

UNIVERSITÀ DEGLI STUDI DI PADOVA  
Dip. di Biomedicina Comparata ed Alimentazione

Corso di laurea magistrale in Biotecnologie per l'alimentazione

## **ANALISI METAGENOMICA DEL MICROBIOTA INTESTINALE DI PRIMATI (RED COLOBUS)**

Relatrice

Prof.ssa Barbara Cardazzo

Correlatori

Dott. Massimo Pindo

Dott.ssa Carlotta De Filippo

Fondazione E. Mach

Laureanda:

Chiara Dallago

Matricola n. 1035921

ANNO ACCADEMICO 2014/2015



“It is not the strongest of species that survive  
Or the most intelligent, but the one most  
Responsive to change.”

Darwin (1859)



## Sommario

RIASSUNTO .....	VII
ABSTRACT .....	IX
1. INTRODUZIONE .....	1
1.1 Metagenomica .....	1
1.1.1 Le scoperte di un mondo di microorganismi. ....	1
1.1.2 La nascita della metagenomica .....	2
1.1.3 Piattaforme Next Generation Sequencing .....	3
1.2 Il gene ribosomiale 16S rRNA .....	7
1.3 Microbiota intestinale .....	10
1.3.1 I colobi rossi ( <i>Procolobus gordonorum</i> ) .....	15
2. SCOPO DEL LAVORO .....	21
3. MATERIALI E METODI .....	23
3.1 Campionamento .....	23
3.2 Estrazione DNA genomico totale .....	24
3.3 Amplificazione PCR della regione 16S rDNA .....	25
3.4 Elettroforesi su gel d'agarosio .....	27
3.5 Purificazione amplificati con biglie magnetiche .....	28
3.6 Lab on a Chip (LOC) .....	30
3.7 Analisi Spettrofotometrica .....	30
3.8 Analisi Fluorimetrica .....	31
3.9 PCR quantitativa .....	31
3.9.1 Principio analitico della PCR in real time .....	31
3.9.2 Profilo della reazione e <i>crossing point</i> .....	31
3.9.3 La curva standard .....	32
3.9.4 Il metodo della derivata seconda massima .....	32
3.9.5 Allestimento PCR .....	33
3.10 PCR in Emulsione .....	34
3.10.1 Recupero delle Biglie .....	37
3.10.2 Arricchimento delle Biglie .....	38
3.10.3 Conta delle Biglie .....	38
3.10.4 Recupero delle biglie arricchite .....	39
3.11 Pirosequenziamento .....	41
4. RISULTATI E DISCUSSIONE .....	43
4.1 Estrazione del DNA genomico .....	43
4.2 Produzione degli ampliconi .....	45
4.3 Purificazione e validazione degli ampliconi .....	46
4.4 Quantificazione assoluta degli ampliconi .....	48

4.5 PCR in emulsione e pirosequenziamento.....	51
4.6 Analisi bioinformatica.....	53
4.6.1 Pre-processamento delle reads e assegnazione tassonomica .....	53
4.6.2 $\alpha$ – diversity .....	54
4.6.3 $\beta$ – diversity .....	59
4.6.4 Analisi metabolica (KEGG pathway) .....	60
5. CONCLUSIONI .....	65
BIBLIOGRAFIA .....	67
RINGRAZIAMENTI .....	79

## RIASSUNTO

Lo studio del microbiota intestinale è un argomento di attuale interesse nel panorama scientifico internazionale a seguito della sua importanza in molteplici processi fisiologici. È stato infatti dimostrato il suo coinvolgimento nella protezione contro agenti patogeni, nell'educazione del sistema immunitario e nella modulazione dello sviluppo del tratto gastrointestinale.

La ricerca in questo campo ha avuto una rapida ascesa grazie al superamento dei limiti tecnologici della microbiologia classica mediante le piattaforme di sequenziamento di nuova generazione che hanno permesso lo studio simultaneo di comunità di microrganismi (metagenomica) ad alta risoluzione.

Questo elaborato di tesi si inserisce in un progetto di ricerca sullo studio del microbiota intestinale nei primati, i colobi rossi (*Procolobus gordonorum*), mediante pirosequenziamento 454 (Roche). Lo studio è stato eseguito confrontando 31 campioni fecali di altrettanti individui provenienti da due differenti foreste (Mwanihana n=15, e Magombera n=16) localizzate in due zone distinte della Tanzania. I primati della foresta Mwanihana vivono all'interno del Parco Nazionale dei monti Udzungwa in un habitat incontaminato, mentre il secondo gruppo è localizzato nella foresta Magombera, soggetta agli effetti dell'antropizzazione e della degradazione ecologica crescente nell'area circostante. I campioni fecali raccolti nelle due foreste sono stati sottoposti ad estrazione del DNA genomico batterico e sottoposti all'amplificazione della regione ipervariabile V1-V3 del gene 16S rRNA. L'analisi metagenomica ha previsto la preparazione delle librerie, la PCR in emulsione e infine il pirosequenziamento 454. Dall'analisi bioinformatica delle sequenze (344.938 reads in totale) è risultata una significativa riduzione dell' $\alpha$ -diversity nel microbiota degli animali che vivono nella foresta disturbata dall'azione dell'uomo (Magombera) rispetto a quella degli animali che vivono nella foresta di Mwanihana. L'analisi della  $\beta$ -diversity ha rivelato una significativa separazione fra le due popolazioni microbiche con diverse abbondanze relative, mostrando differenze significative per le famiglie delle Ruminococcaceae (41,9% in Ma e 35,7% in Mw) e Lachnospiraceae (18,1% in Ma e 15,2% in Mw). Da questo risultato si può dedurre che gli individui appartenenti alla foresta Mwanihana, all'interno del parco naturale, non essendo sottoposti ad antropizzazione né ad altri tipi di disturbo, godono di una maggiore biodiversità del loro microbiota intestinale. La minor ricchezza nel microbiota dei campioni appartenenti alla foresta Magombera rispecchia, invece, la condizione di degradazione dell'ecosistema in cui sono inseriti gli individui di questa popolazione.



## ABSTRACT

Changes in the functional diversity of the microbial community populating the gastrointestinal tract (gut microbiota) have been shown to affect development, health, fitness and evolutionary trajectory in both humans and other animals. A number of drivers are known to influence gut microbiota composition such as diet and phylogeny, evolutionary history and interactions with gut macroparasites. In addition, the host's environment may also determine interspecific and interpopulation differences. Environmental degradation and fragmentation is of particular concern to conservationists: degradation tends to lead to a decline in species richness of both flora and fauna, while fragmentation may also lead to unviable population sizes and hence, local extinction. If habitat loss also leads to changes in gut microbiota, this could have knock-on effects on individual health, and increase the probability of such extinctions. Udzungwa red colobus monkeys are most threatened taxonomic group of primates in Africa. Primarily arboreal and highly sensitive to hunting and habitat destruction, this primate provides a useful model for investigating the effect of habitat fragmentation on gut microbiota diversity and function. Using barcoded 454 amplicon pyrosequencing of the 16S bacterial ribosomal RNA gene of 31 non-invasive fecal samples, we investigated the phylogenetic bacterial diversity of the gut microbiota in seven social groups of free ranging red colobus (*Procolobus gordonorum*) inhabiting two geographically distinct forests located in the Udzungwa Mountains of Tanzania. To understand its variability across habitats with different levels of human disturbances, we analyzed samples from either fragmented (n=16) or unfragmented (n=15) forests. Ruminococcaceae and Lachnospiraceae dominated in all individuals (mean 55%), which reflects the need to extract energy from fibre-rich diets consisting mainly of leaves. Analysis of genus composition shows a marked diversification of red colobus gut microbiota across habitats, with significant differences between groups living in disturbed versus undisturbed forests. We found an increase of biodiversity in the gut microbiota of individuals living in the largest and more undisturbed forest range, correlating with the higher diversity of available food sources. Such geographic variation in microbiota composition may be associated with the diversity of food plants in natural versus human-modified habitat, in turn requiring changes in metabolic pathways to digest xenobiotic molecules.



# 1. INTRODUZIONE

I batteri sono una presenza ubiquitaria ed essenziale in natura, anche se molte volte non si ha la consapevolezza della vasta gamma di processi in cui sono coinvolti. Questi microorganismi sono la forma di vita più diffusa sulla terra, e non solo, sono anche la forma di vita più antica presente sul nostro pianeta. Infatti, si ipotizza siano state forme di vita unicellulari molto simili agli attuali procarioti, dotate di requisiti minimi essenziali per poter essere considerate viventi, cioè la capacità di utilizzare materiali presenti nell'ambiente per mantenere la propria struttura, organizzazione e potersi riprodurre, ad aver dato origine alle prime forme di vita circa 3.5 miliardi di anni fa. Nonostante vengano sottovalutati, essi svolgono molte funzioni fondamentali, sia singolarmente o come comunità in simbiosi con altri organismi negli ambiti più diversi.

I batteri sono coinvolti, ad esempio, nella trasformazione di numerosi prodotti ad uso alimentare come pane, birra, vino, ma anche in processi come quello di azoto fissazione, reazione di cui hanno l'esclusiva, a partire da simbiosi a livello radicale di alcune piante; alcuni di loro sono in grado di raggrupparsi e dar vita al *quorum sensing* che gli permette di espletare delle funzioni che non avrebbero singolarmente (come la fluorescenza di alcuni pesci), altri sono in grado di scindere molecole inquinanti complesse, e molto altro ancora. Ovviamente oltre alla vastità di effetti benefici che sono stati elencati esistono anche gli effetti negativi, ovvero tutte quelle patologie infettive di origine batterica.

## 1.1 Metagenomica

### 1.1.1 Le scoperte di un mondo di microorganismi.

L'uomo ha sfruttato le capacità dei microorganismi fin dall'antichità anche se non conosceva il loro aspetto e nemmeno la loro esistenza. Pare che le tracce più arcaiche di microorganismi trovati coinvolti nei processi fermentativi, quale la realizzazione del vino, risalgono al 5000 a.C. in Mesopotamia.

Altro esempio è la realizzazione del sistema fognario durante l'impero romano, che permise di evitare il contatto con batteri patogeni abbondanti nei liquami, anche se allora si ignorava il perché.

Lo studio dei microorganismi è iniziato non moltissimo tempo fa, risale a dopo la metà del 1600 con il perfezionamento di microscopi ottici preesistenti da parte di Robert Hooke a cui si deve la scoperta delle cellule, ma più nello specifico ad Antoni van Leeuwenhoek che è stato il primo a scoprire

l'esistenza di batteri. Un grande contributo è stato dato indubbiamente da Pauster (ben 200 anni dopo Robert Hooke) considerato il fondatore della microbiologia moderna. Le sue tesi dimostrarono come in un ambiente sterile non è possibile la formazione di microorganismi *ex novo*.

Fu poi Rober Koch, nel 1876 che per primo riuscì a coltivare un microorganismo (*Bacillus Anthracis*) fuori dall'organismo animale e a descrivere il suo ciclo vitale. Koch, inoltre, fu il primo a descrivere il ruolo di un agente patogeno nell'insorgenza di malattie infettive e da queste osservazioni nacquero i postulati di Koch, tutt'ora in uso.

Negli anni successivi sono state sviluppate una serie di tecniche e colorazioni (es. colorazioni di Gram) che hanno permesso lo studio approfondito di questa nuova disciplina sconosciuta fino a pochi anni prima. Un grande contributo è stato fornito da Carl Woese, che nel 1967 identificò il dominio degli Archea che sono stati così separati dai Bacteria. Woese utilizzò tecniche di filogenesi molecolare applicata all'RNA 16S ribosomiale (Woese *et al.*, 1990, 1978 e 1977). Parte da qui l'era dell'analisi molecolare attraverso lo studio della sequenza del gene che codifica la subunità ribosomiale minore (16S) oggi considerata la sequenza principe per la classificazione di batteri e Archea.

Solo una decina di anni dopo grazie alle tecniche di amplificazione (PCR), è stata possibile una rapida accelerazione nello studio dei microorganismi, mediante tecniche molecolari.

La svolta è avvenuta grazie alla messa a punto della prima reazione di sequenziamento da parte di Frederick Sanger (1977), nonché all'applicazione delle tecniche di PCR al sequenziamento del DNA.

### 1.1.2 La nascita della metagenomica

La metagenomica è una branca della genomica che studia simultaneamente una comunità complessa di microorganismi presenti in un campione, evitando la crescita su terreni selettivi. Infatti la crescita su terreni di coltura permette di identificare solo l'1-3% dei microorganismi realmente presenti in campioni naturali, perdendo il 97-99% dell'informazione (Gordon, 2012; Hugenholtz *et al.*, 1998), a causa delle loro particolari condizioni di crescita, come nutrienti specifici e condizioni di anaerobiosi. Il superamento delle tecniche di microbiologia classica per identificare i microorganismi prevede l'uso del DNA genomico totale estratto da una matrice di varia natura e il sequenziamento del 16S rDNA. In questo modo è possibile studiare le comunità batteriche residenti in habitat differenti. L'analisi metagenomica che permette l'analisi simultanea di una comunità di microorganismi ha reso possibile identificare ogni singolo microorganismo facente parte della comunità, ma soprattutto, permette di studiare come i vari microorganismi interagiscono tra loro e con l'ambiente (ecologia microbica) e, non in ultimo, studiare le funzioni specifiche di quella comunità.



Figura 1.1 – Rappresentazione di campi di studio in cui vi è un interesse di tipo metagenomico.

I campi di applicazione della metagenomica sono numerosi: ecologico, agro-alimentare, igienico-sanitario, zootecnico. La maggior parte di questi studi tende a descrivere le comunità batteriche presenti, le interazioni tra di esse e con l'ambiente circostante o l'organismo che le ospita, al fine di comprendere il significato dei complessi processi biologici di cui fanno parte. In Figura 1.1 si possono vedere alcuni esempi di ambienti interessati da studi metagenomici.

### 1.1.3 Piattaforme Next Generation Sequencing

I primi segni di una rivoluzione tecnologica nel mondo del sequenziamento sono apparsi nel 2005 con una pubblicazione decisiva sulla tecnica di sequenziamento di sintesi sviluppata da 454 Life Science (Margulies, *et al.*, 2005) e il protocollo di sequenziamento multiplex polony del George Church's lab (Shendure, *et al.*, 2005). Entrambi i gruppi hanno usato una strategia che riduceva enormemente il volume di reazione mentre veniva esteso il numero di reazioni di sequenziamento. La strategia implicava il sequenziamento in parallelo di svariate centinaia di migliaia di frammenti sia su supporti di vetro che in strati di agarosio, in tal modo è stato possibile miniaturizzare le reazioni di sequenziamento, un enorme incremento in confronto alla produttività di un sequenziatore capillare. Inoltre la possibilità di evitare la costruzione di librerie mediante le tecniche di clonaggio riduceva drasticamente i costi ed i tempi necessari al sequenziamento dei genomi di grandi dimensioni.

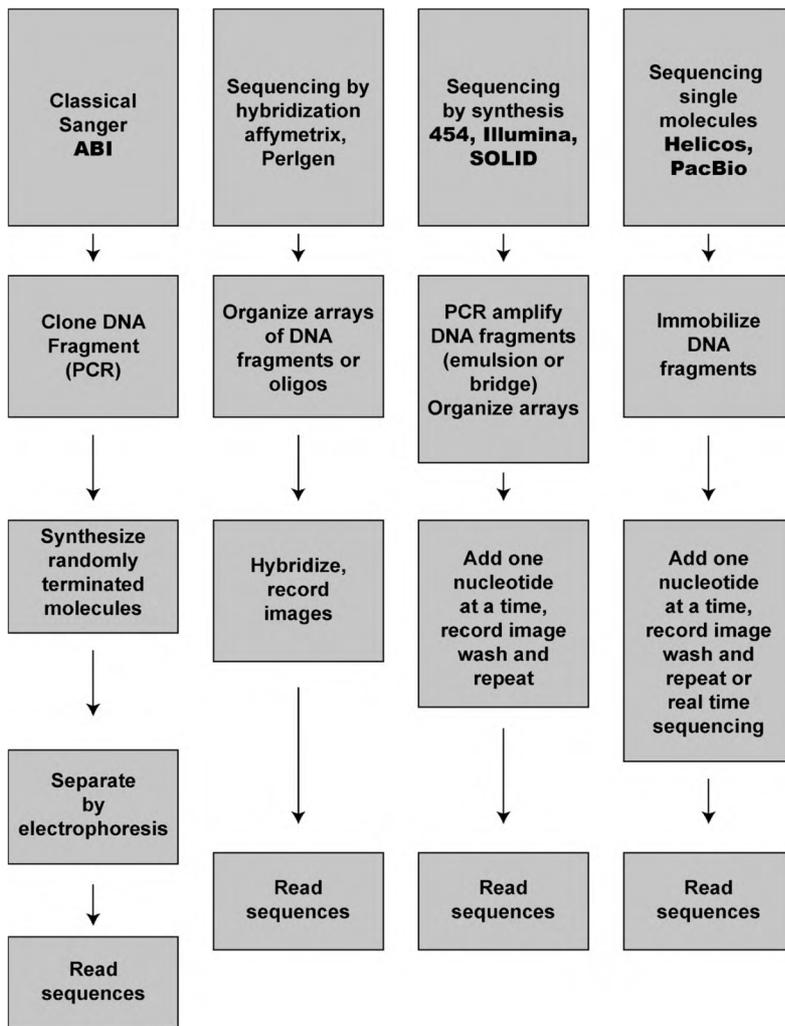


Figura 1.2 – Flow chart dei differenti metodi di sequenziamento (Delsenya et al., 2010).

Le tecniche di sequenziamento next generation sono in continua evoluzione e sono state descritte in numerose e recenti review (Shendure et al., 2008 e Munroe et al., 2010).

Il sequenziamento di sintesi è alla base di queste nuove tecnologie e prevede che i frammenti di DNA siano amplificati in clusters, denaturati e distribuiti in microarrays o in micropiastre che sono introdotte in una cella di flusso dove hanno luogo le reazioni di sequenziamento. Un primer è esteso ciclicamente da uno o alcuni nucleotidi alla volta e la sequenza è letta ad ogni step della sintesi del DNA.

Questa strategia differisce dal metodo di Sanger in cui un'intera gamma di copie parziali di molecole di DNA sono prima sintetizzate e poi analizzate.

La Figura 1.2 rappresenta un flow chart dei differenti metodi, anche se il sequenziamento da ibridazione è una tecnologia che rispetto alle altre ad oggi è limitata.

I vari metodi differiscono nella strategia usata per amplificare le sequenze, nella chimica usata e nella lunghezza delle reads. Hanno però in comune la possibilità di sequenziare fino a svariati milioni di frammenti di DNA in parallelo (Delsenya et al., 2010).

### Sanger “contro” NGS

Come già accennato i vantaggi derivati dalle piattaforme di NGS sono molteplici, prima di tutto l'aumento di produttività, le macchine più moderne arrivano a una produttività maggiore di 100 Gbp per corsa, a confronto di 70 – 100 kbp delle prime macchine di sequenziamento automatiche (Metzker, 2010). Questa capacità è stata incrementata grazie al sequenziamento in parallelo di centinaia di migliaia o milioni di campioni. Il secondo vantaggio è dato dall'eliminazione delle fasi di clonaggio

antecedenti il sequenziamento, ora sostituiti con una semplice ligazione degli adattatori specifici e successiva amplificazione, quest'ultima eliminata a sua volta in alcune tecniche di sequenziamento più recenti che permettono così di ridurre ulteriormente gli errori introdotti dalla PCR. La riduzione del volume di reazione e la possibilità di sequenziare in parallelo hanno anche ridotto i volumi dei reagenti richiesti e quindi i costi. Tutte insieme le tecnologie NGS hanno semplificato la strategia di sequenziamento, ridotto gli artefatti, aumentato la velocità alla quale può essere sequenziato un genoma e ridotto il costo di sequenziamento di molti ordini di grandezza, permettendo così di analizzare molti più campioni.

Dall'altra parte però troviamo anche per queste tecnologie all'avanguardia degli svantaggi e delle limitazioni che non ne permettono l'uso per ogni tipo di campione. Il primo punto a sfavore è la produzione di sequenze corte, questo implica la produzione di errori dal momento che nel genoma si incontrano sequenze ripetute e si ha la necessità di aumentare la copertura per assicurarsi di non compiere errori. Altra questione è la costruzione delle librerie, sebbene sia stata superata la fase di clonaggio con la produzione di librerie a inserto variabile, esistono ancora dei passaggi fondamentali, quali: frammentazione, legame con adattatori, purificazione, amplificazione PCR, ecc.; durante i quali è possibile l'introduzione di errori. Altre limitazioni sono gli errori e gli artefatti durante il sequenziamento, ogni metodo ha le sue debolezze, per esempio con la tecnologia 454 gli errori si verificano quando si analizzano corti omopolimeri perché l'incorporazione multipla dello stesso nucleotide potrebbe verificarsi ad ogni corsa invece della singola incorporazione. Lo stesso punto debole si ritrova anche nella più recente tecnologia Ion Torrent. In conclusione ogni metodo ha le sue peculiarità positive e negative e deve essere accuratamente scelto in base alla sequenza da analizzare e al risultato che si vuole ottenere.

#### Pirosequenziamento 454 Roche

Il pirosequenziamento è un metodo che si basa sul dosaggio del pirofosfato liberato in seguito all'attacco di un dNTP al filamento polimerizzato. I passaggi fondamentali di questa tecnica sono i seguenti:

- PREPARAZIONE della LIBRERIA. Per gli studi di metatassonomica il primo step consiste in un'amplicazione PCR della regione di DNA d'interesse. Durante questa fase si utilizzano dei primer speciali, detti *fusion primer*, che oltre alla sequenza target per il gene di interesse, vi sono legati due adattatori per il sequenziamento 454 Roche "A" e "B", una breve sequenza chiave (key) e una sequenza "barcode" identificativa del campione (MID) (vedi Fig. 1.3). Una volta ottenuti gli ampliconi è necessario effettuare il pooling dei campioni in esame creando un'unica libreria. Tenendo in considerazione i MID utilizzati in amplificazione, il numero di reads per campione che si vuole ottenere dalla corsa di sequenziamento e la concentrazione finale, i campioni sono miscelati in modo equimolare fino ad ottenere un unico campione.

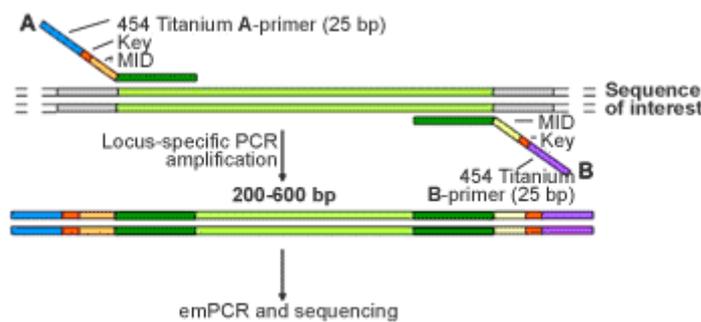
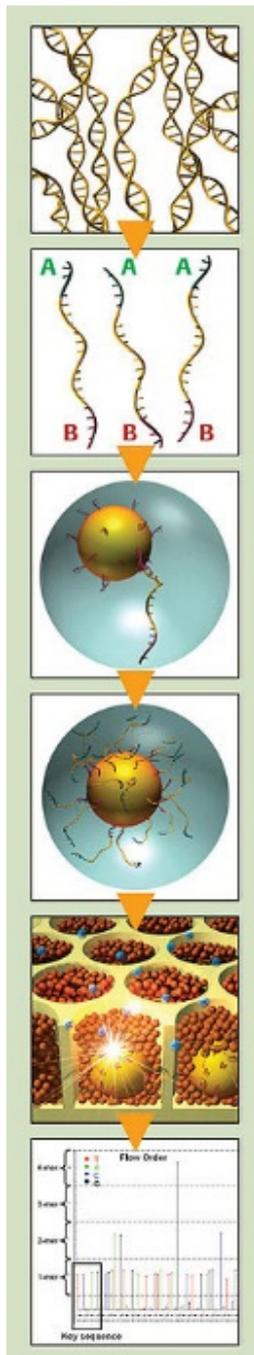


Figura 1.3 – Sequenze dei "fusion primer" e modalità di azione.

- EmPCR (PCR IN EMULSIONE). La libreria di ampliconi è sottoposta ad una seconda amplificazione in una miscela d'olio che, emulsionata, crea delle microgocce che costituiscono dei micro reattori all'interno dei quali è contenuta una sola Capture Bead, legata attraverso l'adattatore "B" ad una sola molecola di DNA stampo, e tutti i reagenti per l'amplificazione: al termine del processo in ogni goccia saranno presenti da 10 a 50 milioni di copie di dsDNA per biglia (Fig. 1.4). Dopo il recupero dall'emulsione, per eliminare le biglie che non hanno formato il legame con il DNA, alla reazione sono aggiunti un primer di arricchimento biotinilato (Enrichment Primer) complementare all'adattatore "A" e delle biglie magnetiche ricoperte di streptavidina. Le biglie senza DNA sono eliminate separandole da quelle legate alle biglie magnetiche (che hanno formato un legame streptavidina-biotina con il primer di arricchimento e quindi con le Capture Beads) con un magnete. Avviene quindi la denaturazione del legame tra il doppio filamento con il distacco del primer di arricchimento legato alle biglie magnetiche e il recupero delle Capture Beads con il DNA, che a questo punto è a singolo filamento.



- **PIROSEQUENZIAMENTO.** Il pirosequenziamento, prevede il caricamento sulla superficie di una piastra di vetro (Pico Titer Plate) delle biglie cariche di DNA, assieme agli enzimi sulfonilasi e luciferasi (Enzyme Beads), le PPiase (PPiase Beads), che eliminano il pirofosfato inorganico (PPI), e le Packing Beads, le quale stabilizzano e mantengono tutti i componenti del sistema immobilizzati entro i pozzetti della piastra PTP durante il sequenziamento. La piastra PTP è costituita da 1,6 milioni di pozzetti ciascuno progettato in modo tale da poter contenere una sola Capture Bead. Ciclicamente sulla superficie della PTP è aggiunto uno dei quattro nucleotidi per volta, laddove avviene la reazione di polimerizzazione dell'acido nucleico da parte della polimerasi segue il rilascio di PPI inorganico. Il PPI e l'ADP aggiunto vengono successivamente convertiti dall'ATP sulfonilasi in ATP, che fornisce l'energia necessaria all'enzima luciferasi per ossidare la luciferina e generare un segnale luminoso. Poiché l'aggiunta volta per volta dei nucleotidi è nota, è possibile determinare la sequenza del DNA o RNA in base all'intensità del segnale luminoso, proporzionalmente al numero di acidi nuclei incorporati, rilevato da una CCD camera. I nucleotidi non incorporati sono lavati e sono rimpiazzati da un secondo nucleotide e l'ordine: incorporazione – registrazione – lavaggio è ripetuto ciclicamente con i quattro nucleotidi fino a quando il primer è stato esteso sufficientemente. L'intensità del segnale è proporzionale al numero di nucleotidi che sono stati incorporati dalla DNA polimerasi (Delsenya *et al.*, 2010). Sebbene le macchine iniziali potessero leggere solo circa 110 bp, oggi le macchine moderne con un design e una chimica migliorata possono generare fino a 1,2 milioni di reads lunghe 800-1000 bp.

