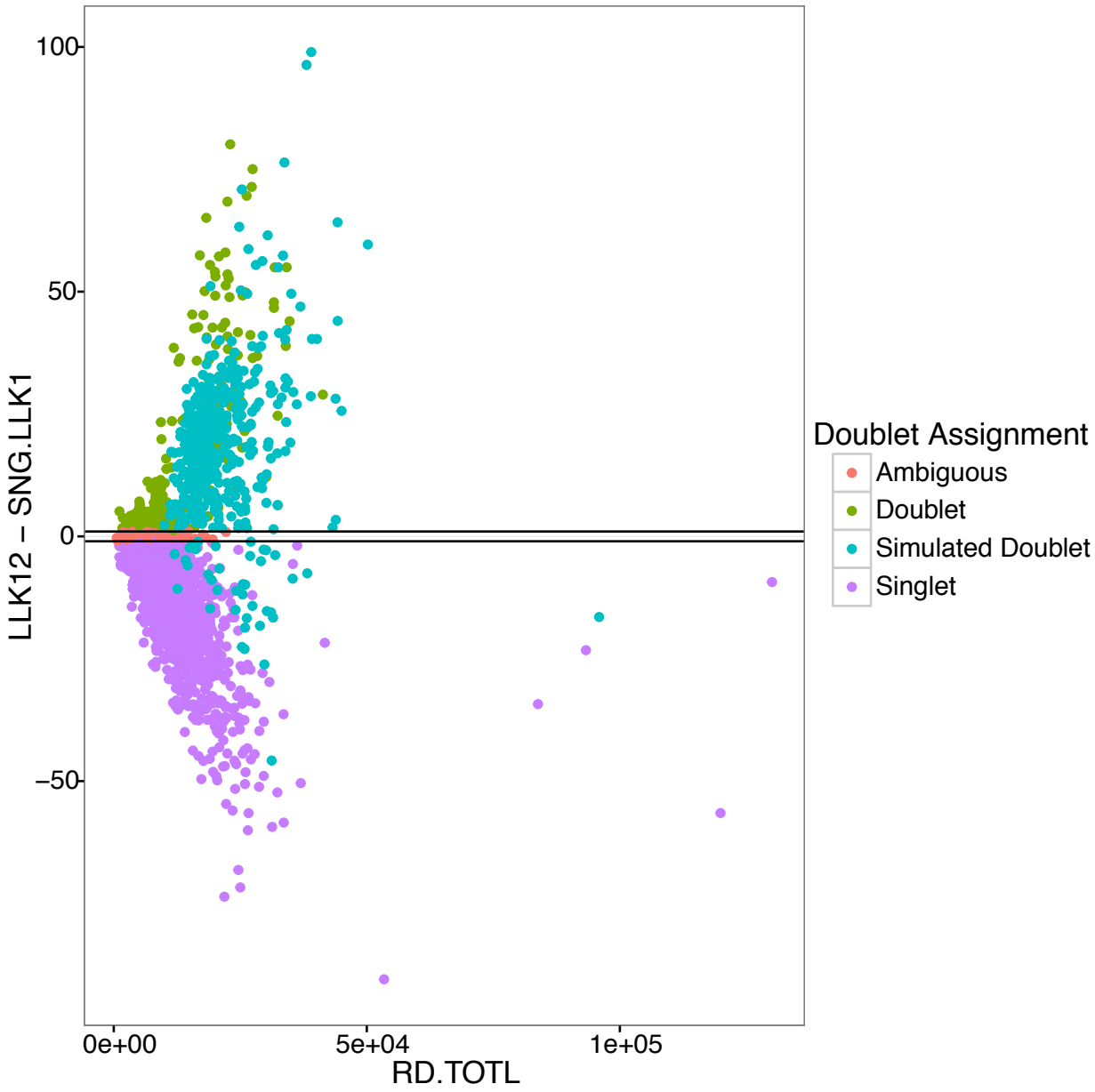
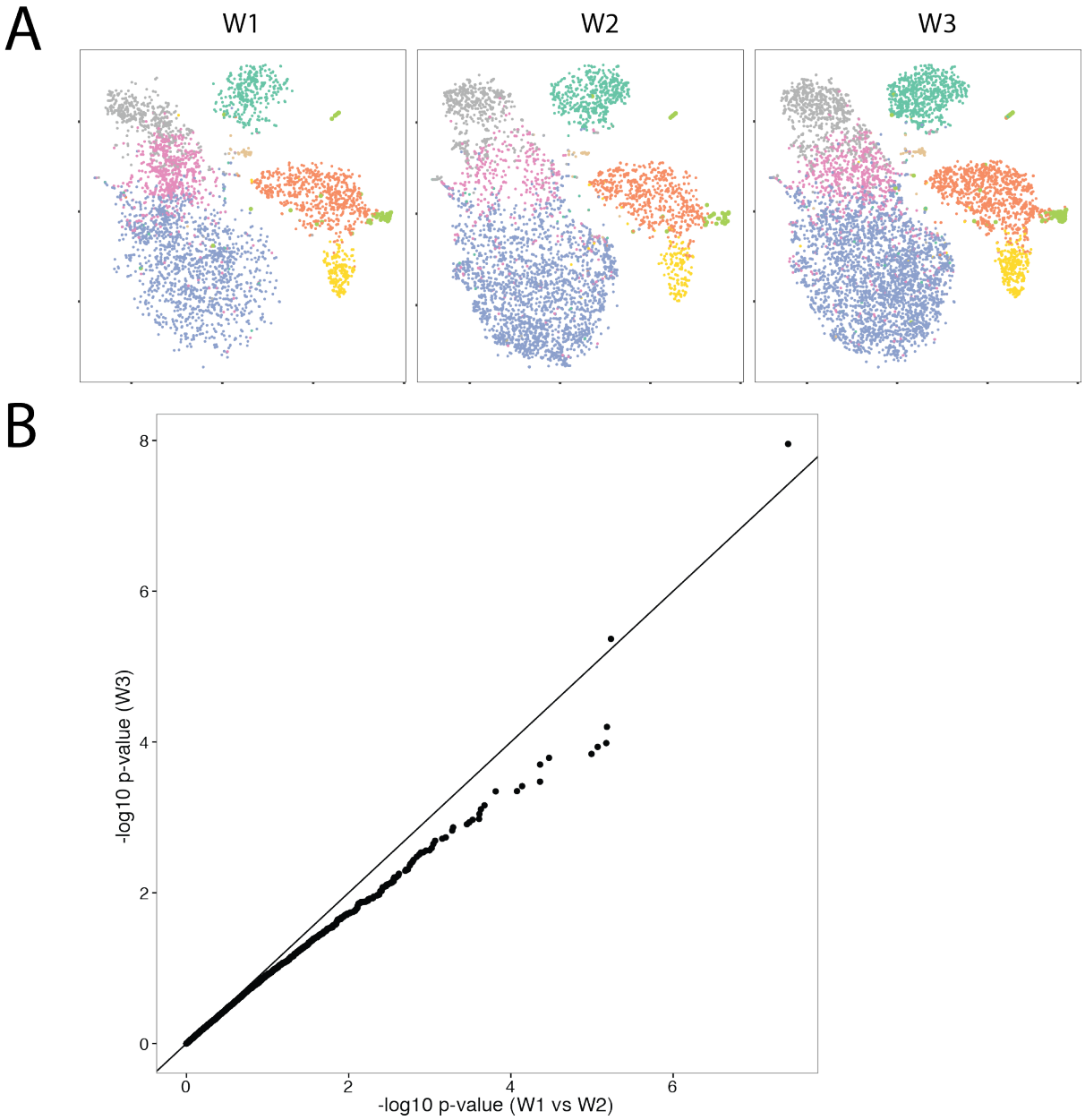
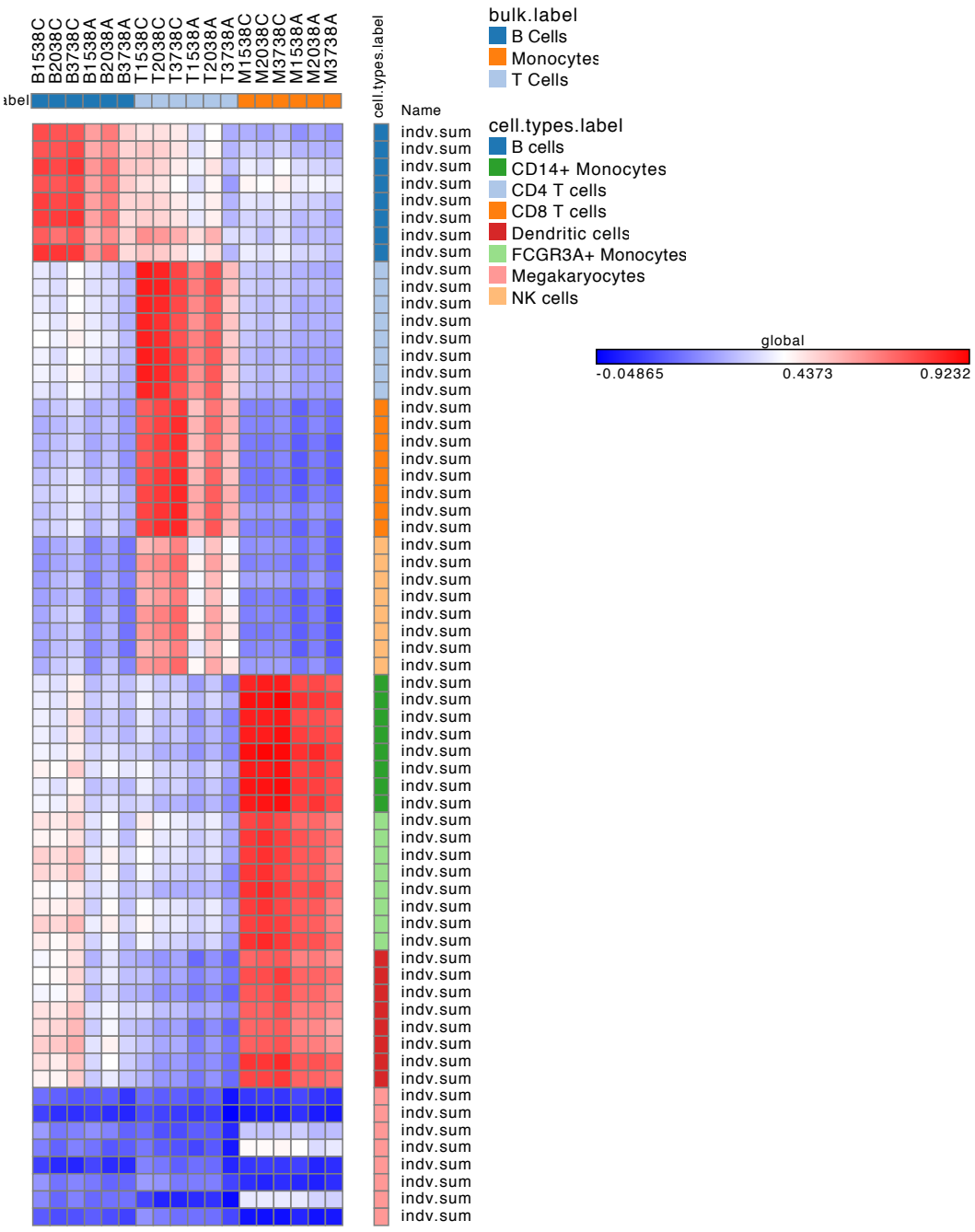
Supplementary Information



