International Journal of Basic Science and Technology
October 2023, Volume 9, Number 4, Pages 178 – 188 http://www.ijbst.fuotuoke.edu.ng/178
https://doi.org/10.5555/UXUB2355 ISSN 2488-8648

Comparative Studies on Manual and Computational Techniques for Gel Electrophoresis Image Analysis

[1]Adebesin, O.A., [1]Adesoye, A.J., [2]Fowora, M.A., [3]Mairiga, J.P., [1]Amusa, O.D., [4]Obioma, P.C., [4]Onyeaghasiri, F.U., [1]Dagunduro, E.T., [1]Martins, V.N., [1]Ejuetueyin, O.E and [1]Ojo, S.G.

[1] Department of Cell Biology and Genetics, University of Lagos, Akoka, Lagos State, Nigeria
[2]Nigeria Institute of Medical Research, Yaba, Lagos State, Nigeria
[3] Department of Biochemistry and Molecular Biology, Nasarawa State University Keffi, Nasarawa State, Nigeria
[4] FOWN Biotechnology Ltd, Yaba, Lagos State, Nigeria.

Article Information

Article # 01006
Received: 9th July 2023
1st Revision:11h August2023
2nd Revision:15th Sept 2023
Acceptance:23rd Sept 2023
Available online:
4th October 2023.

Key Words
RAPD, Image Analysis, Nigeria, Manual, Polymorphism, Gel Electrophoresis

.

Abstract

Manual analysis of gel electrophoresis images is a significant research barrier and due to significant constraints, digital technologies have been applied to accelerate the processing and interpretation of gel images. This study aimed to evaluate DNA polymorphism in selected animal and environmental microbes using a manual and computational program for the gel image analysis. To ascertain and compare the accuracy of the analysis performed manually and computationally, Random amplified polymorphic DNA-PCR (RAPD-PCR) was performed, and gel electrophoresis was run on PCR products obtained from Staphylococcus spp cultured from livestock nasal swabs and Bacillus spp cultured from soil and wastewater samples from different abattoirs in Ogun State, Nigeria. Gel image analysis was done manually and computationally with two software programs, PyElph and GelAnalyzer. Results obtained from both analyses were compared, and some discrepancies were observed as the computerized analysis was not in total concordance with the manual visual analysis. The software programs were compared using several parameters: image quality enhancement, visual band, lane detection and definition, data clustering, and dendrogram analysis. PyElph had approximately 60% accuracy in all the features tested, although when compared to manual gel visualization for bands and lanes calculation; there was an inaccuracy with PyElph while GelAnalyzer had a lower percentage of accuracy and lesser functionalities. This research has comprehensively analyzed the most commonly used noncommercial software programs available to researchers and students. This will help identify the features that need to be modified to develop a more effective system for the computational analysis of gel images.

***Corresponding Author**: Adebesin, O.A.;: oadebesin@unilag.edu.ng

Introduction

The ability to isolate, size, and visualise DNA molecules is significant to today's rapid advancement in molecular biology. Gel electrophoresis is the most commonly utilized method for this purpose (Khakabimamaghani *et al*., 2013).

Protein fragments, RNA or DNA can be separated via gel electrophoresis, based on their molecular weights, by causing them to migrate through a substrate, such as a polyacrylamide gel, while subjected to an electric field (Kaabouch *et al*., 2007). After electrophoresis, the stained gel is visualised on a UV transilluminator and photographed using a Gel doc or digital camera. The resulting gel image comprises lanes (vertical columns), representing fragments of DNA, RNA, or protein, with molecular weight sorted horizontal bands in each lane (Labyed *et al*., 2010). The nature of the experiment informs on the type of analysis that the gel electrophoresis images are subjected to (Ahmed, 2021).

Determination of genetic variation and polymorphisms among organisms can be accomplished by analysing the Random Amplified Polymorphic DNA (RAPD) gel image, which involves

observing the difference in band pattern of the amplified DNA molecules through band scoring using the binomial (1-0) matrix and dendrogram analysis to establish relatedness of the samples (Chatterjee and Raval, 2019). The manual scoring, before the analysis of gel electrophoresis images can be time-consuming, error-prone, and monotonous as the images frequently contain anomalies such as impulse noise, large smears, and image faintness (Abeykoon *et al*., 2015). Due to the immense limitations of manual analysis, processing speed and gel image analysis have been enhanced using digital technologies. A few tools have been developed for gel electrophoresis image analysis, although some of them, which are commercial, do not fulfil all user requirements, and the freeware programs are plagued with complexity coupled with constrained options. (Alnamoly *et al*., 2020). Some commercial software including Dolphin 1D, EzQuant, Gel-Pro Analyzer, Gel-Quant, Intelligent Quantifier, Image Lab, Image Studio, LabImage, Molecular Imaging, myImage Analysis, Un-Scan-it, and Ultraquant cannot generate phylogenetic trees based on the gel image weights. Furthermore, free tools, such as ImageJ, Gel QuantandL aneruler cannot produce dendrograms. According to a survey conducted by Herras *et al*. (2016), much of the effective software currently existing is commercially available but, due to the cost is, not easily accessible to students and self-funded researchers. Many of these free tools have inconsistencies in their analysis, which necessitates a comprehensive comparison of their characteristics.

