



IZSAM G. CAPORALE  
TERAMO



Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica

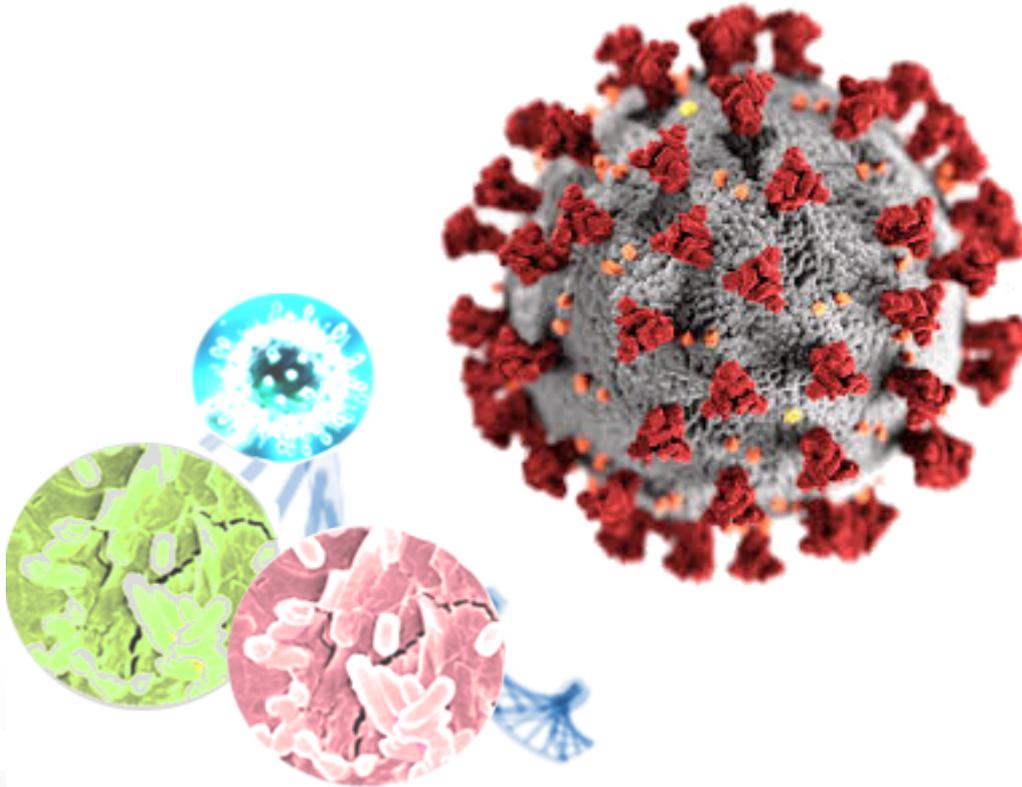
**CMDBuild for the management of  
genomic sequences of microorganisms  
in national systems and European projects**

Adriano Di Pasquale



**CMDBuild DAY, 28 October 2020**

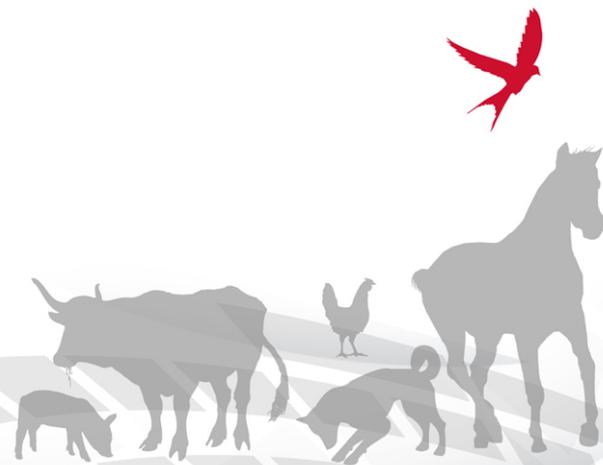
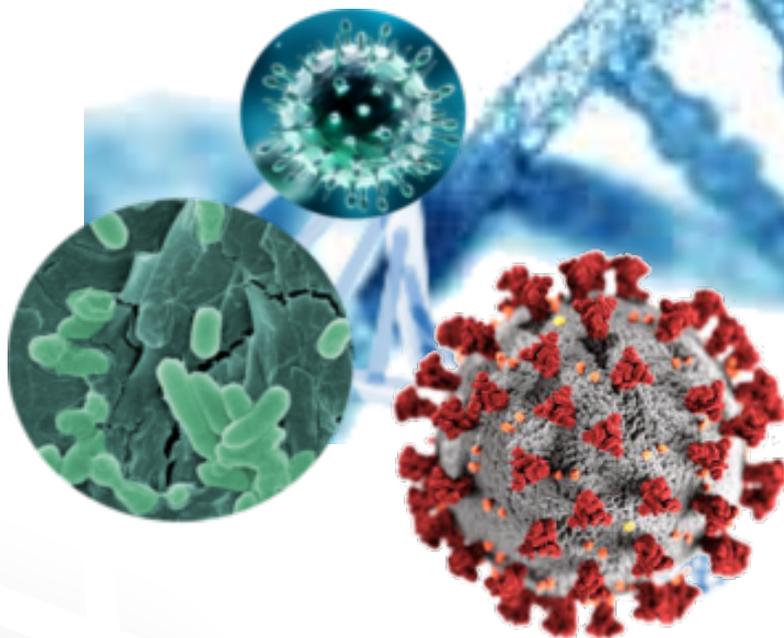
# Microrganisms



