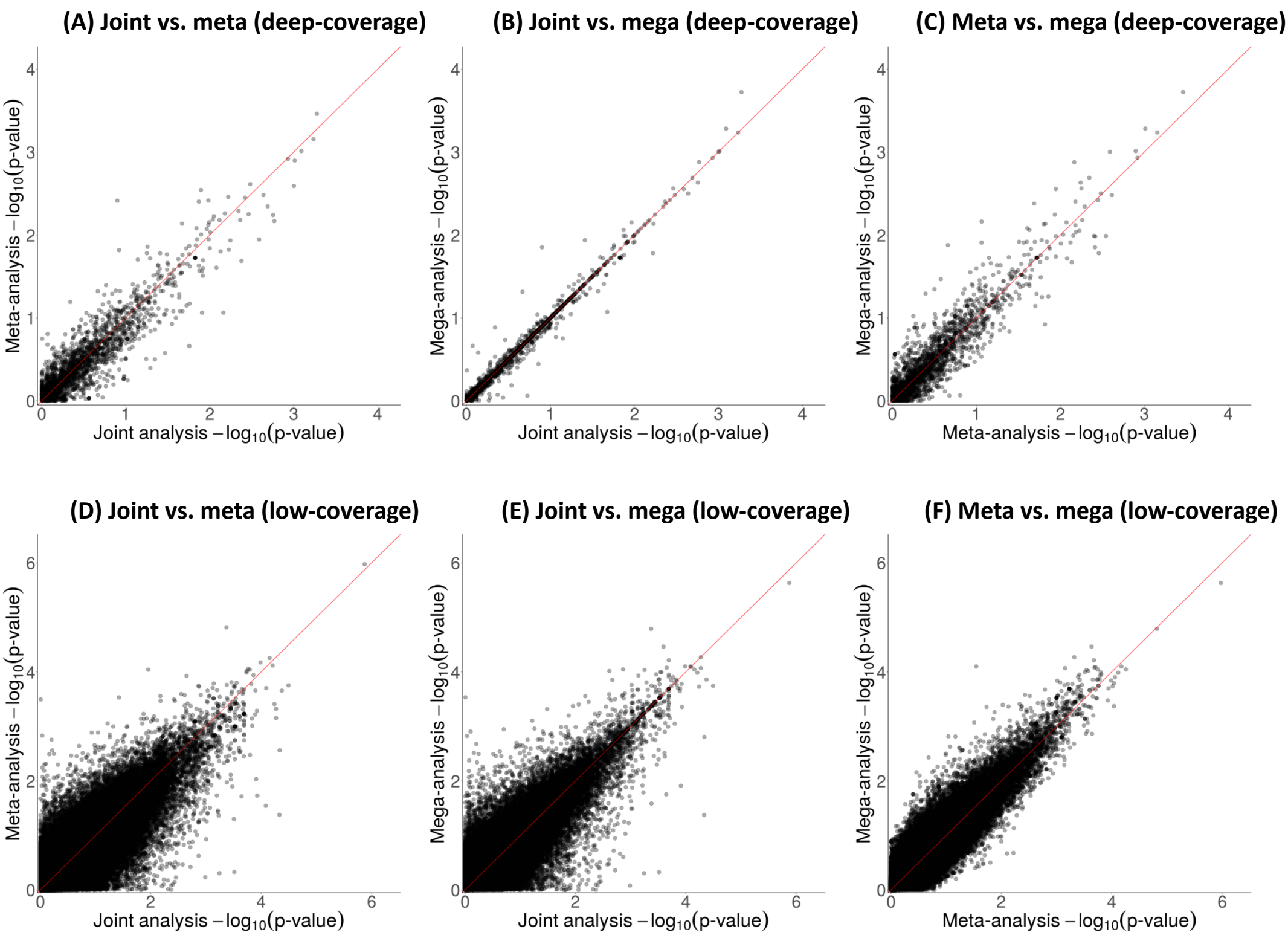
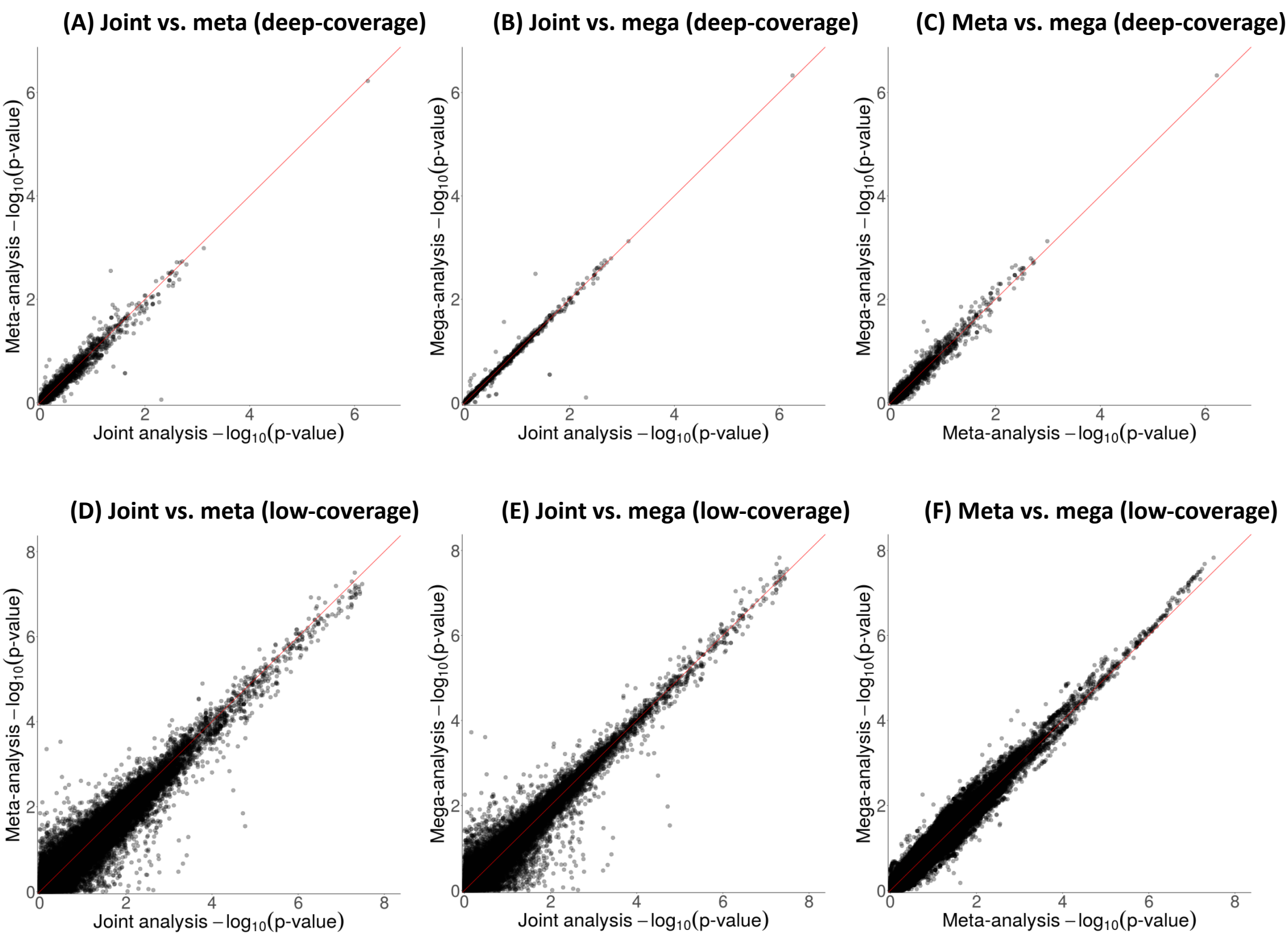
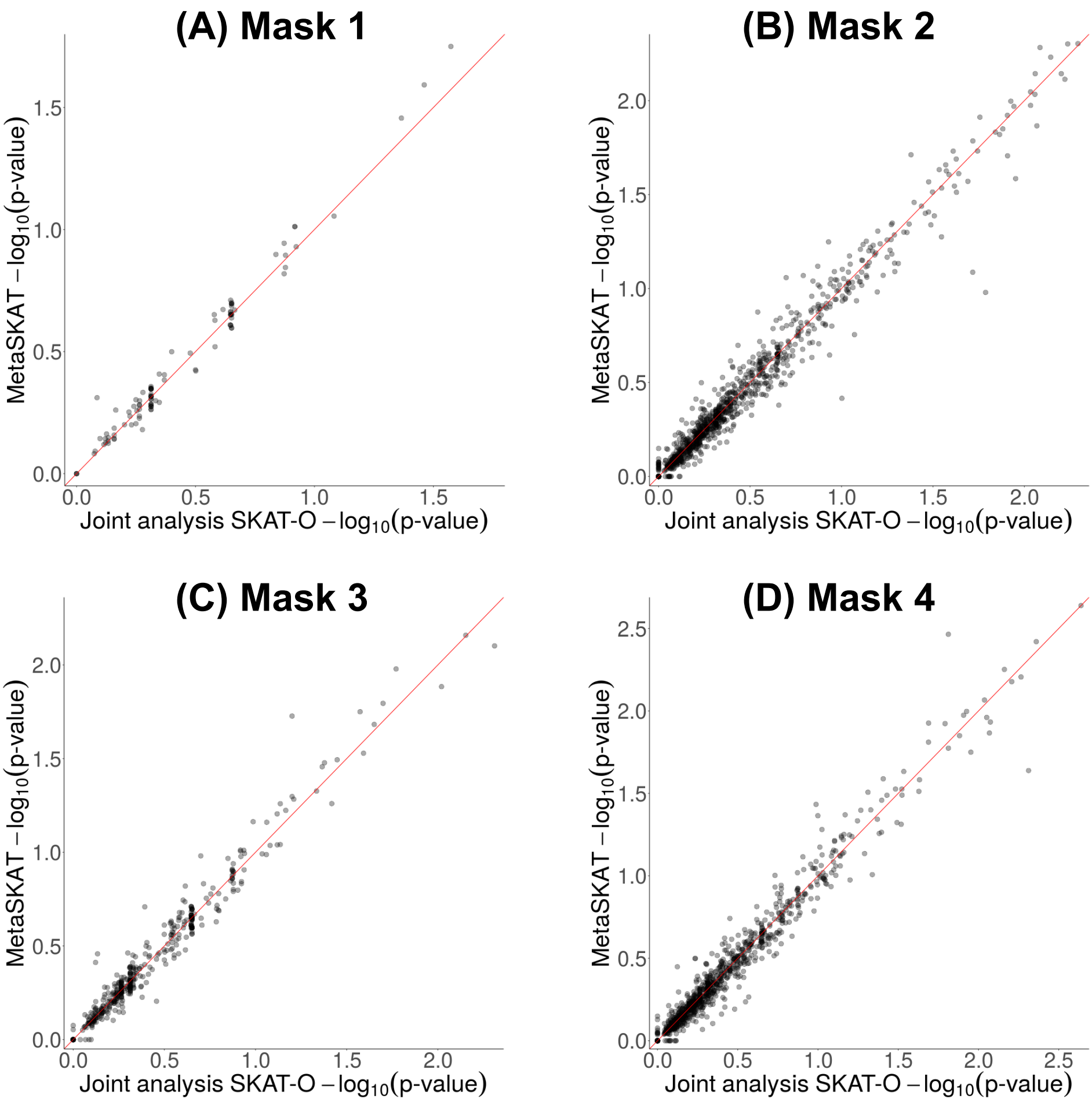


