NORDIC SOCIETY

OF HUMAN GENETICS AND PRECISION MEDICINE

ABSTRACT BOOK 2022

"PRECISION MEDICINE RESEARCH AND IMPLEMENTATION: REBOOTING IN THE NORDICS POST-COVID"

3-4 NOVEMBER · COPENHAGEN · DENMARK

NSHG-PM

Index

P20	Genome Wide Association Study of Clinically Predicted Suicide Liability	Anders	Jespersen
P21	Genome-Wide Association Study Identify New Genetic Determinants of Resting Heart Rate	Marie	Klevjer
P22	Global Long COVID Host Genetics Iniative Identifies FOXP4 locus as the First Genetic Risk Factor Associated to Long COVID	Minttu	Marttila
P23	Heterogeneity of Unknown Auto-Inflammatory Diseases	Sara	Garcia
P24	HLA Associations Between Epstein-Barr Virus and Multiple Sclerosis	Jesse	Huang
P25	Identification of Novel Genomic Risk Loci Shared Between Common Epilepsies and Major Psychiatric Disorders	Naz	Karadag
P26	Identifying Causal Genetic Variants in Diabetes- and Adiposity-Associated Genes Using CRISPR and Isogenic Cell Lines	Sebastian	Kalamajski
P27	Implementation of Polygenic Risk Score Guided Breast Cancer Prediction in Norwegian Data	Eivind	Hovig
P28	Improved Rare and Low-Frequency Variant Imputation with the Novel SISu 4.2 Finnish Population-Specific Genotype Imputation Reference Panel.	Kalle	Pärn
P29	Mendelian Randomization Study of Birth Weight and Risk of Psychiatric Disorders Later in Life	Xiaoping	Wu
P30	Multi-Omic Profiling of Spontaneous Migraine Attacks Treated with a Triptan	Lisette	Kogelman
P31	Multi-PGS Enhances Polygenic Prediction: Weighting 937 Polygenic Scores	Clara	Albiñana
P32	Parental Transmitted and Non-Transmitted Alleles Recapitulate Known Biology of Neonatal Jaundice	Pol	Sole-Navais
P33	Patterns of Genetic Overlap Across Mental Disorders and Braits Beyond Genetic Correlation	Guy	Hindley
P34	Polygenic Health Index, General Health, and Pleiotropy: Sibling Analysis and Disease Risk Reduction	Laurent	Tellier
P35	Population-Level Study of CNVs and Their Associated Risk of Psychiatric Disorders in a Danish Case-Cohort	Morteza	Vaez
P36	Publicly Available Privacy-preserving Benchmarks for Polygenic Prediction	Menno	Witteveen
P37	Quantifying The Assortative Mating of Psychiatric Disorders and Its Consequences in Denmark	Saeid	Rasekhi Dehkordi
P38	Swiss National SARS-CoV-2 Genome and Variant Surveillance Program	Marc	Friedli
P39	The Role of Depression and its Treatment in Antihypertensive Medication Adherence and Persistence	Hanna	Kariis
P40	Time-Varying Effects are Common in Genetic Control of Gestational Duration	Karin	Ytterberg
P41	Ultralow-cost Genomic Population Screening	Ulrik	Stoltze

[1] A LARGE META-ANALYSIS ACROSS FOUR NORDIC COUNTRIES REVEALS SEVERAL SIGNIFICANT GENETIC ASSOCIATIONS FOR EARLY ONSET MAJOR DEPRESSIVE DISORDER

John Shorter,[1], Joëlle Pasman,[2], Andreas Jangmo,[3], Siim Kurvits,[4], Joeri Meijsen,[5], Oleksandr Frei,[3], Bayram Akdeniz,[3], Ole Andreassen,[3], Patrick Sullivan,[6], Thomas Werge,[5], Kelli Lehto,[4], Martin Tesli,[3], Lu Yi,[2], Alfonso Buil,[7]

[1] Mental Health Centre Sct. Hans, Mental Health Services, Institute of Biological Psychiatry, Roskilde, Denmark, [2] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, [3] Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway, [4] Estonian Biobank, Institute of Genomics, Tartu University, [5] Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Mental Health Services Copenhagen, Roskilde, Denmark, [6] Departments of Genetics and Psychiatry, University of North Carolina at Chapel Hill, Nc, USA, [7] Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Mental Health Services Copenhagen, Roskilde, Denmark.

Background

Genetic predisposition contributes substantially to the risk for major depressive disorder (MDD). Based on familial data, it has been estimated that 30-40% of individual differences in MDD can be explained by genetic factors. Early onset depression was recently shown to be heritable, and has been associated with more severe comorbidities, including psychosis and suicidality. However, not much is known about what contributes to individual genetic variation of early onset MDD. Here, we use information from national healthcare registries in countries with ethnically homogeneous populations to study the genetic associations with early onset MDD.

Methods

We examined MDD and early onset MDD across four large Nordic (Denmark, Estonia, Norway and Sweden) cohorts, which included over 100,000 cases of MDD. The powerful national healthcare registries within these countries allows for highly detailed analysis on timing and trajectory of disease progression. Early onset cases were classified as individuals with a specialist treatment contact for MDD at or before age 25. We exclude cases with a lifetime occurrence of bipolar disorder, schizophrenia, and schizoaffective disorder. Controls are individuals with no lifetime history of MDD or a mood or psychotic disorder that are matched

based on country, gender, and age. To ensure a uniform analysis across cohorts, we employed the TRYGGVE2/CorMorMent pipeline that has packaged genome wide association studies (GWAS) and post-GWAS tools into singularity containers.

Results

MDD and early onset MDD have highly significant genetic correlations between the different Nordic cohorts. We also replicate previous findings from the UK biobank on heritability of early onset MDD in our meta-analysis. We observe at least three genomic regions that are associated with early onset MDD. MAGMA gene-based tests also implicate several genes with patterns of brain expression. Surprisingly, there is little gene overlap with the Psychiatric Genetics Consortium's study on MDD and early onset MDD in the Nordic cohorts.

Conclusions

This is the largest study on shared genetic features across the Nordic countries for early onset depression. The identification of genetic variants that are associated with early onset MDD, but not regular MDD may facilitate our understanding of the severe outcomes (psychosis and suicidality) that are co-occurring with early onset MDD. Future analyses on polygenic risk score comparisons and validation of gene candidates identified in the GWA study are currently underway.

Conflict of interest

None.

[2] A GENOME-WIDE CROSS-TRAIT ANALYSIS BETWEEN SEX HORMONE-BINDING GLOBULIN AND RHEUMATOID ARTHRITIS

Yuan Jiang,[1], Qianwen Liu,[1], Lars Aflredsson,[2], Lars Klareskog,[3], Ingrid Kockum,[1], Xia Jiang,[1]

[1]Karolinska Institutet, Department of Clinical Neuroscience, Solna, Stockholm, Sweden, [2] Karolinska Institutet, Department of Clinical Neuroscience, Institute of Environmental Medicine, Solna, Stockholm, Sweden, [3] Karolinska Institutet, Department of Medicine, Solna. Stockholm, Sweden

Objectives

Our study aims to investigate an intrinsic link underlying sex hormone-binding globulin (SHBG) and rheumatoid arthritis (RA), which remains inconclusive in observational settings.

Methods

Summary statistics were collected from the largest GWAS(s) on SHBG ($N_{overall}$ = 370,125; N_{men} = 180,726; N_{women} = 189,473), SHBG adjusted for BMI ($SHBG_{adj}BMI$; $N_{overall}$ = 368,929; N_{men} = 180,094; N_{women} = 188,908), and RA (N_{case} = 14,361; $N_{control}$ = 43,923). A genome-wide cross-trait analysis was performed to quantify global and local genetic correlation, to identify pleiotropic loci, and to infer causal relationship.

Results

Among the overall population, a significant global genetic correlation was observed for $SHBG_{adj}BMI$ and RA (r_g = 0.13, P = 2.0×10^{-4}) which was further supported by local signals (6p21.33-p21.32, 6p21.32, 7q31.33-q32.1, and 22q11.21-q11.22). A total of thirty-four independent pleiotropic SNPs were identified, with corresponding genes enriched in the heart, liver, and pancreas tissues. A putative causal association of $SHBG_{adj}BMI$ on RA was demonstrated (OR = 1.29, 95%CIs = 1.06-1.57) with no reverse causality observed. In sex-specific analysis, most findings were replicated in both sexes. Moreover, a comparable $SHBG_{adj}BMI$-RA causal association was only observed in women (OR = 1.31, 95%CIs = 1.04-1.64) but not in men (OR = 1.06, 95%CIs = 0.87-1.29).

Conclusion

Our cross-trait analysis suggests an intrinsic, as well as a sex-specific, link underlying SHBG and RA, providing novel insights into disease etiology.

Keywords

Sex hormone-binding globulin, rheumatoid arthritis, genome-wide cross-trait analysis, Mendelian randomization analysis

Conflict of interest

None.

[3] A GWAS OF SPEECH ACOUSTICS IDENTIFIES VARIANTS ASSOCIATED WITH VOICE PITCH AND CARDIOVASCULAR TRAITS

Rosa Gisladottir,[1], Patrick Sulem,[1], Kári Stefánsson,[1]

[1] Decode Genetics, Iceland

Introduction

The genetic basis of the human vocal system is largely unknown, as are the sequence variants that give rise to individual differences in voice and speech. Using the largest sample of speech recordings with corresponding genotypes to date, we investigated the genetics of voice and vowel acoustics in 12,901 individuals. Among the acoustic measures we analyzed is *voice pitch*, i.e., how deep or high the voice sounds (known as fundamental frequency or *F0*), and *voice pitch variability* (standard deviation of *F0*), which captures intonation in speech. We also included vowel measures known as formants (labelled *F1*, *F2*, and so on), the vocal tract resonant frequencies that play a key role in distinguishing vowels. These voice and vowel acoustics are important means of interpersonal communication, both at the linguistic and non-linguistic levels, with clear social relevance and potential evolutionary significance. As a bonus, such measures are also useful from a clinical perspective, as they are sensitive to neurological factors affecting speech.

Methods

Participants performed a combination of speech elicitation tasks. Acoustic analysis was performed in Praat and Octave. We inverse normal transformed the acoustic measures separately for each sex and adjusted for age, BMI, and height. We explored how voice pitch and vowel acoustics correlate with anthropometric, physiological, and cognitive traits. We then estimated their SNP-based heritability using LD score regression and performed a genome-wide association study (GWAS).

Results

Voice pitch (in reading) is linked to anthropometric and physiological traits, such as lean body mass (r = -0.08, $P = 5.9\times10^{-19}$) and systolic blood pressure (r = 0.06, $P = 1.4\times10^{-8}$). Voice pitch variability in reading correlates mainly with cognition, including Letter Fluency (r = 0.17, $P = 9.2\times10^{-76}$) and the *g* factor (r = 0.14, $P = 1.8\times10^{-53}$). Vowel formants correlate both with anthropometrics and cognition. The highest SNP-based heritability was observed for

voice pitch in reading (h^2-SNP = 0.17) and variability of voice pitch in reading (h^2-SNP = 0.13). Vowel formant *F2* is also heritable, most notably *F2* in the average of all vowels (h^2-SNP = 0.14). We discovered a common intronic variant, rs11046212-T in *ABCC9* on chromosome 12 (allele frequency 47.7%), that is associated with higher voice pitch in reading (β = 0.11 SD, $P = 2.6\times10^{-18}$, Figure 1). The effect of rs11046212-T is not different between the sexes (Phet = 0.62). Through colocalization analysis of correlated variants ($r^2 > 0.8$) we found that the *ABCC9* variants associated with higher voice pitch are also associated with greater pulse pressure (meta-analysis; β = 0.16 SD, $P = 5.1\times10^{-17}$), reduced ascending aortic area (UK Biobank; β = -0.06 SD, $P = 4.69\times10^{-16}$), and reduced *ABCC9* expression in adrenal gland (GTEx v8; effect size = -0.64 SD, $P = 3.4\times10^{-16}$, Figure 2).

Conclusions

We discovered the first locus for voice pitch; common sequence variants in *ABCC9* that also associate with adrenal gene expression and cardiovascular traits. Rare gain-of-function mutations in *ABCC9* cause Cantú syndrome, a disorder characterized by excess hair growth and cardiovascular, craniofacial and skeletal abnormalities. The results highlight a link between vocal signals and health-related traits, with potential implications for theories on the sexual selection of voice pitch. More generally, by showing that voice and vowel acoustics are influenced by genetics, we have taken important steps towards understanding the genetics and evolution of the human vocal system.

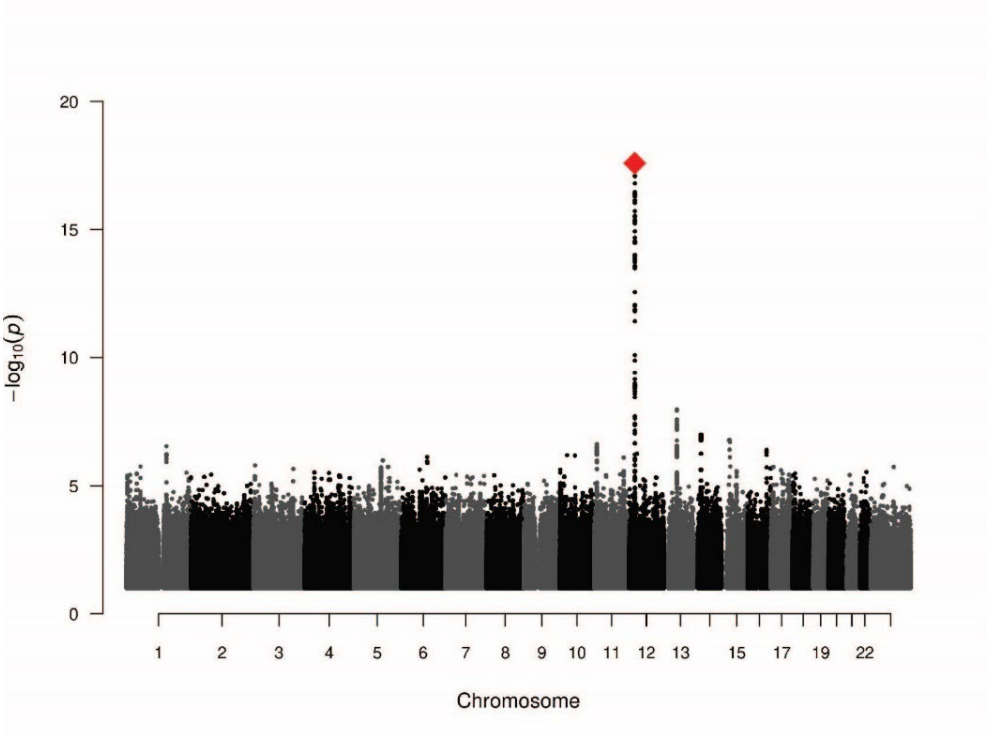


Figure 1. Association with voice pitch. Manhattan plot of the association results for voice pitch in reading (Median *F0*).

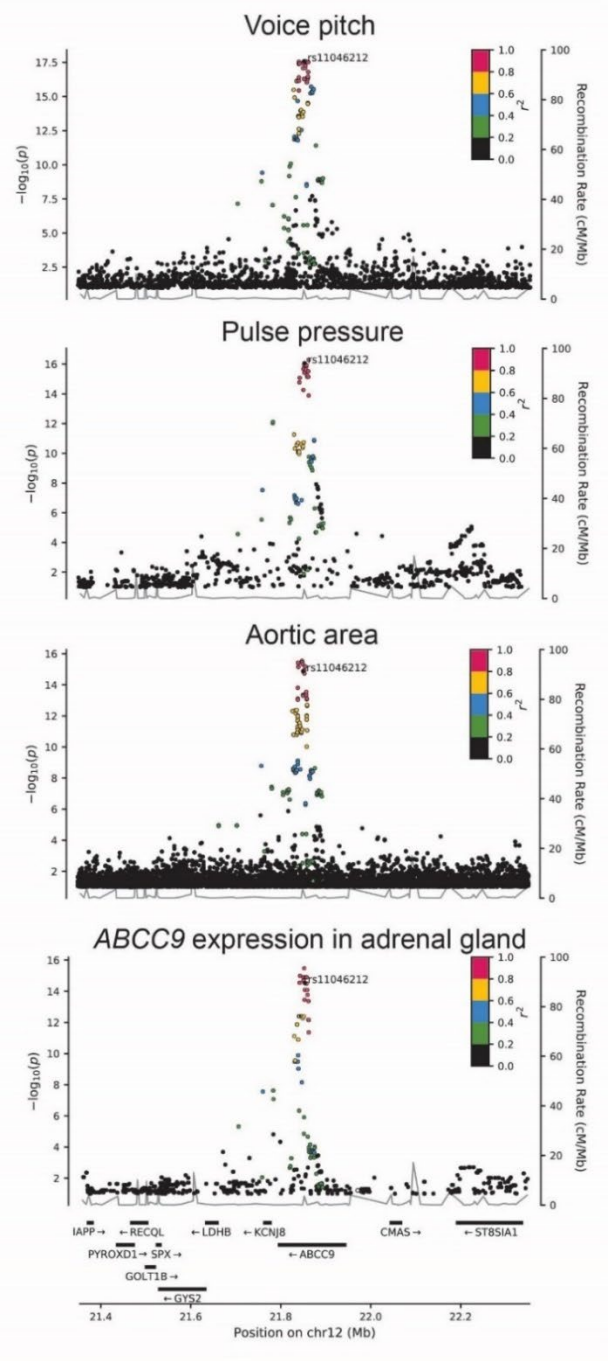


Figure 2. Regional plots of *ABCC9* variants associated with voice pitch, pulse pressure, aortic area and *ABCC9* expression in adrenal gland.

Conflict of interest

None.

[4] A NOVEL REFERENCE ARCHITECTURE FOR MULTI-PARTY FEDERATION: ENABLING JOINT ANALYSIS OF LARGE-SCALE CLINICAL-GENOMIC DATA ACROSS DISTRIBUTED TRUSTED RESEARCH ENVIRONMENTS

Thorben Seeger,[1], Craig Davison,[2], Ingrid Knarston,[2], Bruno Vieira,[2], Ismael Peral,[2], Sangram Keshari Sahu,[2], Rafael Zochling,[2], Eleni Kyriakou,[2], Nate Raine,[2], Rosanna Fennessy,[3], Chiara Bacchelli,[2], Parker Moss,[4], Maria Chatzou Dunford,[2], Pablo Prieto Barja,[2], Mark Avery,[5], Serena Nik-Zainal,[6]

[1] Lifebit, London, United Kingdom, [2] Lifebit Biotech Ltd, [3] University of Cambridge, [4] Genomics England, [5] Eastern Academic Health Science Network, Cambridge University Health Partners, [6] Nihr Cambridge Biomedical Research Centre

Introduction

Enabling researchers access to the vast amounts of linked clinical and genomic data available is crucial to answer the world's most pertinent research questions. Current real world genomic and clinical data is predominantly siloed to protect data privacy and due to strict data regulatory frameworks. Siloed data sets are notoriously difficult for researchers to access and analyse the data. Establishing a secure computing environment which holds data, often referred to as Trusted Research Environments (TREs), enables researchers to access that data for analysis while still securing sensitive, identifiable, patient data. However, within this current model, joint analysis of multiple TREs is not possible and data cannot be moved to a centralised repository without compromising data security. Multi-party federation solves this problem by enabling TREs to securely communicate with each other, increasing potential for novel discoveries. Here, we present what we believe to be the first UK test case of multi-party federation.

Methods

The DARE UK-backed multi-party trusted research environment federation consortium, which includes University of Cambridge, NIHR Cambridge Biomedical Research Centre, Genomics England, Eastern AHSN, Cambridge University Health Partners and Lifebit, developed a novel reference architecture to define how federated analysis can be performed on large-scale clinical-genomic data across distributed TREs. The project utilised and expanded

on Lifebit's Platform technology to enable joint querying and analysis to address defined research questions.

Results

This novel demonstration of multi-party federation securely bridged the TREs of a UK leading research institute, NIHR Cambridge BRC and a public-sector clinical research endeavour, Genomics England. Lifebit deployed its TRE Platform across each of these large scale data resources and developed novel Application Programming Interface technology to enable TRE communication, in addition to a scalable airlock system for secure export into and out of the TRE. All novel technology for this project is fully open-source and developed in alignment with Global Alliance for Genomics and Health (GA4GH) standards. Lifebit's cloud-based Platform allows *in situ* analysis to be run on each TRE individually and then an aggregate of the results from both TREs can be generated in a 'safe-haven'. We successfully deployed the Lifebit platform to bridge these TREs, demonstrating this in a live test case showing joint analysis of data from both TREs without the data leaving the respective locations.

Conclusions

This groundbreaking study and use case have contributed novel insights for future multi-party federated analysis - demonstrating best practice reference architecture, the necessary open-source technology and data governance recommendations going forward for how TREs can communicate to allow joint analysis.

Conflict of interest

Authors include employees of Lifebit Biotech Ltd (Lifebit). MC is CEO, PP is the CTO, TS is CBDO and NR is the Head of Data Custodians for Lifebit. Genomics England and NIHR Cambridge Biomedical Research Centre are Lifebit clients.

[5] A SYSTEMATIC MAPPING OF THE GENOMIC AND PROTEOMIC VARIATION ASSOCIATED WITH MONOGENIC DIABETES

Ksenia Kuznetsova,[1], Jakub Vašíček,[2], Dafni Skiadopoulou,[1], Janne Molnes,[1], Miriam Udler,[3], Alisa Manning,[3], Stefan Johansson,[4], Pål Njølstad,[4], Marc Vaudel,[5]

[1] University of Bergen, Center for Diabetes Research, Department of Clinical Science, Bergen, Norway, [2] University of Bergen, Bergen, Center for Diabetes Research, Department of Clinical Science, Bergen, Norway, [3] Broad Institute, Massachusetts General Hospital, Boston, MA, United States, [4] University of Bergen, Haukeland University Hospital, Center for Diabetes Research, Department of Clinical Science, Bergen, Norway, [5] University of Bergen, Norwegian Institute of Public Health, Center for Diabetes Research, Department of Clinical Science, Bergen, Norway

Introduction

Monogenic diabetes melitus (MD) is a group of mendelian diseases with autosomal dominant inheritance. Several genes like *HNF1A* or *HNF4A* are robustly associated with MD and others are under active scrutiny. In this work we created a database of protein-coding genetic variants associated with MD and mapped their consequences on the proteins they encode. Charting the protein sequences possibly associated with MD will contribute to understanding of the mechanisms of the disease and help to find targets for personalized diagnostics.

Methods

As a main source of data we have used the ClinVar archive of variants associated with human phenotypes. From ClinVar we exported all the variants associated with three phenotypes: "MODY", "Monogenic diabetes", and "Neonatal diabetes". We complemented them with the variants from the review by Rafique et al., published in 2021 that, in turn, implemented literature data mining of the MODY-associated variants. The variants were mapped to the GRCh38 human genome assembly using Ensembl. To decide on the variant consequences to retain, we have analyzed the distribution of variants with known clinical significance. The variants from the categories "missense variant", "frameshift variant", "stop gained", "stop lost", and "start lost", "splice donor", and "splice acceptor" variant, "inframe deletion" and "inframe insertion" have been added to the final database for translation. The possibly encoded protein sequences were obtained by, first, mapping the variant coordinates to transcripts and cDNA sequences using the Ensembl annotation. All the possible isoforms have been included. Then,

the sequences have been translated to proteins. Variant annotation and mapping was achieved by querying the Ensembl and PubMed APIs and using the Vcfanno and Biopython libraries.

Results

A reference table was created from Ensembl by extracting all the variants in the exons of the genes reported to be associated with MD. For the purpose of standardization, all the variants associated with MD from other sources were mapped to this table when possible. After mapping, the variants underwent filtering by the consequence type in order to retain only those variants predicted to affect protein sequences. As a result, 1,360 variants from ClinVar were consolidated in the reference table. We were also able to map 350 variants from Rafique et al. and 211 of them overlapped with the ClinVar dataset. The resulting mapped variants were supplemented with 12 variants manually annotated from Rafique et al., 403 unmapped variants from ClinVar that passed the consequence type filtering and 118 indels from ClinVar. The resulting database consisted of 2,032 genetic variants, *1,947* of which were successfully translated to variants in protein sequences.

Conclusions

A database of variant protein sequences was created for the products of genetic variants associated with monogenic diabetes. All the variants reported to be associated with MD have been included in the database. The database contains variants with three levels of clinical significance: confidently pathogenic, ever reported as pathogenic, and other variants. The sources of information on the genetic variation are not unified which makes mapping of the variants challenging. An automated and reproducible pipeline for variant mapping has been developed and is available for public use. All the MD-associated variants have been translated into protein products and can be compared to the canonical protein sequences. This will help predict the effect of genetic variation on the resulting protein structure and function.

Conflict of interest

None.

[6] ACCURACY OF HAPLOTYPE ESTIMATION AND WHOLE GENOME IMPUTATION AFFECTS COMPLEX TRAIT ANALYSES AND EQUITY IN IMPLEMENTATION OF PRECISION MEDICINE INITIATIVES

Vivek Appadurai,[1], Andrew Schork,[2], Anders Rosengren,[3], Jonas Bybjerg-Grauholm,[4], Morten Krebs,[5], Thomas Werge,[6], Ipsych Consortium,[7]

[1]Institut for Biologisk Psykiatri, Denmark, [2]Institute for Biological Psychiatry, Copenhagen, Denmark, [3]Institut for Biologisk Psykiatri, Roskilde, Denmark, [4]Ssi, Copenhagen, Denmark, [5]Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Roskilde, Denmark, [6]Copenhagen University Hospital, Institute of Biological Psychiatry, Copenhagen, Denmark, [7]Ipsych, Denmark

Background/Objectives

Sample recruitment for research consortia, hospitals, biobanks, and personal genomics companies span years, necessitating genotyping in batches, using different technologies. As marker content on genotyping arrays varies systematically, integrating such datasets is non-trivial and its impact on haplotype estimation (phasing) and whole genome imputation, necessary steps for complex trait analysis, remains under-evaluated.

Methods

Using the iPSYCH consortium dataset, comprising 130,438 individuals, genotyped in two stages, on different arrays, we evaluated phasing and imputation performance across multiple phasing methods and data integration protocols.

Results

While phasing accuracy varied both by choice of method and data integration protocol, imputation accuracy varied mostly between data integration protocols. We demonstrate an attenuation in imputation accuracy within samples of non-European origin, highlighting challenges to studying complex traits and an equitable implementation of precision medicine initiatives in diverse populations.

Conclusion

Imputation errors can modestly bias association tests and reduce predictive utility of polygenic scores. This is the largest, most comprehensive comparison of data integration

approaches in the context of a large psychiatric biobank.

Grants

iPSYCH is supported by the Lundbeck foundation

Dr. Schork is supported by Lundbeck foundation fellowship R335-2019-2318.

Dr. Appadurai is supported by Lundbeck foundation postdoc grant R380-2021-1465.

Conflict of interest

None

[7] ACTIONABLE GENOTYPES IN A LARGE-SCALE STUDY IN ICELAND

Brynjar O. Jensson[1], M.Sc., Gudny A. Arnadottir[1,2], M.Sc., Hildigunnur Katrinardottir[1], M.Sc., Run Fridriksdottir[1], M.Sc., Hannes Helgason[1], Ph.D., Asmundur Oddsson[1], Ph.D., Gardar Sveinbjornsson[1], M.Sc., Hannes P. Eggertsson[1], Ph.D., Gisli H. Halldorsson[1,10], M.Sc., Bjarni A. Atlason[1], B.Sc., Hakon Jonsson[1], Ph.D., Gudjon R. Oskarsson[1], Ph.D., Arni Sturluson[1], Ph.D., Sigurjon A. Gudjonsson[1], B.Sc., Gudmundur A. Thorisson[1], Ph.D., Florian Zink[1], Ph.D., Kristjan H.S. Moore[1,3], M.A., Gunnar Palsson[1], M.Sc., Asgeir Sigurdsson[1], B.Sc., Adalbjorg Jonasdottir[1], B.Sc., Aslaug Jonasdottir[1], M.Sc., Magnus K. Magnusson[1,2], M.D., Anna Helgadottir[1], Ph.D., Valgerdur Steinthorsdottir[1], Ph.D., Julius Gudmundsson[1], M.Sc., Simon N. Stacey[1], Ph.D., Rafn Hilmarsson[2,4], M.D., Isleifur Olafsson[5], M.D., Ph.D., Oskar Th. Johannsson[6], M.D., David O. Arnar[1,2,7], M.D., Jona Saemundsdottir[1], B.Sc., Olafur Th. Magnusson[1], Ph.D., Gisli Masson[1], Ph.D., Bjarni V. Halldorsson[1,8], Ph.D., Agnar Helgason[1,3], Ph.D., Hreinn Stefansson[1], Ph.D., Ingileif Jonsdottir[1,2,9], Ph.D., Hilma Holm[1], M.D., Thorunn Rafnar[1], Ph.D., Unnur Thorsteinsdottir[1,2], Ph.D., Daniel F. Gudbjartsson[1,10], Ph.D., Kari Stefansson[1,2*], M.D., Ph.D., Patrick Sulem[1*], M.D.

[1]deCODE genetics/Amgen, Inc., Reykjavik, Iceland. [2]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. [3]Department of Anthropology, University of Iceland, Reykjavik, Iceland. [4]Department of Urology, Landspitali-The National University Hospital of Iceland, Reykjavik, Iceland. [5]Department of Clinical Biochemistry, Landspitali-The National University Hospital of Iceland, Reykjavik, Iceland. [6]Department of Oncology, Landspitali-The National University Hospital of Iceland, Reykjavik, Iceland. [7]Department of Medicine, Landspitali-The National University Hospital of Iceland, Reykjavik, Iceland. [8]School of Science and Engineering, Reykjavik University, Reykjavik, Iceland. [9]Department of Immunology, Landspitali-The National University Hospital of Iceland, Reykjavik, Iceland. [10]School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. [*]Equal contributions

Introduction

The American College of Medical Genetics and Genomics (ACMG) recommends reporting actionable genotypes in 73 genes (ACMG v3.0, May 2021) associated with diseases for which preventive or therapeutic measures are available. The prevalence and impact of actionable genotypes in these genes in a large-scale study is not yet known. Using manual review of

sequence variants and their associated evidence detected in a large set of whole-genome sequenced Icelanders, we aimed to assess the fraction of individuals carrying an actionable genotype in one or more of the ACMG v3.0 genes.

Methods

We assessed the prevalence of coding and splice variants in genes on the ACMG v3.0 list in whole-genomes of 57,933 Icelanders. With manual assessment of previous reports of pathogenicity in ClinVar, the frequency of the variants and their associations with diseases, we assigned pathogenicity to all reviewed variants to create a manually curated set of actionable genotypes. We compared this set of genotypes to the ones detected when using narrow criteria, previously described by another group as 'strict'. We assessed the impact of the detected actionable genotypes on lifespan by evaluating both survival to death (hazard ratio) and the association with number of years lived up to death as a quantitative trait.

Results

Based on manual curation of 4,405 sequence variants in the ACMG v3.0 genes, we identified 235 actionable genotypes in 53 genes. 2,306 whole-genome sequenced Icelanders (4.0%) carried at least one actionable genotype. Only 1,650 Icelanders (2.9%) carried an actionable genotype according to the previously defined narrow criteria. We found that carrying an actionable genotype in gene linked to cancer gene reduces survival by 3 years and we identified 13 genotypes that associate with shorter lifespan. Additionally, we identified actionable genotypes in the Icelandic population in genes that are currently not on the ACMG v3.0 list.

Conclusions

According to the ACMG v3.0 recommendations, 1 in 25 Icelanders carried an actionable genotype. Manually curating a set of sequence variants using large-scale association data enables accurate classification and can increase the number of detected carriers of actionable genotypes by 40% compared to previous criteria, which can significantly impact societal disease burden.

Conflict of interest

[8] ADULT: AN EFFICIENT AND ROBUST TIME-TO-EVENT MODEL FOR GWAS

Bjarni Vilhjalmsson,[1], Emil Pedersen,[2], Esben Agerbo,[3], Oleguer Plana-Ripoll,[4], Jette Steinbach,[5], Morten Krebs,[6], David Hougaard,[7], Thomas Werge,[6], Merete Nordentoft,[8], Anders Børglum,[4], Katherine Musliner,[9], Andrea Ganna,[10], Andrew Schork,[6], Preben Mortensen,[4], John Mcgrath,[4], Florian Privé,[4]

[1] Aarhus Universitet, National Centre for Register-Based Research, Aarhus V, Denmark, [2] Aarhus University, National Centre for Register-Based Research, Aarhus, Denmark, [3] Aarhus University, National Centre for Register-Based Research, Aarhus V, Denmark, [4] Aarhus University, Denmark, [5] Aarhus University, Aarhus V, Denmark, [6] Copenhagen University Hospital– Mental Health Services Cph, Denmark, [7] Statens Serum Institut, Denmark, [8] Copenhagen University Hospital– Mental Health Services Cph,, Denmark, [9] Aarhus University Hospital-Psychiatry, Denmark, [10] University of Helsinki, Finland

Introduction

Proportional hazards models have previously been proposed to analyze time-to-event phenotypes in genome-wide association studies (GWAS). While proportional hazards models have many useful applications, their ability to identify genetic associations under different generative models where ascertainment is present in the analyzed data is poorly understood. This includes widely used study designs such as case-control and case-cohort designs (e.g., the iPSYCH study design) where cases are commonly ascertained.

Methods

Here we propose the age-dependent liability threshold model (ADuLT), first introduced as the underlying model for the LT-FH++ method, as an alternative approach for time-to-event GWAS. We benchmark ADuLT with four alternative approaches using simulated data under both a proportional hazards generative model and the age-dependent liability threshold model and consider varying degrees of case ascertainment. The four methods compared to are a linear regression GWAS, a computationally efficient Cox regression for GWAS (SPACox), as well as standard Cox regression with and without inverse probability weights. We then perform GWAS in on four psychiatric disorders, ADHD, Autism, Depression, and Schizophrenia, in the IPSYCH data using ADuLT, linear regression, and a computationally efficient Cox regression.

Results

We find all Cox regression methods to underperform when cases are strongly ascertained (cases are oversampled by a factor larger than 5), regardless of the generative model used. In contrast, we found ADuLT to be robust to case-control ascertainment, while being much faster to run. We then applied the three most computationally efficient methods to the iPSYCH case-cohort sample, which has a strong case-ascertainment. Summarizing across all four mental disorders, ADuLT found 20 independent genome-wide significant associations, while linear regression GWAS found 17 and SPACox found 8, consistent with our simulation results.

Conclusions

As more genetic data are being linked to electronic health records, robust GWAS methods that can make use of age-of-onset information have the opportunity to increase power in analyses. We find that ADuLT to be a robust time-to-event GWAS method that performs on par with or better than Cox-regression GWAS, both in simulations and real data analyses of four psychiatric disorders. ADuLT has been implemented in an R package called LTFHPlus and is available on GitHub https://emilmip.github.io/LTFHPlus/.

Conflict of interest

BJV is a member of the Scientific Advisory Board for Allelica.