# Pathogens

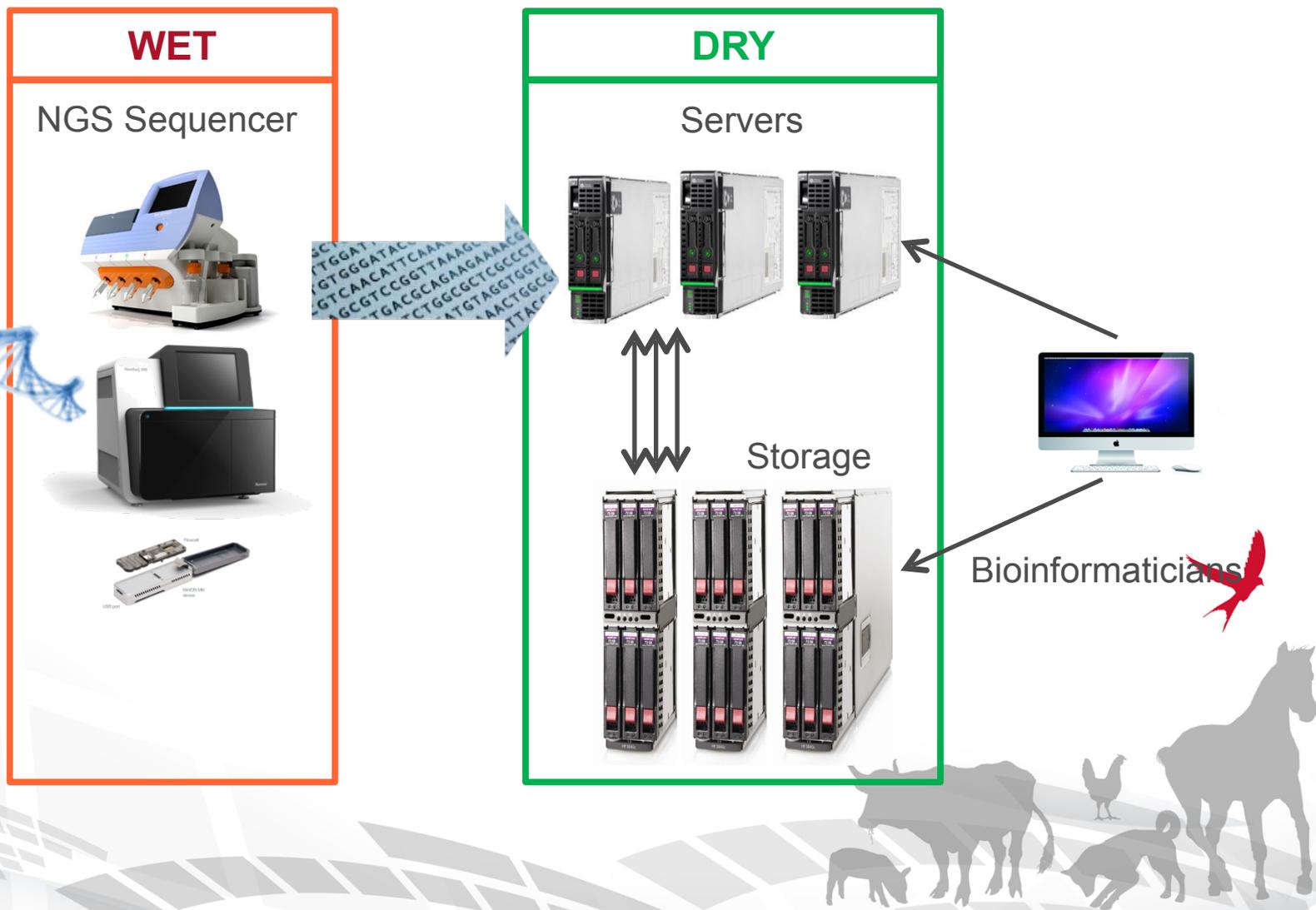


# DNA



# Next Generation Sequencing

Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica

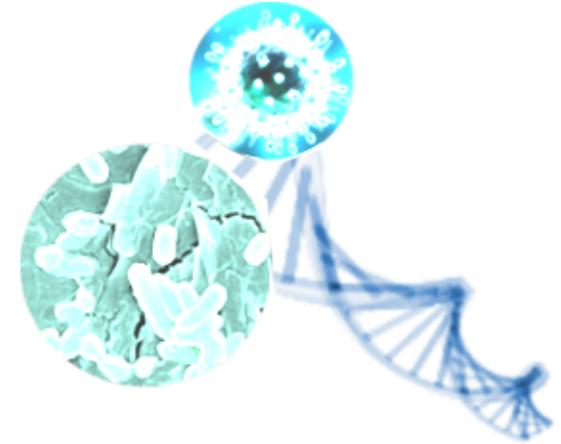


# WGS data



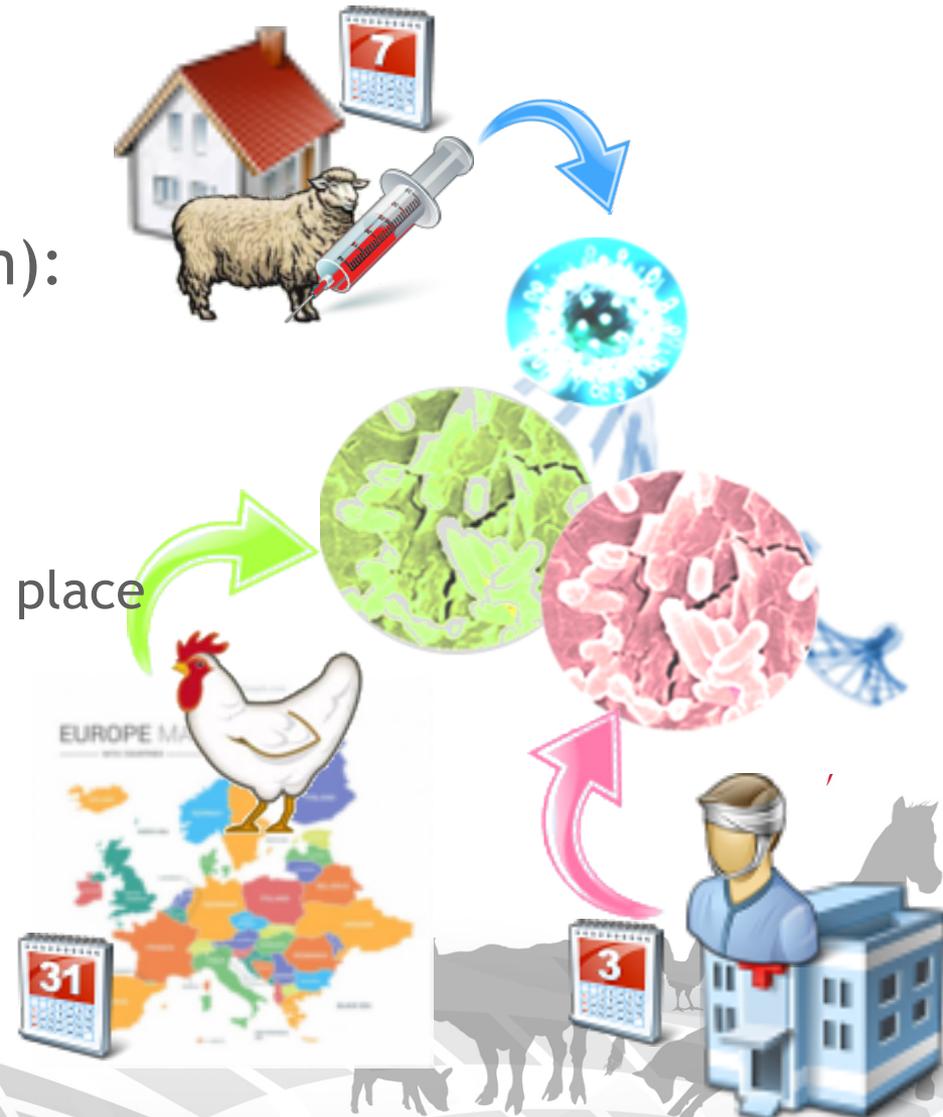
- Analysis results

- Raw data
- De novo assembly
- Typing (e.g., MLST, cgMLST)



# Metadata: epi-data

- What: pathogen species
- Matrix (source of isolation):
  - cheese, blood, water
- When, Where:
  - sample collection date and place



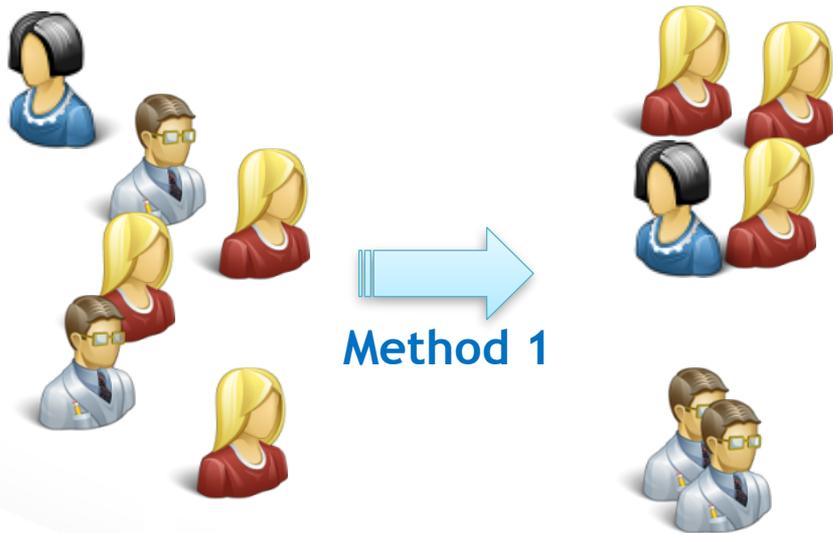
# Use case 1: Data collection

 Centro di Referenza Nazionale  
per **Sequenze Genomiche di**  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica



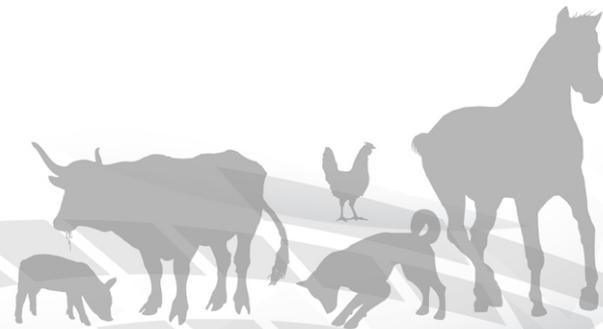
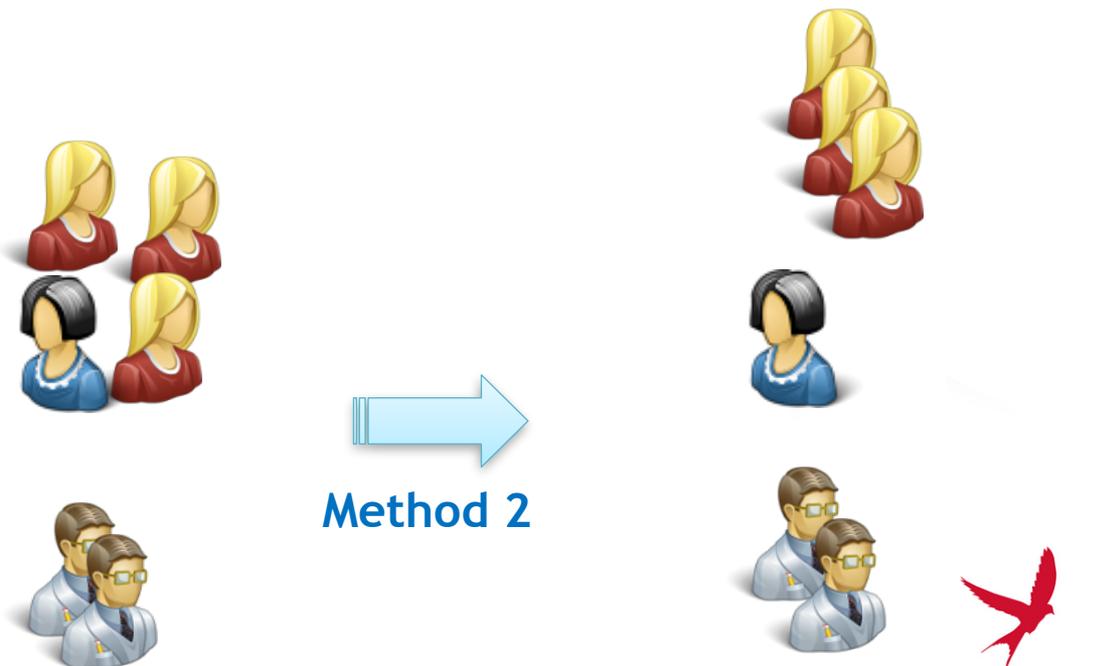
# Use case 2: Molecular Typing

 Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# Use case 2: Molecular Typing

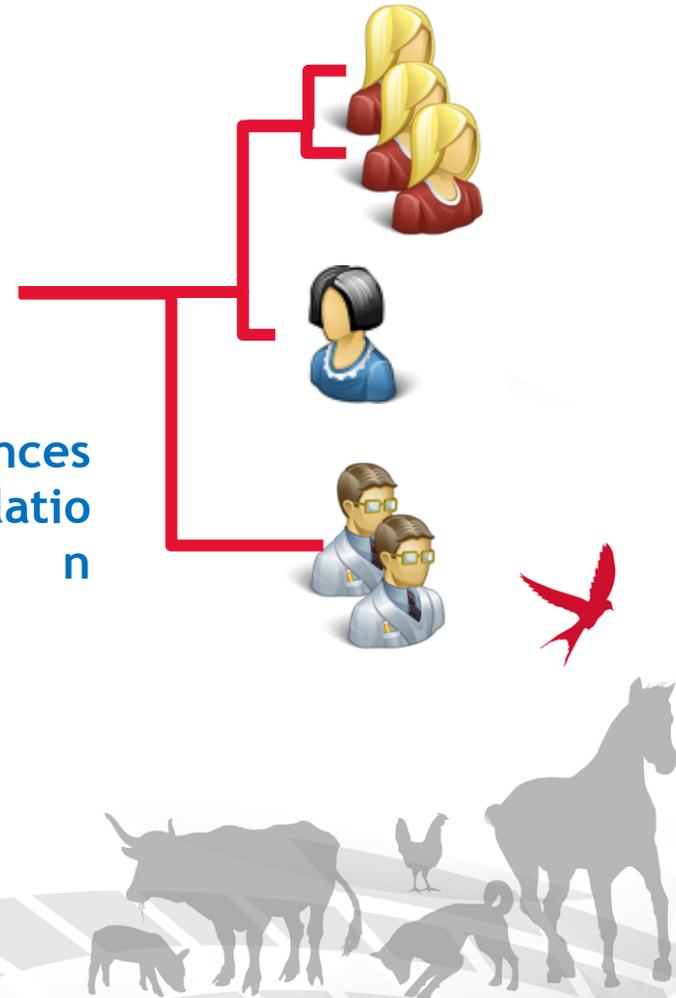
 Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica



# Use case 2: Molecular Typing

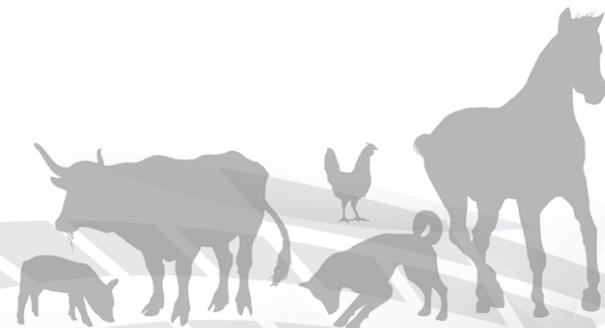
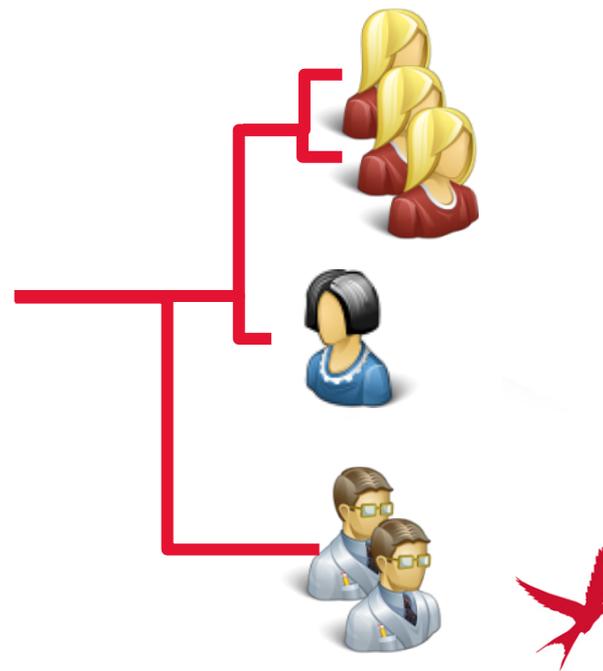
 Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica

Distances  
Calculatio  
n



# Use case 3: Clustering.

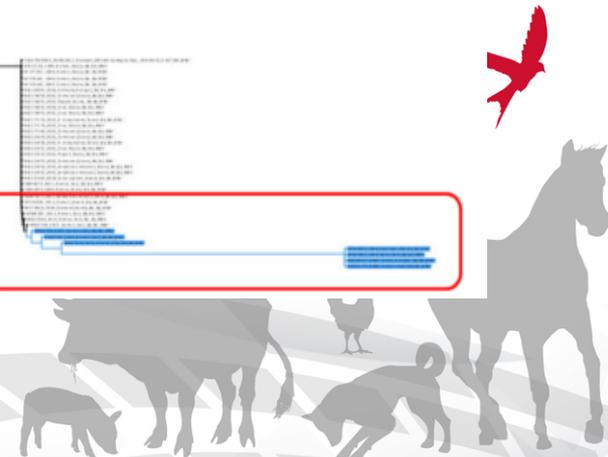
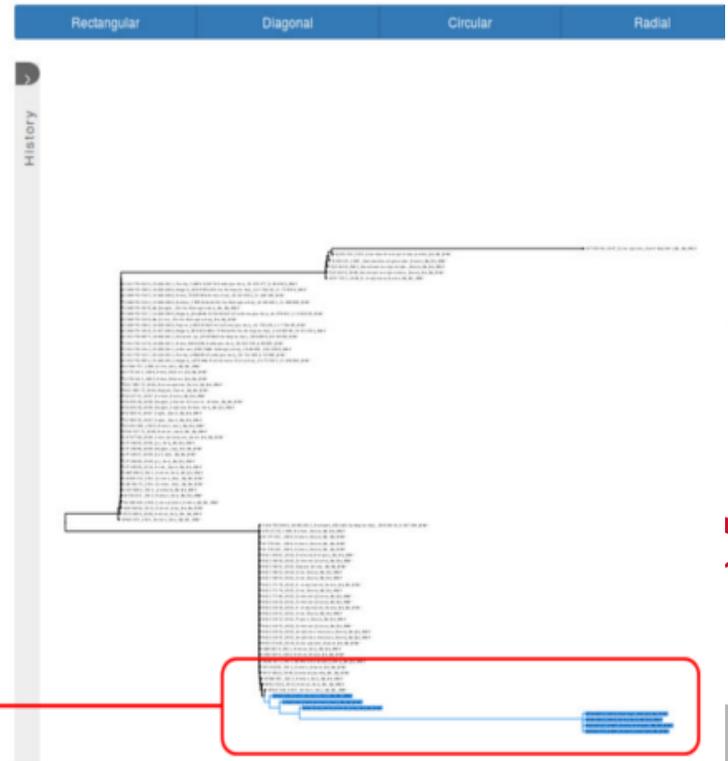
## Combining WGS data and metadata



# Use case 3: Clustering.

## Combining WGS data and metadata

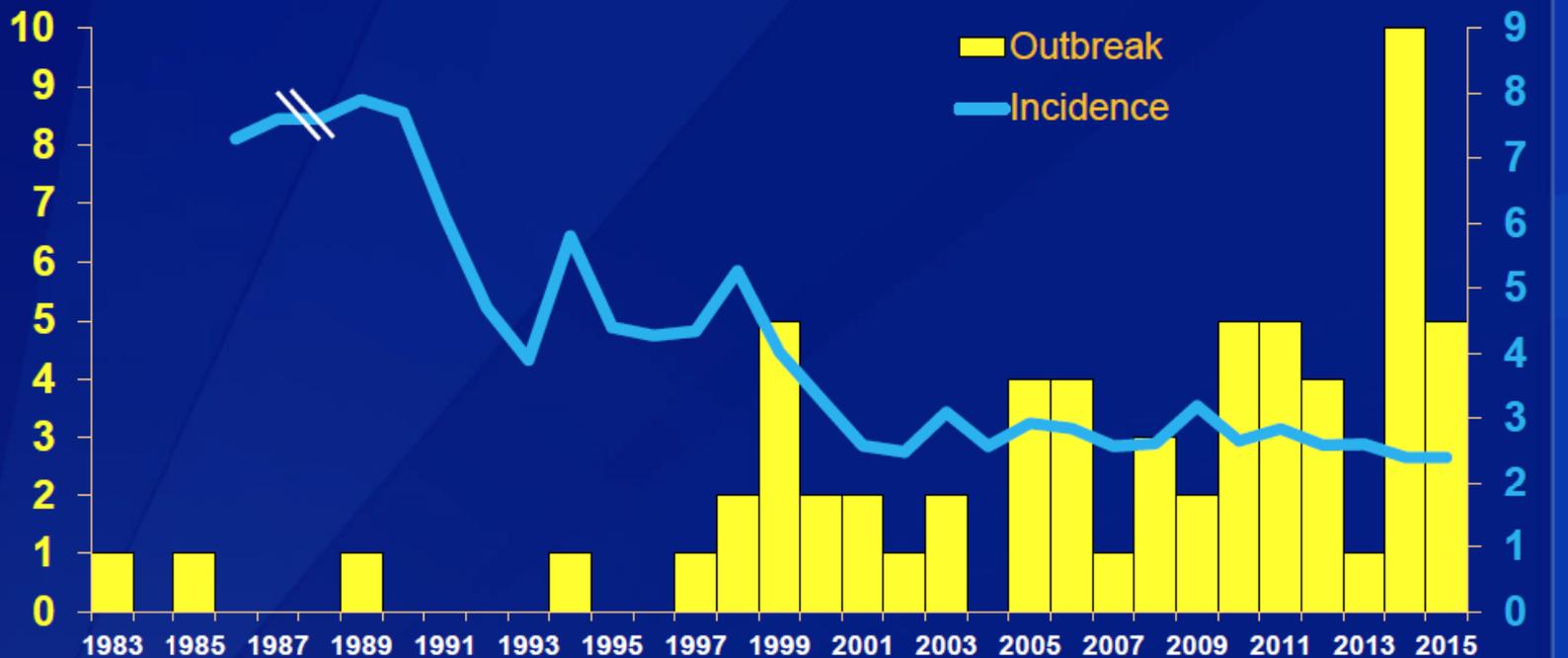
Phylogenetic Map of West Nile Fever



## Listeria Outbreaks and Incidence, 1983-2015

No. outbreaks

Incidence  
(per million pop)



**Era**  
Outbreaks per year  
Median cases per  
outbreak

**Pre-PulseNet**

0.3  
69

**Early  
PulseNet**

2.3  
11

**Listeria  
Initiative**

2.6  
5.5

**WGS**

7.5  
3.5

# GENPAT (GENomics of PATHogens)

## National Reference Centre for Whole Genome Sequencing of microbial pathogens: database and bioinformatic analysis

23-8-2017

GAZZETTA UFFICIALE DELLA REPUBBLICA ITALIANA

Serie generale - n. 196

### MINISTERO DELLA SALUTE

DECRETO 30 maggio 2017.

