastrekkoding

Arter

Det ble totalt observert 48 dinoflagellat-arter med lysmikroskop og metastrekkoding ved 5 m i 2017. Av disse var åtte av artene felles for de to metodene, 21 av artene var unike for lysmikroskop og 18 var unike for metastrekkoding.

Av de 21 unike dinoflagellat-artene kun observert i lysmikroskop var 15 av artene innen slektene *Dinophysis*, *Tripos* eller *Protoperidinium*. *Dinophysis*-artene som forekommer i Oslofjorden viser ekstremt liten til ingen interspesifikk variasjon i 18S rRNA-genet (Edvardsen mfl. 2003; Raho mfl. 2013). Markøren V4-regionen i 18S rRNA-genet som ble brukt i denne studien er derfor uegnet om man ønsker skille *Dinophysis*-sekvenser til art. Det mitokondrielle *cox1*-genet er foreslått som en bedre markør for taksonomisk klassifisering av *Dinophysis*-arter (Raho mfl., 2013).

Av de 2 342 beskrevne dinoflagellat-artene i ALGAEBASE har omtrent 18% fått sitt 18S rRNA-gen sekvensert (Mordret mfl., 2018). I referansedatabasen PR² mangler det sekvenser

for alle de *Tripos*-artene som her ble kun observert i lysmikroskop (Guillou mfl., 2013). *Protoberidinium*-arter er heterotrofe og derfor vanskelig å ha i kultur, mange av sekvensene er fra isolerte enkeltceller. Derfor er det er veldig få sekvenser klassifisert til *Protoberidinium* i PR². Det er den representative sekvensen for OTUene blir søkt opp mot PR² i den bioinformatiske-pipelinen. Den referansesekvensen som har best identitet med den representatives sekvensen blir gitt som taksonomisk klassifisering til OTUen. I PR² er det inkludert mange miljøsekvenser som ikke er taksonomisk klassifisert. Sannsynligheten er stor for at mange av våre OTUer muligens som tilhører *Protoberidinium* har blitt klassifisert til en miljøsekvens eller "Ubestemt dinoflagellat". Ingen av de mest abundante dinoflagellat-OTUene i protistdatasettet ble fylogenetisk plassert som *Protoberidinium*-arter. Metoden viser derimot at ukjente dinoflagellat-OTUer kan bli plassert til orden.

Med metastrekkoding ble det påvist 18 taksa som ikke ble observert i lysmikroskop fordelt på 16 forskjellige slekter. Som nevnt i innledning kan arter være vanskelig å gjenkjenne i lysmikroskop på grunn av liten størrelse eller etter bruk av Lugols. Ved en gjennomgang av artene detektert med metastrekkoding viste det at omtrent halvparten av artene er nakne dinoflagellater. Mange nakne dinoflagellater mangler harde organiske skall og cellene kan bli deformert og ugjenkjennelige når preservert i Lugols (Gómez mfl., 2015; Hansen, 1991). Av de resterende ikke-nakne dinoflagellat-artene var det ingen arter som normalt er mindre enn 15-20 µm. Iallfall fire av de ikke-nakene dinoflagellatene påvist med metastrekkoding er kjent for å danne cyster. Cyster kan være vanskelig å bestemme til art i lysmikroskop, siden cyster fra samme art kan ha forskjellig morfologi og cyster fra forskjellig arter kan være tilsynelatende like (Bravo mfl., 2006). Én av artene påvist med metastrekkoding var *Euduboscquella crenulata* som parasitt i ciliater og kan derfor heller ikke forventes å bli observert i lysmikroskop, selv om den er synlig inni ciliaten den parasitterer (Coats mfl., 2012).

Taksa

I den første sammenligningen av dinoflagellat-taksa (figur 3.14 og tabell S.6) ble det påvist 160 unike taksa sammelagt med de to metodene lysmikroskop og metastrekkoding. Av disse var det kun 8 taksa som var felles for begge metoder. Kun med metastrekkoding ble det påvist 45 taksa innen den parasittiske ordenen Syndiniales.

Det er mange feilkilder i den første sammenligningen av dinoflagellat-taksa og de resultatene vil derfor ikke bli diskutert videre her.

Diversitet forsvinner i den brukte definisjonen av et taksa

I denne studien ble det påvist 567 OTUer som ble taksonomisk klassifisert til dinoflagellater. Av disse var 401 av OTUer tilstedte på 5m dyp i 2017. I den første sammenligningen av antall dinoflagellat-taksa påvist med de to metodene ble ett taksa definert til navnet gitt som nærmeste taksonomiske klassifisering i de to metodene.

Med denne definisjonen ble de 401 dinoflagellat-OTUene i protistdatasettet fordelt på 106 unike taksa. Dette var fordi flere OTUer ble klassifisert til nøyaktig samme referanse-sekvens eller til referanse-sekvenser som var taksonomisk klassifisert til samme taksa. Disse ble dermed definert som ett taksa. Med metastrekkoding ble f.eks. seks ulike OTUer taksonomisk klassifisert til nøyaktig samme referansesekvens (FJ402945 *Tripos candelabrus*), som utgjør kun ett taksa i taksa-definisjonen brukt i sammenligningen i denne studien. En av disse seks *T. candelabrus*-OTUene ble som forklart tidligere omklassifisert til *T. muelleri*. De resterende fem OTUene kan potensielt være noen av de andre *Tripos*-artene observert i lysmikroskop, men fordi det mangler referansesekvenser på disse artene i referansedatabasen PR² har vi ikke hatt mulighet til å påvise de med bruk av kun den databasen.

Videre ble én OTU klassifisert til *Heterocapsa nei/rotundata* og med lysmikroskop ble både *Heterocapsa niei* og *Heterocapsa rotundata* registrert. Alle tre ble i dette studiet talt som unike taksa. Vi kan anta at det er skrivefeil i PR² hvor *nei* skulle vært *niei*, siden det ikke finnes en art som heter *Heterocapsa nei*. Det er en viss sannsynlighet for at *Heterocapsa*-OTUen er en av de overnevnte artene. Siden sekvensene til disse tre ikke ble inspisert i et alignment kan vi ikke si at *Heterocapsa*-OTUen sannsynligvis tilhører *Heterocapsa rotundata* eller *Heterocapsa niei* og de tre taksa forble derfor unike i første sammenligningen.

Det forsvant sannsynligvis mange taksa med den brukte definisjonen i begge metoder. I lysmikroskop ble mange celler som ikke ble klassifisert til art fordelt i grupper etter størrelsesorden som "Ubestemte tekate 20-40 µm". Det er mulig at celler som var morfologisk og taksonomisk ulike ble talt i samme størrelsesorden og da kun registrert som étt taksa. Det ble totalt definert 9 størrelsesorden-taksa med lysmikroskop, 15 om man inkluderer *Gyrodinium/Gymnodinium* og *Protoperidinium* spp. Vi kan anta at flere unike taksa ligger skjult i disse gruppene registrert med.

Med metastrekkoding ble sekvenser med 98% likhet slått sammen til én OTU før sekvensene ble taksonomisk klassifisert. Dette innebærer at det kun var den representative sekvensen for

OTUen som ble klassifisert. Flere taksa kan derfor også ligge skjult i sekvenser som inngår i hver enkelt OTU. Mordret og medarbeidere (2018) viste at omtrent 2/3 av dinoflagellat-diversiteten gikk tapt ved klustering av V4-sekvenser med 98%-likhet. Til slutt, siden noen nært beslektede arter kan ha identisk V4-region i 18 rRNA-et (Edwardsen mfl., 2003) kan en OTU representerer flere arter.

Andre feilkilder i sammenligningen

Mengden sjøvann man analyserer er forskjellig med de to metoden lysmikroskop og metastrekkoding som er et bias i denne sammenligningen, i tillegg til antall prøver som ble sammenlignet.

Normalt i overvåkingen i Indre Oslofjord er det celler fra 25 mL vann som observeres med lysmikroskop. I dette studiet var det et mål å filtrere 2 L vann på hvert Sterivex-filter. Det ble ikke alltid målt volumet vann som gikk igjennom filteret og andre ganger ble filteret tett før 2 L var filtrert. Vannprøven kunne vært pre-filtrert (som blir påpekt i seksjon 4.5) for å unngå store celler, som ikke bare dominerer prøven etter sekvensering, men også kan bidra til at filteret blir fortere tett. Selv om det hadde blitt filtrert 2 L på hvert filter er det ikke nødvendigvis 2 L sjøvann som sekvenseres og analyseres. Mengden kan minke gjennom de molekylære forberedelsene før sekvensering. Når DNAet blir ekstrahert frigjøres alt DNA i cellene og det blandes i en antatt homogen løsning. Når man tar ut en liten del av denne løsningen (her 2.5 μ L) til PCR kan den potensielt inneholde DNA-sekvenser fra alle organismene i den 2 L vannprøven, men vi vet det ikke med sikkerhet. Det blir i seksjon 4.6 foreslått et PCR-forsøk for å undersøke hvor homogen DNA-løsningen er etter ekstraksjon.

I lysmikroskop celletelldataene er taksalisten basert på celler observert i 18 prøver (Norli & Eikrem, 2018), mens i våre data er de basert på de 11 prøvene tatt i 2017. Det er også litt usikkert om lysmikroskopprøvene er samlet på 4 m eller 5 m. I celletellelisten står det 5 m, men i rapporten står det 4 m (Norli & Eikrem, 2018). Jeg har i denne studien valgt å si at lysmikroskopdataene er fra 5 m, siden det er celletellelisten jeg har brukt som data for sammenligning.

Fordel med metastrekkoding

Metastrekkoding har en stor fordel som lysmikroskop ikke har, sekvensene i prøvene kan fylogenetisk analyseres. Dette gjør at sekvenser som ikke taksonomisk klassifiseres til art i en referansedatabase, allikevel kan bli klassifisert ved å bli fylogenetisk plassert nær en art i et

fylogenetisk tre. Flere av de mest abundante dinoflagellat-OTUene med stor sikkerhet plasser nær kjente arter, som OTU 4 – *Heterocapsa rotundata* og OTU 10 – *Pelagodinium bei*. Dette kan være for omfattende arbeid i overvåkningssammenheng, men muligheten er der.

Siden det er en så liten del av DNA-ekstraktet som blir brukt til analyse av prøven, gjør det at prøven kan bli analysert igjen om det skulle være ønskelig. I den farten molekylære metoder utvikles i dag, kan våre resultater her være utdatert om kort tid. Det som er igjen av våre prøver er per i dag lagret i en fryser på Institutt for biovitenskap på Blindern. Det blir anbefalt av protist-arbeidsgruppe (ProWG CBOL) at slike prøver skal bli lagret i en DNA-bank eller i en museums samling, slik at prøvene er tilgjengelig for fremtidige analyser (Pawlowski mfl., 2012).

For å oppsummere kan våre resultatene tyde på at de to metodene er komplementære til hverandre. Metoden lysmikroskop fanger opp arter som ennå ikke er sekvensert eller er vanskelig å skille ved bruk av markøren brukt i dette studiet. Metastrekkoding gir mulighet til å detektere uavhengig av livsstadiet, om cellen har blitt deformert, skille cyster til art og detektere små arter, til og med parasitter.

4.4 Ingen signifikant forskjell mellom dypene

Våre resultater viste ingen signifikant forskjell mellom de to dypene, 0-2 m og 5 m, hverken i hydrografiske forhold eller protist diversitet og abundans. Men noen små forskjeller så vi allikevel.

Forskjeller

Ved sammenligning av antall OTUer ved de to dypene i protistdatasettet var 17 OTUer særegne for 0-2 m og 75 OTUer som var særegne for 5 m. At det ble påvist flere særegne OTUer ved 5 m kan forklares med at det var flere prøver som ble tatt på 5m enn ved 0-2 m. Flere prøver vil fange opp mer diversitet og differansen i antall prøver er en stor feilkilde i sammenligningen.

Kun 12 av de totalt 92 særegne OTUene ble klassifisert til art i referansedatabasen PR². Av disse var fem av artene tilstede ved begge dyp, ved at OTUer fra det andre dypet også ble klassifisert til samme art. Det var bare to OTUer som ble klassifisert med høy sikkerhet (>99% identitet med referansesekvens) til art og begge var bare tilstede på 0-2 m. De to OTUene tilhørte ciliaten *Pseudocohnilembus persalinus* og dinoflagellaten *Oxyrrhis marina*.

Pseudocohnilembus persalinus er en ciliat som er i Korea kjent for å parasittere flyndren *Paralichthys olivaceus* (Kim mfl., 2004). Frittlevende *Pseudocohnilembus*-arter beiter på bakterier og i kultur svømmer nedover om de blir sultet, før de tilslutt kan danne cyster (Fenchel, 1990). Hvorfor denne tilsynelatende bunnlevende ciliaten bare blir observert i overflaten er vanskelig å forklare.

Oxyrrhis marina er en heterotrof dinoflagellat (Goldman mfl., 1989; Hansen mfl., 1996). Den er en euryhalin art, som vil si at den er i stand til å tåle store vekslinger i vannets saltholdighet. Den *Oxyrrhis marina*-arten som finnes i Østersjøen tåler et saltholdighetsspekter fra veldig fortynnet brakkvann til en saltholdighet på ca. 40 (Fenchel, 2005). En mulig preferanse for lavsaltholdig vann kan derfor ikke forklare hvorfor *O. marina* kun finnes i overflaten.

Hansen og medarbeidere (1996) viste at *O. marina* fortrekker å beite på haptofyten *Emiliania huxleyi* framfor annet fytoplankton. Som kan forklare hvorfor vi i prøven fra august 2017 så en mye høyere relativ abundans sekvenser som tilhørte rekken Haptophyta ved 5 m enn i overflaten.

Prøvene tatt i april 2017 viste en mye høyere relativ abundanse av sekvenser som tilhørte diatoméer i overflaten enn ved 5 m. Ved 5 m var den relative abundansen sekvenser som tilhørte dinoflagellater høyere enn ved 0-2m. Sprangsjiktet denne måneden lå på 4 m, og vi kan spekulere at den store mengden dinoflagellater trives godt under sprangsjiktet, som fører til at beitepresset på diatoméene i overflaten er mindre. En annen mulig forklaring kan være at vi i prøvetakningen har fanget opp en eller flere kjeder av *Chaetoceros* sp. (som var dominerende på 0-2 m i april 2017), mens i prøvene fra 5 m har vi ikke fanget opp like mange eller lange kjeder. Dette ville resulterert i færre sekvenser. En siste forklaring kan være PCR-feil. Det som oppformerer tidlig under PCR kan dominere prøven, siden hver PCR-syklus får mer og mer templat av det oppformerte produktet som kan brukes i neste syklus. På den annen siden har vi tatt høyde for skjev oppformering i PCR ved å ha tre PCR-replikaer per prøve, som gjør PCR-feil mindre sannsynlig.

Generelt var den ingen signifikant forskjell i prøvene mellom de to dypene, hverken i OTU-rikhet, diversitet eller jevnhet. Mulige forklaringer på dette kan være at dypene 0-2m og 5m er så nær hverandre at omrøring i de øvre vannmassene fører til at protistsamfunnet er tilnærmet homogent.

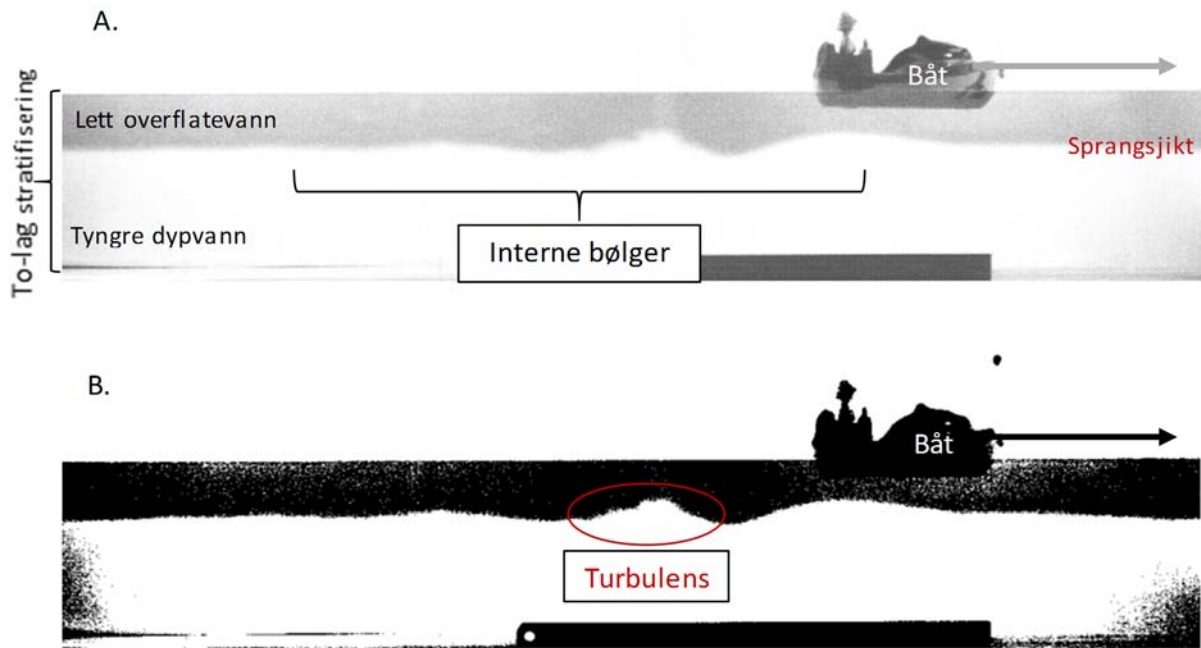
Omrøring i de øvre vannmassene

En type dyp omrøring av vannmassene forårsaket av vind er langmuir-sirkulasjon (Langmuir, 1938). Langmuir-sirkulasjon er en forbigående omrøring av de øvre vannmassene. Denne sirkulasjonen er ofte synlig i vannskorpen som parallelle «vind-rader» som består av bobler/skum og flytende detritus som ligger orientert i vindretning. Under vannoverflaten dannes celler med strømhvirvler, vertikalt 90 grader på vindretningen. Langmuir-sirkulasjon kan oppstå raskt, innen få ti-talls minutter om vindhastigheten er over 3 m/s (Talley, 2011). Den vertikale omrøringen når da vanligvis ned til 4-6 m dyp (Talley, 2011), men også dypere om vindhastigheten er større. Alle månedene, foruten juli 2016, hadde minst en dag (i snitt 4.2 dager) der vindforholdene lå til rette for at langmuir-sirkulasjon kunne oppstått.

Om sommeren er ikke vindhastigheten høy, men andre faktorer som temperaturendringer i atmosfæren bidrar til vertikal omrøring i de øvre vannmassene (Moum & Smyth, 2001). Om dagen varmes de øvre vannmassene av solinnstråling. Om natten synker temperaturen i lufta, som fører til vannoverflaten avgir varme og avkjøles. Avkjølingen av vannoverflaten induserer dannelsen av konveksjonsceller, der det øvre vannet synker ned til vann med lik temperatur og overflatevannet erstattes med vann fra underliggende lag.

En annen ting som skaper vertikal omrøring i de øvre vannmassene er turbulens i tidevannet (Reun mfl., 2018). Vannmassene er utsatt for en omrøring hver gang de går ut og inn av fjorden. I tillegg er det flere undersjøiske terskler i Indre Oslofjord. Tidevannet presses over disse tersklene på vei inn eller ut av fjorden. Dette kan skape indre bølger i vannsøylen som forplanter seg innover fjorden (Baalsrud & Magnusson, 2002) – «lee bølger» – som kan føre til turbulens mellom de stratifiserte vannlagene (Farmer & Dungan Smith, 1980).

De øvre vannmassene påvirkes ikke bare av naturlige fysiske krefter, men også menneskeskapte. Ekman beskrev i sin doktoravhandling i 1904 «dead water», et fenomen som oppstår når et lag ferskere vann ligger over et saltere (Mercier mfl., 2011; Walker, 1991). Når en båt kjører i stratifiserte vannmasser skaper ikke båten bare bølger i overflaten, men også indre bølger i det saltere og tettere vannlaget under det ferskere overflate laget (figur 4.2). Dannelsen av de indre bølgene krever energi, som blir overført fra for eksempel motorkreftene i en båtpropell. De indre bølgene kan skape turbulens i og mellom de to vannlagene. Fritidsbåter kan derfor bidra til ikke bare omrøring i vannmassene over sprangsjiktet, men også mellom.



Figur 4.2. Dead water-effekten demonstrert ved å dra en leke-båt over et kar med vann stratifisert i to lag, et farget lavt-saltholdig lett vannlag over et ufarget salt og tungt vannlag. A og B viser nøyaktig samme bilde hvor (B) har høyere kontrast enn (A). Figur er modifisert fra Mercier mfl., 2011.

Båter som går dypere i vannsøylen enn sprangsjiktet i vil også føre til blanding i og mellom vannlagene. Oslofjorden er en trafikkert fjord, med både tankskip og ferjer som går ut og inn av fjorden omtrent daglig.

Vannmassenes sirkulasjon er kompleks og det er flere faktorer som spiller inn enn de få som her er gjort rede for (se Moum og Smyth 2001 for flere faktorer). I det øvre blandingslaget er vannmassene utsatt for stadig omrøring på grunn av flere fysiske krefter som påvirker samtidig og innen korte tidsrom. Datasimuleringer har vist at lateral røring av vannmasser kan gruppere fytoplanktonet horisontalt flekkvis (phytoplankton patchiness på engelsk) (Abraham, 1998). Dette er også vist med satellittbilder (Gower mfl., 2003; Martin mfl., 2003).

I våre data kan vi ikke se en sammenheng mellom dybden på det øvre blandingslaget og forskjell i prøvene (ordinasjonsavstand) ved de to dypene. Størst ordinasjonsavstand så vi i september 2016, da det øvre blandingslaget var på 10 m. Flere protister har god egenbevegelse eller andre tilpassninger som gjør at de kan regulere sin plassering i vannsøylen. Det er vanlig å finne et dyp der fytoplanktontettheten er høy, kalt det dype klorofyll maksimum. I metastrekkodingstudie til Gran-Stadniczeňko og medarbeidere (2018) sammelignet de prøver fra overflaten med prøver fra det dype klorofyll maksimum. De kunne påvise en signifikant

forskjell i antall OTUer men ikke i diversitet eller jevnhet. Vi kan ut ifra dette anta at vi ikke har tatt prøver fra klorofyll maksimum og at det øvre blandingslaget har vært blandet. Det ble i vårt studiet vist at protistsamfunnet på de to dypene ikke er nevneverdig forskjellig selv om man fjerner dinoflagellatene fra sammenligningen.

4.5 Metodekritikk

I denne studien fulgte vi samme prosedyre for innsamling av vannprøver som brukes under Ocean Sampling Day (Kopf mfl., 2015). Denne metoden har vært lett å gjennomføre og det viktigste har vært å unngå kontaminering av prøver. Dette kan unngås ved bruk av hansker, sørge for at utstyr er godt vasket og rent før bruk.

Ved å se på DNA og ikke RNA, risikerer vi å sekvensere ikke-levende DNA. Om vi hadde brukt RNA kunne vi muligens unngått at celler med stort genom dominerer prøvene (Not, mfl., 2009). Om vi skulle sett på RNA krever dette en annen håndtering av filterne, som rask nedkjøling i flytende nitrogen etter filtrering for å forhindre degradering av RNAet. Dette er en prosedyre som gjør metoden vanskeligere å gjennomføre og målet har vært å undersøke om dette er en metode som lar seg enkelt bruke i overvåkning. DNA er mer robust og tåler mer også etter isolering, som gjør at DNA er mer egnet enn RNA.

Pre-filtrering

I det opprinnelige datasettet tilhørte 12,5% av OTUene og 39% av sekvensene dyr, planter eller sopp. Det var spesielt to OTUer innen klassen Maxillopoda som var meget abundant. For å fjerne store zooplankton kunne vi prefiltrert vannprøvene, som er vanlig prosedyre i flere studier på protist-diversitet (Egge mfl., 2015a; Lovejoy mfl., 2006; Massana, 2011; Massana mfl., 2015; Not mfl., 2004; Seenivasan, 2013).

Våre vannprøver ble filtrert direkte på et 0.22µm filter, som fanget opp alt plankton større enn dette. Størrelsen til protister spenner over et stort spekter, hvor majoriteten av marine pelagiske protister faller innen størrelsesordenen pico- til mikroplankton (0.20-200 µm) (Azam & Malfatti, 2007; Brewin mfl., 2010; Takahashi & Bienfang, 1983). Størrelsen på prefiltreringssilen kan bestemmes etter hvilken størrelsesorden av protister man ønsker å studere, men man må være bevisst at en prefiltrering kan fjerne store arter og arter som danner lange kjeder (Gran-Stadniczeňko mfl., 2018), som diatomeene *Skeletonema* sp. (Takabayashi mfl., 2006; Throndsen mfl., 2007) og *Chaetoceros* sp. (Fryxell, 1978; Throndsen mfl., 2007).

V4-regionen i 18S rRNA-genet som markør

Markøren har i denne studien vist seg å fungere godt til å skille store gruppene, men begrenser klassifisering til art innen noen slekter som *Dinophysis*. CBOL-protist arbeidsgruppen foreslår en tosteg strekkodingsmetode, der man først skiller de store gruppene ved bruk av V4-regionen som markør for deretter benytte en rekke gruppe-spesifikk markør for klassifisering på artsnivå (Pawlowski mfl., 2012).

Primere

Alle supergruppene som ble forventet å bli påvist var tilstede, utenom Excavata. I supergruppen Excavata finner vi rekken Euglenozoa (Heyden mfl., 2004) og en vanlig euglenoid-art i Oslofjorden er *Eutreptiella braarudii* (Thronsen mfl., 2007). Det ble talt *Eutreptiella*-arter i lysmikroskop fra overvåkningsprøver fra mars - mai 2017, tatt på samme lokalitet som våre prøver (Norli & Eikrem, 2018). At vi ikke fanget opp denne arten kan bero på primerene som ble brukt ikke var generelle eller spesifikke nok. Gruppen Foraminifera ble heller ikke påvist. Foraminifera tilhører supergruppen Rhizaria, som fordi de har en mye lenger V4-region enn andre eukaryote har vist seg vanskelig å fange opp ved bruk av universale eukaryote 18S rRNA-primere (Pawlowski, 2000; Pawlowski mfl., 2003).

OTU

Et viktig problem er hvordan å definere en art i metastrekkoding-data (Pawlowski mfl., 2016). For øyeblikket er den molekylære ekvivalenten for en art en OTU (også kalt MOTU – Molecular Operational Taxonomic Unit) (Godfray mfl., 2004). Vi definerte en OTU til sekvenser med 98% likhet. I andre studier er 95% (Caron mfl., 2009), 97% (Balzano mfl., 2015) eller 99% (Hu mfl., 2016) brukt som terskelverdier. For protister er det vanskelig å bestemme en universell terskelverdi for en OTU, siden protister har evolusjonært utviklet seg i forskjellig hastighet (Pernice mfl., 2013). En mulighet for å overkomme dette problemet er å bruke forskjellige grenseverdier avhengig av hvilken gruppe protister man ønsker å studere (Pawlowski mfl., 2016).

Referansedatabasen PR²

En av de største utfordringene vi møtte på var begrenset tilgang på sekvenser som var taksonomisk klassifisert i referansedatabasen PR². Databaser blir daglig forbedret med flere

sekvenser og kuratert taksonomisk klassifisering. Bruk av flere referansedatabaser kunne vært en mulighet.