[9] A TRANSCRIPTIONAL SCORE FOR STRATIFICATION OF PATIENTS WITH ACUTE INFECTION

Eddie Cano-Gamez,[1], Katie Burnham,[2], Cyndi Goh,[3], Alice Allcock,[1], Zunaira Malick,[1], Lauren Overend,[1], Andrew Kwok,[1], David Smith,[1], Hessel Peteres-Sengers,[4], David Antcliffe,[5], Stuart Mckechnie,[6], Brendon Scicluna,[4], Tom van der Poll,[4], Anthony Gordon,[7], Charles Hinds,[8], Emma Davenport,[9], Julian Knight,[1]

[1]University of Oxford, Wellcome Centre for Human Genetics, Oxford, United Kingdom, [2] Wellcome Sanger Institute, Welcome Genome Campus, Cambridge, United Kingdom, [3] University of Oxford, The Jenner Institute, Oxford, United Kingdom, [4] Centre for Experimental and Molecular Medicine, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands, [5] Imperial College London, Division of Anaesthesia, Pain Medicine, and Intensive Care; Department of Surgery and Cancer, London, United Kingdom, [6] Oxford University Hospitals, NHS Foundation Trust, Oxford, United Kingdom, [7] Imperial College London, Division of Anaessthesia, Pain Medicine and Intensive Care; Department of Surgery and Cancer, London, United Kingdom, [8] William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University, London, United Kingdom, [9] Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, United Kingdom

Introduction

Severe infections can lead to sepsis, a dysregulated host immune response accompanied by life-threatening organ dysfunction. Early identification of individuals at elevated risk of severe sepsis is a key clinical challenge. It has been previously shown that sepsis patients can be stratified into subgroups at different mortality risks based on their gene expression profiles. Importantly, these subgroups show differential responses to corticosteroid treatment, which makes them relevant for personalized medicine. However, stratification remains time consuming, and we lack a single reproducible framework for patient classification. Thus, methods with quicker turn-around and more generalizable statistical frameworks are needed to bridge the gap between research and clinical application.

Methods

We integrated whole blood gene expression data from 3,149 sepsis and healthy volunteer samples, comprising microarray, RNA-sequencing, and qRT-PCR measurements from six

different cohorts. We used these data to train SepstratifieR, a suite of machine learning models which can classify patients into subgroups and estimate a quantitative sepsis response score (SRSq) reflective of each patient's degree of immune dysfunction. SepstratifieR is based on 7 and 19-gene signatures, can be combined with multiple gene expression profiling technologies, and is publicly available as an R package. We deployed SepstratifieR in bacterial and viral sepsis, H1N1 Influenza, and COVID-19. Furthermore, we used mediation analysis to integrate SRSq scores with organ dysfunction and clinical outcome measures.

Results

We demonstrated that SepstratifieR can robustly identify individuals at different levels of immune risk and reliably predict SRSq scores for each individual. SRSq was significantly associated with organ failure and mortality in sepsis, influenza, and COVID-19. Elevated SRSq scores correlated with upregulation of neutrophil pathways and suppression of T cell function and antigen presentation. Integration of SRSq with clinical measures further revealed that death secondary to infection was mediated by systemic organ failure and shock in sepsis, but by respiratory failure in COVID-19.

Conclusions

Our method enables better stratification of patients with acute infection and could be used to aid patient monitoring, guide targeted interventions, and inform clinical trial design. Importantly, SRSq scores are generalizable across infections and amenable to point-of-care testing with quick turn-around such as qRT-PCR, thus advancing personalized medicine in infection.

Conflict of interest

None.

[10] A NORDIC NETWORK TO PROMOTE PRECISION MEDICINE RESEARCH

Hakon Heimer

NSHG-PM Secretariat, Denmark

Introduction: The Nordic countries have advantages for biomedical research and its translation to public health via national health systems (Njølstad, 2019, Nature Genetics 51:6). The Nordic Society of Human Genetics and Precision Medicine (NSHG-PM) was founded to establish a Nordic framework for precision medicine (PM) research into common human diseases, particularly through integrated analyses using multiple large-scale datasets.

Methods: NSHG-PM identifies stakeholders, research collaborations, cohorts and biobanks, and published research in the Nordic region. We convene scientific meetings, workshops, and webinars, and our working groups draft white papers and articles to facilitate meetings and communication with stakeholders. Our social media platforms maintain active communication.

Results: NSHG-PM has developed relationships with 100+ Nordic stakeholder organizations. In addition to Society conferences, our highlighted activities include:

- WORKSHOP: "Nordic Precision Medicine: Cooperation in Enabling Research and Education" (2020). Co-organized with the Nordic Deans of Health and Medical Faculties. 50+ leaders of academia, national and regional policymaking, research funding.

- WORKSHOP: "Vision for a Nordic Precision Medicine Initiative: Meeting of the Nordic Funders of Biomedical Research" (2021). Co-organized with Novo Nordisk Foundation. Attendees included major private foundations and Nordic research councils and innovation funds.

- CONFERENCE, VIRTUAL: "COVID-19 – Response and Research in the Nordic Countries" (2021) First research meeting on pandemic in Nordic region, drew 25+ speaker and 350+ registrants.

- ONLINE COURSE: "Personalized medicine from a Nordic perspective" (2022) First Nordic personalized medicine course (2022) for university medical students, featuring 30 videos by 35 contributing lecturers. Co-created with University of Copenhagen, University of Iceland, and Nordic Deans Education Working Group. https://www.coursera.org/lecture/personalised-

medicine-from-a-nordic-perspective/introduction-to-personalised-medicine-from-a-nordic-perspective-saedis-R1JE7

Conclusion: These activities have demonstrated the great enthusiasm for Nordic collaboration on precision and personalized medicine. However, they also reveal the need to remove barriers to the secure use of data for research that can translated into innovation, prevention, and clinical practice.

[11] DEVELOPING RESOURCES FOR GENOMIC MEDICINE IN SWITZERLAND

Katrin Mannik,[1], Jacques Fellay,[2], Ioannis Xenarios,[3], Keith Harshman,[1]

[1]Health 2030 Genome Center, Geneva, Switzerland, [2] Epfl, Health 2030 Genome Center, Lausanne, Switzerland, [3] University of Lausanne, Health 2030 Genome Center, Geneva, Switzerland, Lausanne, Switzerland

Introduction

Switzerland resembles the Nordic countries in many aspects, including population size, GDP per capita and high civic participation and trust in society and government. Switzerland is globally known also as a country with high quality and impactful research (measured in e.g., publication numbers or citations per capita) in clinical and life sciences in general. However, while the Nordic countries are at the forefront of personalized medicine, Switzerland remained, until this year, among the very few last countries in Europe without any national genomics strategy or a coordinated large-scale plan to improve healthcare in the country through the use of innovations in genomics. This lack of coordinated effort focusing on genomic medicine is also reflected in the low numbers of Swiss samples and authorships contributed to international efforts in precision medicine.

In May 2022, the National Steering Board of the government-funded initiative Swiss Personalized Health Network (SPHN) approved the first national strategy to facilitate human genomic research and accelerate the integration of genomics into healthcare practice. The endorsed concept consists of two complimentary pillars to be fulfilled in 2022-2024. First, sequencing a pilot cohort of a national reference genomic dataset (referred to as the "Genome of Switzerland") to demonstrate the feasibility of genome data production and sharing at scale, serve as a population-based resource for future research, and provide strong basis for building the public trust required for genome-based healthcare approaches. Second, setting up an infrastructure (Swiss Federated Genomics Network; SwissFGN) optimized to support and coordinate genomic data generation, processing, and exchange in a scalable and sustainable manner. The SwissFGN will be built as a semi-decentralized infrastructure with three hubs that will provide services to the Swiss biomedical community: an Expert Genomics Hub, a Central Genomic Data Repository, and a Data Coordination Hub.

The role of the genomics lead and central hub in this national initiative has been mandated to the Health 2030 Genome Center (H2030GC). The H2030GC is the genomics arm of the Health 2030 initiative partnered by the Swiss Federal Institute of Technology in Lausanne (EPFL) and universities and university hospitals of Geneva, Lausanne and Bern. The H2030GC is an open, non-profit platform established in 2018 with the aim to promote genomic research and medicine through four pillars. First, clinical grade (ISO15189 accredited) generation and processing of human whole-exome-, genome-, and full transcriptome sequencing data at scale, to address the needs of clinical and research community across diverse clinical areas. Second, continuous evaluation of new technologies and implementation of certified analytical pipelines and medical devices for genomic and transcriptomic analyses (e.g., automated AI-powered variant interpretation and reporting; low-pass WGS for population and predictive genomics studies; privacy-preserving technologies for variant exploration and polygenic risk score calculations). Third, providing means for multi-omic approaches by integrating genomic and transcriptomic analyses of clinical samples with other large-scale molecular data types, such as proteomics and metabolomics in the Swiss Multi-Omics Center (Personalized Health and Related Technologies strategic focus area of the ETH Domain). And finally, genomic data life cycle governance enabling secure storage, use and re-purposing of genomic and transcriptomic data for research and clinical applications.

The proposed poster presentation will introduce the emerging genomic medicine initiatives at the Swiss national level and the role of the Health 2030 Genome Center as an enabler for the genomic research and medicine community.

Conflict of interest

None.

[12] DISCOVERING THE GENETICS OF KIDNEY FUNCTION DECLINE USING LARGE-SCALE GENOMIC DATA

Humaira Rasheed,[1], Mathias Gorski,[2], Alexander Teumer,[3], Cristian Pattaro,[4], Anna Köttgen,[5], Florian Kronenberg,[6], Iris Heid,[7]

[1]K. G. Jebsen Center for Genetic Epidemiology, Ntnu, Norwegian University of Science and Technology, Trondheim, Norway, Department of Public Health and Nursing, Department of Public Health and Nursing, Trondheim, Norway, [2] University Hospital Regensburg, Nephrology, Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany, Dept of Genetic Epidemiology, University Hospital Regensburg, Regensburg, Germany, [3] University Medicine Greifswald, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany, Institute for Community Medicine, Greifswald, Germany, [4] Eurac Research, Institute for Biomedicine (Affiliated With the University of Lübeck), Bolzano, Italy, Italy, [5] Albert-Ludwigs-Universität Freiburg, Innere Medizin IV, Nephrologie und Allgemeinmedizin, Universitätsklinikum, Institute of Genetic Epidemiology, Department of Biometry, Epidemiology and Medical Bioinformatics, Faculty of Medicine and Medical Center–university of Freiburg, Freiburg, Germany, Germany, [6] Medizinische Universität Innsbruck, Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria, Department für Medizinische Genetik, Molekulare und Klinische Pharmakologie, Innsbruck, Austria, [7] Dept of Genetic Epidemiology, University Hospital Regensburg, Dept of Genetic Epidemiology, Helmholtz Zentrum München, Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany, Münich, Germany

Introduction

Chronic kidney disease (CKD) is a significant health problem as every 10th adult person globally is suffering from this disease and the disease is becoming the 5th global cause of death by 2040. Estimated glomerular filtration rate (eGFR) is a commonly used clinical marker of kidney function and its value <60 ml/min/1.73m^2 is an indicator of CKD. Progressive eGFR-decline can lead to kidney failure, necessitating dialysis or transplantation. Hundreds of loci from genome-wide association studies (GWAS) for eGFR help to explain its variability at one point in time, but the contribution of these or other loci to eGFR-decline remains largely unknown.

Methods

Using the largest collection of longitudinal eGFR data to date, we performed GWAS for annual eGFR-decline and meta-analysed 62 studies where eGFR was assessed at two time points (minimum 3 months apart) in all 343,339 individuals. A meta-analysis restricted to those with diabetes mellitus (n=37,375) and CKD (n=26,653) cases was performed to quantify SNP effects on eGFR decline in these high-risk subgroups. Most studies included in our analysis were population based (76%) and of European ancestry (74%). We also explored different covariate adjustments and applied gene prioritization approaches to identify genes in associated loci with compelling biological evidence. As CKD is an age-dependent disease, we further explored the SNP-age interactions for the identified loci. Finally, we tested the cumulative impact (using a weighted genetic risk score) of eGFR-decline associated variants on adverse clinical outcomes including end-stage renal disease and acute kidney injury.

Results

Twelve genome-wide significant independent variants for eGFR-decline unadjusted or adjusted for eGFR-baseline (11 novel, one known for this phenotype), including nine variants robustly associated across models, were identified. All loci for eGFR-decline were known for cross-sectional eGFR, and thus distinguished a subgroup of eGFR-associated loci. Seven of the nine variants showed variant-by-age interactions on eGFR cross sectionally in an independent set of 350,000 individuals, which linked genetic associations for eGFR-decline with age-dependency of variants associated with eGFR cross-sectionally. The high-risk subgroups showed two- to four-fold greater genetic effects on eGFR-decline than the total sample. Five variants associated also with eGFR-decline mapped to genes with functional evidence (*UMOD, SPATA7, GALNTL5, TPPP*). An unfavorable versus favorable nine-variant genetic profile showed increased risk odds ratios of 1.35 for kidney failure (95% confidence intervals 1.03-1.77) and 1.27 for acute kidney injury (95% confidence intervals 1.08-1.50) in over 2000 cases each, with matched controls).

