ThermoFisher SCIENTIFIC
The world leader in serving science

TECHNICAL NOTE

GeneMapper™ PG Software: Creating a New Model

Introduction

GeneMapper™ PG is probabilistic genotyping software which enables users to estimate the Number of Contributors (NOC), deconvolute, and perform likelihood ratio (LR) calculations on mixed evidence samples. The software achieves this through a detailed statistical model of the STR kit used to run the samples. This technical note provides instructions on how to create a new model in GeneMapper™ PG Software v1.0.

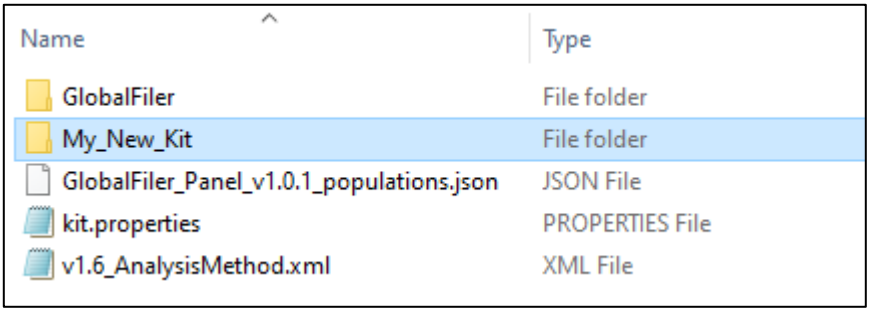
Creating a New Model

In GeneMapper™ PG Software, models are organized in folders according to the STR kit and capillary electrophoresis instrument that was used to generate the data. By default, models are located in C:\Thermo Fisher Scientific\GeneMapper PG\panels. Within this folder there is a folder for each STR kit that has a model. Within these 'kit' folders are sub folders for instrument types. Within these instrument folders are the model files themselves.

To create a new model for a new kit:

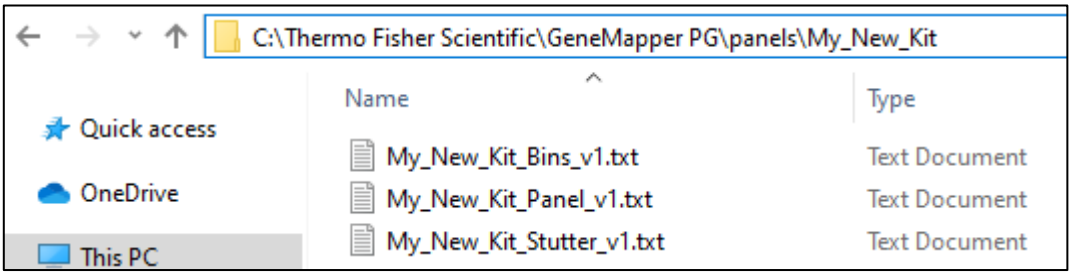
1. Close GeneMapper™ PG Software, if it is already open.

2. Create a new folder for the new kit in C:\Thermo Fisher Scientific\GeneMapper PG\panels. For example, 'My_New_Kit'

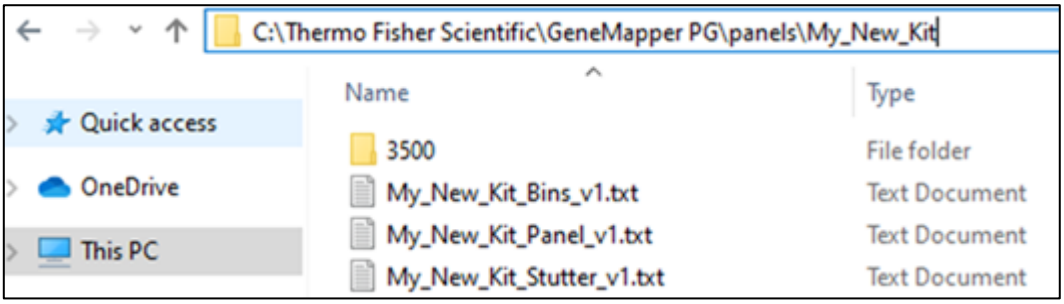


3. Retrieve the GeneMapper™ ID-X panel, bin and stutter definition files for your kit. By default, these are saved at C:\Applied Biosystems\GeneMapperID-X\Panels on a computer that has GeneMapper™ ID-X installed. For further information on retrieving these files, refer to GeneMapper™ ID-X help documentation.[1]

4. Place the GeneMapper™ ID-X panel, bin and stutter files into the folder created in Step 1.



5. Open the folder for another kit in the ‘panels’ folder (e.g. the default ‘GlobalFiler’ folder) and make a copy of an instrument folder from inside that other kit folder. Paste a copy into the new folder created in Step 1. For example, copy the ‘3500’ folder from C:\Thermo Fisher Scientific\GeneMapper PG\panels\GlobalFiler into C:\Thermo Fisher Scientific\GeneMapper PG\panels\My_New_Kit. Repeat this step as needed if multiple models for multiple instrument types are required for the new kit.



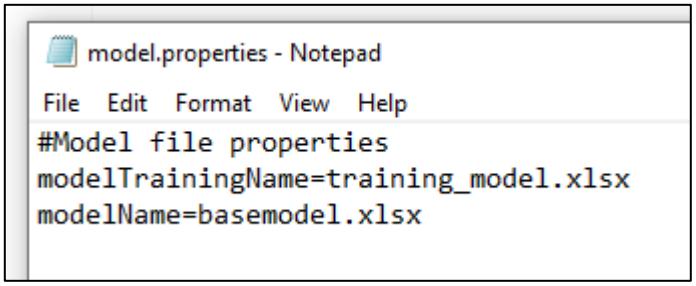
6. Open the new instrument folder made in Step 5 (for example, the folder called '3500')

7. Delete all files in this folder except for:
    a. basemodel.xlsx
    b. model.properties
    c. dropout.csv
    d. variability.csv

Note: The dropout.csv and variability.csv files are not used by the software, but the software will still check for their existence. Keep the unchanged copies from the other kit in this folder.

---

8. The model.properties file contains the name for the model itself and for the 'training' copy of the model that is used if Training mode is active in GeneMapper™ PG Software. These names can be kept at the default values or can be edited in a text editor if desired.



9. Open the basemodel.xlsx file and the new 'panel'.txt file (i.e. the file called 'My_New_Kit_Panel_v1.txt' in the example above) in Microsoft Excel.

10. Switch to the 'Statistical Model' tab in basemodel.xlsx and compare the two files side-by-side. Edit values in basemodel.xlsx to match the properties of the new kit. That is:
    a. Change the name of the Model (in Cell A1 of the 'Statistical Model' tab in basemodel.xlsx) to match the name of the panel for the new kit.
    b. Remove any loci from basemodel.xlsx that are not in the new kit.
    c. Add new rows to basemodel.xlsx for any loci that are in the new kit but are not in basemodel.xlsx.

d. Ensure that the ‘Color’ and ‘Repeat Size’ values in basemodel.xlsx match those listed in the 'B' and 'F' columns of the panel file for every locus.
e. Ensure that the ‘Size from’ and ‘Size To’ values in basemodel.xlsx match those listed in the 'C' and 'D' columns of the panel file for every locus.
f. Set the ‘PAT’, ‘#Inherited Alleles’ and ‘Type of locus’ values in basemodel.xlsx to appropriate values for each locus.

Note: It is recommended for PAT (peak analytical threshold) to determined by each laboratory by internal validation.

---

11. Enter appropriate data for the new kit on each of the tabs in basemodel.xlsx. The tab names are:
    - Back Stutter
    - Double Back Stutter
    - Forward Stutter
    - Half Back Stutter
    - Half Forward Stutter
    - Shoulder
    - Noise
    - Peak Height Ratio
    - Dropout model X cycles

    Check the following items for these tabs:

    - Ensure the correct loci are listed: add or remove loci as needed to match the new kit.
    - For all the stutter data tabs:
        - Ensure that the correct alleles for each locus are listed: check the bins in the ‘bins’ file (e.g. the file called ‘My_New_Kit_Bins_v1.txt’ in the example above) to ensure that these match. All alleles listed in the model must be in the bins file.
        - The top row for each locus (marked ‘(Locus)’) is the mean stutter value for the entire locus. This value is used for filtering if the allele-specific value is not supplied for any allele in the bin set.
        - Column ‘E’ (named, for example, ‘Back Stutter Cap % (5 stddev)’) is the value that is used by the software for filtering of stutter at that locus and allele. Ensure that the desired values are used for each entry.

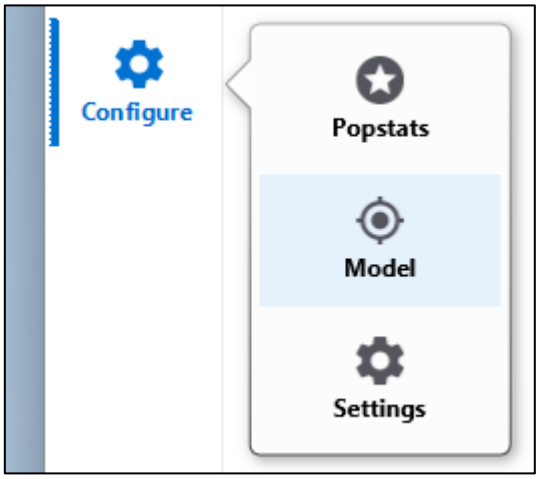
- The data in the Noise and Peak Height Ratio Tabs is listed in bands of peak height. The value used by the software depends on the average peak height in the sample to be analyzed. As for the stutter data, the value in column ‘E’ (the ‘Max’ value) is the one used for filtering by the software.
- For the tab described above as ‘Dropout model X cycles’, X represents the number of PCR cycles used to generate the data. Replace ‘X’ with the relevant number of PCR cycles for the new model and ensure that this matches the number of cycles used to generate the unknown samples that will be analyzed with the new model.

Note: The validity of the data entered when a new model is created is the responsibility of the user. We recommend that appropriate testing and validation is performed when developing a new model.

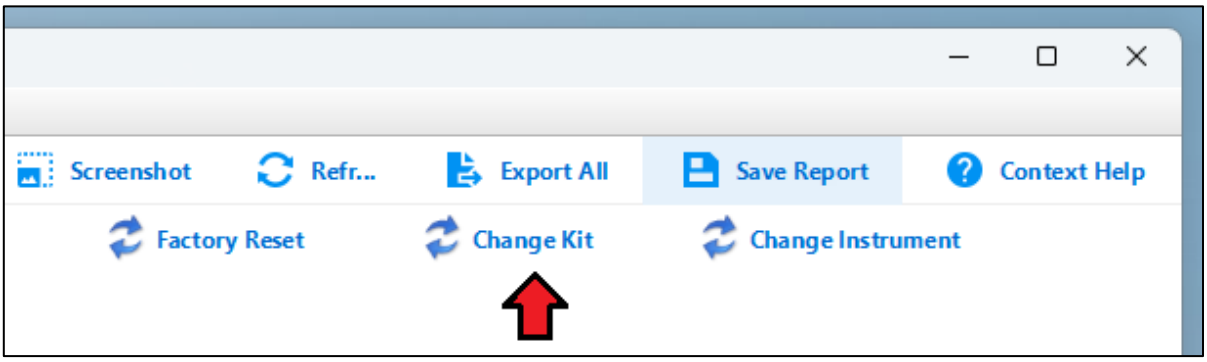
---

12. Once you have the desired data in all the tabs, save and close the basemodel.xlsx file.

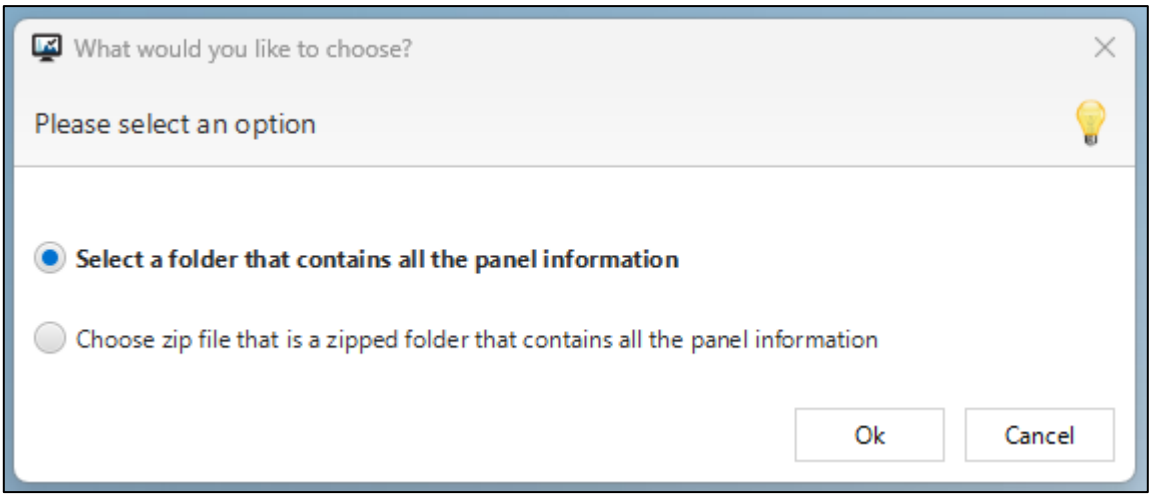
13. To use the new model, start GeneMapper™ PG Software and select ‘Configure’ → ‘Model’.



14. Select ‘Change Kit’.

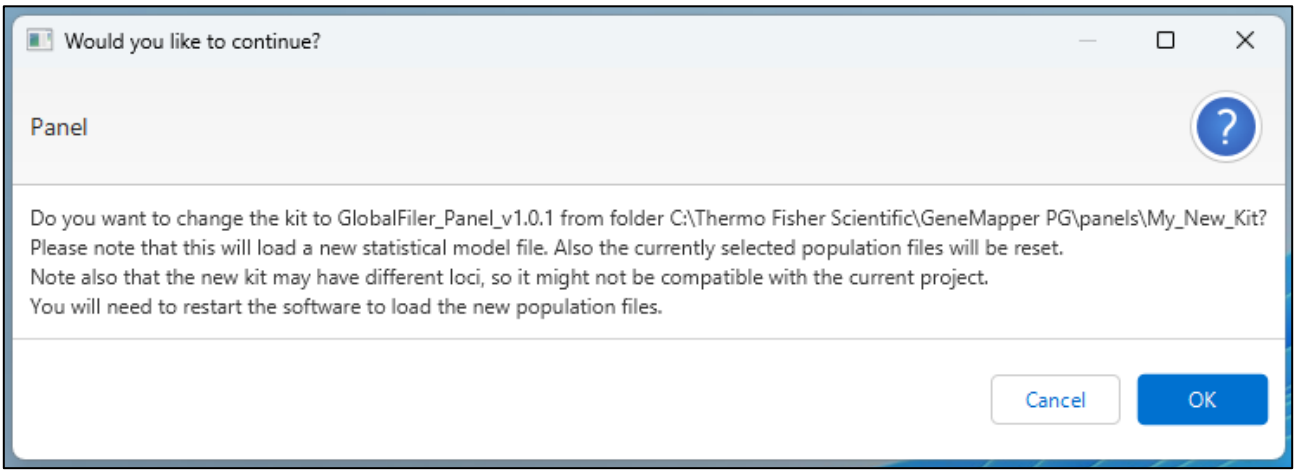


15. Click on ‘Select a folder that contains all the panel information’ in the window that appears. Click ‘Ok’.

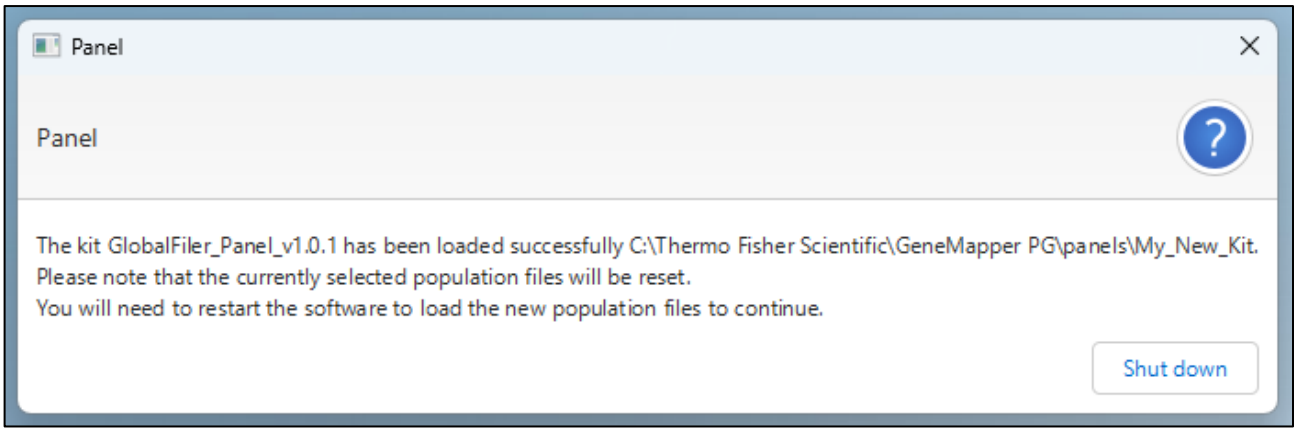


16. Choose the folder you created in Step 1 in the window that appears. Click ‘Select Folder.’

17. If the files have been correctly configured, you will see a confirmation message like the example below. Click ‘Ok’ to proceed.



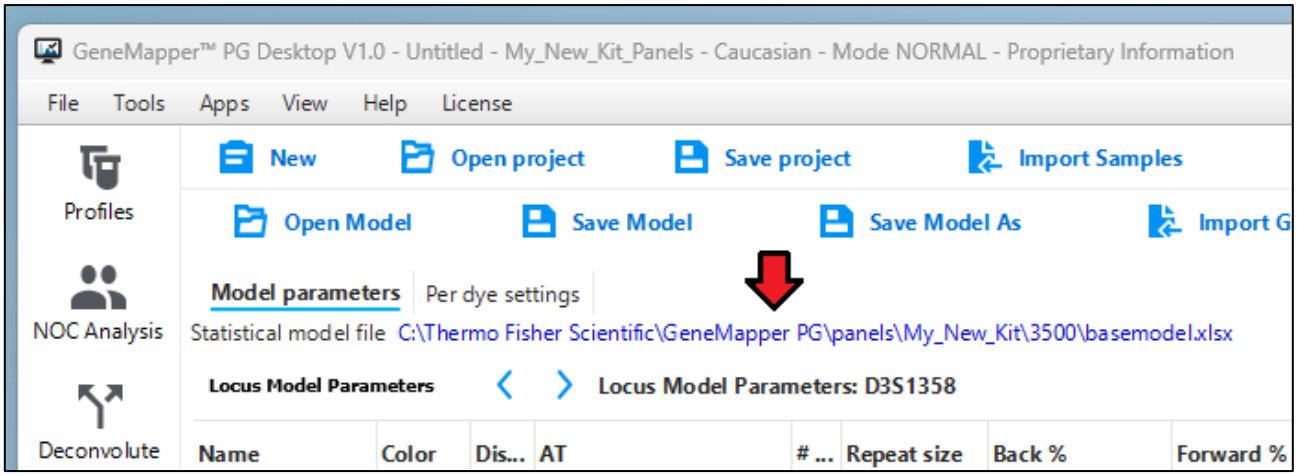
18. You will then get a message confirming that the new model is loaded and prompting you to restart the software. Click ‘Shut down’ to close the software.



19. Restart the software.

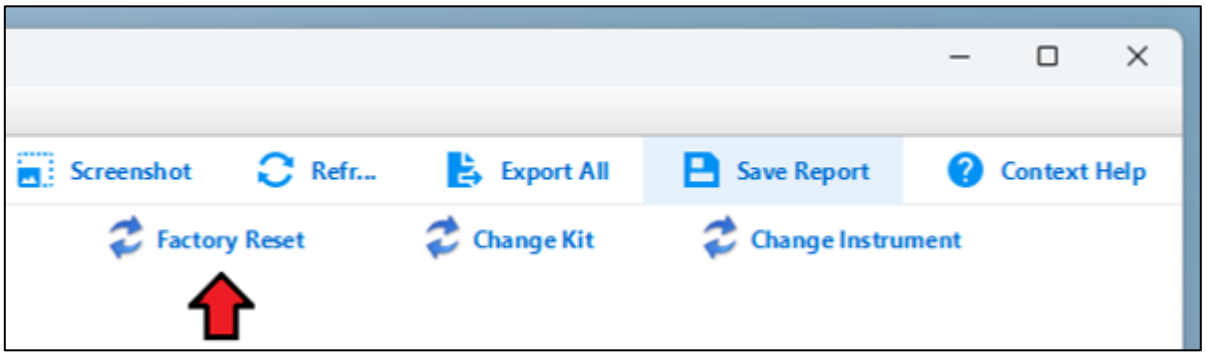
20. Navigate to the model screen again (select ‘Configure’ → ‘Model’).

21. Confirm that the correct data has been loaded and that the name of the new folder/model is visible in the ‘Statistical model file’ link at the top of the screen.



22. Check that the population data is set appropriately (select ‘Configure’ → ‘Popstats’). The new kit may require new allele frequencies to be uploaded to provide data for all markers in the kit. Refer to the in-software Help for guidance on managing population data if needed.

23. An optional final step that can be performed is to make a copy of the new basemodel.xlsx file and name the copy ‘basemodel_original.xlsx’. This file will be used if the ‘Factory Reset’ button on the model configuration screen is ever used. If this is clicked, the model will revert to the settings in basemodel_original.xlsx.



Note that after first use, the software will automatically make a 'basemodel.xlsx.json' file in the new model folder. This is used by the software for caching only. Any changes made to the basemodel.xlsx file, either directly in the file or via the GeneMapper™ PG Software UI, will be used by the software. The user does not need to interact with the json file.

24. The new model is now ready for use.

References

1. GeneMapper ID-X v1.5 Basic Features. 100031701. Revision B. Available from Thermo Fisher Scientific here.

Revision History

Revision	Date	Description
A00	05 SEP 2025	Initial publication