

PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 1 of 28

1. **PURPOSE:** Genomic sequences generated by PulseNet participating laboratories for PulseNet surveillance are processed using the BioNumerics software and a two-phase analysis approach utilizing the RefID database (PND20) and the organism-specific databases (PND05). Fully analyzed sequences can be used for local and national cluster detection. The purpose of this document is to outline the standardized procedures for local cluster detection in the organism-specific databases and querying and matching against the national database.
2. **SCOPE:** This procedure applies to whole genome sequence data generated by PulseNet participating laboratories for PulseNet surveillance.
3. **TABLE OF CONTENTS – HYPERLINKS TO THE PROCEDURE**
 - 3.1. [Local Cluster Detection Using a 60- Or 120-Day cgMLST Dendrogram](#)
 - 3.2. [Local Cluster Detection Using Fast Character Matching \(FCM\) – Compare a Sequence Or a Group of Sequences Against the Entire Local Database](#)
 - 3.3. [Local Cluster Detection Based on Allele Codes](#)
 - 3.4. [Query the National Database and Download Bundles](#)
 - 3.5. [Match Against Server](#)
 - 3.6. [Generate Reports](#)
 - 3.7. [Appendix PND06-1. PulseNet Settings for Creating cgMLST or wgMLST Dendrograms](#)
 - 3.8. [Appendix PND06-2. Instructions to Customize the View in the Local Database](#)
4. **DEFINITIONS/ACRONYMS:**
 - 4.1. **7-gene MLST:** Multi-Locus Sequence Typing scheme based on seven conserved housekeeping genes.
 - 4.2. **Allele:** One of two or more alternative forms of a gene that arise by mutation and are found at the same place on a chromosome.
 - 4.3. **Allele code:** A downloadable sequence type name, assigned in the national database, that is based on comparing cgMLST profiles using single linkage clustering and hierarchical similarity thresholds. Composed of multiple digits, e.g., LMO1.0 – 5.1.1.2.5.1 where each digit corresponds to a specific similarity threshold (allele difference). The more digits two sequences have in common from left to right, the more closely related they are.
 - 4.4. **Analysis Certified:** An individual who is certified for checking the quality, performing analysis and uploading WGS data to the PulseNet National Database and NCBI using BioNumerics.
 - 4.5. **BioNumerics:** Analysis software used by PulseNet, developed by Applied Maths (Sint-Martens-Latem, Belgium).
 - 4.6. **CDC:** Centers for Disease Control and Prevention.
 - 4.7. **cgMLST:** Core Genome Multi-Locus Sequence Typing
 - 4.8. **Cluster (outbreak) code:** A downloadable name assigned to a detected cluster of cases in the national database that populates the “Outbreak” field. Follows the format *year-month-location-organism-cluster number*. For example: 2005GAEXH-1 was the first STEC O157 cluster detected in Georgia in May 2020; 2005MLJPX-2 was the second multi-state *Salmonella* Typhimurium cluster detected in May 2020.

PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 2 of 28

- 4.9. **Core genome:** Genes shared by a vast majority (e.g., >99% of strains *Listeria monocytogenes* strains) of strains of the same species.
- 4.10. **FCM:** Fast Character Match. A database screening tool in BioNumerics that can be used to quickly search for matches in the local or national database for the sequence(s) of interest.
- 4.11. **Organism-specific Database:** A BioNumerics database, v 7.6 or higher, used for comparing isolate sequences for surveillance. Part of the standard PulseNet workflow.
- 4.12. **PN:** PulseNet.
- 4.13. **PulseNet Central:** PulseNet team at CDC comprising of the Database Unit (PulseNet@cdc.gov) and the Next Generation Subtyping Methods Unit (PulseNetNGSlab@cdc.gov).
- 4.14. **QC:** Quality Control.
- 4.15. **RefID Database:** A BioNumerics database, v 7.6 or higher, used for quality control of raw sequence data, assembly of sequences, contamination detection and species identification. Part of the standard PulseNet workflow.
- 4.16. **SOP:** Standard Operating Procedure.
- 4.17. **WGS:** Whole Genome Sequencing.
- 4.18. **wgMLST:** Whole Genome Multi-Locus Sequencing Typing.

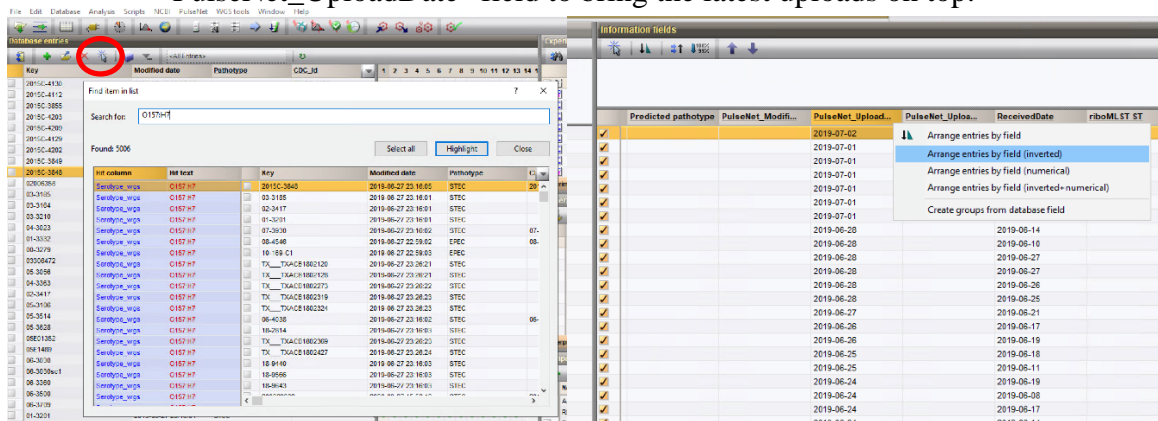
5. RESPONSIBILITIES:

- 5.1. Analysis certified PulseNet public health laboratory personnel perform local cluster detection and matching and querying against the national database using the BioNumerics 7.6 or higher organism-specific surveillance databases. The detected clusters are reported through appropriate channels.
- 5.2. PulseNet Central Database Unit personnel perform cluster detection at the national level and report clusters through appropriate channels.

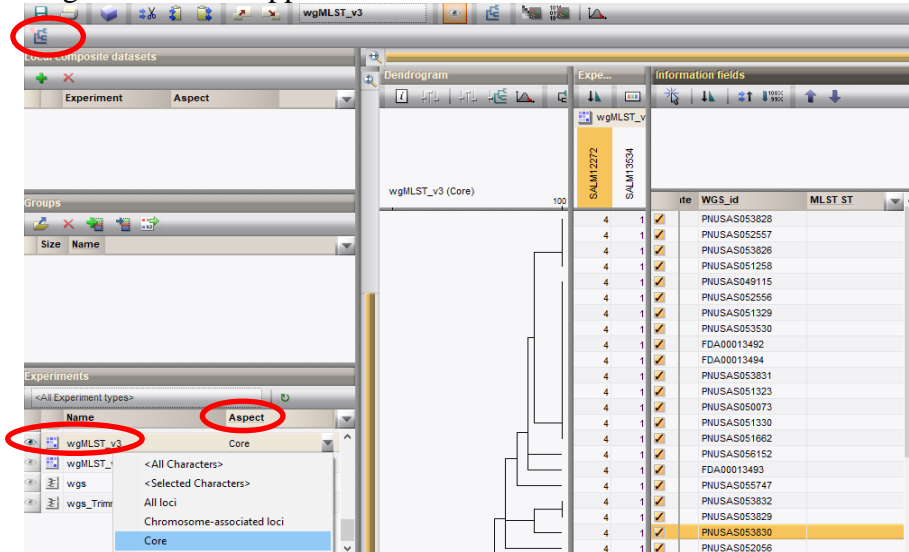
6. PROCEDURE:

6.1. Local Cluster Detection Using a 60- Or 120-Day cgMLST Dendrogram.

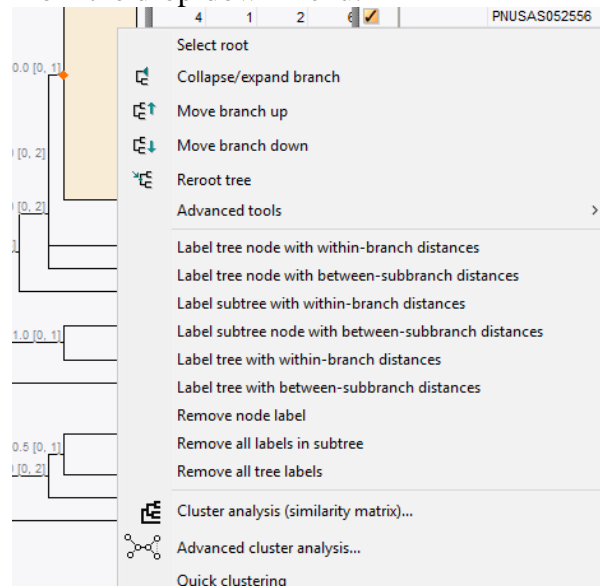
6.1.1. Find and select the relevant entries by querying the database using the “Find entries in list” icon and/or by arranging the “PulseNet_UploadDate” field. For example, you can find all entries of a certain serotype by using the “Find entries in list” functionality, select and open them in a comparison and then arrange the “PulseNet_UploadDate” field to bring the latest uploads on top.



- 6.1.2. Create a comparison by clicking on the green plus in the “Comparisons” panel in the main BioNumerics window or pressing alt+c on the keyboard. The “Comparison” window will open.
- 6.1.3. In the “Experiments” panel in the “Comparison” window, choose “Core” or “Core loci” (depending on the organism database) from the “Aspect” drop-down menu. and click on the eye next to “wgMLST”.
- 6.1.4. Click on the “Cluster analysis using current settings” icon to create a cgMLST dendrogram. **NOTE:** this quick cluster tool creates a dendrogram using the PulseNet settings outlined in Appendix PND06-1.

