

## Frequently Asked Questions about PulseNet

### Definitions

- **PFGE:** Pulsed-field gel electrophoresis—the laboratory method used to create DNA “fingerprints” for analyzation by PulseNet personnel
- **Molecular Epidemiology:** the application of molecular biology to epidemiologic investigations. In bacterial foodborne disease investigations, molecular epidemiology = bacterial subtyping using a molecular method, e.g. PFGE
- **PulseNet Laboratory:** a laboratory that performs PFGE using the approved CDC protocols and receives support funding via the Epidemiology and Laboratory Capacity (ELC) or Emerging Infections Program (EIP) federal grants
- **PulseNet Area Laboratory:** Laboratory, designated by CDC, which has agreed to assume responsibility for additional PulseNet duties for laboratories within their support zone. The current Area laboratories include MA, MI, MN, TX, UT, VA, WA, and CDC
- **Isolate:** unique culture from a single person, animal, food, or the environment; one patient may have several isolates, and one isolate may have up to three PulseNet PFGE patterns (see below) associated with it.
- **Restriction Endonuclease:** an enzyme that is used to “cut” the DNA in specified places. These enzymes are used in PulseNet to produce the PFGE patterns. The enzymes associated with each organism are listed below:

Enzyme Order*	<i>E. coli, Salmonella, Shigella</i>	<i>Listeria</i>	<i>Campylobacter</i>
Primary	<i>XbaI</i>	<i>AscI</i>	<i>SmaI</i>
Secondary	<i>BlnI</i>	<i>ApaI</i>	<i>KpnI</i>
Tertiary	<i>SpeI</i>	<i>SmaI</i>	

\*enzyme order is the order in which enzymes are run

- **Real-time subtyping/submission:** the processing by PFGE, analysis, and submission of PulseNet USA tracked isolates to the National databases within 48 hours of their receipt by the participating lab. The PulseNet USA Task Force and PulseNet USA Steering Committee recommend that all isolates of *Listeria* and *E. coli* O157:H7 be PFGE subtyped in “real-time.”
- **PFGE pattern:** the DNA “fingerprint” that is produced by running an isolate through PFGE. An isolate can be restricted by up to three enzymes via the PulseNet protocols.
- **PulseNet Pattern name:** a designation given to each unique PFGE pattern by PulseNet Central. The format is as follows: XXXYYY.##### where XXX is the 3-character code for the serotype; YYY is the 3-character code for the enzyme used; ##### is the unique number given by PulseNet Central. In no way is this number associated with time, frequency or genetic relatedness.
- **Automated Pattern Naming:** a script that constantly runs in the PulseNet databases that uses an algorithm to name patterns. The pattern names that are given either have “@” or “&” after their name. These pattern names are used as a tool of the database team and are **NEVER** be reported.
- **PulseNet Central:** the PulseNet Database Administration Team located at CDC. Each organism has its own database and each person working at *PulseNet Central* is assigned to a particular database.
- **BioNumerics:** the image analysis software used by PulseNet participants to compare PFGE patterns.
- **Dendrogram:** a branching diagram representing a hierarchy of categories based on degree of similarity or number of shared characteristics. In PulseNet, dendrograms are produced by the BioNumerics software to exhibit the statistical similarity between patterns.
- **PulseNet Database:** a database of PFGE patterns produced and analyzed by using PulseNet standardized protocols and customized PulseNet scripts
- **PulseNet Client Scripts:** program that sets up BioNumerics database(s) in a PulseNet-compatible format.

- **Local PulseNet Database:** a PulseNet database maintained locally within the PulseNet participating lab
- **National PulseNet Database:** a PulseNet database housed and maintained at CDC, Atlanta
- **Match:** (n.) 2 or more patterns that are indistinguishable to one another; a “2-enzyme match” would be 2 or more isolates that have patterns that are indistinguishable by the primary and secondary enzymes. (v.) the unofficial word for “is indistinguishable from;” e.g. the pattern from Isolate A “matches” the pattern from Isolate B.
- **Cluster Search:** a search for matching PFGE patterns uploaded in the past 60 (*E. coli*, *Salmonella*, *Shigella*, and *Campylobacter*) or 120 days (*Listeria*). The PFGE patterns are arranged in a UPGMA dendrogram based on the Dice (band based) coefficient using cluster analysis tools provided in the BioNumerics software.
- **PulseNet Cluster:** a group of isolates with the same serotype with indistinguishable PFGE DNA patterns by one or more restriction enzyme
- **Local PulseNet Cluster:** a cluster of matching PFGE patterns found within a 60 or 120-day window found within one state or geographical area
- **Multi-state PulseNet Cluster:** cluster of matching PFGE patterns found within a 60 or 120-day window found in more than one state
- **PulseNet Listserv (WebBoard):** a closed, unmoderated, electronic web conference used for communication between PulseNet participants. The PulseNet Listserv is open to all laboratory staff at PulseNet participating laboratories. Epidemiologists at these laboratories, when approved by the PulseNet laboratory contact, US food regulatory staff, and PulseNet International representatives may also have access.
- **QA/QC:** Quality Assurance and Quality Control
- **Analysis-certified:** an individual who is certified in BioNumerics gel analysis
- **Gel-certified:** an individual or laboratory that is certified in laboratory methods for PFGE and image acquisition

### Detecting and Following Clusters Locally

- Q What role do state and local public health laboratories play in PulseNet?
- A State and local public health laboratories perform real-time subtyping (PFGE), analysis, and submission of PulseNet patterns using PulseNet standardized protocols for molecular subtyping and data analysis. If the labs are unable to perform PFGE on the isolates in a timely manner, they are expected to submit them to either CDC or their area laboratory. PulseNet participating laboratories are also expected to alert PulseNet Central of possible clusters, head up or assist in cluster/outbreak investigations, and respond to the PulseNet listserv within 48 hours of initial postings.
- Q How does a PulseNet lab detect a cluster in their location?
- A For each organism, a PulseNet lab can perform a local cluster search every time they analyze data and input new patterns into their local PulseNet database. However, this may be something that is done on a weekly basis, depending on the receipt of isolates and local schedule of the lab. Once they have found 2 or more patterns that match in their local database, it is their duty to verify the band markings and make sure the dates fall in line with the 60 or 120-day criteria.
- Q Can a PulseNet lab detect matching patterns in other states, or a multi-state cluster?
- A Yes. If a PulseNet lab is analysis-certified they have the ability to connect to the national PulseNet databases located at CDC and see if their local patterns match other patterns from other states. If they are not analysis-certified, they may still use the information provided in PulseNet listserv postings by other labs to see if their patterns match on-going clusters.
- Q Once a cluster is detected by a PulseNet lab, who is contacted?
- A The PulseNet lab should contact PulseNet Central and post their findings to the PulseNet listserv. As for who the lab contacts within their state, this is up to their “in-house” procedures.
- Q What parameters indicate that a PulseNet lab should post a cluster to the PulseNet listserv?

- A For organisms like *E. coli* O157:H7 and *Listeria*, where time is of the essence, all clusters of 2+ are to be reported to the PulseNet listserv as soon as possible. With other organisms (*Salmonella*, *Shigella*, *Campylobacter*), additional information may be required (i.e. commonality of serotype, commonality of pattern, background epi information). Regardless, it is the laboratory's or PulseNet Central's decision whether or not a cluster should be reported.
- Q Can state and local laboratories give and report official pattern names?
- A No. Only personnel at PulseNet Central at CDC can name and report official pattern names. The state and local labs can designate the patterns a local name, but the CDC is responsible for assigning official names.

### Detecting Clusters at the National Level

- Q How does PulseNet Central detect possible clusters?
- A The database manager at CDC performs a search in the respective national database and all patterns that have been submitted within the appropriate number of days are reviewed. The database manager uses the BioNumerics software to perform a cluster analysis of the first enzyme data and create a dendrogram. Patterns that cluster together are investigated further.
- Q What are the limitations that keep PulseNet Central from detecting **all** possible clusters of foodborne bacteria?
- A **Lack of real-time submission by participating labs:** if participating PulseNet labs are unable to submit their data in real-time (shortage in staff, lack of resources, lack of cooperation from local hospitals, etc.) this will keep PulseNet Central from knowing about clusters. PulseNet Central cannot detect clusters for which the data is not present. In these cases, reports from local epidemiologists are usually the source of information about potential clusters or outbreaks.
- A **Incorrect or missing data:** if data has been submitted with incorrect or missing dates or missing information such as serotype, this can hinder a cluster being detected.
- A **Poor or incorrect band marking:** if patterns have been submitted without following the PulseNet guidelines for band marking, patterns may not cluster. Although the BioNumerics software allows for quite a bit of automation, it is ultimately up to the human eye to decide and mark bands. The marking of these bands is used by the software to perform cluster analyses.
- A **Lack of submissions of all isolates:** if laboratories only submit a sampling of their PFGE data, the numbers submitted may not be enough to warrant an investigation by PulseNet Central. Submission of 1 or 2 patterns in a cluster search involving 2,000 patterns (average submission in 60 days in the *Salmonella* database) will not be enough for further analysis.
- A **Humans detect clusters—there is no automated system:** although there has been work done to create an “automated” cluster detection system, the work is far from done. Therefore, PulseNet Central must still rely on manual cluster searches performed by humans. Although it is rare, there are clusters that do not get detected by PulseNet Central. It is because of this, that PulseNet Central relies on its partnership with its state and local public health participants to perform their duties of timely submissions, local cluster searches, and postings to the PulseNet listserv.

### PulseNet Central: from Detecting Clusters to Cluster Analysis

- Q Once a cluster is detected by PulseNet Central, what are the next steps?
- A **Verify Band Assignments:** because the bands of each pattern are still marked and verified using the human eye, and because there are over 100 individuals possibly marking bands and submitting results, not all patterns will be marked the same. It is the duty of PulseNet Central to verify that all band markings are correct, especially if a cluster is suspected.
- A **Check dates:** although PulseNet Central requests that Isolation Date (sometimes collection date) and Received Date (date the lab receives the isolate in the lab) be entered into the database, this information is often not available. Without this information, PulseNet Central must rely on the Upload Date. This date is automatically filled when the patterns and corresponding information are uploaded to the database. Just because something was uploaded on a date that

falls within the 60-day window does not mean it is recent. For example, something with an upload date of May 2, 2005 could have a Received Date of February 2, 2005. Therefore, this pattern would fall outside of the 60-day window.

- A **Name Patterns:** in order to check for frequency in the database and to see if the cluster pattern has been seen in the database before, PulseNet Central must give the pattern a name. This process is discussed further in “Other Reports and Requests” below.
- A **Determine Pattern Frequency in Database:** once a pattern has been named, PulseNet Central determines the number of times this pattern has been seen in the database and compares this frequency to the top 10 patterns in the database.
- A **Name Cluster:** once it is determined that a PulseNet cluster has indeed been detected, the cluster is given a name. The nomenclature is as follows: YYMMLLSSS-#xxx where YY and MM are the 2-digit year and 2-digit month in which the cluster was detected; LL is the LabID where the cluster was first seen (ML is used for multi-state); SSS is the serotype (for *Salmonella* and *Shigella*); # is the cluster number that occurred in that location in that time; xxx includes other information about the cluster and is not always present (ml for multi-state, c for cluster, etc.).  
**Example: 0405MAJEG-2** is the second *Salmonella* Enteritidis outbreak in Massachusetts, first seen in May 2004.

