FDA on how it uses WGS
Plus, Cost – Benefit - Budget

"Whole Genome Sequencing (WGS) for food safety and its uses in prevention and response of foodborne outbreaks",

The Pathogen Surveillance in Agriculture, Food and the Environment (PATH-SAFE) Programme conference

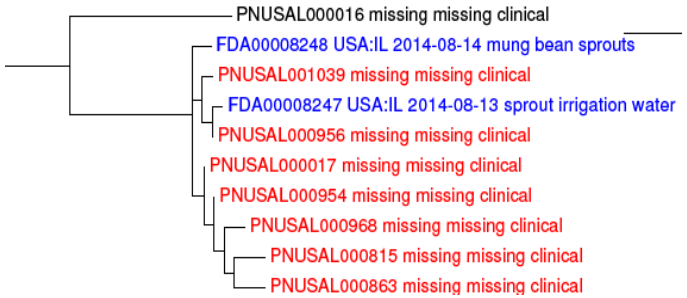
London UK

Feb. 28th and 29th, 2024

Marc Allard PhD, Ruth Timme PhD, and Eric Stevens PhD.
FDA Center for Food Safety and Applied Nutrition
Marc.allard@fda.hhs.gov

New Field: Genomic Epidemiology

Genomic Signal

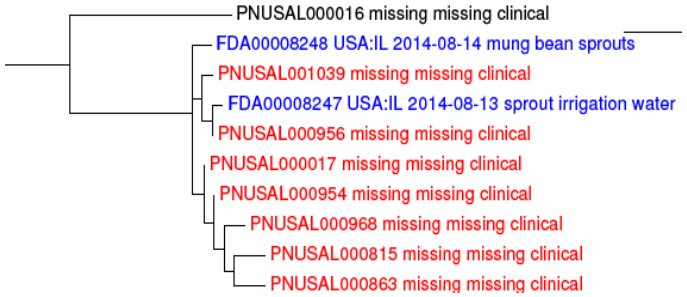


Epidemiological Signal

Investigation

Gen Epi regulatory communication

ORA / LFFM-funded laboratories upload genomic data to NCBI



ORA Office of Regulatory Affairs
Regulatory arm of FDA
Inspections, sequencing

OAO: watches for signal

CORE and OC: to communication to CDC, USDA-FSIS and State DOH + Ag

OAO Office of Analytics and Outreach
Data interpretation and Risk assessment
CORE Outbreaks, OC Compliance

CORE and OC:

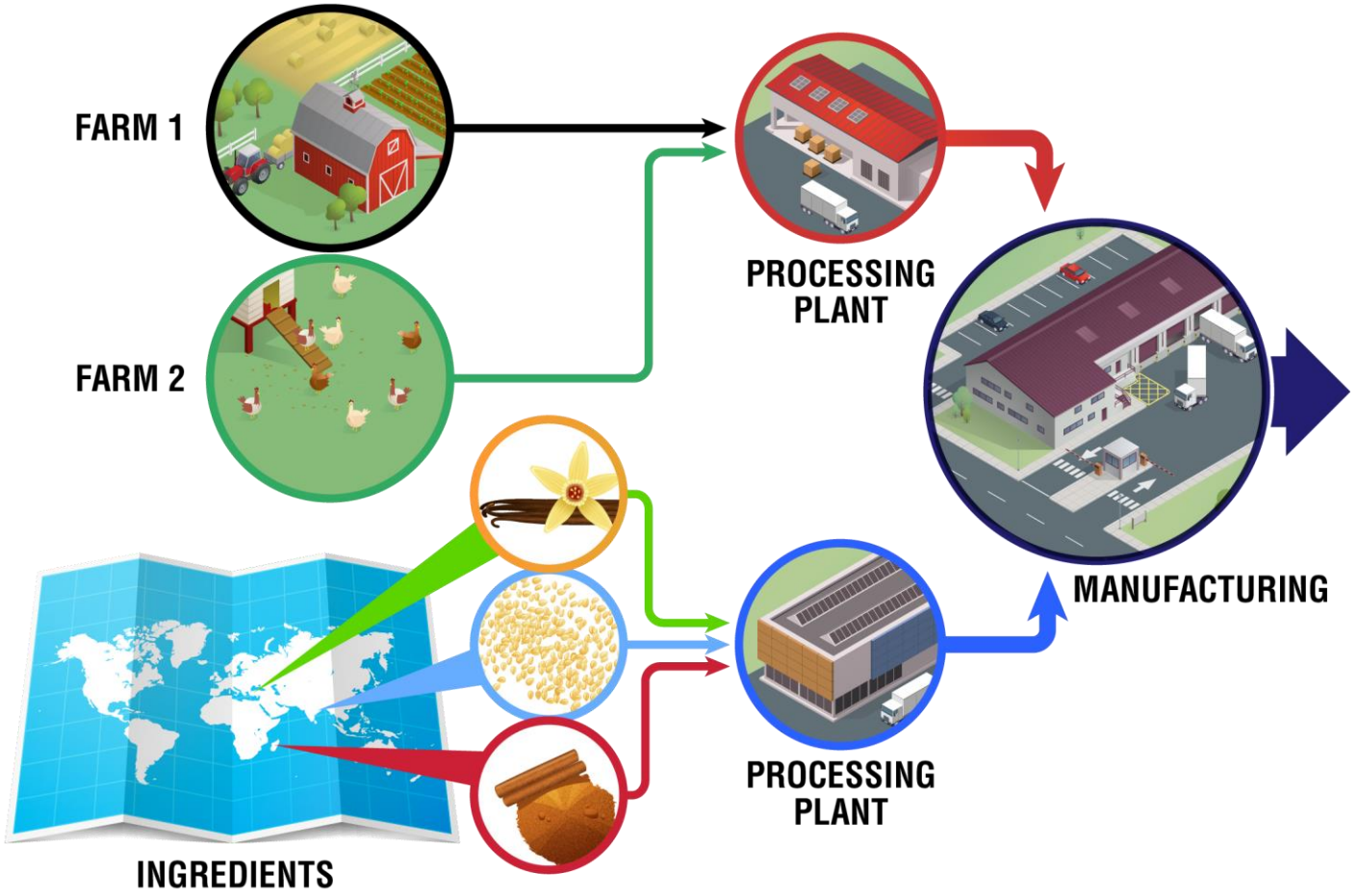
Facility/farm Inspections or product testing by ORA Consumer Safety Officers

OFS Office of Food Safety, policy
OIE Office of International Engagement
ORS Office of Regulatory Science, research

Current FDA workflow works for all pathogens and all genomic methods, collected under FDA surveillance and inspection activities (virus, parasite, shellfish, filth, supplement botanicals).

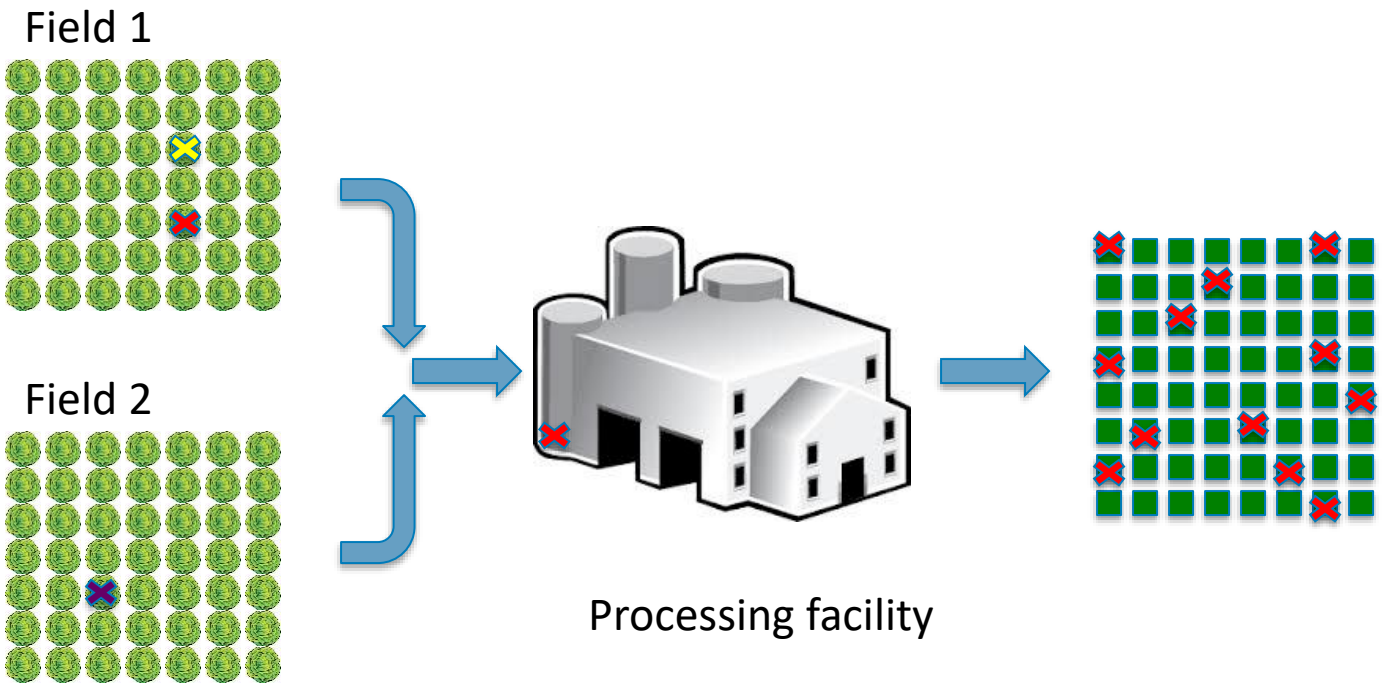
Identifying an Outbreak Vehicle:
Trace Forward and Trace Backward

U.S. FOOD & DRUG ADMINISTRATION

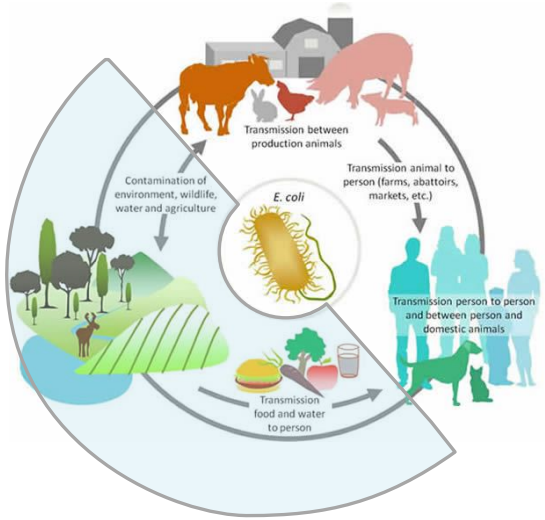


U.S. FOOD & DRUG ADMINISTRATION

Identifying an Outbreak Vehicle:
Determining Resident or Transient pathogen



FDA's GenomeTrakr program

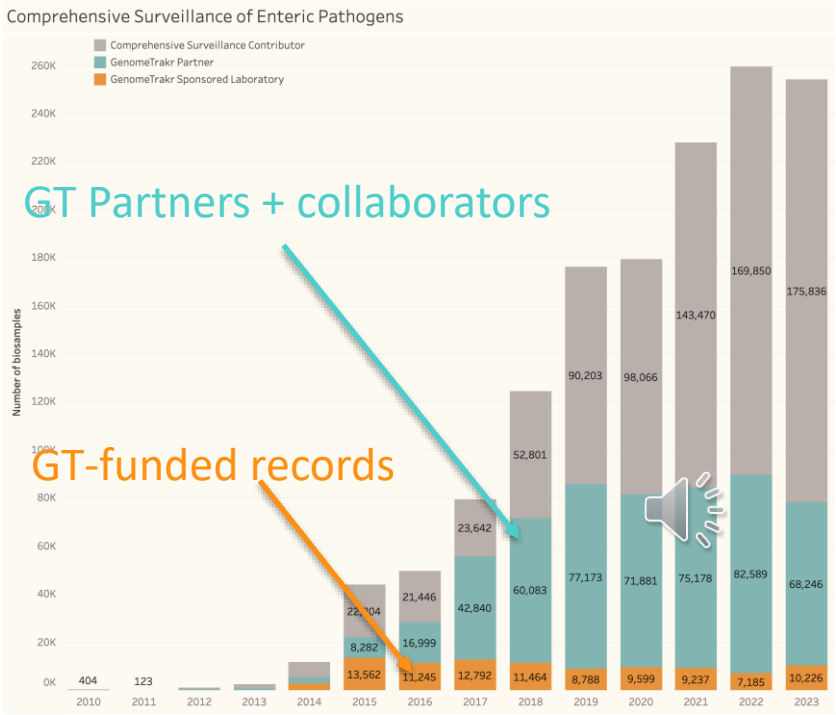


- Sequencing the genomes of foodborne pathogens found in food, food processing facilities, farm environment, water, etc.
- Collaborate with other US agencies and international counterparts to integrate our data with genomic data collected from animals and human clinicals – data made public in real-time.
- Clustering at NCBI Pathogen Detection helps FDA identify causes of foodborne outbreaks and identify other events, like harborage.

45 funded laboratories:

Numerous GT partners and collaborators:

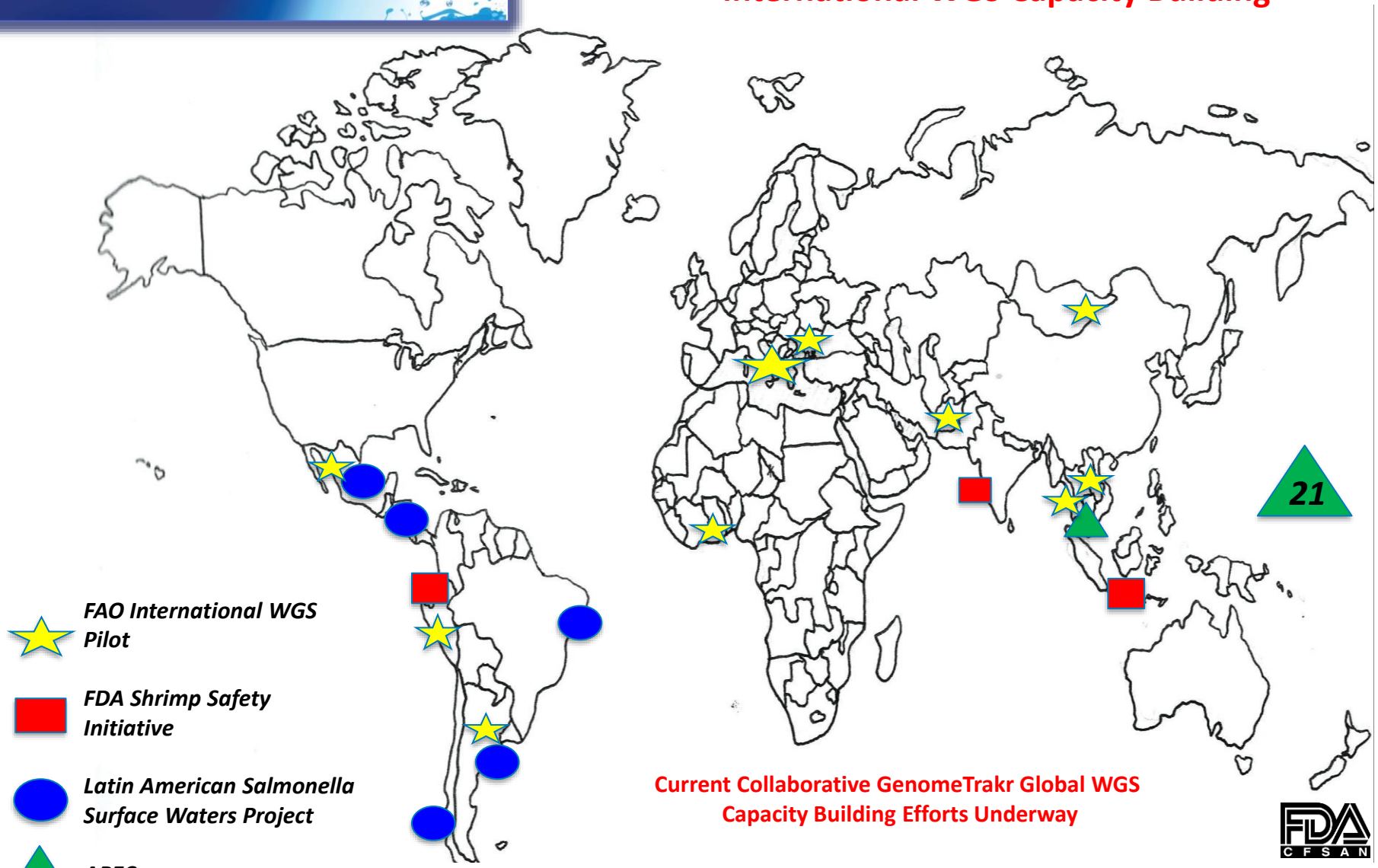
- Public health, Ag, and academic labs
- US agency partners: GenFS and others
- International counterparts



GLOBAL WATER & FOOD SAFETY SUMMIT

Increase environmental sampling across the US, and abroad.
Meeting Nov. 19 - 21, 2019 College Park, MD and FAO, Rome, 2024

International WGS Capacity Building



Current Collaborative GenomeTrakr Global WGS Capacity Building Efforts Underway

FDA CFSAN

Best Practices from GenomeTrakr


BMC Part of Springer Nature

One Health Outlook

Home About Articles Submission Guidelines

Review | Open Access | Published: 19 October 2020

Optimizing open data to support one health: best practices to ensure interoperability of genomic data from bacterial pathogens

Ruth E. Timme, William J. Wolfgang, Maria Balkey, Sai Laxmi Gubbala Venkata, Robyn Randolph, Marc Allard & Errol Strain