**Istituzione di un Centro di referenza nazionale nel settore veterinario.**

#### IL MINISTRO DELLA SALUTE

Vista la legge 23 giugno 1970, n. 503, recante «Ordinamento degli Istituti zooprofilattici sperimentali»;

Visto il decreto legislativo 30 giugno 1993, n. 270, recante «Riordinamento degli Istituti zooprofilattici sperimentali, a norma dell'art. 1, comma 1, lettera h) della legge 23 ottobre 1992, n. 421» e, in particolare, l'art. 2, comma 3, lettera l);

Visto il decreto del Ministro della sanità 16 febbraio 1994, n. 190, concernente «Regolamento recante norme per il riordino degli Istituti zooprofilattici sperimentali, in attuazione dell'art. 1, comma 5, del decreto legislativo 30 giugno 1993, n. 270», pubblicato nella *Gazzetta Ufficiale* 23 marzo 1994, n. 68;

Visto il decreto del Ministro della sanità 4 ottobre 1999, recante «Centri di referenza nazionali nel settore veterinario», pubblicato nella *Gazzetta Ufficiale* 23 dicembre 1999, n. 300;

Visto il decreto legislativo 30 marzo 2001, n. 165, e successive modificazioni, recante «Norme generali nell'ordinamento del lavoro alle dipendenze delle amministrazioni pubbliche»;

Visto il decreto legislativo 28 giugno 2012, n. 106, recante «Riorganizzazione degli enti vigilati dal Ministero della salute, a norma dell'art. 2 della legge 4 novembre 2010, n. 183»;

Vista l'istanza e la relativa documentazione presentata ai sensi del citato decreto ministeriale 4 ottobre 1999 dall'Istituto zooprofilattico sperimentale dell'Abruzzo e del Molise,

Ritenuto, pertanto, di istituire il Centro di referenza nazionale per le Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di bioinformatica;

Decreta:

#### Art. 1.

1. Presso l'Istituto zooprofilattico sperimentale dell'Abruzzo e del Molise è istituito il Centro di referenza nazionale - Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di bioinformatica.

#### Art. 2.

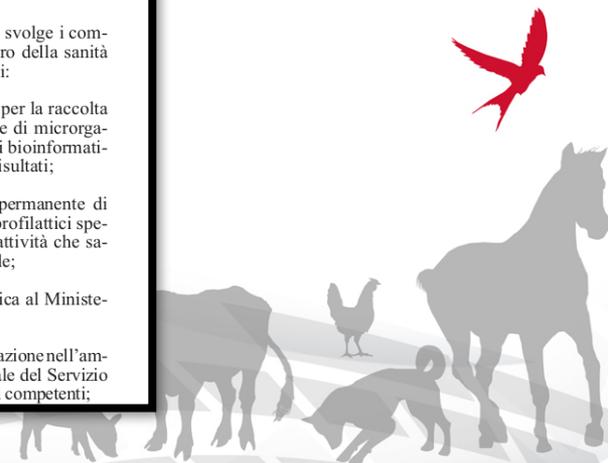
1. Il Centro di referenza, di cui all'art. 1 svolge i compiti di cui all'art. 2 del decreto del Ministro della sanità 4 ottobre 1999, nonché le seguenti funzioni:

a) realizza una piattaforma nazionale per la raccolta e conservazione delle sequenze genomiche di microrganismi patogeni e per l'esecuzione di analisi bioinformatiche, l'archiviazione e la condivisione dei risultati;

b) realizza un sistema strutturato e permanente di referenti all'interno dei singoli Istituti zooprofilattici sperimentali ai fini del coordinamento delle attività che saranno poste in essere sul territorio nazionale;

c) fornisce assistenza tecnico-scientifica al Ministero della salute ed alle autorità competenti;

d) cura l'organizzazione di corsi di formazione nell'ambito delle proprie competenze per il personale del Servizio sanitario nazionale e di altri operatori di Enti competenti;





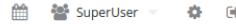


 Centro di Referenza Nazionale  
per **Sequenze Genomiche di  
microrganismi patogeni:**  
banca dati e analisi di Bioinformatica

# GENPAT Home page



BioinfoDB - SuperUser.



## Navigation

- Codici
- Metadati
- Motivi/Richiedenti
- Esami**
- Checks
- Altro
- Reports
- Pipeline
- Download
- Upload
- </> Centro di Referenza Nazionale...
- </> Genpat-Typing
- </> Filtro Campioni per Data Ulti...
- Scheduler
- All items

Centro di Referenza Nazionale per Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di bioinformatica

Il Ministero della Salute, con Decreto del 30 maggio 2017 (G.U.R.I. n. 196 del 23 agosto 2017), ha attivato il **Centro di Referenza Nazionale per Sequenze Genomiche di microrganismi patogeni: banca dati e analisi di bioinformatica** presso la sede centrale dell'IZSAM.

Il Centro svolge le seguenti funzioni:

- realizza una piattaforma nazionale per la raccolta e conservazione delle sequenze genomiche di microrganismi patogeni e per l'esecuzione di analisi bioinformatiche, l'archiviazione e la condivisione dei risultati;
- realizza un sistema strutturato e permanente di referenti all'interno dei singoli Istituti zooprofilattici sperimentali ai fini del coordinamento delle attività che saranno poste in essere sul territorio nazionale;
- fornisce assistenza tecnico-scientifica al Ministero della salute ed alle autorità competenti;
- cura l'organizzazione di corsi di formazione nell'ambito delle proprie competenze per il personale del Servizio sanitario nazionale e di altri operatori di Enti competenti;
- promuove le attività di ricerca nel settore di competenza;
- mette in atto ogni altra utile attività attinente alle proprie competenze ivi comprese la collaborazione e il coordinamento con altre amministrazioni ed Enti del settore.

**Responsabile:** Cesare Cammà

Per visualizzare la pagina WIKI del sistema clicca [qui](#).

Motivo attivo: **non hai nessun Motivo Attivo.**

 [Vai ai Genpat](#)

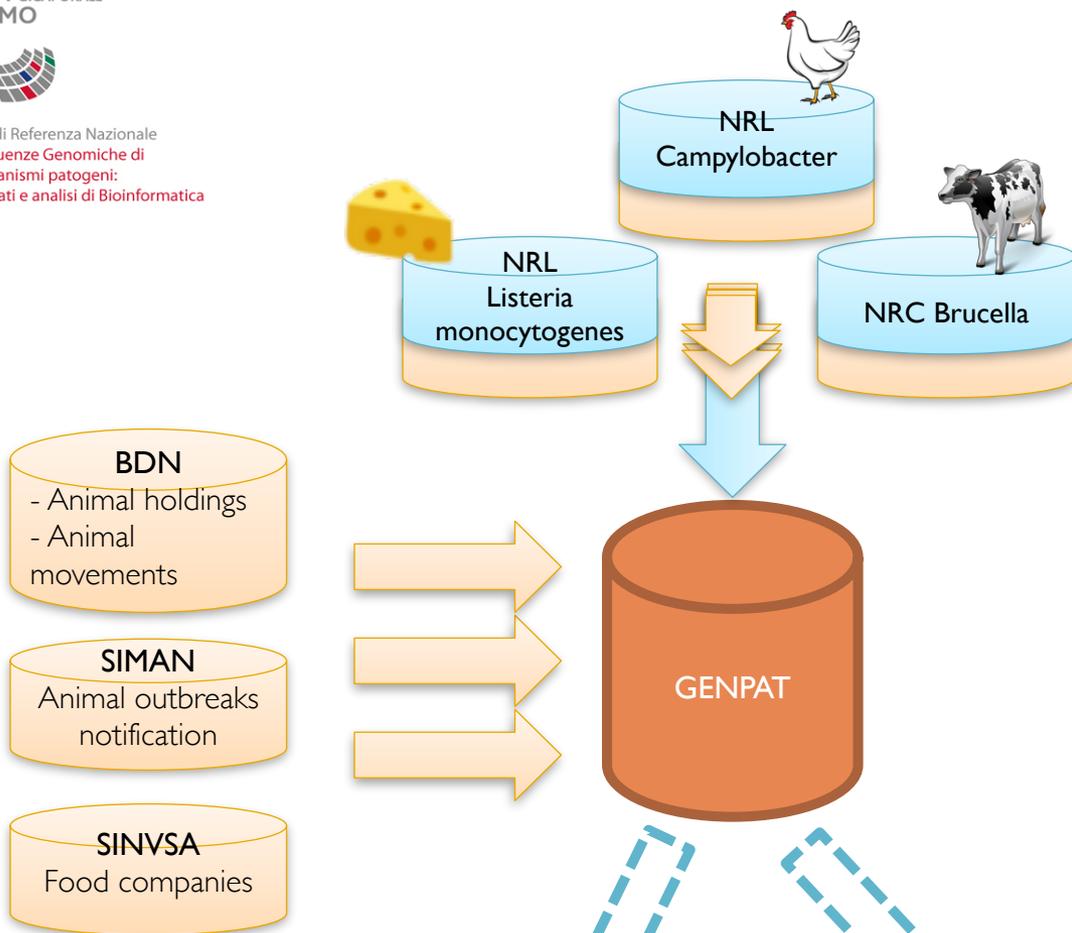
 [Vai ai Risultati](#)

 [Vai al Report Genpat-Esami](#)

 [Vai alla pagina Attiva Motivo](#)



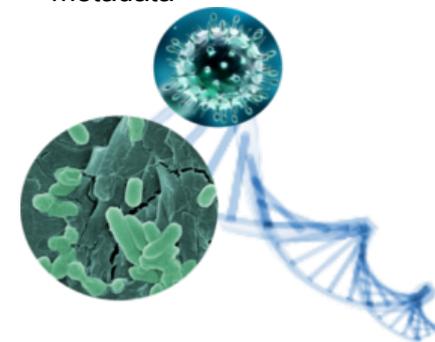
# GENPAT Architecture



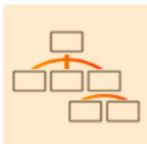
WGS



metadata



# CMDBuild features



Database modeling



Editor and workflow engine



Editor and report engine



Dashboard configuration



Webservices and connectors



Data Import / Export through files



Documents archive



GIS and BIM



Custom interfaces



Mobile APP



Relationship paradigm



History



Email management



Barcode e QR code



Self-Service portal



Visual Graph



Scheduler



User Profiling and Multitenant



Administration Module



Data Management Module

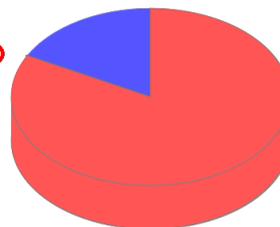


## Report Check

### Check Taxa

Risultati	Totale
Others	7684
Results	1601

Others (7.684) Results (1.601)

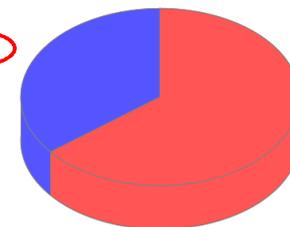


campioni con coverage verticale  
compreso tra 75 e 120  
rispetto la specie assegnata

### Check Assembly

Risultati	Totale
Others	5980
Results	3305

Others (5.980) Results (3.305)

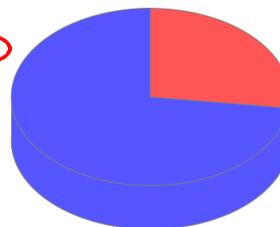


Campioni con al massimo 100  
contig ottenuti mediante  
assembly denovo

### Check Species

Risultati	Totale
Others	2506
Results	6779

Others (2.506) Results (6.779)

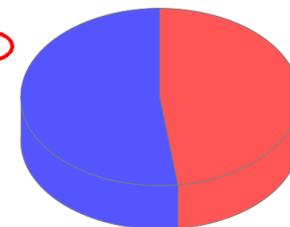


campioni con con copertura  
del genoma di riferimento  
>80% (coverage orizzontale)

### Check Reads

Risultati	Totale
Others	4501
Results	4888

Others (4.501) Results (4.888)

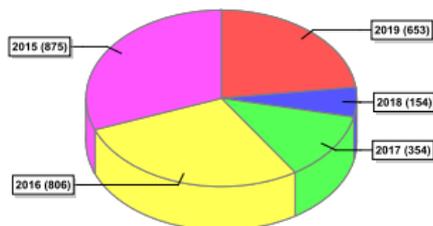


campioni con lunghezza  
reads tra 125 e 150 e con un  
Q30 >85 dopo il trimming

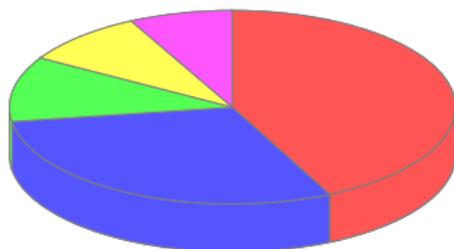
## Report Statistiche Specie

Specie: *Listeria monocytogenes*

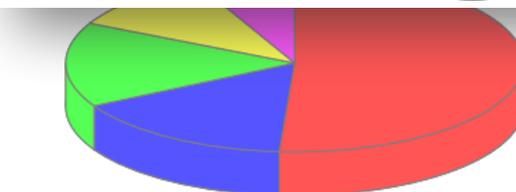
### Campioni Suddivisi per Anno e per Specie



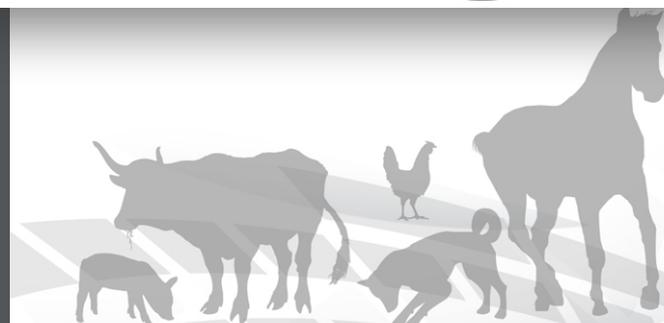
### Campioni Suddivisi per Host e per Specie



SUINO (566) UOMO (394) BOVINO (144) SALMONE (119) (99)



TAMPONE AMBIENTALE (SPONGE BAGS) (838) SALAME (261)  
SANGUE (253) PESCE AFFUMICATO (173) SALSICCIA (116)





Centro di Riferenza Nazionale  
per Sequenze Genomiche di  
microorganismi patogeni:

# GENPAT Features

Cards Campione Genpat

+ Add card Campione Genpat Search... [Icons]

Codice	Specie	Host	Matrice	Data Prel...	Punto Prelievo	Sequence Type	Richiedente
cc8 mag 2018							

Attributes Relations

Choose an attribute

Specie

Operator Equals

Data Prelievo

Operator Greater than

Cards Campione Genpat

+ Add card Campione Genpat Search... [Icons] CC8 2018 [Icons]

Codice	Specie	Host	Matrice	Data Prel...	Punto Prelievo	Sequence Type	Richiedente
[x]	[x]	[x]	[x]	[x]	[x]	[x]	[x]

Lancia Pipeline

Motivo

Tipologia template \*  
Grape Tree

Metodo Analisi Grape Tree

Inizio Porzione Vcf

CSV per Dati dei Reference

File CSV con solo Codici Campioni (senza colonna Risultati)

CSV campioni

DS Esami

Assembly QC

Fine Porzione Vcf

Scarica Template di esempio del File CSV per Dati dei Reference per il template di Informazioni Porzioni Vcf

### Informazioni File CSV per Dati dei Reference

Il Template di esempio del File CSV per Dati dei Reference contiene i campi necessari all'analisi a reperire i dati dei reference. Una spiegazione su come compilare questo file si trova nella pagina Wiki al seguente [link](#).

