

Biodiversità Molecolare: Aspetti di base e nuove tecnologie



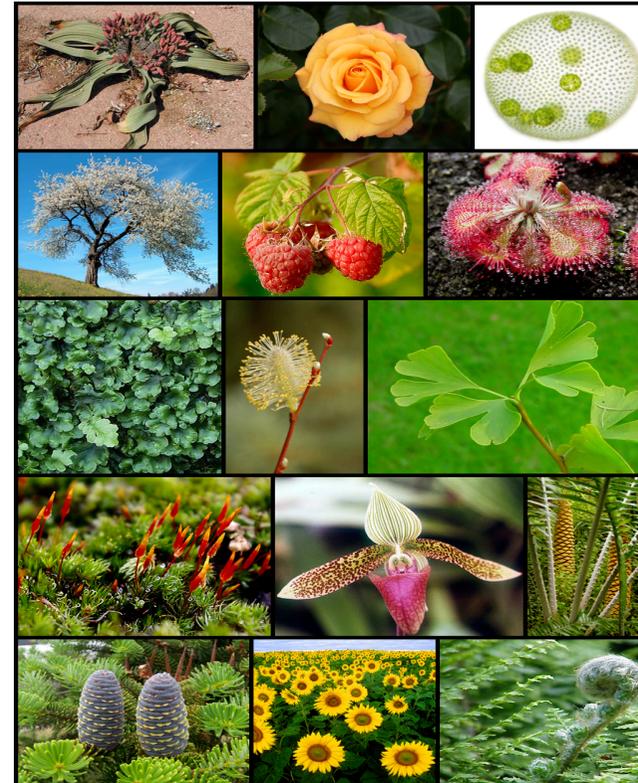
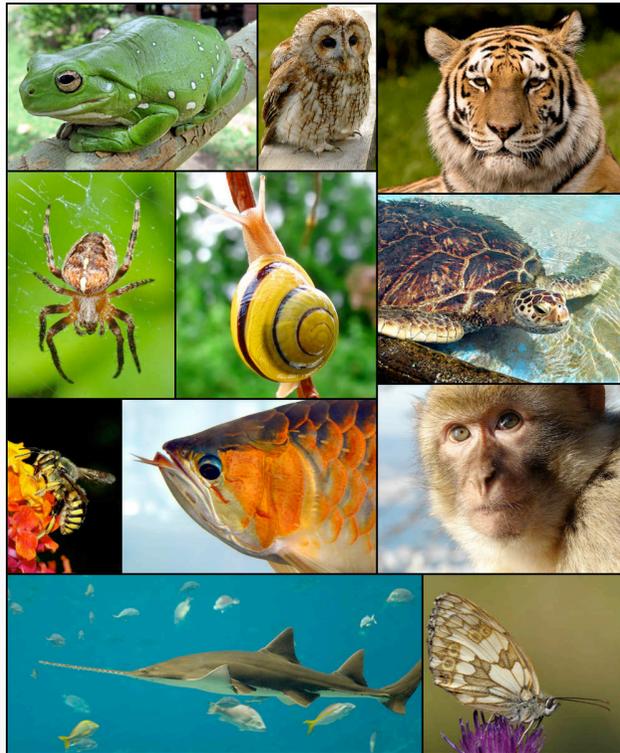
Graziano PESOLE

Università di Bari & IBBE-CNR, Bari, Italy

XVI Convegno Nazionale ANISN "Innovazione didattica e scelte sostenibili per lo sviluppo del territorio", Bari, 10 Settembre 2013

