Figure S1. Principal components analysis cluster analysis. PC, principal components.

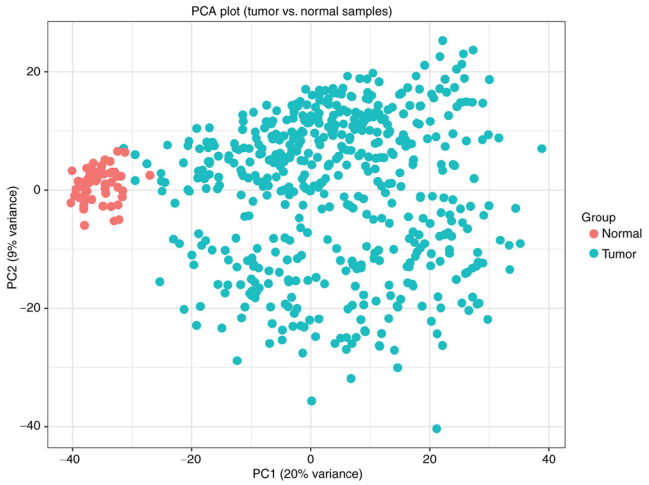

Figure S2. Scatter plots showing the correlation analysis between exposure factors and outcomes. The horizontal axis represents the effect of SNPs on exposure and the vertical axis represents the effect of SNPs on outcomes. The colored lines represent the fitting results of different MR algorithms. A positive slope indicates a risk factor, while a negative slope indicates a safety factor. SNP; single nucleotide polymorphisms; LUAD, lung adenocarcinoma; IVW, inverse variance weighted; OR, odds ratio; MR, Medelian randomization.

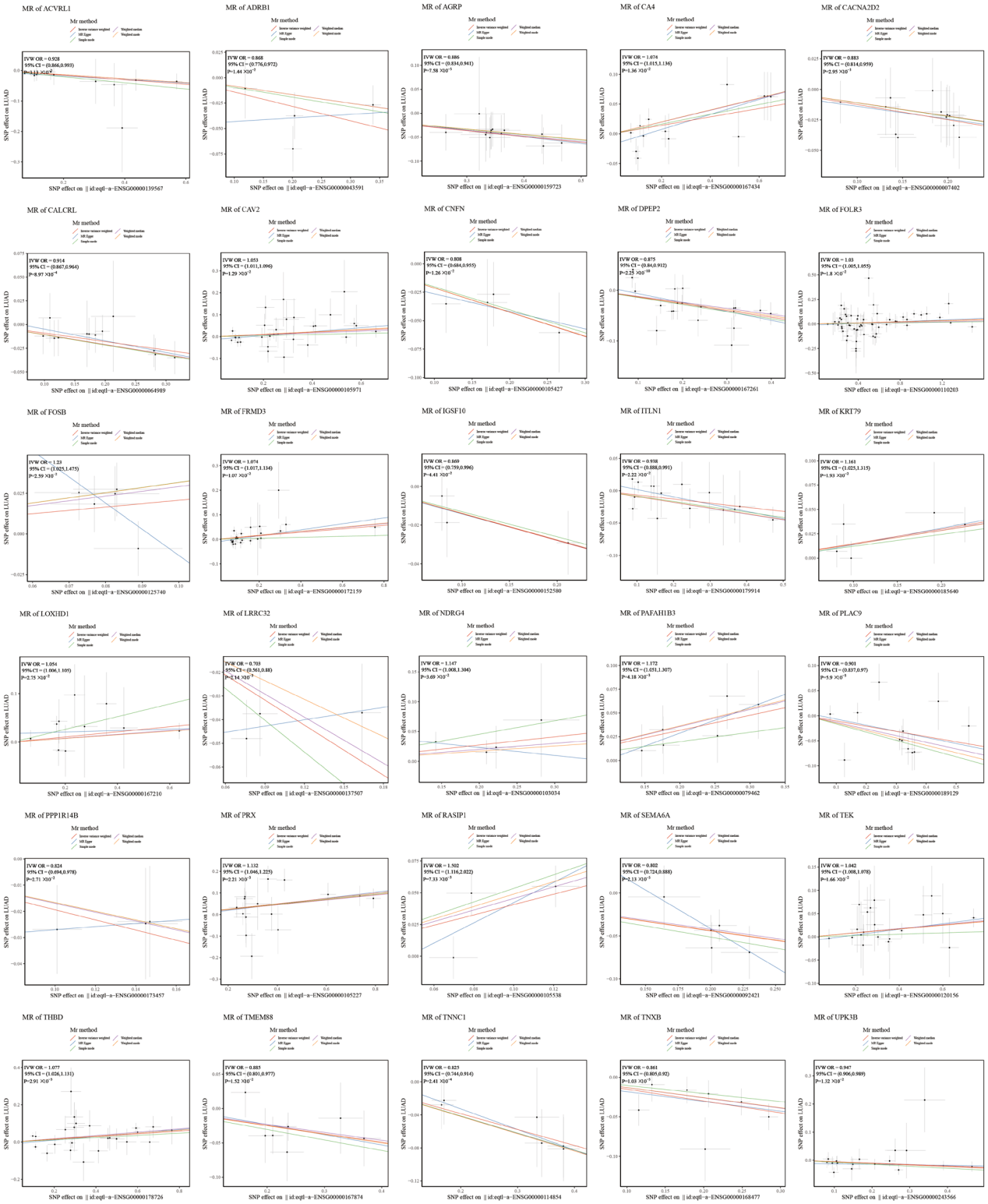

Figure S3. Forest plots displaying the instrumental variable outcome effect estimation. Each horizontal solid line reflects the result estimated by a single SNP using the Wald ratio method. A solid line completely to the left of 0 indicates that an increase in exposure factors can reduce the risk of outcome variables; a solid line completely to the right of 0 indicates that an increase in exposure factors increases the risk of outcome variables; lines crossing 0 indicate non-significant results. MR, Medelian randomization; SNP, single nucleotide polymorphisms.

