PulseNet international: the missing link between PFGE and WGS

Ibrahim-Elkhalil M. Adam

Dept. of Zoology, faculty of Science, University of Khartoum, Sudan

Email: abukhalil.mohamed@gmail.com

ORCID: 0000-0002-4677-4533

Abstract

DNA-based surveillance of bacterial diseases has been using pulsed field gel electrophoresis (PFGE) since 1996. Currently, the international surveillance network (PulseNet international) is turning toward whole genome sequencing (WGS). The letters ATCGs of WGS are compared using several sequence alignment methods. While patterns of horizontal lines of PFGE profiles are being compared using relative positioning of bands within a range of tolerance. A recently suggested image analysis algorithm and a deployed database (geltowgs.uofk.edu) collectively invented a promising method for comparing PFGE to *in-silico* obtained digestion models (DMs) derived from WGS. The database requires a parameter that determines PFGE resolution. Here, the author suggests a new method for calculating this factor. Epidemiological and molecular conclusions returned by the database are evaluated.

Methodology: two PFGE profiles representing *Xba*I digests of *E. coli* and *Salmonella enterica* analyzed by the suggested image analysis software were submitted to the database after calculating resolution of PFGE using Dice percentage of difference between the closest PFGE bands in length. *E. coli* and *Salmonella enterica* test subjects were compared to 489 and 401 DMs respectively. The three data sets returned were analyzed.

Results and discussion: according to PFGE evaluation criteria; a single digestion model is possibly related to *E. coli* test subject. It belonged to the same serovar. No epidemiologically related DM was shown for *S. enterica* test subject. Conclusions mentioned earlier could never be made ignoring co-migration. Standardization of both; suggested image analysis and database algorithms will deepen our understanding of bacterial epidemiology.

Introduction

In order to reveal genomic fingerprints of pathogenic bacteria in cases of suspected outbreaks, DNA digest is obtained using restriction enzymes with rare recognition sequence across chromosomes of interest. DNA fragments are separated using pulsed-field gel electrophoresis (PFGE) (1). This technology was initiated and being used for this purpose in USA since 1996 (2). An international consortium of laboratories is growing ever since, building PulseNet international (PNI) (3,4). Currently, 83 countries around the world have already joined PNI (5). Other phenotypic characteristics such as virulence associated mutations and antibiotic resistance profiles of isolates must be investigated using different methods (6). Due to recent sophistication in DNA sequencing technology, PNI started to shift from PFGE to whole genome sequencing (WGS) of bacterial isolates (4). Identification of strains, valuable information about their antibiotic resistance capabilities and of course genetic relationships to other isolates of previous outbreaks are provided by WGS (7).

In case of shifting to WGS without a method that links PFGE data accumulated during the last 26 years to WGS results, valuable information about bacterial population genetics, population dynamics and evolution will be lost. Taking into account that > 800,000 PFGE record are available in PulseNet network of USA alone (2). Shifting to WGS will take time for all participant laboratories across 83 countries to upgrade their infrastructure for WGS (8). Consequently, valuable epidemiological information will be also lost.

The simple logic behind PFGE is that number and distribution of recognition sequence of the restriction enzyme used determine the final profile (9). While whole chromosome and plasmid DNA sequences revealed by WGS can also be digested *in-silico*, resulting in digestion models (DMs). It may seems possible to link PFGE to WGS data. A recently published scientific article (10) identified the following obstacles for doing so;

1. DNA fragments those have the same or very close molecular weight (up to 5% (11)) will resolve in a single PFGE band (Co-migration) (12).
2. Length of PFGE bands are estimated based on a DNA ladder's exponential correlation between retention factors (rF) and band sizes (13).
3. Fragment length of *in-silico* digested WGS are exact numeric values, not simply estimations like the case in PFGE.
4. Typable PFGE fragments are limited by largest and smallest band sizes of the DNA ladder used (9).

Authors of the same article suggested an image analysis algorithm that quantitatively reveal number of co-migrated fragments across PFGE profiles. They named the new parameter "factor of co-migration (FCM)" (10). In order to assess reliability of FCM calculations, they developed "WGSToGel". Which are simulation algorithms that predict PFGE profiles and FCMs from *in-silico* digestions of whole chromosome sequences. They deployed geltowgs.uofk.edu; a freely accessible online database. It compares PFGE-FCM image analysis results to *in-silico* digested whole chromosome sequences (10).

