



IZSAM G. CAPORALE
TERAMO




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

Centro di Referenza Nazionale Sequenze Genomiche di Microrganismi Patogeni: Banca Dati e Analisi di Bioinformatica

Cesare Cammà






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

GENPAT Decreto

Istituito dal
MINISTERO DELLA SALUTE
con
Decreto del 30 maggio 2017
(G.U.R.I. n. 196 del 23 agosto 2017)



Perchè una piattaforma nazionale




L'utilizzo del WGS in maniera decentrata (senza un repository nazionale) porterebbe a una **dispersione dei dati** e a una riduzione delle capacità di **sorveglianza nazionale**

- Fissando gli standard di qualità dei dati prodotti mediante NGS verrebbe garantita una **armonizzazione delle analisi bioinformatiche**
- Possibilità di **integrare** i dati WGS con informazioni presenti in altri database nazionali




Obiettivi

- 
- Realizzare una **piattaforma nazionale** per la raccolta e conservazione delle **sequenze genomiche** di microrganismi patogeni, per l'esecuzione di analisi bioinformatiche, l'archiviazione e la condivisione dei risultati;
 - Realizzare un **sistema strutturato e permanente di referenti** all'interno dei singoli Istituti zooprofilattici sperimentali (IIZZSS), Istituto Superiore della Sanità (ISS) ed altri Enti di ricerca su indicazione del Ministero della Salute ai **fini del coordinamento delle attività che saranno poste in essere sul territorio nazionale**;

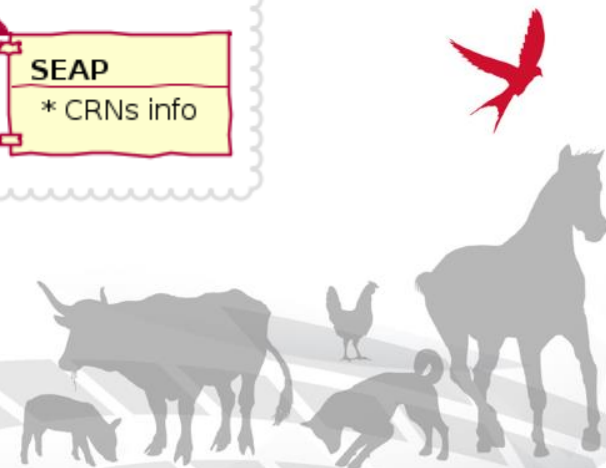
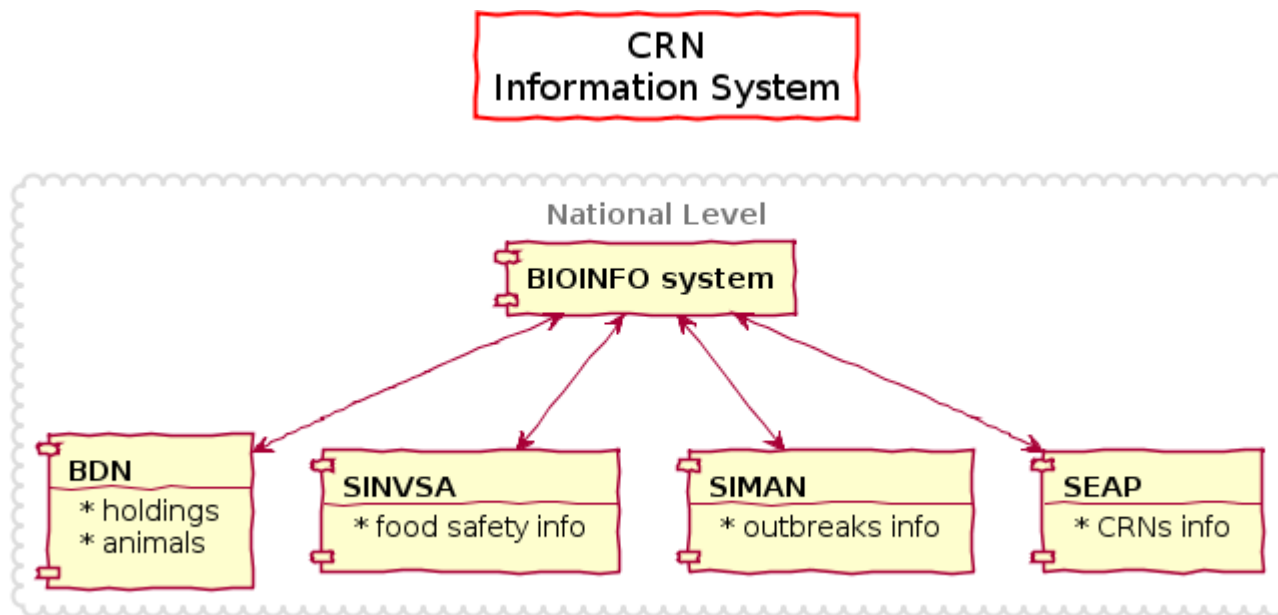


Obiettivi

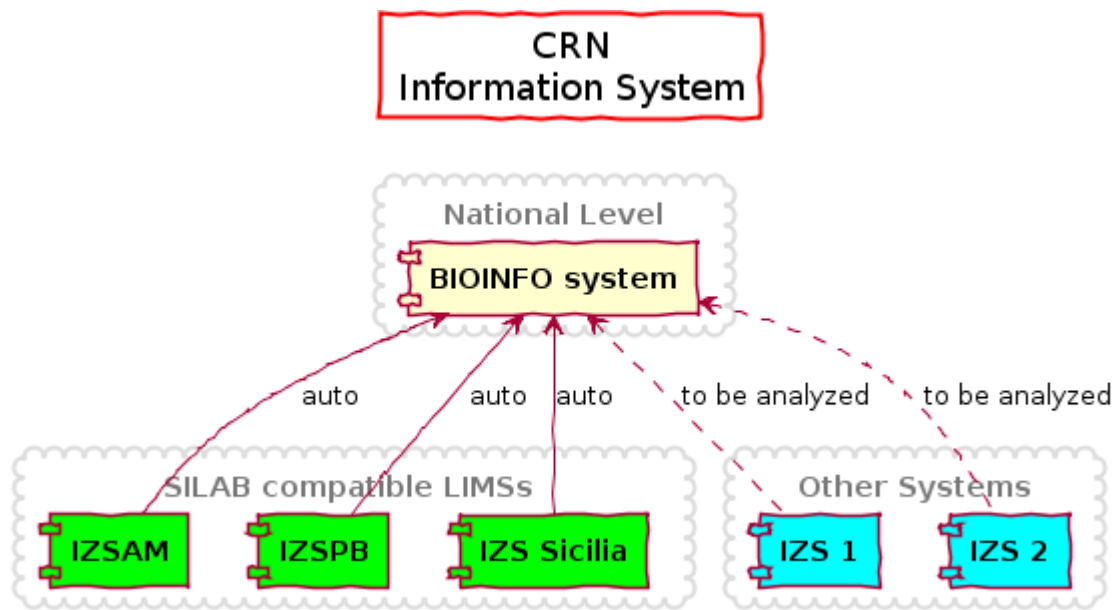
- 
- **Fornire assistenza tecnico-scientifica al Ministero della Salute ed alle autorità competenti;**
 - Promuovere le attività di ricerca nel settore di competenza;
 - Curare l'organizzazione di corsi di formazione nell'ambito delle proprie competenze per il personale del Servizio sanitario nazionale e di altri operatori di Enti competenti (in collaborazione con il Comitato tecnico per la Genomica S.I.Di.L.V.)
 - Mettere in atto ogni altra utile attività attinente alle proprie competenze ivi comprese la collaborazione e il coordinamento con altre amministrazioni ed Enti del settore.