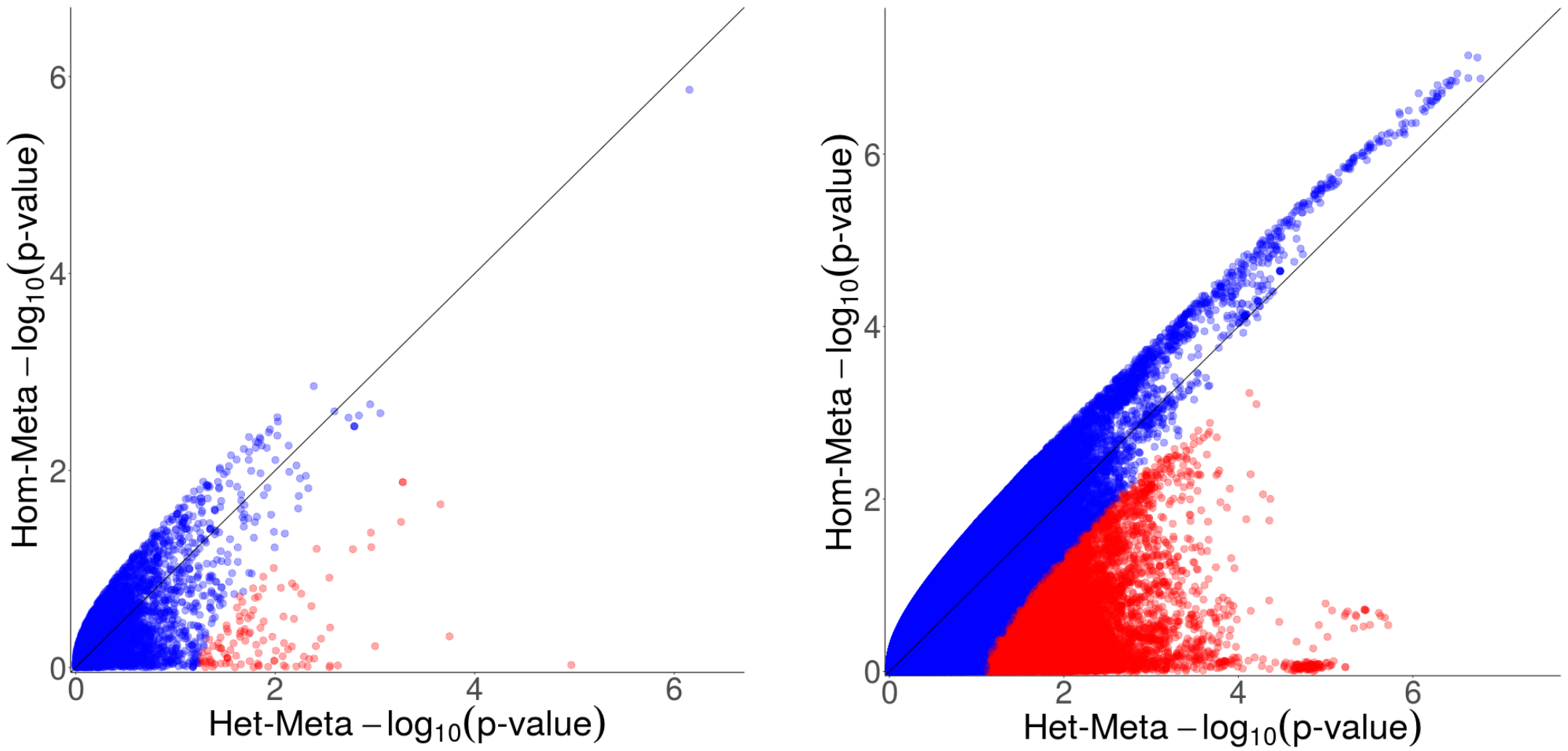
Supplementary Figure 1. Comparison of single-variant association test p-values between joint and single study calling strategies for low-frequency (MAF 0.5-5%) SNVs in (A-C) deep-coverage (~82X) exome sequence data and (D-F) low-coverage (~5X) genome sequence data. *Joint* refers to joint analysis of the joint callset, *meta* refers to fixed-effects meta-analysis of single-study summary statistics, and *mega* refers to joint analysis of the union callset (mega-analysis).