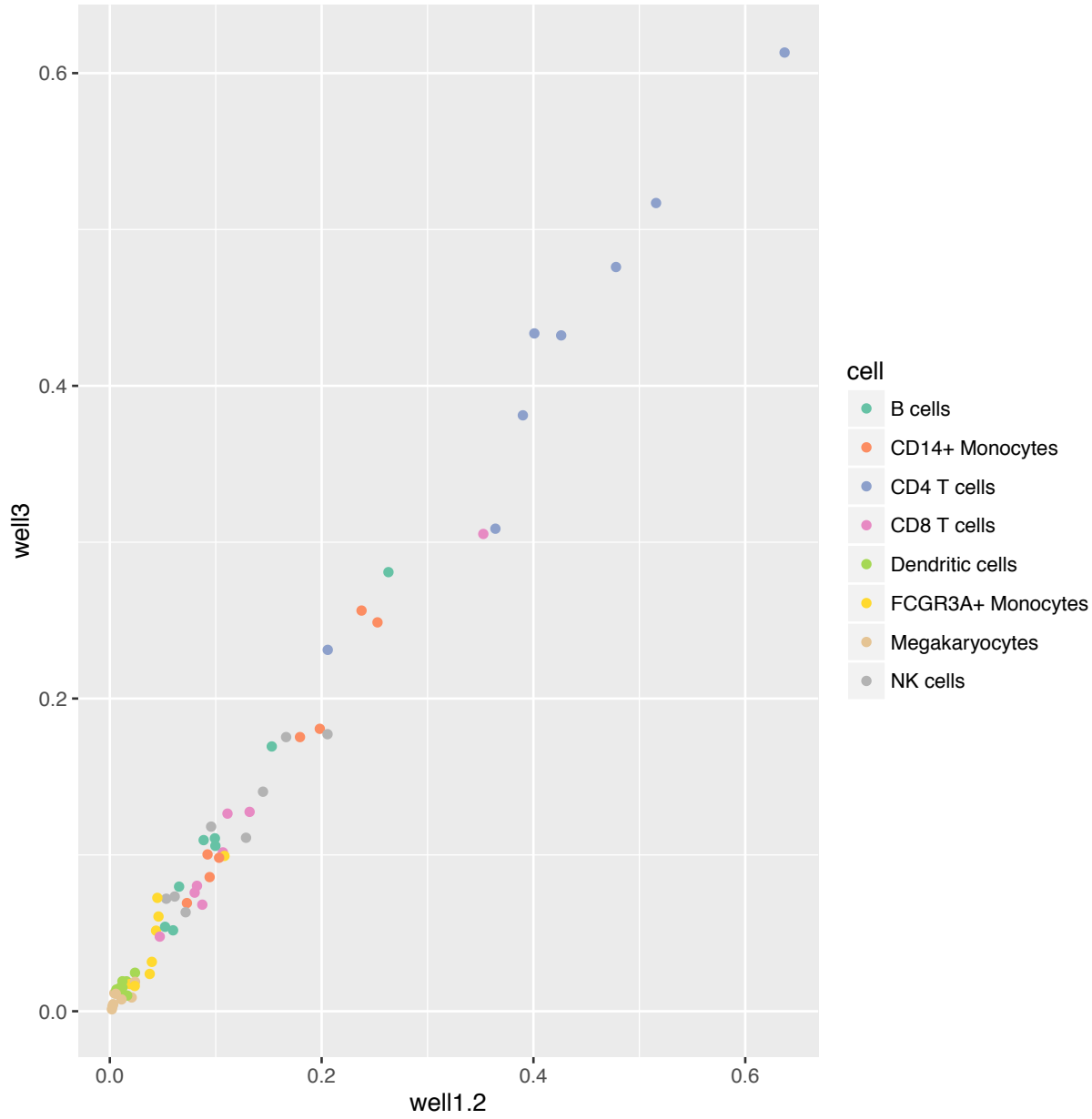
Supplementary Figure 1 – Detection boundaries of likelihood differences between mixture model for doublets (LLK12) and singlets (SNG.LLK1).



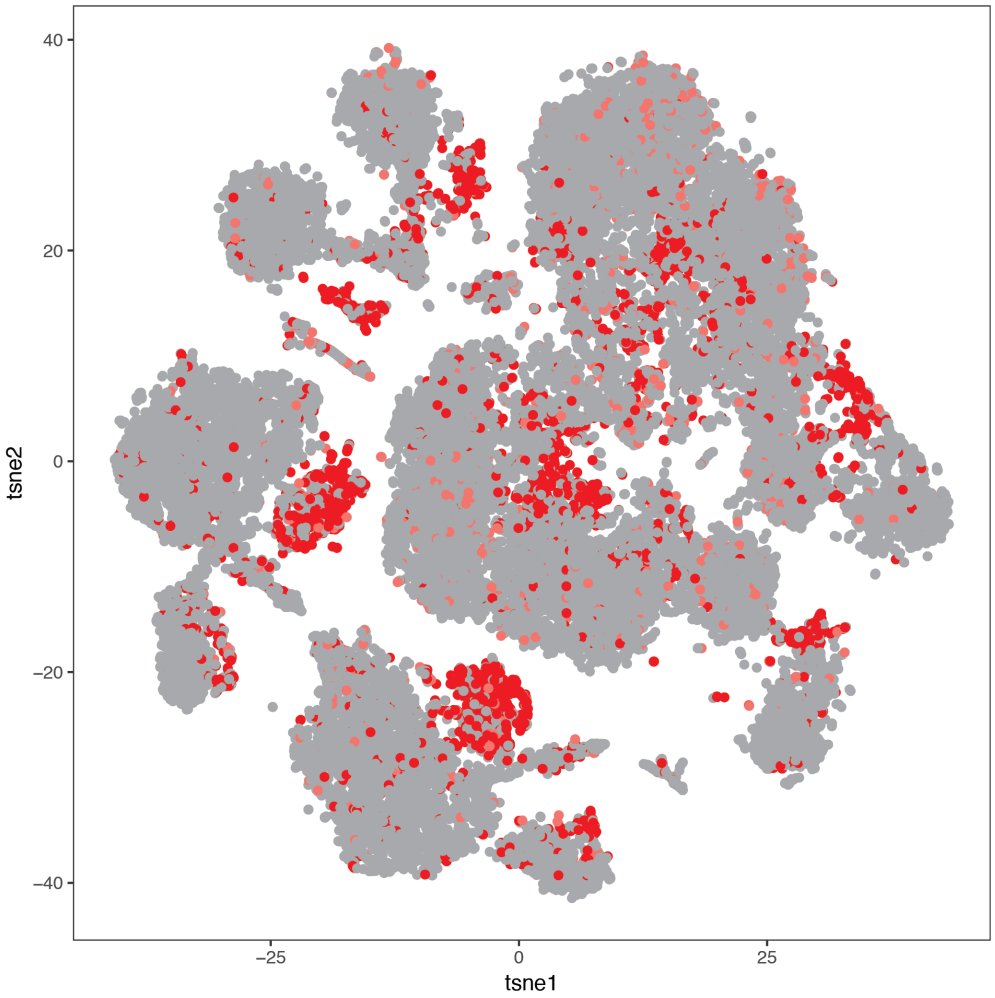
Supplementary Figure 2 – A. TSNE plot for wells W1, W2 and W3 separately. B. QQ plot comparing differential expression analysis between W1 and W2 with W3 for the same individuals.