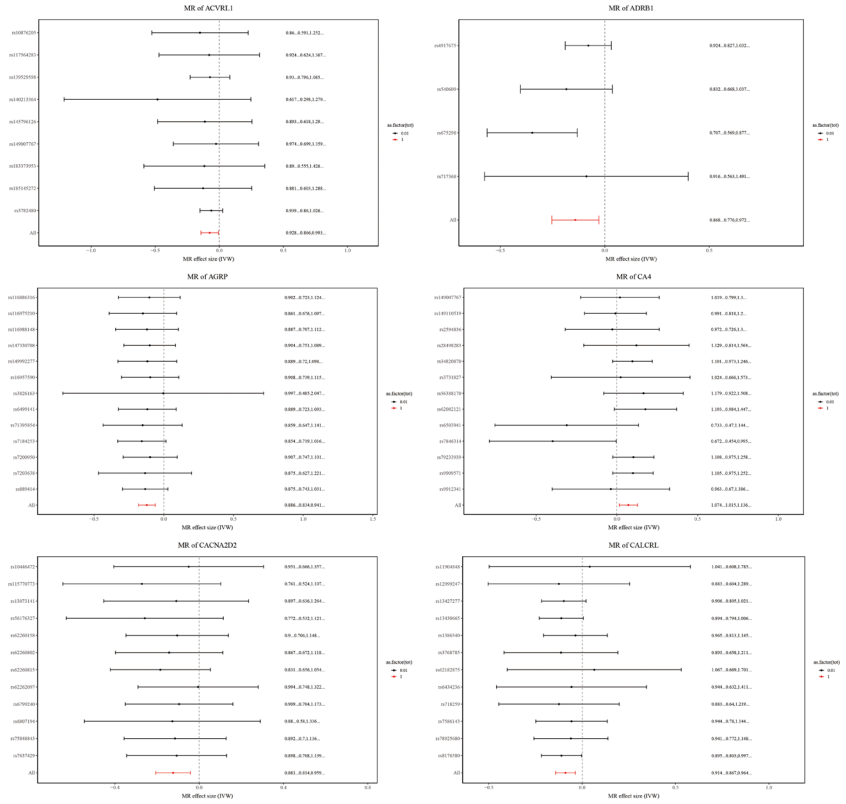

Figure S4. Forest plots of the MR analysis for candidate genes. Each black dot and horizontal line represent the effect size and 95% CI of an individual SNP associated with the gene exposure. The vertical dashed line indicates the null effect (0). Color Coding: The red dot and red horizontal bar at the bottom of each plot represent the overall pooled causal effect estimate calculated using the IVW method. A causal association is considered statistically significant if the 95% CI of the IVW estimate does not overlap with the null line. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

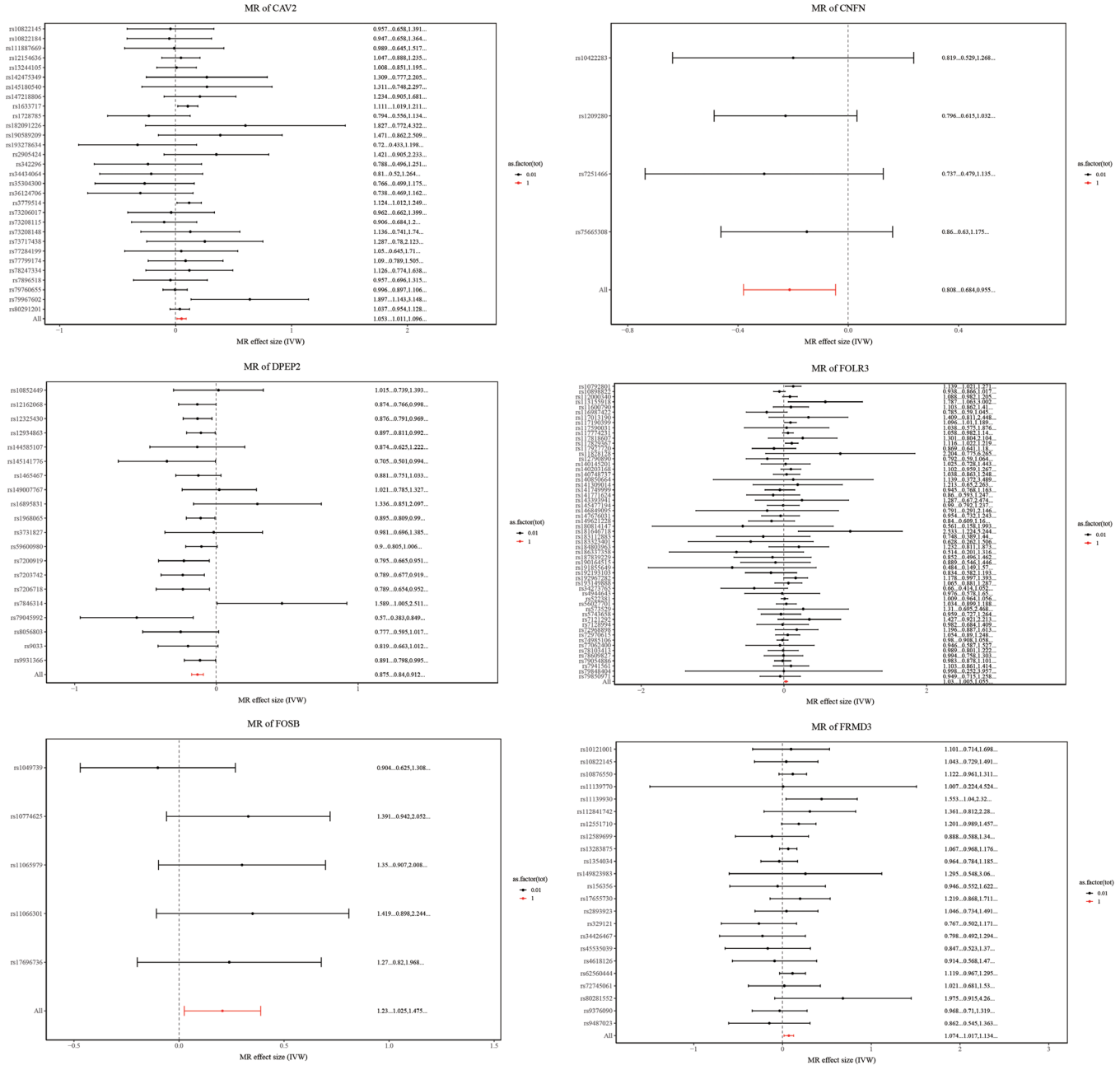

Figure S5. Forest plots of the MR analysis for additional candidate genes. Each black dot and horizontal error bar represent the effect size and 95% CI for an individual SNP associated with the gene exposure. The vertical dashed line denotes the null effect (0). Color coding: The red dot and horizontal bar at the bottom of each plot represent the overall pooled causal effect calculated via the IVW method. A significant causal relationship is indicated when the 95% CI of the IVW estimate does not cross the vertical null line. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

