

Infinium® Genotyping Data Analysis

A guide for analyzing Infinium genotyping data using the Illumina GenomeStudio® Genotyping Module.

INTRODUCTION

Illumina DNA Analysis BeadChips using the Infinium Assay provide researchers unparalleled genomic access and accuracy for analyzing genetic variation. Illumina's portfolio of powerful products offers many benefits to researchers. In particular, BeadChips provide highly multiplexed sample analysis of 300,000 to more than 1,000,000 markers simultaneously. The density of arrayed features used for this high level of parallel analysis also allows for a high level of intrinsic assay redundancy. Each locus is assayed with an average 12- to 18-fold redundancy for each sample, yielding a highly robust intensity estimate that enables highly accurate genotype calls. Another strong contributor to data quality is outlier removal in the averaging of raw intensity data.

Each SNP is analyzed independently to identify genotypes. Genotypes are called by comparing customer-generated data with those in the supplied cluster file. Genotype calls are generally highly accurate and unambiguous for high quality samples. However, accuracy is highly sample dependent. When samples do not perform as expected, experimenters can choose to reprocess (requeue) these samples to confirm or potentially improve results, or to systematically exclude poorly performing samples from the project prior to recalling genotypes. Additionally, some samples (e.g., samples that have been whole-genome amplified) may not fit standard cluster positions well. A custom-generated cluster file may improve the call rate and accuracy for these aberrant cases.

This technical note describes the method used by the Illumina FastTrack Genotyping Services Group for analyzing Infinium genotyping data with the Illumina GenomeStudio Genotyping Module to optimize call rates. Analysis begins with preliminary sample quality evaluation to determine which samples may require reprocessing or removal. If a custom cluster file is required, clustering should be done after removal of failed or suboptimal samples. After selected samples have been reprocessed and included in the analysis, sample quality

can be evaluated in more detail. If required, individual locus analysis, where loci are identified by sorting on per-locus metrics such as call rate or cluster separation, can optimize data quality before generating a final report.

GENCALL SCORE CUTOFF

The first GenomeStudio parameter that should be optimized to obtain the highest genotyping accuracy is the gencall score cutoff, or no-call threshold. The gencall score is a quality metric calculated for each genotype (data point), and ranges from 0 to 1. Gencall scores generally decrease in value the further a sample is from the center of the cluster to which the data point is associated. The no-call threshold is the lower bound for calling genotypes relative to its associated cluster. Illumina FastTrack Genotyping Project Managers typically use a no-call threshold of 0.15 with Infinium data. This means that genotypes with a gencall score less than 0.15 are not assigned genotypes because they are considered to be too far from the cluster centroid to make reliable genotype calls. They are instead assigned a "no call" for that locus. No calls on successful DNAs at successful loci contribute to lowering the call rate for the overall project. The standard 0.15 no-call threshold for Infinium data was determined using projects with trio and replicate information and evaluating a range of no-call thresholds to optimize the call rate without compromising reproducibility or Mendelian consistency. The no-call threshold in GenomeStudio can be changed by selecting **Tools | Options | Project | No-call Threshold**. Sample and SNP statistics must be recalculated after an adjustment to the no-call threshold.

IDENTIFYING AND REPROCESSING POORLY PERFORMING SAMPLES

There are two options to ensure optimal analysis with samples that have not performed well. A sample may be reprocessed in an attempt to improve the results, or the sample can be excluded prior to reclustering. An initial assessment of DNA sample quality, as described below, can indicate which samples, if any, need to be repro-

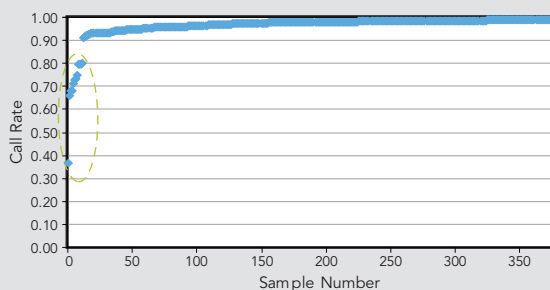
cessed. This procedure should be used after the Infinium Genotyping Assay is complete and all BeadChips are imaged in a project.

- 1) Open a new project in the GenomeStudio GT Module.
- 2) Load your data from a sample sheet or from LIMS using the SNP manifest file (*.bpm) and standard cluster file (*.egt) that correspond to the Infinium product used. For instructions on loading data from a sample sheet, refer to the GenomeStudio GT Module User Guide (Part #11319113).
- 3) After the GenomeStudio tables are populated with data click the **Calculate** button (🔢) above the Samples Table. This updates the average call rate in the Call Rate column for each sample across all loci. Average call rate expectations for standard human products are typically greater than 99%, but you should refer to specific call rate specifications for the relevant product.
- 4) To identify problematic samples, create scatter plots of call rate as a function of sample number (Figure 1) and 10% GenCall (10% GC) score as a function of sample call rate (Figure 2). To create graphs, use the graphing functions in GenomeStudio, or export the Sample Table data and reopen it in another graphing program. Poorly performing samples can be identified as having low sample call rates and low 10% GC scores. Samples with low call rates and 10% GC scores that are outliers from the main population should be considered for reprocessing.

