Supplemental information

Mendelian randomization and pathway analysis demonstrate shared genetic associations between lupus and coronary artery disease

Jessica Kain, Katherine A. Owen, Miranda C. Marion, Carl D. Langefeld, Amrie C. Grammer, and Peter E. Lipsky
**Figure S1. Analysis of SNP-predicted genes associated with both SLE and CAD.** Related to Figures 1 and 2. 

A) Venn diagram of overlap between multi-ancestral SLE- and CAD-associated \((p<10^{-6})\) SNPs. 

B) Venn diagram of overlap between SNP-predicted genes derived from regulatory elements (T-Genes), eQTL analysis (E-Genes), coding regions (C-Genes), and proximity within 5kb (P-Genes). 

C) Application of S-LDSC using summary statistics for SLE, CVD, and CAD GWAS to estimate the heritability of the 189 SNP-predicted genes (top panel) and 135 SNP-predicted proteins (lower panel) from STRINGdb. Bar color indicates coefficient significance. 

D) Application of LDSC to estimate the genetic correlation between SLE and CAD or CVD. 

E) PPI network consisting of 135 putative protein-coding genes. Functional and cell-type enrichments for each cluster were determined using BIG-C (black labels) and I-scope (red labels), respectively. Black labels over colored shadings represent shared BIG-C functional annotations for the clusters they surround.
Figure S2. Bidirectional MR summaries between SLE and CAD. Related to Figure 2 and Table S3. Scatter plots showing GWAS effect size estimates on the exposure (x-axis) and outcome (y-axis) with each dot representing a SNP and lines representing MR-estimates of SLE on CAD, MI and IS (A) and in the reverse direction, with CAD or MI as exposure and SLE as the outcome (B). MR-IVW and MR-Egger heterogeneity test results (Q-value) indicate whether significant heterogeneity was detected (asterisks, p<0.05), which does not necessarily indicate biased causal estimates. MR-Egger intercept indicates whether significant (asterisks, p<0.05) directional horizontal pleiotropy was detected, which usually indicates biased causal estimates. N.s., not significant.
Figure S3. SLE-associated SNPs on chromosome 6 account for the majority of negative causal effects on CAD by SSMR. Related to Figure 3. 

A-B) Forest plots (beta ± standard error) of the top 25 (by absolute value of causal estimates) positive (A) and negative (B) causal SNPs identified by SSMR using the Wald-ratio method. 

C-D) Pie charts illustrating the distribution of 119 positive (C) and 234 negative (D) causal SLE SNPs on CAD. 

E-F) Cluster metastructures for the 498 (E) 557 (F) predicted genes from positive and negative causal SNPs identified by single-SNP MR. Metastructures are based on PPI networks, clustered using MCODE and visualized in Cytoscape. Node gradient shading is proportional to intra-cluster connectivity, cluster size indicates number of genes per cluster and edge weight indicates inter-cluster connections. Functional and cell-type enrichments for each cluster were determined using BIG-C (black labels) and I-scope (red labels), respectively. Bold black labels over colored shadings represent shared functional annotations for the clusters they surround.
Figure S4. MR analyses for positive and negative causal SNPs determined by SSMR. Related to Figure 3 and Table S3. A) Forest plots (beta ± standard error) of the 80 positive (A) and 96 negative (B) causal non-HLA SNPs identified by SSMR using the Wald ratio method, ordered by absolute value of causal estimates.
Figure S5. Analysis of HLA SNP-predicted genes associated with both SLE and CAD. Related to Figure 3. A) Forest plot showing GWAS effect sizes ± standard error for 30 HLA SNPs significantly (p<10^{-6}) associated with both SLE (red) and CAD (blue). B) PPI network consisting of 69 putative protein-coding genes predicted from the 30 HLA SNPs. Functional and cell-type enrichments for each cluster were determined using BIG-C (black labels) and I-scope (red labels), respectively. Black labels over colored shadings represent shared functional annotations for the clusters they surround. C) Gene set enrichments for each cluster were determined using IPA and EnrichR. P-values are from Fisher’s exact test that measures the significance of overlap between analysis-ready genes in each cluster and genes within an annotation, with red shading proportional to significance of each enrichment.
Figure S6. Positive and negative causal estimates for PPI-based clusters using MR-IVW. Related to Figures 4 and 5. (A-B) PPI-based MR-IVW (beta ± standard error) using these the 46 (A) and 67 (B) clusters of SLE SNP-derived IVs in CAD, MI, IS, cardiomyopathy, and atrial fibrillation GWAS. For results, grey indicates insignificant (p > 0.05), dark red and red, positive causal at p < 0.00075 and p < 0.05, respectively; dark blue and blue, negative causal p < 0.00075 and p < 0.05, respectively by IVW.
Figure S7. Expected vs. observed MR-IVW causal estimates corresponding to random vs. PPI-based SNP-to-gene modules. Related to Figure 5. A) Schematic illustrating the Monte Carlo Simulations for expected MR results using random sets of Immunochip-derived SNP-to-Gene modules. B-D) Histograms representing the proportion of insignificant (p>0.05, gray), positive causal (p<0.05, red), negative causal (p<0.05, blue), positive causal (p<0.00075, dark red), and negative causal (p<0.00075, dark blue) results with respect to number of SNPs used as IVs for SLE-exposure on CAD corresponding to (B) the 46 SLE-derived clusters and (C) the comprehensive 67 SLE-derived clusters and (D) over 50,000 random sets of Immunochip-derived SNP-to-Gene modules.
| Figure | Analysis | Purpose | Main findings |
|--------|----------|---------|---------------|
| 1      | MR-multiple methods | Exploratory analyses using an expanded SNP repertoire from multiple sources to examine estimated associations between SLE and CAD. | Results suggest a net-positive association for SLE on CAD. Provide justification to confirm estimated association. |
| 2      | MR-multiple methods, S-LDSC, pathway analysis | Confirmatory analyses using highly curated IVs, S-LDSC to determine heritability and pathway analysis to examine pathways underlying SLE and causal of CAD. | Majority of MR methods show positive causal estimate on CAD and MI, predicted genes capture SLE heritability and pathways reflect immune and CVD dysfunction. |
| 3      | SSMR, S-LDSC, pathway analysis | Orthogonal MR approach to identify single SLE SNPs with positive or negative estimates on CAD. | SSMR identifies positive and negative causal SNPs that capture significant SLE heritability and predict a number of underlying causal and protective pathways. |
| 4      | PPI-based MR, S-LDSC, pathway analysis | Development of PPI-based MR (also outlined in the graphical abstract). | PPI-based MR identifies 46 groups of SNPs for use as IVs based on cluster membership. Application of MR-IVW identifies clusters with positive and negative causal estimates. |
| 5      | PPI-based MR, S-LDSC, pathways analysis | PPI-based MR using a larger, comprehensive network. | Identification of 67 SNP sets as IVs. Application of MR-IVW identifies clusters with positive and negative causal estimates. |
| 6      | MR-IVW | PPI-based MR IV validation after accounting for pleiotropy and LD. | Application of LD clumping to cluster-derived SNPs followed by MR-IVW to confirm clusters with positive and negative causal estimates. |
| 7      | Drug matching | Identify new therapeutic interventions. | Pathways linked to positive causal clusters predict novel therapies for managing the inflammatory environment contributing to CAD in SLE. |

Table S10. Summary of major findings. Related to STAR Methods.