Manual analysis of gel electrophoresis images has proven to be a major bottleneck in research, leading to the development of several innovative technologies capable of resolving the gel image analysis problem. However, some of this existing software has one or more limitations. The most typical limitations of existing software are ergonomic design and accuracy in band and lane detection. Electrophoresed gel image analysis provides researchers with the information needed to understand genetic variations and polymorphisms and detect band pattern differences across different DNA, RNA and protein samples. Hence biological materials can reveal valuable information in some molecular biology applications through electrophoresed and digitised gel images (Maramis and Delopoulos, 2010). For example, relationships can be established between samples via comparison of their DNA patterns on the gel images, by scoring the presence (1) or absence (0) of polymorphic bands in individual lanes, as is the nature of randomly amplified polymorphic DNA (RAPD) markers. Due to the drawback of manual analysis, information technology has been applied to hasten the processing and analysis of gel images. The workflow used by most of this software includes Image optimization, Lane and band detection and the Construction of a Phylogenetic tree. These would be used as criteria for comparison between the two non-commercialised software captured in this study which will aid in equipping researchers with the right information on the software needed for their respective gel analysis and will aid in modifying these existing analytical tools.

Gel Analyzer software is employed in the identification and correct comparison of the multitude of discrete bands in gels for DNA polymorphism patterns in organisms. This program has a broad multicomponent capability to analyse the images of gels resolved fragments of DNA, RNA or proteins. The Software program consists of four modules: image analysis, functional analysis, a track-dependent comparison of spectrum bands, a statistical analysis of the results, data presented in tables, and exported in Excel format, but devoid of gel image phylogenetic analysis (Heras *et al*., 2016). PyElph is a free Python-based software for gel image analyses, utilized in various molecular biology and genetics applications. The software can analyse genetic variants in DNA molecules from various animals or populations (Pavel and Vasile, 2012). It analyses DNA genetic marker gel image patterns and builds phylogenetic trees based on the information in the gel image. This study is aimed at examining the functionalities offered by the most common open-source programs used to analyze gel images as compared to manual analysis to determine the phylogenetic relationship of selected animal / environmental microbes.

Materials and Methods

Selection of Tools

An electronic search for software solutions for gel image analysis was conducted by compiling a list of keywords and search queries. The keywords were used to discover software applications for gel electrophoresis analysis using the Google Scholar search engine and PubMed databases, as suggested by Heras *et al.*, 2016. In the search procedure, the search query term “Gel Image Analysis Software” returned about thirty-six existing software, which was narrowed down to the two computational tools used for this study, using recommendation and accessibility as criteria. These computational tools include PyElph (version 1.4) and GelAnalyzer (version 19.1), released in 2013 and 2010, respectively.

Sample Identification

In total, 44 *staphylococcal* isolates recovered from 162 nasal swabs of healthy animals (cattle, ram, goat and pig) at livestock farms, and thirty (30) suspected

Bacillus spp (catalase positive, haemolysis positive and contained endospore) were also recovered from the 60 environmental (soil and wastewater) samples, all from abattoirs in Ogun State, Nigeria, were used for the study

Genomic DNA Extraction and RAPD-PCR

The genomic DNA was extracted from the presumptive *Staphylococcusspp* and *Bacillus sp*p. harvested from their respective freshly growing cultures on nutrient agar, using the boiling method as described by Dashti *et al*. (2009) and Jena Bioscience Bacteria DNA Kit (Germany), according to the manufacturer's instructions respectively.

RAPD-PCR Analysis.

Extracted DNA from each isolate was randomly amplified using primer OPC 04 (CCGCATCTAC) / RAPD 06 (GTAGACCCGT) via Polymerase Chain Reaction (PCR) in a thermal cycler (Pielterthermal, MJ Research Series). The PCR reaction mixture consisted for each fragment; 0.3 µL10pmol primer (Inqaba, South Africa),4µL 5X FIREPol Blend Master mix (Solis Biodyne, Estonia) and 2µL extracted genomic DNA. The volume of the PCR mixture was adjusted to 20µL with sterile PCR-grade water. The cycling parameters included initial denaturation temperature at 95ºC for 15 min, followed by 35 cycles of denaturation at 95ºC for 1 min, annealing at 30ºC for 1 min and extension at 72ºC for 2 min, thereafter-final extension at 72ºC for 10 min. The amplicons resolved on 2% agarose gel stained with ethidium bromide, were viewed under a UV trans illuminator with digital image acquisition. The 100bp DNA ladder (Solis Biodyne) was used as a DNA molecular weight marker.

Gel Image Analysis

The two software; GelAnalyzer and PyElph, were employed in gel analyses. The gel images generated were also analyzed manually, to determine polymorphism, based on amplified DNA band products that appeared on the gel by scoring reproducible, resolved and non-ambiguous bands as 1 for presence and 0 for absence of a fragment, the binary data matrix generated was tabulated for further analysis. Cluster analysis was done on RAPD molecular data to determine the grouping pattern and genome inference of animal / environmental isolates was calculated by Jaccard's coefficient similarity, using PAST (4.08) application.

Criteria for Evaluation

The criteria used for the evaluation of the different computational tools were split into six categories, five of which were as suggested by Heras *et al*., 2016:

1. General features: This includes the acceptance of specific image formats and the ability to save a work session on the computational tool.
2. Image pre-processing: In this category, the edit option (e.g. crop, rotate or flip) and quality enhancement were reviewed; contrast and brightness adjustment or performing gamma correction of gel images was checked.
3. Lane detection: In this category, automatic detection of lanes and the ability to include and exclude lanes manually, and modification of detected lanes via width and position adjustment were checked.
4. Band detection. In this category, the automatic detection of bands and the ability to add and delete bands manually, modify the detected lanes and adjust their width and position were checked.
5. Comparison and Analysis: This category gathers details on how the different systems construct dendrograms from computed similarity matrices. Also studied dendrograms presentation to the user.
6. Accuracy of the analysis: To prove the accuracy of the analysis performed by each computational tool, the Random amplified polymorphic DNA experiment was performed, and Gel Electrophoresis was run on the PCR products.