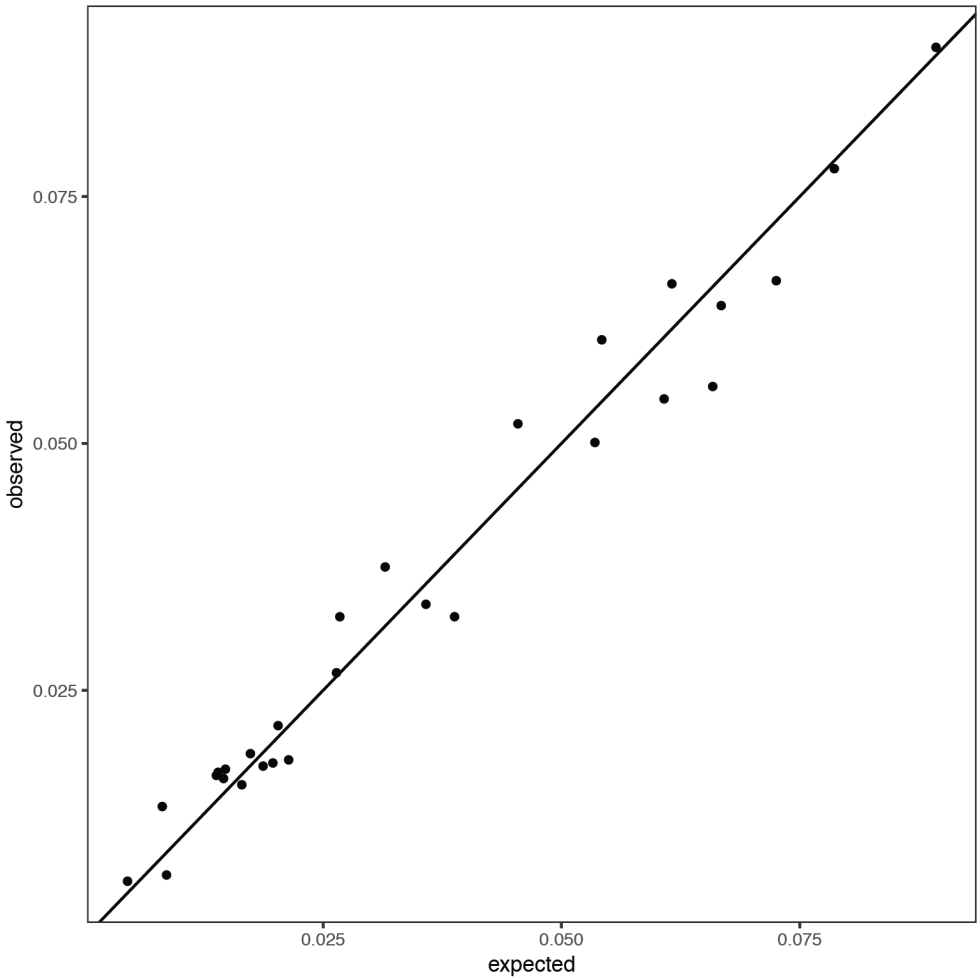
Supplementary Figure 3 - Cell-type specific expression of single cells versus bulk. Average expression over single cells for each cell type (y-axis) versus the bulk expression of B cells, CD4+ T cells and monocytes (x-axis). Single cell averages were estimated for 8 individuals. Bulk expression was generated for 6 individuals.