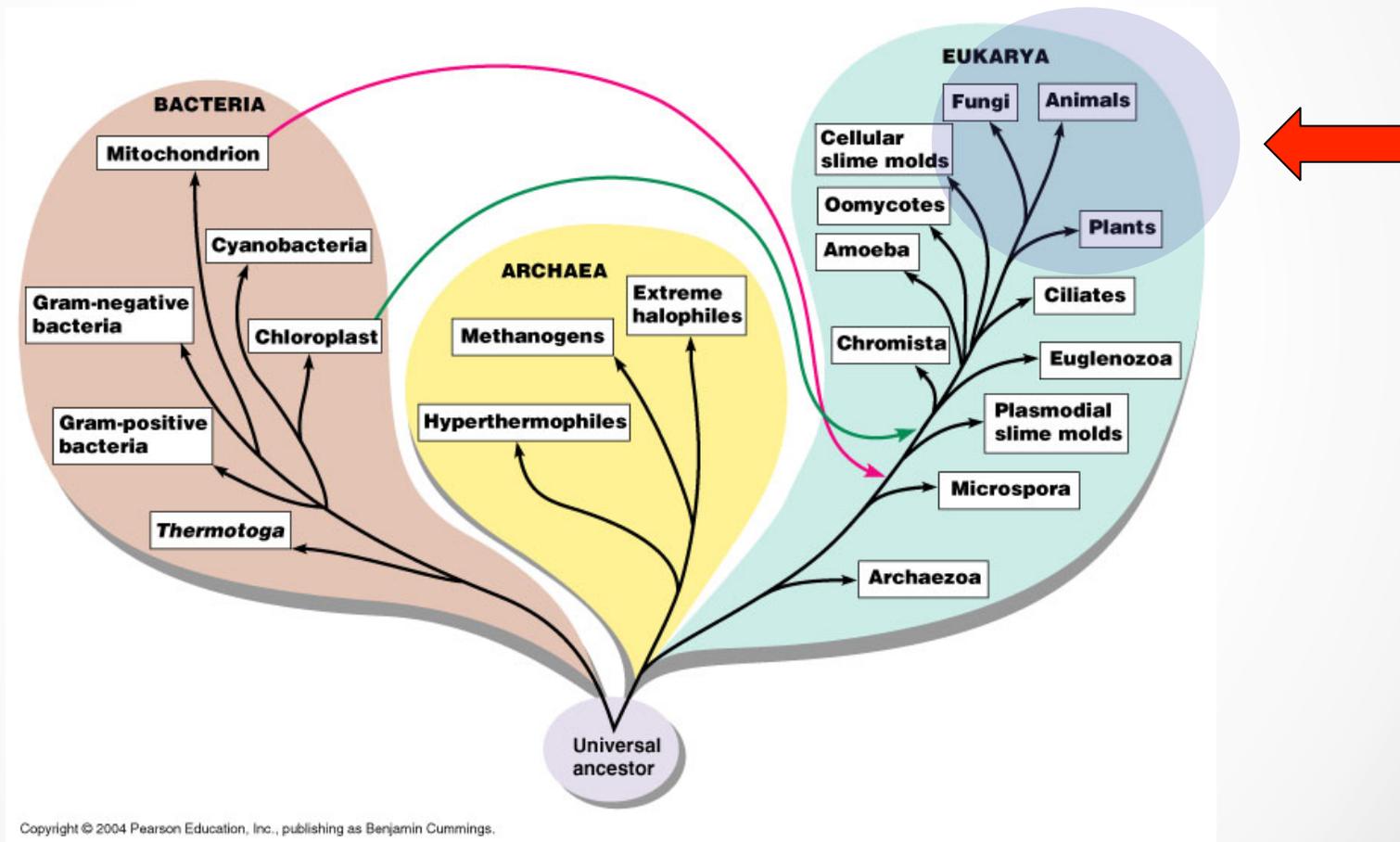
La Biodiversità

L'insieme di tutti organismi viventi, geneticamente distinti, e dei corrispondenti ecosistemi costituisce quella che viene definita la Biodiversità del nostro pianeta. Quindi la Biodiversità include tutte le forme di variabilità biologica: a livello di gene, specie, habitat ed ecosistema.



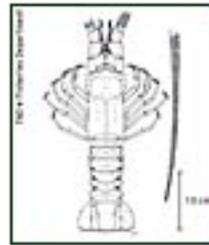
Biodiversità a livello di specie

La maggior parte della biodiversità del pianeta è costituita da organismi microscopici (batteri, virus e eucarioti unicellulari), e la gran parte di essi non è stata ancora classificata.

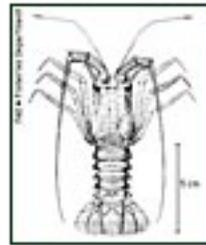


Come si riconosce una specie?

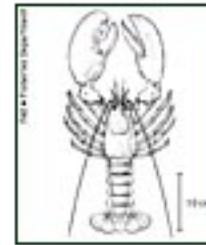
Tipicamente, almeno nell'ambito di animali e piante, una specie può essere riconosciuta, e quindi classificata, sulla base delle sue caratteristiche morfologiche. Allo stesso modo le specie possono essere raggruppate in livelli tassonomici superiori.



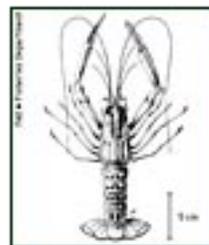
African spear
lobster



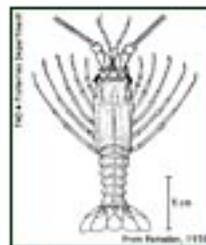
American blunthorn
lobster



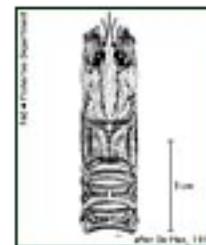
American lobster



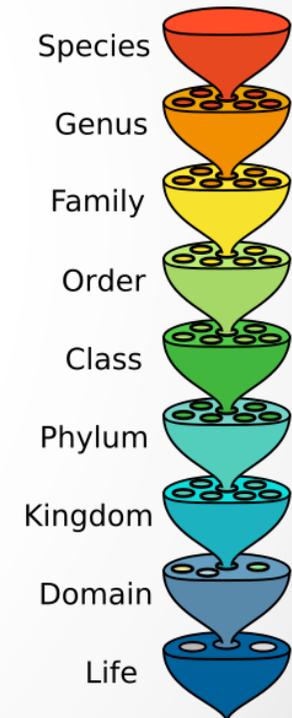
Andaman lobster



Arabian whip
lobster



Arafura lobster



Come si riconosce una specie?

Tuttavia, è possibile osservare un elevato livello di variabilità morfologica nell'ambito della stessa specie.



Come si riconosce una specie?

... e specie differenti possono apparire quasi identiche a livello morfologico (es. *Perichares* sp.)

P. adela



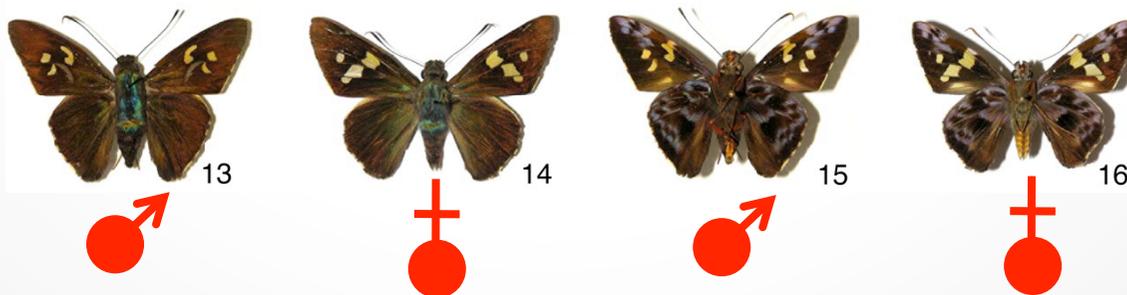
P. poaceaphaga



P. geonomaphaga



P. prestoeaphaga



Un test del DNA per l'identificazione delle specie?

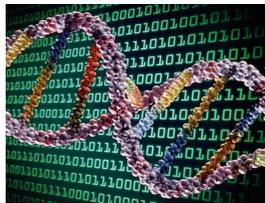


Un approccio "oggettivo" per la caratterizzazione e classificazione delle specie è quello di utilizzare l'informazione racchiusa nel materiale genetico (DNA) che come sappiamo è specifico per ciascun individuo e specie. Mentre le differenze in termini di sequenza nucleotidica tra i genomi di individui della stessa specie sono molto piccole (di norma $<1\%$), quelle tra specie diverse sono proporzionali alla divergenza evolutiva, ma comunque significativamente maggiori delle differenze intraspecifiche (di norma $\geq 3\%$).

