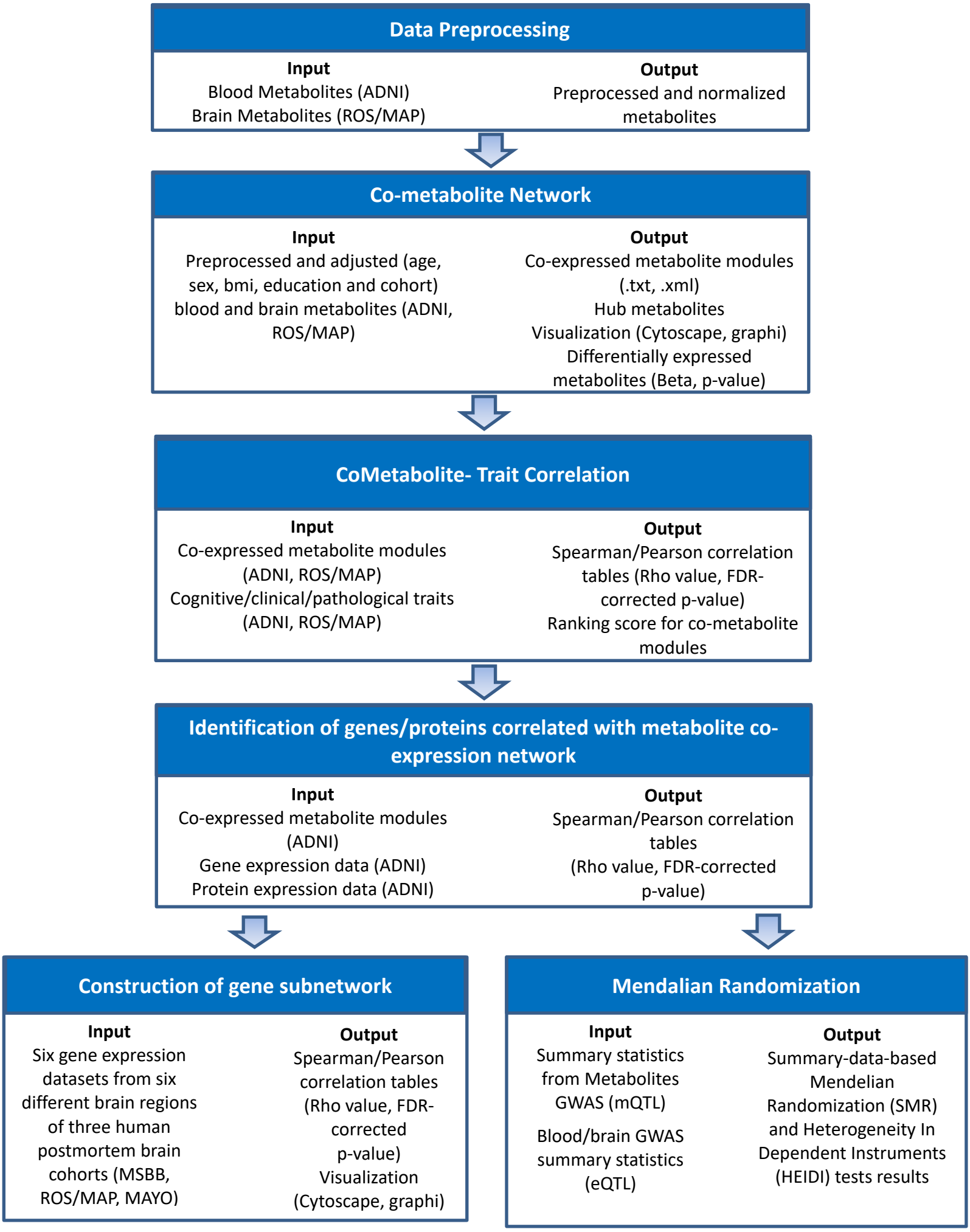
Data Preprocessing
Input
Blood Metabolites (ADNI)
Brain Metabolites (ROS/MAP)
Output
Preprocessed and normalized metabolites
Co-metabolite Network
Input
Preprocessed and adjusted (age, sex, bmi, education and cohort) blood and brain metabolites (ADNI, ROS/MAP)
Output
Co-expressed metabolite modules (.txt, .xml)
Hub metabolites
Visualization (Cytoscape, graphi)
Differentially expressed metabolites (Beta, p-value)
CoMetabolite- Trait Correlation
Input
Co-expressed metabolite modules (ADNI, ROS/MAP)
Cognitive/clinical/pathological traits (ADNI, ROS/MAP)
Output
Spearman/Pearson correlation tables (Rho value, FDR-corrected p-value)
Ranking score for co-metabolite modules
Identification of genes/proteins correlated with metabolite co-expression network
Input
Co-expressed metabolite modules (ADNI)
Gene expression data (ADNI)
Protein expression data (ADNI)
Output
Spearman/Pearson correlation tables
(Rho value, FDR-corrected p-value)
Construction of gene subnetwork
Input
Six gene expression datasets from six different brain regions of three human postmortem brain cohorts (MSBB, ROS/MAP, MAYO)
Output
Spearman/Pearson correlation tables (Rho value, FDR-corrected p-value)
Visualization (Cytoscape, graphi)
Mendalian Randomization
Input
Summary statistics from Metabolites GWAS (mQTL)
Blood/brain GWAS summary statistics (eQTL)
Output
Summary-data-based Mendelian Randomization (SMR) and Heterogeneity In Dependent Instruments (HEIDI) tests results

Figure S1. Experimental design and study workflow: Metabolomics data (Blood Metabolites: ADNI: CN:362, SMC:94, MCI:764, AD:298; and Brain Metabolites: ROS/MAP; CN:50, MCI:30, AD:25) were preprocessed to limit the potential for false-positive findings. To this end, missing values were imputed using minimum value imputation (half of the plate-specific limit of detection), single measurement outliers were winsorized to 3 standard deviations from the global mean, and multi-variate sample outliers were removed using Mahalanobis distance. The imputation had no significant influence on metabolite associations with AD biomarker profiles. All metabolites were adjusted for age, sex, BMI, education, and cohort. A metabolite co-expression network for the ADNI blood expression data was constructed using Multiscale Embedded Gene co-Expression Network Analysis (MEGENA). Co-expressed metabolite modules were ranked based on the overall strength of such correlations of the metabolite modules with clinical/cognitive CSF (tau and Aβ) and imaging (FDG and AV45 PET) biomarker variables. Since ADNI participants have a great number of matched gene and protein expression data, correlation analysis was performed only using ADNI data to identify significant genes and proteins highly associated with co-expressed metabolites. Then, co-expression gene subnetwork was constructed using six gene expression datasets from six different brain regions of three human postmortem brain cohorts, including the Mount Sinai Brain Bank (MSBB) AD RNA-seq (4 cortex region), the Religious Order Study (ROS) and the Rush Memory and Aging Project (MAP) datasets (ROSMAP) RNA-seq (DLPFC) and MAYO Clinic. Lastly, Summary-data-based Mendelian Randomization (SMR) and Heterogeneity In Dependent Instruments (HEIDI) tests were conducted to explore likely causal paths that link gene expression to metabolite concentrations using summary statistics from our metabolites GWAS and GTEx v7 eQTL summary data, GTEx-brain eQTL summary data, Cardiogenics study, eQTLGen Consortium, and eQTL ADNI blood.