Results

Sample identification

Table 3a: Staphylococcus. spp isolate sample classification

Sample Number	*Source*	*Location*	*Classification Name*
1-6	Cattle	Kara, Ogun state	*Bos taurus*, 1-6
7-19	Goat	Kara, Ogun state	*Capra hircus* 1-13
20-26	Pig	Ota, Ogun state	*Sus scrofa domesticus* 1-7
27-28	Pig	Oju-Ore, Ogun state	*Sus scrofa domesticus* 8-9
29-44	Ram	Kara, Ogun state	*Ovis aries* 1-16

T

Table 3b: Bacillus spp. isolate sample classification

Sample Number	*Source*	*Classification Name*
1	Wastewater from Ijoko	Ijoko Wastewater 1
2-5	Soil from Ota	Ota Soil 1-4
6-7	Soil from Oju-Ore	Oju-ore Soil 1-2
8-14	Soil from Kara	Kara Soil 1-7
15-19	Wastewater from Kara	Kara Wastewater 1-5
20-23	Soil from the Tollgate	Tollgate Soil 1-4
24-25	Wastewater from the Tollgate	Tollgate Wastewater 1-2

Gel Electrophoresis Analysis

Six gel images as shown in plates 1-3 were generated, three for each primer, two for the *Staphylococcus spp* (1 - 44) samples and one for the Bacillus. spp, (45 – 79) samples.

Manual Analysis of the Gel Images

The software systems examined in this study allowed for automatic lane and band detection, but initial readings of some gel images highlighted disparities between the software programs, which may be due to variances in the lane and band detection threshold sensitivity. Hence to determine the optimal number of lanes and bands present, gel images were first evaluated manually, and the defined patterns were used as benchmarks for the automated approach.

RAPD of *Staphylococcus* isolates

Primer OPC 04

Genomic DNA of *Staphylococcus* spp (samples 1 – 25) randomly amplified using primer OPC 04, resulted in an average number of 23 visible bands across the first 25 samples. From the 1-0 matrix table made, it was inferred that 14 bands were polymorphic, ranging in molecular weight of 200 to 1200 bp, and about 6 bands were monomorphic with molecular weight ranging from 400 to 2000 bp. The six monomorphic bands were obtained from *B.taurus* 1 and 2 (Cattle), *S. domesticus* 1, 3 and 4 (Pig). It was also observed that B. *taurus* 1 had two distinct monomorphic bands. The labelled gel images are shown in Plate 1. While the second gel (samples 26 – 44) gave an average number of 16 visible bands across the last 19 samples. From the 1-0 matrix table made, it was inferred that 11 bands were polymorphic, ranging in molecular weight of 200 to 1200 bp, and about 4 bands were monomorphic with molecular weight ranging from 400 to 1500 bp. The six monomorphic bands were obtained from *O. aries* 5, 8 and *S. domesticus* 7. It was also observed that *S. domesticus* 7 had two distinct monomorphic bands. The labelled gel images are shown in Plate1.

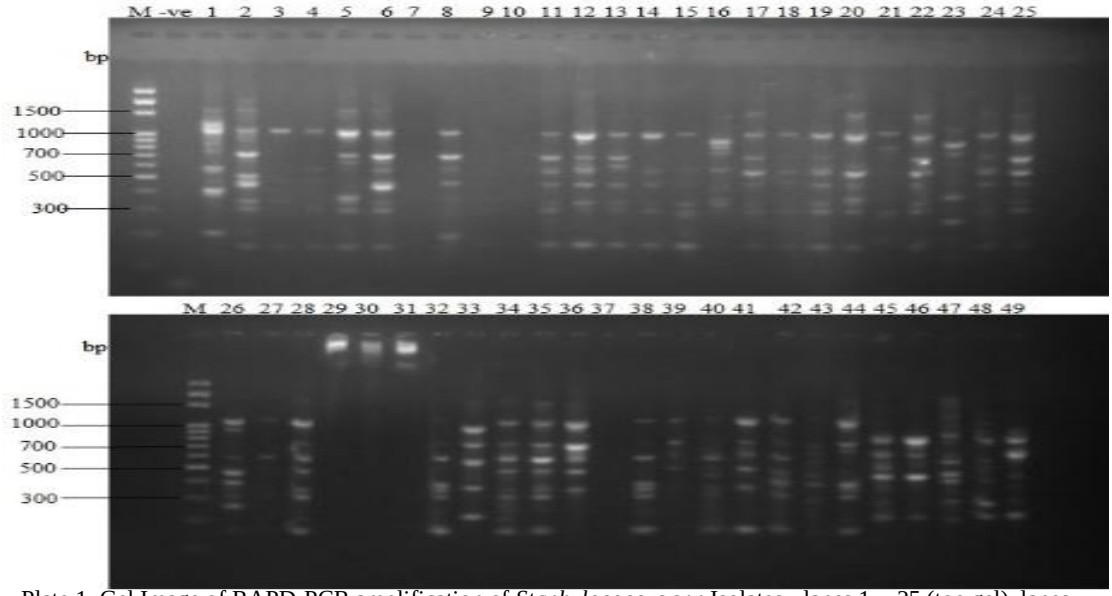


Plate 1. Gel Image of RAPD-PCR amplification of *Staphylococcus spp* Isolates - lanes 1 - 25 (top gel), lanes 26– 44 (bottom gel) and *Bacillus spp* isolates (lanes 45-49) using the OPC 04 Primer