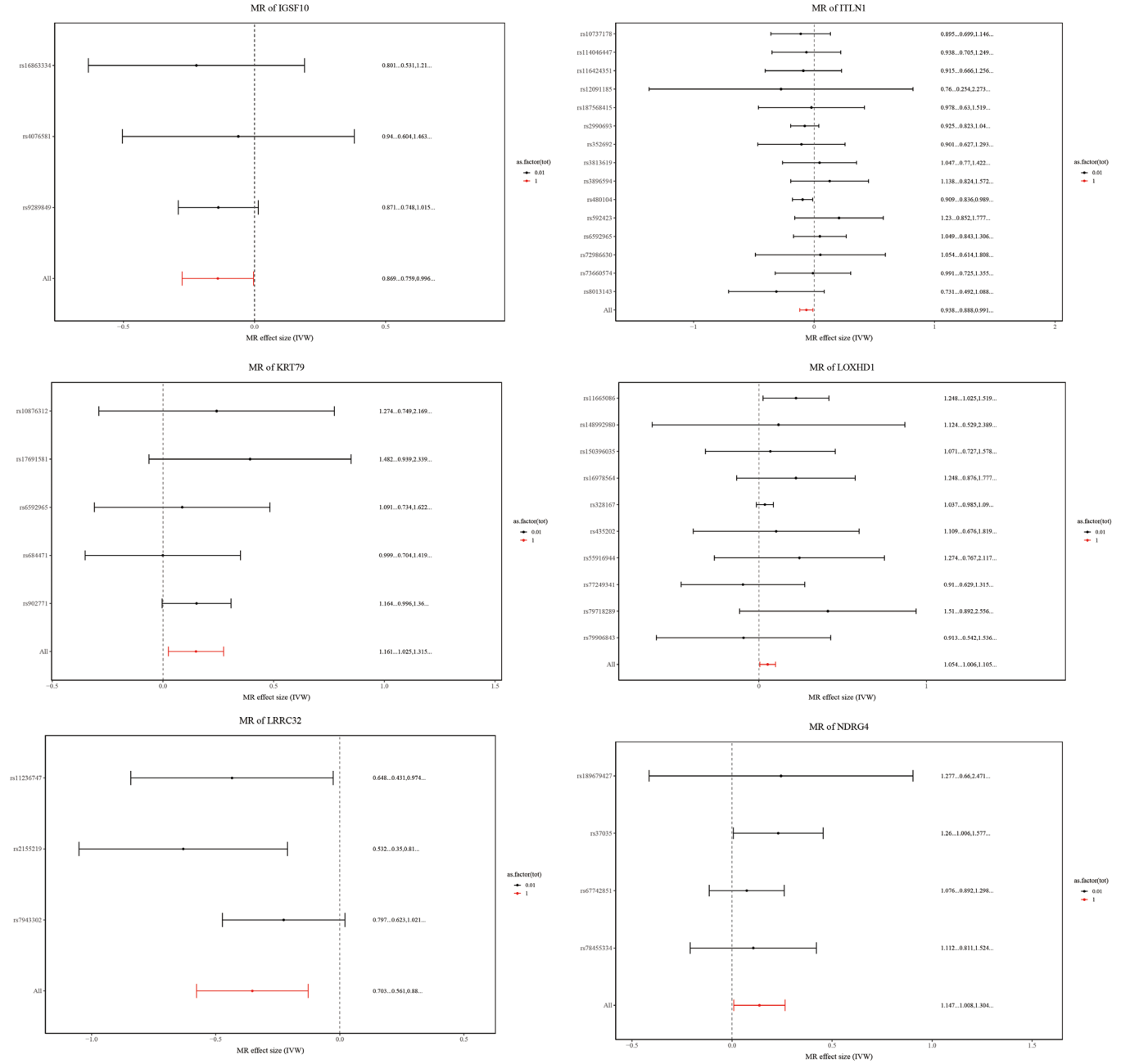

Figure S6. Forest plots of the MR analysis for additional candidate genes. Each black dot and horizontal error bar represent the effect size and 95% CI for an individual SNP associated with the respective gene exposure. The vertical dashed line represents the null effect (0). Color coding: The red dot and horizontal bar at the bottom of each plot represent the pooled causal effect estimate calculated using the IVW method. A causal link is considered statistically significant if the 95% CI of the IVW estimate does not overlap with the vertical null line. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