Conclusion

We provide the to-date largest longitudinal GWAS data resource, genetic loci, and prioritized genes for kidney function decline. Our findings reveal important insights into the age-dependency of kidney function genetics. Besides, this novel GWAS catalog of eGFR-decline is a valuable resource for future Mendelian Randomization studies.

Conflict of interest

None.

[13] DO POLYGENIC RISK SCORES IMPROVE OUR ABILITY TO PREDICT RECURRENT HOSPITAL ADMISSION IN PATIENTS WITH BIPOLAR DISORDER?

Katherine Musliner,[1], Jette Steinbach,[2], Fernando Goes,[3], Peter Zandi,[4], Preben Mortensen,[2], Bjarni Vilhjalmsson,[2]

[1]Aarhus University Hospital, Aarhus University, Department of Affective Disorders, Aarhus, Denmark, [2] Aarhus University, Denmark, [3] Johns Hopkins School of Medicine, United States, [4] Johns Hopkins School of Public Health, United States

Introduction

Polygenic risk scores are of great interest to the scientific community, however their utility as diagnostic or decision aids in a clinical psychiatric setting has yet to be established. Our goal in this project is to evaluate the extent to which polygenic risk scores can improve prediction of recurrence in patients with bipolar disorder above and beyond what is achievable through clinical and demographic risk factors alone.

Methods

Data were obtained from the iPSYCH2015 cohort, which includes all individuals diagnosed with bipolar disorder in Danish psychiatric hospitals from 1995-2015. Our study sample included 2,989 patients (64% female) with at last one bipolar diagnosis (F30-F31) and no prior record of schizophrenia spectrum disorders (F20-F29). Recurrence was defined as a subsequent treatment period for bipolar disorder beginning at least 8 weeks after the end date of the first bipolar episode. Potential non-genetic predictor variables were selected based on a review of the literature and included gender; age at first bipolar diagnosis; setting at point of entry into psychiatric care (inpatient, outpatient or emergency); treatment duration (days); polarity (manic, depressive, mixed, other, unspecified); psychotic symptoms; prior or concurrent comorbid psychiatric diagnoses; prior intake of antidepressants or antipsychotics; occupation; education; parental income; geographic setting (urban vs. rural); and family history of schizophrenia or mood disorders. Polygenic risk scores for bipolar disorder, major depression, schizophrenia, and ADHD were generated using a meta-PRS method combining internally and externally trained components. Individual associations between potential predictor variables and recurrence were assessed using cox proportional hazards models with time to recurrence as the outcome variable. Patients were followed from 8 weeks after the end date of their first

bipolar episode until the start of their first recurrent episode, death, emigration, or December 31, 2016, whichever came first. Prediction models with and without polygenic risk scores were developed using Royston & Palmer flexible parametric survival analysis. Models were adjusted for the first 20 principal components.

Results

947 (32%) of patients experienced ≥1 recurrent hospital contact(s) for bipolar disorder. The following predictors were selected for inclusion in the multivariable prediction model: gender (Hazard Ratio=1.20, 95% Confidence Intervals=1.04-1.38); polarity (depressive [1.58, 1.28-1.95], mixed [1.35, 1.04-1.75], other [1.36, 1.02-1.82], unspecified [1.37, 1.11-1.69]); psychotic smptoms (1.53, 1.19-1.98); prior or concurrent substance abuse disorders (1.26, 1.05-1.51); inpatient (1.53, 1.30-1.79) and emergency (0.74, 0.57-0.96) treatment settings at point of entry into psychiatric care; prior antipsychotic prescription (1.29, 1.11-1.49); family history of depression (1.20, 1.04-1.37); and treatment duration for the first month after initial diagnosis (1.05, 1.01-1.09). PRS for bipolar disorder was statistically significant in the model adjusted only for principal components (1.07, 1.01-1.05) but not the multivariable model (HR=1.04, 0.96-1.13). Comparison of model performance measures show that the model including PRS did not differ significantly from the model without PRS.

Conclusions

Our results suggest that including polygenic risk scores does not improve our ability to predict recurrence in patients with bipolar disorder. Future studies should investigate whether polygenic risk scores can help predict other outcomes in patients with psychiatric disorders.

Conflict of interest

None.

[14] EMPOWERING GENOMICS WITH PROTEOMICS: THERAPEUTIC TARGET AND BIOMARKER DISCOVERY LEVERAGING LARGE POPULATION COHORTS LIKE THE UK BIOBANK

Cynthia Lawley,[1], Anna Lejon,[2], Philippa Pettingill,[2], Klev Diamanti,[2], Lotta Wik,[2], Niklas Nordberg,[2], John Broberg,[2], Johan Björkesten,[2], Erika Assarsson,[2], Sara Henriksson,[2], Ida Grundberg,[2], Christina Westerberg,[2], Adam Falck,[2], Martin Lundberg,[2], Lasse Folkersen,[3], Anders Malarstig,[4]

[1]Olink, Population Health, San Francisco, United States, [2] Olink, Sweden, [3] Nucleus Genomics, United States, [4] Karolinska Institutet, Sweden

Introduction

Proteins are the business molecules of the cell and are primary targets for clinical trial candidates with effects highly influenced by genetic variability. Understanding the dynamics of the human proteomics is crucial for identifying biomarkers to be used as measurable indicators for disease severity and progression, patient stratification, and drug development. Protein data also amplify power to detect genetic supported therapeutic target discovery.

Methods

The Proximity Extension Assay (PEA) is a technology that translates protein information into actionable insights across large samples sizes in both healthy and disease samples. The high-throughput nature of the assay is enabled by linking protein-specific antibodies to DNA-encoded tags that can be read out on a next generation sequencer.

Results

We have combined the PEA technology described above with automated sample preparation and a high-throughput sequencing readout for parallel measurement of ~3,000 proteins for up to 384 samples at a time, generating over 1 million data points per run.

Conclusions

Characterizing the proteome alongside genetic and clinical data enables a pQTL framework to not only validate known clinical targets and identify new clinical targets but to also suggest repurposing opportunities of existing therapies for new indications. Here we will summarize

results where proteomics is impacting large population health studies (e.g., UK Biobank, SCALLOP) to advance epidemiology and precision medicine.

Conflict of interest

CL, AL, PP, KD, LW, NN, JBr, JBj, EA, SH, IG, CW, WL, AF, ML, LF, AM are full time employees of Olink.

[15] ESTIMATING THE GENETIC RELATIONSHIP BETWEEN PSYCHIATRIC DISORDERS AND CARDIOVASCULAR DISEASES USING GENOTYPIC DATA, ALONG WITH THE EXTENSIVE GENEALOGIES AND NATIONAL PATIENT REGISTRIES OF DENMARK AND SWEDEN.

Joeri Meijsen,[1], Kejia Hu,[2], Richard Zetterberg,[3], John Shorter,[3], David Howard,[4], Thomas Werge,[3], Fang Fang,[2], Alfonso Buil,[3]

[1] Mental Health Services Copenhagen, Institute of Biological Psychiatry, Roskilde, Denmark, [2] Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden, [3] Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Mental Health Services Copenhagen, Roskilde, Denmark, [4] Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

Introduction

Individuals with a psychiatric disorder experience, on average a shorter life expectancy of 15 to 20 years than unaffected individuals. This reduced life expectancy is mainly due to somatic comorbidities such as cardiovascular diseases (CVD). The comorbidity between both disorders has been extensively studied using epidemiological data. For instance, a recent study using Danish national registries (5.9 million persons) observed a substantial absolute lifetime risk for CVD after a psychiatric disorder diagnosis. However, these results do not show how biological and environmental risk factors contribute to this observed comorbidity. In this study we integrate national registries from Denmark and Sweden with large genetic data (iPSYCH) to examine the genetic overlap between 6 psychiatric- and 14 cardio-metabolic disorders. We use 1.) polygenic scores (PGS), 2.) LD-score regression, and 3.) genetic correlations, derived using national patient registries and complete genealogical structure of both Denmark and Sweden.

Methods

To investigate the overlap between the genetic causes of CVD and psychiatric disorders, we estimated the association between 6 psychiatric diagnoses and 14 CVD PGS in the iPSYCH 2012 case-cohort study using logistic regression analysis. We evaluated if the CVD PGS associations were independent and not confounded by pleiotropy by including the psychiatric PGS into the regression models. LD-score regression was applied to estimate the genetic correlation between iPSYCH 2012 and CVD GWAS summary statistics. Finally, we compared these

results against tetrachoric- and cumulative incidence based genetic correlations, derived using individuals born in the same iPSYCH 2012 birth window and their family members from the extensive population records and complete genealogical structure of both Denmark and Sweden.

Results

We observed that all CVD PGS were associated with an increased risk for ADHD (OR:1.05-1.25). We also observed fewer, smaller associations, between CVD PGS and other psychiatric outcomes. Two negative PGS associations were observed for anorexia, suggesting a reduced risk for anorexia with an increased genetic liability for type-2 diabetes and peripheral artery disease. These associations remained mostly unchanged when accounting for psychiatric pleiotropy in the logistic regression model. LD-score regression estimates were comparable to the PGS results. Finally, we observed that register based genetic correlations are generally smaller or not different from 0 compared to LDSC or PGS estimates.

Conclusions

In this study we observed mixed results, with substantial shared genetic effects between CVD and psychiatric disorders when using GWAS summary statistics (PGS associations and LDSR), but small to negligible effects when using register based genetic correlations. We hypothesize that LDSC and PGS estimates can be inflated due to potential biases in GWAS studies, i.e., 1) these methods do not account for indirect effects on the phenotype and 2) there is a hidden overrepresentation of psychiatric cases in the CVD GWAS summary statistics. In contrast, our observed register based genetic correlations may represent a more accurate estimate as these are derived using the entire national population and family information in two countries. Until we test our hypothesis, we conclude that there is a mild genetic overlap between ADHD and some CVD. However, we warn about overinterpretation of genetic correlations based on GWAS summery statistics without being aware of potential biases driven those correlations.

Conflict of interest

None.

[16] EVALUATION OF FINE MAPPING APPROACHES BASED ON BAYESIAN LINEAR REGRESSION MODELS

Merina Shrestha,[1], Palle Duun Rohde,[2], Peter Sørensen,[3]

[1] Aarhus University, Center of Quantitative Genetic and Genomics, Aarhus, Denmark, [2] Aalborg University, Aalborg University, Genomic Medicine, Department of Health Science and Technology, Aalborg, Denmark, [3] Aarhus University, Aarhus University, Center for Quantitative Genetics and Genomics, Aarhus, Denmark

Introduction

Understanding how genetic variation shapes human complex traits and diseases is essential for drug target development and the future of precision medicine. Identifying the causal genetic variants underlying phenotypic variation has been at the forefront of human genetic research the last two decades, and much effort has been put into development of novel statistical genetic models that more accurately can identify genetic variants associated with trait variation and disease risk.

Bayesian linear regression (BLR) models have been proposed as a unified framework for accurate identification of associated genetic variants, so-called genetic fine mapping. BLR models account for the underlying genetic architecture of the traits by specifying different prior distribution for the marker effects allowing the true genetic signal to be heterogeneously distributed over the genome.

BLR models used for fine mapping have been specialized to focus on the markers that have the largest chance of being causal. As the effect of a causal variant may be distributed across multiple markers due to linkage disequilibrium (LD) fine mapping approaches consider set of markers in a genomic region rather than individual markers. Several association statistics have been developed to identify marker sets that explain a significant proportion of the genetic variance.

The aim of this study was to investigate to what extend mapping power and false discoveries of the BLR models depends on the choice of 1) prior for marker effects, 2) marker set definition, and 3) marker set association statistics.