PR² inneholder noen feil som gjør arbeid med store datasett litt vanskeligere. Slik som at *Syndiniales* var på taksonomisk nivå klasse og ikke orden.

De 22 påviste protistgruppene tilhørte 8 supergrupper, som ble definert etter det taksonomiske klassifiseringssystemet i PR² (Guillou mfl., 2013). I PR² finnes det 17 supergrupper (se tabell 1 i Guillou mfl., 2013) som gjør sammenligning med annen litteratur noe forvirrende. Supergruppene kan ansees som hovedgreinene i det eukaryote treet. De har gjennomgått en betydelig omorganisering de siste 20 årene (Burki, 2014; Burki mfl., 2007) og vi kan forvente flere omveltninger i fremtiden. Taksonomisk klassifisering er ofte oppe til debatt ettersom molekylære studier stadig avslører nytt slektskap mellom grupper (Massana mfl., 2004; Nikolaev mfl., 2004; Seenivasan, 2013). Stamenopiles, Alveolata og Rhizaria som her omtales som tre supergrupper, blir i annen litteratur klassifisert som superrekker som samlet utgjør supergruppen «SAR» (Adl mfl., 2012; Burki, 2014; Burki mfl., 2007; He mfl., 2016). Supergruppen Apusozoa blir i annen litteratur ikke omtalt som en supergruppe, men trolig en parafyletisk søstergruppe til Opisthokonta som er sammen med Amoebozoa foreslått å utgjøre supergruppen Unikonta (Cavalier-Smith & Chao, 2010). Det er viktig at referansedatabaser er oppdatert og det er nødvendig å etterstrebe et universelt klassifiseringssystem.

Tolkning av antall sekvenser

Et annet kontroversielt aspekt ved bruk av metastrekkoding er hvordan tolke de kvantitative sekvensdataene. Noen økologiske studier viser at et er uoverensstemmelse mellom antallet celler talt i lysmikroskop og den relative abundansen i sekvenser (Egge mfl., 2013; Medinger mfl., 2010). Det er derfor et behov for flere studier som undersøker sammenhengen mellom antallet celler den relative abundansen sekvenser til forskjellige grupper protister. Jo mer kunnskap vi får om dette, jo nærmere kommer vi utviklingen av en prosedyre for behandling og tolkning av metastrekkoding-data.

Det ble det ikke direkte undersøkt om metoden vi brukte er repliserbar. Siden det ble tatt reserveprøver for hver prøve, ville det vært fullt mulig å ekstrahere DNA fra en eller flere av disse som kontroll av metoden. På den annen side viste våre resultater at prøvene tatt på 0-2 m og 5 m ikke var signifikant forskjellig. Dette kan indikerere at metoden er repliserbar.

4.6 Fremtidige studier og konklusjon

Datasettet som vi sitter igjen med etter denne studien er stort. Det inneholder et enormt potensiale for videre undersøkelser. Jeg har kun studert dette datasettet i lys av spørsmålene jeg stilte. Jo mer jeg jobbet med og ble kjent med dette datasettet, jo flere muligheter så jeg at dette datasettet har. Dette var mye av grunnen til at jeg ønsket å legge hele datasettet ut på figshare. Andre øyne vil kanskje ønske å analysere andre ting enn det jeg foreslår under.

4.6.1 Hadde jeg hatt mer tid med datasettet ville jeg

Undersøkt diversiteten på nytt

I den bioinformatiske pipelinen vi brukte ble OTUene taksonomisk klassifisert etter de hadde blitt klustret på 98% likhet. Ved å taksonomisk klassifisere sekvensene før man klustrer kan man unngå at taksa blir gjemt i OTUen. Dette kalles fylo-basert metode. Sekvensene blir først blir søkt opp mot en referansedatabase og taksonomisk merket, deretter slått sammen etter hvor like merkelappene er (Schloss & Westcott, 2011), ikke hvor like sekvensene er som ble gjort med våre data.

Det hadde vært interessant å sammenlignet resultatene våre med resultatene vi hadde fått fra å bruke en fylo-basert metode.

Analysert dataene i overvåkningsøyemed

Hva sier protistsammensetningen og de hydrografiske målingene i våre rå-data om eutrofiering og klimaendringer? Jeg ville samlet overvåkningsrapporter og trendanalyser som allerede er gjort for Indre Oslofjord, undersøkt indikatorarter for eutrofiering og klimaendringer og satt det i sammenheng med hvordan fjorden har utviklet seg spesielt de siste 50 årene.

Sett på parasitt-vert relasjoner

Det ble i våre data påvist en stor del OTUer som tilhørte parasitter. Jeg ville undersøkt om parasitter har en regulerende effekt på protistsamfunnets og om de er med på å påvirke sesongvariasjon i protistsamfunnet.

I pipelinen ble litt over 6 000 sekvenser som potensielt tilhørte 16S rRNA-genet i bakterier filtrert bort (tabell S.2). Med bakterie-sekvensene i tillegg til det opprinnelige datasettet kan man se på parasitt-vert relasjoner som også inkluderer bakterier, dyr og sopp.

Sammenligne med lysmikroskopdata fra samme prøver

For hver dato vi var ute og samlet prøver ble det i tillegg samlet inn prøver til analyse i lysmikroskop. Ph.d student Elisabeth Lundsør vil i sitt prosjekt telle celler fra disse prøvene og det hadde vært spennende å undersøkt hennes resultater sammen med våre. Påviser de to metodene de samme taksa når det er samme vannprøvene som blir analysert?

Sett på arter

Den høye OTU-diversiteten påvist om vinteren er forslått i diskusjonen å bero på tilførsel av arter som kommer med vannfornøyelse eller fra dypere vannlag. Ved å studere litteratur på artene OTUene i vinterprøvene ble klassifisert til, kan man faktasjekke det overnevnte argumentet ved å potensielt oppdage dypvannsarter.

Fylogenetiske studier

Mange av OTUene i protistdatasettet ble f.eks. kun klassifisert til klasse. Som 107 unike OTUer som tilhørte klassen Dinophyceae. Fylogenetiske analyser kan potensielt plassere OTUene nærme kjente arter og gi mer informasjon om protistsammensetningen.

Om det er mye man ikke kan fylogenetisk plassere kan man kanskje få et innblikk i hva som ikke enda er oppdaget i Oslofjorden.

Sett på sjeldne OTUer

De aller fleste OTUene (88%) påvist i denne studie var sjeldne (<0.05% sekvenser per OTU). Det er her blitt fokusert på OTUer med for det meste høy relativ abundans, som innebærer at det er veldig lite av diversiteten som er blitt fremstilt og diskutert.

Jeg ville fjernet alle abundante OTUer fra datasettet. Deretter analysert og fremstilt mange av dataene på nytt for å undersøkt hvordan samfunnet så ut om man kun så på det sjeldne OTUene.

Nye arter for Norge/Oslofjorden

Vi undersøkte ikke om noen av de klassifiserte OTUene representerte arter nye for Norge eller Oslofjorden. Dette ville jeg undersøke ved å gjennomgå artslistene for Norge og Oslofjorden og sammenlignet med våre funn.

Homogenitet i DNA-ekstraktet

Vi har fortsatt de opprinnelige 30 ekstraherte DNA-prøvene i fryseren på Blindern. Det var kun en liten del av prøven tatt ut til videre metastrekkoding-analyse. Ved å kjøre nøyaktig samme

prosedyre på en ny del av den resterende prøven, kan man undersøke om resultatene viser det samme. Gjør de det, vil det bekrefte at det er 2 L vann vi analyserer. Gjør de ikke det, er det rom for forbedring av prosedyren. Dette tror jeg er viktig for å undersøke hvordan best mulig utvikle en standard prosedyre som kan brukes globalt.

Av våre prøver hadde mange en så høy DNA-konsentrasjon ($>5 \text{ ng}/\mu\text{L}$) at de ble fortynnet før PCR (tabell S.1) De tre PCR-replikaene var fra samme fortynnede prøve. Jeg mistenker at man mister DNA-diversitet i fortynning, spesielt fra arter som forekommer i lav abundans. Ved å lage flere fortynninger fra samme prøve f.eks. 5, for deretter kjøre PCR med tre replikaer på hver av de 5 fortynningene kan man undersøke om diversiteten er forskjellig i de 5 prøvene. Det er viktig å da inkludere minst ett replika på en av fortynningene som kontroll.

Hva er OTU 29?

OTU 29 var den mest abundante OTUen i hele protistdatasettet og ble ikke klassifisert nærmere enn til en *Gyrodinium* sp. Så hvilken art er det egentlig?

Under celledtelling i overvåkingen blir det observert en *Gyrodinium* som minner om *Gyrodinium helveticum*, som det i perioder er mye av (Wenche Eikrem, personlig kommunikasjon, 27 september 2018).

Dette underbygger at det i praksis skal være mulig å isolere enkelt celler fra ufikserte vannprøver prøver, oppformere 18S rRNA-genet og sekvensere. Celler som mulig representerer OTU 29 burde i tillegg observeres i lysmikroskop og elektron mikroskop for morfologisk karakterisering.

4.6.2 Hvordan drive overvåking i framtiden

Metode

Jeg vil anbefale en kombinasjon av metastrekkoding og lysmikroskop. Våre resultater har vist at metastrekkoding har et enormt potensiale for å undersøke arter som ikke kan observeres i lysmikroskop. Det er også en metoden som vil gjøre det lettere å registrere nye innvaderende arter på et tidlig tidspunkt.

Metoden metastrekkoding har også sine begrensinger, spesielt for kvantitative analyser. En kombinasjon av disse metodene vil gi et helhetlig og realistisk bilde av protistsamfunnet sammensetning og biomasse. Ved å allerede nå starte overvåking med begge metoder, samler man data som kan brukes til å utvikle en best mulig prosedyre for fremtidens overvåking.

Man må også stille spørsmålet om det i overvåkning er viktig å identifisere til art. For noen giftige arter er det åpenbart det, men for de resterende? Hvis det er tilstrekkelig å ha kjennskap til biomassen av fytoplanktonet, kan en kombinasjon av klorofyll a målinger sammen med satellittovervåkning være en mulig metode (Nezlin mfl., 2010). Man kan også bruke qPCR for å påvise f.eks. giftige alger (Galluzzi mfl., 2004; Zhu mfl., 2005) eller biomassen til enkelte diatoméer (Godhe mfl., 2008).

Dyp

Jeg vil anbefale å avslutte overvåkning på 0-2 m. Dette er ressurser som heller kan bli brukt på å f.eks. ta flere prøver gjennom året og starte opp med metastrekkoding i tillegg til lysmikroskop.

4.6.3 Konklusjon

Denne studien har vist at metastrekkoding er en metode som er godt egnet som et supplement i den marine miljøovervåkingen. Den kan per i dag ikke erstatte den tradisjonelle overvåkingen med lysmikroskop, men kan bidra på områder der lysmikroskopi begrenses. Metastrekkoding er en metode som er tidkrevende og krever molekylær- og bioinformatisk-kompetanse. Men med gode prosedyrer og velfungerende pipeliner kan dette bli et verktøy som de fleste kan håndtere. Metoden metastrekkoding begrenses i dag hovedsakelig av referansedatabaser brukt til taksonomisk klassifisering av sekvensene i prøven og det er viktig å velge en markør som fanger opp det man ønsker å undersøke. Arbeide for å forbedre og oppdatere referansedatabaser er nødvendig, sammen med at flere arter identifiseres og sekvenseres. Ved å starte opp med et overlapp av begge metoder kan man på best grunnlag utvikle prosedyrer for fremtidens overvåking.

Resultatene våre kan bidra til bedre forståelse av protistsamfunnets sammensetning gjennom året i kystnære områder og de påviste protistenes globale utbredelse. Jeg håper dette er starten på et enormt datasett som det vil bli bygget på i lang tid fremover.

Litteraturliste

- Abraham, E. R. (1998). The generation of plankton patchiness by turbulent stirring. *Nature*, **391**:577–580.
- Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology*, **71**:8966–8969.
- Adl, S. M., Leander, B. S., Simpson, A. G. B., Archibald, J. M., Anderson, O. R., Bass, D., ... Spiegel, F. (2007). Diversity, nomenclature, and taxonomy of protists. *Systematic Biology*, **56**, 684–689.
- Adl, S. M., Simpson, A. G. B., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., ... Taylor, M. F. J. R. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology*, **52**:399–451.
- Adl, S. M., Simpson, A. G. B., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., ... Spiegel, F. W. (2012). The revised classification of eukaryotes. *The Journal of Eukaryotic Microbiology*, **59**:429–493.
- Allredge, A. L., & Silver, M. W. (1988). Characteristics, dynamics and significance of marine snow. *Progress in Oceanography*, **20**:41–82.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.
- Alverson, A. J., Beszteri, B., Julius, M. L., & Theriot, E. C. (2011). The model marine diatom *Thalassiosira pseudonana* likely descended from a freshwater ancestor in the genus *Cyclotella*. *BMC Evolutionary Biology*, **11**:125.
- Amato, A., Kooistra, W. H. C. F., Levialedi Ghiron, J. H., Mann, D. G., Pröschold, T., & Montresor, M. (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist*, **158**:193–207.
- Anderson, D. M., Kulis, D. M., Sullivan, J. J., & Hall, S. (1990). Toxin composition variations in one isolate of the dinoflagellate *Alexandrium fundyense*. *Toxicon*, **28**:885–893.
- Armbrust, E. V. (2009). The life of diatoms in the world's oceans. *Nature*, **459**:185–192.
- Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., ... Rokhsar, D. S. (2004). The Genome of the diatom *Thalassiosira Pseudonana*: Ecology, evolution, and metabolism. *Science*, **306**:79–86.

- Arrigo, K. R., Perovich, D. K., Pickart, R. S., Brown, Z. W., Dijken, G. L. van, Lowry, K. E., ... Swift, J. H. (2012). Massive phytoplankton blooms under Arctic sea ice. *Science*, **336**:1408–1408.
- Artsdatabanken.no. (2018). Besøkt november 2018.
- Azam, F., & Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology*, **5**:782–791.
- Azevedo, C. (1989). Fine structure of *Perkinsus atlanticus* n. sp. (Apicomplexa, Perkinsea) parasite of the clam *Ruditapes decussatus* from Portugal. *The Journal of Parasitology*, **75**:627–635.
- Baldauf, S. L. (2008). An overview of the phylogeny and diversity of eukaryotes. *Journal of Systematics and Evolution*, **46**:263–273.
- Balzano, S., Abs, E., & Leterme, S. C. (2015). Protist diversity along a salinity gradient in a coastal lagoon. *Aquatic Microbial Ecology*, **74**:263–277.
- Bates, S. S., Garrison, D. L., & Horner, R. A. (1998). Bloom dynamics and physiology of domoic-acid-producing *Pseudo-nitzschia* species. *Physiological Ecology of Harmful Algal Bloom*, 267–292.
- Berge, J., Amundsen, R., Gitmark, J., Ledang, H., Lømsdal, E., Magnusson, J., ... Holth, T. (2015). *Overvåking av Indre Oslofjord i 2014* (NIVA rapport No. 6833–2015). Norsk institutt for vannforskning. Hentet fra: [file:///C:/Users/Karoline/Downloads/6833_IndreOslofjord2014%20\(1\).pdf](file:///C:/Users/Karoline/Downloads/6833_IndreOslofjord2014%20(1).pdf)
- Berney, C., Fahrni, J., & Pawlowski, J. (2004). How many novel eukaryotic «kingdoms»? Pitfalls and limitations of environmental DNA surveys. *BMC Biology*, **2**:13.
- Bienfang, P. K., Harrison, P. J., & Quarmby, L. M. (1982). Sinking rate response to depletion of nitrate, phosphate and silicate in four marine diatoms. *Marine Biology*, **67**:302.
- Bolch, C. J. S., Negri, A. P., & Hallegraeff, G. M. (1999). *Gymnodinium microreticulatum* sp. nov. (Dinophyceae): a naked, microreticulate cyst-producing dinoflagellate, distinct from *Gymnodinium catenatum* and *Gymnodinium nolleri*. *Phycologia*, **38**:301–313.
- Boutrup, P. V., Moestrup, Ø., Tillmann, U., & Daugbjerg, N. (2016). *Katodinium glaucum* (Dinophyceae) revisited: proposal of new genus, family and order based on ultrastructure and phylogeny. *Phycologia*, **55**:147–164.
- Bouvier, T. C., & Giorgio, P. A. del. (2002). Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. *Limnology and Oceanography*, **47**:453–470.
- Boyd, C., & Gradmann, D. (2002). Impact of osmolytes on buoyancy of marine phytoplankton. *Marine Biology*, **141**:605–618.
- Brainerd, K. E., & Gregg, M. C. (1995). Surface mixed and mixing layer depths. *Deep Sea Research Part I: Oceanographic Research Papers*, **42**:1521–1543.

- Bravo, I., Garcés, E., Diogène, J., Fraga, S., Sampedro, N., & Figueroa, R. I. (2006). Resting cysts of the toxigenic dinoflagellate genus *Alexandrium* in recent sediments from the Western Mediterranean coast, including the first description of cysts of *A. kutnerae* and *A. peruvianum*. *European Journal of Phycology*, **41**:293–302.
- Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., & Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, **221**:1472–1483.
- Brown, E. A., Chain, F. J. J., Zhan, A., MacIsaac, H. J., & Cristescu, M. E. (2016). Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. *Diversity and Distributions*, **22**:1045–1059.
- Bruno, S. F., Staker, R. D., Sharma, G. M., & Turner, J. T. (1983). Primary productivity and phytoplankton size fraction dominance in a temperate North Atlantic estuary. *Estuaries*, **6**:200–211.
- Braarud, T., & Nygaard, I. (1967). *Fytoplankton*. (Delrapport nr. 4). Oslo: Universitet i Oslo, Institutt for Marin Biologi.
- Bråte, J., Krabberød, A. K., Dolven, J. K., Ose, R. F., Kristensen, T., Bjørklund, K. R., & Shalchian-Tabrizi, K. (2012). Radiolaria associated with large diversity of marine alveolates. *Protist*, **163**:767–777.
- Buck, K. R., & Garrison, D. L. (1988). Distribution and abundance of choanoflagellates (Acanthoecidae) across the ice-edge zone in the Weddell Sea, Antarctica. *Marine Biology*, **98**:263–269.
- Burki, F. (2014). The Eukaryotic Tree of Life from a global phylogenomic perspective. *Cold Spring Harbor Perspectives in Biology*, **6**:a016147.
- Burki, F., Inagaki, Y., Bråte, J., Archibald, J. M., Keeling, P. J., Cavalier-Smith, T., ... Shalchian-Tabrizi, K. (2009). Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. *Genome Biology and Evolution*, **1**:231–238.
- Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjæveland, Å., Nikolaev, S. I., Jakobsen, K. S., & Pawlowski, J. (2007). Phylogenomics reshuffles the eukaryotic supergroups. *PloS one*, **2**:e790.
- Burreson, E. M., & Ragone Calvo, L. M. (1996). Epizootiology of *Perkinsus marinus* disease of oysters in Chesapeake Bay, with emphasis on data since 1985. *Journal of Shellfish Research*, **15**:17–34.
- Baalsrud, K., & Magnusson, J. (2002). *Indre Oslofjord Natur og miljø*. Oslo: Fagrådet for vann- og avløpsteknisk samarbeid i indre Oslofjord.
- Calbet, A., & Landry, M. R. (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography*, **49**:51–57.

- Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**:1972–1973.
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., ... Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**:1621–1624.
- Caron, D. A., Countway, P. D., Jones, A. C., Kim, D. Y., & Schnetzer, A. (2012). Marine protistan diversity. *Annual Review of Marine Science*, **4**:467–493.
- Caron, D. A., Countway, P. D., Savai, P., Gast, R. J., Schnetzer, A., Moorthi, S. D., ... Jones, A. C. (2009). Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Applied Environmental Microbiology*, **75**:5797–5808.
- Cavalier-Smith, T. (1993). Kingdom Protozoa and its 18 phyla. *Microbiological Reviews*, **57**:953–994.
- Cavalier-Smith, T., & Chao, E. E.-Y. (2010). Phylogeny and evolution of Apusomonadida (Protozoa: Apusozoa): New genera and species. *Protist*, **161**:549–576.
- Cavalier-Smith, T., & Chao, E. E.-Y. (2003a). Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist*, **154**:341–358.
- Cavalier-Smith, T., & Chao, E. E.-Y. (2003b). Phylogeny of Choanozoa, Apusozoa, and other Protozoa and early eukaryote megaevolution. *Journal of Molecular Evolution*, **56**:540–563.
- Cavalier-Smith, T., Chao, E. E.-Y., & Oates, B. (2004). Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont *Phalansterium*. *European Journal of Protistology*, **40**:21–48.
- Chambouvet, A., Laabir, M., Sengco, M., Vaquer, A., & Guillou, L. (2011). Genetic diversity of Amoebophryidae (Syndiniales) during *Alexandrium catenella/tamarensis* (Dinophyceae) blooms in the Thau lagoon (Mediterranean Sea, France). *Research in Microbiology*, **162**:959–968.
- Chan, A. T. (1980). Comparative physiological study of marine diatoms and dinoflagellates in relation to irradiance and cell size. ii. relationship between photosynthesis, growth, and carbon/chlorophyll a ratio. *Journal of Phycology*, **16**:428–432.
- Ciancio, A., Scippa, S., & Cammarano, M. (2001). Ultrastructure of trophozoites of the gregarine *Lankesteria ascidia* (Apicomplexa: Eugregarinida) parasitic in the ascidian *Ciona intestinalis* (Protochordata). *European Journal of Protistology*, **37**:327–336.
- Coats, D. W., Adam, E., Gallegos, C., & Hedrick, S. (1996). Parasitism of photosynthetic dinoflagellates in a shallow subestuary of Chesapeake Bay, USA. *Aquatic Microbial Ecology*, **11**:1–9.
- Coats, D. W., Bachvaroff, T. R., & Delwiche, C. F. (2012). Revision of the Family Duboscquellidae with Description of *Euduboscquella crenulata* n. gen., n. sp.

- (Dinoflagellata, Syndinea), an intracellular parasite of the ciliate *Favella panamensis* Kofoid & Campbell. *Journal of Eukaryotic Microbiology*, **59**:1–11.
- Conley, D., Schelske, C., & Stoermer, E. (1993). Modification of the biogeochemical cycle of silica with eutrophication. *Marine Ecology Progress Series*, **101**:179–192.
- Corliss, J. O. (1984). The kingdom Protista and its 45 phyla. *Biosystems*, **17**:87–126.
- Costanza, R., d'Arge, R., Groot, R. de, Farber, S., Grasso, M., Hannon, B., ... Belt, M. van den. (1997). The value of the world's ecosystem services and natural capital. *Nature*, **387**:253–260.
- Daugbjerg, N., Andreasen, T., Happel, E., Pandeirada, M. S., Hansen, G., Craveiro, S. C., ... Moestrup, Ø. (2014). Studies on woloszynskioid dinoflagellates VII. Description of *Borghiella andersenii* sp. nov.: light and electron microscopy and phylogeny based on LSU rDNA. *European Journal of Phycology*, **49**:436–449.
- de Salas, M. F., Bolch, C. J. S., Botes, L., Nash, G., Wright, S. W., & Hallegraeff, G. M. (2003). *Takayama* gen. nov. (Gymnodiniales, Dinophyceae), a new genus of unarmored dinoflagellates with sigmoid apical grooves, including the description of two new species1. *Journal of Phycology*, **39**:1233–1246.
- de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., ... Probert, I. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, **348**:1261605.
- DeAngelis, M. M., Wang, D. G., & Hawkins, T. L. (1995). Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Research*, **23**:4742–4743.
- Diaz, R. J., & Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. *Science*, **321**:926–929.
- Dogde, J. D. (1982). *Marine Dinoflagellates of the British Isles*. London: HM Stationery Office.
- Doughman, S. D., Krupanidhi, S., & Sanjeevi, C. B. (2007). Omega-3 Fatty Acids for Nutrition and Medicine: Considering Microalgae Oil as a Vegetarian Source of EPA and DHA. *Current Diabetes Reviews*, **3**:198-203
- Drebes, G., Kühn, S. F., Gmelch, A., & Schnepf, E. (1996). *Cryothecomonas aestivalis* sp. nov., a colourless nanoflagellate feeding on the marine centric diatom *Guinardia delicatula* (Cleve) Hasle. *Helgoländer Meeresuntersuchungen*, **50**:497–515.
- Duinker, A., Lunestad, B. T., Roiha, I. S., & Måge, A. (2016). *Nasjonalt tilsynsprogram for produksjon av skjell og andre bløtdyr – prøver analysert i 2015: Kjemiske forurensende stoffer og mikroorganismer*. (Skjell rapport nr. 978-82-91065-26-7). Oslo: Nasjonalt institutt for ernærings- og sjømatforskning.
- Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics*, **14**:755–763.

- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**:2460–2461.
- Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, **10**:996–998.
- Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, **31**:3476–3482.
- Edgar, R., Domrachev, M., & Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, **30**:207–210.
- Edwardsen, B., Eikrem, W., Green, J. C., Andersen, R. A., Moon-van der Staay, S. Y., & Medlin, L. K. (2000). Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia*, **39**:19–35.
- Edwardsen, B., Eikrem, W., Throndsen, J., Sáez, A. G., Probert, I., & Medlin, L. K. (2011). Ribosomal DNA phylogenies and a morphological revision provide the basis for a revised taxonomy of the Prymnesiales (Haptophyta). *European Journal of Phycology*, **46**:202–228.
- Edwardsen, B., & Paasche, E. (1992). Two motile stages of *Chrysochromulina Polylepis* (prymnesiophyceae): morphology, growth, and toxicity. *Journal of Phycology*, **28**:104–114.
- Edwardsen, B., Shalchian-Tabrizi, K., Jakobsen, K. S., Medlin, L. K., Dahl, E., Brubak, S., & Paasche, E. (2003). Genetic variability and molecular phylogeny of *Dinophysis* species (dinophyceae) from norwegian waters inferred from single cell analyses of rDNA1. *Journal of Phycology*, **39**:395–408.
- Egge, E. S., Bittner, L., Andersen, T., Audic, S., Vargas, C. de, & Edwardsen, B. (2013). 454 Pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. *PloS one*, **8**:e74371.
- Egge, E. S., Eikrem, W., & Edwardsen, B. (2015). Deep-branching novel lineages and high diversity of Haptophytes in the Skagerrak (Norway) uncovered by 454 pyrosequencing. *Journal of Eukaryotic Microbiology*, **62**:121–140.
- Egge, E. S., Johannessen, T. V., Andersen, T., Eikrem, W., Bittner, L., Larsen, A., ... Edwardsen, B. (2015a). Seasonal diversity and dynamics of haptophytes in the Skagerrak, Norway, explored by high-throughput sequencing. *Molecular Ecology*, **24**:3026–3042.
- Eikrem, W., & Throndsen, J. (1990). The ultrastructure of *Bathycoccus* gen. nov. and *B. prasinus* sp. nov., a non-motile picoplanktonic alga (Chlorophyta, Prasinophyceae) from the Mediterranean and Atlantic. *Phycologia*, **29**:344–350.
- Eikrem, W., & Throndsen, J. (1998). Morphology of *Chrysochromulina leadbeateri* (Prymnesiophyceae) from northern Norway. *Phycologia*, **37**:292–299.

- Ellegaard, M., Daugbjerg, N., Rochon, A., Lewis, J., & Harding, I. (2003). Morphological and LSU rDNA sequence variation within the *Gonyaulax spinifera*-*Spiniferites* group (Dinophyceae) and proposal of *G. elongata* comb. nov. and *G. membranacea* comb. nov. *Phycologia*, **42**:151–164.
- Eppley, R. W. (1972). Temperature and phytoplankton growth in the sea. *Fishery Bulletin*, **70**:1063–1085.
- Espeland, G., & Throndsen, J. (1986). Flagellates from Kilsfjorden, southern Norway, with description of two new species of Choanoflagellida. *Sarsia*, **71**:209–226.
- Evans, K. M., Wortley, A. H., & Mann, D. G. (2007). An assessment of potential diatom “barcode” genes (cox1, rbcL, 18S and ITS rDNA) and their Effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist*, **158**:349–364.
- Falkowski, P. G., Barber, R. T., & Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science*, **281**:200–206.
- Farmer, D. M., & Dungan Smith, J. (1980). Tidal interaction of stratified flow with a sill in Knight Inlet. *Deep Sea Research Part A. Oceanographic Research Papers*, **27**:239–254.
- Fenchel, T. (1990). Adaptive significance of polymorphic life cycles in Protozoa: responses to starvation and refeeding in two species of marine ciliates. *Journal of Experimental Marine Biology and Ecology*, **136**:159–177.
- Fenchel, T. (2005). Cosmopolitan microbes and their ‘cryptic’ species. *Aquatic Microbial Ecology*, **41**:49–54.
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, **281**:237–240.
- Figuerola, R. I., Bravo, I., & Garcés, E. (2005). Effects of nutritional factors and different parental crosses on the encystment and excystment of *Alexandrium catenella* (Dinophyceae) in culture. *Phycologia*, **44**:658–670.
- Fryxell, G. A. (1978). Chain-forming diatoms: three species of Chaetoceraceae. *Journal of Phycology*, **14**:62–71.
- Gall, L. L., & Saunders, G. W. (2010). DNA barcoding is a powerful tool to uncover algal diversity: a case study of the Phylloporaceae (Gigartinales, Rhodophyta) in the Canadian flora. *Journal of Phycology*, **46**:374–389.
- Galluzzi, L., Penna, A., Bertozzini, E., Vila, M., Garcés, E., & Magnani, M. (2004). Development of a real-time pcr assay for rapid detection and quantification of *Alexandrium minutum* (a dinoflagellate). *Applied Environmental Microbiology*, **70**:1199–1206.
- Gentekaki, E., & Lynn, D. H. (2009). High-level genetic diversity but no population structure inferred from nuclear and mitochondrial markers of the peritrichous ciliate

- Carchesium polypinum* in the Grand River Basin (North America). *Applied Environmental Microbiology*, **75**:3187–3195.
- Godfray H. C. J., Knapp S., & Blaxter Mark L. (2004). The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **359**:669–679.
- Godhe, A., Asplund, M. E., Härnström, K., Saravanan, V., Tyagi, A., & Karunasagar, I. (2008). Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Applied Environmental Microbiology*, **74**:7174–7182.
- Gold, K., Pfister, R. M., & Liguori, V. R. (1970). Axenic cultivation and electron microscopy of two species of choanoflagellida. *The Journal of Protozoology*, **17**:210–212.
- Goldman, J. C., Dennett, M. R., & Gordin, H. (1989). Dynamics of herbivorous grazing by the heterotrophic dinoflagellate *Oxyrrhis marina*. *Journal of Plankton Research*, **11**:391–407.
- Gómez, F., López-García, P., Takayama, H., & Moreira, D. (2015). *Balechina* and the new genus *Cucumeridinium* gen. nov. (Dinophyceae), unarmored dinoflagellates with thick cell coverings. *Journal of Phycology*, **51**:1088–1105.
- Gower, J. F. R., Denman, K. L., & Holyer, R. J. (1980). Phytoplankton patchiness indicates the fluctuation spectrum of mesoscale oceanic structure. *Nature*, **288**:157–159.
- Graham, L. E., Graham, J. M., & Wilcox, L. W. (2009). *Algae* (2. utg.). San Francisco, CA, Person Education, Inc.
- Gran-Stadniczeňko, S., Egge, E., Hostyeva, V., Logares, R., Eikrem, W., & Edvardsen, B. (2018). Protist diversity and seasonal dynamics in skagerrak plankton communities as revealed by metabarcoding and microscopy. *Journal of Eukaryotic Microbiology*, [epub. før print].
- Gran-Stadniczeňko, S., Šupraha, L., Egge, E. D., & Edvardsen, B. (2017). Haptophyte diversity and vertical distribution explored by 18s and 28s ribosomal rna gene metabarcoding and scanning electron microscopy. *Journal of Eukaryotic Microbiology*, **64**:514–532.
- Gu, H., Liu, T., Vale, P., & Luo, Z. (2013). Morphology, phylogeny and toxin profiles of *Gymnodinium inusitatum* sp. nov., *Gymnodinium catenatum* and *Gymnodinium microreticulatum* (Dinophyceae) from the Yellow Sea, China. *Harmful Algae*, **28**:97–107.
- Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., ... Worden, A. Z. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environmental Microbiology*, **10**:3349–3365.
- Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., ... Christen, R. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote

- Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, **41**:D597–D604.
- Haeckel, E. (1866). *Generelle morphologie der organismen*. Berlin: Reimer.
- Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'Agrosa, C., ... Watson, R. (2008). A global map of human impact on marine ecosystems. *Science*, **319**:948–952.
- Hansen, F., Witte, H., & Passarge, J. (1996). Grazing in the heterotrophic dinoflagellate *Oxyrrhis marina*: Size selectivity and preference for calcified *Emiliana huxleyi* cells. *Aquatic Microbial Ecology*, **10**:307–313.
- Hansen, G., & Daugbjerg, N. (2004). Ultrastructure of *Gyrodinium spirale*, the type species of *Gyrodinium* (Dinophyceae), including a phylogeny of *G. dominans*, *G. rubrum* and *G. spirale* deduced from partial LSU rDNA sequences. *Protist*, **155**:271–294.
- Hansen, P. (1991). Quantitative importance and trophic role of heterotrophic dinoflagellates in a coastal pelagial food web. *Marine Ecology Progress Series*, **73**:253–261.
- Harrison, P. J., Conway, H. L., Holmes, R. W., & Davis, C. O. (1977). Marine diatoms grown in chemostats under silicate or ammonium limitation. III. Cellular chemical composition and morphology of *Chaetoceros debilis*, *Skeletonema costatum*, and *Thalassiosira gravida*. *Marine Biology*, **43**:19–31.
- Hasle, G. R. (1950). Phototactic vertical migration in marine dinoflagellates. *Oikos*, **2**:162175.
- Havforskningsinstituttet. (2007). Kyst og havbruk 2007. Hentet fra: https://www.hi.no/filarkiv/2007/03/1.4_Overvakning_av_alger_langs_norskekysten.pdf/nb-no
- He, D., Sierra, R., Pawlowski, J., & Baldauf, S. L. (2016). Reducing long-branch effects in multi-protein data uncovers a close relationship between Alveolata and Rhizaria. *Molecular Phylogenetics and Evolution*, **101**:1–7.
- Hebert P. D. N., Cywinska A., Ball S. L., & deWaard J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**:313–321.
- Heyden, S. V. D., Chao, E. E., Vickerman, K., & Cavalier-Smith, T. (2004). Ribosomal RNA phylogeny of bodonid and diplomonad flagellates and the evolution of Euglenozoa. *Journal of Eukaryotic Microbiology*, **51**:402–416.
- Hill, M.O. (1979). *Decorana – a fortran program for detrended correspondance analysis and reciprocal averaging*. Cornell University, Ithaca, New York.
- Hill, M. O., & Gauch, H. G. (1980). Detrended Correspondence Analysis: an improved ordination technique. I E. van der Maarel (Red.), *Classification and Ordination: Symposium on advances in vegetation science, Nijmegen, The Netherlands, May 1979* (s. 47–58). Dordrecht: Springer Netherlands.

- Holligan, P. M., Fernández, E., Aiken, J., Balch, W. M., Boyd, P., Burkill, P. H., ... Wal, P. van der. (1993). A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North Atlantic. *Global Biogeochemical Cycles*, **7**:879–900.
- Howarth, R. W. (1988). Nutrient limitation of net primary production in marine ecosystems. *Annual Review of Ecology and Systematics*, **19**:89–110.
- Hu, Y. O. O., Karlson, B., Charvet, S., & Andersson, A. F. (2016). Diversity of Pico- to Mesoplankton along the 2000 km Salinity Gradient of the Baltic Sea. *Frontiers in Microbiology*, **7**:679.
- Huang, C., & Qi, Y. (1997). The abundance cycle and influence factors on red tide phenomena of *Noctiluca scintillans* (Dinophyceae) in Dapeng Bay, the South China Sea. *Journal of Plankton Research*, **19**:303–318.
- Hugenholtz, P. (2003). Chimeric 16S rDNA sequences of diverse origin are accumulating in the public databases. *International journal of systematic and evolutionary microbiology*, **53**:289–293.
- Ichinomiya, M., Honda, M., Shimoda, H., Saito, K., Odate, T., Fukuchi, M., & Taniguchi, A. (2007). Structure of the summer under fast ice microbial community near Syowa Station, eastern Antarctica. *Polar Biology*, **30**:1285–1293.
- Jellett, J. F., Stewart, J. E., & Laycock, M. V. (1995). Toxicological evaluation of saxitoxin, neosaxitoxin, gonyautoxin II, gonyautoxin II plus III and decarbamoylsaxitoxin with the mouse neuroblastoma cell bioassay. *Toxicology in Vitro*, **9**:57–65.
- Jeong, H. J., Yoo, Y. D., Kim, J. S., Seong, K. A., Kang, N. S., & Kim, T. H. (2010). Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. *Ocean Science Journal*, **45**:65–91.
- Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., ... Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, **16**:1245–1257.
- Jaanus, A., Toming, K., Hällfors, S., Kaljurand, K., & Lips, I. (2009). Potential phytoplankton indicator species for monitoring Baltic coastal waters in the summer period. I J. H. Andersen & D. J. Conley (Red.), *Eutrophication in Coastal Ecosystems: Towards better understanding and management strategies Selected Papers from the Second International Symposium on Research and Management of Eutrophication in Coastal Ecosystems, 20–23 June 2006, Nyborg, Denmark* (s. 157–168). Dordrecht: Springer Netherlands.
- Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering*, **96**:317–323.
- Kang, N. S., Jeong, H. J., Moestrup, Ø., Shin, W., Nam, S. W., Park, J. Y., ... Noh, J. H. (2010). Description of a new planktonic mixotrophic dinoflagellate *Paragymnodinium shiwhaense* n. gen., n. sp. from the coastal waters off Western Korea: morphology, pigments, and ribosomal DNA gene sequence. *Journal of Eukaryotic Microbiology*, **57**:121–144.

- Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, **30**:772–780.
- Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., ... Gray, M. W. (2005). The tree of eukaryotes. *Trends in Ecology & Evolution*, **20**:670–676.
- Kelley, D., & Richards, C. (2018). Océ. Hentet fra: <https://cran.r-project.org/web/packages/oce/vignettes/oce.html>
- Kilias, E., Kattner, G., Wolf, C., Frickenhaus, S., & Metfies, K. (2014). A molecular survey of protist diversity through the central Arctic Ocean. *Polar Biology*, **37**:1271–1287.
- Kim, S. M., Cho, J. B., Lee, E. H., Kwon, S. R., Kim, S. K., Nam, Y. K., & Kim, K. H. (2004). *Pseudocohnilembus persalinus* (Ciliophora: Scuticociliatida) is an additional species causing scuticociliatosis in olive flounder *Paralichthys olivaceus*. *Diseases of Aquatic Organisms*, **62**:239–244.
- King, N., Westbrook, M. J., Young, S. L., Kuo, A., Abedin, M., Chapman, J., ... Rokhsar, D. (2008). The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature*, **451**:783–788.
- Kircher, M., Stenzel, U., & Kelso, J. (2009). Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biology*, **10**:R83.
- Klaveness, D., Shalchian-Tabrizi, K., Thomsen, H. A., Eikrem, W., & Jakobsen, K. S. (2005). *Telonema antarcticum* sp. nov., a common marine phagotrophic flagellate. *International Journal of Systematic and Evolutionary Microbiology*, **55**:2595–2604.
- Koike, K., & Takishita, K. (2008). Anucleated cryptophyte vestiges in the gonyaulacalean dinoflagellates *Amylax buxus* and *Amylax triacantha* (Dinophyceae). *Phycological Research*, **56**:301–311.
- Kopf, A., Bicak, M., Kottmann, R., Schnetzer, J., Kostadinov, I., Lehmann, K., ... Glöckner, F. O. (2015). The ocean sampling day consortium. *GigaScience*, **4**:27.
- Kristiansen, S., Farbot, T., & Naustvoll, L.-J. (2001). Spring bloom nutrient dynamics in the Oslofjord. *Marine Ecology Progress Series*, **219**:41–49.
- Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, **29**:115–129.
- Lahaye, R., Bank, M. van der, Bogarin, D., Warner, J., Pupulin, F., Gigot, G., ... Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences*, **105**:2923–2928.
- Langmuir, I. (1938). Surface Motion of Water Induced by Wind. *Science*, **87**:119–123.
- Lawrence, C., & Menden-Deuer, S. (2012). Drivers of protistan grazing pressure: seasonal signals of plankton community composition and environmental conditions. *Marine Ecology Progress Series*, **459**:39–52.

- Leander, B. S. (2007). Molecular phylogeny and ultrastructure of *Selenidium serpulae* (Apicomplexa, Archigregarinia) from the calcareous tubeworm *Serpula vermicularis* (Annelida, Polychaeta, Sabellida). *Zoologica Scripta*, **36**:213–227.
- Leander, B. S., Lloyd, S. A. J., Marshall, W., & Landers, S. C. (2006). Phylogeny of marine gregarines (Apicomplexa) — *Pterospora*, *Lithocystis* and *Lankesteria* — and the Origin(s) of Coelomic Parasitism. *Protist*, **157**:45–60.
- Lein, T. E. (1999). A newly immigrated red alga (‘*Dasysiphonia*’, Dasyaceae, Rhodophyta) to the Norwegian coast. *Sarsia*, **84**:85–88.
- Lin, S., Zhang, H., Hou, Y., Zhuang, Y., & Miranda, L. (2009). High-level diversity of dinoflagellates in the natural environment, revealed by assessment of mitochondrial *cox1* and *cob* genes for dinoflagellate DNA barcoding. *Applied Environmental Microbiology*, **75**:1279–1290.
- Litaker, R. W., Vandersea, M. W., Kibler, S. R., Reece, K. S., Stokes, N. A., Lutzoni, F. M., ... Tester, P. A. (2007). Recognizing dinoflagellate species using its rDNA sequences. *Journal of Phycology*, **43**:344–355.
- Liu, H., Probert, I., Uitz, J., Claustre, H., Aris-Brosou, S., Frada, M., ... Vargas, C. de. (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proceedings of the National Academy of Sciences*, **106**:12803–12808.
- Logares, R. (2017). *ramalok/amplicon_processing: Workflow for Analysing MiSeq Amplicons based on Uparse*. Zenodo. Hentet fra: <https://doi.org/10.5281/zenodo.259579>
- Logares, R., Audic, S., Santini, S., Pernice, M. C., Vargas, C. de, & Massana, R. (2012). Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *The ISME Journal*, **6**:1823–1833.
- Logares, R., Shalchian-Tabrizi, K., Boltovskoy, A., & Rengefors, K. (2007). Extensive dinoflagellate phylogenies indicate infrequent marine–freshwater transitions. *Molecular Phylogenetics and Evolution*, **45**:887–903.
- Longhurst, A., Sathyendranath, S., Platt, T., & Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research*, **17**:1245–1271.
- López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., & Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, **409**:603–607.
- Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., ... Jackson, J. B. C. (2006). Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science*, **312**:1806–1809.
- Lovejoy, C., Massana, R., & Pedrós-Alió, C. (2006). Diversity and distribution of marine microbial eukaryotes in the arctic ocean and adjacent seas. *Applied Environmental Microbiology*, **72**:3085–3095.

- Lund, J. W. G., Kipling, C., & Le Cren, E. D. (1958). The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. *Hydrobiologia*, **11**:143–170.
- Lynn, D. H. (2012). Ciliophora. I *eLS*. John Wiley & Sons, Ltd. Chichester. Hentet fra: <https://doi.org/10.1002/9780470015902.a0001966.pub3>
- MacGillivray, M. L., & Kaczmarska, I. (2011). Survey of the efficacy of a short fragment of the *rbcL* gene as a supplemental DNA barcode for diatoms. *Journal of Eukaryotic Microbiology*, **58**:529–536.
- Marin, B., Klingberg, M., & Melkonian, M. (1998). Phylogenetic relationships among the Cryptophyta: Analyses of nuclear-encoded SSU rRNA sequences support the monophyly of extant plastid-containing lineages. *Protist*, **149**:265–276.
- Martin, A. P. (2003). Phytoplankton patchiness: the role of lateral stirring and mixing. *Progress in Oceanography*, **57**:125–174.
- Massana, R., Castresana, J., Balague, V., Guillou, L., Romari, K., Groisillier, A., ... Pedros-Alio, C. (2004). Phylogenetic and ecological analysis of novel marine Stramenopiles. *Applied and Environmental Microbiology*, **70**:3528–3534.
- Massana, R. (2011). Eukaryotic picoplankton in surface oceans. *Annual Review of Microbiology*, **65**:91–110.
- Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., ... de Vargas, C. (2015). Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing: Protist diversity in European coastal areas. *Environmental Microbiology*, **17**:4035–4049.
- Massana, R., Guillou, L., Díez, B., & Pedrós-Alió, C. (2002). Unveiling the organisms behind novel eukaryotic ribosomal DNA sequences from the ocean. *Applied Environmental Microbiology*, **68**:4554–4558.
- McDevit, D. C., & Saunders, G. W. (2009). On the utility of DNA barcoding for species differentiation among brown macroalgae (Phaeophyceae) including a novel extraction protocol. *Phycological Research*, **57**:131–141.
- Medina, M., Collins, A. G., Silberman, J. D., & Sogin, M. L. (2001). Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proceedings of the National Academy of Sciences*, **98**:9707–9712.
- Medinger, R., Nolte, V., Pandey, R. V., Jost, S., Ottenwalder, B., Schlotterer, C., & Boenigk, J. (2010). Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Molecular Ecology*, **19**:32–40.
- Mercier, M., Vasseur, R., & Dauxois, T. (2011). Resurrecting dead-water phenomenon. *Nonlinear Processes in Geophysics*, **18**:193–208.

- Mertens, K. N., Dale, B., Ellegaard, M., Jansson, I.-M., Godhe, A., Kremp, A., & Louwye, S. (2011). Process length variation in cysts of the dinoflagellate *Protoceratium reticulatum*, from surface sediments of the Baltic–Kattegat–Skagerrak estuarine system: a regional salinity proxy. *Boreas*, **40**:242–255.
- Miljødirektoratet. (2017). Kommunalt avløp. Hentet 20 jan. 2019 fra: <http://www.miljostatus.no/tema/ferskvann/kommunalt-avlop/>
- Millette, N. C., Pierson, J. J., Aceves, A., & Stoecker, D. K. (2017). Mixotrophy in *Heterocapsa rotundata*: A mechanism for dominating the winter phytoplankton. *Limnology and Oceanography*, **62**:836–845.
- Millie, D. F., Schofield, O. M., Kirkpatrick, G. J., Johnsen, G., Tester, P. A., & Vinyard, B. T. (1997). Detection of harmful algal blooms using photopigments and absorption signatures: A case study of the Florida red tide dinoflagellate, *Gymnodinium breve*. *Limnology and Oceanography*, **42**:1240–1251.
- Moestrup, Ø., Hakanen, P., Hansen, G., Daugbjerg, N., & Ellegaard, M. (2014). On *Levanderina fissa* gen. & comb. nov. (Dinophyceae) (syn. *Gymnodinium fissum*, *Gyrodinium instriatum*, *Gyr. uncatenum*), a dinoflagellate with a very unusual sulcus. *Phycologia*, **53**:265–292.
- Moniz, M. B. J., & Kaczmarek, I. (2010). Barcoding of diatoms: nuclear encoded ITS Revisited. *Protist*, **161**:7–34.
- Mordret, S., Piredda, R., Vaultot, D., Montresor, M., Kooistra, W. H. C. F., & Sarno, D. (2018). dinoref: A curated dinoflagellate (Dinophyceae) reference database for the 18S rRNA gene. *Molecular Ecology Resources*, **18**:974–987.
- Moum, J. N., & Smyth, W. D. (2001). Upper ocean mixing processes. I *Encyclopedia of Ocean Sciences* (s. 3093–3100). Corvallis, OR, Academic Press.
- Naturmangfoldloven. (2009). *Lov om forvaltning av naturens mangfold (naturmangfoldloven)*. Hentet fra: <https://lovdata.no/dokument/NL/lov/2009-06-19-100>
- Naustvoll, L., Norli, M., Selvik, J., & Walday, M. (2018). *Overvåking av Ytre Oslofjord 2014-2018. Tilførsler og undersøkelser i vannmassene i 2017*. (Fagrappport nr. 7274–2018). Oslo: Norsk institutt for vannforskning. Hentet fra: http://ytre-oslofjord.no/wp-content/uploads/2017/07/2016_%C3%85rsrapport-for-2016_Overv%C3%A5kning-2014-2018_Rapport.pdf
- Nezlin, N. P., Polikarpov, I. G., Al-Yamani, F. Y., Subba Rao, D. V., & Ignatov, A. M. (2010). Satellite monitoring of climatic factors regulating phytoplankton variability in the Arabian (Persian) Gulf. *Journal of Marine Systems*, **82**:47–60.
- Nikolaev, S. I., Berney, C., Fahrni, J. F., Bolivar, I., Polet, S., Mylnikov, A. P., ... Pawlowski, J. (2004). The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proceedings of the National Academy of Sciences*, **101**:8066–8071.

- Nikolenko, S. I., Korobeynikov, A. I., & Alekseyev, M. A. (2013). BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics*, **14**:S7.
- Ning, S.-B., Guo, H.-L., Wang, L., & Song, Y.-C. (2002). Salt stress induces programmed cell death in prokaryotic organism *Anabaena*. *Journal of Applied Microbiology*, **93**:15–28.
- Norconsult. (2017). *Overvåkning av Indre Oslofjord 2016* (Vedleggsrapport nr. 5145099–04). Oslo: Fagrådet for vann- og avløpsteknisk samarbeid i Indre Oslofjord. Hentet fra: file:///C:/Users/Karoline/Downloads/OvervkingavindreOslofjord2016_B02.pdf
- Norconsult. (2018). *Overvåking av Indre Oslofjord 2017* (Vedleggsrapport nr. 5145099–08). Oslo: Fagrådet for vann- og avløpsteknisk samarbeid i Indre Oslofjord. Hentet fra: file:///C:/Users/Karoline/Downloads/Overvking_Indre_Oslofjord_2017Norconsultgjeldende%20(1).pdf
- Norli, M., & Eikrem, W. (2018). *Årsovervåking med FerryBox - Indre Oslofjord 2017 - Datarapport* (Datarapport nr. 7266–2018). Oslo: Norsk institutt for vannforskning. Hentet fra: file:///C:/Users/Karoline/Downloads/rsovervking_2017_7266-2018NIVAFerryBox%20(1).pdf
- Not, F., Campo, J. del, Balagué, V., Vargas, C. de, & Massana, R. (2009). New insights into the diversity of marine picoeukaryotes. *PloS one*, **4**:e7143.
- Not, F., Latasa, M., Marie, D., Cariou, T., Vaultot, D., & Simon, N. (2004). A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Applied Environmental Microbiology*, **70**:4064–4072.
- Not, F., Massana, R., Latasa, M., Marie, D., Colson, C., Eikrem, W., ... Simon, N. (2005). Late summer community composition and abundance of photosynthetic picoeukaryotes in Norwegian and Barents Seas. *Limnology and Oceanography*, **50**:1677–1686.
- Not, F., Siano, R., Kooistra, W. H. C. F., Simon, N., Vaultot, D., & Probert, I. (2012). Chapter one - Diversity and ecology of eukaryotic marine phytoplankton. I G. Piganeau (Red.), *Advances in Botanical Research* (Bd. 64, s. 1–53). Cambridge, MA: Academic Press.
- Not, F., Zapata, M., Pazos, Y., Campaña, E., Doval, M., & Rodríguez, F. (2007). Size-fractionated phytoplankton diversity in the NW Iberian coast: a combination of microscopic, pigment and molecular analyses. *Aquatic Microbial Ecology*, **49**:255–265.
- Nærings-og fiskeridepartementet. (2018, oktober 23). Et statsbudsjett for hav [Taleartikkel]. Hentet fra: <https://www.regjeringen.no/no/aktuelt/et-statsbudsjett-for-hav/id2616740/>
- Okamoto, N., & Inouye, I. (2005). The Katablepharids are a distant sister group of the Cryptophyta: A proposal for Katablepharidophyta divisio nova/Kathablepharida phylum novum based on SSU rDNA and beta-tubulin phylogeny. *Protist*, **156**:163–179.

- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., ... Wagner, H. (2018). *Vegan: Community Ecology Package. R package version 2.5-3*. Hentet fra: <https://cran.r-project.org/web/packages/vegan/vegan.pdf>
- Park, K.-I., & Choi, K.-S. (2001). Spatial distribution of the protozoan parasite *Perkinsus* sp. found in the Manila clams, *Ruditapes philippinarum*, in Korea. *Aquaculture*, **203**:9–22.
- Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., & Esling, P. (2016). Protist metabarcoding and environmental biomonitoring: Time for change. *European Journal of Protistology*, **55**:12–25.
- Pawlowski, J. (2000). Introduction to the molecular systematics of Foraminifera. *Micropaleontology*, **46**:1–12.
- Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., ... Vargas, C. de. (2012). CBOL protist working group: Barcoding eukaryotic richness beyond the Animal, Plant, and Fungal kingdoms. *PloS Biology*, **10**:e1001419.
- Pawlowski, J., & Burki, F. (2009). Untangling the Phylogeny of Amoeboid Protists. *Journal of Eukaryotic Microbiology*, **56**:16–25.
- Pawlowski, J., Holzmann, M., Berney, C., Fahrni, J., Gooday, A. J., Cedhagen, T., ... Bowser, S. S. (2003). The evolution of early Foraminifera. *Proceedings of the National Academy of Sciences*, **100**:11494–11498.
- Paxinos, R., & Mitchell, J. G. (2000). A rapid Utermöhl method for estimating algal numbers. *Journal of Plankton Research*, **22**:2255–2262.
- Peperzak, L., Colijn, F., Gieskes, W. W. C., & Peeters, J. C. H. (1998). Development of the diatom- *Phaeocystis* spring bloom in the Dutch coastal zone of the North Sea: the silicon depletion versus the daily irradiance threshold hypothesis. *Journal of Plankton Research*, **20**:517–537.
- Pernice, M. C., Logares, R., Guillou, L., & Massana, R. (2013). General patterns of diversity in major marine microeukaryote lineages. *PloS one*, **8**:e57170.
- Pielou, E. C. (1966). Species-diversity and pattern-diversity in the study of ecological succession. *Journal of Theoretical Biology*, **10**:370–383.
- Pierce, R., & Turner, J. (1992). Ecology of plankton ciliates in marine food webs. *Reviews in Aquatic Sciences*, **6**:139-181.
- Piredda, R., Tomasino, M. P., D'Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., ... Zingone, A. (2017). Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS Microbiology Ecology*, **93**:fiw200.
- Piwosz, K., & Pernthaler, J. (2010). Seasonal population dynamics and trophic role of planktonic nanoflagellates in coastal surface waters of the Southern Baltic Sea. *Environmental Microbiology*, **12**:364–377.

- Pizay, M.-D., Lemée, R., Simon, N., Cras, A.-L., Laugier, J.-P., & Dolan, J. R. (2009). Night and day morphologies in a planktonic dinoflagellate. *Protist*, **160**:565–575.
- Prokopowich, C. D., Gregory, T. R., & Crease, T. J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome*, **46**:48–50.
- Paasche, E. (2005). *Forelesninger i marin biologi, botanisk del*. (Undervisningskompendium). Oslo: Universitetet i Oslo.
- Paasche, E., & Erga, S. R. (1987). Næringssalter og planktonvekst i Oslofjorden. *VANN*, **2**:199-206.
- Paasche, E., & Kristiansen, S. (1982). Ammonium regeneration by microzooplankton in the Oslofjord. *Marine Biology*, **69**:55–63.
- Paasche, E., & Østergren, I. (1980). The annual cycle of plankton diatom growth and silica production in the inner Oslofjord. *Limnology and Oceanography*, **25**:481–494.
- Paasche, E., (1990). Oppblomstringer av skadelig planktonalger i sjøvann. *Blyttia*, **48**:15–19.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, **41**:D590–D596.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Hentet fra: <http://www.R-project.org/>.
- Raho, N., Rodriguez, F., Reguera, B., & Marin, I. (2013). Are the mitochondrial *cox1* and *cob* genes suitable markers for species of *Dinophysis* Ehrenberg? *Harmful Algae*, **28**:64–70.
- Reun, T. L., Favier, B., & Bars, M. L. (2018). Parametric instability and wave turbulence driven by tidal excitation of internal waves. *Journal of Fluid Mechanics*, **840**:498–529.
- Reuter, J. A., Spacek, D., & Snyder, M. P. (2015). High-Throughput Sequencing Technologies. *Molecular cell*, **58**:586–597.
- Ribeiro, C. G., dos Santos, A. L., Marie, D., Brandini, F. P., & Vaulot, D. (2018). Small eukaryotic phytoplankton communities in tropical waters off Brazil are dominated by symbioses between Haptophyta and nitrogen-fixing cyanobacteria. *The ISME Journal*, **12**:1360–1374.
- Riebesell, U. (2004). Effects of CO₂ enrichment on marine phytoplankton. *Journal of Oceanography*, **60**:719–729.
- Riemann, L., Steward, G. F., & Azam, F. (2000). Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Applied Environmental Microbiology*, **66**:578–587.
- Rinde, E., Bjørge, A., Eggereide, A., & Tufteland, G. (1998). *Kystøkologi: Den ressursrike norskekysten*. Oslo: Universitetsforlaget AS.

- Rost, B., & Riebesell, U. (2004). Coccolithophores and the biological pump: responses to environmental changes. I H. R. Thierstein & J. R. Young (Red.), *Coccolithophores: From Molecular Processes to Global Impact* (s. 99–125). Berlin: Springer Berlin Heidelberg.
- Rueness, J. (2006). Bestemmelsesnøkler til rødalger. Oslo: Universitet i Oslo.
- Ryther, J. H., & Dunstan, W. M. (1971). Nitrogen, phosphorus, and eutrophication in the coastal marine environment. *Science*, **171**:1008–1013.
- Salomon, P. S., Janson, S., & Granéli, E. (2003). Multiple species of the dinophagous dinoflagellate genus *Amoebophrya* infect the same host species. *Environmental Microbiology*, **5**:1046–1052.
- Saubrekka, K.; Bjørneklett, H. K. (2019): Metabarcoding_inner_Oslofjord_2016_2017.xlsx. figshare. Datasett. DOI: 10.6084/m9.figshare.7637036.v2
- Saunders G. W. (2005). Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**:1879–1888.
- Scamardella, J. M. (1999). Not plants or animals: a brief history. *Internatl microbiology*, **2**:207–216
- Schirmer, M., Ijaz, U. Z., D'Amore, R., Hall, N., Sloan, W. T., & Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Research*, **43**:e37–e37.
- Schloss, P. D., & Westcott, S. L. (2011). Assessing and improving methods used in Operational Taxonomic Unit-based approaches for 16S rRNA gene sequence analysis. *Applied and Environmental Microbiology*, **77**:3219–3226.
- Schmitter, R. E. (1971). The fine structure of *Gonyaulax polyedra*, a bioluminescent marine dinoflagellate. *Journal of Cell Science*, **9**:147–173.
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., ... Consortium, F. B. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, **109**:6241–6246.
- Searles, R. B. (1980). The strategy of the red algal life history. *The American Naturalist*, **115**:113–120.
- Seenivasan, R. (2013). *Identification of a new eukaryotic phylum 'Picozoa' in the oceanic environment* (Doktoravhandling, Universitetet i Köln, Köln, Tyskland). Hentet fra: <http://www.uni-koeln.de/>
- Seenivasan, R., Sausen, N., Medlin, L. K., & Melkonian, M. (2013). Picomonas judraskeda Gen. Et Sp. Nov.: The first identified member of the Picozoa phylum nov., a widespread group of picoeukaryotes, formerly known as 'Picobiliphytes'. *PloS one*, **8**:e59565.

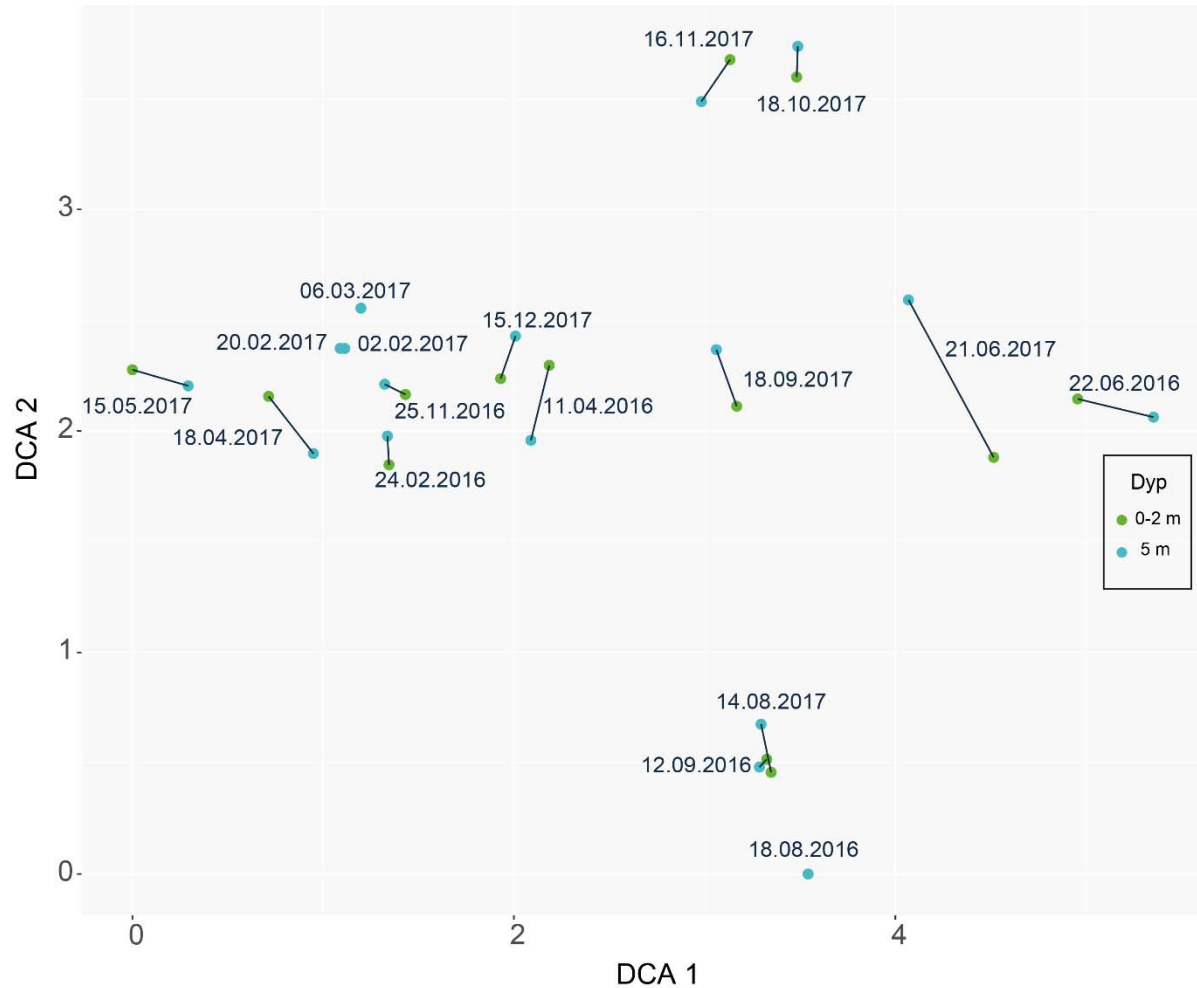
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**:379–423.
- Sherr, E. B., & Sherr, B. F. (1994). Bacterivory and herbivory: Key roles of phagotrophic protists in pelagic food webs. *Microbial Ecology*, **28**:223–235.
- Siano, R., & Arnaud-Haond, S. (2016). Metabarcoding: a tool to accelerate biodiversity assessments? Hentet fra: <https://slideplayer.com/slide/12021955/>
- Sims, D., Sudbery, I., Illott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*, **15**:121–132.
- Skjoldahl H. R., Dundas I. (1991) The *Chrysochromulina* bloom in the Skagerrak and the Kattegat in May-June 1988: Environmental conditions, possible causes and effects. (Cooperative research report nr. 175), København: International Council for the Exploration of the Sea.
- Skjoldal, H. R., Sætre, R., Færnø, A., Misund, O. A., & Røttingen, I. (Red.). (2004). *The Norwegian sea ecosystem*. Trondheim: Tapir Academic Press.
- Skovgaard, A., & Daugbjerg, N. (2008). Identity and systematic position of *Paradinium poucheti* and other *Paradinium*-like parasites of marine copepods based on morphology and nuclear-encoded SSU rDNA. *Protist*, **159**:401–413.
- Smayda, T. J., & Boleyn, B. J. (1966). Experimental observations on the flotation of marine diatoms. ii. *Skeletonema costatum* and *Rhizosolenia setigera*. *Limnology and Oceanography*, **11**:18–34.
- Smayda, T. J., & Boleyn, B. J. (1966a). Experimental observations on the flotation of marine diatoms. iii. *Bacteriastrum hyalinum* and *Chaetoceros lauderi*. *Limnology and Oceanography*, **11**:35–43.
- Smetacek, V. (1981). The annual cycle of protozooplankton in the Kiel Bight. *Marine Biology*, **63**:1–11.
- Spero, H. J. (1987). Symbiosis in the planktonic foraminifer, *Orbulina universa*, and the isolation of its symbiotic dinoflagellate, *Gymnodinium béii* sp. nov. *Journal of Phycology*, **23**:307–317.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**:1312–1313.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, **21**:2045–2050.
- Takabayashi, M., Lew, K., Johnson, A., Marchi, A. L., Dugdale, R., & Wilkerson, F. P. (2006). The effect of nutrient availability and temperature on chain length of the diatom, *Skeletonema costatum*. *Journal of Plankton Research*, **28**:831–840.

- Takahashi, M., & Bienfang, P. K. (1983). Size structure of phytoplankton biomass and photosynthesis in subtropical Hawaiian waters. *Marine Biology*, **76**:203–211.
- Talley, L. D. (2011). Descriptive Physical Oceanography: An Introduction. I Talley, L. D., Pickard, G. L., Emery, W. J., & Swift, J. H. *Descriptive Physical Oceanography*, (6 utg., s. 187-221). Cambridge, MS: Academic Press.
- Tang, E. P. Y. (1996). Why do dinoflagellates have lower growth rates? *Journal of Phycology*, **32**:80–84.
- Taylor, A. H., Harbour, D. S., Harris, R. P., Burkill, P. H., & Edwards, E. S. (1993). Seasonal succession in the pelagic ecosystem of the North Atlantic and the utilization of nitrogen. *Journal of Plankton Research*, **15**:875–891.
- Taylor, F. J. R., Hoppenrath, M., & Saldarriaga, J. F. (2009). Dinoflagellate diversity and distribution. I W. Foissner & D. L. Hawksworth (Red.), *Protist Diversity and Geographical Distribution* (s. 173–184). Dordrecht: Springer Netherlands.
- Tennekes, M., & Ellis, P. (2017). treemap: Treemap Visualization (Versjon 2.4-2). Hentet fra: <https://CRAN.R-project.org/package=treemap>
- Thaulow, H., & Faafeng, B. (2013). *Indre Oslofjord 2013 – status, trusler og tiltak*. (NIVA-rapport nr. 6593). Oslo: Norsk institutt for vannforskning. Hentet fra: <https://brage.bibsys.no/xmlui/handle/11250/194149>
- Triki, H. Z., Laabir, M., Moeller, P., Chomérat, N., & Kéfi Daly-Yahia, O. (2016). First report of goniodomin A production by the dinoflagellate *Alexandrium pseudogonyaulax* developing in southern Mediterranean (Bizerte Lagoon, Tunisia). *Toxicon*, **111**:91–99.
- Thomsen, H. A., Buck, K. R., Bolt, P. A., & Garrison, D. L. (1991). Fine structure and biology of *Cryothecomonas* gen. nov. (Protista incertae sedis) from the ice biota. *Canadian Journal of Zoology*, **69**:1048–1070.
- Thomsen, H. A., & Larsen, J. (1992). Loricated choanoflagellates of the Southern Ocean with new observations on cell division in *Bicosta spinifera* (Thronsen, 1970) from Antarctica and *Saroeca attenuata* Thomsen, 1979, from the Baltic Sea. *Polar Biology*, **12**:53–63.
- Thronsen, J., Hasle, G. R., & Tangen, K. (2007). *Norsk kystplankton flora*. Oslo: Almatel forlag AS.
- Tillmann, U., & Hoppenrath, M. (2013). Life Cycle of the pseudocolonial dinoflagellate *Polykrikos kofoidii* (Gymnodiniales, Dinoflagellata). *Journal of Phycology*, **49**:298–317.
- Trobajo, R., Mann, D. G., Clavero, E., Evans, K. M., Vanormelingen, P., & McGregor, R. C. (2010). The use of partial cox1, rbcL and LSU rDNA sequences for phylogenetics and species identification within the *Nitzschia palea* species complex (Bacillariophyceae). *European Journal of Phycology*, **45**:413–425.

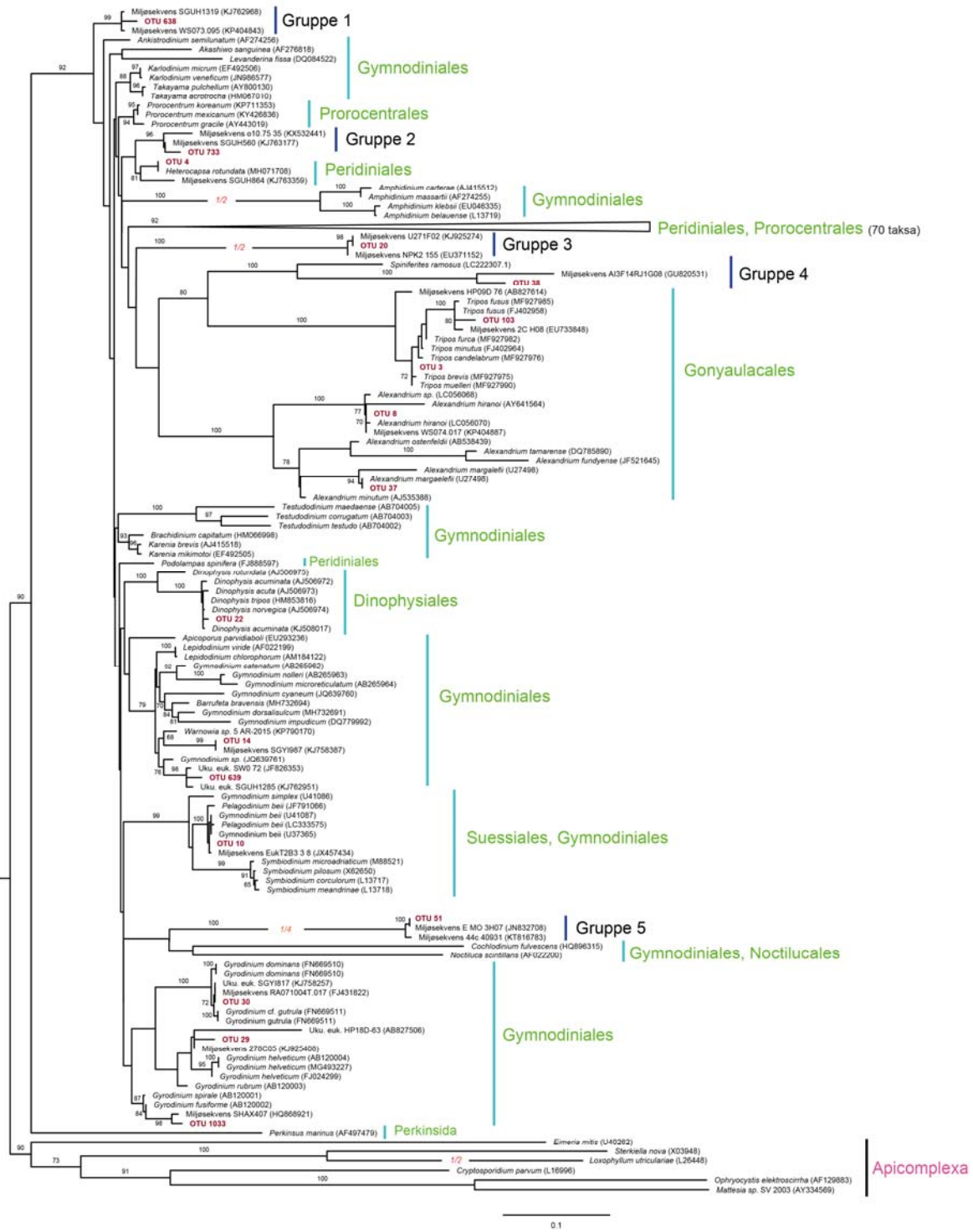
- Uitz, J., Claustre, H., Gentili, B., & Stramski, D. (2010). Phytoplankton class-specific primary production in the world's oceans: Seasonal and interannual variability from satellite observations. *Global Biogeochemical Cycles*, **24**:GB3016.
- Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., ... Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, **25**:929–942.
- van Son, T. C., & Halvorsen, R. (2014). Multiple parallel ordinations: The importance of choice of ordination method and weighting of species abundance data. *Sommerfeltia*, **37**:1–37.
- Vannforskriften. (2007). *Forskrift om rammer for vannforvaltning (vannforskriften)*. Hentet fra: <https://lovdata.no/dokument/SF/forskrift/2006-12-15-1446>
- Vences M., Thomas M., Bonett R. M., & Vieites D. R. (2005). Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**:1859–1868.
- Walday, M., Gitmark, J., Naustvoll, L., & Selvik, J. (2017). *Overvåking av Ytre Oslofjord 2014-2018. Årsrapport for 2016*. (NIVA rapport nr. 7169–2017). Oslo: Norsk institutt for vannforskning. Hentet fra: https://brage.bibsys.no/xmlui/bitstream/handle/11250/2450321/7169-2017_web.pdf?sequence=2&isAllowed=y
- Walker, J. M. (1991). Farthest North, Dead Water and the Ekman Spiral. *Weather*, **46**:158–164.
- Wang, X., Lim, H. J., & Son, A. (2014). Characterization of denaturation and renaturation of DNA for DNA hybridization. *Environmental Health and Toxicology*, **29**:e2014007.
- Whittaker, R. H. (1969). New concepts of kingdoms of organisms. *Science*, **163**:150–160.
- Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, **21**:213–251.
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., ... RStudio. (2018). ggplot2: Create elegant data visualisations using the grammar of graphics (Versjon 3.1.0). Hentet fra: <https://CRAN.R-project.org/package=ggplot2>
- Wood, S. (2018). mgcv: Mixed GAM computation vehicle with automatic smoothness estimation (Versjon 1.8-26). Hentet fra: <https://CRAN.R-project.org/package=mgcv>
- Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., ... Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, **314**:787–790.
- Yang, C., & Albright, L. (1992). Effects of the harmful diatom *Chaetoceros concavicornis* on respiration of rainbow trout *Oncorhynchus mykiss*. *Diseases of Aquatic Organisms*, **14**:105–114.

- Yasumoto, T., Murata, M., Oshima, Y., Sano, M., Matsumoto, G. K., & Clardy, J. (1985). Diarrhetic shellfish toxins. *Tetrahedron*, **41**:1019–1025.
- Yoon, H. S., Price, D. C., Stepanauskas, R., Rajah, V. D., Sieracki, M. E., Wilson, W. H., ... Bhattacharya, D. (2011). Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science*, **332**:714–717.
- Yu, S., Wang, Y., Rédei, D., Xie, Q., & Bu, W. (2013). Secondary structure models of 18S and 28S rRNAs of the true bugs based on complete rDNA sequences of *Eurydema maracandica* Oshanin, 1871 (Heteroptera, Pentatomidae). *ZooKeys*, **319**:363–377.
- Zaiko, A., Samuiloviene, A., Ardura, A., & Garcia-Vazquez, E. (2015). Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. *Marine Pollution Bulletin*, **100**:53–59.
- Zhan, A., Hulák, M., Sylvester, F., Huang, X., Adebayo, A. A., Abbott, C. L., ... MacIsaac, H. J. (2013). High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods in Ecology and Evolution*, **4**:558–565.
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, **30**:614–620.
- Zhu, F., Massana, R., Not, F., Marie, D., & Vaulot, D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiology Ecology*, **52**:79–92.
- Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., & Gemeinholzer, B. (2015). Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular Ecology Resources*, **15**:526–542.
- Zimmermann, J., Jahn, R., & Gemeinholzer, B. (2011). Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms Diversity & Evolution*, **11**:173.

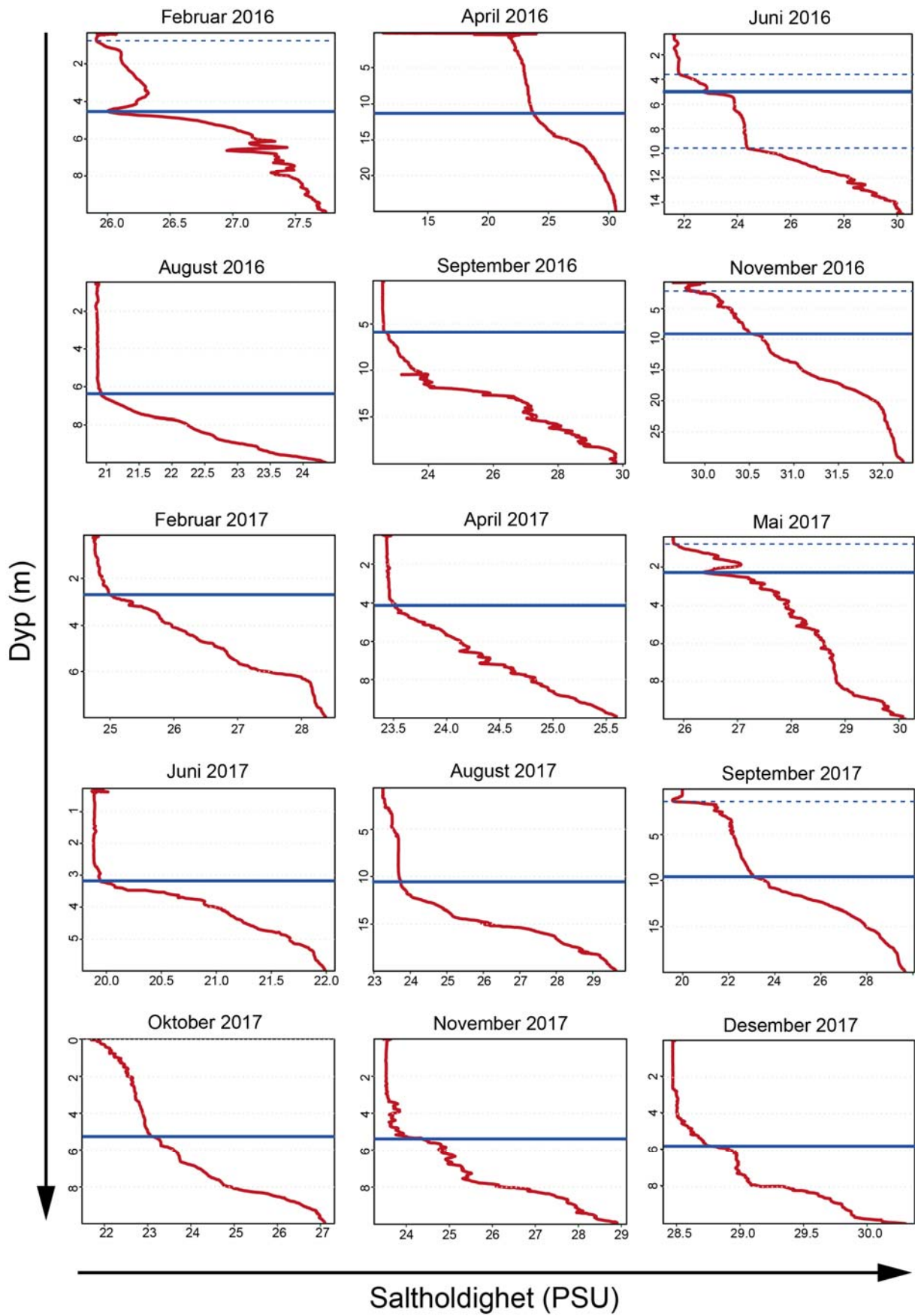
Supplementære figurer



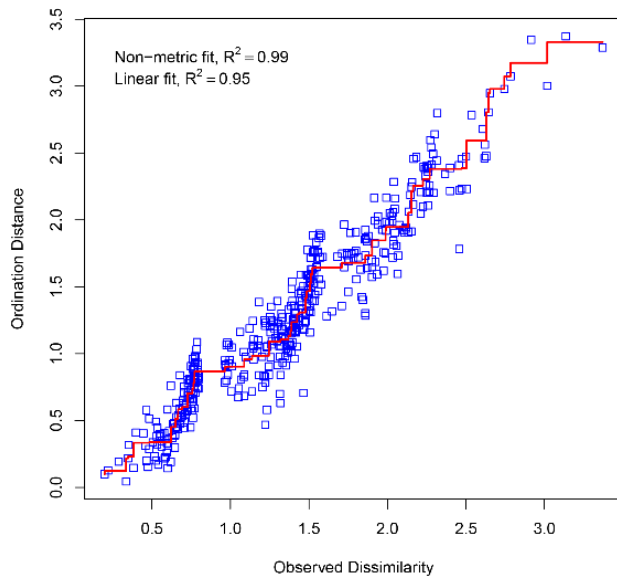
Figur S.1. DCA-ordnasjon utført på subsamplet protistdatasett. Prøver tatt samme dato er knyttet sammen med linje. DCA-ordnasjon viser tunge-effekt.



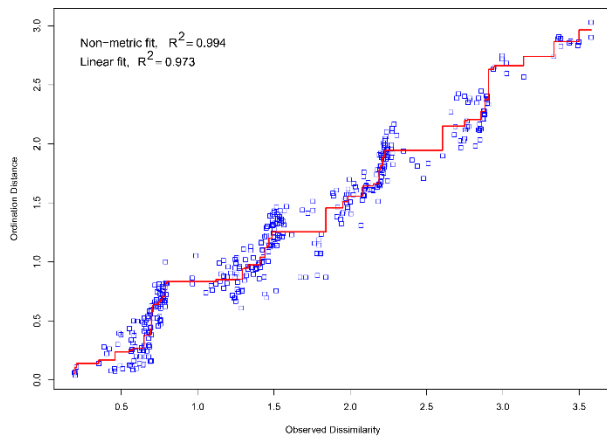
Figur S.2. Fylogenetisk tre basert på 18S rDNA over de mest abundante og frekvente Dionoflagellat OTUer i Protist- og subsamlet datasett. Treet er konstruert med RAxML i superdatamaskinen Albel. Tallene forbundet med nodene er bootstrapverdier for RAxML, bootstrapverdier under 65% vises ikke. OTUene er uthøvet mørkrøde. Hvilken orden de monofyletiske grenene tilhører er makert i grønn skrift. Grener som er forkortet er markert med hvilket forhold de er forkortet (brøk i rødt).



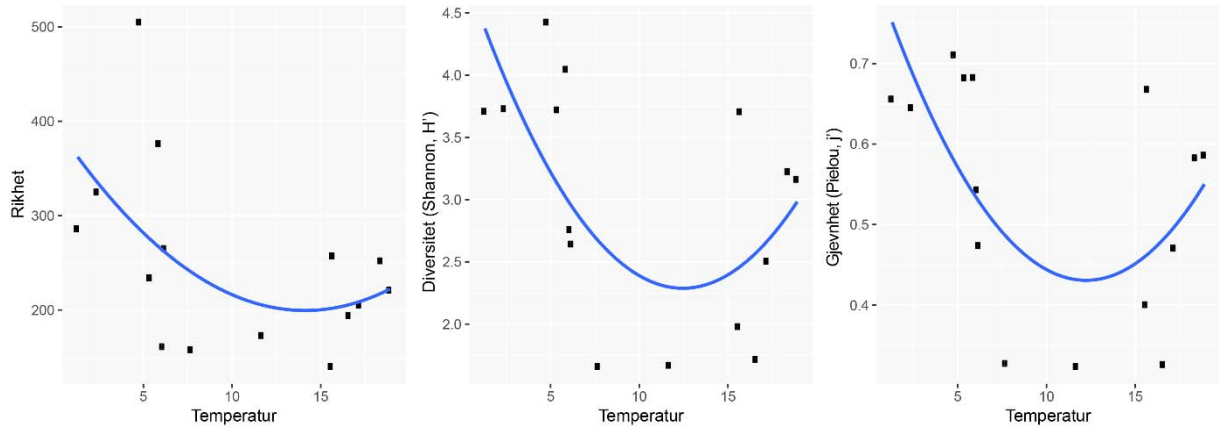
◀Figur S.3.. Hvert diagram viser hvordan saltholdigheten (x-akse) endrer seg med dyp (Y-akse). Heltrukket blå linje markerer haloklinen og striplet blå linje marker mildere spransjikt i saltholdighet. Basert på CTD-målinger fra prøvedagene fra februar 2016 til desember 2017 (mangler CTD-data for 2 februar og mars 2017).



Figur S.4. Shepards-diagram tilhørende gnmads-ordinasjon utført på subsamplet protistdatasett.



Figur S.5. Shepards-diagram tilhørende gnmads-ordinasjon utført på subsamplet protistdatasett der dinoflagellater var ekskludert.



Figur S.6. Indekser for rikhet, mangfold og jevnhet mot temperatur-data, fra 5 meters dyp. Tilpasset GAM (generalisert additiv modell).

Dato:	Ordinasjons-avstand:	Øvre blandingslag:
Sep. 2017	—————	10 m
Jun. 2017	—————	3 m
Mai 2017	—————	2 m
Sep. 2016	—————	6 m
Aug. 2017	—————	10 m
Jun. 2016	—————	5 m
Nov. 2016	—————	9 m
Okt. 2017	—————	5 m
Nov. 2017	—————	5 m
Feb. 2016	—————	4.5 m
Apr. 2017	—————	4 m
Apr. 2016	—————	12 m
Des. 2017	—————	6 m

Figur S.7. Ordinasjonsavstand rangert fra størst til minst og dypet til det øvre blandingslaget (i meter) med tilhørende måned og år som prøven er tatt.

Supplementære tabeller

Tabell S 1. Tabellen viser prøvene, hvilken dato hver enkelt prøve ble tatt, fra hvilket dyp og vannvolum som fra denne prøven ble filtrert på Sterivex-filretet. Tabellen inkluderer også resultatet fra den første (1. pre-PCR) og siste (2. post-PCR) Qubit-måling. Prøvene er tatt av EL = Elisabeth Lundsør eller H&K = Hulda K. Bjørneklett og Karoline Saubrekka

Prøve:	Dato:	Dyp (m):	Vann filtrert (L):	CTD-målinger	1. DNA konsentrasjon (ng/μL):	2. DNA konsentrasjon (ng/μL):	Prøve tatt av:
1.	24.02.16	0-2	-	x	10.1	129.0	EL
2.	24.02.16	5	-	x	10.5	145.0	EL
3.	11.04.16	0-2	-	x	8.46	186.0	EL
4.	11.04.16	5	-	x	14.4	139.0	EL
5.	22.06.16	0-2	1.95	x	21.4	145.0	H&K
6.	22.06.16	5	1.875	x	22.6	89.2	H&K
7.	18.08.16	5	1.2	x	13.4	146.0	EL
8.	12.09.16	0-2	2	x	15.9	96.0	EL
9.	12.09.16	5	2	x	18.4	139.0	EL
10.	25.11.16	0-2	-	x	10.5	158.0	EL
11.	25.11.16	5	-	x	22.2	78.0	EL
12.	02.02.17	5	2	-	8.22	181.0	EL
13.	20.02.17	5	2	x	11.7	151.0	EL
14.	06.03.17	5	1	-	9.52	41.0	EL
15.	18.04.17	0-2	0.35	x	5.86	76.8	EL
16.	18.04.17	5	1	x	9.2	70.4	EL
17.	15.05.17	0-2	1	x	43.6	47.6	H&K
18.	15.05.17	5	0.8	x	20.4	93.4	H&K
19.	21.06.17	0-2	2	x	1.27	76.6	H&K
20.	21.06.17	5	2	x	13.2	140.0	H&K
21.	14.08.17	0-2	2	x	1.84	64.6	H&K
22.	14.08.17	5	2	x	13.7	131.0	H&K
23.	18.09.17	0-2	2	x	29.0	152.0	H&K
24.	18.09.17	5	2	x	14.7	90.0	H&K
25.	18.10.17	0-2	2	x	17.5	132.0	H&K
26.	18.10.17	5	2	x	4.72	182.0	H&K
27.	16.11.17	0-2	2	x	1.59	68.2	H&K
28.	16.11.17	5	2	x	9.18	125.0	H&K
29.	15.12.17	0-2	2	x	12.7	120.0	H&K
30.	15.12.17	5	2	x	4.48	123.0	H&K

Tabell S.2. For hver prøve kan man lese hvor mange sekvenser prøven bestod av etter hver de 4 første stegene i pipelinen.

Prøve:	Dato:	Dyp:	Fra sek.senter	BAYES-HAMMER	PEAR	USEARCH	HMMR
1.	24.02.2016	0-2	374 682	372 471	361 152	195 102	194 776
2.	24.02.2016	5	415 983	413 545	402 464	229 026	228 683
3.	11.04.2016	0-2	343 767	341 989	333 004	183 468	182 772
4.	11.04.2016	5	423 511	421 093	409 614	223 982	223 253
5.	22.06.2016	0-2	317 263	315 227	309 118	174 448	174 150
6.	22.06.2016	5	469 646	467 233	454 637	234 614	234 283
7.	18.08.2016	5	94 646	93 899	92 350	51 433	51 401
8.	12.09.2016	0-2	173 318	172 206	169 222	95 540	95 312
9.	12.09.2016	5	414 557	412 391	405 387	235 753	235 382
10.	25.11.2016	0-2	162 895	161 789	157 420	89 392	91 150
11.	25.11.2016	5	253 290	251 707	246 268	140 522	140 377
12.	02.02.2017	5	150 530	149 387	146 440	82 214	82 165
13.	20.02.2017	5	333 230	331 342	324 535	185 585	185 158
14.	06.03.2017	5	300 100	298 268	290 892	160 982	160 892
15.	18.04.2017	0-2	298 549	296 936	292 402	167 736	167 645
16.	18.04.2017	5	172 686	171 568	169 147	97 039	96 991
17.	15.05.2017	0-2	303 675	302 311	299 457	184 694	184 662
18.	15.05.2017	5	265 034	263 820	259 393	152 769	152 729
19.	21.06.2017	0-2	449 302	447 644	442 250	256 005	255 647
20.	21.06.2017	5	609 061	606 330	590 932	329 432	329 280
21.	14.08.2017	0-2	458 687	456 881	448 758	269 190	267 865
22.	14.08.2017	5	339 426	337 666	332 254	184 460	184 036
23.	18.09.2017	0-2	336 260	334 165	324 697	138 320	138 260
24.	18.09.2017	5	427 614	425 238	413 240	228 182	228 047
25.	18.10.2017	0-2	179 800	177 607	166 966	44 863	44 836
26.	18.10.2017	5	372 294	370 353	362 279	203 866	203 788
27.	16.11.2017	0-2	288 427	287 257	283 564	162 541	162 449
28.	16.11.2017	5	360 433	358 995	354 536	213 175	213 038
29.	15.12.2017	0-2	381 913	380 351	371 213	214 043	213 623
30.	15.12.2017	5	392 507	390 834	383 046	220 928	220 504
		SUM:	9 863 086	9 810 503	9 596 637	5 349 304	5 343 154

Tabell S.3. Tabellen viser antallet sekvenser i hver prøver etter de var ferdig filtrert i pipelinen (Rådata), etter manuell filtrering av OTUer som kun forkom i en prøve eller hadde totalt <10 sekvenser (Opprinnelige datasettet), etter manuell fjerning av ikke-protist-gruppene Fungi, Metazoa, Opisthokona og Streptophyta (Protistdatasettet) og etter alle prøvene ble subsamplet til likt antall sekvenser (subsamplet).

Prøve:	Dato:	Dyp (m):	Rådata	Opprinnelige datasettet	Protistdatasettet	Subsamplet protistdatasett
1.	24.02.2016	0-2	184 535	184 357	168 932	10 530
2.	24.02.2016	5	214 765	214 522	183 493	10 530
3.	11.04.2016	0-2	176 639	176 418	125 711	10 530
4.	11.04.2016	5	213 663	213 539	137 910	10 530
5.	22.06.2016	0-2	165 024	164 754	122 615	10 530
6.	22.06.2016	5	224 197	223 799	177 061	10 530
7.	18.08.2016	5	47 304	47 183	29 680	10 530
8.	12.09.2016	5	223 433	223 159	109 865	10 530
9.	12.09.2016	0-2	89 108	88 984	57 786	10 530
10.	25.11.2016	0-2	82 401	81 927	66 960	10 530
11.	25.11.2016	5	129 864	129 323	113 637	10 530
12.	02.02.2017	5	75 258	75 147	56 123	10 530
13.	20.02.2017	5	177 047	176 695	111 500	10 530
14.	06.03.2017	5	148 065	147 945	60 158	10 530
15.	18.04.2017	5	81 764	81 658	42 543	10 530
16.	18.04.2017	0-2	144 390	144 113	132 062	10 530
17.	15.05.2017	0-2	178 972	178 911	172 791	10 530
18.	15.05.2017	5	143 415	143 208	88 358	10 530
19.	21.06.2017	5	317 962	316 434	118 965	10 530
20.	21.06.2017	0-2	250 703	250 539	198 987	10 530
21.	14.08.2017	5	170 748	170 500	78 640	10 530
22.	14.08.2017	0-2	249 018	248 509	164 123	10 530
23.	18.09.2017	0-2	131 664	131 603	10 530	10 530
24.	18.09.2017	5	222 631	222 100	37 576	10 530
25.	18.10.2017	0-2	41 958	41 920	22 336	10 530
26.	18.10.2017	5	195 430	195 156	83 763	10 530
27.	16.11.2017	5	205 581	205 431	90 647	10 530
28.	16.11.2017	0-2	159 436	159 330	81 555	10 530
29.	15.12.2017	0-2	203 506	203 041	114 457	10 530
30.	15.12.2017	5	209 350	208 641	129 510	10 530
		SUM:	5 057 831	5 048 846	3 088 274	315 900

Tabell S.4 Tabellen viser hvor mange OTUer og sekvenser som tilhører hver taksonomiske gruppe i de tre datasettene (Opprinnelige-, Protist- og Subsamplet datasett). Gruppene markert i gult, er de som fjernes fra det Originale datasettet og de gjenværende danner da Protist datasettet.

SUPERGRUPPE	Gruppe	OPPRINNELIGE- OG PROTIST- DATASETT		SUBSAMPLET PROTISTDATASETT	
		OTUer	Sekvenser	OTUer	Sekvenser
ALVEOLATA	Apicomplexa	13	2325	12	677
	Ciliophora	124	132 536	117	12 171
	Dinoflagellata	567	2 000 859	534	203 100
	Perkinsea	3	64	3	9
AMOEBOZOA	Amoebozoa	2	30	1	3
APUSOZOA	Apusozoa	4	64	4	9
ARCHAEPLASTIDA	Chlorophyta	70	140 025	67	12 698
	Rhodophyta	22	50 124	22	4 496
	Streptophyta	7	462	-	-
HACROBIA	Centroheliozoa	5	1 407	4	198
	Cryptophyta	10	137 956	10	12 660
	Haptophyta	28	46 074	28	4 573
	Katablepharidophyta	6	16 034	6	1 628
	Picozoa	12	30 812	12	2 852
	Telonemia	16	13 771	16	1 967
OPISTHOKONTA	Choanoflagellida	28	11 652	28	1 004
	Fungi	57	7 479	-	-
	Mesomycetozoa	7	4 757	7	398
	Metazoa	132	1 952 516	-	-
	Opisthokonta	1	115	-	-
RHIZARIA	Cercozoa	97	49 363	94	6 281
	Radiolaria	12	2 621	10	176
STRAMENOPILES	Ochrophyta	190	333 956	182	37 599
	Stramenopiles	9	4 900	9	729
	Het. Stramenopiles	160	107 880	148	12 576
UKL. EUKARYOTA	Uklas. Eukaryota	2	1 064	2	96
TOTALT:	Opprinnelige datasett	1 584	5 048 846		
	Protistdatasettet	1 387	3 088 274		
	Subsamplet			1315	315900

Tabell S.5. Mest abundante og frekvente Dinoflagellat-OTUer fra Protist- og subsamplet protistdatasett. OTUenes taksonomiske Blast-treff i referansedatabasen PR². Blast-søk gjort mauelt mot NCBI, der to av de beste treffene er inkludert i tabellen. "minus" er Blast-søk mot NCBI hvor uklutiverte sekvenser fra miljøprøver er ekskludert. Alle har e-verdi = 0. Rekkfølgen på tabellen er fra mest til minst abundant i protistdatasettet.

		Bit-score	Query cover	Ident.	Accession
OTU 29					
PR²	Dinophyceae art	752	100%	99%	KJ925408
NCBI	Uncultured eukaryote clone 278C05	752	100%	99%	KJ925408
	Uncultured dinoflagellate clone HP18D-63	752	100%	99%	AB827506
minus	<i>Gyrodinium helveticum</i>	747	100%	99%	FJ024299
	<i>Gyrodinium helveticum</i>	741	100%	99%	MG493227
OTU 4					
PR²	<i>Heterocapsa</i> sp.	774	100%	100%	JX661031
NCBI	<i>Heterocapsa rotundata</i>	774	100%	100%	MH071708
	Uncultured eukaryote clone SGUH864	774	100%	100%	KJ763359
OTU 3					
PR²	<i>Tripos candelabrus</i>	763	100%	99%	FJ402945
NCBI	Uncultured dinoflagellate clone HP09D-76	774	100%	100%	AB827614
minus	<i>Tripos muelleri</i>	771	99%	100%	MF927990
	<i>Tripos brevis</i>	769	100%	99%	MF927975
OTU 10					
PR²	Suessiales art	774	100%	100%	JX457434
NCBI	<i>Pelagodinium beii</i>	774	100%	100%	LC333575
	Uncultured eukaryote clone EukT2B3_3_8	774	100%	100%	JX457434
minus	<i>Gymnodinium beii</i>	774	100%	100%	U41087
OTU 8					
PR²	<i>Alexandrium hiranoi</i>	771	100%	100%	LC056068
NCBI	Uncultured eukaryote clone WS074.017	771	100%	100%	KP404887
minus	<i>Alexandrium hiranoi</i>	771	100%	100%	LC056070
	<i>Alexandrium</i> sp.	771	100%	100%	LC056068
OTU 14					
PR²	Dinophyceae art	774	100%	100%	KJ758398
NCBI	Uncultured eukaryote clone SGYI987	774	100%	100%	KJ758387
minus	<i>Warnowia</i> sp.	747	100%	99%	KP790170
	<i>Gymnodinium</i> sp.	741	100%	99%	JQ639761
OTU 1033					
PR²	Dinophyceae art	754	100%	99%	HQ868921
NCBI	Uncultured eukaryote clone SHAX407	754	99%	99%	HQ868921
minus	<i>Heterocapsa rotundata</i>	741	100%	99%	KY980394
OTU 103					
PR²	<i>Tripos fusus</i>	752	100%	99%	FJ402958
NCBI	<i>Tripos fusus</i>	752	100%	99%	FJ402958

	Uncultured eukaryote clone 2C_H08	752	100%	99%	EU733848
OTU 22					
PR²	<i>Dinophysis acuminata</i>	769	100%	99%	KJ508017
NCBI	<i>Dinophysis acuminata</i>	769	100%	99%	KJ508017
	<i>Dinophysis tripos</i>	769	100%	99%	HM853816
OTU 20					
PR²	Dinophyceae art	776	100%	110%	KJ925274
NCBI	Uncultured eukaryote clone 271F02	776	100%	100%	KJ925274
	Uncultured eukaryote clone NPK2_155	763	100%	99%	EU371152
OTU 733					
PR²	Dinophyceae art	758	100%	99%	KJ763177
NCBI	Uncultured eukaryote clone o10.75_35	758	100%	99%	KX532441
	Uncultured eukaryote clone SGUH560	758	100%	99%	KJ763177
minus	<i>Podolampas spinifera</i>	730	100%	98%	FJ888597
	<i>Gyrodinium rubrum</i>	730	100%	98%	AB120003
OTU 30					
PR²	<i>Gyrodinium</i> sp.	774	100%	100%	KJ758257
NCBI	Uncultured eukaryote clone SGYI817	774	100%	100%	KJ758257
	Uncultured Gymnodiniales clone RA071004T.017	774	100%	100%	FJ431822
minus	<i>Gyrodinium</i> cf. <i>guttrula</i>	769	100%	99%	FN669511
	<i>Gyrodinium dominans</i>	769	100%	99%	FN669510
OTU 638					
PR²	Dinophyceae art	758	100%	99%	KJ763538
NCBI	Uncultured eukaryote clone WS073.095	758	100%	99%	KP404843
	Uncultured eukaryote clone SGUH1319	758	100%	99%	KJ762968
minus	<i>Prorocentrum mexicanum</i>	736	100%	98%	KY426836
	<i>Prorocentrum koreanum</i>	736	100%	98%	KP711353
OTU 37					
PR²	<i>Alexandrium margaelefii</i>	767	100%	99%	U27498
NCBI	<i>Alexandrium margaelefii</i>	767	100%	99%	U27498
	<i>Alexandrium minutum</i>	725	100%	98%	AJ535388
OTU 51					
PR²	Dinophyceae art	767	100%	100%	JN832710
NCBI	Uncultured eukaryote clone E_MO_3H07	767	100%	100%	JN832708
	Uncultured eukaryote clone 44c_40931	721	93%	100%	KT816783
OTU 639					
PR²	<i>Gymnodinium</i> sp.	756	100%	99%	KJ762951
NCBI	Uncultured eukaryote clone SGUH1285	756	100%	99%	KJ762951
	Uncultured alveolate clone SW0_72	756	100%	99%	JF826353
minus	<i>Barrufeta bravensis</i>	728	100%	98%	MH732694
	<i>Gymnodinium dorsalisulcum</i>	728	100%	98%	MH732691

Tabell S.6. Liste over dinoflagellat-taksa og tilhørende taksonomi, basert på protistdatasettet = MB og lysmikroskop celledata = LM (Nori & Eikrem, 2018). Uthevet med sort er taksa registrert i begge datasett. Uthevet i blått er taksa kun tilstedet i 2016 (6 stk.) og uthevet i grønt er taksa kun tilstedet ved 2 meter i grønt (1 stk.) i protistdatasettet. Tilstedeværelse er kun registrert (med x) for prøver tatt ved 5 meters dyp i 2017 ved prøvetakningsstasjonen DK1, i Indre Oslofjord. Taksonomisk klassifisering er kontrollert mot algaebase.org. Taksa med <90% identitet med referansesekvens i referanse databasen PR² er markert med * (2 stk.).

KLASSE	ORDEN	FAMILIE/GRUPE	SLEKT/KLADE	TAKSA/ART	MB	LM	
DINOPHYCEAE	Amphidinales	Amphidiniaceae	Amphidinium	<i>Amphidinium sphenoides</i>		x	
	Dinophysiales	Amphidomataceae	cf. Azadinium	<i>Azadinium</i> spp.		x	
		Dinophysiaceae	Dinophysis	<i>Dinophysis acuminata</i>	x	x	
				<i>Dinophysis acuta</i>		x	
				<i>Dinophysis norvegica</i>		x	
				<i>Dinophysis rotundata</i>		x	
				<i>Dinophysis tripos</i>		x	
				<i>Dinophysis</i> spp.		x	
				<i>Diplopsalis</i> -gruppen		x	
				<i>Diplopsalis</i> spp.		x	
				cf. <i>Diplopsalis bomba</i>		x	
			Oxyphysiaceae	Phalacroma	<i>Phalacroma rotundatum</i>	x	
		Gonyaulacales	Ceratiaceae	Ceratium	<i>Ceratium</i> sp.	x	
				Tripos	<i>Tripos candelabrus</i>	x	
					<i>Tripos concilians</i>	x	
					<i>Tripos furca</i>	x	x
					<i>Tripos fusus</i>	x	x
					<i>Tripos horridus</i>		x
					<i>Tripos limulus</i>	x	
					<i>Tripos lineatum</i>		x
					<i>Tripos longipes</i>		x
					<i>Tripos muelleri</i>		x
					<i>Tripos tenuis</i>	x	
				<i>Tripos</i> spp.		x	
		Goniodomataceae	Alexandrium	<i>Alexandrium fundyense</i>	x		
				<i>Alexandrium hiranoi</i>	x		
				<i>Alexandrium margalefii</i>	x		

		<i>Alexandrium ostenfeldii</i>		x
		<i>Alexandrium pseudogoniaulax</i>		x
Gonyaulaceae	Amylax	<i>Amylax buxus</i>		x
	Gonyaulax	<i>Gonyaulax polygramma*</i>		
		<i>Gonyaulax spinifera</i>		x
		<i>Goniaulax</i> spp.		x
		<i>Gonyaulax verior</i>		x
	Lingulodinium	<i>Lingulodinium polyedra</i>		x
		<i>Lingulodinium polyedrum</i>		x
Protoceratiaceae	Protoceratium	<i>Protoceratium reticulatum</i>		x
Pyrophacaceae	Fragilidium	<i>Fragilidium subglobosum</i>		x
Brachidiniaceae	Karenia	cf. <i>Karenia mikimotoi</i>		x
	Karlodinium	cf. <i>Karlodinium veneficum</i>		x
Gymnodiniaceae	Akashiwo	<i>Akashiwo sanguinea</i>		x
	Balechina	<i>Balechina pachydermata</i>		x
	Gymnodinium	<i>Gymnodinium aureolum</i>		x
		<i>Gymnodinium catenatum</i>		
		<i>Gymnodinium dorsalisulcum</i>		x
		<i>Gymnodinium microreticulatum</i>		x
		<i>Gymnodinium</i> sp.		x
	Gyrodinium	<i>Gyrodinium fusiforme</i>		x
		<i>Gyrodinium</i> sp.		x
		cf. <i>Gyrodinium spirale</i>		x
		<i>Gyrodinium/Gymnodinium</i> 20-40 µm		x
		<i>Gyrodinium/Gymnodinium</i> 40-60 µm		x
		<i>Gyrodinium/Gymnodinium</i> 60-80 µm		x
		<i>Gyrodinium/Gymnodinium</i> >80 µm		x
	Lepidodinium	<i>Lepidodinium chlorophorum</i>		x
	Margalefidinium	<i>Margalefidinium fulvescens</i>		x
		<i>Margalefidinium polykrikoides</i>		x
	Nusuttodinium	<i>Nusuttodinium amphidinioides</i>		
	Paragymnodinium	<i>Paragymnodinium shiwhaense</i>		
	Polykrikos	<i>Polykrikos geminatus</i>		
		<i>Polykrikos kofoidii</i>		x
		<i>Polykrikos tanit</i>		x
	Spiniferodinium	<i>Spiniferodinium palustre</i>		x
	Levanderina	<i>Levanderina fissa</i>		

Peridinales	Warnowiaceae	Warnowia	<i>Warnowia</i> sp.	x
	Amphidiniopsidaceae	Amphidiniopsis	<i>Amphidiniopsis dragescoi</i>	x
		Islandinium	<i>Islandinium tricingulatum</i>	x
		Amphidomataceae	Amphidoma	<i>Amphidoma languida</i>
	Blastodiniaceae	Blastodinium	<i>Blastodinium mangini</i>	x
	Heterocapsaceae	Heterocapsa	<i>Heterocapsa nei/rotundata</i>	x
			<i>Heterocapsa niei</i>	x
			<i>Heterocapsa rotundata</i>	x
			<i>Heterocapsa cf. triquetra</i>	x
			<i>Heterocapsa</i> sp.	x
	Oxytoxaceae	Oxytoxum	<i>Oxytoxum gracile</i>	x
	Podolampadaceae	Podolampas	<i>Podolampas bipes</i>	x
		Protoperidinium	<i>Protoperidinium bipes</i>	x
			<i>Protoperidinium brevipes</i>	x
			<i>Protoperidinium conicum</i>	x
			<i>Protoperidinium depressum</i>	x
			<i>Protoperidinium divergens</i>	x
			<i>Protoperidinium monovelum</i>	x
			<i>Protoperidinium pallidum</i>	x
		<i>Protoperidinium pellucidum</i>	x	
		<i>Protoperidinium punctulatum*</i>	x	
		<i>Protoperidinium steinii</i>	x	
		<i>Protoperidinium</i> spp. 20-40 µm	x	
		<i>Protoperidinium</i> spp. 40-60 µm	x	
Thoracosphaeraceae	Stoeckeria	<i>Stoeckeria algicida</i>	x	
		Ukjent	x	
Prorocentrales	Prorocentrales	Prorocentrum	<i>Prorocentrum cf. balticum</i>	x
			cf. <i>Prorocentrum balticum</i>	x
			<i>Prorocentrum micans</i>	x
			<i>Prorocentrum cf. minimum</i>	x
			<i>Prorocentrum</i> sp.	x
Suessiales	Borghiellaceae	Borghiella	<i>Prorocentrum triestinum</i>	x
			<i>Borghiella tenuissima</i>	x
	Suessiaceae	Ansabella	<i>Ansabella granifera</i>	x
		Biecheleria	<i>Biecheleria brevisulcata</i>	x
			<i>Biecheleria tirensis</i>	x
	Biecheleriopsis	<i>Biecheleriopsis adriatica</i>	x	

Tabell S.7. Dominante OTUer i hver enkelt prøve tatt i 2016 og 2017. Dominant er her definert til >10% av sekvensene i prøven. Tabellen inneholder OTU-nummer (OTU), andelen sekvenser OTUen hadde i prøver oppgitt i prosent (%) og OTUens taksonomiske klassifisering i referansedatabasen PR². Basert på subsamplet protistdatasett.