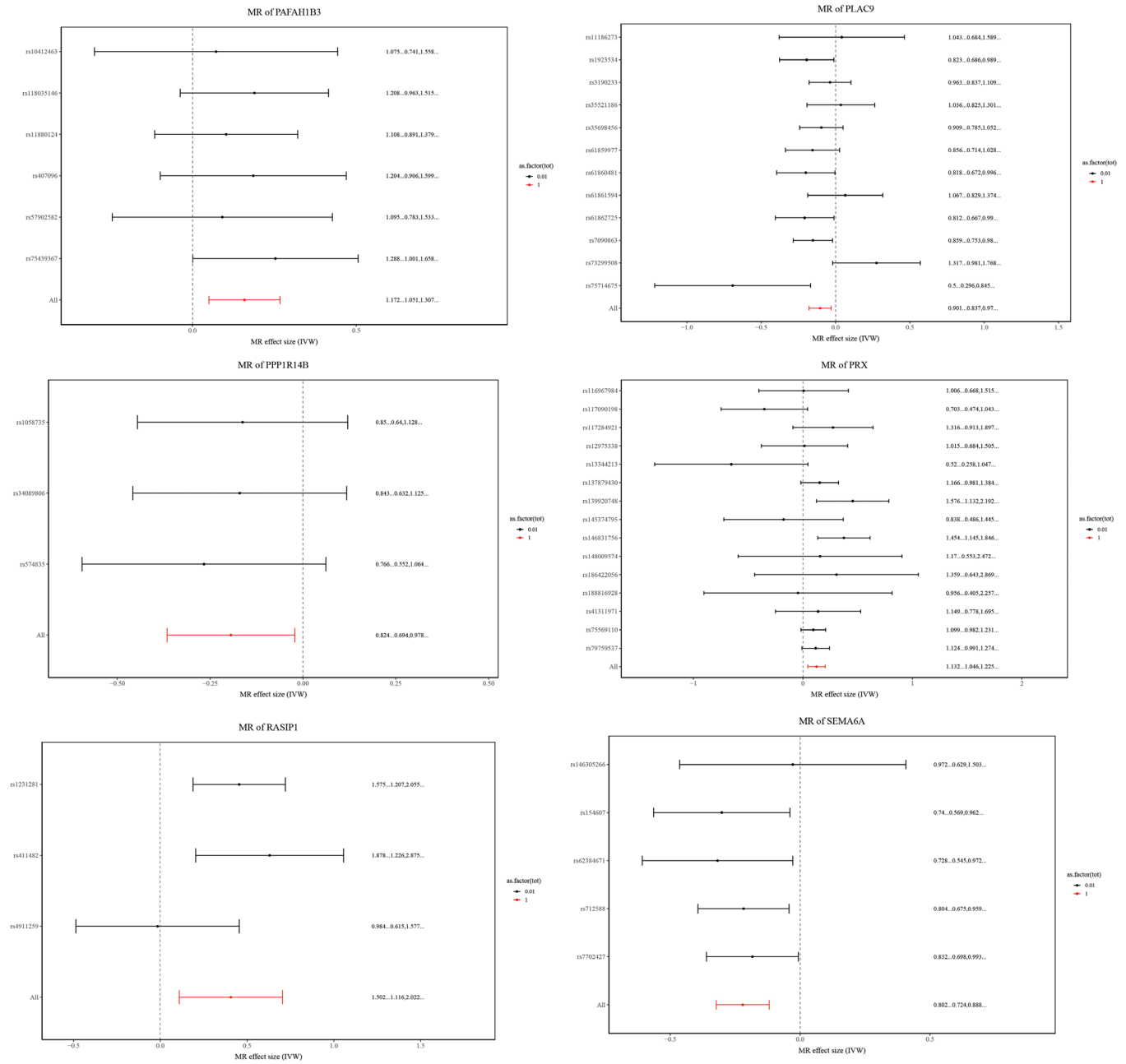

Figure S7. Forest plots of the MR analysis for additional candidate genes. Each black dot and the accompanying horizontal line represent the effect size and 95% CI for an individual SNP serving as an instrumental variable for the respective gene exposure. The vertical dashed line indicates the null effect (0). Color Coding: The red dot and horizontal bar at the bottom of each plot represent the overall causal effect estimate calculated using the IVW method. A causal relationship is considered statistically significant if the 95% CI of the IVW estimate does not overlap with the vertical null line. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