https://doi.org/10.5555/UXUB2355 ISSN 2488-8648

Primer RAPD 06

Genomic DNA of *Staphylococcus spp* (samples 1-25) randomly amplified by using primer RAPD 06, resulted in an average number of 15 visible bands across the first 23 samples. From the 1-0 matrix table made, it was inferred that 11 bands were polymorphic, ranging in molecular weight of 300 to 1200 bp, and about 4 bands were monomorphic with molecular weight ranging from 200 to 1500 bp. Four monomorphic bands were obtained from *B. taurus* 1 and *C. hircus* 11 and 12. The labelled gel images are shown in Plate 2. while randomly amplified DNA from samples 26-44 resulted in an average number of 15 visible bands across the second set of 21 samples. From the 1-0 matrix table, it was inferred that 11 bands were polymorphic, ranging in molecular weight from 300 to 1200 bp, and about 4 bands were monomorphic with molecular weight ranging from 200 to 1500 bp. The four monomorphic bands were obtained from *B.taurus* 1 and *C.hircus* 11 and 12. The labelled gel images are shown in Plate 2.

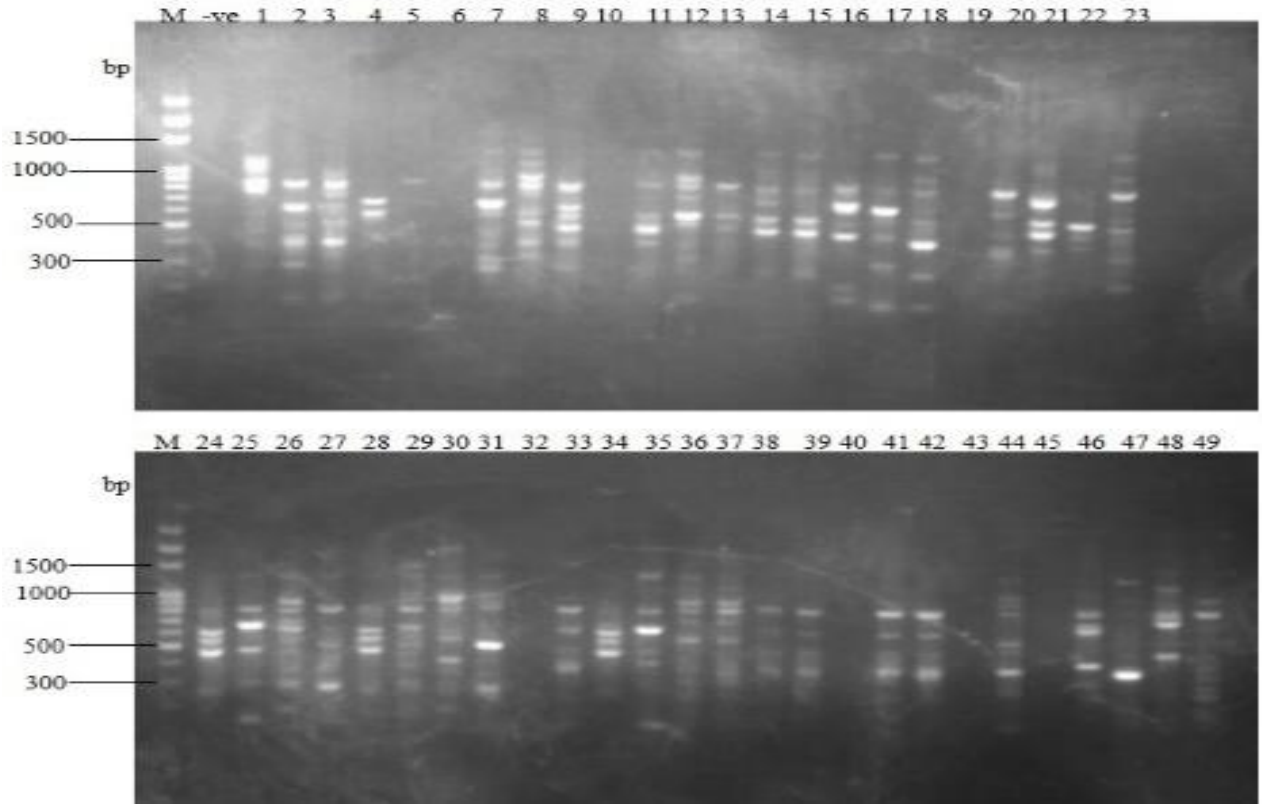


Plate2. Gel Image of RAPD-PCR amplification of *Staphylococcus* spp Isolates - lanes 1 - 25 (top gel), lanes 26– 44 (bottom gel) and *Bacillus spp* isolates (lanes 45-49) using the OPC 06 Primer

RAPD of *Bacillus* isolates

Primer OPC 04

Genomic DNA of *Bacillus spp*. randomly amplified using primer OPC 04, revealed an average number of 18 visible bands across the 25 samples. From the 1-0 matrix table made, it was inferred that 13 bands were polymorphic, ranging in molecular weight from 200 to 1200 bp, and about 5 bands were monomorphic with molecular weight ranging from 100 to 1700 bp. The five monomorphic bands obtained were from Ijoko wastewater, Kara soil, Ota soil, and Oju ore soil samples. The labelled gel images are shown in Plate 3a.

