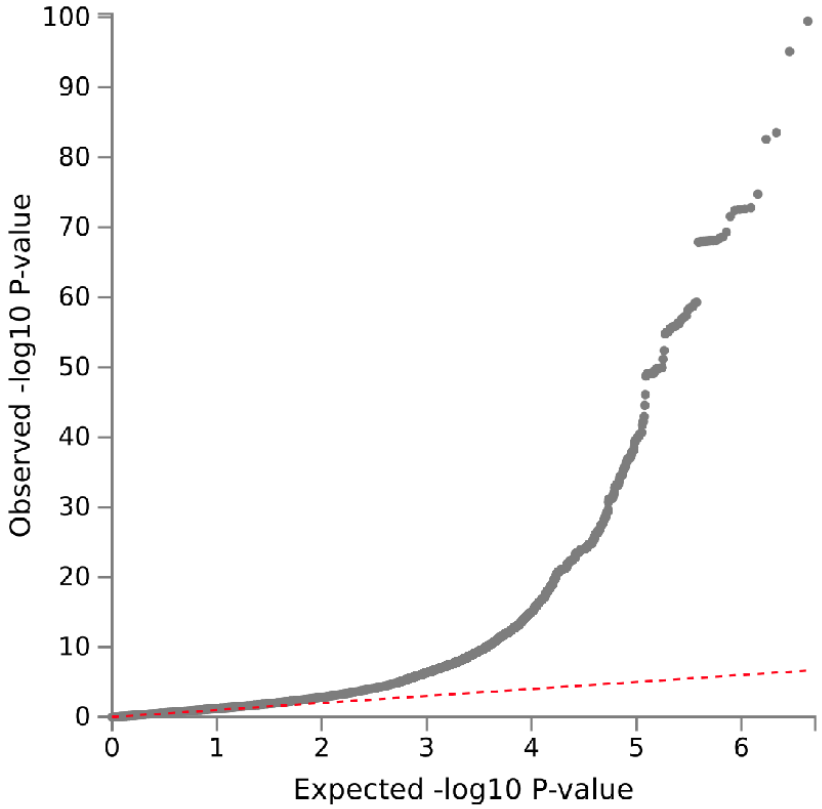
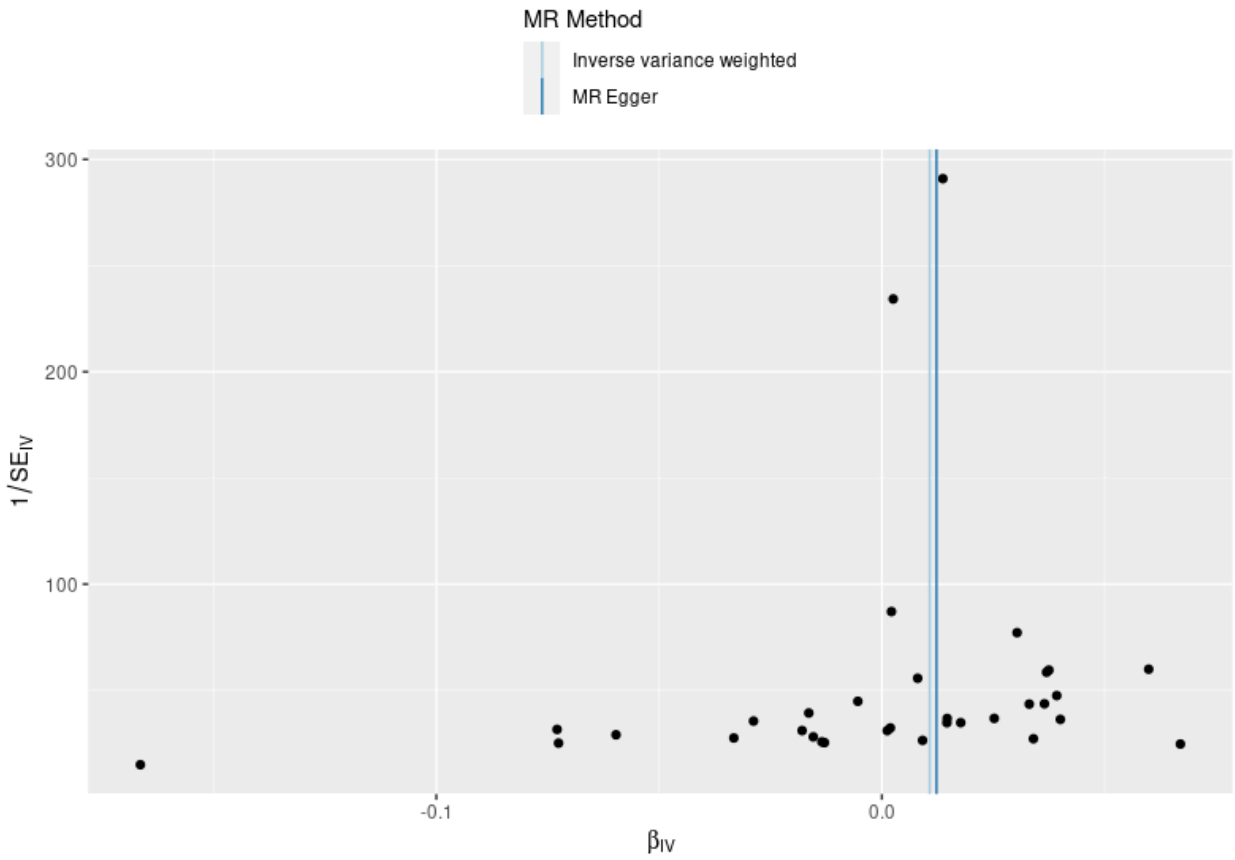
Uncovering Novel Genetic Loci and Biological Pathways Associated with Age-Related Cataracts through GWAS Meta-Analysis

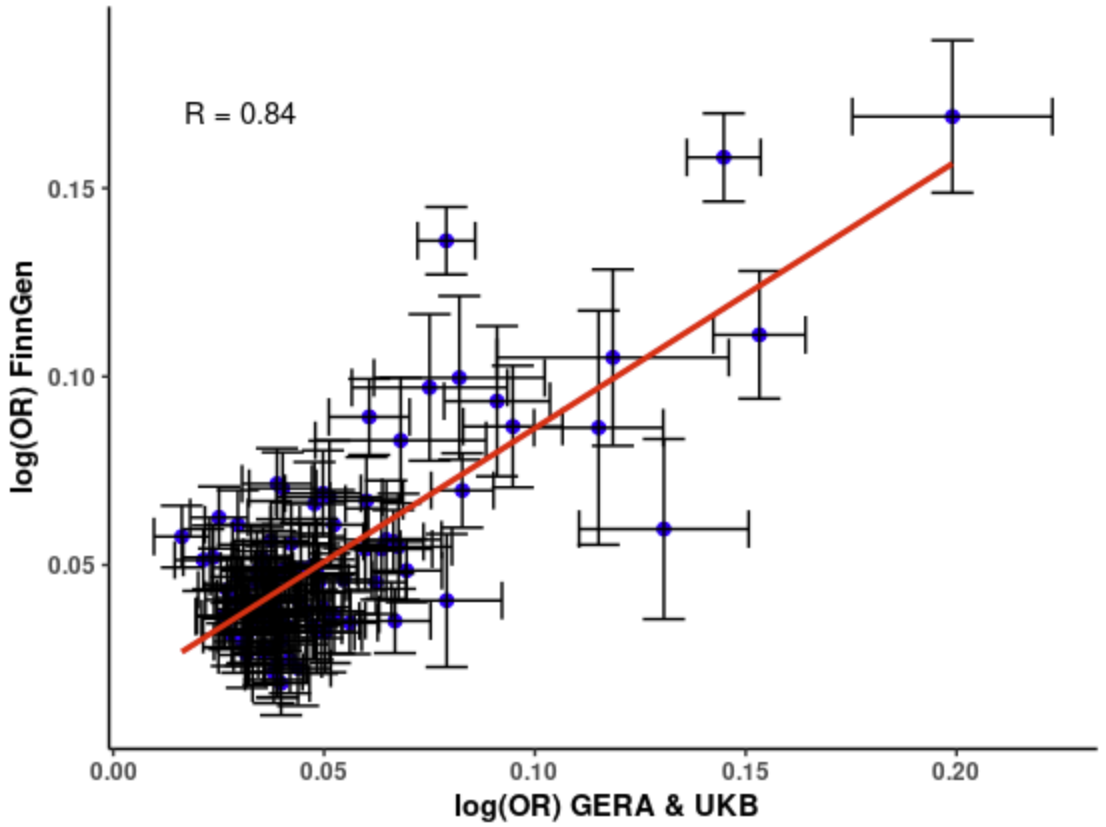
Diaz-Torres *et al.*



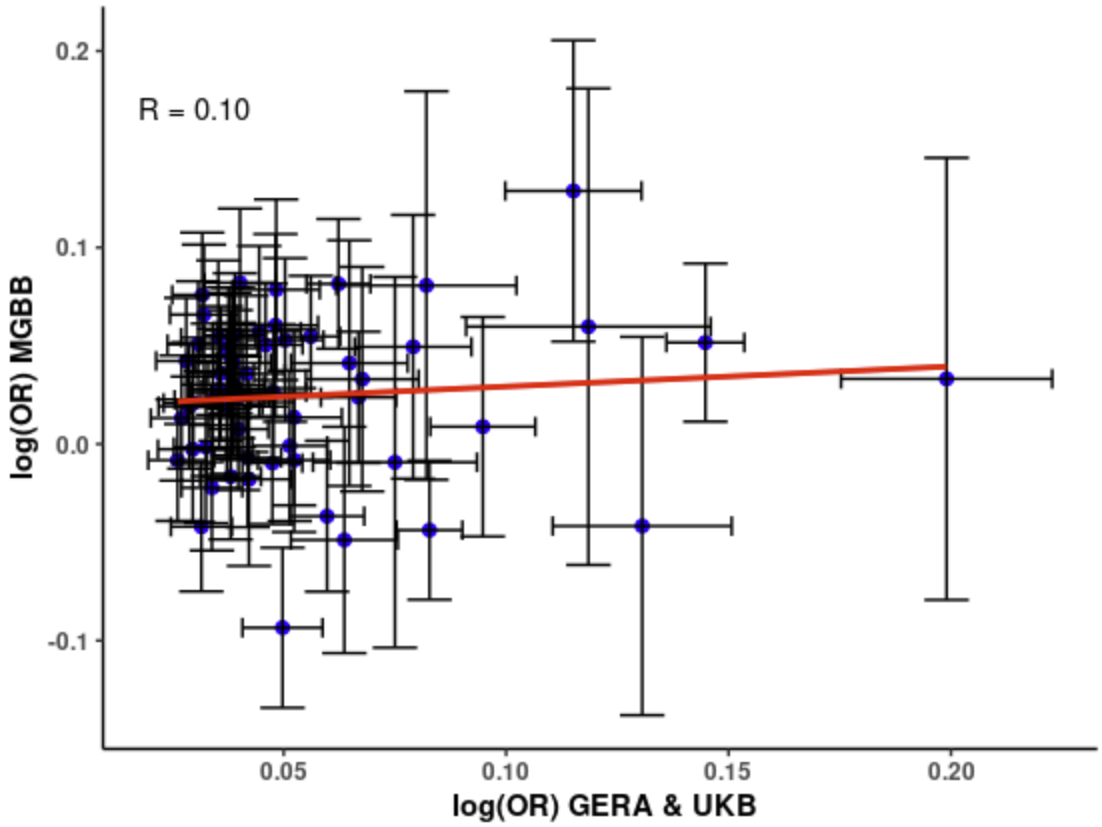
Supplementary Figure 1. Q-Q plot based on the GWAS meta-analysis, which combines data from five cohorts, showed a genetic inflation factor of 1.14 and an LD-score intercept of 1.04 (se=0.009).



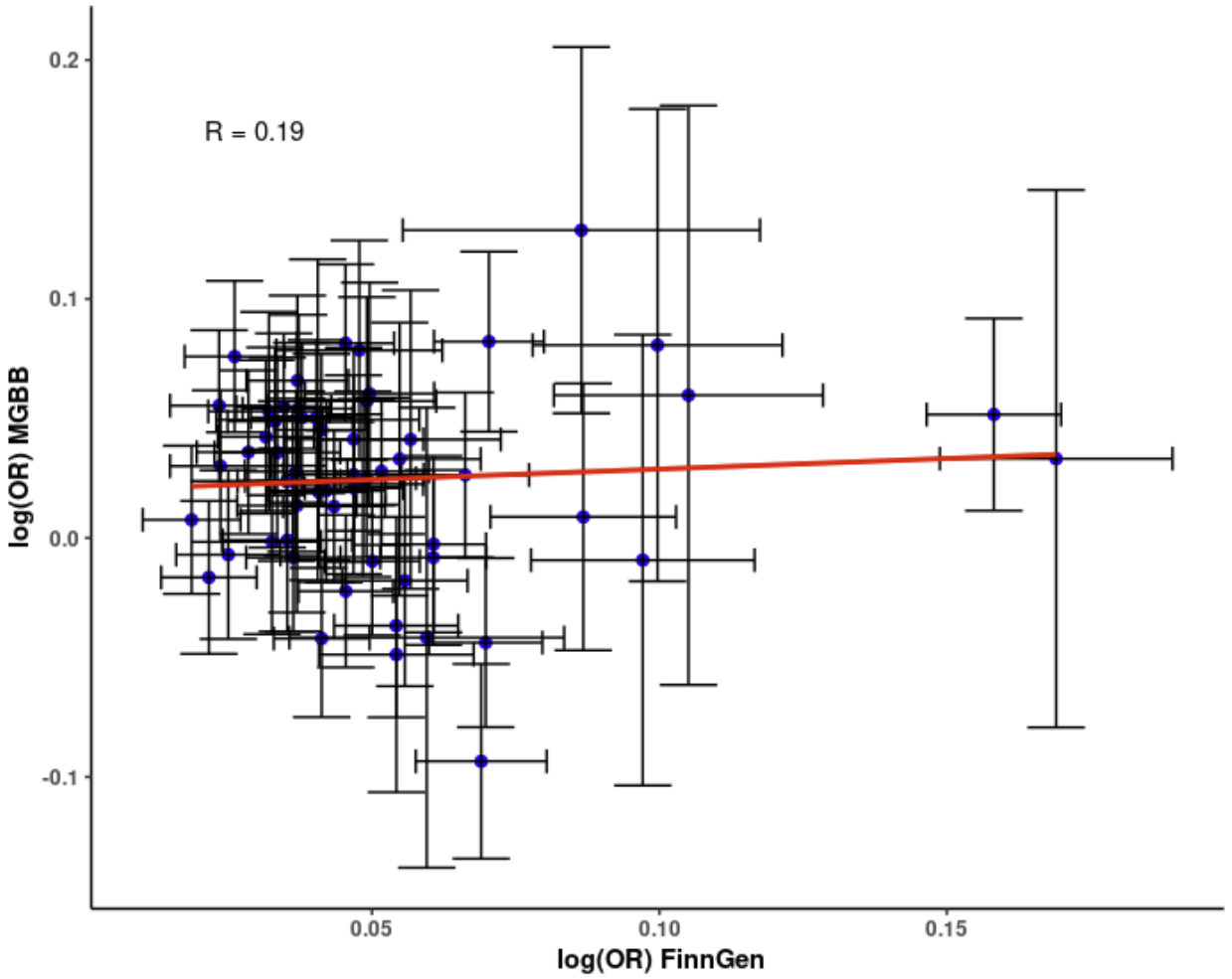
Supplementary Figure 2. Funnel plot based on the MR analysis of the association between cataracts and type 1 diabetes.



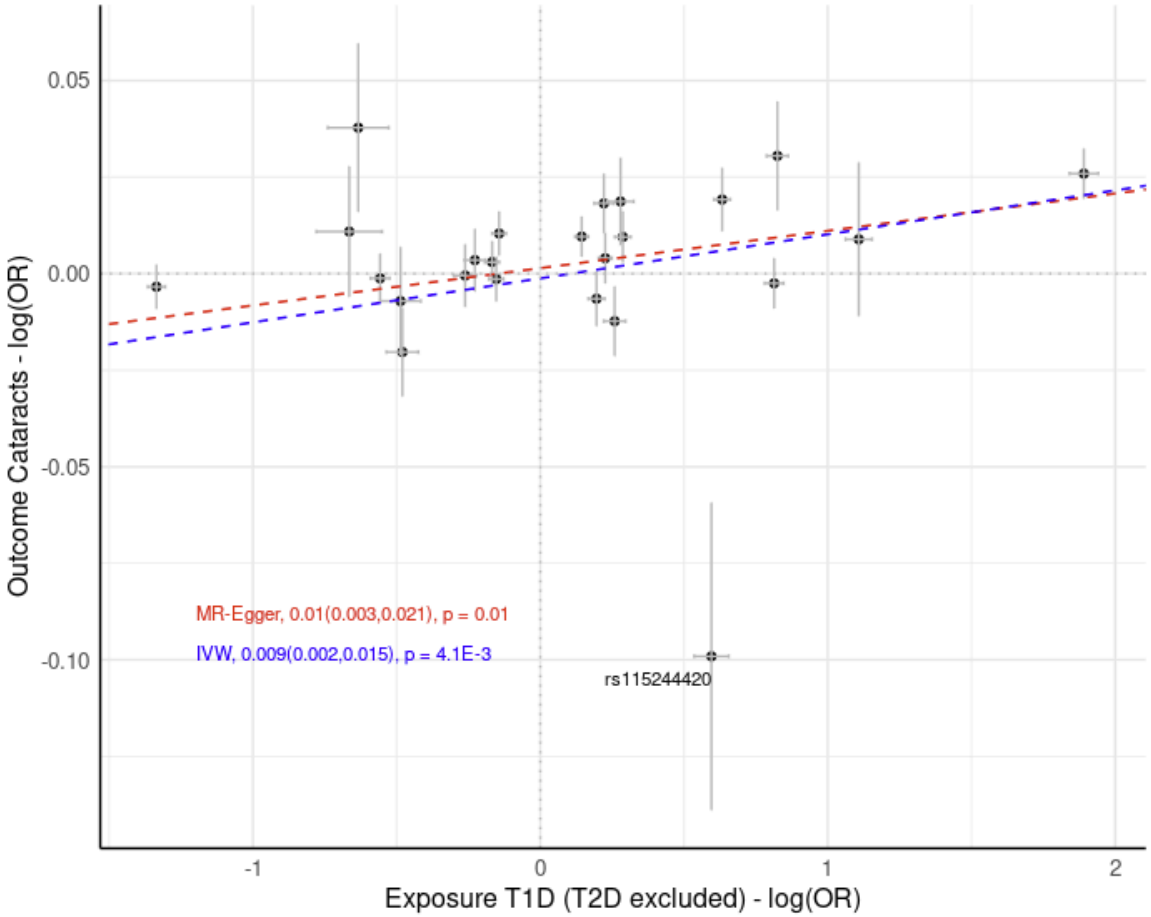
Supplementary Figure 3. Correlation of allele effect estimates, log(OR), between FinnGen and the previous cataract GWAS metanalysis that includes GERA and UK Biobank, using independent and genome-wide significant loci (N=101) and weighting variants according to the square standard error of their effect estimates. Alleles that maintained an increased effect for the larger study (GERA & UKB) were selected and FinnGen alleles were aligned to match the same allele. Effects estimates (center points) are presented as logarithms of odd ratios (**log**(OR)) and black crosses represent a 95% confidence interval.



Supplementary Figure 4. Correlation of allele effect estimates, log(OR), between the Mass General Brigham Biobank (MGBB) and the previous cataract GWAS metanalysis that includes GERA and UK Biobank, using independent and genome-wide significant loci (N=101) and weighting variants according to the square standard error of their effect estimates. Alleles that maintained an increased effect for the larger study (GERA & UKB) were selected and MGBB alleles were aligned to match the same allele. Given the smaller sample size of MGBB (approximately 3k cases) in comparison with the previous meta-analysis (approximately 68k cases), larger confidence intervals and lower statistical power were expected. Nevertheless, the correlation still maintains a consistent positive correlation of effects between studies. Effects estimates (center points) are presented as logarithms of odd ratios (**log**(OR)) and black crosses represent a 95% confidence interval.



Supplementary Figure 5. Correlation of allele effect estimates, log(OR), between the Mass General Brigham Biobank (MGBB) and FinnGen, using independent and genome-wide significant loci (N=101) and weighting variants according to the square standard error of their effect estimates. Alleles that maintained an increased effect for the larger study (FinnGen) were selected and MGBB alleles were aligned to match the same allele. Given the smaller sample size of MGBB (approximately 3k cases) in comparison with FinnGen (approximately 51k cases), larger confidence intervals and lower statistical power were expected. Nevertheless, the correlation still maintains a consistent positive correlation of effects between studies. Effects estimates (center points) are presented as logarithms of odd ratios (**log**(OR)) and black crosses represent a 95% confidence interval.



Supplementary Figure 6. Effect of variants (N=24) associated with Cataracts (Outcome) and type 1 Diabetes (Exposure). All the variants that have been associated with type 2 Diabetes were excluded. Effects are presented as logarithms of odd ratios (log(OR)) and grey crosses represent a 95% confidence interval. The dashed blue line shows the inverse variance weighted (IVW) fit and the red dashed line shows the MR-Egger-fit. Effects estimates (center points) are presented as logarithms of odd ratios (**log**(OR)) and grey crosses represent a 95% confidence interval.