However, if there are samples expected to have large amounts of loss of heterozygosity (LOH) or copy number variation (CNV), such as tumor or WGA samples, there may be outliers in call rate and 10% GC for biological reasons, not because of poor DNA quality. Therefore, these samples are not good candidates for reprocessing.

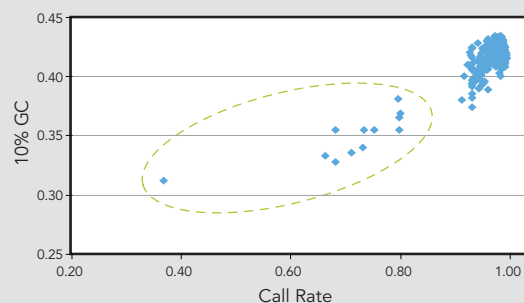
If you observe call rates below 99% across a majority of samples in human projects, the sample intensities measured might not match the intensities from the standard cluster file provided by Illumina. To test for this possibility, recluster SNPs on good samples only and recalculate the average call rate. Save this new preliminary cluster file with a meaningful name (e.g., Preliminary_Hap550_Project Name.egt). If the sample call rate is still low after reclustering SNPs on samples with the highest call rates and 10% GC scores only, there could be a systematic non-biological issue affecting the data quality.

FIGURE 1: SCATTER PLOT OF CALL RATES ACROSS A SET OF SAMPLES



Viewing the call rate of all samples quickly reveals poorly performing samples with abnormally low call rates (green circle).

FIGURE 2: SCATTER PLOT OF 10% GC SCORES COMPARED TO CALL RATES OF A SAMPLE SET



Poorly performing samples are obvious outliers from the majority of samples when 10% GC Score is plotted against sample call rate (green circle).

DETERMINING FINAL SAMPLE SET

After poorly performing samples have been reprocessed, the new requeued samples should be screened in GenomeStudio for the best performing instance of each sample. The higher quality instance of each sample (across all SNPs) should be retained for analysis. If both instances of a sample perform poorly, the entire sample should be excluded from the project prior to generating final reports. Below is the procedure for including only the instance of each sample that performs best.

- 1) Load reprocessed DNA samples into the existing GenomeStudio project by either selecting **File | Update Project from LIMS** or **File | Load Additional Samples**.
- 2) Select the appropriate sample sheet and data repository.

- 3) After all additional samples have been included, click the **Calculate** button.
- 4) Select **Analysis | Exclude Samples by Best Run**. This automatically excludes the lower quality queued samples throughout the entire project.
- 5) To determine whether any additional samples should be excluded from the project, reevaluate samples as described in the previous section.
- 6) Once the GenomeStudio project contains only final and unique samples, update SNP statistics.

LOCUS ANALYSIS AND RECLUSTERING

Standard cluster files provided with Infinium products identify expected intensity levels of genotype classes for each SNP. Comparing sample intensities to this cluster file is usually sufficient for generating extremely high quality data for human projects. Standard cluster files for standard human Infinium products are created using a diverse set of over 100 samples from the Caucasian (CEU), Asian (CHB+JPT), and Yoruban (YRI) HapMap populations, and therefore should incorporate much of the genetic diversity in these populations. If the genetic diversity of a sample set is well represented by the standard cluster file and sample call rates are high, it is appropriate to use this standard file.

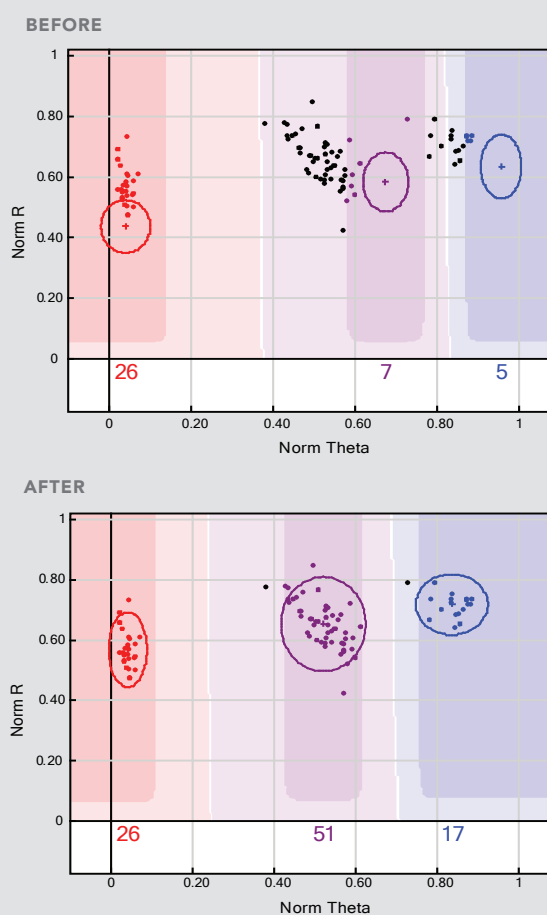
However, in some situations, sample intensities might not overlay perfectly onto the standard cluster positions. Reclustering some or all SNPs can optimize GenomeStudio's ability to call genotypes and results in higher overall call rates (Figure 3). In some cases, reclustering is not helpful and the locus must be zeroed (Figure 4). Reclustering using samples from a particular project will redefine cluster positions, making them more representative of the samples in the project.

