

# Public Health Laboratory Recruitment Guide: Genomic Epidemiologist



## Summary

The genomic epidemiologist, often called molecular epidemiologists, will be responsible for leading program initiatives focusing on the integration of genomic sequencing data and epidemiologic metadata to better understand disease transmission and assist with pathogen surveillance and outbreak investigations. The genomic epidemiologist will apply a wide understanding of next generation sequencing (NGS) methodologies, pathogen genomics, data analysis and integration to generate reports and visualizations and disseminate results to relevant partners. There are various uses for sequencing data, including surveillance, outbreak investigation, and clinical diagnostics, which involve careful interpretation by genomic epidemiologists. Depending on the specific sequencing purpose, investigative questions asked, or the epidemiological application of genomic interpretations, genomic epidemiologists will work and share results with a wide variety of partners.

The genomic epidemiologist can be situated within the laboratory or within the health department depending on your organizational structure. Access to jurisdictional surveillance data systems may dictate which organization is best suited to employ the genomic epidemiologist. There is value in having a genomic epidemiologist sitting within the laboratory because their foundational knowledge in genomics, microbiology and epidemiology allows for this role to serve as a bridge between laboratorians and epidemiologists, as well as foster more robust bi-directional communication and information sharing. Building a stronger partnership between the laboratory and epidemiology departments can increase information accessibility, support comprehensive outbreak investigations, and improve analytical turnaround for public health action. In addition to building partnerships within the health department, the genomic epidemiologist may also work to strengthen relationships with other partners, including state and local public health agencies, healthcare facilities, infection preventionists, outside laboratories and federal partners.

Interested applicants are likely to come from different academic backgrounds. An applicant who majored in public health epidemiology, while strong in data analysis and information sharing, may lack biological knowledge and need to rely heavily on laboratory staff for determining biological plausibility. At the other end of the spectrum are candidates with a foundation in biological sciences who have practiced epidemiology

## Executing Bioinformatic Pipelines

There are many factors that should be considered when asking genomic epidemiologists to execute bioinformatic pipelines. The degree to which this is a job requirement or expectation depends on the personnel resources to complement the genomic epidemiologist position in the laboratory or health department. Depending on the pathogen sequenced and the laboratory's regulatory requirements (e.g., CLIA), genomic epidemiologists may not be able to run pipelines for results. Genomic epidemiologists may have varied expertise for executing bioinformatics pipelines. A genomic epidemiologist with limited bioinformatics experience may be expected to follow standard operating procedures to run previously constructed (preconstructed) bioinformatics pipelines through graphic user interface (i.e., Terra.bio). A genomic epidemiologist with more bioinformatics experience may be comfortable executing pipelines independently at the command line interface and troubleshooting common errors. Genomic epidemiologists may run bioinformatic pipelines to compare results and to deepen their understanding of results produced by collaborating with bioinformaticians. Although some genomic epidemiologists may be comfortable executing code, their knowledge and skillsets are different from a bioinformatician. As such, these roles should largely remain separate.

during their career. These candidates may understand the biological significance of the data (e.g., phylogenetic analysis, antimicrobial resistance, etc.) and communicate well with clinical laboratories, but may not possess efficient statistics, data analysis, and report generation skills. Candidates with a degree in genomic or molecular epidemiology were educated with a mix of microbiology, molecular biology and epidemiology course work, giving them strong technical knowledge of molecular diagnostic testing and the epidemiological results that can be inferred from that data. Managers should consider hiring a candidate without the most ideal degree if an epidemiologist possesses other job competencies, as many training resources are available for on-the-job workforce development.

Enclosed are position descriptions, based on the various competencies around sequencing, epidemiologic investigations, data analysis and report generation, in which genomic epidemiologists should be proficient. In addition to an understanding of genomics and epidemiologic skillsets, having experience in training others and communicating across different areas of expertise, including both scientific and non-scientific communities, are important for genomic epidemiologists to succeed in a public health laboratory setting.

All materials in this guide are meant to be a template for a job posting but can and should be modified to fit your specific needs and resources. Download a word version of this guide for easy editing.

## Resource Requirements to Support This Position

Genomic epidemiologists work with a variety of data types, file sizes and analytical software which require IT infrastructure in place to support their work. Prior to recruiting applicants for this position, you are encouraged to review some of the below resources that are essential to their job duties:

- High bandwidth (internet network upload and download speeds and switch architecture to support large file transfers) and capability and capacity to easily transfer data and analytic results between the interested parties.
- Access to pertinent laboratory data, epidemiologic data and metadata for thorough analyses.
- Access to required computational and analytical platforms (Terra.bio, CLC Workbench, Nextflow Tower, Datapult, etc.) and relevant workgroups.
- Access to software packages to support data visualization and curation (R, SAS, SQL, Tableau, PowerBI, etc.).

## Helpful Tips for Hiring Process

- Include both laboratory and epidemiology departments on the interview panel.
- Making this position remote eligible or hybrid remote will encourage a wider talent pool of qualified applicants.

# Roles of a Genomic Epidemiologist by Department

## In the Laboratory

- Works closely with bioinformaticians for pipeline development and data analysis.
- Cross-trains with bioinformaticians and data analysts to ensure report-ready file structures and staff continuity for surge needs.
- Shares computational resources with bioinformaticians and data analysts.
- Prioritizes samples for sequencing, avoiding unnecessary sequencing.
- Models sampling strategies during peak periods (e.g., COVID).
- Understands and communicates data outputs and quality concerns.
- Rapidly gains insights into data quality and its impact.

## In the Epidemiology Department

- Trains epidemiologists in transmission dynamics and sampling design.
- Accesses case reports to provide insights into epidemiological variables.
- Utilizes historic data for comprehensive investigations.
- Interfaces with stakeholders for data and analysis discussions.
- Accesses essential informatics systems and data linkages.
- Coordinates the return of data from other laboratories and jurisdictions.

## In Both Departments

- Laboratory and epidemiology departments collaborate due to separate agency structures.
- Laboratory focuses on genomic approaches, relatedness thresholds, and report interpretation by collaborating with bioinformaticians for pipeline evaluation.
- Epidemiology acts as a liaison with local health departments for WGS requests.
- Genomic epidemiologists in each department may have access to different data systems that can facilitate better data sharing and linkage.
- Facilitates inter-agency communication and collaboration.

## Position Description

Program/Department: Public Health Laboratory/Health Department

Position Title: Genomic Epidemiologist

Reports to: Director / Division Director / Manager

Previous Incumbent: **None (New position)**

**Download an editable version  
of the position description  
and competencies.**

## Job Position Summary

The genomic epidemiologist will be responsible for analyzing genomic data to support disease surveillance and outbreak investigation efforts. The incumbent is expected to apply analytical and statistical skills to create data visualizations and reports of genomic results. This includes but is not limited to appropriately selecting and/or running bioinformatics pipelines, performing phylogenetic analyses to delineate clusters of genetically related cases and identifying trends for pathogen surveillance. Through the integration of case interview information, laboratory sequencing results and outbreak surveillance data, genomic epidemiologists translate these findings in a meaningful way to prevent disease transmission and inform public health response. Effective and efficient communication skills, basic knowledge of microbiology and bioinformatics, in-depth knowledge of genomics and epidemiologic best practices, statistical analysis experience, training and outreach are necessary requirements for this position. The individual in this position will also work to strengthen partnerships with other public health agencies, health department divisions, medical providers and laboratories.

## Essential Job Duties

- Assess and advise on how sequencing and analytics may be used to address local and state investigations of different infectious pathogens; be knowledgeable of the scope of services that can be supported by the state public health laboratory and those which may be available through regional and national partnerships.
- Integrate sequencing data and metadata to identify trends for pathogen surveillance, plasmid exchange, and prediction of phenotypic characteristics (virulence, antimicrobial resistance).
- Conduct phylogenetic analyses to assess case relatedness and infer outbreak sources.
- Create and generate reports (e.g., public dashboards, surveillance reports, situational reports) and genomic interpretations from NGS data to share within the health department and with partners to support surveillance efforts and outbreak investigations.
- Acts as a liaison to the laboratory and epidemiology departments to facilitate information sharing and advance the use of genomic data in public health decision making.
- Educate, train and provide guidance to in-house staff on the interpretation of genomic results and epidemiologic analysis.

## Job Position Competencies

The competencies listed below can and should be adjusted to your jurisdiction's specific needs. The percentage of time spent within these competencies may change based on how this role will operate in your organization and should be adapted accordingly.

### Data Quality Assessment and Data Management (20%)

- 1. Consult with bioinformaticians to review the quality of consensus genomes.**
  - Understand quality metrics generated from bioinformatic analyses at designated checkpoints.
  - In consultation with bioinformaticians, inform their review of sequencing read data with epidemiologic context to help differentiate between coinfection, contamination and recombination.
- 2. Develop and evaluate sampling schemes to accurately estimate strain prevalence.**
  - Identify and develop a framework that estimates the appropriate number of sequencing samples needed to detect variants and to estimate variant prevalence among specific populations.
  - Implement innovative data management techniques for genomic data, metadata and pipeline parameters.
  - Track genomic consistency and organize metadata through the implementation of data management tools.
  - Ensure consistency of sequencing results by tracking pipeline parameters.
- 3. Data cleaning of associated metadata.**

Consolidate data sets and standardize data formats for downstream analysis.
- 4. Store, access, analyze, share, and archive data according to jurisdiction regulations and policy.**
  - Store data across multiple clouds and on premises information systems.
  - Archive and store data in accordance with retention schedules and regulatory requirements.
  - Maintain data security, confidentiality and privacy.