Figura 1.4 – Sequenziamento 454, le fasi di preparazione della libreria, PCR in emulsione, caricamento PTP e sequenziamento.

## 1.2 Il gene ribosomiale 16S rRNA

Il gene ribosomiale 16S rRNA, rappresentato in Fig. 1.5, è costituito da 10 regioni conservate e da 9 regioni ipervariabili (Fig. 1.6), è sottoposto a un basso tasso di evoluzione ed è presente in tutti i batteri. A causa delle restrizioni cui è soggetta la struttura dell'rRNA, come conseguenza del fatto che deve assumere una struttura secondaria definita e deve interagire con diverse proteine per formare un

ribosoma funzionale, il tasso di variazione delle sequenze dei geni che codificano per l'rRNA è di gran lunga inferiore a quello degli altri geni, di conseguenza è possibile determinare le relazioni filogenetiche su ampie distanze evolutive. Questo tipo di RNA funziona, quindi, come un orologio molecolare e consente accurate determinazioni delle distanze filogenetiche.

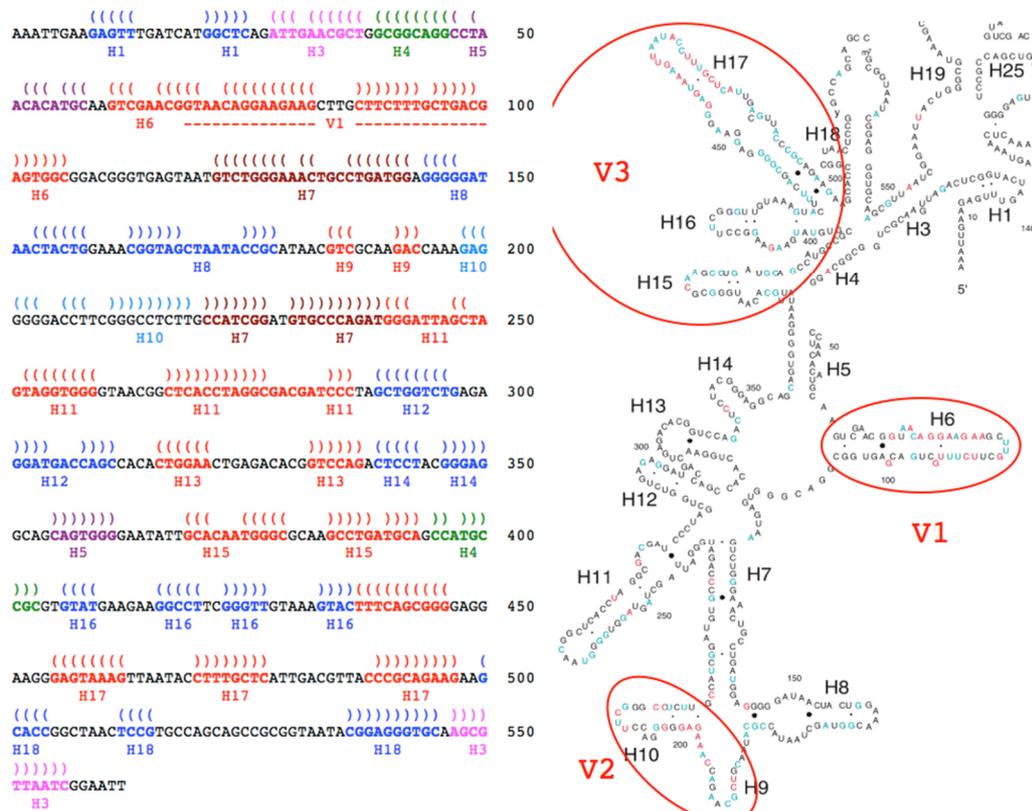


Figura 1.5 – Rappresentazione delle regioni ipervariabili V1, V2 e V3 della sequenza genica 16S rRNA.

Il prodotto genico del 16S rRNA è una regione che costituisce la subunità minore 30S dei ribosomi dei procarioti (Woese *et al.*, 1977).

La scelta dell'uso del gene 16S rRNA come marcatore filogenetico per esaminare la diversità microbica e per identificare e classificare i microorganismi nasce dalla difficoltà nella coltivazione della maggior parte dei microorganismi presenti in ambienti naturali.

Da quando Lane *et al.* (1985) per primi hanno descritto l'uso del gene 16S rRNA per l'identificazione e la classificazione dei microorganismi non coltivabili dell'ambiente, l'amplificazione mediante PCR, il clonaggio e il sequenziamento sono state le tecnologie primarie usate per identificare i procarioti provenienti da vari ecosistemi. Ad oggi sono state depositate in un database 3 milioni di sequenze batteriche e 54 mila sequenze di Archea della regione 16S (RDP Release 11, Update 3, settembre 2014 <http://rdp.cme.msu.edu/index.jsp>).

In filogenesi una “Operational Taxonomic Unit” (OTU) consiste in una definizione operativa tassonomica di una specie o di un gruppo di specie, spesso usata quando sono disponibili solo dati di sequenza del DNA, basandosi sulla similarità di sequenza. L’assegnazione tassonomica si basa sulla sequenza delle 9 regioni ipervariabili del gene 16S rRNA (Fig. 1.6). Specie, genere, famiglia e phylum sono convenzionalmente definiti con un valore di distanza filogenetica di 0.03, 0.05, 0.10 e 0.20 rispettivamente, basandosi sulla lunghezza intera (circa 1540 bp) della sequenza del gene 16S rRNA (Schloss & Handelsman, 2004).

Tale metodo può però creare, talvolta, incertezza nella corretta assegnazione tassonomica delle OTUs soprattutto se le differenze tra le sequenze geniche di diversi 16S rRNA non sono distribuite in modo eterogeneo lungo il gene 16S rRNA, ma sono concentrate soprattutto in alcune delle nove regioni ipervariabili (Stackebrandt & Goebel, 1994). Questo può accadere se l’analisi viene effettuata selettivamente solo su alcune regioni ipervariabili. È stato infatti dimostrato che alcune delle 9 regioni ipervariabili sono più informative di altre (Youssef *et al.*, 2009; Yu & Morrison, 2004) consentendo una assegnazione tassonomica più affidabile di altre (Liu *et al.*, 2007, 2008; Wang *et al.*, 2007).

Sempre più spesso per l’analisi globale dell’abbondanza e della diversità presente in microbiomi complessi sono utilizzate le piattaforme NGS ed in particolare il pirosequenziamento 454 (Claesson *et al.*, 2009; Sogin *et al.*, 2006; Kober *et al.*, 2009; Youssef *et al.*, 2009). Le singole regioni ipervariabili così sequenziate, sono raggruppate in OTUs, calcolando lo stesso valore di distanza convenzionale che si usa per la sequenza integrale del gene 16S. In alcuni studi recenti, sono state comparate tra loro le singole regioni ipervariabili confrontandole con l’intera sequenza del gene per stimare le abbondanze relative rilevate (Claesson *et al.*, 2009; Dethlefsen *et al.*, 2008; Huse *et al.*, 2008). È stato dimostrato che la scelta della regione ipervariabile incide significativamente sulla valutazione dell’abbondanza relativa delle OTUs. Engelbrektson *et al.*, (2010) hanno mostrato anche che l’analisi delle regioni V1 - V2 (circa 350 bp) e la regione V8 producono differenti omogeneità di OTUs di uno stesso campione raccolto da termiti.

Poiché la capacità delle piattaforme NGS sono in progressivo sviluppo, ad oggi si possono raggiungere, con la tecnologia GS FLX+ di Roche fino a 800bp, coprendo un numero di regioni ipervariabili sequenziate sempre più esteso. Pertanto, c’è una cruciale necessità di identificare le regioni ipervariabili più adatte, a seconda della matrice da analizzare, che possono fornire analisi affidabili del microbiota (Kim M. *et al.*, 2011).

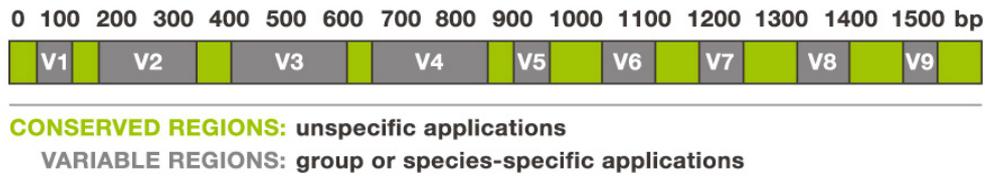


Figura 1.6 – Schema della suddivisione delle regioni ipervariabili all'interno della sequenza genica del 16S rRNA.

Quindi, è ormai provato che le diverse regioni ipervariabili non forniscono sempre gli stessi risultati se applicate a campioni biologici differenti. In ogni matrice, a seconda della comunità batterica vanno identificate le regioni ipervariabili più informative. Nello specifico, i campioni fecali forniscono una maggiore informazione tassonomica andando a studiare la regione ipervariabile V1-V3. Inoltre, nei database pubblici sono presenti per la maggior parte sequenze che corrispondono alla regione V1-V3 rispetto alle regioni a valle. Così, le sequenze parziali che corrispondono a questa regione avranno più sequenze nel database per essere comparate, facilitando di molto l'analisi filogenetica.

### 1.3 Microbiota intestinale

I mammiferi, come gli altri vertebrati, vivono in associazione co-evolutiva con la pleora di microorganismi che si trovano in una varietà di microambienti tissutali diversi: sulla pelle si rileva la presenza di Actinobacteria, Firmicutes e Proteobacteria; nella cavità orale di Bacteroidetes, Firmicutes, Fusobacteria e Proteobacteria; nella tratto respiratorio di Bacteroidetes, Firmicutes e Proteobacteria; nel tratto gastro-intestinale (GI) di Bacteroidetes, Proteobacteria, Actinobacteria e Firmicutes e nel tratto urogenitale di Firmicutes (specie in genere *Lactobacillus*) (Wang *et al.*, 2014). L'insieme di microorganismi in un particolare habitat è chiamato microbiota. I genomi collettivi di tutti i microorganismi in un microbiota sono definiti microbioma (Cryan *et al.*, 2012).

Nonostante esista una notevole variabilità inter-individuale nel microbiota intestinale, sembra che ci sia un equilibrio che conferisce benefici per la salute e che un'alterazione di batteri benefici possa influenzare negativamente lo stato di salute (Cryan *et al.*, 2011).

Il numero di specie stimato varia, ma è generalmente accettato che il microbioma adulto sia costituito da più di 1000 specie (Ley *et al.*, 2006), ovvero  $10^{14}$  cellule batteriche, che sono, nel complesso, circa 10 volte più numerose del totale delle cellule somatiche umane (Belkaid e Naik, 2013). Il microbiota batterico colonizza praticamente qualsiasi superficie del corpo umano, l'organo più colonizzato è il tratto gastrointestinale dove la quantità e la diversità del microbiota aumenta dallo stomaco al intestino tenue e al colon (Wang *et al.*, 2014); solo nel colon si trova oltre il 70% di tutti i

microorganismi presenti nell'organismo umano: si tratta di un ambiente dominato da batteri anaerobi, soprattutto anaerobi stretti, ma anche da virus, protozoi, archea e funghi (Dinan *et al.*, 2012 e Sekirov *et al.*, 2010). La composizione del microbiota è estremamente variegata ed è influenzata da fattori temporali e spaziali. L'intestino fetale dei mammiferi è sterile, e la colonizzazione inizia immediatamente dopo la nascita ed è influenzata dall'alimentazione, dal trasferimento materno, dagli stimoli ambientali e dall'uso di antibiotici (Sekirov *et al.*, 2010) (Fig. 1.7).

Durante il primo anno di vita, la composizione microbica dell'intestino dei mammiferi è relativamente semplice e varia notevolmente tra i diversi individui (Valles *et al.*, 2012).

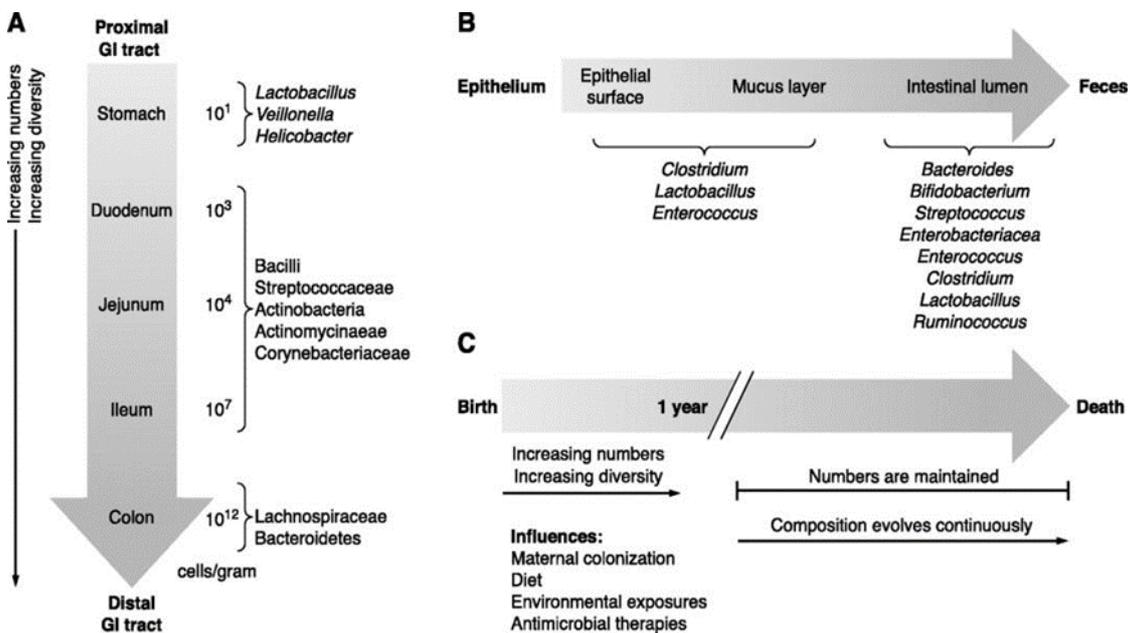


Figura 1.7 - Composizione del microbiota intestinale dal punto di vista spazio-temporale (da Sekirov *et al.*, 2010).

Il numero e la diversità di anaerobi stretti aumenta in base della dieta e l'ambiente e, dopo 1 anno di età, è già evidente la formazione un complesso microbioma simile a quello adulto (Cryan *et al.*, 2011). Un terzo del microbiota intestinale è comune alla maggior parte degli individui, mentre due terzi sono specifici per ognuno. La composizione del microbiota, quindi, evolve durante la vita, dalla nascita alla vecchiaia, ed è il risultato di diversi fattori ambientali (Sekirov *et al.*, 2010).

Negli adulti i phyla Bacteroidetes e Firmicutes di solito sono quelli più abbondanti del microbiota, mentre Actinobacteria, Proteobacteria e Verrucomicrobia, sono in genere costituenti minori del microbiota intestinale (Belkaid e Naik, 2013). Gli Archaea metanogeni (principalmente *Methanobrevibacter smithii*), gli eucarioti (soprattutto i lieviti) e i virus (principalmente fagi) sono normalmente presenti nell'intestino dei mammiferi, anche se però in minima percentuale.

Il microbiota nel tratto gastrointestinale nei mammiferi, non è omogeneo. Il numero di cellule batteriche varia da  $10^1$ - $10^3$  batteri per grammo di contenuto nello stomaco e nel duodeno, da  $10^4$  a  $10^7$  batteri per grammo nel digiuno e da  $10^{11}$ - $10^{12}$  cellule per grammo nel colon (Lozupone *et al.*, 2012). Approcci molecolari recenti forniscono nuovi metodi per l'identificazione delle specie batteriche non coltivabili in laboratorio, in particolare, mediante sequenziamento delle regioni ipervariabili del 16S RNA ribosomiale.

Il database delle specie batteriche si è arricchito enormemente, infatti ad oggi sono state depositate 3.019.928 16S sequenze di rRNAs. (<http://rdp.cme.msu.edu>)

Il consorzio MetaHIT ha pubblicato un primo documento che descrive l'insieme dei dati di metagenomica, ottenuti utilizzando la tecnologia di sequenziamento Illumina, contenente più di 200 sequenze del microbiota intestinale ottenute da campioni fecali di 124 individui europei, tra adulti sani, sovrappeso, obesi e pazienti con malattia infiammatoria intestinale (IBD) (Qin *et al.*, 2010). L'analisi funzionale ha identificato un insieme di geni che sembrano essere essenziali per la sopravvivenza dei batteri nell'intestino. Questi includono geni housekeeping tipici e geni che possono codificare per prodotti coinvolti nell'adesione alle proteine o nella raccolta degli zuccheri che si trovano nel sangue o nelle cellule epiteliali. I geni che codificano per queste funzioni possono essere essenziali per il funzionamento dell'ecosistema microbico intestinale nel suo complesso, ma possono essere codificati da differenti specie batteriche. Recentemente, il consorzio MetaHIT (Arumugam *et al.*, 2011), ha descritto la variazione del microbioma in soggetti provenienti da vari paesi occidentali, definendo tre specifici enterotipi. Gli autori hanno esaminato i metagenomi intestinali di 22 individui europei e hanno combinato i risultati con i dati già esistenti provenienti da individui giapponesi e americani. Gli autori affermano che i microbiomi analizzati non clusterizzano in base al paese o continente di origine, l'indice di massa corporea (BMI), l'età o il sesso, ma si possono raggruppare in base alle abbondanze relative delle specie dominanti, definendo tre enterotipi. Ciascuno di questi tre enterotipi è identificabile in base alla variazione dei livelli di uno dei tre generi batterici maggiormente abbondante: *Bacteroides* (enterotipo 1), *Prevotella* (enterotipo 2) e *Ruminococcus* (enterotipo 3). Gli autori affermano che gli enterotipi appaiono complessi, probabilmente non sono definiti da abitudini alimentari e non possono essere semplicemente spiegati come proprietà intrinseche dell'ospite, come l'età o come l'BMI, anche se esistono marcatori funzionali. Tale categorizzazione risulta comunque discutibile soprattutto perché non tiene in considerazione popolazioni diverse da quelle occidentali.

Le abitudini alimentari sono considerate uno dei principali fattori che contribuiscono alla diversità del microbiota intestinale umano. Secondo un recente lavoro, mediante la tecnica del pirosequenziamento, sono state catalogate tutte le specie batteriche, anche quelle non coltivabili, delle feci di bambini sani di età dagli uno ai sei anni che vivono in un villaggio rurale africano in Burkina Faso, le cui condizioni di vita sono molto simili a quelle dell'uomo del Neolitico al momento della

scoperta dell'agricoltura e dell'allevamento. La dieta di questi bambini è prevalentemente vegetariana, composta soprattutto da miglio, sorgo e vegetali. I risultati sono stati quindi confrontati con quelli ottenuti da una popolazione di bambini della stessa età che vivono a Firenze e che hanno una dieta tipicamente occidentale. I bambini del Burkina Faso presentano un microbiota con maggiore biodiversità, ricco di batteri che sono in grado di digerire la cellulosa, i quali, come controparte sono in grado di restituire come prodotto finale composti benefici per il nostro intestino quali il butirrato, un potente antinfiammatorio naturale. Inoltre, nonostante questi bambini vivano in condizioni igieniche precarie e siano soggetti a un tasso elevato di malattie infettive, presentano un ridotto numero di batteri potenzialmente patogeni quali *E. Coli*, *Shigella*, *Salmonella* che sono presenti nelle feci dei bambini fiorentini. Il microbiota dei bambini fiorentini invece è caratterizzato da specie batteriche che sono state spesso ritrovate negli obesi. Lo studio apre quindi notevoli spunti per capire il sottile equilibrio fra ecosistema microbico intestinale e stato di salute dell'ospite, al fine di comprendere l'origine delle malattie tipiche del mondo occidentale: obesità, allergie, malattie infiammatorie croniche intestinali, problemi sempre più crescenti e pressanti nella nostra società occidentale e in notevole incremento in età pediatrica (De Filippo *et al.*, 2010).

Il microbiota intestinale insieme con il corredo genetico, coopera con la fisiologia di un organismo (Belkaid e Naik, 2013). Il microbiota proprio di ogni individuo è stabilito subito dopo la nascita e deriva almeno in parte dalla madre. Il microbiota è sufficientemente stabile anche a fronte di cambiamenti nella dieta, la genetica dell'ospite, l'età, il sesso, interventi chirurgici e altri fattori ambientali. La maturazione del microbiota umano è un esempio dell'evoluzione ecologica: i numerosi microorganismi che popolano l'intestino umano, formano una comunità complessa dal punto di vista ecologico, e influenzano la normale fisiologia e la predisposizione alla malattia attraverso le attività metaboliche collettive e l'interazione con l'ospite. È ormai dimostrato che il microbiota intestinale ha molte funzioni benefiche, come la sintesi della vitamina K, la digestione delle fibre alimentari e la regolazione delle risposte infiammatorie (Tlaskalova *et al.*, 2011).

Gioca un ruolo importante nella resistenza ai patogeni, sia attraverso l'interazione diretta con i batteri patogeni che influenzano il sistema immunitario, sia quando il microbiota è in una condizione di equilibrio, tenendo sotto controllo la crescita di microorganismi patogeni e la conseguente invasione nelle cellule a livello dell'epitelio intestinale (Russell *et al.*, 2009).

La compromessa funzionalità dell'intestino umano è stata collegata a varie patologie come l'obesità, le malattie cardiovascolari, le patologie infiammatorie croniche intestinali, la sindrome metabolica, il diabete, le infezioni e le patologie neuronali (Cani *et al.*, 2008) in cui il microbiota è risultato giocare un ruolo chiave.

Per comprendere la co-evoluzione dei mammiferi e delle loro comunità microbiche, Ley e colleghi hanno condotto un'analisi basata sulle sequenze dei 16S rRNA batterici presenti nel microbiota fecale

di esseri umani e di 59 specie di mammiferi che vivono in due zoo diversi e in cattività. I risultati indicano che nei mammiferi, sia la dieta che la filogenesi, influenzano l'aumento di diversità batterica da carnivoro a onnivoro a erbivoro. Sembra che il microbiota intestinale dell'uomo che vive con uno stile di vita attuale, sia riconducibile a quello dei primati onnivori (Ley *et al.*, 2008a).

In un altro studio, Ley *et al.* (2008b), hanno classificato il microbioma delle diverse specie di mammiferi (compresa la specie umana) all'interno di categorie, divise in base alle loro diverse abitudini alimentari (Fig. 1.8) differenziandoli in erbivori, onnivori, carnivori.

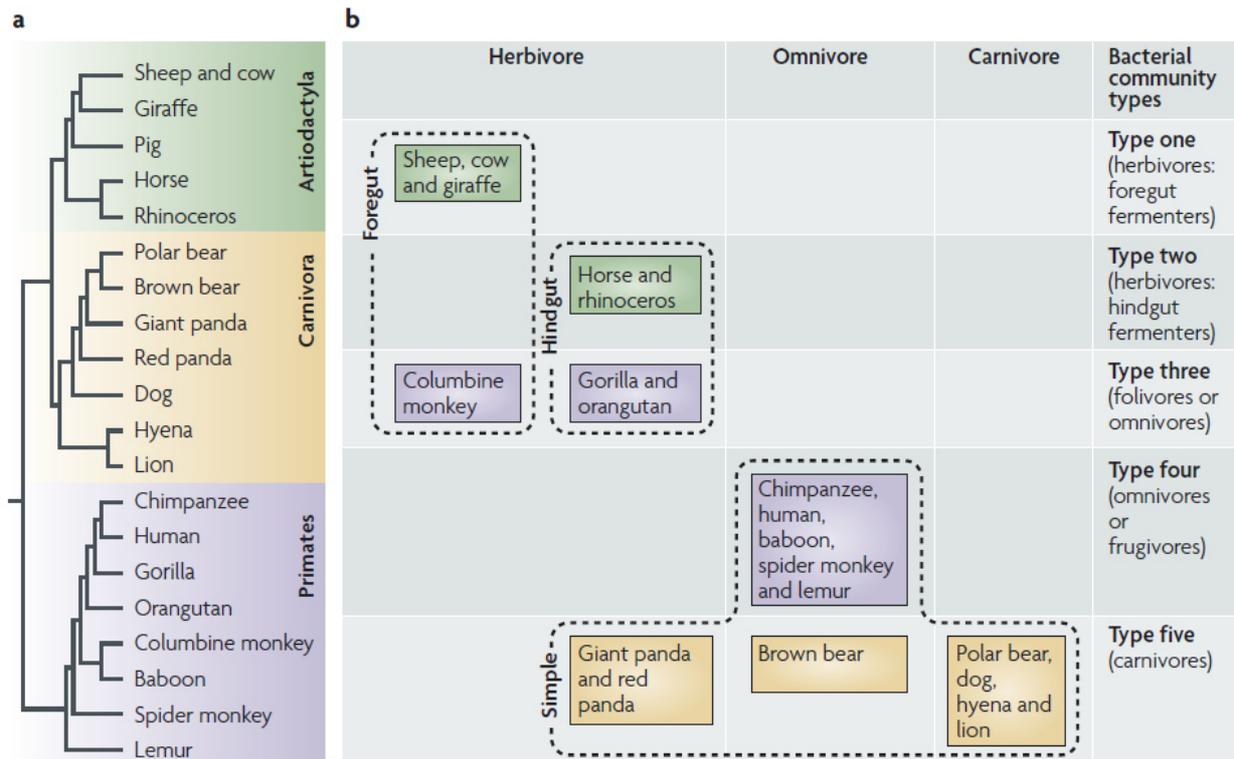


Figura 1.8 – Fattori che sono condivisi dai microbiota intestinali dei mammiferi. Rappresentazione schematica dell'impatto della filogenesi degli ospiti, la morfologia dell'intestino e la dieta sulla composizione di comunità batteriche fecali (Ley *et al.*, 2008b).

Da tale studio è emerso che animali con la stessa dieta hanno una vicinanza filogenetica, a livello di microbiota intestinale, maggiore rispetto ad animali con diete differenti. Inoltre è possibile effettuare una suddivisione anche a livello di morfologia gastrointestinale, che rispecchia la suddivisione in base alla filogenesi dei microorganismi e alla dieta, questo sta ad indicare una possibile coevoluzione tra la morfologia gastro intestinale e i microorganismi in relazione al tipo di dieta.

A livello di phylum rappresentativi tra campioni intestinali di vertebrati, incluso l'intestino umano, troviamo che i Firmicutes e i Bacteroidetes sono numericamente i phyla dominanti. Sebbene altri phyla possono coprire una grande porzione delle sequenze trovate in alcuni ospiti (per esempio,

Actinobacteria nelle pecore), sono soprattutto i Firmicutes e i Bacteroidetes quelli più comuni e ubiquitari (Fig. 1.9).

