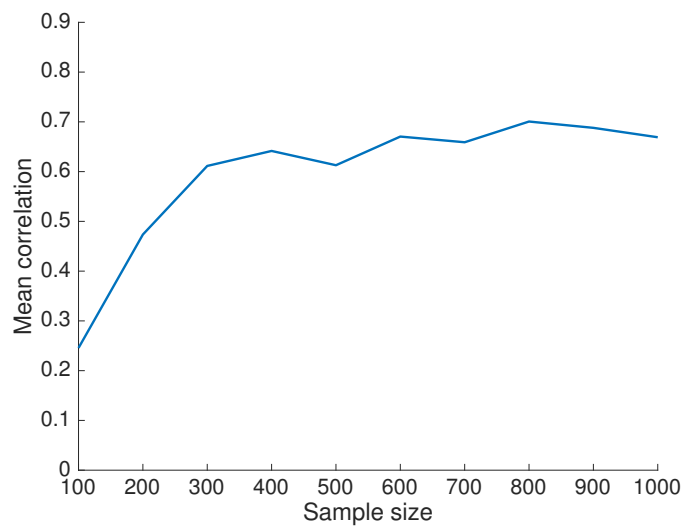**Supplementary Information For:**

A Bayesian Framework for Estimating Cell Type Composition from DNA Methylation Without the Need for Methylation Reference

Elior Rahmani, Regev Schweiger, Liat Shenhav, Eleazar Eskin and Eran Halperin.

Figure S1: The mean absolute correlation of BayesCCE's estimates with the cell types in simulated data ($k = 6$) as a function of the sample size.
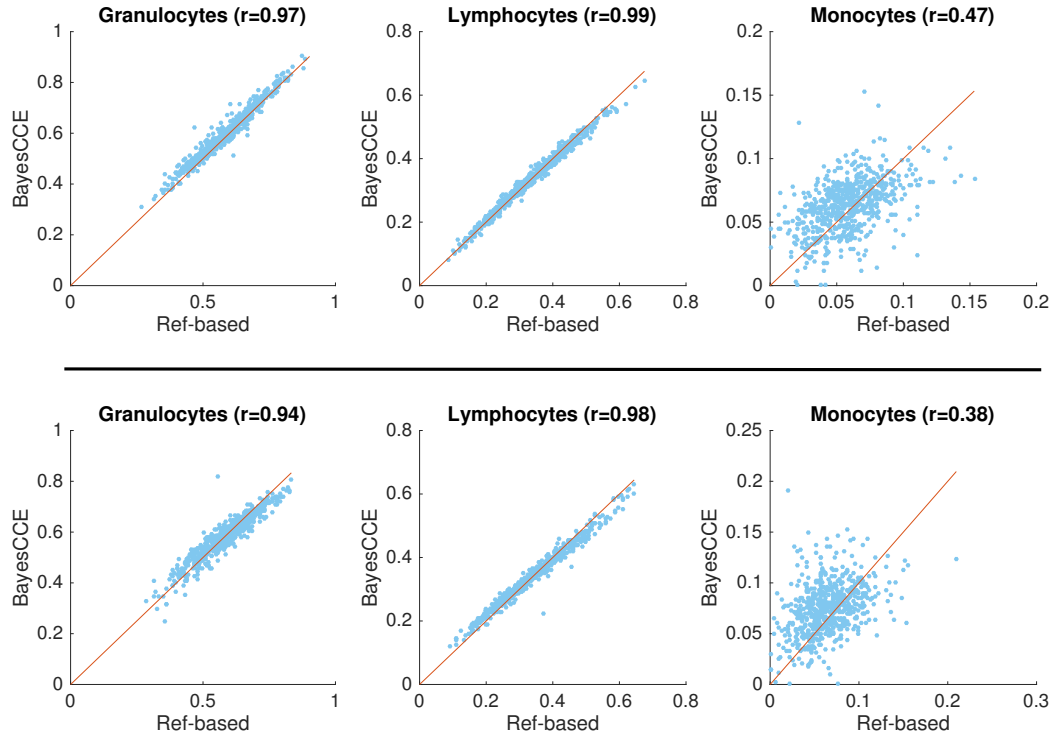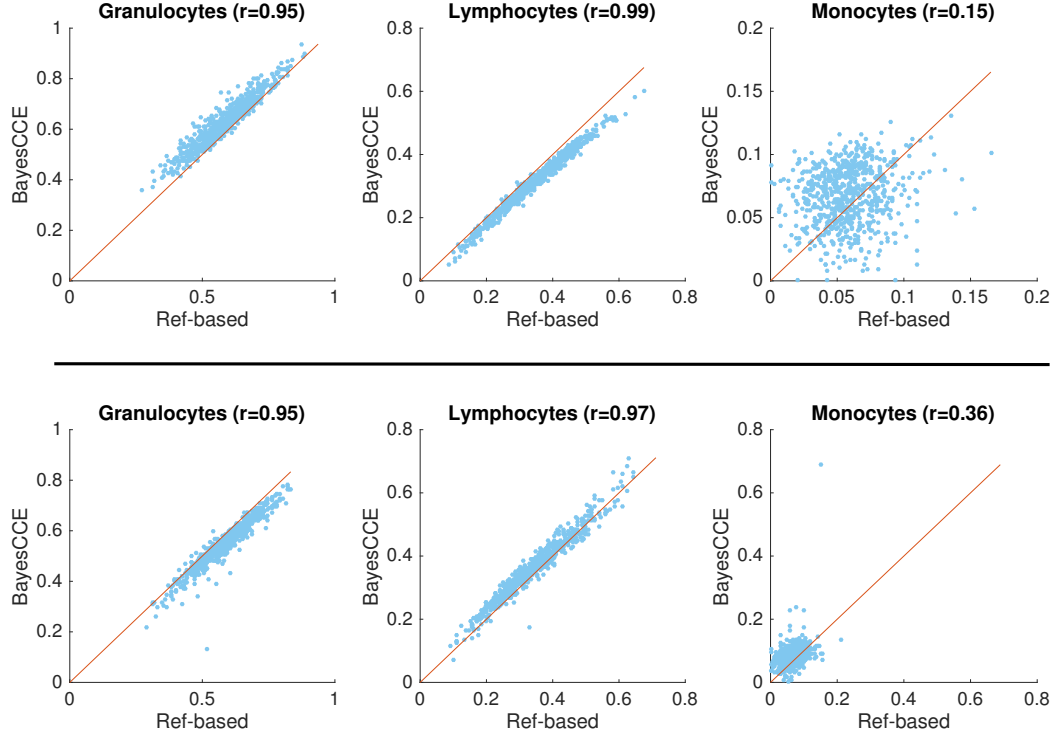
Figure S2: BayesCCE captures cell type proportions under the assumption of three cell types in the data ($k = 3$): granulocytes, lymphocytes and monocytes, and assuming known cell counts for randomly selected 5% of the samples in the data. Top panel: the results for the Liu et al. data. Bottom panel: the results for the Hannum et al. data. All correlations were calculated while excluding the samples with assumed known cell counts.