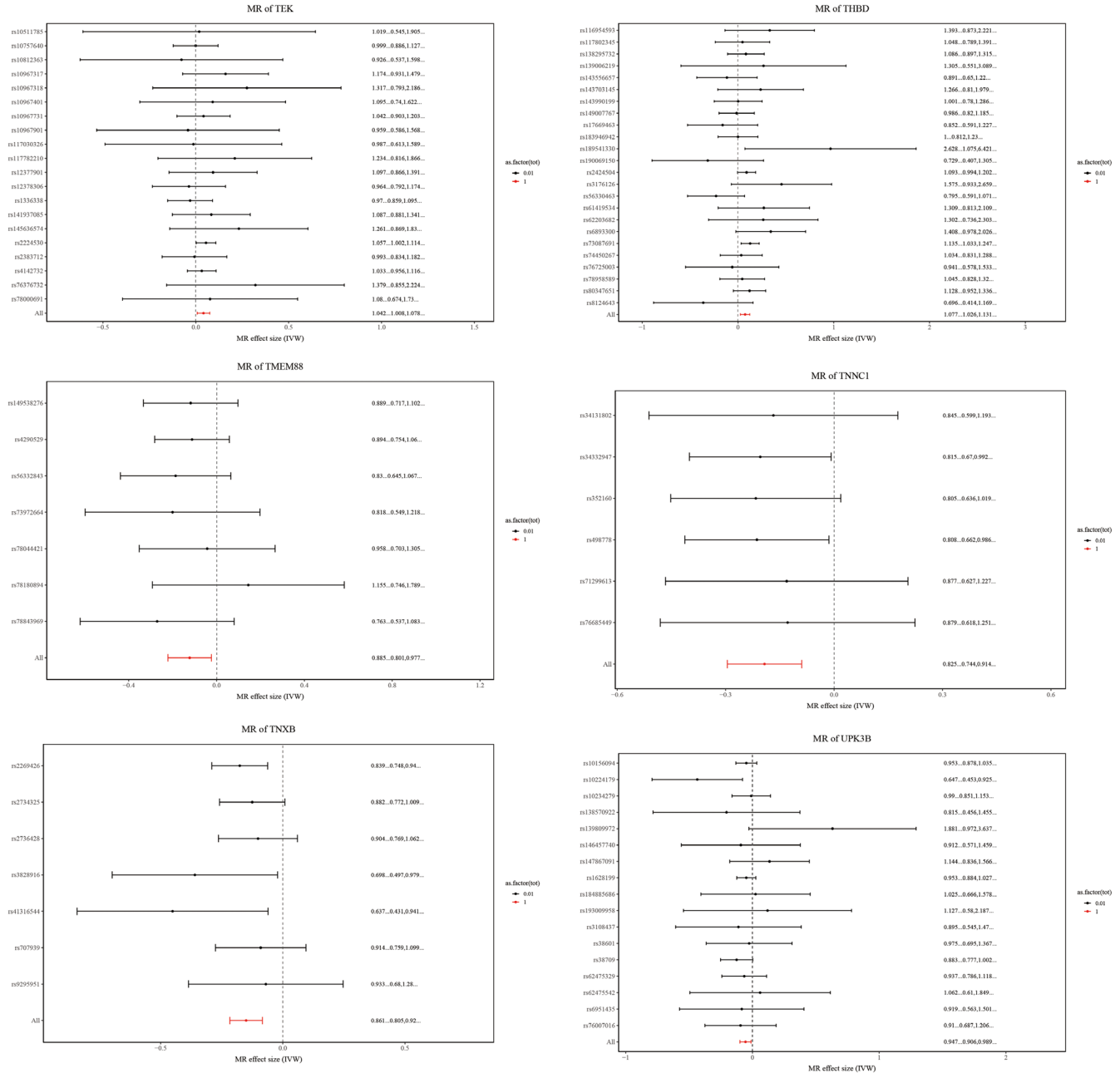

Figure S8. Funnel plot illustrating whether the analysis has randomness and conforms to Mendel's second law for random grouping. The horizontal axis represents the β value of the instrumental variable and the vertical axis represents the reciprocal of the standard error of the instrumental variable.

Figure S9. Leave-one-out analysis forest plot showing the significant changes caused by each instrumental variable (SNP) on the outcome. The inverse variance weighted method is used for analysis, with the horizontal axis representing the effect values of each SNP locus on the outcome through exposure factors, the vertical axis representing SNP loci and the red line representing the overall effect value. SNP, single nucleotide polymorphisms.

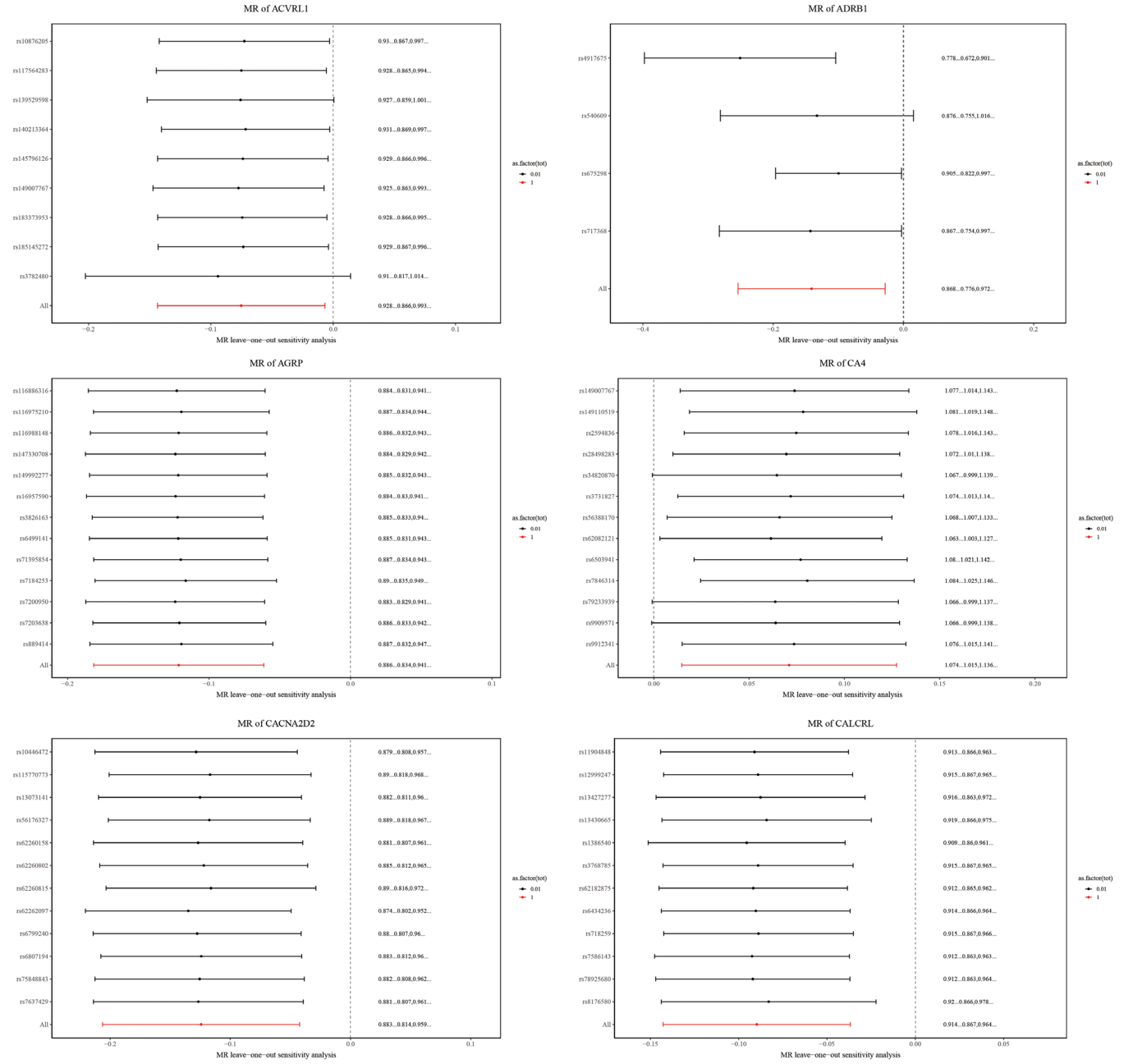

Figure S10. Leave-one-out sensitivity analysis of the Mendelian randomization (MR) for candidate genes. Each black dot and horizontal error bar represent the recalculated IVW estimate and its 95% CI after the systematic removal of each specific SNP listed on the vertical axis. The vertical dashed line indicates the null effect (0). Color Coding: The red dot and horizontal line at the bottom of each plot represent the overall pooled causal effect estimate calculated using all available instrumental variables. The consistency of the results after individual SNP exclusion indicates that the overall causal inference is not disproportionately driven by any single outlier SNP. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

