

The American Journal of Human Genetics, Volume 107

Supplemental Data

**Cancer PRSweb: An Online Repository
with Polygenic Risk Scores for Major Cancer Traits
and Their Evaluation in Two Independent Biobanks**

Lars G. Fritsche, Snehal Patil, Lauren J. Beesley, Peter VandeHaar, Maxwell Salvatore, Ying Ma, Robert B. Peng, Daniel Taliun, Xiang Zhou, and Bhramar Mukherjee

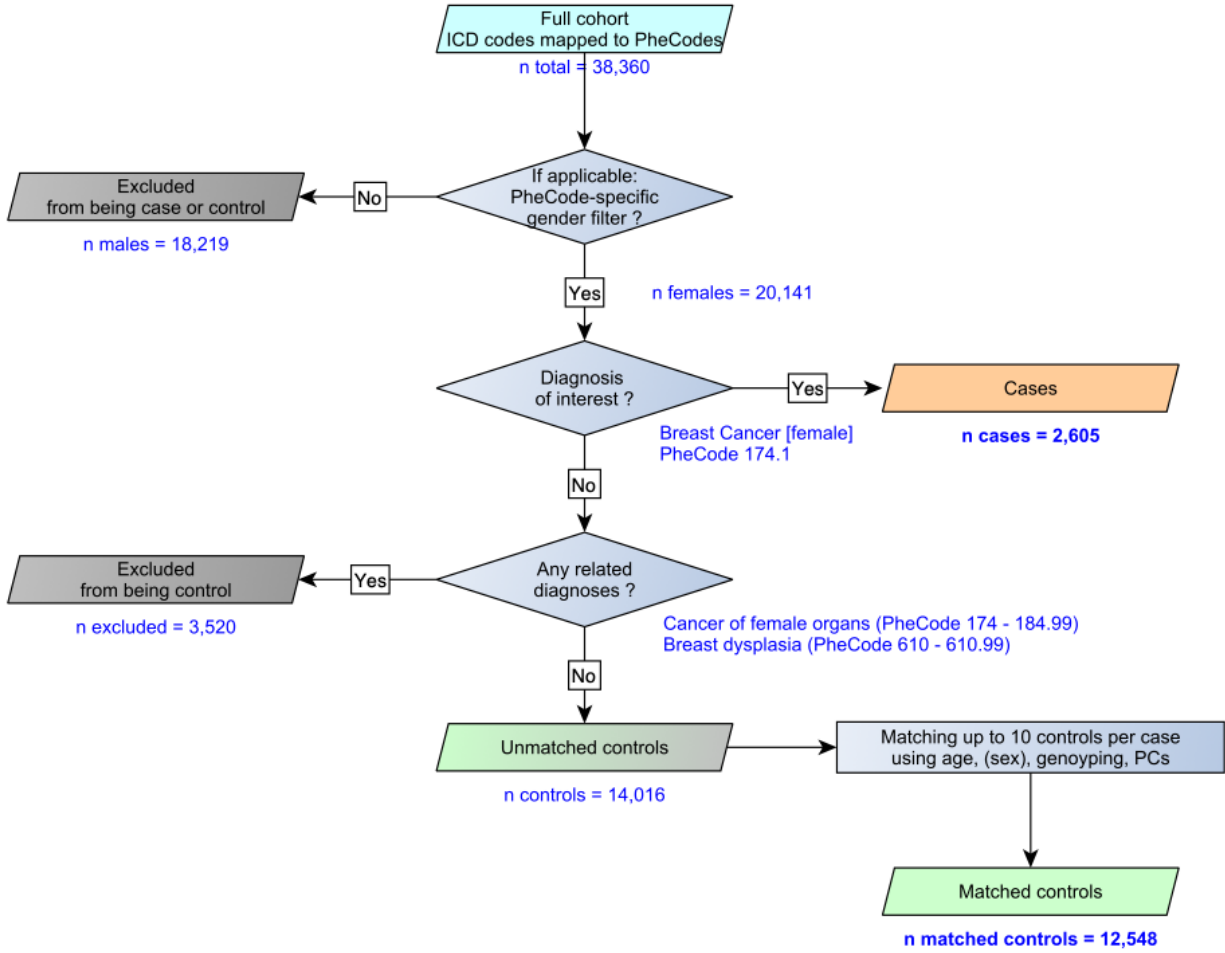


Figure S1 Generation of matched case control studies using PheCode-based phenomes. The example for “Breast Cancer [female]” (PheCode 174.1) of the MGI cohort is shown in blue.

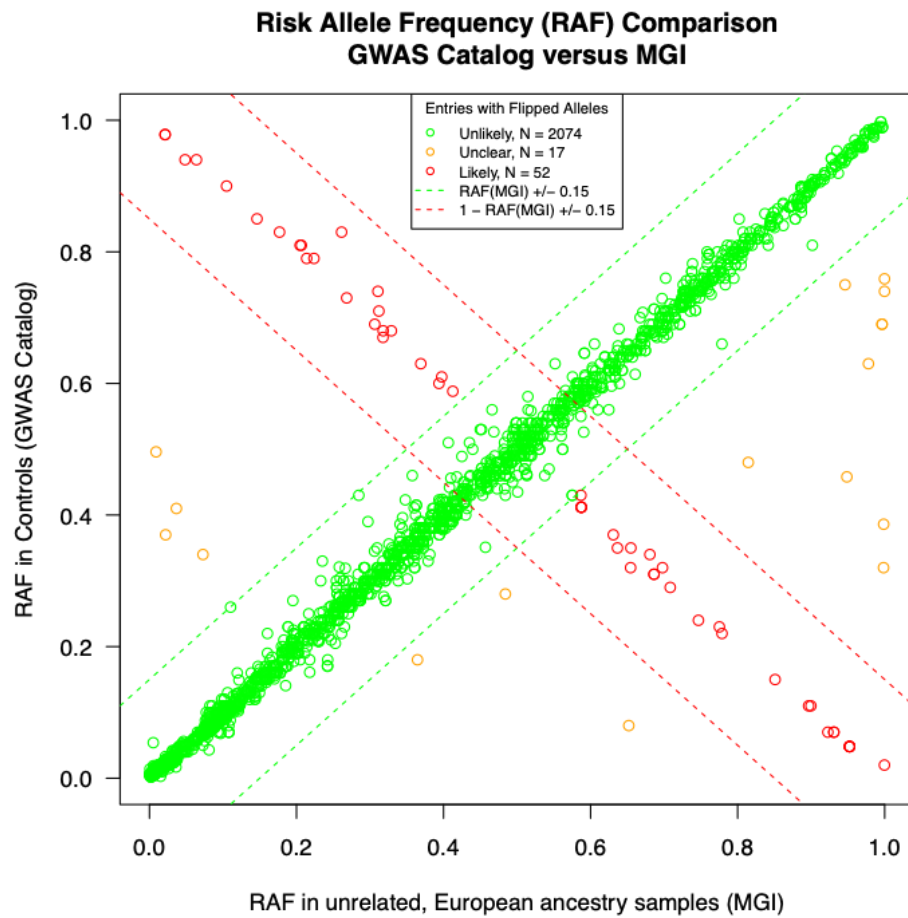


Figure S2. Comparison of the Risk Allele Frequencies in the GWAS Catalog vs. MGI. Each frequency comparison is coded as unlikely (green, $n = 2074$), unclear (orange, $n = 17$), or likely (red, $n = 52$) allele flip.

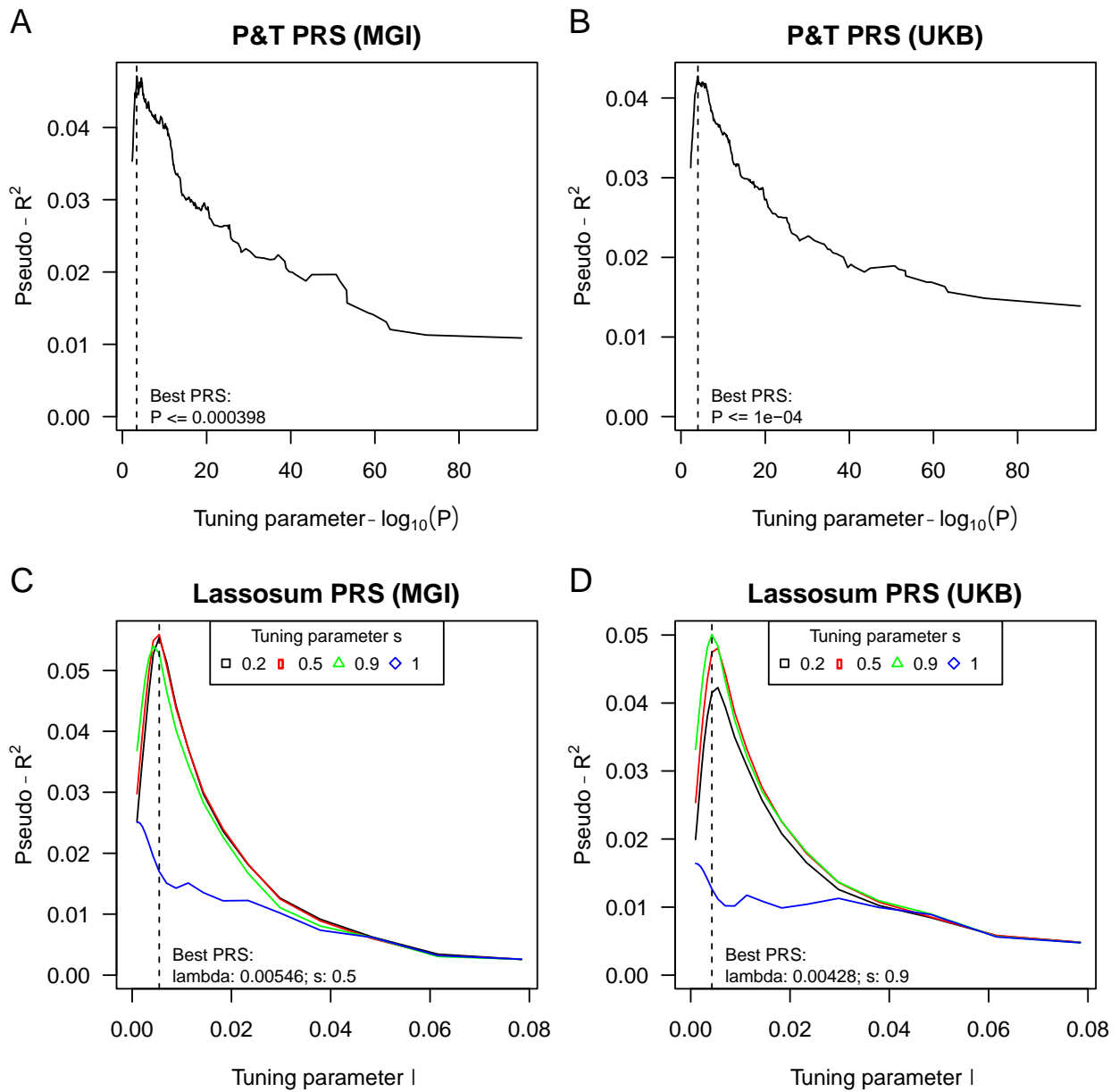


Figure S3. Tuning Parameter Optimization. Tuning parameter optimization for the large GWAS based breast cancer PRS with Lassosum (A & B) and “P&T” approach (C & D) for MGI (A & C) and UKB (B & D).

GWAS Hits versus P&T

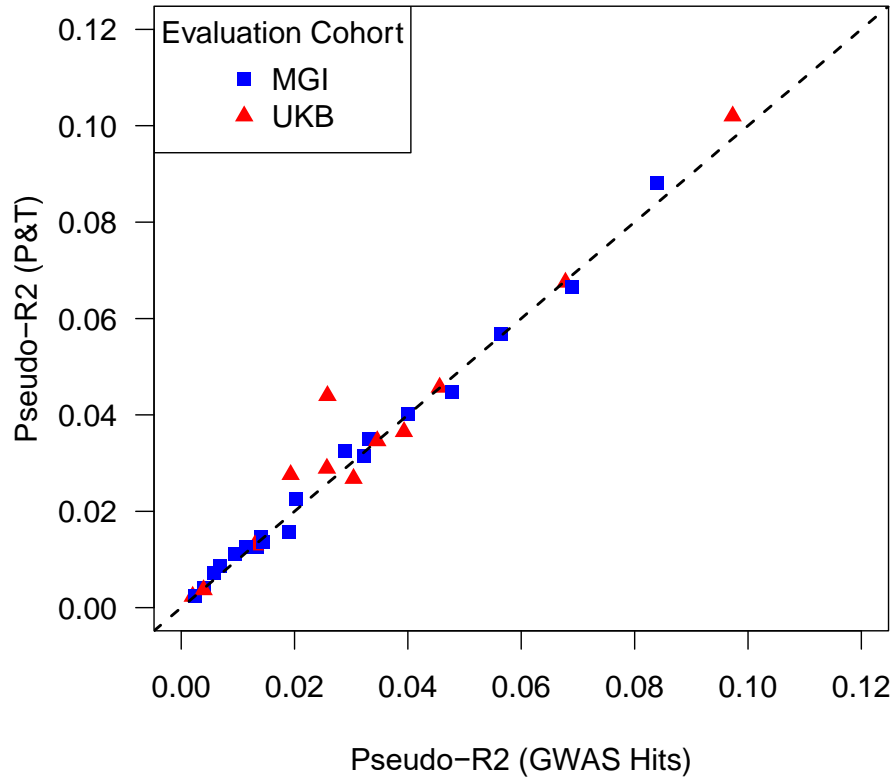


Figure S4. Performance comparison between P&T and GWAS Catalog hits-based PRS. Pairwise comparison of the two PRS methods P&T and “GWAS hits” ($P \leq 5 \times 10^{-8}$) using GWAS Catalog entries as input. Pseudo- R^2 values of 31 PRS for 21 cancer traits (19 MGI PRS and 12 UKB PRS) are shown. Dashed line: identity line.

Comparison PRS Methods P&T versus Lassosum

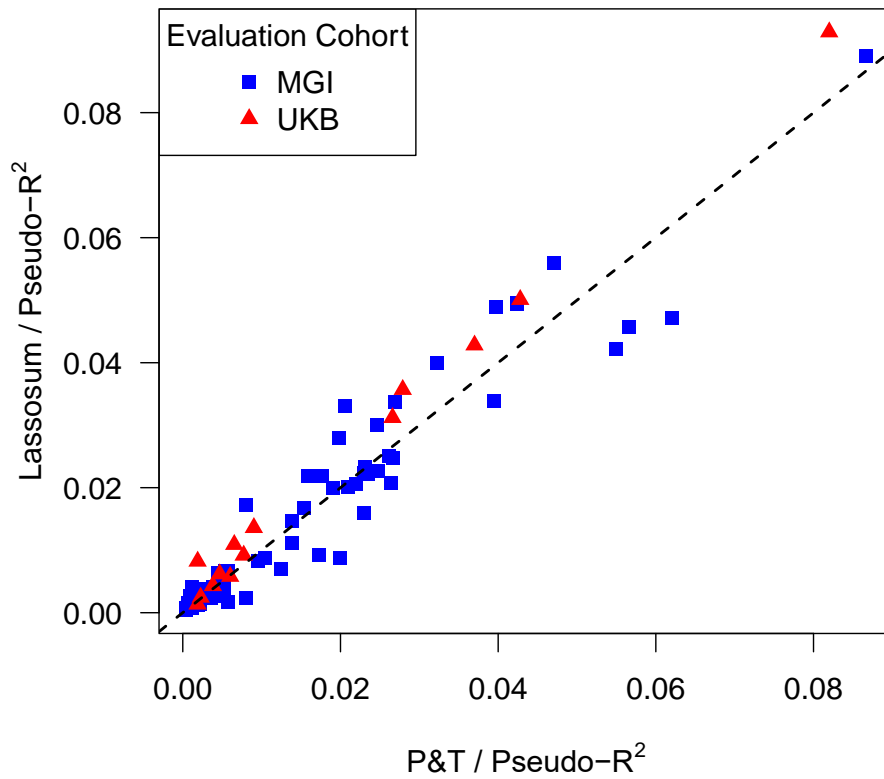


Figure S5. PRS performance comparison “P&T versus Lassosum”. Pairwise comparison of the two PRS methods P&T and Lassosum using pseudo-R². 47 GWAS sources where P&T and Lassosum-based PRS were positively and nominally significant associated with their cancer trait in MGI (blue; 37 PRS) and UKB (red; 10 PRS) are shown. Dashed line: identity line.