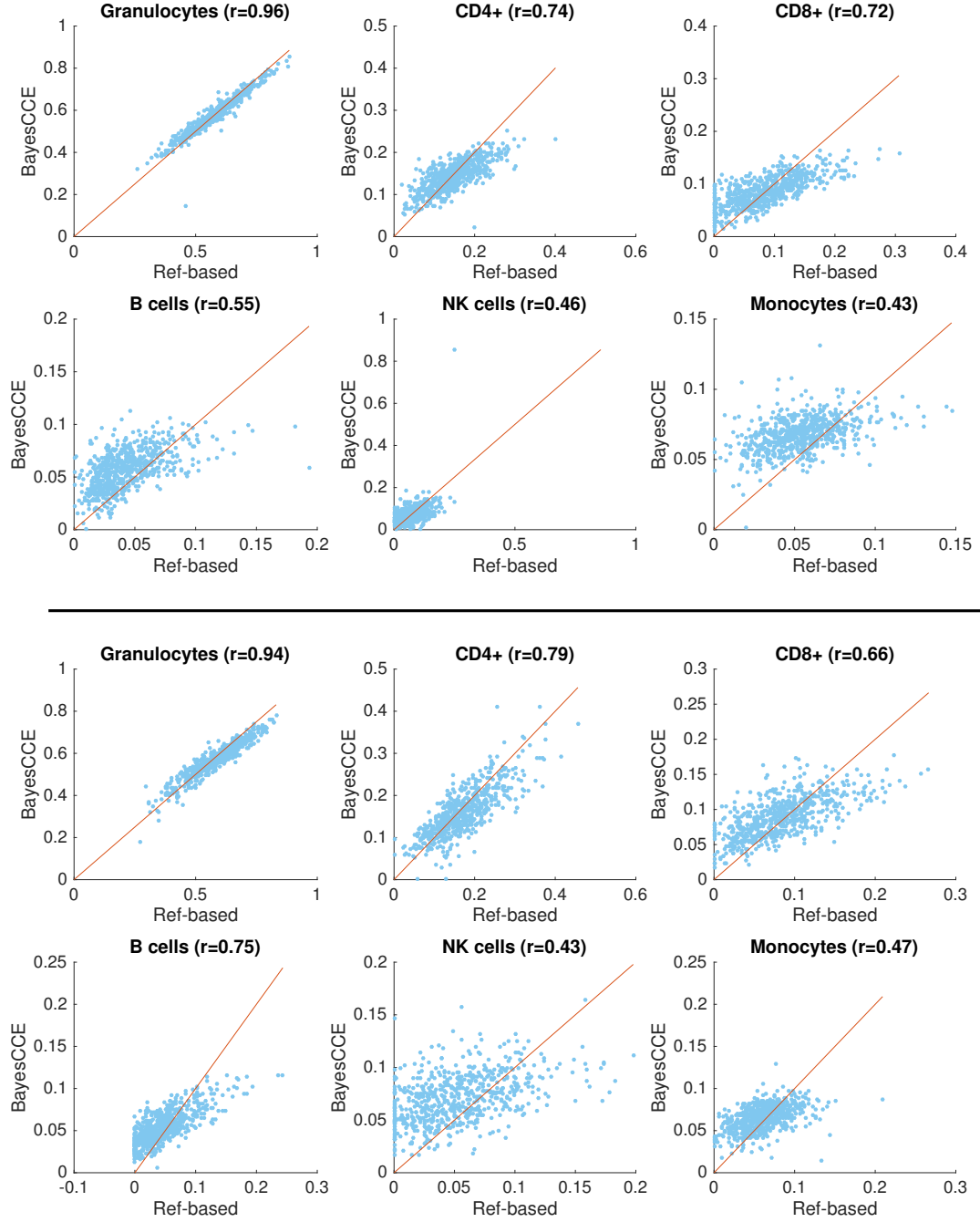
Figure S3: BayesCCE captures cell type proportions under the assumption of six cell types in the data ($k = 6$): granulocytes, four subtypes of lymphocytes (CD4+, CD8+, B cells and NK cells) and monocytes, and assuming known cell counts for randomly selected 5% of the samples in the data. Top panel: the results for the Liu et al. data. Bottom panel: the results for the Hannum et al. data. All correlations were calculated while excluding the samples with assumed known cell counts.
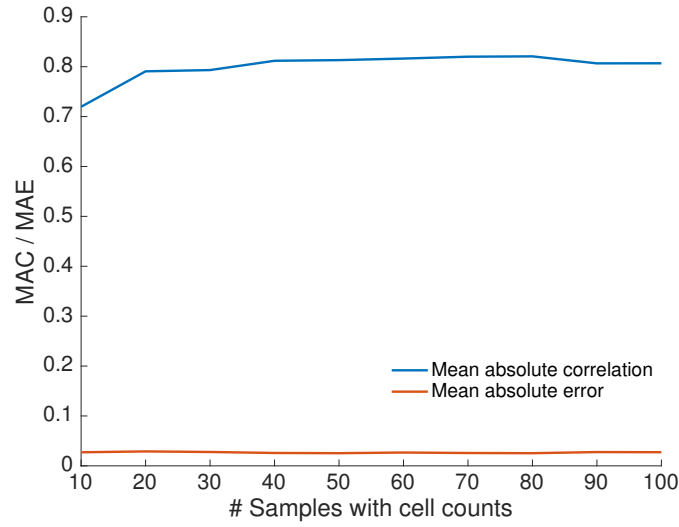
Figure S4: The mean absolute correlation (MAC) and mean absolute error (MAE) of BayesCCE's estimates with the cell type proportions in simulated data ($k = 6, n = 650$) as a function of the number of samples with known cell counts. The MAC and MAE values were calculated while excluding the samples with assumed known cell counts.

Figure S5: BayesCCE captures cell type proportions under the assumption of three cell types in the data ($k = 3$): granulocytes, lymphocytes and monocytes, and using cell counts and methylation levels of a group of samples from external data. Top panel: the results for the Liu et al. data, using cell counts and methylation for a randomly selected 5% of the samples in the Hannum et al. data. Bottom panel: the results for the Hannum et al. data, using cell counts and methylation for a randomly selected 5% of the samples in the Liu et al. All correlations were calculated while excluding the samples with assumed known cell counts.
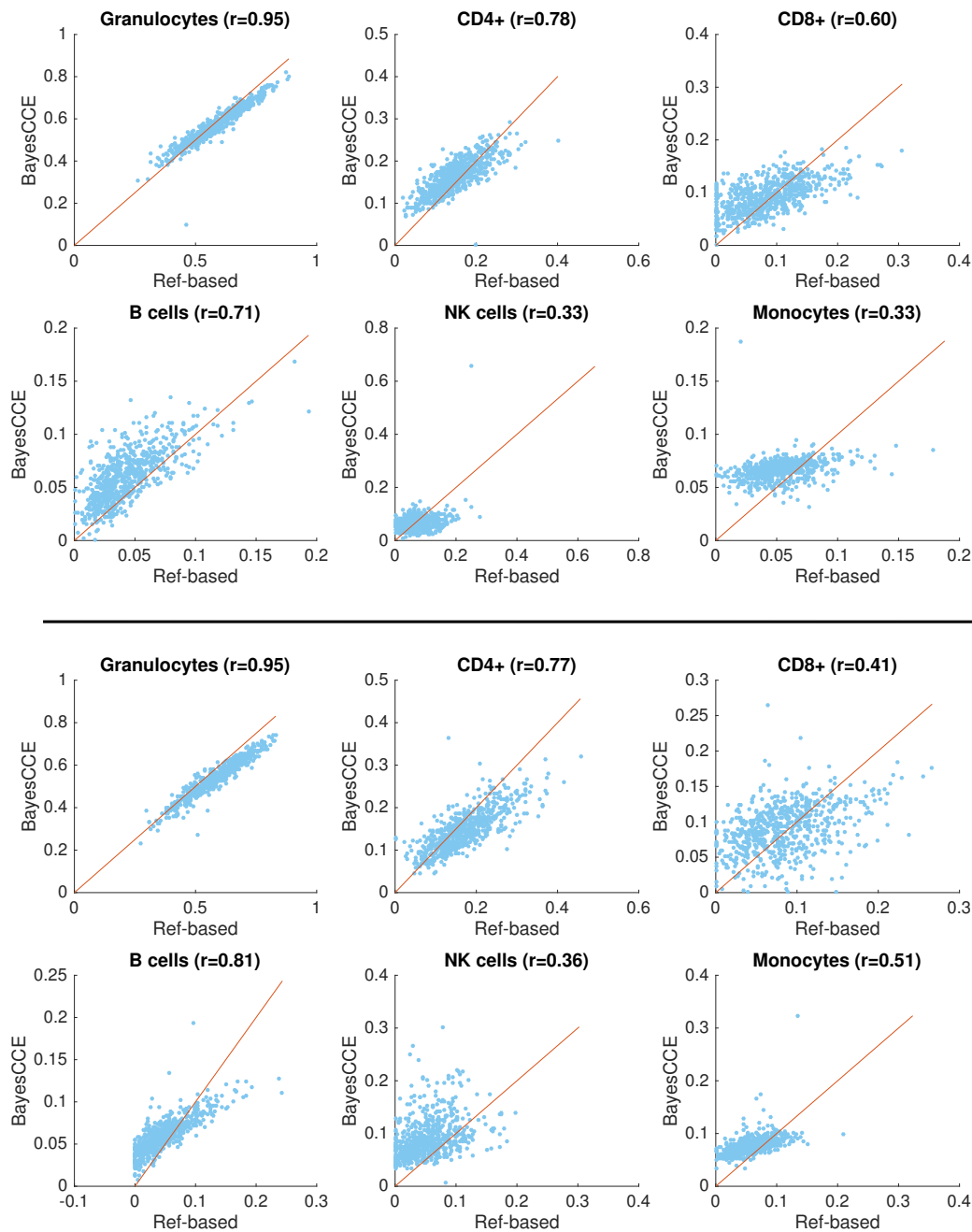
Figure S6: BayesCCE captures cell type proportions under the assumption of six cell types in the data ($k = 6$): granulocytes, four subtypes of lymphocytes (CD4+, CD8+, B cells and NK cells) and monocytes, and using cell counts and methylation levels of a group of samples from external data. Top panel: the results for the Liu et al. data, using cell counts and methylation for a randomly selected 5% of the samples in the Hannum et al. data. Bottom panel: the results for the Hannum et al. data, using cell counts and methylation for a randomly selected 5% of the samples in the Liu et al. All correlations were calculated while excluding the samples with assumed known cell counts.