Genotyping results

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, Biosearch Technologies™ at the end of your genotyping service project.

Data format

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

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.

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, orange 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).

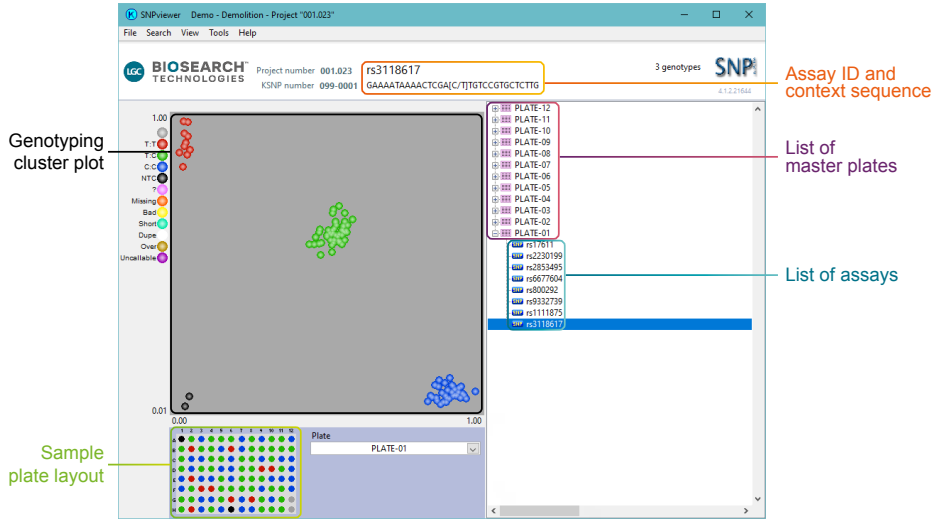


Figure 1. A typical genotyping project as viewed in SNPviewer. The master plates are listed in the tree on the right hand side, and can be expanded to show the assays that have been run on each plate of DNA samples. The data point colours correspond to the different genotype calls.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Opening a results file in Excel/similar

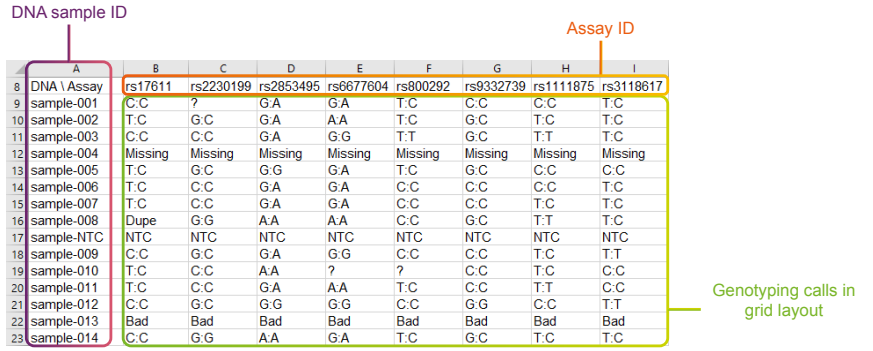
When opening the CSV file in Excel or other spreadsheet software packages, the genotyping data will be displayed in a list format. The data includes the master plate name, the genotyping call, the assay name, the DNA sample name, and the well location. Table 1 provides a detailed explanation of the abbreviations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

Abbreviation	Meaning
SNP	Your assay reference name or ID
Plate	Your plate reference name or 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 No Template Controls (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 or 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 (listed per plate).
Polyploid Het1 *	The number of samples that are assigned this polyploid heterozygous genotype (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 dark yellow colour.
Polyploid Het2 *	The number of samples that are assigned this polyploid heterozygous genotype (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 dark green colour.
Polyploid Het3 *	The number of samples that are assigned this polyploid heterozygous genotype (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 dark red colour.
Polyploid Het4 *	The number of samples that are assigned this polyploid heterozygous genotype (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 dark purple colour.
Allele Y%	Frequency of allele Y listed by plate
Allele X%	Frequency of allele X 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 Y	Lists the base or sequence assigned to allele Y
Allele X	Lists the base or sequence assigned to allele X

Table 1. Abbreviations used in the 'Statistics' section of Biosearch Technologies'' genotyping results file. * The Polyploid Het genotypes (#1-4) are optional and can be used when analysing genotyping results for polyploid organisms. If your genotyping project does not use polyploid genotypes then there will be no Polyploid Het samples in the results.

Excel grid files

In addition to the full results file (CSV, to be opened in conjunction with SNPviewer), Biosearch Technologies 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
8	DNA \ Assay	rs17611	rs2230199	rs2853495	rs6677604	rs800292	rs9332739	rs1111875	rs3118617
9	sample-001	C:C	?	G:A	G:A	T:C	C:C	C:C	T:C
10	sample-002	T:C	G:C	G:A	A:A	T:C	G:C	T:C	T:C
11	sample-003	C:C	C:C	G:A	G:G	T:T	G:C	T:T	T:C
12	sample-004	Missing	Missing	Missing	Missing	Missing	Missing	Missing	Missing
13	sample-005	T:C	G:C	G:G	G:A	T:C	G:C	C:C	C:C
14	sample-006	T:C	C:C	G:A	G:A	C:C	C:C	C:C	T:C
15	sample-007	T:C	C:C	G:A	G:A	C:C	C:C	T:C	T:C
16	sample-008	Dupe	G:G	A:A	A:A	C:C	G:C	T:T	T:C
17	sample-NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC
18	sample-009	C:C	G:C	G:A	G:G	C:C	C:C	T:C	T:T
19	sample-010	T:C	C:C	A:A	?	?	C:C	T:C	C:C
20	sample-011	T:C	C:C	G:A	A:A	T:C	C:C	T:T	C:C
21	sample-012	C:C	G:C	G:G	G:G	C:C	G:G	C:C	T:T
22	sample-013	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
23	sample-014	C:C	G:G	A:A	G:A	T:C	G:C	T:C	T:C

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 (orange) and the DNA samples are listed vertically (purple). Genotyping (allele) calls are detailed within the grid.

If you need any assistance, please email our technical support team at techsupport@lgcgroup.com.

Integrated tools. Accelerated science.

@LGCBiosearch | biosearchtech.com

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC