



IZZAM G. CAPORALE  
TERAMO

  *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

## NGS su ceppi di *Listeria monocytogenes* isolati da matrici alimentari

Cesare Cammà  
Massimiliano Orsini





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Next Generation Sequencing

## WGS

per *Listeria monocytogenes*?





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Whole Genome Sequencing

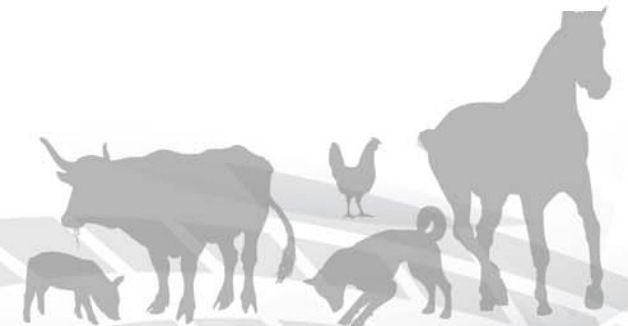


Scientific Colloquium Series

**EFSA Scientific Colloquium N°20**

*Use of Whole Genome Sequencing (WGS) of food-borne pathogens for  
public health protection*

**16-17 June 2014 | Parma, Italy**





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Tipizzazione molecolare dei patogeni trasmessi dagli alimenti



- Indagini sui focolai di tossinfezione alimentare
- Studio di ceppi con caratteristiche di elevata virulenza
- Analisi di ceppi che possiedono resistenza multipla agli antibiotici
- EFSA in collaborazione con ECDC, EURL ha in corso un progetto a livello europeo per la condivisione dei dati sulla tipizzazione molecolare (TESSy)





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

**PFGE**

## Gold standard per studi di epidemiologia molecolare

Metodo sviluppato circa 30 anni fa

- A volte i risultati sono di difficile interpretazione
- In qualche caso è necessario un maggiore potere discriminatorio tra ceppi che presentano lo stesso pulsotipo

**I continui sviluppi tecnologici in ambito NGS hanno portato a un aumento della produttività e ad una diminuzione dei costi**





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Next Generation Sequencing



In USA CDC e FDA hanno sviluppato diverse iniziative per diffondere l'uso del WGS in ambito microbiologico allo scopo di:

- Confrontare ceppi isolati con quelli presenti nel database
- Verificare se alcuni ceppi formano un cluster (appartengono allo stesso focolaio)
- Controllare se esiste somiglianza tra ceppi isolati dai casi clinici, dal cibo, dall'ambiente





**Listeria monocytogenes**  
Laboratorio Nazionale di Riferimento

# FDA Genome Tracker

## Basic Data Flow for Global WGS Public Access Databases

### DATA ACQUISITION

Sequence and upload genomic and geographic data



Other distributed sequencing networks



### DATA ASSEMBLY, ANALYSIS, AND STORAGE

International Nucleotide Sequence Database Collaboration (INSDC)

Shared Public Access Databases

- NCBI – National Center for Biotechnology Information
- EMBL – European Molecular Biology Laboratory
- DDBJ – DNA Databank of Japan



### PUBLIC HEALTH APPLICATION AND INTERPRETATION OF DATA

- Find clinical links
- Identify clusters
- Conduct traceback
- Develop rapid methods
- Develop culture independent tests
- Develop new analytical software