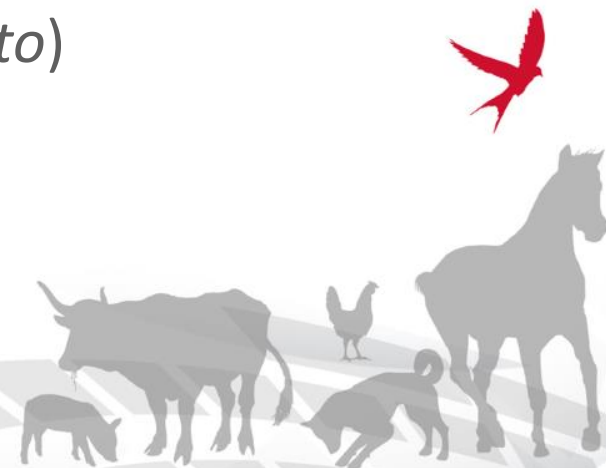
• Interoperabilità



Automatic import from LIMSs systems



- 
- Quali microrganismi sono prioritari
 - Decisione su come si alimenta il sistema
 - Ad ogni sequenziamento? (*sconsigliabile*)
 - Durante piani di sorveglianza? (*possibile*)
 - Durante focolai o epidemie? (*consigliato*)





- Uniformare i protocolli di analisi

- I protocolli di sequenziamento
- I Workflow di analisi bioinformatica
- Criteri di inclusione

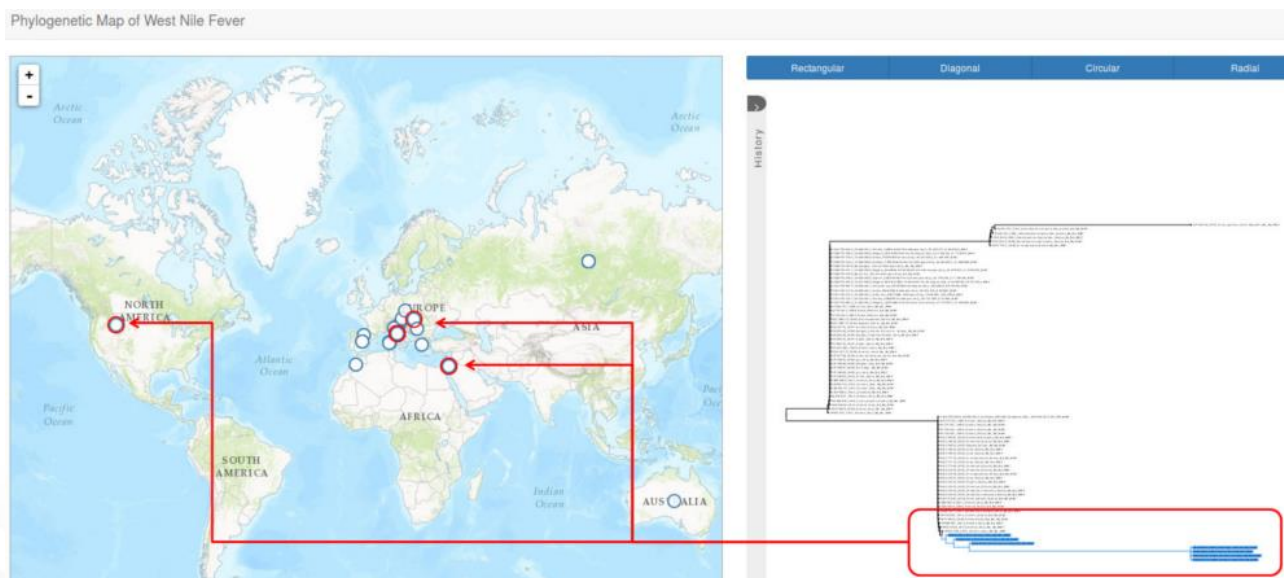
- Definire lista dei termini

- Nomenclature (es, wg-MLST)
- Codifiche di meta-dati (es, geo-localizzazione, specie, materiale e matrice)
- Analoghi ai “catalogs” di EFSA

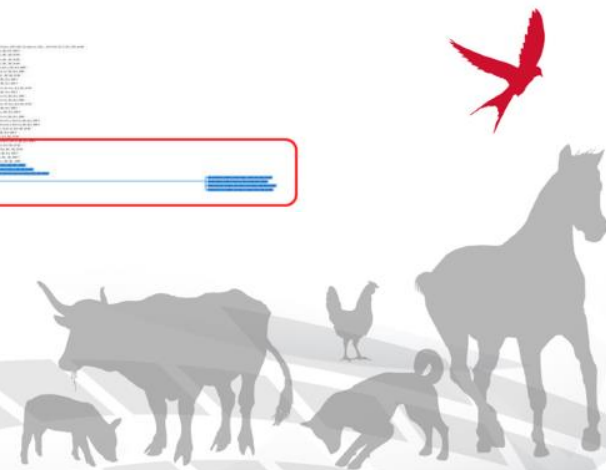


• Funzionalità e reportistica

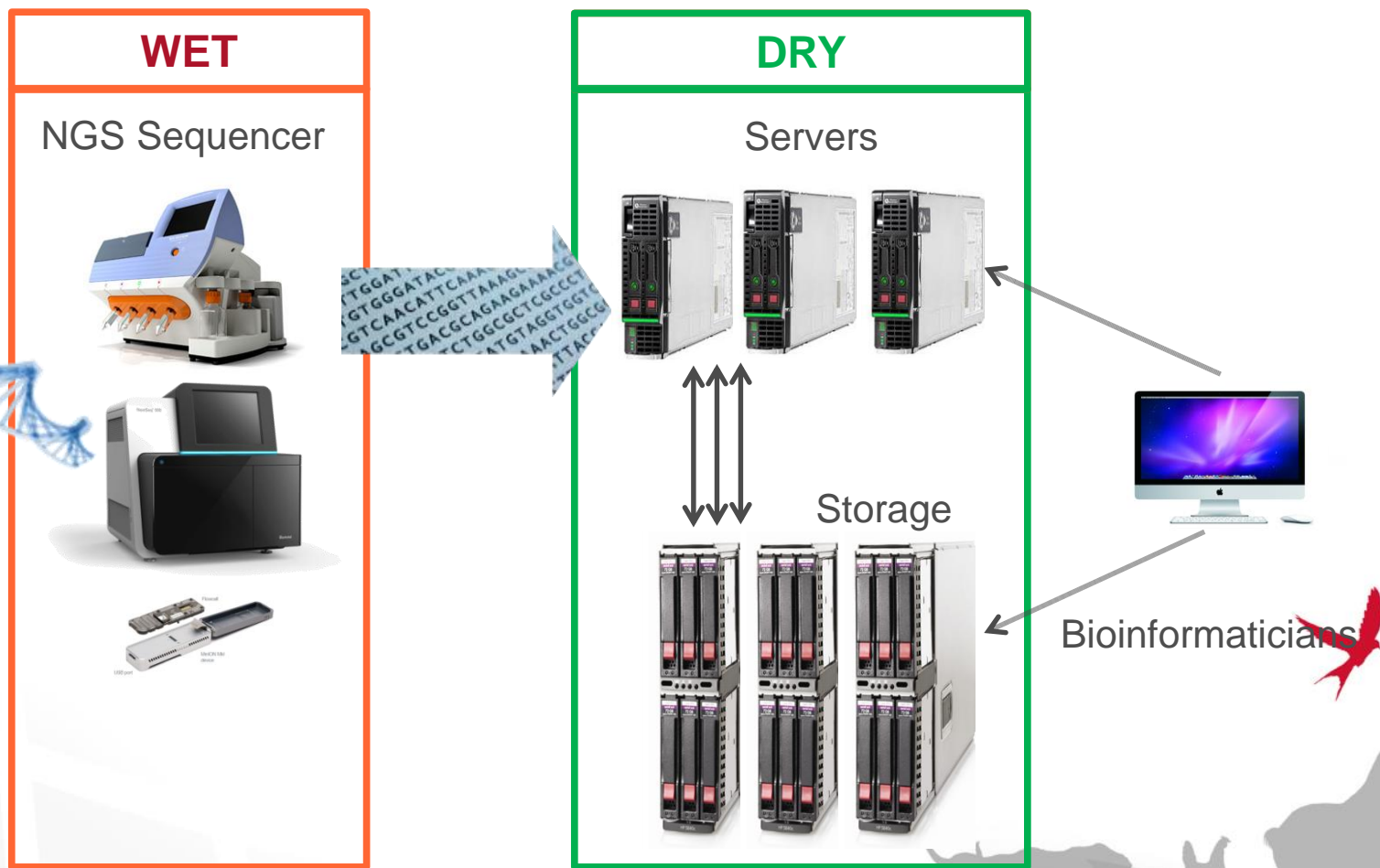
- Esempio: analisi filogenetiche integrate con dati geografici e/o delle movimentazioni degli animali



