Processing step	Recommendation	Comments
(I) Gene information re-annotation	Re-annotator package or BioMart online tool.	Re-annotator allows updating the information for all probes (standard Agilent and custom) but requires specific software; BioMart is simple to use and does not require additional software, however is limited to standard Agilent probes.
(II) Data filtering	Exclude probes that did not exceed the background in at least *k*% of samples.	Increases data reliability, however the threshold (***k*%)** is still an arbitrary choice.
(III) Probe selection	RNA-seq: filter genes based on correlation to RNA-seq expression measures; select the probe with highest correlation to RNA-seq gene expression where data are available; Alternatively, selecting probes with the highest DS can be generalised for the full gene set of AHBA.	RNA-seq provides an external validation to the microarray data, after removing probes with low correlation (e.g., $\rho < 0.2$). RNA-seq filtering is limited to genes that are present in both datasets, but DS probe selection can be used for remaining genes given its high correlation with RNA-seq based selection.
(IV) Sample assignment	Generate donor-specific parcellations; assign samples separately for left/right, cortex/subcortex; apply a distance threshold.	Donor-specific parcellation more accurately represents individual anatomical variation and leads to more accurate sample assignment. Subject-specific parcellations require manual inspection. A distance threshold of 2mm strikes a balance between sample retention and close anatomical correspondence.
(V) Data normalisation	Evaluate relative expression levels for each gene within each sample; evaluate relative expression levels of each gene within each brain by applying SRS normalization.	SRS normalisation is robust to outliers and is directly comparable between donors.
(VI) Gene filtering	Differential stability and/or hypothesis-based gene selection	In the absence of a specific hypothesis, selection of genes based on DS reduces donor-specific variance and focuses on brain-relevant genes.
(VII) Accounting for spatial effects	Depending on research question, regression, mean subtraction for each class of region pairs, or spatially constrained randomisation can be implemented.	Regression and mean subtraction are simple but apply only to analyses of inter-regional transcriptional coupling (CGE). Spatially constrained null models can be tailored for specific research questions, including analyses of regional expression. For whole-brain analyses, consider normalising cortical and subcortical measures separately prior to distance correction.