11/2014

State, Local, Federal, and Foreign Public Health Agencies

Academia/Industry



© 2011 Disney/Pixar

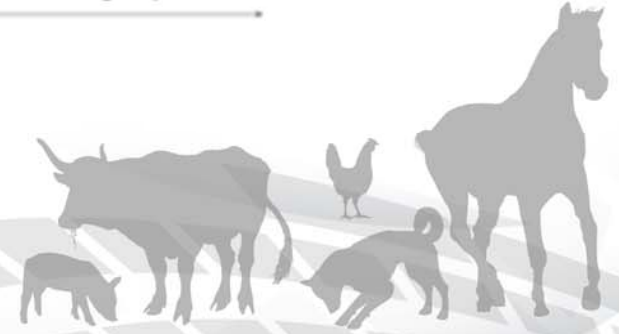
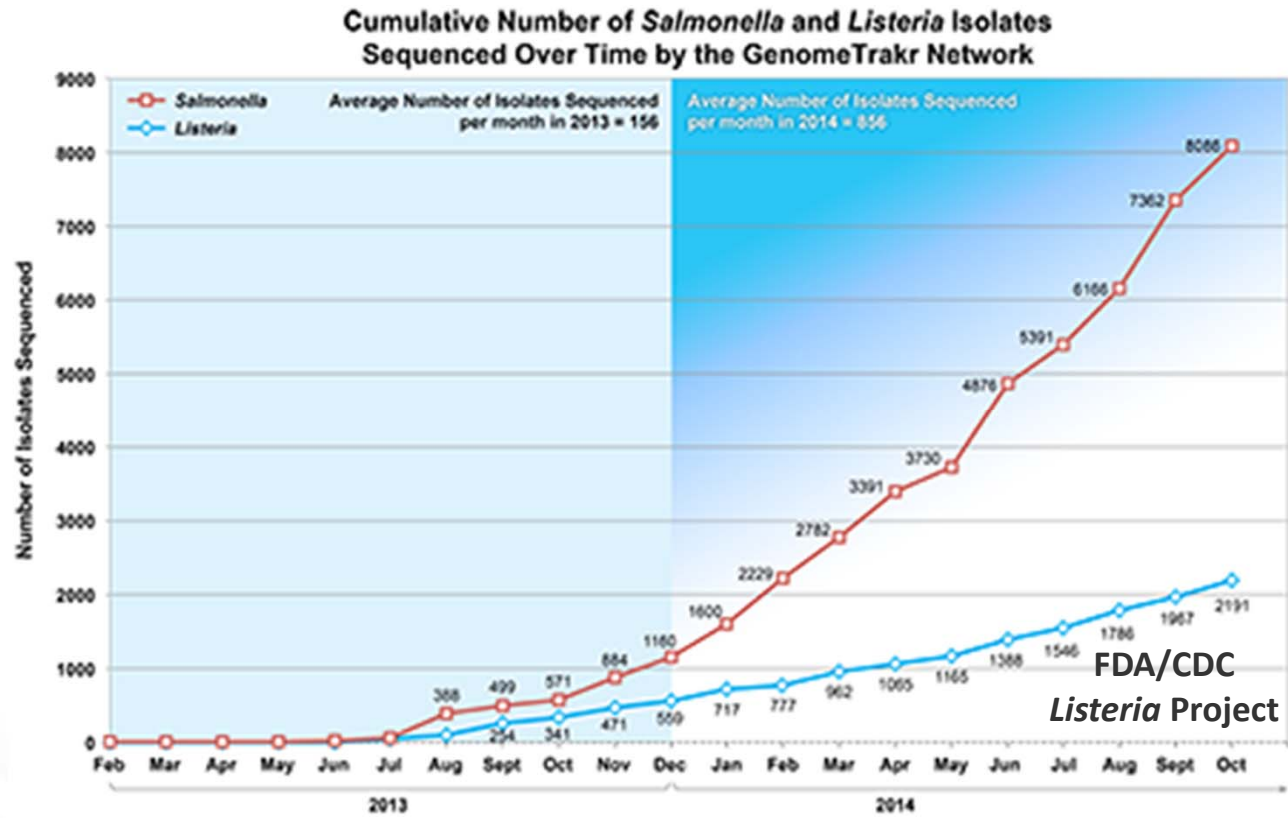
To the infinity... and beyond!!





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Quanti ceppi sequenziati?