http://mapserver.izs.it/gis_phylown/



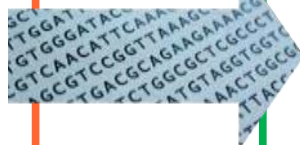
Logical Architecture



Logical Architecture

WET

NGS Sequencer



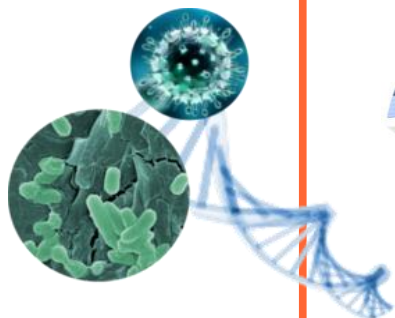
DRY



> 400 core,
> 2TB RAM



> 120TB





Samples
data

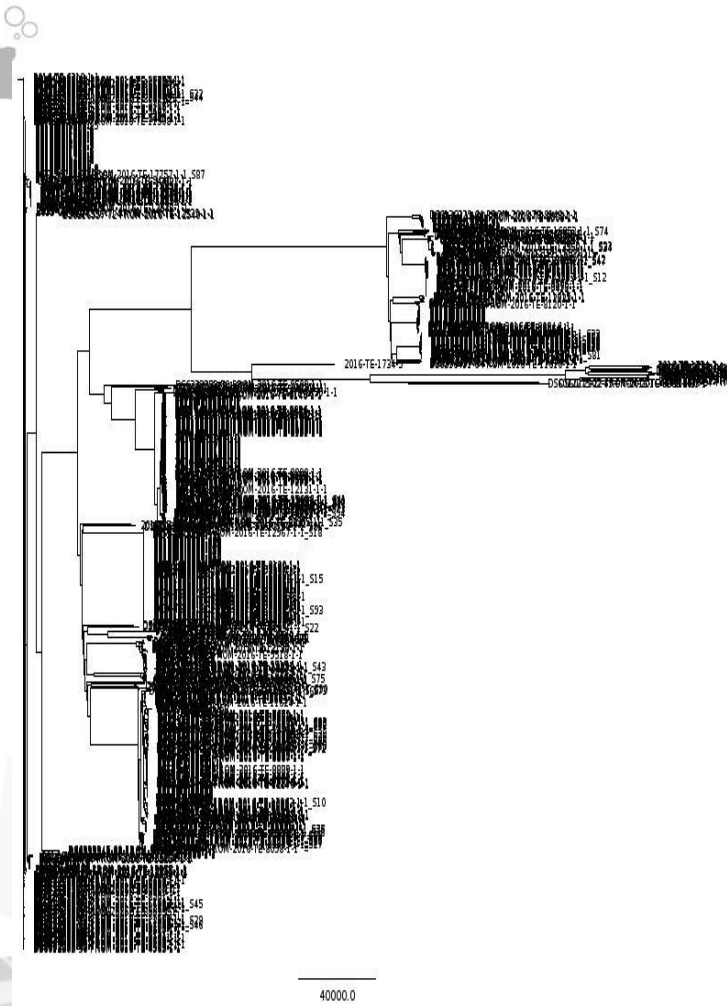


results



Epidemia listeriosi Marche 2015/2016

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



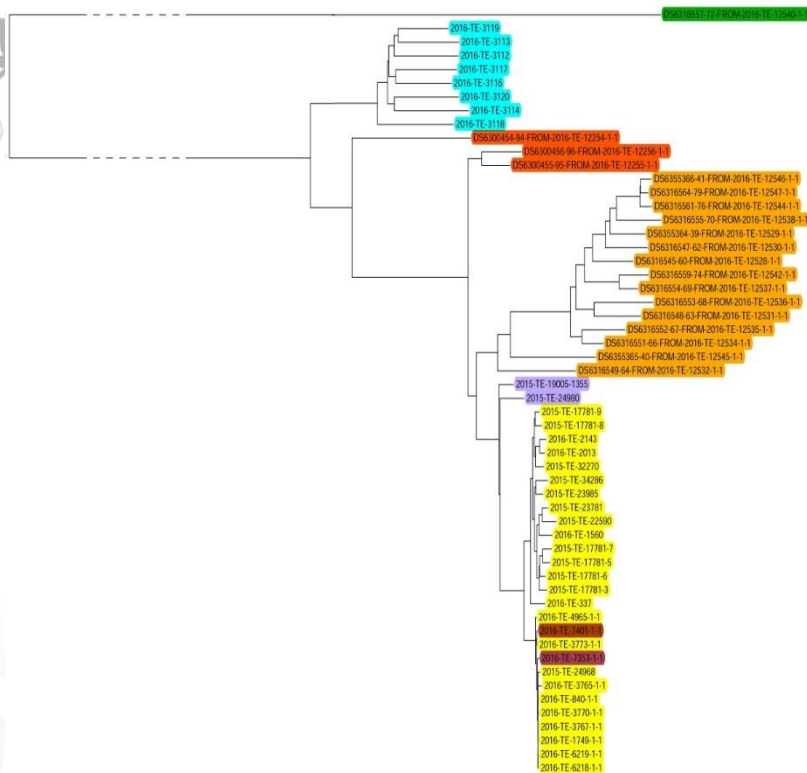
>800 ceppi di *L. monocytogenes* isolati da
alimenti e ambiente
>30 ceppi clinici

Pipeline di Assemblaggio Ibrido
>99% coverage orizzontale (ref: EGD-e)
Albero NJ basato su matrice di SNPs

Inclusione nel cluster epidemico < 25 SNPs



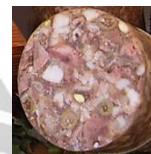
Centro di Referenza Nazionale
per Sequenze Genomiche di
microorganismi patogeni:
banca dati e analisi di Bioinformatica




Isolati di *L. monocytogenes* europei
con stesso profilo PFGE e/o stesso ST

Listeria monocytogenes 1/2a
MLST Sequence Type 7
PFGE Pulsotipo: Ascl.0183 Apal.0063
Ascl.0060 Apal.0063

- Is
- Outbreak strains** ST7 - Ascl.0183 Apal.0063 PFGE profile
 - Italian strains** ST7 - Ascl.0060. Apal.0063 PFGE profile
 - Strains from NL** ST7 - Ascl.0183 Apal.0063 PFGE profile
 - Strains from FR** ST7
 - Strains from AT** ST7
 - Strain from FR** CC7
 - Foodborne outbreak strains** ST7 - Ascl.0183 Apal.0063 PFGE profile



Epidemia listeriosi Marche

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



CRN: immediato confronto a livello nazionale delle sequenze e con altre banche dati internazionali

- *avvenuti a seguito di scambi telefonici (internazionali) oppure a posteriori (per alcuni isolati clinici)*

- Omogeneità del dato finale

- *assemblaggio, annotazione, criteri di inclusione*

- Utilizzo di nomenclatura minima per metadati

- Es: dato minimo di localizzazione, centroide della provincia oppure Comune, anonimizzazione dell'azienda.

- Utilizzo di Dizionari:

- Es: taleggio, Gorgonzola piccante, Brie → *SoftCheese*. Vantaggio per ricerca e confronto





È possibile creare un sistema nazionale per la identificazione e lo studio di focolai ?