In this study, the author suggests a new method for estimating the parameter critical co-migration threshold (CCT) required by GelToWGS database. Test subjects are previously published PFGE-FCM results of two widely used PFGE markers:

1. *Xba*I digestion profile of *E. coli* O157:H7 strain G5244 (14).
2. *Xba*I digestion of *S. enterica* serotype Braenderup (strain H9812) suggested by Hunter and her team (10,15).

Epidemiological conclusions returned by the database were analyzed. Detailed descriptions of their suggested methods are cited (10). The author will also discuss a qualitative approach to analyze GelToWGS results.

Methodology

Test subjects:

The two PFGE markers were chosen for this study because are they are widely used across PNI laboratories (15). Factor of co-migration using exponential correlation between single-fragment bands and pixel densities (FCM-ECSB) is the image analysis algorithm suggested by Adam IE and his colleagues (10). It depends on the observation that; across a PFGE profile, pixel densities (PDs) reduces as band size do (16). Warner and Onderdonk reported that higher PDs shown by smaller band sizes are caused by co-migration (17) . FCM-ECSB simply calculate FCMs by calculating hypothetical values for bands with "odd PDs" based on other bands of the profile (10).

FCM-ECSB results of *E. coli* strain G5244 were reported in Supplementary material 1 (S1) of the cited article (10). It is a screen-recorded video available online here. FCM-ECSB results of *S. enterica* strain H9812 were also reported by the same author and his colleagues (10).

Selecting CCT and marker errors:

The author suggests selection of the lowest value of Dice %age of difference (between the closest band sizes across a given PFGE profile) as the CCT. Accordingly; for *S. enterica* test CCT was calculated between PFGE bands #10 (173.4 kbp) and #11 (167.1 kbp). $\mathrm{CCT} = \left(\frac{173.4-167.1*2}{173.4+167.1}\right)\%$ This results in 3.7. Similarly, CCT for *E. coli* test calculated between bands #19 (40.5 kbp) and #20 (36.3 kbp) resulted in 5.4687. PFGE marker errors for both test subjects as reported in mentioned citation (10) i.e, 0.007 and 0.002 for *S. enterica* and *E. coli* respectively.

GelToWGS database

FCM-ECSB result for each test was uploaded in MS Excel workbook. Corresponding parameters (digestion models, CCT and marker errors) were supplied to the system. GelToWGS database server returns three data sets for each test. These data sets are available with this publication in supplementary materials 1 and 2 for *E. coli* and *S. enterica* tests respectively.

Results

At selected CCTs for both test subjects, DMs those showed at least 5 matches with each test were 50.1% and 40.65% for *E. coli* and *S. enterica* respectively. Results have clearly shown that total number of fragments for each species-although digested by the same restriction enzyme- were highly different; *Xba*I in *E. coli* chromosomes result in significantly higher number of typable DNA fragments than *S. enterica* (7,465 compared to 3,020). Effect of co-migration was seriously high on isolate relationships. For both tests, raw number of query matches was higher than typable matches. This will make no sense if FCM of each PFGE band is unknown. But when taking into account FCMs of test subject, number of actual matches was reasonable **Table (1)**.

Significant numbers of matches were returned by GelToWGS database for both test subjects. But according to PFGE comparison criteria, a single DM in *E. coli* test was possibly related (differed by 6 fragments). Results of *S. enterica* test did not show any significant relationship according to the same

criteria (18). Details of *E. coli* and *S. enterica* results are in Supplementary material 1 (S1) and Supplementary material 2 (S2) respectively.

Table (1): PFGE-FCM results and corresponding parameters submitted to GelToWGS database and summary of results returned by the database.

Test data submitted to GelToWGS database				Compared to	Summary of results returned by GelToWGS database server					
					Models having ≥ 5 matches			Matches with test subjects		
Test subject	Total number of DNA fragments (PFGE-FCM)	Critical co-migration thresholds	PFGE marker error		No. of *In-silico* digested chromosomes	%age From total	typable fragments	Total query matches	Excluded due to co-migrations	Actual matches
*Xba*I digestion profile of *E. coli* O157:H7 strain G5244	25	5.4687 %	0.007	489 *Xba*I –digested chromosome sequences of *E. coli*	245	50.1%	7,465	9,721	3,423	6,298
*Xba*I digestion of *S. enterica* serotype Braenderup (strain H9812)	20	3.7 %	0.002	401 *Xba*I –digested chromosome sequences of *S. enterica*	163	40.65%	3,020	3,542	1,296	2,246

*Xba*I digestion profile of *E. coli* O157:H7 strain G5244

Out of 489 DM, 245 (50.1%) has shown ≥ 5 matches with the 25 DNA fragments of test subject. Total number of fragments among the 245 DM is 7,465. When considering number of matches; total query matches were 9,721 which is even more than typable ones. Out of total query matches, 3,423 were excluded due to co-migration. Consequently, actual matches reduced to 6,298. These results suggest that DMs of *E. coli* stored in the database shows a high degree of overall similarity and 15.63% of matches mislead PFGE evaluation.

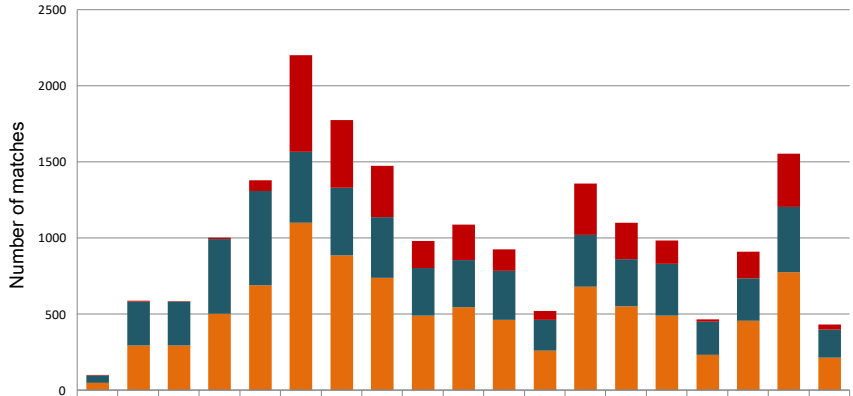
In fact, top 26 DMs (with the exception of two clones) those have shown high similarity with test subject belongs to the same serotype (O157:H7). Among this number, sum of actual matches was 812 (12.89% of total matches). Number of these typable DNA fragment per each DM ranged from 23-34. Number of matched fragments with test data (19 PFGE band representing 25 DNA fragments) ranged from 19-14 within the O157:H7 serotype. According to currently adopted criteria for PFGE evaluation (19), a single DM is considered possibly related to test subject; PFGE profiles differed by 6 DNA fragments. It is the strain 3384 with NCBI-GeneBank accession number CP017440.1 which showed Dice coefficient of 0.667, Supplementary material 1 (S1).

Epidemiological information as retrieved from NCBI-BioSample database (Key: SAMN04447623) show that it was collected in USA Kansas state in 2003. It showed enterohemorrhagic activity (EHEC). Epidemiological information about strain G5244 in NCBI-BioSample database returned a single result (BioSample ID: SAMN05360176). Collection date is missing. Collected in USA by the CDC. The sample also showed EHEC activity. No WGS data is available.

When considering possible occurrences of each PFGE band across DMs; band #6 with length of 244.1 kbp (representing a single DNA fragment) occurs in 419 out of 489 DM. query matches with this fragment ranged from 0 to 4. But when considering actual matched; band #5 (276.9 kbp representing two fragments) occurs the most. PFGE band #1 is very rare. It occurred in 48 and 50 times as actual and query match respectively **Fig. (1).**

Bands #19, #16 and #12 were also relatively rare. Interestingly, exclusion due to co-migration was very rare across PFGE bands those showed co-migration. Namely; bands #2 to #5 and #16. Single-fragment bands #1, #12 and #19 seem to be relatively rare across tested DMs. Not only because of their relatively rare occurrence, but also by their very few exclusions due to co-migrations **Fig. (1)**.

Fig. (1): Total details of matches with each PFGE band across *Xba*I digestion profile of *E. coli* O157:H7 strain G5244 returned by GelToWGS.