ÅR	MND.	0 - 2 meter			5 meter		
		OTU	ANDEL I PRØVE	KLASSIFISERT TIL	OTU	ANDEL I PRØVE	KLASSIFISERT TIL
2016	Feb.	4	12%	<i>Heterocapsa</i> sp.	29	11%	<i>Dinophysis</i> sp.
					14	17%	<i>Dinophysis</i> sp.
	Apr.	3	12%	<i>Tripos candelabrus</i>	4	21%	<i>Heterocapsa</i> sp.
		4	15%	<i>Heterocapsa</i> sp.			
	Juni	3	11%	<i>Tripos candelabrus</i>	8	66%	<i>Alexandrium hiranoi</i>
		8	36%	<i>Alexandrium hiranoi</i>			
Aug.				38	23%	<i>Dinophysis</i> sp.	
				4	11%	<i>Heterocapsa</i> sp.	
Sep.	37	11%	<i>Alexandrium margalefii</i>	4	52%	<i>Heterocapsa</i> sp.	
	4	18%	<i>Heterocapsa</i> sp.				
Nov.	29	17%	<i>Dinophysis</i> sp.	29	11%	<i>Dinophysis</i> sp.	
2017	2 feb.				29	19%	<i>Dinophysis</i> sp.
					15	10%	<i>Thalassiosira</i> sp.
	20 feb.				29	17%	<i>Dinophysis</i> sp.
					15	13%	<i>Thalassiosira</i> sp.
	Mars				15	34%	<i>Thalassiosira</i> sp.
	April	29	28%	<i>Dinophysis</i> sp.	29	29%	<i>Dinophysis</i> sp.
					1 033	12%	<i>Dinophysis</i> sp.
	Mai	13	18%	<i>Chaetoceros</i> sp.	56	18%	MAST
		29	93%	<i>Dinophysis</i> sp.	29	67%	<i>Dinophysis</i> sp.
	Juni				3	34%	<i>Tripos candelabrus</i>
		10	63%	Suessiales-art	10	39%	Suessiales-art
	Aug.	4	49%	<i>Heterocapsa</i> sp.	4	32%	<i>Heterocapsa</i> sp.
		10	13%	Suessiales-art			
		19	11%	Chlorodendrales-art			
	Sep.				27	10%	<i>Emiliana</i> sp.
		51	14%	<i>Dinophysis</i> sp.	116	11%	<i>Dinophysis</i> sp.
Okt.	3	21%	<i>Tripos candelabrus</i>	3	15%	<i>Tripos candelabrus</i>	
	3	67%	<i>Tripos candelabrus</i>	3	64%	<i>Tripos candelabrus</i>	
Nov.				103	10%	<i>Tripos fusus</i>	
	3	58%	<i>Tripos candelabrus</i>	3	46%	<i>Tripos candelabrus</i>	
Des.				103	12%	<i>Tripos fusus</i>	
	733	10%	<i>Dinophysis</i> sp.	4	13%	<i>Heterocapsa</i> sp.	
3	10%	<i>Tripos candelabrus</i>					

Tabell S.8. OTUer med 100% identitet mot referansesekvens i referansedatabasen PR².

	REKKER	KLASSE	KLASSIFISERT TIL	OTU		
ALVEOLATA	Ciliophora	Spirotrichea	<i>Pelagostrobilidium neptuni</i>	97		
			<i>Strombidium capitatum</i>	229		
			<i>Laboea strobila</i>	108		
			<i>Schmidingerella quequenensis</i>	382		
			<i>Tintinnopsis tubulosoides</i>	433		
			<i>Helicostomella subulata</i>	1 422		
	Dinoflagellata	Dinophyceae	<i>Alexandrium hiranoi</i>	8		
			<i>Gonyaulax spinifera</i>	216		
			<i>Lingulodinium polyedra</i>	1 551		
			<i>Protoceratium reticulatum</i>	268		
<i>Polykrikos kofoidii</i>			344			
ARCHAEPLASTIDA			Chlorophyta	Chlorophyceae	<i>Chlamydomonas parkeae</i>	534
	Mamiellophyceae	<i>Bathycoccus prasinus</i>			16	
		<i>Ostreococcus mediterraneus</i>		66		
		<i>Mamiella gilva</i>		227		
		Prasino-Clade-V		<i>Pycnococcus provasolii</i>	114	
	Pyramimonadophyceae			<i>Pyramimonas disomata</i>	131	
	Rhodophyta	Florideophyceae	<i>Ceramium rubrum</i>	284		
			<i>Dasya baillouviana</i>	271		
			<i>Leptosiphonia schousboei</i>	590		
			<i>Polysiphonia fucoides</i>	318		
		<i>Rhodomela confervoides</i>	371			
HACROBIA	Cryptophyta	Cryptophyceae	<i>Teleaulax amphioxeia</i>	11		
	Haptophyta	Prymnesiophyceae	<i>Calyptrosphaera sphaeroidea</i>	172		
	Katablepharidophyta	Katablepharidaceae	<i>Leucocryptos marina</i>	52		
OPIS.	Mesomycetozoa	Ichthyosporea	<i>Pseudoperkinsus tapetis</i>	528		
RHIZARIA	Cercozoa	Filosa-Thecofilosea	<i>Ebria tripartita</i>	47		
STRAMENOPILES	Ochrophyta	Bacillariophyceae	<i>Licmophora gracilis</i>	482		
			<i>Licmophora paradoxa</i>	903		
			<i>Synedra hyperborea</i>	671		
			<i>Cylindrotheca closterium</i>	238		
			<i>Pleurosigma intermedium</i>	1 394		
			Bolidophyceae	<i>Triparma strigata</i>	119	
				<i>Triparma pacifica</i>	312	
				Coscinodiscophyceae	<i>Coscinodiscus radiatus</i>	291
					<i>Leptocylindrus aporus</i>	88
				<i>Leptocylindrus convexus</i>	708	
		Dictyochophyceae	<i>Florenciella parvula</i>	99		
			<i>Pseudopedinella elastica</i>	158		
			Mediophyceae	<i>Attheya longicornis</i>	690	
				<i>Cerataulina pelagica</i>	36	
			<i>Chaetoceros setoense</i>	72		
			<i>Chaetoceros tenuissimus</i>	139		
			<i>Chaetoceros socialis</i>	259		
			<i>Chaetoceros muellerii</i>	228		
			<i>Chaetoceros pumilum</i>	304		

	<i>Chaetoceros decipiens</i>	332
	<i>Ditylum brightwellii</i>	462
	<i>Minutocellus polymorphus</i>	327
	<i>Porosira pseudodenticulata</i>	269
	<i>Skeletonema marinoi</i>	21
	<i>Thalassiosira pseudonana</i>	226
	<i>Thalassiosira delicatula</i>	704
Pelagophyceae	<i>Aureococcus anophagefferens</i>	737

Tabell S.9. Welch Two Sample t-test for indeksene til rikhet, Shannon diversitet og jevnhet for dypene 0-2 meter og 5 meter. Tabeller hentet fra R.

Welch Two Sample t-test

data: indexes\$R_2m and indexes\$R_5m
t = -0.92361, df = 23.87, p-value = 0.3649
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-111.47709 42.56306
sample estimates:
mean of x mean of y
214.3077 248.7647

data: indexes\$S_2m and indexes\$S_5m
t = -0.72334, df = 22.435, p-value = 0.4769
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.0455124 0.5043371
sample estimates:
mean of x mean of y
2.732295 3.002882

data: indexes\$J_2m and indexes\$J_5m
t = -0.65096, df = 21.818, p-value = 0.5219
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.16003055 0.08359646
sample estimates:
mean of x mean of y
0.5059308 0.5441479

Tabell S.10. Dinoflagellat-arter som med stor sikkerhet (<99% identitet med referansedatabasen PR²) er påvist med metastrekkoding i denne studien. Uthevet i grått er arter som også er observert i lysmikroskop (Norli & Eikrem, 2018), uthevet i rosa er arter som er nakne (mangler harde organiske skall) og uthevet i grønt er arter som også er observert i lysmikroskop under ett annen navn (uthevet i synonymmer).

	Str. (µm)	Informasjon	Synonym	Kilde	Hvor observert i kilde.
<i>Akashiwo sanguinea</i>	45-80	Påvist med begge metoder.	- <i>Gymnodinium sanguineum</i> (Hirasaka 1922)	- Thronsdén mfl., 2007	Kosmopolitt, vanlig i Oslofjorden.
<i>Alexandrium fundyense</i>	>20	Forskjellig morfologi på cyster. Giftig	- <i>Gonyaulax catenella</i> (Whedon og Kofoid, 1936) - <i>Alexandrium catenella</i> (Balech, 1985)	- Bravo mfl., 2006 - Figueroa mfl., 2005	Middelhavet
<i>Alexandrium hiranoi</i>	>20	Mulig cyster?			
<i>Alexandrium margalefii</i>	>20	Påvist med begge metoder.		- Bravo mfl., 2006	Middelhavet
<i>Alexandrium ostenfeldii</i>	40-50	Cyster. Giftig	- <i>Goniodoma ostenfeldii</i> (Paulsen 1904) - <i>Protogonyaulax ostenfeldii</i> (Fraga og Sánchez 1985)	- Thronsdén mfl., 2007	Vanlig langs hele norskekysten
<i>Amylax buxus</i>	20-30			- Koike & Takishita, 2008	Japan
<i>Balechina pachydermata</i>	50-100	Fenotypisk plastistet – cingulum flytter seg. Naken.	- <i>Gymnodinium pachydermatum</i> (Koifoid og Swezy, 1921)	- Gómez mfl., 2015	Middelhavet
<i>Borghiella tenuissima</i>	20	Ferskvannsort. Flat celleform. Naken.	- <i>Gymnodinium tenuissimum</i> (Lauterborn, 1989), - <i>Woloszynskia tenuissima</i> (R. H. Thompspm 1950)	- Daugbjerg mfl., 2014	Skottland

<i>Dinophysis acuminata</i>	>20	Påvist med begge metoder.				
<i>Euduboscquella crenulata</i>	50-60 Synlig inni ciliat i lysmikroskop	Intracellulær parasitt i ciliaten <i>Favella panamensis</i>			- Coats mfl., 2012	Assawoman Bay, Ocean City, Maryland, USA
<i>Fragilidium subglobosum</i>	20-60 (snitt 50)	Beiter på Tripos spp.	- <i>Helgolandinium subglobosum</i> (Stosch 1969)		- Skovgaard & Daugbjerg, 2008	Kattegat, Danmark
<i>Gonyaulax spinifera</i>	25-50	Mobile cyster.	- <i>Peridinium spiniferum</i> (Claparède og Lachmann 1859)		- Ellegaard mfl., 2003	Kattegat, Danmark (i kultur)
<i>Gymnodinium microreticulatum</i>	15-25	Naken.			- Bolch mfl., 1999 - Gu mfl., 2013	Newcastle Harbour, Australia. Gulehavet, kina
<i>Kapelodinium vestifici</i>	15-30	Naken.	- <i>Gymnodinium vestifici</i> (Schütt 1895) - <i>Amphidinium extensum</i> Wulff, 1919)		- Boutrup mfl., 2016	Østersjøen, Kattegat
<i>Levanderina fissa</i>	30-50	Naken.	- <i>Gymnodinium fissum</i> (Levander 1894) - <i>Gyrodinium fissum</i> (Kofoid og Swezy, 1921)		- Moestrup mfl., 2014	Østersjøen
<i>Lingulodinium polyedra</i>	20-40 (35-55, Thronsdens mfl. 2007)	Mobile cyster. Giftig. Påvist med begge metoder under forskjellig navn.	- <i>Gonyaulax polyedra</i> (F. Stein 1883) - <i>Lingulodinium polyedrum</i> (Dodge)		- Figueroa mfl., 2005 - Thronsdens mfl., 2007	Vanlig langs sør- og vestkysten, Skagerrak, Nordsjøen, Norskehavet.
<i>Noctiluca scintillans</i>	0.5-1.5 mm (Skrivefeil i Thronsdens mfl. 2007, µm)	Naken.			- Thronsdens mfl., 2007	
<i>Paragymnodinium shiwhaense</i>	10-20	Naken.			- Kang mfl., 2010	Korea

<i>Phalacroma rotundatum</i>				- <i>Dinophysis rotundata</i> (Claparède og Lachmann 1859)			
<i>Polykrikos kofoidii</i>	40-60	Påvist med begge metoder under forskjellig navn.	Naken. Cyster. Giftig.	-	Tillmann & Hoppenrath, 2013	Skottland (i kultur)	
<i>Protoceratium reticulatum</i>	25-55	Cyster. Giftig.	-	- <i>Peridinium reticulatum</i> (Claparède og Lachmann 1859) - <i>Gonyaulax grindley</i> (Reinecke 1967)	- Mertens mfl., 2011 - Throndsen mfl., 2007	Kystnær til estuarin- langs hele norske kysten til Svalbard	
<i>Protoperidinium conicum</i>	70-90			- <i>Peridinium divergens</i> (Gran 1900) - <i>Peridinium</i> (Ostenfeld og Schmidt 1902)	- Throndsen mfl., 2007	Primært arktisk-kystnær. Sporadisk langs norskekysten om sommeren	
<i>Protoperidinium depressum</i>	100-200	Påvist med begge metoder.		-	- Throndsen mfl., 2007		
<i>Spatulodinium pseudonociluca</i>	100-200	Naken.		- <i>Gynodinium pseudonociluca</i> (Pouchet 1885) - <i>Spatulodinium</i> (Chacon og Chacon 1976)	- Throndsen mfl., 2007	Sporadisk langs hele norskekysten, vanlig om høsten	
<i>Tripes candelabrus</i>		Påvist med begge metoder.		Ble etter MAFFT-alignment sjekk bestemt til <i>Tripes muelleri</i> .			
<i>Tripes fusus</i>		Påvist med begge metoder.					

Appendiks

Appendiks 1.

Protokoll for DNA ekstraksjon fra Sterivex-filter med bruk av Qiagen Sterivex Kit.

Denne protokollen er en modifisert protokoll for å ekstrahere DNA fra Sterivex-filter ved å bruke Qiagen DNeasy Power Water Sterivex Kit. Denne protokollen krever ikke "Tube extenders", "VacConnector", "VacValve" eller steg 18-26 i Quick-Start Protokollen som følger med kittet.

Før start:

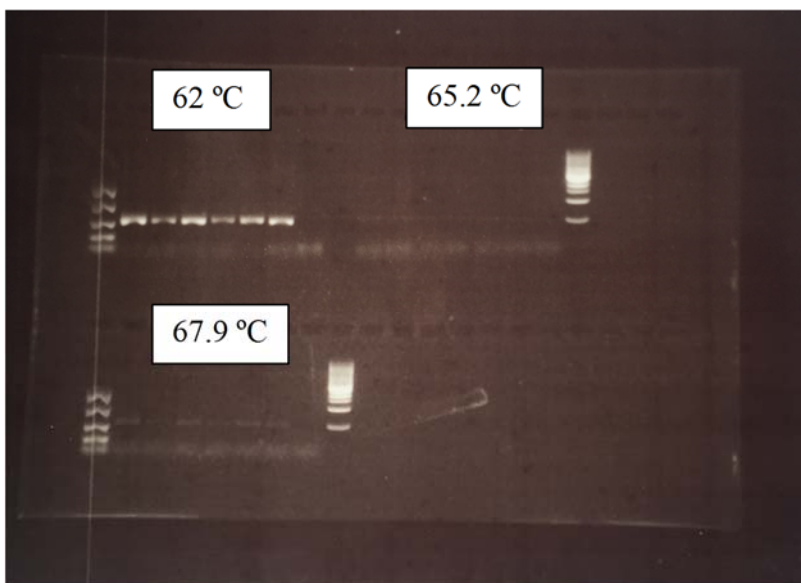
- Les Quick-Start protokollen og sikkerhetsprotokoller ved UiO.
- Oppbevar alle løsningene ved romtemperatur Etter blandning av ST1B-løsning, sett den i kjøleskap ved 2-8°C.
- Vask arbeidsbenk med 20% deconex-løsning.
- Still inn to inkubatorer til 65°C og 90°C.
- Tin filterne på is. Sørg for at de er helt tømt for sjøvann. Om ikke, trekk ut resterende vann med en sprøyte.
- Protokollen utføres i romtemperatur.

Protokoll:

1. Plasser MBL- og MR-løsning i inkubator, ved 65°C, til de skal brukes.
2. Hell ST1A løsning i ST1B pulver og bland ved å riste.
3. Forsegl filterets utløp enten med tilhørende hette eller tack-it heftemasse.
4. Tilsett 900 µL ST1B-løsning gjennom filterets innløp ved å bruke en tynn 1 mL-pipette. Plasser pipettespissen forbi innløpet, like over filterets hvite membran. Tilt filtret slik at luft slipper ut innløpet, samtidig som væsken sakte presses inn.
5. Forsegl innløpet og plasser filteret på vortexeren med innløpet pekende utover. Vortex i 5 min med minimumshastighet (500). Roter filteret 180 grader, når det fortsatt er festet til vortexseren med innløpet pekende ut og vortex 5 min med minimumshastighet.
6. Tilsett 900 µL varm (65°C) MBL-løsning i filteret.
7. Forsegl filteret, pakk det inn i aluminiumsfolie og inkuber ved 90°C i 5 min (ikke varmere eller lenger).
8. Fjern folie og avkjøl filteret i 2 min ved romtemperatur. Vortex (innløp ut) i 5 min med maksimumshastighet (2100).
9. Trekk ut lysat fra filter ved bruk av en 3 mL sprøyte. Overfør lysat til Power Bead rør (blått lokk).

10. Vortex Power Bead tube (lokk ut) i 5 min ved maks hastighet.
11. Sentrifuger Power Bead tube i 1 min/4000g.
12. Overfør kun supernatant fra Power Bead Tube til 2.2 mL rør med en mikropipette.
13. Tilsett 300 μ L IRS-løsning, vortex raskt og inkuber ved 4°C i 5 min. Deretter vortex igjen raskt og sentrifuger i 1 min/13 000 rpm. Overfør supernatant til 5 mL rør.
14. Tilsett 3 mL varm (65°C) MR-løsning og vortex tuben i noen sekunder.
15. Overfør 750 μ L lysat til MB Spin kolonner (som står plassert i tilhørende rør). Bruk to Spin kolonner per prøve. Sentrifuger i 1 min/8000 rpm. Kast den filtrerte væsken og gjenta prosessen ved å bruke den samme kolonnen til alt lysatet er filtrert på kolonnen (tre runder).
16. Tilsett 750 μ L etanol til Spin kolonnen og sentrifuger i 1 min/8000 rpm. Kast den filtrerte væsken.
17. Tilsett 750 μ L PW til Spin kolonnen og sentrifuger i 1 min/8000 rpm. Kast den filtrerte væsken.
18. Tilsett 750 μ L etanol til Spin kolonnen og sentrifuger i 3 min/4000 rpm. Kast den filtrerte væsken og flytt Spin kolonnen over i et nytt rent rør. La kolonnen lufttørke til filteret ikke lenger lukter etanol (minimum 5 min).
19. Tilsett 20 μ L EB-løsning i midten av Spin kolonnen og inkuber i 1 min. Sentrifuger i 1 min/10000 rpm. **Ikke kast den filtrerte løsningen,** men overfør den til den samme kolonnen og sentrifuger igjen 1 min/10000 rpm.
20. Bland de to løsningene fra de to Spin kolonnene som var brukt på samme prøve.
21. Merk tubene og legg de i fryseren.

Appendiks 2.



Figur A.1. Agarosegel. Bilde er tatt av agarosegelen som viser resultatet etter test av temperaturgradienten 62°C, 65.2°C og 67.9°C i PCR. Det er tydelige bånd under 62°C.

Appendiks 3.

TAE x 50 Stock solution

Fremstilt av kjemi-ingeniør Sissel Irene Brubak

Bland:

- 242 g Tris-acetat
- 57.1 mL iseddik (glacial acetic acid)
- 100 mL 0.5 M EDTA (pH 8.0)

- Tilsett dH₂O til totalvolum blir 1 L

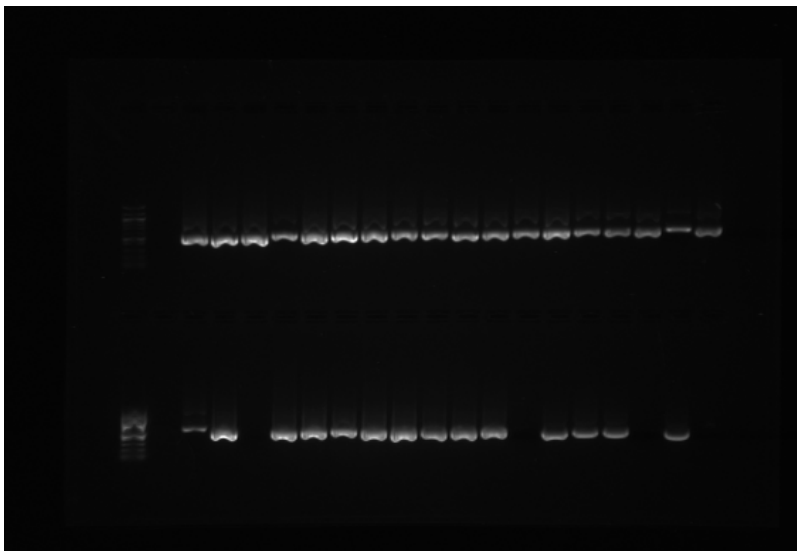
- Autoklaver.

TAE x1 buffer

Bland:

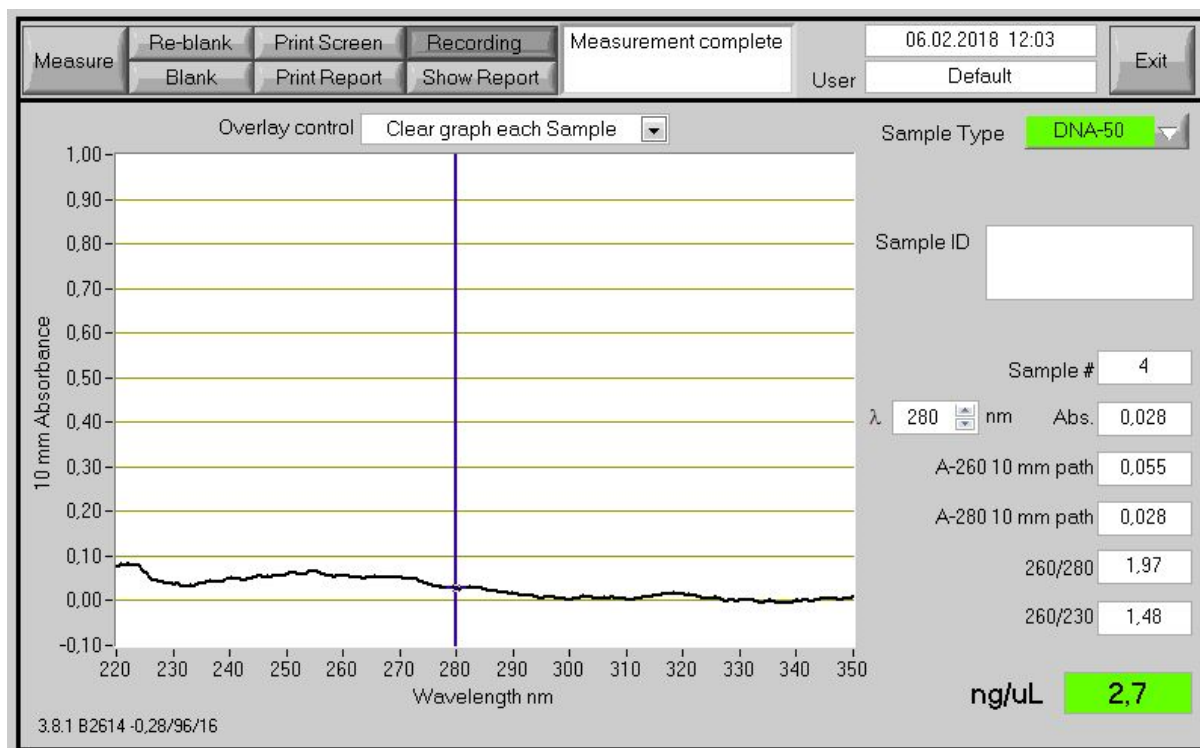
- 20 mL TAE x50 Stock solution
- 980 mL dH₂O

Appendiks 4.



Figur A.2. Agarosegel. Bilde er tatt av agarosegelen der det ble testet at alle prøvene inneholdt DNA-fragmenter med lik lengde. Størrelsemarkørene er plassert helt til høre, begge rader. De kolonnen som ikke viser bånd, er dette fordi de er blanke-kontroller (3 stk.).

Appendiks 5.



Figur A.3. Nanodrop-resultater. Bilde viser resultatet hvor blandeprøven ble testet for urenheter (260/280) og mengde DNA (ng/ μ L) med Nanodrop, før den ble sendt til Norsk Sekvenseringsenter.

Appendiks 6

```
#####  
## RAMI OROS PIPELINE i ABEL ##  
#####
```

```
#Hver gang(dag) når vi starter arbeid i ABEL, ber vi om ql og in  
ql og in --account=ui o --ntasks=1 --mem-per-cpu=4G --time=05:00:00
```

```
#og loader disse modulene.  
module purge  
module load amplicon_processing/1.5
```

```
## i hver forløkke som det her står *S*, så er dette fordi alle våre prøvemapper har et navn  
som inneholder S.
```

```
## LAGER ARBEID SOMRÅDET OG HENTE SEKVENSER
```

```
#lager directory som vi skal jobbe i, kaller det DK1.  
mkdir DK1  
cd DK1
```

```
#henter sekvensene vi skal arbeide med  
wget --user=edvardsenb-18s1 --ask-password https://hts-  
nonsecure.web.sigm2.no/180306_M01132.Project_EdvardsenB-18S1-2018-02-14/ --accept "*.tar" --  
recursive --no-directories --no-parent
```

```
#skriver inn passord tilsendt fra sekvenseringsenteret.
```

```
#sjekker hva vi har i mappen  
ls
```

```
#filen som ender på .gz, skal vi unzippe med tar -xvzf (filnavn)  
tar -xvzf (filnavn)
```



```

#Nå har vi en unzi ppa-mappe, med alle prøvemappene i. Gir mappen nytt navn som nå nå
arbeidsområdet.

#Noter hva FILNAVNENE heter.

#####
# Starter UPARSE pipeline #
#####

## STEG 1. SEPARERE FILER INN I MAPPER SOM HAR PRØVENAVNET.

#_L001 er i alle filnavnene, filene får også nye navn etter alt som står før _L001.
for i in $(awk -F"_L001" '{print $1}' <(ls *.fastq.gz) | sort | uniq); do mkdir $i; mkdir
$i/Raw; mv $i*.fastq.gz $i/Raw/.; done

## STEG 2. og 3. BAYESHAMMER OG SLÅ SAMMEN FOROVER- OG REVERS SEKVENSER (PEAR og USEARCH)

#lager slurmskript for begge jobbene
#slurmfil vist i nano
#####
# GNU nano 2.0.9 File: ampli con.slurm
#
# #! /bin/sh
# #SBATCH --job-name=spade_error
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=24:00:00
# #SBATCH --ntasks-per-node=4
# #SBATCH --mem-per-cpu=4G
# source /cluster/bin/jobsetup
#
# module load ampli con_processing/1.5
#
# STR="$(pwd | sed 's/^\.*\///')
#
# spades.py --only-error-correction -1 Raw/*_R1_* -2 Raw/*_R2_* -o $STR.corrected;
# pear2 -j 24 -n 200 -o $STR -f $STR.corrected/corrected/*_R1_* -r
$STR.corrected/corrected/*_R2_*;
#
# gunzip *.corrected/corrected/*_R1_*
# gunzip *.corrected/corrected/*_R2_*
#
# usearch -fastq_mergepairs $STR.corrected/corrected/*_R1_* -reverse
$STR.corrected/corrected/*_R2_* -fastout
# $STR.usearch.merged.fastq -fastq_maxdifs 10 -fastq_maxdq_maxdifs 10 -fastq_minovlen 16;
#####

#Kjører slurmskrip:
for i in $(ls -d *S*); do cp ampli con.slurm $i; cd $i; sbatch ampli con.slurm; cd ..; done

#Denne jobben vil kjøre til i morgen.
#sjekker at alt er ok

grep -c ^@MO */*fastq | grep -w 'usearch\|assembled'

#ser:
#####
# 38_S38/38_S38.assembled.fastq: 361152
# 38_S38/38_S38.usearch.merged.fastq: 237748
# 39_S39/39_S39.assembled.fastq: 333004
# 39_S39/39_S39.usearch.merged.fastq: 222535
# ...
#####

## STEG 4. KVALITETS FILTRERING
# har satt minimumlengde på sekvensene til 100 og maksimumforventningfeil =1

#PEAR
for i in $(ls -d *S*); do cd $i; usearch -fastq_filter *assembled.fastq -fastq_maxee 1 -
fastq_minlen 100 -fastaout $i.pear.longname.fna ;cd ..; done

#USEARCH
for i in $(ls -d *S*); do cd $i; usearch -fastq_filter *usearch.merged.fastq -fastq_maxee 1 -
fastq_minlen 100 -fastaout $i.usearch.longname.fna ;cd ..; done

#for å sammenligne sekvensene som har passert kvalitetskontrollen for PEAR og USEARCH
grep -c ">" */*longname*

#####
#38_S38/38_S38.pear.longname.fna: 50258
#38_S38/38_S38.usearch.longname.fna: 43735

```

```

#39_S39/39_S39.pear.longname.fna: 45861
#39_S39/39_S39.usearch.longname.fna: 40445
#...
#####

## STEG 5. FÖRENKLE FASTE OVERSKRIFTENE.
#gjør dette for pear og usearch

for i in $(ls -d *S*); do cd $i; simplifyFastaHeaders.pl $i.pear.longname.fna $i.seqnum
$i.pear.fna $i.map; cd ..; done
for i in $(ls -d *S*); do cd $i; simplifyFastaHeaders.pl $i.usearch.longname.fna $i.seqnum
$i.usearch.fna $i.map; cd ..; done

## STEG 6. SETTER SEKVENSER I SAMME RETNING (5'-3') vha. HMM
#6.1 Gjør om fastafilene slik at den blir indeksert, så HMM kan lese

#pear
for i in $(ls -d *S*); do cd $i; cdbfasta $i*pear.fna ; cd ..; done

#usearch
for i in $(ls -d *S*); do cd $i; cdbfasta $i*usearch.fna ; cd ..; done

#6.2
#6.2.1
#Lager to slurm-skript, et for pear og et for usearch.
#####
# GNU nano 2.0.9 File: hmm_pear.slurm
#
#!/bin/sh
# #SBATCH --job-name=hmm_pear
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=10:00:00
# #SBATCH --ntasks=16
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load ampli con_processing/1.5
#
# STR=$(pwd | sed 's/^.*\///')
# HMM3=/cluster/software/VERSIONS/ampli con_processing/1.5/mi tag/mi tags_extraction_protocol -
#master/mi TAGs_extraction_protocol /HMM3
#
# rna_hmm3.py -i $STR.pear.fna -o $STR.pear.rRNA -m ssu,lsu -k bac,arc,euk -p 16 -L $HMM3
#####
#####
# GNU nano 2.0.9 File: hmm_usearch.slurm
#
#!/bin/sh
# #SBATCH --job-name=hmm_usearch
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=10:00:00
# #SBATCH --ntasks=16
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load ampli con_processing/1.5
#
# STR=$(pwd | sed 's/^.*\///')
# #HMM3=/cluster/software/VERSIONS/ampli con_processing/1.5/mi tag/mi tags_extraction_pr#otocol -
#master/mi TAGs_extraction_protocol /HMM3
#
# rna_hmm3.py -i $STR.usearch.fna -o $STR.usearch.rRNA -m ssu,lsu -k bac,arc,euk -p #16 -L
#$HMM3
#####
#####
#kjører jobbene
for i in $(ls -d *S*); do cp hmm* $i; cd $i; sbatch hmm_usearch.slurm; sbatch hmm_pear.slurm;
cd ..; done

## 6.2.2
#kjører kommandoen under igjen om vært logget av
HMM3=/cluster/software/VERSIONS/ampli con_processing/1.5/mi tag/mi tags_extraction_protocol -
master/mi TAGs_extraction_protocol /HMM3

for i in $(ls -d *S*); do cd $i; parse_rna_hmm3_output.pl $i.pear.rRNA ; cd ..; done
for i in $(ls -d *S*); do cd $i; parse_rna_hmm3_output.pl $i.usearch.rRNA ; cd ..; done

#vi bestemmer oss for å bare jobbe videre med pear. Legger derfor alle usearch i en egen mappe
vi kalle USEARCH

```

```

for i in $(ls -d *S*); do cd $i; mkdir USEARCH; mv *usearch* USEARCH; cd ..; done

## 6.2.3
#lager slurmskriptet extract_rrna.slurm
#####
# GNU nano 2.0.9      File: extract_rrna.slurm
#
# #!/bin/sh
# #SBATCH --job-name=extract
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=24:00:00
# #SBATCH --ntasks=1
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
# module load amplicon_processing/1.5
#
# extract_rrna_seqs.pl *.rRNA.parsed 1 250
#####
#kjører jobben
for i in $(ls -d *S*); do cp extract_rrna.slurm $i; cd $i; sbatch extract_rrna.slurm; cd ..; done

#når jobben er ferdig
#sjekker at alle mappene er der, nr. Ser det er mindre av 16S.
ls *S*/*18S* -alh
ls *S*/*16S* -alh

## STEG 7. GI NYTT NAVN I UPARSE FORMAT
#fester identiteter på sekvensene (legger til prøvenavnet til sekvensoverskriften)
#bruker bare 18S
for i in $(ls -d *S*); do cd $i; cp *18S_rRNA $i.rRNA.fna ;cd ..; done
for i in $(ls -d *S*); do cd $i; sed "-es/^>\(.*)/>\1; barcodelabel=$i;/" < $i.rRNA.fna >
$i.rRNA.pear.uparse.fna ;cd ..; done

## STEG 8.
# Så langt har vi jobbet på hver enkelt prøve hver for seg, men nå er det på tide å slå de
sammen til en fil
# Det kan vi gjøre fordi alle sekvensene nå har navn etter den prøven de er fra
# Etter dette steget gjøres alle kommandoer på det sammenslåtte datasettet, og man trenger
ikke lenger å bruke for-løkker.
###slår resultatene sammen til fasta fil med alle sekvensene i

mkdir ../concatenated_sekvenser_same_dirction
for i in $(ls -d *S*); do cd $i; cp $i.rRNA.pear.uparse.fna
../concatenated_sekvenser_same_dirction ;cd ..; done
cat ../concatenated_sekvenser_same_dirction/*rRNA.pear.uparse.fna >
../concatenated_sekvenser_same_dirction/all_sekvenser_5_3dir_pear_UPARSEfmt.fna
mv ../concatenated_sekvenser_same_dirction/all_sekvenser_5_3dir_pear_UPARSEfmt.fna .
rm -rf ../concatenated_sekvenser_same_dirction

## STEG 9
#hopper over 9.1
## 9.2 DEREPLICASJON

usearch -derep_fulllength all_sekvenser_5_3dir_pear_UPARSEfmt.fna -fastaout
all_sekvenser_5_3dir_pear_UPARSEfmt_dereplicated.fna -si zeout

#####
#####
#usearch -derep_fulllength all_sekvenser_5_3dir_pear_UPARSEfmt.fna -fastaout
#all_sekvenser_5_3dir_pear_UPARSEfmt_dereplicated.fna -si zeout
#usearch v9.2.64_i86linux32, 4.0Gb RAM (65.9Gb total), 32 cores
#(C) Copyright 2013-16 Robert C. Edgar, all rights reserved.
#http://drive5.com/usearch
#
#License: sabry.razick@usit.uio.no
#
#00:25 2.8Gb 100.0% Sekvensing all_sekvenser_5_3dir_pear_UPARSEfmt.fna
#
#WARNING: Max OMP tsekvenser 1
#
#00:43 3.0Gb 100.0% DF
#00:43 3.1Gb 5619953 seqs, 1338884 uniques, 1053880 singletons (78.7%)
#00:43 3.1Gb Min size 1, median 1, max 107429, avg 4.20
#00:51 3.0Gb 100.0% Writing all_sekvenser_5_3dir_pear_UPARSEfmt_dereplicated.fna
#bash-4.1$
#####
#####
### Max OMP tsekvenser WARNING: Har ikke noe å si, så denne overser vi.

```

```

## 9.3 SETTER VARIABLELER
#Pear Merging
sekvenser_derep_pear=all_sekvenser_5_3dir_pear_UPARSEfmt_derepliated.fna
sekvenser_pear=all_sekvenser_5_3dir_pear_UPARSEfmt.fna

## 9.4 TETTHET OG KASTER SINGLE
#setter den med flest øverst og kaster single sekvenser
usearch -sortbysize $sekvenser_derep_pear -fastaout pear_sorted_sekvenser.fa -minsize 2

## 9.5 OTU CLUSTERING ved bruk av UPARSE og de-novo kimærsekk.
####Vi må bestemme oss for % likhet i OTU

#lager slurmskript for alle tre i nano

##99% clustering
#usearch -cluster_otus pear_sorted_sekvenser.fa -thsekvens 8 -otu_radius_pct 1 -otus
otus99_pear_repset.fa
#
#####
# GNU nano 2.0.9 File: cluster99.slurm
#
# #!/bin/sh
# #SBATCH --job-name=99cluster
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=140:00:00
# #SBATCH --ntasks=8
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load amplicon_processing/1.5
#
# usearch -cluster_otus pear_sorted_sekvenser.fa -thsekvenser 8 -otu_radius_pct 1 -otus
otus99_pear_repset.fa
#####

##98% clustering
#usearch -cluster_otus pear_sorted_sekvenser.fa -thsekvens 8 -otu_radius_pct 2 -otus
otus98_pear_repset.fa
#
#####
# GNU nano 2.0.9 File: cluster98.slurm
#
# #!/bin/sh
# #SBATCH --job-name=98cluster
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=140:00:00
# #SBATCH --ntasks=8
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load amplicon_processing/1.5
#
# usearch -cluster_otus pear_sorted_sekvenser.fa -thsekvenser 8 -otu_radius_pct 2 -otus
otus98_pear_repset.fa
#
#####

##97% clustering
#usearch -cluster_otus pear_sorted_sekvenser.fa -thsekvens 8 -otus otus97_pear_repset.fa
#
#####
# GNU nano 2.0.9 File: cluster97.slurm
#
# #!/bin/sh
##SBATCH --job-name=97cluster
##SBATCH --account=ui o
##SBATCH --output=slurm-%j.base
##SBATCH --time=140:00:00
##SBATCH --ntasks=8
##SBATCH --mem-per-cpu=4G
#
#source /cluster/bin/jobsetup
#
#module load amplicon_processing/1.5
#

```

```

#usearch -cluster_otus pear_sorted_sekvenser.fa -thsekvenser 8 -otus otus97_pear_repset.fa
#####

#lager ett nytt dir til å ha disse jobbene i, flytter slurmskripta dit, og siste jobben
mkdir CLUSTER
mv cluster* CLUSTER
mv pear_sorted_sekvenser.fa CLUSTER/

#kjører jobbene
sbatch cluster99.slurm
sbatch cluster98.slurm
sbatch cluster97.slurm

# VIKTIG! Når jobben er ferdig:
#sjekker og skriver opp ANTALL OTU'er (ved 100%), dette skal brukes videre i steg 9.8
#####
#-bash-4.1$ more slurm-21204274.base
#Starting job 21204274 ("97cluster") on c17-[30-33],c18-29 at Mon May 7 11:14:57 CEST 2018
#usearch v9.2.64_i86linux32, 4.0Gb RAM (65.9Gb total), 32 cores
#(C) Copyright 2013-16 Robert C. Edgar, all rights reserved.
#http://drive5.com/usearch
#
#License: sabry.razi@usit.uio.no
#
#00:11 53Mb 16.2% 892 OTUs, 510 chimeras
#00:22 56Mb 32.4% 1221 OTUs, 1004 chimeras
#00:32 56Mb 45.1% 1462 OTUs, 1729 chimeras
#00:43 57Mb 55.1% 1631 OTUs, 2338 chimeras
#00:54 58Mb 63.0% 1816 OTUs, 3152 chimeras
#01:04 59Mb 70.3% 1930 OTUs, 3726 chimeras
#01:15 59Mb 77.6% 2093 OTUs, 4379 chimeras
#01:26 60Mb 85.2% 2246 OTUs, 4956 chimeras
#01:37 61Mb 92.1% 2392 OTUs, 5682 chimeras
#01:47 61Mb 99.0% 2551 OTUs, 6332 chimeras
#01:48 61Mb 100.0% 2567 OTUs, 6406 chimeras
#
#WARNING: Option -thsekvenser not used #ikke noe å bry seg om
#####

more slurm-21204275.base
#####
#Starting job 21204275 ("98cluster") on c17-[30-32],c18-29 at Mon May 7 11:17:23 CEST 2018
#02:27 63Mb 100.0% 2771 OTUs, 10408 chimeras
#####

more slurm-21204276.base
#####
#Starting job 21204276 ("99cluster") on c17-[6,12,18,28] at Mon May 7 11:22:40 CEST 2018
#06:18 65Mb 100.0% 2956 OTUs, 13826 chimeras
#####

## 9.7 NAVNSETTER OTU-SEKVENSER

scripts=/cluster/software/VERSIONS/amplicon_processing/1.5/scripts
python $scripts/fastanumber.py otus99_pear_repset.fa OTU_ > otus99_pear_repset_clean.fa
python $scripts/fastanumber.py otus98_pear_repset.fa OTU_ > otus98_pear_repset_clean.fa
python $scripts/fastanumber.py otus97_pear_repset.fa OTU_ > otus97_pear_repset_clean.fa

## 9.8 MAP SEKVENSER (INCLUDING SINGLETONS) TILBAKE TIL OTUS

#Går inn i nano og lager slurmskript for alle jobbene.
#HUSK!!! Endre maxrejects til det antall OTUer som blir laget i job 9.5. Endre også ntasks til
16 og legg inn kommandoen for hver jobb, se under.

#Disse to kommandolinjene må også inn i jobbskripte:
# sekvenser_derep_pear=all_sekvenser_5_3dir_pear_UPARSEfmt_dereplicated.fna
# sekvenser_pear=all_sekvenser_5_3dir_pear_UPARSEfmt.fna

#99%
#usearch -usearch_global $sekvenser_pear -db otus99_pear_repset_clean.fa -strand plus -id 0.99
-uc map99.pear.uc -maxhits 1 -maxaccepts 20 -maxrejects 2956 -thsekvenser 16
#
#####
# GNU nano 2.0.9 File: mapsekvenser99.slurm
#
##!/bin/sh
##SBATCH --job-name=99mapsekvenser
##SBATCH --account=ui o
##SBATCH --output=slurm-%j.base
##SBATCH --time=140:00:00
##SBATCH --ntasks=16
##SBATCH --mem-per-cpu=4G
#

```

```

#
# source /cluster/bin/jobsetup
#
# module load amplicon_processing/1.5
#
# sekvenser_derep_pear=all_sekvenser_5_3dir_pear_UPARSEfmt_derepliated.fna
# sekvenser_pear=all_sekvenser_5_3dir_pear_UPARSEfmt.fna
#
# usearch -usearch_global $sekvenser_pear -db otus99_pear_repset_clean.fa -strand plus -id
0.99 -uc map99.pear.uc -maxhits 1 #-maxaccepts 20 -maxrejects 2956 -thsekvenser 16
#####

#98%
#usearch -usearch_global $sekvenser_pear -db otus98_pear_repset_clean.fa -strand plus -id 0.98
-uc map98.pear.uc -maxhits 1 -maxaccepts 20 -maxrejects 2771 -thsekvenser 16
#
#####
#
# GNU nano 2.0.9          File: mapsekvenser98.slurm
#
# #!/bin/sh
# #SBATCH --job-name=98mapsekvenser
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=140:00:00
# #SBATCH --ntasks=16
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load amplicon_processing/1.5
#
# sekvenser_derep_pear=all_sekvenser_5_3dir_pear_UPARSEfmt_derepliated.fna
# sekvenser_pear=all_sekvenser_5_3dir_pear_UPARSEfmt.fna
#
# usearch -usearch_global $sekvenser_pear -db otus98_pear_repset_clean.fa -strand plus -id
0.98 -uc map98.pear.uc -maxhits 1 #-maxaccepts 20 -maxrejects 2771 -thsekvenser 16
#
#####

#97%
#usearch -usearch_global $sekvenser_pear -db otus97_pear_repset_clean.fa -strand plus -id 0.97
-uc map97.pear.uc -maxhits 1 -maxaccepts 20 -maxrejects 2567 -thsekvenser 16
#
#####
#
# GNU nano 2.0.9          File: mapsekvenser97.slurm
#
# #!/bin/sh
# #SBATCH --job-name=97mapsekvenser
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=140:00:00
# #SBATCH --ntasks=16
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load amplicon_processing/1.5
#
# sekvenser_derep_pear=all_sekvenser_5_3dir_pear_UPARSEfmt_derepliated.fna
# sekvenser_pear=all_sekvenser_5_3dir_pear_UPARSEfmt.fna
#
# usearch -usearch_global $sekvenser_pear -db otus97_pear_repset_clean.fa -strand plus -id
0.97 -uc map97.pear.uc -maxhits 1 #-maxaccepts 20 -maxrejects 2567 -thsekvenser 16
#
#####

#flytter inn begge, all_sekvenser_5_3dir_pear_UPARSEfmt og
all_sekvenser_5_3dir_pear_UPARSEfmt_derepliated filene inn i CLUSTER-mappa.
mv all_sekvenser_5_3dir_pear_UPARSEfmt.fn CLUSTER
mv all_sekvenser_5_3dir_pear_UPARSEfmt_derepliated.fn CLUSTER

#kjører jobbene;
sbatch mapsekvenser97.slurm
sbatch mapsekvenser98.slurm
sbatch mapsekvenser99.slurm

#disse tar tid.

## 9.9 LAGE OTU TABELLER

scripts=/cluster/software/VERSIONS/amplicon_processing/1.5/scripts

```

```

#99%
python $scripts/uc2otutab.py map99.pear.uc > otu_table99.pear.txt

#sjekker OTU tabel
wc otu_table99.pear.txt
grep -c ">" otus99_pear_repset_clean.fa

#98%
python $scripts/uc2otutab.py map98.pear.uc > otu_table98.pear.txt

wc otu_table98.pear.txt #n-1 = ant. OTUs
grep -c ">" otus98_pear_repset_clean.fa

#97%
python $scripts/uc2otutab.py map97.pear.uc > otu_table97.pear.txt

wc otu_table97.pear.txt
grep -c ">" otus97_pear_repset_clean.fa

## STEG 10. TAKSONOMISK KLASSIFISERING
#dette tar tid, lag slurmskript for alle jobbene

#disse to kommandolinjene skal i alle skripta
silva132=/usr/t/abel/u1/anderkkr/blastdb/SILVA_132/SILVA_132_blastDB
pr2=/usr/t/abel/u1/anderkkr/blastdb/PR2/pr2_version_4.10.0

# 99% clustering
blastn -db $silva132 -query otus99_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 -max_target_seqs 1 -eval ue 0.0001 -out blastn_vs_SILVA_v132_eval ue10mi n4_99per_pear - num_thsekvenser 16
blastn -db $pr2 -query otus99_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 - max_target_seqs 1 -eval ue 0.0001 -out blastn_vs_pr2_eval ue10mi n4_99per_pear -num_thsekvenser 16

#98%
blastn -db $silva132 -query otus98_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 -max_target_seqs 1 -eval ue 0.0001 -out blastn_vs_SILVA_v132_eval ue10mi n4_98per_pear - num_thsekvenser 16
blastn -db $pr2 -query otus98_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 - max_target_seqs 1 -eval ue 0.0001 -out blastn_vs_pr2_eval ue10mi n4_98per_pear -num_thsekvenser 16

# 97% clustering
blastn -db $silva132 -query otus97_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 -max_target_seqs 1 -eval ue 0.0001 -out blastn_vs_SILVA_v132_eval ue10mi n4_97per_pear - num_thsekvenser 16
blastn -db $pr2 -query otus97_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 - max_target_seqs 1 -eval ue 0.0001 -out blastn_vs_pr2_eval ue10mi n4_97per_pear -num_thsekvenser 16

####eksempel på slurm-skript til disse tre jobbene
#
#####
# #!/bin/sh
# #SBATCH --job-name=99BLAST
# #SBATCH --account=ui o
# #SBATCH --output=slurm-%j.base
# #SBATCH --time=140:00:00
# #SBATCH --ntasks=16
# #SBATCH --mem-per-cpu=4G
#
# source /cluster/bin/jobsetup
#
# module load amplicon_processing/1.5
#
# silva132=/usr/t/abel/u1/anderkkr/blastdb/SILVA_132/SILVA_132_blastDB
# pr2=/usr/t/abel/u1/anderkkr/blastdb/PR2/pr2_version_4.10.0
#
# blastn -db $silva132 -query otus99_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 -max_target_seqs 1 -#eval ue 0.0001 -out blastn_vs_SILVA_v132_eval ue10mi n4_99per_pear - num_thsekvenser 16
# blastn -db $pr2 -query otus99_pear_repset_clean.fa -outfmt '6 std qlen' -perc_identity 75 - max_target_seqs 1 -eval ue #0.0001 -out blastn_vs_pr2_eval ue10mi n4_99per_pear -num_thsekvenser 16
#####
#send inn jobbene
sbatch BLAST97.slurm

```

```
sbatch BLAST98. slurm
sbatch BLAST99. slurm
```

Appendiks 7.

R Skript

```
#pakker som ble burkt:
library(grid)
library(reshape2)
library(treemap)
library(rlang)
library(ggplot2)
library(vegan)
library(egg)
library(RColorBrewer)
source("https://biocductor.org/biocLite.R")
biocLite("systemPipeR")
library("systemPipeR")
library(Lattice)
library(MASS)
library(stats)
library(oce)
library(tables)

#####
## CTD-filer ##
#####

ctd.file <- "DK1_24022016_oks.cnv"
ctd <- sekvens.table(ctd.file, skip=219) #hopper over de første 219 radene.

names(ctd) <- c(
  "depth", # Depth [salt water, m]
  "temp", # Temperature [ITS-90, deg C]
  "sal", # Salinity, Practical [PSU]
  "sigT", # Density [sigma-t, Kg/m^3]
  "flSP", # Fluorescence, Seapoint
  "flWL", # Fluorescence, WET Labs WETstar [mg/m^3]
  "turb", # Turbidity, Seapoint [FTU]
  "svCM", # Sound Velocity [Chen-Millero, m/s]
  "Ox (%sat)", #prosent Oksygen saturasjon
  "Ox (ml/L)", #mengde O2 i ml/L
  "Ox (mg/L)", #mengde O2 i mg/L
  "flag") # flag: 0.000e+00

#Runder av til hver meter.
table(round(ctd$depth))

par(mfrow=c(1, 1))
plot(table(round(ctd$depth)))
median(table(round(ctd$depth))) # Ant målinger/meter

z <- round(ctd$depth)
ctd.mean <- aggregate(ctd[, 2:12], by=list(z), mean)
names(ctd.mean)[1] <- "depth"

#for å sjekke at alt ser fint ut.
par(mfrow=c(1, 3))
plot(-depth ~ temp, type="l", data=ctd.mean)
points(-depth ~ temp, pch=".", col=2, data=ctd)
plot(-depth ~ sal, type="l", data=ctd.mean)
points(-depth ~ sal, pch=".", col=2, data=ctd)
plot(-depth ~ sigT, type="l", data=ctd.mean)
points(-depth ~ sigT, pch=".", col=2, data=ctd)
#skriv ut fil
write.csv(file = "ctd.mean24022016.csv", ctd.mean)

#####
## CTD ggplot ##
#####

#Figur 3.6

gr <- sekvens.table("ctd_2_5b.txt", header = TRUE)
gr$date <- factor(gr$date, levels=c("Feb. 16", "Apr. 16", "Jun. 16", "Aug. 16", "Sep. 16",
"Nov. 16", "Feb. 17", "Apr. 17", "Mai. 17", "Jun. 17", "Aug. 17", "Sep. 17", "Okt. 17", "Nov. 17",
"Des. 17"))

#temp
gt <- ggplot(gr, aes(x = date, y = temp)) +
```



```

geom_point(aes(shape = depth, colour=depth)) +
geom_line(aes(group = depth, linetype = depth, colour=depth)) +
theme(axis.text.x = element_text(angle=45, vjust=0.5,
colour="darkgrey"), legend.position="none") +
ylab("Temperatur (C)") +
xlab("Dato")

#sal tholdi ghet
gs <- ggplot(gr, aes(x = date, y = sal)) +
  geom_point(aes(shape = depth, colour=depth)) +
  geom_line(aes(group = depth, linetype = depth, colour=depth)) +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"),
  legend.position="none") +
  xlab("Dato") +
  ylab("Salinitet")

#fluorescense
gf2 <- ggplot(gr, aes(x = date, y = flSP)) +
  geom_point(aes(shape = depth, colour=depth)) +
  geom_line(aes(group = depth, linetype = depth, colour=depth)) +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey")) +
  xlab("Dato") +
  ylab("Flourescens (mg/m^3)")

#Oksygen
go <- ggplot(gr, aes(x = date, y = Ox...sat.)) +
  geom_point(aes(shape = depth, colour=depth)) +
  geom_line(aes(group = depth, linetype = depth, colour=depth)) +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"),
  legend.position="none") +
  xlab("Dato") +
  ylab("Oksygensaturasjon (%)")

#tetthet
gth <- ggplot(gr, aes(x = date, y = sigT)) +
  geom_point(aes(shape = depth, colour=depth)) +
  geom_line(aes(group = depth, linetype = depth, colour=depth)) +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"),
  legend.position="none") +
  xlab("Dato") +
  ylab("Tetthet")

