Manual

KlusterCaller and Kraken Plate Data Import Wizard

For Research Use Only. Not for use in diagnostic procedures.

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Manual

KlusterCaller and Kraken Plate Data Import Wizard

Contents

1. Introduction 3
2 Before you start 3
3 Importing a data file 4
  3.1 Using KlusterCaller 4
  3.2 Using Kraken 8
4 The plate reader configuration wizard – specific examples 10
  4.1 Import of 'List format #1' 12
  4.2 Import of 'List format #2' 23
  4.3 Import of 'Grid format' 35
5 Summary 46
6 Further support 46

For Research Use Only. Not for use in diagnostic procedures.

Manual

KlusterCaller and Kraken Plate Data Import Wizard

1. Introduction

This document is intended as a guide to importing genotyping data into KlusterCaller/Kraken using the inbuilt 'Plate reader configuration wizard'. Data from both qPCR instruments and plate readers can be imported into the software, and the wizard will enable the user to teach the software to recognise their specific file format. Once the plate reader configuration wizard has been successfully completed/run, the software will be able to directly import all future files of this format.

2. Before you start

Before attempting to import your raw data into KlusterCaller/Kraken, it is essential to save the data in a format that the software can recognise. Acceptable file types are:

- Comma separated values file (.csv)
- Text/Tab-delimited (.txt)
- MS Excel 97-2003 (.xls)
- Data file (.dat) (N.B. File type specific to BMG OMEGA & PHERAstar plate readers)

If you are using a qPCR instrument, you will need to export the data from the qPCR instrument software and save it as one of the above file types. It is not possible to import data directly from a qPCR instrument software file (e.g. CFX Manager file, SDS 2.4 file). If your raw data is in MS Excel (post-2003), you can 'Save As' and choose one of the file formats listed above.

IMPORTANT: when saving your plate reader file(s) for import into KlusterCaller/Kraken, **do not** use dots (.) in the middle of the filename. Use of dots prevents the file from being imported correctly. For example "PLATE1234.01.01.2024.csv" will not be imported correctly whereas "PLATE1234_01_01_2024.csv" will import without an issue.

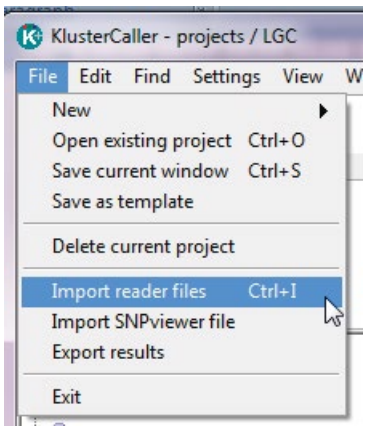
For Research Use Only. Not for use in diagnostic procedures.

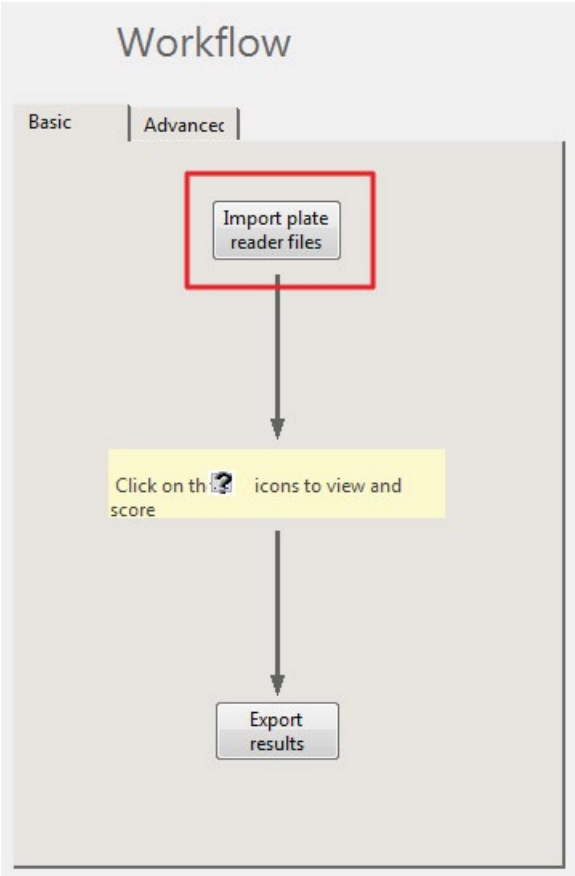
3. Importing a data file

To import a new format of data file into either KlusterCaller (section 3.1) or Kraken (section 3.2), you will first need to navigate to the 'plate reader configuration wizard'.

3.1 Using KlusterCaller

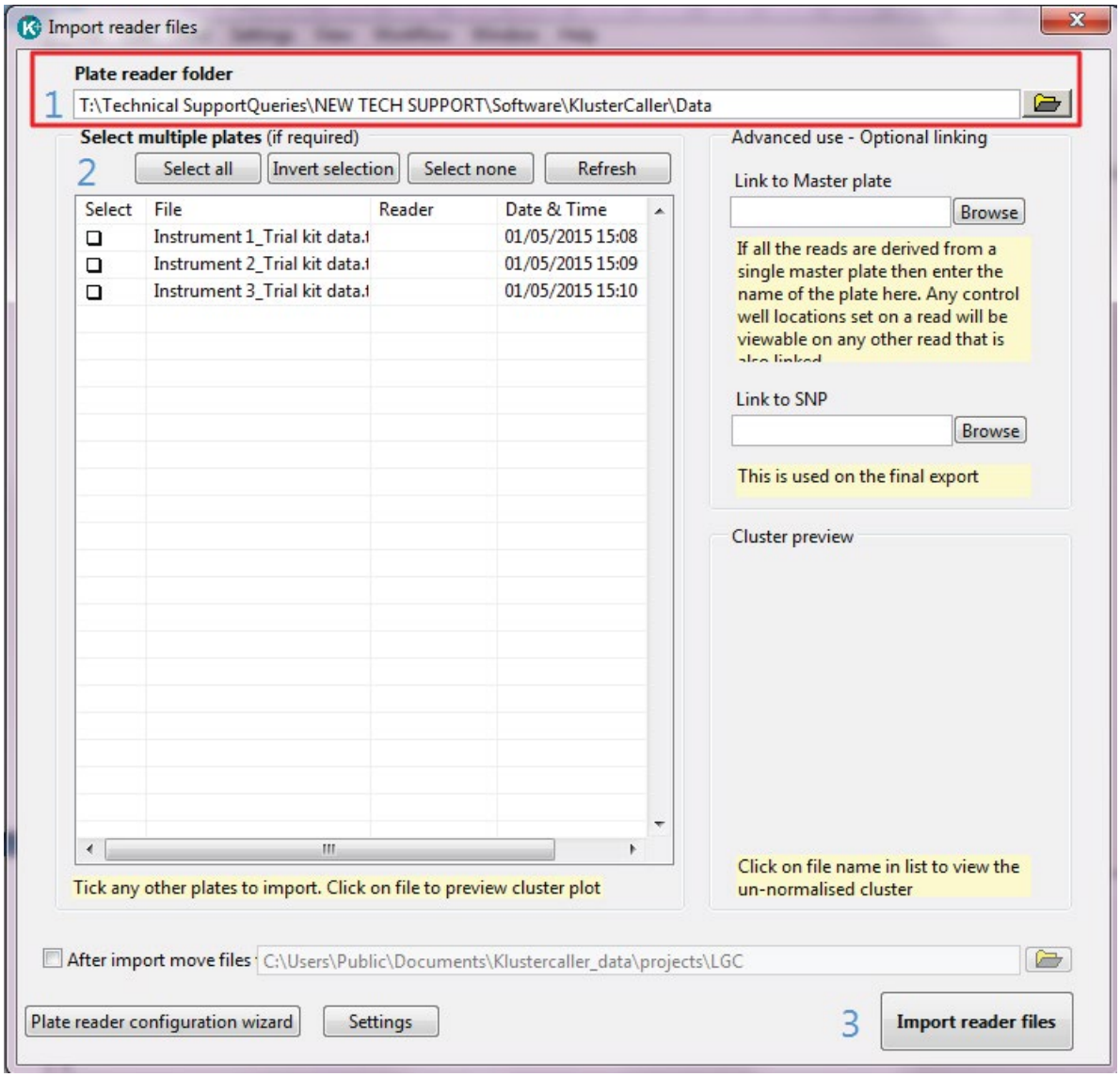
a. Click on the 'File' menu and select 'Import Reader Files' or click on the 'Import Plate Reader Files' button in the 'Workflow' view.





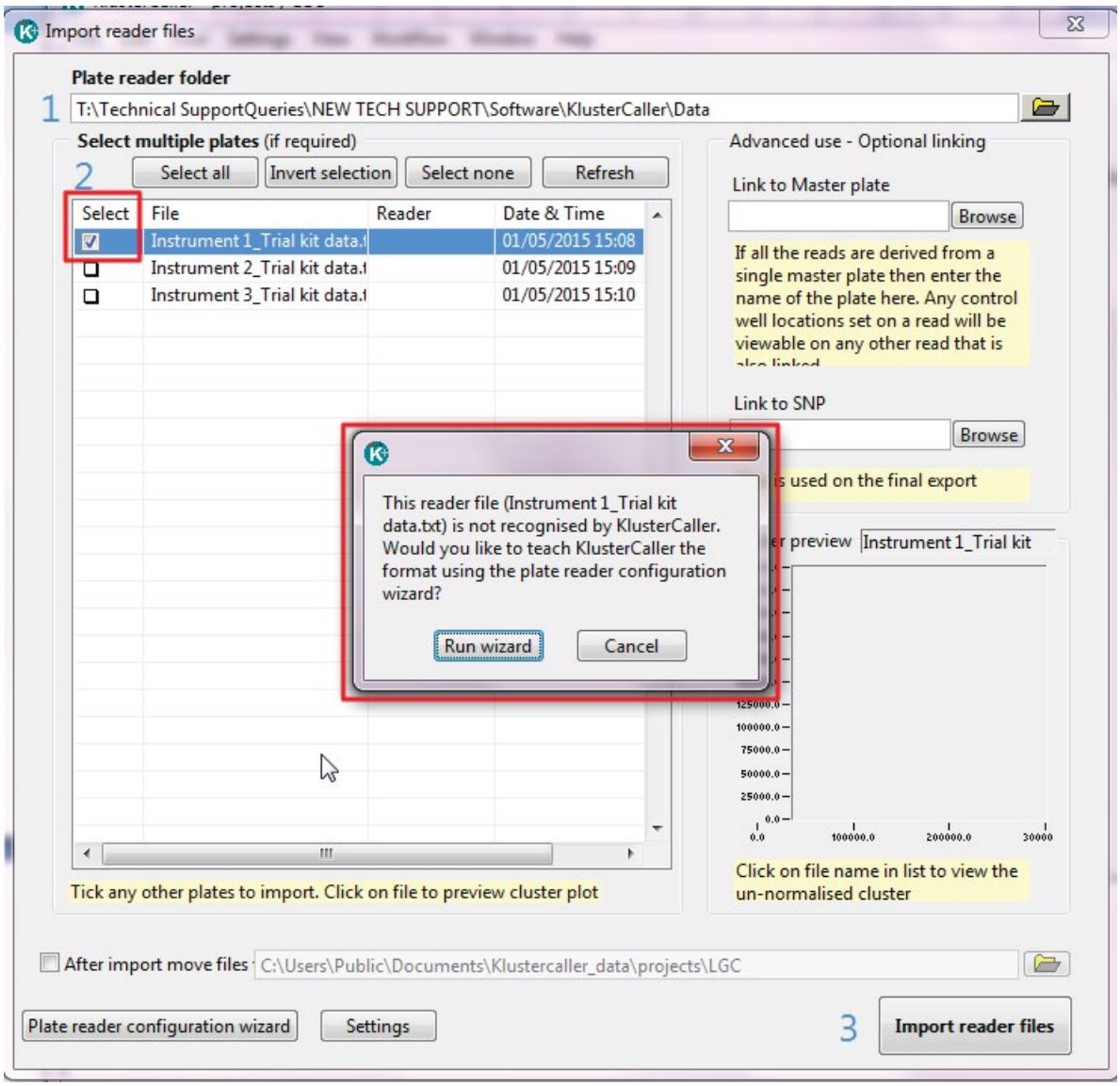
For Research Use Only. Not for use in diagnostic procedures.

b. In the ‘Import reader files’ window, navigate to the folder that contains the raw data file and press OK.



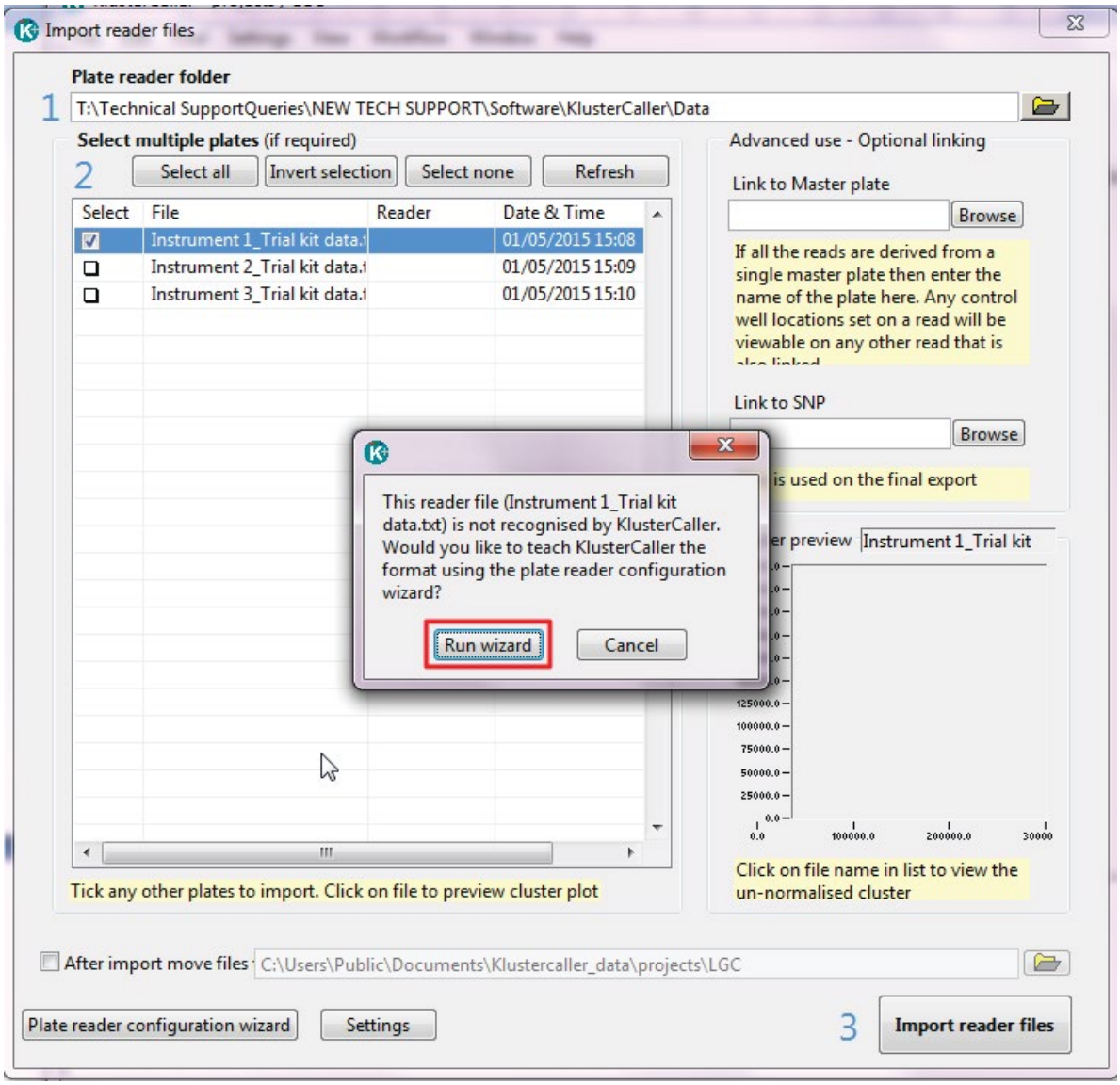
For Research Use Only. Not for use in diagnostic procedures.

c. Tick the box next to the file that you wish to import. If this file type is not recognised by KlusterCaller, you will be prompted to teach KlusterCaller the format using the plate reader configuration wizard.



For Research Use Only. Not for use in diagnostic procedures.

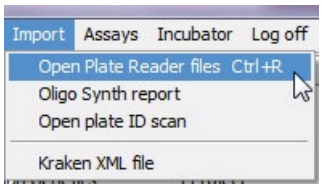
d. Click the ‘Run Wizard’ button to start the wizard.



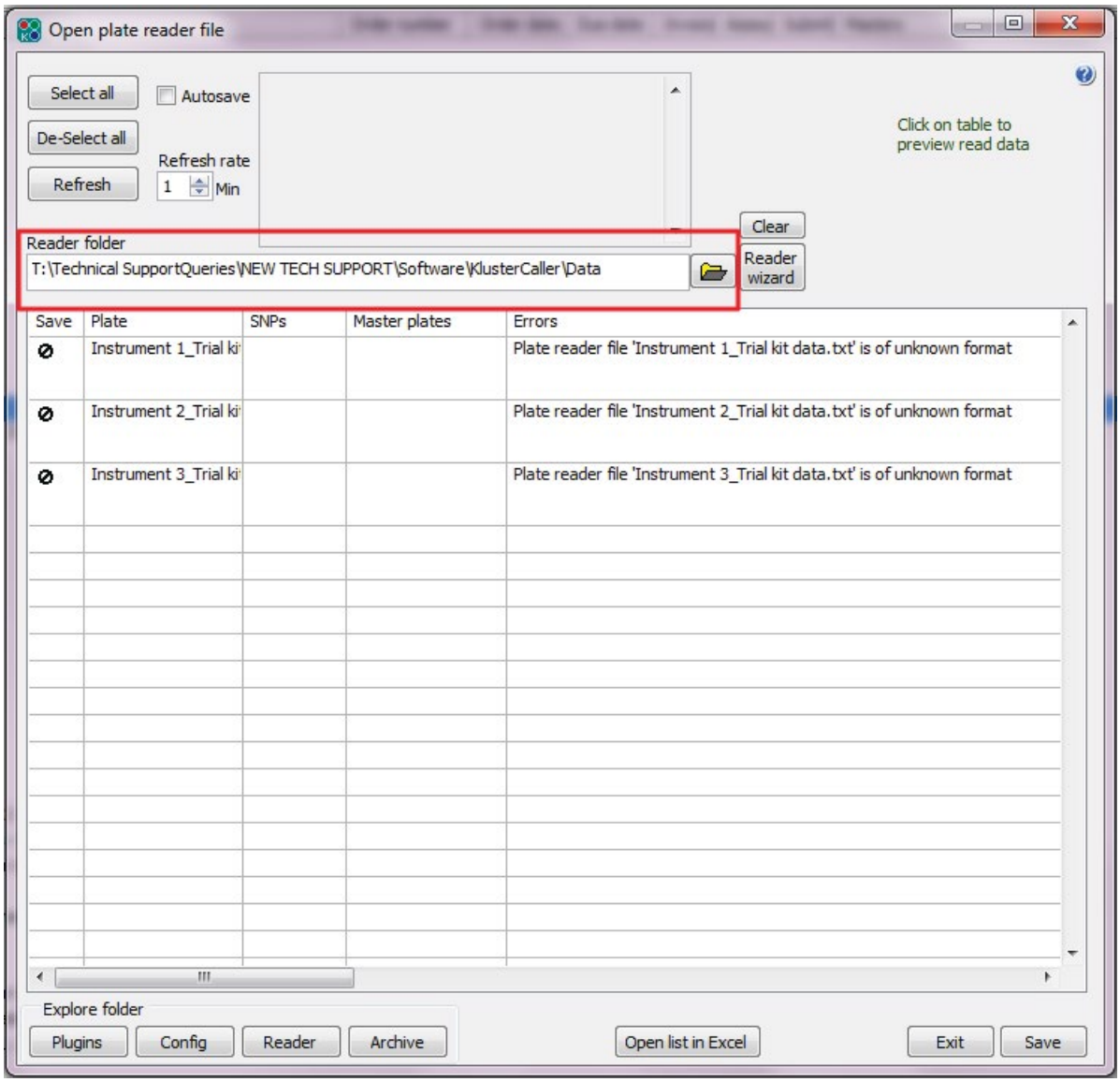
For Research Use Only. Not for use in diagnostic procedures.

3.2 Using Kraken

a. Click on the ‘Import’ menu and select ‘Open Plate Reader Files’.

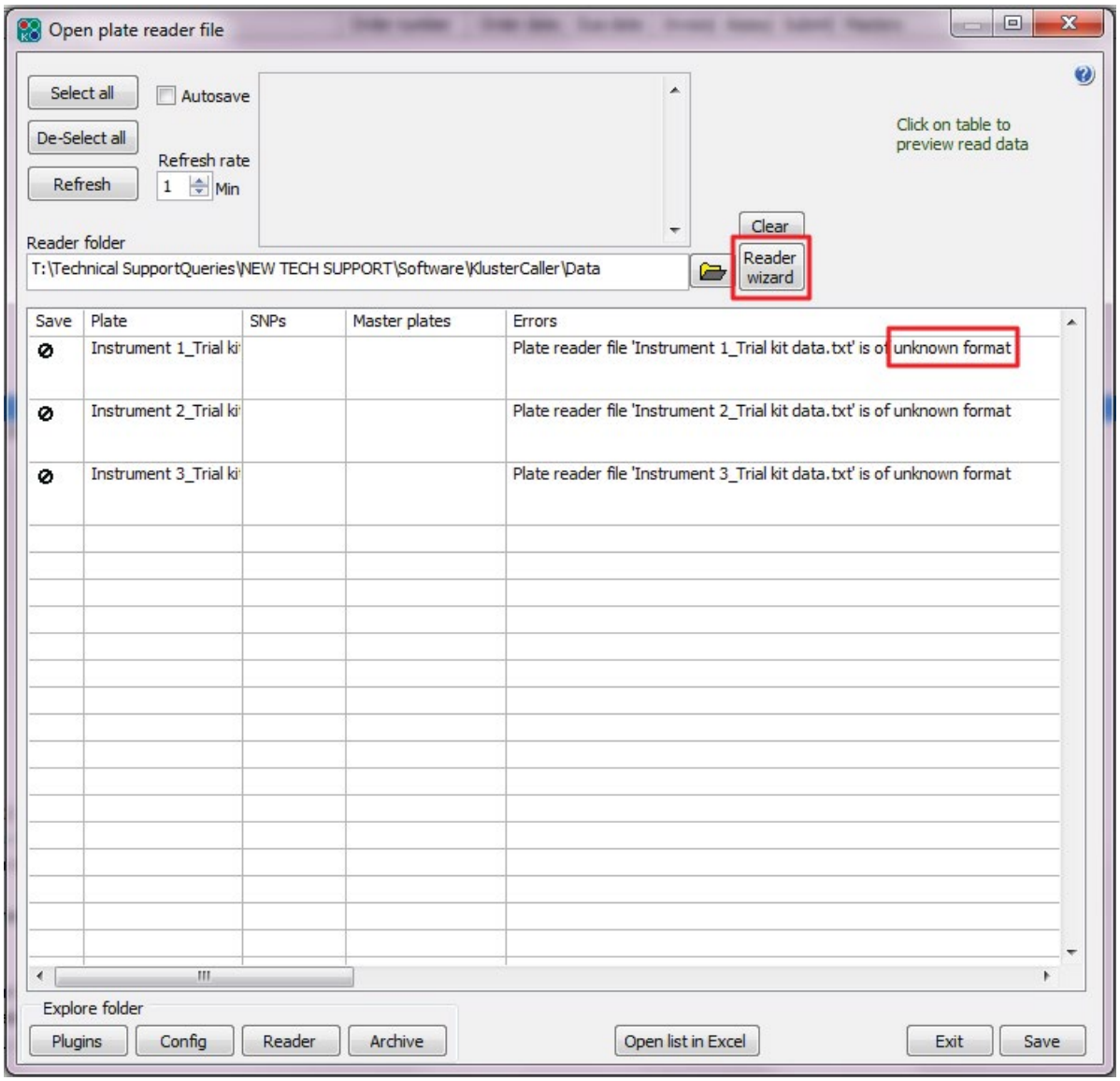


b. In the ‘Open plate reader file’ window, navigate to the folder that contains the raw data file and press ok.



For Research Use Only. Not for use in diagnostic procedures.

c. If the file type is not recognised (listed as ‘unknown format’), click on the ‘Reader Wizard’ button to open the plate reader configuration wizard.

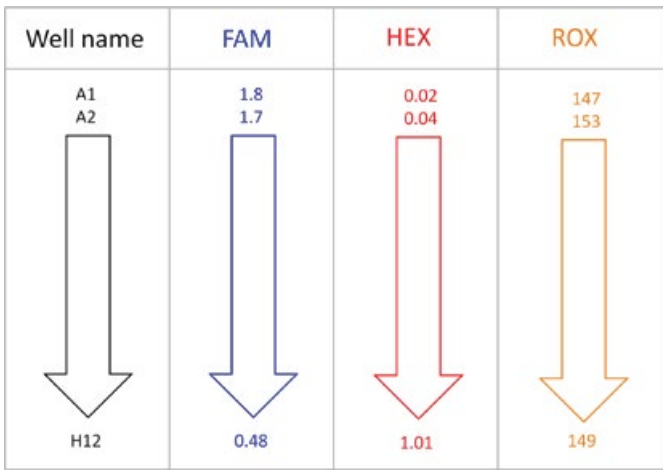


For Research Use Only. Not for use in diagnostic procedures.

4. The plate reader configuration wizard – specific examples

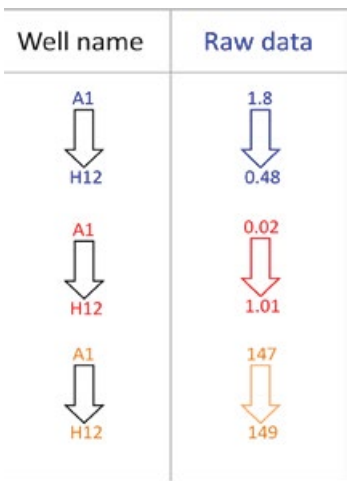
There are three main file formats for raw KASP genotyping data that can be imported into KlusterCaller or Kraken. These formats are:

a) List format #1: this file format typically contains the raw genotyping data in four columns. One column contains the well ID information (i.e. A01, A02), and then the additional three columns contain the raw fluorescence values for FAM, HEX (VIC) and ROX for each well respectively. It is not essential to read ROX, but the ROX values will enable your data to be normalised to correct for pipetting variations.



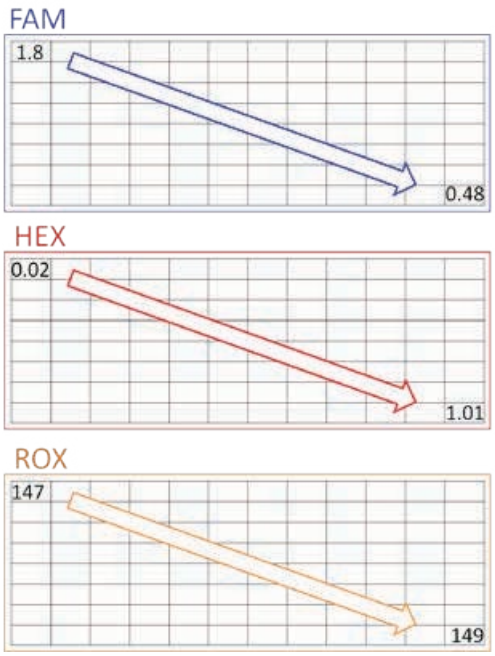
For data files of this type, there may be additional columns within the file containing other run specific information – these can remain in the file, but will not be used for data analysis.

b) List format #2: this file format typically contains the raw genotyping data in two columns. One column contains the well ID information (i.e. A01, A02), and then the second column contains the raw fluorescent values for FAM, then HEX (VIC), and then ROX respectively. The well ID information will be repeated in the first column to correspond with the data for each of the fluors.



For Research Use Only. Not for use in diagnostic procedures.

c) Grid format: this file format typically contains the raw genotyping data in a plate layout or grid format. There will be one grid for each set of fluorescent values, so a dataset containing FAM, HEX and ROX reads will consist of three grids of data.



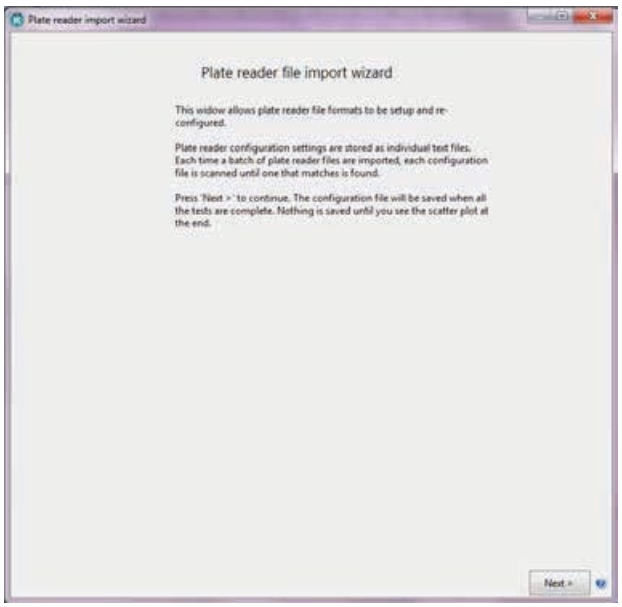
This document will outline how to import these three different file formats in the following sections:

a) List format #1: Section 4.1
b) List format #2: Section 4.2
c) Grid format: Section 4.3

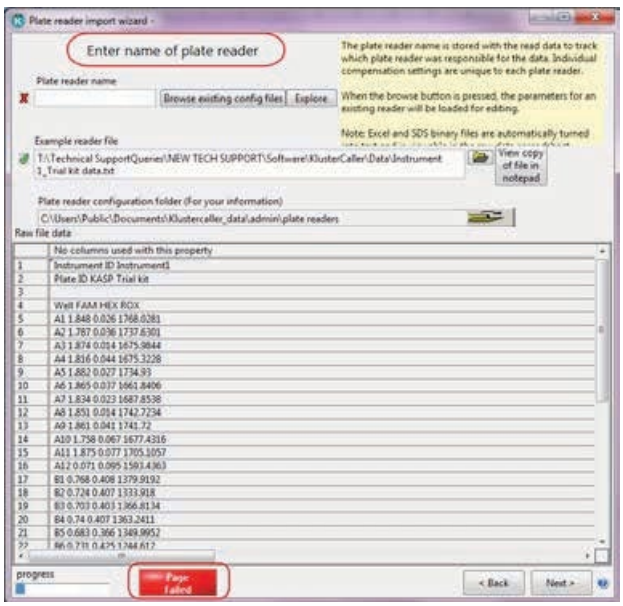
For Research Use Only. Not for use in diagnostic procedures.