Primer RAPD 06

Genomic DNA of *Bacillus spp*. randomly amplified using primer RAPD 06, revealed an average number

of 14 visible bands across the 18 samples. From the 1-0 matrix table made, it was inferred that 13 bands were polymorphic, ranging in molecular weight from 200 to 1500 bp, and one band was monomorphic with a molecular weight of about 1000 bp. The monomorphic band was obtained from the Tollgate soil samples. The labelled images are shown in Plate 3b.

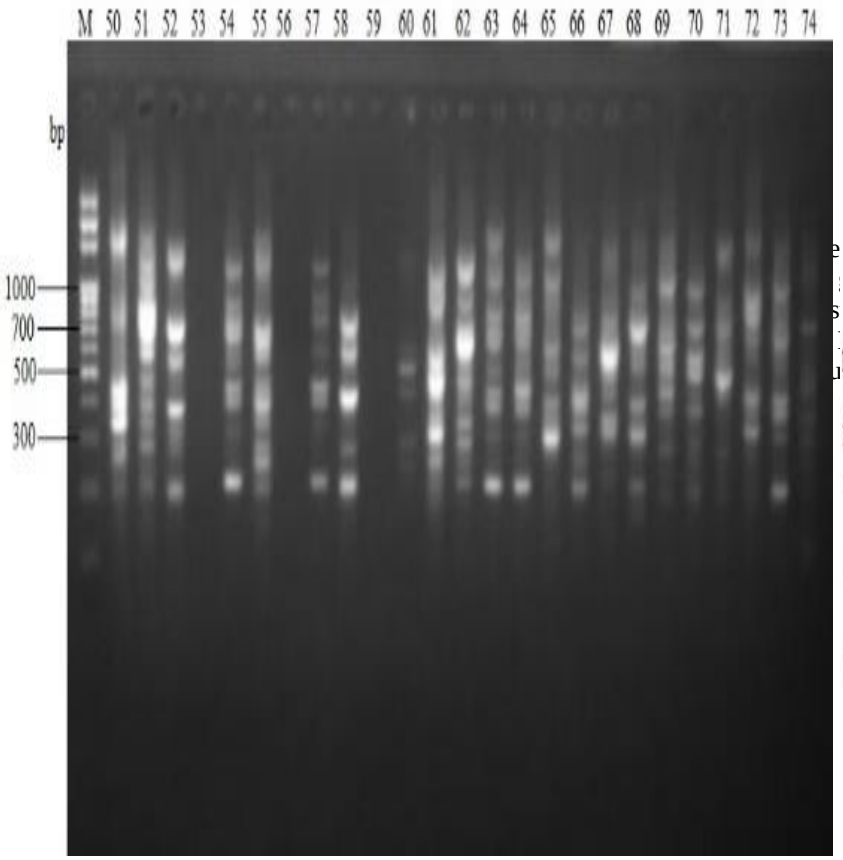


Plate 3a. Gel image of RAPD-PCR amplicons of *Bacillus spp* isolated from waste water and soil samples of abattoirs in Ogun State, Nigeria

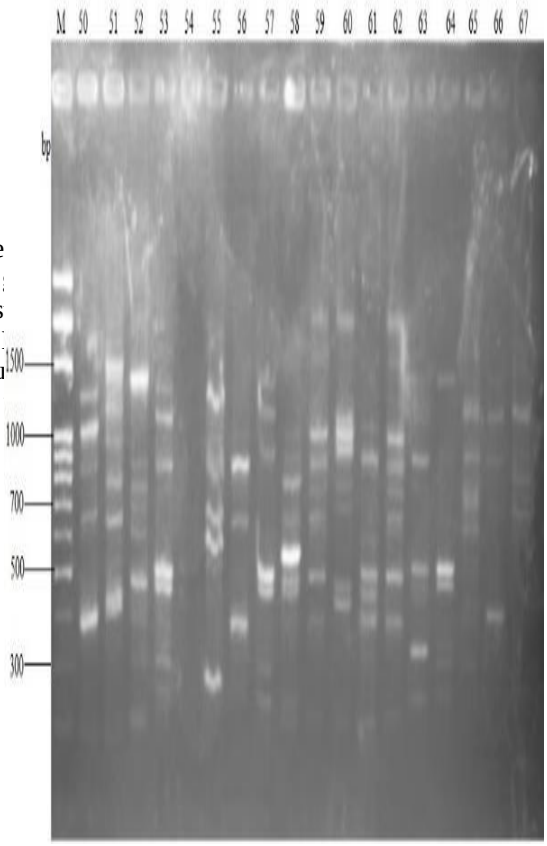


Plate 3b. Gel image of RAPD-PCR amplicons of *Bacillus spp* isolated from waste water and soil samples of abattoirs in Ogun State, Nigeria