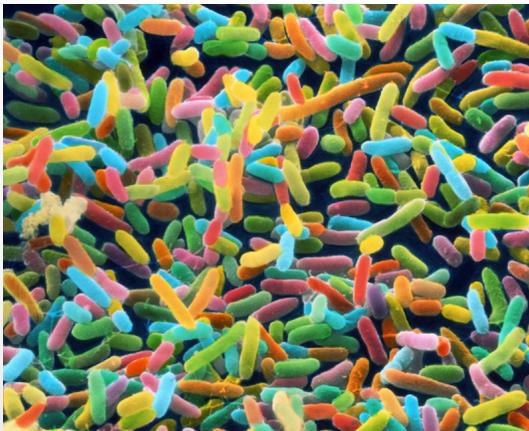
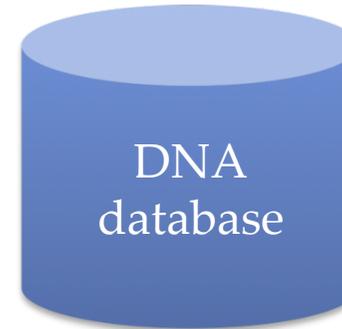
Dato che i genomi di moltissime specie sono stati completamente caratterizzati (<http://www.genomesonline.org>) nota la sequenza di un nuovo genoma (o più ragionevolmente di una sua frazione più o meno estesa, es. **DNA Barcode**) è possibile, attraverso una analisi comparativa, stabilire la specie di appartenenza o almeno il gruppo tassonomico di riferimento.

Qualora il campione in esame contenga più specie, come nel caso di campioni di cui si voglia esaminare la popolazione microbica, l'analisi delle sequenze di DNA presenti nel campione può dare indicazioni oggettive sulla popolazione microbica del campione a livello qualitativo e quantitativo. L'analisi genetica per l'identificazione delle specie risulta quindi indipendente dalla definizione delle caratteristiche morfologiche, che in molti casi (es. batteri) non sono facilmente riconoscibili.

Un test del DNA per l'identificazione delle specie?

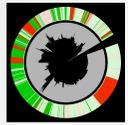


Analisi comparativa

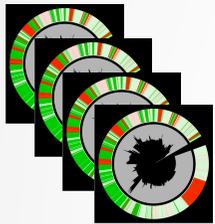


Assegnazione di specie
o gruppo tassonomico

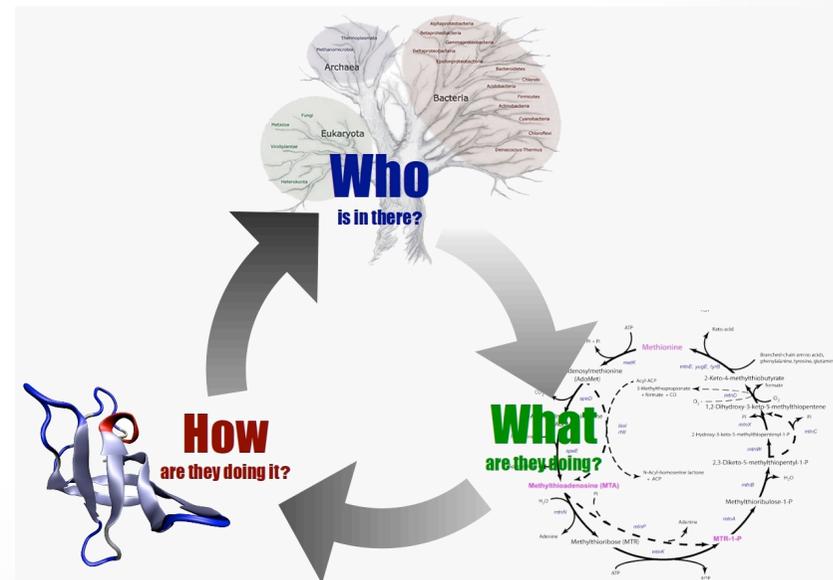
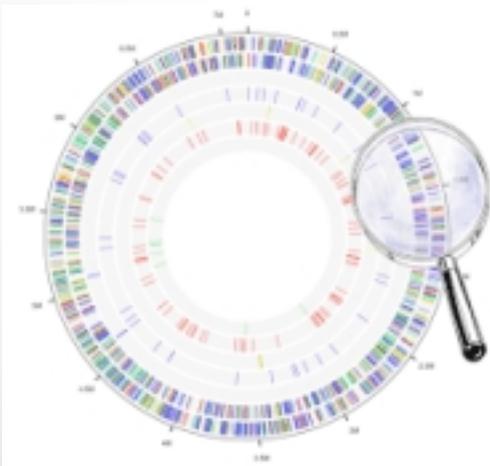
Genomica e Metagenomica



Genomica: studio del patrimonio genetico di uno specifico organismo per comprenderne i meccanismi alla base delle sue caratteristiche fisiologiche e funzionali

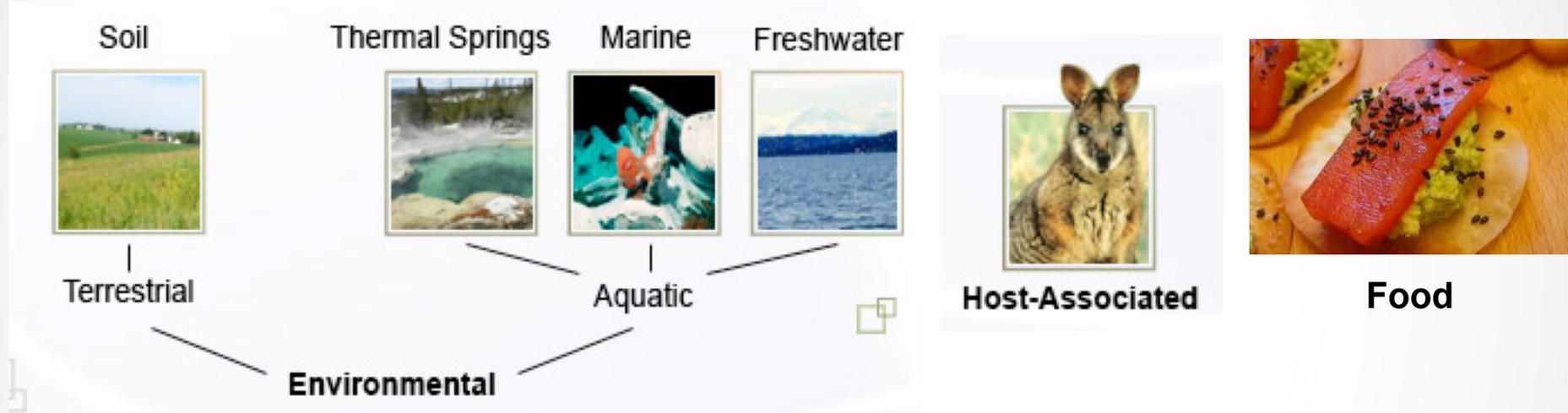


Metagenomica: studio del patrimonio genetico complessivo di una comunità di organismi per comprendere le caratteristiche fisiologiche e funzionali della intera popolazione anche in relazione all'ambiente in cui essa vive. Lo studio della metagenomica si rivolge prevalentemente al mondo dei microrganismi.



Metagenomica: definizione e obiettivi

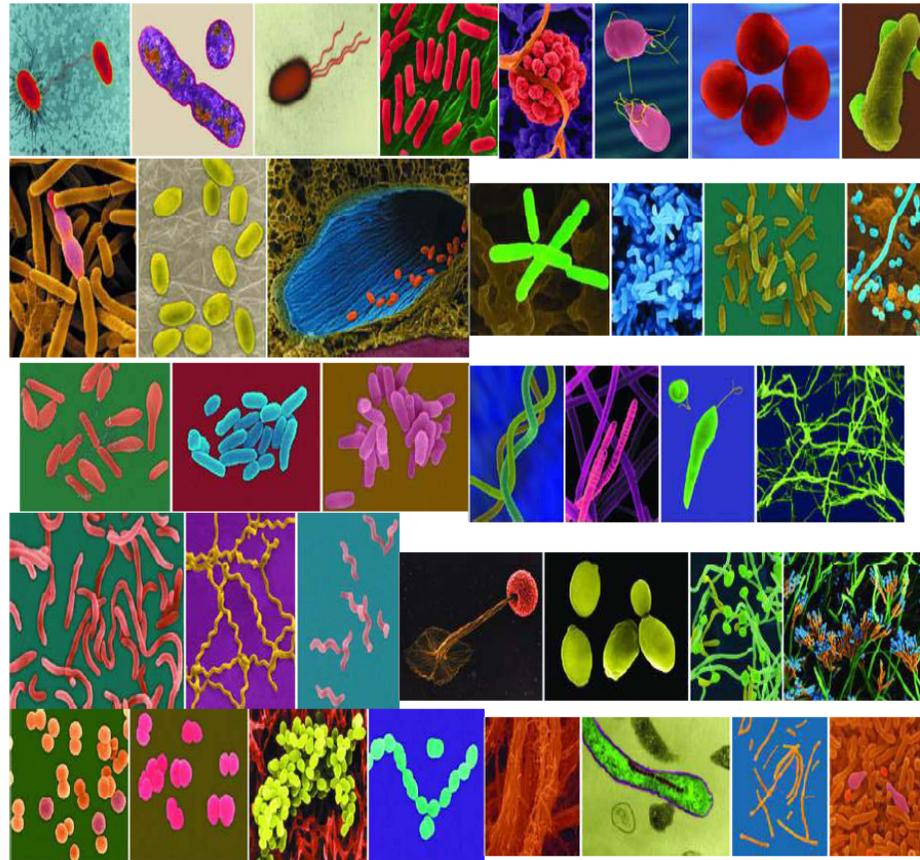
Un ambiente (es. acqua, suolo, etc.) o uno specifico organismo è tipicamente popolato da una comunità più o meno complessa di microorganismi.



- L'insieme del materiale genetico presente in un campione, definito **metagenoma**, sarà quindi rappresentativo delle specie che lo popolano. Analogamente il **metatrascrittoma** sarà rappresentativo delle sue funzioni biologiche.
- L'assunto base della Metagenomica è quindi che la **Biodiversità** tassonomica e funzionale di qualunque ambiente è pienamente rappresentata dal materiale genetico in esso presente.

Metagenomica microbica

Fino ad oggi è stato possibile studiare solo una piccolissima parte (~ 1%) delle specie microbiche, ovvero quelle coltivabili in laboratorio. La metagenomica permette di svelare il **100% del patrimonio genetico di un ambiente** e quindi di caratterizzarne la biodiversità delle comunità microbiche.



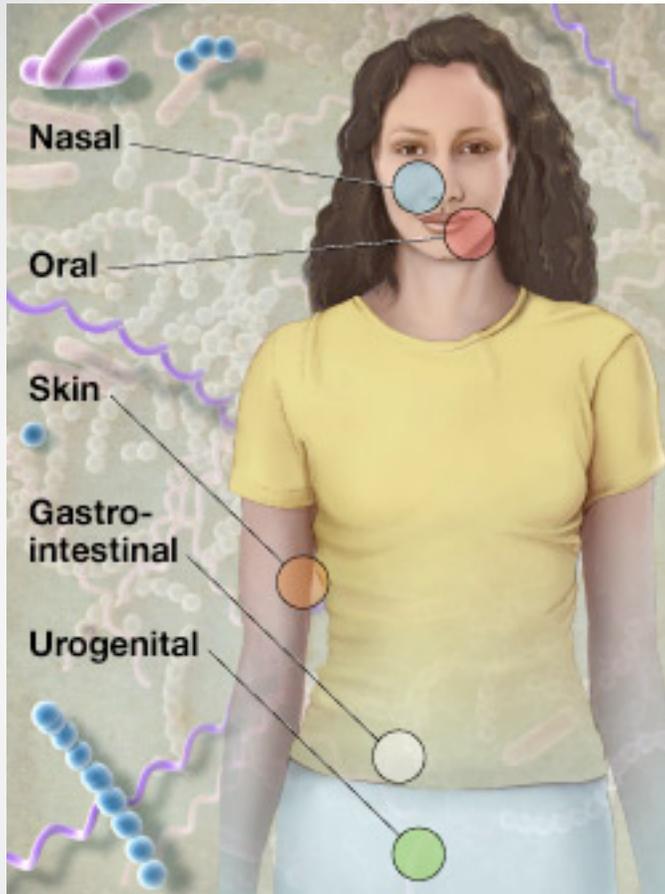
Metagenomica microbica

Il microbiota, soprattutto quello residente in ambiente estremi, costituisce un enorme serbatoio di preziose funzioni biologiche, che potranno essere utilizzate in processi e prodotti biotecnologici, una volta identificati i geni o le vie metaboliche responsabili.



- Risanamento dei siti tossici di rifiuti su scala planetaria
- Produzione di nuovi agenti e percorsi terapeutici e preventivi.
- Generazione di energia e di sviluppo delle fonti rinnovabili di energia (ad esempio, il metano e l'idrogeno).
- Produzione di catalizzatori chimici, reagenti ed enzimi per migliorare l'efficienza dei processi industriali.
- Gestione del biossido di carbonio ambientale, legato ai cambiamenti climatici.
- Individuazione degli organismi che causano malattie e il monitoraggio della sicurezza dei rifornimenti di cibo e acqua.
- Uso di batteri geneticamente modificati come sensori viventi (biosensori) per rilevare sostanze chimiche nocive nel suolo, aria, o acqua.
- Comprensione dei sistemi specializzati utilizzati dalle popolazioni microbiche a vivere in ambienti naturali specifici.
- Studio delle reti trofiche in diversi sistemi ecologici, e dei meccanismi per la diffusione di specie aliene e invasive.

Il Microbioma umano



Anche l'uomo è un ambiente estremamente complesso. Il nostro corpo è costituito da circa 10^{13} cellule, ma contiene un numero dieci volte maggiore (circa 10^{14}) di cellule batteriche. Il cosiddetto "microbiota umano" ha una profonda influenza sulla fisiologia dell'organismo, sulla nutrizione, e risulta cruciale per la nostra salute. Difatti, esso fornisce nutrienti e vitamine, coadiuva la risposta alle infezioni e la detossificazione da diverse sostanze tossiche.

La **Metagenomica** ora rende possibile la caratterizzazione della composizione e della dinamica di popolazione della comunità microbica umana, e le interazioni cooperative (o antagoniste) con le cellule e i tessuti umani.



Next-Generation Sequencing (NGS)

L'avvento delle piattaforme di sequenziamento di nuova generazione (NGS, Next Generation Sequencing) rappresenta una rivoluzione epocale nella ricerca biomolecolare. Il crescente volume di dati che è possibile produrre, e i costi sempre più contenuti, aprono prospettive prima non immaginabili in moltissimi ambiti di ricerca e offrono numerosissime applicazioni.



Roche / 454 Genome Sequencer FLX titanium
(800 bp, 800 Mb / run)



Ion Proton



Illumina / Solexa
Genetic Analyzer HiSeq 2000
(150x2 bp, 600 Gb / run)



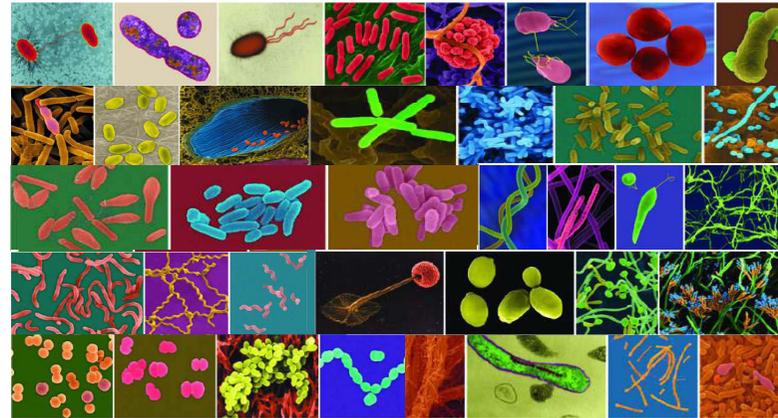
PacBio



Applied Biosystems
SOLiD 4 System™
(100x2 bp, 400 Gb / run)

Metagenomica e NGS

Le piattaforme di sequenziamento di nuova generazione rappresentano uno strumento eccezionale per analizzare in tempi rapidi e costi accessibili, a seguito della produzione di milioni di sequenze nucleotidiche, una grande varietà di campioni in modo da ampliare in modo straordinario le nostre conoscenze sulla loro biodiversità tassonomica e funzionale.



L'approccio metagenomico non richiede l'isolamento e la messa in coltura di ogni singola specie, e la sua identificazione su basi biochimiche o morfologiche. Esso inoltre rende possibile l'esplorazione di un gran numero di ambienti e condizioni diversi.

Una grande varietà di campioni può essere analizzata (es. acqua, suolo, sedimenti, campioni bentonici, contenuto di organi cavi, feci, etc.) ottenendo una rappresentazione pressoché esaustiva della composizione tassonomica e delle proprietà funzionali di popolazioni complesse. Naturalmente, l'interpretazione delle sequenze prodotte si basa sulle conoscenze pregresse di carattere genetico o biochimico-funzionale.

Approcci sperimentali per la Metagenomica



Metagenomica Target-oriented (basata su amplicon-sequencing)

Sequenziamento massivo e in parallelo di una specifica regione target (es. rRNA 16S o ITS) ottenuta attraverso amplificazione specifica utilizzando primer universali per un ampio raggruppamento tassonomico (es. batteri o funghi)



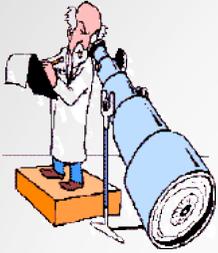
Metagenomica shot-gun

Sequenziamento shotgun del DNA (o RNA) totale estratto da un dato campione.

Il primo approccio è particolarmente indicato se si focalizza l'indagine su specifici raggruppamenti tassonomici, per cui sono disponibili primer adeguati per amplificare una specifica regione target in un ampio raggruppamento tassonomico. Il secondo approccio, non focalizzato, consente di identificare specie per cui non si dispone di marker universali (es. virus).