	582.6 (1)	445.2 (2)	357.8 (2)	302.1 (3)	276.9 (2)	244.1 (1)	237.3 (1)	223.5 (1)	185.3 (1)	176.3 (1)	157.1 (1)	125.7 (1)	102 (1)	68.8 (1)	60.3 (1)	53.5 (2)	46.6 (1)	40.5 (1)	36.3 (1)
Order in PFGE profile	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Excluded due to co-migration	2	8	4	9	70	633	446	339	177	233	140	57	340	240	153	14	176	348	34
Actual matches	48	287	289	493	620	468	442	399	314	312	323	203	340	311	339	219	280	429	182
Query matches	50	295	293	502	690	1101	888	738	491	545	463	260	680	551	492	233	456	777	216

X-axis shows PFGE band sizes in kbp and corresponding number of DNA fragments represented by each band (FCMs).

*Xba*I digestion of *S. enterica* serotype Braenderup (strain H9812)

163 DM out of the total 401 (40.65%) showed ≥ 5 matches with test subject, taking into account its 17 PFGE bands represent 20 different co-migrated DNA fragments. These fragments were compared to 3,020 fragments of DMs. Total query matches were 3,542 which is more than typable fragments of DMs. When considering co-migration in query fragments, the number dropped to 2,246 (74.37%). These findings suggest that co-migration reduces accuracy of PFGE by 24.64% **Table (1)**.

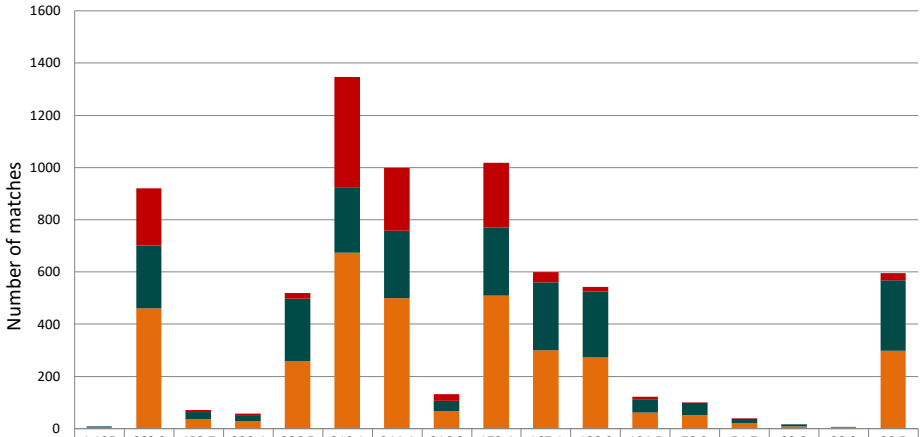
When considering top matched individual DMs; top three matches belong to different serovars. Match number 4 belong to the same Braenderup serovar. Among top 25 matches; matched fragments ranged from 9 to 12 while number of typable fragments per DM ranged from 23 to 29 and its sum is 250 fragments out of 2,246 (11.13 % from total). Dice coefficient of variation ranged from 0.5789 to 0.4615. Matches with the same serovar (Braenderup) comprised only 11 fragments (0.49 % from total) **Table (1)** and supplementary material 2 (S2).

Epidemiological information for previously mentioned DM (retrieved from NCBI-BioSample database Key: SAMN06044997) show that it belong to strain SA20026289. No more data is available.

When we consider presence of each PFGE band across DMs; actual matches with bands #6, #2, #9, #7 and #17 comprised 1,515 (67.54%) out of 2,246. Effect of co-migration is highest (425 out of 673) in band #6 (310.1 kbp representing a single DNA fragment) followed by #9, #7 and #2. In contrast, #16, #1, #15 and #14 have vastly less occurrences across DMs tested. Interestingly, #15 and #13 actually

represent 3 and 2 different co-migrated fragments respectively. Exclusions due to co-migration for both bands were 0 and 2 respectively **Fig. (2)** And supplementary material 2 (S2).

Fig. (2): Total number of matches for each individual PFGE bands across *Xba*I digestion of *S. enterica* serotype Braenderup (strain H9812) returned by GelToWGS database.



	1,135 (1)	668.9 (1)	452.7 (1)	398.4 (1)	336.5 (1)	310.1 (1)	244.4 (1)	216.9 (1)	173.4 (1)	167.1 (1)	138.9 (1)	104.5 (1)	78.2 (2)	54.7 (1)	33.3 (3)	28.8 (1)	20.5 (1)
Order in PFGE profile	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Excluded due to co-migration	0	219	7	6	21	425	242	26	249	39	18	13	2	3	0	0	26
Actual matches	4	241	28	22	238	248	257	40	260	260	253	48	48	17	8	3	271
Query matches	4	460	35	28	259	673	499	66	509	299	271	61	50	20	8	3	297

X-axis shows PFGE band sizes in kbp and corresponding number of DNA fragments represented by each band (FCMs).

Detailed comparison between test subject and the DM that belongs to the same serovar (Braenderup strain SA20026289), reveal an error in GelToWGS comparison algorithm. A single query fragment might be considered a match more than once. Band #9 and #10 (173.4 and 167.1 kbp respectively) of test subject (at 3.7 CCT) had query fragment of 168.336 kbp counted twice as query FCM. In spite of this error, final conclusion of actual matches is not affected, obviously because both test fragments represent a single DNA fragment according to FCM-ECSB calculations **Fig.(3)**. Interestingly, bands #1, #16, #3 and #4 which showed rare occurrences across DMs were all present in this DM which belongs to the same serovar as test subject **Fig.(3)**.

Fig. (3): Details of comparing PFGE-FCM results of *Xba*I digestion of *S. enterica* serotype Braenderup (strain H9812) to i*n-silico* obtained *Xba*I-digested whole chromosome sequence of *S. enterica* that belong to the same serotype (strain SA20026289). NCBI-GeneBank accession number CP022490.1. Results were returned by GelToWGS database.

No.	PFGE band size	PFGE FCM	Upper limit (from PFGE)	*XbaI In-silico* digestion of CP022490.1	Lower limit (from PFGE)	Query band size	Query FCM	Number of matched fragments
1	1135000	1	1195722.5	1084426	1074277.5	1084426	1	1
2	668900	1	704686.15	663587	633113.85	663587	1	1
3	452700	1	476919.45	452619	428480.55	452619	1	1
4	398400	1	419714.4	406846	377085.6	406846	1	1
5	336500	1	354502.75	248198	318497.25		0	0
6	310100	1	326690.35	216259	293509.65		0	0
7	244400	1	257475.4	209952	231324.6	248198	1	1
8	216900	1	228504.15	176565	205295.85	213105.5	2	1
9	173400	1	182676.9	168336	164123.1	172450.5	2	1
10	167100	1	176039.85	160594	158160.15	164465	2	1
11	138900	1	146331.15	149542	131468.85	146180	1	1
12	104500	1	110090.75	146180	98909.25		0	0
13	78200	2	82383.7	127473	74016.3	82087	1	1
14	54700	1	57626.45	112940	51773.55		0	0
15	33300	3	35081.55	86582	31518.45		0	0
16	28800	1	30340.8	82087	27259.2		0	0
17	20500	1	21596.75	59442	19403.25	20354.5	2	1
				38509				
				37688				
				30698				
				25895				
				20992				
				19717				
				6517				
				2603				
				633				

Summary of GelToWGS results (GeneBank: CP022490.1)	
Dice coefficient of variation	0.523809524
Identification	Salmonella enterica serovar Braenderup Strain SA20026289
Accession Number	CP022490.1
Number of query fragments	22
Total PFGE FCMs	20
Epidemiological relatedness	Unrelated

Each match is indicated by a different color. Notice that DNA fragments of *in-silico* digestion, which corresponds PFGE band #9 considered as a match with both PFGE bands #9 and #10 (indicated by overlapping pink and dark red). Query FCMs for both corresponding PFGE bands were set to 2. DNA fragments indicated by Red backgrounds are non-typable by PFGE.

Discussion

Since GelToWGS database is built to bridge the gap between PFGE and WGS, results shown above reveal reasonable conclusions under currently adopted criteria of PFGE evaluation. A single DM of *E. coli* test was possibly related to PFGE test subject. Both of them belong to the same serovar (O157:H7). *S. enterica* test showed no relationship. We argue that number of DMs stored in the database is small; hence it is not representative to the diversity of these species. This claim is especially true for *S. enterica* test. That a single DM within the same serovar was shown. It showed no epidemiological relationship. For *E. coli* test; a significant number of DMs within serovar O157:H7 is stored in GelToWGS database. It is important to clarify that these conclusions were made based on DNA fragments in both; PFGE images and *In-silico* digested WGSs. Algorithmic error shown in **Fig. (3)** is obviously serious and should be corrected.

Since GelToWGS algorithms are configured to run a simulation based on the assumption that DMs were run under the same wet-lab PFGE conditions, reasonability of results would never be achieved without considering co-migration in both PFGE and DMs. It is important to deeper evaluate FCM-ECSB algorithm in terms of sequencing PFGE bands after DNA extraction from gel. If such test is done for an isolate that already been sequenced, it would be more informative. Alongside, if GelToWGS database is upgraded to store sequences of each fragment (instead of length only), PFGE typing will be upgraded to a new level. Such steps will allow epidemiologists to evaluate PFGE and *In-silico* digested WGS in terms of qualitative

approach; matches may be linked to chromosomal loci of the fragment. Common fragments within serotypes and pulsotypes may become determinants of these categories **Fig. (2)**. Rare matches coupled with high overall number of matches will indicate degrees of relatedness described by Kam and his team **Fig. (1 and 2)**.

Regarding our suggested method of selecting CCT, the question raised by authors of GelToWGS still needs an answer which is "does Dice percentage of differences reflects behavior of DNA fragments during PFGE?" The answer is crucial not only to described method in this article, but also to WGSToGel simulation algorithms upon which the claim that 35% of PFGE fragments co-migrate (10) is made.

When we consider findings of WGSToGel simulations reported by the same publication (statistics from supplementary material 4) (10). *Xba*I DMs of *E. coli* and *S. enterica* (64 and 70 DM respectively) suggest that recognition sequence of the enzyme (…TCTAGA…) (20) occurs in similar frequency across both species **Table (2).** However, co-migration occurs more frequently in *E. coli* DMS. When considering limitations of PFGE; averages and standard deviations clearly reflect effects of marker size limitation and co-migrations (reduction from total number of fragments compared to both; PFGE-typable and expected bands) **Table (2)**. Considering FCM-ECSB results of test subjects, sum of FCM represents typable fragments, while number of bands correspond expected ones. FCM-ECSB results are within the ranges of maximum and minimum values of WGSToGel simulations; 19 bands representing 25 fragments for *E. coli* strain G5244 and 17 bands showing 20 fragments in *S. enterica* strain H9812 **Table (2)**.

Table (2): Summary statistics of WGSToGel simulations of *Xba*I digestions of whole chromosome sequences of *E. coli* and *S. enterica.*

PFGE simulation	Statistics	%age of typable size	Total Number Of Fragments	PFGE-typable Fragments	Expected PFGE bands
*Xba*I digestion model of *E. coli* whole chromosome sequences (64 in total)	Sum	N/A	3609	2419	1379
	Mean ±SD	97.3045 ±3.4307	56 ±41	38 ±18	21 ± 5
	Max	99.692	161	82	32
	Min	82.451	17	4	4
*Xba*I digestion model of S. enterica whole chromosome sequences (70 in total)	Sum	N/A	3412	2426	1551
	Mean ±SD	95.498 ±9	49 ±21	35 ±12	22 ±6
	Max	99.845	73	63	42
	Min	66.694	18	16	12

Data in this table were derived from a supplementary material of previously published article by IEM Adam and his team. The article described WGSToGel simulations of several PFGE protocols.

References

1. Mcclelland M, Jones R, Patel Y, Nelson M. Restriction endonucleases for pulsed field mapping of bacterial genomes. Nucleic Acids Res. 1987;15(15):5985–6005.

2. Tolar B, Joseph LA, Schroeder MN, Stroika S, Ribot EM, Hise KB, et al. An Overview of PulseNet USA Databases. Foodborne Pathog Dis. 2019;16(7):457–62.

3. Swaminathan B, Gerner-Smidt P, Ng L-K, Lukinmaa S, Kam K-M, Rolando S, et al. Building PulseNet International: An Interconnected System of Laboratory Networks to Facilitate Timely Public Health Recognition and Response to Foodborne Disease Outbreaks and Emerging Foodborne Diseases. Foodborne Pathog Dis. 2006;

4. Nadon C, Walle I Van, Gerner-smidt P, Campos J, Chinen I, Concepcion-acevedo J. PulseNet International : Vision for the implementation of whole genome sequencing (WGS) for global food- borne disease surveillance. Euro Surveill. 2017;22(23):1–12.

