Mar 17, 2021

Calculating number of Co-Migrated DNA Fragments across Pulsed Field gel electrophoresis (PFGE) profiles: Image analysis algorithm

Ibrahim El-khalil Adam[1]

[1]department of zoology, faculty of Science, University of Khartoum

In Development This protocol is published without a DOI.

Ibrahim El-khalil Adam

SUBMIT TO PLOS ONE

ABSTRACT

The author describes a new image analysis algorithm that enables identification of how many DNA fragments co-migrate during PFGE. The method is named factor of co-migration based on exponential correlation between single-fragment bands and their pixel densities "FCM-ECSB".

EXTERNAL LINK

https://www.researchgate.net/publication/348389960_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

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

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: 52–67. doi:10.1101/2020.07.05.188623

ATTACHMENTS

diagram.pdf JAMBR-139 final manuscript.pdf

EXTERNAL LINK

https://www.researchgate.net/publication/348389960_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

PROTOCOL CITATION

Ibrahim El-khalil Adam 2021. Calculating number of Co-Migrated DNA Fragments across Pulsed Field gel electrophoresis (PFGE) profiles: Image analysis algorithm. **protocols.io** https://protocols.io/view/calculating-number-of-co-migrated-dna-fragments-ac-bteknjcw

MANUSCRIPT CITATION please remember to cite the following publication along with this protocol

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: 52–67. doi:10.1101/2020.07.05.188623

KEYWORDS

Outbreak investigation, Food-borne disease, PFGE, WGS, PulseNet international

CREATED

Mar 17, 2021

LAST MODIFIED

Mar 17, 2021

PROTOCOL INTEGER ID

48300

GUIDELINES

It is well known that in order to estimate a DNA band size in any electrophoresis technique, a molecular weight marker is run alongside test DNA. The same rule applies for our method. Factor of co-migration is defined as 'The number of DNA fragments that appear as a single band in the gel due to resolution limitations of the gel or simply, because more than one band have the same length'.

Image analysis algorithm of this protocol is named " factor of co-migration based on exponential correlation between single-fragment bands and their pixel densities (FCM-ECSB)"

DISCLAIMER:

The described FCM-ECSB method is patented to the first and second authors of the cited article. Accordingly, the use of any one of previously mentioned methods and/or algorithms without permission for upgrading image analysis software or any bioinformatics tool, creating another web or any offline software that consume any of the methods/algorithms is considered a financial conflict of interests. However, It is OK to use this protocol by means of described methods.
The full article is available

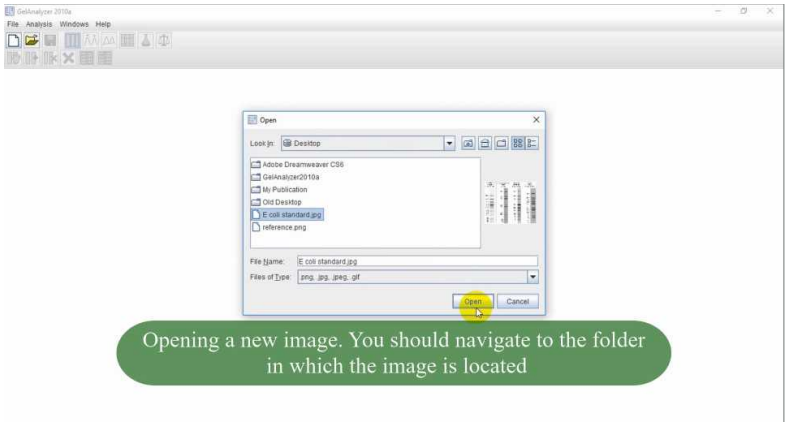
BEFORE STARTING

This protocol requires gel image analysis software. The program should be able to provide band intensity profiles alongside band sizes and marker exponential equation (correlation coefficient: R^2).
For this demonstration, GelAnalyzer 2010a will be used. Make sure that you have java runtime environment (JRE) is installed on your system. for more information on how to use this software, please refer to its user manual here: gelanalyzer.com/downloads/users_manual_2010.pdf
For FCM-ECSB calculations, raw pixel density data alongside band sizes should be transferred to a statistical package (Paleontological Statistics: PAST will be used) in order to obtain the correlation and regression equation for calculating expected pixel density (EPD). Finally, MS Excel will be used to calculate the expected values of FCM.

