Appendix E: Contrast QC for SNP 6.0

Contrast QC is the per sample Quality Control test metric for SNP 6.0 intensity data (CEL files). When all steps of the assay are working as expected, the Contrast QC is typically greater than 0.4. As an added flag for potentially problem data sets, check that the proportion of samples that fall below the 0.4 threshold are less than 10%, and the average of the samples that pass this 0.4 test are greater than or equal to 1.7. If the proportion falling below 0.4 is greater than 10%,or the average of the passing samples is at or below 1.7, then sample quality and process should be closely examined for possible issues.

The Contrast QC is a metric that captures the ability of an experiment to resolve SNP signals into three genotype clusters. It uses a static set of 10,000 randomly chosen SNP 6.0 SNPs, measuring the difference between peaks in "Contrast" distributions (Fig 217) produced by homozygote genotypes, and the valleys they share with the heterozygote peak, and takes the smaller of the two values. In poor quality experiments the homozygote peaks are not well-resolved from the heterozygote peak and the difference values approach zero. Contrast QC values are also computed for Contrast distributions produced by a static set of 20K randomly chosen SNPs on Nsp fragments only and a static set of 20K randomly chosen SNPs on Sty fragments only. These are called Contrast QC (Nsp) and Contrast QC (Sty); respectively. If the absolute difference between these two values is greater than two, this is evidence that that a sample may have worked properly with one enzyme set, but not with the other, and the Contrast QC value is adjusted to zero to reflect this problem. These Contrast QC values are well correlated with the higher Call Rates and concordance achieved when calls are subsequently made with Birdseed (versions 1 or 2). The correlation between Birdseed accuracy and Birdseed Call Rate is also very high. As an extra guard against the inclusion of any outlier samples that pass through the Contrast QC filter, it is a good idea to reject samples that are notable outliers in terms of their Birdseed Call Rate. When using Birdseed (v1), clustering larger batches of samples will improve the performance of the algorithm. The algorithm improvements in Birdseed v2 allow you to cluster by plate with the same performance as clustering larger batches of samples.
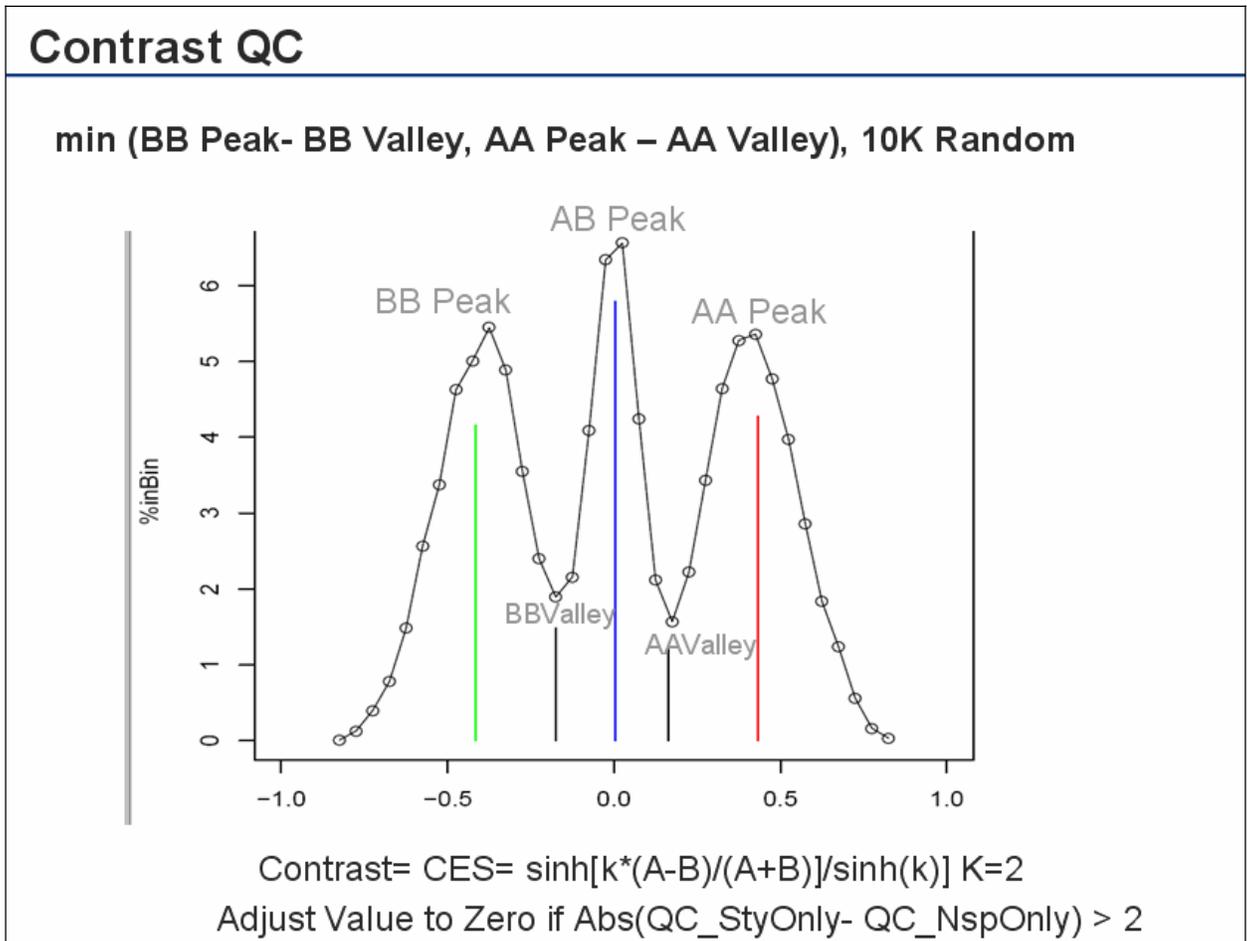
Figure 217. Distribution of Contrast Values. The X axis is the Contrast Value about which a bin of size 0.02 is centered. The Y axis is the %of SNPs (10000 random autosomal GW 6 SNPs) whose Contrast values fall within the bin. Contrast = sinh[K*(A-B)/(A+B)]/sinh(K)], K=2, A and B are the summary values for probes covering the A and B alleles; respectively (see http://www.affymetrix.com/support/technical/whitepapers/brlmm_whitepaper.pdf).

The Contrast QC is adjusted to zero if abs[Contrast QC (Nsp)- Contrast QC (Sty)] > 2