

Investigating the Utility of Pharmaceutically Informed, Partitioned Type-2 Diabetes Mellitus Polygenic Scores in the UK Biobank

Luke Marques, Prof. Cathryn M. Lewis, Prof. Gerome Breen, Dr. Jonathan Coleman

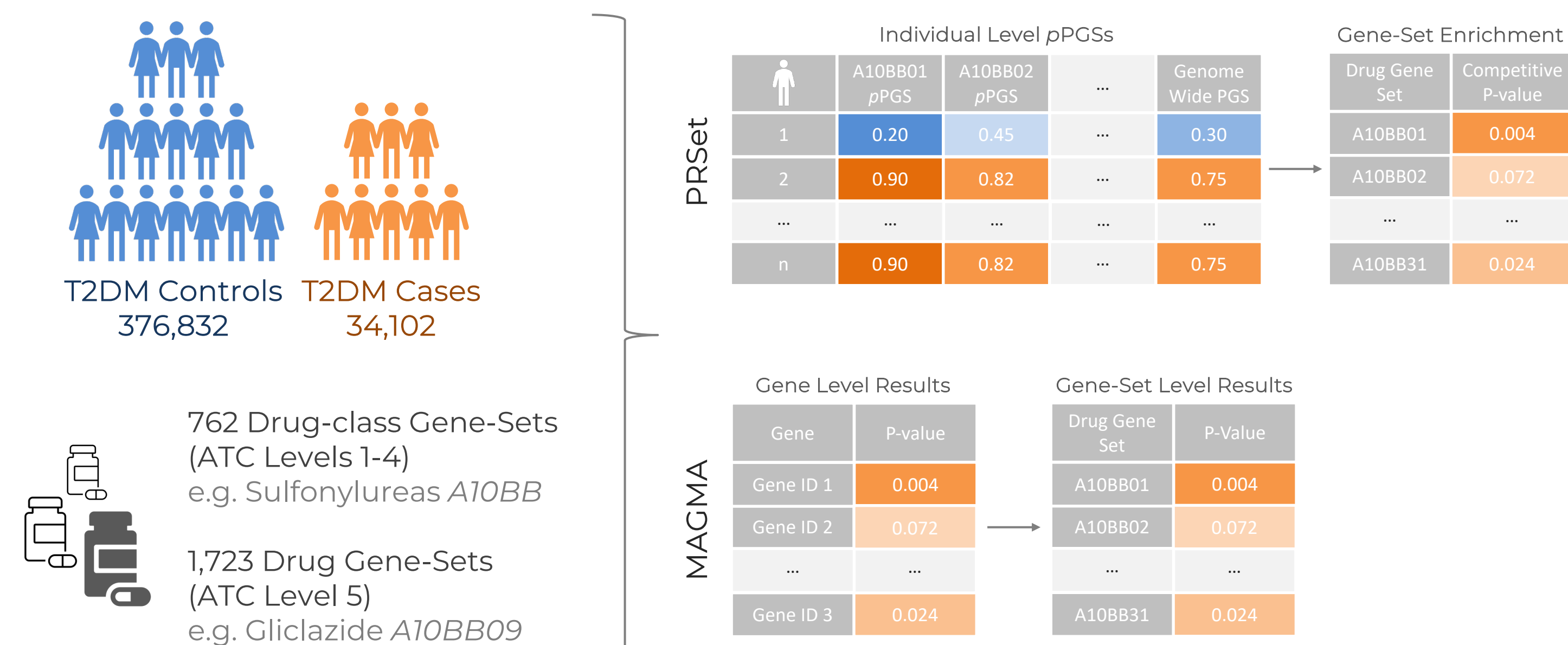
Introduction

Inspiration

- Cost of developing novel pharmaceuticals is high, driven by high failure rates.
- MAGMA gene-set analysis: opportunities for drug repurposing in schizophrenia.^{1,2}
- MAGMA provides group level results, but no individual level information.

Goals

- Assess the utility of partitioned polygenic scores (pPGSs), generated using PRSet³, for identifying enriched target-indication pairs, for Type-2 Diabetes Mellitus (T2DM).
- Compare results to existing pathway analysis methods.
- Suggest future directions for pharmaceutically informed pPGSs.



Methodology

UK Biobank (UKBB)

- Over 500,000 individuals, aged 40-69 at recruitment (2006 – 2010).
- Standard sample and SNP QC, removed relatives, filtered for European descent.
- Identified cases (n=34,102) and controls (n=376,832) for T2DM, removing diagnoses at ages under 35 to avoid Type-1 diabetes contamination.

Partitioned Polygenic Scores (pPGS)

- Generated pPGSs for T2DM using PRSet, and extension of PRSice-2.
- pPGSs based on 762 and 1,723, previously developed, drug-class and drug gene-sets.²
- Drug gene-sets mapped to WHO Anatomical Therapeutic Chemical (ATC) codes.
- Competitive p-value of association between pPGSs and T2DM status obtained via PRSet's competitive enrichment testing.
- Competitive p-values bounded by the permutations (M = 10,000) performed comparing pPGSs against randomized scores from background set.

$$P_{Competitive} = \frac{\sum_{m=1}^M I(P_{Null} < P_{Observed}) + 1}{M+1}$$

MAGMA Pathway Analysis

- MAGMA's competitive pathway analysis used as comparator to PRSet's enrichment.¹
- Summary statistics for T2DM in UKBB, used by MAGMA, computed using Regenie.⁴
- Drug-class analysis differs from PRSet – tests compare drug-gene sets within one ATC class to those not present in that class.
- Provides additional gene-level results showing GWAS signal captured by each gene.

Results

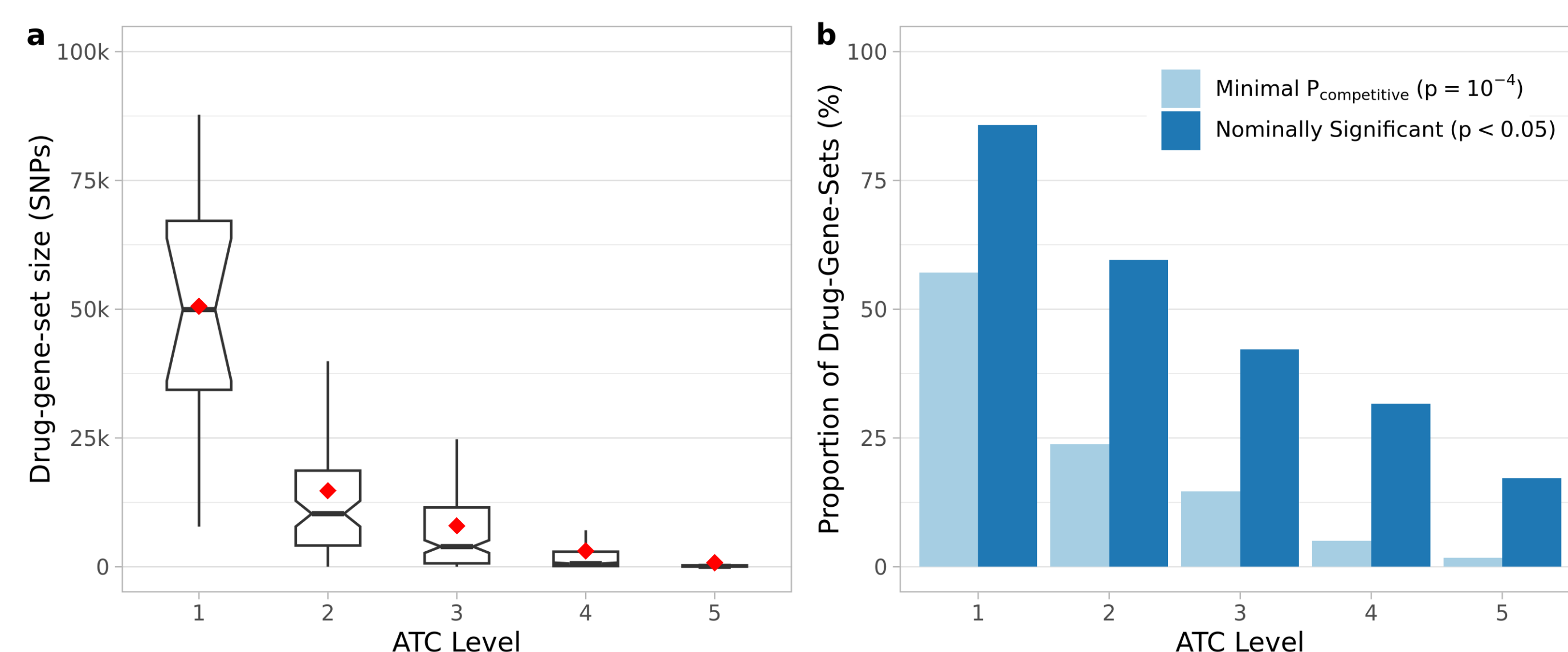


Figure 1. (a) Distributions of gene-set size across ATC levels. (b) Proportion of gene sets showing nominal or minimal competitive p-values from PRSet's competitive analysis.

- PRSet: No gene-sets reached statistical significance for enrichment due to bounded competitive p-value.
- 108 gene-sets reached minimal competitive p-value achievable ($p = 10^{-4}$).
- Inflation of competitive p-values for broader sets observed, despite PRSet's control for set size differences in its competitive testing methodology.
- Inflation could be due to increased residual LD in larger sets compared to the null set, as genes in pathways likely to have non-random gene proximities.

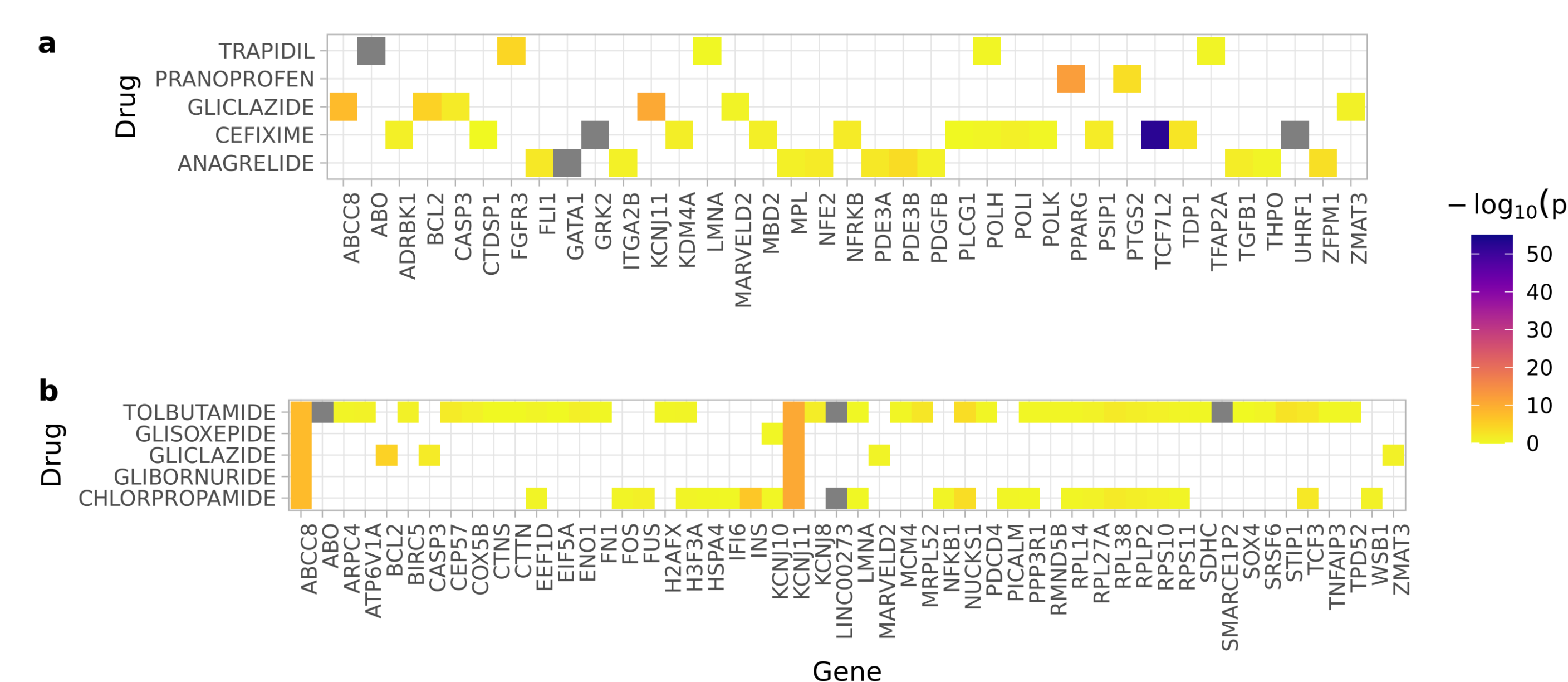


Figure 3. MAGMA gene-level association results for (a) enriched drug gene-sets and (b) drug gene-sets belonging to the sulfonylurea drug-class. Grey genes were missing from MAGMA analyses.

- MAGMA gene-level results highlight key genes driving pathway enrichment.
 - Gliclazide's enrichment is mainly due to two genes: ABBC8 and KCNJ11, which encode subunits of the sulfonylurea 1 receptor.⁵
 - Variants in these genes are linked to increased sensitivity to sulfonylureas.⁷
- Sulfonylurea drug sets also derive enrichment from ABBC8 and KCNJ11 primarily.
- Pranoprofen primarily driven by PPARG: encodes the PPARY transcription factor.
 - Increased expression is associated with increased risk of T2DM development.⁸
 - Main target of T2DM treatment drug-class thiazolidinediones.

Conclusion

- PRSet competitive testing for enriched drug/drug-class gene-sets not supported by existing methods.
- Due to methodology of PRSet's competitive testing, PRSet may not be suitable for high-throughput enrichment analyses for discovery purposes.
- MAGMA gene level results point toward the clinical relevance for these gene-sets, and pharmaceutical pPGSs demonstrate an ability to capture genetic heterogeneity.
- No gene effect annotation in drug sets limits ability to infer directionality from pPGSs.
- Use of pharmaceutically informed pPGSs in further stratifying clinically relevant outcomes, or for improving on the predictive performance of the T2DM genome-wide PGS's for treatment response⁹ remains for further investigation.

Acknowledgements & Conflicts of Interests

Research funded by The Medical Research Council Doctoral Training Partnership (MRC DTP) iCASE Programme and the PhD industrial partner UCB.

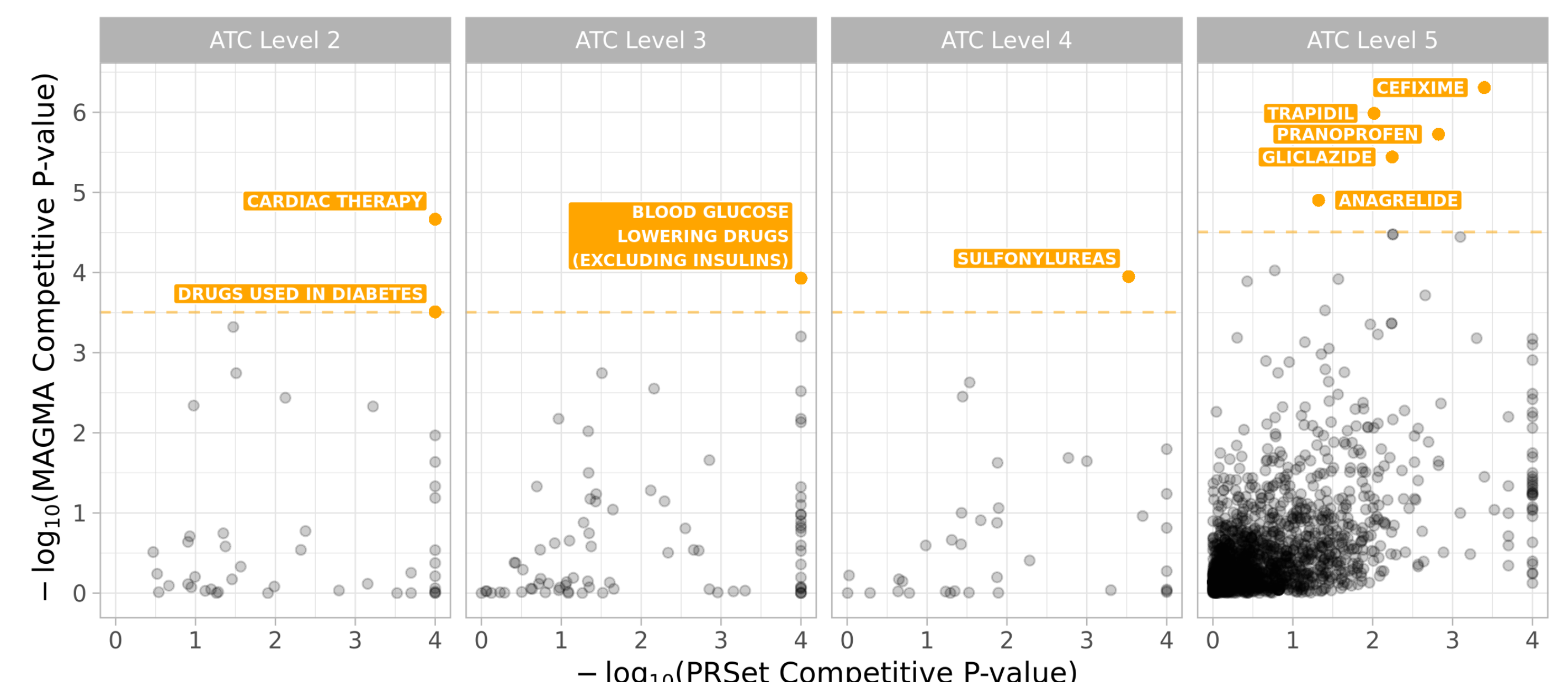


Figure 2. Comparison of PRSet and MAGMA competitive pathway analysis for drug/drug-class gene-sets. Bonferroni corrected significance thresholds are marked by red dashed lines.

- MAGMA's pathway analysis demonstrated greater specificity.
- 5 drug and 3 drug-class gene-sets with statistically significant enrichment for T2DM, including Sulfonylureas at ATC Level 4.
- One approved T2DM treatment gene-set enriched: Gliclazide, a sulfonylurea.
- Sulfonylurea drug class result is supported by existing literature, identifying insulin secretagogues as enriched for T2DM signal through similar methods.⁵
- Enriched sets from MAGMA do not achieve minimal p-value from PRSet, suggesting PRSet not supported by existing methods, in this context.

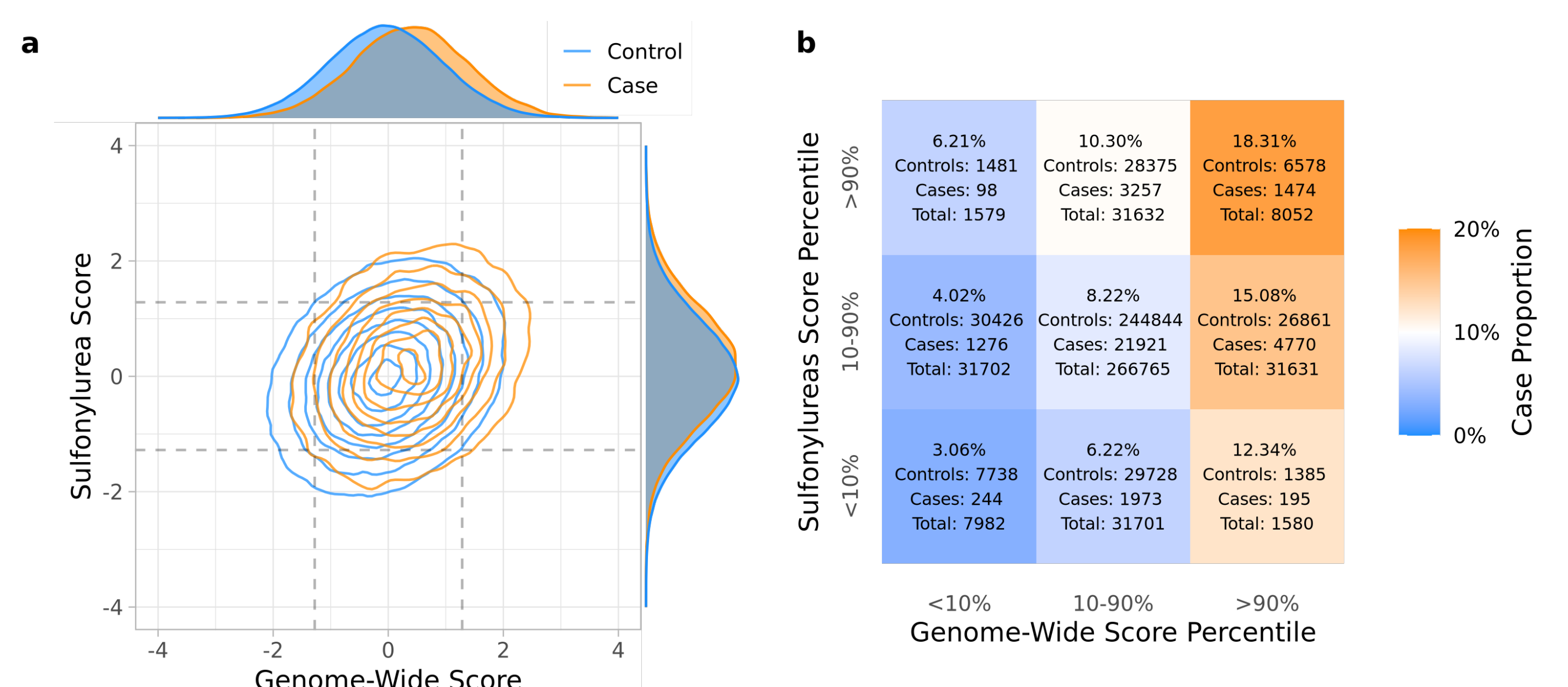


Figure 4. (a) Sulfonylurea and genome-wide T2DM PGS's distributions and (b) T2DM prevalence across sulfonylurea partitioned and genome-wide PGS (Z-scores) deciles.

- Some individuals exhibit uneven distribution of risk alleles across the sets, and when compared to the T2DM genome-wide PGS.
- Individuals in the top decile for both the sulfonylurea pPGS and genome-wide PGS were of higher risk for T2DM than those only in the top decile of the genome-wide score.
- Such pPGSs could be used to further stratify individuals with regards to T2DM clinically relevant outcomes, such as treatment response.