### PulseNet Central: Reporting Clusters

- Q What parameters indicate that the cluster be posted to the PulseNet listserv?
  - A **Commonality of Serotype/Pattern:** if there is an increase in the number of isolates from a particular serotype that is less common (a serotype that is not among the 10 most common serotypes in *Salmonella*, for example), this cluster information is sent to the appropriate people and posted to the PulseNet listserv. For *E. coli* or *Shigella*, this could be anything other than *E. coli* O157:H7, or *Shigella sonnei*. For *E. coli* O157:H7 and *Listeria*, commonality of pattern is not as important—most all clusters of 3+ will get reported.
  - A **Epidemiological Information (if known):** If epidemiological information is provided with the isolates, this information can be used to preliminarily decide if a cluster should be reported. For *S. sonnei*, for example, if all the matching patterns are from people older than 5 years of age, this may indicate something more than a daycare outbreak.
- Q Who is contacted when a cluster is detected by PulseNet Central?
  - A The labs that are involved in the cluster are first sent an email, usually with a report that has been put into Excel. The PulseNet liaison is sent a similar email to distribute to the appropriate epidemiologist(s). As the cluster/outbreak progresses, the database manager will contact all involved parties (lab, state epi, CDC epi, USDA, etc.) directly.
- Q How can we determine if a certain number of “matches” is significant for a particular organism (what does it mean if a pattern is “common” or “uncommon”)?
  - A PulseNet Central currently creates quarterly reports that list the most “common” serotypes and their patterns. These percentages are often very low (<5%), due to the diversity of PFGE patterns. When the frequency of a pattern is determined, it is compared to the percentages of the most common patterns. There are some patterns in every database that are known to appear sporadically throughout the US, all during the year. These are not usually posted to the PulseNet listserv. PulseNet Central is currently working on a program that will look at the number of weekly submissions for a pattern up to 2 years prior to the current date. This will allow for a more accurate determination of a true “rise” in the occurrence of a pattern.

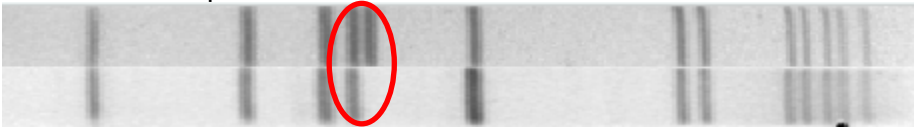
### PulseNet Central: Other Reports and Requests

- Q How long does it take to name a pattern?
  - A The time it takes to name a pattern differs for each database; however the process is the same for each database. The following example is naming of a pattern from a cluster search for the *Salmonella* database (the largest).

1. Open Database: 30 sec	2. 60-day cluster search: 20 sec	3. Organize by serotype: 1 sec
4. Dendrogram: 4 min, 20 sec	5. Open images: 4 min, 40 sec	6. ID possible clusters: 15 min
7. Edit poor images: 5 min	8. Check/fix band markings: 30 min. (this can take some time due to many images that may have to be edited)	9. Pick out representative pattern from cluster: 30 sec
10. Compare to Unique Pattern List: 5 min	11. Name/confirm pattern: 2 min	12. <b>Average total time: ~1 hr, 12 min</b>

Q What is a one-band difference?

A This is an example of a one-band difference:



A

Q When looking at a dendrogram of a group of PFGE patterns, why aren't one-band differences or patterns with a high percentage of similarity (~90%) included in initial cluster reports? Does this vary by organism or enzyme?

A One-band differences are so common in large databases that virtually all submissions would be part of a cluster. Because most of the isolates included in the PulseNet databases are due to foodborne and not person-to-person transmission, genetic mutations are not expected.

A **Exceptions:** Because many shigellosis outbreaks are due to person-to-person contact, isolates with similar patterns may be included. During the course of an outbreak investigation, epidemiology may keep someone in the investigation that has a close, but distinguishable pattern to the outbreak pattern.

Q When should 2<sup>nd</sup> enzyme be run?

A 2<sup>nd</sup> enzyme is routinely performed for *Listeria* and most *E. coli* isolates. During an outbreak investigation, 2<sup>nd</sup> enzyme information can be useful to increase the discriminatory power of PFGE, confirm or separate results obtained with primary enzyme, and can be used in light of limited epidemiological data.

Q How are things reported if 2<sup>nd</sup> enzyme matches and 1<sup>st</sup> enzyme does not?

A PulseNet Central does not further investigate those clusters that only match by 2<sup>nd</sup> enzyme. The criterion is to look at 1<sup>st</sup> enzyme results and then 2<sup>nd</sup> enzyme.

Q What does it mean when a known outbreak is associated with more than one pattern?

A **Poor quality gels:** this can prevent the clear analysis of a pattern and therefore it is unclear if the patterns are truly indistinguishable. In this case, the isolates are usually sent to CDC for rerunning.

A **Multiple strain contamination:** some sources can be contaminated with multiple strains

A **Possible strain mutation:** this can occur during person-to-person passage

Q Can PFGE patterns be used to determine serotype?

A No. However, PFGE can be used to determine that an isolate is **not** a specific serotype. Certain serotypes tend to have certain types of patterns (i.e. bands falling within certain parts of the gel)

### Quality Assurance and Quality Control (QA/QC)

Q How can epidemiologists know that their lab produces quality PulseNet data?

A Most all new lab personnel are trained either at CDC or by approved PulseNet personnel in the running of PFGE and analysis of images in BioNumerics. This training is tested by a stringent certification program and is maintained through annual proficiency testing. If a lab is having problems, they can receive troubleshooting assistance from their area lab and/or CDC.

A PulseNet Central has also put together a QA/QC Manual with Standard Operating Procedures (SOPs) to be used by the participating laboratories.

## **PulseNet Database Manager Duties**

### **I. Managing the database**

- Analyze gels that are received from public health laboratories, USDA, FDA, and international affiliates. (labs that are not certified send gels & information to CDC to be analyzed & uploaded)
- Compare & name patterns (pattern confirmation is only done at CDC)
- Conduct national active surveillance of bacterial foodborne and diarrheal disease by detecting clusters
- Database maintenance (delete duplications, manually input info from older gels, confirm pattern names)
- Maintain national unique pattern list (running list of unique patterns that must be updated daily)
- Extensively test any new software or script developments
- Add/maintain additional serotypes and serotype codes to the Master Scripts and National databases

### **II. Working with other PulseNet state labs**

- Complete requests made by PulseNet participants
- Request isolates to be sent to CDC
- Request 2<sup>nd</sup> enzyme runs for all necessary isolates
- Provide support to other labs (troubleshoot gels and database issues)

### **III. Working with epidemiologists**

- Report significant clusters to the CDC epidemiologists
- Alert epis to new matches within a cluster; produce weekly cluster log
- Provide epis with updated spreadsheets with demographic isolate information for clusters
- Communicate with epis about patterns related to clusters (give frequencies, histories, etc)
- Develop improved communication tools with epi section (isolate request forms, etc.)
- Respond to epi requests/Assist with manuscripts

### **IV. Communication using PulseNet listserv**

- Maintain PulseNet listserv conferences by posting responses (conferences relay cluster information)
- Provide summaries/updates/responses to clusters by posting to PulseNet listserv
- Delete/archive old PulseNet listserv conferences

### **V. Training & Proficiency Testing**

- Organize and conduct at least biannual trainings to teach PulseNet participants from USDA, FDA, public health laboratories, and international laboratories ways to successfully develop PFGE gels, analyze PFGE gels, upload PFGE gels, and detect potential clusters
- Visit national and international sites throughout the year for trainings, meetings, and presentations
- Review/Sign lab certifications and PT test results
- Organize & lead training courses for lab protocols and database software
- Conduct training for new EISO officers and new staff dealing with CDC Epidemiology department

### **VI. Additional activities**

- Complete Freedom of Information Act (FOIA) requests
- Write Standard Operating Procedures (SOPs) for all duties related to participating PulseNet laboratories
- Prepare presentations, abstracts, publications, and posters on PulseNet and BioNumerics for annual update meeting and professional conferences
- Participate in outside projects (collaborate with groups such as NARMS, FOODNET, etc)
- Provide regular demos to visitors to CDC (professors, government officials, students, etc)
- Review & discuss Epidemiology and Laboratory Capacity (ELC) grant requests
- Produce PulseNet newsletter (4 issues a year)