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



FOOD SAFETY: Patricia M. Griffin, Section Editor

Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation

Brendan R. Jackson,¹ Cheryl Tarr,¹ Errol Strain,² Kelly A. Jackson,¹ Amanda Conrad,¹ Heather Carleton,¹ Lee S. Katz,¹ Steven Stroika,¹ L. Hannah Gould,¹ Rajal K. Mody,¹ Benjamin J. Silk,¹ Jennifer Beal,² Yi Chen,² Ruth Timme,² Matthew Doyle,² Angela Fields,² Matthew Wise,¹ Glenn Tillman,³ Stephanie Defibaugh-Chavez,⁴ Zuzana Kucerova,¹ Ashley Sabol,¹ Katie Roache,¹ Eija Trees,¹ Mustafa Simmons,³ Jamie Wasilenko,³ Kristy Kubota,⁵ Hannes Pouseele,⁶ William Klimke,⁷ John Besser,¹ Eric Brown,² Marc Allard,² and Peter Gerner-Smidt¹

¹Centers for Disease Control and Prevention, Atlanta, Georgia; ²Food and Drug Administration, College Park, Maryland; ³US Department of Agriculture, Food Safety and Inspection Service, Athens, Georgia; ⁴US Department of Agriculture, Food Safety and Inspection Service, Washington D.C.; ⁵Association of Public Health Laboratories, Silver Spring, Maryland; ⁶Applied Maths, Sint-Martens-Latem, Belgium; and ⁷National Institute for Biotechnology Information, National Institutes of Health, Bethesda, Maryland



È possibile creare un sistema nazionale per la identificazione e lo studio di focolai ?

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

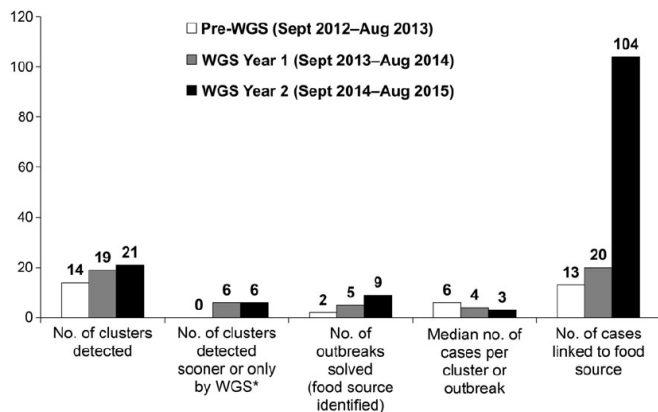


Figure 1. Listeriosis clusters detected and outbreaks solved before and after implementation of real-time whole-genome sequencing (WGS) of *Listeria monocytogenes* isolates from patients, food, and the environment, United States, September 2012–August 2015. *Cluster detection sooner or only by WGS, as compared with pulsed-field gel electrophoresis.

Table 1. Whole-Genome Sequencing Analyses by High-Quality Single-Nucleotide Polymorphism and Whole-Genome Multilocus Sequence Typing of *Listeria monocytogenes* Isolates From Patients, Foods, and the Environment in the 8 Multistate Outbreaks of Listeriosis Solved in the United States During September 2012–August 2015

WGS Year ^a	No. of Cases	Implicated Food	No. of Food and Environmental Isolates	Maximum No. of Differences Among Outbreak-Related Isolates (Median)		Reference
				hqSNPs	Alleles by wgMLST	
Pre-WGS	6	Soft cheese (French-style)	2	8 (1)	5 (1.5)	[36]
Year 1	8	Soft cheese (Latin-style)	37	18 (7.5)	19 (5)	[37]
Year 1	5	Mung bean sprouts	36	25 (6)	14 (5)	[31]
Year 1	2	Stone fruit	46	44 (2)	43 (10)	[34]
Year 2	9	Soft cheese (Latin-style)	7	21 (16)	22 (14)	[38]
Year 2	35	Caramel apples	32	Clade 1: 18 (10) Clade 2: 10 (4) ^b	Clade 1: 11 (3) Clade 2: 8 (3) ^b	[32]
Year 2	4	Ice cream (from production facility 1)	144	31 (11) ^c	30 (15)	[29]
Year 2	5	Ice cream (from production facility 2)	51	67 (17) ^d	32 (15)	[29]
Year 2	34	Soft cheese (Middle Eastern and European-style)	9	38 (16)	26 (13)	[30]

Abbreviations: hqSNPs, high-quality single-nucleotide polymorphisms; wgMLST, whole-genome multilocus sequence typing; WGS, whole-genome sequencing.


^a Pre-WGS year is September 2012–August 2013, Year 1 is September 2013–August 2014, and Year 2 is September 2014–August 2015.

^b Clades were median of 7494 hqSNPs and 1635 alleles different.

^c Four additional food isolates were up to 277 hqSNPs and 239 alleles from others from facility 1.

^d Up to 387 hqSNPs and 265 alleles between isolates from facilities 1 and 2 of producer A. Cases linked to the 2 facilities are considered part of the same outbreak.



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

È possibile creare un sistema nazionale per la identificazione e lo studio di focolai ?



RÉPUBLIQUE FRANÇAISE
MINISTÈRE
DE L'AGRICULTURE
DE L'AGROALIMENTAIRE
ET DE LA FORÊT

Real-Time Whole-Genome Sequencing for Surveillance of *Listeria monocytogenes*, France

Alexandra Moura,¹ Mathieu Tourdjman,¹ Alexandre Leclercq,¹ Estelle Hamelin, Edith Laurent, Nathalie Fredriksen, Dieter Van Cauteren, Hélène Bracq-Dieye, Pierre Thouvenot, Guillaume Vales, Nathalie Tessaud-Rita, Mylène M. Maury, Andreea Alexandru, Alexis Criscuolo, Emmanuel Quevillon, Marie-Pierre Donguy, Vincent Enouf, Henriette de Valk, Sylvain Brisse,² Marc Lecuit²



È possibile creare un sistema nazionale per la identificazione e lo studio di focolai ?


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


Table 3. Source identification of cgMLST cluster-alerts for surveillance of *Listeria monocytogenes*, France*

cgMLST cluster-alert size, no. human isolates	No. (%)				
	Clusters, n = 119	Human isolates, n = 770	Food/environment isolates, n = 1,973	Clusters with confirmed source	Cluster-alerts with identified source resulting in withdrawal or recall
Small, 1	37 (31)	37 (5)	145 (7)	10 (27)	3 (30)
Medium, 2–5	73 (61)	185 (24)	123 (6)	9 (12)	5 (55)
Large, >5	9 (8)	89 (12)	83 (4)	3 (38)	0

*cgMLST, core genome multilocus sequence typing.



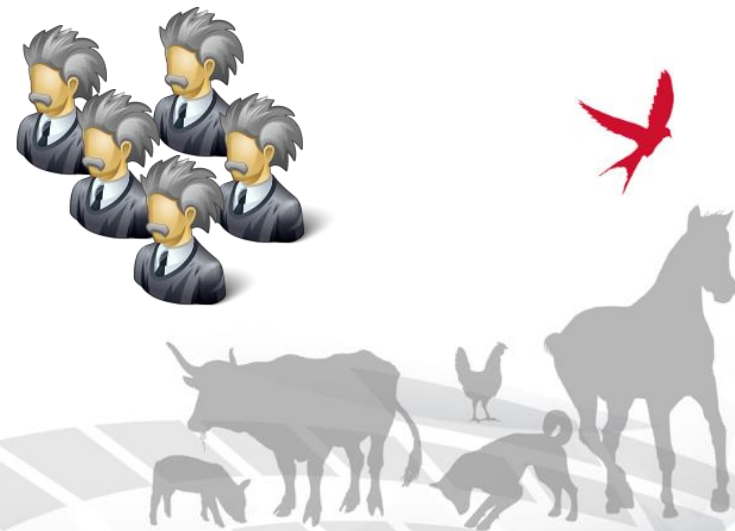
Piano GENPAT per il 2018

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

- 
- Costruzione **piattaforma nazionale**



- Definizione **sistema strutturato e permanente di referenti** per il coordinamento delle attività;





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

Grazie per l'attenzione

