







1.00

0.95

000

0.80

0.85

0.90 Confidence score





SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1: Accuracy versus dropped rate plots (ADPs) comparing CRLMM and BRLMM for each HapMap dataset are shown in separate plots. The curves are further stratified between heterozygotes and homozygotes as determined by the HapMap gold standard. A) 269 Affymetrix high-quality XBA arrays. B) 269 Affymetrix high-quality HIND arrays. C) 270 Affymetrix high-quality STY arrays. D) 152 Broad Institute first-pass STY arrays. E) 95 Affymetrix first-pass STY arrays. F) 270 Affymetrix high-quality NSP arrays. G) 152 Broad Institute first-pass NSP arrays. H) 95 Affymetrix high-quality NSP arrays.

Supplemental Figure 2: Confidence measure versus accuracy plots (CAPs) comparing CRLMM and BRLMM for each HapMap dataset are shown in separate plots. The curves are further stratified between heterozygotes and homozygotes as determined by the HapMap gold standard. A) 269 Affymetrix high-quality XBA arrays. B) 269 Affymetrix high-quality HIND arrays. C) 270 Affymetrix high-quality STY arrays. D) 152 Broad Institute first-pass STY arrays. E) 95 Affymetrix first-pass STY arrays. F) 270 Affymetrix high-quality NSP arrays. G) 152 Broad Institute first-pass NSP arrays. H) 95 Affymetrix first-pass NSP arrays.

Supplemental Figure 3: The drop rate versus confidence measure are shown for CRLMM and BRLMM. The curves are further stratified between heterozygotes and homozygotes as determined by the HapMap gold standard. A) 269 Affymetrix high-quality XBA arrays. B) 269 Affymetrix high-quality HIND arrays. C) 270 Affymetrix high-quality STY arrays. D) 152 Broad Institute first-pass STY arrays. E) 95 Affymetrix first-pass STY arrays. F) 270 Affymetrix high-quality NSP arrays. G) 152 Broad Institute first-pass NSP arrays. H) 95 Affymetrix first-pass NSP arrays.

Supplemental Figure 4: Validity of drop rates using called data. Drop rates stratified by homozygote and heterozygote calls are plotted against the non-stratified drop rates. The stratification is performed with HapMap gold standard calls (lines) and algorithm

derived calls (triangles and crosses). That the curves derived from HapMap and algorithm results are by and large equivalent demonstrates call rates for homozygotes and heterozygotes from algorithms are valid. Thus in the end, by also using the confidence metric of CRLMM, one can achieve a data set in which all calls are both above a prespecified accuracy threshold and have equivalent drop rates between homozygotes and heterozygotes. A) 269 Affymetrix high-quality XBA arrays. B) 269 Affymetrix high-quality HIND arrays. C) 270 Affymetrix high-quality STY arrays. D) 152 Broad Institute first-pass STY arrays. E) 95 Affymetrix first-pass STY arrays. H) 95 Affymetrix high-quality NSP arrays. G) 152 Broad Institute first-pass NSP arrays.

Supplemental Figure 5: ADPs are plotted as in Supplemental Figure 1 except that the drop rate is forced to be the same for both heterozygous and homozygous calls. A) 269 Affymetrix high-quality XBA arrays. B) 269 Affymetrix high-quality HIND arrays. C) 270 Affymetrix high-quality STY arrays. D) 152 Broad Institute first-pass STY arrays. E) 95 Affymetrix first-pass STY arrays. F) 270 Affymetrix high-quality NSP arrays. G) 152 Broad Institute first-pass NSP arrays. H) 95 Affymetrix first-pass NSP arrays.