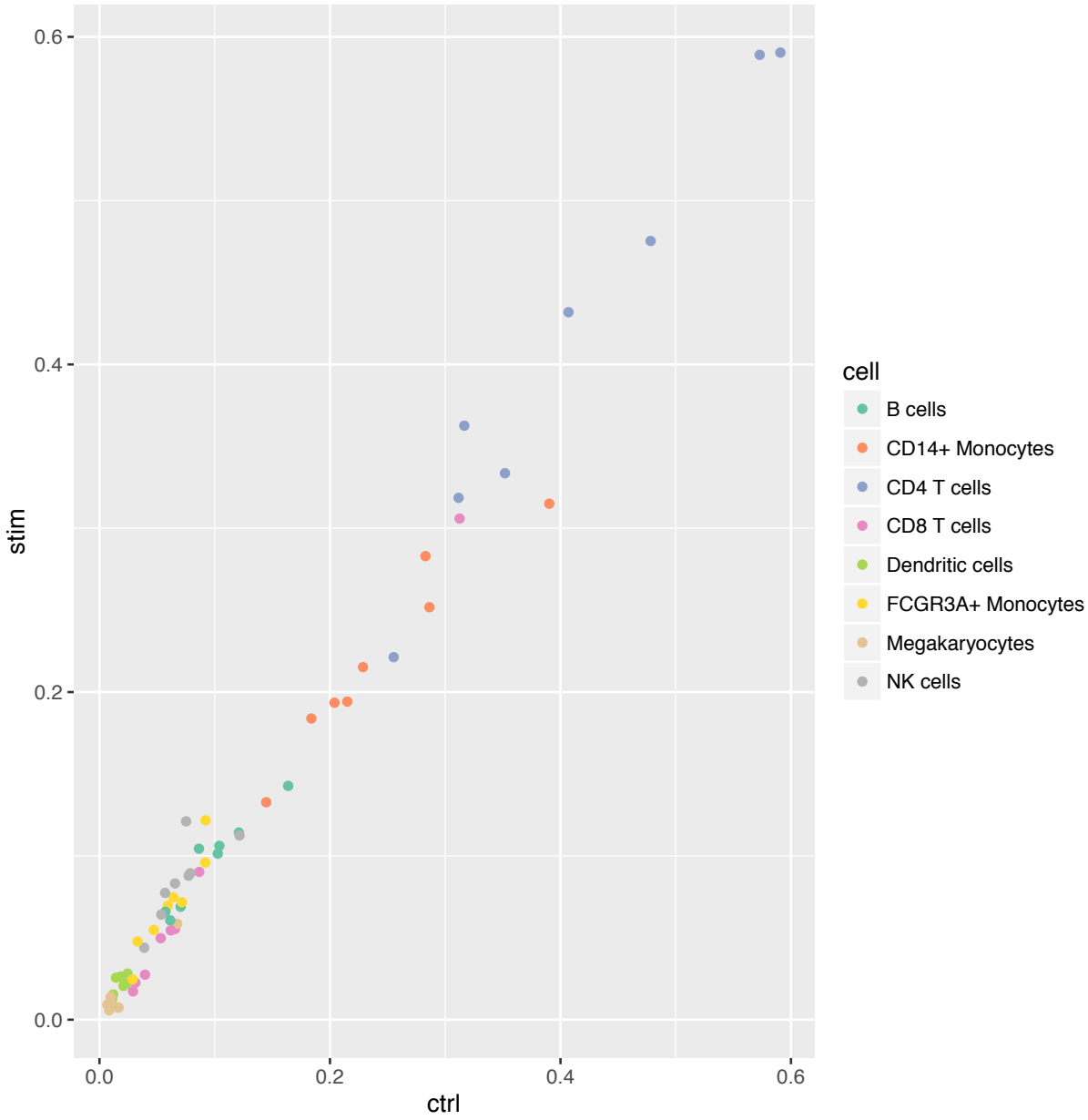
Supplementary Figure 4 – Estimates of cell type proportion from three well experiment of unstimulated PBMCs. Estimates of cell type proportion for wells W1 and W2 (x-axis) versus estimates of cell type proportion for the same individuals in well W3 (y-axis).



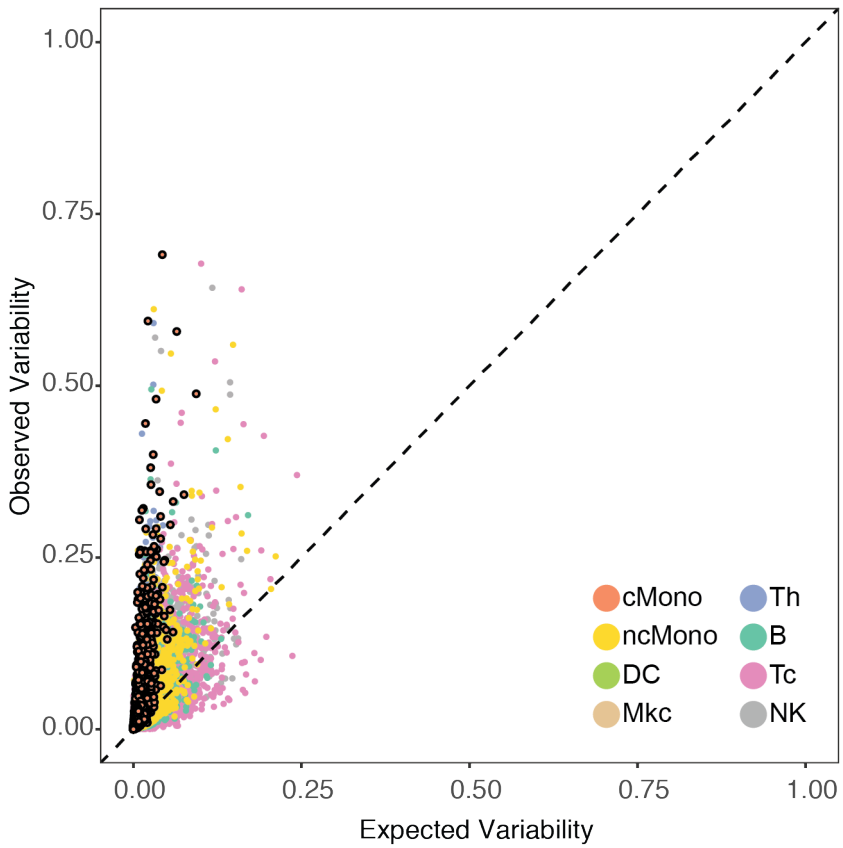
Supplementary Figure 5 – Detected doublets from the IFN-beta stimulation experiment. Red (confidently called doublets). Pink (ambiguous). Gray (singlets).



Supplementary Figure 6 – Estimated versus observed doublet rates for unstimulated cells. Expected doublet rates for each pair of individuals (i,j) are calculated as pi * pj, where pi and pj are the proportion of singlets individuals i and j.



Supplementary Figure 7 - Estimates of cell type proportion from unstimulated and IFN-beta stimulated PBMCs. Estimates of cell type proportion for unstimulated cells (x-axis) versus stimulated cells (y-axis) for 8 individuals.



Supplementary Figure 8 – Expected versus observed variability in cell populations. Variance of average expression over single cells across 8 individuals (y) versus variance of average expression over single cells across 8 synthetic replicates matching cell number for each individual (x).

Supplementary Table 1. Single cell differential expression analysis (see Excel attachment)

Supplementary Table 2. Pathway analysis of differentially expressed genes (see Excel attachment)

Supplementary Table 3. ***Conditional probability $P(b_{ij}|e_{ij}g_i)$ of read b_{ij} given true genotype g_i, and read error e_{ij}***

True Genotype g_i	**Base Calling Error Event e_{ij}**	**Pr(b_{ij} =A)**	**Pr(b_{ij} =B)**	**Pr(b_{ij} =E)[b]**
g_i = AA[a]	e_{ij} = 0	1	0	0
	e_{ij} = 1	0	1/3	2/3
g_i = AB[a]	e_{ij} = 0	1/2	1/2	0
	e_{ij} = 1	1/6	1/6	2/3
g_i = BB[a]	e_{ij} = 0	0	1	0
	e_{ij} = 1	1/3	0	2/3

[a]AA, AB, BB: A allele homozygote, heterozygote, and B allele homozygote
[b]E: alleles other than A or B; assumes four possible alleles (bases)

(from Jun, G. et al. Detecting and Estimating Contamination of Human DNA Samples in Sequencing and Array-Based Genotype Data. *The American Journal of Human Genetics*, Vol. 91 839-848 (2012).)