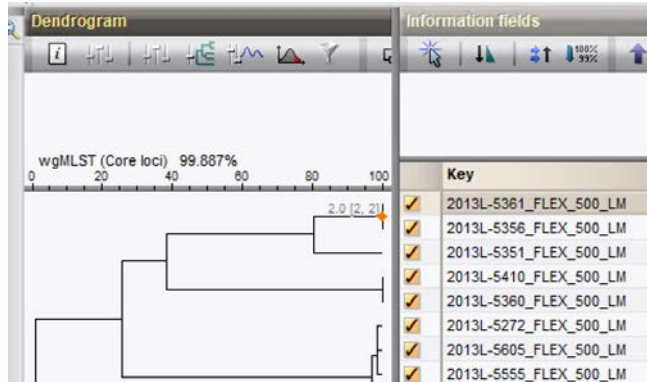


- 6.1.5. To view allele differences, right-click on the appropriate node and select the desired option from the drop-down menu.

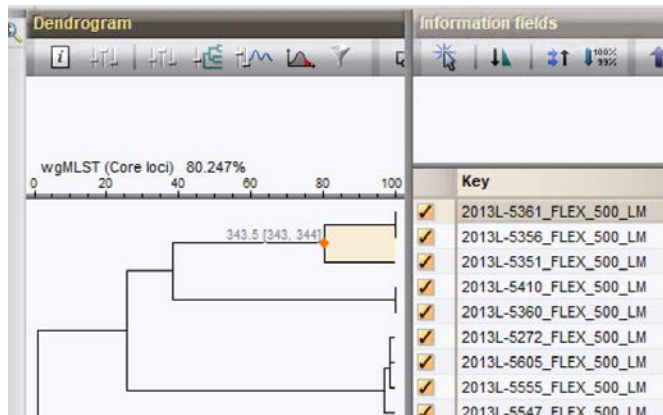


- 6.1.5.1. The following labeling options are most useful:

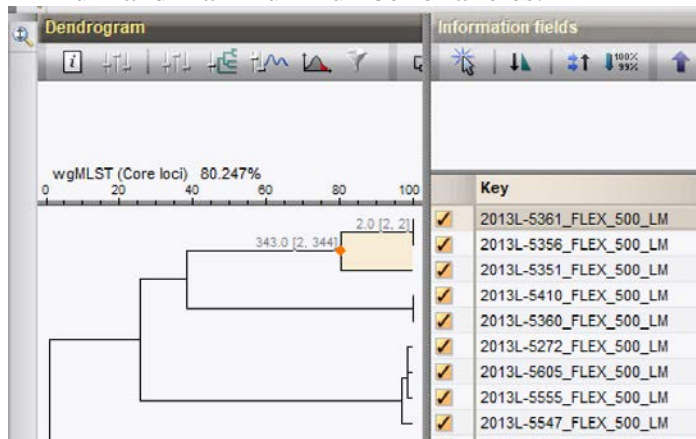
6.1.5.1.1. “Label tree node with within-branch distances”: labels the node with the median allele difference and the range for the minimum and maximum number of alleles detected within node.



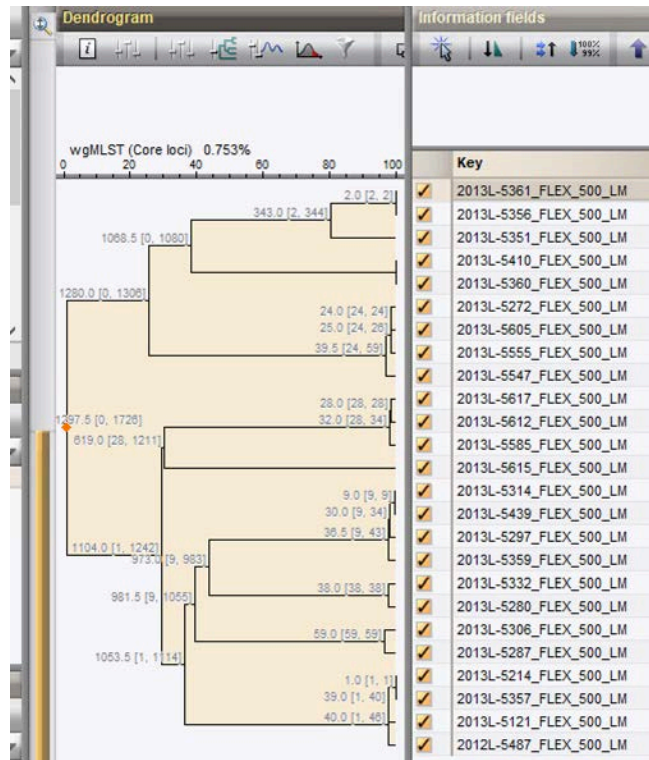
6.1.5.1.2. “Label tree node with between-subbranch distances”: labels the node with multiple subbranches with the median allele difference and the range for the minimum and maximum number of alleles detected between the subbranches within the node.



6.1.5.1.3. “Label subtree with within-branch distances”: labels each branch of the subtree with the median allele difference and the range for the minimum and maximum number of alleles.



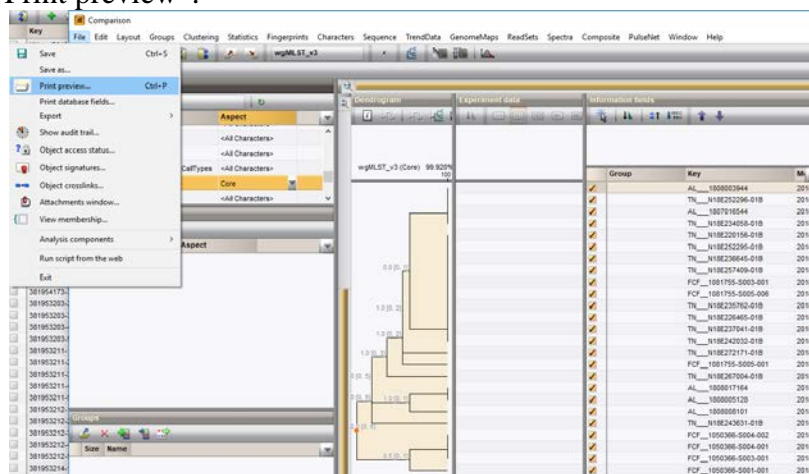
6.1.5.1.4. “Label tree with within-branch distances”: labels each branch in the tree with the median allele difference and the range for the minimum and maximum number of alleles.



6.1.6. Exporting a dendrogram.

6.1.6.1. If the allele calls are shown, turn off the allele calls by clicking on the eye next to the “wgMLST” experiment.

6.1.6.2. In the “Comparison” window from the “File” drop-down menu, choose “Print preview”.

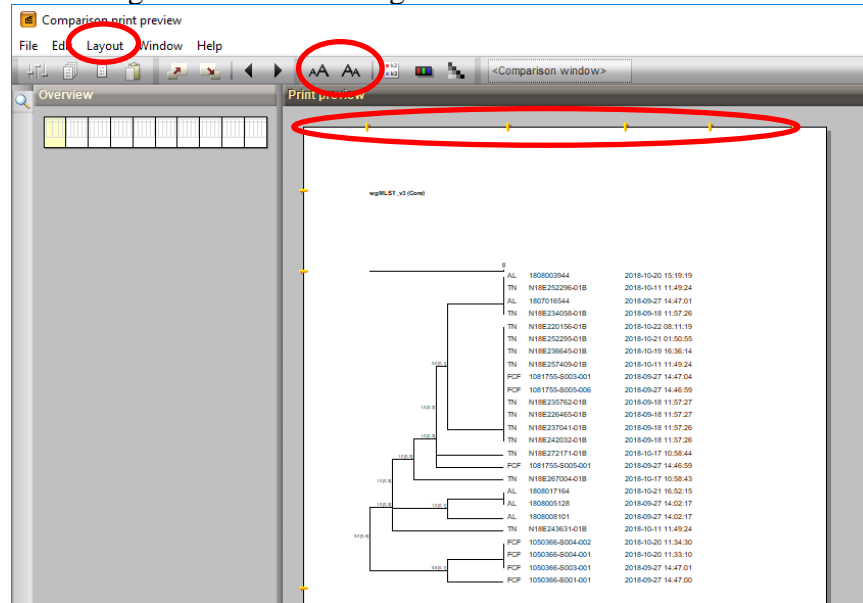


6.1.6.3. In the “Comparison print preview” window, you can customize the view:

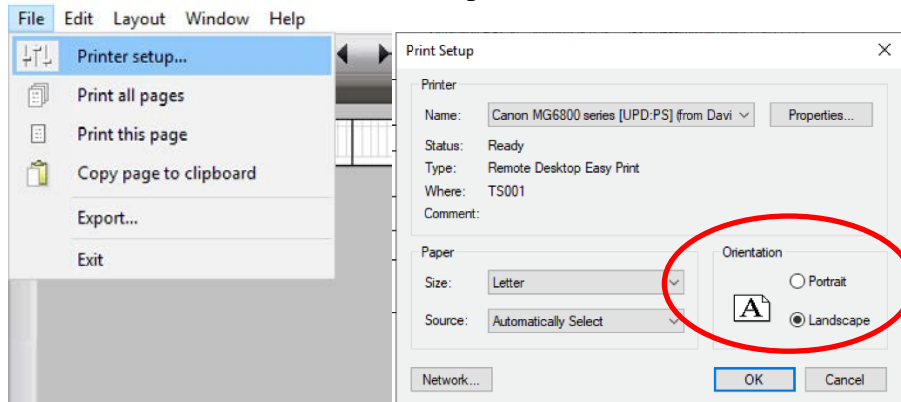
6.1.6.3.1. Move the yellow nodes to hide or display fields.

6.1.6.3.2. Add the field headers by choosing “Show field names” from the “Layout” drop-down menu.

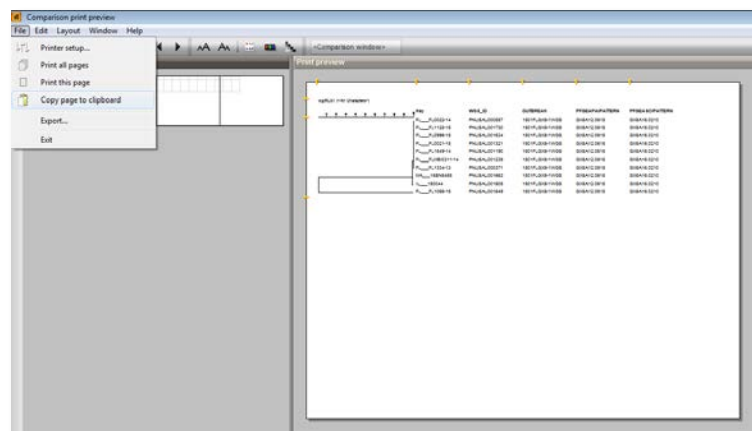
6.1.6.3.3. Enlarge or reduce the image size.