IMAGE PROCESSING

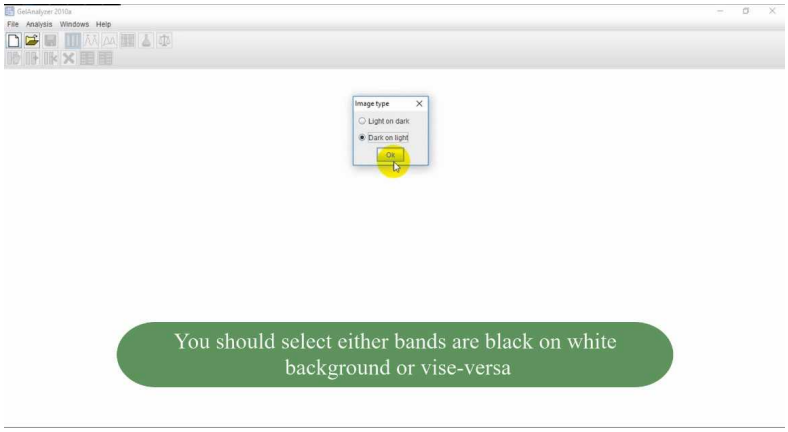
1 Launch the program

1.1 Open PFGE image of interest. You should navigate to containing folder first.

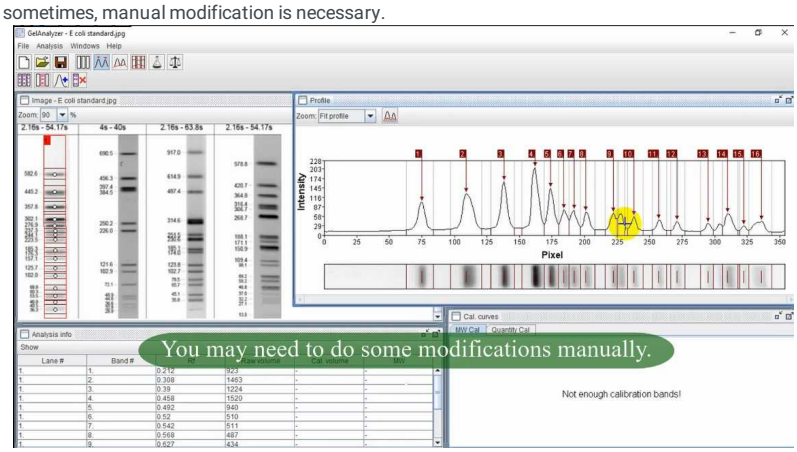


1.2

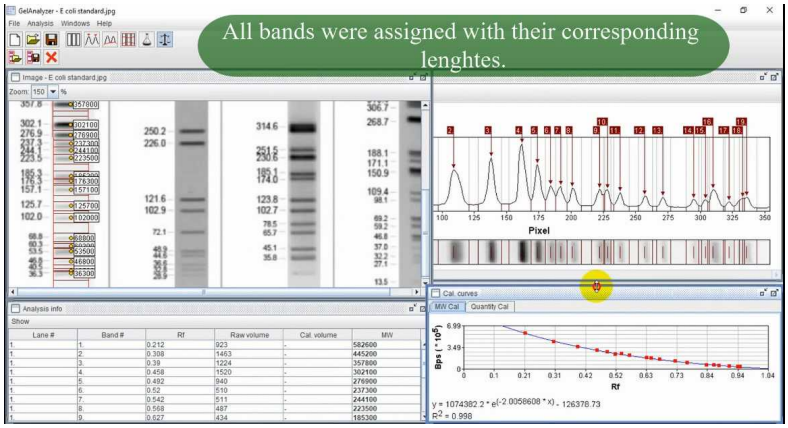
It is critical to tell the program if DNA bands are white in a black background or otherwise.



1.3 Define lanes and then bands.
sometimes, manual modification is necessary.

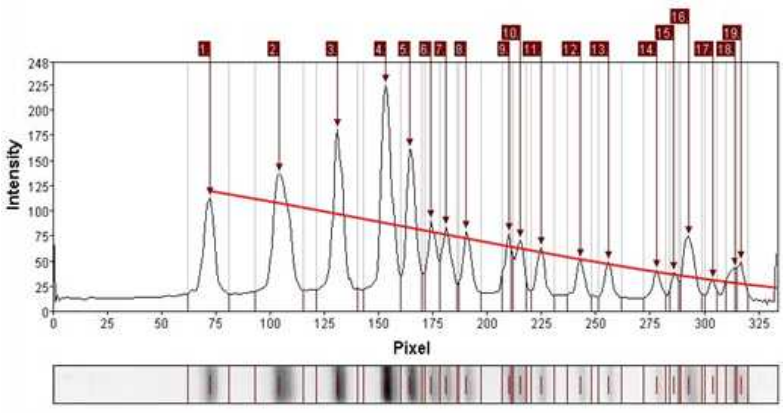


1.4 Select the lane that shows the DNA ladder. then indicate each band with its corresponding size. (**Sizes are in kilo base-pair (kbp). you should multiply each value by 1000 to get accurate results**).



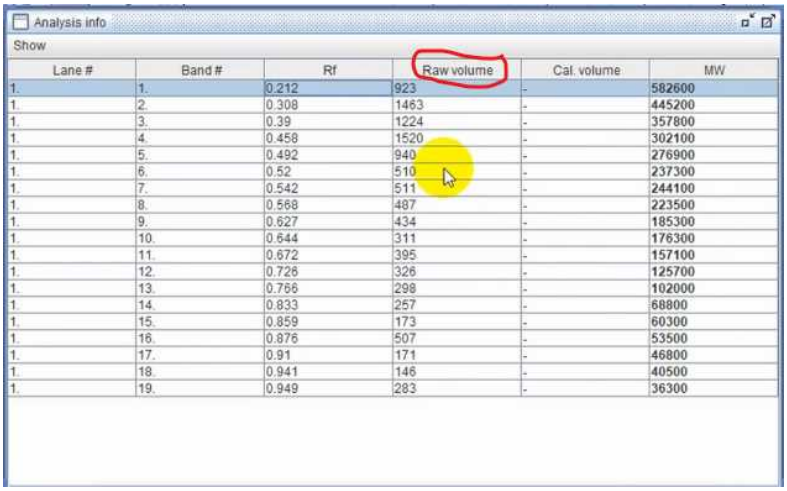
2 CALULATING EXPECTED PIXEL DENSITIES FOR CO-MIGRATED DNA FRAGMENTS

This step is critical. it is based on end-user's estimation. Keep in mind that co-migrated fragments show high pixel densities (Intensity profiles). band intensities should reduce as you go down the profile (toward smaller band sizes).



The red line roughly indicate expected band intensities across the whole profile.

2.1 Copy all the data from "Analysis info." from GelAnalyzer2010a. Press Ctrl + A to copy data.



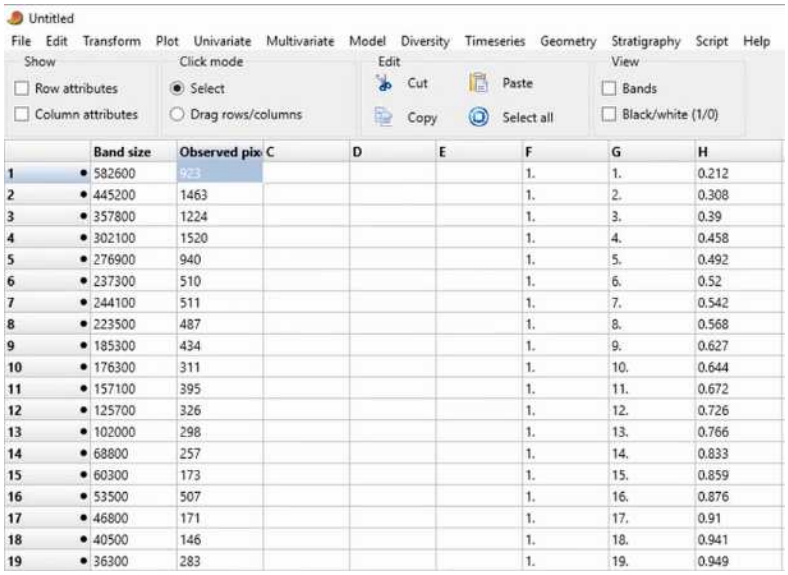
Analysis info

Show

Lane #	Band #	Rf	Raw volume	Cal. volume	MW
1.	1.	0.212	923	-	582600
1.	2.	0.308	1463	-	445200
1.	3.	0.39	1224	-	357800
1.	4.	0.458	1520	-	302100
1.	5.	0.492	940	-	276900
1.	6.	0.52	510	-	237300
1.	7.	0.542	511	-	244100
1.	8.	0.568	487	-	223500
1.	9.	0.627	434	-	185300
1.	10.	0.644	311	-	176300
1.	11.	0.672	395	-	157100
1.	12.	0.726	326	-	125700
1.	13.	0.766	298	-	102000
1.	14.	0.833	257	-	68800
1.	15.	0.859	173	-	60300
1.	16.	0.876	507	-	53500
1.	17.	0.91	171	-	46800
1.	18.	0.941	146	-	40500
1.	19.	0.949	283	-	36300

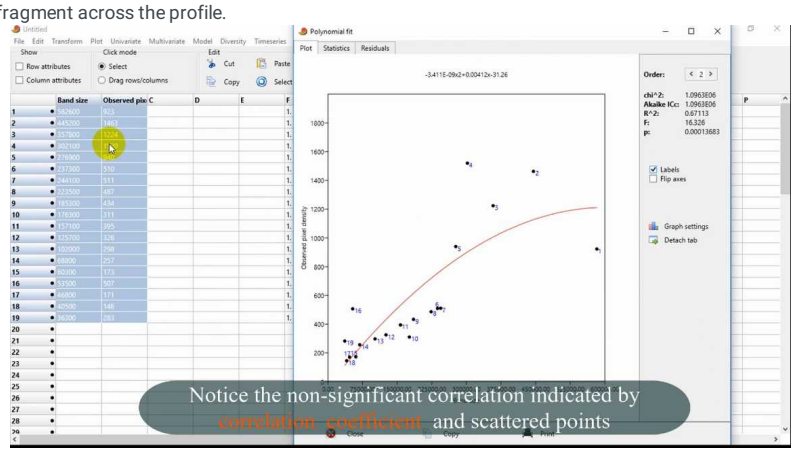
Notice that "Raw volumes" are intensity profiles corresponding each band size "MW".

2.2 Open PAST statistics. The goal is to get a polynomial equation with significantly high R^2 value. Past the data copied from GelAnalyzer2010a and past it into PAST statistics. Change column attributes to indicate (Band size) and (Pixel density).



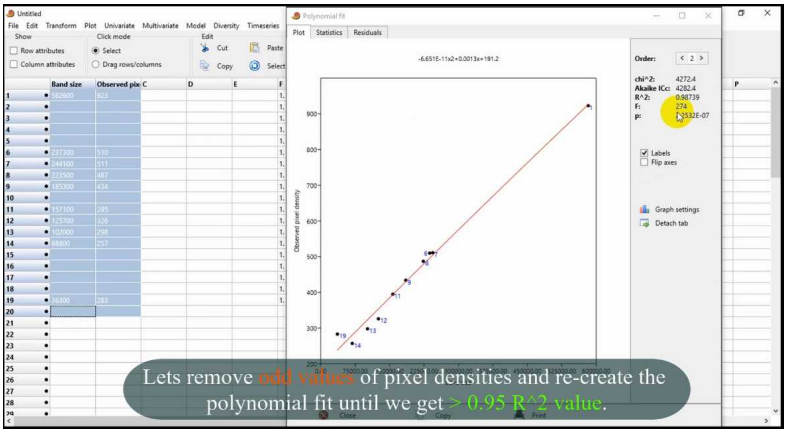
	Band size	Observed pix	C	D	E	F	G	H
1	582600	923				1.	1.	0.212
2	445200	1463				1.	2.	0.308
3	357800	1224				1.	3.	0.39
4	302100	1520				1.	4.	0.458
5	276900	940				1.	5.	0.492
6	237300	510				1.	6.	0.52
7	244100	511				1.	7.	0.542
8	223500	487				1.	8.	0.568
9	185300	434				1.	9.	0.627
10	176300	311				1.	10.	0.644
11	157100	395				1.	11.	0.672
12	125700	326				1.	12.	0.726
13	102000	298				1.	13.	0.766
14	68800	257				1.	14.	0.833
15	60300	173				1.	15.	0.859
16	53500	507				1.	16.	0.876
17	46800	171				1.	17.	0.91
18	40500	146				1.	18.	0.941
19	36300	283				1.	19.	0.949

2.3 Create a preliminary polynomial fit. R^2 value at this point roughly indicates number of co-migrated fragment across the profile.



Notice that R2 value is 0.67113.

2.4 Refer to step 2 (remove values above the red line). Recreate the polynomial fit and check R^2 value. Remove band sizes and their corresponding pixel densities from both columns (**According to ODD values of pixel densities)**.



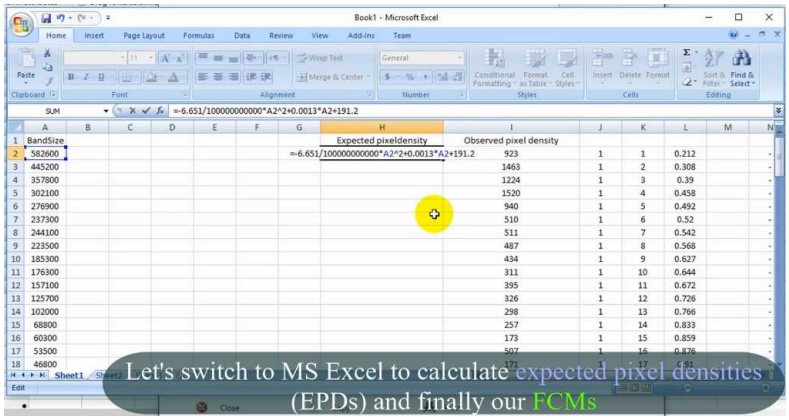
R2 now is 0.98739. Bands removed; 2 to 5, 10 and15 to 18.

The polynomial equation we need is:

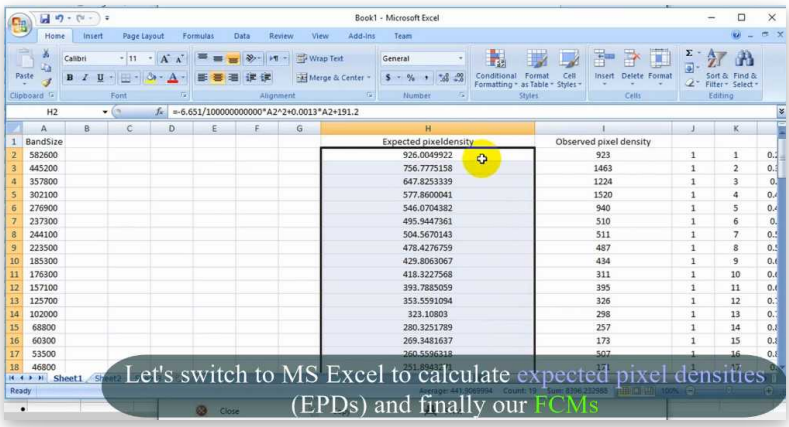
$$EPD = -6.651E-11 * BS^2 + 0.0013 * BS + 191.2$$

EPD: expected pixel density, BS: band size

3 DENOTATION OF ALL BAND SIZES INTO TE EQUATION (from step 2.4):
This step requires the same two columns from step 2.2. Open Microsoft Office™ Excel. Past the two previously mentioned columns. Start denotation of column "Band size" into the equation. Use an empty column.

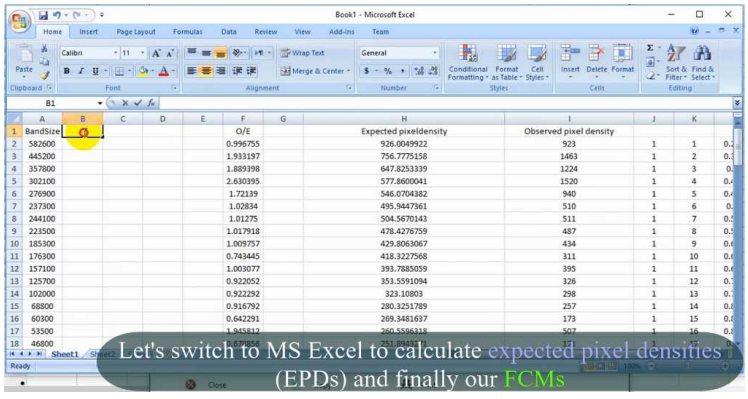


Denotation process.

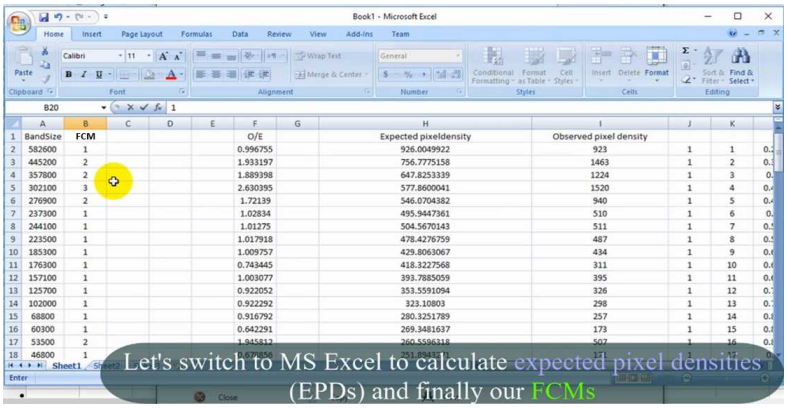


Expected pixel densities are ready. notice how close expected and observed PDs are close for bands number 1,8, 9 and 11. we can safely conclude that they represent single DNA fragments. While expected PD of band number 2 is half the observed one.

3.1 In order to calculate approximate number of DNA fragments represented by each PFGE band, we must simply divide observed PDs by the corresponding ones. column is named "O/E" indicating observed/expected.



3.2 FCMs are simply truncated values of O/E column.



Here is a screen-recorded video showing the entire protocol in action :