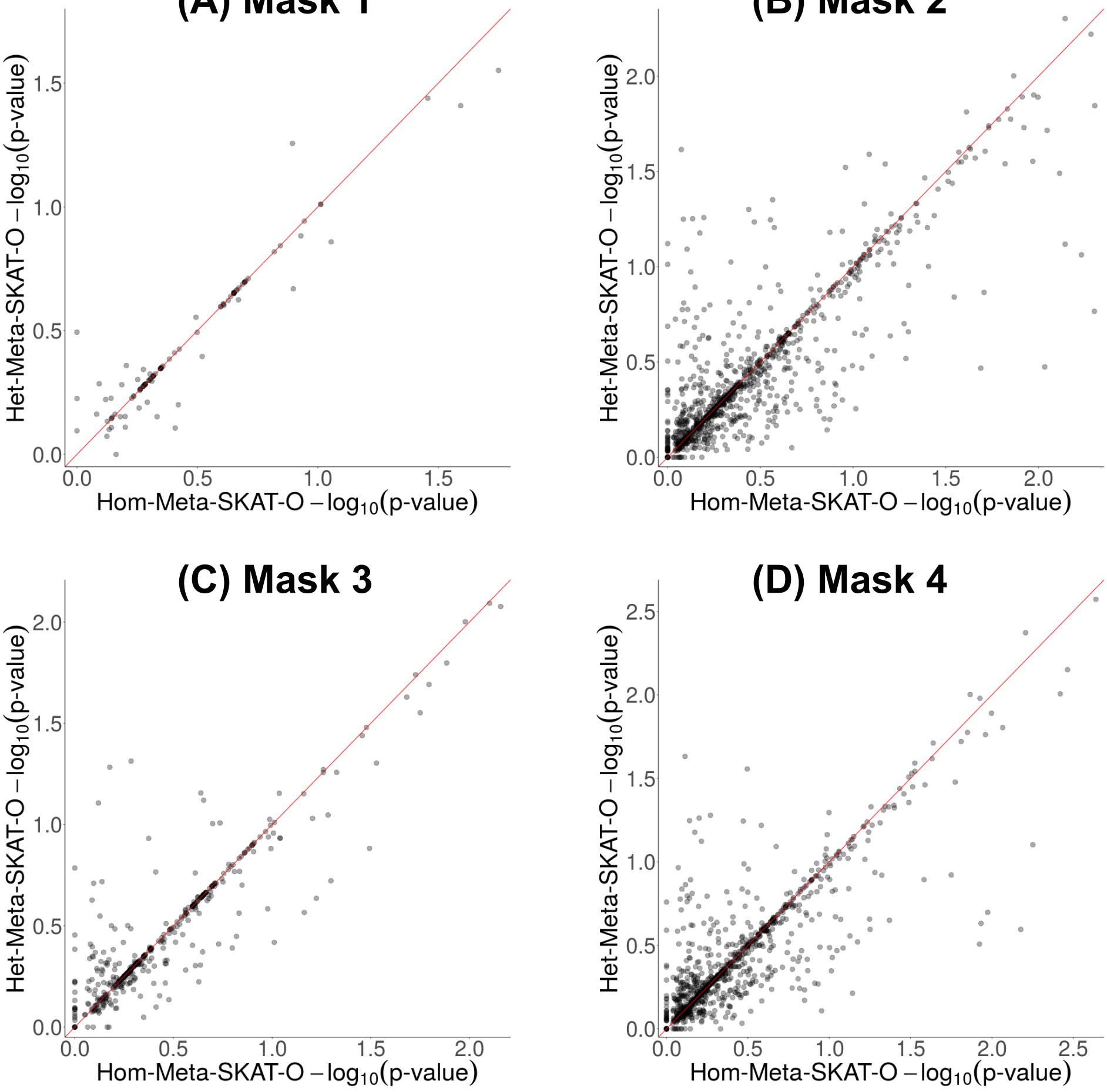
Supplementary Figure 2. Comparison of single-variant association test p-values between joint and single study calling strategies for common (MAF >5%) SNVs in (A-C) deep-coverage (~82X) exome sequence data and (D-F) low-coverage (~5X) genome sequence data. *Joint* refers to joint analysis of the joint callset, *meta* refers to fixed-effects meta-analysis of single-study summary statistics, and *mega* refers to joint analysis of the union callset (mega-analysis).