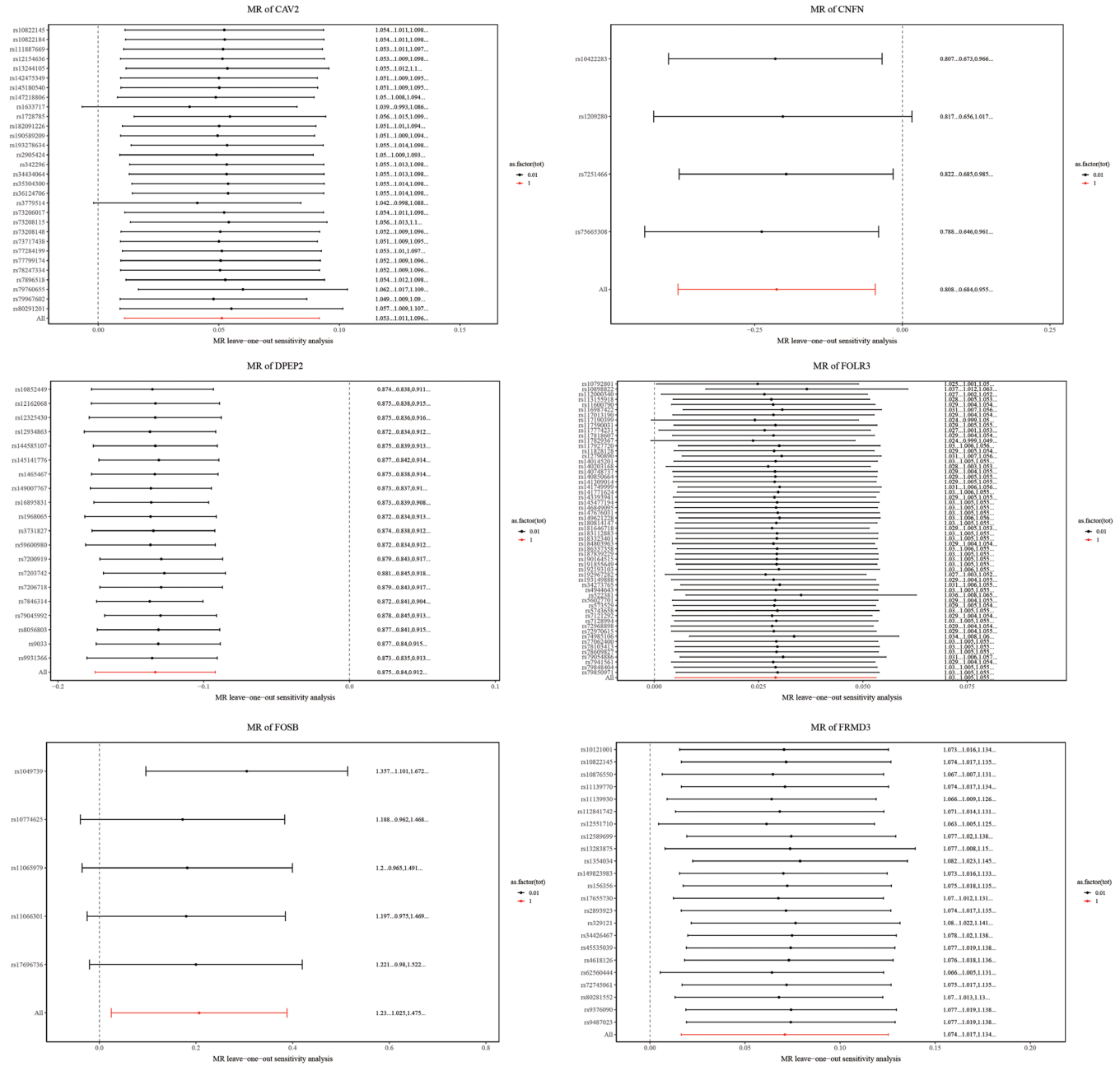

Figure S11. Leave-one-out sensitivity analysis of the MR for additional candidate genes. Each horizontal black line represents the recalculated IVW estimate and its 95% CI when the specific SNP listed on the vertical axis is excluded from the analysis. The vertical dashed line indicates the null effect (0). Color Coding: The red dot and red horizontal bar at the bottom of each panel represent the final pooled causal effect estimate using all available instrumental variables. The overall consistency of the effect estimates across individual SNP exclusions demonstrates that the causal inference for these genes is stable and not disproportionately influenced by any single SNP. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

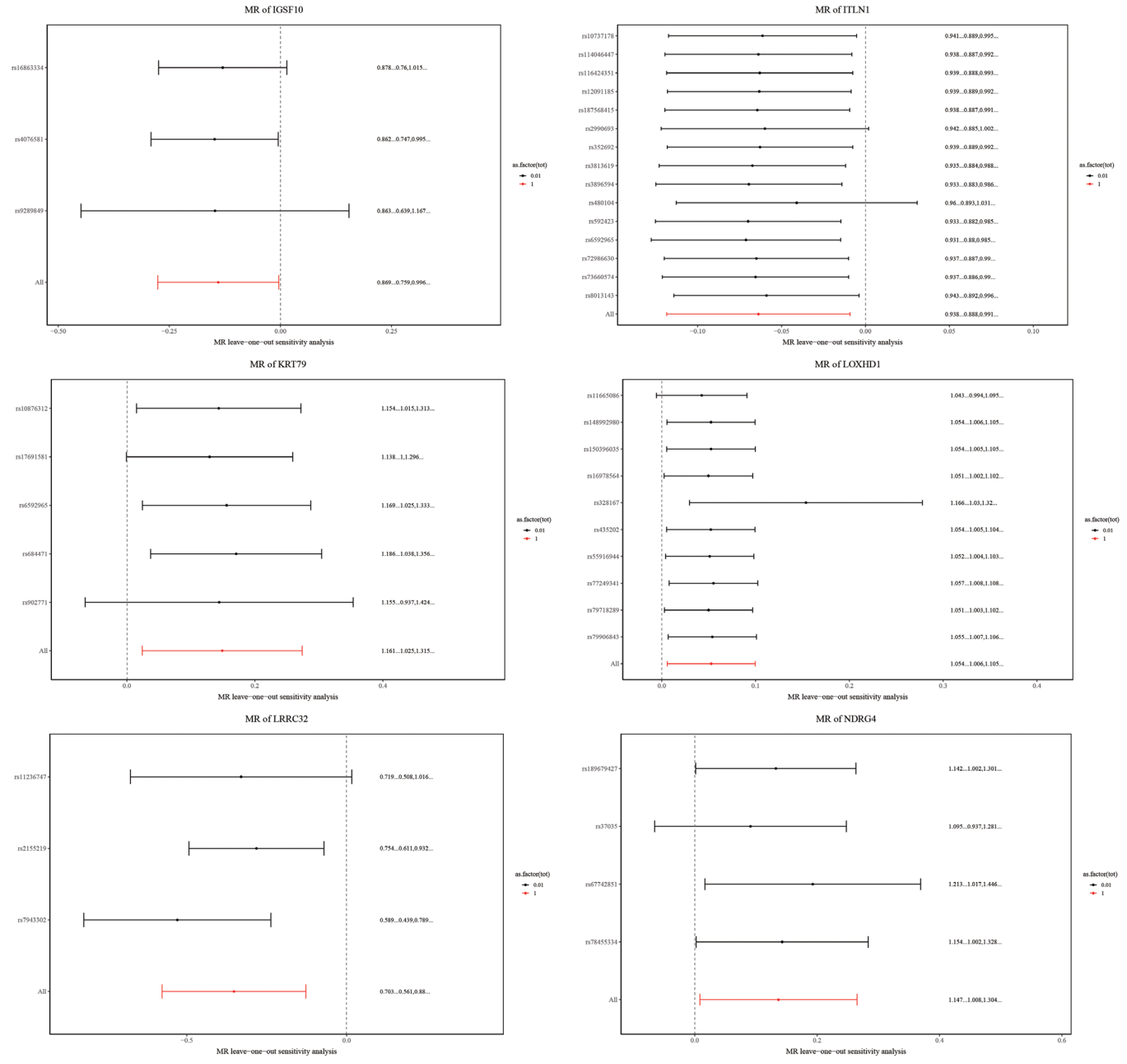

Figure S12. Leave-one-out sensitivity analysis of the MR for additional candidate genes. Each horizontal black line represents the recalculated IVW estimate and its 95% CI upon the systematic exclusion of each individual SNP listed on the vertical axis. The vertical dashed line indicates the null effect (0). Color Coding: The red dot and red horizontal bar at the bottom of each panel represent the final pooled causal effect estimate calculated using all available instrumental variables. The consistency of results across individual SNP exclusions indicates that the overall causal inference for these genes is robust and is not disproportionately driven by any single outlier SNP. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