```

```
ggarrange(gt, gs, gth, go, gf2, ncol = 3, nrow = 2)
```

```
#####
## TREEMAP ##
#####
```

```
#Figur 3.1 og 3.5
```

```
all <- sekvens.table(file="treemap.m.meta2", header = TRUE)
as.data.frame(all)
```

```
windows() #Åpener et nytt vindu for plottet da R studio ikke viser det godt nok.
vplayout <- function(x, y) viewport(layout.pos.row = x, layout.pos.col = y)
grid.newpage()#Lager et plot fra start på en ny blank, hvis ikke overlapper den sist plot.
pushViewport(viewport(layout = grid.layout(1, 2)))
```

```
treemap(all,
  index=c("Supergruppe", "Gruppe"),
  vsize="OTUs",
  vcolor=y,
  type="index",
  palette = "MinPalett",
  vp = vplayout(1,1))
treemap(all,
  index=c("Supergruppe", "Gruppe"),
  vsize="Sekvenser",
  vcolor= y,
  type="index",
  palette = "MinPalett",
  vp = vplayout(1,2))
```

```
#####
## Histogram (sekvenser per OTU) ##
#####
```

```
#Figur 3.2
```

```

sum <- sekvens.table("sum_ant_otuer.txt", header = T)
plot (sum)

p<-ggplot (data=sum, aes(x=OTU_id, y=sum)) +
  geom_bar(stat="identity") +
  ylim (0,1000)

#####
## Rarefaction-kurver ##
#####

#Figur 3.3

data <- sekvens.table("rarkurve_før_sub.txt", header = TRUE, sep="\t")
data2 <- data[,-1]
tdata <- t(data2)
raremax <- min(rowSums(tdata))

col <- c("black", "darkred", "forestgreen", "orange", "blue", "yellow", "hotpink")
lty <- c("solid", "dashed", "longdash", "dotted")
pars <- expand.grid(col = col, lty = lty, stringsAsFactors = FALSE)
head(pars)

raremax2 <- min(rowSums(t(data2)))
Srare <- rarefy(t(data2), raremax)
plot(specnumber(t(data2)), Srare, xlab = "Observert antall OTUer", ylab = "Rarefied antall
OTUer")
abline(0, 1)
out <- rarecurve(t(data2), step = 20, sample = raremax, col = "blue", cex = 0.6)

Nmax <- sapply(out, function(x) max(attr(x, "Sample")))
Smax <- sapply(out, max)
plot(c(1, max(Nmax)), c(1, max(Smax)), xlab = "Sekvenser per prøve",
      ylab = "Antall OTUer", type = "n")
abline(v = raremax)
for (i in seq_along(out)) {
  N <- attr(out[[i]], "Sample")
  with(pars, lines(N, out[[i]], col = col[i], lty = lty[i]))
}

#####
## Subsampling ##
#####

data <- sekvens.table("Protist.txt", row.name = 1, header = TRUE)
y <- as.data.frame(t(data))

#finn laveste nummer av sekvenser i samplene
min(rowSums(y))

# Subsampling
otu_rrarefy<-rrarefy(y, sample = 10530) # min(rowSums(otu_trans))
otu_rrarefy2<-rrarefy(y, sample = 10530) # min(rowSums(otu_trans))
otu_rrarefy3<-rrarefy(y, sample = 10530)
otu_rrarefy4<-rrarefy(y, sample = 10530)
cor.test(otu_rrarefy, otu_rrarefy2)
cor.test(otu_rrarefy, otu_rrarefy3)
cor.test(otu_rrarefy, otu_rrarefy4)
cor.test(otu_rrarefy2, otu_rrarefy3)
cor.test(otu_rrarefy2, otu_rrarefy4)
cor.test(otu_rrarefy3, otu_rrarefy4)

write.csv(otu_rrarefy, "Subsamplet")

#####
## Barplot over antall OTUer ##
#####

#Figur 3.9

#0-2m

ant2 <- sekvens.table("ant0TU2m.txt", header = TRUE, row.names = 1)
]

otu2 <- melt(as.matrix(ant2), useNA = "ifany")
names(otu2) <- c("Gruppe", "Dato", "Ant.OTUer")

plototu2 <- ggplot (otu2, aes(x=Dato, y=Ant.OTUer, fill=Gruppe, useNA = "ifany")) +

```

```

coord_cartesian(ylim = c(0, 580))+
geom_bar(stat="identity", na.rm = FALSE)+
scale_fill_manual(values = "MinPal ett", name="Gruppe")

#5m
ant5 <- sekvens.table("ant0TU5m.txt", header = TRUE, row.names = 1)

otu5 <- melt(as.matrix(ant5), useNA = "ifany")
names(otu5) <- c("Gruppe", "Dato", "Ant. 0TUer")

plototu5 <- ggplot(otu5, aes(x=Dato, y=Ant.0TUer, fill=Gruppe, useNA = "ifany")) +
  coord_cartesian(ylim = c(0, 580))+
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"))+
  geom_bar(stat = "identity")+
  scale_fill_manual(values = "MinPal ett", name="Gruppe")

ggarrange(plototu2, plototu5, ncol = 1, nrow = 2)

#####
## Andel sekvenser (%) ##
#####

#figur 3.10

#0-2m

data <- sekvens.table("prob.sekvenser.2m.NA", header = T, row.names = 1)
head(data)
colSums(data)
pro <- prop.table(as.matrix(data), margin = 2)
sum(pro)
bptot <- pro*100

#write.table(bptot, "prosent_sekvenser_2m.txt", col.names= TRUE)

tx <- melt(as.matrix(bptot))
head(tx)
names(tx) <- c("Gruppe", "Dato", "value")

pbp2 <- ggplot(tx, aes(x=Dato, y=value, ymin=0, ymax=value, fill = Gruppe)) +geom_bar(stat =
"identity") +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"))+
  ylab("Andel sekvenser (%)") +
  xlab("") +
  scale_fill_manual(values = "MinPal ett", name="Gruppe") +
  ggtitle("0-2 m")

#5m

data5 <- sekvens.table("prob.sekvenser.5m", header = T, row.names = 1)
head(data5)
colSums(data5)
pro5 <- prop.table(as.matrix(data5), margin = 2)
sum(pro5)
bptot5 <- pro5*100

#write.table(bptot5, "prosent_sekvenser_5m.txt", col.names= TRUE)

tx5 <- melt(as.matrix(bptot5))
head(tx5)
names(tx5) <- c("Gruppe", "Dato", "value")

pbp5 <- ggplot(tx5, aes(x=Dato, y=value, ymin=0, ymax=value, fill = Gruppe)) +geom_bar(stat =
"identity") +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"))+
  ylab("Andel sekvenser (%)") +
  xlab("Dato") +
  scale_fill_manual(values = "MinPal ett", name="Gruppe") +
  ggtitle("5 m")

#legge plotta i samme bildevindu
windows()
ggarrange(pbp2, pbp5, ncol = 1, nrow = 2)

#linjeplot
plot <- ggplot(trx5m, aes(x=Dato, y=Prosent, colour=Gruppe, group=Gruppe)) +
  geom_line(na.rm = TRUE)

```

```

#####
## Heatmap ##
#####

#Figur 3.11

myPalette2 <- brewer.pal(6, "Blues")

#5m

heat5 <- sekvens.table("heat5_sum.txt", header = TRUE, row.names = 1, sep = "\t")

ma5 <- prop.table(as.matrix(heat5), margin=2)
pro5 <- ma5*100
write.table(pro5, "%fordeling_klasse_heat_5m", col.names = T)

h5 <- melt(pro5)
colnames(h5) <- c("Klasse", "Dato", "Prosent")
h5[h5==0] <- NA

heatmap5 <- ggplot(h5, aes(Dato, Klasse)) +
  geom_tile(aes(fill=Prosent)) +
  geom_text(aes(fill=h5$Prosent), label = round(h5$Prosent, 2), size=2) +
  scale_fill_gradientn(colours = myPalette2, na.value="white", space = "Lab", guide =
"colourbar") +
  theme(axis.text.y = element_text(size = 8), axis.text.x = element_text(size = 8, angle=90),
axis.title.y=element_blank(), legend.position="none")
#element_text(size = 8)
#rownames(h5) <- c("SUM", "Feb.", "Apr.", "Jun.", "Aug.", "Sep.", "Nov", "Feb.(t)",
"Feb.(s)", "Mar.", "Apr.", "Mai", "Jun.", "Aug.", "Sep.", "Okt.", "Nov.", "Des.")

#0-2m

heat2 <- sekvens.table("heat2_sum.txt", header = TRUE, row.names = 1, sep = "\t")

ma2 <- prop.table(as.matrix(heat2), margin=2)
pro2 <- ma2*100
write.table(pro2, "%fordeling_klasse_heat_2m", col.names = T)

h2 <- melt(as.matrix(pro2))
colnames(h2) <- c("Klasse", "Dato", "Prosent")
h2[h2==0] <- NA

heatmap2 <- ggplot(h2, aes(Dato, Klasse)) +
  geom_tile(aes(fill=Prosent)) +
  geom_text(aes(fill=h2$Prosent), label = round(h2$Prosent, 2), size=2) +
  scale_fill_gradientn(colours = myPalette2, na.value="white", space = "Lab", guide =
"colourbar") +
  theme(axis.text.y = element_text(size = 8), axis.text.x = element_text(size = 8, angle=90),
axis.title.y=element_blank(), legend.position="none")

#ggarrange(heatmap2, heatmap5, ncol = 2, nrow = 1)

#####
## % TROFI ##
#####

#Figur 3.12

#0-2m

tro2m <- sekvens.table("trofi 2m.NA", header = T, row.names = 1)
head
colSums(tro2m)
ptro2m <- prop.table(as.matrix(tro2m), margin = 2)
sum(ptro2m)
bptro2m <- ptro2m*100

trx2m <- melt(as.matrix(bptro2m))
head(trx2m)
names(trx2m) <- c("Trofi", "Dato", "Prosent")

pt <- ggplot(trx2m, aes(x=Dato, y=Prosent, ymin=0, ymax=Prosent, fill = Trofi)) +geom_bar(stat =
"identity") +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"))+
  ylab("Relativ abundans (%)") +
  xlab("") +
  scale_fill_manual((values = "MyPalette"), name="Gruppe") +
  ggtitle("0-2 m")

```

```

#5m

tro5m <- sekvns.table("trofi5m.NA", header = T, row.names = 1)
head
colSums(tro5m)
ptro5m <- prop.table(as.matrix(tro5m), margin = 2)
sum(ptro5m)
bptro5m <- ptro5m*100

trx5m <- melt(as.matrix(bptro5m))
head(trx5m)
names(trx5m) <- c("Trofi", "Dato", "Prosent")

pt5 <- ggplot(trx5m, aes(x=Dato, y=Prosent, ymin=0, ymax=Prosent, fill = Trofi)) +geom_bar(stat =
"identity") +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"))+
  ylab("Relativ abundans (%)") +
  xlab("") +
  scale_fill_manual(values = "MyPallett"), name="Gruppe") +
  ggtitle("5 m")

ggarrange(pt2, pt5, ncol = 1, nrow = 2)

#linjeplot
plot <- ggplot(trx5m, aes(x=Dato, y=Prosent, colour=Trofi, group=Trofi)) +
  geom_line(na.rm = TRUE)

#####
## Venn diagram ##
#####

#Figur 3.14, figur 3.15 og figur 3.17

df.venn <- sekvns.table("chi.txt", header = TRUE)
head(df.venn)

MB <- as.character(df.venn[df.venn$Metastrekking==1,1])
LM <- as.character(df.venn[df.venn$Lysmikroskop==1,1])

venn <- list(mb=MB, lm=LM)

vennset_all <- overlap(venn, type="vennsets")
vennplot(vennset_all, mymain = "Venndiagram over delte taksa", line=c("cyan", "yellow"))

#####
## CTD sammenlignet ##
#####

#Figur 3.16

CTD <-sekvns.csv("ctd.all.mean_b", sep = "\t", dec = ",")

summary(CTD)

#endre dato til dato, dag i året og år, hver for seg.
CTD$date = as.Date(CTD$date, "%d.%m.%Y")
CTD$DoY = as.numeric(format(CTD$date, "%j"))
CTD$Year = as.numeric(format(CTD$date, "%Y"))

#hente ut målingene fra de to dypene
CTD5 = subset(CTD, subset = (depth == 5))
CTD02 = subset(CTD, subset = (depth <= 2))

#Temperatur
CTD2t = aggregate(temp~Date, data = CTD02, FUN = mean)

dt <- as.data.frame(cbind(CTD5$temp, CTD2t$temp))
t1 <- ggplot(dt, aes(V1, V2)) +
  geom_point() +
  stat_smooth(method = lm, se = FALSE)+
  xlab("5m")+
  ylab("2m")+
  ggtitle("Temperatur (C)")

summary(lm(V2 ~ V1, dt))

#Salinitet
CTD2s = aggregate(sal~Date, data = CTD02, FUN = mean)

ds <- as.data.frame(cbind(CTD5$sal, CTD2s$sal))
s2 <- ggplot(ds, aes(V1, V2)) +

```

```

geom_point() +
stat_smooth(method = lm, se = FALSE)+
xlab("5m")+
ylab("2m")+
ggtitle("Salinitet")

summary(lm(V2 ~ V1, ds))

#Tetthet
CTD2sT = aggregate(sigT-Date, data = CTD02, FUN = mean)
dsT <- as.data.frame(cbind(CTD5$sigT, CTD2sT$sigT))
th5 <- ggplot(dsT, aes(V1, V2)) +
  geom_point() +
  stat_smooth(method = lm, se = FALSE)+
  xlab("5m")+
  ylab("2m")+
  ggtitle("Tetthet (kg/m^3)")

summary(lm(V2 ~ V1, dsT))

#Oksygen
CTDox = aggregate(Ox...sat.-Date, data = CTD02, FUN = mean)
dox <- as.data.frame(cbind(CTD5$Ox...sat, CTDox$Ox...sat))
o4 <- ggplot(dox, aes(V1, V2)) +
  geom_point() +
  stat_smooth(method = lm, se = FALSE)+
  xlab("5m")+
  ylab("2m")+
  ggtitle("Oksygensaturasjon (%)")

summary(lm(V2 ~ V1, dox))

#Flourescens
CTD2f = aggregate(flSP-Date, data = CTD02, FUN = mean)
df <- as.data.frame(cbind(CTD5$flSP, CTD2f$flSP))
fSP <- ggplot(df, aes(V1, V2)) +
  geom_point() +
  stat_smooth(method = lm, se = FALSE)+
  xlab("5m")+
  ylab("2m")+
  ggtitle("Flourescens")

summary(lm(V2 ~ V1, fSP))

windows()
ggarrange(t1, s2, th5, o4, fSP, ncol = 3, nrow = 2)

#sjekke salinitet og tetthet mot hverandre for de to dypene
summary(lm(dsT$V1 ~ ds$V1))
summary(lm(dsT$V2 ~ ds$V2))

#####
## Linjeplot ##
#####

#Figur 3.10 C og figur 3.12 C

library(reshape2) #melt-funksjonen
library(ggplot2)

#bruker samme inputfil som %sekvens-barplot.
data5 <- sekvens.table("prob.sekvenser.5m", header = T, row.names = 1)
pro5 <- prop.table(as.matrix(data5), margin = 2)
sum(pro5)
bptot5 <- pro5*100

line <- melt(as.matrix(bptot5))
names(line) <- c("Gruppe", "Dato", "Sekvenser")

plot <- ggplot(line, aes(x=Dato, y=Sekvenser, colour=Gruppe, group=Gruppe))+
  geom_line() +
  theme(axis.text.x = element_text(angle=45, vjust=0.5, colour="darkgrey"))+
  scale_fill_manual(values = "MyPallett", name="Gruppe")

#####
## gnmds ##
#####

#Figur 3.18 og figur 3.22

```

```

df <- sekvens.table("ordinasjon.txt", row.name = 3, header = TRUE)
ctd <- sekvens.table("ctd.txt", header = TRUE)
ctd2 <- ctd[, -c(7:9, 11:13)]

data <- cbind(ctd2, df)
data2 <- data[, -c(1,2)]

y <- data[, -c(1:9)]

# Lage Bray-Curtis ulikhets-matrise:
dist.y <- vegdist(y, method="bray")

# Erstatte avstander (B-C > 0.8 by geodesic avstander)
geodist.y <- isomapdist(dist.y, epsilon=0.8)

k=2 #k = ant. Dimensjoner i ordinasjonen
mds <- NULL #definerer et tomt objekt
# Lager 1000 "mds"er fra første starts konfigurasjon og sette dem inn i det # mds-objektet
for(i in 1:1000)
{mds[[i]] <- monoMDS(geodist.y, matrix(c(runif(dim(y)[1]*k)), nrow=dim(y)[1]),
                    k=2, model = "global", maxit=20000, smin = 1e-7, sfgrmin = 1e-7)}

# mds-objektet er nå en liste med 1000 "Sub-objekter", som igjen er
# listet.
# tar ut stress-verdiene som en vektor
# stress er gitt som det 22 elementet i hver "sub-objekt"-liste.
mds.stress <- unlist(lapply(mds, function(v){v[[22]]}))

# rangere stressverdiene
order(mds.stress)
# lagre rangeringen som en vektor
ordered <- order(mds.stress)

# finne stressverdien til løsningen med lavest og nest lavest stress-
# verdier
mds.stress[ordered[1]]
mds.stress[ordered[2]]

# skalere aksene til half change units og varimax rotation med postMDS
mds.best <- postMDS(mds[[ordered[1]]], geodist.y, pc = TRUE, halfchange = TRUE, threshold = 0.8)
mds.secbest <- postMDS(mds[[ordered[2]]], geodist.y, pc = TRUE, halfchange = TRUE, threshold =
0.8)
par(mfrow=c(1,1))

# Procrustes sammenligning
procrustes(mds.best, mds.secbest, permutations=999)
protest(mds.best, mds.secbest, permutations=999)

plot(procrustes(mds.best, mds.secbest, permutations=999))

# Lage variabler for GNMDS for 1 og 2 akse
gnm1 <- mds.best$points[, 1]
gnm2 <- mds.best$points[, 2]

# få miljøvariablene inn i plottet
ctd.var <- ctd2[, -c(1,2)]
df <- data.frame(gnm1, gnm2)
fit <- envfit(df, ctd.var, 999, na.rm = TRUE)
plot(fit, arrow.mul = 0.75, p=0.05, col=1, cex=0.75)

fit.df <- data.frame(scores(fit, display = "vectors"))

mds.z <- envfit(scores(mds.best, display="sites", choices=1:2, origin=TRUE)[, 1:2], ctd.var, 999,
na.rm = TRUE)
plot(gnm1, gnm2, type="n")
plot(mds.z, arrow.mul=1.6, col=3, add=T, cex=0.7)

# plottet
ggplot(data, aes(x= gnm1, y= gnm2, colour=Dyp, label=Dato, group=Dato)) +
  geom_point(size=2) +
  geom_text(aes(label=Dato), hjust=0, vjust=0) +
  geom_line() +
  geom_segment(data=fit.df, aes(x=0, xend=gnm1, y=0, yend=gnm2),
              arrow = arrow(length = unit(0.5, "cm")), colour="grey", inherit.aes=FALSE) +
  geom_text(data=fit.df, aes(x=gnm1, y=gnm2, label=row.names(fit.df)), size=5,
inherit.aes=FALSE) +
  coord_fixed()

# Shepard-diagram (stressplot)
stressplot(mds.best, dist.y)

## BOKSPLOT
x <- dato$Dato

```

```

plot(gnmds1 ~ factor(dato$Dyp))
summary(lm(gnmds1 ~ factor(dato$Dyp)))

plot(gnmds2 ~ factor(dato$Dyp))
summary(lm(gnmds2 ~ factor(dato$Dyp)))

#####
## dca ##
#####

#supplementær figur S. 1.

dca.Y<-decorana(y, i resc=4)
decorana(veg = y, i resc = 4)
head (dca.Y)

dca.Y1<-scores(dca.Y, display="sites", choices=1:4, origin=FALSE)[, 1]
dca.Y2<-scores(dca.Y, display="sites", choices=1:4, origin=FALSE)[, 2]

ggplot(dato, aes (x= dca.Y1, y= dca.Y2, colour=Dyp, label=Dato, group=Dato)) +
  geom_point(size=2) +
  geom_text(aes(label=Dato), hjust=0, vjust=0) +
  geom_line()

plot(dca.Y1 ~ factor(dato$Dyp))
summary(lm(dca.Y1 ~ factor(dato$Dyp)))

plot(dca.Y2 ~ factor(dato$Dyp))
summary(lm(dca.Y2 ~ factor(dato$Dyp)))

#####
## Korrelasjons-test ##
#####

cor.test(gnmds1, dca.Y1, method = "k") #evt "s" = Spearman
cor.test(gnmds2, dca.Y2, method = "k")

#####
## Shannon, Rikhet og Jevnhet ##
#####

#figur 3. 21

#2m

sh2 <- sekvens.table("sh_2m.txt", header = T, row.names = 1)
sh2_o <- sh2[, -(which(colSums(sh2) == 0) )]
# fjern OTUs ingen verdi

### SHANNON ###
H2<-diversity(sh2_o, index="shannon", base = exp(1))
write.table(H2, file="shannon_index_2m.txt", sep="\t")

#### Rikhet ####
S2<-specnumber(sh2_o)
write.table(S2, file="rikhet_index_2m.txt", sep="\t")

### Jevnhet/Pielou ###
J2<-H2/1og(S2)
write.table(J2, file="pielou_index_2m.txt", sep="\t")

#5m

sh5 <- sekvens.table("sh_5m.txt", header = T, row.names = 1)
sh5_o <- sh5[, -(which(colSums(sh5) == 0) )]

### SHANNON ###
H<-diversity(sh5_o, index="shannon", base = exp(1))
write.table(H, file="shannon_index_5m.txt", sep="\t")

#### Rikhet ####
S<-specnumber(sh5_o)
write.table(S, file="rikhet_index_5m.txt", sep="\t")

### Jevnhet/Pielou ###
J<-H/1og(S)
write.table(J, file="pielou_index_5m.txt", sep="\t")

### SAMMENLIGNER DE TO DYP ###

indexes <- sekvens.table(file="indexes.txt", header=TRUE)

#T tests for Indexes
t.test(indexes$R_2m, indexes$R_5m)
t.test(indexes$S_2m, indexes$S_5m)

```



```

t.test(indexes$J_2m , indexes$J_5m)

mi <- melt(indexes)
write.table(mi, "miindexestest.txt", col.names= TRUE)

# Fikser tabell i Excel. Med overskriftene:
# var month index depth value

min <- sekvens.tabl e(file="mi indexes2.txt", header=TRUE)
head(min)

min$depth <- factor(min$depth, levels=c("2m", "5m"))
min$index <- factor(min$index, levels=c("R", "S", "J"))
min$month <- factor(min$month, levels=c("Feb. 24. 2016", "Apr. 11. 2016", "Jun. 22. 2016",
"Aug. 18. 2016", "Sep. 12. 2016", "Nov. 25. 2016", "Feb. 02. 2017", "Feb. 20. 2017", "Mar. 06. 2017",
"Apr. 04. 2017", "Mai. 15. 2017", "Jun. 21. 2017", "Aug. 14. 2017", "Sep. 18. 2017", "Okt. 18. 2017",
"Nov. 16. 2017", "Des. 15. 2017"))

summaryBy(value ~ depth + index , min, FUN=c(median))

bwpalette<- c("white", "white") # Set up a black/white palette

d1 <- ggplot(min, aes(x = depth, y = value, fill = depth)) + scale_fill_manual(values =
bwpalette)+
  geom_boxplot() +
  facet_wrap(~ index, scales = "free")+
  theme_bw()+
  theme(axis.text.x = element_text(color="black", size=8), axis.text.y =
element_text(color="black", size=8), legend.title = element_text(colour="white",
size=14, face="bold"), legend.text = element_text(colour="black", size=8)) +ylab("")+
  ylab("") + xlab("")

d2 <- ggplot(min, aes(x = month, y = value)) +
  geom_point(aes(shape = depth), size=1)+scale_shape(solid = FALSE) +
  geom_smooth(aes( group = depth, method = "gam", linetype = depth), color = "black", size
=0.5, se = FALSE) +
  facet_wrap(~index, scales = "free") +
  theme_bw() +
  theme(axis.text.x = element_text(color="black", size=8, angle=90), axis.text.y =
element_text(color="black", size=8), legend.title = element_text(colour="white",
size=14, face="bold"), legend.text = element_text(colour="black", size=8)) +ylab("") +xlab("")

windows()
ggarrange(d1, d2)

#####
## indekser vs. miljøvariabel ##
#####

#Supplementær figur S.6

shCTD <- sekvens.tabl e("shannonCTD.txt", header = TRUE, row.names = 1, dec=",")
str(shCTD)

### Shannon og temp###
st <- ggplot(shCTD, aes(x=temp, y=S)) +
  geom_point() +
  stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, se=FALSE)+
  xlab("Temperatur") +
  ylab("Shannon indeks")

###Rikhet og temp###
sr <- ggplot (shCTD, aes(x=temp, y=R)) +
  geom_point() +
  stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, se=FALSE)+
  xlab("Temperatur") +
  ylab("Rikhet")

###Jevnhet og temp###
sj <- ggplot (shCTD, aes(x=temp, y=J)) +
  geom_point() +
  stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, se=FALSE)+
  xlab("Temperatur") +
  ylab("Fordeling")

windows()
ggarrange (sr, st, sj, ncol = 3)

#ANOVA TEST, KURVE MOT LINEÆR og null###
summary(m1 <- lm(S ~ temp , data=na.omit(shCTD)))
summary(m2 <- lm(S ~ temp + I(temp^2), data=na.omit(shCTD)))
summary (m3<- lm(S ~ 1, data = na.omit(shCTD)))

anova(m1, m2)
anova (m2, m3)
AIC (m1, m2)

```

```

AlC(m2,m3)
### Dette ble gjort for R og J også ###
#####
## Vind ##
#####

#Figur 3.8

library(scales)

vind <- read.table("vind.txt")

vind$Date = as.Date(vind$V1, "%Y-%m-%d")

gt <- ggplot(vind, aes(x = Date, y = V3)) +
  geom_point() +
  scale_x_date(limits = as.Date(c("2016-02-01", "2017-31-12"))) +
  geom_hline(yintercept = 3) +
  theme(axis.text.x = element_text(colour="darkgrey")) +
  xlab("Date") +
  ylab("Vind (m/s)")

#####
## Saltholdighetsprofil ##
#####

#Supplementær figur S.3

ctd1 <- read.ctd("DK1_24022016_oks.cnv")
pyncocline <- ctdTrim(ctd1, "range",
  parameters=list(item="pressure", from=0, to=10))
plotProfile(pyncocline, which="density+N2", yaxs = "i")

#####
## Øvre blandingslag ##
#####

#Figur 3.7

uml <- read.table("mld.txt", header = TRUE)

uml$Date = as.Date(uml$Date, "%Y-%m-%d")

gt <- ggplot(uml, aes(x = Date, y = MLD)) +
  geom_point() +
  geom_line() +
  scale_x_date(limits = as.Date(c("2016-02-01", "2017-31-12"))) +
  geom_hline(yintercept = 5) +
  theme(axis.text.x = element_text(colour="darkgrey")) +
  xlab("Date") +
  ylab("Dyp (m)") +
  scale_y_reverse()

#####
## ANOSIM ##
#####

dato <- read.table("ordinasjon_input_dato_dyp.txt", row.name = 3, header = TRUE)
data <- read.table("ordinasjon_input.txt", row.name = 1, header = TRUE)
maned <- read.table("ordinasjon_input_dato_dyp_maned.txt", row.names = 4, header = TRUE)

y <- as.data.frame(t(data))
x <- y[, -c(1:2)]

p <- anosim(x, grouping = dato$Dyp, permutations = 999, distance = "bray") #,
strata=dato$Dyp)
plot(p)
summary(p)

```