### Data Analysis and Report Generation (50%)

- 1. Identify new or existing genomic epidemiology analysis tools.**

Demonstrate knowledge of and utilize current best-practice tools for quality control, metadata incorporation, phylogenetic and geospatial analysis and report generation.
- 2. Deploy existing pipelines and perform required analysis.**

Run or request analysis by preconstructed bioinformatics pipelines to generate genome assemblies and consensus.
- 3. Utilize data visualization tools and develop reports for leadership and partners.**
  - Understand and account for limitations of inferences drawn for genomic data related to sequencing quality, sampling schemes, etc.
  - Develop visualizations to share analyzed data and examine data linkages.
- 4. Interact directly with other scientific staff and public health partner data users to ensure appropriate interpretation of results in a scientific context.**

## Workforce Training (10%)

Training the public health workforce on key concepts and applications for genomic epidemiology to aid integration for surveillance and infectious disease programs is a key component of advancing the use of genomics for disease control and prevention. This outreach could include specific disease programs within the health department, to local health department epidemiologists and environmental health, and additionally to laboratorians to build the workforce that will prepare and interpret reports for partners.

### 1. Training design

- Develop needs assessment questions and documents.
- Set up reasonable overall training goals and timeline based on the evaluation of a needs assessment.
- Integrate principles of adult learning for use in designing training.
- Collaborate with subject matter experts to gather content, focusing on use-case examples.

### 2. Training content development

- Identify necessary training topics and tools or software based on the learning objectives.
- Assemble instructional materials including knowledge principles, data, online resource, demo, etc.
- Set up needed resources (i.e., facilities, computers, software, software licenses, etc.) prior to the training.
- Integrate multiple types of training materials into training design.
- Tailor training topics and materials to the knowledge differences of various learners.

### 3. Training implementation

- Apply principles of learning to training implementation and delivery.
- Use the most effective presentation tools and techniques.
- Answer training-related questions raised by learners in a timely manner.
- Use examples, exercise, questions, etc., to increase the knowledge absorbance of learners.

### 4. Training evaluation

- Develop tools and questions to evaluate the overall training, as well as learner knowledge and skill development. Integrate training goal assessment into the training evaluation.
- Track and compile evaluation data into summative training reports.
- Evaluate and analyze training reports. Summarize the participants' achievement of training objectives and the training's overall strengths and weaknesses.
- Develop an improvement plan based on results of the training report evaluation.

## Communication (20%)

### 1. Communication techniques

- Deploy formal written and oral communication strategies.
- Apply logical structure to written communications.
- Tailor and translate epidemiologic discipline-specific terms, concepts, results, etc., into accessible information for non-epidemiologic audiences in both written and oral communications.
- Consume questions, requests and other communication from non-epidemiologic audiences and transfer the information into actionable items.

### 2. Communication technology

- Utilize technology to communicate information to internal and external partners.
- Use designated technology for sharing information.
- Revise and improve communication technology when needed for maximum communication efficiency.

### 3. Communication professionalism

- Display professional demeanor in all situations with professional peers.
- Assist with process improvement for communication of non-clinical (non-CLIA) sequencing results, as well as routing sequencing requests, between the laboratory and health department investigators.
- Interpret existing data sharing and public release policies, as well as develop new procedures regarding the use of genomic data for epidemiologic purposes, with supervisory oversight.
- Determine authorized and select essential information to share based on communication recipients.

### 4. Prepare professional written reports

- Prepare memos and internal reports as needed.
- Help create abstracts and scientific manuscripts, oral presentations for meetings, conferences and other training materials.

## Qualifications

### Minimum Qualifications

- Bachelor's degree plus two years of relevant experience or advanced degree in Epidemiology, Public Health, Molecular Biology, Microbiology or a related field.
- Experience with database management systems, e.g., population-based disease metadata, electronic laboratory results.
- Familiarity with next generation sequencing technologies and other molecular testing methodologies.
- Knowledge of bioinformatics software packages, bioinformatics analysis and interpretation of sequencing results.
- Experience preparing and presenting highly technical training and communicating scientific information to audiences of varied professions.
- Excellent organization, oral and written communication skills.

## Preferred Qualifications

- Advanced degree in Epidemiology, Public Health, Molecular Biology, Microbiology or related field.
- Experience with public health surveillance for a variety of pathogens based on public health genomic data.
- Experience with phylogenetic analysis, database use and management of sequencing and bioinformatics data
- Experience creating reports to summarize and visually represent genomic surveillance data.
- Experience with advanced computing resources (e.g. high-performance computing).
- Experience working with existing genomic epidemiology tools (e.g., MicrobeTrace, Nextstrain, Microreact, etc.).
- Experience working in the command-line and proficiency in programming languages.