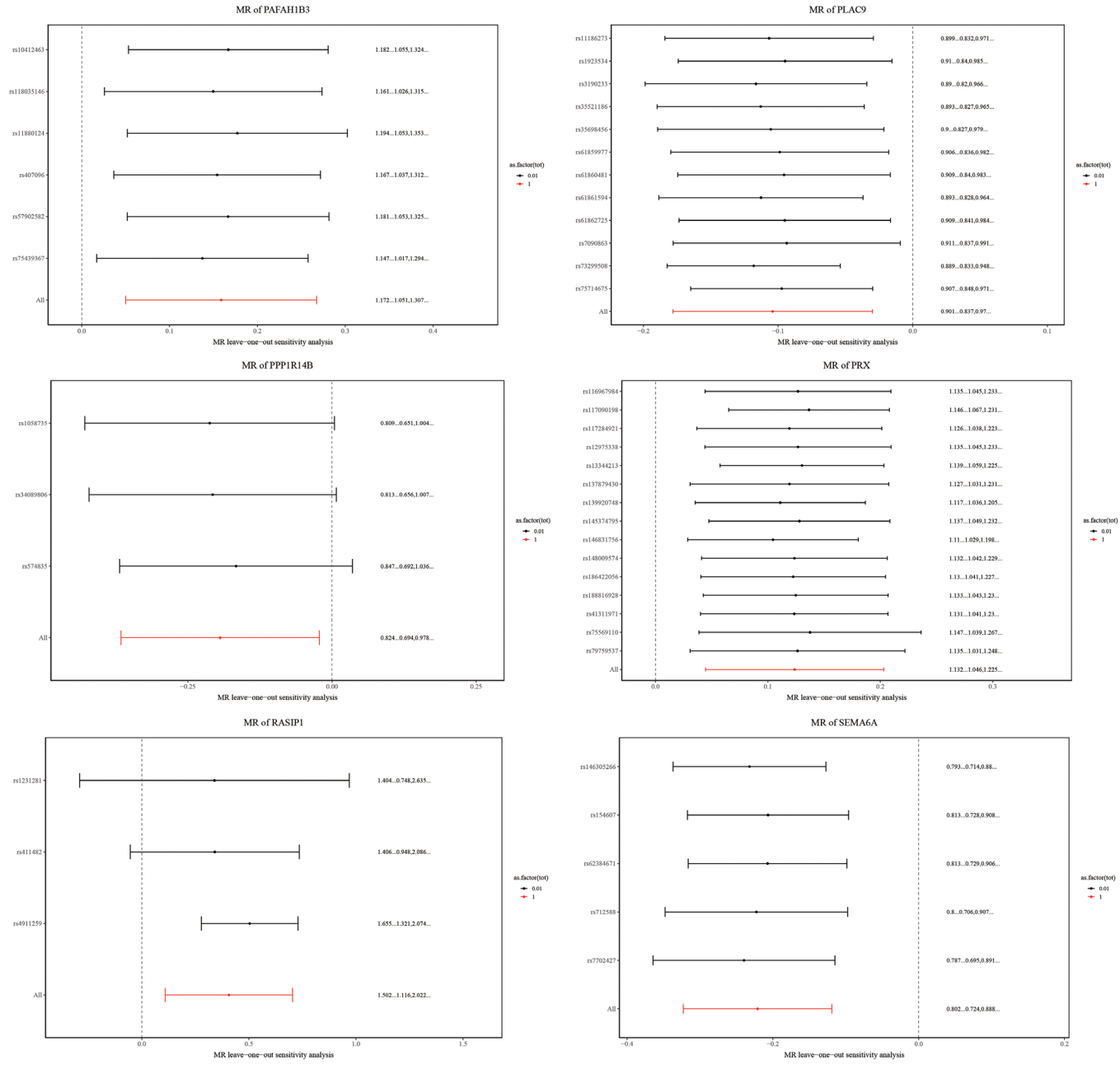

Figure S13. Leave-one-out sensitivity analysis of the MR for additional candidate genes. Each horizontal black line and its error bar represent the recalculated IVW estimate and the corresponding 95% CI after the systematic removal of the specific SNP listed on the vertical axis. The vertical dashed line indicates the null effect (0). Color Coding: The red dot and red horizontal bar at the bottom of each panel represent the final pooled causal effect estimate calculated using all instrumental variables. The consistency of results across individual SNP exclusions demonstrates that the overall causal inference for these genes is not driven by any single influential SNP. CI, confidence interval; IVW, inverse variance weighted; LUAD, lung adenocarcinoma; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

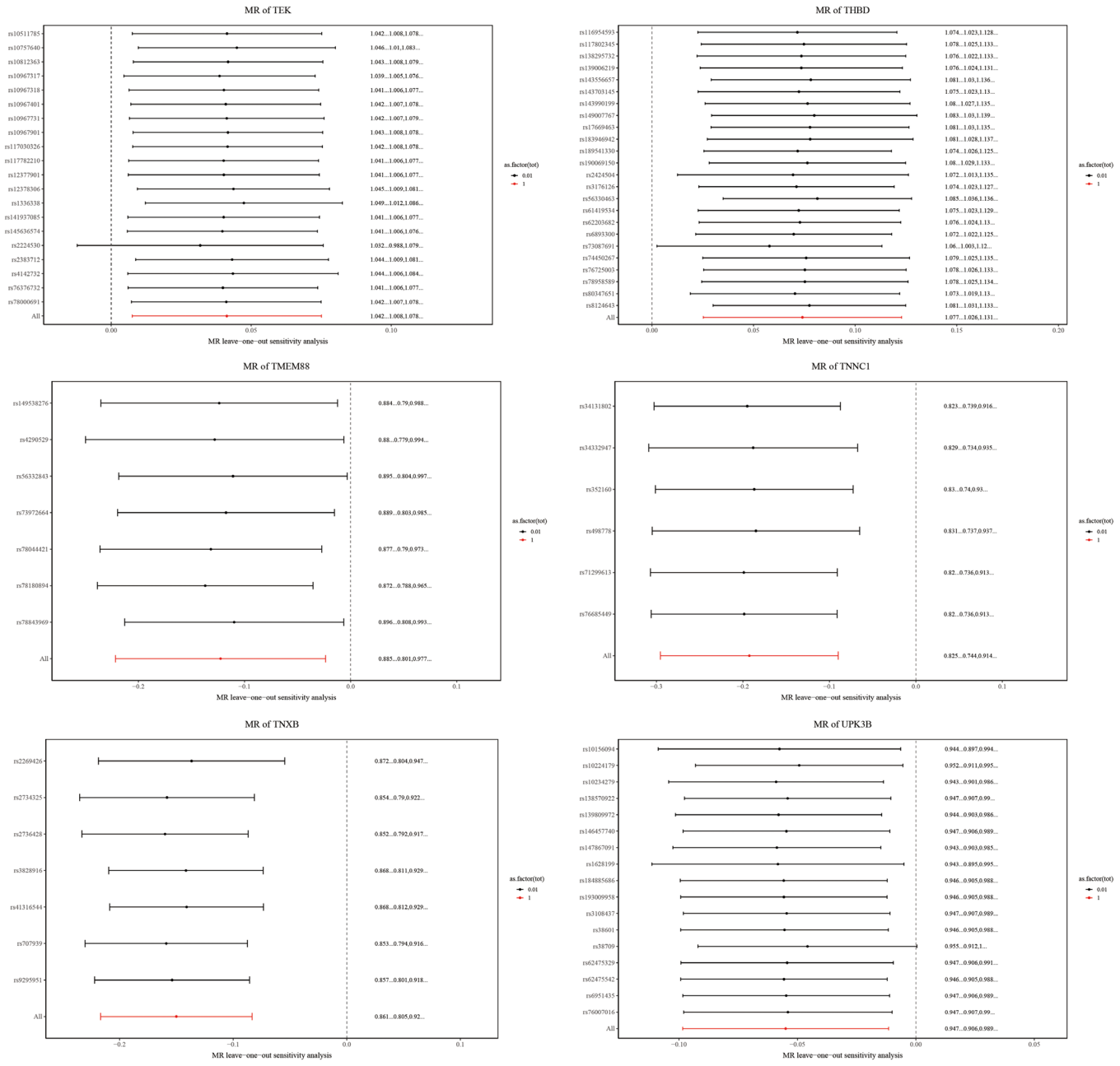

Figure S14. (A) The violin plot represents the distribution of nFeature-RNA (number of genes detected per cell), nCount-RNA (sum of the expression levels of all genes detected per cell) and percentage mt (proportion of mitochondrial genes detected per cell) in the sample. (B) The first 2,000 highly variable genes after UMAP dimensionality reduction are displayed as highly variable genes. (C) Linear dimensionality reduction scree plot, with latitude values on the horizontal axis and standard deviation on the vertical axis, with scatter points representing dimensional distribution.(D) UMAP representation using linear dimensionality reduction to obtain the optimal linear dimension value for cell clustering, which can be used to divide subsequent cell groups. (E) The expression of marker genes in each cell type. PC, principal components; LUAD, lung adenocarcinoma; NK, natural killer; UMAP, Uniform Manifold Approximation and Projection.