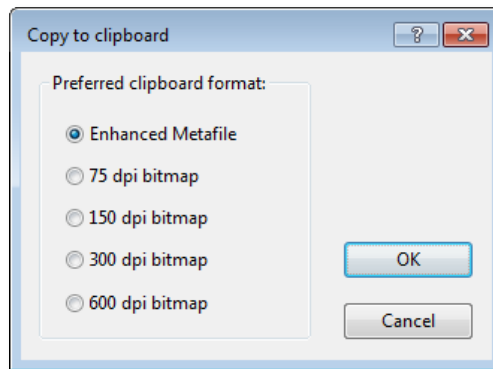
6.1.6.3.4. You can switch to landscape orientation by choosing “Printer setup” from the “File” drop-down menu. In the “Print Setup” window under “Orientation”, select “Landscape”.



6.1.6.4. Once the image is set, choose “Copy page to clipboard” from the “File” drop-down menu.



6.1.6.5. “Copy to clipboard” window will open. Choose “Enhanced metafile” and click “OK”.



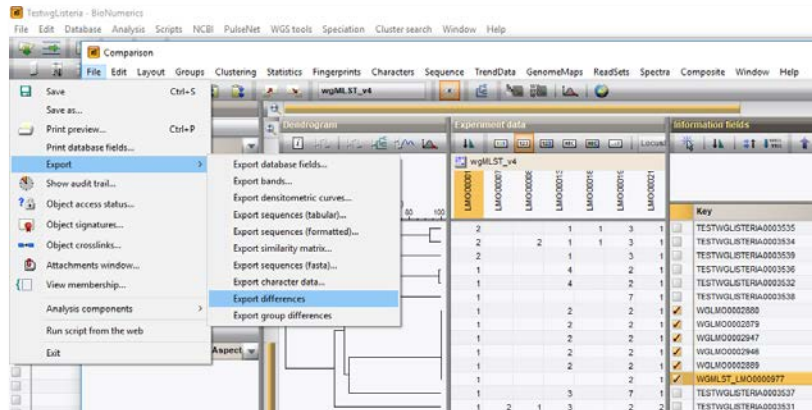
6.1.6.6. The image can now be pasted into PowerPoint or Word.

6.1.6.7. Steps 6.1.6.4. – 6.1.6.6. need to be repeated for each page you wish to copy.

6.1.7. View allele differences entry by entry in a similarity (allele) matrix.

6.1.7.1. In the “Comparison” window, select the entries you want to include in the matrix.

6.1.7.2. From the “File” drop-down menu, choose “Export” and from the appearing drop-down menu, choose “Export differences”.



6.1.7.3. An Excel file with the similarity matrix opens displaying pairwise allele differences between the strains.

| | WGLMO0002879 | WGLMO0002880 | WGLMO0002889 | WGLMO0002946 | WGLMO0002947 | WGLMLST_LMO0000977 |
|--------------------|--------------|--------------|--------------|--------------|--------------|--------------------|
| WGLMO0002879 | 0 | 0 | 1 | 1 | 1 | 4 |
| WGLMO0002880 | 0 | 0 | 1 | 1 | 1 | 4 |
| WGLMO0002889 | 1 | 1 | 0 | 0 | 0 | 5 |
| WGLMO0002946 | 1 | 1 | 0 | 0 | 0 | 5 |
| WGLMO0002947 | 1 | 1 | 0 | 0 | 0 | 5 |
| WGLMLST_LMO0000977 | 4 | 4 | 5 | 5 | 5 | 0 |

6.1.8. View allele differences between clades in a similarity (allele) matrix.

6.1.8.1. In the “Comparison” window, select the entries belonging to the same clade.

6.1.8.2. In the “Groups” panel, click “Create new group from selection” or press ctrl+g on the keyboard and in the appearing “Create new group” window, enter the name of the group, e.g., “Clade 1”.

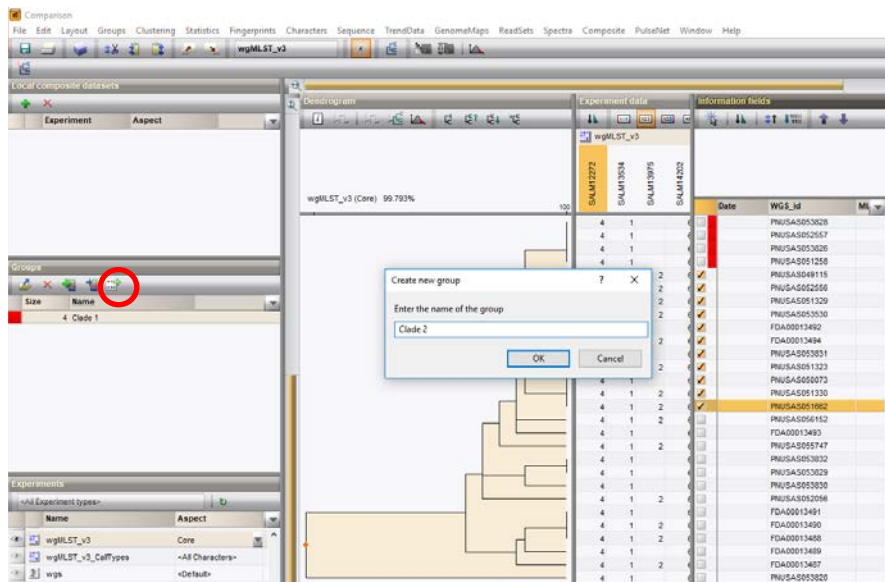
PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

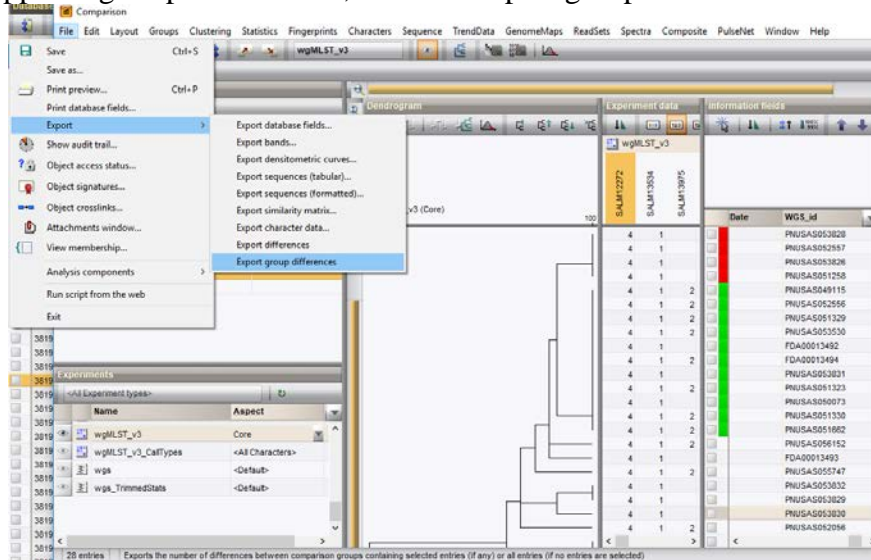
Page 8 of 28



6.1.8.3. Repeat steps 6.1.8.1. and 6.1.8.2. for the remaining clades.

6.1.8.4. Select the entries in the clades of interest.

6.1.8.5. From the “File” drop-down menu, choose “Export” and from the appearing drop-down menu, choose “Export group differences”.



6.1.8.6. An Excel file with the similarity matrix opens displaying the pairwise median allele differences between the clades and the minimum and maximum ranges for each pairwise comparison.

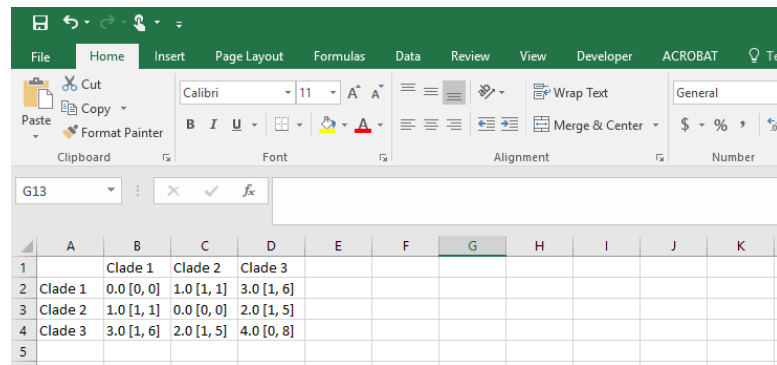
PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 9 of 28



| | A | B | C | D | E | F | G | H | I | J | K |
|---|---------|------------|------------|------------|---|---|---|---|---|---|---|
| 1 | | Clade 1 | Clade 2 | Clade 3 | | | | | | | |
| 2 | Clade 1 | 0.0 [0, 0] | 1.0 [1, 1] | 3.0 [1, 6] | | | | | | | |
| 3 | Clade 2 | 1.0 [1, 1] | 0.0 [0, 0] | 2.0 [1, 5] | | | | | | | |
| 4 | Clade 3 | 3.0 [1, 6] | 2.0 [1, 5] | 4.0 [0, 8] | | | | | | | |
| 5 | | | | | | | | | | | |

6.1.9. Keep species/serotype-specific comparisons saved for the past 60 or 120 days and update them by adding new entries and removing old (past 60/120 days) entries.

6.1.9.1. To save a comparison:

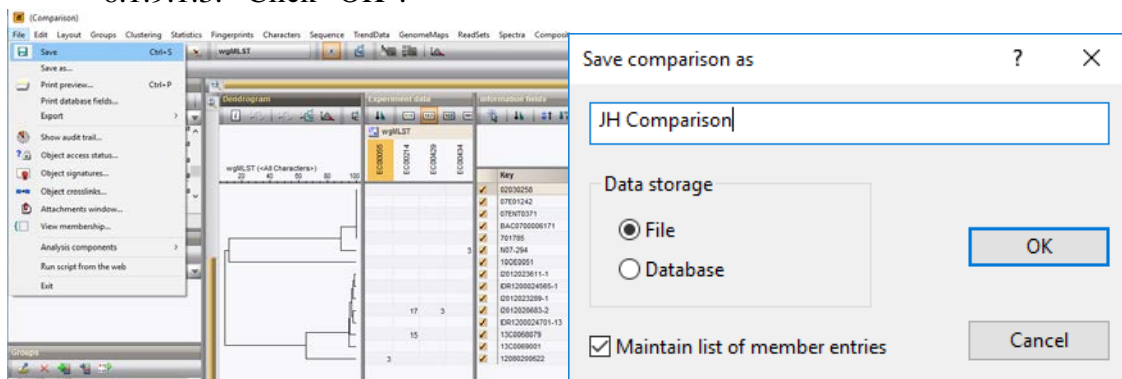
6.1.9.1.1. From the “File” drop-down menu, choose “Save as”. “Save comparison as” window appears.

6.1.9.1.2. Name the comparison and select the data storage option.

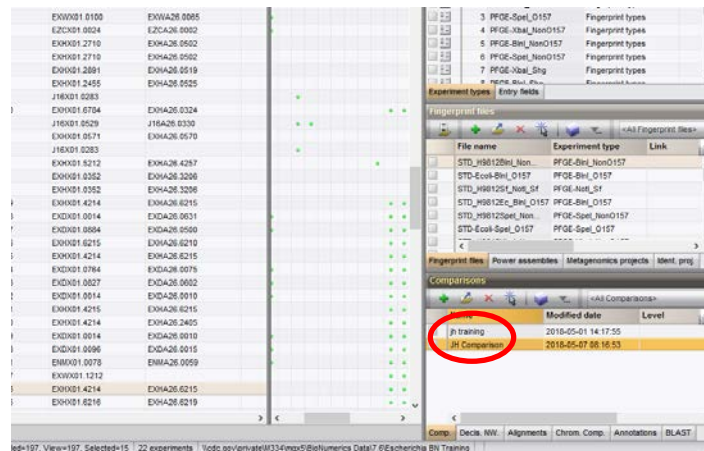
6.1.9.1.2.1. Choosing the “File” option means that only the creator of the comparison can see it and work on it in a shared database.

6.1.9.1.2.2. Choosing the “Database” option will save the comparison within the relational database so other users of the shared database can see it and work on it.

6.1.9.1.3. Click “OK”.



6.1.9.1.4. Saved comparisons are listed in the “Comparisons” panel in the main window.

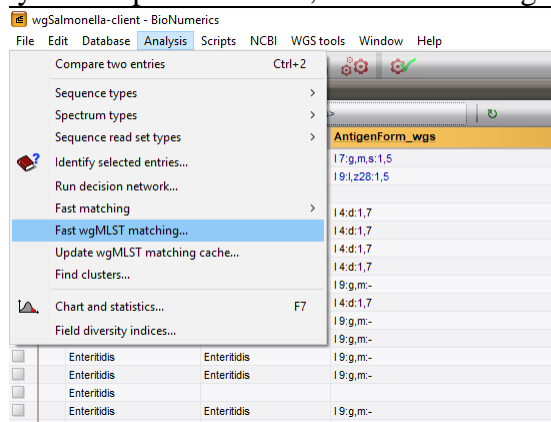


- 6.1.9.2. To add new entries to the comparison:
 - 6.1.9.2.1. Select the entries to be copied to the comparison.
 - 6.1.9.2.2. In the main BioNumerics window, click “Copy selected entries to clipboard”.
 - 6.1.9.2.3. In the target comparison window, click “Paste selection”.
- 6.1.9.3. To remove old entries from the comparison:
 - 6.1.9.3.1. In the comparison window, select entries to be removed.
 - 6.1.9.3.2. Click “Cut selection (ctrl+x)”. A confirmation window appears. Click “Yes”. **NOTE:** the cut entries will remain selected in the database. To de-select, press F4.



6.2. Local Cluster Detection Using Fast Character Matching (FCM) – Compare a Sequence Or a Group of Sequences Against the Entire Local Database.

- 6.2.1. In the main BioNumerics window, select the entries of interest.
- 6.2.2. From the “Analysis” drop-down menu, choose “Fast wgMLST matching”.



- 6.2.3. In the “Fast character matching settings” window, define the parameters:
 - 6.2.3.1. Use the default values for “Input data”:
 - 6.2.3.1.1. Experiment type: wgMLST
 - 6.2.3.1.2. Character view: Core
 - 6.2.3.2. Quality filtering: check the box for “Enforce minimum locus presence of” if you only want to compare against sequences that passed the minimum core percentage threshold. Enter 95.0 for *Listeria* and 85.0 for *Escherichia*, *Salmonella* and *Campylobacter*. Leave the box unchecked if you also want to

include sequences that may not have passed the core percentage present.

NOTE: in an outbreak situation for timeliness of results it may sometimes be necessary to include lower quality sequences to the comparison while waiting for repeat sequencing.

6.2.3.3. The search can be restricted by defining values for “Reference filtering”:

6.2.3.3.1. Check the box for “Use date restriction” if you want to narrow down the search for a specific date range, e.g., the past 60 or 120 days.

6.2.3.3.2. Check the box for “Use field value restriction” if you want to narrow down the search based on a specific field, e.g., serotype or allele code.

6.2.3.4. “Match parameters” – type in the allelic distances you are interested in for primary and secondary matches.

6.2.3.5. “Output” – define how you want to see the results displayed:

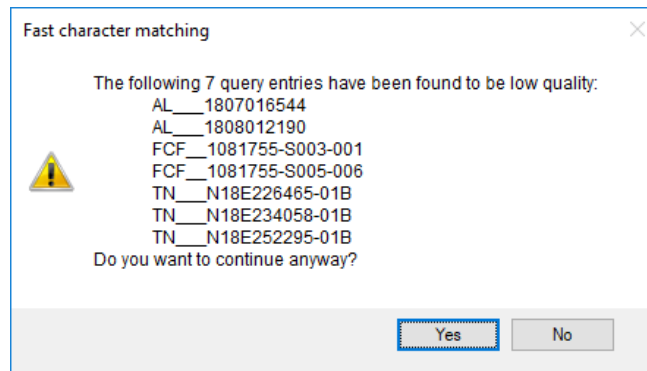
6.2.3.5.1. Update your original selection with either primary hits (entries that are within the primary match threshold defined under “Match parameters”) or all hits (entries that are within the secondary match threshold under “Match parameters”).

6.2.3.5.2. Create a comparison with either primary hits or all hits.

6.2.3.5.3. Open a match window with either primary hits or all hits.

6.2.3.6. Click “OK”.

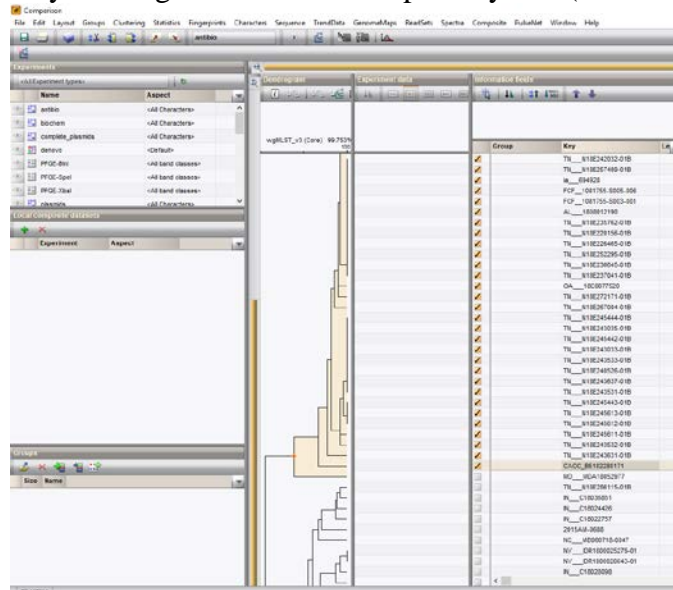
6.2.4. A message will be displayed if entries found during the matching are of poor quality, e.g., percent core <95/85%. Click “Yes” to continue. The poor quality entries will be included in the output.



6.2.5. Viewing the FCM output in a comparison. **NOTE:** this example is based on the FCM settings from the above screenshot.

6.2.5.1. All hits (within 10 alleles) are displayed in the comparison.

6.2.5.2. Only the original selection and primary hits (within 5 alleles) are selected.



6.2.6. Viewing the output in the FCM window. **NOTE:** this example is based on the FCM settings from the above screenshot.

6.2.6.1. The “Entries” panel in the top lists the original entries chosen to be compared.

6.2.6.2. The “Matches” panel in the bottom displays the results of the FCM. All hits (within 10 alleles) are listed but only primary hits (within 5 alleles) are selected. **NOTE:** the original selection is also listed in the bottom panel because they are being matched against themselves.

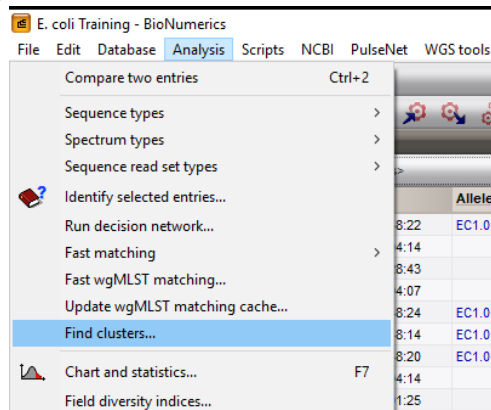
6.2.6.3. The “Distance” field in the bottom panel shows the allele differences between each entry in the top panel vs. the entries in the bottom panel.

| Key | Modified date | cdc_id | AntigenForm_wgs | Serotype_wgs |
|-----------------------|---------------------|--------|-----------------|--------------|
| AL__1807018544 | 2018-09-27 14:47:01 | | | |
| AL__1808012190 | 2018-09-27 14:46:59 | | | |
| FCF__1081755-S003-001 | 2018-09-27 14:47:04 | | | |
| FCF__1081755-S005-006 | 2018-09-27 14:46:59 | | | |
| TN__N18E220156-01B | 2018-10-22 08:11:19 | | | |
| TN__N18E226465-01B | 2018-09-18 11:57:27 | | | |
| TN__N18E234058-01B | 2018-09-18 11:57:26 | | | |
| TN__N18E235762-01B | 2018-09-18 11:57:27 | | | |
| TN__N18E239645-01B | 2018-10-19 16:36:14 | | | |
| TN__N18E252295-01B | 2018-10-21 01:50:55 | | | |
| TN__N18E257409-01B | 2018-10-11 11:49:24 | | | |

| Distance | Key | Modified date | cdc_id | AntigenForm_wgs | Serotype_wgs |
|----------|--------------------|---------------------|--------|-----------------|--------------|
| 2.000 | TN__N18E267004-01B | 2018-10-17 10:58:43 | | | |
| 2.000 | TN__N18E245442-01B | 2018-09-18 11:57:27 | | | |
| 2.000 | TN__N18E243531-01B | 2018-09-18 11:57:26 | | | |
| 2.000 | TN__N18E243033-01B | 2018-09-18 11:57:26 | | | |
| 3.000 | TN__N18E245443-01B | 2018-09-18 11:57:27 | | | |
| 3.000 | TN__N18E245611-01B | 2018-10-21 15:36:50 | | | |
| 3.000 | TN__N18E243532-01B | 2018-09-18 11:57:26 | | | |
| 3.000 | TN__N18E245612-01B | 2018-09-18 11:57:26 | | | |
| 6.000 | MD__MDA18052977 | 2018-10-19 16:53:12 | | | |
| 9.000 | 2015AM-0688 | 2018-11-15 09:36:56 | | | |
| 9.000 | IN__C18036851 | 2018-10-20 14:45:22 | | | |
| 10.000 | WI__18MP005406 | 2018-10-11 13:32:54 | | | |
| 10.000 | IN__C18024426 | 2018-10-21 20:55:15 | | | |
| 10.000 | TN__N18E286115-01B | 2018-10-11 13:32:53 | | | |
| 10.000 | NC__MB060718-0047 | 2018-10-11 13:32:50 | | | |
| 10.000 | IN__C18022757 | 2018-10-11 13:32:54 | | | |
| 10.000 | ... | ... | | | |

6.3. Local Cluster Detection Based on Allele Codes. NOTE: see PND05 (PulseNet standard operating procedure for the organism-specific database workflow) step 6.9. for how to download allele codes from the national database.

6.3.1. In the main BioNumerics window, choose “Find clusters” from the “Analysis” drop-down menu.



6.3.2. In the “Found clusters” window, define the parameters. **NOTE:** the below recommendations are for *Listeria*.

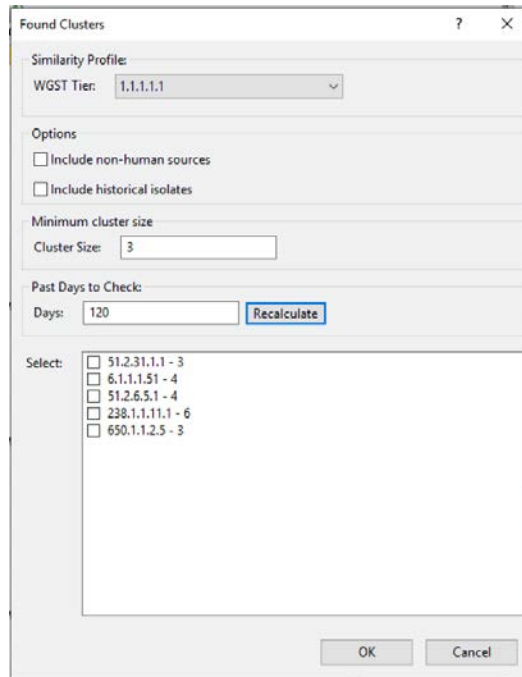
6.3.2.1. Under “Similarity profile” define the “WGS Tier”: 1.1.1.1.1 (up to the 5th digit equals to a difference of approximately 7 alleles).

6.3.2.2. Under “Options”, you can choose to include non-human isolates or historical isolates.

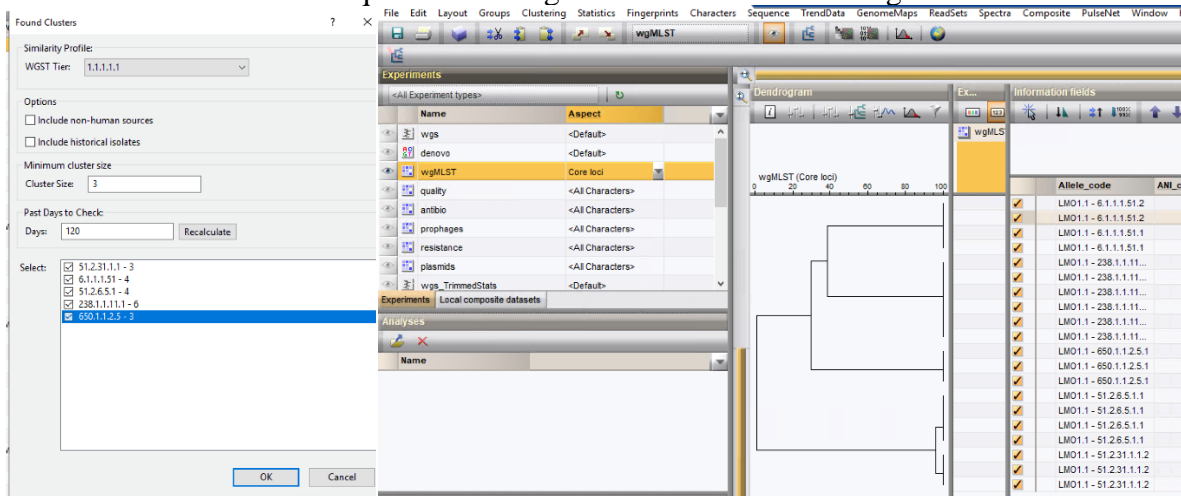
6.3.2.3. Define the “Minimum cluster size”: nationally 3; locally 2 may be more appropriate depending on the size of jurisdiction.

6.3.2.4. Define the “Past days to check”: 120.

6.3.2.5. Found clusters are displayed with the number of cases in the “Select” panel.



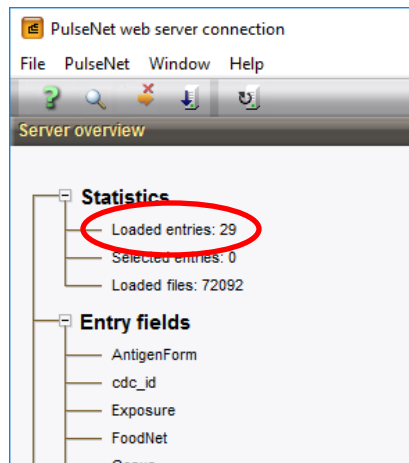
6.3.3. Select the cluster(s) of interest by checking the box next to it and click “OK” to open a comparison window of the entries belonging to the cluster(s). Select the entries and use the quick clustering tool to view the clustering.



6.4. Query the National Database and Download Bundles.

6.4.1. Authenticate to the PulseNetWGS firewall and connect to the national database to the number of days you want to include in your search.

6.4.2. The “PulseNet web server connection” window will open and display the number of entries you are connected to.



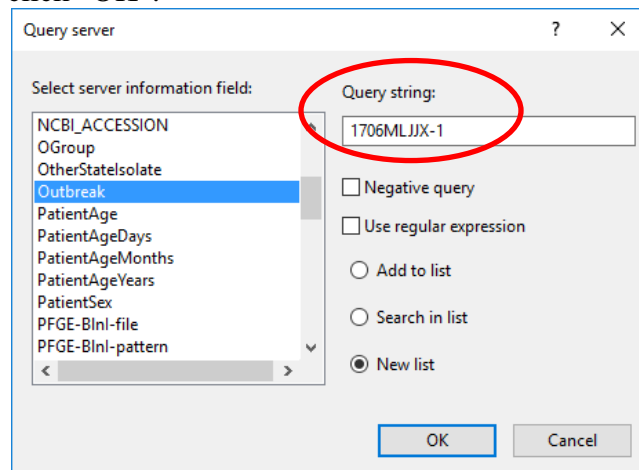
6.4.3. To perform a basic query:

6.4.3.1. Click on the “Query the server using the generic query form” icon (magnifying glass) to begin your query.



6.4.3.2. The “Query server” window will open. Choose the search field from the “Select server information field” list and enter the “Query string” for that information field, e.g., an outbreak code.

6.4.3.3. Select the appropriate output (“Add to list”, “Search in list” or “New list”) and click “OK”.



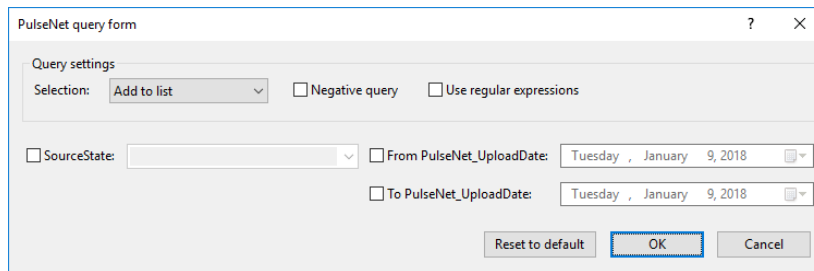
6.4.4. To perform an advanced query:

6.4.4.1. In the “PulseNet web server connection” window, click on “Query the server using the query form” (green question mark).

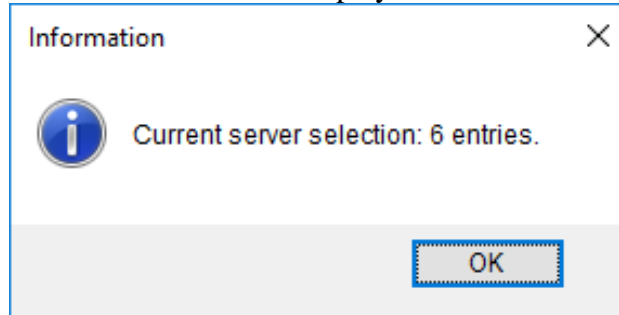


6.4.4.2. The “PulseNet query form” window will open. You can query entries from a particular state and/or from a particular upload date range.

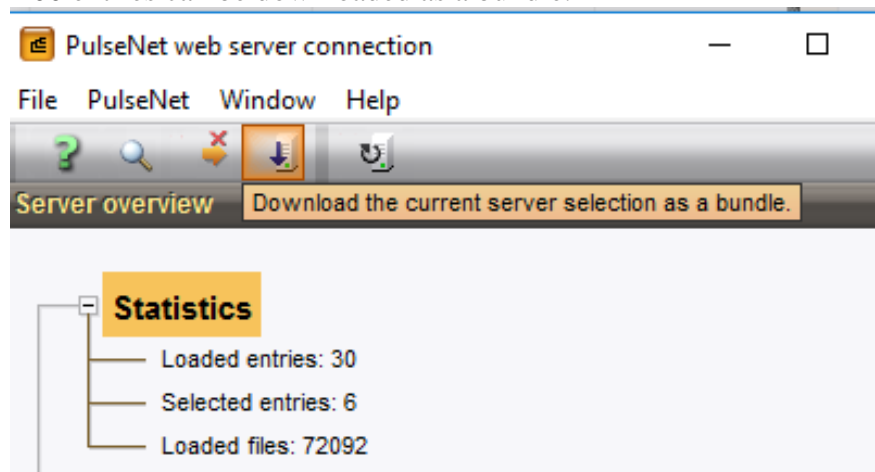
6.4.4.3. Select the appropriate output (“Add to list”, “Search in list” or “New list”) under “Query settings” and click “OK”.



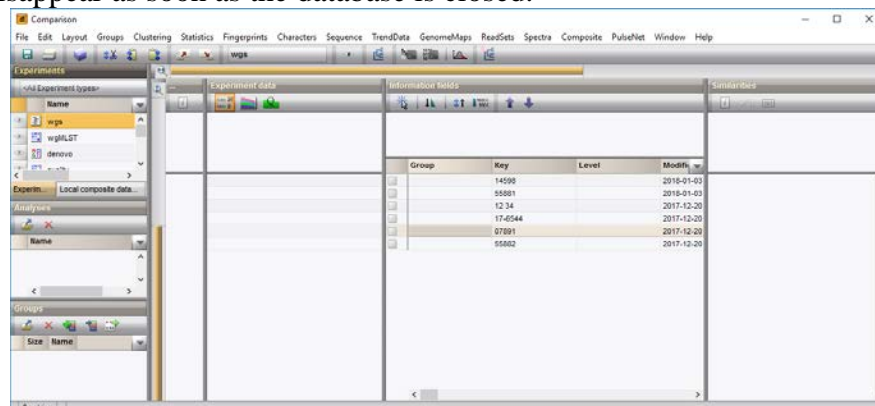
6.4.5. The information window will display the current server selection. Click “OK”.



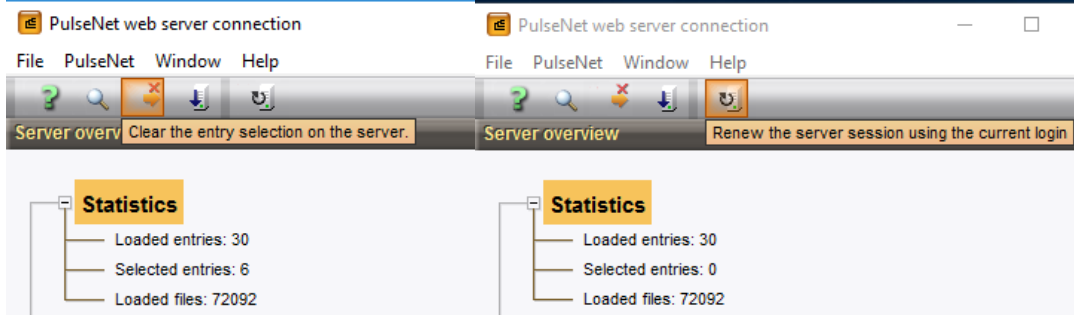
6.4.6. Click “Download the current server selection as a bundle”. **NOTE:** a maximum of 100 entries can be downloaded as a bundle.



6.4.7. A comparison widow showing the downloaded entries opens automatically in your local database. **NOTE:** the download is temporary. The downloaded entries disappear as soon as the database is closed.



6.4.8. If performing multiple different searches, you can clear the previously selected entries by clicking “Clear the entry selection on the server” between the searches. You can refresh the server connection by clicking “Renew the server session using the current login”.

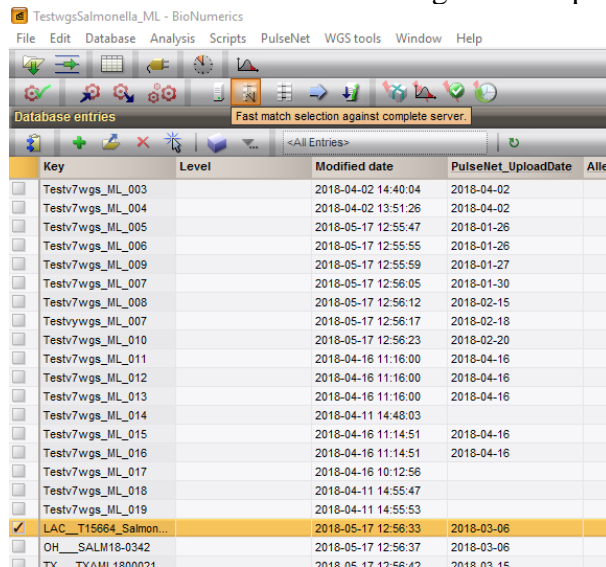


6.5. Match Against Server.

6.5.1. Authenticate to the PulseNetWGS firewall and connect to the national database.

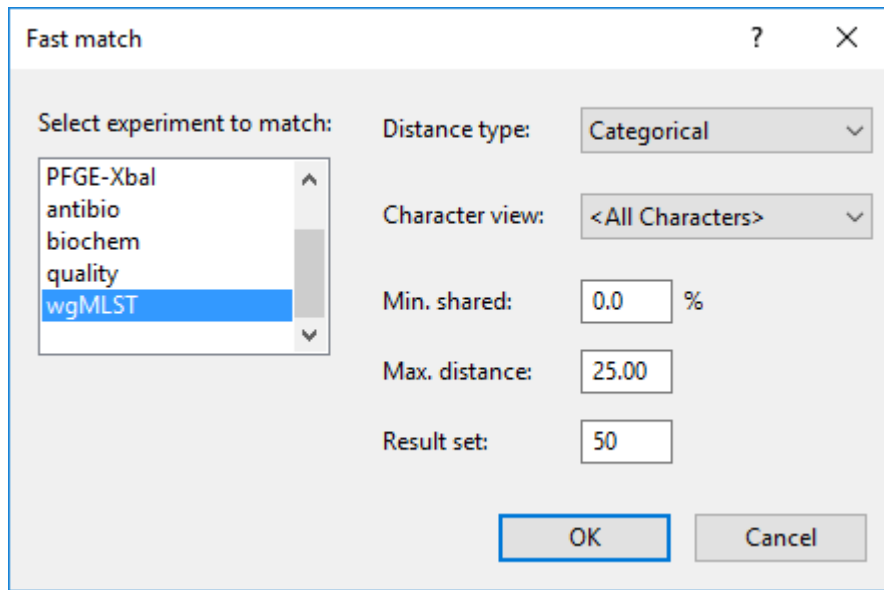
NOTE: the “Match against server” functionality will match against the entire database regardless of the number of days you choose to connect to.

6.5.2. In the main BioNumerics window, select the entry/entries to match against the server and click on “Fast match selection against complete server”.



6.5.3. The “Fast match” window will open. Define the match parameters:

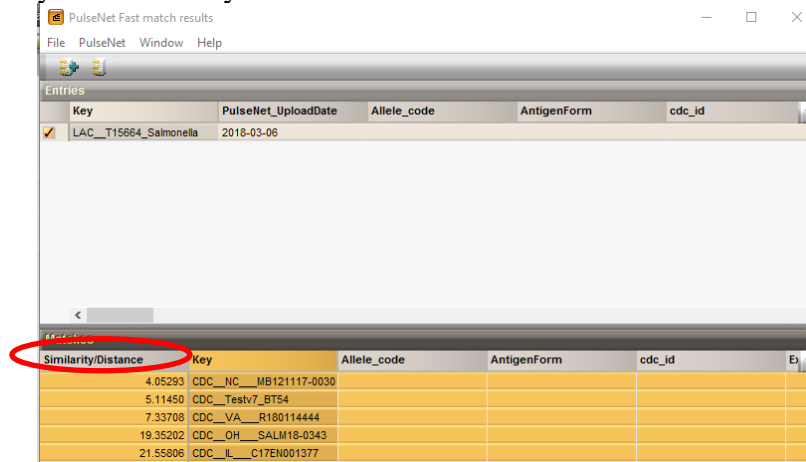
- 6.5.3.1. Experiment type to match: wgMLST
- 6.5.3.2. Distance type: categorical
- 6.5.3.3. Character view: cgMLST
- 6.5.3.4. Max distance: the maximum number of allele differences you are interested in. For example, choosing 25 here will mean that you will get results from the national database that are within 25 alleles of your local sequence.
- 6.5.3.5. Result set: The maximum number of matches to be displayed for your query. For example, choosing 50 here will mean that up to 50 closest matches will be displayed.
- 6.5.3.6. Click “OK”



6.5.4. The “PulseNet fast match results” window will open.

6.5.4.1. The “Entries” panel in the top will show your local entry/entries being matched against the server.

6.5.4.2. The “Matches” panel in the bottom will show the closest matches in the national database. The “SimilarityDistance” field lists the allele differences to your local entry.



6.5.4.3. To view the match results locally:

6.5.4.3.1. Highlight the entries on the “Matches” panel.

6.5.4.3.2. Click “Add selected matches to the server selection” (icon with the green plus) or “Replace server selection with selected matches” (icon without the green plus).

PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 20 of 28

| Group | Key | Modified date | Additional Allele codes | Antigen/AntigenForm_wgs | IsolateDate | LabID | MLST_ST | NA/RMS-Cl | NCBI_Accession | Outbreak | PatientAge | Age |
|-------|-----------------------------|---|--------------------------|-------------------------|---------------|-------|----------|-------------|----------------|----------|------------|-----|
| 1 | MN_101901629 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.3.3 | 14c1.2 | 1/28/2016 MN | ST19 | | SAMN1981824 | 1812M_IPK-1 | | 31 | |
| 2 | MN_101901763 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.3.3 | 14c1.2 | 3/29/2016 MN | ST19 | | SAMN1981829 | 1812M_IPK-1 | | 27 | |
| 3 | MN_101901367 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.3.3 | 14c1.2 | 1/25/2016 MN | ST19 | | SAMN1981869 | 1812M_IPK-1 | | | |
| 4 | MN_101900295 | 6/26/2019 1:00 | | 14c1.2 | 3/6/2016 MN | ST19 | | SAMN1984179 | 1812M_IPK-1 | | | |
| 5 | MN_101900859 | 6/26/2019 1:00 | | 14c1.2 | 2/12/2016 MN | ST19 | | SAMN1984183 | 1812M_IPK-1 | | | |
| 6 | MN_101900914 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.3.1 | 14c1.2 | 12/28/2016 MN | ST19 | | SAMN1984185 | 1812M_IPK-1 | | 10 | |
| 7 | MN_101900295 | 6/26/2019 1:00 | | 14c1.2 | 3/6/2016 MN | ST19 | | SAMN1984166 | 1812M_IPK-1 | | | |
| 8 | MN_101901723 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.3.3 | 14c1.2 | 12/29/2016 MN | ST19 | | SAMN1975287 | 1812M_IPK-1 | | | |
| 9 | MN_101901882 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.3.2 | 14c1.2 | 11/18/2016 MN | ST19 | | SAMN1984066 | 1812M_IPK-1 | | | |
| 10 | 2012 MN_E01200652 | 6/26/2019 0:39 | | 14c1.2 | 3/3/2016 MN | | | SAMN1981812 | 1280VA_IPK-1 | | 34 | |
| 11 | new APH1-15-12281_scaf10ids | 10/29/2015 17:00 do not up | SALML10_24.5.34.1.2.3.1 | | | | | SAMN1208277 | | | | |
| 12 | new APH1-15-02596_scaf10ids | 10/29/2015 17:00 do not up | SALML10_24.5.34.1.2.3.1 | | | | | SAMN1208276 | | | | |
| 13 | 2012 WA_17034 | 6/26/2019 1:29 | | 14c1.2 | 8/14/2012 WA | | END21063 | SAMN1981895 | 1280VA_IPK-1 | | 61 | |
| 14 | new APH1-16-02227_scaf10ids | 10/29/2015 17:00 do not up | SALML10_24.5.34.1.2.3.2 | | | | | SAMN1208278 | | | | |
| 15 | CO_161207003 | 10/29/2015 15:09 | SALML10_24.5.34.1.2.3.1 | 14c1.2 | 11/4/2016 CO | ST19 | | SAMN1985153 | 1812M_IPK-1 | | 10 | |
| 16 | WA_20652 | 10/29/2015 15:21 | SALML10_24.5.34.1.2.1.3 | 14c1.2 | 2/18/2016 WA | ST19 | END21052 | SAMN1179170 | 1812M_IPK-1 | | | |
| 17 | NE_NF11_180763 | 10/29/2015 15:28 | SALML10_24.5.34.1.2.1.3 | 14c1.2 | 12/7/2016 NE | | | SAMN1977465 | 1812M_IPK-1 | | 6 | |
| 18 | MO_23023 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.1.3 | 14c1.2 | 11/27/2016 MO | ST19 | | SAMN1947307 | 1812M_IPK-1 | | 18 | |
| 19 | VA_R39C26291 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.1.3 | 14c1.2 | 2/19/2019 VA | ST19 | | SAMN1135912 | 1812M_IPK-1 | | 11 | |
| 20 | MO_23709 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.1.3 | 14c1.2 | 12/2/2016 MO | ST19 | | SAMN1969379 | 1812M_IPK-1 | | 4 | |
| 21 | new IAC_W8310_Salmoneilla | 6/26/2019 0:35 do not upload already in NDR | | | 4/11/2016 | | | SAMN1158166 | 1812M_IPK-1 | | | |
| 22 | VA_R39C2498 | 10/29/2015 15:26 | SALML10_24.5.34.1.2.1.3 | 14c1.2 | 2/10/2019 VA | ST19 | | SAMN1135917 | 1812M_IPK-1 | | 30 | |
| 23 | NE_N0304460 | 10/29/2015 15:23 | SALML10_24.5.34.1.2.1.8 | 14c1.2 | 10/28/2016 NE | ST19 | | SAMN1969380 | 1812M_IPK-1 | | 28 | |
| 24 | new APH1-16-01007_scaf10ids | 10/29/2015 17:00 do not up | SALML10_24.5.34.1.2.1.1 | | | | | SAMN1984081 | | | | |
| 25 | IN_C8HD1950 | 10/29/2015 15:18 | SALML10_24.5.34.1.2.1.13 | 14c1.2 | 8/31/2019 IN | ST19 | | SAMN1152489 | 1812M_IPK-1 | | | |
| 26 | TX_74AM180375 | 10/29/2015 15:41 | SALML10_24.5.34.1.2.1.12 | 14c1.2 | 10/29/2016 TX | ST19 | | SAMN1148574 | 1812M_IPK-1 | | 1/11/1900 | |
| 27 | KY_15_R069818_0002 | 10/28/2015 15:19 | SALML10_24.5.34.1.2.1.4 | 14c1.2 | 3/16/2016 KY | ST19 | | SAMN1187381 | 1812M_IPK-1 | | 31 | |

6.6.2. Export a dendrogram: see step 6.1.6.

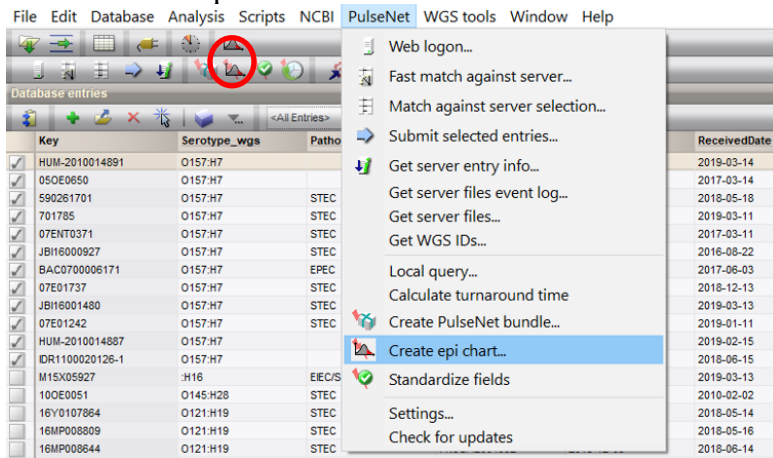
6.6.3. Export a similarity (allele) matrix: see step 6.1.7.

6.6.4. Export resistance or virulence information: see SOP PND05 step 6.6.5.

6.6.5. Create and export a frequency chart (epi curve).

6.6.5.1. In the main BioNumerics window, select the entries to be included in the chart.

6.6.5.2. From the “PulseNet” drop-down menu, select “Create epi chart” or click on the “Create epi field chart from the selected entries” icon.

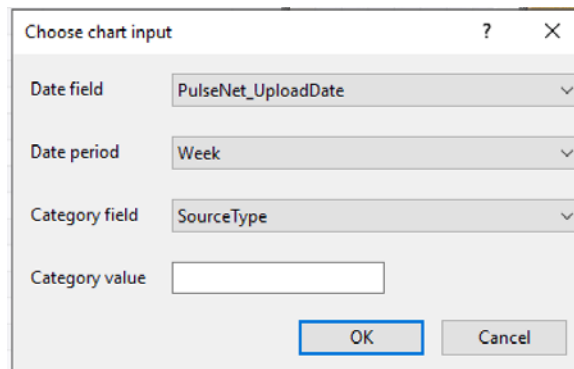


6.6.5.3. In the “Choose chart input” window, choose the input parameters.

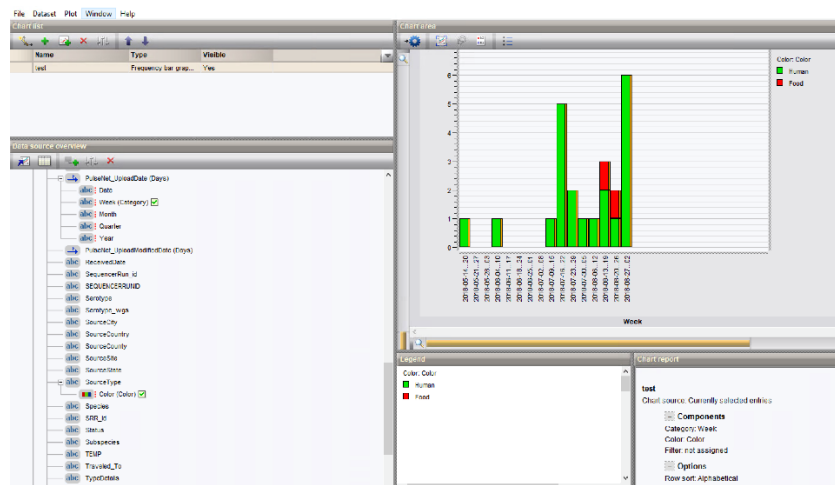
6.6.5.3.1. Date field: PulseNet_UploadDate

6.6.5.3.2. Date period: select the appropriate time interval for the chart X-axis.

6.6.5.3.3. Category field and category value are optional, e.g., by choosing “SourceType” for category field, the human and non-human isolates will be color-coded on the chart.

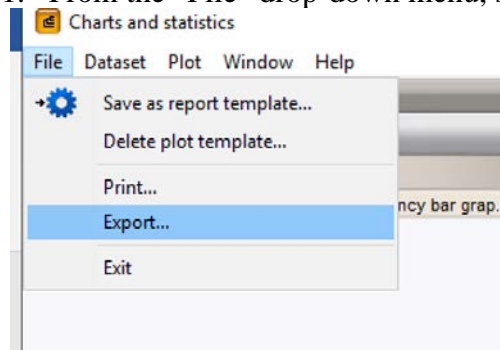


6.6.5.3.4. Click “OK”. The “Charts and statistics” window will open displaying the chart.



6.6.5.4. Export the chart:

6.6.5.4.1. From the “File” drop-down menu, select “Export”.



6.6.5.4.2. In the appearing “Export image” window, you can define the file type to be exported, e.g., if you choose “Metafile to clipboard” you can paste the chart into PowerPoint or Word. Click “OK”.

PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

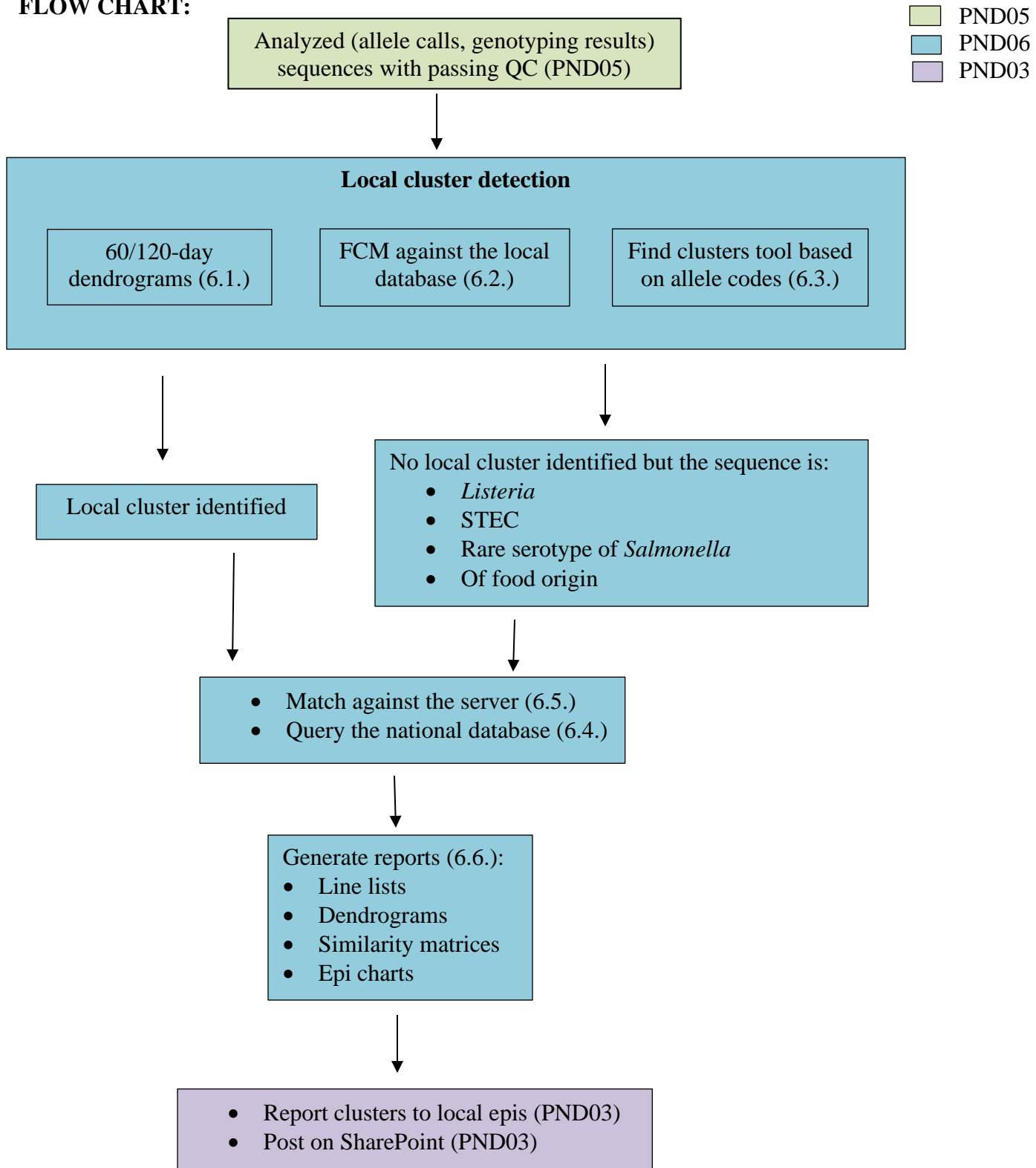
Ver. No. 01

Effective Date:

Page 22 of 28



7. FLOW CHART:



PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 24 of 28

8. REFERENCES:

9. CONTACTS:

9.1. CDC PulseNet Database Team Inbox: PulseNet@cdc.gov

10. AMENDMENTS:

10.1. 06/15/2020 – New Document

PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 25 of 28

11. APPROVAL SIGNATURES:

Approved By: _____ Date: _____
PulseNet QA/QC Personnel

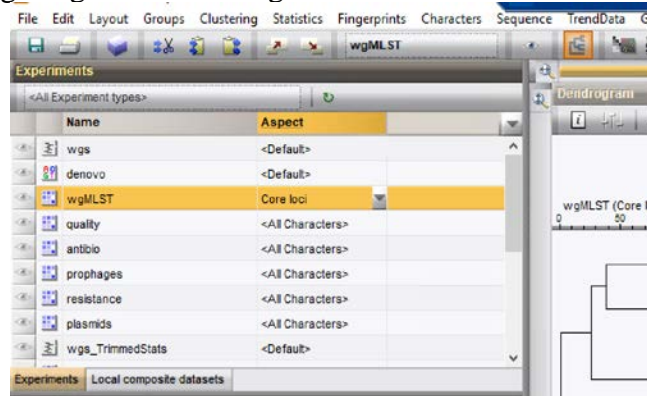
Approved By: _____ Date: _____
PulseNet Outbreak Detection and Surveillance Unit Chief

Approved By: _____ Date: _____
PulseNet Next Generation Subtyping Methods Unit Chief

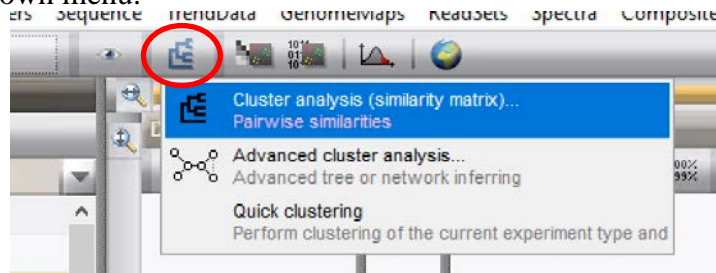
Approved By: _____ Date: _____
PulseNet Reference Outbreak Surveillance Team Lead

Appendix PND06-1. PulseNet Settings for Creating cgMLST or wgMLST Dendrograms

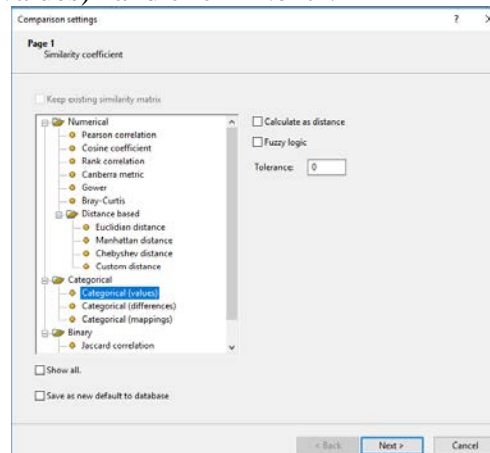
1. In the “Comparison” window, highlight “wgMLST” in the “Experiments” panel, choose “Core loci” or “Core” (depending on the organism database) from the “Aspect” drop-down menu if creating a cgMLST dendrogram or leave the default as “All characters” if creating a wgMLST dendrogram.



2. Click on the clustering icon and choose “Cluster analysis (similarity matrix)” from the drop-down menu.



3. “Comparison settings/Page1/Similarity coefficient” window will open. Choose “Categorical (values)” and click “Next”.



4. In the “Comparison settings/Page2/Cluster analysis” window, choose “UPGMA” for “Method” and click “Finish”.

PULSENET STANDARD OPERATING PROCEDURE FOR LOCAL CLUSTER DETECTION AND MATCHING AGAINST THE NATIONAL DATABASE

Doc. No. PND06

Ver. No. 01

Effective Date:

Page 27 of 28

Comparison settings ? X

Page 2
Cluster analysis

Method

- UPGMA
- Ward
- Neighbor Joining
- Single linkage
- Complete linkage
- Create graph

Degeneracy handling

- Enable degeneracy handling

Secondary criterion: Do not use

Degeneracy: Do not calculate

Cut off above: %

Calculate error flags Calculate cluster cutoff

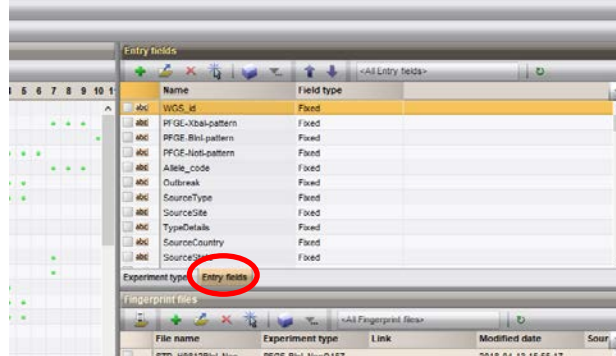
Branch quality: Do not calculate

Dendrogram name: wgMLST (Core)

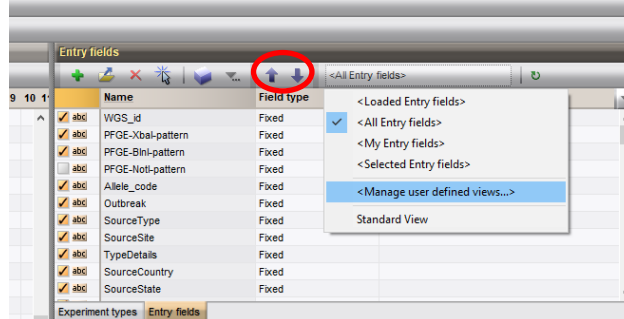
< Back Finish Cancel

Appendix PND06-2. Instructions to Customize the View in the Local Database

1. In the top right panel of the BioNumerics main window, select the “Entry fields” tab.



2. Check the boxes next to the information fields you wish to include in your customized view. If you want the fields in any particular order, use the up and down arrows to move the fields around.
3. Click on the drop-down menu at the top of the “Entry fields” panel and choose “Manage user defined views”. The “Manage Entry field views” window opens.



4. Click “Add”. The “New Entry fields view” window opens.
5. Enter the name of the new layout, leave all other settings as default and click “OK”.
6. Back at the “Manage Entry field views” window, you can make the new layout you just saved your start up view by choosing it from the “Startup view for all levels” drop-down menu. Click “Exit”.