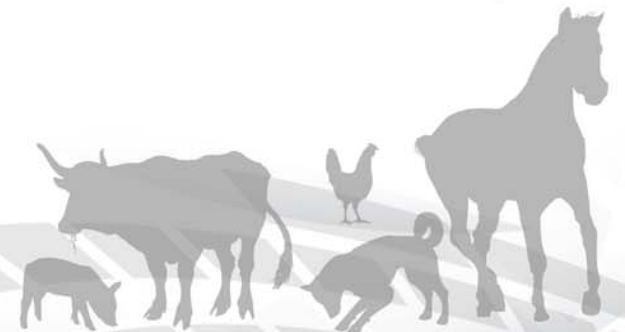




 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# A cosa serve WGS?

- **Epidemiologia molecolare**
  - Rintracciare la fonte di infezione
  - Gestire i focolai
  - Valutare la persistenza dei patogeni nell'ambiente
  - Monitorare la presenza di patogeni emergenti





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# A cosa serve WGS?

## Genomica funzionale

- Studio dei fattori di virulenza
- Ricerca di geni di resistenza agli antibiotici
- Interazione ospite/patogeno

## Diagnosi molecolare

- Sviluppo di nuovi metodi diagnostici rapidi e specifici



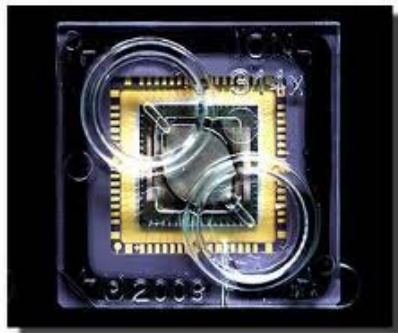


 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# ion torrent



by *life* technologies™

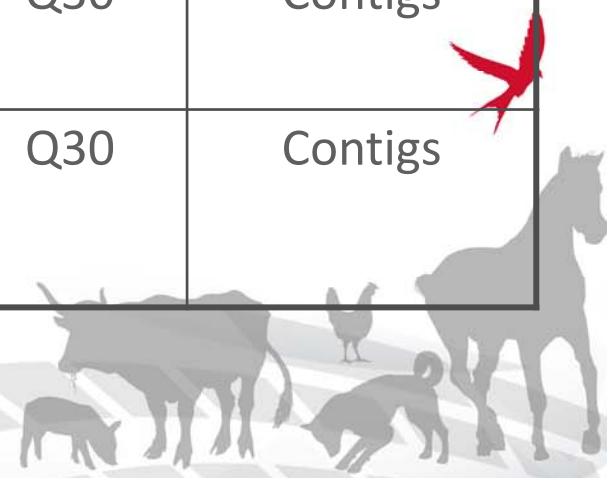


# NGS wet lab



# NGS su ceppi di *Listeria monocytogenes*

n. isolati	Matrice	Piattaforma NGS	Coverage	Qualità	Status
1	Isolato clinico	Illumina	>>400X	Q40	Genoma annotato
11	Alimenti ambiente	Illumina	>>400X	Q40	Contigs
4	Ceppi di referenza	Ion Torrent	20X	Q30	Contigs
64	Ambiente gorgonzola	Ion Torrent	20X	Q30	Contigs

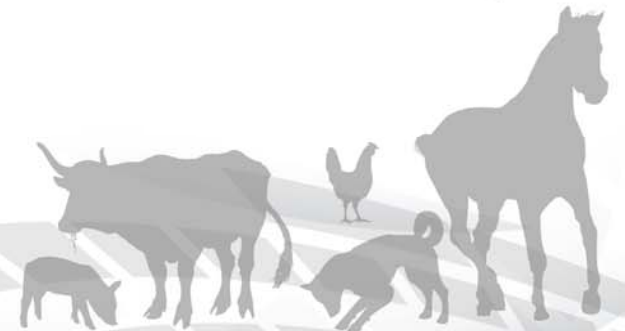




 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Risultati preliminari

- **Un ceppo isolato da un caso clinico (2008)**
  - Genoma completo e annotato manualmente
  - Inviato a GenBank
  
- **11 ceppi ambientali e da alimenti**
  - Genomi parziali assemblati (coverage >99%)
  - Annotazione automatica

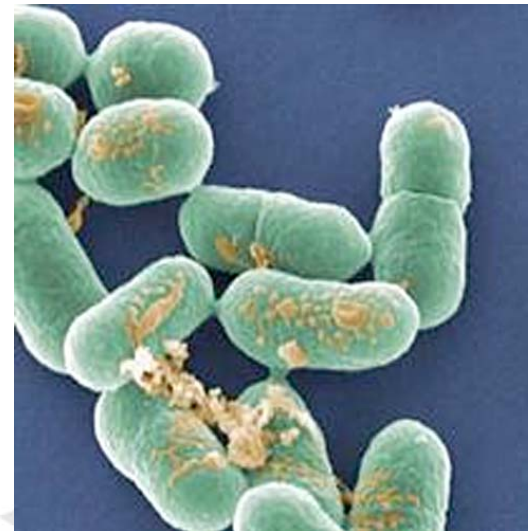


# *Listeria monocytogenes* e gorgonzola



Collaborazione tra il Consorzio per la tutela del formaggio gorgonzola e il Laboratorio Nazionale di Riferimento per *Listeria monocytogenes*

- Definire le misure di contenimento in grado di eliminare o ridurre *L.monocytogenes* nelle aziende di lavorazione
- Determinare il livello di rischio di listeriosi per il consumatore





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# *Listeria monocytogenes* e gorgonzola

- Azienda a conduzione familiare
- Stabilimento grande che svolge attività di stagionatura per conto di altre aziende



# *Listeria monocytogenes* e gorgonzola

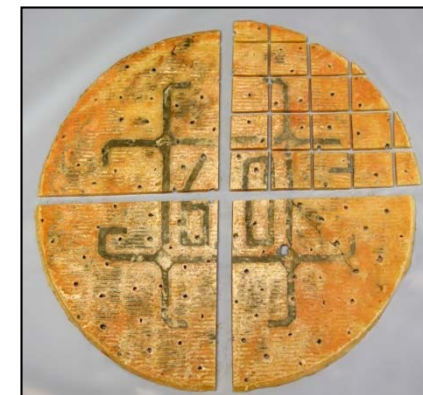


## Analisi ambientali:

- superfici a contatto con il gorgonzola (utensili e macchinari)
- superfici non a contatto (canaline di scarico del siero e tombini)

## Analisi del prodotto finito:

- prelievo di campioni dalla crosta e della pasta







*Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Analisi dei campioni

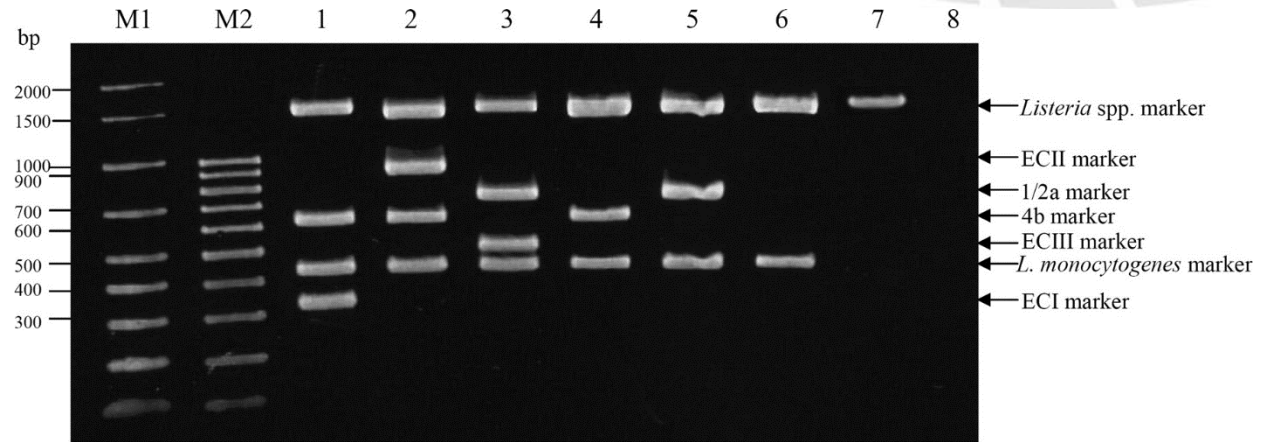


Isolamento

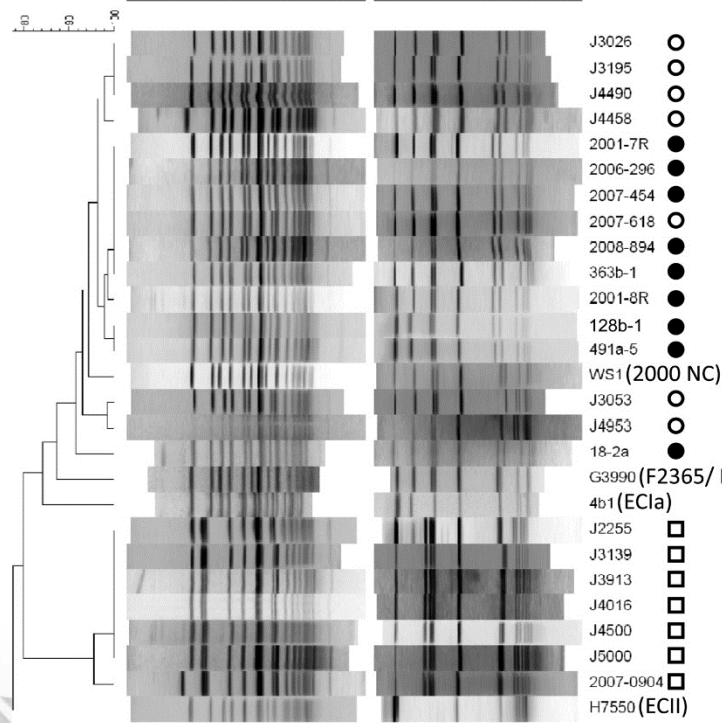
PCR

Serotyping

PFGE (*ApaI*, *AscI*)

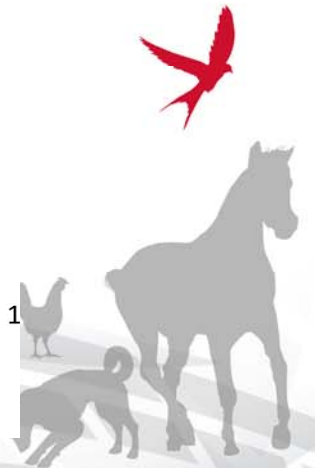


PFGE - *ApaI* / PFGE - *AscI*



Groups 2 and 3

Group 1





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Risultati preliminari NGS *L. monocytogenes* in gorgonzola

## 64 isolati

- Sequenziamento completo del genoma
- Clusterizzati con un genoma di riferimento
- SNP analisi
- Analisi gene by gene