4.1 Import of ‘List format #1’ data using the plate configuration wizard

- If your data file is not recognised, you will be prompted to run the plate configuration wizard. After pressing the ‘Run Wizard’ button, the ‘plate reader import wizard’ window will open.

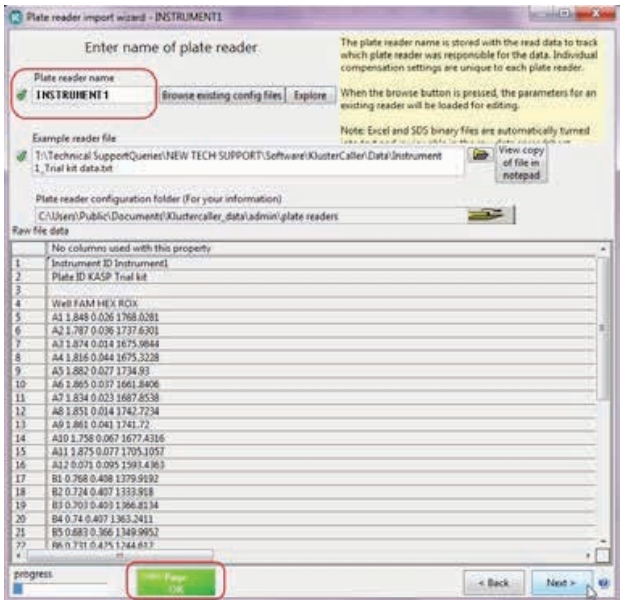


- Press ‘Next’.
- You will then be prompted to enter the name of your plate reader. Until you do this, the page will be marked as failed.

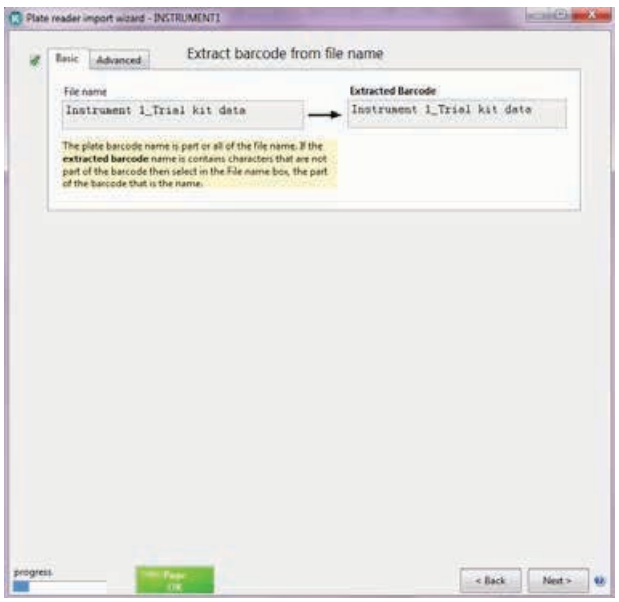


For Research Use Only. Not for use in diagnostic procedures.

- Input an appropriate name (in this case, ‘Instrument1’), and press ‘Next’.

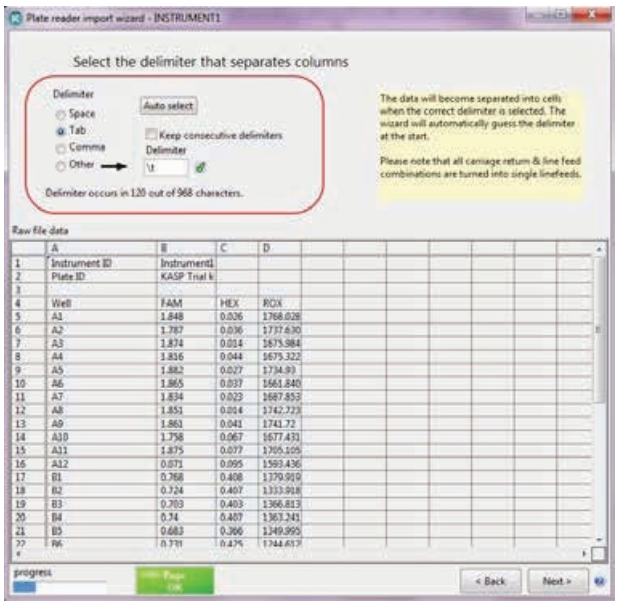


- The wizard will then extract the file name as the identifier for the data. You do not need to edit anything in this window. Press ‘Next’.

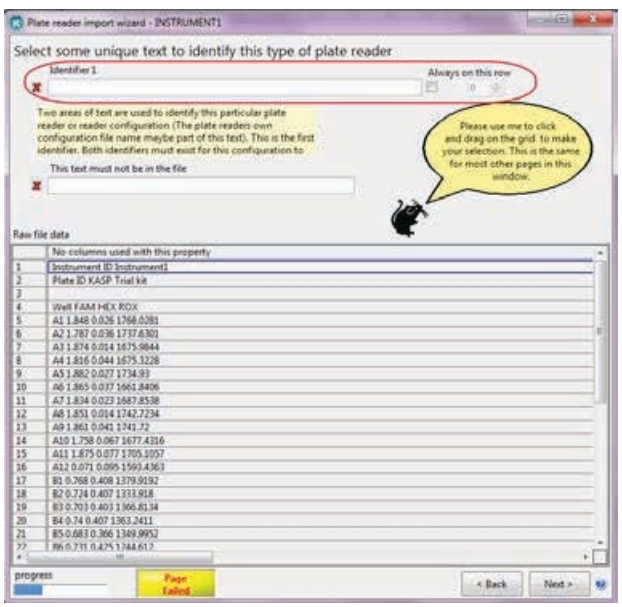


For Research Use Only. Not for use in diagnostic procedures.

- The wizard will auto-select the delimiter that separates the columns in your file. The delimiter can also be manually changed if the wizard has chosen the incorrect delimiter. In this example, it is 'Tab'. By selecting 'Tab' from the list, the preview of the data shows the data split into columns. If you have selected the incorrect delimiter, the data in the preview will not be split into columns. Ensure that you have selected the appropriate delimiter, and then press 'Next'.

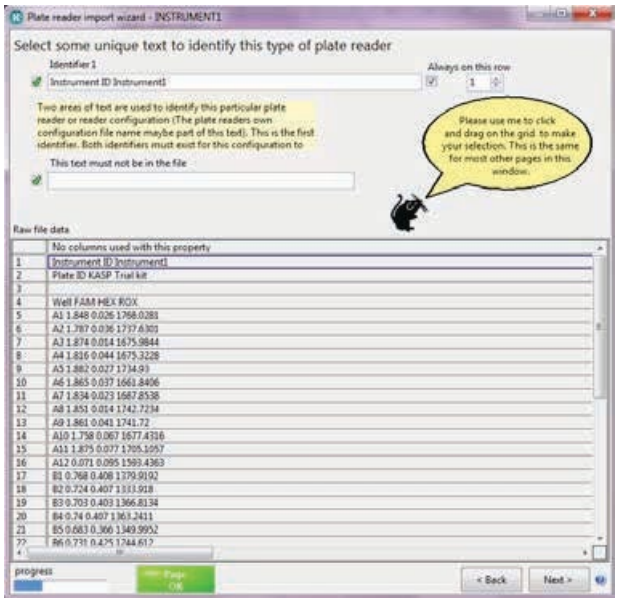


- The wizard then requires some unique text that appears in your data file. This text must also be present in all future files of this type, so text such as the individual plate ID will not be suitable. In this example, the instrument ID appears in row 1, and will appear in this position in all data files generated using this instrument – it is therefore suitable as the 'unique text'.

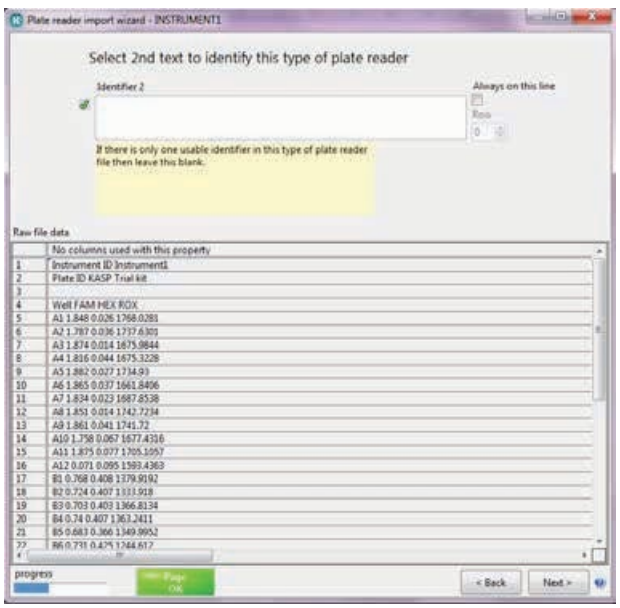


For Research Use Only. Not for use in diagnostic procedures.

- Click on the ‘unique text’ in the file preview. The text will appear in the ‘Identifier 1’ box, and the red cross will become a green tick. Press ‘Next’.

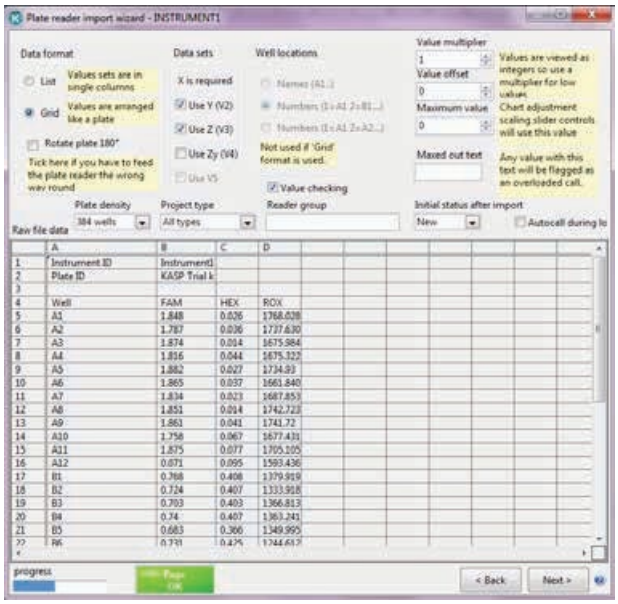


- There is then the option to identify a second set of text as an additional identifier for the file type, but it is not necessary to include this. Either select an appropriate identifier, or leave the box blank, and then press ‘Next’.

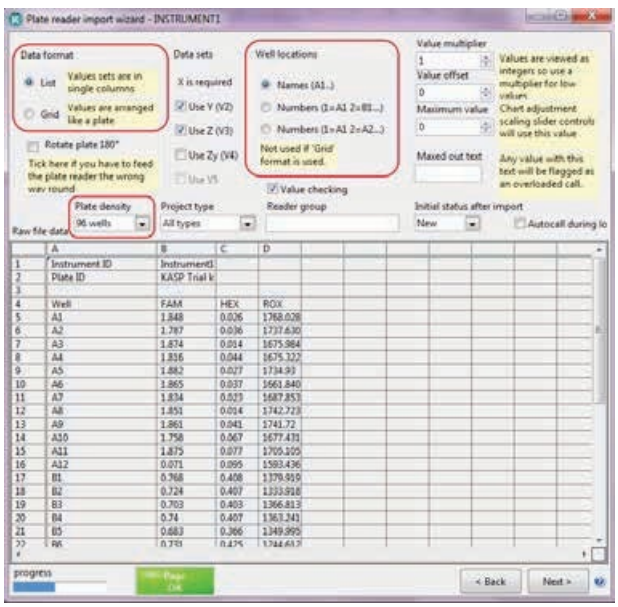


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information about the layout of the data in your file. The window will default to the settings below, but will require adjustment according to the specifics of your file format.

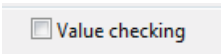


- For this file format, the following sections require editing:
  - Data format: this should be changed from ‘Grid’ to ‘List’ as the raw data are formatted in lists
  - Well locations: if applicable, this should be changed to ‘Names (A1..)’ as each well is identified in the list by its actual name. In some cases, well names may be identified by numbers and you will need to select an alternative option.
  - Plate density: this should be reviewed to ensure that the correct plate type is selected i.e. 96 wells or 384 wells.

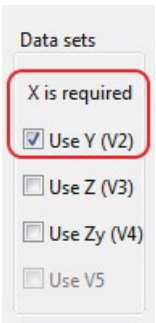


For Research Use Only. Not for use in diagnostic procedures.

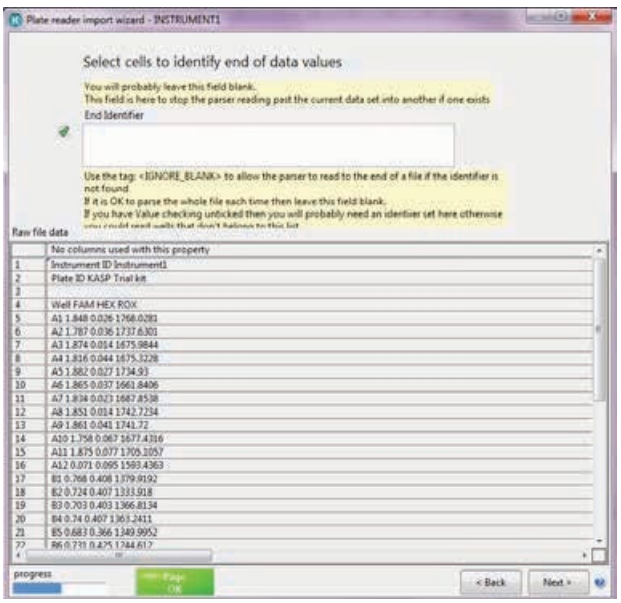
- If your data file does not contain data for a full plate (i.e. all 96 or 384 wells), it is essential to untick the ‘Value checking’ box.



- If your data file only contains FAM and HEX data (i.e. you have not read ROX), you will also need to edit the ‘Data sets’ section. In this case, only X and Y data sets should be selected.

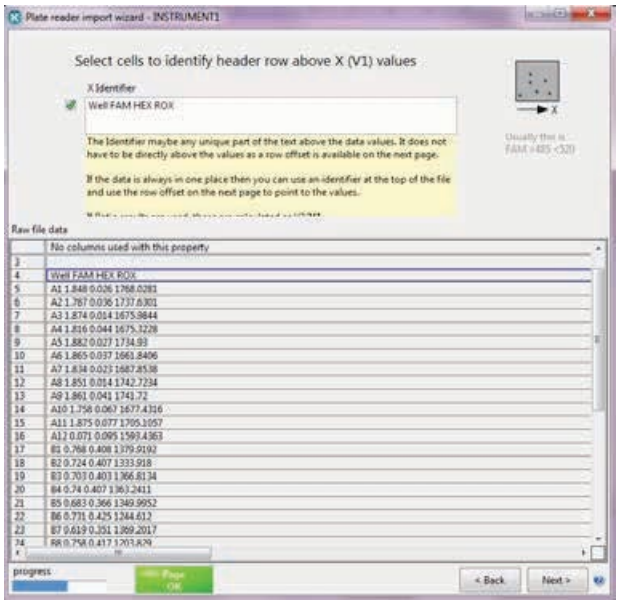


- Once all of the relevant sections have been edited, press ‘Next’.
- The next window asks you to select cells that identify the end of your data values. This is not required, and you can just press ‘Next’.

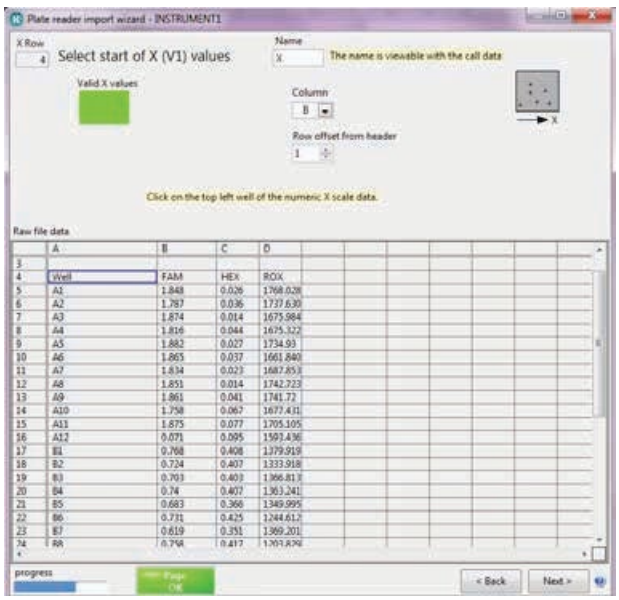


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires a reference to the header row in the data file that appears above the X data values (X = FAM data). Click on the header row (in this example it is row 4), and the contents will appear in the ‘X identifier’ box at the top of the window. Press ‘Next’.

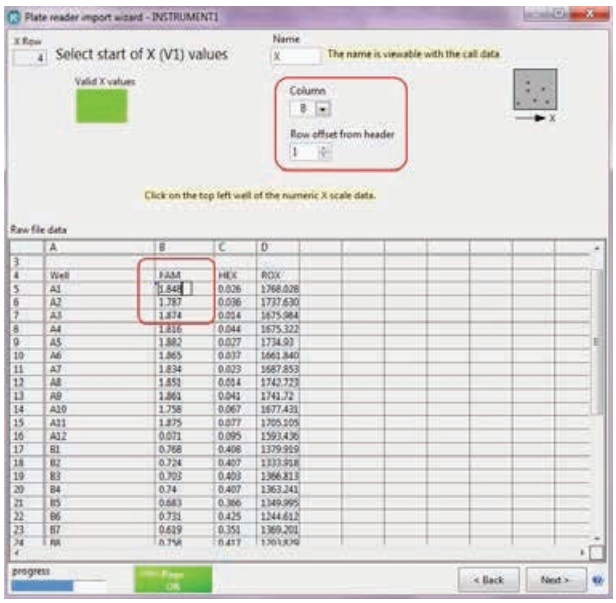


- The wizard then requires information regarding the position of the start of the X data values in the raw data file.

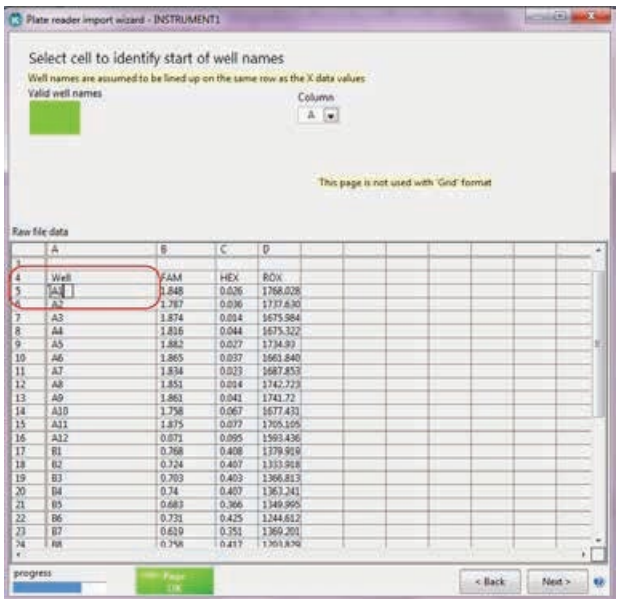


For Research Use Only. Not for use in diagnostic procedures.

- In this example, the X (FAM) data appears in column B so click on the relevant cell to identify the first X data value (in this case cell B5). Press ‘Next’.

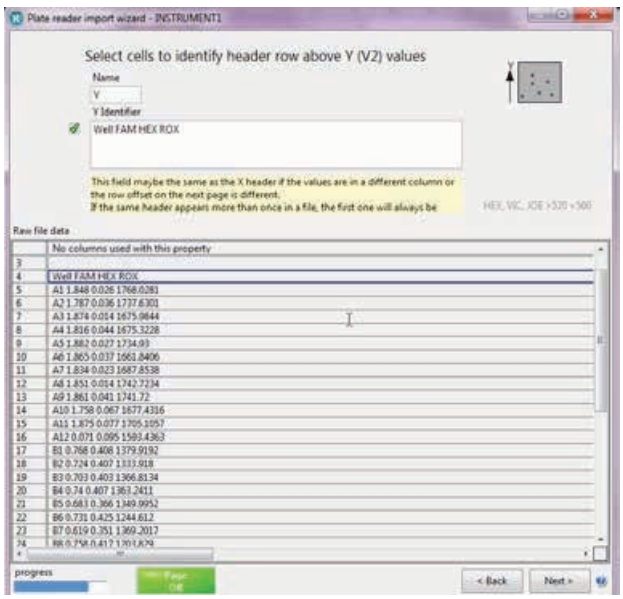


- The wizard then requires information regarding the position of the list of well names. Click on the first well name in the raw data file preview (in this case, it is cell A5). Press ‘Next’.

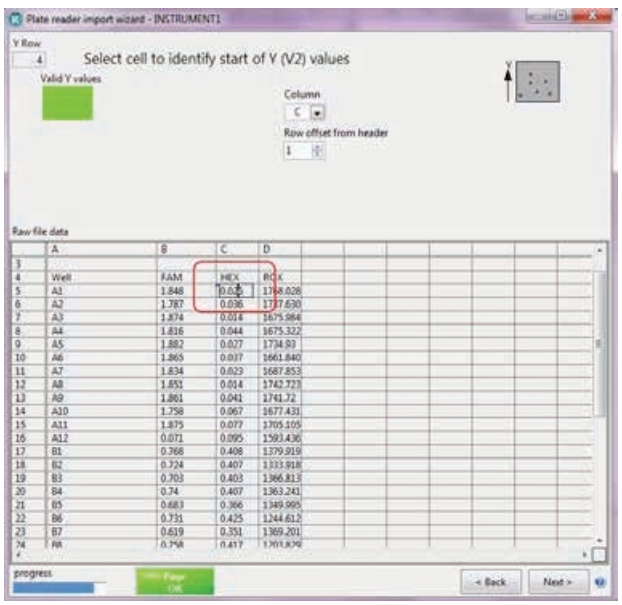


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires a reference to the header row in the data file that appears above the Y data values (Y = HEX data). Click on the header row (in this example it is row 4), and the contents will appear in the ‘Y identifier’ box at the top of the window. Press ‘Next’.

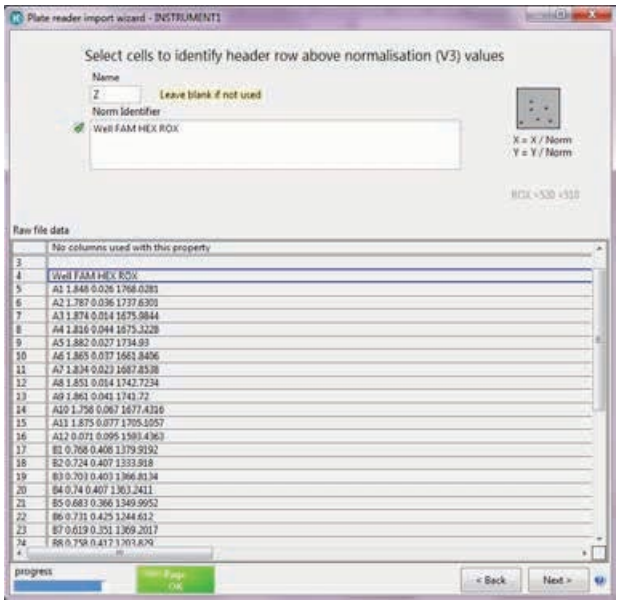


- The wizard then requires information regarding the position of the start of the Y data values in the raw data file. In this example, the Y (HEX) data appears in column C so click on the relevant cell to identify the first Y data value (in this case cell C5). Press ‘Next’.

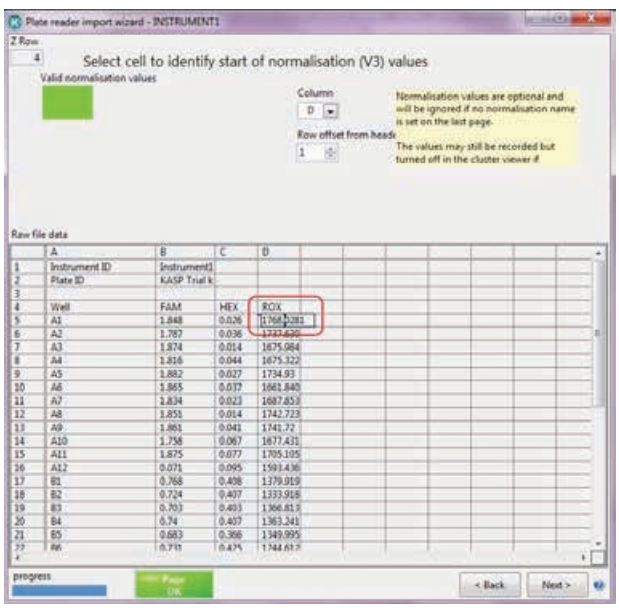


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires a reference to the header row in the data file that appears above the Z data values (Z = ROX data). Click on the header row (in this example it is row 4), and the contents will appear in the ‘Norm identifier’ box at the top of the window. Press ‘Next’.

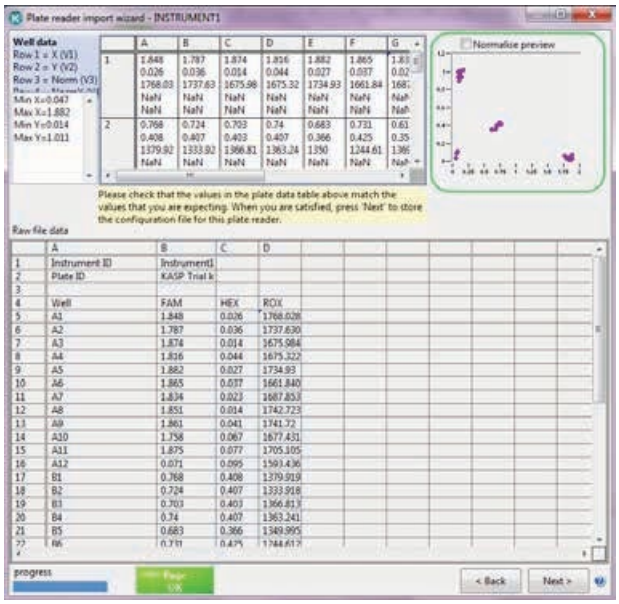


- The wizard then requires information regarding the position of the start of the Z data values in the raw data file. In this example, the Z (ROX) data appears in column D so click on the relevant cell to identify the first Z data value (in this case cell D5). Press ‘Next’.

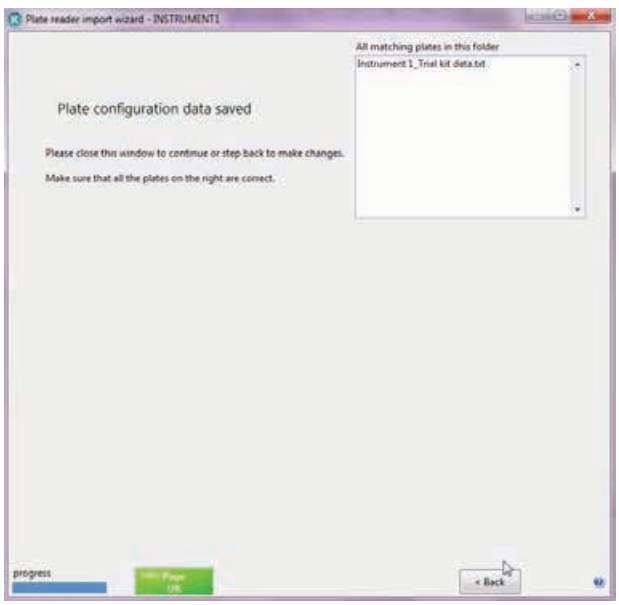


For Research Use Only. Not for use in diagnostic procedures.

- The wizard will now show a preview of your data, both as numerical values in a plate layout format and as a cluster plot. If this data looks as expected (i.e. A1 data is shown in the A1 grid layout at the top of the window), click ‘Next’.

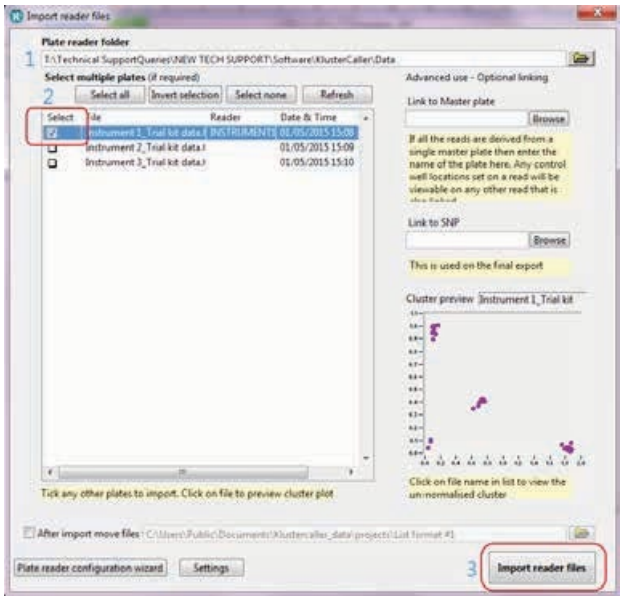


- The wizard will then open a confirmation window, stating that the configuration file for this data format has been saved. All raw data files that are recognised by this configuration file will be listed in the top right hand corner. Click on the red cross to close the wizard.



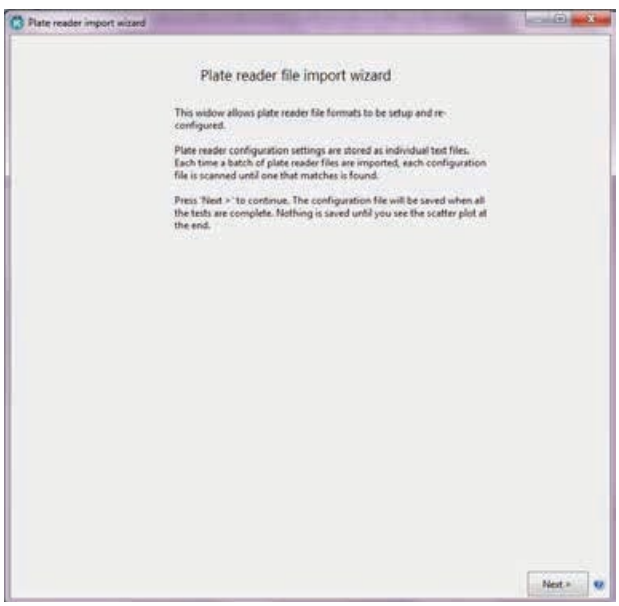
For Research Use Only. Not for use in diagnostic procedures.

- KlusterCaller will then return to the ‘Import reader files’ window. Place a tick in the box next to the relevant raw data file – this should now be recognised by the software. A preview cluster plot will also be visible. Click on the ‘Import reader files’ button to import this data file to your KlusterCaller project.



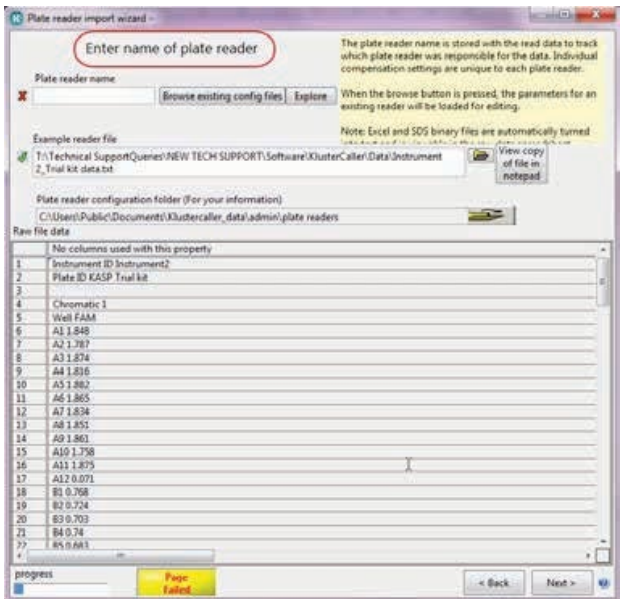
4.2 Import of ‘List format #2’ data using the plate configuration wizard

- If your data file is not recognised, you will be prompted to run the plate configuration wizard. After pressing the ‘Run Wizard’ button, the ‘plate reader import wizard’ window will open.

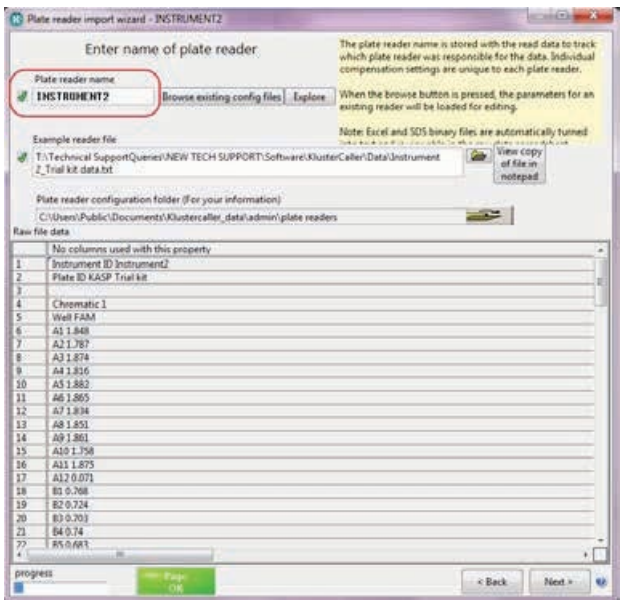


For Research Use Only. Not for use in diagnostic procedures.

- Press ‘Next’.
- You will then be prompted to enter the name of your plate reader. Until you do this, the page will be marked as failed.

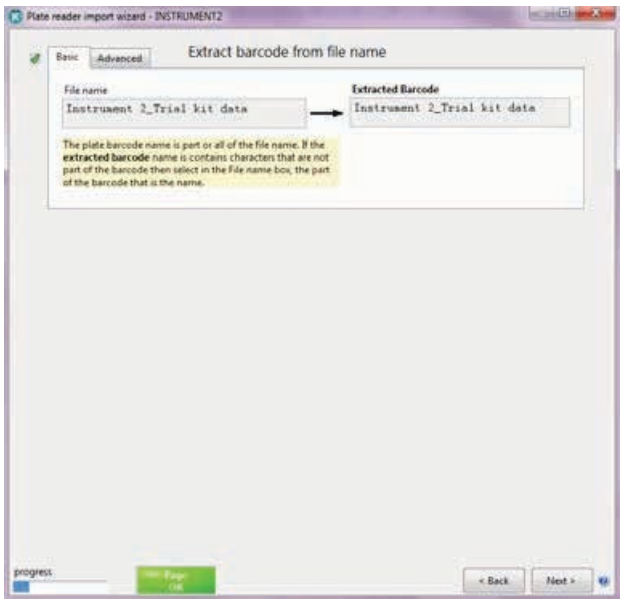


- Input an appropriate name (in this case, ‘Instrument2’), and press ‘Next’.

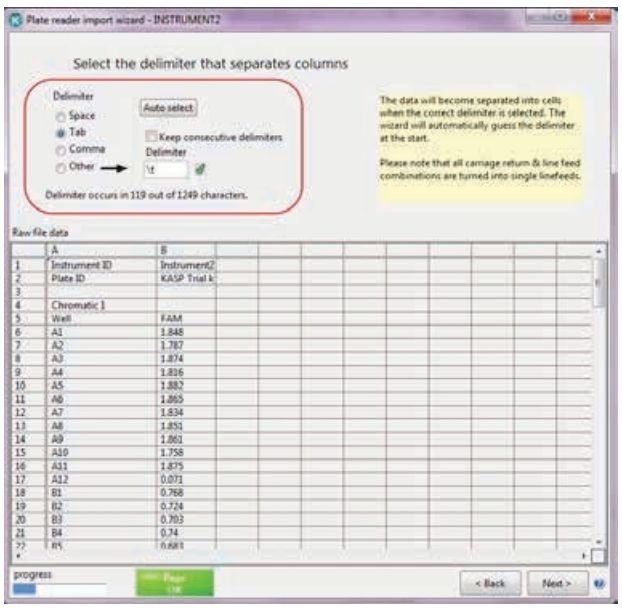


For Research Use Only. Not for use in diagnostic procedures.

- The wizard will then extract the file name as the identifier for the data. You do not need to edit anything in this window. Press ‘Next’.

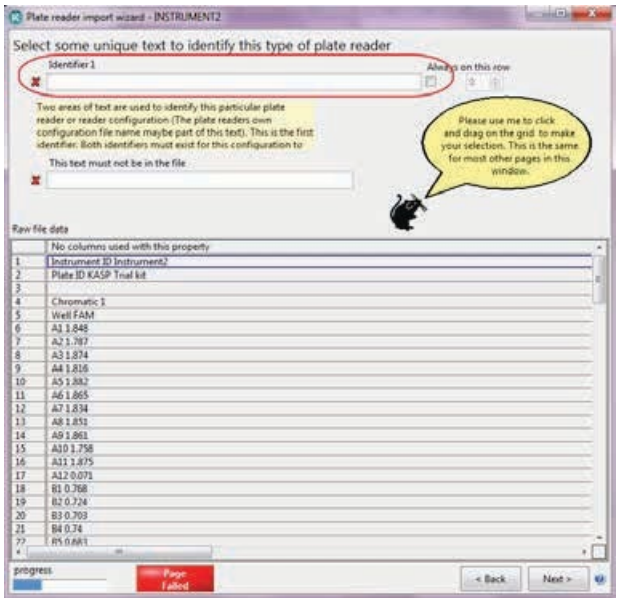


- The wizard will auto-select the delimiter that separates the columns in your file. The delimiter can also be manually changed if the wizard has chosen the incorrect delimiter. In this example, it is ‘Tab’. By selecting ‘Tab’ from the list, the preview of the data shows the data split into columns. If you have selected the incorrect delimiter, the data in the preview will not be split into columns. Ensure that you have selected the appropriate delimiter, and then press ‘Next’.

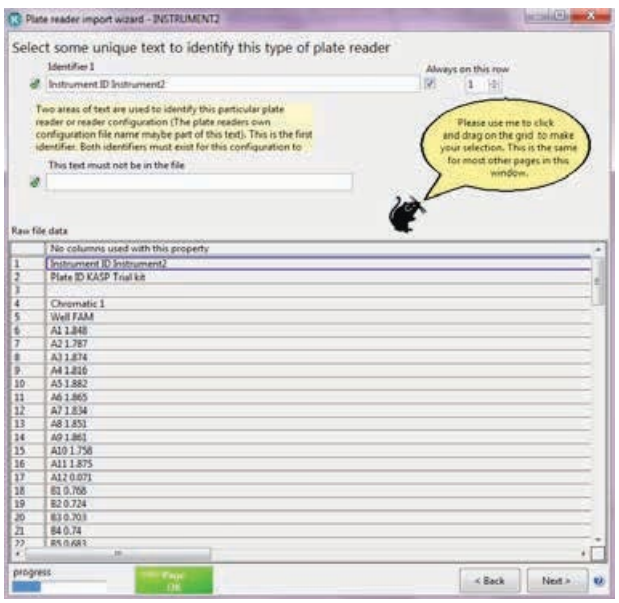


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires some unique text that appears in your data file. This text must also be present in all future files of this type, so text such as the individual plate ID will not be suitable. In this example, the instrument ID appears in row 1, and will appear in this position in all data files generated using this instrument – it is therefore suitable as the ‘unique text’.

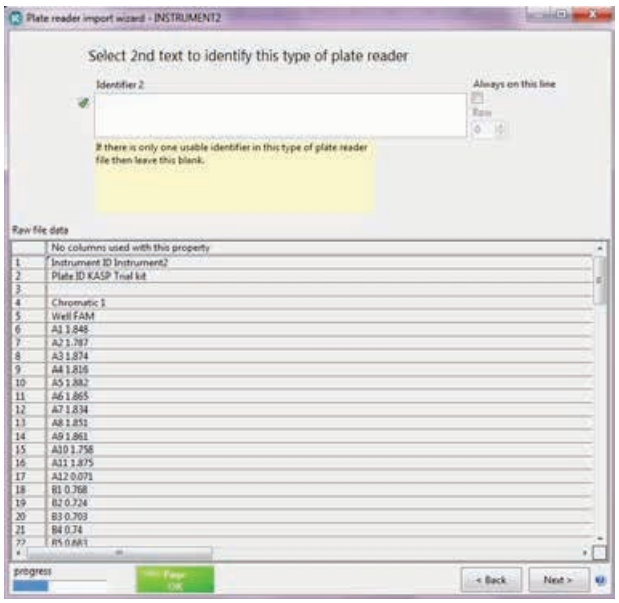


- Click on the ‘unique text’ in the file preview. The text will appear in the ‘Identifier 1’ box, and the red cross will become a green tick. Press ‘Next’.

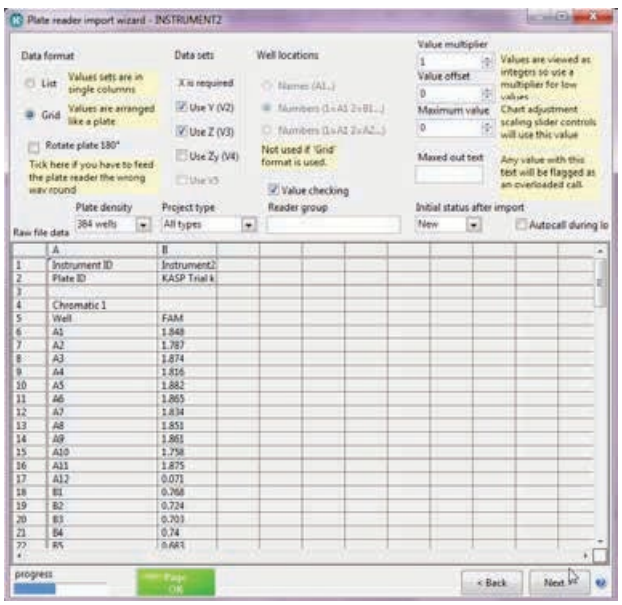


For Research Use Only. Not for use in diagnostic procedures.

- There is then the option to identify a second set of text as an additional identifier for the file type, but it is not necessary to include this. Either select an appropriate identifier, or leave the box blank, and then press ‘Next’.

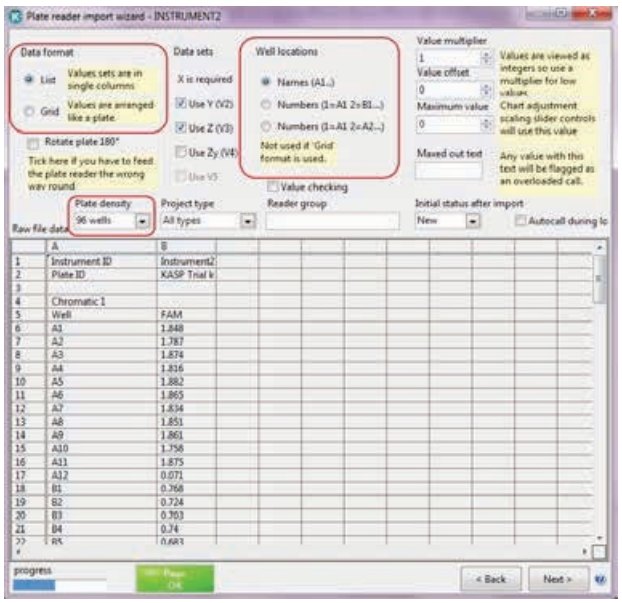


- The wizard then requires information about the layout of the data in your file. The window will default to the settings below, but will require adjustment according to the specifics of your file format.

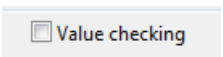


For Research Use Only. Not for use in diagnostic procedures.

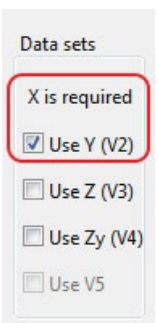
- For this file format, the following sections require editing:
  - Data format: this should be changed from ‘Grid’ to ‘List’ as the raw data is formatted in a list.
  - Well locations: if applicable, this should be changed to ‘Names (A1..)’ as each well is identified in the list by its actual name. In some cases, well names may be identified by numbers and you will need to select an alternative option.
  - Plate density: this should be reviewed to ensure that the correct plate type is selected i.e. 96 wells or 384 wells.



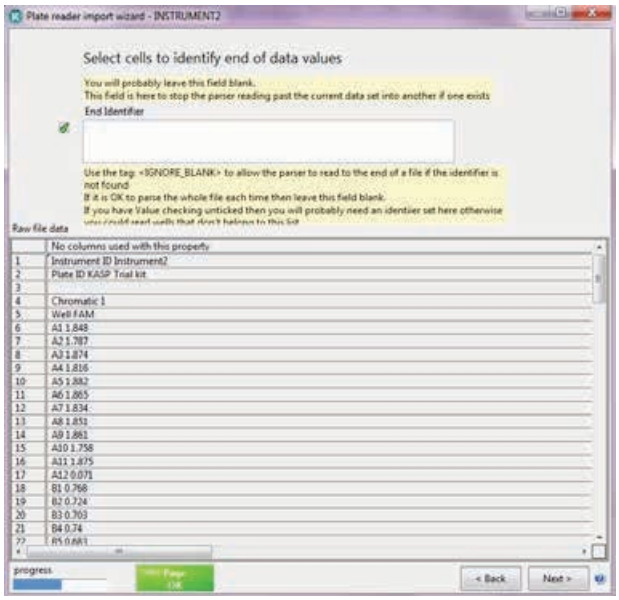
- If your data file does not contain data for a full plate (i.e. all 96 or 384 wells), it is essential to untick the ‘Value checking’ box.



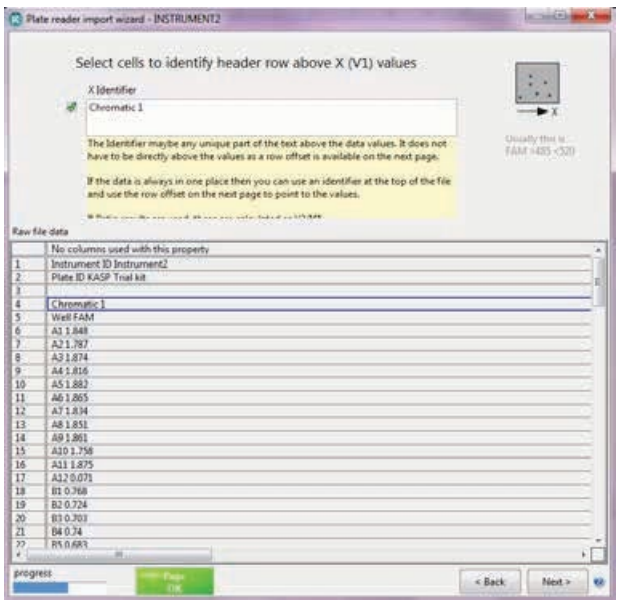
- If your data file only contains FAM and HEX data (i.e. you have not read ROX), you will also need to edit the ‘Data sets’ section. In this case, only X and Y data sets should be selected.



- Once all of the relevant sections have been edited, press ‘Next’.
- The next window asks you to select cells that identify the end of your data values. This is not required, and you can just press ‘Next’.

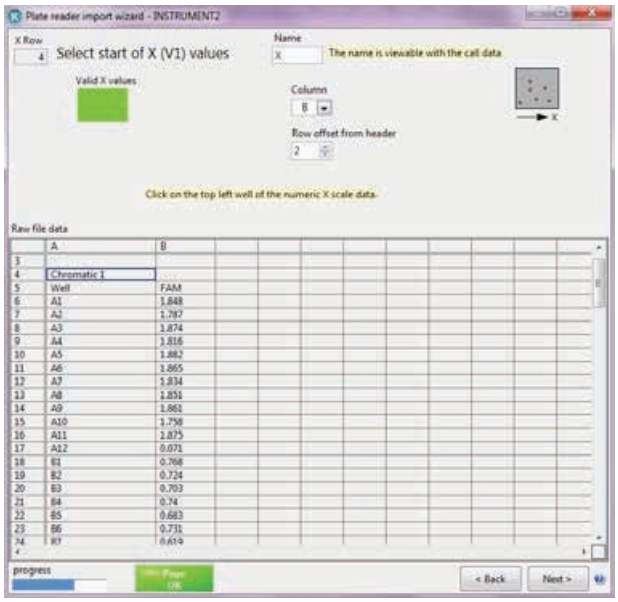


- The wizard then requires a reference to the header row in the data file that appears above the X data values (X = FAM data). Click on the header row (in this example, row 4 can be used), and the contents will appear in the ‘X identifier’ box at the top of the window. Press ‘Next’.

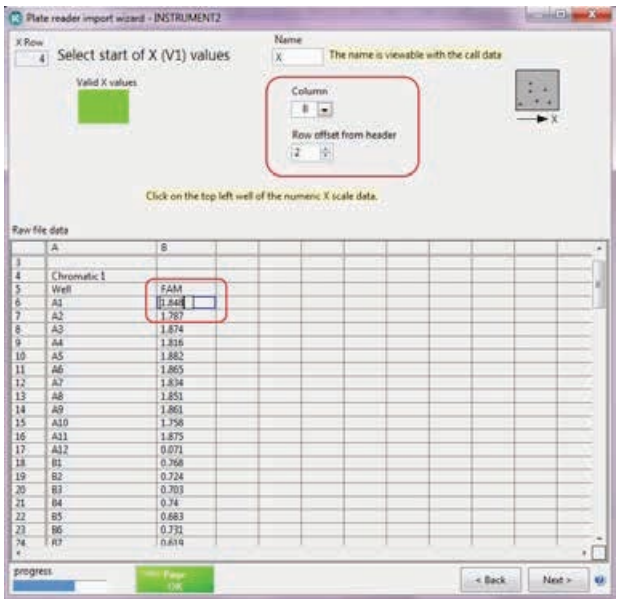


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information regarding the position of the start of the X data values in the raw data file.

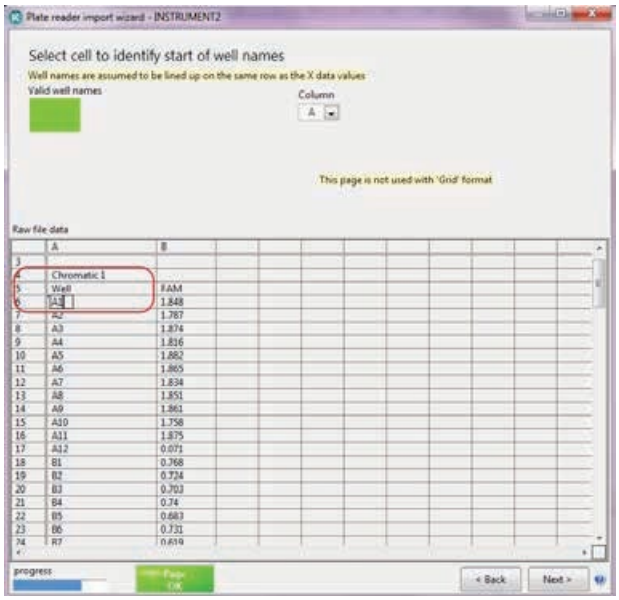


- In this example, the X (FAM) data appears in column B so click on the relevant cell to identify the first X data value (in this case cell B6). Press ‘Next’.

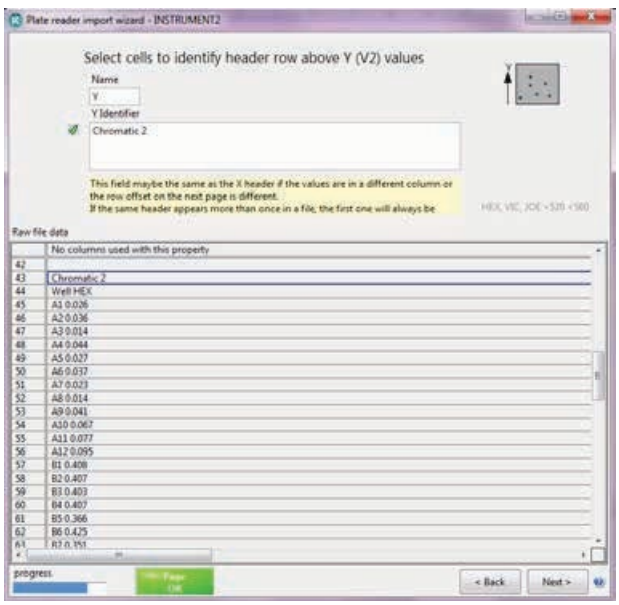


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information regarding the position of the list of well names. Click on the first well name in the raw data file preview (in this case, it is cell A6). Press ‘Next’.

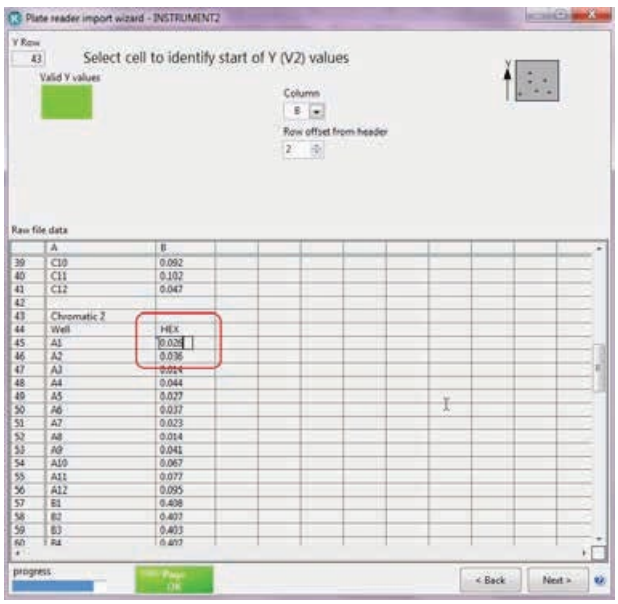


- The wizard then requires a reference to the header row in the data file that appears above the Y data values (Y = HEX data). Scroll down the raw data file and click on the header row (in this example it is row 43), and the contents will appear in the ‘Y identifier’ box at the top of the window. Press ‘Next’.

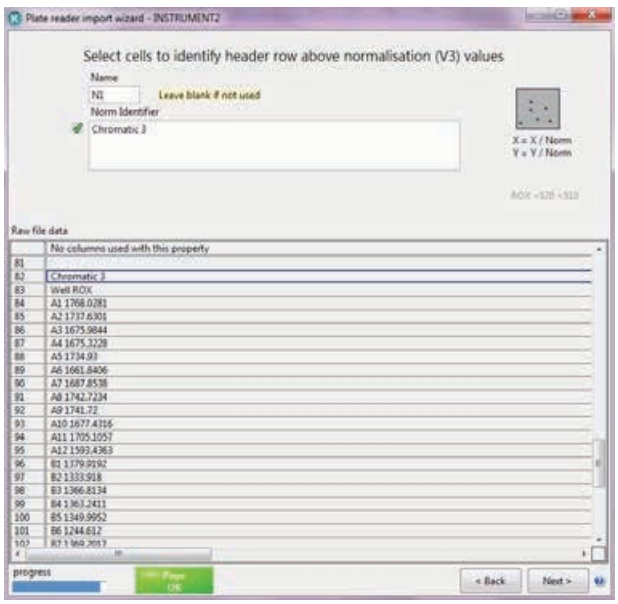


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information regarding the position of the start of the Y data values in the raw data file. In this example, the Y (HEX) data appears in column B, but starts on row 45, so click on the relevant cell to identify the first Y data value (in this case cell B45). Press ‘Next’.

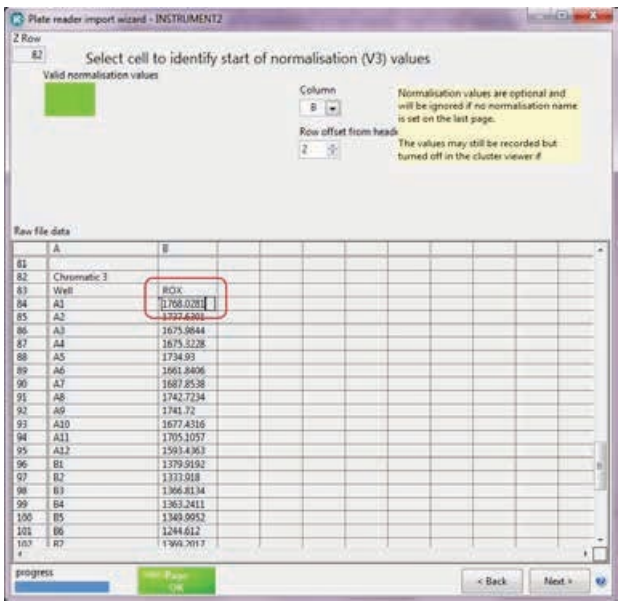


- The wizard then requires a reference to the header row in the data file that appears above the Z data values (Z = ROX data). Scroll down the raw data file and click on the header row (in this example it is row 82), and the contents will appear in the ‘Norm identifier’ box at the top of the window. Press ‘Next’.

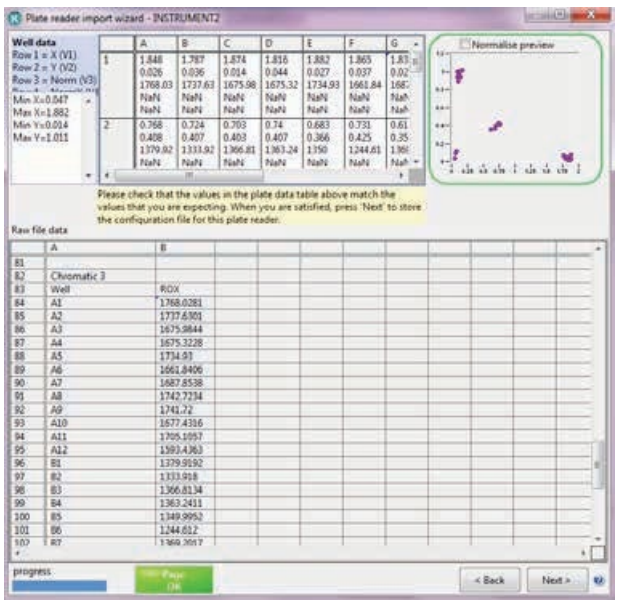


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information regarding the position of the start of the Z data values in the raw data file. In this example, the Z (ROX) data appears in column B, but starts on row 84, so click on the relevant cell to identify the first Z data value (in this case cell B84). Press ‘Next’.

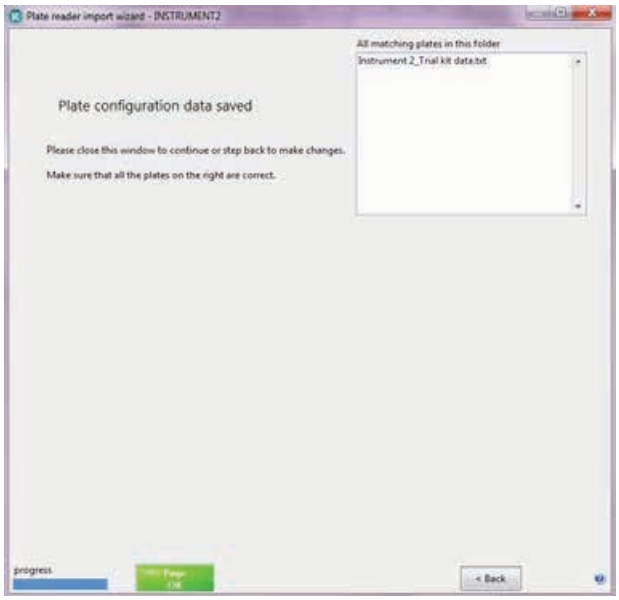


- The wizard will now show a preview of your data, both as numerical values in a plate layout format and as a cluster plot. If this data looks as expected (i.e. A1 data is shown in the A1 grid layout at the top of the window), click ‘Next’.

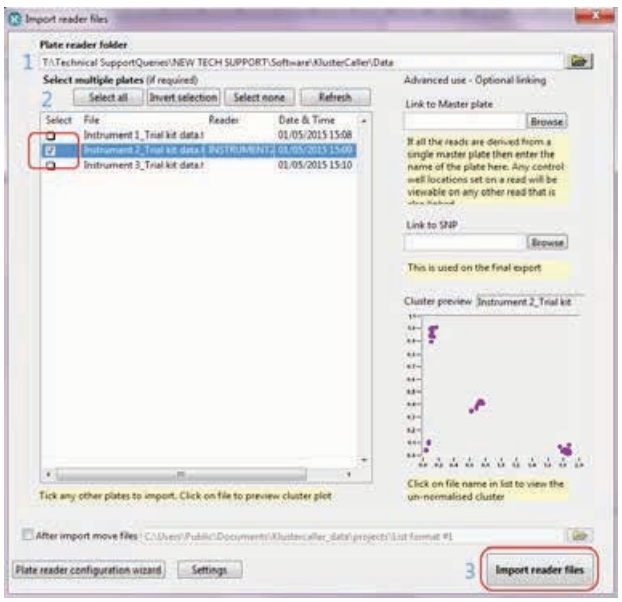


For Research Use Only. Not for use in diagnostic procedures.

- The wizard will then open a confirmation window, stating that the configuration file for this data format has been saved. All raw data files that are recognised by this configuration file will be listed in the top right hand corner. Click on the red cross to close the wizard.



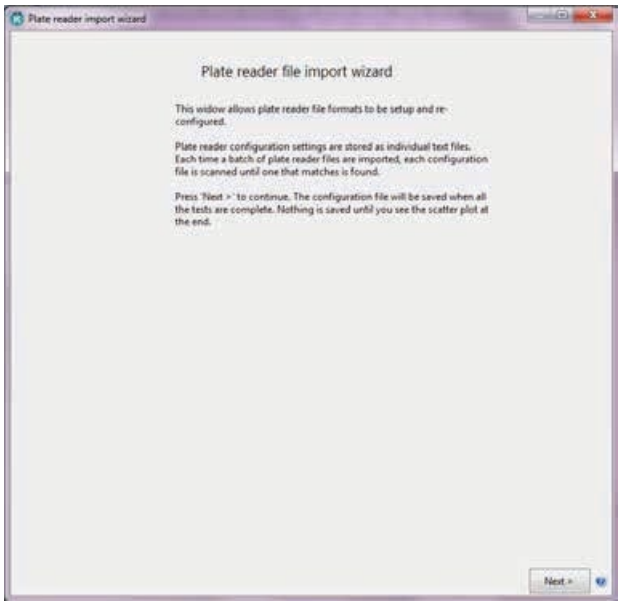
- KlusterCaller will then return to the ‘Import reader files’ window. Place a tick in the box next to the relevant raw data file – this should now be recognised by the software. A preview cluster plot will also be visible. Click on the ‘Import reader files’ button to import this data file to your KlusterCaller project.



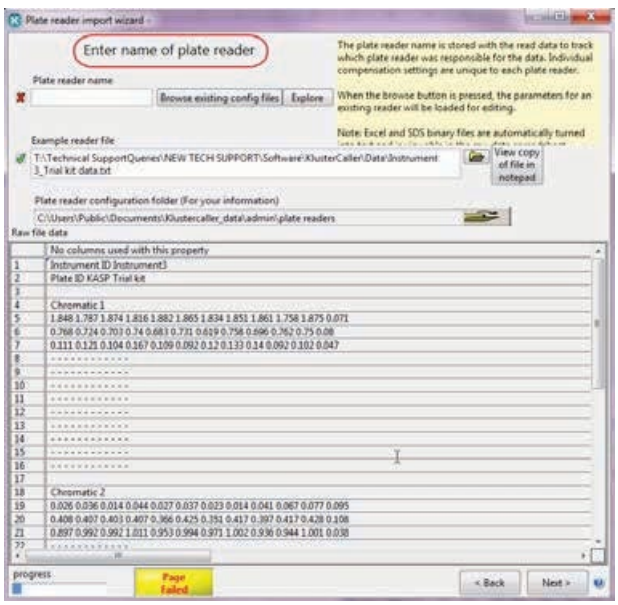
For Research Use Only. Not for use in diagnostic procedures.

4.3 Import of ‘Grid format’ data using the plate configuration wizard

- If your data file is not recognised, you will be prompted to run the plate configuration wizard. After pressing the ‘Run Wizard’ button, the ‘plate reader import wizard’ window will open.

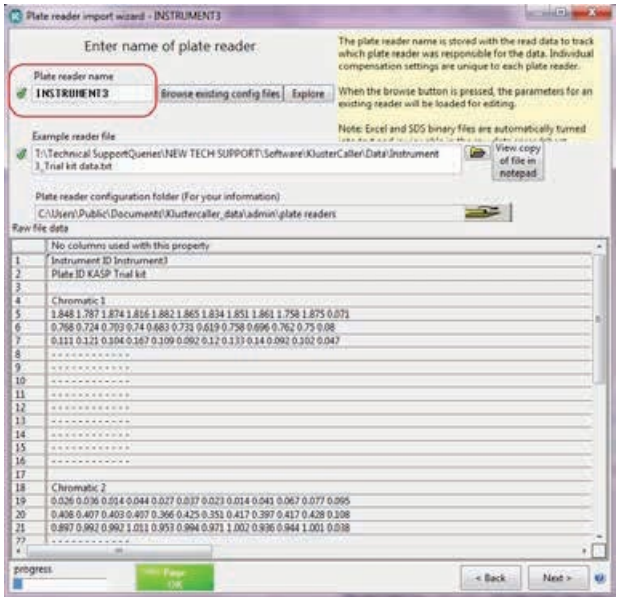


- Press ‘Next’.
- You will then be prompted to enter the name of your plate reader. Until you do this, the page will be marked as failed.

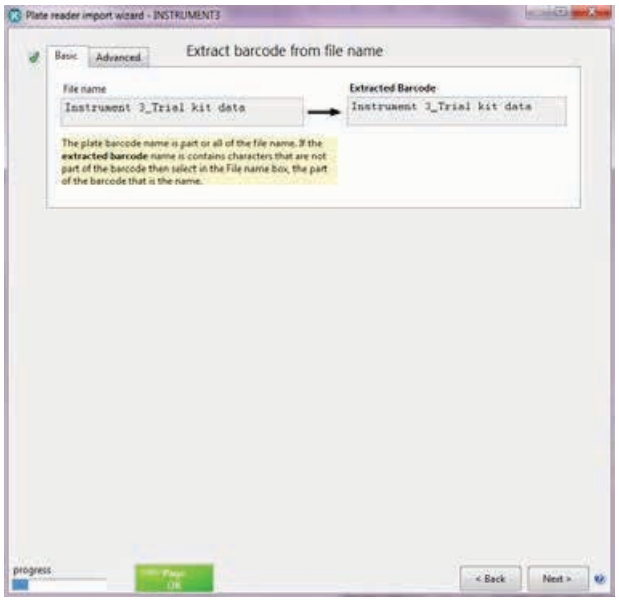


For Research Use Only. Not for use in diagnostic procedures.

- Input an appropriate name (in this case, ‘Instrument3’), and press ‘Next’.

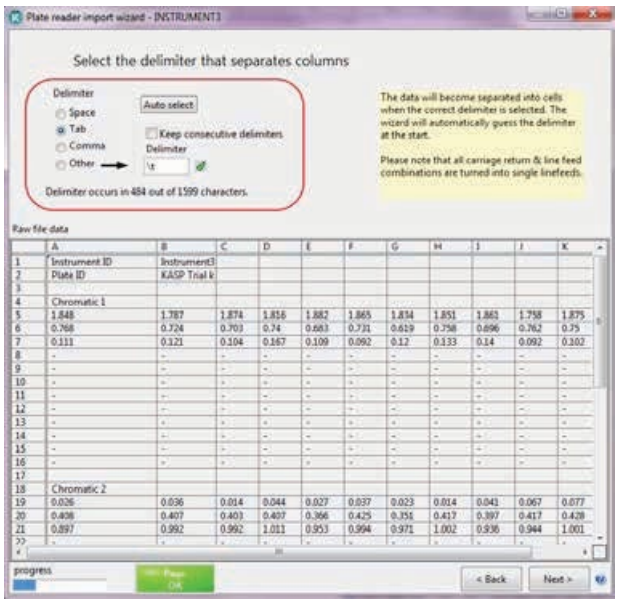


- The wizard will then extract the file name as the identifier for the data. You do not need to edit anything in this window. Press ‘Next’.

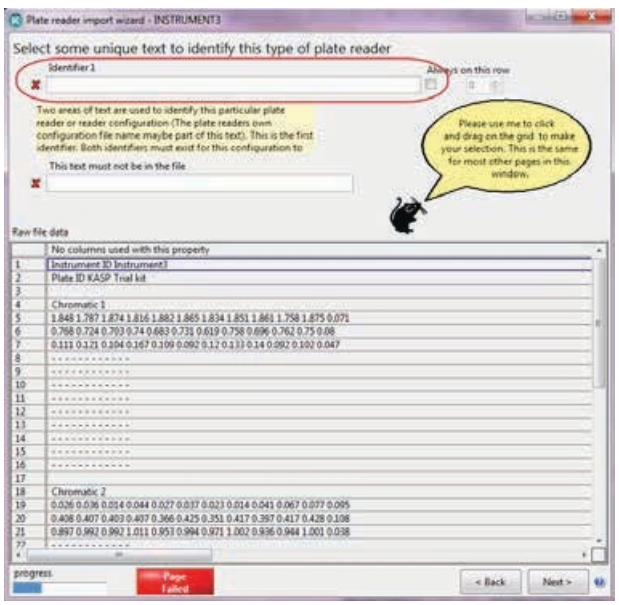


For Research Use Only. Not for use in diagnostic procedures.

- The wizard will auto-select the delimiter that separates the columns in your file. The delimiter can also be manually changed if the wizard has chosen the incorrect delimiter. In this example, it is 'Tab'. By selecting 'Tab' from the list, the preview of the data shows the data split into columns. If you have selected the incorrect delimiter, the data in the preview will not be split into columns. Ensure that you have selected the appropriate delimiter, and then press 'Next'.

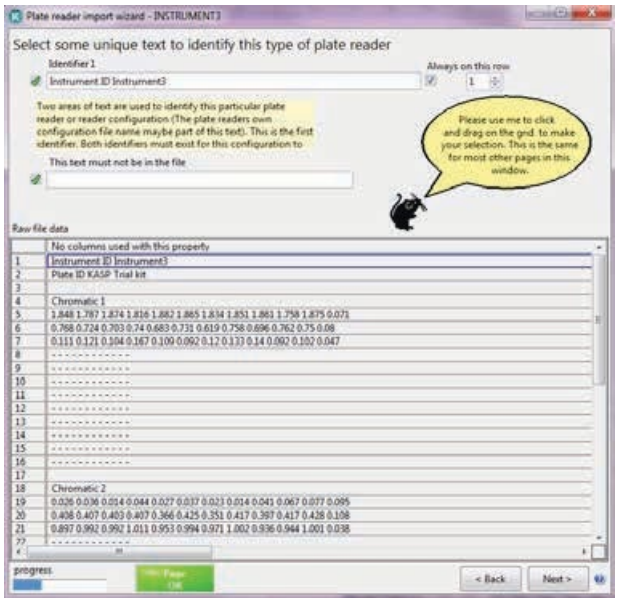


- The wizard then requires some unique text that appears in your data file. This text must also be present in all future files of this type, so text such as the individual plate ID will not be suitable. In this example, the instrument ID appears in row 1, and will appear in this position in all data files generated using this instrument – it is therefore suitable as the 'unique text'.

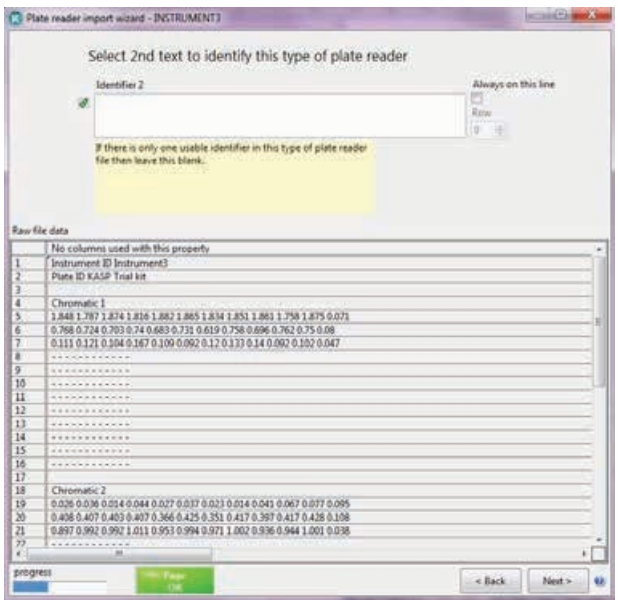


For Research Use Only. Not for use in diagnostic procedures.

- Click on the ‘unique text’ in the file preview. The text will appear in the ‘Identifier 1’ box, and the red cross will become a green tick. Press ‘Next’.

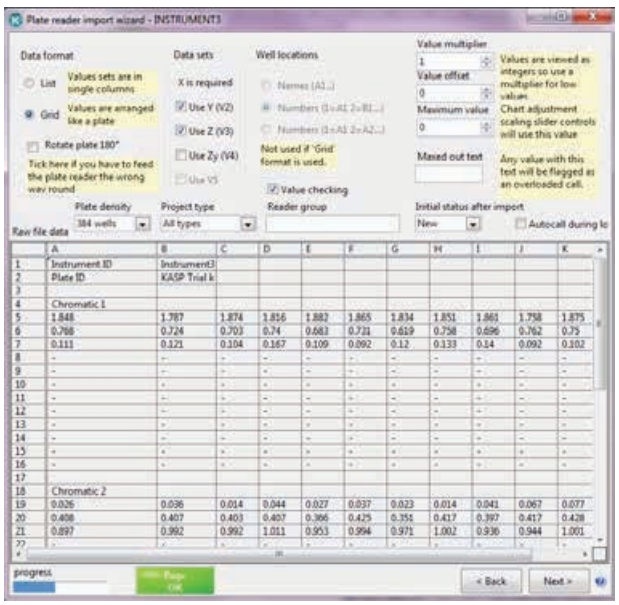


- There is then the option to identify a second set of text as an additional identifier for the file type, but it is not necessary to include this. Either select an appropriate identifier, or leave the box blank, and then press ‘Next’.

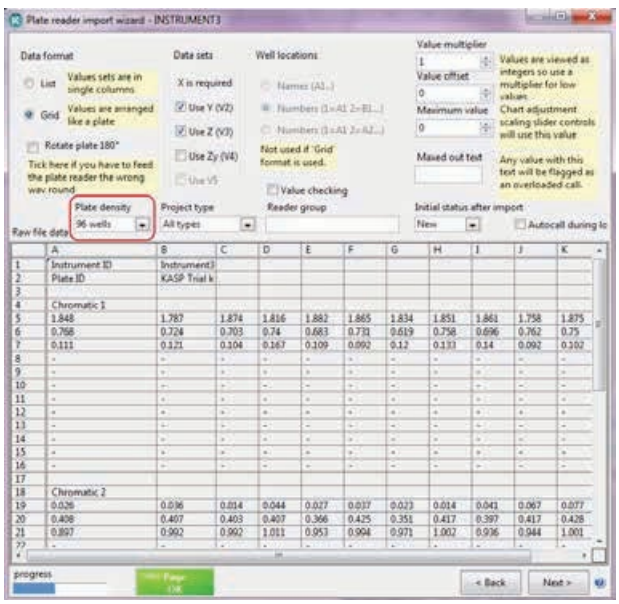


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information about the layout of the data in your file. The window will default to the settings below, but will require adjustment according to the specifics of your file format.

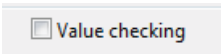


- For this file format, the following section requires editing:
  - Plate density: this should be reviewed to ensure that the correct plate type is selected i.e. 96 wells or 384 wells.

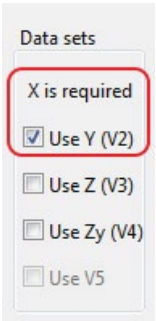


For Research Use Only. Not for use in diagnostic procedures.

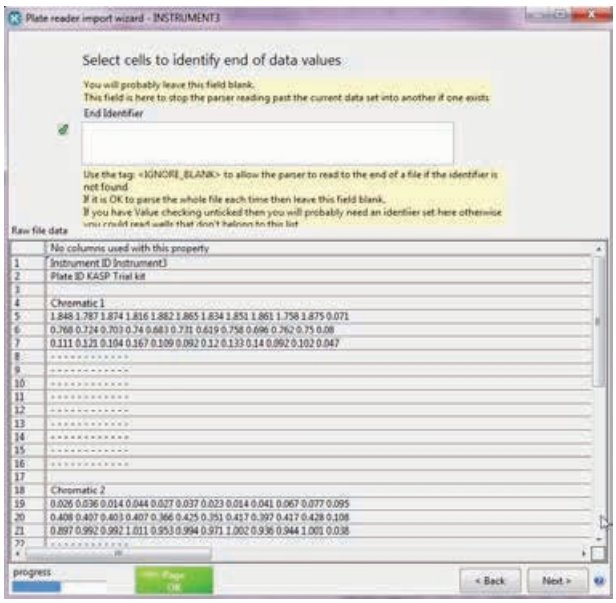
- If your data file does not contain data for a full plate (i.e. all 96 or 384 wells), it is essential to untick the ‘Value checking’ box.



- If your data file only contains FAM and HEX data (i.e. you have not read ROX), you will also need to edit the ‘Data sets’ section. In this case, only X and Y data sets should be selected.

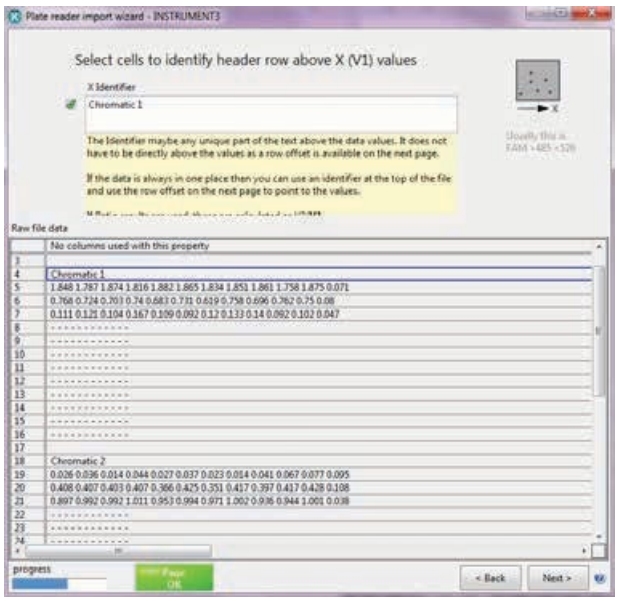


- Once all of the relevant sections have been edited, press ‘Next’.
- The next window asks you to select cells that identify the end of your data values. This is not required, and you can just press ‘Next’.

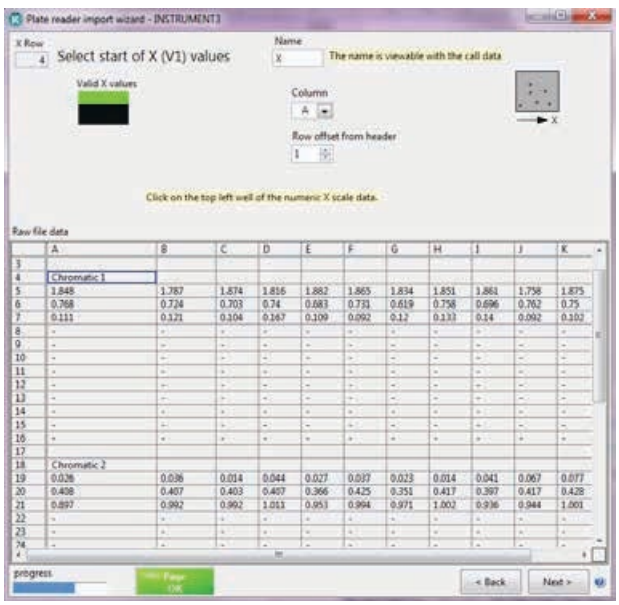


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires a reference to the header row in the data file that appears above the X data values (X = FAM data). Click on the header row (in this example, row 4 can be used), and the contents will appear in the ‘X identifier’ box at the top of the window. Press ‘Next’.

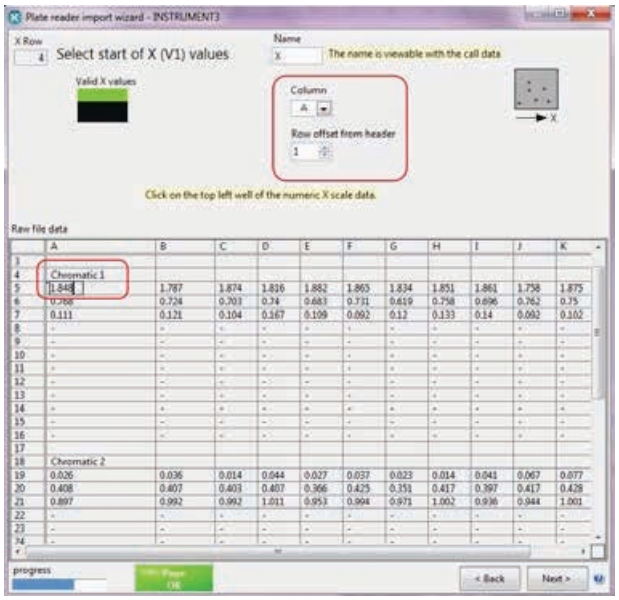


- The wizard then requires information regarding the position of the start of the X data values in the raw data file.

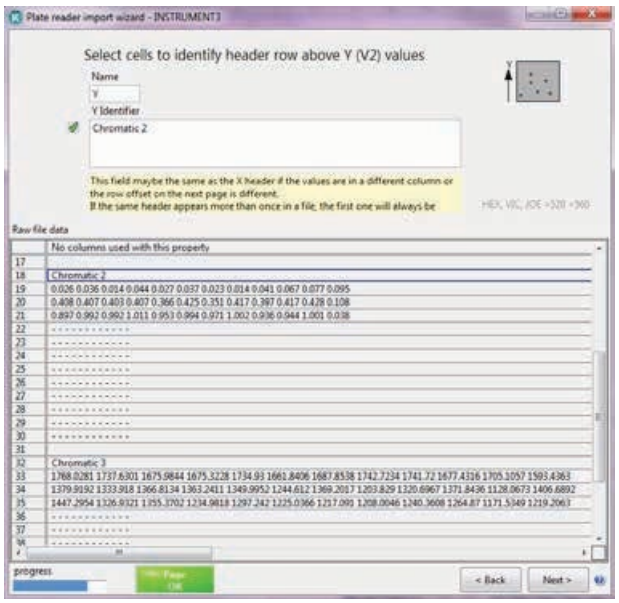


For Research Use Only. Not for use in diagnostic procedures.

- In this example, the X (FAM) data starts in cell A5 so click on this cell to identify the first X data value. Press ‘Next’.

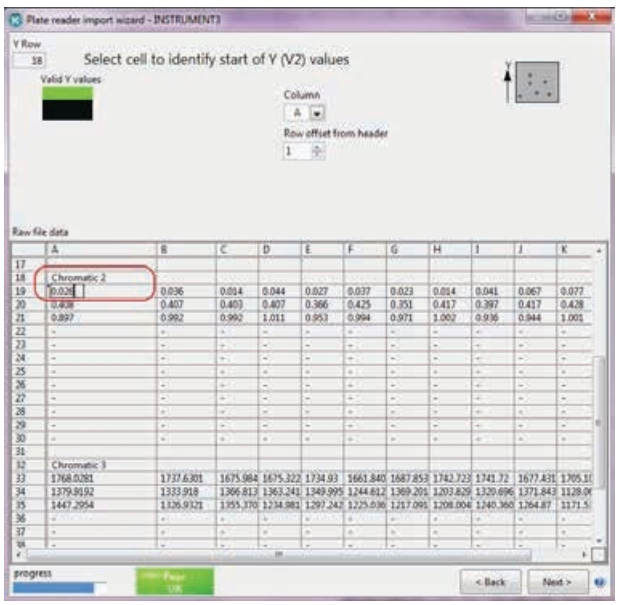


- The wizard then requires a reference to the header row in the data file that appears above the Y data values (Y = HEX data). Scroll down the raw data file and click on the header row (in this example it is row 18), and the contents will appear in the ‘Y identifier’ box at the top of the window. Press ‘Next’.

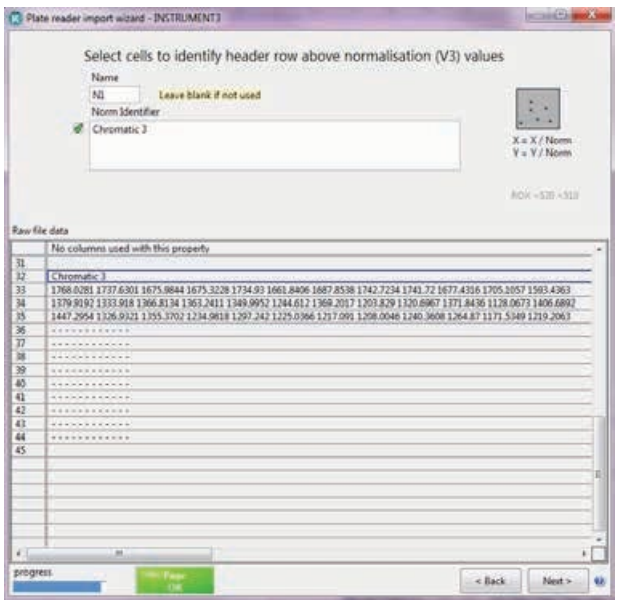


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information regarding the position of the start of the Y data values in the raw data file. In this example, the Y (HEX) data starts in cell A19 so click on this cell to identify the first Y data value. Press ‘Next’.

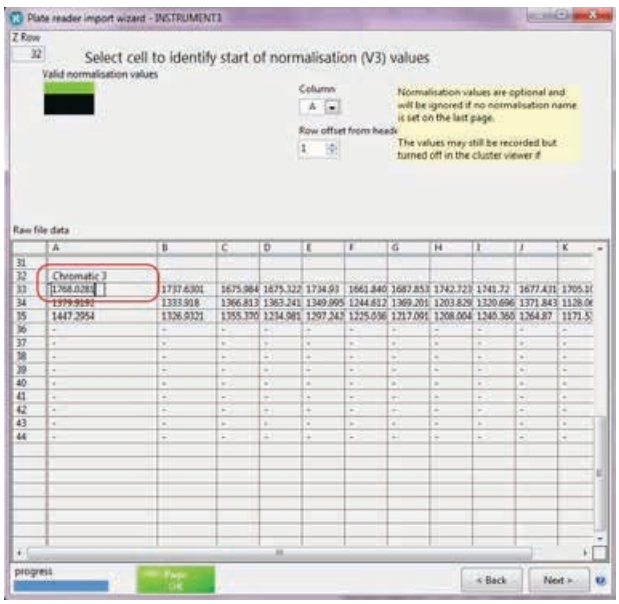


- The wizard then requires a reference to the header row in the data file that appears above the Z data values (Z = ROX data). Scroll down the raw data file and click on the header row (in this example it is row 32), and the contents will appear in the ‘Norm identifier’ box at the top of the window. Press ‘Next’.

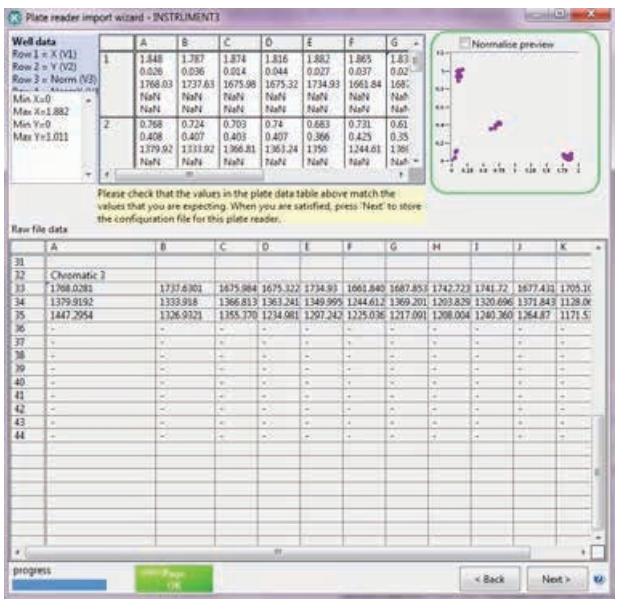


For Research Use Only. Not for use in diagnostic procedures.

- The wizard then requires information regarding the position of the start of the Z data values in the raw data file. In this example, the Z (ROX) data starts in cell A33 so click on this cell to identify the first Z data value. Press ‘Next’.

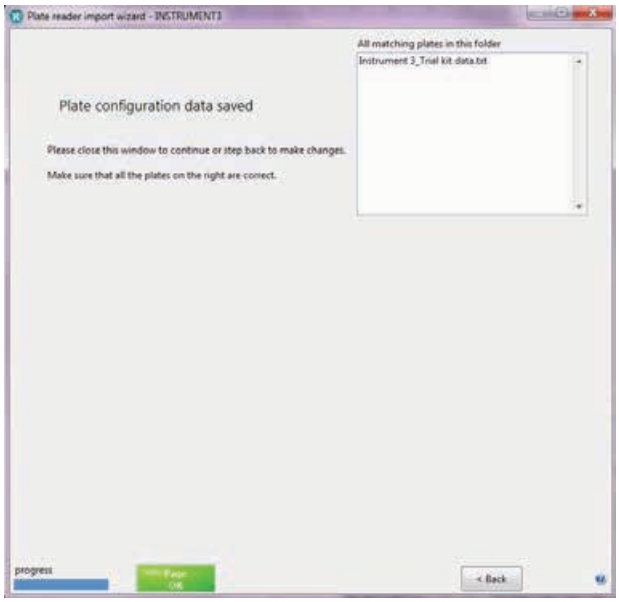


- The wizard will now show a preview of your data, both as numerical values in a plate layout format and as a cluster plot. If this data looks as expected (i.e. A1 data is shown in the A1 grid layout at the top of the window), click ‘Next’.

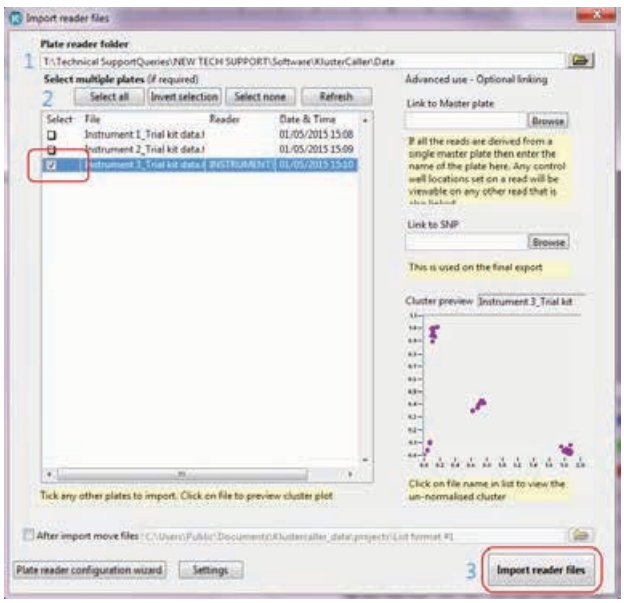


For Research Use Only. Not for use in diagnostic procedures.

- The wizard will then open a confirmation window, stating that the configuration file for this data format has been saved. All raw data files that are recognised by this configuration file will be listed in the top right hand corner. Click on the red cross to close the wizard.



- KlusterCaller will then return to the 'Import reader files' window. Place a tick in the box next to the relevant raw data file – this should now be recognised by the software. A preview cluster plot will also be visible. Click on the 'Import reader files' button to import this data file to your KlusterCaller project.



For Research Use Only. Not for use in diagnostic procedures.

5. Summary

In this document, we have outlined how to import the three most common formats of genotyping data using the plate configuration wizard. Using these examples, the same process can be applied to other data formats if your file does not match these exactly.

The KlusterCaller manual can be accessed from the Help menu within KlusterCaller and contains a lot of basic information about how to start using KlusterCaller for analysis of your genotyping data.

Should you need any further assistance or have a question not addressed in the manual please the technical support team (contact details below).

6. Further support

If you require additional support, please contact our technical support team at techsupport@lgcgroup.com.

For Research Use Only. Not for use in diagnostic procedures.

For Research Use Only. Not for use in diagnostic procedures.

@LGCBiosearch | biosearchtech.com

All trademarks and registered trademarks mentioned herein are the property of their respective owners. All other trademarks and registered trademarks are the property of LGC and its subsidiaries. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any retrieval system, without the written permission of the copyright holder. © LGC Limited, 2024. All rights reserved. GEN/1177/EK/0924

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC