**Option 1**

Aims

* Identify GxG effects on biomarker concentration to inform future drug development and identify subgroup effects for existing therapies
* Improve prediction of biomarker concentration using interaction terms

Results

* vGWAS QC
* GxG testing to identify interaction effects
* Association of GxG effects with disease (MI, heart failure, CKD, LD, T2D, Gout, )
* For disease associated GxG perform colocalization with eQTL data
* Replication of disease interaction effects (? independent cohort) + sibs
* Improvements in variance explained by use of gxg variants
* GxE – colocalize main effect on gene with eQTL data. MVMR to estimate GxE effect?
* Describe the occurrence of vQTLs at lipid drug target loci and compare results with RCT evidence
* Colocalize lipid vQTLs with e/pQTL data and discuss findings in relation to drug development

Discussion

* Utilities in drug target assessment
* Comparison with OSCA
  + OSCA is better powered with non-normality than B-P but has higher T1E
  + Effect size cannot be described by linear effect & cannot produce an effect with P=0
  + OSCA requires rounded dosage values and preadjusted phenotypes
  + CPU performance?

Figure. Manhattan plots of biomarker GWAS using Breusch-Pagan test

Figure. Gene-by-gene interaction analyses of biomarker vQTLs

Figure. Gene-by-environment interaction analyses of biomarker vQTLs

Supplementary Figure. Power to detect SNP-interaction effects using heteroscedasticity testing under simulation



Φ, Interaction effect size relative to main effect. Inflation factor, sample size relative to the size required to detect the main effect with 80% power. Normal, distribution with mean of 0 and variance of 1. Lognormal, distribution with mean of 0 and variance of 1. T-dist, distribution with 4 degrees of freedom. OSCA, OmicS-data-based Complex trait Analysis. SNP, single-nucleotide polymorphism simulated with minor allele frequency of 0.4 in Hardy-Weinberg equilibrium. All simulations had a fixed main effect detectable with 80% power when the sample size inflation factor was equal to 1. Simulation was produced with 200 repetitions. Sample size inflation factor of 1 was set to 200 observations. Error bars represent the 95% confidence interval.

Supplementary Figure. Type I error of Breusch-Pagan test in the presence of outliers and low minor allele frequency

Normal, distribution with mean of 0 and variance of 1. Lognormal, distribution with mean of 0 and variance of 1. T-dist, distribution with 4 degrees of freedom. Mixed normal, distribution produced with 90% N(0,1) and 10% N(5,1). MAF, minor allele frequency. A, MAF=0.01. B, MAF=0.05. C, MAF=0.1. D, MAF=0.2. Simulation was produced with 1000 repetitions and 1000 observations. Dashes represent the 95% confidence interval.

Supplementary Figure. Type I error of OSCA test in the presence of outliers and low minor allele frequency

Normal, distribution with mean of 0 and variance of 1. Lognormal, distribution with mean of 0 and variance of 1. T-dist, distribution with 4 degrees of freedom. Mixed normal, distribution produced with 90% N(0,1) and 10% N(5,1). MAF, minor allele frequency. A, MAF=0.01. B, MAF=0.05. C, MAF=0.1. D, MAF=0.2. Simulation was produced with 1000 repetitions and 1000 observations. Dashes represent the 95% confidence interval.

Supplementary Figure. Type I error of Breusch-Pagan test with main effect following variable transformation

Normal, distribution with mean of 0 and variance of 1. Lognormal, distribution with mean of 0 and variance of 1. T-dist, distribution with 4 degrees of freedom. Mixed normal, distribution produced with 90% N(0,1) and 10% N(5,1). A, Log transformation. B, Square root. C, Inverse-rank normal transformation. D, Cube root. Simulations were produced with 1000 repetitions and 1000 observations; all experiments were set to have a main effect detectable with 80% power. The SNP was simulated to have 10% allele frequency. Dashes represent the 95% confidence interval.

Supplementary Figure. Type I error of OSCA test with main effect following variable transformation

Normal, distribution with mean of 0 and variance of 1. Lognormal, distribution with mean of 0 and variance of 1. T-dist, distribution with 4 degrees of freedom. Mixed normal, distribution produced with 90% N(0,1) and 10% N(5,1). A, Log transformation. B, Square root. C, Inverse-rank normal transformation. D, Cube root. Simulations were produced with 1000 repetitions and 1000 observations; all experiments were set to have a main effect detectable with 80% power. The SNP was simulated to have 10% allele frequency. Dashes represent the 95% confidence interval.

Supplementary Figure. Histogram of biomarker distribution

Diagram, engineering drawing

Description automatically generated

SD units.

Supplementary Figure. QQ plot of biomarker distribution

A picture containing background pattern

Description automatically generated

SD units.

Supplementary Figure. QQ plots of biomarker GWAS using Breusch-Pagan test