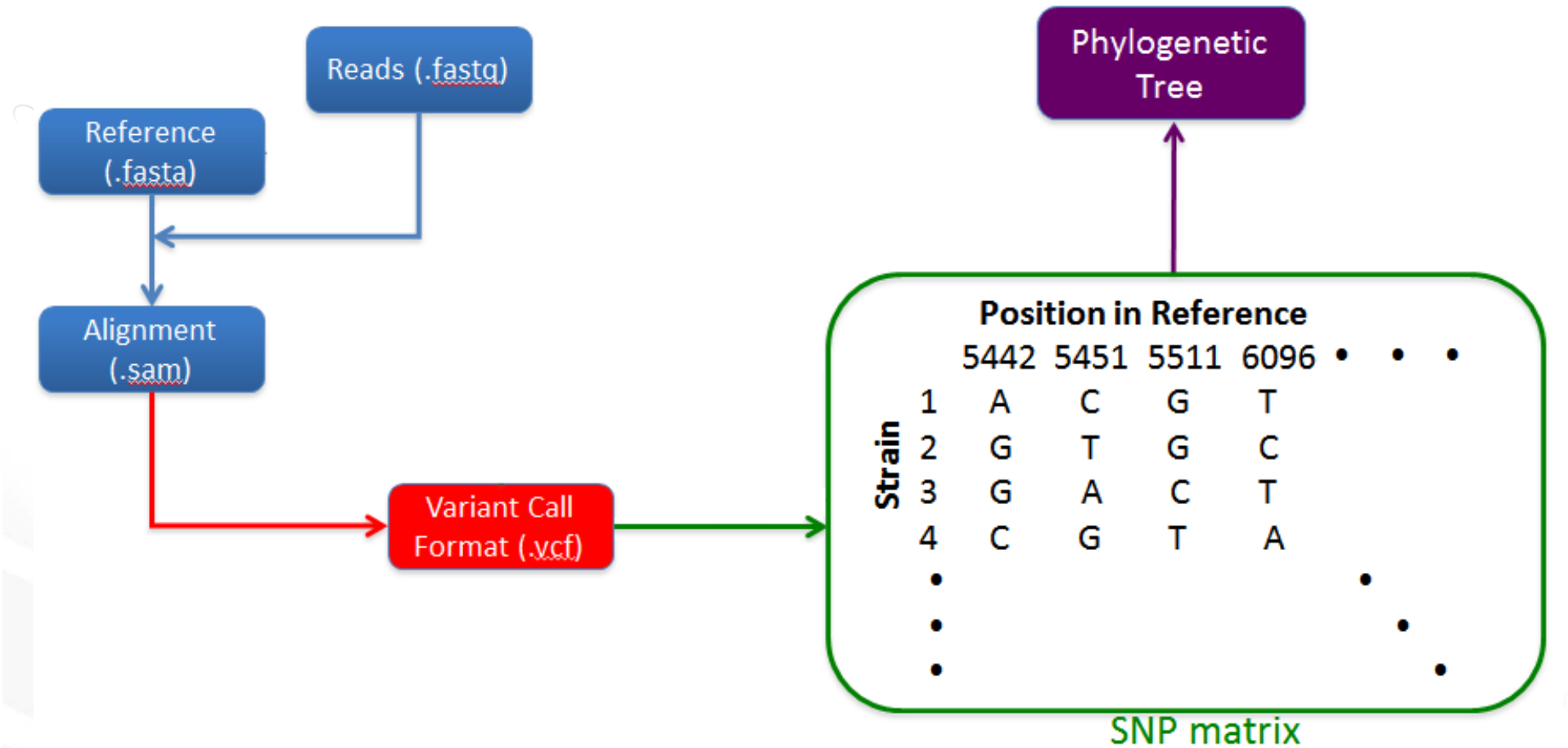
 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Procedura analisi di SNPs

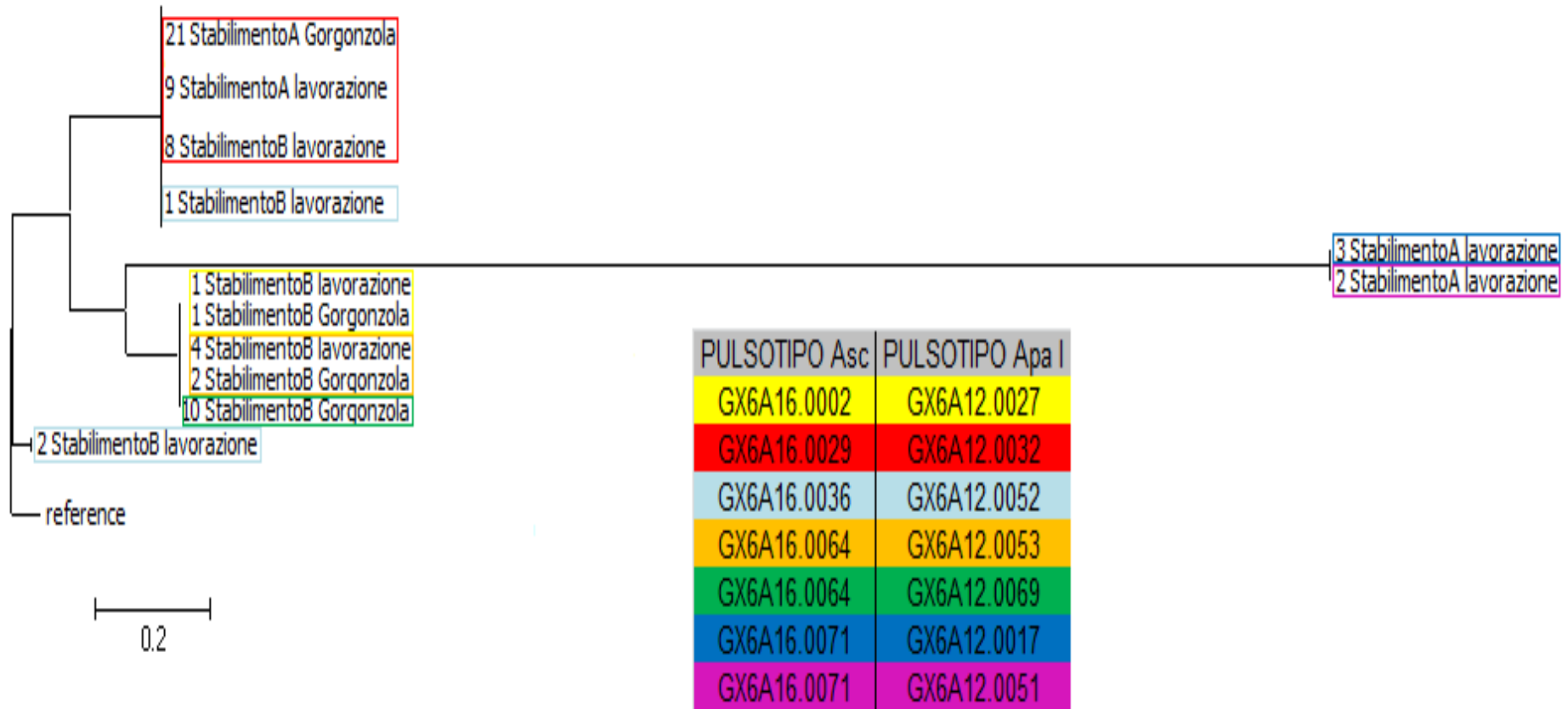
- Assemblaggio delle reads contro un genoma di referenza mediante Bowtie 2
- Ricerca di SNP ad alta qualità (in base al coverage e alla qualità del nucleotide)
- Creazione di una matrice di SNP che verrà utilizzata per costruire un nuovo allineamento in cui sono presenti solamente le posizioni che variano rispetto al reference
- Costruzione di un albero di massima verosimiglianza (Maximum Likelihood) a partire dall'allineamento utilizzando MEGA6



# SNPs Pipeline



# Analisi filogenetica





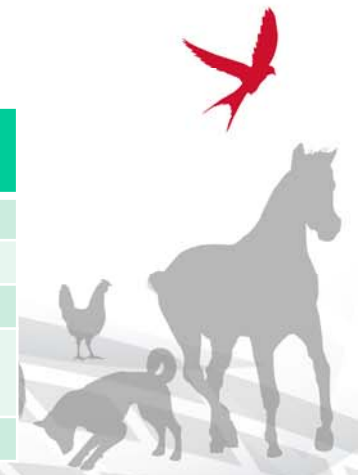
 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Prospettive future

## Studio sulla presenza di ceppi persistenti

- Campionamenti nell'azienda a conduzione familiare
- Due sessioni di indagine ambientale e suoi prodotti in tutte le aree di produzione dello stabilimento (a distanza di pochi giorni) .
- Prelievi da effettuare tra un processo di sanificazione completo e quello successivo.
- Le due sessioni prevedono campionamenti negli stessi identici punti

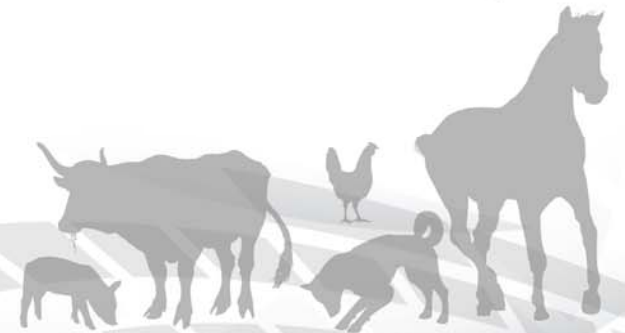
Area lavorazione	Campioni su superfici a contatto	Campioni su superfici non a contatto	Campioni su prodotto
Caseificazione	10	10	
Salatura	10	10	
Stagionatura	10	10	30*
Taglio / confezionamento	10	10	
<b>TOTALE</b>	<b>40</b>	<b>40</b>	<b>30</b>





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Flusso di lavoro NGS





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Principali punti critici nell'analisi dati NGS



- Abbondanza e qualità delle reads
- Obiettivi dell'analisi
- Confronto dei risultati con quelli ottenuti mediante altri metodi
- Scelta degli strumenti bioinformatici







 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Quali metodi di analisi?



## Assemblaggio

- Controllo qualità (trimming)
- De novo
- Mapping
- Ibrido

## Clustering

- Con assemblaggio
- Senza assemblaggio
- Altro (Kmer, tag...)





# Pipeline di analisi dati NGS

PGM reads

AlienTrimmer

Trimmed reads

MIRA 4.0rc4

Assembly

Fill gaps, correct errors, etc.

Final sequence

RNAmmer

rRNA

Glimmer

CDS

tRNAscan-SE

tRNA

Blast → UniProt, COGs...  
HMMER → TIGRFAM, PFAM...

Annotation





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Annotazione genomica



## Identificazione di geni e altri elementi funzionali

- **Metodi di annotazione**
  - Completamente automatica
  - Manuale
  - Validazione
- **Richiede il supporto da parte di «esperti» del microrganismo**





 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Hardware

## • La piattaforma NGS e la sua produttività influiscono su:

- Velocità di trasferimento dei dati
- Capacità dell' hard disk
- RAM
- CPU
- Software e Sistema operativo

## • Hardware o Cloud?

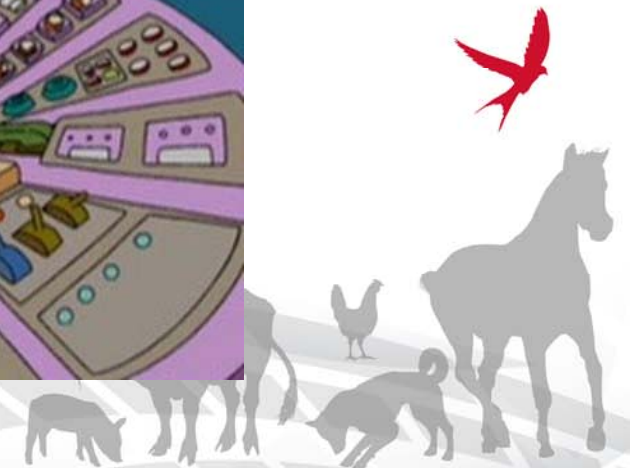




 *Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento

# Il bioinformatico

## La soluzione dei problemi ?



# Ringraziamenti



IZSAM G. CAPORALE  
TERAMO



*Listeria monocytogenes*  
Laboratorio Nazionale di Riferimento



IZSAM G. CAPORALE  
TERAMO



COVEPI  
Centro di Riferenza Nazionale

**CED**



**Gruppo Genomica**  
**Reparto Ricerca e Sviluppo Biotecnologie**

- Massimo Ancora
- Maurilia Marcacci
- Massimiliano Orsini
- Iolanda Mangone
- Alfreda Tonelli
- Cesare Cammà