https://doi.org/10.5555/UXUB2355 ISSN 2488-8648

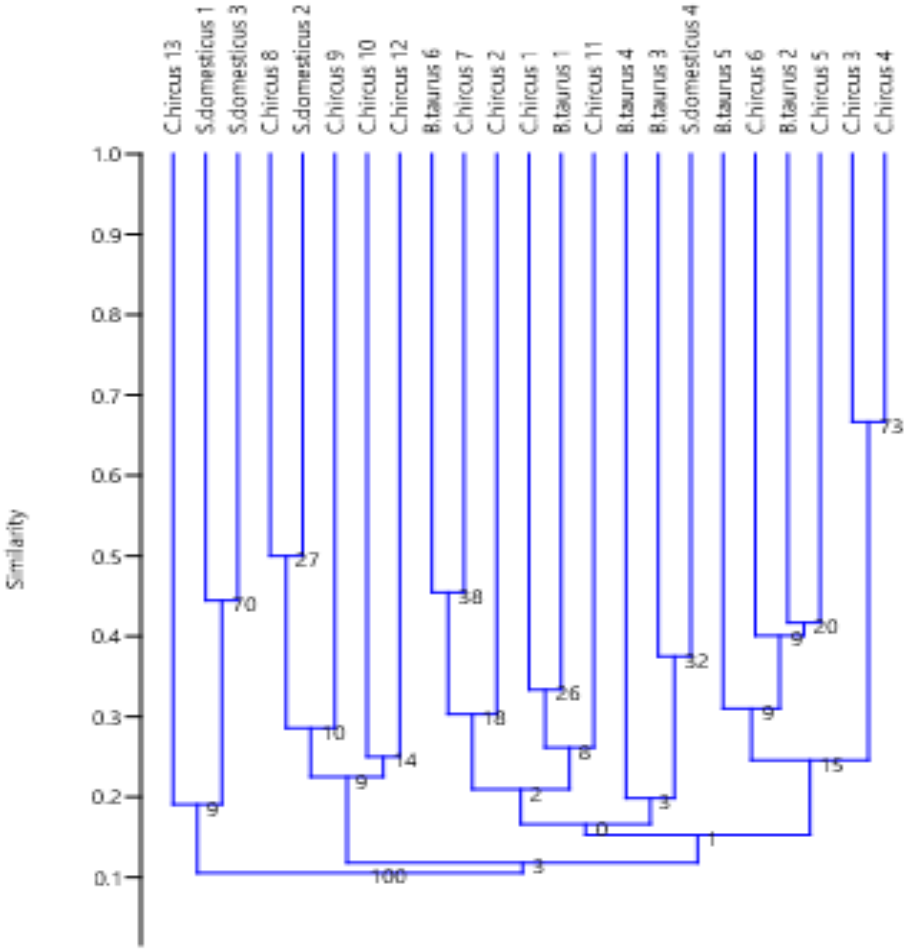


Figure 1a. Phylogenic relationship of *Staphylococcus spp* isolated from Livestock in Ogun State, Nigeria

https://doi.org/10.5555/UXUB2355 ISSN 2488-8648

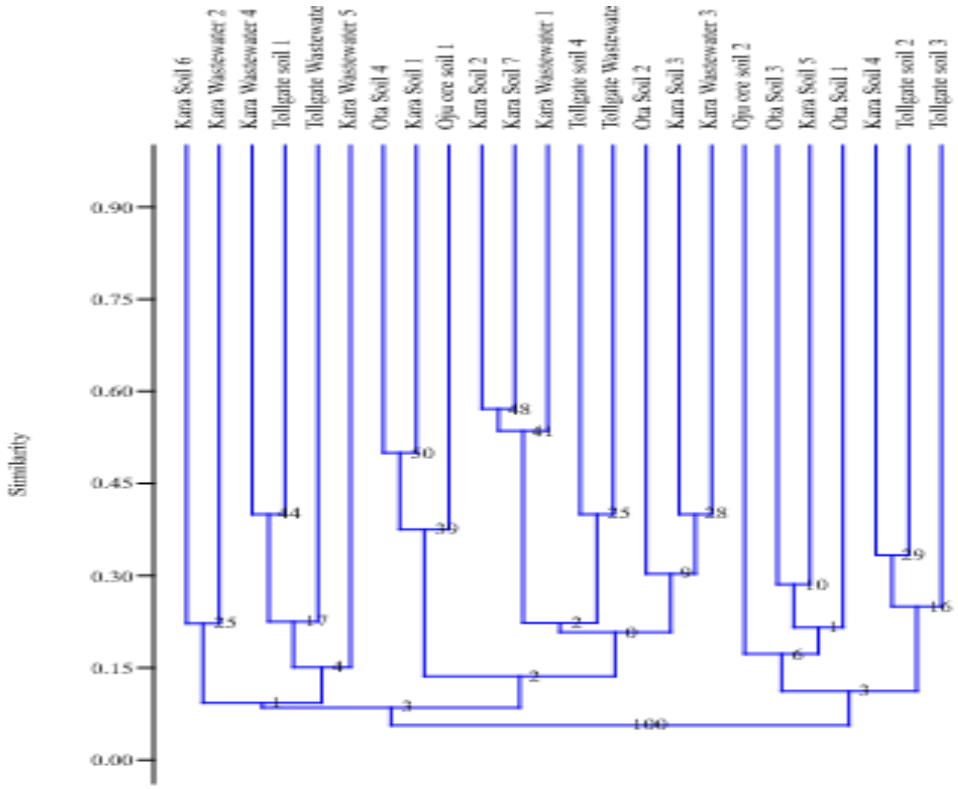


Figure 1b. Phylogenic Relationship of *Bacillus spp* isolated from Soil and Effluents of Abattoirs in Ogun State, Nigeria

Computational Analysis

Pyelph was used to generate phylogenetic trees using the 1-0 matrix result and from the results generated, it was deduced that the *Staphylococcus spp*. from some livestock from Kara, Ogun state formed clusters with livestock from Ota, and Oju-ore in Ogun state. It was also deduced that the *Bacillus spp*. obtained from the wastewater and soil samples from Kara and Oju-ore, had a strong correlation.

Comparative Analysis of the Computational Tools

The images were analysed using the two software programs, and their workflow and output were compared. IMAGE FORMAT: Pyelph works with images in a standard format, e.g. tiff or jpeg, while GelAnalyer does not. The tiff format is a widely used format for gel images and biological images.

Dendrogram Techniques Analysis

This compares the techniques used to analyse the gel images on the different software programs. PyElph analysed the different gel images using different techniques including Unweighted Pair Group Method with Arithmetic (UPGMA), Neighbour-Joining, Weighted Pair Group Method with Averaging (WPGMA), Single Linkage and Complete Linkage.
Table 3a-c highlights the differences between the two computerized tools using the different criteria explained above.

Table 3a: Image Format Comparison among the Computational Tools

	Bmp	gif	Jpg	Jpeg	Png	tiff/tif
GelAnalyzer	✓	✓	✓	✓	✓	✗
PyElph	✓	✓	✓	✓	✓	✓

Table 3b: Band and Lane Detection Technique Comparison among the Computational Tools

	Automatic Band and Lane Detection	Addition /Deletion of Bands and Lanes	Background Removal	The threshold for Band Detection	Band Matching across lanes	Band scoring matrix
GelAnalyzer	✓	✓	✓	✗	✗	✗
PyElph	✓	✓	✓	✓	✓	✓

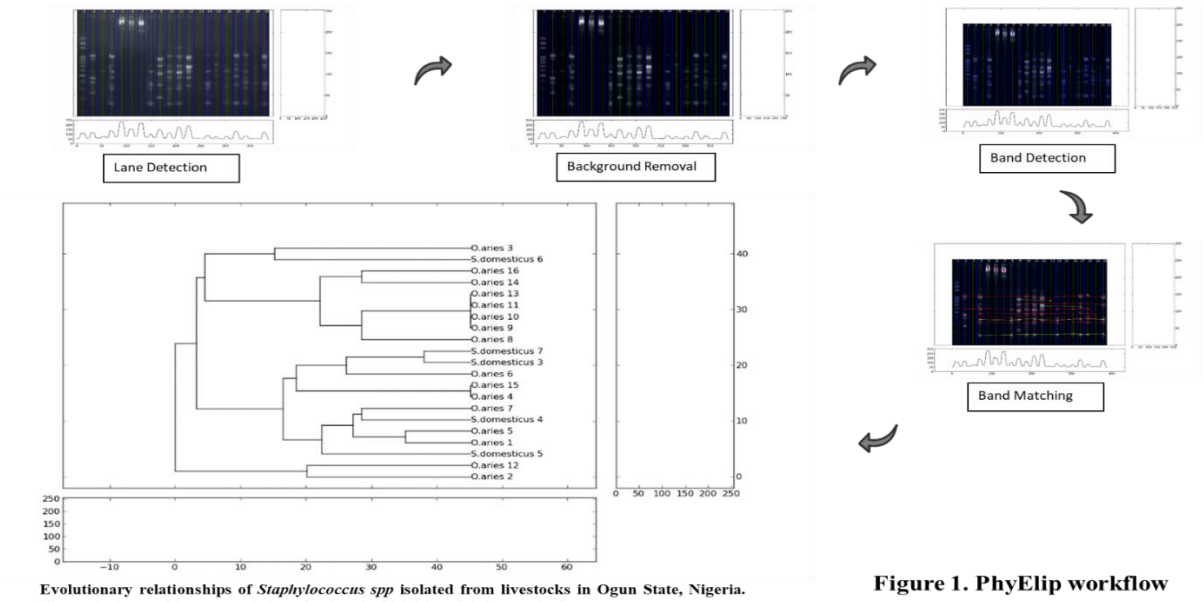
Table 3c: Dendrogram Analysis Techniques Comparison among the Computational Tools

	Dendrogram support	UPGMA	Single linkage	Neighbour-joining	Complete linkage
GelAnalyzer	✗	✗	✗	✗	✗
PyElph	✓	✓	✓	✓	✓

Accuracy of Computational Gel Image Analysis

The gel images from this study were analysed using the PyElph software, and the results were consistent with the results obtained from manual gel analysis. This software program automatically detected approximately 60% of the lanes, with the remaining lanes being manually added. It also effectively detected around 50% of the bands before to manual impute. Based on these data, the software program has been validated to be quite accurate. The workflow process used by this software is highlighted in Figure 1. Whereas the gel images analysed using GelAnalyzer software revealed that the automatic lanes and bands detection had less than 40% accuracy because it did not capture the majority of the lanes and bands; thus, manual inclusion of these dimensions had to be implemented. The molecular weight calculation across bands was also conducted manually with the aid of the 100 bp DNA ladder. When compared (Tables 4) to the analysis required for the gel images, the functionalities given by this software were minimal. The analysis performed on this computational tool was found to have low validity and accuracy

.



Evolutionary relationships of *Staphylococcus spp* isolated from livestocks in Ogun State, Nigeria. The evolutionary history was inferred using the Neighbor-Joining method

Figure 1. PhyElip workflow process

Table 4: Concordance among manual and computerised analyses of RAPD Gel Images

Primers	Samples	Manual Band and Lane Detection	PyElph (Automatic Detection)	GelAnalyzer (Automatic Detection)
OPC 04	*Staphylococcus spp* (Set 1)	Band Average- 23 Lanes- 25	Band Average- 18 Lanes- 22	Band Average- 12 Lanes- 6
	Staphylococcus spp (Set 2)	Band Average- 16 Lanes- 19	Band Average- 16 Lanes- 12	Band Average- 6 Lanes- 10
	Bacillus spp.	Band Average- 18 Lanes- 25	Band Average- 15 Lanes- 23	Band Average- 8 Lanes- 4
RAPD 06	*Staphylococcus spp* (Set 1)	Band Average- 15 Lanes- 23	Band Average- 12 Lanes- 21	Band Average- 10 Lanes- 7
	Staphylococcus spp (Set 2)	Band Average- 15 Lanes- 21	Band Average- 12 Lanes- 15	Band Average- 5 Lanes- 10
	Bacillus spp.	Band Average- 14 Lanes- 18	Band Average- 11 Lanes- 13	Band Average- 9 Lanes- 8

Discussion

The visual interpretation of gel images is an important technique in molecular and genetic studies. This approach is hampered, however, by the inherent difficulties of correct and proper analysis, the time-consuming process of mathematical analysis, as well as the manual construction of dendrograms. Computerized gel analysis has the advantage of rapid analysis and construction of dendrograms, but it is expensive and requires proper training in electronic data processing (Rementeria *et al*., 2001). This has brought about the need for quality, easy-to-use, and inexpensive software for gel image analysis, which may be achieved after adequately evaluating the existing free software widely used by researchers. In this study, two non-commercial, widely used gel image analysis software were compared using various fixed parameters that some researchers denoted as important in comparing computerised programs (Rementeria *et al.,* 2001; Heras *et al.,* 2016). These characteristics included image quality, visual band and lane detection and definition, data clustering, dendrogram analysis, and accuracy of the gel image analysis. Although some discrepancies between these two programs were observed as expected, some general similarities were also highlighted. Both systems needed the user to make selections at various stages of the analysis, such as parameter selection and definition of the molecular marker/ladder in each gel image. Another vital point is that the images must go through a pre-processing phase to be enhanced for accurate analysis before performing a gel analysis. These systems have also been seen to be mostly reliant on the user's input, making the tag automatic gel analysis redundant (Alnamoly *et al*., 2020). When compared to manual analysis, Py Elph displayed inaccurate band detection and low-quality image preprocessing whereas, GelAnalyzer on the other hand, had a lower percentage of accuracy in band and lane detection with lesser functionalities. The inability to automatically detect bands and lanes in low-quality images showed its lack of proper image enhancement feature (Heras *et al*., 2016; Alnamoly *et al*., 2020) In general, free tools implement the basic functionality for analysing gel images. However, they lack features to obtain more accurate results and offer fewer options needed for analysis when compared to commercially available tools (Heras *et al*., 2016). Commercial tools have been seen to enhance the basic functionality with features that improve the precision of studies and the user experience (Alnamoly *et al*., 2020). The limitations of the software programs evaluated highlight the necessity to develop more noncommercial, accurate and user-friendly software capable of analysing gel pictures while using appropriate measures.

This research has comprehensively analysed the most commonly used non-commercial software programs available to researchers and students. This will help identify the features that need to be modified to develop a more effective system for the computational analysis of gel images. Further studies of the existing commercial software programs are required to provide researchers with the necessary knowledge to develop an operational computational tool for gel image analysis.

References

Abeykoon, A., Dhanapala, M., Yapa, R. and Sooriyapathirana, S. (2015). An automated system for analyzing agarose and polyacrylamide gel images. *Ceylon Journal of Science (Biological Sciences*), 44(1): 45-54.

Ahmed, N. (2021). EgyGene GelAnalyzer4: a powerful image analysis software for one dimensional

gel electrophoresis. *Journal of Genetic Engineering and Biotechnology,* 19(1): 18-20.

Alnamoly, M., Alzohairy, A. and El-Henawy, I. (2020). EGBIOIMAGE: A Software Tool for Gel Images Analysis and Hierarchical Clustering. *IEEE Access,* 8(1): 10768-10781.

Chatterjee, S. and Raval, I. (2019). Pathogenic Microbial Genetic Diversity with Reference to Health. *Microbial Diversity in the Genomic Era,* 2(1): 559-577.

Dashti, N., Khanafer, M., El-Nemr, I., Sorkhoh, N., Ali, N. and Radwan, S. (2009). The potential of oil-utilizing bacterial consortia associated with legume root nodules for cleaning oily soils. *Chemosphere,* 74(10): 1354-1359.

Heras, J., Dominguez, C., Mata, E., Pascual, V., Lozano, C., Torres, C. and Zarazaga, M. (2016). A survey of tools for analysing DNA fingerprints. *Briefings in Bioinformatics,*17(6): 903-911.

Kaabouch, N., Schultz, R., Milavetz, B. and Balakrishnan, L. (2007). An analysis system for DNA gel electrophoresis images based on automatic thresholding and enhancement. *IEEE International Conference on Electro/Information Technology,* pp 26-31.

Khakabimamaghani, S., Najafi, A., Ranjbar, R. and Raam, M. (2013). GelClust: A software tool for gel electrophoresis images analysis and dendrogram generation. *Computer Methods and Programs in Biomedicine,* 111(2): 512-518.

Labyed, Y., Kaabouch, N., Schultz, R., Singh, B. and Milavetz, B. (2010). An Improved 1-D Gel Electrophoresis Image Analysis System. *Advances in Experimental Medicine and Biology,* 680: 609-617.

Maramis, C. and Delopoulos, A. (2010). Efficient Quantitative Information Extraction from PCRRFLP Gel Electrophoresis Images. *Proceedings of the 2010 20th International Conference on Pattern Recognition,* 2(1): 2560–2563.

Pavel, A. and Vasile, C. (2012). PyElph - a software tool for gel images analysis and phylogenetics. *BMC Bioinformatics,* 13(1): 40-49.

Rementeria, A., Gallego, L., Quindos, G. and Garaizar, J. (2001). Comparative evaluation of three commercial software packages for analysis of DNA polymorphism patterns. *Clinical Microbiology and Infection,* 7(6): 331-336.