Supplementary Figure 3. Comparison of gene-based association test p-values between joint and single study calling strategies in deep-coverage (~82X) exome sequence data. *MetaSKAT* refers to homogeneous effects Meta-SKAT-O test implemented in the MetaSKAT R package. Mask 1: protein-truncating SNVs; Mask 2: Mask1+missense SNVs with MAF<1%; Mask 3: Mask1+SNVs predicted deleterious by all algorithms (Polyphen2-HumDiv, PolyPhen2-HumVar, LRT, Mutation Taster, and SIFT); Mask 4: Mask1+SNVs with MAF<1% predicted deleterious by at least one algorithm.

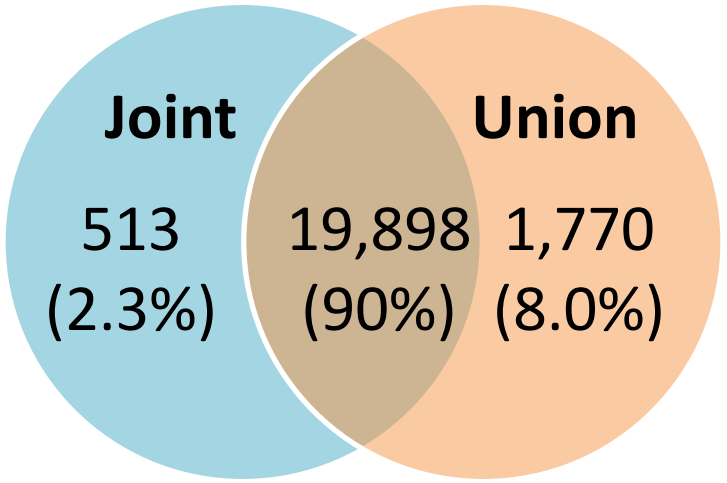


Supplementary Figure 4. Comparison of trans-ethnic meta-analysis (Het-Meta) using MR-MEGA and fixed-effects meta-analysis (Hom-Meta) using METAL for (A) deep-coverage (~82X) exome sequence data and (B) low-coverage (~5X) genome sequence data. Red points denote variants whose heterogeneity in genetic effects is correlated with ancestry (p-value<0.05) while blue points denote variants whose heterogeneity is not correlated with ancestry (p-value≥0.05).

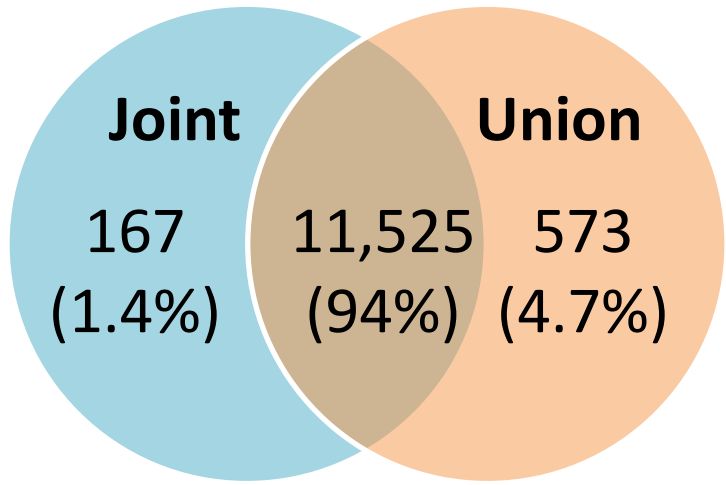


Supplementary Figure 5. Comparison between gene-based meta-analysis assuming homogeneous genetic effects between single-study cohorts (Hom-Meta-SKAT-O) and gene-based meta-analysis assuming heterogeneous genetic effects (Het-Meta-SKAT-O) in deep-coverage (~82X) exome sequence data. Mask 1: protein-truncating SNVs; Mask 2: Mask1+missense SNVs with MAF<1%; Mask 3: Mask1+SNVs predicted deleterious by all algorithms (Polyphen2-HumDiv, PolyPhen2-HumVar, LRT, Mutation Taster, and SIFT); Mask 4: Mask1+SNVs with MAF<1% predicted deleterious by at least one algorithm.

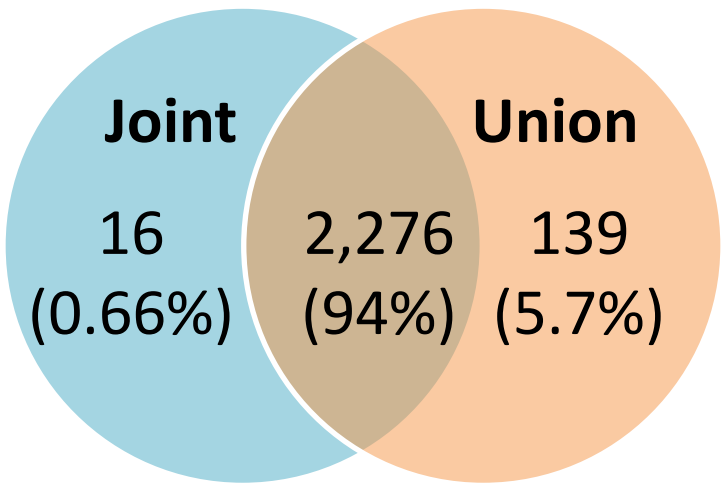
A) Deep-coverage, rare



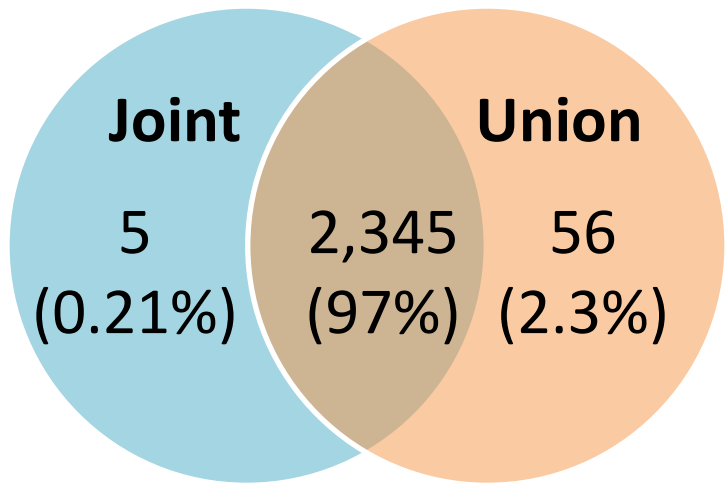
B) Low-coverage (coding), rare



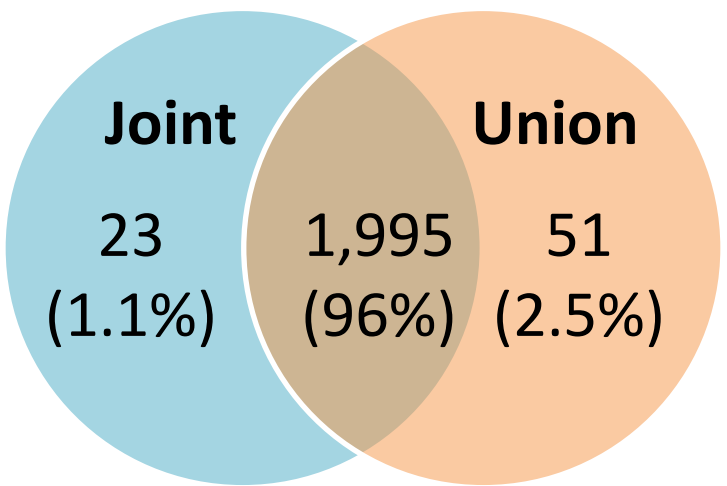
C) Deep-coverage, low-freq.



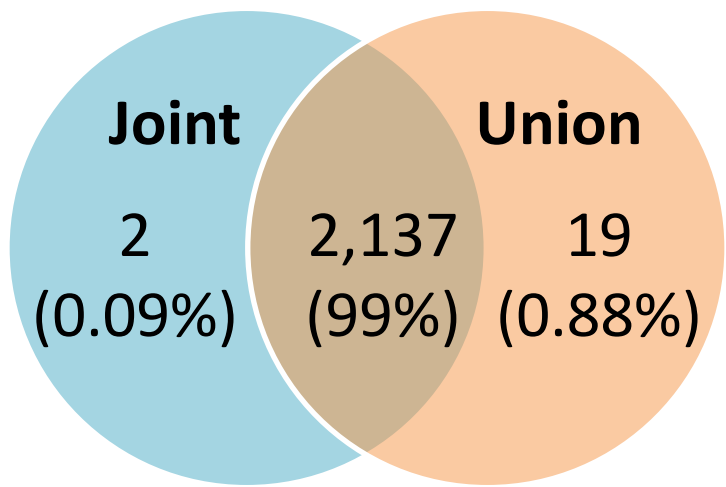
D) Low-coverage (coding), low-freq.



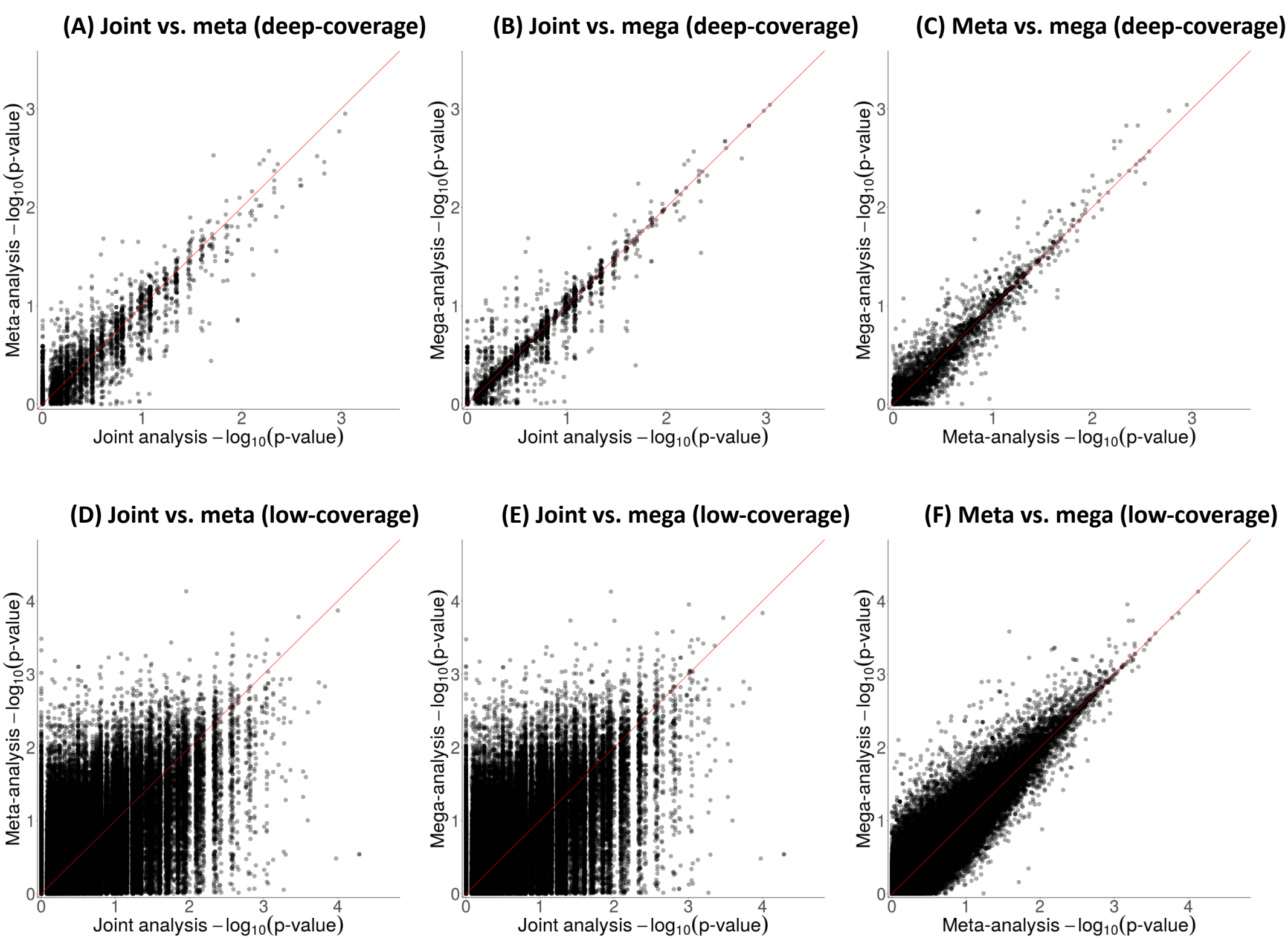
E) Deep-coverage, common



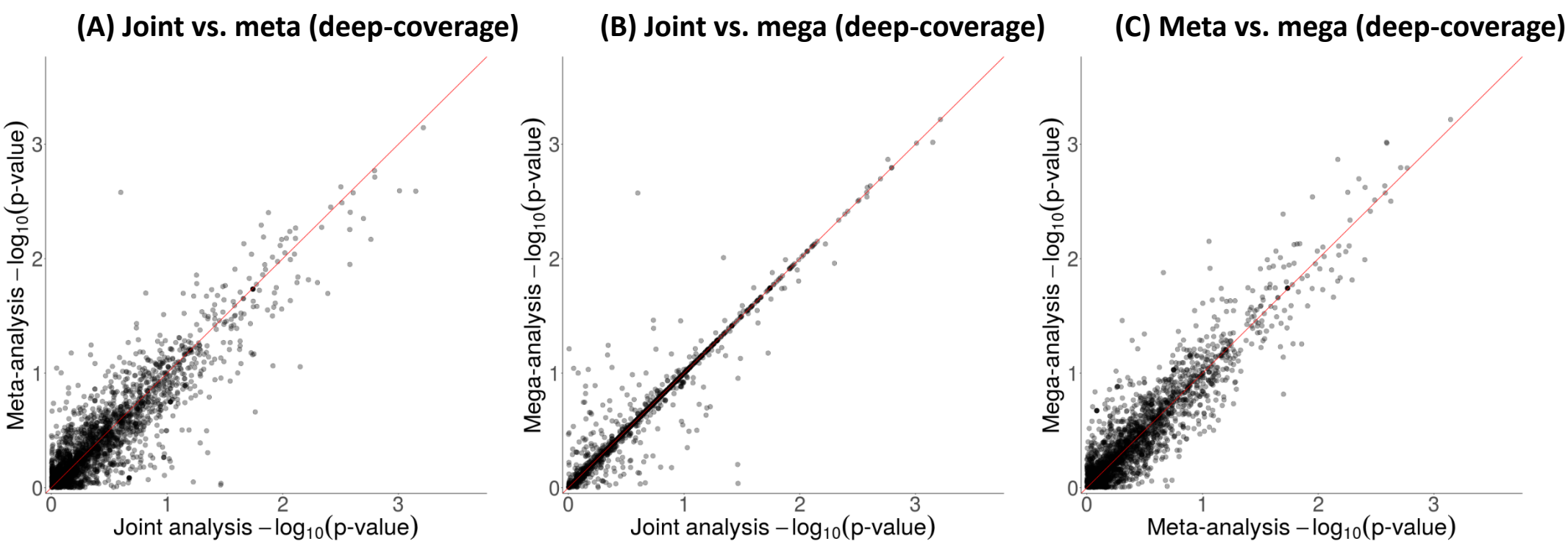
F) Low-coverage (coding), common



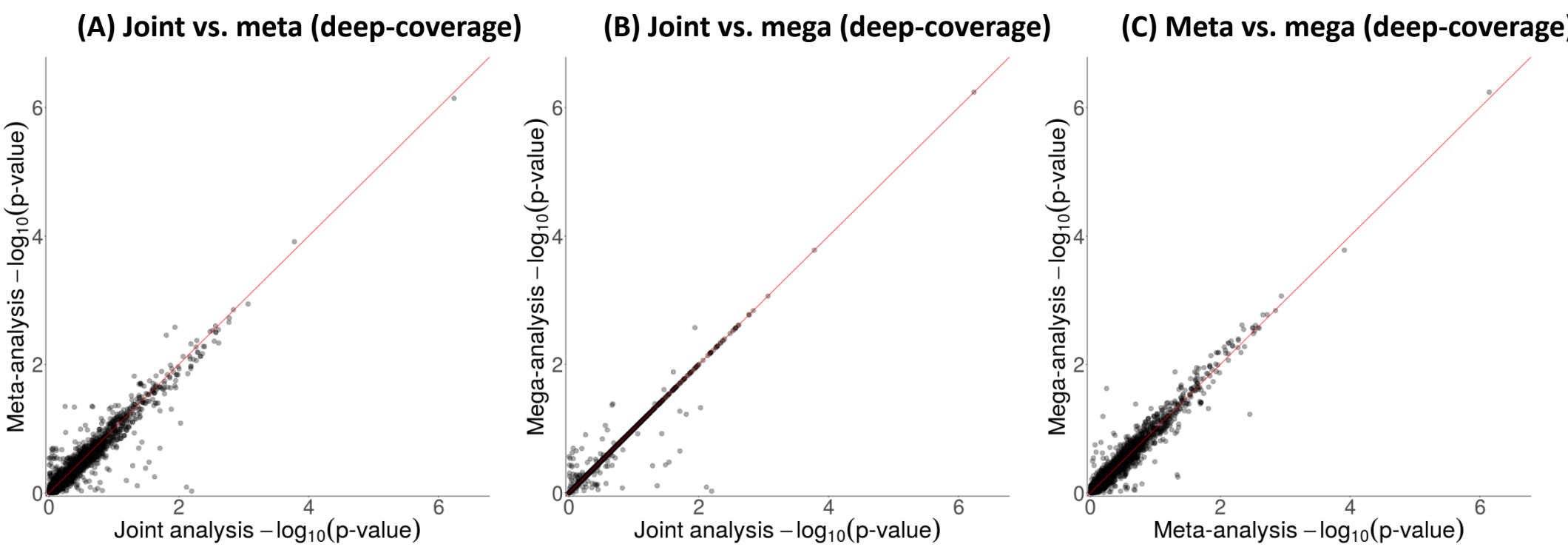
Supplementary Figure 6. GATK Pipeline. Comparison of variant detection between joint and single study calling strategies for rare (MAF<0.5%), low-frequency (MAF 0.5-5%), and common (MAF>5%) SNVs in deep-coverage (~82X) exome sequence data and low-coverage (~5X) genome sequence data restricted to coding regions.



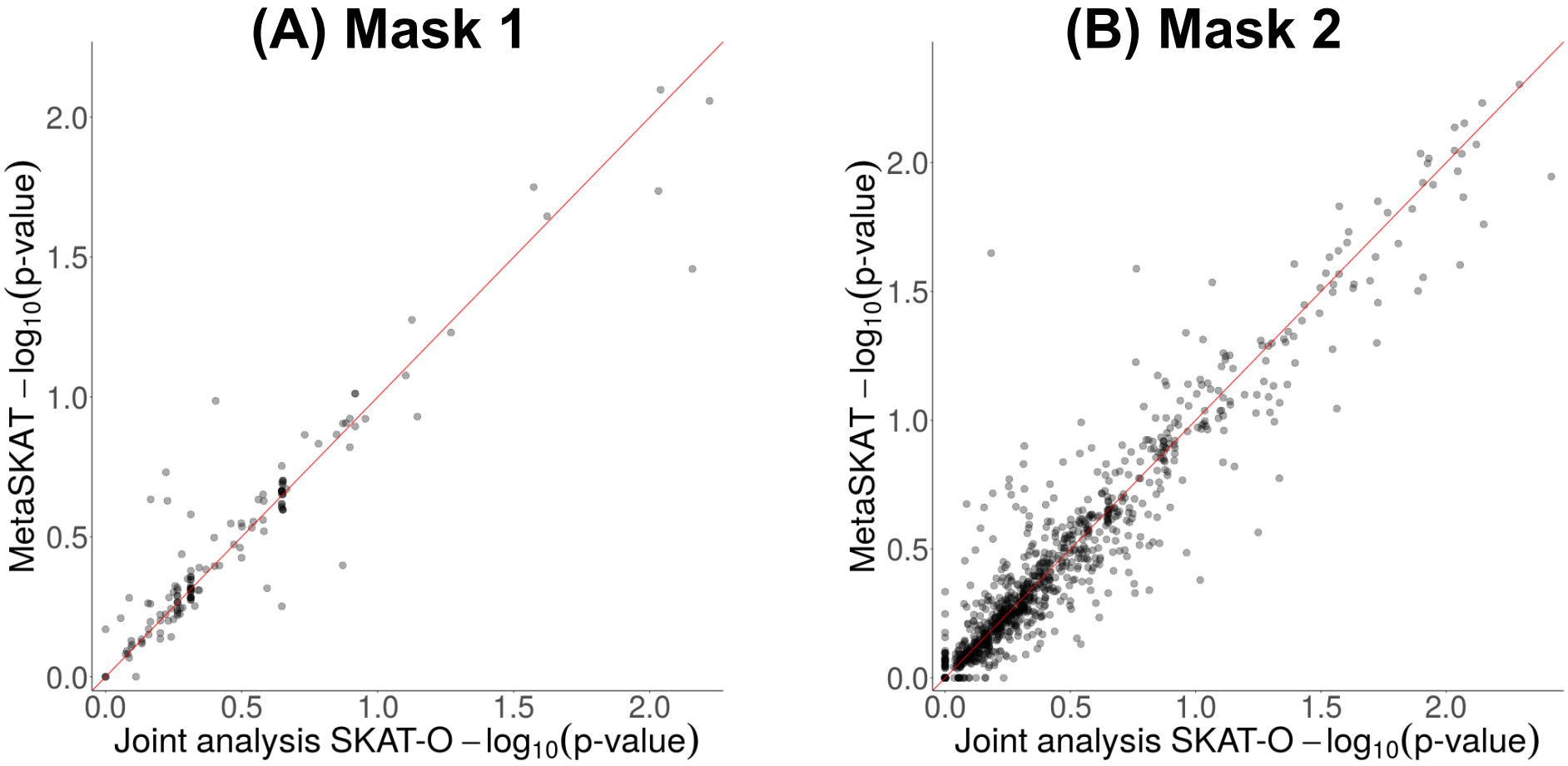
Supplementary Figure 7. GATK pipeline. Comparison of single-variant association test p-values between joint and single study calling strategies for rare (MAF<0.5%) SNVs in (A-C) deep-coverage (~82X) exome sequence data and (D-F) low-coverage (~5X) genome sequence data. *Joint* refers to joint analysis of the joint callset, *meta* refers to fixed-effects meta-analysis of single-study summary statistics, and *mega* refers to joint analysis of the union callset (mega-analysis).



Supplementary Figure 8. GATK pipeline. Comparison of single-variant association test p-values between joint and single study calling strategies for low-frequency (MAF 0.5-5%) SNVs in (A-C) deep-coverage (~82X) exome sequence data and (D-F) low-coverage (~5X) genome sequence data. *Joint* refers to joint analysis of the joint callset, *meta* refers to fixed-effects meta-analysis of single-study summary statistics, and *mega* refers to joint analysis of the union callset (mega-analysis).



Supplementary Figure 9. GATK pipeline. Comparison of single-variant association test p-values between joint and single study calling strategies for common (MAF >5%) SNVs in (A-C) deep-coverage (~82X) exome sequence data and (D-F) low-coverage (~5X) genome sequence data. *Joint* refers to joint analysis of the joint callset, *meta* refers to fixed-effects meta-analysis of single-study summary statistics, and *mega* refers to joint analysis of the union callset (mega-analysis).



Supplementary Figure 10. GATK pipeline. Comparison of gene-based association test p-values between joint and single study calling strategies in deep-coverage (~82X) exome sequence data. *MetaSKAT* refers to homogeneous effects Meta-SKAT-O test implemented in the MetaSKAT R package. Mask 1: protein-truncating SNVs; Mask 2: Mask1+missense SNVs with MAF<1%; Mask 3: Mask1+SNVs predicted deleterious by all algorithms (Polyphen2-HumDiv, PolyPhen2-HumVar, LRT, Mutation Taster, and SIFT); Mask 4: Mask1+SNVs with MAF<1% predicted deleterious by at least one algorithm.