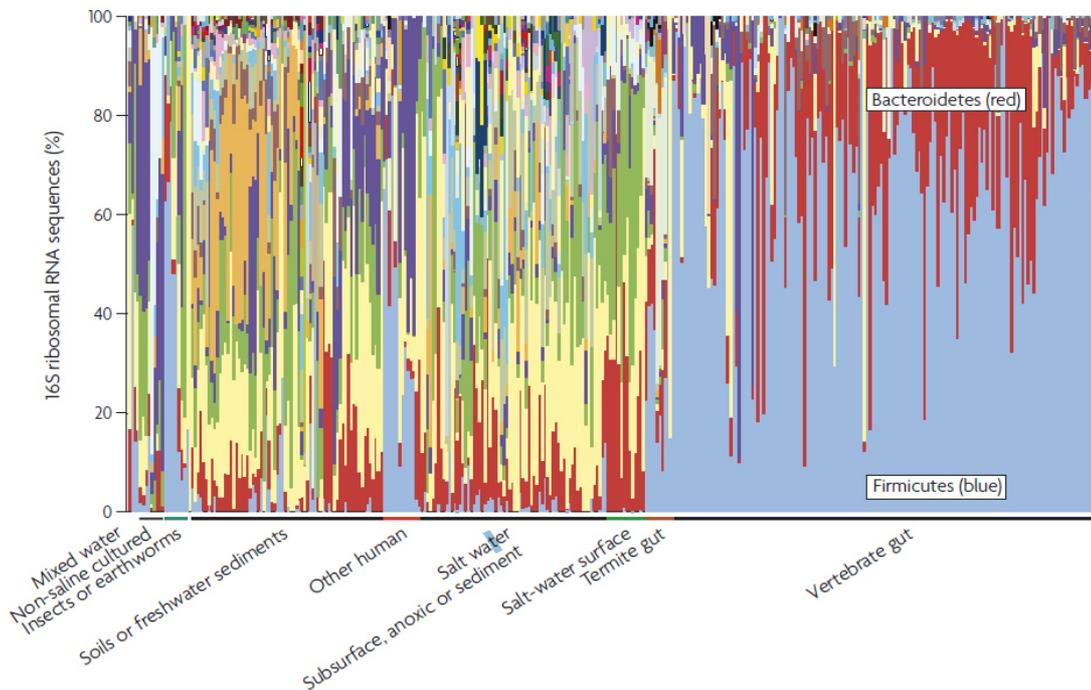


Figura 1.9 – Abbondanza relativa dei phyla nei campioni. Le barre nel grafico mostrano la proporzione delle sequenze in ogni campione che può essere classificato a livello di phylum (Ley et al., 2008b).

Un caso di morfologia gastro – intestinale molto interessante, in cui ogni comparto si è specializzato per una precisa funzione è quella presente nei ruminanti, in cui la fermentazione avviene nella parte anteriore dell'intestino e sembra essersi evoluta indipendentemente negli erbivori. Animali che hanno un tipo di fermentazione simile sono ad esempio gli Artiodattili (come capre, pecore, renne, giraffe e antilopi). Ma non solo, anche bradipi, scimmie colobine, alcuni roditori e due famiglie di marsupiali possiedono una fermentazione nella parte anteriore dell'intestino.

### 1.3.1 I colobi rossi (*Procolobus gordonorum*)

Le scimmie Colobinae sono gli unici tra i primati ad essersi adattati a una dieta che consiste principalmente di foglie, evolvendo un complesso fermentativo nello stomaco e nella parte anteriore dell'intestino (Map of Life - "Foregut fermentation in mammals" university of Cambridge, [http://www.mapoflife.org/topics/topic\\_573\\_Foregut-fermentation-in-mammals/](http://www.mapoflife.org/topics/topic_573_Foregut-fermentation-in-mammals/)).

I colobi rossi appartengono al genere *Piliocolobus*, a loro volta appartenenti alle Old World monkeys. All'interno del genere *Piliocolobus* vi sono numerosi colobi tra cui i colobi rossi: *Procolobus gordonorum* (o *Piliocolobus gordonorum*). Questa specie è endemica delle montagne Udzungwa in Tanzania (URC – Udzungwa red colobus) ed è considerata una specie in via di estinzione dalla IUCN Red

List delle specie minacciate, e la sua distribuzione è limitata al frammento di foresta nelle montagne Udzungwa del sud della Tanzania e una piccola parte di foresta vicino alla Kilombero Valley (Rodgers & Homewood, 1982; Struhsaker *et al.*, 2004; Struhsaker, 2005; IUCN: Struhsaker *et al.*, 2008).

In assenza di ambienti degradati e sottoposti a bracconaggio, la popolazione dei colobi rossi può prosperare attraverso un ampio gradiente ecologico (foreste pluviali, zone più o meno montuose). Nowak (2007) ha dimostrato che la variabilità comportamentale dei colobi rossi permette loro di sopravvivere anche entro habitat severamente degradati. Sfortunatamente la sopravvivenza a lungo termine di popolazioni in zone degradate è poco probabile (Struhsaker *et al.*, 2004).

Le montagne Udzungwa fanno parte della catena montuosa dell'Arco Orientale, la quale si estende dal sud del Kenya attraverso la Tanzania dell'est e sud – centrale (Lovett & Wasser, 1993). I monti dell'Arco Orientale sono considerati uno dei 25 *hotspot* di biodiversità del mondo, definito da Myers *et al.* (2000) come un'area contenente almeno 1.500 delle 300.000 specie di piante al mondo classificate come endemiche. Infatti le montagne dell'Arco Orientale e le foreste costiere del Kenya e della Tanzania hanno la più alta densità di piante endemiche e specie vertebrate per unità di area al mondo (Myers *et al.*, 2000). Questo alto livello di endemismo può essere attribuito sia all'età delle montagne (più di 30 milioni di anni) che all'influenza climatica che ha supportato storicamente una copertura di foreste in questa regione (Lovett & Wasser, 1993).

All'interno di questo ricco complesso naturalistico la popolazione colobi rossi si stima sia compresa tra i 10.000 e i 20.000 individui (Rodgers & Homewood, 1982; Dinesen *et al.*, 2001; Struhsaker, 2005). La popolazione è frammentata in numerose sottopopolazioni con molte popolazioni sorvegliate, contenenti meno di 2.000 individui (Rodgers & Homewood, 1982).

Mentre metà della popolazione di colobi rossi si trova tra il Parco Nazionale dei monti Udzungwa (UMNP: 1990 km<sup>2</sup> non tendendo conto della copertura delle foreste), l'altra metà potrebbe essere persa per la degradazione delle foreste o la completa distruzione di queste (Struhsaker 2005, 2010). Le foreste al di fuori dal Parco Nazionale, incluse le riserve, sono fortemente degradate a causa della deforestazione, della conversione all'agricoltura intensiva, di incendi incontrollati, di produzione di carbone e di bracconaggio (Struhsaker 2005, 2010; Marshall 2008a,b; Rovero *et al.*, 2012). Dati derivanti da censimenti degli animali indicano che la densità dei colobi nelle foreste fuori dal Parco sono particolarmente colpite dal bracconaggio e dalla degradazione delle foreste, con pericolo di estinzione in alcune zone più a rischio (Rovero *et al.*, 2012).

Negli scorsi decenni molti habitat al di fuori dal Parco Nazionale dei monti Udzungwa sono stati quasi distrutti, tra cui le foreste Lukonga e Kalunga (Struhsaker, 2010). Anche quasi metà della foresta Magombera è stata distrutta negli anni '80 (Rodgers & Homewood 1982; Struhsaker 2005, 2010; Marshall 2008b) e la restante parte è in corso di degradazione a causa del disboscamento, di incendi volontari e del bracconaggio (Marshall 2008b, per. obs.).

### Foresta Magombera (270 m slm)

La foresta di Magombera è particolarmente importante in una prospettiva di conservazione ed è un punto focale di studio, infatti contiene la più alta densità conosciuta di colobi rossi:  $149.2 / \text{km}^2$ , basata su una grandezza di popolazione stimata da Marshall (2008a), e possiede al suo interno una vasta gamma di habitat, dalla pianura alle foreste pluviali montuose che racchiudono una notevole biodiversità. È una piccola foresta nella valle Kilombero vicino al villaggio di Katulukila (Fig. 1.10) con un grande apporto di acqua. Negli anni '60 era una riserva del governo ed è stata divisa in due parti dalla ferrovia Tazara. Quando è stata costruita la ferrovia, molti dei più grandi alberi sono stati tagliati per ottenere legname. In seguito la foresta è stata ulteriormente degradata da costruzioni di vario genere che illegalmente hanno occupato questa zona. L'area della riserva era circa  $11 \text{ km}^2$ , ma negli anni '80 il governo ha escluso la parte della foresta a nord della ferrovia. Questa porzione è stata in poco tempo deforestata e sostituita con campi agricoli, portando ad una riduzione di circa  $6 \text{ km}^2$  (Decker, 1994). La rimanente foresta è stata destinata all'annessione alla riserva Selous Game Reserve, ma non è chiaro se è successo davvero.

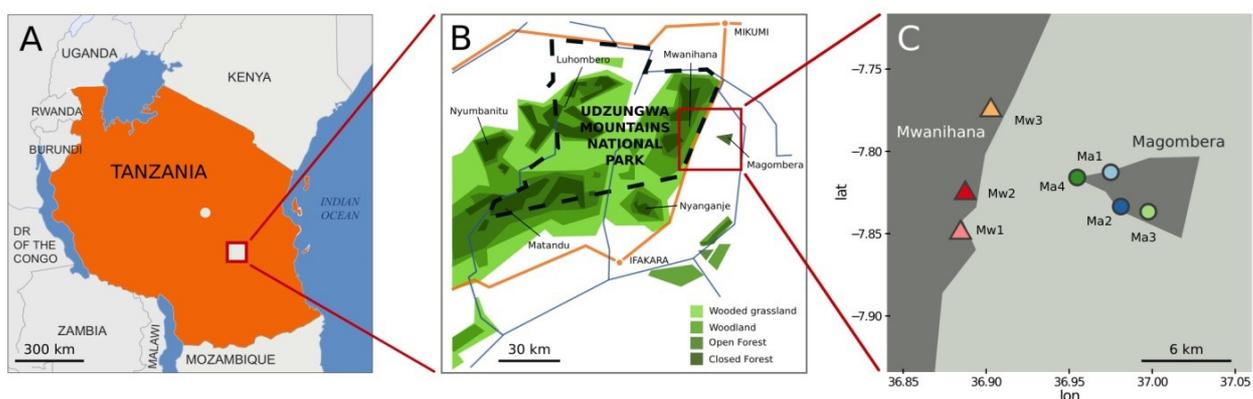


Figura 1.10 - Localizzazione geografica del Parco Nazionale dei monti Udzungwa (A, B), delle foreste Magombera e Mwanihana (C).

### Foresta Mwanihana (300 – 2080 m slm)

La foresta Mwanihana è situata interamente all'interno del Parco Nazionale delle montagne dell'Udzungwa (UMNP). È interessata da piogge annuali di 2000 – 2500 mm. È composta da una mosaico di diversi tipi di vegetazione, inclusa la pianura, la media montagna, la montagna e le montagne superiori, un mix di foreste sempreverdi e semidecidue con una copertura in altezza di 30 – 40 m negli esemplari più vecchi. Foreste di "Miombo" (termine Swahili che indica il genere

*Brachystegia spp.*) si presentano nelle aree più basse e asciutte, mentre i bamboo e le distese erbose si trovano ad altitudini maggiori (Lovett, 1993; Lovett *et al.*, 1988; Lovett & Pocs, 1993).

Negli anni '60 e '70 si sono verificate alcuni eventi di disboscamento, quando era una riserva del governo. L'area disturbata era prevalentemente situata lungo la parte est, vicino alla strada tra Kitadu e Kiberge. Invece, la parte della foresta rimasta più intatta è nella parte a ovest e lontana dalla strada (Struhsaker *et al.*, 2004).

## Dieta

I colobi rossi sono principalmente folivori ma anche granivori (Struhsaker 1975; Oates *et al.*, 1977; Marsh 1981; Mturi 1991; Siex 2003; Harris & Chapman 2007) con una dieta costituita per lo più da foglie giovani intere, semi non maturi, e piccioli di foglie mature. Come tutti i colobini, possiedono uno stomaco compartimentato, simile come struttura a quello dei ruminanti, atto a digerire per maggior tempo possibile il cibo, in modo da poter scindere la cellulosa in zuccheri semplici e metabolizzabili. Questo tipo di organizzazione fa parte della strategia di fermentazione digestiva di queste scimmie in cui i batteri producono enzimi quali l'idrolasi, capace di scindere il legame  $\beta$  dell'emicellulosa e della cellulosa, componenti della parete cellulare delle cellule vegetali (Kay & Davies 1994). Come tutti i vertebrati, i colobini non producono cellulasi, enzima necessario per rompere le cellulose ed emicellulose, rendendo questi batteri criticamente importanti nella loro fisiologia (Lambert 1998, 2007). Quando le emicellulose e le cellulose sono degradate dalla via fermentativa, gli acidi grassi volatili (anche chiamati acidi grassi a catena corta) sono rilasciati e subito disponibili come risorsa energetica (Kay & Davies 1994; Lambert 1998).

In relazione ai ruminanti che hanno una fermentazione ad alta efficienza, studi sulla fermentazione dei colobini dimostrano che essi assorbono acidi grassi volatili nel loro stomaco in maniera rapida come i ruminanti (Kay & Davies 1994).

Le diverse porzioni di pianta assunte dai primati, quali foglie giovani o mature, semi e frutti, non sono equivalenti a livello nutrizionale. Un criterio per la selezione del cibo messo in atto è in base alla percentuale più alta di proteine e lipidi (grassi) (Lambert 2007). Per quanto riguarda i colobi rossi non sempre vale questo criterio di selezione, hanno una versatilità abbastanza ampia sulle specie vegetali assunte. Tendono a preferire foglie giovani, più nutrienti in termini di contenuto di proteine e fibre, ma una buona parte della dieta è anche basata su foglie mature (Mowry *et al.*, 1996, Chapman & Chapman 2002).

Il bilanciamento dei nutrienti permette di massimizzare i consumi e l'acquisizione dei nutrienti, attraverso multiple fonti di cibo, e riducendo l'impatto di particolari composti secondari

nell'assorbimento dei nutrienti e nella digestione (Freeland & Janzen 1974; Westoby 1974; Glander 1982; Waterman & Kool 1994; Chapman & Chapman 2002; Shimada 2006; Struhsaker 2010).

Da una recente tesi di dottorato (Steel PhD, 2012) sui colobi rossi dell'Udzungwa è emerso che le foglie giovani intere costituiscono la percentuale maggiore della loro dieta (37.5%), mentre il secondo cibo più comune è rappresentato da frutti e semi non maturi (18.3%) e il terzo da parti di foglie mature (picciolo) (15.3%).

### Il microbiota intestinale nei primati

Gli studi che, ad oggi, sono stati compiuti sui primati hanno mostrato che la filogenesi dell'ospite e i fattori ambientali come la dieta giocano ruoli importanti nella formazione del microbiota. Come riportato in uno studio di Ochman *et al.* (2010) sul microbiota di chimpanzè (*Pan troglodytes*), bonobo (*Pan paniscus*) e gorilla (*Gorilla gorilla*) i profili delle comunità batteriche riassumono la filogenesi dell'ospite. Un lavoro di Frey *et al.* (2006) sui gorilla selvatici e Degnan *et al.* (2012) sugli scimpanzè selvatici hanno evidenziato l'esistenza di una stabilità temporale del microbiota dei primati, supportando l'ipotesi che il microbioma gastrointestinale potrebbe essere acquisito precocemente, e poi raggiungere lo stato di stabilità delle comunità microbiche, che sono resistenti alle perturbazioni (Degnan *et al.*, 2012; Koenig *et al.*, 2011; Turnbaugh *et al.*, 2009).

Per spiegare i processi di come le comunità batteriche si strutturano e insediano all'interno dell'organismo si utilizzano due "modelli" solitamente utilizzati nella dinamica ospite-parassita (Banks *et al.*, 2009; Kliks, 1990). Il "modello heirloom" sostiene la tesi che la formazione del microbiota sia da attribuirsi principalmente alla trasmissione filogenetica, senza tenere in considerazione la componente ambientale, mentre il "modello souvenir" supporta l'ipotesi di una derivazione ambientale del microbiota.

Un recente studio di McCord *et al.* (2014) dimostra che processi specie-specifici influenzano molto la struttura delle comunità batteriche e che il fattore geografico gioca un ruolo minore. Supportando la tesi che il microbiota è da vedersi orientato verso il paradigma "heirloom" rispetto a quello "souvenir". Nel suddetto studio sono stati confrontati il microbiota di primati filogeneticamente vicini (red colobus e black and white colobus) ma che vivono sia in foreste incontaminate che in foreste discontinue alternate a zone antropizzate. È stato visto come, nonostante la vicinanza filogenetica tra le due specie, la similarità nella fisiologia digestiva e la similarità nella dieta, c'è un'ampia separazione tra il microbiota dei red colobus e dei black and white colobus (Chapman & Pavelka, 2005; Chivers, 1994). Gli autori ipotizzano che questa differenza sia riconducibile alla diversità a livello fisiologico dell'apparato digestivo. I colobi rossi, infatti, hanno un compartimento dello stomaco aggiuntivo, il presaccus (Groves, 2007), che, sebbene sia una componente anatomica della parte anteriore

dell'apparato gastro intestinale, potrebbe avere effetti a monte sul microbiota intestinale. In conclusione, in questo studio è stato dimostrato che il microbiota dei primati che vivono in zone frammentate e degradate non è poi così diverso rispetto al microbiota di quelle specie di primati che vivono in habitat intatti, rinforzando l'idea che la filogenesi sia il fattore fondamentale capace di determinare il microbiota dei primati. C'è però da aggiungere che questo studio non possiede quella risoluzione necessaria per rilevare effetti di cambiamento nel microbiota ad un livello più dettagliato. Infatti, è stata utilizzata una tecnologia (ARISA) con bassa risoluzione rispetto alle tecnologie di next generation sequencing, fornendo solo uno screening preventivo.

Di parere contrastante è invece lo studio di Amato *et al.* (2013) in cui è stato ipotizzato che primati residenti in zone qualitativamente diverse (più o meno degradate) sviluppano microbioti differenti. Ciò accade perché gli animali che occupano zone degradate generalmente si nutrono di cibi differenti rispetto a quelli in zone indisturbate (Kamler *et al.*, 2007; Flaherty *et al.*, 2010; Abbas *et al.*, 2011; Chaves *et al.*, 2012).

I cambiamenti nella composizione dei taxa microbici, e conseguentemente i geni metabolici dipendenti da tali taxa descritti in diversi habitat, dimostrano gli effetti di come la composizione del microbioma gastrointestinale influenzi la fisiologia dei primati. Ad esempio il butirrato è la prima fonte di energia per l'epitelio del colon ed è riconosciuto come il promotore di numerosi effetti fisiologici (Donohoe *et al.*, 2011; Grootaert *et al.*, 2011; Hu *et al.*, 2010; Rawluszko *et al.*, 2011; Sharma & Singh, 2011). I microorganismi che producono butirrato, come *Butirvibrio spp.* e *Butiricoccus spp.*, sono i più abbondanti nei primati che occupano le foreste pluviali sempreverdi, mentre sono meno rappresentati negli esemplari cresciuti in cattività.

I risultati riportati da Amato *et al.* (2013) dimostrano come l'abbondanza relativa, la diversità microbica e la composizione del microbiota cambia in base all'habitat.

Quindi è evidente che la degradazione dell'habitat e conseguenti differenze nella dieta influenzano la fisiologia dell'ospite attraverso cambiamenti del suo microbiota intestinale.

Questo studio indica che i primati presenti in ambienti degradati sono sottoposti prima di tutto al rischio di una riduzione delle risorse disponibili che a sua volta comporta una riduzione di biodiversità delle specie microbiche intestinali (Amato *et al.*, 2013). Questa riduzione nella diversità del microbiota intestinale può avere come effetto la perdita di alcune funzioni metaboliche, rendendo il microbiota meno efficiente, meno resistente ai disturbi e più suscettibile alle invasioni da parte dei patogeni (Rosenfeld, 2002; Costello *et al.*, 2012).

## 2. SCOPO DEL LAVORO

Il presente elaborato di tesi si inserisce in un progetto di ricerca del Gruppo di ricerca Nutrizione e Nutrigenomica della Fondazione Edmund Mach di San Michele all'Adige svolto in collaborazione con il Museo delle Scienze di Trento, sezione di Biodiversità Tropicale. Il soggetto della ricerca è lo studio del microbiota intestinale nei primati colobi rossi (*Procolobus Gordonorum*) provenienti dalle montagne Udzungwa in Tanzania sud orientale. Il lavoro sperimentale consiste nell'analisi metagenomica dei campioni fecali di scimmie appartenenti a due insediamenti localizzati in differenti foreste (Mwanihana e Magombera). Il primo insediamento si trova in una foresta protetta, all'interno del Parco Nazionale dei monti Udzungwa (Mwanihana) mentre il secondo fa parte di una foresta circoscritta da insediamenti umani e sottoposta a fenomeni di disturbo a causa della crescente antropizzazione (Magombera). L'analisi del microbiota permette di indentificare la sua composizione batterica e in particolare eventuali differenze legate a fattori ambientali quali la disponibilità e qualità del cibo.

I campioni oggetto di questo elaborato sono 31 provenienti dalle foreste Mwanihana e Magombera sui quali è stata svolta l'estrazione del DNA, l'amplificazione del gene 16S rDNA ed il suo sequenziamento mediante piattaforma Roche 454. Mentre le successive analisi di bioinformatica sono state eseguite dal gruppo di ricerca di Metagenomica Computazionale della Fondazione Edmund Mach.



## 3. MATERIALI E METODI

### 3.1 Campionamento

I campioni fecali utilizzati in questo elaborato di tesi fanno parte del progetto “Primagut” condotto dalla Dr.ssa Claudia Barelli, raccolti durante due periodi di siccità (gennaio – febbraio e agosto – settembre). Gli animali erano semi-abituati agli umani, così i campioni fecali sono stati facilmente raccolti, sebbene non è stato possibile assegnare un sesso ai diversi campioni. Per evitare potenziali duplicati dello stesso campione, ognuno dei 7 gruppi è stato campionato solo una volta e durante un singolo evento di defecazione, ed è stato segnalato simultaneamente ogni singolo differente materiale fecale sul terreno.

Sono stati collezionati 31 campioni (in media 4,4 campioni per gruppo, intervallo: 3 – 7), 15 campioni provenienti dai tre gruppi residenti nella foresta Mwanihana (MW) e 16 campioni da quattro gruppi occupanti la foresta Magombera (MA) all’interno degli Udzungwa (7°40’S a 8°40’S e 35°10’E a 36°50’E, Fig. 1.10). Le due foreste campionate rappresentano un ampio contrasto nel tipo di habitat, nel gradiente di elevazione, nella protezione e nel livello di degradazione.

I campioni sono stati conservati in RNA $\text{later}^{\text{®}}$  (Lifetechnologies, Thermo Fisher Scientific) vista l’impossibilità di utilizzarli immediatamente per l’estrazione o congelarli.

I campioni sono stati raccolti senza contatto diretto con gli animali e sotto permesso approvato da “Tanzania Commission for Science and Technology” (COSTECH) e “Tanzania Wildlife Research Institute” (TAWIRI) e “Tanzania National Parks” (TANAPA). La procedura per la raccolta dei dati aderisce ai requisiti legali e si attiene alle leggi che governano la ricerca nel campo naturale della Tanzania.

La composizione dei campioni è rappresentata in Tabella 3.1.

Mwanihana		Magombera	
ID CAMPIONE	ID Gruppo	ID CAMPIONE	ID Gruppo
Mw1	C3	MA03_01	MA03
Mw9	C3	MA03_02	MA03
Mw12	C3	MA03_03	MA03
Mw2	C3	MA03_04	MA03
Mw7	NHH	MA03_05	MA03
Mw8	NHH	MA06_01	MA06
Mw10	NHH	MA06_02	MA06
Mw11	NHH	MA06_03	MA06
Mw13	NHH	MA06_04	MA06
Mw14	NHH	MA06_05	MA06
Mw15	NHH	MA20_01	MA20
Mw3	St	MA20_03	MA20
Mw4	St	MA20_04	MA20
Mw5	St	MA20_23	MA20
Mw6	St	MA24_52	MA24
		MA25_01	MA25

Tabella 3.1 – Corrispondenza campioni – gruppo – foresta. C3 = Campite 3, St = Sanje trail, NHH = near HondoHondo.

Le analisi riportate in questo elaborato di tesi si riferiscono unicamente ai 15 campioni provenienti dalla foresta Mwanihana (Tab. 3.1), mentre i rimanenti campioni sono stati analizzati precedentemente. Le analisi bioinformatiche sono state effettuate su tutti i 31 campioni totali.

### 3.2 Estrazione DNA genomico totale

L'estrazione del DNA è stata eseguita mediante l'uso di FastDNA SPIN kit for feces (MP Biomedicals, CA, USA) seguendo le specifiche del produttore con alcune modifiche.

La quantità di materiale di partenza per una corretta estrazione è di 500 mg, che devono essere depositi in un tubo Lysing Matrix E da 2,00 mL. All'interno dei tubini è presente una miscela di particelle di ceramica e silice progettati per lisare meccanicamente in modo efficiente tutti i microorganismi presenti nelle feci.

La procedura prevede, poi, l'aggiunta di buffer sodio fosfato e PLS per una prima lisi. In seguito si aggiunge il buffer MT e nuovamente il buffer sodio fosfato e si esegue una omogeneizzazione. I reagenti di lisi e pre-lisi hanno la capacità di proteggere e solubilizzare gli acidi nucleici, facilitare

l'estrazione di DNA genomico, e minimizzare la contaminazione da proteine, RNA, acido umico o polifenoli.

Successivamente alla lisi, i campioni sono centrifugati 5 minuti a 14000 giri per far sedimentare i detriti e la matrice di lisi. Da questo momento in poi le fasi prevedono il lavaggio, purificazione e recupero del DNA d'interesse. Per ottenere questo risultato è necessario eseguire un primo lavaggio, a cui segue il trasferimento dei campioni in provette munite di filtri SPIN che sono a base di silice GeneClean®. All'interno di tali provette si eseguono una serie di lavaggi, centrifughe ed eliminazione del surnatante, fino all'ultimo passaggio dove viene staccato il DNA d'interesse dal filtro mediante il TES.

### 3.3 Amplificazione PCR della regione 16S rDNA

Prima di procedere con l'amplificazione è necessario stabilire la qualità e la quantità del DNA estratto mediante l'uso di uno spettrofotometro. Lo strumento utilizzato è il Thermo scientific NanoDrop 8000 UV – Vis e permette di ottenere la concentrazione del campione in base all'assorbanza degli acidi nucleici e di verificare la presenza di diversi contaminanti quali: proteine, composti secondari e sali.

La reazione a catena della polimerasi (PCR) (Protocollo di reazione Tab. 3.4 e Programma PCR Tab. 3.5) è stata allestita per l'amplificazione della regione specifica di DNA dei campioni estratti: l'intervallo V1-V3 della regione ipervariabile del 16S ribosomiale, mediante l'uso di primer specifici (27-Forward: 5' - AGAGTTTGATCMTGGCTCAG-3' e 533-Reverse: 5'-TTACCGCGGCTGCTGGCAC – 3'). Il primer forward contiene: l'adattatore 454-A, specifico per la chimica "Lib – L" del metodo di sequenziamento "One – Way reads" (454 Life Science Corp., CT, United States); la sequenza TCAG, la sequenza barcode (Multiple Identifier, MID) specifica per ogni campione e la sequenza target specifica 27F (Tab. 3.2). Il primer reverse contiene l'adattatore 454-B, la sequenza TCAG e la sequenza target specifica 533R (Tab. 3.3).

Campione	MID	Lib –L Adaptor A	Key sequence	Sequenza MID	27F
Mw1	F8M8	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	CTCGCGTGTC	AGAGTTTGATCMTGGCTCAG
Mw2	F8M15	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	ATACGACGTA	AGAGTTTGATCMTGGCTCAG
Mw3	F8M17	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	CGTCTAGTAC	AGAGTTTGATCMTGGCTCAG
Mw4	F8M2	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	ACGCTCGACA	AGAGTTTGATCMTGGCTCAG
Mw5	F8M37	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	TACACACACT	AGAGTTTGATCMTGGCTCAG
Mw6	F8M40	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	TACGCTGTCT	AGAGTTTGATCMTGGCTCAG
Mw7	F8M41	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	TAGTGTAGAT	AGAGTTTGATCMTGGCTCAG
Mw8	F8M73	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	TGTCGTCGCA	AGAGTTTGATCMTGGCTCAG
Mw9	F8M43	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	TCGCACTAGT	AGAGTTTGATCMTGGCTCAG
Mw10	F8M46	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	TGACGTATGT	AGAGTTTGATCMTGGCTCAG
Mw11	F8M58	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	CGTACAGTCA	AGAGTTTGATCMTGGCTCAG
Mw12	F8M88	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	ATCTACTGAC	AGAGTTTGATCMTGGCTCAG
Mw13	F8M59	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	CGTACTCAGA	AGAGTTTGATCMTGGCTCAG
Mw14	F8M60	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	CTACGCTCTA	AGAGTTTGATCMTGGCTCAG
Mw15	F8M61	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG	CTATAGCGTA	AGAGTTTGATCMTGGCTCAG

Tabella 3.2 – Elenco dei campioni e relativo nome del MID associato, sequenza completa degli adattatori, key, MID e primer Fw.

Lib – L Adaptor B	Key sequence	533R
CCTATCCCCTGTGTGCCTTGGCAGTC	TCAG	TTACCGGGCTGCTGGCAC

Tabella 3.3 - Sequenza completa adattatore, key e primer Rw

Forward primer (Primer A):  
5'-CCATCTCATCCCTGCGTGTCTCCGACTCAG-(MID)-(template-specific sequence)-3'

Reverse primer (Primer B):  
5'-CCTATCCCCTGTGTGCCTTGGCAGTCACG-(template-specific sequence)-3'

Figura 3.1 - Sequenza nucleotidica dei Primer Fw e Rw preceduti da MID, key e adattatore.

**Protocollo di reazione:**

Rreagenti	Volume: $\mu\text{l}$ per una reazione	Concentrazione
FastStart High Fidelity Taq <sup>®</sup> polymerase	0.25	5 U/ $\mu\text{l}$
FastStart HiFi reaction Buffer con 18 mM MgCl <sub>2</sub>	2.5	10 X
FastStart HiFi PCR Nucleotide Mix	0.5	10 mM di ogni dNTP
Primer F	1	10 $\mu\text{M}$
Primer R + MID	1	10 $\mu\text{M}$
Molecular Biology Grade Water	13.75	
DNA	6	20 ng

Tabella 3.4 – Protocollo di reazione con concentrazioni e volumi dei reagenti necessari per l'amplificazione PCR del 16S per un volume finale di 25  $\mu\text{l}$ .

*Programma PCR:*

94°C x 3'	} x 25 cicli	Denaturazione iniziale
94°C x 15''		Denaturazione
60°C x 45''		Annealing
72°C x 1'		Estensione
72°C x 8'		Estensione finale
4°C → ∞		

Tabella 3.5 - Programma di reazione della PCR.

### 3.4 Elettroforesi su gel d'agarosio

L'amplificazione PCR solitamente è seguita da un'analisi degli amplificati che permette di valutare l'avvenuta reazione e l'eventuale presenza di prodotti aspecifici o dimeri di primer. Questo passaggio è definito dalla corsa elettroforetica dei campioni su gel d'agarosio.

L'allestimento dell'analisi prevede la preparazione del gel alla concentrazione dell'1% di agarosio e 1,5% di EtBr (1 g di agarosio e 1,5 $\mu\text{l}$  di EtBr in 100 ml di TAE, Tris-Acetil-EDTA 1X). Il bromuro di etidio (EtBr) è un intercalante degli acidi nucleici che permette la visualizzazione del DNA in quanto, quando

sottoposto a luce ultravioletta, emette fluorescenza. In questo modo è possibile visualizzare lo stato della migrazione del DNA che forma delle bande intrappolandosi nelle maglie del gel. La migrazione è possibile mediante applicazione di un potenziale elettrico alla base della vaschetta in cui è inserito il gel e il buffer di corsa TAE 1X.

I campioni sono caricati su gel utilizzando 5µl di DNA e 2µl di 6X DNA Loading Dye (Thermo scientific, USA).

Per capire la dimensione in bp delle bande dei campioni si utilizza un marcatore di peso molecolare (Ladder FastRuler® Middle Range DNA) caricato nel gel di agarosio, questo si separa all'interno delle maglie del gel formando delle bande con peso predefinito (Fig. 3.2).

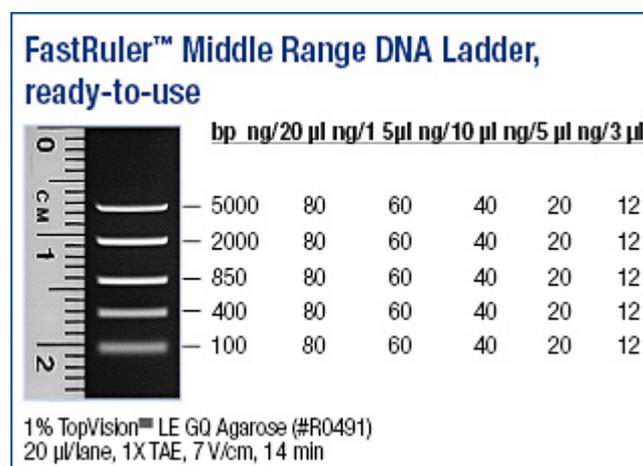


Figura 3.2 - Schema del bandeggio dato dal Marcatore di peso Molecolare FastRuler® Middle Range utilizzato per rapportare le bande risultanti da elettroforesi su gel dei campioni.

### 3.5 Purificazione amplificati con biglie magnetiche

I prodotti di PCR sono sottoposti a purificazione usando biglie magnetiche Agencourt AMPure XP (Beckman Coulter, USA) per eliminare eventuali prodotti aspecifici (primer, dimeri di primer) che in fase di sequenziamento potrebbero competere con i frammenti di interesse.

Il protocollo originale Roche/454 "Amplicon Library Preparation Method Manual - GS FLX Titanium Series" alla sezione "Library Purification" è stato modificato in alcune sue parti. È riportato di seguito il protocollo di esecuzione.

1. Aliquotare 22,5µl di acqua sterile in ciascuna provetta da 1,5ml
2. Trasferire 22,5 µl di campione in ciascuna delle provette precedentemente preparate
3. Aliquotare in ciascun tubino 45 µl di Ampure Beads XP dopo averle agitate al vortex per almeno un minuto (rapporto campione/biglie = 1:1 anziché 1:1,6)
4. Vortex per 5'', spin e incubare a TA (temperatura ambiente) per 10'

5. Trasferire le provette sul MPC (Magnetic Particle Collector) e incubare a TA per 5'
6. Senza toccare il pellet sulla parete posteriore della provetta eliminare tutto il surnatante
7. Togliere le provette dal MPC e aliquotare in ciascuna 25  $\mu$ l di buffer TE 1X
8. Vortex per 5'' (o finché il pellet non sarà completamente risospeso), spin e incubare sul MPC per 2' a TA
9. Aliquotare in nuove provette 25  $\mu$ l di acqua
10. Trascorso il tempo di incubazione, senza toccare il pellet sulla parete posteriore del tubino, trasferire i 25  $\mu$ l di surnatante nei tubini precedentemente preparate
11. Aliquotare in ciascuna proveta 50  $\mu$ l di Ampure Beads XP
12. Vortex per 5'', spin e incubare a TA per 10'
13. Trasferire le provette sul MPC e incubare a TA per 5'
14. Senza toccare il pellet sulla parete posteriore della provetta, eliminare tutto il surnatante
15. Lasciando le provette sul MPC, aliquotare 200  $\mu$ l di etanolo 70%
16. Vortex 5'' (il pellet non si risospenderà completamente, ma si può comunque procedere)
17. Rimettere le provette sui MCP e incubare 1' a TA
18. Eliminare tutto il surnatante, facendo attenzione a non toccare il pellet sulla parte posteriore della provetta
19. Ripetere i passaggi di lavaggio dal punto 15 al punto 18
20. Eliminare ogni traccia di etanolo dalle provette (anche la più piccola gocciolina sul fondo)
21. Brevissimo spin e rimettere le provette sul MPC per eliminare le ultime gocce
22. Mettere il MPC con le provette aperte nel forno alla temperatura di 35°C per 2'' circa. Il tempo necessario per l'asciugatura dei campioni varia molto, a seconda di fattori molto diversi tra loro, come ad esempio la temperatura dell'ambiente lavorativo e l'umidità in esso presente, la presenza o meno di gocce di etanolo sul fondo della provette, la concentrazione di DNA e la tipologia di campione. ATTENZIONE: il pellet di biglie non deve presentare crepe, segno di una asciugatura eccessiva. Controllare il pellet con luce per verificarne il grado di asciugatura.
23. Una volta asciutte chiudere i tubini e rimetterli sul MPC (i pellet non si asciugano mai nello stesso momento, quindi le provette i cui pellet si asciugano prima di altri, vanno tolte dal forno e poste ben chiuse su un rack a TA)
24. Aliquotare in ciascuna provetta 20  $\mu$ l di buffer TE 1X
25. Vortex fino a che il pellet non è completamente risospeso, spin brevissimo e incubare sul MPC a TA per 2'
26. Facendo attenzione a non toccare il pellet e a non aspirare nessuna biglia, pipettare 19  $\mu$ l del surnatante in provette pulite (da 1,5 ml)

### 3.6 Lab on a Chip (LOC)

Per *Lab-on-a-chip* si intendono le tecnologie di microfluidica, che sfruttano il principio del movimento



Figura 3.3 – Chip High Sensitivity (Agilent Technologies, Germany).

dei liquidi attraverso microstrutture per mezzo di campi elettrici o pressione/vuoto; tutto questo avviene all'interno di microchip (Fig. 3.3). Queste tecnologie offrono la possibilità di ridimensionare le tecniche analitiche dalla scala del laboratorio a quella di un chip: “Virtually any biochemical testing that can be done in a laboratory can theoretically be done on a chip” (da: “Agilent 2100 Bioanalyzer 2100 Expert User’s Guide - Edition: May 2005”). Le analisi LOC si basano sullo stesso principio della separazione delle molecole attraverso

l'elettroforesi su gel tradizionale, ma con una riduzione drastica del tempo di separazione e della quantità di campione utilizzata. Con lo strumento Agilent 2100 Bioanalyzer si possono ottenere informazioni sulla dimensione e quantificazione di campioni di DNA, RNA e proteine in formato digitale. Il chip è in grado di analizzare 11 campioni in una sola corsa. I pozzetti sono collegati tra di loro tramite una rete di microcanali in vetro che durante la preparazione del chip vengono riempiti con una matrice polisaccaridica e molecole intercalanti fluorescenti.

Una volta che i pozzetti e i canali sono stati caricati, il chip diventa un vero e proprio circuito elettrico integrato, le molecole dotate di carica come il DNA e l'RNA sono sottoposte ad una differenza di potenziale che le fa migrare. Esse si separano in base alle dimensioni (peso molecolare) e alla carica elettrica, grazie alla presenza della matrice che come il gel, funge da setaccio. I frammenti più corti migreranno quindi più velocemente di quelli più lunghi.

Le molecole fluorescenti si intercalano tra i filamenti di DNA e questi complessi vengono eccitati da un laser emettendo una fluorescenza che viene rilevata. I dati sono poi elaborati dal software dello strumento in immagini e in elettroferogrammi. Con l'aiuto di un marcatore che contiene frammenti di dimensioni e concentrazioni note, viene costruita una curva standard del tempo di migrazione in relazione alla lunghezza dei frammenti. Dal tempo di migrazione misurato per ogni frammento nel campione, se ne calcola la taglia.

### 3.7 Analisi Spettrofotometrica

L'analisi eseguita mediante l'uso dello spettrofotometro NanoDrop 8000 è utilizzata per misurare l'assorbanza di un campione a diverse lunghezze d'onda. A seconda della lunghezza d'onda utilizzata si possono misurare differenti classi di molecole: acidi nucleici (DNA e RNA), proteine, metaboliti secondari, sali. In base ai rapporti tra queste diverse molecole è possibile valutare la qualità di un campione di DNA e al tempo stesso calcolarne approssimativamente la concentrazione.

### 3.8 Analisi Fluorimetrica

Questo tipo di quantificazione permette di ottenere la concentrazione dei campioni utilizzando un principio diverso da quello del NanoDrop 8000. In questo caso si utilizza uno spettrofluorimetro, il Thermo Scientific NanoDrop 3300. Il PicoGreen® (Invitrogen, USA) è un fluoroforo intercalante degli acidi nucleici ed è utilizzato per la quantificazione di DNA a doppio filamento (dsDNA). Lo svantaggio principale delle spettroscopie a UV per la quantificazione del dsDNA è il contributo di segnali provenienti da DNA a singolo filamento (ssDNA) e altri contaminanti, come proteine e buffer di estrazione. Il reagente PicoGreen® elude tali contributi da sostanze interferenti esibendo un massimo di emissione a 530 nm quando si lega specificatamente al dsDNA (PicoGreen® non legato mostra fluorescenza minima nella soluzione).

### 3.9 PCR quantitativa

Con la PCR quantitativa è possibile sapere con esattezza il numero di molecole presenti nei diversi campioni attraverso l'utilizzo di una real time PCR e la costruzione di una curva standard.

#### 3.9.1 Principio analitico della PCR in real time

La PCR in real time si basa sull'utilizzo di una molecola fluorescente, il *SYBR Green I* (Roche), intercalante del DNA che emette un segnale luminoso rilevato in tempo reale ad ogni ciclo di amplificazione dallo strumento LightCycler® 480 (Roche, Mannheim, Germany). Le fasi della PCR sono quelle di una PCR tradizionale: infatti, durante la fase di annealing, i primer si ibridano al filamento del DNA target e formano DNA a doppio filamento nel quale si intercala il *SYBER Green I* che nella fase di estensione continua a inserirsi tra i filamenti di neo-sintesi aumentando il segnale fluorescente che viene misurato a 530 nm quando, al termine di questa crescita, raggiunge il suo valore massimo, ovvero quando il DNA è diventato a doppio filamento e la quantità massima del *SYBR Green I* si è intercalata. Ci sono diversi metodi per quantificare copie di DNA con la PCR in real time, con caratteristiche e utilità differenti, ma in questo contesto si prenderà in esame solo il metodo utilizzato e cioè quello della quantificazione assoluta.

#### 3.9.2 Profilo della reazione e *crossing point*

Il profilo della reazione, in un'analisi di quantificazione assoluta, contiene 3 fasi:

– una fase di *lag* iniziale

- una fase esponenziale (*log-linear*)
- una fase di *plateau* finale.

La fase iniziale (o fase *background*) inizia con l'avvio dell'amplificazione e dura fino a quando il segnale fluorescente del prodotto di PCR supera il livello della fluorescenza di fondo. Il ciclo durante il quale avviene il passaggio è detto “*crossing point*” (Cp) del campione; infatti il prodotto di PCR è osservabile solamente nel momento in cui il numero di molecole prodotte oltrepassa il limite di rilevazione della reazione (al Cp, sono presenti nella reazione approssimativamente da  $10^{11}$  a  $10^{12}$  molecole). Alla fase iniziale, succede la fase esponenziale che termina quando l'efficienza della reazione diminuisce e la reazione entra in fase di *plateau*.

### 3.9.3 La curva standard

Per determinare la quantità di molecole di un campione ignoto, è utilizzata una curva standard, costruita misurando la fluorescenza di almeno 5 campioni standard a concentrazione nota (Tabella 3.6).

La curva standard è ottenuta dall'interpolazione dei valori di fluorescenza misurata nei campioni standard.

<b>Nome standard</b>	<b>Copie/<math>\mu</math>l di DNA a doppio filamento</b>
Std1	5,00E + 07
Std2	5,00E + 06
Std3	5,00E + 05
Std4	5,00E + 04
Std5	5,00E + 03
Std6	5,00E + 02

Tabella 3.6 – Descrizione della quantità di copie di DNA per  $\mu$ l di Standard.

### 3.9.4 Il metodo della derivata seconda massima

Per determinare la concentrazione del campione il software dello strumento LightCycler® 480 fornisce diversi metodi di analisi. Quello utilizzato è il metodo della derivata seconda massima. In breve, esso identifica il Cp di un campione come il punto in cui la curva di fluorescenza del campione ha il massimo valore della derivata seconda (Figura 3.4) (da “LightCycler® 480 Instrument Operator’s Manual”).

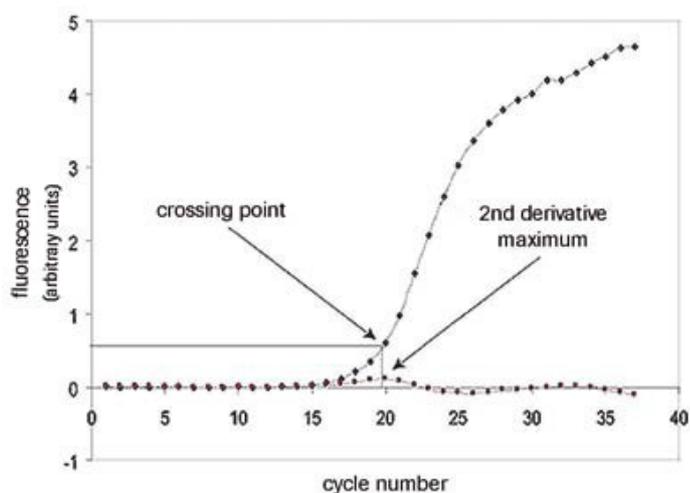


Figura 3.4 - Esempio di curva di reazione con Cp e derivata seconda massima, in ascissa è presente il numero di cicli di amplificazione in ordinata le unità di fluorescenza da “LightCycler® 480 Instrument Operator’s Manual”).

### 3.9.5 Allestimento PCR

Di seguito è riportata la sequenza di operazioni per la preparazione della piastra di PCR utilizzando il kit KAPA library quantification 454 Titanium (Kapabiosystems, Capetown, South Africa), descritta in “Technical Data Sheet - KAPA Library Quantification Kits For Roche 454 GS FLX platform”, dove è possibile reperire tutte le informazioni nello specifico.

La piastra utilizzata è la LightCycler® 480 Multiwell Plates. Lo schema di caricamento dei pozzetti è il seguente:

Reagente	µl per pozzetto		Campione		Standard		Controllo negativo
KAPA SYBER FAST qPCR Master contenente Primer Premix	12	+	4 µl	o	4 µl	o	4 µl
Molecular Biology Grade Water	4						EB o Molecular Biology Grade Water

Tabella 3.7 – Schema di caricamento della piastra per la RT – PCR, i reagenti sono dispensati in ogni pozzetto, a questi va aggiunto o il campione o lo standard o il controllo negativo.

È consigliabile caricare i campioni e gli standard in triplicato. I campioni sono preparati in due diluizioni diverse per avere un margine maggiore di rientrare nell’intervallo definito dagli standard. I campioni sono diluiti con il Dilution Buffer (EB), preparato alla concentrazione di 10 mM Tris – HCl a pH 8.0 + 0,05% Tween 20. La preparazione delle diluizioni è descritta sotto:

1:10 → 2 µl DNA + 18 µl EB

1:100 → 2 µl [1:10] + 18 µl EB

1:500 → 5 µl [1:100] + 20 µl EB

1:2.000 → 5 µl [1:500] + 15 µl EB

- Preparare prima di iniziare tutte le provette necessarie e siglarle.
- Dispensare in ognuna 18 µl di EB con il dispensatore automatico, tranne in quelle che conterranno le diluizioni 1:500 e 1:2000.
- Procedere con l'aggiunta del campione ricordandosi di mescolare bene (vortex) ogni volta che si passa da una diluizione all'altra.

Al termine dell'allestimento coprire la piastra con LightCycler® 480 Sealing Foil, per evitare l'evaporazione e caricare la piastra sullo strumento seguendo il profilo termico riportato in Tabella 3.8.

*Programma PCR:*

95°C x 5'	} x 35 cicli	Denaturazione iniziale
95°C x 30''		Denaturazione
60°C x 45''		Annealing – Estensione – Acquisizione dati

*Tabella 3.8 - Protocollo di amplificazione per PCR in real time (per frammenti di dimensioni superiori a 700 bp, è necessario aumentare il tempo di annealing/estensione a 90 sec.).*

### 3.10 PCR in Emulsione

Per realizzare un sequenziamento di qualità i protocolli Roche consigliano di arricchire le biglie con i frammenti di DNA in una percentuale dell'8%. Per ottenere questa percentuale è necessario trovare il giusto rapporto tra il numero di molecole e il numero di biglie da arricchire, attraverso una titolazione. Questa operazione consiste nell'eseguire una prova iniziale utilizzando differenti rapporti tra molecole di DNA e numero di biglie. Una volta determinato il rapporto migliore, questi campioni sono processati secondo il protocollo di sequenziamento prescelto. Un elemento fondamentale per la buona riuscita della titolazione è il corretto calcolo dei volumi di campione da aggiungere alle biglie per ottenere il numero di molecole per biglia corretto. La formula che permette di calcolare i µl di campione è:

$\mu\text{l di campione} = (\text{n}^\circ \text{ di copie per biglia} * \text{n}^\circ \text{ di biglie per Small Volume}) / \text{concentrazione libreria}$

Per valutare il risultato della titolazione e scegliere quindi i parametri corretti per l'esecuzione della vera e propria PCR in emulsione, vengono attuate alcune verifiche. Durante i passaggi intermedi è

eseguito un primo controllo del numero di biglie recuperate dall'emulsione e un secondo verificando la quantità di biglie arricchite con DNA, tramite lo strumento conta particelle *Cell and Particle Counter* (Beckman Coulter, USA). Queste informazioni sono elaborate in una tabella che aiuta a valutare i risultati della titolazione.

La PCR in emulsione consente di amplificare la libreria in maniera controllata, giocando sulla concentrazione di DNA in reazione. In particolare la emPCR-SV (PCR in emulsione-Small Volume), può essere definita come una curva di taratura: sono allestite tre reazioni di PCR in emulsione, uguali in tutto tranne che per il numero di molecole di DNA di partenza. Il DNA è legato a delle biglie di silice grandi circa 15 µm; è allestita la PCR in emulsione; le biglie, al termine della PCR, sono recuperate e arricchite; l'arricchimento permette di ottenere, dal totale, soltanto le biglie che portano una o più molecole di DNA attaccate. Le percentuali di arricchimento finale oscillano in genere tra 5 e 20%. Un arricchimento ottimale è di circa il 8-10%. Generalmente sono allestiti tre punti di titration a 0,1; 0,25 e 0,5 molecole per biglia, ma ovviamente possono variare in base al campione e al tipo di libreria (shotgun o ampliconi).

Di seguito sono illustrati i passaggi descritti dal protocollo "emPCR Amplification Manual – Lib-A SV - GS FLX Titanium Series – June 2013":

Punti da testare: 0.1 – 0.25 – 0.5

➔ Per ogni punto calcolare la quantità di DNA da prelevare: preparare le diluizioni.

$$\mu\text{l di Library per punto} = \frac{\text{molecole per biglia} \times 2 \cdot 10^6 \text{ biglie per punto (n° biglie)}}{\text{Concentrazione Library (molecole/ } \mu\text{l)}}$$

Se necessario preparare una diluizione della library in modo tale da ottenere un volume compreso tra 1 e 10 µl.

$$\text{Punto 0,1} = (0,1 \times 2 \cdot 10^6) / 1 \cdot 10^5 = 2 \mu\text{l}$$

$$\text{Punto 0,25} = (0,25 \times 2 \cdot 10^6) / 1 \cdot 10^5 = 5 \mu\text{l}$$

$$\text{Punto 0,5} = (0,5 \times 2 \cdot 10^6) / 1 \cdot 10^5 = 10 \mu\text{l}$$

➔ Preparare l'olio: 3 provette, disporle nei supporti appositi del TissueLyser. Far partire lo strumento per 2 min a 25 Hz (programma: P1).

➔ Preparare la Mock. Per ogni tubino ne servono 290 µl (per 3 tubini:  $V_T = 870 \mu\text{l}$ ). Lo stock è 5X, quindi prelevare 174 µl di Mock 5X + 696 µl H<sub>2</sub>O.

Aggiungere 290 µl di Mock appena preparata ad ogni tubino di olio. Invertire i tubini alcune volte per mescolare bene e posizionare nuovamente sul TissueLyser per 5 min a 25 Hz (P2).

→ Preparare la MIX 1X:

Mol. Bio. Grade Water	75 µl
emPCR Additive	93,75 µl
5X Amplification Mix	48,75 µl
Amplification Primer	14,375 µl
emPCR Enzyme Mix	12,5 µl
PPIase	0,3125 µl
Totale	244,6875 µl

*Tabella 3.9 - Schema di preparazione della Mix di reazione per la emPCR.*

- Preparare il Wash Buffer: serve per attivare le biglie, le prepara a legare il DNA. Lo stock è diluito 10X quindi sarà necessario diluirlo 10 volte.
- Preparare le BIGLIE. Vortex dei tubini in cui sono contenute, trasferire 70 µl di DNA Capture Beads per ogni punto ( $70 \times 3 = 210 \mu\text{l}$ ) in una provetta nuova. Ogni 70 µl corrispondono 2.000.000 di biglie, quante ne occorrono per punto. Centrifugare x 10 sec da entrambi i lati per far sedimentare bene le biglie, rimuovere il surnatante. Effettuare 2 lavaggi con 1 ml di Wash Buffer. Ogni volta vortex e centrifuga da entrambi i lati e rimuovere il surnatante. Risospendere le biglie in 40 µl di Wash Buffer per punto ( $40 \times 3 = 120 \mu\text{l}$ ). Pesare la provetta (precedentemente calcolata la tara), fare la differenza e dividere il peso del contenuto in 3, dispensare tale quantità in ogni punto. Distribuire in 3 provette nuove, spin e rimuovere il surnatante.
- Unire alle provette contenenti le biglie il DNA e la MIX. Aliquotare 215 µl di Mix in ogni punto.
- Successivamente, il contenuto delle provette appena preparate (biglie + DNA + Mix) deve essere trasferito (no vortex, solo risospendere con il puntale) nei tubini contenenti olio e Mock (precedentemente siglati). Invertire i tubini alcune volte per mescolare. Caricarli sul TissueLyser a 15 Hz per 5 min (P3). Questo passaggio effettua l'emulsione. Ogni biglia dovrebbe essere incorporata in una bolla di olio così da permettere una PCR controllata.
- Ora bisogna dispensare il contenuto di ogni tubino in una piastra PCR da 96 pozzetti (Fig. 3.5) mediante un pipetta automatica tarata su 100 µl (si riempiono circa 10 pozzetti). Dispensare con un gesto deciso ma non troppo violento perché l'emulsione potrebbe rompersi. Lasciare una fila vuota tra un punto e l'altro (che servirà per il recupero) e ricordarsi di siglare la piastra. Chiudere i pozzetti utilizzati con i tappi appositi e fare attenzione a tenere la piastra in posizione orizzontale senza agitarla.

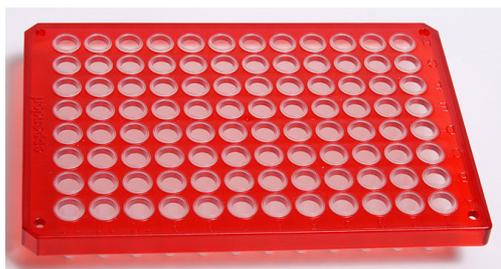


Figura 3.5 - Piastra di caricamento della emPCR.

➔ Caricare la piastra nel termociclatore. Programma di amplificazione:

- 1x 4 minuti a 94°C
- 50x 30 secondi a 94°C
- 4.5 minuti a 58°C
- 30 secondi a 68°C
- ∞ ➔ 10°C

### 3.10.1 Recupero delle Biglie

Portare a temperatura ambiente l'Enanching Fluid XT e l'Annealing Buffer XT.

Preparare 6 provette (3 per il recovery e 3 per l'arricchimento), siglarle e pesarle annotando il peso.

Preparare 3 siringhe (siglarle), 3 aghi (diametro 16) e 3 filtri. Preparare le siringhe avvitate all'ago e una vaschetta di isopropanolo (rompe l'emulsione) vicino ad una cappa aspirante.

- Lavaggi iniziali
  - 1) Prelevare 100 µl di isopropanolo con la multicanale e dispensarli nei primi 12 pozzetti. Prelevare con la siringa solo dai primi 10 pozzetti (quelli con il campione) e ogni 3-4 rivolgere la siringa verso l'alto, aspirare e riportare lo stantuffo il più vicino possibile alla siringa senza far fuoriuscire niente. Ripeto questa operazione altre 2 volte (per ottenere 3 lavaggi). L'ultima volta si preleva anche dal pozzetto 11 e 12. Al termine aspirare il più possibile, togliere l'ago e inserire il filtro al suo posto.
  - 2) Ripetere l'operazione per ogni punto.
- Lavaggi delle biglie
  - 1) Vortex della siringa con il filtro verso il basso per 5 sec e poi rimuovere il contenuto (fino in fondo).
  - 2) Invertire la siringa, riempire fino a 3 ml di aria e aspirare 3 ml di isopropanolo, vortex per 5 secondi e eliminare il contenuto tramite la siringa. Ripetere 3 volte questo passaggio.
  - 3) Invertire la siringa, riempire fino a 3 ml di aria e aspirare 3 ml di Enanching Buffer, vortex per 5 secondi e eliminare il contenuto (come prima). Ripetere per 3 volte.

- 4) Invertire la siringa, riempire fino a 3 ml di aria e aspirare 1 ml di Enanching Buffer, vortex per 3 secondi, poi girare subito la siringa con il filtro verso l'alto e nel frattempo portare il pistone della siringa il più lontano possibile dal filtro (senza farlo uscire). Con la siringa in questa posizione togliere il filtro e riporre il contenuto della siringa nella provetta corrispondente, rimettere il filtro. Ripetere questa operazione 2 volte.
- 5) Centrifugare le provette da una parte e dall'altra ed eliminare il surnatante.
- 6) Ripetere il punto 4) e 5) altre 2 volte e la terza solo un passaggio di enanching.

### 3.10.2 Arricchimento delle Biglie

Preparare il blocco - fornello acceso su 65°C.

Preparare la Melt Solution nella quantità desiderata in una proporzione di 125 µl di NaOH (soda) in 9,875 ml di Molecular Grade Water.

N.B. Dal momento che si aggiunge la soda al campione le operazioni non devono durare più di 10 min, altrimenti potrebbe degradare il DNA. La soda serve per denaturare il dsDNA.

- 1) Aggiungere 1 ml di Melt Solution per provetta e vortex. Incubare per 2 minuti a temperatura ambiente. Spin – rotate – spin ed eliminare il surnatante. Ripetere questo passaggio.
- 2) Aggiungere 1 ml di Annealing Buffer XT (è un coadiuvante, neutralizza la soda e aiuta a legare il primer il DNA) per punto e vortex. Spin – rotate – spin ed eliminare il surnatante. Ripetere questo passaggio.
- 3) Risospendere in 200 µl di Annealing Buffer.
- 4) Pesare le provette e fare la differenza con le tare prese all'inizio.

### 3.10.3 Conta delle Biglie

Prelevare 3 µl della provetta (per prelevare il giusto numero di biglie occorre agitare bene la provetta e prelevare a metà della profondità), dispensarle nell'apposito bicchierino per la conta contenente il buffer adeguato. Effettuare la conta e dal numero ottenuto applicare tale formula:

$$\% \text{ Bead Recovery} = \frac{\text{Numero di biglie recuperate}}{\text{Biglie totali}} \times 100$$

La percentuale che si ottiene dovrebbe essere compresa tra l'80 e il 90%.

- 5) Spin – rotate – spin ed eliminare il surnatante.
- 6) Aggiungere 30 µl di Annealing Buffer e 3 µl di primer per tubino.
- 7) Posizionare le provette nel blocco termico a 65°C per 5 min (il primer si attacca).

- 8) Spostare le provette in ghiaccio per 2 min (blocca l'annealing del buffer).
- 9) Aggiungere 500 µl di Enanching Fluid per tubino e vortex. Spin – rotate – spin ed eliminare il surnatante. Ripetere nuovamente questo passaggio.
- 10) Aggiungere in 800 µl di Enanching Fluid per tubino e vortex.

Preparazione delle biglie magnetiche:

- Vortex del tubino contenente l'Enrichment Beads per 1 min per risospendere il contenuto.
- Prelevare 20 µl di biglie per punto, in questo caso 20 x 3 = 60 µl e dispensarli in una nuova provetta.
- Posizionare la provetta sul Magnetic Particle Concentrator (MPC) e aspettare per 5 min.
- Eliminare il surnatante.
- Aggiungere 500 µl di Enanching Fluid, vortex e posizionare nuovamente sul MPC.
- Invertire il supporto alcune volte per far aderire al meglio le biglie al magnete.
- Eliminare il surnatante.
- Ripetere questo passaggio di lavaggio un'altra volta.
- Dopo aver eliminato il surnatante risospendere nel volume iniziale di biglie prelevate (20 x 3), cioè 60 µl.

Arricchimento di biglie che hanno legato il DNA:

- Aggiungere 20 µl di biglie per campione, ogni volta vortex e cambiare il puntale per ogni campione.
- Vortex e posizionare le provette sul Rotator a temperatura ambiente per 5 minuti.
- Mettere le provette sul MPC per 5 minuti e nel frattempo tirare giù dal tappo le gocce.
- Eliminare il surnatante da ogni tubino facendo attenzione a non toccare il pellet.
- Fare dei lavaggi con Enanching Fluid XT fino a quando non si vedranno più biglie bianche scendere, i lavaggi sono fatti come segue:
  - o Aggiungere 500 µl di Enanching Fluid XT per tubino,
  - o Rimuovere le provette dal MPC e vortex,
  - o Riposizionare i tubini sul PMC, invertire il PMC alcune volte, rimuovere le gocce dal tappo e metterle nella provetta,
  - o Eliminare il surnatante facendo attenzione a non smuovere il pellet sulla parete,
  - o Ripetere da 3 a 6 volte fino a quando non si vedrà più il fantasma di biglie bianche.

#### 3.10.4 Recupero delle biglie arricchite

N.B. i passaggi con la soda non devono superare i 10 minuti di tempo.

Nei passaggi seguenti si utilizzano le provette che si sono preparate in precedenza siglate e tarate.

1. Aggiungere 500 µl di Melt solution, vortex e posizionare le provette sul MPC, invertire alcune volte il PMC,
2. Prelevare e dispensare all'interno della provetta le eventuali goccioline sul tappo,
3. Quando il pellet è formato trasferire il surnatante nella nuova provetta,
4. Far scivolare 3 gocce di soda sul pellet con le biglie magnetiche, si vedranno scendere dal pellet le biglie bianche, prelevarle e trasferirle in un nuovo tubino (quella appena riempita nel passaggio precedente),
5. Ripetere dal punto 1. al punto 4.
6. Spin – rotate – spin ed eliminare il surnatante,
7. Aggiungere 500 µl di Annealing Buffer XT (serve per neutralizzare la soda) per provetta e vortex per 5 sec,
8. Spin – rotate – spin ed eliminare il surnatante,
9. Ripetere i passaggi 7. e 8. altre 2 volte,
10. Aggiungere 100 µl di Annealing Buffer XT per ogni provetta,
11. Pesare le provette e fare la differenza con le tare prese all'inizio delle provette nuove.

Ora si procede con la conta. Si eseguono i medesimi passaggi descritti al punto 3.10.3 e la formula che si utilizza è la seguente:

$$\% \text{ Bead Enrichment} = \frac{\text{Numero di biglie arricchite} \times 100}{\text{Biglie totali}}$$

I valori percentuali di biglie arricchite se sono compresi tra il 5 e il 15% sono indicativi di librerie che solitamente danno un buon risultato in sequenziamento.

Con la titolazione si ottiene un valore “in piccolo” (SmallVolume) che rispecchia quello che dovrebbe accadere “in grande” (Medium o Large Volume): ci permette quindi di capire quale sia la concentrazione di DNA corretta da inserire in reazione per ottenere un numero corretto di biglie “cariche e funzionanti” da caricare sul sequenziatore.

In base al numero di reads che si vogliono ottenere si decide se procedere, dopo la titolazione, con l'amplificazione in emulsione Small Volume, Medium Volume o Large Volume (Tabella 3.10). I protocolli di tutte e tre le procedure sono forniti dal produttore (“emPCR Amplification Manual – LibA MV (o LV) – GS FLX Titanium series – June 2013”) e prevedono una PCR in emulsione e successivo recupero e arricchimento delle biglie nella quali è avvenuta amplificazione. A tali biglie contenenti l'amplicone d'interesse si lega il primer che servirà per il sequenziamento.

Region Size	Regions per PTP Device	Reads per Region (x 10 <sup>3</sup> )
Large	2	> 255
Medium	4	> 100
Medium/Small	8	> 50
Small	16	> 10

Tabella 3.10 – Throughput per librerie ad ampliconi e indicazioni per la maschera di divisione (gasket).

### 3.11 Pirosequenziamento

Dopo essersi assicurati che il numero di biglie arricchite sia del numero desiderato in base al numero di reads che si vogliono ottenere, si può passare al sequenziamento. Questo prevede la preparazione del vetrino PTP con vari strati, tra cui delle biglie che fungono da marker a sequenza nota, si caricano le biglie arricchite con il campione, si prepara la macchina caricando tutti i buffer necessari e si posiziona il vetrino nella camera di lettura. A questo punto è necessario far partire il programma di sequenziamento. Le istruzioni complete per l'esecuzione della procedura sono riportate nel protocollo "Sequencing Method Manual for the GS FLX+ Instrument - GS FLX+ Series — XL+ Kit – June 2013". La descrizione delle reazioni che avvengono durante il pirosequenziamento è riportata nell'Introduzione.



## 4. RISULTATI E DISCUSSIONE

### 4.1 Estrazione del DNA genomico

Per le operazioni che si sono svolte in laboratorio sono stati usati i 15 campioni provenienti dalla foresta Mwanihana, mentre i restanti 16 campioni dalla foresta Magombera erano già stati processati in precedenza.

Per prima cosa i 15 campioni fecali provenienti dalla foresta Mwanihana sono stati sottoposti ad estrazione del DNA genomico totale mediante il kit commerciale FastDNA™ SPIN Kit for Feces (MP Biomedicals, CA, USA) secondo il protocollo del produttore (vedi Materiali e Metodi).

I campioni estratti sono stati sottoposti ad analisi spettrofotometrica mediante Nanodrop 8000 che ha permesso di verificare la qualità e la concentrazione del DNA genomico.

Le rese ottenute sono comprese tra 9 e 25 ng/μl (Fig. 4.1) e per la maggior parte dei campioni si è ottenuta una concentrazione sufficiente per l'amplificazione delle regioni V1-V3 del 16S rDNA e i passaggi successivi. Le analisi spettrofotometriche consentono di valutare il rapporto di assorbanza 260/280, utile per verificare la qualità dell'estratto. Un rapporto superiore o uguale a 1,8 è generalmente buono, se il rapporto fosse più basso potrebbe indicare la presenza di proteine o altri contaminanti. Dalla Tabella 4.1 si evince che quasi tutti i campioni hanno un valore di circa 1,8, solo in pochi casi il valore scende sotto 1,6.

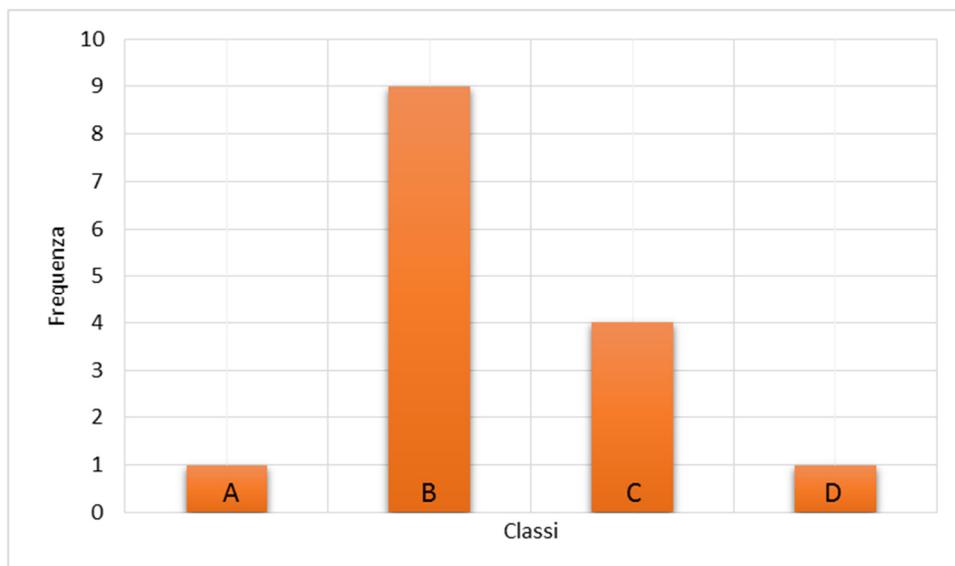


Figura 4.1- Frequenza di distribuzione delle concentrazioni dei campioni di DNA estratti da matrice fecale mediante kit commerciale. Le classi di distribuzione corrispondono a: A=0-9 ng/ul, B=10-19 ng/ul, C=20-29 ng/ul, D=30-39 ng/ul.

ID Campione	Conc. (ng/μl)	A260	A280	260/280
Mw1	10,88	0,22	0,13	1,65
Mw2	11,48	0,23	0,14	1,68
Mw3	16,41	0,33	0,18	1,81
Mw4	24,08	0,48	0,27	1,78
Mw5	16,41	0,33	0,19	1,73
Mw6	32,18	0,64	0,38	1,71
Mw7	14,09	0,28	0,18	1,53
Mw8	25,77	0,52	0,30	1,71
Mw9	10,96	0,22	0,14	1,54
Mw10	19,54	0,39	0,25	1,59
Mw11	14,23	0,28	0,17	1,71
Mw12	11,98	0,24	0,14	1,68
Mw13	9,40	0,19	0,12	1,62
Mw14	24,55	0,49	0,29	1,71
Mw15	15,80	0,32	0,21	1,54

Tabella 4.1 – Media aritmetica dei risultati ottenuti dall'analisi in triplicato di ogni campione con spettrofotometro Nanodrop 8000. Per ogni campione sono riportati: ID campione, MID associato, concentrazione in ng/μl, assorbanza a 260 nm, assorbanza a 280 nm e rapporto assorbanze 260/280.

## 4.2 Produzione degli ampliconi

Basandosi sulle concentrazioni ottenute dal NanoDrop 8000, 3 dei 15 campioni in esame: Mw4, Mw8 e Mw12, sono stati sottoposti ad amplificazione mediante PCR della regione target V1-V3 del gene 16S rDNA (vedi Materiali e Metodi), a due differenti concentrazioni finali di DNA: 20 e 50 ng, e per valutare quale fosse la quantità di DNA minima necessaria.

I prodotti di amplificazione sono poi stati caricati su gel d'agarosio all'1% e sottoposti a corsa elettroforetica. Sono stati caricati 5 µl di amplificato per campione e 2 µl di 6X DNA Loading Dye (Thermo scientific,USA).

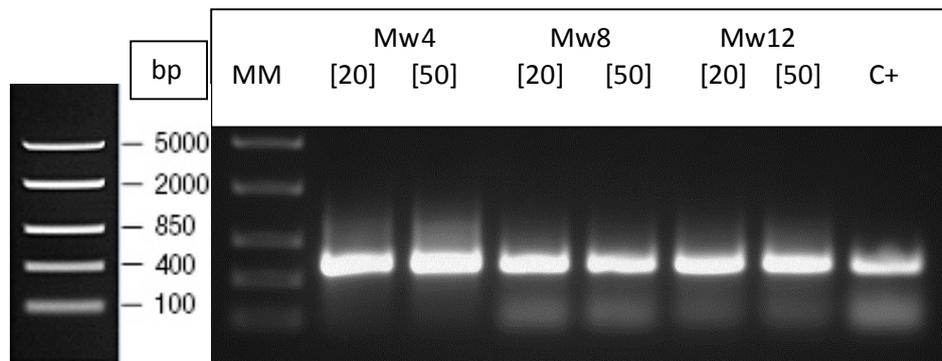


Figura 4.2 – Corsa elettroforetica dei campioni: nel primo pozzetto è stato caricato il marcatore di peso molecolare (MM) (FastRuler™ Middle Range). Successivamente i campioni Mw4, Mw8, Mw12 alle due concentrazioni ed infine il controllo positivo.

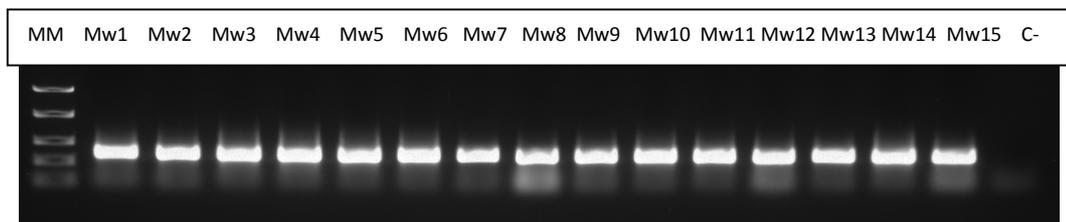


Figura 4.3 – Corsa elettroforetica degli ampliconi V1 – V3 dei campioni. A partire da sx: marcatore di peso molecolare(MM), Mw1, Mw2, Mw3, Mw4, Mw5, Mw6, Mw7, Mw8, Mw9, Mw10, Mw11, Mw12, Mw13, Mw14, Mw15 e controllo negativo.

Il risultato della PCR è stato soddisfacente, tutti i campioni hanno amplificato ad entrambe le concentrazioni (Fig. 4.2). Confrontando le bande ottenute dei campioni con quelle del marcatore di peso molecolare si può ricavare la dimensione della banda. L'analisi elettroforetica su gel d'agarosio ha evidenziato la presenza di bande del peso molecolare atteso (595 pb).

Lo stesso schema di reazione di amplificazione è stato esteso a tutti gli altri campioni utilizzando 20 ng di DNA genomico come template per diminuire eventuali interferenze date dal DNA genomico stesso. Le amplificazioni sono state eseguite in triplicato tecnico e sottoposte ad analisi elettroforetica su gel d'agarosio (vedi Fig. 4.3).

I risultati ottenuti sono buoni e corrispondono al valore atteso di circa 600bp, quindi sono stati uniti i triplicati di ogni campione in un'unica provetta per i passaggi successivi di preparazione dei campioni per il pirosequenziamento.

#### 4.3 Purificazione e validazione degli ampliconi

I prodotti di amplificazione sono stati sottoposti a purificazione mediante biglie magnetiche AMPure Beads (vedi Materiali e Metodi) per eliminare i primer non incorporati, dimeri di primer ed eventuali residui della reazione di PCR.

Per verificare l'avvenuta purificazione e l'esatta dimensione degli amplificati i campioni sono stati analizzati mediante elettroforesi capillare su chip High Sensitivity (Agilent Technologies, Germany) secondo il protocollo del produttore (vedi Materiali e Metodi).

In Figura 4.4 è riportato il profilo della corsa riferita al campione Mw7 e come si può notare dall'elettroferogramma, la lunghezza del frammento è di 581 bp come atteso. Inoltre si può vedere che non ci sono tracce di primer o di dimeri di primer, né prodotti aspecifici. Tutti i campioni sono stati validati mediante elettroforesi capillare High Sensitivity e sottoposti alla successiva quantificazione mediante PCR quantitativa.

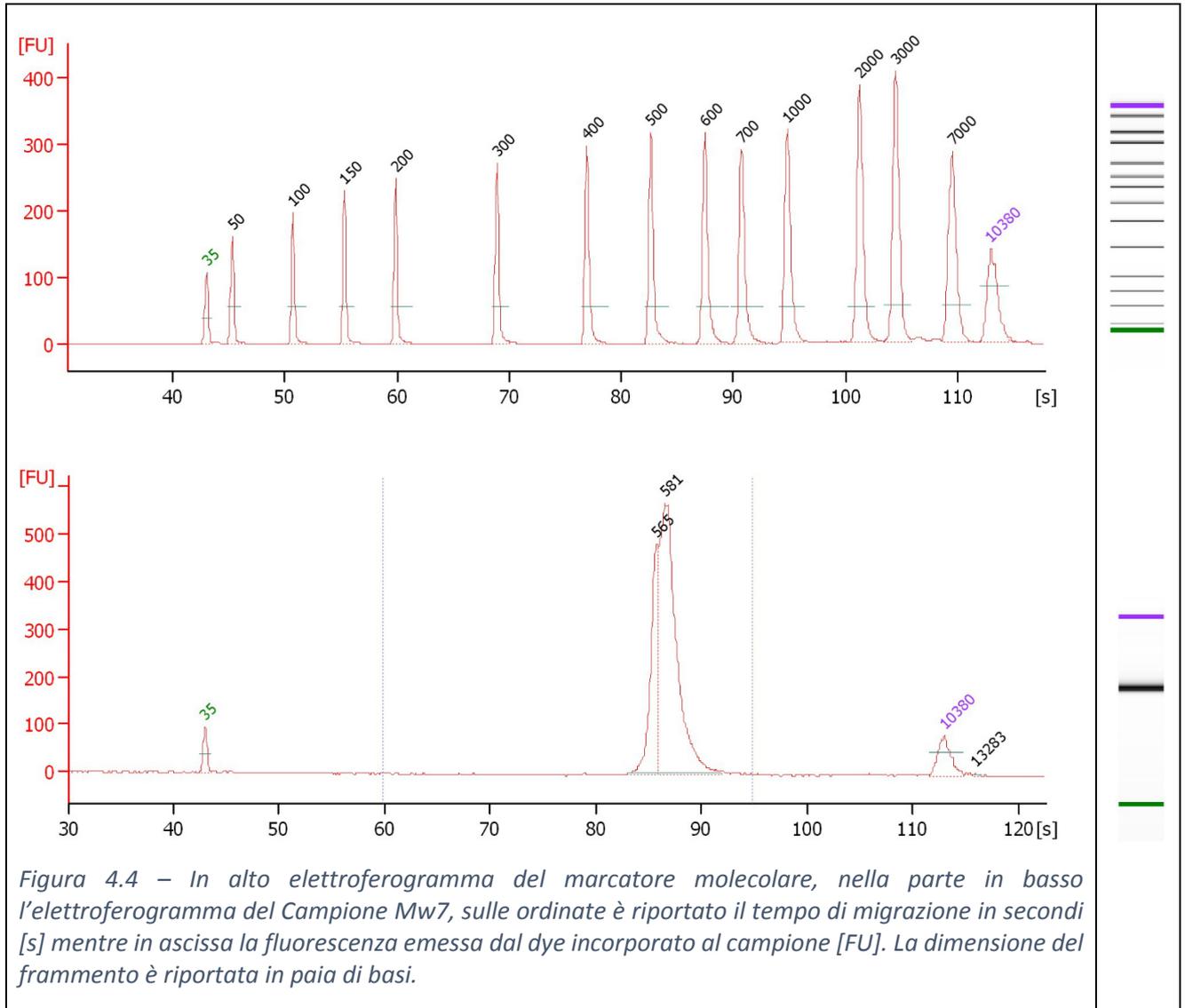


Figura 4.4 – In alto elettroferogramma del marcatore molecolare, nella parte in basso l'elettroferogramma del Campione Mw7, sulle ordinate è riportato il tempo di migrazione in secondi [s] mentre in ascissa la fluorescenza emessa dal dye incorporato al campione [FU]. La dimensione del frammento è riportata in paia di basi.

Gli ampliconi sono stati preventivamente quantificati mediante saggio fluorimetrico con PicoGreen Quant-IT assay (Invitrogen, USA) al fine di valutarne la concentrazione iniziale per definire le diluizioni da utilizzare in real time PCR (vedi Materiali e Metodi). I valori ottenuti rientrano nell'intervallo di concentrazione pari a 20 – 55 ng/μl e i campioni hanno valori più rappresentati che vanno da 29 a 39 ng/μl (Fig. 4.5).

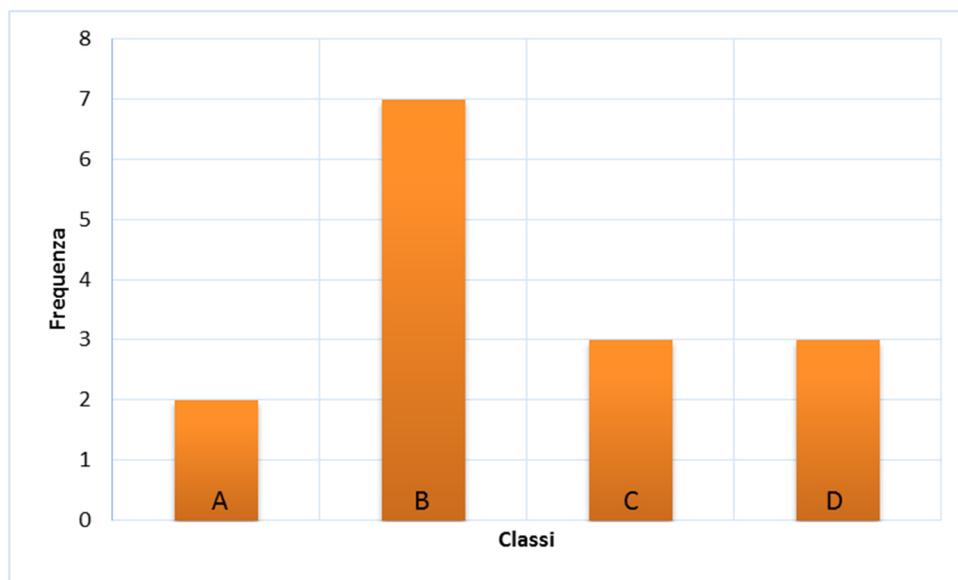


Figura 4.5 - Frequenza di distribuzione delle concentrazioni degli ampliconi risultanti da amplificazione della regione ipervariabile V1 – V3. Le classi di distribuzione corrispondono a: A=0-29 ng/ul, B=30-39 ng/ul, C=40-49 ng/ul, D=50-59 ng/ul.

#### 4.4 Quantificazione assoluta degli ampliconi

La conoscenza dell'esatto numero in molecole di ampliconi è fondamentale per la buona riuscita del pirosequenziamento, infatti una sovra/sotto stima di questo valore porterebbe ad errori di valutazione del rapporto biglie-molecole durante la fase di PCR in emulsione, causando una riduzione della capacità e qualità del sequenziamento (vedi Materiali e Metodi). La PCR quantitativa (qPCR) in real time è considerata il metodo *gold standard* per la quantificazione del numero di molecole amplificabili (da "Technical Data Sheet - KAPA Library Quantification Kits For Roche 454 GS FLX platform").

L'allestimento della qPCR ha previsto innanzitutto la preparazione delle diluizioni dei campioni affinché fossero comprese all'interno dei valori delle curva di taratura (vedi Materiali e Metodi).

Sono state utilizzate due diverse diluizioni del DNA di partenza: 1:500 e 1:2000, allestendo le reazioni in triplicato per tutti i campioni; i dati ottenuti sono stati quindi analizzati mediante il software LightCycler® 4.0 (Roche, Mannheim, Germany).

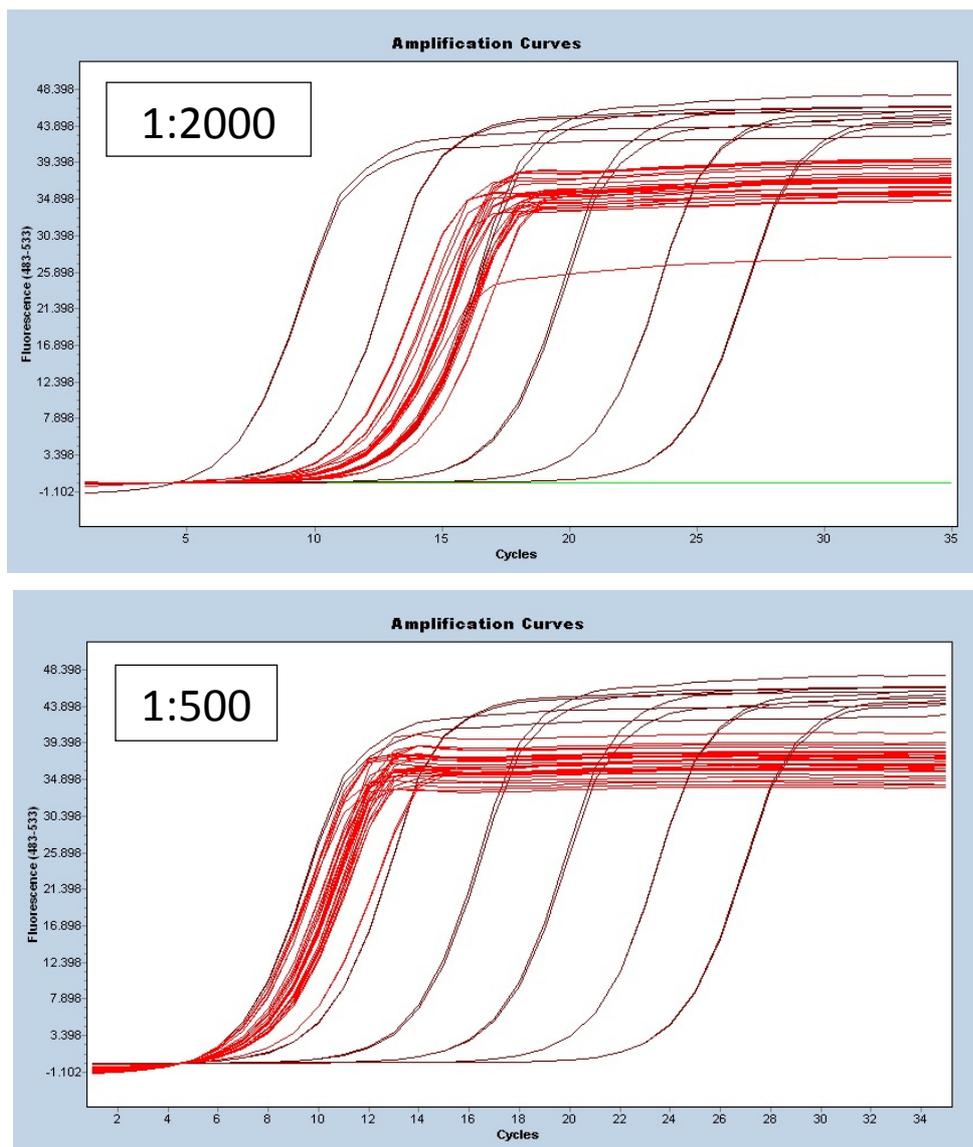


Figura 4.6 - Curve di amplificazione dei campioni (in rosso) diluiti 1:500 (in alto) e 1:2000 (in basso) e degli standard (in marrone), in ordinata sono riportati i cicli di amplificazione mentre in ascissa le unità di fluorescenza. In base al numero di Cp, i campioni nelle due diluizioni, sono in rapporto tra di loro.

Le curve di amplificazione dei campioni in entrambe le diluizioni rientrano ampiamente nei valori della curva di taratura, questo consente al software di poter calcolare il numero di copie di molecole presenti in ogni campione. Successivamente i valori ottenuti dal software sono stati rielaborati per ottenere il volume necessario alla composizione equimolare della libreria di sequenziamento (Tab. 4.2).

Sample ID	Size (bp)	Avg. Copies/rxn	Size adjusted copies/rxn	Copies in library stock (copies/ul)	ul lib TQ	ul TE 1X	[copies/ul]
Mw1 1:500	600	1,61E+07	1,23E+07	1,54E+09	2,0	28,0	1E+08
Mw2 1:500	600	3,98E+07	3,05E+07	3,81E+09	0,8	29,2	1E+08
Mw3 1:500	600	3,27E+07	2,50E+07	3,13E+09	1,0	29,0	1E+08
Mw4 1:500	600	2,54E+07	1,94E+07	2,43E+09	1,2	28,8	1E+08
Mw5 1:500	600	8,28E+06	6,34E+06	7,93E+08	3,8	26,2	1E+08
Mw6 1:500	600	2,33E+07	1,79E+07	2,23E+09	1,3	28,7	1E+08
Mw7 1:500	600	2,15E+07	1,65E+07	2,06E+09	1,5	28,5	1E+08
Mw8 1:500	600	1,74E+07	1,33E+07	1,66E+09	1,8	28,2	1E+08
Mw9 1:500	600	1,71E+07	1,31E+07	1,64E+09	1,8	28,2	1E+08
Mw10 1:500	600	1,93E+07	1,48E+07	1,85E+09	1,6	28,4	1E+08
Mw11 1:500	600	2,32E+07	1,77E+07	2,22E+09	1,4	28,6	1E+08
Mw12 1:500	600	2,94E+07	2,25E+07	2,81E+09	1,1	28,9	1E+08
Mw13 1:500	600	2,08E+07	1,59E+07	1,99E+09	1,5	28,5	1E+08
Mw14 1:500	600	2,33E+07	1,79E+07	2,23E+09	1,3	28,7	1E+08
Mw15 1:500	600	3,17E+07	2,43E+07	3,03E+09	1,0	29,0	1E+08

*Tabella 4.2 – Numero di molecole per reazione ottenute dal software e calcoli che hanno portato ad ottenere l'esatto numero di molecole in ogni campione. Nella prima colonna sono presenti i nomi identificativi dei campioni, a seguire: la lunghezza del frammento d'amplificazione, il numero di molecole per reazione, il numero di molecole relazionato alla dimensione reale del frammento, le copie di DNA per  $\mu$ l. Le ultime tre colonne indicano la quantità in  $\mu$ l da prelevare di ogni campione e di buffer TE 1X per ottenere la concentrazione finale di  $1 \times 10^8$  molecole per  $\mu$ l di ogni campione.*

Nella prima colonna della Tabella 4.2 è riportata la "Size adjusted copies/rxn", il numero di molecole per reazione ottenute dal software (Avg. Copies/rxn) è moltiplicato per la lunghezza dello standard (460 bp) e diviso per la lunghezza del campione stesso (600 bp).

Nella colonna "Copies in library stock (copies/ $\mu$ l)" è riportato il numero effettivo di copie di DNA per  $\mu$ l.

Infine nelle ultime tre colonne è riportato il volume di campione e di buffer TE 1X per ottenere un volume finale di 30  $\mu$ l in cui saranno presenti  $1 \times 10^8$  molecole per  $\mu$ l.

Successivamente tutti i campioni sono stati diluiti alla medesima concentrazione di  $1 \times 10^7$  molecole per  $\mu$ l e riuniti in una libreria equimolare (pool).

Il pool di campioni è stato sottoposto ad una doppia purificazione mediante biglie magnetiche Agencourt AMPure XP (Beckman Coulter, USA) e nuovamente quantificato con la real time PCR (Fig. 4.7), utilizzando gli stessi standard a concentrazione nota e la stessa tipologia di diluizione, vedi Materiali e Metodi. Al termine della quantificazione si ottiene la reale concentrazione del pool, con l'esatta quantità di molecole amplificabili durante la PCR in emulsione, come indicato in Tabella 4.3.

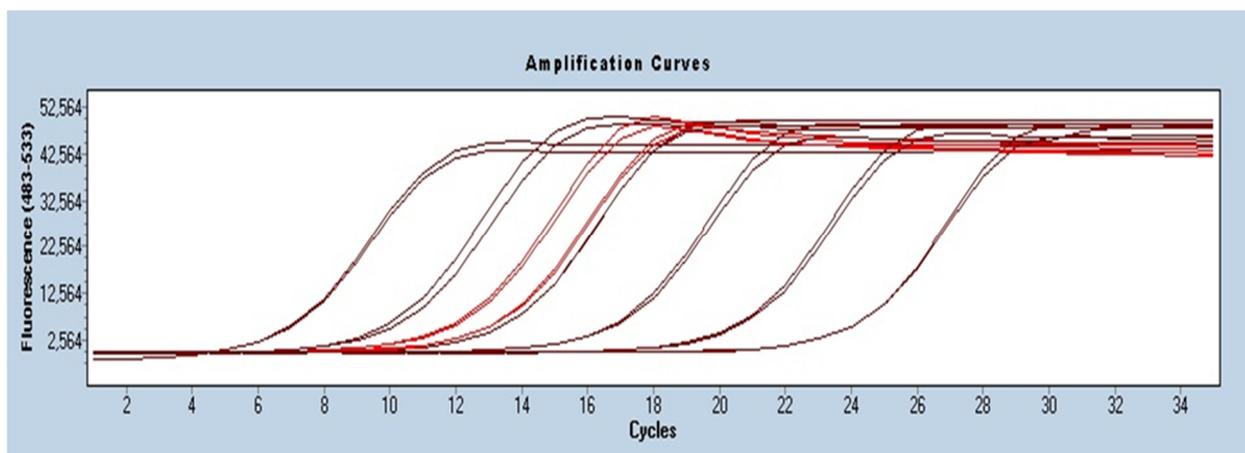


Figura 4.7 - Curva di amplificazione del pool nelle due diluizioni 1:500 e 1:1000.

Dil.	Library	Avg. Size	Copies/rxn	Size adjusted copies/rxn	Copies in library stock (copies/ul)		ul lib TQ	ul TE 1X	[copies/ul]
500	Pool 1:500	600	1,25E+06	9,59E+05	1,20E+08		1,7	18,3	1,00E+07
1000	Pool 1:1000	600	6,02E+05	4,61E+05	1,15E+08	0,962812			

Tabella 4.3 - Numero di molecole per reazione ottenute dal software e i calcoli che hanno potuto ad ottenere l'esatto numero di molecole nel pool. Nella prima colonna sono presenti i fattori di diluizione del pool, a seguire: il nome della libreria, la lunghezza del frammento d'amplificazione in paia di basi, il numero di molecole per reazione, il numero di molecole relazionato alla dimensione reale del frammento, le copie di DNA per  $\mu\text{l}$ , l'efficienza della reazione e le ultime tre colonne indicano i  $\mu\text{l}$  da prelevare di campione e di buffer TE 1X per ottenere la concentrazione finale di  $1 \times 10^8$  molecole per  $\mu\text{l}$ .

#### 4.5 PCR in emulsione e pirosequenziamento

Per determinare il numero corretto di molecole di libreria da utilizzare durante la PCR in emulsione è stata eseguita una titolazione preventiva con tre punti di misura (vedi Materiali e Metodi), i dati ottenuti sono stati rielaborati in una tabella Excel come segue (Tab. 4.4):

Sample	Copies per Bead	Conc beads recovered	Volume beads recovered	Conc beads enriched	Volume beads enriched	Tot beads recovered	Tot beads recovered	Tot beads enriched	Tot beads enriched	% beads	Real Tot beads enriched
		[x 10 <sup>6</sup> /ml]	[μl]	[x 10 <sup>6</sup> /ml]	[μl]		[%]		[%]		[%]
Pool 1	0,1	7,729	222	0,38	108	1715838	85,79	40676	2,03	2,37	2,35
Pool 2	0,25	8,199	217	1,10	106	1779183	88,96	116600	5,83	6,55	6,26
Pool 3	0,5	7,539	215	1,39	111	1620885	81,04	153957	7,70	9,50	9,96

Tabella 4.4 – Valori ottenuti dalla titolazione di tre diversi rapporti tra il numero di copie di libreria e biglie utilizzate: 0.1, 0.25 e 0.5. Rispettivamente sono stati riportati: la concentrazione per ml ed il volume delle biglie recuperate e delle biglie arricchite, il nr. totale e la percentuale di biglie recuperate e biglie arricchite, infine il rapporto tra il nr. di biglie arricchite rispetto al nr. di biglie utilizzate ( $2 \times 10^6$ ) e la percentuale reale di arricchimento ottenuto dal rapporto tra il nr. di biglie arricchite e le biglie recuperate.

I risultati ottenuti sono stati successivamente inseriti in un grafico riportando il numero di copie per biglia utilizzate e le percentuali di arricchimento ottenute, ciò ha consentito la corretta valutazione del dato e la scelta del rapporto più idoneo per il pirosequenziamento, vedi Fig. 4.8.

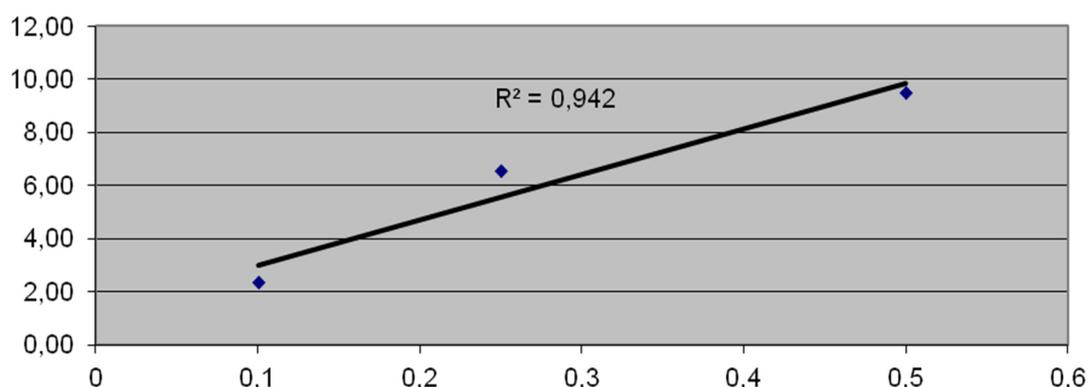


Figura 4.8 – Curva di taratura costruita in base ai valori di copie per biglia utilizzati e le percentuali di arricchimento ottenute. La retta di interpolazione tra i punti permette di stimare le percentuali di arricchimento ottenibili con differenti nr. di copie per biglia.

Per la preparazione delle biglie di sequenziamento la libreria è stata processata utilizzando un maggior numero di biglie, rispetto a quelle utilizzate durante la titolazione, che permette una copertura di reads più alta durante il sequenziamento. Il numero di copie per biglia scelto per il pirosequenziamento è stato pari a 0,7 e ha prodotto un arricchimento dell'8,7% che rientra nell'intervallo compreso tra 5 e 15%, come consigliato dal protocollo Roche.

Le biglie arricchite sono state sottoposte a pirosequenziamento, vedi Materiali e Metodi, producendo 236.507 reads totali, distribuite in maniera omogenea tra i campioni con valori compresi tra i 10.752 del campione Mw3 e i 19.606 del campione Mw9 (Fig. 4.9).

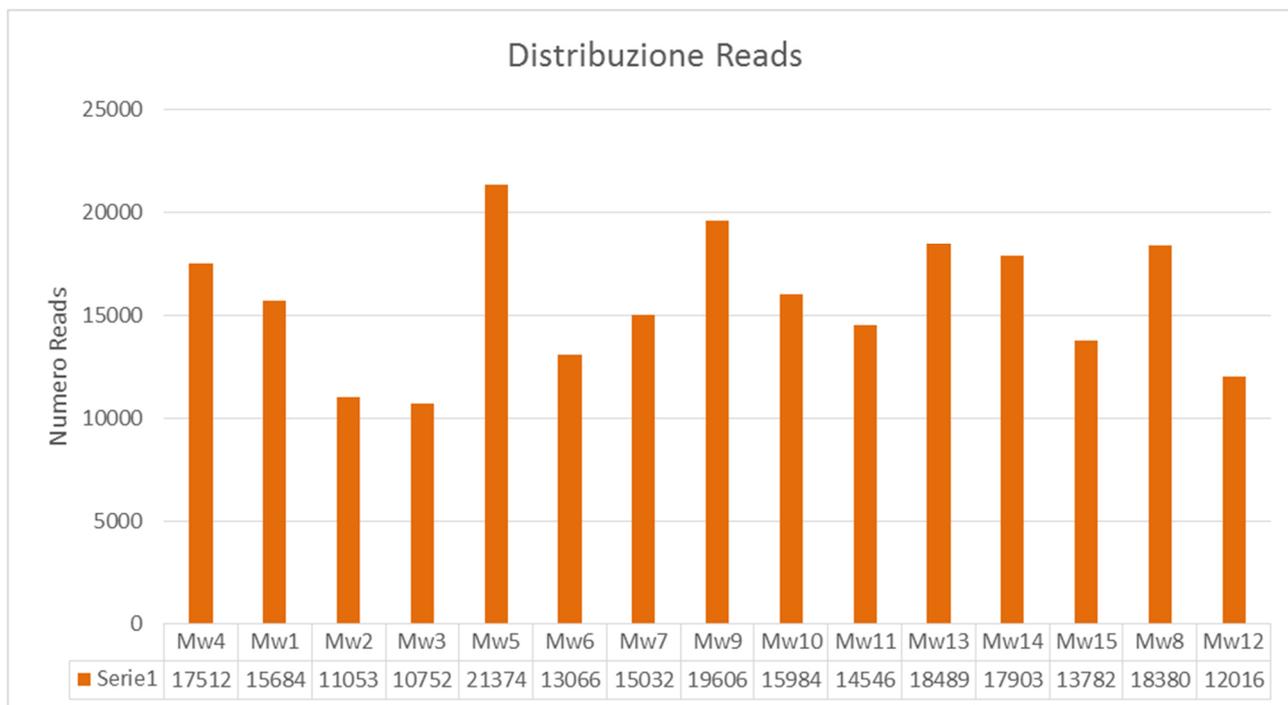


Figura 4.9 - Distribuzione delle reads per ogni singolo campione della foresta Mw.

## 4.6 Analisi bioinformatica

Le analisi bioinformatiche sono state condotte sulle sequenze ottenute dai campioni fecali dei primati di entrambe le foreste (Magombera e Mwanihana).

### 4.6.1 Pre-processamento delle reads e assegnazione tassonomica

Il sequenziamento 454 ha generato un totale di 344.938 reads riunite in un unico file binario in formato \*.sff. A partire da questi dati *raw* (grezzi), il primo passaggio prevede la separazione delle reads appartenenti ad ogni singolo campione, mediante il programma *sfffile* (Roche), il quale legge la sequenza barcode (Multiple IDentifiers) e la assegna in maniera univoca ad ogni campione. In media sono state assegnate 11.127 reads per campione.

In seguito i dati *raw* sono stati sottoposti ad un pre-processamento usando la pipeline software MICCA (versione 0.1, <http://compmetagen.github.io/micca/>): (i) taglio delle sequenze dei primer forward e reverse; (ii) filtraggio partendo dal 3' basato sulla qualità; (iii) eliminazione delle sequenze con meno di 300 nucleotidi.

Successivamente le sequenze con una similarità maggiore del 97% sono state clusterizzate in OTU (Operational Taxonomic Unit) attraverso il programma *micca-otu-denovo*. Ogni OTU è stata classificata tramite il software per l'assegnamento tassonomico RDP (Wang *et al.*, 2007). L'allineamento multiplo

delle sequenze (MSA) è stato condotto tramite il software PyNAST (Caporaso *et al.*, 2010) utilizzando il database Greengenes (DeSantis *et al.*, 2006) come riferimento.

#### 4.6.2 $\alpha$ – diversity

Le OTUs così ottenute sono state utilizzate per l'analisi della  $\alpha$ -diversity che indica il numero di specie presenti in una comunità. Lo studio della  $\alpha$ -diversity può essere analizzata nella sue tre componenti principali: ricchezza, omogeneità e diversità, mediante l'uso di indici statistici che cambiano in relazione al tipo di analisi che si sta svolgendo. In termini generali la  $\alpha$ -diversity possiede un valore elevato quando è presente un alto numero di specie e le loro abbondanze relative sono molto simili (Fig. 4.10).

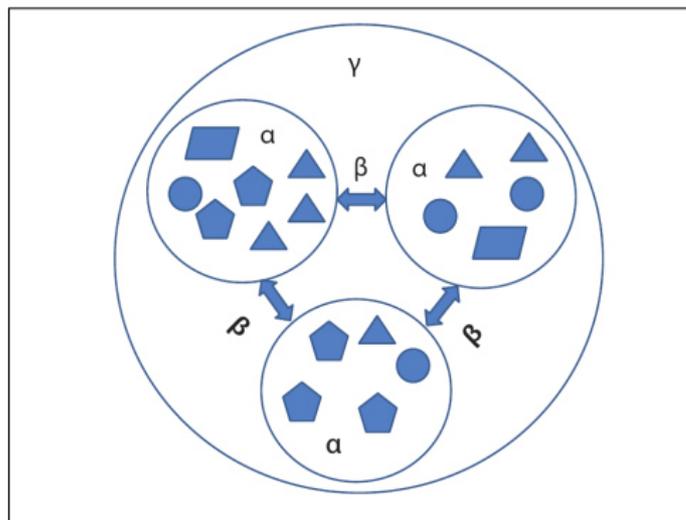


Figura 4.10 – Rappresentazione schematica delle  $\alpha$ ,  $\beta$  e  $\gamma$  – diversity (“Biodiversità in teoria e in pratica” Roberto Cazzolla Gatti, libreria universitaria.it edizioni 2014).

Per prima cosa occorre valutare se la profondità di indagine raggiunta è soddisfacente, quindi è stata creata una curva di rarefazione dalla relazione tra il numero di OTUs identificate ed il numero delle reads prodotte. Nella fase iniziale all'aumentare delle reads aumentano proporzionalmente il numero delle OTUs, una volta identificate tutte le OTUs presenti nel campione si raggiunge la fase di *plateau*. Da questa analisi si evince che per tutti i campioni è stata raggiunta la saturazione delle OTUs disponibili (vedi Figura 4.11), ed è stato determinato il numero di 4.067 reads rappresentativo per rendere tutti i campioni confrontabili.

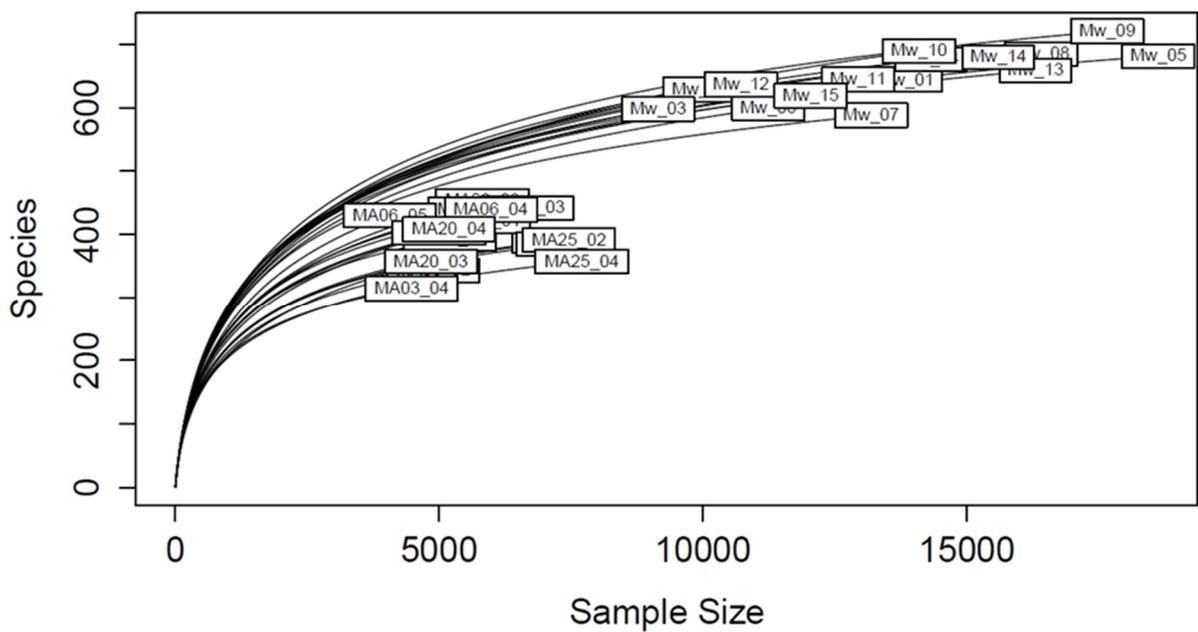


Figura 4.11 – Curva di rarefazione relativa ai 31 campioni analizzati. In ascissa sono indicati il numero delle reads, e in ordinata il numero delle OTUs per ogni campione. Questa analisi permette di valutare se tutte le specie presenti nel campione sono state identificate (species richness).

Le OTUs sono state classificate a diversi livelli tassonomici (Phylum, Classe, Ordine, Famiglia e Genere) nelle due foreste. Le percentuali sono state calcolate a partire dall'abbondanza delle reads suddivise a livello di OTUs e non dei campioni.

Dall'analisi delle abbondanze relative è emerso che, a livello di phylum, le OTUs più rappresentative sono per entrambe le foreste quelle appartenenti ai Firmicutes (75,4% in Ma e 66% in Mw), a cui seguono i Bacteroidetes (12,2% in Ma e 17,5% in Mw). Tutti gli altri phyla hanno percentuali inferiori al 5% e sono rappresentati da Verrucomicrobia (1,7% in Ma e 2% in Mw), Spirochaetes (1,1% in Ma e 2% in Mw) e Proteobacteria (0,4% in Ma e 0,7% in Mw). L'abbondanza relativa dei batteri non classificati va dal 9,1% all'11,7% nelle foreste di Magombera e Mwanihana rispettivamente (vedi Fig. 4.12).

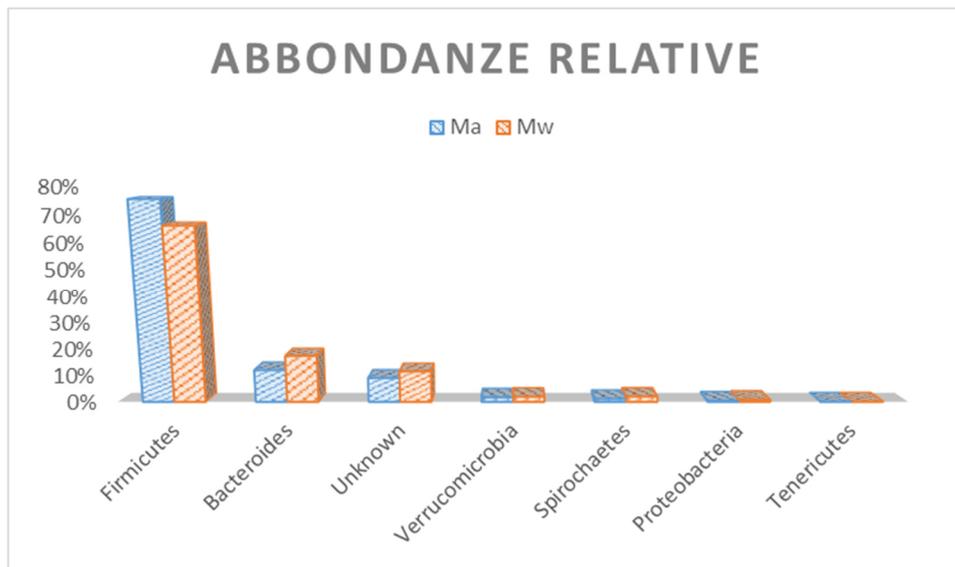


Figura 4.12 – Rappresentazione grafica delle abbondanze relative calcolate a partire dalla media delle abbondanze relative in percentuale e dalle OTUs corrispondenti a livello di phylum.

Parallelamente è stata calcolata l'abbondanza relativa anche a livello di famiglia, dove è emerso che le famiglie maggiormente rappresentate in entrambe le foreste sono quelle delle Ruminococcaceae (41,9% in Ma e 35,7% in Mw) e Lachnospiraceae (18,1% in Ma e 15,2% in Mw), seguite dalle Porphyromonadaceae (6,2% in Ma e 8,5% in Mw), Bacteroidaceae (0,6% in Ma e 1,9% in Mw), Spirochaetaceae (1,1% in Ma e 1,8% in Mw), Prevotellaceae (0,5% in Ma e 0,9% in Mw) e Clostridiales\_Incertae Sedis XIII (0,7% in Ma e 0,4% in Mw) (vedi Fig. 4.13).

Inoltre sono state calcolate le differenze significative in termini di valor medio delle abbondanze relative a livello di famiglia (Fig. 4.13) mediante il test-t di Welch a due campioni (Welch, 1947). Sono state riscontrate differenze significative nelle due foreste relative alle famiglie Bacteroidaceae, Prevotellaceae ( $p < 0,005$  e  $p < 0,05$  rispettivamente) più abbondanti negli individui provenienti dalla foresta Mwanihana, Ruminococcaceae e Clostridiales\_Incertae Sedis XIII ( $p < 0,05$ ) più abbondanti nella foresta Magombera.

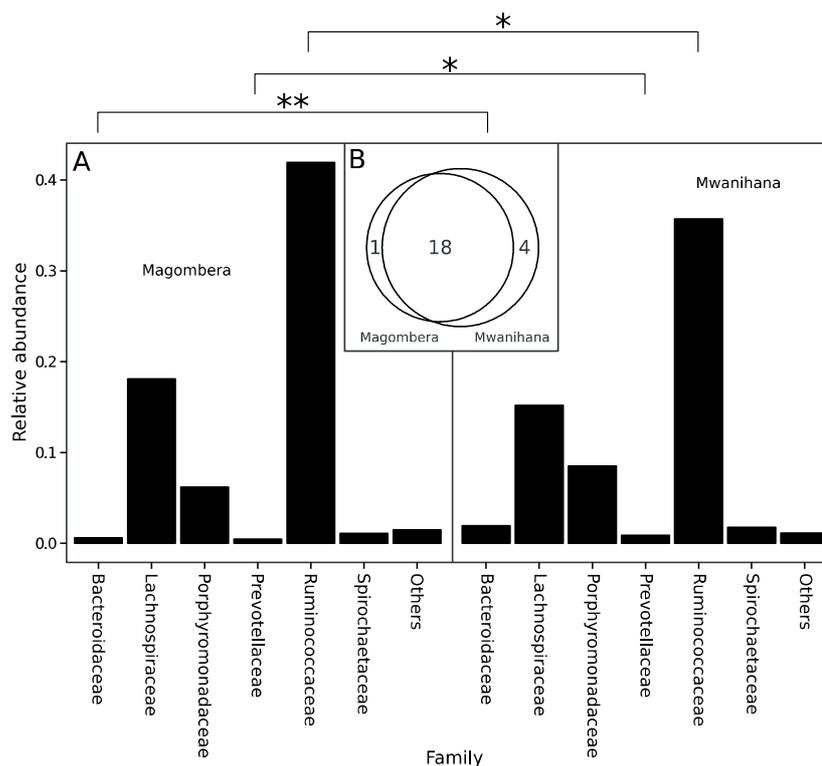


Figura 4.13 – Abbondanze relative delle famiglie componenti il microbiota intestinale, raggruppate in base alle due foreste Magombera e Mwanihana (A). Differenze significative dell'abbondanza relativa sono state trovate in Bacteroidaceae, Prevotellaceae e Ruminococcaceae. \*  $p < 0,05$ ; \*\*  $p < 0,005$  (test - t di Welch a due campioni e p - value). Il diagramma di Venn mostra le famiglie uniche e in comune negli individui delle due foreste (B).

Le famiglie Lachnospiraceae e Ruminococcaceae sono due delle famiglie più abbondanti all'interno del microbiota intestinale dei mammiferi, e sono state associate al mantenimento della salute del microbiota in quanto svolgono un'importante funzione, quale la degradazione di componenti vegetali indigeribili (Biddle *et al.*, 2013). I componenti delle Lachnospiraceae, come *Coproccoccus sp.* e *Roseburia sp.*, sono stati associati alla produzione di butirrato, necessario per la salute del tessuto epiteliale del colon (Barcenilla *et al.*, 2000 e Duncan *et al.*, 2002), ed è stato dimostrato che la loro presenza è tendenzialmente scarsa in malattie infiammatorie dell'intestino (Frank *et al.*, 2007).

Entrambe le famiglie in questione possiedono nel loro corredo genetico famiglie geniche come idrolasi glicosidiche (GH), coinvolte nella degradazione di composti vegetali complessi come cellulose ed emicellulose, altrimenti indigeste per l'ospite. Questi composti sono poi fermentati e convertiti in corte catene di acidi grassi (soprattutto acetato, butirrato e propionato) che possono essere assorbiti e usati come fonte di energia (Biddle *et al.*, 2013).

Possono essere applicati altri indici di ricchezza. Uno di questi è l'indice di Chao 1, che stima il numero assoluto di specie di un campione, basandosi sul numero di specie rare (Schloss & Handelsman, 2007). Un altro indice che può essere utilizzato è l'entropia di Shannon, relativo alla distribuzione delle OTUs all'interno del campione. Questi indici di diversità forniscono informazioni sulla ricchezza e sull'omogeneità ("evenness") delle diverse specie all'interno dei campioni (Schloss & Handelsman, 2007).

Oltre al calcolo del numero di OTUs osservate (Fig. 4.14A) sono stati utilizzati gli indici di Chao 1 (Fig. 4.14B) e di Shannon (Fig. 4.14C).

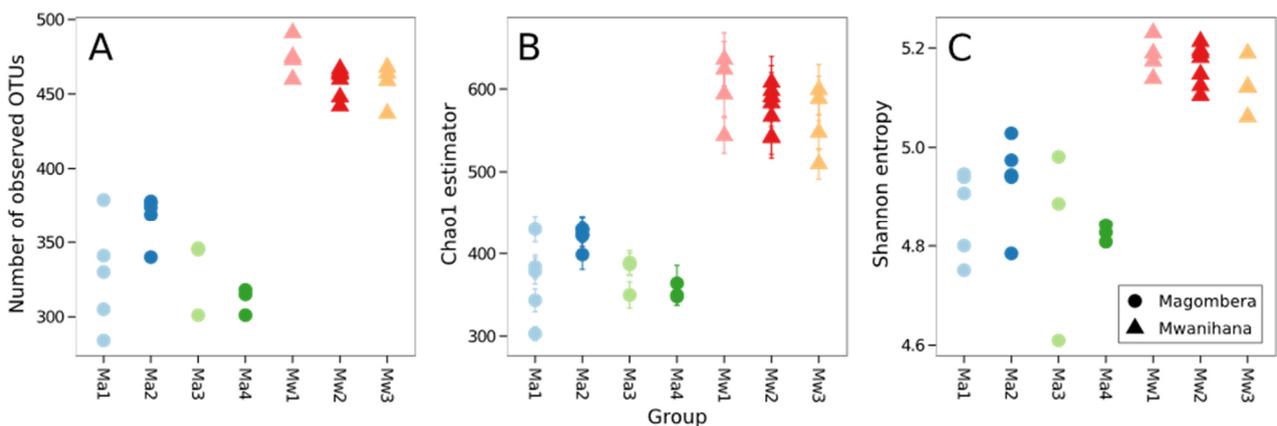


Figura 4.14 – Misura dell' $\alpha$  – diversity. Sono state calcolate tre stime per la valutazione dell' $\alpha$  – diversity: il numero di OTUs osservate (A), l'indice Chao 1 (B), e l'entropia di Shannon (C). I cerchi rappresentano la foresta Magombera e i triangoli la foresta Mwanihana. I simboli sono colorati secondo la suddivisione nei gruppi (Fig. 1.10). La barra d'errore indica l'errore Standard.

Dalla analisi della ricchezza (Figura 4.14 A), si nota che c'è una netta separazione tra le due foreste per quanto riguarda l'abbondanza relativa delle OTUs, è decisamente maggiore nei campioni appartenenti

alla foresta Mwanihana (Wilcoxon rank sum test,  $p=2.3 \times 10^{-6}$ ). Questo risultato è consistente con l'indice di Chao1 (Fig. 4.14B,  $p=6.7 \times 10^{-9}$ ) e di Shannon (Fig. 4.14C,  $p=6.7 \times 10^{-9}$ ). Da questo risultato si può dedurre che gli individui appartenenti alla foresta Mwanihana, all'interno del parco naturale, i quali non sono sottoposti ad antropizzazione né ad altri tipi di disturbo, godono di una maggiore ricchezza batterica all'interno del loro microbiota intestinale. La minor ricchezza nel microbiota dei campioni appartenenti alla foresta Magombera rispecchia, invece, la condizione di degradazione dell'ecosistema in cui sono inseriti gli individui di questa popolazione.

#### 4.6.3 $\beta$ – diversity

Il passo successivo prevede il calcolo della diversità tra i campioni delle due foreste. Questo tipo di misura permette di studiare quanto differiscano tra loro due o più ambienti in termini di struttura di popolazione. Infatti il suo valore aumenta in presenza di ambienti eterogenei, mentre diminuisce negli ambienti omogenei.

A questo scopo sono state impiegate due misure ampiamente utilizzate nella letteratura ecologica, gli indici di dissomiglianza di Bray-Curtis e UniFrac non pesata. Nel primo caso si tratta di un calcolo usato per quantificare la composizione di diversità tra due differenti siti, in questo caso basata sulla conta delle OTUs. La distanza UniFrac, invece, è basata sulle relazioni filogenetiche tra le diverse OTUs.

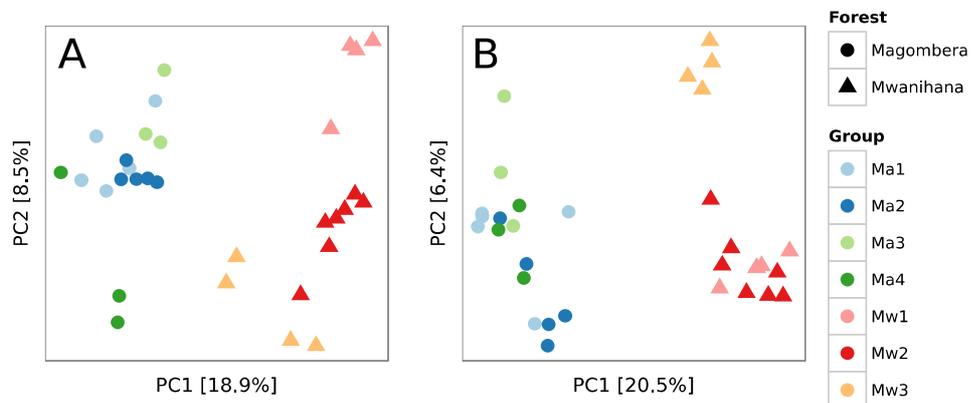


Figura 4.15 – Misura della  $\beta$  – diversity. PCoA della distanza tra i campioni misurata usando la dissimilarità di Bray – Curtis (A). PCoA della distanza tra i campioni misurata usando la distanza UniFrac (B). I cerchi rappresentano la foresta Magombera, mentre i triangoli la foresta Mwanihana. I simboli sono colorati secondo la suddivisione nei gruppi (Fig. 1.10).

È possibile notare una separazione netta tra le popolazioni microbiche delle due foreste rispetto la prima componente (PC1), infatti come mostrato in Fig. 4.15, tutti i componenti della foresta Mwanihana (triangolo) occupano la parte destra del grafico, mentre i campioni appartenenti alla

foresta Magombera (cerchio) sono situati nella porzione sinistra del grafico. A livello della seconda componente (PC2), invece, è possibile notare la differenza anche all'interno dei gruppi delle due foreste, più marcata con la distanza Bray-Curtis (PERMANOVA  $p < 0.001$ ). Soprattutto nella foresta Mwanihana, i tre gruppi Mw1, Mw2 e Mw3, sono ben separati tra loro.

#### 4.6.4 Analisi metabolica (KEGG pathway)

Un'ulteriore analisi che è stata effettuata, è l'individuazione dei pathway metabolici coinvolti nelle popolazioni microbiche identificate mediante un'analisi di inferenza che si basa sulla corrispondenza con famiglie geniche rappresentative delle OTU. Questo passaggio serve per capire più in profondità quale sia il ruolo metabolico dei microorganismi che costituiscono il microbiota intestinale degli individui delle due foreste in relazione alle distanze filogenetiche.

Il programma usato è PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) che compie una predizione della composizione funzionale di un metagenoma usando dati provenienti da geni marcatori e da database di genomi di riferimento (Langille *et al.*, 2013). In questo modo è stato possibile predire le famiglie geniche presenti e in seguito combinarle per valutare la composizione del metagenoma partendo dalle OTUs. È stata creata una tabella di famiglie geniche con un numero specifico per ogni campione. L'uso di KEGG (Kyoto Encyclopedia Genes and Genomics, <http://www.genome.jp/kegg/>) ha permesso di codificare la funzione di ogni famiglia genica mediante la corrispondenza in geni ortologhi.

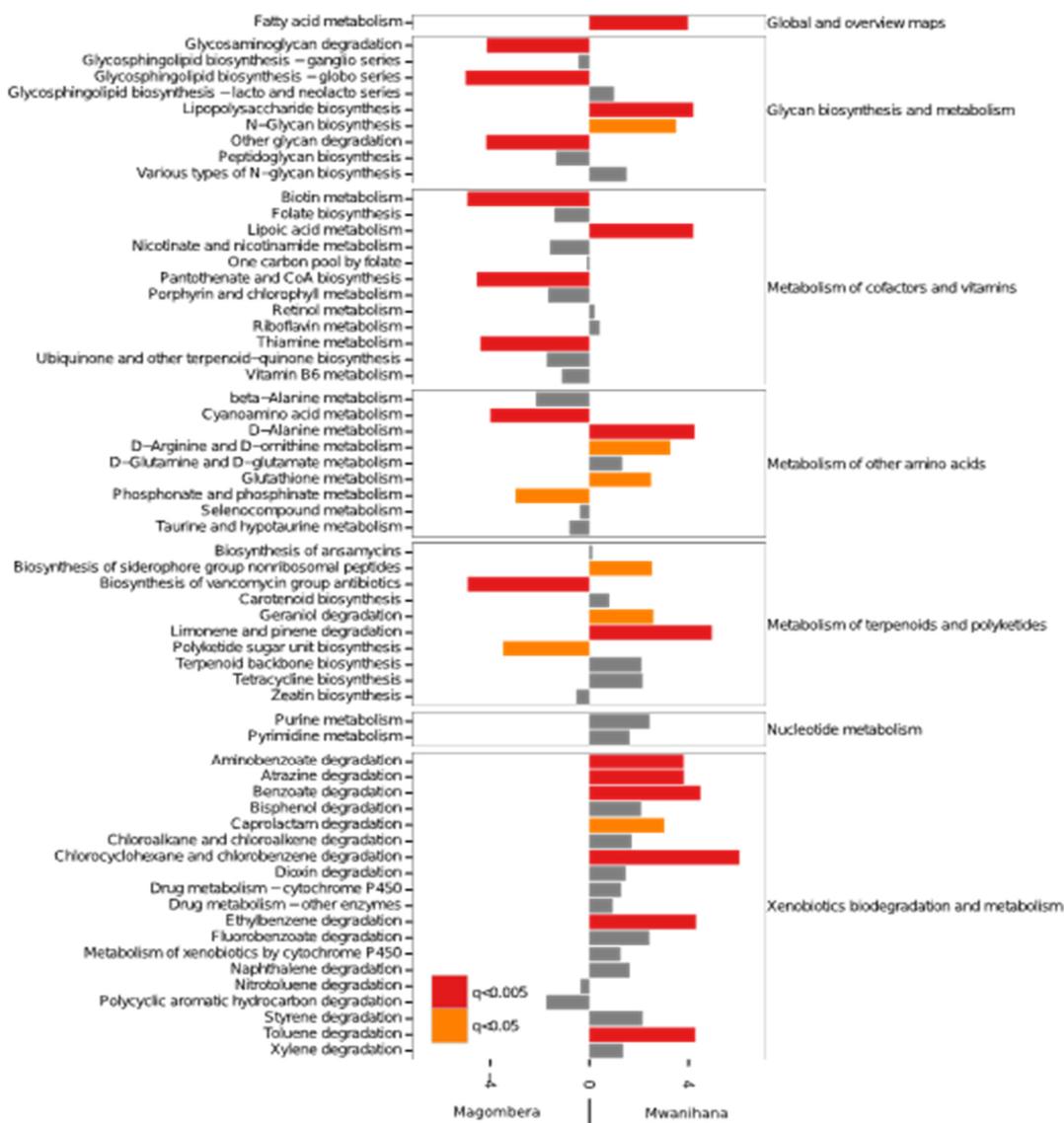


Figura 4.16A – Rappresentazione delle classi metaboliche rappresentative nelle due foreste. È stata applicata la statistica t – test che calcola la differenza tra le medie delle funzioni delle due foreste rapportate alla deviazione standard. Il calcolo statistico è accompagnato al calcolo del livello di significatività di ogni funzione metabolica ed è definito dalle barre colorate: grigio: non significativo, arancio:  $0,005 < q < 0,05$  e rosso:  $q < 0,05$ .

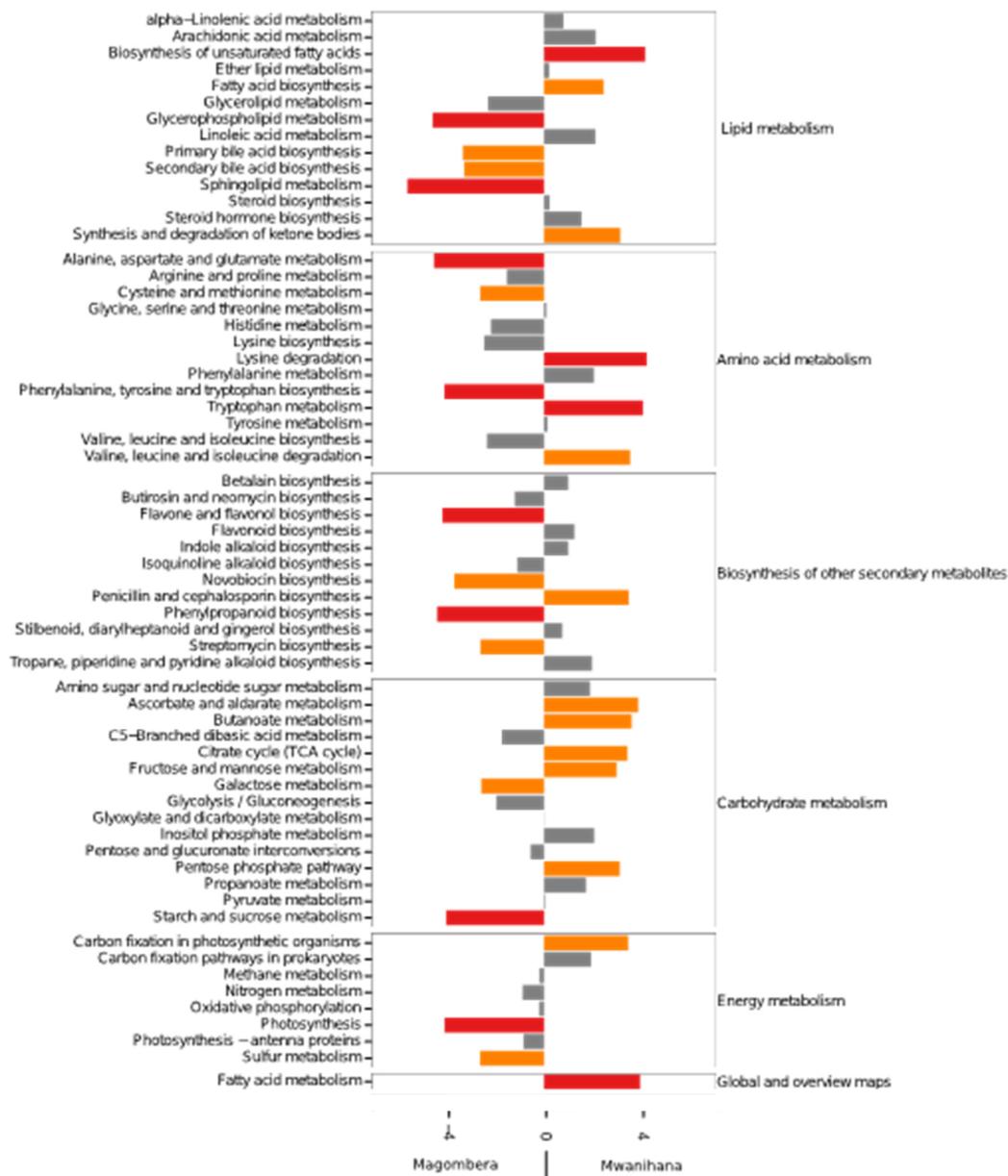


Figura 4.16B – Rappresentazione delle classi metaboliche rappresentative nelle due foreste. È stata applicata la statistica t – test che calcola la differenza tra le medie delle funzioni delle due foreste rapportate alla deviazione standard. Il calcolo statistico è accompagnato al calcolo del livello di significatività di ogni funzione metabolica ed è definito dalle barre colorate: grigio: non significativo, arancio:  $0,005 < q < 0,05$  e rosso:  $q < 0,05$ .

Molte sostanze ingerite con la dieta hanno bisogno dell'intervento di batteri per la loro degradazione, soprattutto in una dieta ricca di vegetali. Per esempio, alcuni carboidrati complessi presenti nelle piante come xilani, glicani, pectine, non sono degradati dagli enzimi dei mammiferi in quanto questi mancano di geni specifici con tale funzione, che viene conferita grazie all'intervento dei batteri che possiedono questo tipo di enzimi (Koropatkin *et al.*, 2012). Nella Figura 4.16B è possibile notare che il

metabolismo dei carboidrati è maggiormente rappresentato nei campioni provenienti dalla foresta Mwanihana, ciò sta ad indicare un'attività metabolica più significativa, probabilmente correlato con il tipo di dieta degli animali che risiedono in quella foresta. I pathway coinvolti sono soprattutto il metabolismo dell'ascorbato e dell'aldarato, il metabolismo del butanoato, il ciclo del citrato, il metabolismo del mannosio e del fruttosio e il pathway del pentoso fosfato.

Invece, i metabolismi più significativamente espressi nei campioni provenienti dalla foresta Magombera sono quelli dei lipidi (soprattutto il metabolismo dei glicerofosfolipidi e degli sfingolipidi) e la biosintesi dei metaboliti secondari (Fig. 4.16B).

Un risultato interessante è dato dal metabolismo e dalla degradazione degli xenobiotici, funzione maggiormente rappresentata nel microbiota intestinale di individui appartenenti alla foresta Mwanihana (degradazione dell'aminobenzoato, dell'altrazina, del benzoato, di clorocicloexani e clorobenzeni, etilbenzene e toluene) (Figura 4.16A). Questo risultato può far pensare a una co-evoluzione della comunità batterica con il suo ospite, capace di attivare processi di detossificazione di composti potenzialmente tossici presenti nelle piante, come ad esempio i tannini, di cui i colobi rossi si cibano.



## 5. CONCLUSIONI

È stato ormai dimostrato in numerosi studi, come il microbiota intestinale influenza lo stato di salute dell'ospite e come può modificarsi in relazione ai fattori ambientali fra cui il tipo di dieta assunta. Quello che non è stato compreso finora, e che questo studio è stato in grado di mettere in luce, è come un'alterazione a livello ambientale, quali l'alterazione dell'habitat naturale in cui gli animali vivono, sia capace di riflettersi direttamente sulle risorse nutrizionali ad essi disponibili e conseguentemente influenzare la fisiologia di specie a rischio di estinzione quali i colobi rossi.

In questo lavoro è stato analizzato il microbiota intestinale di due popolazioni di colobi rossi strettamente correlate a livello geografico, eppure distinte, provenienti da due foreste dei monti Udzungwa in Tanzania (Mwanihana e Magombera), che presentano diversi livelli di biodiversità vegetale e differenti interferenze da parte dell'uomo.

In questo studio è stato osservato che il microbiota intestinale dei colobi rossi provenienti dalle due foreste possono essere facilmente distinti, infatti il microbiota dei campioni provenienti dalla foresta Mwanihana, più grande e incontaminata, mostra una maggiore  $\alpha$ -diversity rispetto a quelli provenienti dalla foresta Magombera, più piccola e più soggetta a disturbo da parte dell'uomo, in quanto è circondata da aree coltivate e da insediamenti umani. L'influenza degli insediamenti umani sul microbiota intestinale dei colobi rossi è confermato dalla correlazione tra la diversità e la ricchezza del microbiota intestinale e la distanza dalle aree antropizzate, indicando come alterazioni ambientali possono alterare il microbiota intestinale di popolazioni animali selvatiche.

L'analisi dell'abbondanza relativa delle OTUs ha mostrato una prevalenza delle famiglie Ruminococcaceae e Lachnospiraceae in entrambi i gruppi analizzati. Queste famiglie sono altamente specializzate nella degradazione di composti vegetali complessi (Flint & Bayer 2008). Questi composti sono infatti fermentati dai batteri sopracitati e convertiti in acidi grassi a catena corta (soprattutto acetato, butirato e propionato) per poi essere assimilati e usati come fonte di energia dall'ospite. Questo risultato è supportato da analisi funzionali, che mostrano un "*core gut microbiome*" in entrambi i gruppi di animali analizzati, ricco in geni coinvolti nell'estrazione di energia dalle fibre indigeribili, nel metabolismo dei grassi, degli zuccheri e delle proteine. L'equilibrio che si instaura all'interno del microbiota intestinale è quindi facilmente alterato da fattori ambientali per quanto riguarda la composizione tassonomica dei microorganismi, mentre per quanto riguarda le potenzialità funzionali si osserva, in entrambe le foreste, la presenza di un "*core gut microbiome*" che garantisce le funzioni metaboliche necessarie quali la digestione dei carboidrati indigeribili (cellulose, xilani etc.). Abbiamo invece osservato una diversità relativa al metabolismo degli xenobiotici, presenti nella dieta folivora, che risulta maggiormente presente, in modo significativo, negli animali che vivono nella foresta incontaminata di Mwanihana.

In conclusione è stato dimostrato che, in habitat degradati la ricchezza e la diversità del microbiota intestinale sono minori. Ciò suggerisce che una importante frazione dei microorganismi presenti nei colobi rossi dell'Udzungwa abbiano un'origine ambientale o comunque influenzata dalle condizioni dell'habitat in cui è inserito l'individuo e dalla relativa disponibilità di risorse alimentari (dieta).

Studi futuri potrebbero esaminare la relazione tra la composizione del microbiota intestinale dei colobi rossi e la loro fisiologia direttamente misurando fattori di benessere che potrebbero essere influenzati dal microbioma come abbondanza e diversità dei microorganismi commensali. Infine, un'attenta valutazione della dieta della popolazione e della fitness di riproduzione potrebbe aiutare a capire come differenti pressioni ambientali determinano la struttura delle comunità del microbiota intestinale e la funzione, e come i cambiamenti conseguenti delle comunità batteriche gastrointestinali a sua volta impattano l'efficienza del metabolismo e la fitness dell'ospite

## BIBLIOGRAFIA

Abbas F., Morellet N., Hewison A. J. M., Merlet J., Cargnelutti B., Lourtet B., *et al.*, 2011. Landscape fragmentation generates spatial variation of diet composition and quality in a generalist herbivore. *Oecologia* 167: 401–411.

Amato K. R., Yeoman C. J., Kent A., Righini N., Carbonero F., Estrada A., Gaskins H. R., Stumpf R. M., Yildirim S., Torralba M., Gillis M., Wilson B. A., Nelson K. E., White B. A., Leigh S. R., 2013. Habitat degradation impacts black howler monkey (*Alouatta pigra*) gastrointestinal microbiomes. *ISME J* 7:1344–1353.

Arumugam M., Raes J., Pelletier E., *et al.*, 2011. Enterotypes of the human gut microbiome. *Nature* 473:174.

Banks J. C., Cary S. C., Hogg I. D., 2009. The phylogeography of Adelie penguin faecal flora. *Environ Microbiol* 11:577–588.

Barcenilla A., Pryde S.E., Martin J.C., Duncan S.H., Stewart C.S., Henderson C., Flint H.J., 2000. Phylogenetic Relationships of Butyrate-Producing Bacteria from the Human Gut. *Appl. Environ. Microbiol.* 66:1654–1661.

Belkaid Y., Naik S., 2013. Compartmentalized and systemic control of tissue immunity by commensals. *Nat Immunol.* 14(7):646-53. doi: 10.1038/ni.2604.

Biddle A., Stewart L., Blanchard J., Leschine S., 2013. Untangling the Genetic Basis of Fibrolytic Specialization by Lachnospiraceae and Ruminococcaceae in Diverse Gut Communities. *Diversity* 5, 627-640; doi:10.3390/d5030627.

Cani P. D., Amar J., Burcellin R., 2008. Role of gut microflora in the development of obesity and insulin resistance following high-fat diet feeding. *Pathol. Biol.*, 56:305-309.

Caporaso J. G., Bittinger K., Bushman F. D., *et al.*, 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics*, 26:266-267.

Chapman C. A., Pavelka M. S., 2005. Group size in folivorous primates: ecological constraints and the possible influence of social factors. *Primates* 46:1–9.

Chapman C. and Chapman L., 2002. Foraging challenges of red colobus monkeys: influence of nutrients and secondary compounds. *Comparative Biochemistry and Physiology Part A*, 133: 861-875.

Chaves O. M., Stoner K. E., Arroyo-Rodriguez V., 2012. Differences in diet between spider monkey groups living in forest fragments and continuous forest in Mexico. *Biotropica* 44: 105–113.

Chivers D. J., 1994. Functional anatomy of the gastrointestinal tract. In: Davies AG, Oates JF, editors. *Colobine monkeys: Their ecology, behaviour and evolution*. Cambridge: Cambridge University Press. 205–228.

Claesson M. J., O'Sullivan O., Wang Q., Nikkila J., Marchesi J. R., Smidt H., de Vos W. M., Ross R. P., O'Toole P. W., 2009. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE* 4, e6669.

Costello E. K., Stagaman K., Dethlefsen L., Bohannan B. J., Relman D. A., 2012. An application of ecological theory toward an understanding of the human microbiome. *Science* 336: 1255–1262.

Cryan J. F., Dinan T. G., 2012. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* 13, 701–712.

Cryan J. F., O'Mahony S. M., 2011. The microbiome—gut—brain axis: from bowel to behavior. *Neurogastroenterol. Motil.* 23, 187—192.

De Filippo C., Cavalieri D., Di Paola M., Ramazzotti M., Poullet B. J., Massart S., Collini S., Pieraccini G., Lionetti P., 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. [www.pnas.org/cgi/doi/10.1073/pnas.1005963107](http://www.pnas.org/cgi/doi/10.1073/pnas.1005963107).

Decker B. S., 1994. Endangered primates in the Selous Game Reserve and an imminent threat to their habitat. *Oryx* 28: 183–190. [www.pnas.org/cgi/doi/10.1073/pnas.1005963107](http://www.pnas.org/cgi/doi/10.1073/pnas.1005963107).

Degnan P. H., Pusey A. E., Lonsdorf E. V., *et al.*, 2012. Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. *Proc Natl Acad Sci USA* 109:13034–13039.

Delsenya M., Hanb B., Hsingc Y., 2010. High throughput DNA sequencing: The new sequencing revolution *Plant Science*. 179: 407–422.

DeSantis T. Z., Hugenholtz P., Larsen N., *et al.*, 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*. 72:5069–5072.

Dethlefsen L., Huse S., Sogin M. L., Relman D. A., 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol*. 6, e280.

Dinan T. G., Cryan J. F., 2012. Regulation of the stress response by the gut microbiota: Implications for psychoneuroendocrinology. *Psychoneuroendocrinology*, doi:10.1016/j.psyneuen.2012.03.007.

Dinesen L., Lehmberg T., Rahner M. C., Fjeldsa J., 2001. Conservation priorities for the forests of the Udzungwa Mountains, Tanzania, based on primates, duikers and birds. *Biol Cons* 99:223–236.

Donohoe D. R., Garge N., Zhang X., Sun W., O'Connell T. M., Bunger M. K., *et al.*, 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* 13: 517–526.

Duncan S.H., Barcenilla A., Stewart C.S., Pryde S.E., Flint H.J., 2002. Acetate utilization and butyryl coenzyme A (CoA): acetate-CoA transferase in butyrate- reducing bacteria from the human large intestine. *Appl. Environ. Microbiol*. 68:5186–5190.

Engelbrektson A., Kunin V., Wrighton K. C., Zvenigorodsky N., Chen F., Ochman H., Hugenholtz P., 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J*. 4, 642–647.

Flaherty E. A., Ben-David M., Smith W. P., 2010. Diet and food availability: Implications for foraging and dispersal of prince of wales northern flying squirrels across managed landscapes. *J Mammal* 91: 79–91.

Flint H. J., Bayer E. A., 2008. Plant cell wall breakdown by anaerobic microorganisms from the mammalian digestive tract. *Annals of the New York Academy of Sciences*, 1125, 280–288.

Frank D.N., Amand A.L.S., Feldman R.A., Boedeker E.C., Harpaz N., Pace N.R., 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. USA* 104:13780–13785.

Freeland W. e Janzen D., 1974. Strategies in herbivory by mammals: the role of secondary compounds. *The American Naturalist*, 108: 269-289.

Frey J. C., Rothman J. M., Pell A. N., *et al.*, 2006. Fecal bacterial diversity in a wild gorilla. *Appl Environ Microb* 72:3788– 3792.

Glander K., 1982. The impact of plant secondary compounds on primate feeding behavior. *Yearbook of Physical Anthropology*, 25: 1-18.

Gordon J.I., 2012. Honor thy gut symbionts redux. *Science* 336, 1251-1253.

Grootaert C., Van de Wiele T., Van Roosbroeck I., Possemiers S., Vercoutter-Edouart A. S., Verstraete W., *et al.*, 2011. Bacterial monocultures, propionate, butyrate and H<sub>2</sub>O<sub>2</sub> modulate the expression, secretion and structure of the fasting-induced adipose factor in gut epithelial cell lines. *Environ Microbiol* 13: 1778–1789.

Groves C. P., 2007. The taxonomic diversity of the Colobinae of Africa. *J Anthropol Sci* 85:7–34.

Harris T. e Chapman C., 2007. Variation in diet and ranging of black and white colobus monkeys in Kibale National Park, Uganda. *Primates*, 48: 208-221.

Hu X., Xu C., Zhou X., He B., Wu L., Cui B., *et al.*, 2010. Sodium butyrate protects against myocardial ischemia and reperfusion injury by inhibiting high mobility group box 1 protein in rats. *Biomed Pharmacother*; e-pub ahead of print 25 Sept 2010; doi: 10.1016/j.biopha2010.09.005.

Hugenholtz P., Goebel B. M., Pace N.R., 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. *Journal of Bacteriology*. 180 (18):4765–4774.

Huse S. M., Dethlefsen L., Huber J. A., Mark Welch D., Relman D. A., Sogin M. L., 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* 4, e1000255.

Kamler J. F., Ballard W. B., Wallace M. C., Gipson P. S., 2007. Diets of swift foxes (*Vulpes velox*) in continuous and fragmented prairie in Northwestern Texas. *Southwest Nat* 52: 504–510.

Kay R. e Davies G., 1994. Digestive physiology. In A.G. Davies and J.F. Oates (Eds.), *Colobine Monkeys: Their Ecology, Behaviour, and Evolution*. Cambridge: Cambridge University Press.

Kim M., Morrison M., Yu Z., 2011. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *Journal of Microbiological Methods* 84: 81–87.

Kliks M. M., 1990. Helminths as heirlooms and souvenirs—a review of New-World paleoparasitology. *Parasitol Today* 6:93–100.

Koenig J. E., Spor A., Scalfone N., *et al.*, 2011. Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci US A*108:4578–4585.

Koropatkin N. M., Cameron E. A., Martens E. C., 2012. How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol.* 10(5):323-35. doi: 10.1038/nrmicro2746. Review. PubMed PMID: 22491358; PubMed Central PMCID: PMC4005082.

Krober M., Bekel T., Diaz N.N., Goesmann A., Jaenicke S., Krause L., Miller D., Runte K.J., Viehover P., Puhler A., Schluter A., 2009. Phylogenetic characterization of a biogas plant microbial community integrating clone library 16S-rDNA sequences and metagenome sequence data obtained by 454-pyrosequencing. *J. Biotechnol.* 142, 38–49.

Lambert J., 1998. Primate digestion: interactions among anatomy, physiology, and feeding ecology. *Evolutionary Anthropology*, 7.1: 8-20.

Lambert J., 2007. Primate nutritional ecology: feeding biology and diet at ecological and evolutionary scales. In C.J. Campbell, A. Fuentes, K.C. MacKinnon, M. Panger, and S.K. Bearder (Eds.), *Primates in Perspective*. Oxford: Oxford University Press.

Lane D. J., Pace B., Olsen G. J., Stahl D. A., Sogin M. L., Pace N. R., 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. USA.* 82, 6955–6959.

Langille M. G. I., Zaneveld J., Caporaso J. G., *et al.*, 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, 31:14–821.

Ley R. E., Hamady M., Lozupone C., *et al.*, 2008a. Evolution of mammals and their gut microbes. *Science* 320:1647–1651.

Ley R. E., Lozupone C. A., Hamady M., Knight R., Gordon J. I., 2008b. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* 6:776–788.

Ley R. E., Peterson D. A., Gordon J. I., 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848.

Liu Z., DeSantis T. Z., Andersen G. L., Knight R., 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Res.* 36, e120.

Liu Z., Lozupone C., Hamady M., Bushman F. D., Knight R., 2007. Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res.* 35, e120.

Lovett J. C., & Pocs T., 1993. Assessment of the Condition of the Catchment Forest Reserves, a Botanical Appraisal, Catchment Forestry Report 93. 3, Dar es Salaam, Tanzania.

Lovett J. C., 1993. Eastern Arc moist forest flora. In Lovett J. C. e Wasser S. K. (eds.), *Biogeography and Ecology of the Rain Forests of Eastern Africa.*, Cambridge University Press, Cambridge, UK, 33–55.

Lovett J. C., Bridson D. M., Thomas D. W., 1988. A preliminary list of the moist angiosperm flora of Mwanihana Forest Reserve, Tanzania. *Ann. Mo. Bot. Gard.* 75: 874–888.

Lovett J. C., Pocs T., 1993. Assessment of the Condition of the Catchment Forest Reserves, a Botanical Appraisal, Catchment Forestry Report 93. 3, Dar es Salaam, Tanzania.

Lozupone C. A., Stombaugh J. I., Gordon J. I., Jansson J. K. & Rob Knight R., 2012. Diversity, stability and resilience of the human gut microbiota doi:10.1038/nature11550.

Margulies M., Egholm M., Altman W. E., Attiya S., Bader J. S., Bemben L. A., Berka J., Braverman M. S., Chen Y. J., Chen Z., Dewell S. B., Du L., Fierro J. M., *et al.*, 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380. doi:10.1038.

Marsh C., 1981. Diet choice among red colobus (*Colobus badius rufomitratu*s) on the Tana River, Kenya. *Folia Primatologica*, 35: 147-178.

Marshall A.R. 2008 a. Ecological Report on Magombera Forest. Unpublished report to WWFTPO.

Marshall A.R. 2008 b. Udzungwa Forest Project Summary for EAZA June 2008. Report for Flamingo Land Ltd., presented to the European Association of Zoos and Aquaria. [www.york.ac.uk/res/celp/people/associates/marshall](http://www.york.ac.uk/res/celp/people/associates/marshall).

McCord A. I., Chapman C. A., Weny G., Tumukunde A., Hyeroba D., Klotz K., *et al.*, 2014. Fecal Microbiomes of Non-Human Primates in Western Uganda Reveal Species-Specific Communities Largely Resistant to Habitat Perturbation. *American Journal of Primatology* 76:347–354.

Metzker M.L., 2010. Sequencing technologies. The next generation. *Nat. Rev. Genet.* 11: 31–46.

Mowry C., Decker B., e Shure D., 1996. The role of phytochemistry in dietary choices of Tana River red colobus monkeys (*Procolobus badius rufomitratu*s). *International Journal of Primatology*, 17.1: 63-84.

Mturi F., 1991. The feeding ecology and behavior of the red colobus monkey (*Procolobus badius kirkii*). Ph.D. dissertation. University of Dar es Salaam: Dar es Salaam, TZ.

Munroe D. J., Harris T.J.R., 2010. Third-generation sequencing fireworks at Marco Island. *Nat. Biotechnol.* 28: 426–428.

Myers N., Mittermeier, R. A., Mittermeier C. G., de Fonseca G. A. B., Kent J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Nowak K., 2007. Behavioural flexibility and demography of *Procolobus kirkii* across floristic and disturbance gradients. Ph.D. dissertation. Downing College of The University of Cambridge: Cambridge, U.K.

Oates J., Swain T., Zantovska J., 1977. Secondary compounds and food selection by *Colobus* monkeys. *Biochemical Systematics and Ecology*, 5: 317-321.

Ochman H., Worobey M., Kuo C. H., 2010. Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol* 8:e1000546.

Qin J., Li R., Raes J., *et al.*, 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010; 464:59-65.

Rawluszko A. A., Krokowicz P., Jagodzinski P. P., 2011. Butyrate induces expression of 17 $\beta$ -hydroxysteroid dehydrogenase type 1 in HT29 and SW707 colorectal cancer cells. *DNA Cell Biol* 30: 661–669.

Rodgers W. & Homewood K., 1982. Biological values and conservation prospects for the forests and primates of the Uzungwa Mountains, Tanzania. *Biological Conservation*. 24: 285-304.

Rosenfeld J. S., 2002. Functional redundancy in ecology and conservation. *Oikos* 98: 156.

Round J. L., Mazmanian S. K., 2009. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 9:313.

Rovero F., Mtui A. S., Kitegile A. S., Nielsen M. R., 2012. Hunting or habitat degradation? Decline of primate populations in Udzungwa Mountains, Tanzania: An analysis of threats. *Biological Conservation* 146: 89–96.

Russell J. A., Moreau C. S., Goldman-Huertas B., *et al.*, 2009. Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. *Proceedings of the National Academy of Sciences, USA*, 106, 21236–21241.

Sanger F., Nicklen S., Coulson A. R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*, 74 (1): 5463-5467.

Schloss P. D. & Handelsman J., 2007. The last word: books as a statistical metaphor for microbial communities. *Annu. Rev. Microbiol.* 61:23-34.

Schloss P. D., Handelsman J., 2004. Status of the microbial census. *Microbiol. Mol. Biol. Rev.* 68, 686–691.

Sekirov I., Russell S. L., Antunes L. C., Finlay B. B., 2010. Gut microbiota in health and disease. *Physiol. Rev.* 90, 859–904.

Sharma B., Singh N., 2011. Attenuation of vascular dementia by sodium butyrate in streptozotocin diabetic rats. *Psychopharm* 215: 677–687.

Shendure J., Porreca G. J., Reppas N. B., Lin X., McCutcheon J. P., Rosenbaum A. M., Wang M. D., Zhang K., Mitra R. D., Church G. M., 2005. Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome. *Science* 309 (5741): 1728–1732. DOI: 10.1126/science.1117389.

Shendure Li H., 2008. Next generation DNA sequencing. *Nat. Biotechnol.* 26: 1135–1145.

Shimada T., 2006. Salivary proteins as a defense against dietary tannins. *Journal of Chemical Ecology.* 32: 1149-1163.

Siex K., 2003. Effects of population compression on the demography, ecology, and behavior of the Zanzibar red colobus monkey (*Procolobus kirkii*). Ph.D. dissertation. Duke University: Durham, NC.

Sogin M. L., Morrison H. G., Huber J.A., Welch D.M., Huse S.M., Neal P.R., Arrieta J.M., Herndl G.J., 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. USA* 103, 12115–12120.

Stackebrandt E., Goebel B. M., 1994. Taxonomic note: a place for DNA–DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* 44, 846–849.

Steel PhD, 2012. The Effects of Habitat Parameters on the Behavior, Ecology, and Conservation of the Udzungwa Red Colobus Monkey (*Procolobus gordonorum*). Department of Biological Anthropology and Anatomy Duke University.

Struhsaker T. T., Marshall A. R., Detwiler K., Siex K., Ehardt C., Lisbjerg D. D., Butynski T. M., 2004. Demographic Variation Among Udzungwa Red Colobus in Relation to Gross Ecological and Sociological Parameters. *International Journal of Primatology*, 25 (3): 615-658.

Struhsaker T., 1975. *The Red Colobus Monkey*. University of Chicago Press: Chicago.

Struhsaker T., 2005. Conservation of red colobus and their habitats. *International Journal of Primatology*. 26.3: 525-538.

Struhsaker T., 2010. *The Red Colobus Monkeys: Variation in Demography, Behavior, and Ecology of Endangered Species*. Oxford University Press: Oxford.

Struhsaker T., Butynski T., Ehardt C., 2008. *Procolobus gordonorum*. In IUCN 2011, IUCN Red List of Threatened Species, Version 2011.2, [www.iucnredlist.org](http://www.iucnredlist.org). Accessed Feb. 1, 2012.

Traskalová-Hogenová H., Šteřpánková R., Kozáková H., Hudcovic T., Vannucci L., Tuczková L., Rossmann P., Hrnčíř T., Kverka M., Zákostelská Z., Klimesová K., Průbylová J., *et al.*, 2011. The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: contribution of germ-free and gnotobiotic animal models of human diseases. *Cellular & Molecular Immunology* 8, 110–120; doi:10.1038/cmi.2010.67.

Turnbaugh P. J., Hamady M., Yatsunenko T., *et al.*, 2009. A core gut microbiome in obese and lean twins. *Nature* 457:480–484.

Valles S. M., Strong C. A., Buss E. A., Oi D. H., 2012. Non-enzymatic hydrolysis of RNA in workers of the ant *Nylanderia pubens*. *J Insect Sci.* ;12:146. doi: 10.1673/031.012.14601.

Wang Q., Garrity G. M., Tiedje J. M., Cole J. R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267.

Wang Y., Kasper L. H., 2014. The role of microbiome in central nervous system disorders. *Brain Behav. Immun.* <http://dx.doi.org/10.1016/j.bbi.2013.12.015>.

Waterman P. e Kool K., 1994. Colobine food selection and plant chemistry. In A.G. Davies and J.F. Oates (Eds.), *Colobine Monkeys: Their Ecology, Behaviour, and Evolution*. Cambridge: Cambridge University Press.

Welch B. L., 1947. The generalization of "Student's" problem when several different population variances are involved. *Biometrika* 34 (1–2): 28–35. doi:10.1093/biomet/34.1-2.28.MR 19277.

Westoby M., 1974. An analysis of diet selection by large generalist herbivores. *The American Naturalist*, 108: 290-304.

Woese C, Kandler O, Wheelis M, 1990 Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*, 87 (12): 4576-9.

Woese C, Magrum L, Fox G, 1978 Archaeobacteria. in *J Mol Evol*, 11 (3): 245-51.

Woese C., Fox G., 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences* 74(11): 5088–5090.

Youssef N., Sheik C. S., Krumholz L. R., Najjar F. Z., Roe B. A., Elshahed M. S., 2009. Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl. Environ. Microbiol.* 75, 5227–5236.

Yu Z., Morrison M., 2004. Comparisons of different hypervariable regions of rrs genes for use in fingerprinting of microbial communities by PCR-denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 70, 4800–4806.



## RINGRAZIAMENTI

Ringrazio tutte le persone che mi hanno supportata e sopportata questo periodo che seppur stimolante, non è stato privo di difficoltà.

Grazie ai miei genitori, su cui posso sempre contare, immancabilmente pronti per una parola di conforto e di incoraggiamento. A Davide, che mi è sempre stato vicino e ha condiviso con me gioie, perplessità e momenti bui.

Grazie alle mie/miei compagne/i di corso. A Veronica, con la quale riesco a scappare dalla realtà anche solo nel tempo di una telefonata; ad Angelica, che mi pensa ed incoraggia nei momenti giusti; e a Tiziano, sapiente in questioni burocratiche e di segreteria. È sempre un piacere rivederli all'Agripolis!

Grazie a Daniela, Maddalena e a tutti quelli con cui ho collaborato in laboratorio, che mi hanno assistita e hanno saziato ogni mio dubbio e perplessità inerenti alla ricerca, ma anche per le belle chiacchierate.

Grazie a Claudia Barelli che mi ha consentito di collaborare al progetto "Primagut" e a Davide Albanese che ha avuto la pazienza di spiegarmi la parte più ostica dei test statistici di elaborazione dei dati.

Grazie a Massimo, non in qualità di correlatore, hai solo fatto il tuo dovere, ma in qualità di amico e mentore. In questi mesi hai sempre saputo darmi una risposta e la tua determinazione ha dato forza anche a me. Sono stati di particolare aiuto anche le pause caffè e gli articoli dell'Internazionale.

Grazie anche, e soprattutto, a Carlotta, per la fiducia che mi ha dato nel lavoro di laboratorio, lasciandomi spazio per sbagliare, imparare, per le sue correzioni e i suoi preziosi insegnamenti.