5. Hunter SB, Vauterin P, Lambert-Fair MA, Van Duyne MS, Kubota K, Graves L, et al. Establishment of a universal size standard strain for use with the pulsenet standardized pulsed-field gel electrophoresis protocols: Converting the national databases to the new size standard. J Clin Microbiol. 2005;43(3):1045–50.

6. Ozdemir K, Acar S. Plasmid profile and pulsed-field gel electrophoresis analysis of Salmonella enterica isolates from humans in Turkey. PLoS One. 2014;9(5):1–7.

7. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D. Genomic analysis of diversity , population structure , virulence , and antimicrobial resistance in Klebsiella pneumoniae , an urgent threat to public health. 2015;

8. CDC. PulseNet International: On the path to implementing whole genome sequencing for foodborne disease surveillance [Internet]. Cdc.Gov. 2019. Available from: www.cdc.gov/pulsenet

9. Struelens MJ, De Ryck R, Deplano A. Analysis of Microbial Genomic Macrorestriction Patterns by Pulsed-Field Gel Electrophoresis (PFGE) Typing. New Approaches Gener Anal Microb Typing Data. 2001;159–76.

10. Adam I-E, Abdokashif I, Elrashid A, Bayoumi H, Musa A, Abdulgyom E, et al. Novel algorithms for PFGE bacterial typing: Number of co-migrated DNA fragments, linking PFGE to WGS results and computer simulations for evaluation of PulseNet international typing protocols. J Appl Microbiol Res. 2020;3(2):52–67.

11. Van Belkum A, Van Leeuwen W, Kaufmann ME, Cookson B, Forey F, Etienne J, et al. Assessment of resolution and intercenter reproducibility of results of genotyping Staphylococcus aureus by pulsed-field gel electrophoresis of SmaI macrorestriction fragments: A multicenter study. J Clin Microbiol. 1998;36(6):1653–9.

12. Goering R V. Pulsed field gel electrophoresis: A review of application and interpretation in the molecular epidemiology of infectious disease. Infect Genet Evol [Internet]. 2010;10(7):866–75. Available from: http://dx.doi.org/10.1016/j.meegid.2010.07.023

13. Hunter SB, Vauterin P, Lambert-Fair MA, Van Duyne MS, Kubota K, Graves L, et al. Establishment of a universal size standard strain for use with the pulsenet standardized pulsed-field gel electrophoresis protocols: Converting the national databases to the new size standard. J Clin Microbiol. 2005;43(3):1045–50.

14. Adam I-EM. PFGE to WGS: Illustration of the method and User manual [Internet]. 1st ed. Khartoum: University of Khartoum; 2018. 9 p. Available from:

http://41.67.20.20/Documentation/GelToWGS.pdf

15. Hunter SB, Vauterin P, Lambert-Fair MA, Van Duyne MS, Kubota K, Graves L, et al. Establishment of a universal size standard strain for use with the pulsenet standardized pulsed-field gel electrophoresis protocols: Converting the national databases to the new size standard. J Clin Microbiol. 2005;43(3):1045–50.

16. Kai MK, Luey CKY, Parsons MB, Cooper KLF, Nair GB, Alam M, et al. Evaluation and validation of a PulseNet standardized pulsed-field gel electrophoresis protocol for subtyping Vibrio parahaemolyticus: An International Multicenter Collaborative study. J Clin Microbiol. 2008;46(8):2766–73.

17. Warner JE, Onderdonk AB. Method for optimizing pulsed-field gel electrophoresis banding pattern data. J Mol Diagnostics [Internet]. 2003;5(1):21–7. Available from: http://dx.doi.org/10.1016/S1525-1578(10)60447-3

18. Tenover FC, Arbeit RD, Goering R V., Mickelsen PA, Murray BE, Persing DH, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed- field gel electrophoresis: Criteria for bacterial strain typing. J Clin Microbiol. 1995;33(9):2233–9.

19. van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect. 2007;13(SUPPL. 3):1–46.

20. XbaI [Internet]. NebCutter|New England BioLabs Inc. [cited 2021 Mar 7]. Available from: https://international.neb.com/products/r0145-xbai#Product Information