Genotyping results data from LGC

Upon completion of your genotyping service project, your project manager will provide you with your results in a standard format

Overview

The purpose of this document is to provide an overview of the genotyping data files that you will receive from LGC at the end of your genotyping service project.

Data format

LGC genotyping results files are provided in a CSV (comma separated value) format. This file is designed to be viewed with LGC's SNPviewer application, which is available as a free download from the LGC website.

The CSV file can also be opened with most spreadsheet software and programs that can read plain text documents (e.g. MS Excel or Notepad).

SNPviewer

SNPviewer provides a simple means to graphically view the results as genotyping cluster plots. A Cartesian plot is generated using the FAM and HEX fluorescent values; FAM is plotted on the X-axis and HEX is plotted on the Y-axis.

Genotyping calls displayed in SNPviewer cannot be edited

NB. To be able to download SNPviewer you must first create a user account on our webpage. Click here to download SNPviewer.

Opening a results file in SNPviewer

To open a genotyping results file in SNPviewer, first make sure that you have downloaded and installed SNPviewer and saved the results file to your computer.

1. Open SNPviewer by clicking on the icon
2. A new window will appear that will allow you to select your results file. Choose the correct file and click 'OK'.
3. SNPviewer will now display the results from the file. Figure 1 shows a typical genotyping project as viewed in SNPviewer. The area on the right hand side of the SNPviewer window is used to select assay and master plates, and the corresponding cluster plots are displayed on the left hand side of the SNPviewer window.
4. When an assay name is selected from the list, the context sequence of the DNA region is displayed at the top of the SNPviewer window (Figure 1, yellow box). The genotyping results are reported in terms of this context sequence e.g. if the sequence states [C/T], then the C allele will be reported with FAM (X axis) and the T allele will be reported with HEX (Y axis).

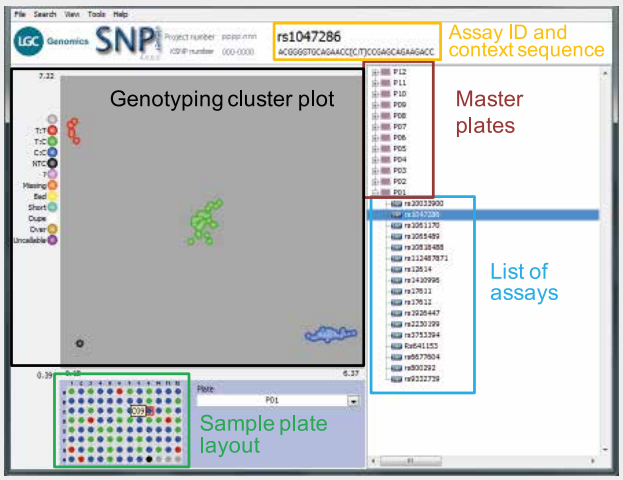

Figure 1. A typical genotyping project as viewed in SNPviewer. The master plates are listed in the tree on the right hand side, and can be expanded to show the assays that have been run on each plate of DNA samples. The data point colours correspond to the different genotype calls.

Opening a results file in Excel / similar

When opening the CSV file in Excel or other spreadsheet software packages, the genotyping data will be displayed in a list format. The data includes the master plate name, the genotyping call, the assay name, the DNA sample name, and the well location. Table 1 provides a detailed explanation of the abbreviations used in the 'Statistics' section of the results file.

Abbreviation	Meaning
SNP	Your assay reference name / ID
Plate	Your plate reference name / ID
X	The number of samples genotyped as homozygous for allele X (listed per plate). These are reported by the FAM fluorophore, plotted on the X-axis and designated by the blue colour.
Het	The number of samples genotyped as heterozygous for allele X and Y (listed per plate). These are reported by both the FAM and HEX fluorophores, plotted on both the X-axis and Y-axis and designated by the green colour.
Y	The number of samples genotyped as homozygous for allele Y (listed per plate). These are reported by the HEX fluorophore, plotted on the Y-Axis and designated by the red colour.
NTC	The number of samples designated as **N**o - **T**emplate – **C**ontrols (listed per plate). These are negative controls and are necessary to assess the performance of the assay(s) and reagents.
Unused	The number of samples that cannot be reliably assigned to a genotype cluster (listed per plate).
Missing	The number of samples that are listed in the submitted plate map but did not arrive / were not present in physical plate therefore cannot be genotyped (listed per plate).
Bad	The number of samples that consistently produce spurious data or do not amplify. These are marked "**bad**" and removed from any further analysis genotyped (listed per plate).
Allele X%	Frequency of allele X listed by plate
Allele Y%	Frequency of allele Y listed by plate
X2	Chi Squared (χ^2) test. This is a statistical test used to assess distribution in multinomial data sets, based on the Hardy-Weinberg equation.
Allele X	Lists the base / sequence assigned to allele X
Allele Y	Lists the base / sequence assigned to allele Y

Table 1. Abbreviations used in the 'Statistics' section of LGC's genotyping data file.

Excel grid files

In addition to the full results file (CSV, to be opened in conjunction with SNPviewer), LGC typically provide an additional 'Grid' file (CSV) that summarises the genotyping results in an easy-to-view format. When opened with Excel, the genotyping results are detailed in a grid, with DNA samples listed vertically and assay names listed horizontally (See Figure 2).

	A	B	C	D	E	F	G	H	I
1	DNA \ assay	rs1234	rs5678	rs9101	rs1121	rs3141	rs5161	rs7181	rs9202
2	Sample 1	T:T	C:C	T:C	G:A	A:A	T:T	G:A	T:A
3	Sample 2	T:T	C:C	T:C	G:A	G:G	T:T	A:A	T:A
4	Sample 3	T:T	C:C	T:C	G:A	G:G	T:A	G:G	T:A
5	Sample 4	T:T	C:C	T:C	G:A	A:A	T:A	G:G	A:A
6	Sample 5	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
7	Sample 6	T:T	C:C	C:C	G:A	A:A	A:A	G:G	T:T
8	Sample 7	C:C	T:C	C:C	G:A	A:A	A:A	G:G	T:A
9	Sample 8	C:C	T:C	C:C	G:A	?	A:A	G:G	A:A
10	Sample 9	C:T	C:C	T:T	A:A	A:A	?	G:G	T:T
11	Sample 10	Missing	Missing	Missing	Missing	Missing	Missing	Missing	Missing
12	Sample 11	Dupe	T:T	T:C	G:G	G:A	T:T	G:G	T:A
13	Sample 12	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC
14	Sample 13	C:T	C:C	C:C	G:A	A:A	T:A	G:G	T:A

Assay ID

DNA sample ID

Genotyping calls in grid layout

Result	Explanation
Genotype:Genotype e.g. T:T or T:A	Genotyping call in the format allele:allele
Bad	Sample removed from analysis e.g. consistently does not amplify
?	Genotype could not be determined
Missing	DNA sample not present in sample plate
Dupe	Sample ID appears within project multiple times with conflicting calls

Figure 2. Genotyping results in the 'Grid' file (CSV), opened in Excel. The assay IDs are listed horizontally (yellow) and the DNA samples are listed vertically (red). Genotyping (allele) calls are detailed within the grid.

If you need any assistance, please contact our technical support team on +44 (0)1992 470 757 **or email us at** tech.support@lgcgroup.com