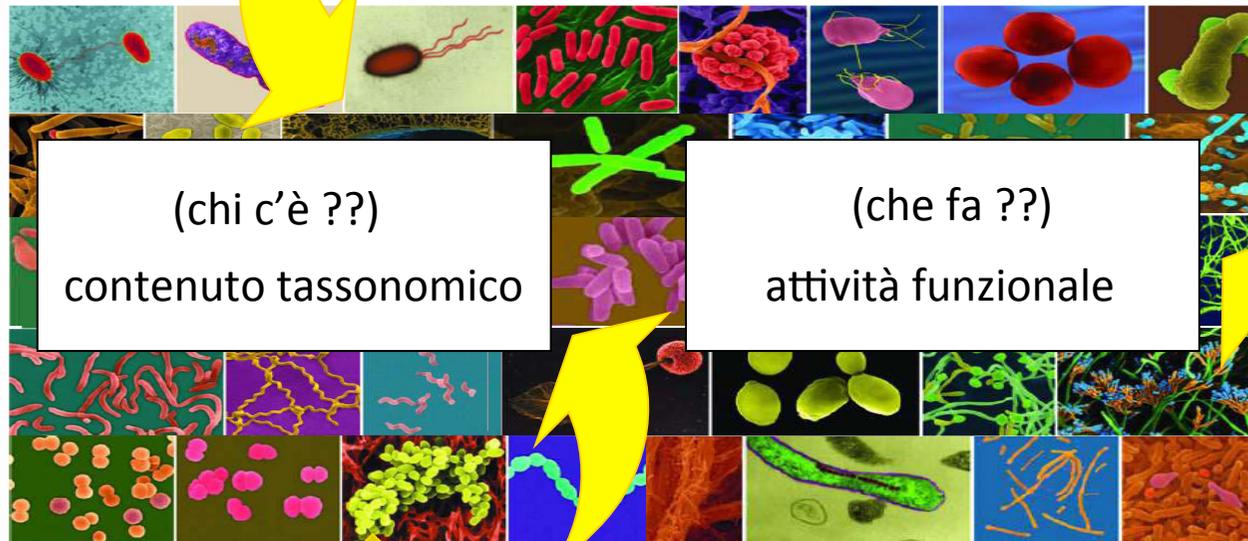
L'analisi **meta-trascrittomica**, consente di ottenere un quadro del profilo di espressione dei geni, e quindi le funzioni biologiche attive in una determinata condizione ambientale.

NGS per la Metagenomica



Shotgun NGS

- identifica specie, geni e attività funzionali di comunità microbiche complesse
- coverage limitato per rilevare specie rare
- costoso sia in termini di sequenziamento che di analisi computazione



Sequenziamento massivamente parallelo di ampliconi

- elevata sensibilità nella identificazione di specie
- poco costoso in termini di sequenziamento che di analisi computazione
- basato su PCR con primer universali
- richiede appropriato database di riferimento (es. RDP per 16S, ITSoneDB per ITS1)
- possibili distorsioni dovute a diversa efficienza di amplificazione in diversi gruppi tassonomici
- nessuna informazione funzionale

Metagenomica target-oriented

E' necessario definire una sequenza target di DNA che sia **ubiquitaria** nel raggruppamento tassonomico di interesse, e che risponda alle seguenti caratteristiche peculiari:



Regioni fiancheggianti altamente conservate che permettano il disegno di primer universali per l'amplificazione di un ampio spettro di specie

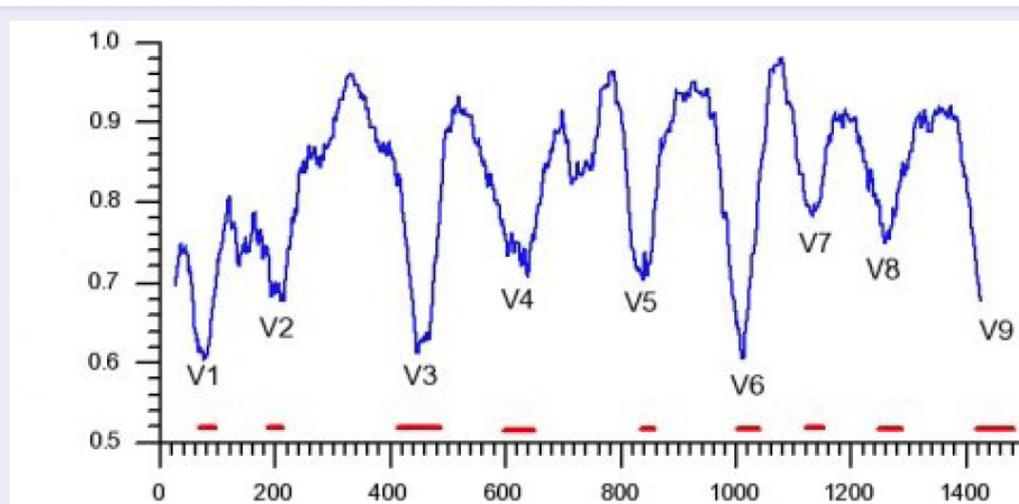


Sufficiente variabilità della regione interna che consenta di discriminare facilmente la variabilità inter-specifica da quella intra-specifica.



Dimensioni dell'amplicone compatibili con la piattaforma NGS utilizzata.

Small subunit 16S ribosomal RNA structure

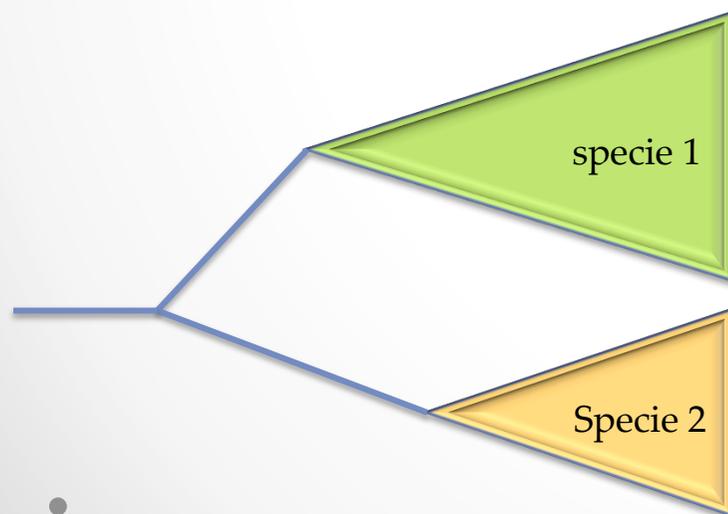


Le regioni variabili dell'rRNA 16S batterico (V1-V9) vengono generalmente utilizzate per l'analisi metagenomica nei batteri.

Metagenomica target-oriented

Differenti regioni target sono state definite per diversi gruppi tassonomici, sulla base delle raccomandazioni del Consorzio sul “Barcode of Life” (CBOL, <http://www.barcodeoflife.org/>).

Taxon	Target
Bacteria	Regioni variabili del 16S rRNA
Fungi	Regioni trascritte interne del cluster degli rRNA (ITS)
Metazoa	Cox1 mitocondriale
Plants	matK e rbcL di cloroplasto



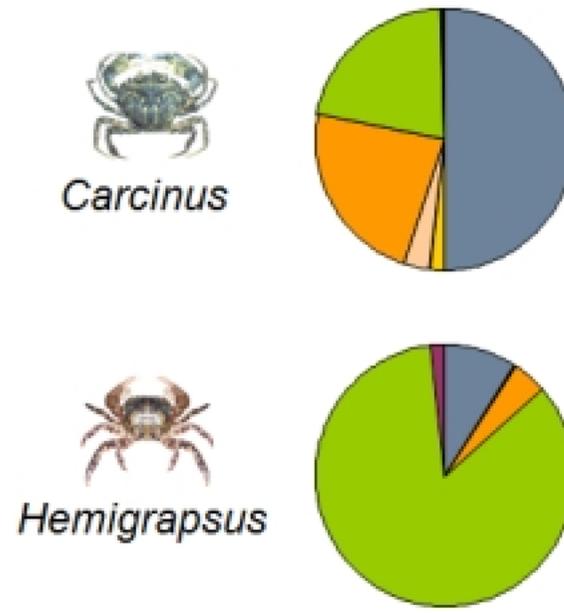
Perchè la regione target (o DNA barcode) sia appropriata, essa deve essere facilmente amplificabile e la minima variabilità interspecifica essere significativamente maggiore della massima variabilità intraspecifica.

Metagenomica target-oriented: Applicazioni

Chi mangia cosa, e quanto?

Una applicazione molto interessante della metagenomica "target-oriented" è la caratterizzazione delle abitudini alimentari di diversi gruppi di organismi, in diverse condizioni. Il processo digestivo rende poco pratico il riconoscimento morfologico della composizione tassonomica della dieta, che può invece essere facilmente ottenuta dall'analisi metagenomica. Questa informazione è cruciale per comprendere le dinamiche di espansione/contrazione delle popolazioni e la suddivisione delle risorse trofiche.

Ad esempio, l'analisi metagenomica del contenuto dell'intestino di due specie di granchi che vivono in uno stesso ambiente può essere molto informativa per comprendere la loro dinamica di popolazione e le relative capacità invasive.



Metagenomica Shotgun

I sequenziamento shotgun di acidi nucleici estratti da campioni ambientali o clinici, può essere condotto a livello di DNA o RNA e può fornire simultaneamente una panoramica della composizione tassonomica e funzionale del microbiota.

Nel caso in cui non si disponga di specifici target (es. virus) questa strategia è l'unica possibile.

In particolare, l'analisi del meta-trascrittoma consente di ottenere la caratterizzazione funzionale della comunità microbica in toto, rivelando i geni espressi e le vie metaboliche e i processi funzionali attivi.



Environmental Genome Shotgun Sequencing of the Sargasso Sea

J. Craig Venter,^{1*} Karin Remington,¹ John F. Heidelberg,³
Aaron L. Halpern,² Doug Rusch,² Jonathan A. Eisen,³
Dongying Wu,³ Ian Paulsen,³ Karen E. Nelson,³ William Nelson,³
Derrick E. Fouts,³ Samuel Levy,² Anthony H. Knap,⁶
Michael W. Lomas,⁶ Ken Nealson,⁵ Owen White,³
Jeremy Peterson,³ Jeff Hoffman,¹ Rachel Parsons,⁶
Holly Baden-Tillson,¹ Cynthia Pfannkoch,¹ Yu-Hui Rogers,⁴
Hamilton O. Smith¹

Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products

Jo Handelsman¹, Michelle R Rondon¹, Sean F Brady², Jon Clardy² and Robert M Goodman¹



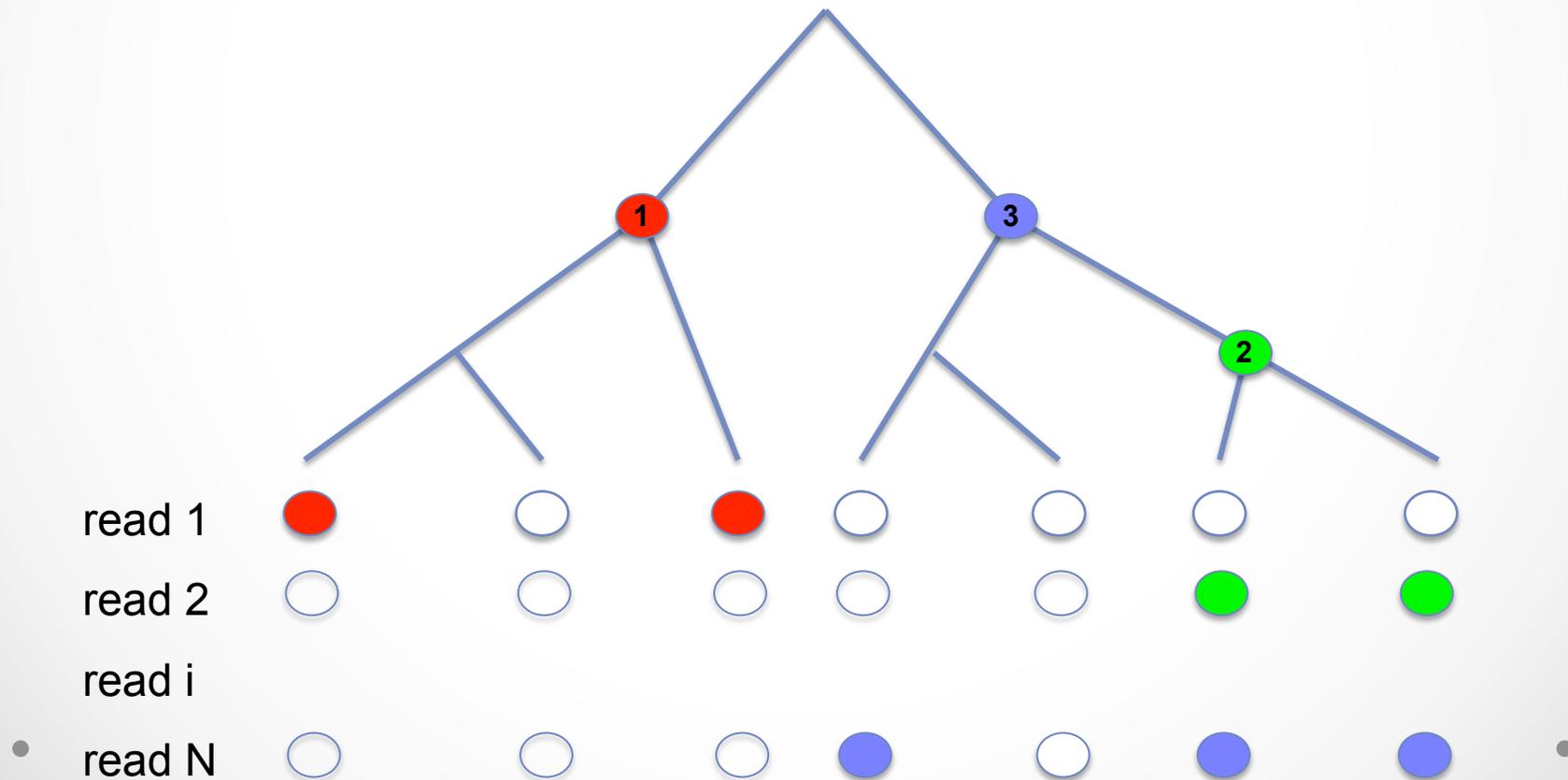
Bioinformatica per la Metagenomica

La analisi metagenomiche basate su approcci NGS producono un enorme volume di dati, corrispondenti a decine di milioni di sequenze nucleotidiche, che richiede strumenti adeguati per il loro **immagazzinamento**, **organizzazione** ed **analisi**. Sono quindi fondamentali **strumenti bioinformatici** appropriati, che tipicamente effettuano il confronto (allineamento) tra le sequenze ottenute (read) e uno o più database di riferimento, in cui le sequenze siano state ben caratterizzate dal punto di vista tassonomico e/o funzionale.



Classificazione tassonomica

Nella determinazione della composizione tassonomica di una collezione di read metagenomiche si assume normalmente che se la % di identità tra una read incognita e una sequenza di riferimento è superiore ad una certa soglia (tipicamente 97%), la stessa read eredita gli attributi tassonomici della sequenza di riferimento. Tuttavia, spesso si osservano ambiguità che possono essere risolte o assegnando la read al cosiddetto "lowest common ancestor" o usando approcci di massima verosimiglianza che normalmente fanno l'assegnazione del genere/specie più probabile.



Prospettive della Metagenomica



Ambiente ed ecologia

per nuove teorie e previsioni sulle interazioni tra comunità microbiche e la qualità dell'ambiente o i cambiamenti climatici

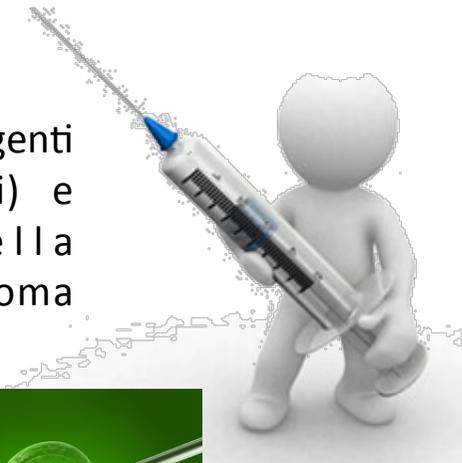


Agricoltura

Per elaborare nuove strategie per la difesa dei raccolti e migliorare la qualità e la sicurezza degli alimenti

Salute

per la diagnosi genetica di agenti infettivi (anche sconosciuti) e studiare l'effetto della composizione del microbioma sull'insorgenza di patologie.



Industria

per l'identificazione di nuove capacità biosintetiche e biocatalitiche delle comunità microbiche al fine di processi e prodotti utili in ambito industriale.

Energia

per sviluppare nuove risorse bioenergetiche più economiche, sostenibili da un punto di vista ambientale e meno soggette alle imprevedibilità del nostro pianeta.



Nodo di Bari – Infrastruttura Avanzata per lo studio della Biodiversità Molecolare

1

Laboratorio di Genomica e high-throughput sequencing

2

Laboratorio di Biologia Funzionale e dei Sistemi

3

Piattaforma Bioinformatica

4

Laboratorio per la gestione e l'analisi della Biodiversità microbica

5

Laboratorio di Biodiversità Vegetale

-3.70

-3.70

Laboratorio virtuale “Biomolecolare”



High-throughput sequencing

Anna Maria D’Erchia	Alessio Valletti
Caterina Manzari	Francesca Mastropasqua
Marinella Marzano	Teresa De Filippis
Giuseppe Sgaramella	Italia Aiello
Marianna Intranuovo	Valentino Pousis

System Biology e Bioenergetica

Anna Atlante

Bioinformatica

Monica Santamaria	Ernesto Picardi
Bruno Fosso	Domenico Simone
Bachir Balech	

System & Quality Management

Francesca De Leo	Giuseppe Sgaramella
------------------	---------------------