

PLINK Input Report Plug-in v2.1.3 for GenomeStudio® Software

1. Introduction

The PLINK Input Report Plug-in is a software plug-in that works with Illumina's GenomeStudio Genotyping module. This plug-in allows you to create input files for the PLINK application, developed by Shaun Purcell of MGH/Broad.

PLINK can be downloaded from <http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml>. Among its many uses, PLINK is designed to efficiently run association analysis on GWAS (Genome-Wide Association Studies) data sets, including eQTL (Expression Quantitative Trait Loci) and similar analysis types.

2. Installing the PLINK Input Report Plug-in

Perform the following steps to install the PLINK Input Report plug-in on your computer:

1. Download the PLINK Input Report Plug-in setup program (*.msi) from the [illumina•connect](http://www.illumina.com/illuminaconnect) website, <http://www.illumina.com/illuminaconnect>, or from the GenomeStudio Portal.
2. Run the PLINK Input Report Plug-in setup program.

The PLINK Input Report Plug-in is installed in the appropriate directory on your workstation.

3. Using the PLINK Input Report Plug-in

Perform the following steps to use the newly-installed plug-in to create a PLINK Input Report from data loaded into GenomeStudio software:

1. In the GenomeStudio application, open an existing genotyping project, or create a new genotyping project.
2. In the GenomeStudio genotyping project, select **Analysis | Reports | Report Wizard**.
The Report Wizard appears.
3. From the Custom Report dropdown list, choose **PLINK Input Report**.
4. If desired, adjust the algorithm input parameters from the GUI. Persistent changes to the input parameters can be achieved by editing the plug-in config file, as described in Section 4 of this document.

The Configuration Parameters table on the next page of this document includes descriptions of the input parameters.

5. Click **OK** to create the report.

A progress bar shows the report creation status. When the report has finished processing, a dialog box asking "**Would you like to view this report?**" appears.

6. Click **Yes** to display a text file showing the location of the output file.

Note: If the file is too large to open in your default Windows text viewer (e.g., NotePad), you can open it using WordPad.

Five files are created:

<projectname>.ped, the LINKAGE format input file

<projectname>.map, the map file, showing base pair and position for each marker

<projectname>.phenotype, a text file which lists the quantitative trait data for each sample.

<projectname>.script, a text file which lists the input parameters to be used by the PLINK executable.

<projectname>.bat, a batch file that runs PLINK with a default set of input parameters listed in the script file.

Notes

--A pedigree file is not needed if the "Mother ID", "Father ID", and "Affected Status" columns appear in the Samples table.

--Only sample IDs that appear in both the pedigree file and the genotyping project are processed by the report.

--Sample gender is determined using the Gender column in the Samples table.

4. Configuring the PLINK Input Report Plug-in

The PLINK Input Report format can be changed via an editable configuration file. When you install the PLINK Input Report Plug-in, the config file is installed on your computer in this location:

<C:\Program Files\Illumina\GenomeStudio\Modules\BSGT\ReportPlugins\PLINKInputReport>

Note: When you modify and save the default config file, changes are preserved for future sessions.

The following tables contain configuration settings for and descriptions of PLINK Input Reports.

Configuration Parameters

Parameter Name	Description	Default Value
PedigreeFile	[Optional] The path to the pedigree file	
TraitDataFile	[Optional] The path to the trait data file. You can create this report manually, or by using the GX Custom Output Report.	
AnalysisType	The type of PLINK analysis for which to create input files	Association
ExePath	The directory path to the PLINK executable file	C:\plink\plink-1.04-dos\plink.exe

Pedigree File Format (Tab-Delimited)

Column Name	Description	Mandatory?
Pedigree	Numeric ID representing the family ID	Yes
Sample ID	Sample ID used in the genotyping project	Yes
Father ID	Sample ID for the father; 0 for no father sample available	Yes
Mother ID	Sample ID for the mother; 0 for no mother sample available	Yes
Affected Status	0 = Unknown	Yes

	1 = Non-Affected 2 = Affected	
Additional Columns	Ignored	No

Trait Data File Format (Tab-Delimited)

The trait data file includes the quantitative trait data with probe IDs in row 1 and sample IDs in column 1.

5. Technical Support

Direct questions about the GenomeStudio Genotyping Module PLINK Input Report Plug-in to Illumina Technical Support at techsupport@illumina.com or +1.800.809.4566.

Direct questions about PLINK to plink@chgr.mgh.harvard.edu.

6. Version History

Build	Date
v2.1.3 for GenomeStudio Software	7/8/2010
<ul style="list-style-type: none"> Improved handling of excluded samples & duplicate sample IDs Improved genetic distance interpolation Added quotes around exe path in batch file Added timestamp to output directory For quantitative traits, output a -9 instead of NaN, per PLINK docs Added detection of "Mother ID", "Father ID", and "Affected Status" columns in samples table, and use those values in auto-created PED file Added ability to use forward strand, configured in GUI or with config parameter "UseForwardStrand" 	
Build	Date
v2.1.1 for GenomeStudio Software	11/17/2009
<ul style="list-style-type: none"> Fixed issue in which GenomeStudio projects with N excluded samples would omit the last N samples in Samples table from the PLINK *.ped and *.phenotype files 	
Build	Date
v2.1.0 for GenomeStudio Software	2/9/2009
<ul style="list-style-type: none"> Initial Release 	



For Research Use Only

© 2008–10 Illumina, Inc. All rights reserved.

Illumina, illuminaDx, Solexa, Making Sense Out of Life, Oligator, Sentrix, GoldenGate, GoldenGate Indexing, DASL, BeadArray, Array of Arrays, Infinium, BeadXpress, VeraCode, IntelliHyb, iSelect, CSPRO, GenomeStudio, Genetic Energy, HiSeq, and HiScan are registered trademarks or trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.