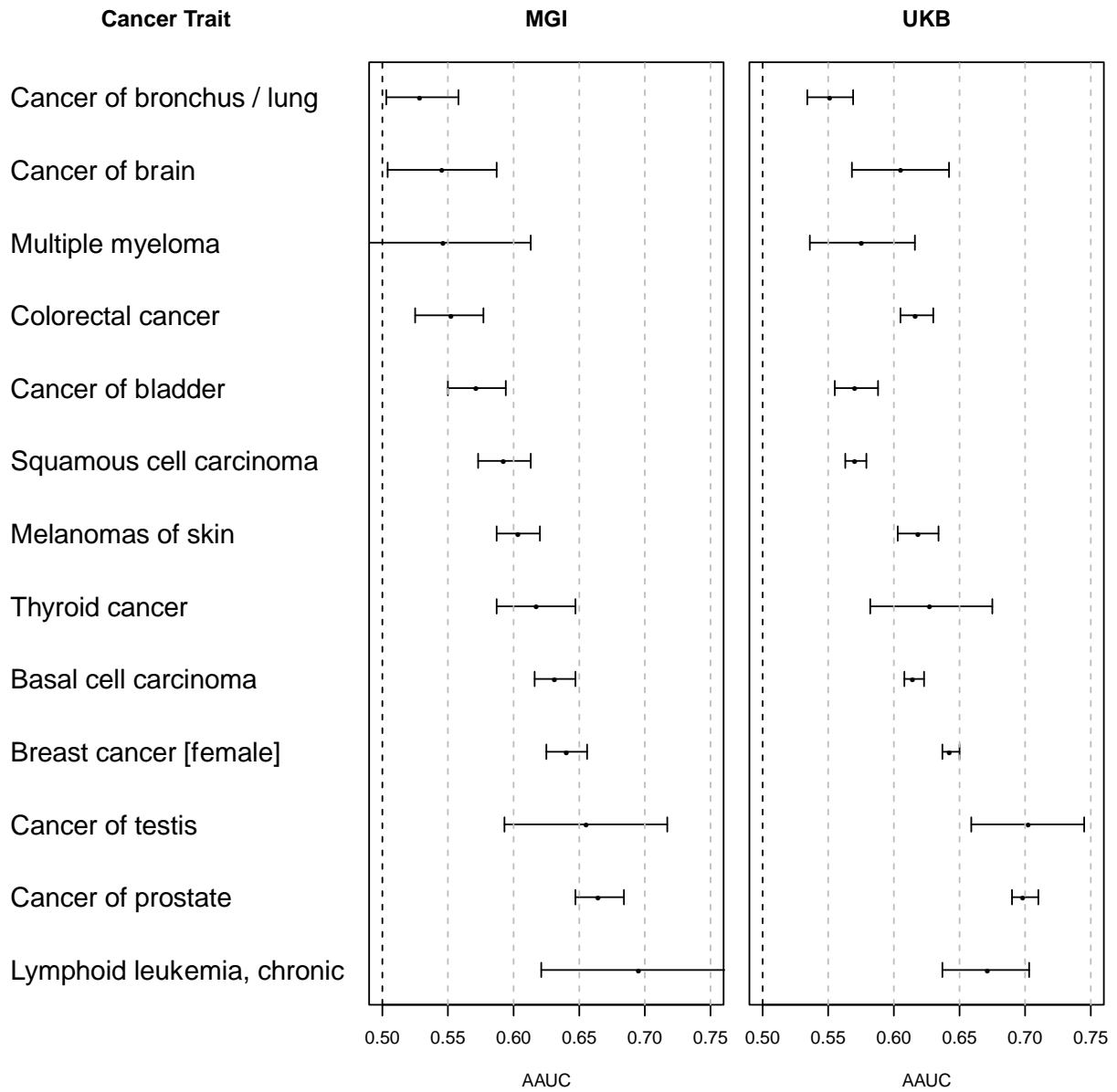


Figure S6. AAUC comparison of PRS between MGI and UKB. AAUC values (dots) and their 95% confidence intervals of PRS for 13 cancers that were present for MGI (left) and UKB (right) are shown.

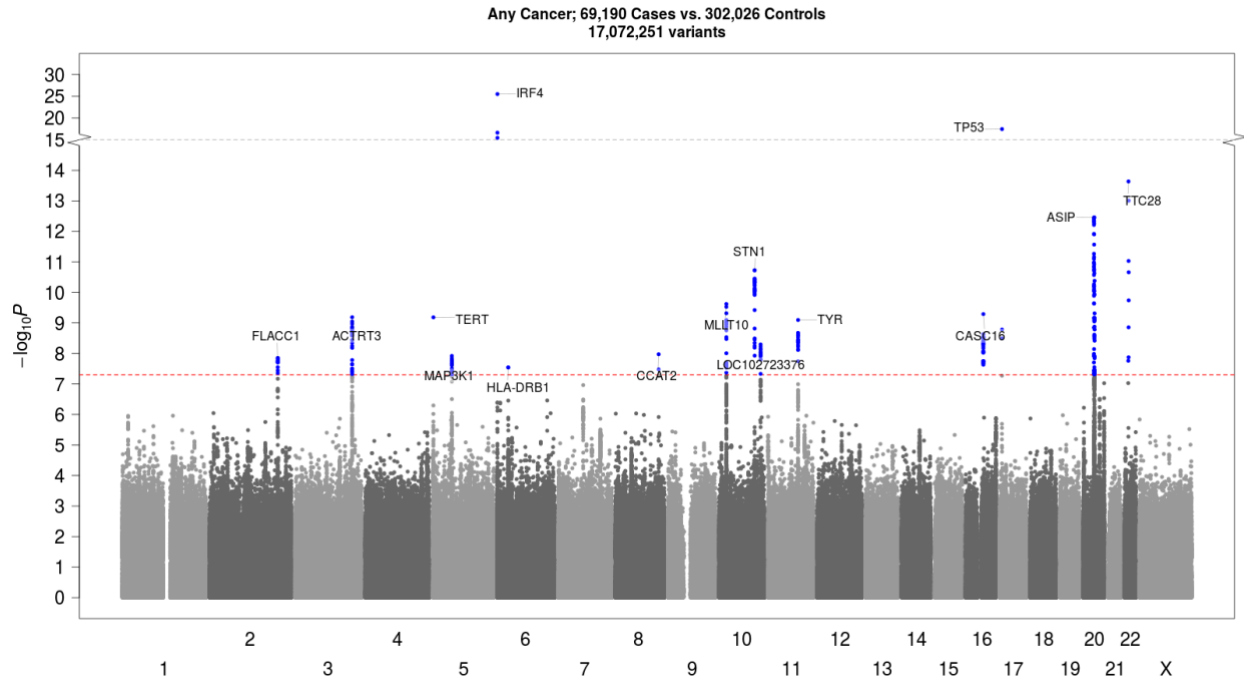


Figure S7. Manhattan plot of the UKB GWAS on 69,190 cases with any cancer versus 302,026 controls. SNPs with $P < 5 \times 10^{-8}$ are highlighted in blue. Candidate loci are named after the nearest gene closest to the strongest signal.

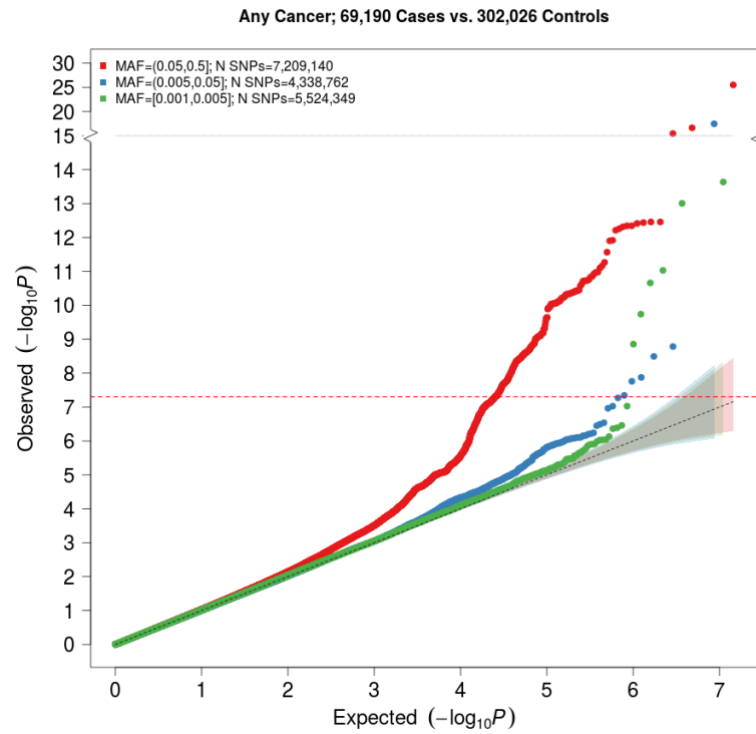


Figure S8. QQ plot of the UKB GWAS on 69,190 cases with any cancer versus 302,026 controls.

Negative log₁₀(P-values) stratified by minor allele frequency (MAF) bins are shown.

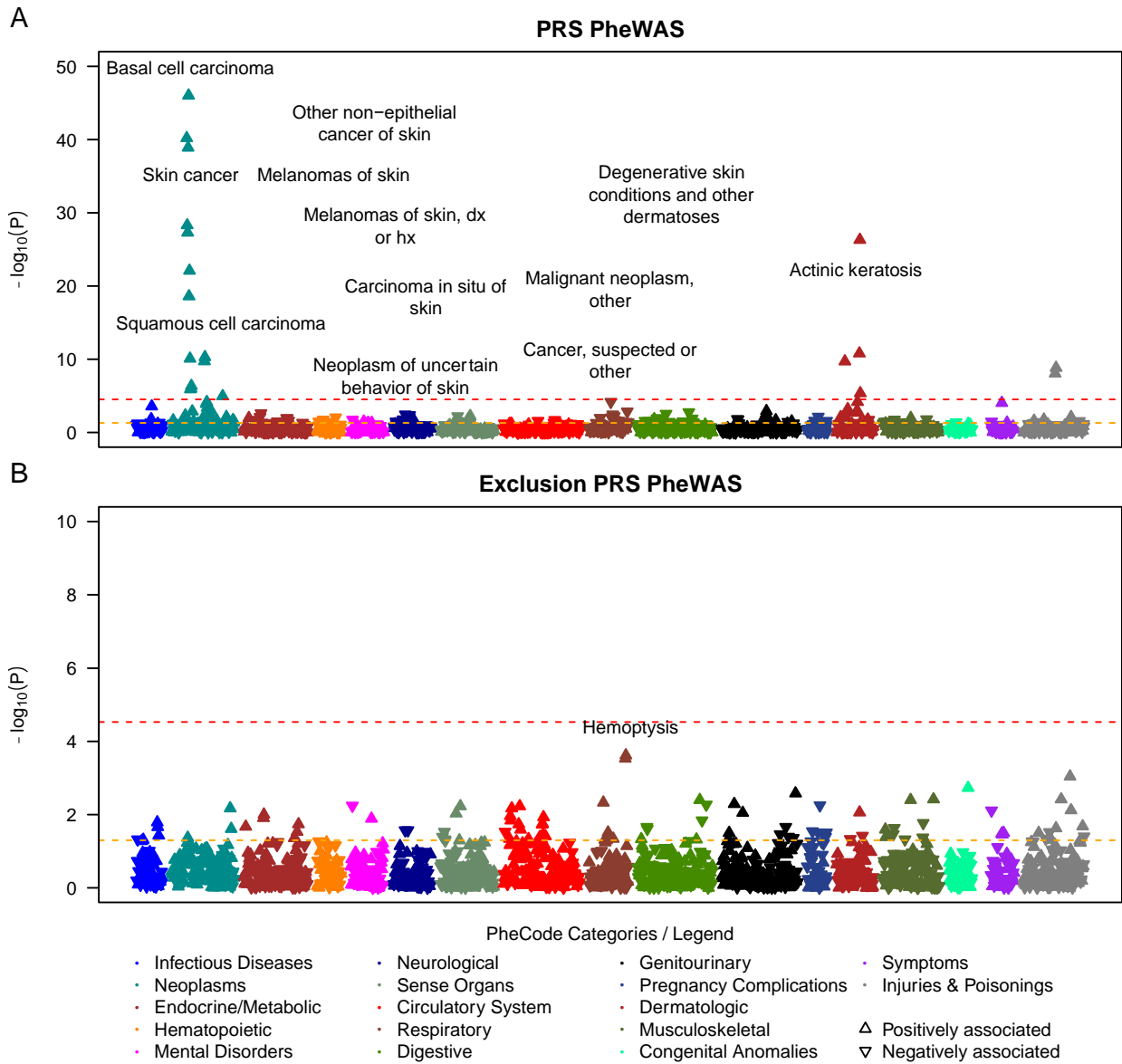


Figure S9. PRS PheWAS plot of the ‘Any Cancer’ lassosum PRS. PRS PheWAS results in MGI before (A) and after (B) excluding 20,751 MGI individuals with ‘any cancer’ are shown.

Supplemental Acknowledgements

Breast Cancer Association Consortium (BCAC)

The breast cancer genome-wide association analyses were supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, the 'Ministère de l'Économie, de la Science et de l'Innovation du Québec' through Genome Québec and grant PSR-SIIRI-701, The National Institutes of Health (U19 CA148065, X01HG007492), Cancer Research UK (C1287/A10118, C1287/A16563, C1287/A10710) and The European Union (HEALTH-F2-2009-223175 and H2020 633784 and 634935). All studies and funders are listed in Michailidou et al (2017).

Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL)

The Prostate cancer genome-wide association analyses are supported by the Canadian Institutes of Health Research, European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175), Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A16563, C5047/A3354, C5047/A10692, C16913/A6135, and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative).

Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) [U19 CA 148537 for ELucidating Loci Involved in Prostate cancer SuscEptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008I] and by Cancer Research UK grant A8197/A16565. Additional analytic support was provided by NIH NCI U01 CA188392 (PI: Schumacher).

We would also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now PCUK), The Orchid Cancer Appeal, Rosetrees Trust, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust.