We conducted the analysis using different configurations of hyper-parameter values, with a range where the prior expectations of $\theta$ and $\delta_i$'s vary from 0.4 to 2.5 and the corresponding prior variances vary from 0.4 to 6.25. For each set of hyper-parameters values, we ran $N = 20,000$ iterations of the MCMC algorithm, storing draws every 10 iterations and ignoring an initial burn-in period of 10,000 iterations (hence, results are shown in terms of 1,000 iterations). Changes in hyper-parameter values (within the employed range) produced negligible differences in posterior inference (see Fig S1.). Therefore, we show numerical results on the basis of a single configuration of hyper-parameter values (all equal to 1). Trace-plots and auto-correlation plots (some of which are shown in Fig S2.) suggest a good mixing of the chains. Regarding the adaptive proposals, we stopped the adaptation after 10,000 iterations, hence the results presented here are based on constant proposal variances.
Figure S1: **Sensitivity to prior hyper parameters.** For the mouse ES dataset analysed on the main paper. Vertical lines represent the 95% high posterior density interval (dot located at the posterior median) for 4 different combinations of hyper parameter values (within the range described in this supplementary text). First and second rows: $\mu_i$ and $\delta_i$ for 3 randomly selected genes, respectively. Third and fourth rows: $\phi_j$ and $s_j$ for 3 randomly selected cells, respectively. Fifth row: $\theta$. All these hyper parameters configurations produced virtually the same posterior inference for all model parameters.
Figure S2: For the mouse ES dataset analysed on the main paper. Trace-plots (left) and auto-correlation plots (right) for $\mu_1$, $\nu_2$, $\theta$, $\delta_1$, $\kappa_2$, $\phi_2$ and $s_2$. Trace-plots suggest convergence of the chains. The chains of the $\kappa_j$’s mix less well, however mixing is substantially improved when presenting the results in terms of the $\phi_j$’s.