One Health Outlook 2, Article number: 20 (2020) | Cite this article

2353 Accesses | 11 Citations | 35 Altmetric | Metrics

Abstract

The holistic approach of One Health, which sees human, animal, plant, and environmental health as a unit, rather than discrete parts, requires not only interdisciplinary cooperation, but standardized methods for communicating and archiving data, enabling participants to easily share what they have learned and allow others to build upon their findings. Ongoing work by NCBI and the GenomeTrakr project illustrates how open data platforms can help meet the needs of federal and state regulators, public health laboratories, departments of agriculture, and universities. Here we describe how microbial pathogen surveillance can be transformed by having an open access database along with Best Practices for contributors to follow. First, we describe the open pathogen surveillance framework, hosted on the NCBI platform. We cover the current community standards for WGS quality, provide an SOP for assessing your own sequence quality and recommend QC thresholds for all submitters to follow. We then provide an overview of NCBI data submission along with step by step details. And finally, we provide curation guidance and an SOP for keeping your public data current within the database. These Best Practices can be models for other open data projects, thereby advancing the One Health goals of Findable, Accessible, Interoperable and Re-usable (FAIR) data.

GO FAIR

Standard metadata required for interoperability

- FAIR = Findable, accessible, interoperable, reusable

protocols.io

Public, version-controlled protocols

- GenomeTrakr workspace:

Galaxy Trakr

Open-access analysis platform

- "MicroRunQC": QC workflow for microbial pathogens

DDBJ INSDC IAC ICM NCBI ENA/EBI

Open data repository for hosting genome + metadata

- Enables public/private collaboration

Timme, R.E., Wolfgang, W.J., Balkey, M. *et al. One Health Outlook* **2,** 20 (2020).
https://doi.org/10.1186/s42522-020-00026-3
https://www.protocols.io/workspaces/genometrakr1/publications

One Health Enteric package:

US Interagency Collaboration for Genomics for Food and Feed Safety (Gen-FS)

National Center for Biotechnology Information (NCBI)
Centers for Disease Control and Prevention (CDC)
Food and Drug Administration (FDA)
U.S. Department of Agriculture (USDA)

OHE package scope:

CORE attributes
- Isolate identifiers
- Collected by
- Date of collection
- Geographic location
- Sampling purpose
- Sampling device
- Project name
- IFSAC category
- Source type
- sequenced by

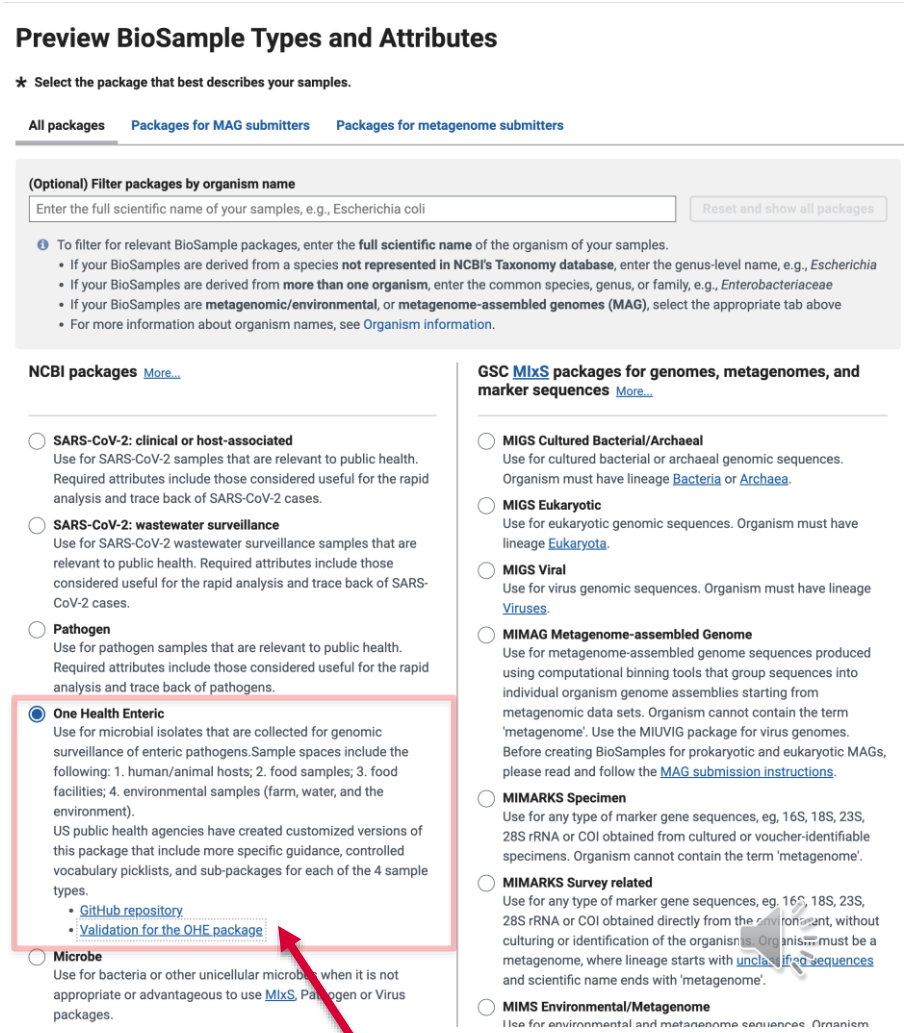
Human/animal host
- Host
- Host disease
- Host sex + age
- Host tissue sampled
- Animal environment
- Antimicrobials in food
- Animal housing system

Food samples
- Geographic origin
- Intended consumer
- Collection site description
- Food product type
- Food source
- Food processing types
- Food preservation process
- Food cooking process
- Food additives
- Food contact surface
- Food container wrapping
- Food quality date

Food facility
- Facility type
- Building setting
- Food processed
- Facility location
- Monitoring zone
- Indoor sampling surface
- Surface material
- Surface material cond.
- Surface orientation
- Surface temperature
- Biocide used
- Animal intrusion

Farm and Environment
- Environmental material
- Farm type
- Plant growth medium
- watering method
- Relative loc of sample
- Fertilizer administration
- Food cleaning process
- Sanitizer used
- Farm equip. used
- Water samples
- Extreme weather event
- Mechanical damage

Generic template available at NCBI B :



New links to GenomeTrakr resources!