All or a subset of loci can be reclustered to generate a custom cluster file. To recluster all SNPs in the GenomeStudio project select **Analysis | Cluster All SNPs**. To recluster a subset of SNPs, highlight relevant rows of the SNP Table, right-click the SNP Table, and select **Cluster Selected SNPs**. An alternative method is to use filters to identify and select SNPs that should be excluded from reclustering. Consult the GenomeStudio Framework User Guide (Part #11318815) for information about how to create filters for various metrics in the Full Data Table, SNP Table, or Samples Table. Generating a custom cluster file will necessitate a thorough evaluation of newly defined SNP cluster positions to ensure the accuracy of

the genotype data (see Evaluating and Editing SNP Cluster Positions later in this document).

An important consideration in the decision to recluster is that the GenomeStudio clustering algorithms require approximately 100 samples to produce cluster positions that are representative of the overlying population. Therefore, in projects with fewer than 100 different samples, it is best to use the standard cluster file for calling genotypes.

FIGURE 3: EXAMPLE OF A SUCCESSFULLY RECLUSTERED SNP



Three distinct clusters exist but are poorly represented by the standard cluster file. This leads to an initial SNP call frequency of 39.6% for this locus (before). After reclustering, the cluster positions better represent the data and yield a call frequency of 99.7% (after).

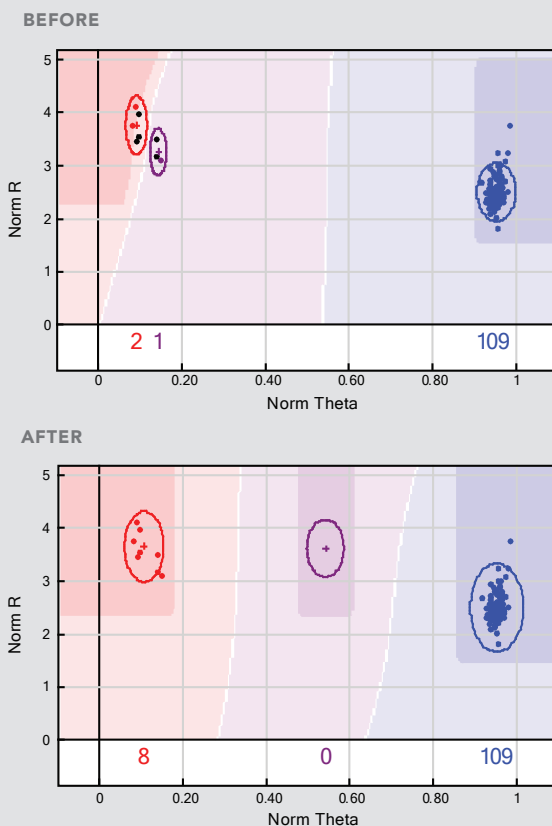
FIGURE 4: EXAMPLE OF AN UNSUCCESSFULLY RECLUSTERED SNP



At this locus, project samples form diffuse clusters and do not correlate well with the standard cluster positions (before). The genotype calls are unreliable because of poor cluster separation. Reclustering on this locus does not improve data quality. If not zeroed automatically by the algorithm (as above), this locus should be zeroed based on the low score for cluster separation.

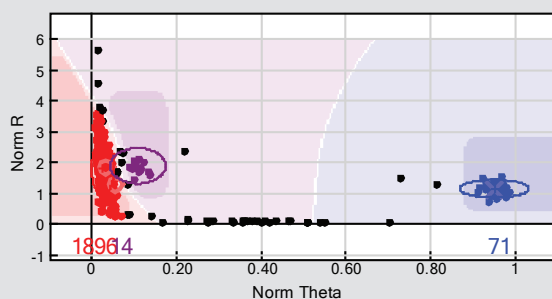
Special consideration must be given to mitochondrial (mtDNA) and Y chromosome (Y-chr) SNPs. The GenomeStudio clustering algorithms do not automatically accommodate loci that lack heterozygous clusters, so you must manually edit mtDNA (Figures 5 and 6) and Y-chr SNPs (Figures 7 and 8) following reclustering. Before reclustering all SNPs, evaluate the mtDNA and Y-chr SNPs to decide whether the predefined Illumina standard cluster positions are appropriate for the current project samples. If appropriate, exclude these loci from reclustering.

FIGURE 5: A SUCCESSFULLY EDITED MITOCHONDRIAL SNP



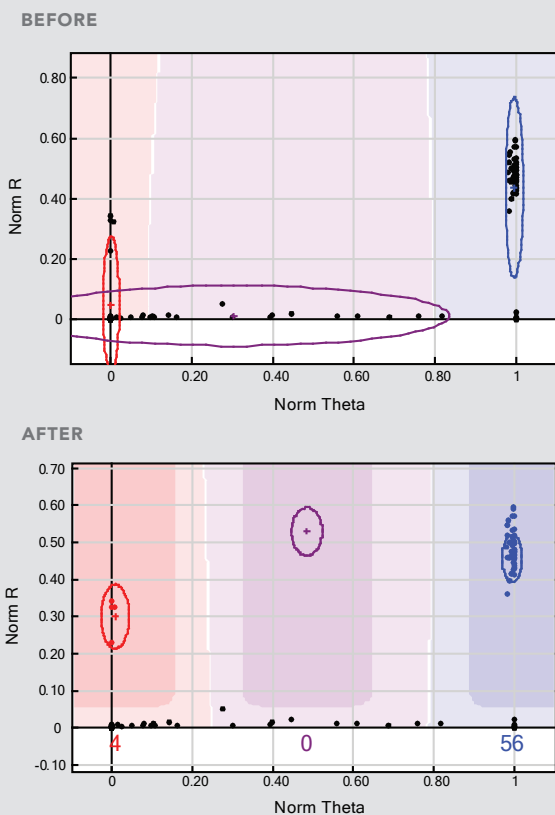
After reclustering, the AA cluster is incorrectly divided into two clusters (before) and must be manually edited (after).