Methods

The power of the BLR fine mapping approach was assessed using simulations. We simulated ten replicated quantitative traits with an underlying genetic architecture ranging from low (GA1) to moderate (GA2) polygenicity. The traits were simulated using UK Biobank (UKB) genotyped markers (533,679 SNPs) for 335,744 White British Unrelated individuals. We simulated a trait with heritability of 30%, proportion of causal SNPs, 0.1%, 1% and 5%, chosen randomly among the genotyped markers. Marker effects and standard errors was obtained by linear regression of phenotypes on each genotype. BLR models were fitted using the marker effects and a sparse LD matrix constructed from the UKB genotypes. Different prior distributions of the marker effects were specified ranging from low (BayesC) to moderate (BayesR and BayesA) to high (BayesN) polygenicity.

Marker sets were constructed based on two LD thresholds ($r^2 > 0.7$ and $r^2 > 0.5$), and marker set association statistics was defined as the sum of genetic variance (T_{VAR}), and the sum of posterior inclusion probability (T_{PIP}). The power of the different BLR models, marker sets and association statistics was estimated using F1 classification score.

Results

The highest average F1 classification score (avg. F1) for T_{VAR} was observed for BayesC and BayesR priors in both GA1 scenario: 0.686 and 0.684 and GA2 scenario: 0.661 and 0.660, respectively (*p-value* < 0.001). A slightly lower avg. F1 of 0.646 in GA1 and 0.621 in GA2 was observed for BayesA prior. The lowest avg. F1 of 0.228 in GA1 and 0.167 in GA2 was observed for the BayesN prior. For T_{PIP}, BayesC performed best with avg. *F1* of 0.596 in GA1 and 0.555 in GA2 compared to Bayes R 0.057 in GA1 and 0.043 in GA2 at the stringent cut-off of significance (*p-value* < 1×10^{-5}). However, BayesC and BayesR showed similar avg. *F1* of 0.681 and 0.684 respectively in GA1, and 0.652 and 0.659 in GA2 at *p-value* < 0.001. In general, a higher LD threshold ($r^2 > 0.7$) resulted in higher avg. *F1* across all marker effect priors and marker set association statistics in the GA1 and GA2 scenarios.

Conclusions

Our preliminary results suggest that the power of BLR fine mapping is influenced by how the marker sets are defined, the prior distributions, and type of marker set association statistic. Additional ways to define marker sets and type of association statistics will be explored.

Conflict of interest

None.

[17] FAMILIAL CO-AGGREGATION OF MULTIPLE SCLEROSIS AND AUTOIMMUNE DISEASES - A SWEDISH POPULATION-BASED STUDY

Qianwen Liu,[1], Thomas Frisell,[2], Yuan Jiang,[3], Pernilla Stidh,[2], Klementy Shchetynsky,[2], Lars Alfredsson,[4], Ingrid Kockum,[5], Xia Jiang,[2], Ali Manouchehrinia,[2]

[1] Karolinska Institutet, Department of Clinical Neuroscience, Stockholm, Sweden, [2] Karolinska Institutet, [3] Karolinska Institutet, Department of Clinical Neuroscience, Solna, Stockholm, Sweden, [4] Karolinska Institutet, Inst of Environmental Medicine, Stockholm, Sweden, [5] Karolinska Instituet, Clinical Neuroscience, Stockholm, Sweden

Introduction

Previous studies have associated multiple sclerosis (MS) with several autoimmune diseases (ADs). However, comprehensive quantification of such association using well-powered nation-wide registers is currently lacking. To further estimate if MS share familial risk factors with ADs, we performed a population-based familial co-aggregation study.

Methods

In this case-control study nested within prospectively recorded Swedish national registers, we identified 26,086 MS cases in the Swedish MS Register or with at least 3 International Classification of Diseases (ICD) codes in the National Patient Register (NPR), and controls (1:10) matched on age, sex, and residential area from the General Population Register. First-degree relative pairs (N=111,386 for cases and N=1208,923 for controls) with their ICD codes for 22 ADs including type 1 diabetes, autoimmune thyroid disease, reactive arthritis, rheumatoid arthritis, ankylosing spondylitis, polyarteritis nodosa, giant cell arteritis, idiopathic thrombocytopenic purpura, systemic lupus erythematosus, systematic sclerosis, polymyalgia rheumatica, pemphigoid, psoriasis, vitiligo, acute disseminated encephalomyelitis, other acute widespread myelin damage, inflammatory polyneuropathy, myasthenia gravis, inflammatory bowel diseases, celiac disease, sarcoidosis, and glomerulonephritis were obtained by linkage of the Multi-Generation Register and the NPR. We estimated the odds ratios (ORs) of familial co-aggregation of MS and ADs using logistic regression. We modelled the familial co-aggregation combining all first-degree relatives and stratified by kinships. We performed the analysis for overall ADs of interest as well as for specific AD types.

Results

Familial co-aggregation was found between MS and any investigated ADs in overall first-degree relatives (OR=1.09, 95%CI=1.07−1.10), and was not modified by type of first-degree relatives. Among specific ADs, we observed the greatest familial co-aggregation of other acute widespread myelin destruction and MS across kinships, with ORs varying from 3.56 to 7.06, reflecting the validity of our study design. Ten of the 21 remaining ADs, including acute disseminated encephalomyelitis, celiac disease, inflammatory bowel diseases, type 1 diabetes, reactive arthritis, autoimmune thyroid diseases, psoriasis, polymyalgia rheumatica, systemic lupus erythematosus, and idiopathic thrombocytopenic purpura, showed modest familial co-aggregation with MS in at least one of the kinships, with ORs ranging from 1.10 to 1.88. These findings remain statistically significant after multiple testing correction.

Conclusions

We found moderate evidence of familial factors contributing to the co-aggregation of MS and other ADs. Our findings provided etiological insights into the shared genetic and environmental factors underlying autoimmunity, as well as highlighted having a first degree relative with ADs is a risk factor for MS, which may inform consultation in clinical practice.

Conflict of interest

None.

[18] GENETIC AND PHENOTYPIC ASSOCIATIONS BETWEEN THYROID AND REPRODUCTIVE HEALTH TRAITS

Jéssica Figueredo,[1], Kristi Krebs,[1], Triin Laisk,[1], Reedik Mägi,[1]

[1] *Institute of Genomics, University of Tartu, Estonian Genome Centre, Tartu, Estonia*

Introduction

The thyroid is a vital endocrine gland essential for the metabolism, growth, and development of the human body. Together with the hypothalamus and pituitary, thyroid composes the HPT axis, which synthesizes and releases the thyroid hormones (T3 and T4) responsible for regulating a variety of biological functions, including those related to reproductive health. When the thyroid hormone production is unbalanced, this can lead to hypothyroidism or hyperthyroidism, having a significant impact on reproductive hormones, such as sex-hormone binding globulin and testosterone levels, and can also affect fertility, and menstrual and sexual function. Since thyroid disorders result from the interaction of many genetic variants and are considered polygenic, the question remains whether the associations between thyroid and reproductive health are purely due to the interactions between the regulatory axes, or if in some cases shared genetic causes might also play a role. Investigating the shared genetics and genetic pleiotropy between thyroid function and reproductive health, would contribute to the understanding on the correlation between those traits.

Methods

We conducted large scale genetic analyses of thyroid traits (hypothyroidism, hyperthyroidism and TSH measurement levels) in up to 744,345 European ancestry individuals from different cohorts (Estonian Biobank, UK Biobank, FinnGen, HUNT, MGI and ThyroidOmics Consortium) to explore their associations with reproductive health traits, both on phenotypic and genetic level. For the genetic analyses assessment, we performed GWAS for each trait with the EstBB data and GWAS meta-analyses using GWAS summary statistics from the other biobanks/consortiums. Candidate genes for each genomic risk locus were determined using the following criteria: 1) if the gene is the nearest gene of the variant; 2) if there was significant (posterior probability ≥ 0.8) colocalisation between GWAS signal and gene eQTL; 3) if the variant function was missense; and 4) if the gene was associated with a thyroid related mouse phenotype or with monogenic thyroid diseases in humans. To evaluate the correlation between thyroid diseases and reproductive health phenotypes: (1) we checked the association

between the candidate genes with mouse phenotypes or monogenic diseases related to reproductive health; (2) performed genetic correlation analyses between our data and several GWAS summary statistics for reproductive hormones and phenotypes; and (3) performed a phenotype vs phenome-wide-association analysis with EstBB individual level data sex-stratified.

Results

In total, 200 genome-wide significant associations were obtained for the selected thyroid phenotypes, where 33/200 loci are novel in the context of thyroid biology, and 5/200 loci were previously reported to be associated with both thyroid and reproductive health. Using a complex candidate gene mapping approach, we identify numerous genes (228) as likely causal in the thyroid analyses that also have a clear role in reproductive health, providing evidence for pleiotropic effects. Our genetic correlation analysis showed that hypothyroidism correlated negatively with age at menopause and sex hormone (sex hormone binding globulin and testosterone) levels, and positively with genital tract disorders. Most of the genetic correlation results are in line with previous findings, however, some correlations, such as with female genital prolapse and redundant prepuce, phimosis and paraphimosis, need more investigation. In addition, phenotype-level analysis showed considerable overlap between thyroid phenotypes and genital tract disorders in women.

Conclusions

Our study confirms that thyroid function and reproductive health share common genetic determinants, and we provide the first roadmap to understanding the shared genetics and genetic pleiotropy between the traits, both in men and women.

Conflict of interest

None.

[19] GENOME WIDE ASSOCIATION STUDIES OF THE KYNURENINE PATHWAY

Chinh Nguyen,[1]

[1] *University of Bergen, Department of Clinical Science, Bergen, Norway*

Introduction

The kynurenine pathway of tryptophan catabolism plays a key role in physiological processes such as neuronal modulation, aging, and immunity. The pathway produces several bioactive catabolites and represents a functional convergence of amino acid auxotrophy and immune regulation. Nevertheless, genetic factors associated with kynurenine products and pathway regulation are largely uncharacterized and previous research has primarily used the candidate gene approach. Here we present results of genome-wide association studies (GWAS) between common variants and plasma levels of nine metabolites and seven biomarker-ratios of the kynurenine pathway using a community-based cohort, the Hordaland Health Study (HUSK).

Methods

We used data from 4871 unrelated participants of Norwegian ancestry (HUSK) and performed GWAS between single nucleotide polymorphisms (SNPs) and plasma levels of nine metabolites and seven metabolite-ratios measured by targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS). The sixteen metabolic traits include tryptophan (Trp) and its eight catabolites kynurenine (Kyn), kynurenic acid (KA), hydroxykynurenine (HK), hydroxy anthranilic acid (HAA), anthranilic acid (AA), xanthurenic acid (XA), picolinic acid (Pic), quinolinic acid (QA) as well as the seven biomarker-ratios: inflammaging-associated ratio Kyn/Trp (KTR), neuromodulation (KA/QA, Pic/QA), vitamin B6 status (HKr, HKr_a and HKr_b) and pathway-regulation ratio Kyn/KA. We conducted cross-phenotype analysis by interrogating the identified metabolite quantitative loci (mQTLs) using Phenoscanner and SNiPA databases for previously published phenotypic trait-GWAS and performed colocalization analysis of biomarker-associated loci with cis-eQTL effects by using the largest blood-eQTL database (eQTLGen).

Results

We identify 23 independent mQTLs (p < 5E-8) using 7,5 million SNPs (MAF>=1%) and 16 metabolic traits. Functional enrichment of these SNPs confirms the involvement of the

pathway in immune regulation and suggests unrecognized links with fatty acid oxidation, phase II transformation, urate levels and efflux processes. Our study provides a comprehensive genetic catalogue for relevant enzymes and transporters of the pathway, key factors responsible for pathway regulation, biomarker-associated variants, and emerging regulators of amino acid metabolism-immune crosstalk.

Conclusions

This is the first metabolomic GWAS to reveal genetic determinants of a large panel of important tryptophan catabolites and metabolite-ratios in relation to disease risks including inflammaging, neuromodulation and vitamin B6 status. These metabolite-associated loci belong to not only pathway reactions but also important physiological processes including lipid metabolism, phase II transformation, efflux, metabolic-immune crosstalk, and immune homeostasis. This study provides important insights into relevant genetic factors influencing levels of circulating kynurenines and opens opportunities for therapeutic interventions targeting key regulatory nodes of the pathway.

Conflict of interest

None.

[20] GENOME WIDE ASSOCIATION STUDY OF CLINICALLY PREDICTED SUICIDE LIABILITY

Anders Jespersen,[1], Oleguer Plana-Ripoll,[2], Jette Steinbach,[1], Clara Albiñana,[1], John Mcgrath,[1], Preben Mortensen,[1], Florian Privé,[1], Esben Agerbo,[1], Bjarni Vilhjálmsson,[1]

[1] Aarhus University, Ncrr, Denmark, [2] Aarhus University, Department of Clinical Epidemiology, Denmark

Introduction

Suicide accounts for 1 in every 100 deaths worldwide, resulting in more than 700,000 deaths annually. Globally the number of suicides has decreased by 36% in the last 20 years whilst the Americas have seen a 17% increase in deaths by suicide. Although there is considerable evidence that attempted suicide and death by suicide have distinct etiologies, they share risk factors, which include previous suicide attempt, depression, and alcohol use disorders. Heritability estimates of suicidal behavior range from 17% to 55% suggesting a significant genetic component to suicide risk. However, as suicides are fortunately relatively uncommon, genetic studies of suicidal behaviour suffer from insufficient number of genotyped cases and are generally underpowered.

Methods

In this study, we aim to address this limitation by leveraging Danish electronic health care records to perform a GWAS of clinically predicted suicide liability. We will achieve this by using health registers containing all hospital admissions, diagnoses, and prescriptions of the Danish population, to clinically predict death by suicide. In order to do this, we will first use a Danish population sample comprising 323,000 individuals, 869 of whom died by suicide, to train a penalized regression. The resulting weights will be used to clinically predict suicide liability in an independent sample of 134,000 genotyped individuals. We then perform a GWAS of the clinically predicted suicide liability and compare the genetic overlap with the most recent case-control GWAS of death by suicide (Docherty et al. 2020), as well as genetic correlations between clinically predicted suicide liability and major psychiatric disorders.

Conclusions

We expect results from this study to shed light on the contribution of common genetic variants to the clinically predicted risk of death by suicide. The study will demonstrate the level of

genetic similarity between clinically predicted death by suicide and observed death by suicide. Finally, this method of leveraging electronic health records for clinical prediction to increase power in genetic studies could pave the way for genetic studies of phenotypes otherwise suffering from low numbers of genotyped cases.

Conflict of interest

None.

[21] GENOME-WIDE ASSOCIATION STUDY IDENTIFY NEW GENETIC DETERMINANTS OF RESTING HEART RATE

Marie Klevjer,[1], Ada Nilsen Nordeidet,[2], Anja Bye,[3]

[1] Ntnu, Department of Medicine and Health, Trondheim, Norway, [2] Norwegian University of Science and Technology (Ntnu), Ntnu, Department of Circulation and Medical Imaging, Trondheim, Norway, [3] Norwegian University of Science and Technology, Ntnu, Circulation and Medical Imaging, Trondheim, Norway

Introduction

Resting heart rate (RHR) is an independent risk factor for cardiovascular disease (CVD), both for healthy and diseased individuals. RHR is defined as the number of times your heart beats at rest and is an objective measurement easily accessible for clinicians and individuals themselves. In general, a lower RHR implies a more efficient heart function and better cardiovascular fitness. The genetic contribution of RHR is estimated to be up to ~65% in twin-studies. Previous genome-wide association studies (GWAS) have identified up to 437 independent loci associated with RHR.

Methods

We performed GWAS of RHR using BOLT-LMM on two study populations, The health study in Trøndelag (HUNT) and UK Biobank (UKBB), including 69,155 and 481,312 participants, respectively. A meta-analysis of the two GWASes was performed using METAL, with 44,789,347 single nucleotide polymorphisms (SNPs), with minor allele count ≥ 10. We also conducted sex-specific analyses that will be used to look for differences in the genomic architecture between sexes. Functional mapping and annotation of the GWAS results will be conducted using FUMA.

Results

In the meta-analysis of the total population, we identified 403 independent SNPs significantly associated with RHR, including 51 SNPs that not previously have been reported to be associated with RHR. In the male population, there were 168 independent significant SNPs, and in the female population there were 219 independent significant SNPs. These SNPs were significantly related to 311, 145, and 219 genes in the total, male, and female population, respectively. In the total population, there was a significant overrepresentation of genes related to

the atrial appendage of the heart, the left ventricle, the gastroesophageal junction, esophagus muscularis, uterus, sigmoid colon, and skeletal muscle. In the male population, the associated genes were significantly overrepresented in the atrial appendage of the heart, and in the left ventricle. In the female population, the associated genes were significantly overrepresented in the atrial appendage of the heart, left ventricle, and skeletal muscle.

Conclusions

We have conducted the largest GWAS on RHR to date. We have verified 352 previously identified SNPs and have discovered 51 novel SNPs significantly associated with RHR in the total population. By performing sex-specified analysis we hope to shed light over the biological differences between sexes.

Conflict of interest

None.

[22] GLOBAL LONG COVID HOST GENETICS INIATIVE IDENTIFIES FOXP4 LOCUS AS THE FIRST GENETIC RISK FACTOR ASSOCIATED TO LONG COVID

Minttu Marttila, [1]

[1] *University of Helsinki, Helsinki, Finland*

Global Long COVID Host Genetics Initiative Identifies FOXP4 locus as the first genetic risk factor associated to long COVID

Marttila, M. [1]*, Lammi, V.*[1]*, Jones, S.E.*[1]*, Nakanishi, T.*[2,3]*, Ollila, H.M.* [1,4] *Long COVID Host Genetics Initiative, 1 Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland, 2 McGill University, Montreal, Canada, 3 Kyoto University, Kyoto, Japan, 4 Broad Institute of MIT and Harvard, MA, USA*

Introduction

While most people recover from COVID-19 within 3 months, a substantial proportion suffer for long term condition or so called "long COVID". Most common symptoms include fatigue, cognitive dysfunction, muscle soreness and breathlessness, although symptoms vary widely. To understand the disease etiology and genetic liability behind long COVID we build a global working group under the COVID-19 Host Genetics Initiative.

Materials and Methods

At present (data freeze 3) 24 studies from 16 countries representing 6 genetic ancestry groups contributed data to our meta-analysis. Long COVID was captured using questionnaires, with long COVID being identified as 'any diagnosis that cannot be explained by alternative diagnosis' 3 months after the onset of COVID-19. In studies with electronic health record (EHR) data, specific diagnosis codes for post COVID condition (ICD-10 U09 or SNOMED) were used.

Results and discussion

We performed GWAS meta-analysis comparing long COVID cases (6450) to population controls (46208) and identified a genetic risk locus on chromosome 6 upstream of the *FOXP4* gene (rs9367106, P=7,3 e-9). Our early findings indicate that *FOXP4* may affect successibility to long COVID. *FOXP4* is an RNA gene located in the short arm of chromosome 6,

coding for FOXP4 protein. It belongs to the FOX transcription factor family. COVID-19 associated variants in *FOXP4* have been shown to modify gene expression in the lung.

Conflict of interest

None.

[23] HETEROGENEITY OF UNKNOWN AUTO-INFLAMMATORY DISEASES

Sara Ludovico Garcia[1], Andres Jimenez Kaufmann[1], Isabella Friis Jørgensen[1], Immunome Project Consortium for Autoinflammatory Disorders (ImmunAID); Søren Brunak[1]

[1]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Introduction

Systemic autoinflammatory diseases (SAIDs) are a group of rare conditions where there is an inappropriate inflammatory response by the body. These comprise monogenic familial disorders, where genetic mutations have been identified, and phenotype-based systemic inflammatory disorders. Most SAIDs are characterized by recurrent episodes of unexplained fevers, and the average age of onset varies from infancy to mid-childhood[1]. The Immunome project consortium for AutoInflammatory Disorders (ImmunAID)[2] aims to build the world's largest multi-omics cohort on SAIDs and improve the diagnosis, medical care, and quality of life of these patients, through a joint international effort of clinical centers and data analysts. One of the biggest challenges is inflammation of unknown origin (IUO). A lot of patients with IUO remain undiagnosed after several years and few develop complications or die from empirical treatment[3]. In this study, we aim to subgroup these patients into more homogeneous groups by finding specific differential biological mechanisms and/or pathways.

Methods

Liquid chromatography mass spectrometry-based proteomics and whole-genome sequencing data were analyzed for 12 and 24 patients, respectively, with IOU. First, ComBat[4] was used to correct the batch effect on the proteomics data, and the expression profile of this group of patients was analyzed. For each trio (N=4) and duo (N=6), de-novo variants were investigated using SLIVAR[5].

Results

In these preliminary results, it was not detected de-novo variants in this group of patients. As for the proteomics data, several proteins involved in blood clotting were highly expressed, such as prothrombin, plasminogen, and antithrombin-III (AT), which apart from being involved in blood clotting, also elicits distinct anti-inflammatory signaling responses[6].

Conclusions

There are few studies on IUO, and to our knowledge, this will be the first one trying to better understand biological mechanisms in IUO, and if this is associated with different clinical manifestations. Our goal is to, later, validate these results in a Danish dataset[7], where we have access to previous diagnoses of patients with SAIDs. Furthermore, we aim to integrate multi-omics data (proteomics, genomics, and metagenomics), which is available for the ImmunAID cohort, using known methods such as MOFA+[8]. Finally, it is important to mention that, on average, up to five inadequate treatments are given to SAID patients before a correct diagnosis, thus, given the high heterogeneity found even within a known disease, this pipeline can later be applied to other diseases included in the ImmunAID cohort. This could help to elucidate and/or discover new markers specific to each disease.

(1) Vanderschueren, S.; Del Biondo, E.; Ruttens, D.; Van Boxelaer, I.; Wauters, E.; Knockaert, D. D. C. Inflammation of Unknown Origin versus Fever of Unknown Origin: Two of a Kind. Eur. J. Intern. Med. **2009**, 20 (4), 415–418. https://doi.org/10.1016/j.ejim.2009.01.002.

(2) IMMUNAID Immunome project consortium for AutoInflammatory Disorders. https://www.immunaid.eu/ (accessed 2022-10-13).

(3) Béra, S.; Jamilloux, Y.; Gerfaud-Valentin, M.; Durupt, S.; Nove-Josserand, R.; Lega, J.-C.; Durieu, I.; Hot, A.; Sève, P. Inflammation of Unknown Origin: Evaluation and Prognosis of 57 Cases. J. Clin. Med. **2021**, 11 (1), 32. https://doi.org/10.3390/jcm11010032.

(4) ComBat-seq: batch effect adjustment for RNA-seq count data | NAR Genomics and Bioinformatics | Oxford Academic. https://academic.oup.com/nargab/article/2/3/lqaa078/5909519 (accessed 2022-10-13).

(5) Effective variant filtering and expected candidate variant yield in studies of rare human disease | npj Genomic Medicine. https://www.nature.com/articles/s41525-021-00227-3 (accessed 2022-10-13).

(6) Schlömmer, C.; Brandtner, A.; Bachler, M. Antithrombin and Its Role in Host Defense and Inflammation. Int. J. Mol. Sci. **2021**, 22 (8), 4283. https://doi.org/10.3390/ijms22084283.

(7) Schmidt, M.; Schmidt, S. A. J.; Sandegaard, J. L.; Ehrenstein, V.; Pedersen, L.; Sørensen, H. T. The Danish National Patient Registry: A Review of Content, Data

Quality, and Research Potential. Clin. Epidemiol. **2015**, 7, 449–490. https://doi.org/10.2147/CLEP.S91125.

(8) Argelaguet, R.; Arnol, D.; Bredikhin, D.; Deloro, Y.; Velten, B.; Marioni, J. C.; Stegle, O. MOFA+: A Statistical Framework for Comprehensive Integration of Multi-Modal Single-Cell Data. Genome Biol. **2020**, 21 (1), 111. https://doi.org/10.1186/s13059-020-02015-1

[24] HLA ASSOCIATIONS BETWEEN EPSTEIN-BARR VIRUS AND MULTIPLE SCLEROSIS.

