

UNIVERSITÀ DEGLI STUDI DI PISA

**Corso di Laurea Magistrale in
Conservazione ed Evoluzione**



Prova finale

**Confronto tra approccio tradizionale e next-generation
sequencing per la caratterizzazione delle comunità di ciliati**

Candidato

Alessia Rossi

Relatori

Prof. Giulio Petroni

Dott. Vittorio Boscaro

Anno Accademico 2014/2015

INDICE

Introduzione

1. I ciliati	pag. 3
1.1 Ecologia dei ciliati	pag. 6
2. Caratterizzazione delle comunità di ciliati	pag. 8
3. Scopo del lavoro	pag. 11

Materiali e Metodi

1. Area di studio	pag. 12
2. Raccolta e trattamento dei campioni	pag. 14
3. Caratterizzazione delle comunità di ciliati	pag. 14
3.1 Primo approccio: identificazione morfologica e sequenziamento Sanger	pag. 15
3.1.1 Compilazione delle check-list	pag. 15
3.1.2 Allestimento colture e identificazione molecolare	pag. 17
3.1.3 Analisi filogenetiche	pag. 21
3.1.4 Analisi statistiche	pag. 23
3.2 Secondo approccio: Next-Generation Sequencing	pag. 25
3.2.1 Estrazione del DNA e preparazione delle <i>libraries</i> Illumina	pag. 25
3.2.2 Analisi delle sequenze, ricerca delle OTU e affiliazione tassonomica	pag. 29
3.2.3 Analisi <i>pipeline</i> 2 bis	pag. 31
3.3 Confronto dei due approcci	pag. 31

Risultati e Discussione

1. Primo approccio: identificazione morfologica e sequenziamento Sanger	pag. 32
1.1 Numero di morfotipi	pag. 32
1.2 Composizione tassonomica	pag. 34
1.3 Dati molecolari	pag. 39
1.4 Analisi filogenetiche	pag. 41
1.5 Analisi statistiche e biogeografia	pag. 46

2. Secondo approccio: Next-Generation Sequencing	pag. 50
2.1 Numero di <i>contig</i> e di OTU	pag. 50
2.2 Analisi <i>pipeline 2 bis</i>	pag. 55
2.2.1 Composizione tassonomica	pag. 58
3. Confronto dei due approcci	pag. 61
3.1 Numero di OTU	pag. 61
3.2 Composizione tassonomica	pag. 65
3.3 Analisi statistiche	pag. 69
<i>Conclusioni</i>	pag. 73
<i>Bibliografia</i>	pag. 74
<i>Appendice</i>	pag. 84

Introduzione

1. I ciliati

I ciliati sono organismi eucariotici unicellulari o coloniali appartenenti all'ampio ed eterogeneo gruppo dei protisti, di dimensioni variabili tra circa 10 μm e 4,5 mm.

Da sempre riconosciuti come un gruppo a sé stante grazie alla loro particolare morfologia, oggi sono considerati un phylum, Ciliophora (Doflein 1901), compreso nel clade degli Alveolata insieme ai phyla Dinoflagellata e Apicomplexa (Adl et al., 2012).

I ciliati possiedono tre apomorfie distintive:

- presenza di ciglia e di una complessa infraciliatura nella regione corticale
- dimorfismo nucleare
- fenomeno della coniugazione

Le ciglia sono morfologicamente identiche al flagello eucariotico, ma più corte rispetto alla dimensione della cellula e presenti in elevato numero sulla sua superficie. Derivano dai cinetosomi, corpuscoli basali con annessi tre distinti associati fibrillari la cui disposizione e struttura sono molto variabili e rappresentano un criterio utile per la classificazione dei ciliati ad elevato rango tassonomico.

La ciliatura è suddivisa in una parte somatica, la cui funzione è legata essenzialmente al movimento e all'adesione al substrato, ed una orale, il cui ruolo è quello di convogliare il cibo verso il citostoma (sede della fagocitosi) e permetterne l'ingestione.

Tutti i ciliati presentano le ciglia almeno in una fase del ciclo vitale. Gli organismi appartenenti alla sottoclasse dei Suctoria (classe Phyllopharyngea) in passato sono stati considerati da alcuni autori organismi non ciliati poiché possiedono le ciglia solo durante la loro fase dispersiva natante.

Tutti i ciliati sono eterocarioti, hanno cioè due diversi tipi di nucleo: uno o più nuclei somatici, i macronuclei (MAC), e uno o più nuclei germinativi, i micronuclei (MIC). I MAC, solitamente presenti in una o due copie, contengono solo una porzione del genoma, in particolare i geni che regolano la vita vegetativa della cellula. Questa parte di genoma è però

replicata più volte ed i MAC vengono definiti ampliploidi. La forma del macronucleo è molto variabile e viene utilizzata come carattere diagnostico per identificare alcune specie.

I MIC, presenti solitamente in più copie rispetto ai MAC, sono nuclei diploidi con funzione sessuale coinvolti nei processi di meiosi e coniugazione. La forma è quasi sempre sferica o rotonda, e sono molto più piccoli dei MAC.

Il terzo carattere distintivo dei ciliati è il fenomeno della coniugazione, un tipo di ricombinazione sessuale che non implica però la riproduzione (definita come incremento del numero di organismi). Durante tale processo, due cellule sessualmente compatibili, dette coniuganti, si fondono temporaneamente a livello del citostoma per scambiare MIC gametici (aploidi) e rimescolare così il patrimonio genetico.

Il phylum Ciliophora è riconosciuto come monofiletico anche dagli studi filogenetici molecolari basati sulle sequenze del gene codificante per l'rRNA 18S (Lynn, 2008).

Lo schema di classificazione utilizzato in questo lavoro segue la monografia di Denis Lynn del 2008 ("The ciliated Protozoa").

Il phylum Ciliophora è suddiviso in due subphyla, Postciliodesmatophora e Intramacronucleata. Postciliodesmatophora comprende solo due classi, Karyolicea e Heterotrichea, accomunate dalla presenza di un particolare annesso ciliare, il postciliodesma. Tutti i ciliati compresi all'interno del subphylum Intramacronucleata sono invece caratterizzati dall'utilizzo di microtubuli intramacronucleari durante il processo di divisione del macronucleo. Questo subphylum comprende 9 classi: Spirotrichea, Armophorea, Litostomatea, Phyllopharyngea, Nassophorea, Colpodea, Plagiopylea, Prostomatea e Oligohymenophorea.

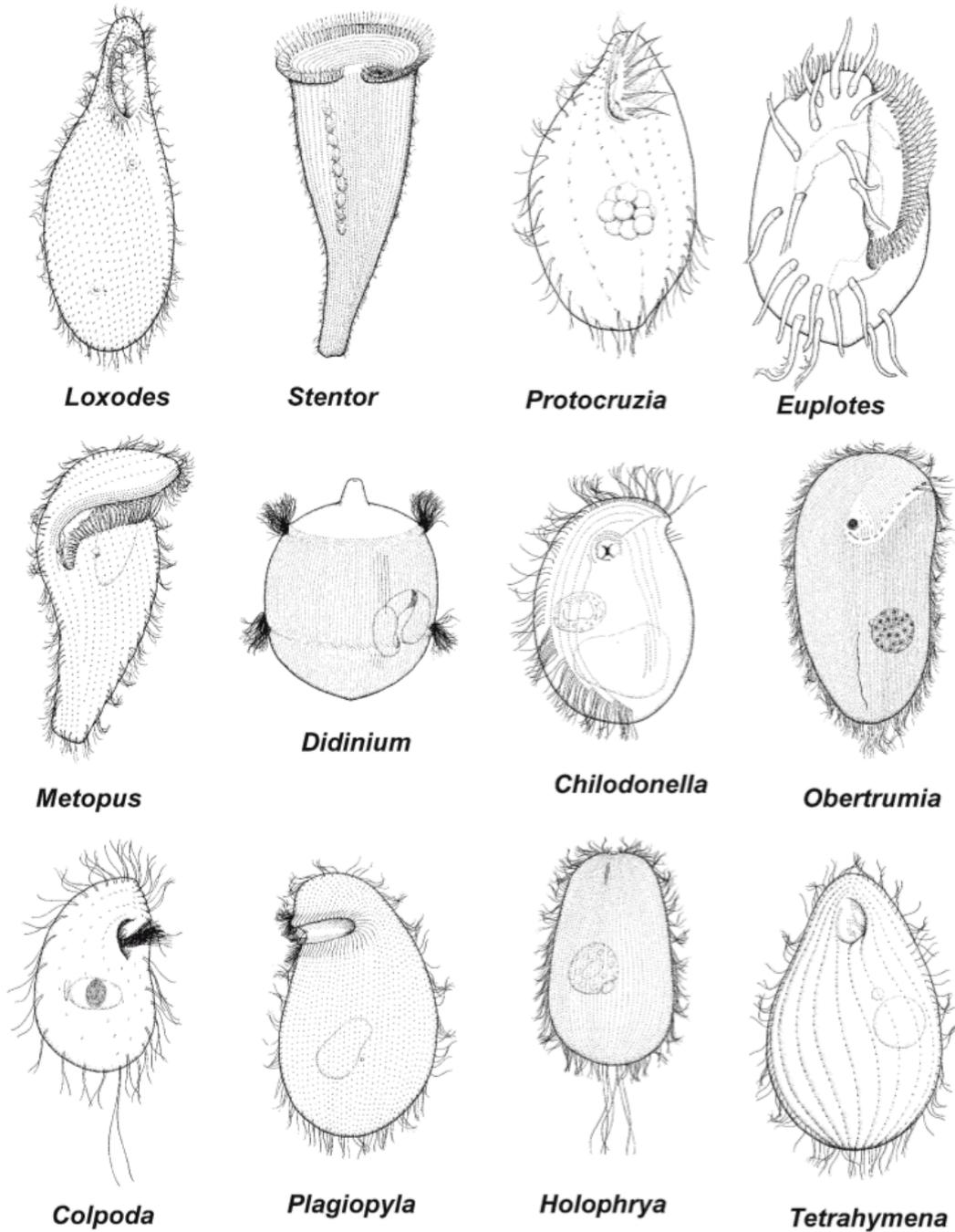


Figura 1. Illustrazioni di generi rappresentativi di ognuna delle 11 classi del Phylum Ciliophora: *Loxodes* – Classe Karyorelictea; *Stentor* – Classe Heterotrichea; *Protocruzia*, *Euplotes* – Classe Spirotrichea; *Metopus* – Classe Armophorea; *Didinium* – Classe Litostomatea; *Chilodonella* – Classe Phyllopharingea; *Obertruria* – Classe Nassophorea; *Colpoda* – Classe Colpodea; *Plagiopyla* – Class Plagiopylea; *Holophrya* – Classe Prostomatea; *Tetrahymena* – Classe Oligohymenophorea (da Lynn, 2008).

1.1 Ecologia dei ciliati

I protisti ciliati occupano una grande varietà di habitat e sono distribuiti dall'equatore fino ai poli. Si conoscono circa 4500 specie di ciliati a vita libera, ma la stima del numero totale di specie esistenti è molto maggiore (Foissner et al., 2007).

Si possono trovare negli oceani, dalle profondità marine (Hausmann et al., 2002) fino ai ghiacci antartici (Song & Wilbert, 2000), negli ambienti di acqua dolce come laghi (Esteban et al., 2000) e fiumi (Foissner, 1997b), nelle acque salmastre fino alle lagune ipersaline (Yasindi et al., 2002). I ciliati sono presenti anche nell'ambiente terrestre, principalmente nel suolo e nel muschio (Ryan et al., 1989; Foissner, 1998). La maggior parte dei ciliati sono organismi a vita libera, ma esistono numerose specie endo- o ectosimbionti di metazoi, come i ciliati commensali del rumine (Foissner, 1999) e le specie che colonizzano le appendici dei crostacei o le branchie dei bivalvi.

I ciliati sono organismi eterotrofi che si nutrono di una vasta gamma di microrganismi; alcune specie presentano inoltre al loro interno endosimbionti algali da cui traggono beneficio grazie ai prodotti della fotosintesi (mixotrofia).

Le "prede" dei ciliati possono essere incredibilmente varie, da virus a batteri, ad alghe unicellulari fino ad altri ciliati e piccoli metazoi.

I ciliati giocano un ruolo cruciale nelle catene trofiche degli ambienti acquatici; costituiscono un importante anello nel trasferimento di energia dai batteri ai livelli trofici superiori poiché sono uno dei maggiori gruppi di organismi batterivori e a loro volta sono preda di altri protozoi, invertebrati (es. piccoli crostacei), ma anche larve di pesci (Finlay & Esteban, 1998; Fukami et al., 1999). I ciliati batterivori sono ritenuti particolarmente importanti nella regolazione della biomassa batterica (Foissner, 1988); ad esempio sono molto utili nel mantenere la "qualità" dei reflui degli impianti di depurazione in quanto possono ridurre la densità batterica fino ad un decimo di quella iniziale (Madoni, 2003).

I ciliati rispondono molto velocemente ai cambiamenti ambientali grazie al loro elevato tasso di divisione, all'elevata sensibilità cellulare e alla posizione che occupano all'interno della catena trofica (Foissner 1997a). Per questo sono stati utilizzati anche come bioindicatori dello stato di qualità delle acque (Madoni & Zangrossi 2005; Jiang et al., 2007).

La struttura delle comunità di ciliati e la loro relativa biomassa sono determinate largamente dallo stato trofico del bacino idrico nel quale si trovano; i sistemi acquatici oligotrofici sono di solito caratterizzati da una ridotta diversità in numero di specie e da una minore biomassa rispetto ai sistemi ipereutrofici (Beaver & Crisman 1989); inoltre è stato osservato come l'incremento della concentrazione dei nutrienti favorisca maggiormente lo sviluppo dei ciliati batterivori di piccole dimensioni (<35 μm) a discapito dei taxa più grandi (Nakano et al. 1998).

Negli ultimi decenni, molti studi sono stati dedicati alla ricerca dei fattori che influenzano la biomassa e la struttura delle comunità di ciliati; oltre alla disponibilità di cibo legata allo stato trofico di un bacino, anche la temperatura (Levinsen et al., 2000) e la concentrazione di ossigeno disciolto (Guhl et al., 1994) sembrano esercitare una certa influenza sulla distribuzione dei ciliati nell'ambiente.

La capacità di percepire le alterazioni ambientali e di rispondere in modo rapido e appropriato è di vitale importanza per ogni organismo e in particolare per quelli unicellulari (Corliss & Esser 1974). Numerose specie di ciliati hanno sviluppato all'interno del loro ciclo vitale un particolare stadio di quiescenza definito "ciste". Le cisti rappresentano uno stadio di inattività che nella maggior parte dei casi ha funzione di protezione e dispersione; durante l'incistamento la cellula subisce drastici cambiamenti morfologici e fisiologici come la diminuzione del volume cellulare, la riorganizzazione del comparto citoplasmatico, la formazione di una parete parzialmente permeabile intorno alla cellula composta da tre o più strati, e cambiamenti a livello dei nuclei e degli organelli (Gutierrez et al., 1998; Lynn, 2008). Molti ciliati possono sopravvivere alle condizioni ambientali avverse, come disseccamento, mancanza di cibo, abbassamento della temperatura o variazione delle specifiche richieste di ossigeno proprio grazie alla formazione di cisti di resistenza; al ritorno delle condizioni ambientali idonee tali organismi potranno poi ricostituirsi (Foissner et al. 2007). Tale processo gioca un importantissimo ruolo ecologico nella distribuzione dei ciliati nell'ambiente e può aver contribuito enormemente alla loro diffusione.

2. Caratterizzazione delle comunità di ciliati

Le comunità di ciliati sono da sempre studiate utilizzando approcci morfologici ai quali, negli ultimi anni, si sono affiancati approcci combinati morfologici-molecolari o esclusivamente molecolari.

Rispetto a numerosi altri gruppi di protisti, i ciliati sono stati molto studiati con metodi morfologici poiché la maggior parte di essi sono relativamente grandi, morfologicamente diversi e quindi ben distinguibili fra di loro. Di conseguenza, esiste una cospicua letteratura sui caratteri che distinguono le diverse morfospecie ed in particolare su quali di essi siano più utili nell'identificazione. L'osservazione in vivo permette di individuare numerosi caratteri diagnostici come forma e dimensione della cellula, posizione e morfologia generale del citostoma, distribuzione delle ciglia, presenza di pigmenti e/o endosimbionti algali, e modalità di movimento. Inoltre, quando si studiano campioni ambientali è preferibile effettuare su di essi un'osservazione prolungata nel tempo combinata con livelli diversi di arricchimento del campione stesso, in modo da permettere a quella parte di biodiversità "nascosta" (ciliati incistati e specie rappresentate da pochi individui) di svilupparsi e rendersi così visibile. Fenchel e collaboratori (1997), oltre che arricchire il campione ambientale con diversi tipi di nutrimento, hanno sottoposto aliquote dello stesso campione a tutta una serie di trattamenti mirati a modificare le condizioni abiotiche iniziali (modifiche della temperatura, della concentrazione di ossigeno disciolto, della salinità etc.). Lo scopo degli autori era quello di creare microhabitat con differenti condizioni ecologiche favorevoli per lo sviluppo dei ciliati rari e/o incistati.

La sola osservazione dei campioni in vivo non permette però di identificare la maggior parte dei taxa a livello specifico; per tale scopo è necessario sottoporre le singole cellule a particolari tecniche di colorazione che mettano in evidenza tutta una serie di strutture (nuclei, ciglia, strutture infraciliari) che hanno un alto potere di discriminazione. Le tecniche più utilizzate sono la reazione di Feulgen, l'impregnazione argentea e il protargolo.

Sia l'osservazione in vivo che quella fatta su cellule fissate richiedono molto tempo e la presenza di personale specializzato.

Più recentemente, grazie all'espansione delle tecniche molecolari, sono stati pubblicati numerosi lavori di caratterizzazione delle comunità microbiche che utilizzano solo tali metodologie. Le tecniche molecolari più utilizzate applicate alle comunità di ciliati sono il clonaggio (Lara et al., 2007) e il T-RFLP (Dopheide et al., 2009).

Tali approcci si basano sull'estrazione del DNA totale dai campioni ambientali (metagenoma) e sulla successiva amplificazione di un marcatore molecolare dai microrganismi di interesse; il marcatore più utilizzato in questi tipi di studi è il gene codificante per l'rRNA 18S.

La struttura di tale gene, caratterizzata da regioni altamente conservate e da regioni ipervariabili, lo rende idoneo per valutare la diversità genetica all'interno di comunità microbiche e per stabilire le relazioni filogenetiche tra i diversi organismi. Le regioni conservate hanno permesso inoltre l'individuazione di primers universali che sono utilizzati nelle reazioni di amplificazione del DNA (Backer et al., 2003).

Trattandosi di uno dei marcatori filogenetici più utilizzati, è disponibile *on line* un enorme database di sequenze riferibili a organismi ciliati, che permette di identificare le sequenze ottenute da lavori di caratterizzazione esclusivamente molecolare, molto spesso a livello di genere e specie.

Negli ultimi anni la tecnologia NGS (Next-Generation Sequencing) ha preso il sopravvento sugli studi di caratterizzazione delle comunità microbiche; tale tecnologia ha il grande vantaggio di poter processare contemporaneamente numerosi campioni, ma soprattutto ha una profondità di sequenziamento molto elevata, definita "*high throughput*", che è in grado di individuare le sequenze anche dei cosiddetti taxa rari, rappresentati da pochi individui. Le più recenti piattaforme possono infatti sequenziare in parallelo fino a milioni di frammenti di DNA. Grazie a questa nuova tecnologia sono state eliminate le fasi di clonaggio precedenti al sequenziamento, sostituite da una semplice ligazione di adattatori ad ogni amplificato. Dall'altra parte però troviamo anche per queste tecnologie all'avanguardia degli svantaggi e delle limitazioni. Ad esempio durante la costruzione delle *libraries*, sebbene sia stata superata la fase di clonaggio, sono richiesti numerosi passaggi durante i quali è possibile l'introduzione di errori; anche il sequenziamento può produrre artefatti. Con tale metodologia inoltre non è possibile ottenere lunghi frammenti di DNA come l'intero gene codificante l'rRNA 18S (circa 1700 pb), e le sequenze ottenute sono quindi relativamente corte

(attualmente, fino a 300 pb, estendibili fino a circa 550 pb attraverso il sequenziamento “*paired-end*”).

Le tecnologie NGS hanno richiesto il contemporaneo sviluppo della bioinformatica, una disciplina che integra competenze informatiche, statistiche e analitiche nell’ambito della biologia, necessaria per la gestione e l’analisi dell’enorme quantità di dati prodotti da tali tecnologie.

L’utilizzo di un approccio esclusivamente molecolare nella caratterizzazione di campioni ambientali, e in particolare le tecnologie NGS, ha individuato la presenza di un’enorme diversità microbica precedentemente non rilevata dai metodi morfologici, anche a livello dei protisti ciliati (Epstein & Lopez-Garcia, 2008; Bachy et al., 2013; Stoeck et al., 2014).

Sebbene parte di tale biodiversità individuata solo dai metodi molecolari esista realmente (taxa rari, ciliati incistati, organismi di piccole dimensioni che sfuggono all’osservazione morfologica), numerosi autori suggeriscono che molte delle sequenze ottenute siano in realtà artefatti che portano quindi ad una sovrastima della reale diversità presente nel campione (Behnke et al., 2011; Stoeck et al., 2014).

Pochi sono gli studi che utilizzano entrambi gli approcci per la caratterizzazione delle comunità di ciliati (Dopheide et al., 2009; Bachy et al., 2013; Stoeck et al., 2014). In particolare finora non è stato pubblicato nessun lavoro che utilizza insieme ad un approccio morfologico la piattaforma NGS Illumina per la caratterizzazione delle comunità di ciliati; tale tecnologia, benché possa sequenziare frammenti lunghi al massimo 550 bp, è fra le tecnologie NGS quella che produce meno errori a livello di sequenziamento.

3. Scopo del lavoro

Nel seguente lavoro di tesi, comunità di ciliati provenienti da numerose tipologie di ambienti dulciacquicoli della provincia di Pistoia sono state indagate con due diversi tipi di approcci: uno morfologico-molecolare di tipo tradizionale ed un secondo che sfrutta la tecnologia del next-generation sequencing, in particolare la piattaforma Illumina Miseq.

Le conclusioni raggiunte con i due approcci sono state confrontate per valutare vantaggi, svantaggi e differenze delle due metodologie. Inoltre con i dati ottenuti da ciascuna caratterizzazione è stata indagata attraverso tecniche multivariate la presenza di *pattern* caratteristici di biodiversità legati alle variabili ambientali.

Materiali e Metodi

1. Area di studio

All'interno della provincia di Pistoia sono state scelte 10 aree dulciacquicole rappresentative di diverse tipologie di habitat, sia di tipo planiziale (<200 m.s.l.m.) che montano (>1200 m s.l.m.) (Figura 2). Le diverse aree non sono in contatto idraulico fra loro o, se in contatto, sono state considerate distinte poiché di tipo lentico o lotico. I campionamenti sono stati effettuati in 7 aree naturali, di cui 4 lentiche e 3 lotiche, e in 3 aree artificiali lentiche. Per ogni area sono state individuate da 1 a 4 stazioni di campionamento ad una profondità di 10-50 cm (Tabella 1). Lo studio è stato ripetuto in tre diversi periodi: autunno 2012 (α), primavera (β) e autunno 2013 (γ) per un totale di 24, 27 e 24 stazioni campionate rispettivamente in ogni stagione; la differenza nel numero di siti campionati dipende dall'area Lago Nero, che si è potuta raggiungere solo durante il campionamento β .

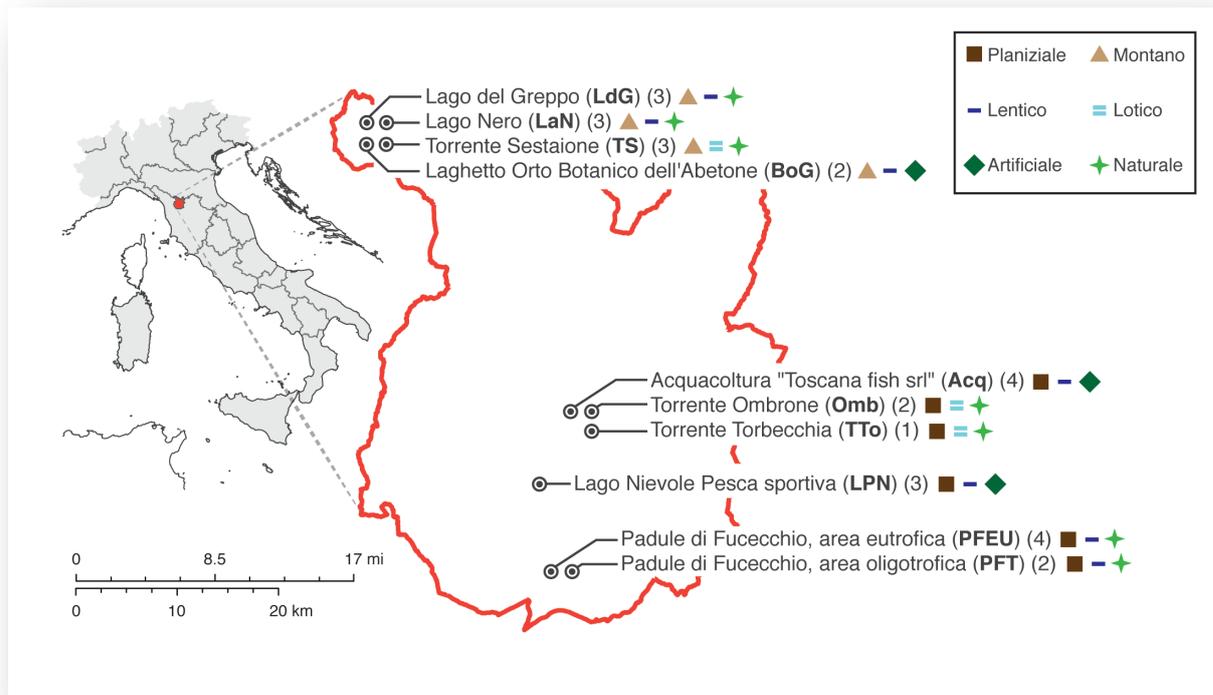


Figura 2. Cartina della provincia di Pistoia con indicate le 10 aree oggetto di studio; i numeri tra parentesi si riferiscono al numero di stazioni campionate in ogni area in ciascuna stagione.

AREA	STAZIONI DI CAMPIONAMENTO				INFORMAZIONI
Padule di Fucecchio zona eutrofica (PFEU)	naturale	lentico	planiziale	4	apporto idrico deriva da acque di risalita di falda e acque piovane
Padule di Fucecchio zona oligotrofica (PFT)	naturale	lentico	planiziale	2	apporto idrico deriva da numerosi corsi d'acqua
Lago del Greppo (LdG)	naturale	lentico	montano	3	laghetto di origine glaciale nella riserva Naturale di Campolino
Lago Nero (LaN)	naturale	lentico	montano	3	laghetto di origine glaciale
Laghetto dell'Orto Botanico (BoG)	artificiale	lentico	montano	2	piccolo specchio d'acqua all'interno dell'Orto Botanico dell'Abetone
Torrente Sestaione (TS)	naturale	lotico	montano	3	la sorgente origina dalle infiltrazioni del lago Nero nel terreno
Torrente Ombrone (Omb)	naturale	lotico	planiziale	2	affluente di destra del fiume Arno
Torrente Torbecchia (TTo)	naturale	lotico	planiziale	1	va ad alimentare le vasche di acquacoltura Toscana Fish srl
Lago Nievole Pesca sportiva (LPN)	artificiale	lentico	planiziale	3	acque provenienti dal torrente Nievole
Acquacoltura Toscana Fish srl (Acq)	artificiale	lentico	planiziale	4	vasche di acquacoltura, acqua proveniente dal torrente Torbecchia

Tabella 1. Le 10 aree oggetto di studio per le quali è indicato il numero di stazioni campionate e alcune caratteristiche ambientali.

2. Raccolta e trattamento dei campioni

In ogni stazione sono stati prelevati dallo stesso punto due campioni (5-10 ml di sedimento umido e 35-40 ml d'acqua) con una provetta Falcon da 50 ml. Un campione è stato mantenuto in vivo, mentre il secondo è stato fissato in loco secondo quanto segue: dopo essere stato invertito più volte per assicurare il mescolamento della colonna d'acqua con il sedimento, è stato diviso in due aliquote; 15 ml di campione sono stati fissati in etanolo al 70%, e una seconda aliquota da 35 ml è stata fissata in soluzione di Bouin al 10%. Questa seconda aliquota, non oggetto del mio lavoro di tesi, è stata fissata per eventuali successive analisi morfologiche.

Composizione della soluzione di Bouin (Skibbe, 1994):

acido picrico: 15 ml

formalina: 5 ml

acido acetico glaciale: 1 ml

In ogni stazione di campionamento sono stati misurati, mediante la sonda Lovibond® SensoDirect 150, parametri abiotici quali temperatura dell'acqua, pH e concentrazione di ossigeno disciolto. Sono state inoltre rilevate per ogni stazione l'altitudine e le coordinate GPS.

3. Caratterizzazione delle comunità di ciliati

La caratterizzazione delle comunità di ciliati si è avvalsa di due differenti approcci; un primo approccio combinato morfologico-molecolare che ha previsto l'utilizzo dei campioni mantenuti in vivo e un secondo approccio esclusivamente molecolare in cui abbiamo utilizzato le aliquote fissate in etanolo al 70%.

3.1 Primo approccio: identificazione morfologica e sequenziamento Sanger

3.1.1 Compilazione delle check-list

Una volta in laboratorio, i campioni mantenuti in vivo sono stati divisi in tre aliquote; dopo aver invertito ogni provetta, 25 ml di campione (acqua e sedimento) sono stati divisi in due piastre Petri denominate aliquota I e II. Il restante campione (aliquota III), contenente la maggior parte del sedimento, è stato mantenuto nella provetta Falcon originale per facilitare la creazione di un ambiente anossico favorevole alle specie anaerobiche e/o microaerofile.

Mediante osservazioni in vivo allo stereomicroscopio (ingrandimento 10-40x), si è provveduto all'individuazione dei diversi morfotipi appartenenti al phylum Ciliophora presenti nelle aliquote I e II di ogni campione. L'identificazione dei morfotipi individuati è stata fatta per confronto, osservando cioè le caratteristiche morfologiche riconoscibili in vivo e confrontandole con le descrizioni dei diversi taxa riportate all'interno di specifici manuali (Curds, 1982; Foissner et al., 1999; Lee et al., 2000; Kreutz & Foissner, 2006; Lynn 2008).

Alcune caratteristiche utilizzate per l'identificazione sono:

- forma generale e dimensione della cellula
- posizione e morfologia del citostoma
- distribuzione delle ciglia
- colorazione della cellula
- movimento della cellula

Per i diversi morfotipi, con l'aiuto dove necessario di un microscopio istologico a contrasto interferenziale Nomarski Leitz (Weitzlar, Germany), è stata individuata la sottoclasse di appartenenza (o la classe, ove una sottoclasse non sia mai stata definita formalmente). Per alcuni è stato inoltre possibile identificare il genere e la specie.

Ogni morfotipo è stato assegnato a una di tre categorie di abbondanza:

- raro (se rappresentato da un minimo di 1 a un massimo di 3 cellule)
- comune (se rappresentato da un minimo di 4 a un massimo di 30 cellule)
- abbondante (se rappresentato da più di 30 cellule)

Dopo aver compilato per entrambe le aliquote di ogni campione la check-list dei taxa appartenenti al gruppo dei ciliati osservati a tempo zero (t_0), cioè il giorno stesso del

campionamento, la Petri I è stata differenziata attraverso l'aggiunta di una cariossido di riso bollita. L'arricchimento con il riso serve a sostenere la crescita batterica e di conseguenza quella dei ciliati batterivori e delle popolazioni da essi dipendenti.

Il contenuto delle due piastre Petri e della provetta originale è stato osservato più volte nel tempo, in modo da individuare la possibile presenza di quei taxa che in un primo momento potrebbero non essere stati osservati poiché presenti in concentrazioni non rilevabili o sottoforma di cisti. Per ogni campione si è provveduto quindi a ulteriori quattro osservazioni, riportando ogni volta nella corrispondente check-list i nuovi taxa trovati. Le osservazioni sono state fatte secondo il seguente schema cronologico (Figura 3):

Dopo 7 giorni dalla data di campionamento: osservazione delle Petri I e II, aggiornamento della check-list, aggiunta di una cariossido di riso bollita nella Petri II e nella provetta Falcon originale.

Dopo 14 giorni: osservazione delle Petri I e II e aggiornamento della check-list.

Dopo 21 giorni: osservazione dell'aliquota III (campione rimasto nella Falcon originale) e aggiornamento della check-list

Dopo 35 giorni: osservazione delle Petri I, II e della aliquota III e aggiornamento della check-list.

Nelle aliquote II e III il riso è stato aggiunto dopo una settimana rispetto all'aliquota I, per mantenere un apporto di nutrienti evitando però la scomparsa immediata dei taxa competitivamente svantaggiati in ambienti ricchi.

Durante i 35 giorni di osservazione, le tre aliquote relative a ogni campione sono state mantenute in una cella termostata a 19° C con un ciclo di luce/oscurità di 12 ore.

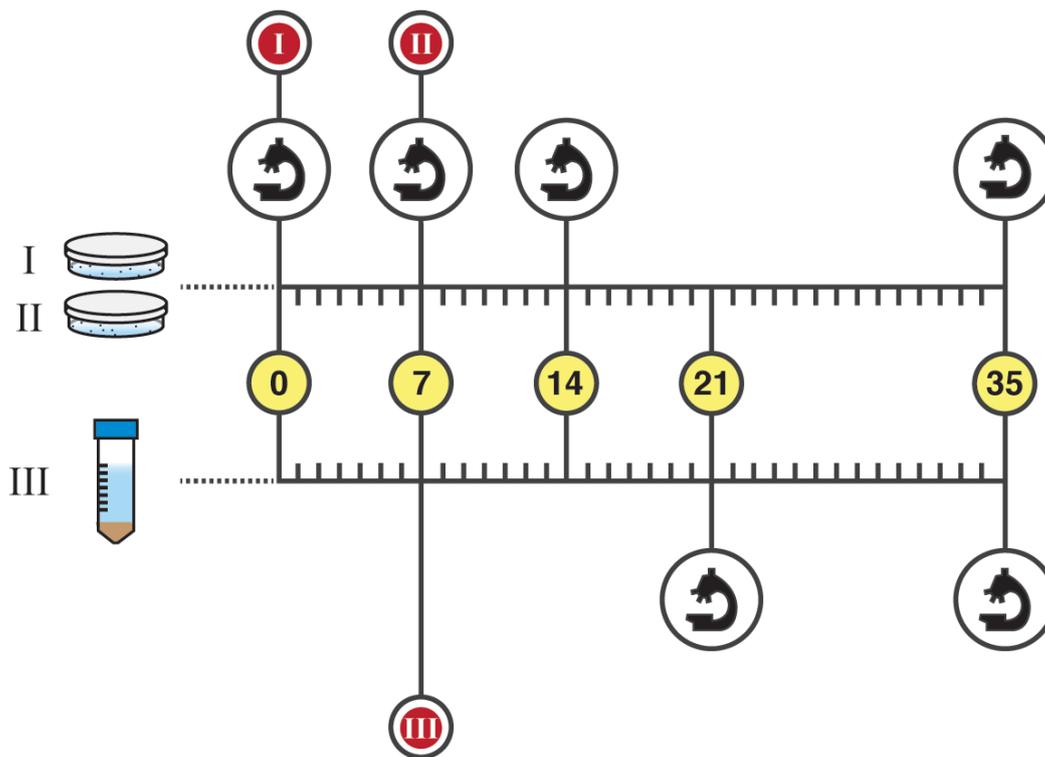


Figura 3. Linea del tempo in cui sono indicate le 5 sessioni di osservazione. I numeri corrispondono ai giorni trascorsi dalla data di campionamento. I pallini rossi indicano l'arricchimento con la cariossidi di riso delle aliquote indicate.

3.1.2 Allestimento colture e identificazione molecolare

Nel corso delle cinque settimane di osservazione, dai campioni sono state selezionate e isolate cellule appartenenti ad alcuni dei morfotipi individuati.

L'isolamento del ciliato dal campione ambientale avviene tramite una micropipetta di vetro ricavata da comuni pipette Pasteur tirate sulla fiamma di un becco Bunsen. Il diametro del capillare della micropipetta deve essere di poco superiore alle dimensioni del ciliato in modo da catturare singole cellule con il minor quantitativo possibile di mezzo circostante. In seguito le cellule isolate sono sottoposte a passaggi di lavaggio in acqua distillata sterile utilizzando vetrini a tre depressioni, in modo da rimuovere eventuali contaminazioni da parte di altri eucarioti.

L'isolamento dei diversi morfotipi ha due scopi principali:

- allestimento di colture
- fissazione di singole o poche cellule in etanolo.

Sono stati eseguiti tentativi di coltura su numerosi morfotipi di ciliati. Dopo aver selezionato e lavato una o più cellule del ciliato d'interesse, queste sono state poste in un terreno liquido costituito da acqua minerale arricchita con diversi tipi di nutrimento che dipendono dall'organismo considerato. Come fonte di cibo è stato utilizzato:

- l'alga verde unicellulare *Dunaliella tertiolecta* (Chlorophyceae, Volvocales)
- la diatomea *Phaeodactylum tricornutum* (Bacillariophyceae, Naviculales)
- mezzo di coltura Cerophyl (composizione riportata in appendice a pag. 84) inoculato con il batterio *Raoultella planticola* (Gammaproteobacteria, Enterobacteriales)
- cariossidi di riso, utilizzata per far aumentare i batteri naturalmente associati all'organismo in coltura (in questi casi, le cellule non venivano sottoposte al trattamento di lavaggio).

Ad ogni coltura è stato somministrato a diverse cadenze temporali, ogni 4-15 giorni, uno o più dei nutrimenti sopracitati a una concentrazione dipendente dalla densità della coltura stessa. Tutte le colture sono state mantenute in una cella termostata a 19° C con un ciclo di luce/oscurità di 12 ore.

L'allestimento delle colture ha come scopo il raggiungimento di colture stabili, che cioè possiedono un ritmo di crescita costante, tale da mettere a disposizione sufficiente materiale per eventuali esperimenti e studi più approfonditi. Nel nostro studio, le colture sono state utilizzate per ottenere dati molecolari che ci aiutassero nell'identificazione degli organismi d'interesse.

Dalle colture divenute stabili sono state prelevate dalle 50 alle 100 cellule; dopo numerosi passaggi di lavaggio in acqua distillata sterile, come precedentemente descritto, le cellule sono state fissate in etanolo al 70% in una provetta eppendorf da 1,5 ml.

L'estrazione del DNA genomico totale (tgDNA) è stata effettuata sulle cellule fissate utilizzando il kit commerciale Nucleo Spin® Plant II (Macherey-Nagel). Il protocollo impiegato è quello per l'estrazione del tgDNA da funghi, leggermente modificato. Il tgDNA risultante è stato eluito in acqua distillata sterile e mantenuto a una temperatura di -20°C.

L'allestimento di una coltura richiede molto tempo ed è possibile attualmente, e con i mezzi a nostra disposizione, solo per una minima parte dei ciliati conosciuti. Per ottenere un numero maggiore di dati molecolari, dove possibile, da 1 a 5 cellule appartenenti allo stesso morfotipo sono state isolate, lavate e fissate direttamente in etanolo al 70% in provette eppendorf da 0,2 ml. Una o poche cellule possono essere infatti utilizzate come materiale su

cui effettuare direttamente la PCR del gene d'interesse, evitando il passaggio intermedio di estrazione del tgDNA.

Il marker più utilizzato negli studi di filogenesi molecolare sui protisti è il gene codificante per l'rRNA 18S. Nel nostro studio la sequenza di tale gene è stata amplificata tramite PCR; le reazioni di amplificazione sono state effettuate con un Thermal Cycler modello C1000TM (BioRad, Hercules, California) utilizzando una miscela di reazione contenente acqua distillata sterile, buffer (exTaq buffer, TaKaRa®), una miscela equimolare dei quattro deossinucleotidi (TaKaRa®), primer forward e reverse e Taq polimerasi (exTaq, TaKaRa®), più un'aliquota della soluzione di DNA estratto (eventualmente diluita o concentrata).

Per quanto riguarda la PCR su singola o poche cellule, al posto dell'aliquota di DNA estratto, è stata utilizzata il materiale fissato in etanolo. La provetta eppendorf da 0,2 ml contenente da 1 a 5 cellule dello stesso morfotipo è stata centrifugata per 10 minuti ad una accelerazione di circa 5000 RCF in modo da fare aderire le cellule al fondo della provetta; quest'ultima è stata poi invertita per eliminare l'etanolo e posta in un essiccatore (SpeedvacTM) finché tutto l'etanolo residuo non è evaporato. A questo punto abbiamo proceduto con la PCR inserendo la miscela di reazione nella provetta.

I profili termici utilizzati, successivi alla fase di denaturazione iniziale di 3 minuti a 94°C, sono diversi per le due tipologie di PCR.

Per la PCR in cui si utilizza un'aliquota di DNA estratto, il seguente profilo è stato ripetuto per 30 cicli:

- denaturazione a 94°C per 30 secondi
- annealing a 50°C per 30 secondi
- estensione a 72°C per 120 secondi

Per la PCR eseguita su singole o poche cellule, è stato utilizzato il seguente profilo termico ripetuto per 40 cicli:

- denaturazione a 94°C per 30 secondi
- annealing a 55°C per 30 secondi
- estensione a 72°C per 120 secondi

Per entrambe le PCR c'è poi una fase di estensione finale a 72°C della durata di 6 minuti.

Per l'amplificazione del gene codificante per l'rRNA 18S sono stati utilizzati il primer forward 18S F9 Euk (Medlin et al., 1988) e il primer reverse 18S R1513 Hypo (Petroni et al., 2002).

In alcuni casi è stato necessario effettuare sugli amplificati poco concentrati una PCR semi-nested che utilizza lo stesso primer forward 18S F9 Euk e il primer reverse 18S R1052 (Rosati et al., 2004) interno alla sequenza del gene d'interesse. L'amplificato di partenza è stato purificato e diluito prima di questa seconda PCR. Il profilo termico che si ripete per 30 cicli è il seguente:

- denaturazione a 94°C per 30 secondi
- annealing a 55°C per 30 secondi
- estensione a 72°C per 90 secondi

Tutte e tre le tipologie di PCR prevedevano la presenza di un controllo negativo, cioè di un campione privo di DNA, necessario per valutare la presenza di contaminazione dei reagenti e della strumentazione utilizzata.

Dopo la PCR, per verificare l'avvenuta amplificazione e stimare sommariamente la qualità e la concentrazione dell'amplificato, si è proceduto con la corsa elettroforetica su gel di agarosio all'1%, densità ottimale per differenziare frammenti di oltre 1000 pb come il nostro marcatore. L'amplificato addizionato con un buffer di carico è stato fatto correre per circa 40 minuti a 150 V in una soluzione di TBE all'1% (Tris, Acido Borico, EDTA) assieme ad un marker con pesi molecolari noti (GeneRuler DNA Ladder Mix). Alla fine della corsa, il gel è stato immerso in una soluzione di buffer TBE e bromuro di etidio (0,5 µg/ml) per circa 30 minuti, per poi essere visualizzato e fotografato attraverso un transilluminatore.

Qualora le bande siano nette, singole e alla giusta altezza (circa 1700 pb per le PCR da singole o poche cellule e da tgDNA e circa 1300 pb per le PRC semi-nested) si procede con la purificazione degli amplificati su colonna (EUROGOLD Cycle-Pure Kit, EuroClone®) in modo da eliminare il più possibile i reagenti in eccesso, residui e dimeri di primers che potrebbero inficiare l'esito della reazione di sequenziamento.

Il sequenziamento è stato effettuato attraverso il metodo Sanger dalla ditta GATC Biotech AG (Cologne, Germany). Per tutti gli amplificati è stato utilizzato il primer di sequenziamento 18S R1052; per alcuni inoltre sono stati sequenziati ulteriori frammenti

utilizzando i primers 18S R536 e 18S F783 (Rosati et al., 2004). I frammenti sono stati poi sovrapposti per ottenere sequenze più complete del gene d'interesse.

Tutti gli elettroferogrammi ottenuti sono stati controllati con il software Chromas Lite (vers. 2.01; McCarthy, 1998). Sono state eliminate le porzioni di sequenza corrispondenti ai primer di amplificazione e le regioni di bassa qualità. Ai doppi picchi incerti (quando il picco secondario ha un'intensità pari alla metà o più di quello principale) sono stati assegnati i corrispondenti simboli di ambiguità secondo la nomenclatura IUPAC. Le sequenze di alta qualità e con una lunghezza maggiore di 500 pb sono state analizzate con il programma BLAST (Basic Local Allignment Search Tool) (Altschul et al., 1990) che, attraverso uno specifico algoritmo, trova all'interno di un'enorme banca dati le sequenze più simili a quella d'interesse e ci dà quindi informazioni riguardo alla più probabile affiliazione.

Le sequenze che davano come risultato del BLAST un organismo ciliato sono state utilizzate per confermare o meno la corrispettiva identificazione morfologica. In taluni casi, sono invece state ottenute sequenze di contaminanti, perlopiù fungini, che sono state scartate.

NOME	VERSO	SEQUENZA	REFERENZE
18S F9 Euk	forward	5'-CTGGTTGATCCTGCCAG-3'	Medlin et al., 1988
18S R1513 Hypo	reverse	5'-TGATCCTTCYGCAGGTTC-3'	Petroni et al., 2002
18S R1052	reverse	5'-AACTAAGAACGGCCATGCA-3'	Rosati et al., 2004
18S R536	reverse	5'-CTGGAATTACCGCGGCTG-3'	Rosati et al., 2004
18S F783	forward	5'-GACGATCAGATACCGTC-3'	Rosati et al., 2004

Tabella 2. Primer utilizzati per le reazioni di amplificazione e sequenziamento.

3.1.3 Analisi filogenetiche

Per due diversi gruppi di ciliati, la sottoclasse Peniculia (subphylum Intramacronucleata, classe Oligohymenophorea) ed il subphylum Postciliodesmatophora, sono state effettuate ricostruzioni filogenetiche attraverso l'utilizzo del pacchetto ARB (Ludwig et al., 2004), per mostrare le relazioni evolutive di alcuni degli organismi osservati.

Le analisi filogenetiche sono state eseguite su un totale di 84 sequenze del gene codificante per l'rRNA 18S appartenenti alla sottoclasse Peniculia, di cui 45 provenienti dal nostro studio e 39 scaricate dal database NCBI GenBank, e su 88 sequenze del gene per l'rRNA 18S

appartenenti al subphylum Postciliodesmatophora, di cui 21 provenienti dal nostro studio e 67 scaricate dal database NCBI GenBank. Le sequenze estratte dal database sono state sottoposte ad un controllo qualità (presenza di siti ambigui, lunghezza della sequenza, anno di pubblicazione). Tutte le sequenze sono poi state importate all'interno del software ARB su sistema operativo Linux. Questo software possiede funzioni avanzate per l'allineamento delle sequenze, ottimizzate soprattutto per i geni dell'RNA ribosomale, permette di creare matrici di caratteri con criteri definiti dall'utente e costruire e visualizzare alberi filogenetici.

Una volta importate, le sequenze sono state allineate automaticamente tramite l'algoritmo del software, per poi affinare l'allineamento manualmente in modo da ottimizzare gli appaiamenti nelle regioni di stem.

Mediante l'utilizzo di filtri programmabili nel software ARB, è possibile escludere delle regioni di allineamento a scelta dell'operatore. In questo modo si possono scartare dall'analisi le regioni agli estremi delle sequenze (la cui lunghezza dipende dai primers utilizzati) e le regioni ipervariabili all'interno della matrice, che potrebbero influenzare l'analisi, collocando alcune specie "*fast-evolving*" in una posizione ambigua. Per entrambi i gruppi, le matrici dei caratteri sono state costruite applicando un filtro che mantiene le colonne in cui almeno uno stesso carattere non-gap è presente almeno nel 5% delle sequenze. Per il gruppo dei *Peniculia* la matrice risultante è costituita da 1181 colonne, mentre quella del gruppo dei *Postciliodesmatophora* da 1114 colonne.

Per stimare in modo accurato il quadro di relazioni evolutive tra le sequenze è necessario scegliere un modello adatto di evoluzione molecolare. La scelta è stata compiuta in base al criterio statistico AIC (Akaike Information Criterion) secondo i calcoli effettuati dal software jModelTest v2.1.4 (Darriba et al., 2012). Il modello utilizzato è il GTR+I+G (*General Time Reversible* + siti invarianti + funzione gamma, da noi approssimata con 4 categorie discrete).

Il metodo utilizzato in questo studio è il *Maximum Likelihood*. Tale metodo cerca l'albero con la massima verosimiglianza, cioè quello che massimizza la probabilità di osservare i dati raccolti. I due alberi filogenetici sono stati calcolati tramite il software PHYML compreso all'interno del pacchetto ARB. La stabilità dei nodi nella topologia degli alberi è stata stimata attraverso la procedura del *bootstrapping* (Felsenstein, 1981) utilizzando 100 pseudorepliche.

3.1.4 Analisi statistiche

Per rappresentare i rapporti fra le comunità di ciliati osservate nei diversi siti di campionamento è stata impiegata una tecnica di ordinamento multivariata.

Il metodo impiegato è l'nMDS (*non-metric Multidimensional Scaling*), che può essere applicato anche a disegni non bilanciati (diverso numero di unità di campionamento per fattore) e che non richiede l'assunzione di normalità. La procedura permette una rappresentazione grafica delle similarità (o dissimilarità) tra campioni in uno spazio bidimensionale (Benedetti-Cecchi et al., 2003).

L'analisi si effettua a partire da una matrice triangolare di similarità (o di dissimilarità) calcolata fra tutte le coppie possibili di campioni; quest'ultima deriva da una precedente matrice $m \times n$, dove m rappresenta il numero di campioni e n il numero di taxa osservati. La similarità (o dissimilarità) tra campioni viene comunemente calcolata tramite l'indice di Bray-Curtis (Bray & Curtis, 1957), misura ampiamente utilizzata in ecologia.

Il fine dell'nMDS applicato alla matrice di similarità è quello di costruire una mappa o "configurazione" dei campioni, in un numero specificato di dimensioni, in modo che l'ordine dei ranghi delle distanze tra i campioni sulla mappa rifletta l'ordine dei ranghi dei valori di similarità (o dissimilarità) presi dalla matrice triangolare; i punti nel piano risultano quindi tanto più vicini tanto maggiore è la loro similarità.

Per valutare la bontà della rappresentazione di un modello nMDS si utilizza una misura di errore chiamata stress; in generale, se il valore di stress è minore di 0,2 si può ritenere l'ordinamento attendibile e tanto migliore quanto più lo "stress" tende a zero.

Con i dati provenienti dalle checklist di ogni campione, sono state costruite quattro diverse matrici d'incidenza (A, B, C, D) che presentano sulle righe i campioni (eccetto i tre campioni provenienti dal lago nero della primavera del 2013) mentre sulle colonne i taxa osservati; i dati sono stati quindi espressi sottoforma di presenza (1) / assenza (0) dei taxa in ogni campione. Per garantire che tutte le colonne fossero mutuamente esclusive, sono state rimosse dalle matrici le colonne con le osservazioni riferite alle categorie "*unknown ciliate*" e "*unknown Oligohymenophorea*". Nella matrice A sono stati riportati i taxa trovati al tempo zero nei diversi campioni, nella matrice B i taxa ritrovati considerando il tempo totale di osservazione di 35 giorni. Le matrici C e D riportano gli stessi dati, ma raggruppando i taxa in categorie più ampie, coincidenti con le sottoclassi.

Mediante il software PAST (versione 2.17c) (PAleontological STatistic, Hammer et al., 2001), le matrici d'incidenza sono state sottoposte a ordinamento nMDS basato su matrici di similarità calcolate con il coefficiente di Bray-Curtis.

3.2 Secondo approccio: Next-Generation Sequencing

3.2.1 Estrazione del DNA e preparazione delle *libraries* Illumina

La caratterizzazione delle comunità di ciliati con un approccio esclusivamente molecolare è stata eseguita su dodici dei campioni fissati in etanolo e conservati a -20°C (Tabella 3).

Le dodici aliquote contenenti 15 ml di campione (acqua e sedimento) fissato in etanolo al 70% sono state sottoposte a estrazione del DNA totale.

Le Falcon sono state centrifugate a circa 5000 RCF per 10 minuti in modo da far precipitare il sedimento in sospensione e gli organismi ciliati. In seguito l'etanolo è stato eliminato e il pellet trasferito nelle provette del kit commerciale PowerSoil® DNA Isolation Kit (MoBio) con cui è stata eseguita l'estrazione. Il kit impiegato è ottimizzato per l'isolamento di DNA da campioni di suolo e per la rimozione dei composti umici presenti nel sedimento, che potrebbero inibire le successive reazioni di amplificazione. La procedura utilizzata è quella descritta nel protocollo del kit eccetto per il passaggio finale di eluizione del DNA, in cui abbiamo utilizzato acqua distillata sterile al posto del buffer in dotazione (soluzione C6).

Campione
β PFEU1
γ PFEU1
β PFEU3
γ PFEU3
β LdG2
γ LdG2
β LdG3
γ LdG3
β Acq1
γ Acq1
β Omb2
γ Omb2

Tabella 3. I dodici campioni utilizzati nel II approccio di caratterizzazione.

La concentrazione del DNA estratto da ogni campione è stata quantificata attraverso un saggio fluorimetrico. Questo metodo di quantificazione prevede l'uso di sonde che sono in grado di legarsi al DNA generando una fluorescenza rilevabile attraverso un fluorimetro. L'intensità dell'emissione è dipendente dalla quantità della molecola contenente il fluoroforo legata con successo al DNA. Il kit utilizzato è il Qubit dsDNA HS Assay (Life Technologies, Invitrogen) con cui è stato possibile rilevare la concentrazione di DNA a doppio filamento. Dopo aver effettuato la quantificazione, ogni campione è stato normalizzato ad una concentrazione finale di 5 ng/μl, necessaria per effettuare le successive analisi.

Per la caratterizzazione molecolare delle comunità di ciliati dei dodici campioni presi in esame, è stato scelto come marcatore la regione V4 del gene codificante per l'rRNA 18S. La scelta è ricaduta su tale marcatore sulla base di studi in silico in cui è stato mostrato come questa regione stimi valori di ricchezza in specie per le comunità di ciliati simili a quelli calcolati utilizzando l'intera sequenza del gene (Dunthorn et al., 2012). Nei ciliati le dimensioni della regione V4 sono molto variabili, da circa 350 a quasi 550 pb.

La preparazione delle *libraries* dei prodotti di PCR di ogni campione per il sequenziamento su MiSeq è stata eseguita seguendo il protocollo Illumina "16S Metagenomic Sequencing Library Preparation" (Figura 4).

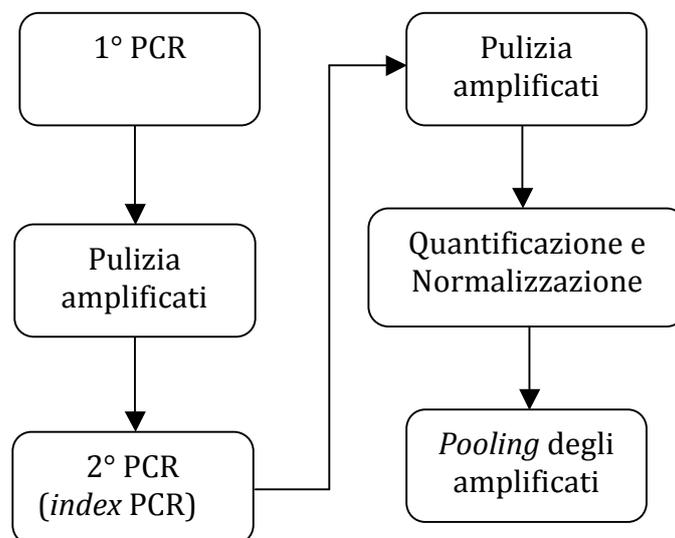


Figura 4. Workflow del protocollo Illumina "16S Metagenomic Sequencing Library Preparation".

Il protocollo prevede una prima reazione di PCR necessaria per amplificare la regione V4 dalle molecole di DNA estratto. Durante questa prima PCR, alle estremità di ogni amplificato viene aggiunto un adattatore Illumina, cioè una specifica sequenza di nucleotidi necessaria per la successiva reazione di PCR.

Per valutare la combinazione di primer che consentisse di amplificare in maniera ottimale la regione V4 dal maggior numero di specie di ciliati a vita libera, sono state effettuate delle prove in silico utilizzando il database di sequenze SSUref_NR_119 SILVA tramite il pacchetto ARB. La miglior coppia di primer trovata è quella rappresentata dal primer forward TAREuk454FWD1 (Stoeck et al., 2010) leggermente modificato e dalla miscela dei 4 primers reverse Cil-r I, Cil-r II e Cil-r III, Cil-r IV (Lara et al., 2007) (Tabella 4). I primer sono stati sintetizzati con l'aggiunta al 5' della corrispondente sequenza dell'adattatore necessario per la compatibilità con gli indici Illumina della seconda PCR.

Adattatore Illumina I + Primer Forward (TAREuk454FWD1 modificato)

5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAGCASC CGCGGTAATWCC-3'

Adattatore Illumina II + Primers Reverse (Cil-r I; Cil-r II; Cil-r III; Cil-r IV)

5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTGATCGTCTTTGATCCCCTTA-3'

5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTAATCGTCTTTGATCCCCTA-3'

5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTGATCGTCTTTGATCCCCTA-3'

5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTGATTGTCTTTGATCCCCTA-3'

Tabella 4. Primers utilizzati nella prima PCR. In grassetto è riportata la sequenza degli adattatori Illumina. Il primer forward utilizzato ha una degenerazione aggiuntiva alla base 18 e una in meno alla base 9 rispetto all'originale di Stoeck et al., 2010.

La reazione di amplificazione è stata effettuata con un Thermal Cycler modello C1000TM (BioRad, Hercules, California) utilizzando per ogni campione una miscela di reazione contenente 2,5 µl di DNA (5 ng/µl), 5 µl di primer forward (1 µM), 5 µl della miscela equimolare dei 4 primers reverse (1 µM) e 12,5 µl di KAPA HiFi HotStart ReadyMix (2X) (Kapa Biosystems), una miscela contenente buffer di reazione, MgCl₂, dNTPs e la DNA polimerasi KAPA HiFi HotStart. Dopo una denaturazione iniziale a 95° C per 3 minuti, è stato ripetuto per 25 cicli il seguente profilo termico:

- denaturazione a 95°C per 30 secondi
- annealing a 55°C per 30 secondi

- estensione a 72°C per 30 secondi

Il passaggio di estensione finale a 72°C ha avuto una durata di 5 minuti.

I prodotti di PCR sono stati purificati usando speciali biglie magnetiche (Agencourt AMPure XP, Beckman Coulter, USA) in grado di legare frammenti di DNA superiori a 100 pb; su questo principio si basa la separazione degli amplificati da eventuali prodotti aspecifici (primer, dimeri di primer), dNTP non incorporati e altri reagenti che vengono così eliminati.

Una seconda PCR definita “*Index PCR*” è stata utilizzata per legare agli amplificati di ogni campione una coppia univoca di indici Illumina (Nextera XT Index kit, Illumina) (Tabella 5); gli indici sono necessari per poter separare le sequenze appartenenti ad ogni campione dopo il sequenziamento, grazie a una breve regione unica (“*barcode*”). Gli indici contengono inoltre un ulteriore adattatore Illumina necessario per il sequenziamento.

Il profilo termico di questa seconda PCR è identico al precedente ma si ripete solo per 8 cicli. Per ogni campione sono stati utilizzati 5 µl di amplificato purificato proveniente dalla prima PCR, 10 µl della specifica coppia di indici, 2 µl della miscela KAPA HiFi HotStart Ready Mix(2X) e 10 µl di acqua distillata sterile.

		Indice 2			
		S501	S502	S503	S504
Indice 1	N701	βPFEU1	γPFEU1	βPFEU3	γPFEU3
	N702	βLdG2	γLdG2	βLdG3	γLdG3
	N703	βAcq1	γAcq1	βOmb2	γOmb2

Tabella 5. Sono mostrate le combinazioni degli indici Illumina utilizzate per identificare i dodici campioni.

Dopo un'ulteriore purificazione mediante biglie magnetiche, i prodotti di amplificazione di questa seconda PCR sono stati quantificati e portati ad una concentrazione di 4 nM.

Da ogni campione sono stati aliquotati 5 µl e mescolati poi in una singola provetta eppendorf (*pooling*).

Le *libraries* di amplificati dei dodici campioni sono state quindi sequenziate con tecnologia Illumina, un metodo di nuova generazione in grado di analizzare simultaneamente milioni di

molecole di DNA. Il sequenziamento Illumina è stato effettuato dalla BMR Genomics (Padova) su piattaforma MiSeq; gli amplificati sono stati letti a partire da entrambe le estremità per circa 300 pb (sequenziamento 2x300 *paired-end* con kit di reagenti MiSeq® v3 da 600 cicli) in modo da ottenere, grazie alla successiva sovrapposizione in silico dei due frammenti, una sequenza contenente l'intera regione V4.

3.2.2 Analisi delle sequenze, ricerca delle OTU e affiliazione tassonomica

I dati ottenuti dal sequenziamento Illumina *paired-end* sono stati forniti in formato fastq e divisi per campione. Il sequenziatore produce dati grezzi (“*reads*”) in un formato che contiene 3 informazioni fondamentali:

- ID (identificatore individuale della *read*)
- sequenza nucleotidica
- stima della qualità di ogni base

A ciascuna base della sequenza è infatti assegnato un punteggio di qualità (“*quality score*”), inversamente correlato alla probabilità che quella base non sia corretta.

Il programma FastQC (vers. 0.11.2) è stato utilizzato per visualizzare i *quality score* dei file fastq.

I dati originali sono stati elaborati con due pacchetti di software: QIIME (“*Quantitative Insights Into Microbial Ecology*”, Caporaso et al., 2010) e USEARCH (Edgar, 2010). Con ciascun programma sono state effettuate due diverse analisi dei dati, una più permissiva e una più restrittiva in termini di filtraggio delle sequenze. In totale sono state quindi eseguite 4 analisi (*pipelines*) in cui abbiamo impostato diversi valori dei parametri qualitativi.

I due programmi hanno previsto numerosi passaggi, attraverso i quali a partire dalle *reads* sono state ottenute per ogni campione le sequenze V4 di buona qualità appartenenti al gruppo dei ciliati e la loro affiliazione tassonomica.

Il sequenziamento *paired-end* ha prodotto per ogni amplificato 2 *reads* di 300pb, che sono stati assemblati per ottenere le sequenze complete (*contig*) della regione V4.

Diversi filtri qualitativi sono stati applicati alla zona di sovrapposizione; è stata considerata la sua lunghezza minima e la massima percentuale di basi discordanti ammesse.

I *contig* ottenuti per ogni campione sono stati sottoposti ad ulteriori filtri qualitativi basati sui *quality score* dei singoli nucleotidi dell'intera sequenza.

Le porzioni di sequenza corrispondenti ai primer utilizzati durante la PCR sono state rimosse, e solo i *contig* con lunghezza rimanente pari o superiore a 250 pb sono stati conservati.

Per studiare la composizione tassonomica dei 12 campioni ci siamo avvalsi del concetto di OTU (*Operational Taxonomic Unit*). In uno studio basato solo su dati di tipo molecolare, una OTU consiste in un insieme di *contig* che presentano fra loro un'elevata similarità di sequenza. Nella nostra analisi abbiamo scelto una soglia di similarità minima del 97%, il valore più spesso incontrato in letteratura.

Per individuare tali OTU, i *contig* sono stati inizialmente raggruppati utilizzando un dataset di riferimento (SILVA 119); in seguito, le sequenze con una similarità minore del 97% con quelle contenute nel dataset sono state confrontate fra loro e raggruppate in OTU "*de novo*".

Dal momento che ogni OTU può contenere molte sequenze simili ma non identiche, la sequenza più abbondante di ognuna è stata scelta come rappresentativa nelle analisi a valle.

L'individuazione di possibili chimere (sequenze spurie contenenti frammenti di diversa derivazione, che si formano durante le PCR) è stata effettuata confrontando le sequenze rappresentative con un database di riferimento privo di sequenze chimeriche; per tutte e 4 le analisi è stato utilizzato lo *script* "*uchime_ref*" all'interno del pacchetto USEARCH.

Oltre alle OTU chimeriche individuate, sono state eliminate da ogni campione anche le OTU costituite da 10 o meno sequenze (*denoising*) in modo da ridurre possibili artefatti e contaminazioni, e contenere la sovrastima della diversità.

L'identificazione tassonomica delle OTU è stata eseguita su QIIME. Ogni sequenza rappresentativa è stata classificata utilizzando il database di riferimento per gli eucarioti SILVA 119; le OTU che avevano una similarità di sequenza minore del 90% con tali referenze sono state rinominate come "*unassigned*".

Per tutte e 4 le analisi sono poi state estratte OTU e *contig* appartenenti al gruppo dei Ciliati.

3.2.3 Analisi *pipeline 2 bis*

Solo l'analisi che ha fornito il minor numero di OTU di ciliati ("*Pipeline 2bis*") è stata presa in considerazione nella seconda parte dello studio.

Per ogni campione è stata indagata la composizione tassonomica in classi e sottoclassi di tali OTU; inoltre le sequenze rappresentative sono state caricate su ARB per controllare la correttezza dall'affiliazione tassonomica delle OTU eseguita da QIIME.

Per ogni campione è stata calcolata la rispettiva curva di rarefazione, per indagare se lo sforzo di sequenziamento sia stato sufficiente a rilevare tutte le possibili OTU di ciliati presenti.

Anche in questo caso, l'nMDS è stato scelto come metodo di ordinamento per rappresentare i rapporti fra le comunità di ciliati. Disponendo dei dati di abbondanza delle OTU in ogni campione, è stato possibile costruire una matrice in cui ciascun elemento rappresentasse l'abbondanza della *i*-esima OTU (intesa come numero di *contig* riconducibili a quella OTU) nel *j*-esimo campione.

Prima di realizzare tale matrice, il numero di sequenze di ogni campione è stato standardizzato ad uno stesso valore, pari al numero di *contig* nel campione meno abbondante. Mediante il software PAST la matrice di abbondanza standardizzata è stata sottoposta a ordinamento nMDS basato su matrici di similarità calcolate con il coefficiente di Bray-Curtis.

3.3 Confronto dei due approcci

Il numero di OTU molecolari di ogni campione sottoposto ad analisi Illumina è stato confrontato con il numero di morfotipi (o OTU morfologiche) trovati mediante il I approccio nei campioni provenienti dalle 12 medesime stazioni. Oltre al numero di OTU sono state confrontate anche le relative composizioni tassonomiche suddivise per classi e sottoclassi.

È stato effettuato anche un confronto dei risultati nMDS ottenuti su due diverse matrici: per il primo approccio è stata costruita una matrice d'incidenza con i dati derivanti solo dalle 12 stazioni a comune, mentre per il II approccio la matrice di abbondanza è stata trasformata in una d'incidenza in modo che l'ordinamento grafico risultante tenesse conto solamente della ricchezza in taxa di ogni campione.

Risultati e Discussione

1. Primo approccio: identificazione morfologica e sequenziamento Sanger

1.1 Numero di morfotipi

Il primo approccio di caratterizzazione effettuato su 75 campioni ha individuato in totale 1031 osservazioni (Tabella 6 riportata in appendice) per una media di 14 diversi morfotipi per campione. γ PFEU1 e γ PFEU2 sono i campioni in cui è stato riscontrato il maggior numero di taxa (32), mentre α PFT2 è quello in cui ne è stato osservato il minor numero (2).

La check-list che prende in considerazione solo i taxa osservati il giorno stesso del campionamento (t_0) (taxa riquadrati in verde nella tabella 6) ha riportato un totale di 589 osservazioni, cioè solo il 57,1% della diversità complessiva osservata in 5 settimane; a t_0 sono stati osservati una media di 8 morfotipi per campione.

La percentuale di taxa osservati esclusivamente in seguito al primo controllo va da un minimo dell'11% (campione β Omb2 dove sono stati osservati solo 2 nuovi taxa su un totale di 17) fino a valori massimi del 100%. Nelle stazioni α TS1, α TTo1 e β PFEU3 a t_0 infatti non era stato individuato nessun organismo ciliato; l'osservazione ripetuta ha invece permesso di ritrovare fino a 7 morfotipi. La percentuale, come appena descritto, varia molto da campione a campione (Figura 5), ma le medie delle diverse aree non differiscono in modo sostanziale tra i diversi fattori (stagione, tipologia di habitat, altitudine etc.).

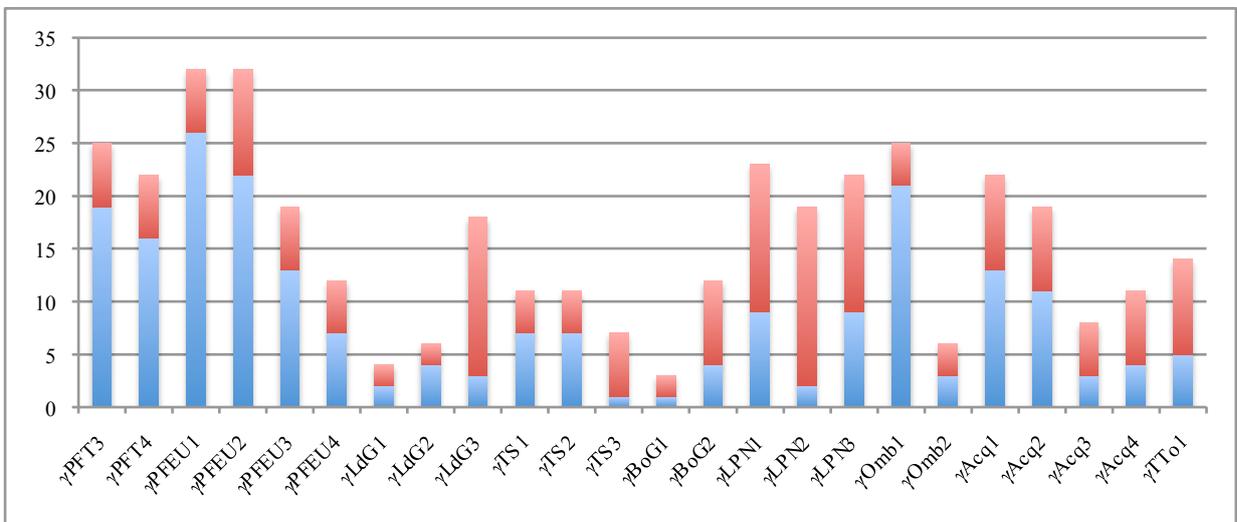
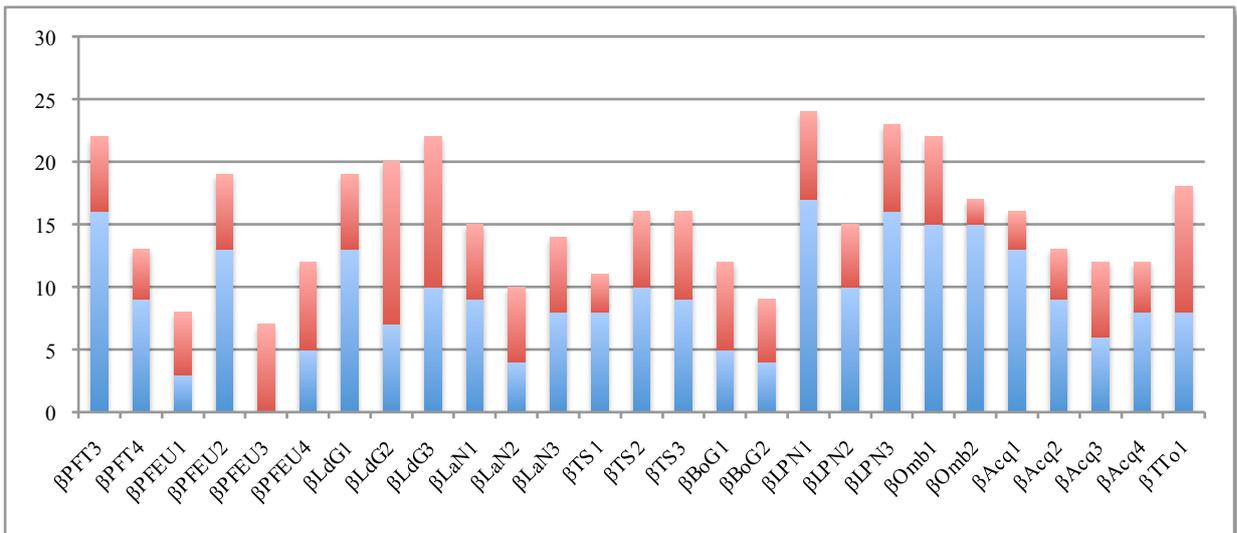
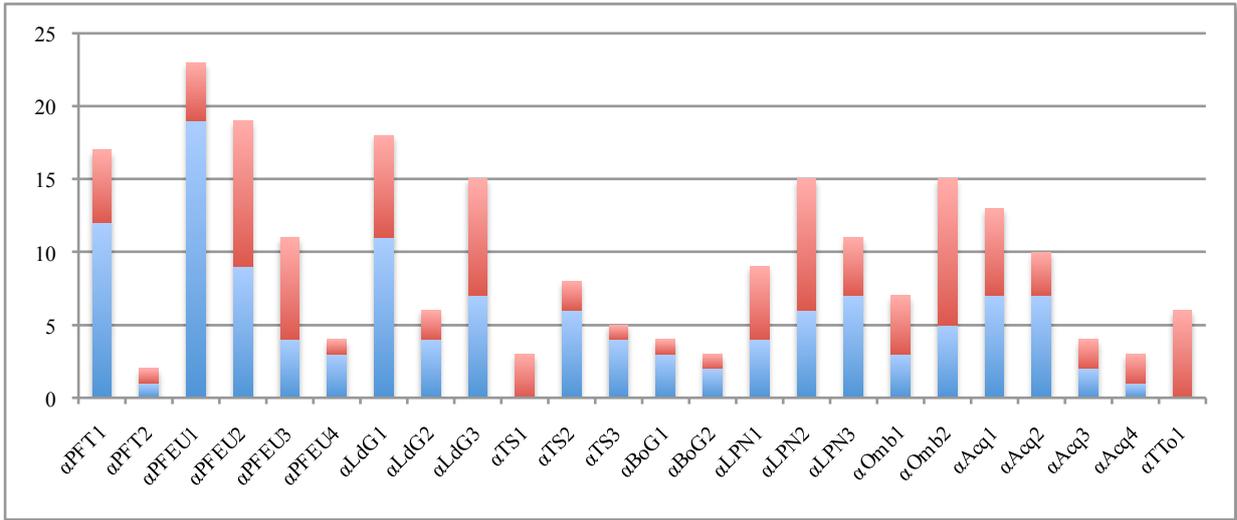


Figura 4. Sulle ordinate è riportato il numero di taxa osservati mentre sulle ascisse sono riportati i 75 campioni. Ogni grafico corrisponde a una stagione di campionamento (α : autunno 2012; β : primavera 2013; γ : autunno 2013). In blu è mostrato il numero di taxa osservati a t0 mentre in rosso i taxa osservati esclusivamente in seguito.

L'osservazione ripetuta nel tempo effettuata sulle tre aliquote di ogni campione ci ha permesso quindi di scoprire una grande parte di biodiversità all'interno delle comunità di ciliati che altrimenti non sarebbe stata individuata. I dati mostrano infatti come grazie a questa metodica il numero di morfotipi trovati mediamente in un campione quasi raddoppi. Questa parte di biodiversità "nascosta" è essenzialmente costituita da organismi ciliati presenti inizialmente nel campione sottoforma di cisti oppure rappresentati da pochi individui in attesa delle condizioni ambientali idonee per moltiplicarsi. Risultati comparabili sono stati ottenuti anche in altri studi (Fenchel et al., 1997), dove aliquote dello stesso campione venivano, oltre che osservate per lungo tempo, sottoposte a diversi trattamenti (arricchimento con numerosi nutrienti, modifiche della temperatura, della concentrazione di ossigeno disciolto etc..) in modo da creare microhabitat con condizioni ecologiche differenti adatte per lo sviluppo dei ciliati rari e/o incistati.

Il cariossido di riso aggiunto alle aliquote dei nostri campioni è servito come risorsa addizionale di nutrimento per i batteri che, aumentando la loro biomassa, hanno favorito l'excistamento di molti dei ciliati presenti e supportato la crescita di altri.

Come abbiamo già detto la parte di biodiversità nascosta non è infatti soltanto costituita da organismi incistati ma anche da quei taxa rappresentati da pochi individui che sfuggono ad una prima ispezione. Questa affermazione può essere sostenuta da osservazioni effettuate su alcuni dei nostri campioni in cui rappresentanti del genere *Paramecium* sono stati trovati solo dopo una settimana dal campionamento. I ciliati del genere *Paramecium* non possiedono infatti nel loro ciclo vitale uno stadio di quiescenza e quindi presumibilmente erano già presenti nel campione, ma rappresentati da pochi individui.

1.2 Composizione tassonomica

L'osservazione morfologica ha individuato in totale 54 differenti categorie tassonomiche (riportate in appendice a pag. 85).

Le classi di ciliati più rappresentate sono quella degli Spirotrichea (36,0% del totale delle osservazioni) e quella degli Oligohymenophorea (28,4%) (Figura 6). Queste due classi comprendono alcuni dei generi più conosciuti come *Euplotes* (Spirotrichea, Hypotrichia), *Paramecium* e *Tetrahymena* (Oligohymenophorea, Peniculia). Entrambe sono accomunate da

un'ampia diversità complessiva, sia per numero di sottoclassi che per numero di specie (Lynn, 2008). Nei nostri campioni la sottoclasse degli Spirotrichea più rappresentata è quella degli Stichotrichia mentre, con i generi *Paramecium* e *Frontonia*, i Peniculia sono la sottoclasse degli Oligohymenophorea osservata più frequentemente.

All'interno della classe dei Prostomatea, che comprende il 10,4% delle osservazioni, quasi la metà appartiene al genere *Coleps* (e affini), mentre i Litostomatea (11% delle osservazioni) sono rappresentati da numerosi morfotipi del gruppo a vita libera degli Haptoria. Il predatore gigante *Homalozoon*, che può arrivare fino ad una lunghezza di 2,5 mm (Lynn, 2008), è stato osservato solo nei campioni provenienti dai bacini artificiali, quali le vasche di acquacoltura e il lago di pesca sportiva, cioè sistemi con un elevato contenuto di nutrienti. I ciliati appartenenti al genere *Didinium* (e affini), noti predatori del più grande *Paramecium*, sono stati osservati solo nei campioni provenienti dalle aree del Padule di Fucecchio.

I taxa più rappresentati all'interno della classe degli Heterotrichea (5% delle osservazioni) sono il genere *Spirostomum* e il genere *Stentor*, di cui *Stentor coeruleus* è la specie più facilmente identificabile per la presenza di una colorazione citoplasmatica verde-azzurra.

Le classi meno rappresentate sono quelle a cui appartengono ciliati anaerobi o microaerofili, che prediligono quindi nicchie ecologiche peculiari: Armophorea (2,7%), Colpodea (0,38%) e Plagiopylea (0,09%). I Karyorelictea sono anch'essi poco rappresentati (1,45%) poiché comprendono solo un genere di ciliati di acqua dolce (*Loxodes*) con poche specie microaerofile.

La numerosa classe dei Phyllopharyngea in questo studio è stata osservata molto di rado (1,2%) poiché la maggior parte dei suoi membri possiede uno stile di vita parassitico o commensale.

I diversi morfotipi appartenenti al gruppo dei Nassophorea sono solitamente rappresentati da popolazioni non abbondanti (Lynn, 2008) e per questo motivo possono sfuggire molto facilmente all'osservazione; nel nostro studio tale classe è stata infatti individuata raramente (1,8% delle osservazioni).

Per l'1,4% delle osservazioni non è stata identificata la classe di appartenenza e quindi sono state categorizzate come "unknown ciliate".

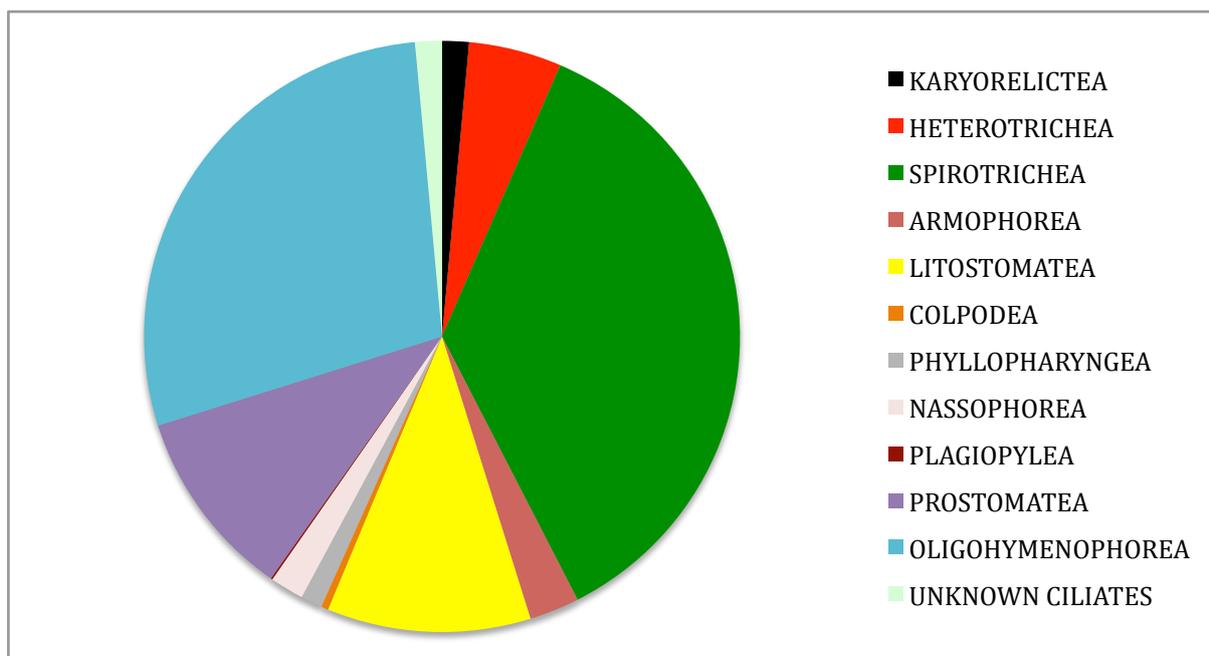


Figura 6. Nel grafico sono rappresentate le osservazioni totali trovate nei 75 campioni divise per classe.

Per la maggior parte delle classi il rapporto t_0/t_T si avvicina alla media calcolata sul totale delle osservazioni (57,1%) (Figura 7). Vi sono però tre eccezioni. La prima riguarda la classe dei Plagiopylea per cui però è stato individuato solo un morfotipo in un'unica stazione; tale taxon è stato osservato dopo due settimane dal campionamento nell'aliquota III del campione α PFEU3.

Le altre due eccezioni riguardano i Karyolicea e gli Armophorea; ad entrambe le classi appartengono ciliati di acqua dolce adattati a basse concentrazioni di ossigeno ma, mentre per i Karyolicea si conoscano solo poche specie appartenenti ad un singolo genere (*Loxodes*; numerosi altri generi appartenenti alla stessa classe sono esclusivamente marini), gli Armophorea sono rappresentati da numerose specie. Dalla nostra analisi è risultato che i rispettivi valori del rapporto t_0/t_T si discostano dalla media in senso opposto: per i Karyolicea l'80% delle osservazioni è stato individuato il giorno stesso del campionamento, mentre per gli Armophorea solo il 32% è stato osservato a t_0 . Questo suggerisce che la forma liberamente natante di *Loxodes* sia abbastanza comune nei campioni, non avendo necessità di condizioni estreme di ossigeno, ma scompaia velocemente probabilmente poiché suscettibile a variazioni anche minime di tali livelli (come accade ad esempio quando il campione originale viene trasferito nelle due piastre Petri). Per gli Armophorea vale la situazione

opposta, ovvero la maggior parte dei morfotipi sono presenti inizialmente come forme di resistenza che si schiudono successivamente, quando si vengono a creare nel campione condizioni anossiche; questi ciliati, eccetto che per una singola osservazione, sono stati riscontrati nelle aliquote III.

Inoltre dalle nostre check-list si può notare come, contrariamente al genere *Loxodes*, gli Armophorea siano stati osservati, eccetto che in un caso (campione α TT01), solo in campioni provenienti da bacini lentici.

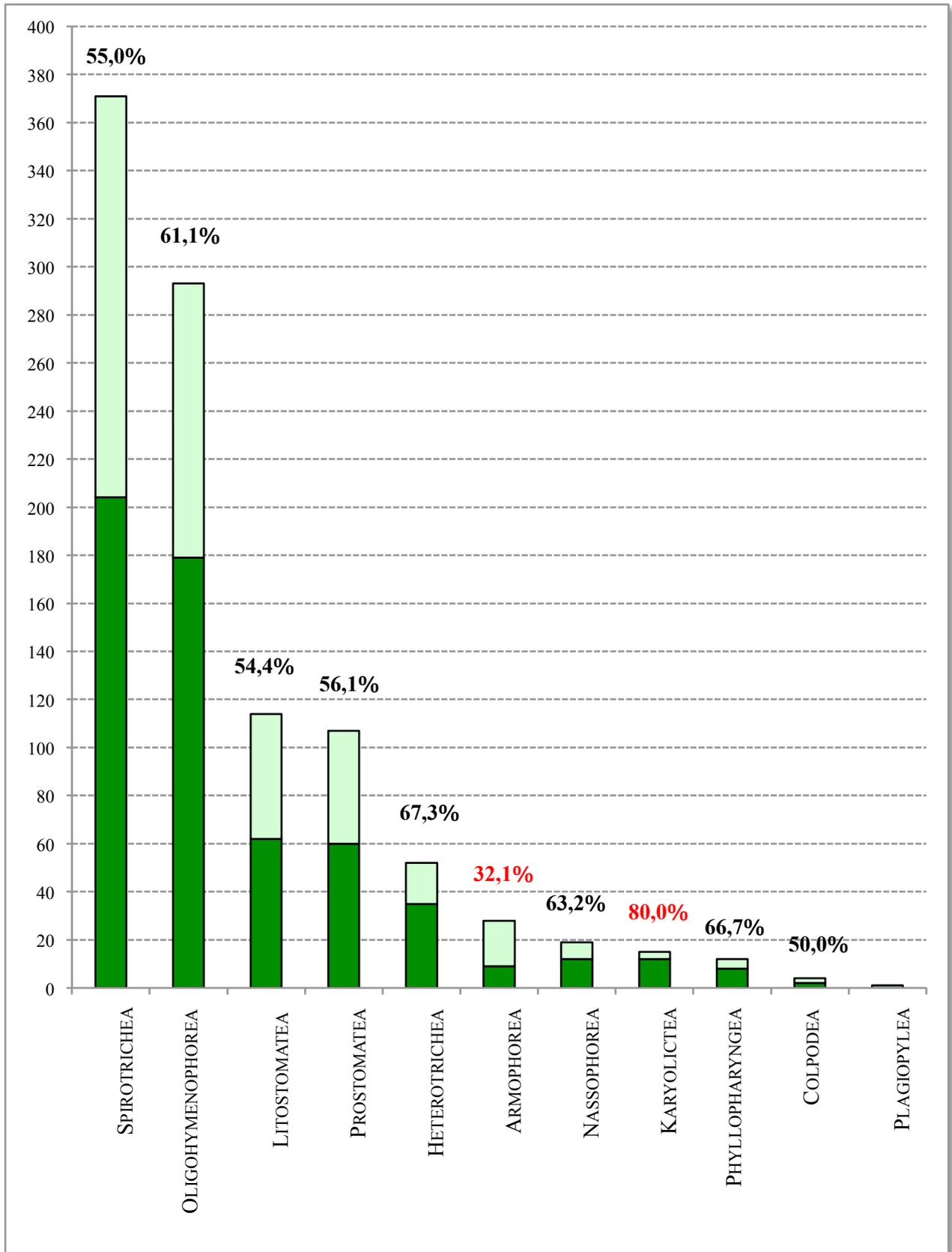


Figura 7. Nel grafico è mostrato il numero di osservazioni totali divise per classe; in verde scuro sono evidenziate le osservazioni trovate a t_0 mentre in verde chiaro quelle individuate esclusivamente in seguito. Le percentuali si riferiscono al rapporto t_0/t_T di ogni classe.

1.3 Dati molecolari

Dalle analisi molecolari sulle cellule fissate in etanolo abbiamo ottenuto un totale di 261 sequenze del gene codificante per l'rRNA 18S; tali sequenze corrispondono al 25,3% delle osservazioni.

Due passaggi in particolare ci hanno impedito di ottenere un maggior numero di dati: l'ottenimento di materiale su cui effettuare le analisi molecolari e l'efficienza delle PCR su cellule.

Meno della metà delle popolazioni osservate sono state fissate (41%). Infatti, solo per i morfotipi più abbondanti è stato possibile prelevare con facilità cellule su cui effettuare direttamente la PCR, anche considerando la perdita di materiale (lisi) durante i passaggi di lavaggio. Non è inoltre possibile, come già accennato, ottenere colture per tutti i ciliati poiché la maggior parte non sono coltivabili.

Il secondo ostacolo riguarda invece l'efficienza stessa della PCR su singola o poche cellule (38,2%) per la quale, omettendo il passaggio di estrazione del DNA, si ha la permanenza nella provetta di possibili inibitori della reazione di amplificazione.

Anche le PCR semi-nested non hanno fornito molti dati (efficienza 22,8%) poiché spesso veniva amplificata insieme alla molecola di interesse anche quella di contaminanti rimasti nonostante i passaggi di lavaggio in acqua distillata.

Un'efficienza maggiore (87,7%) si è avuta con le PCR eseguite sul DNA estratto da coltura; purtroppo però, oltre al fatto che solo una piccola parte dei ciliati è coltivabile, il processo è comunque molto dispendioso sia in termini di tempo che di denaro.

La percentuale di aplotipi unici ottenuti nell'analisi (rispetto al numero totale effettivamente presente nelle comunità studiate) è certamente diverso dal 25%, poiché molte delle popolazioni che abbiamo riportato condividono la stessa sequenza del gene per l'rRNA 18S. Questo è dimostrato dalle molte sequenze identiche che abbiamo ottenuto da popolazioni appartenenti allo stesso taxon trovate in campioni differenti.

Questa "ridondanza" varia con i diversi gruppi tassonomici; ad esempio per la sottoclasse dei Peniculia su un totale di 44 sequenze del gene per l'rRNA 18S provenienti da altrettante popolazioni abbiamo ottenuto solo 19 differenti aplotipi; mentre per la sottoclasse degli Haptoria, su 20 sequenze abbiamo ottenuto fino a 15 diversi aplotipi. Questo indica che per i

Peniculia abbiamo una buona copertura della diversità interna al gruppo mentre per gli Haptoria e altri cladi, tale copertura è inferiore.

I dati molecolari ottenuti sono stati utilizzati per confermare l'identificazione morfologica degli organismi individuati. Su 261 sequenze, solo per 3 osservazioni è stata riscontrata un'incongruenza tra l'identificazione morfologica e quella molecolare. Nel campione β TT01 un organismo identificato come Phyllopharyngea è risultato appartenere alla classe dei Nassophorea nell'analisi della sequenza effettuata con il programma BLAST, mentre per altri due organismi trovati nei campioni α LdG3 e α LdG1 e identificati come Heterotrichea (in particolare *Condylostomides*-like) l'analisi delle rispettive sequenze ha modificato l'affiliazione all'interno della classe degli Armophorea.

Il tasso di errore calcolato è quindi basso, stimato intorno all'1,1% degli organismi identificati molecolarmente.

La possibilità di commettere errori dipende dall'abilità dell'osservatore e dal suo grado di familiarità con i differenti gruppi; per questo abbiamo deciso di utilizzare categorie tassonomiche relativamente ampie per certi morfotipi. Sebbene una maggiore risoluzione sia preferibile, la scala di errore potrebbe aumentare di molto cercando di identificare ogni morfotipo a livello di specie o genere. La letteratura ci mostra che per stabilire un tale livello tassonomico, per la maggior parte dei ciliati sono necessari dettagliati studi morfometrici fatti sia su organismi in vivo che fissati, combinati a indagini molecolari effettuate talvolta su più di un marcatore (Modeo et al., 2003; Rosati et al., 2004; Boscaro et al., 2014). Numerosi studi utilizzano sistemi di categorizzazione simili ai nostri e riconoscono un numero più o meno simile di categorie chiaramente distinguibili che corrispondono a differenti ranghi tassonomici (es. Gaedke & Wickham, 2004).

Grazie a questo studio abbiamo ottenuto per la prima volta le sequenze del gene per l'rRNA 18S di due generi descritti finora solo a livello morfologico: *Linostomella* e *Disematostoma*.

1.4 Analisi filogenetiche

Per due diversi gruppi di ciliati, la sottoclasse Peniculia (subphylum Intramacronucleata, classe Oligohymenophorea) e il subphylum dei Postciliodesmatophora, sono state fatte ricostruzioni filogenetiche. In particolare, le analisi sono state effettuate per mostrare i rapporti filogenetici dei due taxa di ciliati per i quali ad oggi mancava qualsiasi tipo di dato molecolare, i generi *Linostomella* e *Disematostoma*.

Linostomella (Figura 8) è un genere di ciliati di acqua dolce appartenente alla famiglia dei Condylomatidae (subphylum Intramacronucleata, classe Heterotrichea). Si tratta di un organismo di medie dimensioni (circa 150 x 100 µm) che si muove senza sosta nella colonna d'acqua, ha un corpo a forma di sacco e un grande apparato boccale che occupa la metà



Figura 8. *Linostomella* sp. immagine in vivo. La lunghezza della barra mostrata è di 20 µm.

anteriore della cellula; si differenzia dal genere affine *Condylostoma* per l'assenza di cirri apicali.

In totale abbiamo osservato 3 popolazione afferenti al genere *Linostomella*, nelle stazioni α PFT1, α PFEU1 e γ PFEU1; tutti e tre i campioni provengono dal Padule di Fucecchio e le osservazioni sono sempre state riscontrate a t_0 e in nessuno dei successivi check. Per due popolazioni siamo riusciti ad ottenere colture stabili che sono state utilizzate come materiale su cui effettuare l'estrazione del tgDNA per ottenere la sequenza del gene per l'rRNA 18S. Dall'albero filogenetico

(Figura 9) si può vedere come le due sequenze ottenute, che sono identiche tra loro, siano strettamente associate all'unica sequenza presente in banca dati del genere *Condylomatides*. Insieme formano il *sister group* del clade *Condylostoma*, che contiene oltre alle sequenze di tale taxon, anche l'unica sequenza conosciuta di *Chattodinium* (che formalmente appartiene ad una diversa famiglia, Chattonidiidae; Modeo et al., 2006). Tralasciando la posizione ambigua di tale sequenza, già riportata in letteratura (Yan et al., 2015; Fernandes et al., 2015), dall'albero si può evincere come la famiglia Condylomatidae, benché non supportato da un elevato valore di *bootstrap*, possa essere definita un gruppo monofiletico.

Anche le somiglianze morfologiche tra *Linostomella*, *Condylostoma* e *Condylostomides* giustificano il loro raggruppamento, e ulteriori sequenze potrebbero stabilizzare questa parte dell'albero degli Heterotrichea.

Le altre sequenze provenienti dalla nostra analisi confermano la restante topologia dell'albero già descritta in letteratura (Miao et al., 2009, Boscaro et al., 2014; Yan et al., 2015). I tre ordini di Karyolicea ricevono un moderato supporto nell'analisi; va messo però in evidenza che le sequenze del gene per l'rRNA 18S di tale gruppo sono *fast-evolving* e quindi molto difficili da allineare in modo non ambiguo.

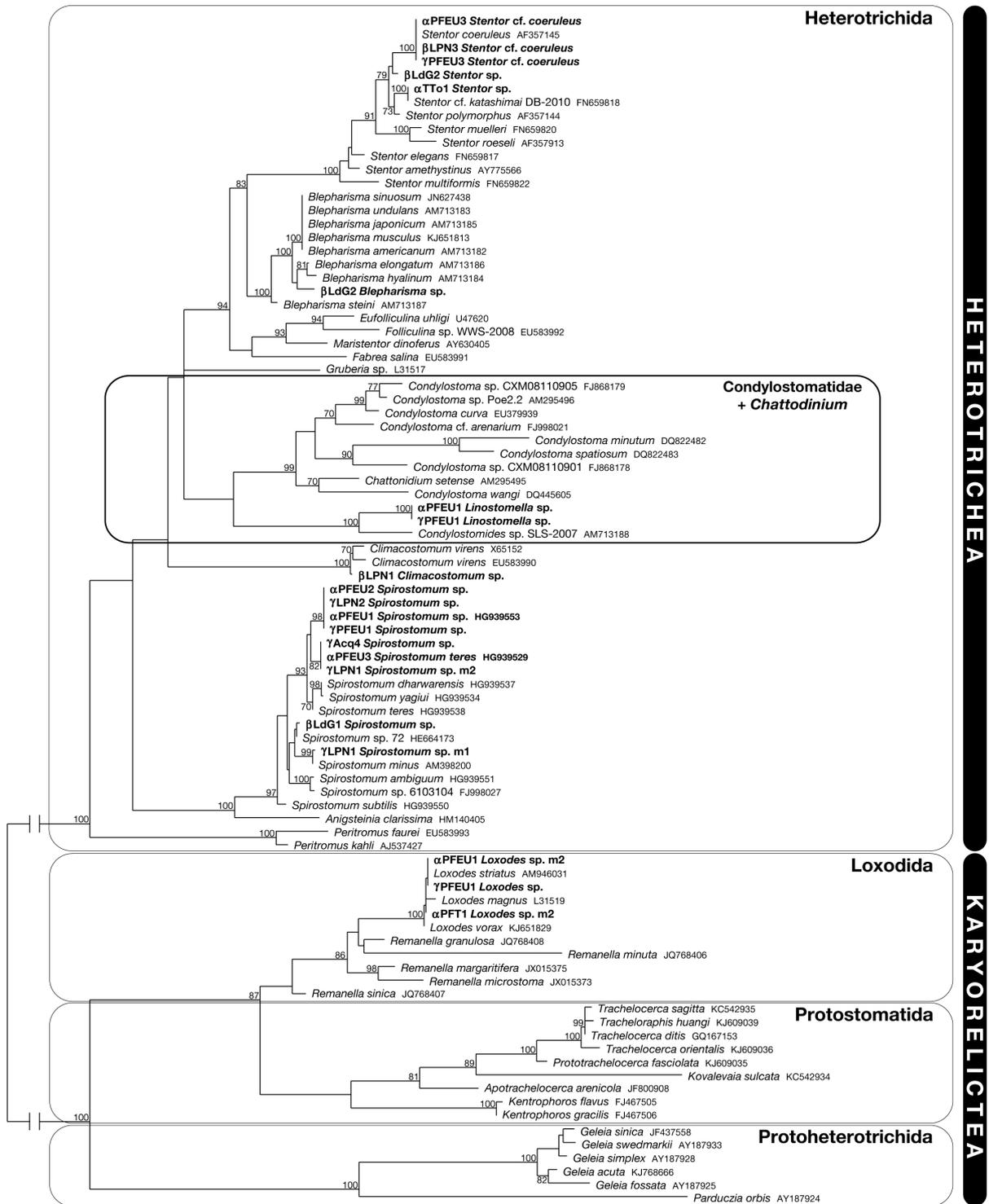


Figura 9. Albero ML del phylum Postciliodesmatophora. La radice è stata messa tra le due classi Heterotrichida e Karyorelictea. Le sequenze ottenute in questo lavoro sono in grassetto. I numeri vicino ai nodi corrispondono ai valori di bootstrap (valori minori di 75 sono stati rimossi). La barra corrisponde a una divergenza stimata del 10%.

Per la sottoclasse dei Peniculia, come si può vedere dall'albero filogenetico (Figura 11), le nostre sequenze coprono una buona parte della diversità esistente interna al gruppo. Il genere *Disematostoma* (Figura 10), rappresentato esclusivamente dalle nostre 3 sequenze con 2 diversi aplotipi, è stato individuato in 5 campioni tutti provenienti dal Padule di Fucecchio. I



Figura 10. *Disematostoma* sp. immagine in vivo. La lunghezza della barra mostrata è di 20 μm .

ciliati appartenenti a tale genere hanno una forma piriforme, con la superficie ventrale piatta e quella dorsale convessa; l'apertura orale si trova a circa un terzo della lunghezza della cellula; come per *Linostomella*, *Disematostoma* è stato ritrovato solo a t_0 e mai più riscontrato in nessuno dei successivi check.

Dall'albero si nota come le 3 sequenze di *Disematostoma* siano strettamente associate all'unica sequenza presente in banca dati del genere *Stokesia*, che formalmente appartiene

però ad una diversa famiglia (Stokesiidae). Insieme ai generi *Frontonia* e *Apofrontonia*, *Disematostoma* appartiene infatti alla famiglia dei Frontoniidae; la ricostruzione filogenetica mostra, come già ampiamente descritto in letteratura (Gao et al., 2008; Fan et al., 2011, 2013), la non monofilia di tale famiglia.

Per quanto riguarda il genere *Frontonia*, di per sé non monofiletico, si può notare come la variabilità molecolare sia più ampia di quella finora presente in banca dati. La maggior parte delle nostre sequenze afferenti a tale morfotipo formano infatti un clade a se stante *sister-group* del clade *Disematostoma-Stokesia*.

In accordo con le analisi molecolari, *Frontonia* è quindi un taxon eterogeneo costituito da numerose linee filetiche separate, che hanno mantenuto una serie di caratteri morfologici primitivi rispetto ad altri generi della sottoclasse Peniculia. Per la famiglia Frontoniidae appare quindi necessaria una revisione tassonomica, che richiede però un gran numero di dati per via delle molte morfospesie descritte del genere *Frontonia* (Fokin, 2008; Fan et al., 2013).

Le altre sequenze di Peniculia provenienti dalla nostra analisi confermano la restante topologia dell'albero rispetto a quella già descritta in letteratura (Boscaro et al., 2012).

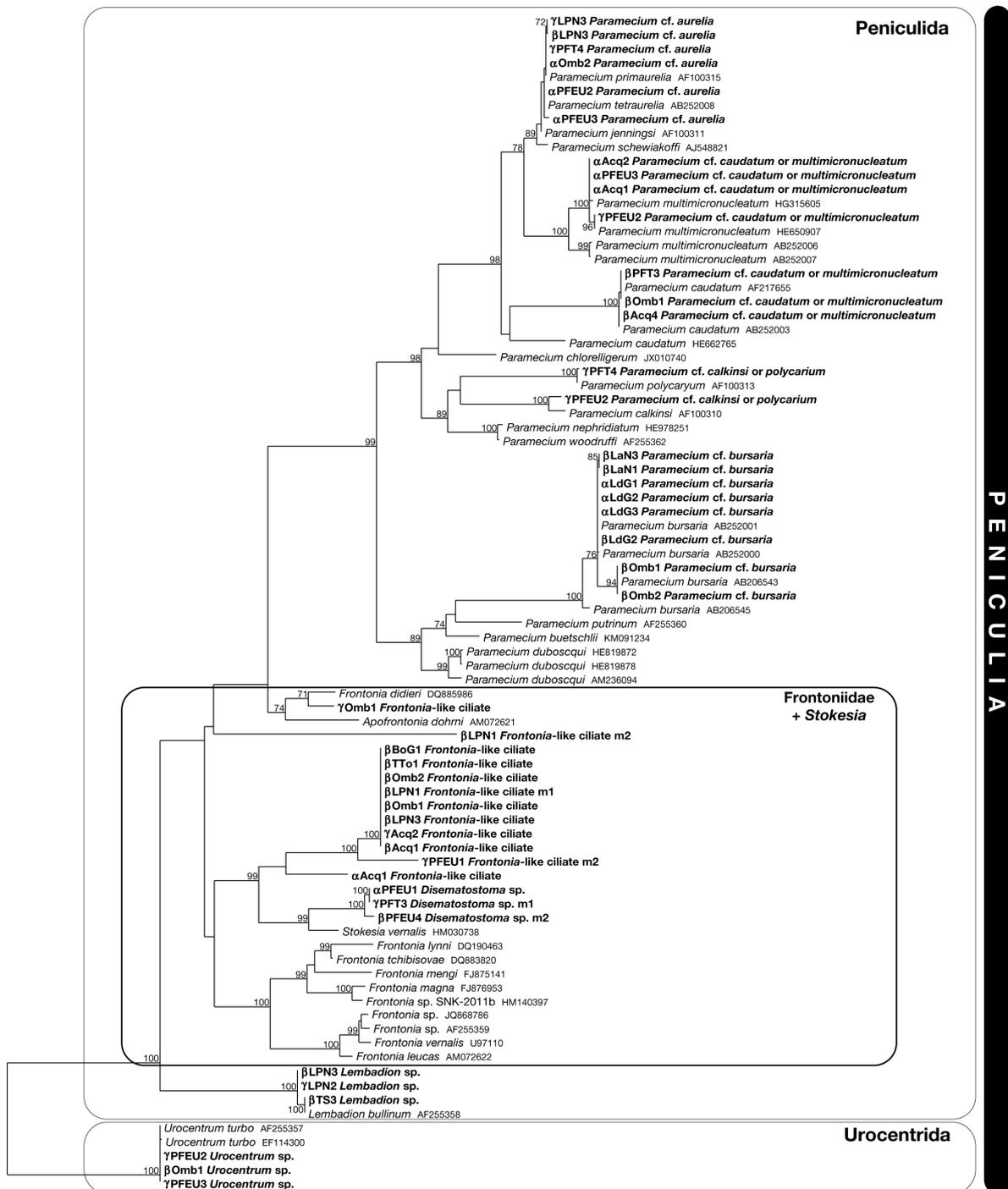


Figura 10. Albero ML della sottoclasse Peniculia (Oligohymenophorea). La radice è stata messa tra i due ordini Peniculida e Urocentrida. Le sequenze ottenute in questo lavoro sono in grassetto. I numeri vicino ai nodi corrispondono ai valori di bootstrap (valori minori di 75 sono stati rimossi). La barra corrisponde a una divergenza stimata del 10%.

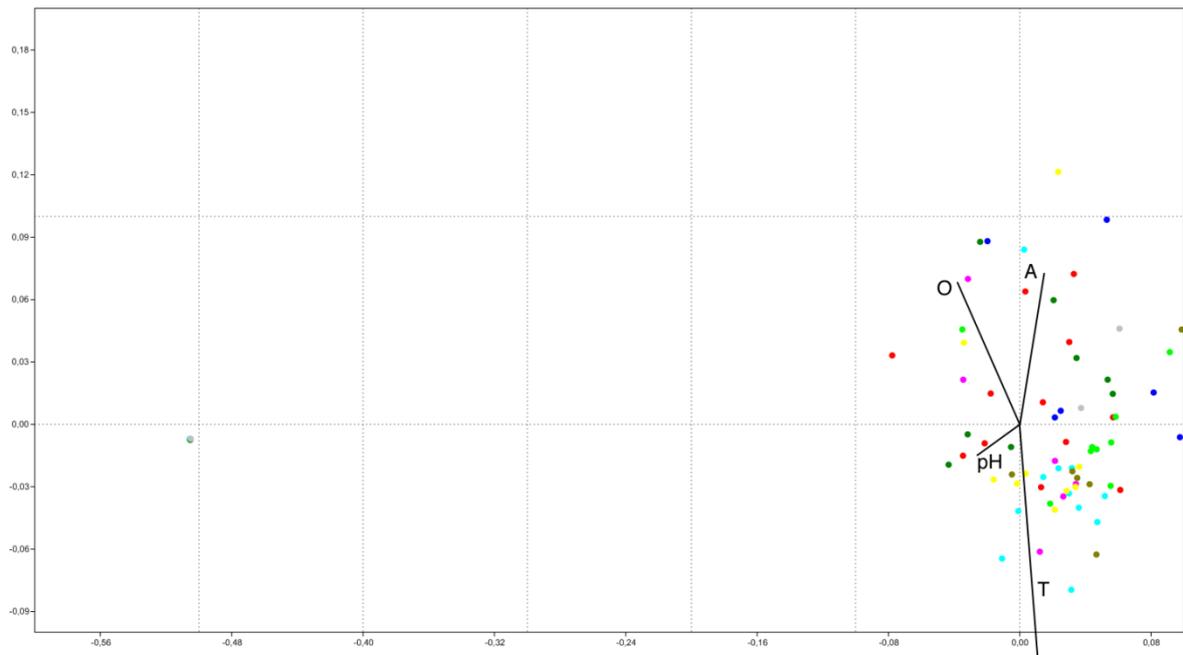
1.5 Analisi statistiche e biogeografia

Per rappresentare i rapporti fra le comunità di ciliati osservate nei diversi siti di campionamento è stata impiegata la tecnica di ordinamento multivariata nMDS utilizzando come dati di partenza matrici di presenza/assenza dei taxa.

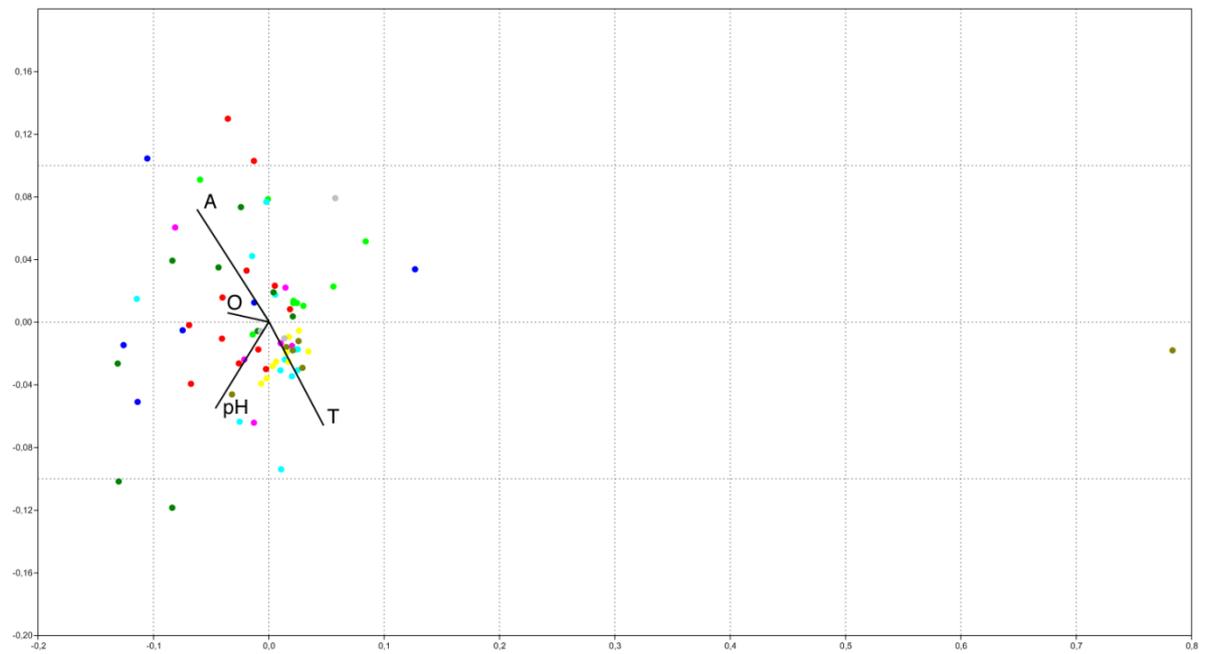
Sia per le matrici costruite con i dati provenienti dalle osservazioni effettuate a t_0 che per quelle calcolate su t_T , e indipendentemente dal sistema di classificazione utilizzato (sistema di classificazione 1: le 54 categorie tassonomiche descritte in precedenza; sistema di classificazione 2: sottoclassi), non è stato possibile evidenziare particolari *pattern* di similarità tra i campioni. Non sono cioè stati osservati cluster corrispondenti a nessuno dei fattori considerati, quali stagione di campionamento (α , β , γ), altitudine (planiziale/montano), idrodinamismo (bacini lentici/bacini lotici), tipo di habitat (naturale/artificiale) o area di campionamento (Figura 12).

Il metodo di ordinamento non è riuscito quindi ad individuare *pattern* caratteristici di biodiversità tra le comunità che condividevano le stesse condizioni ambientali. È possibile che ciò sia dovuto al basso potere di risoluzione tassonomica della nostra analisi morfologica. Inoltre l'analisi multivariata, essendosi basata solo sui dati di presenza/assenza delle singole categorie tassonomiche in ogni campione, manca dell'informazione relativa all'abbondanza dei singoli taxa; purtroppo i dati di abbondanza non sono stati inclusi poiché non è stato effettuato uno studio quantitativo sui taxa identificati.

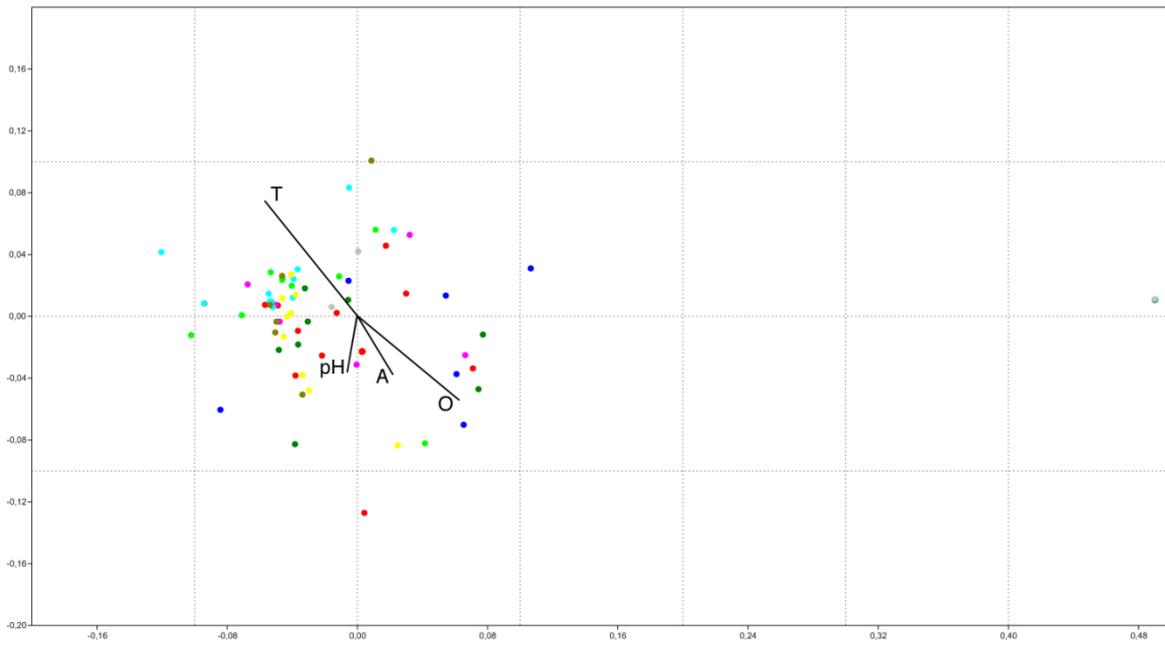
Il risultato ottenuto potrebbe anche essere sintomo di un fenomeno ecologico tipico delle comunità microbiche; gli organismi unicellulari come i ciliati sono significativamente più influenzati dalle caratteristiche ecologiche del loro microhabitat che dai fattori generali dell'area in cui vivono. In altre parole la scala spaziale utilizzata sarebbe troppo grande per osservare specifici *pattern* di biodiversità, poiché le differenze nella composizione tassonomica sono maggiori quando si analizzano comunità appartenenti a diversi microhabitat di una stessa area che fra comunità prelevate da aree diverse.



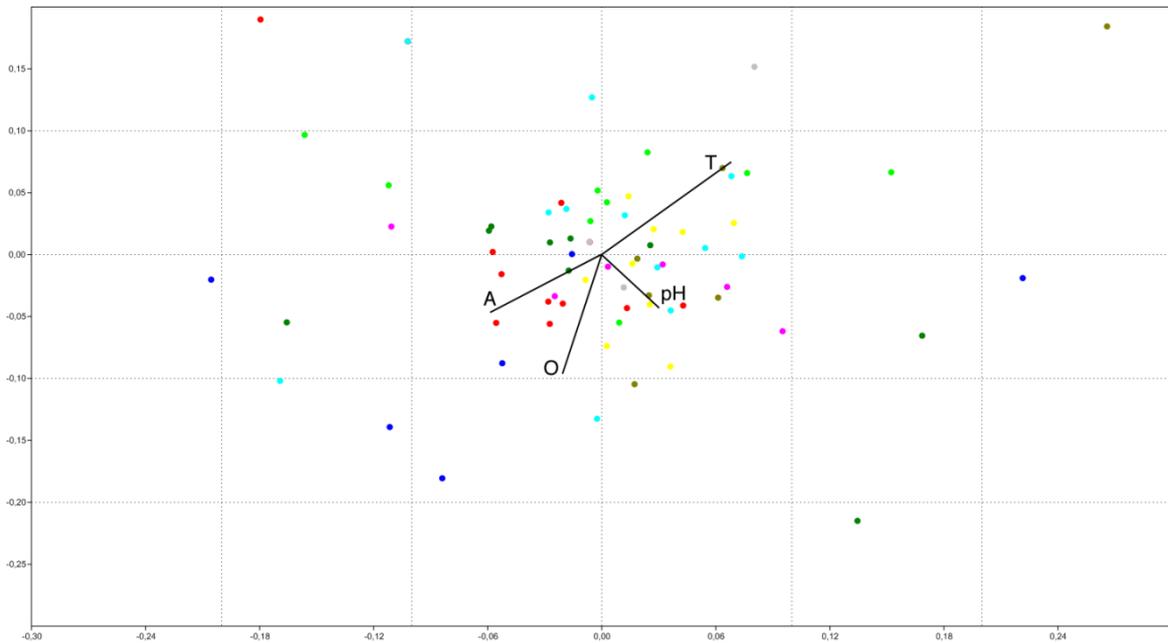
nMDS “sistema di classificazione 1 – osservazioni t_0 ”



nMDS “sistema di classificazione 1 – osservazioni t_T ”



nMDS “sistema di classificazione 2 – osservazioni t_0 ”



nMDS “sistema di classificazione 2 – osservazioni t_T ”

Figura 12. Grafici nMDS calcolati per le 4 diverse matrici. Nei grafici sono mostrati 72 campioni, poiché sono stati esclusi i 3 provenienti dal Lago Nero. Ogni colore corrisponde ad una diversa area di campionamento secondo la legenda sottostante. I vettori mostrano il contributo di ogni variabile ambientale (A=Altitudine; T=Temperatura; O=Ossigeno; pH) all’ordinamento generato.



Figura 13. Legenda colori grafici nMDS

Questo suggerisce quindi che non è possibile prevedere in modo affidabile la composizione di una comunità di ciliati utilizzando descrittori generali dell'area considerata. Quanto appena detto non è però necessariamente vero se si prendono in considerazione singoli taxa; ad esempio nel nostro studio alcuni generi mostrano un modesto grado di endemismo (*Linostomella*, *Disematostoma*, *Chilodontopsis* e i generi affini a *Didinium* sono stati osservati esclusivamente nel Padule di Fucecchio). Altri taxa possono essere influenzati da fattori come la temperatura o il contenuto di nutrienti: ad esempio tutte le popolazioni del genere *Paramecium* sono state osservate solo in ambienti planiziali, eccetto che per una singola specie (*Paramecium bursaria*) trovata anche in due laghi di montagna (Lago del Greppo e Lago Nero). Lo stesso *Paramecium bursaria* è probabilmente sensibile al contenuto di nutrienti poiché, oltre che nei due laghi di montagna di origine glaciale, è stato trovato tra le aree planiziali indagate solo nel Torrente Ombrone, un ambiente lotico che presenta quindi, come i laghi di origine glaciale, una minor concentrazione di nutrienti.

Anche se è difficile da dimostrare in modo assoluto, può anche accadere che certi taxa generalmente molto comuni non siano mai trovati in alcuni biotipi: ad esempio, il genere *Euplotes* molto presente nelle nostre checklist, a volte anche con diverse morfospesie nello stesso campione, non è mai stato osservato nel Laghetto dell'Orto Botanico e nel Torrente Sestaione. Un caso simile si può osservare per il genere *Halteria*, che anche se molto presente nella nostra analisi, non è stato mai osservato nel Torrente Ombrone e nel Lago Nievole di Pesca Sportiva.

Nei grafici nMDS si può notare come la temperatura, l'altitudine e la concentrazione di ossigeno, parametri fra loro correlati, siano le principali variabili che influenzano le comunità di ciliati (vedi lunghezza dei vettori nei 4 grafici).

2. Secondo approccio: Next-Generation Sequencing

2.1 Numero di *contig* e di OTU

L'analisi Illumina effettuata su 12 dei campioni fissati in etanolo ha prodotto un totale di 8.135.868 *paired-end reads*, con una media di 677.989 *reads* a campione. Il valore massimo è stato riscontrato nel campione β Omb2 (989.792), mentre il valore minimo nel campione β PFEU1 (437.893).

I dati grezzi di sequenziamento sono stati elaborati con due diversi pacchetti di software; le analisi "1" e "1bis" sono state eseguite con QIIME mentre la "2" e "2 bis" con il pacchetto USEARCH. Le *pipelines* 1 e 2 hanno utilizzato parametri più permissivi per il filtraggio qualitativo; le *pipelines* 1 bis e 2 bis sono invece quelle più restrittive. In ogni analisi le sequenze sono state filtrate a diversi step; i 5 passaggi principali sono mostrati in figura 14.

Il primo filtro è stato applicato durante l'appaiamento dei *reads*. Nelle *pipeline* 1 e 1 bis effettuate con QIIME, questo è il passaggio in cui è stato scartato il maggior numero di sequenze; in particolare per l'analisi 1 bis è stato eliminato fino al 40,4% dei dati originali. Per le analisi effettuate con USEARCH, durante l'assemblaggio è stato eliminato solo il 7,1% dei *reads* sulla stessa *pipeline*.

Il filtraggio qualitativo delle sequenze basato sui valori dei *quality score* dei singoli nucleotidi dell'intera sequenza è stato il passaggio che ha scartato il maggior numero di *contig* nelle *pipeline* 2 e 2 bis; in particolare nell'analisi più restrittiva durante questo step è stato eliminato un ulteriore 43,6% delle sequenze originali.

In tutte e 4 le *pipeline*, nei passaggi successivi è stata scartata solo una piccola percentuale delle sequenze (*contig*). Durante il terzo step sono state rimosse le sequenze in cui non erano riconoscibili i primer utilizzati durante la prima PCR, ed i *contig* con una lunghezza inferiore a 250 pb (in media l'1,25% dei *contig*).

Una volta identificate le OTU (gruppi di *contig* che condividono almeno il 97% di similarità di sequenza) sono stati applicati due ulteriori filtri. La ricerca delle chimere ha rimosso dal 25 al 28% di OTU rappresentate da sequenze spurie contenenti frammenti di diversa derivazione; tali OTU sono però costituite in media solo dal 6% dei *contig*, e sono quindi

molto meno rappresentate della media. Simili percentuali di sequenze chimeriche sono state trovate anche in lavori di caratterizzazione di comunità fungine da campioni di suolo fatte con tecnologia Illumina (Schimidt et al., 2013). Un recente studio mostra come la maggior complessità tassonomica di un campione ambientale incida sulla presenza di tali sequenze (Fonseca et al., 2012).

Il *denoising* ha influito maggiormente sul numero totale di OTU; con valori simili per le 4 *pipeline*, le OTU costituite da meno di 11 sequenze in ogni campione rappresentano infatti circa il 69,1% delle OTU non chimeriche (ma anche queste OTU sono poco rappresentate, poiché costituiscono solo l'1,6% dei contig conservati fino a questo step dell'analisi).

Per ognuna delle 4 analisi è stato così ottenuto il numero totale di *contig* di buona qualità:

<i>Pipeline</i>	Contig totali di buona qualità	% rispetto ai dati grezzi
1	5.511.766	67,7 %
2	4.775.886	58,7 %
1 bis	4.187.502	51,5 %
2 bis	3.590.679	44,1 %

La maggior parte delle sequenze sono quindi state scartate durante l'assemblaggio dei *reads*, a causa della presenza di una zona di sovrapposizione troppo corta e/o con troppi *mismatch*, e durante il filtraggio qualitativo dove vengono valutate le sequenze in base ai *quality score* delle singole basi.

I primer utilizzati durante la prima PCR sono stati disegnati in modo da amplificare in maniera ottimale la regione V4 dal maggior numero di specie di ciliati a vita libera, ma non come primer esclusivi per tale gruppo. Di conseguenza, molte sequenze ottenute non appartengono a protisti ciliati.

Il numero di *contig* che condividono almeno il 90% di similarità con sequenze del phylum dei ciliati è risultato rispettivamente:

<i>Pipeline</i>	<i>Contig di buona qualità solo di Ciliati</i>	<i>% rispetto ai contig totali di buona qualità</i>
1	2.308.400	41,9 %
2	2.102.926	44,0 %
1 bis	1.677.379	40,0 %
2 bis	1.620.504	45,1 %

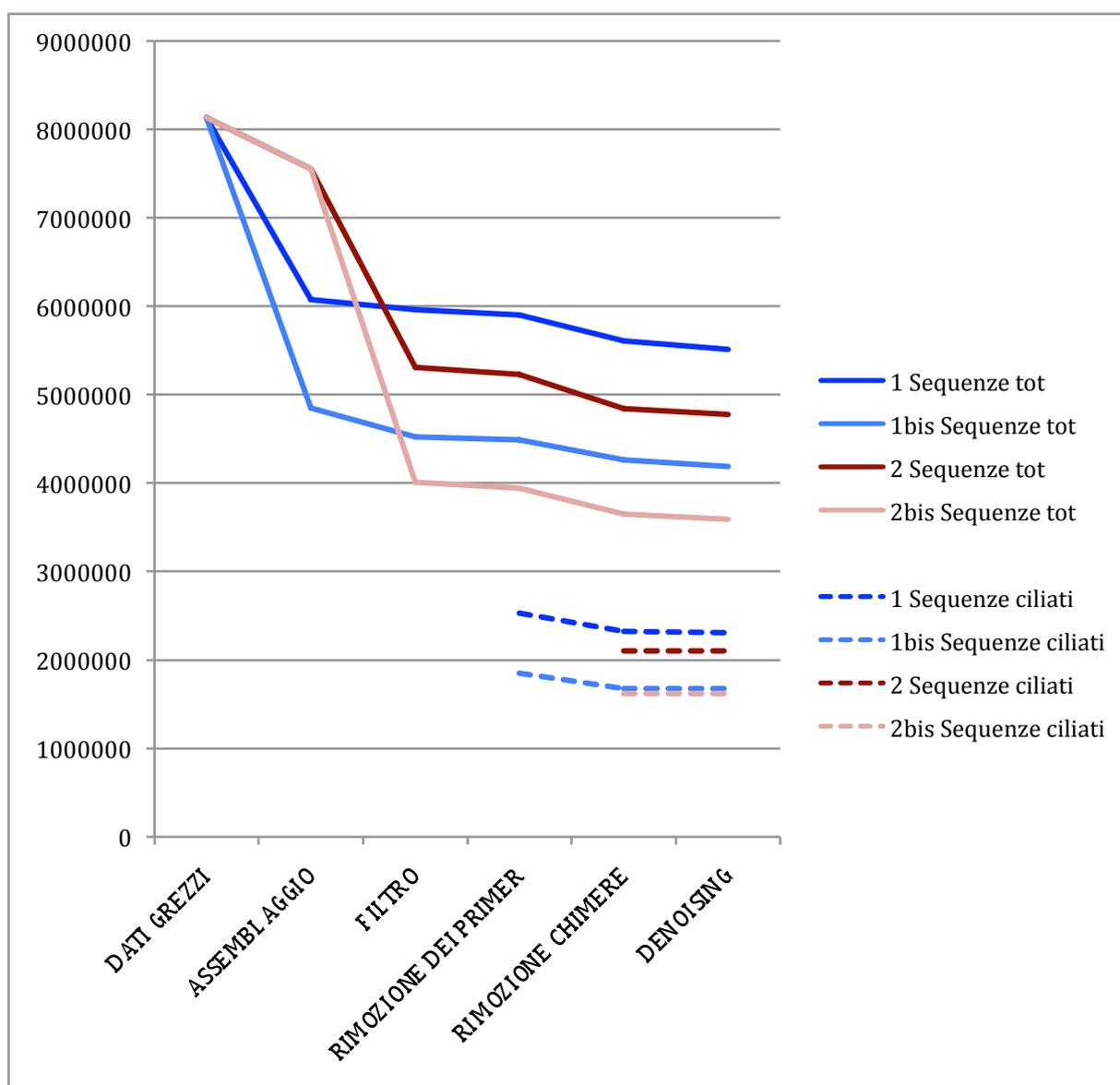


Figura 14. Nella figura è indicato sulle ordinate il numero di sequenze risultanti dopo ogni passaggio.

La percentuale di OTU di ciliati rimosse dal denoising in questo caso differisce fortemente nei valori fra le analisi fatte con i due differenti pacchetti di *software* (come si può notare anche dalla figura 15). Per l'analisi 1 e 1 bis eseguite con QIIME, sono state eliminate rispettivamente il 68,3% e il 55,7% delle OTU (0,58% - 0,40% dei *contig*), mentre per le analisi 2 e 2 bis effettuate con USEARCH, sono state eliminate solo il 24,2% e il 25,6% delle OTU (0,24% - 0,29% dei *contig*).

Probabilmente in queste due ultime analisi, molte delle OTU costituite da meno di 11 sequenze erano già state eliminate nei passaggi precedenti, poiché contenenti *contig* di bassa qualità. I valori assoluti delle OTU finali, infatti, tendono a convergere.

Questo dato conferma in parte quanto viene riportato da alcuni autori. La maggior parte delle OTU costituite da poche sequenze (1-10 *contig*), non rappresenterebbe infatti quella parte di biodiversità definita rara ("*rare biosphere*") (Sogin et al., 2006), cioè taxa rappresentati da pochi individui, ma sarebbe in larga parte frutto di errori durante l'amplificazione e il sequenziamento (Behnke et al., 2011; Bachy et., al 2013; Wang et al., 2013).

Tali errori possono portare ad individuare un numero eccessivo di OTU e quindi ad una sovrastima della reale ricchezza all'interno dei campioni.

L'analisi 2 bis è quella più restrittiva poiché mantiene il minor numero di dati, meno della metà dei *reads* originari (44,1%); inoltre allo stesso tempo è quella che contiene percentualmente il maggior numero di sequenze di organismi ciliati (45,1%).

Aumentando la stringenza dell'analisi, la maggior parte delle sequenze scartate e quindi di minore qualità sono quindi probabilmente afferenti ad organismi non ciliati. Questo dato suggerisce che i primer, pur non essendo esclusivi per i ciliati, hanno fornito sequenze più attendibili per questo gruppo rispetto ad altri.

Il numero di OTU di ciliati fornite rispettivamente da ciascuna analisi è:

<i>Pipeline</i>	OTU di ciliati	% rispetto alle OTU totali
1	1.576	21,0 %
2	876	14,9 %
1 bis	996	17,2 %
2 bis	800	16,4 %

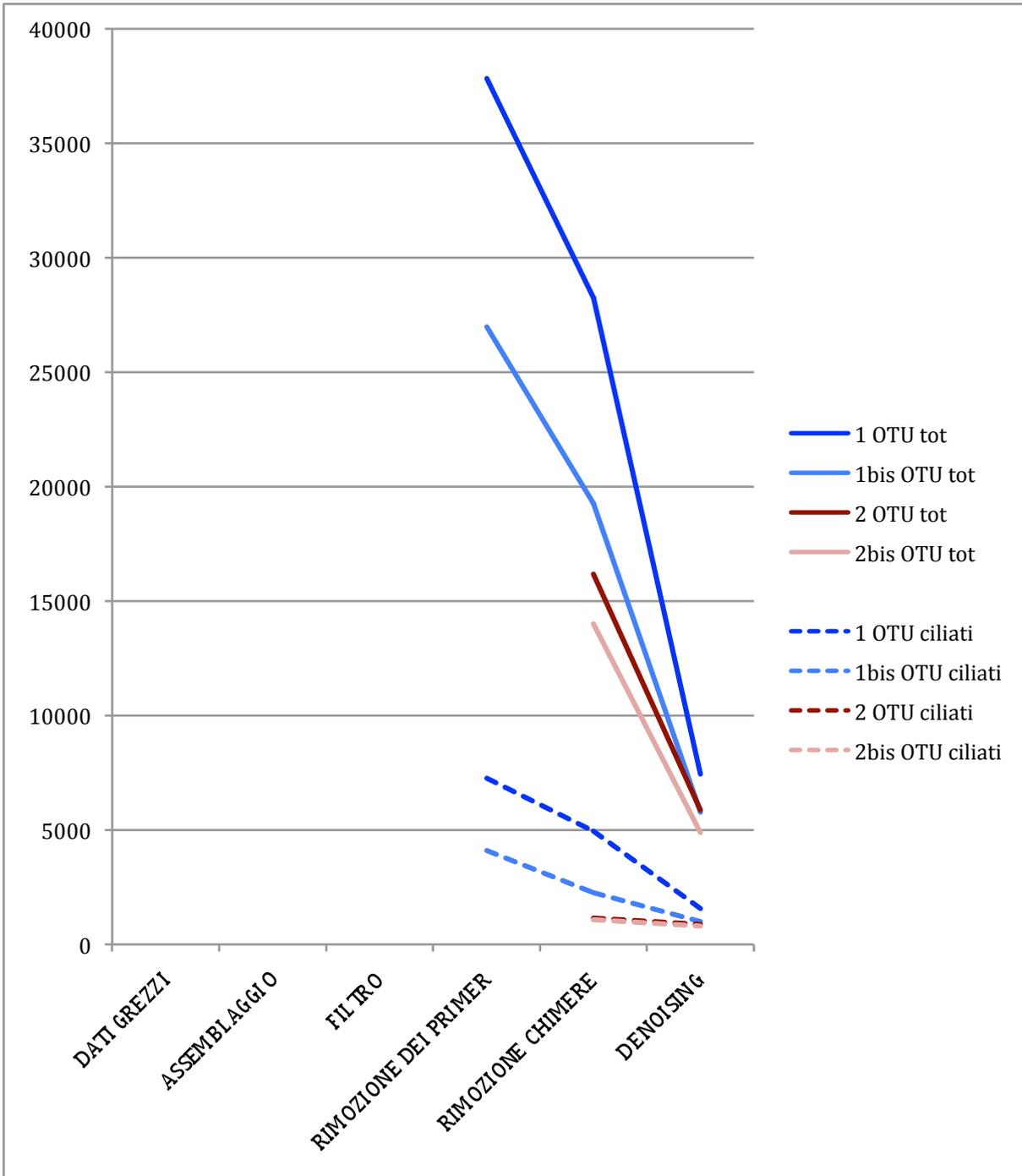


Figura 15. Nella figura è indicato sulle ordinate il numero di OTU risultanti dopo ogni passaggio.

2.2 Analisi *pipeline 2 bis*

Per lo studio della composizione tassonomica abbiamo deciso di utilizzare i dati provenienti dalla *pipeline 2 bis* poiché, oltre a fornire il minor numero di OTU totali (4.881) e in particolare di ciliati (800), si tratta dell'analisi dove la probabilità di trovare errori di sequenziamento è minore. Inoltre, l'analisi comparativa (Figura 15) mostra una tendenza delle diverse *pipeline* a convergere attorno a questo valore; solo la *pipeline 1* ha prodotto un numero di OTU di ciliati considerevolmente maggiore.

Il numero di OTU di ciliati trovate in questa *pipeline* corrisponde al 16,4% delle OTU complessive (percentuali simili sono state trovate anche nelle altre 3 analisi). Queste OTU rappresentano però il 45,1% dei *contig* totali, indicando come in termini di abbondanza i ciliati costituiscano una larga parte dei dati ottenuti.

Come possiamo infatti vedere dal grafico (Figura 16), in quasi tutti i campioni analizzati il gruppo dei ciliati è quello che contiene il maggior numero di *contig*, con una percentuale massima del 63,2% nel campione γ Acq1 e una minima del 28,1% nel campione β LdG2.

Gli altri gruppi di organismi più rappresentati sono rispettivamente:

- il gruppo dei Funghi con il 21,2% dei *contig* totali
- le OTU “*unassigned*” con il 17,3% dei *contig* (quelle cioè che hanno una similarità inferiore al 90% con qualsiasi sequenza contenuta nel database di riferimento SILVA 119 utilizzato per l'identificazione tassonomica)
- il gruppo dei Cercozoa con il 9,2% dei *contig*
- il gruppo dei Dinoflagellata con il 3,0% dei *contig*

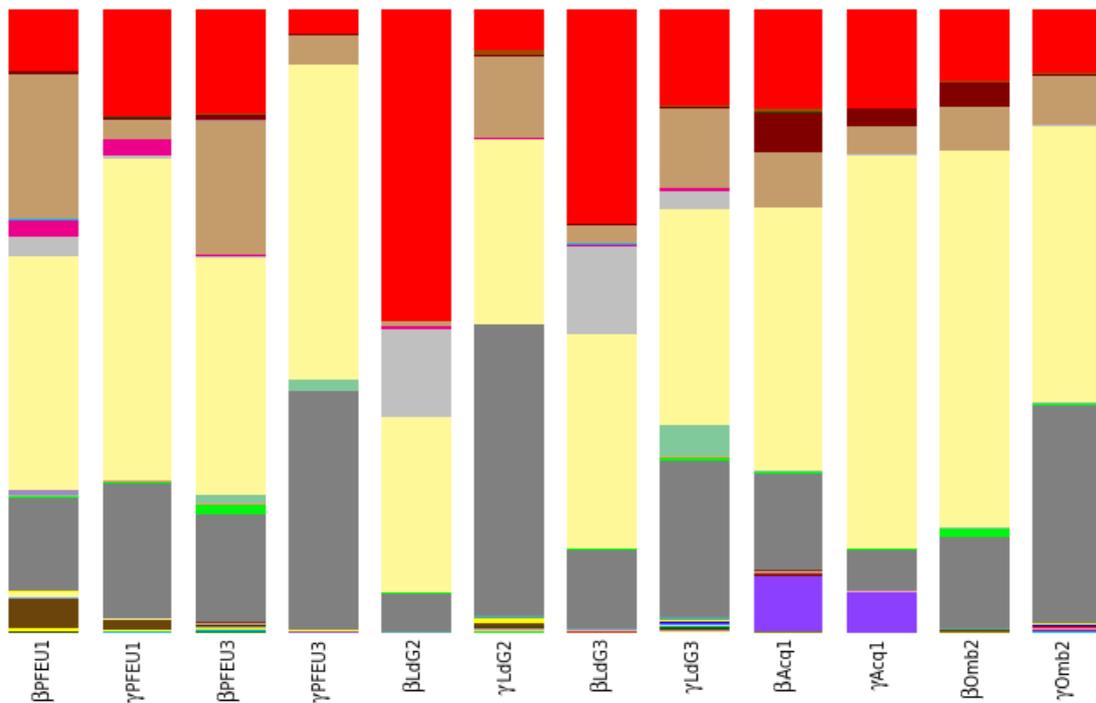


Figura 16. I 12 campioni dell'analisi Illumina per cui è mostrata la percentuale di sequenze di buona qualità per ogni gruppo tassonomico individuato.

Legenda	Tassonomia
	D_0_Eukaryota;D_1_Archaeplastida;D_2_Chloroplastida;D_3_Chlorophyta
	D_0_Eukaryota;D_1_Archaeplastida;D_2_Chloroplastida;D_3_uncultured
	D_0_Eukaryota;D_1_Centrohelida;D_2_Acanthocystidae;D_3_Pterocystis
	D_0_Eukaryota;D_1_Centrohelida;D_2_Heterophryidae;Other
	D_0_Eukaryota;D_1_Cryptophyceae;D_2_Cryptomonadales;D_3_Cryptomonas
	D_0_Eukaryota;D_1_Cryptophyceae;D_2_Kathablepharidae;D_3_uncultured
	D_0_Eukaryota;D_1_Cryptophyceae;D_2_Cryptomonadales;D_3_Cryptomonas
	D_0_Eukaryota;D_1_Cryptophyceae;D_2_Kathablepharidae;D_3_uncultured
	D_0_Eukaryota;D_1_Opisthokonta;D_2_Holozoa;D_3_Choanomonada
	D_0_Eukaryota;D_1_Opisthokonta;D_2_Holozoa;D_3_Ichthyosporia
	D_0_Eukaryota;D_1_Opisthokonta;D_2_Holozoa;D_3_Metazoa
	D_0_Eukaryota;D_1_Opisthokonta;D_2_Nucleotmycea;D_3_Fungi
	D_0_Eukaryota;D_1_Opisthokonta;D_2_Nucleotmycea;D_3_LKM11
	D_0_Eukaryota;D_1_SAR;D_2_Alveolata;D_3_Apicomplexa
	D_0_Eukaryota;D_1_SAR;D_2_Alveolata;D_3_BOLA914
	D_0_Eukaryota;D_1_SAR;D_2_Alveolata;D_3_Ciliophora
	D_0_Eukaryota;D_1_SAR;D_2_Alveolata;D_3_Dinoflagellata
	D_0_Eukaryota;D_1_SAR;D_2_Alveolata;D_3_Protalveolata
	D_0_Eukaryota;D_1_SAR;D_2_Rhizaria;D_3_Cercozoa
	D_0_Eukaryota;D_1_SAR;D_2_Stramenopiles;D_3_Ochrophyta
	D_0_Eukaryota;D_1_SAR;D_2_Stramenopiles;D_3_Peronosporomycetes
	D_0_Eukaryota;Other;Other;Other
	Unassigned;Other;Other;Other

Figura 17. Legenda colori; ogni colore rappresenta un diverso gruppo tassonomico. Tassonomia basata sul database di riferimento per gli eucarioti SILVA 119. Le sequenze che avevano una similarità di sequenza minore del 90% con tali referenze sono state rinominate come "Unassigned".

Il numero di OTU di ciliati trovato in ogni campione varia da un minimo di 115 OTU nel campione β LdG3 ad un massimo di 235 nel campione γ Omb2, con una media di 163,5 OTU a campione. Per quanto riguarda i *contig* di ciliati in ogni campione, si ha un minimo in β PFEU1 con 57.537 e un massimo nel campione β Omb2 con 310.269 *contig*; la media è di circa 135.042 *contig* per campione.

Le curve di rarefazione calcolate per ogni campione (Figura 18) mostrano come lo sforzo di sequenziamento (misurato in numero di contig di buona qualità ottenuti) sia stato più che adeguato. Tutte le curve raggiungono un plateau stabile molto presto, confermando che un numero maggiore di sequenze non avrebbe aumentato il numero di OTU ottenute. Questo ci permette di comparare tra loro in maniera affidabile i campioni.

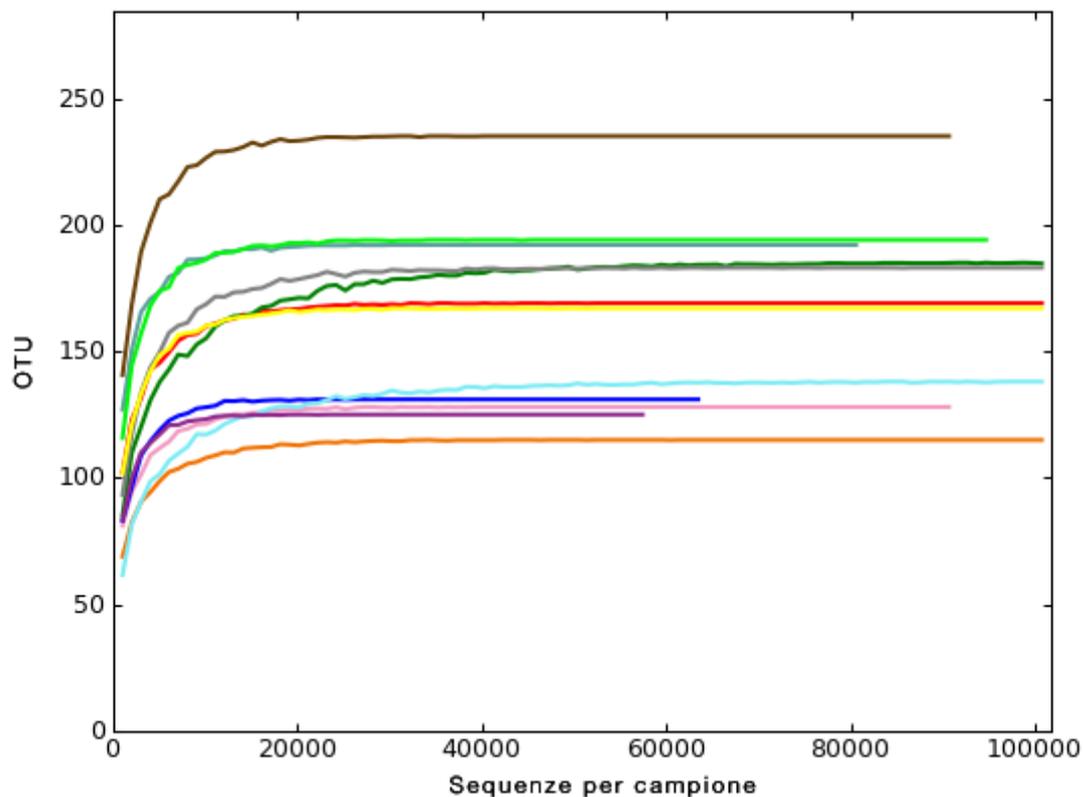


Figura 18. Curve di rarefazione relative ai 12 campioni Illumina. Sulle ordinate sono indicate il numero di OTU di ciliati individuate mentre sulle ascisse il numero di sequenze di ciliati.

■	β Acq1	■	β PFEU1
■	γ Acq1	■	γ PFEU1
■	β LdG2	■	β Omb2
■	γ LdG2	■	γ Omb2
■	β LdG3	■	β PFEU3
■	γ LdG3	■	γ PFEU3

Figura 19. Legenda colori delle curve di rarefazione

2.2.1 Composizione tassonomica

La composizione tassonomica totale delle 800 OTU di ciliati individuate dalla *pipeline 2 bis* e dei rispettivi *contig*, è mostrata nelle figure 20, 21, 22, 23.

Più di un quarto delle OTU (26,4%) appartiene alla classe degli Oligohymenophorea, mentre quella degli Spirotrichea è rappresentata dal 20,7% delle OTU. Le abbondanze di queste classi, misurate in numero di sequenze, sono però maggiori: rispettivamente il 31,2% per gli Oligohymenophorea e il 27% per gli Spirotrichea. Le OTU di queste due classi sono quindi rappresentate da più sequenze rispetto a quelle di altri gruppi come ad esempio i Litostomatea: tale classe, la seconda più rappresentata come numero di OTU (21,5%) è costituita solo dal 16,2% dei *contig*. Lo stesso si può affermare anche per le classi Phyllopharingea (6,1% delle OTU contro il 3,6% dei *contig*), Nassophorea (4,1% delle OTU contro lo 0,55% dei *contig*), e Prostomatea (4,9% delle OTU contro il 2,6% dei *contig*).

Secondo tali risultati possiamo sostenere che le OTU di Oligohymenophorea, Spirotrichea ma anche dei Colpodea, essendo rappresentate da un maggior numero di *contig* rispetto a quelle di altri gruppi, siano le OTU dominanti cioè i taxa rappresentati all'interno dei diversi campioni da un maggior numero di individui.

Occorre però ricordare che le abbondanze rappresentate dal numero di sequenze in questo tipo di analisi sono potenzialmente inficiate da un elevato numero di elementi distortenti. L'alta variabilità nel numero di copie del gene codificante per l'rRNA 18S tra i diversi taxa, ad esempio, può condizionare anche di molto i reali valori di abbondanza di ogni OTU (Dunthorn et al., 2014; Medinger et al., 2010; Stoeck et al., 2014). Come riportato da Prescott

(1994) e Riley & Katz (2001), durante la maturazione del macronucleo dei ciliati non tutte le sequenze vengono amplificate nella stessa misura e i geni per gli RNA ribosomali sono tra quelli che subiscono il maggior numero di replicazioni in misura disomogenea tra le diverse specie; si possono avere centinaia di copie in alcuni taxa (es. Colpodea, Litostomatea, Prostomatea e Oligohymenophorea) e più di mille in altri (es. Phyllopharingea e Spirotrichea).

Anche la dimensione delle cellule può incidere sul numero di copie di tale marcatore (Zhu et al., 2005); dai piccoli Scuticociliatia si passa a organismi come quelli appartenenti ai generi *Stentor* o *Spirostomum* nei quali la grande dimensione (1 o più mm) è accompagnata anche da una maggior quantità di DNA (macronucleo più grande o più macronuclei) e quindi un elevato numero di copie di ciascun gene. Un altro potenziale effetto di disturbo può dipendere anche da un'efficienza non omogenea delle reazioni di amplificazione per taxa diversi.

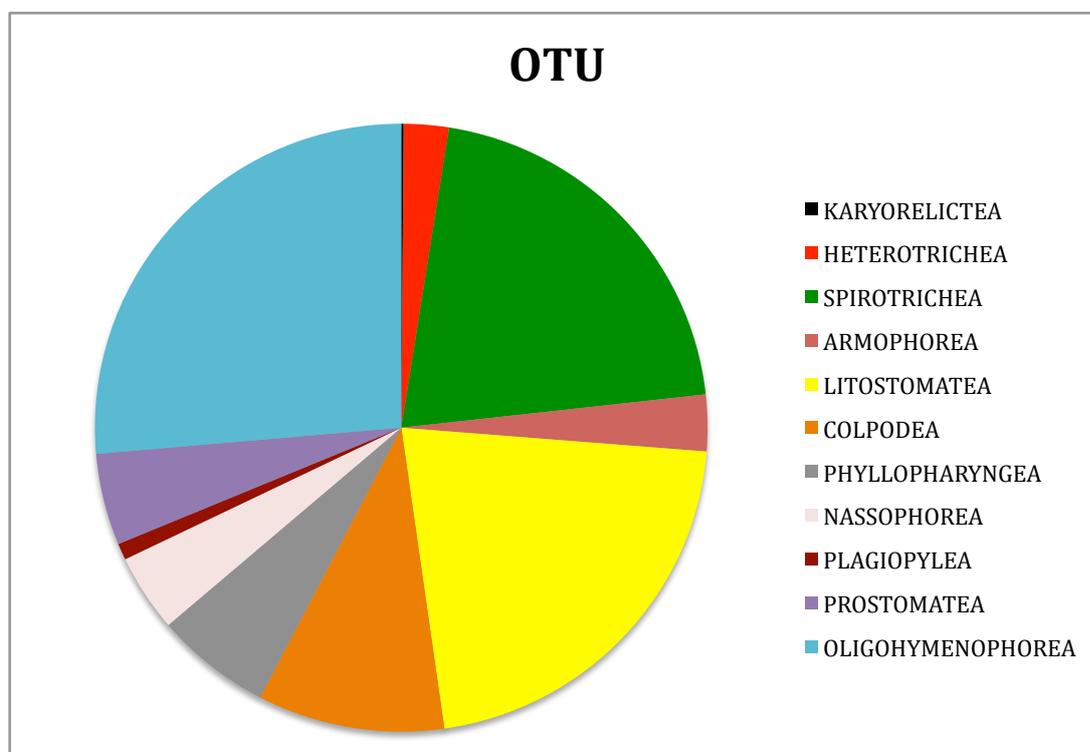


Figura 20. OTU totali di ciliati divise per classe.

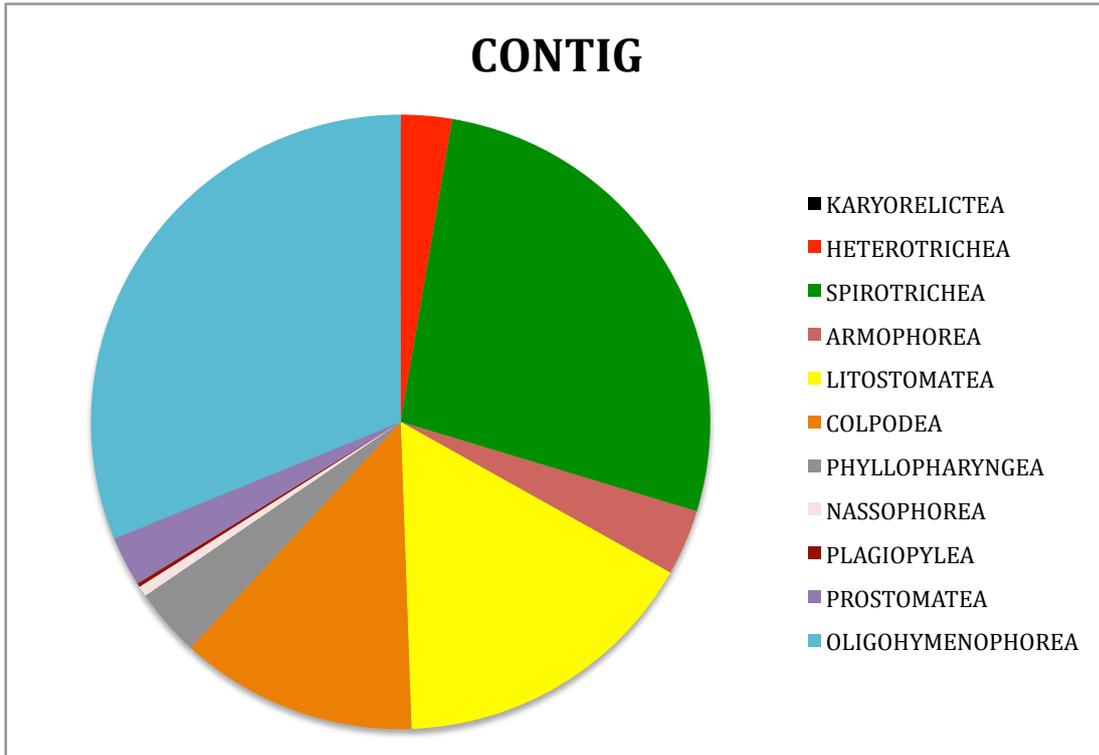


Figura 21. Contig totali di ciliati divisi per classe.

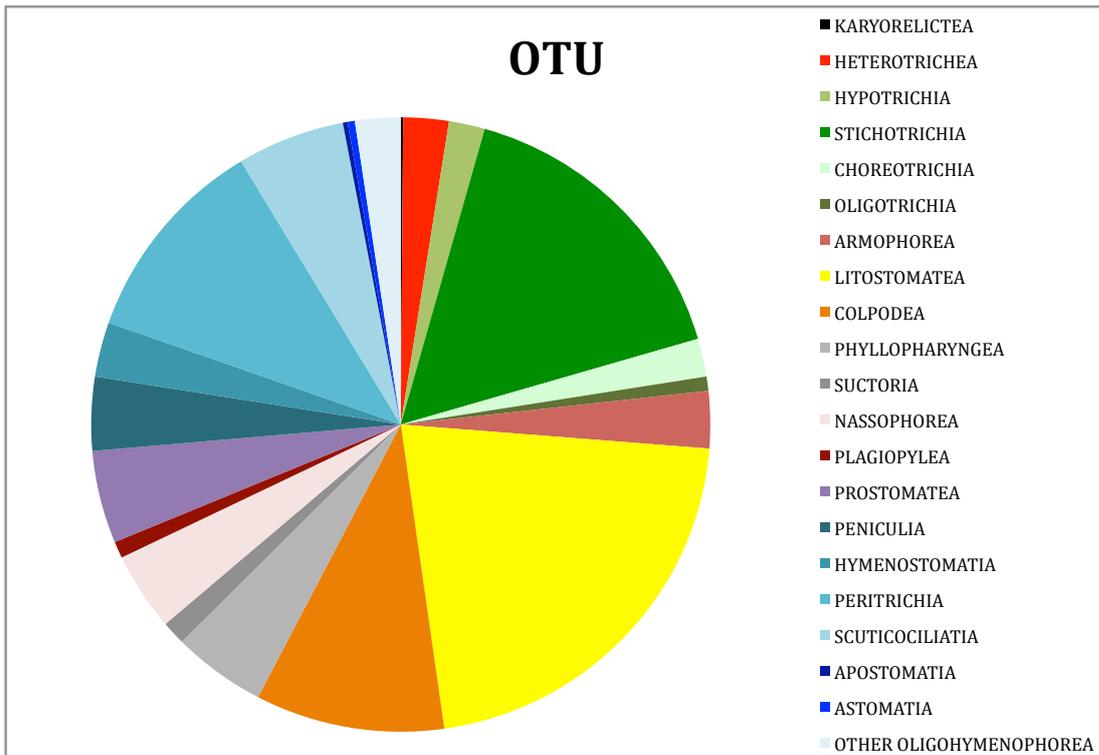


Figura 22. OTU totali di ciliati divise per classe e sottoclasse (dove presente).

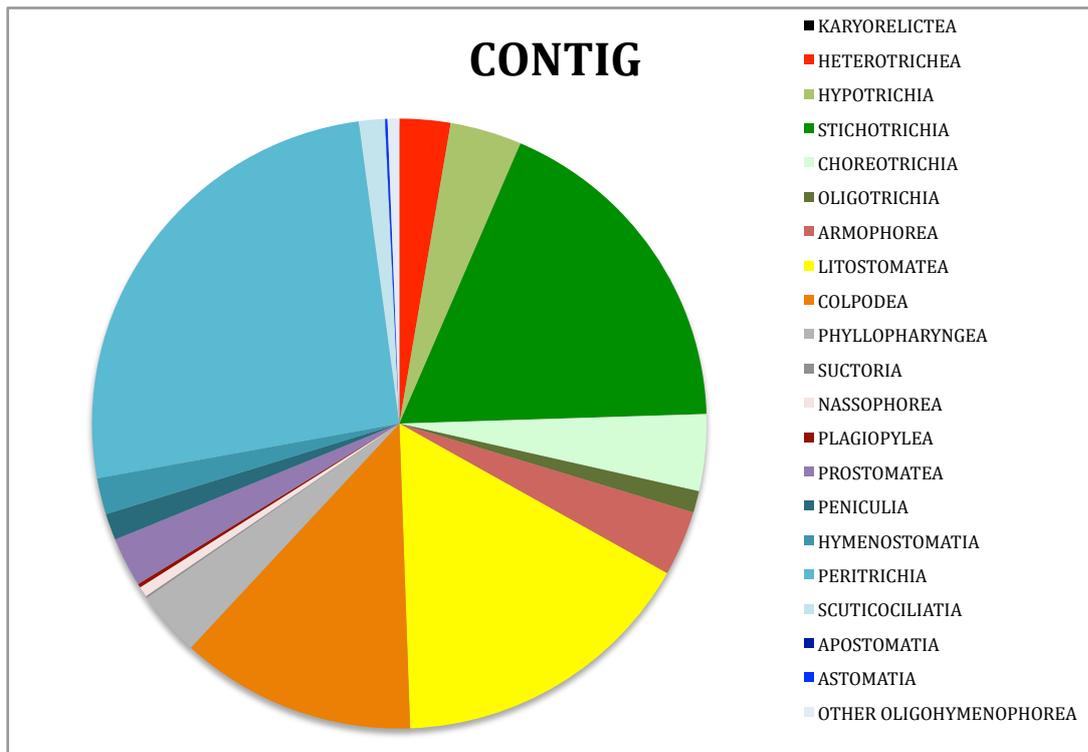


Figura 23. Contig totali di ciliati divise per classe e sottoclasse (dove presente).

3. Confronto dei due approcci

I risultati ottenuti dalla caratterizzazione dei campioni sottoposti ad analisi Illumina sono stati confrontati con i risultati ottenuti mediante il I approccio dai campioni provenienti dalle 12 medesime stazioni di campionamento.

3.1 Numero di OTU

I risultati ottenuti con i due differenti approcci sono stati confrontati in base al numero di OTU riscontrate in ogni campione (Tabella 7); per l'approccio I le OTU corrispondono ai diversi morfotipi osservati in ciascun campione mentre per l'approccio II si parla di OTU molecolari, ovvero di gruppi di *contig* che condividono una similarità di sequenza maggiore o uguale al 97%.

Come possiamo vedere dalla tabella 7 e dalla figura 24, c'è una netta differenza nel numero di OTU trovate con i due diversi metodi, corrispondente a circa un ordine di grandezza. La media del numero di OTU morfologiche è infatti di 16 diversi morfotipi a campione mentre con il secondo approccio si ha una media di 163 diverse OTU a campione.

	OTU morfologiche	OTU Illumina
β PFEU1	8	125
γ PFEU1	32	183
β PFEU3	7	167
γ PFEU3	19	194
β LdG2	20	131
γ LdG2	6	128
β LdG3	22	115
γ LdG3	18	192
β Acq1	16	169
γ Acq1	22	138
β Omb2	17	185
γ Omb2	6	235

Tabella 7. Per ogni campione sono mostrati il numero di morfotipi individuati con il primo approccio e il numero di OTU Illumina trovate con il secondo approccio

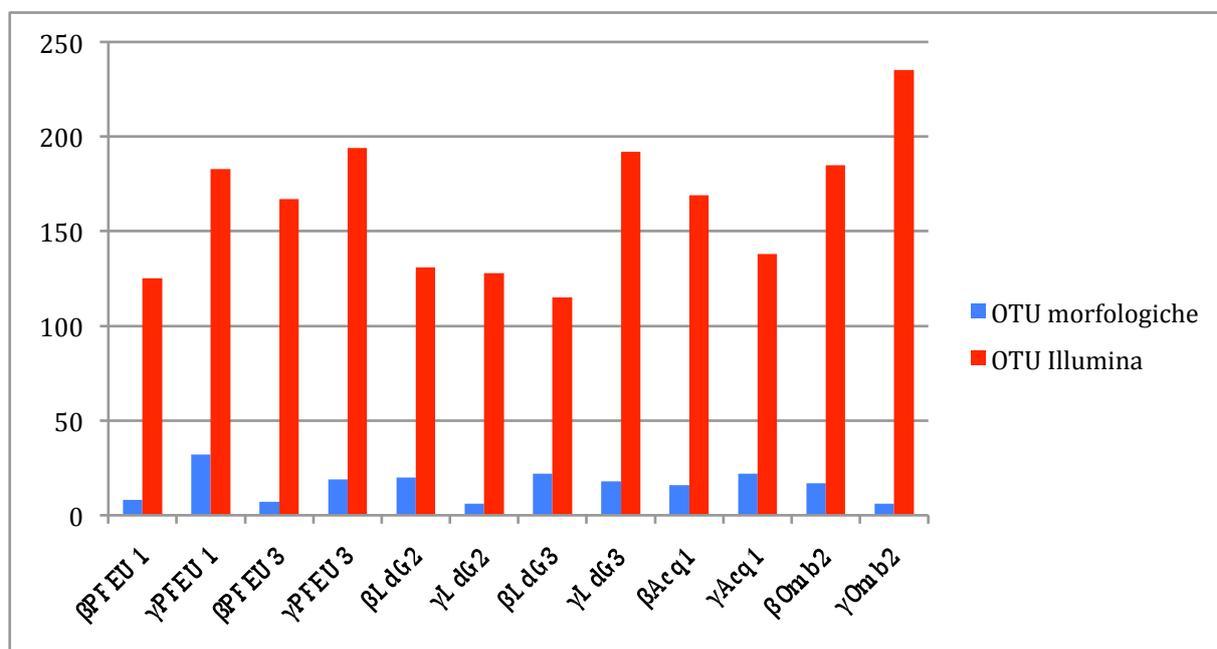


Figura 24.

Non era inaspettato che il sequenziamento Illumina trovasse molti più taxa rispetto al numero individuato tramite il primo approccio; valori simili sono riportati anche in altri studi che utilizzano sia l'analisi NGS (in particolare, la piattaforma 454) che un'analisi morfologica per la caratterizzazione di comunità eucariotiche (Bachy et al., 2013; Medinger et al., 2010; Stoeck et al., 2014).

I motivi di tale discrepanza nel numero di OTU trovate sono molteplici:

- In primo luogo, il potere di risoluzione di un'analisi molecolare che utilizza come marcatore la regione più variabile del gene per l'rRNA 18S, il frammento V4 (Van de Peer et al., 1994), è maggiore rispetto a quello del primo approccio dove i diversi morfotipi vengono individuati in vivo solo attraverso l'osservazione al microscopio. Ad esempio, come verrà successivamente mostrato in dettaglio, per alcuni gruppi tassonomici come gli Heterotrichea il numero di OTU individuate con i due approcci è molto simile poiché le diverse morfospecie sono ben note e relativamente semplici da distinguere anche ad una prima osservazione. Al contrario, per altri gruppi come Stichotrichia, Peritrichia e Litostomatea, i morfotipi sono difficili da discriminare a piccolo ingrandimento; due diverse morfospecie di stichotrichi possono differire anche solo per il numero di cirri e quindi, senza l'utilizzo di specifiche colorazioni o osservazioni al microscopio elettronico, non è semplice distinguerle. Un problema simile riguarda gli organismi di piccole dimensioni, come gli Scuticociliatia (poche decine di micrometri), per i quali la sola osservazione al microscopio non permette talvolta di rilevare la presenza o l'identificazione dei diversi morfotipi.

- Nonostante il primo approccio abbia previsto per ogni campione l'osservazione ripetuta nel tempo di tre diverse aliquote arricchite, probabilmente non tutti i ciliati presenti sottoforma di cisti o rappresentati da pochi individui hanno trovato le condizioni idonee e avuto quindi la possibilità di svilupparsi ed essere così osservati. Tali organismi non individuati con il primo approccio contribuiscono invece al numero di OTU molecolari trovate; infatti i comuni protocolli di estrazione del DNA funzionano anche su organismi incistati (Medinger et al., 2010) e la tecnologia Illumina, avendo un elevato *throughput* in termini di numero di sequenze, è in grado di individuare anche quella parte di biodiversità definita rara, cioè

costituita da taxa rappresentati da pochi individui, che molto spesso sfuggono all'osservazione.

- Inoltre come si può osservare dalle figure 22 e 23, il secondo approccio riesce a rilevare anche la presenza di organismi ciliati simbiotici e/o parassiti come i membri appartenenti al gruppo degli Astomatia, Apostomatia (Oligohymenophorea), dei Suctoria (Phyllopharingea) e dei Clevellandellida (Armophorea). Attraverso il primo approccio, per nessuno dei 12 campioni è stato possibile riscontrare morfotipi appartenenti a tali gruppi; solo per uno dei campioni non indagati con tecnologia Illumina è stato possibile individuare un organismo epibionte appartenente alla sottoclasse dei Suctoria; per la maggior parte di questi ciliati l'individuazione tramite osservazioni al microscopio è molto difficile. In totale l'analisi Illumina ha individuato 16 diverse OTU fra organismi parassiti e simbiotici, distribuite in numerosi campioni.

La maggior quantità di OTU riscontrate grazie al II approccio appare quindi un dato realistico; ciononostante l'elevato numero può essere causato anche da una serie di artefatti:

- come già detto in precedenza, i dati provenienti da analisi NGS (Next-Generation Sequencing) possono contenere numerosi errori, generati durante il processo di amplificazione e di sequenziamento. La formazione di sequenze chimeriche e inserzione o sostituzione di basi sono gli errori riscontrati più frequentemente. Benché i diversi filtri applicati nell'analisi bioinformatica mirino a scartare la maggior parte delle sequenze contenenti tali artefatti, non esiste ancora un protocollo standardizzato per l'analisi di queste tipologie di dati. Non dobbiamo dimenticare che le varie piattaforme NGS, come osservato da Schloss e collaboratori (2011), sono state sviluppate inizialmente per il sequenziamento dei genomi; i tassi di errore per un genoma assemblato sono bassi a causa della sovrapposizione di moltissime *reads* in ciascun punto della sequenza, mentre il tasso di errore per una singola *read* può essere invece anche molto alto. Ne consegue che l'affidabilità di ciascuna sequenza rappresentativa è proporzionale al numero di *contig* appartenenti a quella OTU.

Benché sia stata scelta la pipeline con i filtri più stringenti, la discrepanza nel numero di OTU osservate nei due approcci è comunque rilevante.

Tali errori possono portare ad individuare un numero eccessivo di OTU e quindi ad una sovrastima della reale ricchezza all'interno dei campioni.

- Un altro fattore che può portare ad individuare un maggior numero di OTU è la presenza nei sedimenti di DNA extracellulare rilasciato da organismi ciliati ormai non più presenti. La persistenza di DNA nei sedimenti è riportata da numerosi autori (Nielsen et al., 2007, Vlassov et al., 2007) e data la piccola taglia del marcatore V4 preso in esame nel nostro studio, è molto probabile che una parte del DNA extracellulare presente nel sedimento possa essere stato amplificato e sequenziato.

Anche se l'approccio molecolare ha trovato per tutti e 12 i campioni un maggior numero di OTU, un piccolo numero di morfotipi e sequenze ottenute con il primo approccio non è stato riscontrato nei risultati Illumina. Tali bias potrebbero ricondursi ad un'inefficienza del processo di estrazione del DNA su tali organismi (Delmont et al., 2011; Zhao et al., 2012) oppure ad una amplificazione non ottimale del loro frammento V4, dovuta alla presenza di *mismatch* fra i primer e la sequenza *target* (Sipos et al., 2007; Schloss et al., 2011). Inoltre le sequenze potrebbero essere state amplificate e sequenziate, ma eliminate durante l'analisi bioinformatica poiché di bassa qualità.

3.2 Composizione tassonomica

La composizione tassonomica delle osservazioni totali trovate con il primo approccio deve essere confrontata non con il numero totale di OTU Illumina ma con il numero totale di osservazioni molecolari che è dato dalla somma delle OTU Illumina trovate nei diversi campioni.

Confrontando tali dati (Tabella 8) si può notare come per tutte le classi di ciliati il numero di osservazioni molecolari è sempre maggiore rispetto al numero di osservazioni morfologiche, eccetto per il gruppo dei Karyolicea.

Per tutte le altre classi, benché il primo approccio abbia trovato sempre un numero minore di taxa, il rapporto osservazioni morfologiche/osservazioni molecolari varia notevolmente all'interno dei diversi gruppi tassonomici.

Tale rapporto è molto basso (minore del 10%), oltre che per i Plagiopylea, per i quali attraverso il primo approccio non è stato riscontrato nessun morfotipo a fronte delle 15 osservazioni Illumina, anche per le classi Colpodea, Armophorea, Phyllopharyngea, Nassophorea e Litostomatea.

Per i Colpodea sono state individuate solo 2 popolazioni rispetto alle 210 osservazioni molecolari individuate dall'approccio Illumina. A tale gruppo appartengono numerosi taxa, la maggior parte dei quali presenta nel ciclo vitale uno stadio di quiescenza o resistenza (Lynn, 2008). Visto la grande differenza riscontrata fra numero di osservazioni morfologiche e molecolari, probabilmente molti di questi organismi erano presenti nei diversi campioni sottoforma di cisti che solo l'analisi Illumina è riuscita a rilevare, nonostante i campioni fossero stati osservati per 35 giorni.

Tutte le specie conosciute del gruppo degli Armophorea sono rappresentate da organismi anaerobi o microaerofili, che necessitano quindi di un ambiente povero di ossigeno per il loro sviluppo. Nonostante l'osservazione ripetuta nel tempo e l'arricchimento delle diverse aliquote, il primo approccio ha individuato solo nel campione β LdG2 3 diversi morfotipi; l'analisi Illumina ha ritrovato in totale 39 osservazioni riferibili agli Armophorea, distribuiti in questo ed altri campioni.

Anche per i Nassophorea e i Phyllopharyngea sono state individuate pochissime osservazioni morfologiche rispetto a quelle molecolari. Oltre alla difficoltà di discriminare i diversi morfotipi di tali gruppi solo attraverso l'osservazione in vivo, nella classe dei Phyllopharyngea il primo approccio non è stato in grado di rilevare il gruppo dei Suctoria, costituito da organismi simbionti.

Il terzo gruppo con la maggiore discrepanza tra le osservazioni ottenute con i due approcci è quello dei Litostomatea. La diversità genetica riscontrata per tale classe è elevata, in ogni campione sono state infatti ritrovate in media 34 diverse OTU Illumina. Il pattern che prevede la presenza di molti taxa di Litostomatea, rappresentati da un numero relativamente basso di organismi, era suggerito anche dal primo approccio.

Il rapporto osservazioni morfologiche/osservazioni molecolari è leggermente più alto per gli Oligohymenophora, con valori massimi per la sottoclasse dei Peniculia (38%) e minimi per gli Apostomatia e gli Astomatia, due gruppi costituiti completamente da organismi simbionti individuati solo dall'analisi Illumina.

Per i Prostomatea sono state individuate 17 osservazioni morfologiche fra i diversi campioni rispetto alle 91 trovate dall'analisi Illumina. Per tale gruppo, la discriminazione dei diversi morfotipi è stata problematica; ad esempio, in molti campioni sono stati osservati numerosi individui appartenenti al genere *Coleps* che sono stati però tutti inseriti all'interno di una singola categoria tassonomica (morfotipo) poiché la distinzione delle varie morfospecie di questo genere si basa sulla forma della teca, un carattere visibile solo a maggior ingrandimento.

Le classi dove l'analisi morfologica ha avuto maggior successo sono quelle degli Spirotrichea e degli Heterotrichea: per il primo gruppo con la sottoclasse degli Hypotrichia siamo arrivati ad individuare fino al 65% delle osservazioni molecolari mentre per gli Stichotrichi solo l'11% a causa dell'esistenza di numerose morfospecie difficili da discriminare con la sola osservazione in vivo.

Per quanto riguarda gli Heterotrichea, nonostante la percentuale possa sembrare bassa al 32%, le osservazioni Illumina non individuano ulteriori taxa rispetto a quelli trovati nel primo approccio, ma li riportano in più campioni (almeno in parte, la spiegazione potrebbe essere dovuta a DNA libero presente nei campioni).

Infine l'unica classe per cui l'analisi morfologica ha dato maggiori risultati rispetto all'analisi Illumina, è quella dei Karyollictea.

Il genere *Loxodes*, l'unico taxon conosciuto di acqua dolce appartenente a tale gruppo, con il approccio è stato identificato in tre distinti campioni (β PFEU1, γ PFEU1 e γ PFEU3) per un totale di 3 osservazioni. L'analisi Illumina invece è riuscita a trovare la corrispondente OTU solo in due campioni su tre; tale risultato è causato probabilmente dalla presenza di un *mismatch* tra il primer reverse utilizzato durante la PCR e la regione V4 di tale taxon, che potrebbe aver portato alla non amplificazione della sequenza in uno dei tre campioni.

	Osservazioni morfologiche	Osservazioni molecolari	% Osservazioni morfologiche / osservazioni molecolari
SPIROTRICHEA	74	449	16,5
OLIGOHYMENOPHOREA	52	525	10,0
LITOSTOMATEA	19	406	4,7
PROSTOMATEA	17	91	18,7
HETEROTRICHEA	15	47	31,9
NASSOPHOREA	3	60	5,0
ARMOPHOREA	3	39	7,7
KARYORELICTEA	3	2	150
PHYLLOPHARYNGEA	2	118	1,7
COLPODEA	2	210	0,9
PLAGIOPYLEA	0	15	0
TOTALE	190	1962	9,7

Tabella 8. Sono mostrate il numero totale di osservazioni morfologiche e molecolari divise per classe, trovate nei 12 siti di campionamento rispettivamente grazie al I e al II approccio.

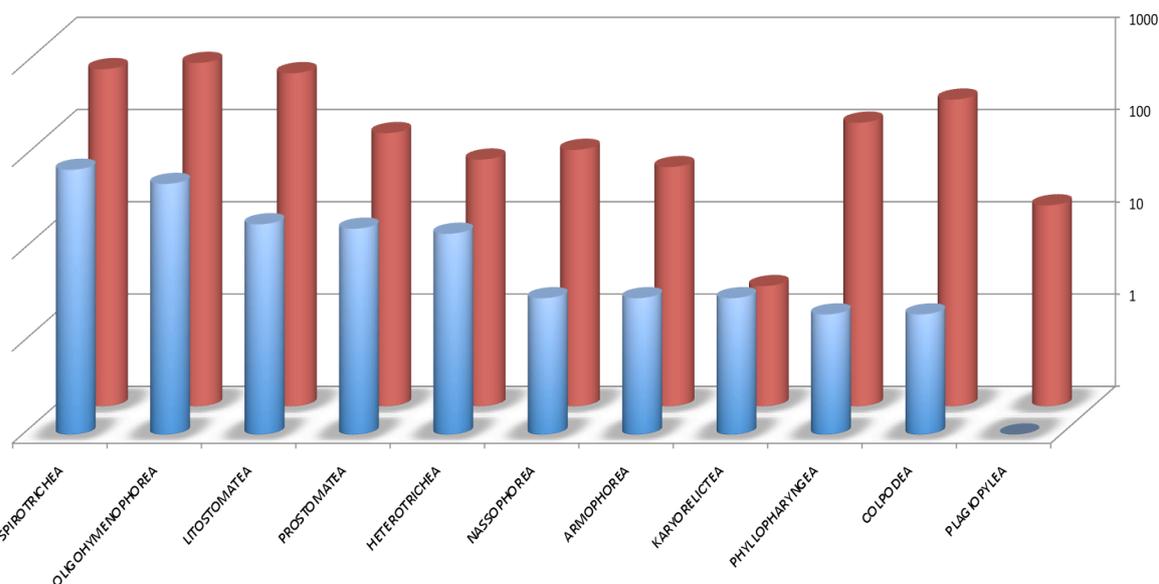


Figura 25. Nel grafico sono mostrate in celeste il totale delle osservazioni morfologiche divise per classe, mentre in rosso le osservazioni molecolari. L'asse delle ordinate è espresso in scala logaritmica.

3.3 Analisi statistiche

Anche per i dati Illumina, l'nMDS è stato scelto come metodo di ordinamento per rappresentare i rapporti fra le comunità di ciliati analizzate con un approccio di caratterizzazione esclusivamente molecolare.

L'ordinamento nMDS generato a partire dalla matrice di abbondanza standardizzata per il numero dei *contig* in ogni campione (Figura 26), ha evidenziato solo un modesto pattern di similarità tra i campioni in base all'area di campionamento.

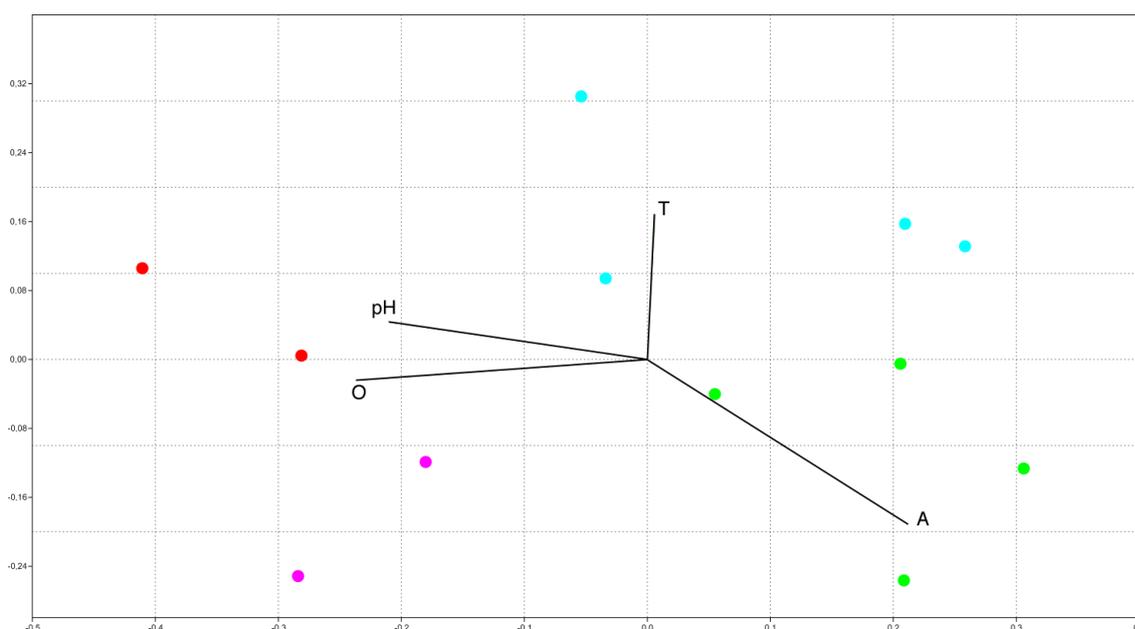
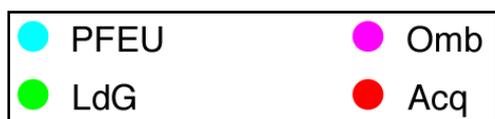


Figura 26. Grafico nMDS calcolato con i dati di abbondanza delle OTU ritrovate nei 12 campioni Illumina. Ogni colore corrisponde ad una diversa area di campionamento secondo la legenda sottostante. I vettori mostrano il contributo delle variabili ambientali all'ordinamento generato.



Con i dati ottenuti dal primo approccio, è stata inoltre costruita una matrice di presenza-assenza che tenesse conto solo dei 12 campioni utilizzati nel confronto (Figura 27); in particolare la matrice utilizzata per la ricostruzione nMDS prende in considerazione la checklist t_T e il sistema di classificazione 1 (tutte le categorie). Per tale matrice di dati, il grafico

nMDS non ha individuato particolari pattern di similarità tra campioni, probabilmente per il basso potere di risoluzione tassonomica dell'analisi.

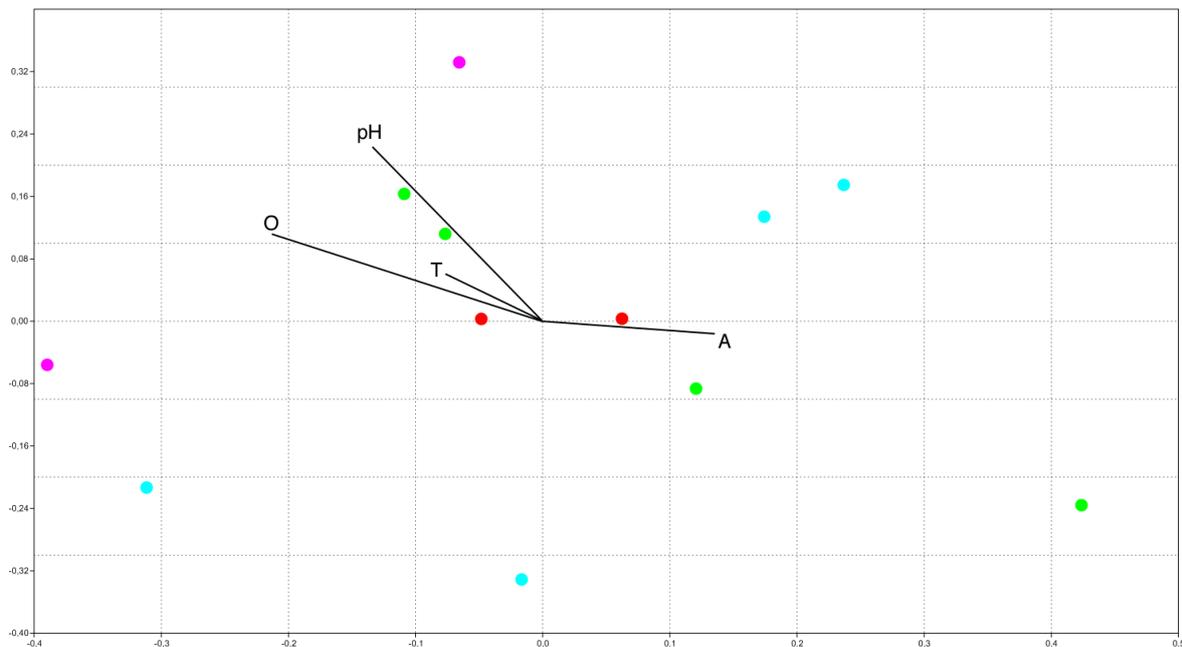


Figura 27. Grafico nMDS calcolato con i dati ottenuti dal primo approccio relativi solo ai 12 campioni utilizzati nel confronto.

Per i campioni analizzati con tecnologia Illumina è stata costruita anche una matrice solo con i dati di presenza/assenza delle singole OTU nei diversi campioni, eliminando cioè le relative abbondanze.

L'ordinamento nMDS generato da tale matrice (Figura 28) raggruppa in maniera evidente campioni provenienti dalle medesime aree di campionamento. Questo risultato appare sorprendente quando si considera che, solitamente, dati quantitativi sulle abbondanze sono considerati più informativi rispetto alla sola presenza/assenza.

Come già esposto in precedenza, però, l'alta variabilità nel numero di copie tra i diversi taxa del gene codificante per l'rRNA 18S, può condizionare anche di molto i reali valori di abbondanza di ogni OTU.

Visto che gli indici, come quello di Bray-Curtis qui utilizzato, sono fortemente influenzati dai dati di abbondanza e le tecniche di ordinamento guidate dai taxa più abbondanti, i grafici generati utilizzando tali dati hanno quindi un valore dubbio (Haegeman et al., 2013).

Per valutare i dati ottenuti da tecnologia NGS in cui si utilizza come marcatore una sequenza con un'alta variabilità nel numero di copie tra i diversi taxa, come i geni ribosomali, Stoeck e collaboratori (2014) suggeriscono proprio l'utilizzo dei soli dati di presenza-assenza.

L'alto potere di risoluzione dovuto all'elevato numero di OTU ottenute con l'analisi NGS ha permesso quindi di identificare pattern geografici precedentemente oscurati dal basso potere di risoluzione dato dai soli morfotipi (Bass et al., 2007; Boenigk et al., 2007).

Inoltre, utilizzando solamente i dati di presenza/assenza, viene data maggiore importanza alle OTU discriminati cioè alle OTU caratteristiche di ogni area, quei taxa presenti solo nei luoghi in cui si realizzano tutte le condizioni ambientali di cui necessitano.

In contrasto con le conclusioni suggerite dall'approccio tradizionale, i dati NGS suggerirebbero quindi un certo livello di similarità fra campioni provenienti dalla stessa area, anche se raccolti in periodi diversi, purché si consideri solo l'identità dei taxa e non le loro abbondanze. Questa differenza può essere dovuta, come già accennato, al diverso livello di risoluzione tassonomica. Un altro elemento potrebbe però essere legato alle differenti domande a cui i due approcci rispondono. Se l'approccio morfologico, infatti, ci permette di individuare le popolazioni "attive" in un determinato momento (o, nel nostro caso, anche quelle pronte a essere stimolate da un'improvvisa variazione nelle condizioni ambientali), la caratterizzazione NGS effettuata su DNA totale individua l'intera biodiversità, sia attiva che latente. In altre parole, il nostro lavoro suggerisce che le condizioni legate al microhabitat determinano la comunità fisiologicamente attiva in un sito, agendo selettivamente su un bacino di biodiversità molto più grande e relativamente omogeneo sull'intera area. Non sembrerebbe però supportata una delle famose ipotesi "estreme" sulla biodiversità microbica, che recita "tutto è ovunque, è l'ambiente a decidere", poiché parrebbe esistere una sostanziale differenza tra la biodiversità totale delle diverse aree. Lavori futuri che utilizzino più campioni per ciascuna area indagata potranno supportare, o contraddire, questa conclusione.

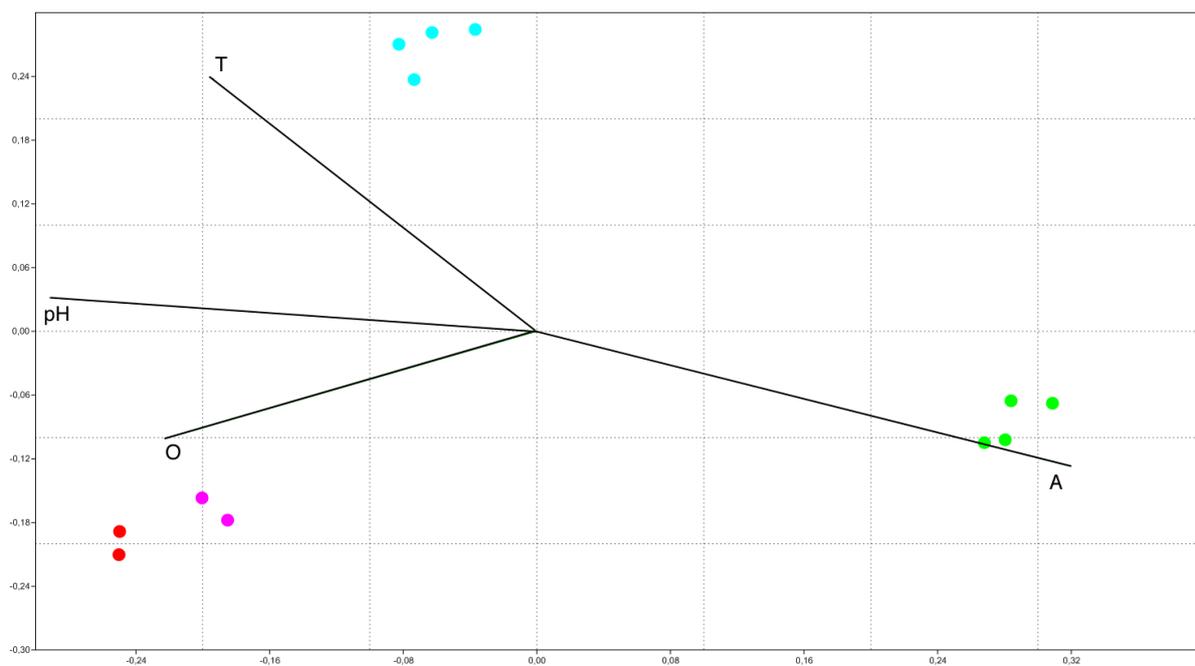


Figura 28. Grafico nMDS calcolato con i dati di presenza/assenza delle OTU ritrovate nei 12 campioni Illumina.

Conclusioni

All'interno della provincia di Pistoia sono state indagate, per quanto riguarda le comunità di ciliati associati al sedimento, 10 aree dulciacquicole rappresentative di diverse tipologie di habitat. La caratterizzazione di tali comunità è stata effettuata con due diversi approcci: uno morfologico-molecolare di tipo tradizionale ed un secondo che sfrutta la tecnologia del next-generation sequencing (NGS).

I dati riportati dal primo approccio mostrano un drastico aumento della biodiversità rilevabile attraverso una prolungata osservazione dei campioni, evidenziando numerose popolazioni presenti inizialmente solo con pochi individui o sottoforma di stadi di quiescenza. I dati molecolari sono stati utilizzati per confermare l'identificazione morfologica dei morfotipi individuati; il tasso di errore riscontrato è relativamente basso, stimato intorno all'1,1% degli organismi identificati molecularmente. I dati molecolari inoltre ci hanno permesso di ottenere per la prima volta le sequenze del gene per l'rRNA 18S di due generi descritti finora solo a livello morfologico: *Linostomella* e *Disematostoma*.

L'analisi Illumina effettuata su 12 dei 75 campioni presi in esame nel primo approccio ha individuato un numero di OTU di ciliati circa dieci volte maggiore. Sebbene questo dato possa essere una sovrastima, questo tipo di approccio molecolare ha dimostrato un enorme potere di risoluzione, e confermato che neppure un'osservazione prolungata è in grado di individuare tutti i taxa di più difficile osservazione.

Sono stati individuati pattern di similarità fra campioni provenienti dalle medesime aree di campionamento solo utilizzando dati di presenza – assenza provenienti dall'analisi Illumina. L'alto potere di risoluzione dovuto all'elevato numero di OTU ottenute con l'analisi NGS ha permesso quindi di identificare pattern geografici precedentemente oscurati dal basso potere di risoluzione dato dai soli morfotipi del primo approccio. I nostri dati potrebbero anche suggerire che le comunità fisiologicamente attive dipendono dalle condizioni del microhabitat, mentre la biodiversità complessiva è influenzata dai fattori del più ampio ecosistema. I dati relativi alle abbondanze molecolari, invece, si sono dimostrati inaffidabili per via di errori sistematici già riportati da altri autori.

Bibliografia

Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, et al., (2012) The revised classification of eukaryotes. *J Eukaryot Microbiol* 59:429-493

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410.

Bachy C, Dolan J, López-García P, Deschamps P, Moreira D (2013) Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *The ISME journal* 7:244–255

Backer GC, Smith JJ, Cowan DA (2003) Review and reanalysis of domain-specific 16S primers. *Journal of Microbiological Methods* 55:541–555

Bass D, Richards TA, Matthai L, Marsh V, Cavalier-Smith T (2007) DNA evidence for global dispersal and probable endemism of protozoa. *BMC Evol Biol* 7:162

Beaver JR & Crisman TL (1989) The role of ciliated protozoa in pelagic freshwater ecosystems. *Microbial Ecology* 17:111–136

Behnke A, Engel M, Christen R, Nebel M, Klein R, Stoeck T (2011) Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ Microbiol* 13:340–349

Benedetti-Cecchi L, Airoidi L, Frascchetti S, Terlizzi A (2003) Metodi sperimentali per la valutazione di influenze antropiche su popolamenti ed ambienti marini costieri In: Gambi MC, Dappiano M (eds), *Manuale di metodologie di campionamento e studio del benthos marino Mediterraneo*. *Biologia Marina Mediterranea*, 10 (Suppl. 1): 485-508.

Boenigk J, Jost S, Stoeck T, Garstecki T (2007) Differential thermal adaptation of clonal strains of a protist morphospecies originating from different climatic zones. *Environ Microbiol* 9:593–603

Boscaro V, Fokin SI, Verni F, Petroni G (2012) Survey of *Paramecium duboscqui* using three markers and assessment of the molecular variability in the genus *Paramecium*. *Molecular Phylogenetics and Evolution* 65:1004–1013

Boscaro V, Carducci D, Barbieri G, Senra MVX, Andreoli I, Erra F, Petroni G, Verni F, Fokin SI (2014) Focusing on genera to improve species identification: revised systematics of the ciliate *Spirostomum*. *Protist* 165:527-541

Bray JR & Curtis JY (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27:25-349

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al., (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336

Corliss JO & Esser SC (1974) Comments on the role of the cyst in the life cycle and survival of free-living protozoa. *Transactions of the American Microscopical Society* 93:578–593

Curds CR (1982) British and other freshwater ciliated protozoa. Cambridge University Press, Cambridge

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772

Delmont TO, Robe P, Cecillon S, Clark IM, Constancias F, Simonet P, et al., (2011) Accessing the soil metagenome for studies of microbial diversity. *Appl. Environ. Microbiol.* 77:13–24

Dopheide A, Lear G, Stott R, Lewis G (2009) Relative diversity and community structure of ciliates in stream biofilms according to molecular and microscopy methods. *Appl Environ Microbiol* 75:5261-5272

Dunthorn M, Klier J, Bunge J, Stoeck T (2012) Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. *J Eukaryot Microbiol* 59:185–187

Dunthorn M, Stoeck T, Clamp J, Warren A, Mahe F (2014) Ciliates and the rare biosphere: a review. *J Eukaryot Microbiol* 61:404-409

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461

Epstein S & Lopez-Garcia P (2008) “Missing” protists: a molecular prospective. *Biodivers Conserv* 17:261–276

Esteban GF, Finlay BJ, Olmo JL, Tyler PA (2000) Ciliated protozoa from a volcanic crater-lake in Victoria, Australia. *J Nat Hist* 34:159–189

Fan X, Chen X, Song W, Al-Rasheid KAS, Warren A (2011) Two novel marine *Frontonia* species, *Frontonia mengi* spec. nov. and *Frontonia magna* spec. nov. (Protozoa; Ciliophora), with notes on their phylogeny based on small-subunit rRNA gene sequence data. *Int J Syst Evol Microbiol* 61:1476-1478

Fan X, Lin X, Liu W, Xu Y, Al-Farraj SA, Al-Rasheid KAS, Warren A (2013) Morphology of three new marine *Frontonia* species (Ciliophora; Peniculida) with note on the phylogeny of this genus. *Eur J Protistol* 49:312-323

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368-376

Fenchel T, Esteban GF, Finlay BJ (1997) Local versus global diversity of microorganisms: cryptic diversity of ciliated protozoa *Oikos* 80:220-225

Fernandes NM, Dias RJP, Schrago CG, Silva-Neto ID (in press) Redescription and phylogenetic position of *Condylostoma arenarium* Spiegel, 1926 (Ciliophora, Heterotrichea) from Guanabara Bay, Brazil. *J Eukaryot Microbiol.* doi: 10.1111/jeu.12228

Finlay BJ & GF Esteban (1998) Freshwater protozoa: biodiversity and ecological function. *Biodiv Conserv* 7:1163–1186

Foissner W (1988) Taxonomic and nomenclatural revision of Sládeček's lists of ciliates (Protozoa: Ciliophora) as indicators of water quality. *Hydrobiologia* 166:1–64

Foissner W (1997a) Protozoa as bioindicator in agroecosystems, with emphasis on farming practices, biocides, and biodiversity. *Agr Ecosyst Environ* 62:93-103

Foissner W (1997b) Faunistic and taxonomic studies on ciliates (Protozoa, Ciliophora) from clean rivers in Bavaria (Germany), with descriptions of new species and ecological notes. *Limnologica* 27:179–238

Foissner W (1998) An updated compilation of world soil ciliates (Protozoa, Ciliophora), with ecological notes, new records, and descriptions of new species. *Eur J of Protistol* 34:195–235

Foissner W (1999) Protist diversity: estimates of the near-imponderable. *Protist* 150:363–368

Foissner W, Berger H, Schaumburg J (1999) Identification and ecology of limnetic plankton ciliates. Bavarian State Office of Water Management, Munich (pp. 793)

Foissner W, Chao A, Katz LA (2007) Diversity and geographic distribution of ciliates (Protista: Ciliophora) *Biodiv Conserv* 17:345–363

Fokin SI (2008) Rediscovery and characterisation of *Frontonia fusca* (Quennerstedt, 1869) Kahl, 1931 (Ciliophora, Peniculia). *Denisia* 23:251-259

Fonseca VG, Nichols B, Lallias D, Quince C, Carvalho GR, Power DM, Creer S (2012) Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. *Nucleic Acids Research* 40:e66

Fukami K, Watanabe A, Fujita S, Yamaoka K, Nishijima T (1999) Predation on naked protozoan microzooplankton by fish larvae. *Mar Ecol Prog Ser* 185:285–291

Gaedke U & Wickham SA (2004) Ciliate dynamics in response to changing biotic and abiotic conditions in a large, deep lake (Lake Constance). *Aquat Microb Ecol* 34:247-261

Gao S, Chen Z, Shao C, Long H, Al-Rasheid KAS, Song W (2008) Reconsideration of the phylogenetic position of *Frontonia*-related Peniculia (Ciliophora, Protozoa) inferred from the small subunit ribosomal RNA gene sequences. *Acta protozool* 47:47-54

Guhl BE, Finlay BJ, Schink B (1994) Seasonal development of hypolimnetic ciliate communities in a eutrophic pond. *FEMS Microbiol Ecol* 14:293-306

Gutiérrez JC, Martín-Gonzalez A, Callejas S (1998) Nuclear changes, macronuclear chromatin reorganization and DNA modifications during ciliate encystment. *European Journal of Protistology* 34:97–103

Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS (2013) Robust estimation of microbial diversity in theory and in practice. *ISME J* 7:1092–1101

Hammer O, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistical software package for education and data analysis. *Paleontologia Electronica* 4: 1-9.

Hausmann K, Hulsmann N, Polianski I, Schade S, Weitere M (2002) Composition of benthic protozoan communities along a depth transect in the eastern Mediterranean Sea. *Deep-Sea Research* 49:1959–1970

Jiang J, Wu S, Shen Y (2007) Effects of seasonal succession and water pollution on the protozoan community structure in an eutrophic lake. *Chemosphere* 66:523–532

Kreutz M & Foissner W (2006) The sphagnum ponds of Simmelried in Germany: a biodiversity hot-spot for microscopic organisms. *Protozoological Monographs Vol. 3*:1-267

Lara E, Berney C, Harms H, Chatzinotas A (2007) Cultivation-independent analysis reveals a shift in ciliate 18S rRNA gene diversity in a polycyclic aromatic hydrocarbon polluted soil. *FEMS Microbiology Ecology* 62:365-373.

Lee JJ, Leedale GF, Bradbury PC (2000) An illustrated guide to the protozoa. 2nd edn, vol 1. Society of Protozoologists, Lawrence, Kansas

Levinsen H, Nielsen TG, Hansen BW (2000) Annual succession of marine pelagic protozoans in Disko Bay, West Greenland, with emphasis on winter dynamics. *Marine Ecology Progress Series* 206:119–134

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhu K, Buchner A, et al., (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* 32:1363–1371

Lynn DH (2008) The ciliated protozoa. Characterization, classification, and guide to the literature, 3rd edn. Springer, Dordrecht

Madoni P (2003) Protozoa as indicators of wastewater treatment efficiency. In: *The Handbook of water and wastewater microbiology* (chap. 22, pp. 361–371). Amsterdam: Elsevier

Madoni P & Zangrossi S (2005) Ciliated protozoa and saprobial evaluation of water quality in the Taro River (northern Italy). *It J Zool* 72:21-25

McCarthy C (1998) Chromas 1.45. Technelysium Pty. Ltd. Helensvale, Queensland, Australia. <<http://www.technelysium.com.au/chromas.html>>

Medinger R, Nolte V, Pandey RV, Jost S, Ottenwalder B, Schloetterer C, Boenigk J (2010) Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Mol Ecol* 19:32–40.

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding region. *Gene* 71:491-499

Miao M, Song W, Clamp JC, Al-Rasheid KAS, Al-Khedhairy AA, Al-Arif S (2009) Further consideration of the phylogeny of some “traditional” heterotrichs (Protista, Ciliophora) of uncertain affinities, based on new sequences of the small subunit rRNA gene. *J Eukaryot Microbiol* 56:244-250

Modeo L, Petroni G, Rosati G, Montagnes DJ (2003) A multidisciplinary approach to describe protists: redescription of *Novistrombidium testaceum* Anigstein 1914 and *Strombidium inclinatum* Montagnes, Taylor, and Lynn 1990 (Ciliophora, Oligotrichia). *J Eukaryot Microbiol* 50:175-189

Modeo L, Rosati G, Andreoli I, Gabrielli S, Verni F, Petroni G (2006) Molecular systematics and ultrastructural characterization of a forgotten species: *Chattonidium setense* (Ciliophora, Heterotrichea). *Proc Jpn Acad Ser B* 82:359-374

Nakano S, Ishii N, Manage PM, Kawabata Z (1998) Trophic roles of heterotrophic nanoflagellates and ciliates among planktonic organisms in a hypertrophic pond. *Aquat Microb Ecol* 16:153–161

Nielsen KM, Johnsen PJ, Bensasson D, Daffonchio D (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* 6:37–53

Petroni G, Dini F, Verni F, Rosati G (2002) A molecular approach to the tangled intragenetic relationships underlying phylogeny in *Euplotes* (Ciliophora, Spirotrichea). *Mol Phylogenet Evol* 22:118-130

Prescott DM (1994) The DNA of ciliated protozoa. *Microbiological Reviews* 58:233–267

Riley JL & Katz LA (2001) Widespread distribution of extensive chromosomal fragmentation in ciliates. *Molecular Biology and Evolution* 18:1372–1377

Rosati G, Modeo L, Melai M, Petroni G, Verni F (2004) A multidisciplinary approach to describe protists: a morphological, ultrastructural, and molecular study on *Peritromus kahli* Villeneuve-Brachon, 1940 (Ciliophora, Heterotrichea). *J Eukaryot Microbiol* 51:49-59

Ryan PG, Watkins BP, Lewis Smith RI, Dastych H, Eicker A, Foissner W, et al., (1989) Biological survey of Robertskollen, western Dronning Maud Land: area description and preliminary species lists. *South African Journal of Antarctic Research* 19:10–20

Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE* 6: e27310

Schmidt PA, Balint M, Greshake B, Bandow C, Rombke J, Schmitt I (2013) Illumina metabarcoding of a soil fungal community. *Soil Biol Biochem* 65:128–132

Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K, Nikolausz M (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiology Ecology* 60:341-350

Skibbe O (1994) An improved quantitative protargol stain for ciliates and other planktonic protists. *Arch Hydrobiol* 130: 339-347

Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, et al., (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* 103:12115–12120

Song W & Wilbert N (2000) Ciliates from Antarctic sea ice. *Polar Biology* 23:212–222

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, Richards TA (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* 19:21–31

Stoeck T, Breiner HW, Filker S, Ostermaier V, Kammerlander B, Sonntag B (2014) A morphogenetic survey on ciliate plankton from a mountain lake pinpoints the necessity of lineage-specific barcode markers in microbial ecology. *Environ Microbiol* 16:430-444

Van de Peer Y, Van den Broeck I, De Rijk P, De Wachter R (1994) Database on the structure of small ribosomal subunit RNA. *Nucleic Acids Res* 22:3488-3494

Vlassov VV, Laktionov PP, Rykova EY (2007) Extracellular nucleic acids. *Bioessays* 29:654–667

Wang J, McLenachan PA, Biggs PJ, Winder LH, Schoenfeld BIK, Narayan VV, Phiri BJ, Lockhart PJ (2013) Environmental bio-monitoring with high-throughput sequencing. *Briefings in Bioinformatics* 14:575-588

Yan Y, Chen X, Chen X, Gao F, Al-Farraj SA, Al-Rasheid KAS (2015) Morphology and molecular phylogeny of three marine *Condylostoma* species from China, including two new ones (Ciliophora, Heterotrichea). *Eur J Protistol* 51:66-78

Yasindi AW, Lynn DH, Taylor WD (2002) Ciliated protozoa in Lake Nakuru, a shallow alkaline–saline lake in Kenya: seasonal variation, potential production and role in the food web. *Archiv für Hydrobiologie* 154:311–325

Zhao F, Xu K, He Y (2012) Application of the Ludox–QPS method for estimating ciliate diversity in soil and comparison with direct count and DNA fingerprinting. *Eur J Soil Biol* 49:112–118

Zhu F, Massana R, Not F, Marie D, Vaultot D (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* 52:79–92

Appendice

Protocollo per la preparazione del mezzo di coltura “Cerophyl”

COMPONENTI:

Wheatgrass medium

Wheatgrass powder	
GSE Vertrieb GmbH	25 g/l

Cerophyl buffer

Na ₂ HPO ₄	79,75 g/l
NaH ₂ PO ₄ x 2H ₂ O	27,1 g/l

Balanced Salt solution

NaCl	20,8 g/l
MgSO ₄ x 7H ₂ O	8 g/l
MgCl ₂ x 6H ₂ O	17 g/l
CaCl ₂	2,7 g/l
KCl	4,6 g/l

Stigmasterolo 5 mg/ml in Etanolo assoluto

Per preparare 1 litro di mezzo di coltura Cerophyl:

H ₂ O distillata	900 ml
Wheatgrass medium	100 ml
Cerophyl buffer	5 ml
Balanced salt solution	5 ml
Stigmasterolo	100 µl

Una volta inoculato con il batterio *Raoultella planticola*, collocare il mezzo in incubatore a 37°C **overnight**. Assicurarsi di lasciare abbastanza aria all'interno del contenitore per permettere una corretta crescita del batterio.

Le 54 categorie tassonomiche, divise per classe e sottoclasse individuate con il primo approccio di caratterizzazione (sono mostrate in grassetto).

KARYORELICTEA

Loxodes spp.

HETEROTRICHEA

Stentor spp.

Stentor coeruleus

Other *Stentor* spp.

Blepharisma spp.

Spirostomum spp.

Linostomella spp.

Climacostomum spp.

SPIROTRICHEA

Hypotrichia

Euplotes spp.

Aspidisca spp.

Other hypotrichs

Stichotrichia

Halteria like ciliates

Oxytricha like ciliates

Uroleptus like ciliates

Other stichotrichs

Oligotrichia

Choreotrichia

Codonella like ciliates

Other choreotrichs

ARMOPHOREA

Metopidae

Brachonella spp.

Other metopids

Caenomorphidae

Caenomorpha spp.

LITOSTOMATEA

Haptoria

Lacrymaria like ciliates
Didinium like ciliates
Enchelys like ciliates
Homalozoon spp.
Actinobolina spp.
Other haptorians

COLPODEA

Colpoda like ciliates
Other colpodeans

PHYLLOPHARYNGEA

Cyrtophoria

Phascolodon like ciliates
Trithigmostoma spp.
Other phyllopharyngean (=cyrtophorian)

Suctoria

PROSTOMATEA

Coleps like ciliates
Other prostomateans

NASSOPHOREA

Chilodontopsis like ciliates
Other (non-microthoracid) nassophorean
Microthoracida

PLAGIOPYLEA

Plagiopyla like ciliates

OLIGOHYMENOPHOREA

Peniculia

Frontonia like ciliates
Disematostoma like ciliates
Paramecium aurelia
Paramecium caudatum or *multimicronucleatum*
Paramecium calkinsi or *polycarium*

Paramecium bursaria

Lembadion spp.

Urocentrum spp.

Other penicules

Hymenostomatia

Colpidium like ciliates

Ophryoglena like ciliates

Other hymenostomes

Peritrichia

Vorticella like ciliates

Other peritrichs

Scuticociliatia

Unknown oligohymenophoreans

			<i>Euplotes</i> sp.	Spirotrichea	Hypotrichia		Some	Some	Some		Some	PCR on extracted DNA	1028	<i>Euplotes elegans</i>	DQ309868	98.5		
			<i>Oxytricha</i> -like ciliate (m2)	Spirotrichea	Stichotrichia		Some		Some			PCR on cell(s)	1527	<i>Stylonychia lemnae</i>	AF164124	99.2		
			<i>Trithigmostoma</i> sp.	Phyllopharyngea	Cyrtophoria			Rare										
			Prostomatean	Prostomatea	/		Some		Some									
			<i>Halteria</i> -like ciliate	Spirotrichea	Stichotrichia					Some								
			Phyllopharyngean ⁸	Nassophorea	?						Rare	PCR on cell(s)	1712	<i>Chilodontopsis</i> sp. FSP2007122601 ⁸	FJ998040	94.6		
γ	TTol		Stichotrich (m1)	Spirotrichea	Stichotrichia	Rare												
			<i>Frontonia</i> -like ciliate	Oligohymenophorea	Peniculia	Rare	Rare											
			Stichotrich (m2)	Spirotrichea	Stichotrichia	Rare				Some		PCR on cell(s)	1049	<i>Notolymena apoaustralis</i> / <i>Paraurostyla weissei</i>	KC430934 / AF164127	99.9		
			<i>Stentor</i> cf. <i>coeruleus</i>	Heterotrichea	/	Rare		Rare										
			<i>Ophryoglena</i> -like ciliate	Oligohymenophorea	Hymenostomatia	Rare												
			Haptorian	Litostomatea	Haptoria		Rare			Rare	Rare							
			<i>Lembadion</i> sp.	Oligohymenophorea	Peniculia			Many		Rare								
			Stichotrich (m3)	Spirotrichea	Stichotrichia			Rare		Rare		PCR on cell(s)	738	<i>Diaxonella trimarginata</i> voucher SC20040818	JQ424833	100.0		
			<i>Vorticella</i> -like ciliate	Oligohymenophorea	Peritrichia				Some									
			<i>Euplotes</i> sp.	Spirotrichea	Hypotrichia			Rare										
			Prostomatean	Prostomatea	/					Rare								
			Hymenostome	Oligohymenophorea	Hymenostomatia						Some	PCR on cell(s)	764	<i>Tetrahymena</i> sp. orphan clone	KJ028514	100.0		
			<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia					Some								
			Scuticociliate	Oligohymenophorea	Scuticociliatia						Many							
Fish Farm "Toscana Fish srl" Tanks	α	Acq1	<i>Coleps</i> -like ciliate	Prostomatea	/	Many	Many			Some		Some	Some					
			<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia	Many	Many			Some		Some	Some	Some				
			<i>Vorticella</i> -like ciliate	Oligohymenophorea	Peritrichia	Some	Rare											
			Oligotrich	Spirotrichea	Oligotrichia	Some				Some		Some	Some					
			Haptorian	Litostomatea	Haptoria	Some	Rare						Some					
			<i>Oxytricha</i> -like ciliate	Spirotrichea	Stichotrichia	Rare	Rare				Some			PCR on cell(s)	1087	<i>Stylonychia notophora</i>	FM209297	100.0
			Peritrich	Oligohymenophorea	Peritrichia	Rare	Rare											
			<i>Frontonia</i> -like ciliate	Oligohymenophorea	Peniculia			Many						PCR on extracted DNA	1705	<i>Frontonia tchibisovae</i> strain QD-pop2	DQ883820	92.9
			<i>Colpidium</i> sp.	Oligohymenophorea	Hymenostomatia				Many		Many	Some		PCR on extracted DNA	1698	<i>Colpidium striatum</i>	HM030739	99.3
			<i>Paramecium</i> cf. <i>caudatum</i> or <i>multimicronucleatum</i>	Oligohymenophorea	Peniculia						Many		Some	PCR on extracted DNA	1710	<i>Paramecium multimicronucleatum</i> strain Pm	HG315606	100.0
			<i>Euplotes</i> sp.	Spirotrichea	Hypotrichia					Some	Some	Some		PCR on extracted DNA	1829	<i>Euplotes muscorum</i>	DQ661046	98.7
			Hymenostome	Oligohymenophorea	Hymenostomatia					Some		Some						
			Stichotrich	Spirotrichea	Stichotrichia							Some						
			α	Acq2		<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia	Some	Some				Some	Many			
			Stichotrich (m1)	Spirotrichea	Stichotrichia	Some	Some					PCR on cell(s)	1197	<i>Holosticha diademata</i>	DQ059583	99.7		
			Hymenostome (m1)	Oligohymenophorea	Hymenostomatia	Some	Some											
			<i>Paramecium</i> cf. <i>caudatum</i> or <i>multimicronucleatum</i>	Oligohymenophorea	Peniculia	Some	Some			Some		PCR on extracted DNA	1168	<i>Paramecium multimicronucleatum</i> strain Pm	HG315606	100.0		
			Oligotrich	Spirotrichea	Oligotrichia		Some											
			<i>Vorticella</i> -like ciliate	Oligohymenophorea	Peritrichia	Some		Some										
			<i>Euplotes</i> sp. (m1)	Spirotrichea	Hypotrichia	Rare		Some		Some		Many	PCR on extracted DNA	1829	<i>Euplotes muscorum</i>	DQ661046	98.7	
			Stichotrich (m2)	Spirotrichea	Stichotrichia		Some		Some									
			Hymenostome (m2)	Oligohymenophorea	Hymenostomatia			Some		Some			PCR on extracted DNA	1207	<i>Glaucomides bromelicola</i>	AJ810077	98.3	
			<i>Euplotes</i> sp. (m2)	Spirotrichea	Hypotrichia				Rare		Some		PCR on extracted DNA	1862	<i>Euplotes daidaleos</i> W3	FR873718	98.9	

α	Acq3	Coleps-like ciliate	Prostomatea	/	Some		Some	Some		Many							
		Haptorian	Litostomatea	Haptoria		Rare											
		Peritrich	Oligohymenophorea	Peritrichia			Some	Some									
		Vorticella-like ciliate	Oligohymenophorea	Peritrichia				Some									
α	Acq4	Vorticella-like ciliate	Oligohymenophorea	Peritrichia	Some		Some			Many							
		Coleps-like ciliate	Prostomatea	/			Some	Some	Some								
		Peritrich	Oligohymenophorea	Peritrichia													
β	Acq1	Aspidisca sp.	Spirotrichea	Hypotrichia	Some	Some	Some			Rare							
		Uroleptus-like ciliate	Spirotrichea	Stichotrichia			Some										
		Vorticella-like ciliate	Oligohymenophorea	Peritrichia			Some										
		Frontonia-like ciliate	Oligohymenophorea	Peniculia	Rare	Rare	Some	Some	Some		Some		PCR on cell(s)	1096	Frontonia didieri strain QD-pop1	DQ885986	94.5
		Oligotrich	Spirotrichea	Oligotrichia			Some	Some	Some								
		Oligohymenophorean	Oligohymenophorea	?			Rare	Rare									
		Oxytricha-like ciliate (m1)	Spirotrichea	Stichotrichia	Rare	Rare	Some	Some					PCR on cell(s)	1531	Stylonychia pustulata	AF396973	98.9
		Euplotes sp.	Spirotrichea	Hypotrichia			Rare	Rare			Some						
		Choreotrich	Spirotrichea	Choreotrichia			Some	Rare									
		Stichotrich (m1)	Spirotrichea	Stichotrichia	Some		Some	Some					PCR on cell(s)	1530	Bistichella variabilis isolate DY-20090926-HE-1-1-14	HQ699895	97.4
		Scuticociliate	Oligohymenophorea	Scuticociliatia			Rare				Some						
		Spirostomum sp.	Heterotrichea	/			Rare										
		Oxytricha-like ciliate (m2)	Spirotrichea	Stichotrichia	Some		Some	Some	Some		Some		PCR on cell(s)	1533	Stylonychia ammermanni	JX885703	100.0
		Stichotrich (m2)	Spirotrichea	Stichotrichia			Some	Some	Some								
Colpidium sp.	Oligohymenophorea	Hymenostomatia						Some		PCR on extracted DNA	1099	Colpidium sp.	KF639910	99.1			
Hymenostome	Oligohymenophorea	Hymenostomatia						Some									
β	Acq2	Frontonia-like ciliate	Oligohymenophorea	Peniculia	Some	Some	Rare	Some		Some							
		Stichotrich (m1)	Spirotrichea	Stichotrichia	Some	Some	Some	Some									
		Aspidisca sp.	Spirotrichea	Hypotrichia	Some	Some	Some		Some			Some	Some				
		Prostomatean	Prostomatea	/	Some		Some	Some									
		Oxytricha-like ciliate	Spirotrichea	Stichotrichia	Some	Some		Some									
		Oligotrich	Spirotrichea	Oligotrichia	Some		Rare	Some	Some		Some						
		Homalozoon sp.	Litostomatea	Haptoria			Rare	Rare									
		Oligohymenophorean	Oligohymenophorea	?			Rare										
		Euplotes sp.	Spirotrichea	Hypotrichia			Rare		Some		Rare	Rare					
		Coleps-like ciliate	Prostomatea	/					Some		Some						Some
		Haptorian	Litostomatea	Haptoria								Rare					
Stichotrich (m2)	Spirotrichea	Stichotrichia								Rare	Some						
Scuticociliate	Oligohymenophorea	Scuticociliatia								Some							
β	Acq3	Frontonia-like ciliate	Oligohymenophorea	Peniculia						Some							
		Stichotrich (m1)	Spirotrichea	Stichotrichia	Rare		Some		Some								
		Choreotrich (m1)	Spirotrichea	Choreotrichia			Some										
		Oligotrich	Spirotrichea	Oligotrichia			Rare	Some			Some						
		Choreotrich (m2)	Spirotrichea	Choreotrichia	Many			Some			Some		PCR on cell(s)	1524	Strobilidium caudatum	AY143573	98.8
		Stichotrich (m2)	Spirotrichea	Stichotrichia			Some		Some		Some		PCR on cell(s)	1437	Anteholosticha monilata	GU942567	99.7
Haptorian	Litostomatea	Haptoria						Some									

		Haptorian (m2)	Litostomatea	Haptoria	Rare														
		<i>Uroleptus</i> -like ciliate	Spirotrichea	Stichotrichia	Rare	Some	Some	Some	Some	Some	PCR on cell(s)	723	<i>Uroleptus</i> sp. INHC148	KF051782	99.9				
		<i>Stentor</i> sp.	Heterotrichea	/	Rare	Rare	Rare	Rare											
		Haptorian (m3)	Litostomatea	Haptoria		Some		Some	Rare	PCR on cell(s)	666	<i>Arcuospathidium cultriforme</i>	DQ411860	99.1					
		<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia		Some	Some	Some											
		<i>Vorticella</i> -like ciliate	Oligohymenophorea	Peritrichia		Some													
		Haptorian (m4)	Litostomatea	Haptoria		Some													
		Stichotrich (m2)	Spirotrichea	Stichotrichia					Many	PCR on cell(s)	1120	<i>Oxytricha granulifera</i> isolate 2	AM412770	100.0					
		Scuticociliate	Oligohymenophorea	Scuticociliatia					Many										
β	LdG2	<i>Paramecium</i> cf. <i>bursaria</i>	Oligohymenophorea	Peniculia	Rare		Some	Some	Rare	Some	PCR on extracted DNA	911	<i>Paramecium bursaria</i> strain CCAP 1660/12	AB252001	100.0				
		<i>Spirostomum</i> sp.	Heterotrichea	/	Some	Some			Some										
		Scuticociliate (m1)	Oligohymenophorea	Scuticociliatia		Rare													
		<i>Frontonia</i> -like ciliate	Oligohymenophorea	Peniculia	Some	Rare	Some	Some											
		Scuticociliate (m2)	Oligohymenophorea	Scuticociliatia															
		<i>Halteria</i> -like ciliate	Spirotrichea	Stichotrichia	Rare	Some	Many	Many	Many	Many									
		Metopid (m1)	Armophorea	/	Rare														
		<i>Stentor</i> sp.	Heterotrichea	/		Rare	Rare	Rare		PCR on cell(s)	1002	<i>Stentor polymorphus</i>	AF357144	98.8					
		Haptorian	Litostomatea	Haptoria				Rare											
		<i>Oxytricha</i> -like ciliate	Spirotrichea	Stichotrichia		Some		Some	Some										
		<i>Vorticella</i> -like ciliate	Oligohymenophorea	Peritrichia		Rare	Some	Some											
		<i>Uroleptus</i> -like ciliate	Spirotrichea	Stichotrichia		Some	Some	Some											
		<i>Blepharisma</i> sp.	Heterotrichea	/						PCR on cell(s)	1690	<i>Blepharisma</i> sp. JJ-2014	KF206427	98.3					
		Metopid (m2)	Armophorea	/					Some										
		Stichotrich (m1)	Spirotrichea	Stichotrichia					Rare										
		Prostomatean	Prostomatea	/					Some										
		<i>Caenomorpha</i> sp.	Armophorea	/					Rare										
		Stichotrich (m2)	Spirotrichea	Stichotrichia					Some	Some									
		Stichotrich (m3)	Spirotrichea	Stichotrichia					Some										
		<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia						Rare									
β	LdG3	<i>Frontonia</i> -like ciliate	Oligohymenophorea	Peniculia	Some	Some		Some											
		<i>Oxytricha</i> -like ciliate (m1)	Spirotrichea	Stichotrichia	Rare	Some		Rare											
		Stichotrich (m1)	Spirotrichea	Stichotrichia	Rare	Some													
		<i>Vorticella</i> -like ciliate	Oligohymenophorea	Peritrichia	Rare	Some	Some	Some											
		Scuticociliate (m1)	Oligohymenophorea	Scuticociliatia	Rare														
		<i>Spirostomum</i> sp.	Heterotrichea	/	Rare	Rare													
		Haptorian (m1)	Litostomatea	Haptoria	Rare	Rare													
		<i>Uroleptus</i> -like ciliate	Spirotrichea	Stichotrichia		Some	Some	Some	Some										
		Choreotrich	Spirotrichea	Choreotrichia		Some													
		<i>Halteria</i> -like ciliate	Spirotrichea	Stichotrichia	Some	Some	Some	Many	Some	Many									
		Haptorian (m2)	Litostomatea	Haptoria		Some	Rare	Rare											
		Haptorian (m3)	Litostomatea	Haptoria		Some	Rare	Rare											
		<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia		Some	Some	Some											
		<i>Oxytricha</i> -like ciliate (m2)	Spirotrichea	Stichotrichia		Rare	Some	Some	Some	PCR on cell(s)	904	<i>Srylonychia bifaria</i>	FM209296	99.6					
		<i>Paramecium</i> cf. <i>bursaria</i>	Oligohymenophorea	Peniculia			Some	Many	Many										
		<i>Trithigmastoma</i> sp.	Phyllopharyngea	Cyrtophoria			Rare												

		Stichotrich	Spirotrichea	Stichotrichia	Some	Some	Rare	Some	Some								
		Metopid (m2)	Armophorea	/	Some												
		<i>Oxytricha</i> -like ciliate	Spirotrichea	Stichotrichia	Rare	Some	Some		Some		PCR on cell(s)	1731	<i>Stylonychia lemnae</i>	AF164124		99.4	
		Choreotrich	Spirotrichea	Choreotrichia		Some											
		<i>Paramecium cf. bursaria</i>	Oligohymenophorea	Peniculia		Some		Some		Some	PCR on extracted DNA	1077	<i>Paramecium bursaria</i> strain CCAP 1660/12	AB252001		99.9	
		<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia		Some											
		Haptorian (m2)	Litostomatea	Haptoria		several			Some		PCR on cell(s)	1086	<i>Lagynophrya acuminata</i> isolate LagAc	JQ723972		98.9	
		<i>Euplotes</i> sp.	Spirotrichea	Hypotrichia		Rare					PCR on cell(s)	1854	<i>Euplotes daidaleos</i> isolate W3	FR873718		99.9	
		<i>Uroleptus</i> -like ciliate	Spirotrichea	Stichotrichia				Some	Some								
β	LaN2	Unknown ciliate	?	?	Some												
		Haptorian (m1)	Litostomatea	Haptoria	Rare												
		<i>Aspidisca</i> sp.	Spirotrichea	Hypotrichia	Rare	Rare		Some	Some								
		Oligotrich	Spirotrichea	Oligotrichia	Some			Some	Some								
		<i>Coleps</i> -like ciliate	Prostomatea	/				Some									
		Stichotrich (m1)	Spirotrichea	Stichotrichia				Some	Some								
		Stichotrich (m2)	Spirotrichea	Stichotrichia				Some	Rare	Some	Some	Some	PCR on cell(s)	1269	<i>Oxytricha granulifera</i> isolate 2	AM412770	99.9
		Hymenostome	Oligohymenophorea	Hymenostomatia					Rare								
		Haptorian (m2)	Litostomatea	Haptoria					Rare								
		<i>Halteria</i> -like ciliate	Spirotrichea	Stichotrichia				Some	Some		Some	Some	PCR on cell(s)	1623	<i>Halteria grandinella</i>	AF194410	99.9
β	LaN3	Metopid	Armophorea	/	Some												
		Prostomatean	Prostomatea	/	Some												
		Unknown ciliate (m1)	?	?	??												
		Unknown ciliate (m2)	?	?	Some												
		Haptorian (m1)	Litostomatea	Haptoria	Rare	Some	Some				PCR on cell(s)	1027	<i>Rimaleptus binucleatus</i> isolate SKS422	KJ680552		99.0	
		<i>Ophryoglena</i> -like ciliate	Oligohymenophorea	Hymenostomatia	Some												
		Haptorian (m2)	Litostomatea	Haptoria	Rare												
		Stichotrich	Spirotrichea	Stichotrichia	Rare				Some								
		<i>Oxytricha</i> -like ciliate	Spirotrichea	Stichotrichia		Rare			Rare								
		<i>Halteria</i> -like ciliate	Spirotrichea	Stichotrichia		Some			Some		Many						
		Haptorian (m3)	Litostomatea	Haptoria					Rare								
		<i>Stentor</i> sp.	Heterotrichea	/					Rare								
		Haptorian (m4)	Litostomatea	Haptoria					Rare								
		<i>Paramecium cf. bursaria</i>	Oligohymenophorea	Peniculia					several	Some	Many	PCR on cell(s)	1576	<i>Paramecium bursaria</i> strain CCAP 1660/12	AB252001		100.0

^a We only considered organisms identified at least at the genus level

^{*} This sequence has been published in Boscaro et al. 2014 as *Spirostomum* sp. PFEU1_Spte (accession number: HG939553)

^{**} This sequence has been published in Boscaro et al. 2014 as *Spirostomum teres* PFEU3_Sm2 (accession number: HG939529)

[‡] Discrepancy between morphology and molecular data discussed in the text

[†] In this case, DNA was extracted from about 100 cysts formed after the organisms was grown in culture

[#] The first blastn hit was actually *Paranassula*, a Nassophorea. However, all other hits in the top 30 list are sequences of *Frontonia* or other Peniculia. We therefore concluded that the *Paranassula* sequence was misidentified, and the phylogenetic analysis corroborated this hypothesis

^{##} The first blastn hit was actually *Ameobophrya*, a parasitic dinoflagellate. However, all other hits in the top 50 list are sequences of prostomateans or other ciliates. We therefore concluded that the *Ameobophrya* sequence was misidentified