FIGURE 6: AN UNSUCCESSFUL MITOCHONDRIAL SNP



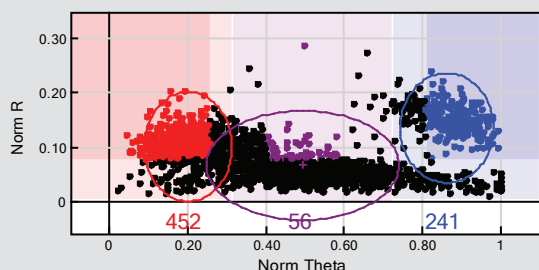
The clustering is ambiguous and genotypes are unreliable. This SNP should be zeroed.

FIGURE 7: A SUCCESSFULLY EDITED Y-CHROMOSOME SNP



Female samples have very low intensities and a broad range of theta values and are incorrectly called AB after reclustering (before). After manually adjusting the cluster positions, female samples are not called, and the two homozygous clusters correctly define male genotypes (after).

FIGURE 8: AN UNSUCCESSFUL Y-CHROMOSOME SNP



The clustering is ambiguous and the genotypes are unreliable. This SNP should be zeroed.

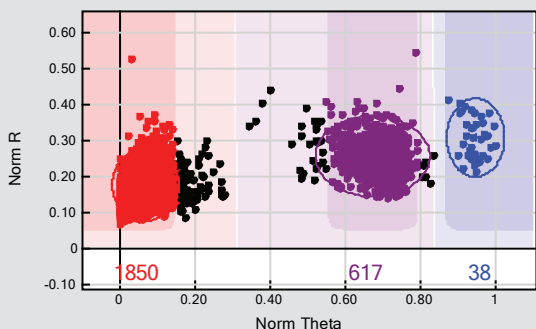
EVALUATING AND EDITING SNP CLUSTER POSITIONS

To identify loci that need to be manually edited or zeroed, evaluate newly reclustered SNPs using the metrics listed in the SNP Table. These metrics are based on all samples for each locus, and thus provide overall performance information for each locus. To find loci that might need to be edited or removed prior to making final reports, it is useful to sequentially sort the SNP Table by columns and screen loci by each quality metric. To determine hard cutoffs and grey zones, sort data by one column at a time and explore values starting at the extremes of the ranges. The hard cutoff should be defined as the level below (or above) which the majority of loci are unsuccessful and should be zeroed. The grey zone should be defined to contain loci that are 80–90% successful and can be improved by manual editing. The upper limit (or lower limit) of the grey zone is the point at which all loci are successful. SNPs falling in the grey zone should be evaluated and either zeroed or manually edited by moving cluster positions. Hard cutoffs and grey zones may not transfer between projects since they are highly dependent on sample quality and generally should be determined for each project.

The choice to manually edit loci should be carefully considered. Any changes made should be consistent with principles of population genetics and prevent introduction of subjective bias into the data set. Additionally, manually editing SNPs can be very time consuming. The following procedure describes a method of evaluating the quality of newly created cluster positions by sequentially sorting the SNP Table by various column statistics. For each step, it may be helpful to determine and record hard cutoff and grey zone thresholds.

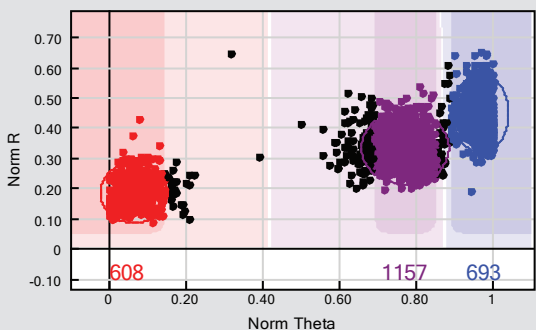
- 1) For projects with 100 samples or more, sort the SNP Table by **Cluster Sep**. **Cluster Sep** measures the separation between the three genotype clusters in the theta dimension and varies from 0–1. Evaluate individual SNPs for overlapping clusters, starting with those having low **Cluster Sep**. If clusters are well separated, the SNP can be manually edited (Figure 9). SNPs with overlapping clusters should be zeroed (Figure 10).

FIGURE 9: EXAMPLE OF A SUCCESSFUL SNP LOCUS



Clusters are sufficiently separated so that clustering is not ambiguous. Manually editing this locus can increase the call frequency.

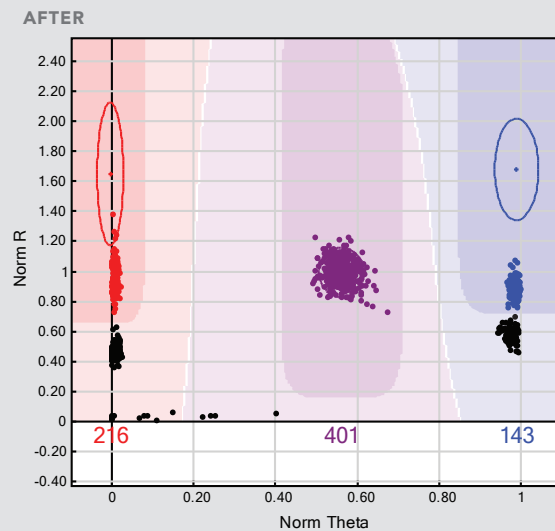
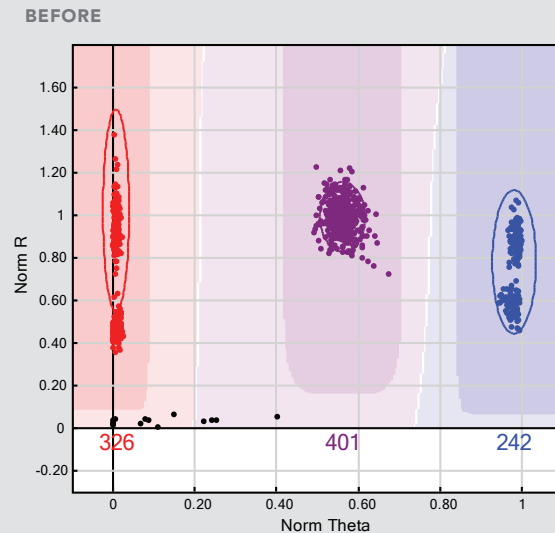
FIGURE 10: EXAMPLE OF AN UNSUCCESSFUL LOCUS



AB and BB clusters overlap and genotyping cannot be considered reliable. This SNP should be zeroed.

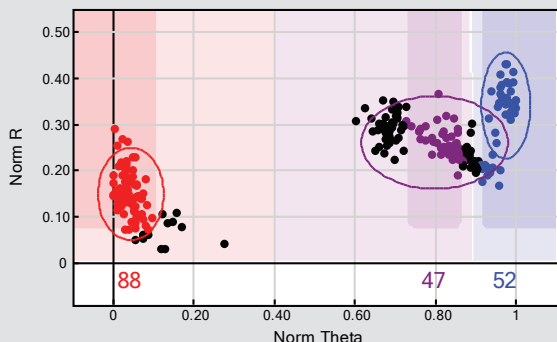
- Sort the SNP Table by call frequency (**Call Freq**). **Call Freq** is the proportion of all samples at each locus with call scores above the no-call threshold. The value varies from 0–1. Evaluate SNPs starting with those having low **Call Freq** values. Zero the SNP if the low call frequency cannot be attributed to a potential biological effect (Figures 11 and 12).

FIGURE 11: A SUCCESSFUL LOCUS WITH LOW CALL FREQUENCY



Although this SNP has a call frequency that falls below expectations, it has a phenotype that may indicate the presence of a deletion or a third polymorphic allele. There are two AA clusters (possibly AA and A/-), an AB cluster, and two BB clusters (possibly BB and B/-). In this case, samples that are not called (the black samples along the x axis) may represent a biologically meaningful failure of those samples (e.g., -/-). We recommend no-calling the A/- and B/- clusters.

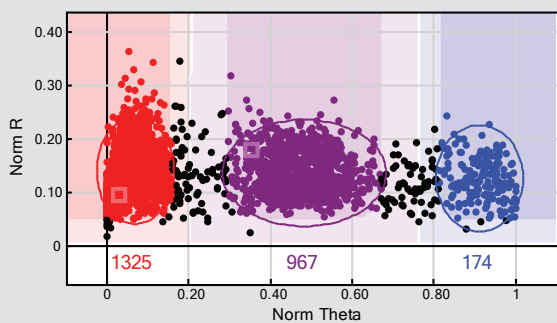
FIGURE 12: AN UNSUCCESSFUL LOCUS WITH LOW CALL FREQUENCY



AB and BB clusters are not distinct and genotyping cannot be considered reliable. This SNP should be zeroed.

- Sort the SNP Table by **AB R Mean**, the mean normalized intensity (R) of the heterozygote cluster. This metric helps identify SNPs with low intensity data and has values increasing from 0. Evaluate SNPs from low to high **AB R Mean** and zero any SNPs where the intensities are too low for genotypes to be called reliably (Figure 13).

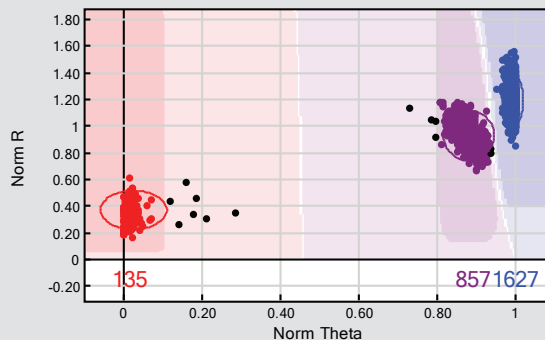
FIGURE 13: SNP WITH LOW INTENSITIES



Intensities (Norm R) are too low for genotypes to be reliably called. This SNP should be zeroed.

- Sort the SNP Table by **AB T Mean**, the mean of the normalized theta values of the heterozygote cluster. This value ranges from 0–1. Evaluate SNPs with **AB T Mean** ranging from 0–0.2 and 1–0.8 (or more, if necessary) to identify SNPs where the heterozygote cluster has shifted toward the homozygotes. Edit the SNP if clusters can be reliably separated (Figure 14); otherwise, zero the locus.

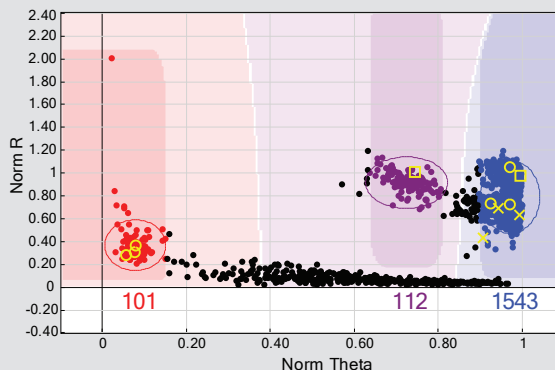
FIGURE 14: A SUCCESSFUL SNP LOCUS WITH HIGH AB T MEAN VALUE



This locus can be manually edited to include more AA samples to improve call frequency. There is no ambiguity in clusters at this locus.

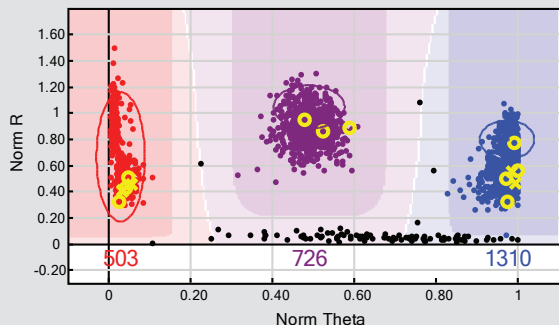
- To identify errors in Mendelian inheritance (MI), sort the SNP Table by columns **P-P-C Errors** and **P-C Errors**, both of which measure deviations from expected allelic inheritance patterns in matched parent and child samples. Values range from zero to two times the maximum number of trios included in the project. In the GenomeStudio SNP Graph, MI errors are displayed as circles (o) for parents and a cross (x) for a child. Sort by this column and evaluate all SNPs with one or more errors. Zero SNPs with both MI errors and ambiguous clusters (Figure 15). SNPs associated with copy number polymorphisms may exhibit P-P-C and P-C errors that are biologically meaningful. Therefore, loci with errors should be zeroed only if the clustering is also ambiguous (Figure 16).

FIGURE 15: AN UNSUCCESSFUL LOCUS WITH MI ERRORS



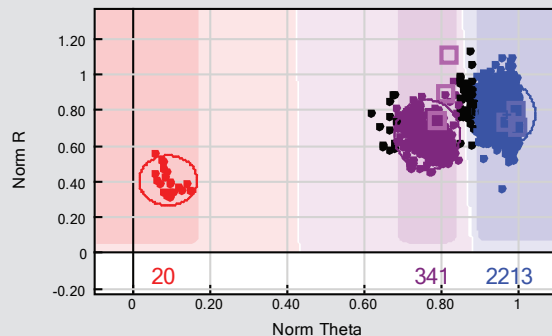
This locus exhibits six MI errors and ambiguous clusters (left BB cluster could be AB). This locus should be zeroed.

FIGURE 16: EXAMPLE OF A SNP WITH ERRONEOUS TRIOS



Clustering is not ambiguous and the locus has a phenotype that may indicate the presence of a chromosomal deletion or third polymorphic allele. This locus might be of special interest.

FIGURE 18: A FAILED LOCUS WITH REPRODUCIBILITY ERRORS

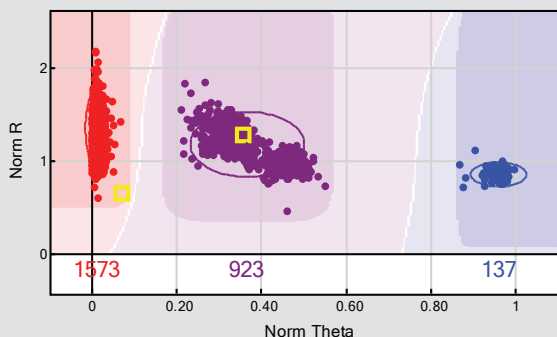


AB and BB clusters overlap. Genotypes cannot be called reliably so this locus should be zeroed.

- 6) The **Rep Errors** column of the SNP Table measures the reproducibility of genotype calls for replicate samples at each SNP and ranges from 0 to the maximum number of replicates included in the project. Sort by this column and evaluate all SNPs with one or more errors (Figures 17 and 18). Reproducibility errors are displayed as squares in the GenomeStudio SNP Graph.

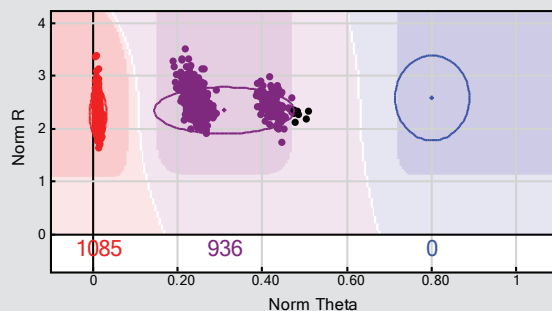
- 7) The SNP Table metric, **Het Excess**, is an indicator of the quantity of excess heterozygote calls relative to expectations based on Hardy-Weinberg Equilibrium. This metric varies from -1 (complete deficiency of heterozygotes) to 1 (100% heterozygotes). Evaluate SNPs with **Het Excess** values less than -0.3 or greater than 0.2 (or more, if necessary). If clusters are unambiguous (Figure 19), edit the SNP and retain it for the final report; otherwise, the SNP should be zeroed (Figure 20). Because males are not expected to be heterozygous for X chromosome loci, the **Het Excess** metric might not be relevant for X-linked SNPs. To evaluate X chromosome SNPs, see step 9.

FIGURE 17: A SUCCESSFUL LOCUS WITH A REPRODUCIBILITY ERROR



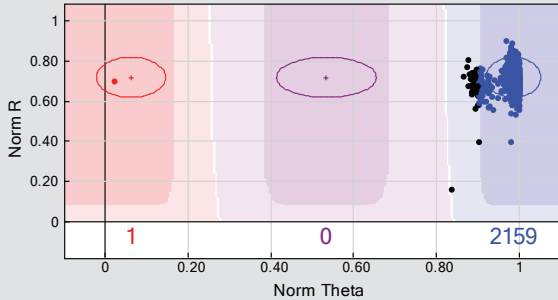
There is one erroneous genotype (the AA replicate should be AB, and could be manually no-called) and the clustering is not ambiguous at this locus.

FIGURE 19: EXAMPLE OF A SNP WITH HIGH HETEROZYGOTE EXCESS



Clusters can be separated and are not ambiguous. If the algorithm minimum cluster separation parameters allow it, this locus can be manually edited to call genotypes correctly.

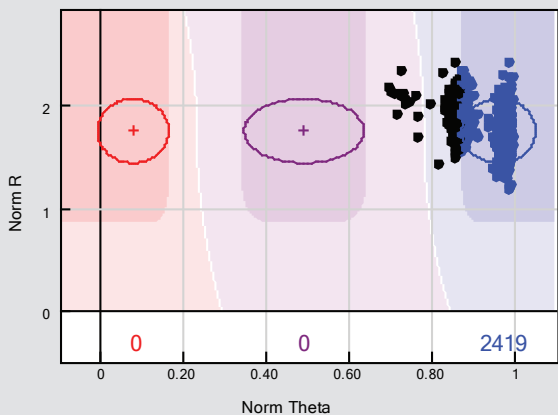
FIGURE 20: EXAMPLE OF SNP WITH A HETEROZYGOTE DEFICIENCY



The clustering is ambiguous (left BB cluster could be AB). This locus should be zeroed.

- 8) The SNP Table column **Minor Freq** measures the SNP minor allele frequency and can help identify loci where homozygotes have been incorrectly identified. **Minor Freq** values vary from 0–1 and all SNPs with **Minor Freq** less than 0.1 should be evaluated to detect false homozygotes. A SNP can be edited unless clusters overlap or cannot be separated unambiguously (Figure 21).

FIGURE 21: EXAMPLE OF A SNP WITH MINOR FREQ OF 0

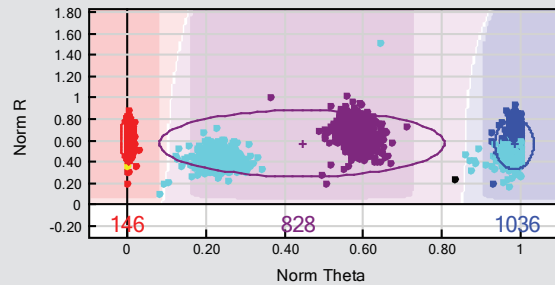


This SNP has been incorrectly clustered with a Minor Freq of 0. In this case, clusters are too close together for genotypes to be called and the locus should be zeroed.

- 9) Because males are not expected to be heterozygous for any X-linked markers, these loci need to be evaluated while taking gender into account. In GenomeStudio, the gender of each sample is estimated using X chromosome SNPs (**Gender Est** column in the Samples Table). Males can be selected

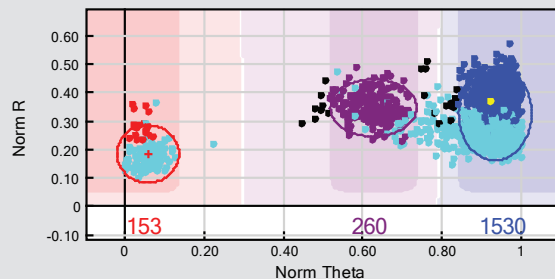
in the Samples Table and marked with a different color on the SNP Graph to easily evaluate SNPs on the X chromosome. If males are called heterozygote at X chromosome SNPs, clusters may need to be manually adjusted so that they are called homozygote (Figure 22). However, if clustering is ambiguous, the SNP should be zeroed (Figure 23).

FIGURE 22: EXAMPLE OF X CHROMOSOME SNP THAT CAN BE SUCCESSFULLY EDITED



The clustering algorithm called males AB at this X chromosome SNP. The males (light blue) that are called heterozygote should be manually edited to be included in the AA cluster.

FIGURE 23: EXAMPLE OF AN UNSUCCESSFUL X LOCUS



Male AB and BB clusters (light blue) overlap and genotyping cannot be considered reliable.

FINAL REPORT

After completing the above data analysis and editing SNP clusters, calculate sample statistics by clicking the **calculator** button in the Samples Table toolbar. If necessary, also update sample reproducibility and heritability statistics by selecting **Analysis | Update Heritability/Reproducibility Errors**. At this point, the GenomeStudio project is final. Save changes using **File | Save Project Copy As**. The newly created cluster file (*.egt) can be exported from the GenomeStudio project by selecting **File | Export Cluster Positions**.

The Final Report Wizard lets you export the genotyping data from GenomeStudio for use in downstream analysis applications. To run the Final Report Wizard, select **Analysis | Reports | Report Wizard** and follow the on-screen instructions to filter the project data for excluded samples, non-zeroed SNPs, by group, or by other attributes. The Final Report Wizard also allows you to include SNP annotation and choose from among a variety of formats to accommodate most genotyping analysis packages.

SUMMARY

By identifying problematic samples and loci in a systematic manner, a project data set can be analyzed comprehensively. This takes full advantage of the robustness of Illumina Genotyping BeadChips in genotyping experiments. The sample workflow described in this document is a method recommended by Illumina to optimize final data quality. By editing loci that are not clustered or called correctly, collected data can be fully utilized. When editing is not possible for unsuccessful loci or samples, excluding them from the data set ensures that the remaining data are of the highest quality.

APPENDIX: DNA REPORT COLUMN DEFINITIONS

#No_Calls: The total number of genotypes in each sample with a GenCall score below the no-call threshold as defined in the project options. Genotypes that are not called are shown on the GenomeStudio SNP Graph as black points falling outside of the darkly shaded regions.

#Calls: The total number of genotypes in each sample with a GenCall score above the no-call threshold.

Call_Freq: Call_Freq is equal to $\#Calls / (\#No_Calls + \#Calls)$. Call_Freq is equivalent to Call Rate in the GenomeStudio Samples Table.

A/A_Freq: For each sample, the number of AA genotype calls divided by #Calls.

A/B_Freq: For each sample, the number of AB genotype calls divided by #Calls.

B/B_Freq: For each sample, the number of BB genotype calls divided by #Calls.

Minor_Freq: If the number of AA calls is less than the number of BB calls for a sample, the frequency for the minor allele A is $[(2 \cdot AA) + AB]$ divided by $[2 \cdot (AA+AB+BB)]$ across all called loci for that sample.

GenCall Score: This score is a quality metric that indicates the reliability of the genotypes called. A GenCall score value is calculated for every genotype and can range from 0.0 to 1.0. GenCall scores are calculated

using information from the sample clustering algorithm. Each SNP is evaluated based on the angle of the clusters, dispersion of the clusters, overlap between clusters, and intensity. Genotypes with lower GenCall scores are located furthest from the center of a cluster and have lower reliability. There is no global interpretation of a GenCall score as it depends on the clustering of samples at each SNP, which is affected by many different variables, including the quality of the samples and loci.

50%_GC_Score: For each sample this represents the 50th percentile of the distribution of GenCall scores across all called genotypes. For SNPs across all samples, this is referred to as the 50%_GC_Score. For samples across all loci, it is referred to as p50GC in the Samples Table.

10%_GC_Score: For each sample this represents the 10th percentile of the distribution of GenCall scores across all called genotypes. For SNPs across all samples, this is referred to as the 10%_GC_Score. For samples across all loci, it is referred to as p10GC in the Samples Table. Note: Call frequency, 50% GenCall score, and 10% GenCall score are useful metrics for evaluating the quality and performance of DNA samples in an experiment.

O/1: GenomeStudio calculates a threshold from the distribution of 10%_GC_Score values across all samples in the DNA report. A '1' is assigned to samples whose 10%_GC_Score is at or above this threshold. A '0' is assigned to samples whose 10%_GC_Score is below this threshold. The equation defining this threshold is $0.85 \cdot 90^{\text{th}}$ percentile of 10%_GC_Score values for all samples in DNA Report (i.e., $0.85 \cdot 90^{\text{th}}$ percentile of column K in the DNA report).

ADDITIONAL INFORMATION

Visit our website or contact us at the address below to learn more about Illumina's DNA Analysis and Software products.

Illumina, Inc.
Customer Solutions
 9885 Towne Centre Drive
 San Diego, CA 92121 USA
 1.800.809.4566 toll-free
 1.858.202.4566 tel
 techsupport@illumina.com
 www.illumina.com

FOR RESEARCH USE ONLY

© 2010 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, and HiSeq are registered trademarks or trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.

Pub. No. 970-2007-005 Current as of 13 January 2010

illumina®