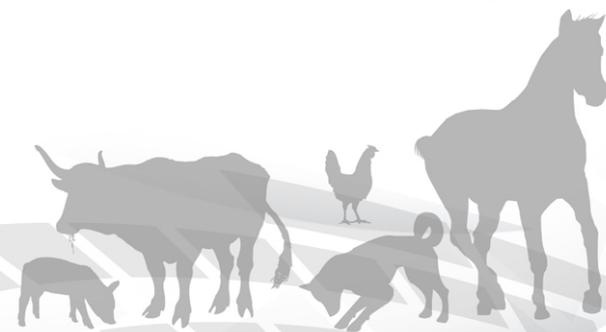
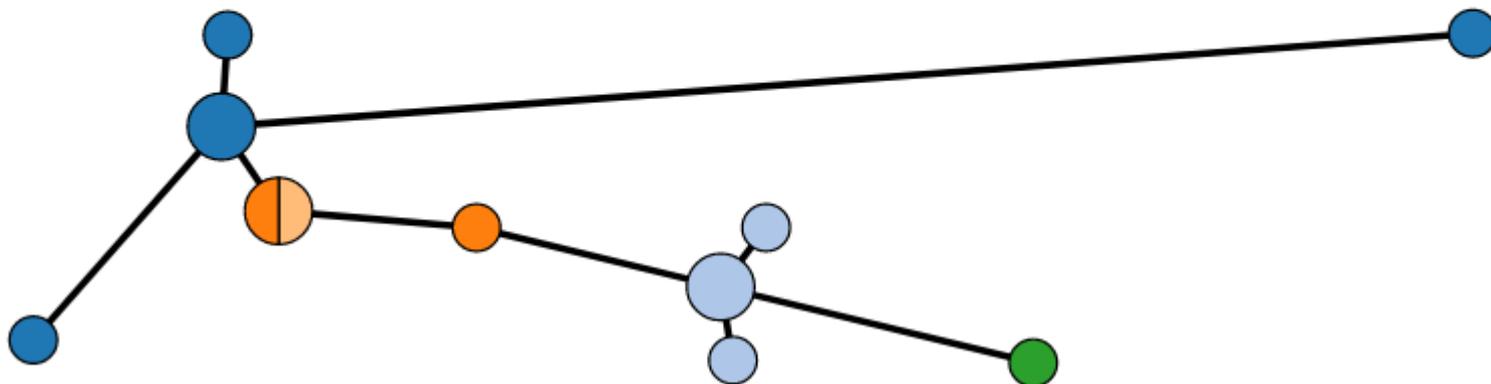
### Risultato Grape Tree

- [Albero MST con GrapeTree - Minimum Spanning Tree in formato newick](#)
- [Albero NJ con GenPAT GIS - Neighbor joining in formato newick](#)
- [Chiamata Allelica](#)
- [Metadati](#)
- [Missing Loci](#)
- [Log cgMLST](#)

# GENPAT Features

## matrice

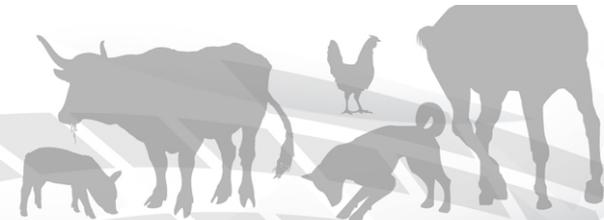
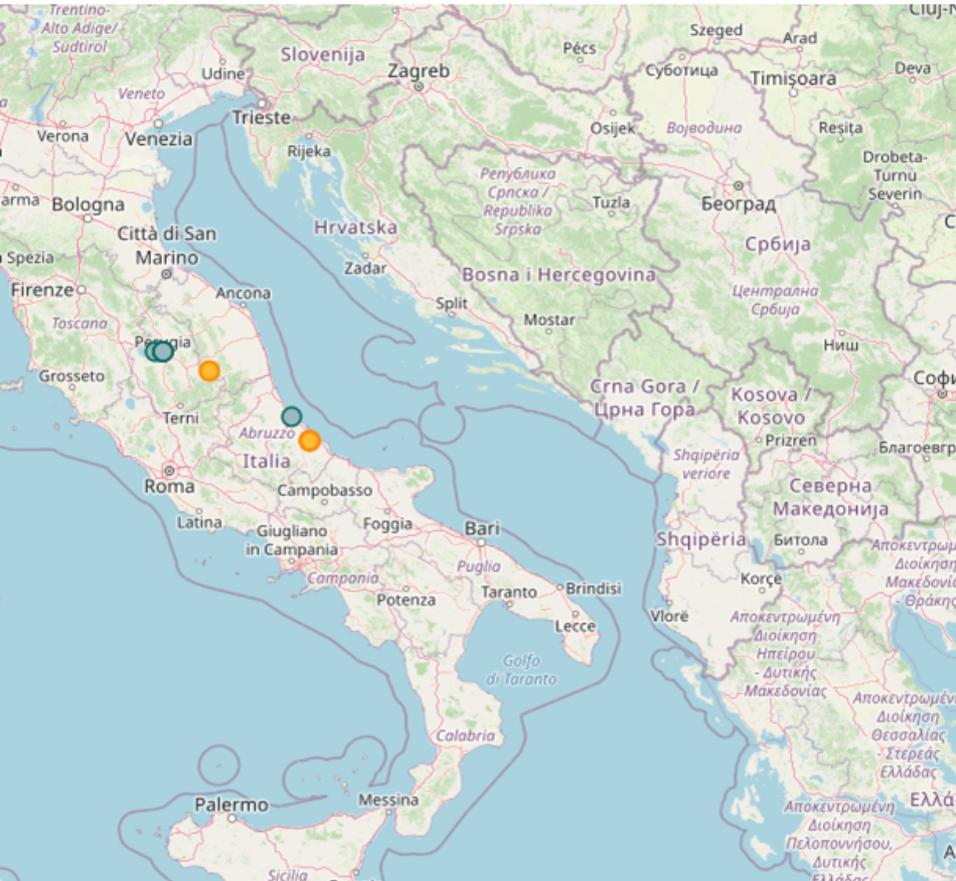
-  PROSCIUTTO CRUDO [5]
-  TAMPONE AMBIENTALE (SPC)
-  SANGUE [2]
-  LIQUOR [1]
-  SOVRACOSCIA [1]





 Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica

# GENPAT Features





Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica

# OHEJP COHESIVE

<https://onehealthhejp.eu>



About ▾ Consortium ▾ Projects ▾ Training ▾ Outcomes ▾ ASM ▾ News ▾ Contacts 

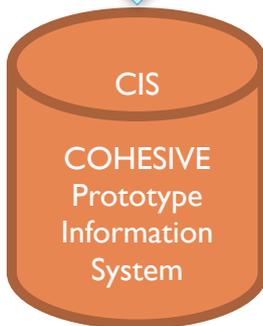
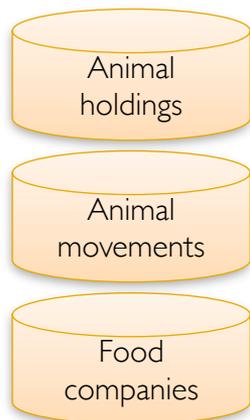
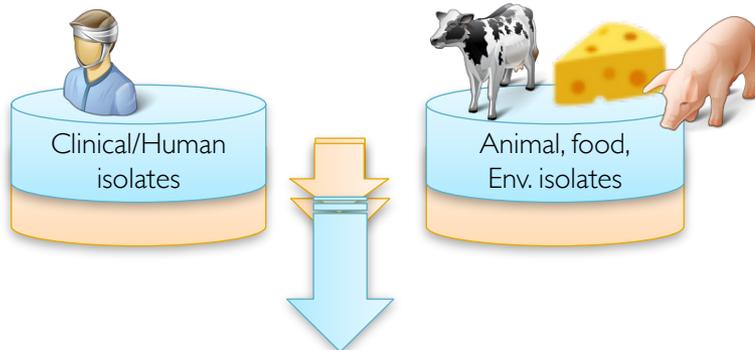
## One Health European Joint Programme

A landmark partnership between 37 partners, including acclaimed food, veterinary and medical laboratories and institutes across Europe and the Med-Vet-Net Association.

[LEARN MORE](#)



# OHEJP COHESIVE



WGS



metadata



Epidemiological system

OUTBREAK INVESTIGATION



SURVEILLANCE



## 4.1.2: COHESIVE prototype DEMO

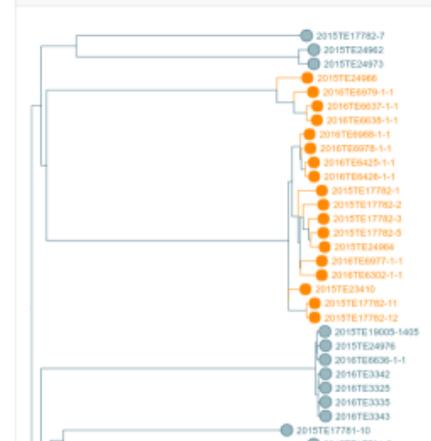
 Centro di Referenza Nazionale  
per Sequenze Genomiche di  
microrganismi patogeni:  
banca dati e analisi di Bioinformatica

### Main systems

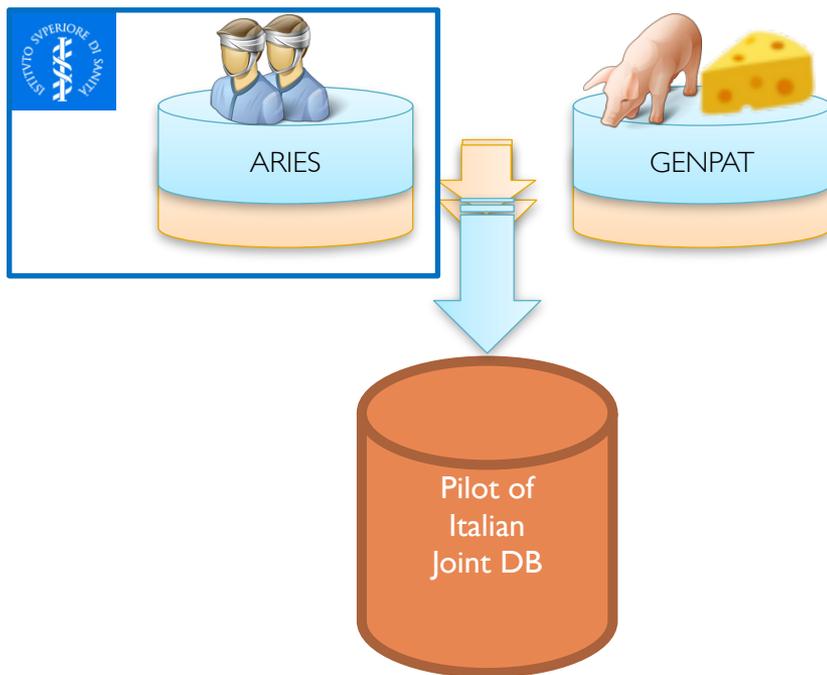
[https://cohesive.izs.it/cis\\_italy](https://cohesive.izs.it/cis_italy) [https://cohesive.izs.it/cis\\_norway](https://cohesive.izs.it/cis_norway) [https://cohesive.izs.it/cis\\_holland](https://cohesive.izs.it/cis_holland)

### GIS module

[http://mapserver.izs.it/gis\\_genpat/](http://mapserver.izs.it/gis_genpat/)



# Italian Ministry of Health Project ARIES/GENPAT integration



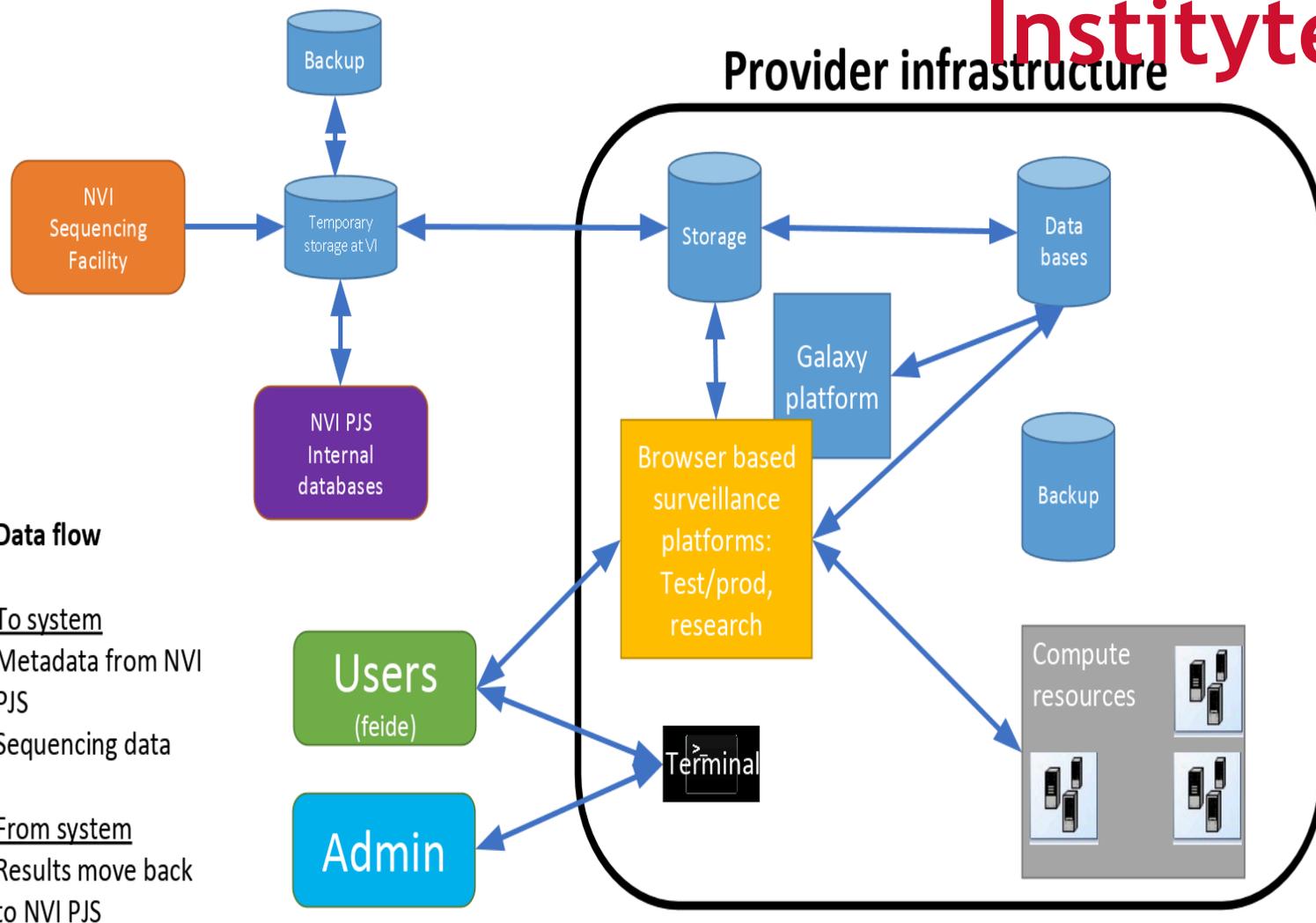
WGS



metadata



# CIS for Norway Veterinary Institute





 Centro di Referenza Nazionale  
per **Sequenze Genomiche di**  
microorganismi patogeni:  
banca dati e analisi di Bioinformatica

**Thank you!**

**<https://genpat.izs.it>**

**Adriano Di Pasquale**  
**[a.dipasquale@izs.it](mailto:a.dipasquale@izs.it)**