US enteric pathogen surveillance

Pathogens

- Academia
- US Federal Agencies
- Public health labs
- Government Research
- Agriculture labs
- Hospitals
- Veterinary Clinics
- Industry

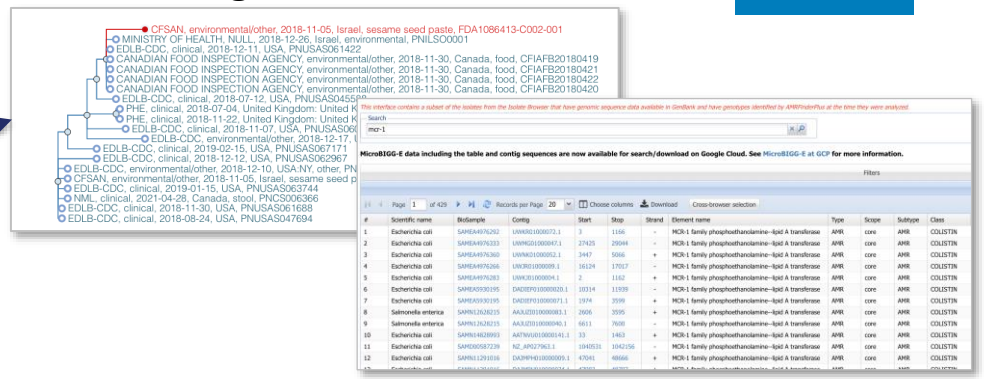
Pathogen DOM

PulseNet

Pathogen DOM

INSDC
Public repository

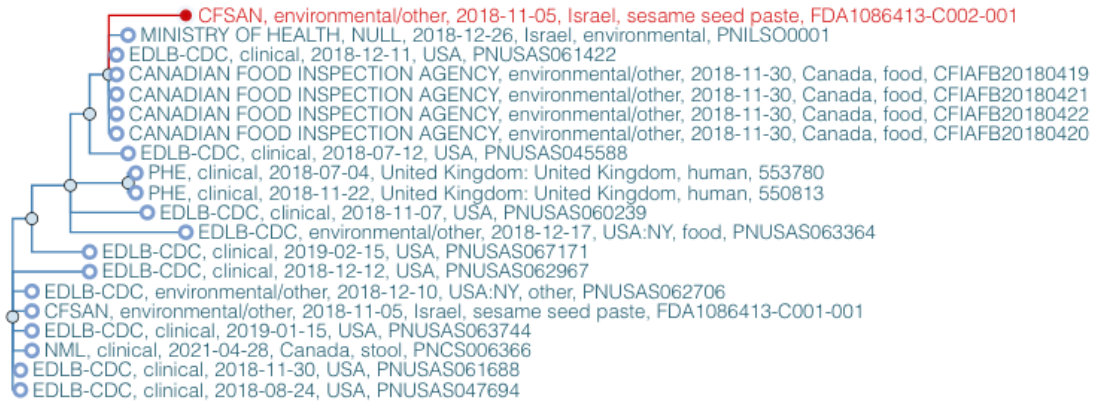
NCBI Pathogen Detection



FDA public dashboards

Salmonella tahini clusters highlight global contribution

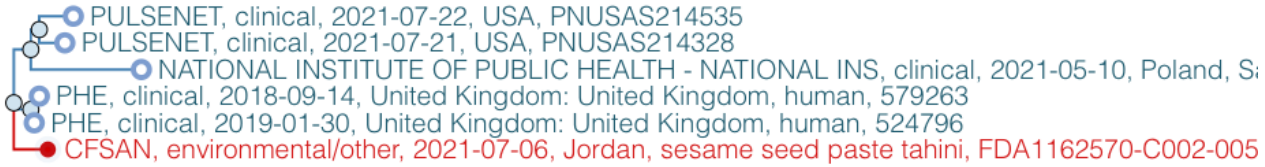
130,935 Clusters currently tracked.



- United Kingdom (PHE)
- United States (GenomeTrakr, PulseNet)
- Canada (CFIA, NLM)
- Israel

- United Kingdom (PHE, GBRU)
- United States (GenomeTrakr, PulseNet)

- United Kingdom
- United States (GenomeTrakr)



- United Kingdom (PHE)
- United States (GenomeTrakr, PulseNet)
- Poland

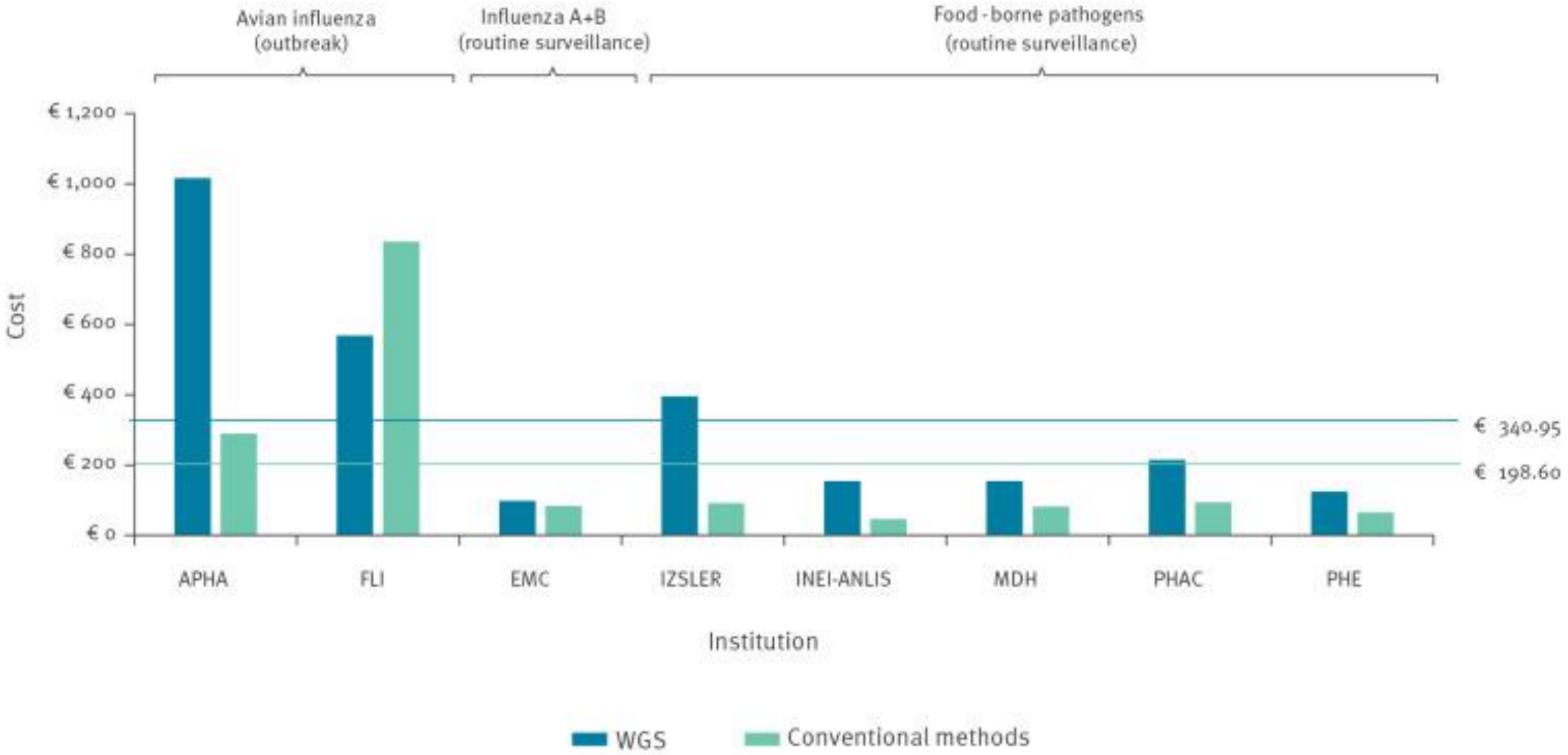
Case study area	Avian influenza (HPAI)		Influenza A+B	Food-borne pathogens[a]				
Institution	APHA (UK)	FLI (DE)	EMC (NL)	IZSLER (IT)	INEI-ANLIS (ARG)	MDH (US)	PHAC (CAN)	PHE (UK)
Outbreak or routine surveillance	Outbreak	Outbreak	Routine surveillance	Routine surveillance	Routine surveillance	Routine surveillance	Routine surveillance	Routine surveillance
Number of samples in reference period	26	30	630	175	320	1,767	8,630	15,791
	in 8 months	3 months	5 months	12 months	12 months	12 months	12 months	12 months
WGS								
Sequencer used	Illumina MiSeq	IonTorrent PGM	Nanopore GridION	Illumina MiSeq	Illumina MiSeq	Illumina MiSeq	Illumina MiSeq	Illumina HiSeq
Batch size for sample processing/sequencing	1–2	6	30	24	12	24	32	Processing: 40 Sequencing: 96
Equipment	€ 58.53	€ 210.71	€ 2.50	€ 163.49	€ 43.02	€ 29.53	€ 75.90	€ 35.23
Consumables	€ 830.97	€ 254.88	€ 33.52	€ 165.37	€ 104.62	€ 104.40	€ 69.75	

Economic impact studies

Overview of per-sample costs of whole genome sequencing vs conventional methods, by cost type, case studies covering a specified reference period between 2016 and 2019 (n = 8 institutes)

INEI-ANLIS Dr Carlos G Malbrán, Buenos Aires, Argentina

Alleweldt et al. Economic evaluation of whole genome sequencing for pathogen identification and surveillance—results of case studies in Europe and the Americas 2016 to 2019. Euro. Surveill. 2021 Mar 4; 26(9): 1900606 **Celine Nadon,



Over-all per-sample costs of whole genome sequencing vs conventional methods, case studies covering a specified reference period between 2016 and 2019 (n = 8 institutes).

Case study institution	IZSLER (IT)	INEI-ANLIS (ARG)	MDH (US)	PHAC (CAN)	PHE (UK)	Average
Cost per sample (WGS)	€ 395.14	€ 154.49	€ 154.51	€ 215.36	€ 124.59	€ 208.82
Cost per sample (conventional methods)	€ 91.87	€ 46.61	€ 81.16	€ 94.29	€ 65.46	€ 75.88
Differential cost of WGS compared with conventional methods	€ 303.27	€ 107.88	€ 73.35	€ 121.07	€ 59.13	€ 132.94
Number of samples per year (*Salmonella*)	110	128	1,010	8,273	10,147	3,934
Total additional costs per year due to the use of WGS	€ 33,360	€ 13,809	€ 74,084	€ 1,001,623	€ 599,992	€ 344,573
Average cost per reported case of salmonellosis	€ 12,124	€ 11,821	€ 13,225	€ 12,174	€ 12,401	€ 12,349
Number of reported cases of salmonellosis that need to be avoided to break even	2.8	1.2	5.6	82.3	48.3	28.0
Number of cases of salmonellosis reported annually[a]	276[b]	758	906	7,665	8,770	4,404
Percentage of total number of reported cases of salmonellosis that need to be avoided to	1.0%	0.2%	0.6%	1.1%	0.6%	0.7%

Results of break-even analysis, whole genome sequencing vs conventional methods, case studies covering a specified reference period between 2016 and 2018 (n = 5 institutes).

Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna, Parma, Italy

Results

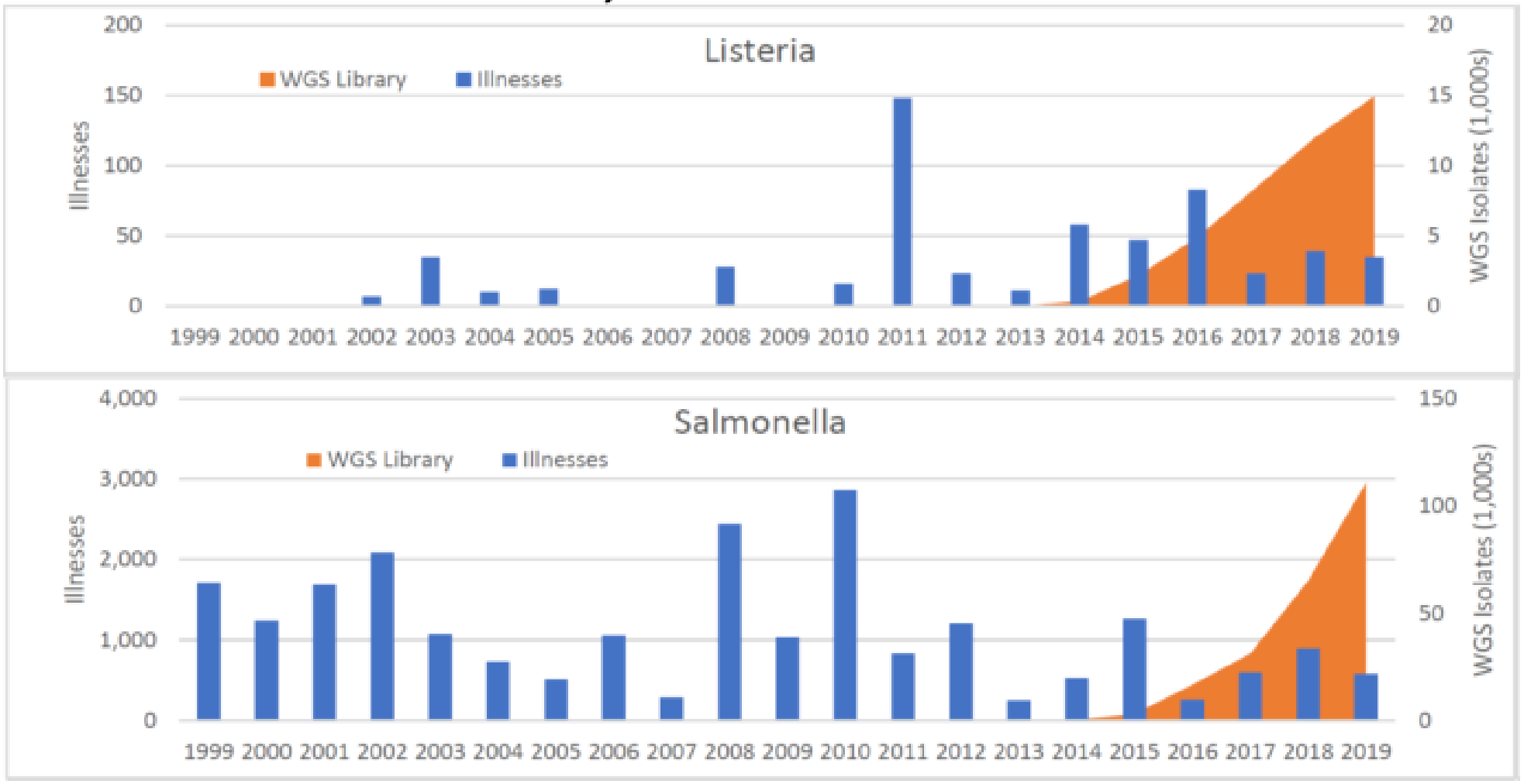
On a per-sample basis, WGS was between 1.2 and 4.3 times more expensive than routine conventional methods. However, WGS brought major benefits for pathogen identification and surveillance, substantially changing laboratory workflows, analytical processes and outbreaks detection and control. Between 0.2% and 1.1% (on average 0.7%) of reported salmonellosis cases would need to be prevented to break even with respect to the additional costs of WGS.

Conclusions

Even at cost levels documented here, WGS provides a level of additional information that more than balances the additional costs if used effectively. The substantial cost differences for WGS between reference laboratories were due to economies of scale, degree of automation, sequencing technology used and institutional discounts for equipment and consumables, as well as the extent to which sequencers are used at full capacity.

Ford et al. Cost of whole genome sequencing for non-typhoidal Salmonella enterica. PLoS ONE 2021; 16(3):e0248561 For Australia break even is 1.9%

Data and Summary Statistics



Brown et al. (2021) An economic evaluation of the Whole Genome Sequencing source tracking program in the U.S. PLoS ONE 16(10): e0258262.

$$SV = \underbrace{[p_x * x - (c_x(x) + c_e(e(WGS))]}_{\text{profit Function}} - \underbrace{[C_I * x * \gamma_I(e(WGS)) * n_I(WGS)]}_{\text{public health externatility function}} - \underbrace{[c_{WGS}(WGS)]}_{\text{implentation cost}}$$

$$I_O = \underbrace{x * \gamma_I(e(WGS))}_{\text{probability outbreak occurs}} * \underbrace{n_I(WGS)}_{\text{number of illnesses in outbreak}} * \underbrace{\alpha_O(WGS)}_{\text{probability illnesses are observed}}$$

$$Y_{p,t} = \beta_0 + \beta_1 WGS\_library_{p,t} + \varepsilon_{p,t}$$

$$Y_{p,t} = \beta_0 + \beta_1 WGS\_library_{p,t} + \beta_2 X_{p,t} + \varepsilon_{p,t}$$

$$Y_{p,t} = \beta_0 + \beta_1 WGS\_library_{p,t} + \beta_2 X_{p,t} + \beta_3 FSMA_t + \varepsilon_{p,t}$$

$$Benefits = \hat{\beta}_1 \, x \, WGS \; Isolates \, x \, Underreporting \; Multiplier \, x \, Monetary \; Loss$$

If you want to do these kinds of calculations, please let our PhD economist talk to yours.

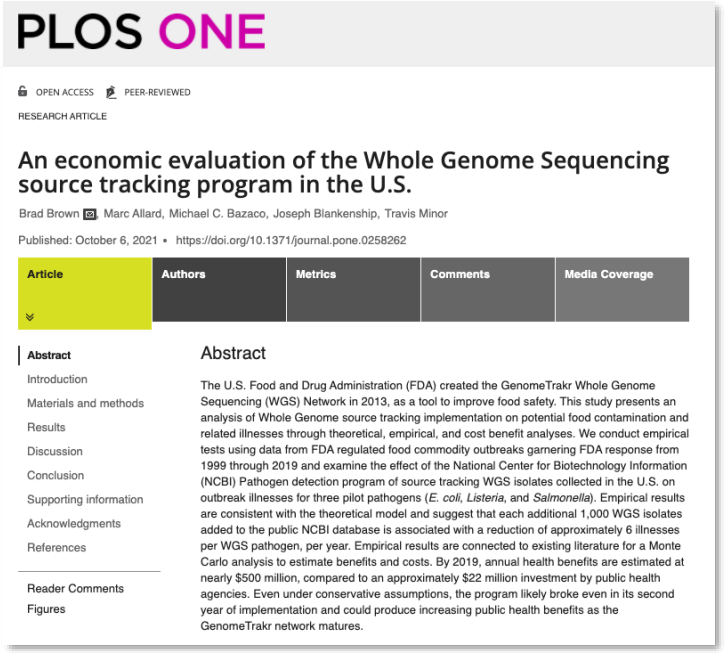
Economic Evaluation of WGS Reduces the Burden of Illness

Total Burden Averted (in millions)

	Listeria	E. coli	Salmonella	Yearly Total	Total 90% CI
2014	$7.43	$0.12	$0.39	$7.94	($2.96 - $13.61)
2015	$50.95	$1.68	$2.83	$55.46	($20.79 - $94.89)
2016	$114.23	$6.13	$14.69	$135.04	($51.03 -$229.39)
2017	$197.39	$15.24	$27.46	$240.09	($90.87 - $406.78)
2018	$280.62	$29.94	$57.30	$367.86	($139.56 - $620.41)
2019	$348.48	$51.03	$97.47	$496.98	($188.62 -$835.92)

Economic Impact

- GenomeTrakr program was likely cost effective by its second year of implementation
- $100 M -> $450 M in net annual health benefits (est. from 2019). >$ Billion estimated benefits.



PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

An economic evaluation of the Whole Genome Sequencing source tracking program in the U.S.

Brad Brown, Marc Allard, Michael C. Bazaco, Joseph Blankenship, Travis Minor

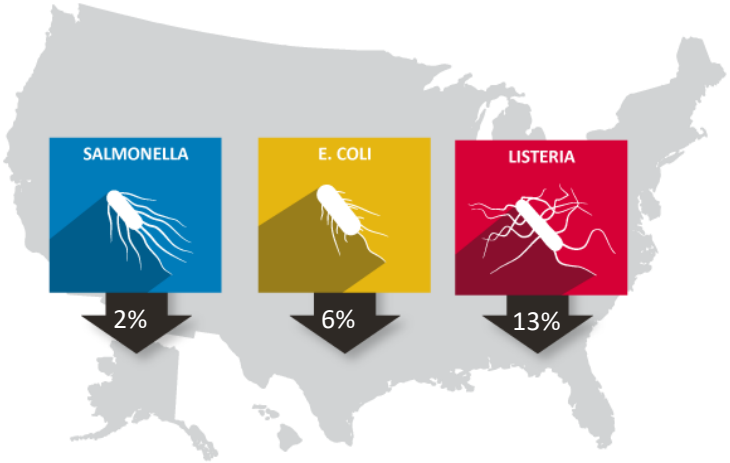
Published: October 6, 2021 • https://doi.org/10.1371/journal.pone.0258262

Article | Authors | Metrics | Comments | Media Coverage

Abstract
Introduction
Materials and methods
Results
Discussion
Conclusion
Supporting information
Acknowledgments
References
Reader Comments
Figures

Abstract

The U.S. Food and Drug Administration (FDA) created the GenomeTrakr Whole Genome Sequencing (WGS) Network in 2013, as a tool to improve food safety. This study presents an analysis of Whole Genome source tracking implementation on potential food contamination and related illnesses through theoretical, empirical, and cost benefit analyses. We conduct empirical tests using data from FDA regulated food commodity outbreaks garnering FDA response from 1999 through 2019 and examine the effect of the National Center for Biotechnology Information (NCBI) Pathogen detection program of source tracking WGS isolates collected in the U.S. on outbreak illnesses for three pilot pathogens (*E. coli*, *Listeria*, and *Salmonella*). Empirical results are consistent with the theoretical model and suggest that each additional 1,000 WGS isolates added to the public NCBI database is associated with a reduction of approximately 6 illnesses per WGS pathogen, per year. Empirical results are connected to existing literature for a Monte Carlo analysis to estimate benefits and costs. By 2019, annual health benefits are estimated at nearly $500 million, compared to an approximately $22 million investment by public health agencies. Even under conservative assumptions, the program likely broke even in its second year of implementation and could produce increasing public health benefits as the GenomeTrakr network matures.



Return on Investment: $10 dollars in averted human health costs for every $1 dollar invested. For each additional 1,000 WGS isolates added to the public NCBI database is associated with a reduction of approximately 6 illnesses per WGS pathogen, per year.

Price et al. 2023 A systematic review of economic evaluations of whole-genome sequencing for the surveillance of bacterial pathogens. Microb Genom 2023; 9(2). Discussion of 9 different economic impact studies.

There were significant variations in the research questions addressed in the various publications yet, most studies demonstrated cost savings due to WGS that were largely attributed to averted cases of infection.

For this benefit to be realized maximally, WGS needs to be employed early in the analytical pipeline. Conversely, delay in the use of WGS reduces the benefits, as early detection of outbreaks enables timely implementation of interventions to interrupt transmission.

More economic evidence of WGS in public health settings is required to foster wider applications of WGS as a surveillance tool in public health.

We dedicate this work to Robert Stones FERA