Jesse Huang,[1], Katarina Tengvall,[2], Izaura Bomfim Lima,[2], Anna Karin Hedström,[2], Angelika Michel,[3], Julia Butt,[3], Nicole Brenner,[3], Alexandra Gyllenberg,[2], Pernilla Stridh,[2], Mohsen Khademi,[2], Sigurgeir Olafsson,[4], Ingileif Jonsdottir,[4], Kári Stefansson,[4], Ingemar Ernberg,[2], Ali Manouchehrinia,[2], Jan Hillert,[2], Lars Alfredsson,[2], Tim Waterboer,[3], Tomas Olsson,[2], Ingrid Kockum,[2]

[1] Karolinska Institutet, Clinical Neuroscience, Solna, Sweden, [2] Karolinska Institutet, Sweden, [3] German Cancer Research Center (Dkfz), Germany, [4] Decode Genetics, Iceland

Introduction

To evaluate host genetic factors that influence serological response against Epstein-Barr virus (EBV) and history of infectious mononucleosis (IM) and cross-evaluate its association with multiple sclerosis (MS) risk and genetic susceptibility.

Methods

IgG antibody levels against EBNA-1 (truncated=aa[325-641], peptide=aa[385-420]) and VCAp18 were determined using bead-based multiplex serology for 8744 MS cases and 7229 population-matched controls. Genome-wide and imputed classical HLA allele association analyses were conducted for EBV antibody levels.

Results

Increased levels of antibodies against VCAp18 (OR=1.74, 95% CI=1.60-1.88) and EBNA-1, particularly the peptide (OR=3.13, 95% CI=2.93-3.35), were associated with an increased risk for MS. Risk increased linearly with rising anti-EBNA-1 IgG levels up to an OR~12. We identified several independent HLA haplotypes associated with EBV serology that overlap with known MS-associated HLA risk alleles (e.g., DRB1*15:01, DRB1*03:01). Weighted genetic risk scores (wGRS) for anti-EBNA-1 IgG levels, particularly the peptide fragment, were strongly associated with MS, while wGRS for MS were also associated with EBV antibody levels to a lesser extent. In contrast, wGRS for anti-VCAp18 IgG levels was not associated with increased MS risk.

Conclusions

Our findings emphasize the increased MS risk associated with anti-EBNA-1 IgG levels, likely related to overlapping HLA risk haplotypes.

Conflict of interest

Tomas Olsson has received lecture and/or advisory board honoraria, and unrestricted MS research grants from Astrazeneca, Biogen, Novartis, Merck, Roche, Almirall and Genzyme.

[25] IDENTIFICATION OF NOVEL GENOMIC RISK LOCI SHARED BETWEEN COMMON EPILEPSIES AND MAJOR PSYCHIATRIC DISORDERS

Naz Karadag,[1], Alexey Shadrin,[2], Kevin S O'connell,[2], Guy Hindley,[3], Zillur Rahman,[4], Nadine Parker,[1], Shahram Bahrami,[2], Vera Fominykh,[5], Weiqiu Cheng,[2], Børge Holen,[2], Silje Alvestad,[6], Erik Taubøll,[7], Nils Eiel Steen,[8], Srdjan Djurovic,[9], Anders M Dale,[10], Oleksandr Frei,[11], Ole A Andreassen,[12], Olav Smeland,[13]

[1]University of Oslo, Norment, Institute of Clinical Medicine, Oslo, Norway, [2] Norment, University of Oslo, Norment, Institute of Clinical Medicine, Oslo, Norway, [3] University of Oslo, University of Oslo, Norment, Oslo, Norway, [4] Oslo University Hospital, University of Oslo, Norment, Institute of Clinical Medicine, Oslo, Norway, [5] University of Oslo, Oslo University Hospital, Norment, Institute of Clinical Medicine, Oslo, Norway, [6] Oslo University Hospital, University of Bergen, Norway, [7] Oslo University Hospital - Rikshospitaletepilepsy Unit, Oslo University Hospital - Rikshospitalet, Oslo, Norway, [8] University of Oslo, [9] Norment, Oslo University Hospital, [10] University of California San Diego, [11] Norment, Center for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway, Norway, [12] Norment, Norway, [13] Institute of Clinical Medicine, University of Oslo, Norment, Oslo, Norway

Introduction

Epilepsy and psychiatric disorders are heritable brain disorders with overlapping clinical features and high comorbidity. However, the causative mechanisms underlying the relationship between these disorders are poorly understood. Here we aimed to identify overlapping genetic loci between epilepsy and psychiatric disorders to gain a better understanding of their comorbidity and shared clinical features.

Methods

Here, we analyzed non-overlapping genome-wide association study (GWAS) data on all epilepsies (n = 44,889), and its subtypes genetic generalized epilepsy (n = 33,446), and focal epilepsy (n = 39,348), and the psychiatric disorders schizophrenia (n = 77,096), bipolar disorder (n = 413,466), depression (n = 500,199), attention-deficit hyperactive disorder (n = 53,293) and autism spectrum disorder (n = 46,350). We analyzed GWAS summary data using the conjunctional false discovery rate (conjFDR) statistical tool to increase power for locus discovery. Identified genetic loci were then functionally annotated using FUMA. To test the validity

of our findings, we performed sign concordance tests between the discovery and independent datasets.

Results

We observed cross-trait enrichment between genetic generalized epilepsy and all psychiatric disorders as well as between all epilepsies and schizophrenia and depression. We identified 39 distinct loci associated with genetic generalized epilepsy, and four loci associated with all epilepsies; totaling 40 distinct loci jointly associated with epilepsy and psychiatric disorders. Of these, 30 loci were novel findings for genetic generalized epilepsy and one locus was novel for all epilepsies. Consistent sign concordance between the discovery and independent datasets for all disorders supports the validity of these findings. We observed mixed concordance of allelic effects in the shared loci, in line with minimal genome-wide genetic correlations. Gene-set analysis for the shared loci between genetic generalized epilepsy and schizophrenia have implicated several processes related to cell cycle regulation, protein phosphatase activity, and membrane and vesicle function. The gene-set analyses for the other groups of loci were underpowered.

Conclusions

The extensive genetic overlap with mixed effect directions between psychiatric disorders and common epilepsies demonstrates a complex genetic relationship between these disorders, in line with their bidirectional relationship, and indicates that overlapping genetic risk may contribute to shared pathophysiological and clinical features between epilepsy and psychiatric disorders.

Conflict of interest

O.A.A. has received speaker's honorarium from Lundbeck and is a consultant for Healthlytix. A.M.D. is a founder of and holds equity interest in CorTechs Labs and serves on its scientific advisory board. He is also a member of the Scientific Advisory Board of Healthlytix and receives research funding from General Electric Healthcare (GEHC). The terms of these arrangements have been reviewed and approved by the University of California, San Diego in accordance with its conflict-of-interest policies. Remaining authors have no conflicts of interest to declare.

[26] IDENTIFYING CAUSAL GENETIC VARIANTS IN DIABETES- AND ADIPOSITY-ASSOCIATED GENES USING CRISPR AND ISOGENIC CELL LINES

MI Huang,[1], Sebastian Kalamajski,[1], Hindrik Mulder,[1], Paul Franks,[1]

[1]Lund University, Clinical Sciences Malmö, Malmö, Sweden

Introduction

Most of the trait-associated single-nucleotide polymorphisms (SNPs) discovered by GWAS studies are representatives of larger haploblocks within which other, potentially causative SNPs may reside. Therefore, pinpointing the SNPs that impose biologically significant changes on gene expression and cell biology requires validation by experimental means. In this endeavor, CRISPR toolkit provides several methods that allow single-nucleotide, scarless genomic editing, allowing allele substitutions in isogenic cell lines. Any allele-dependent effects on cell biology can then be assessed in a controlled genetic background, which could aid in finding the SNPs germane to precision medicine treatment.

Methods

Three common human genetic variants in *MTIF3, PPARGC1A,* and *PPARGC1B,* previously associated with diabetes or adiposity, were investigated for their causality using CRISPR-mediated allele substitution in preadipocytes. Two SNPs (rs8192678, rs10071329) were the only common variants in their haploblocks, while the third SNP (rs67785913) was linked to a GWAS index SNP and deemed potentially causal using a haploblock tiling luciferase reporter screen. The CRISPR-edited preadipocytes, homozygous for the respective SNP alleles, were differentiated, and their gene expression (allele-mediated cis-eQTL) was assessed using qPCRs. Furthermore, the cellular phenotype related to adipocyte or mitochondrial functions was examined.

Results

The alleles at *MTIF3* rs67785913 and at *PPARGC1B* rs10071329 differentially alter the expression of the respective genes, as observed by quantitative PCR. The rs10071329 SNP also affects adipogenic differentiation, norepinephrine effect on lipolysis, and mitochondrial content. The *PPARGC1A* rs8192678 alleles, encoding different amino acids, impact on adipogenic differentiation and adipogenesis in an allele dose-dependent manner. The alleles also

differentially impact the transcriptional activity of the *PPARGC1A*-encoded protein PGC-1alpha.

Conclusions

We have conducted functional genomics studies using different CRISPR methods to test the causality (eQTL) of diabetes- or weight loss-linked genetic variants. By performing allele substitution in human preadipocyte cell lines, we have established a biological function of several trait-linked SNPs in regulating gene expression and cellular function. The translatability of these results could later be tested in genotype-based recall trials, which could aid in establishing novel precision medicine treatments.

Conflict of interest

None.

[27] IMPLEMENTATION OF POLYGENIC RISK SCORE GUIDED BREAST CANCER PREDICTION IN NORWEGIAN DATA

Bayram Cevdet Akdeniz,[1], Morten Mattingsdal,[2], Mev Dominguez Valentin,[3], Oleksandr Frei,[4], Alexey Shadrin,[5], Mikk Puustusmaa,[6], Regina Saar,[6], Siim Sõber,[7], Ole A Andreassen,[8], Peeter Padrik,[6], Eivind Hovig,[9]

[1] Center for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway, Norwegian Centre for Mental Disorders Research (Norment), Division of Mental Health and Addiction, Oslo University Hospital and University of Oslo, Oslo, Norway, Norway, [2] Centre for Coastal Research, Department of Natural Sciences, University of Agder, Kristiansand, Norway, Center for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway, Norway, [3] Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway, Norway, [4] Norment, Center for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway, Norway, [5] Norment, University of Oslo, Norment, Institute of Clinical Medicine, Oslo, Norway, [6] Oü Antegenes, Tartu, Estonia, Estonia, [7] Oü Antegenes, Tartu, Estonia, Norway, [8] Norment, Norway, [9] University of Oslo, Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway, Centre for Bioinformatics, Oslo, Norway

Introduction

Statistical associations of numerous SNPs with breast cancer (BC) have been discovered and robustly replicated as a result of different genome-wide association studies (GWAS). Such findings lead to the development of different risk monitoring methods with Polygenic Risk Score (PRS). A recent study regarding the application of PRS models was done by Antegenes, using data from the Estonian Genome Center (EGC) and UK Biobank (UKB) [medRxiv 2020.08.17.20176263]. Our main aim in this study is to characterize the performance of PRS models presented by Antegenes on Norwegian Data in order to advance development of breast cancer prediction/prevention procedure for Norwegian cohorts.

Methods

We have collected genomic and phenotypic data of 18,000 individuals from a geographically scattered sample set from Norway. Once the imputation of the genomic data was done, then we have calculated PRS of each individual and evaluated the performance with different metrics. In order to perform those analysis, we have used container-based data analysis platform

already developed for CoMorMent (https://github.com/comorment/containers).

DNA samples was extracted and genotyped using the Illumina OmniExpress 24 v 1.1 chip at deCODE genetics. The genotyped dataset with 713,014 SNPs were stored and processed at TSD platform at the University of Oslo (N≈15k). We used DRAGEN-GATK pipeline on whole genome sequencing data from 1368 individuals to create the Norwegian Reference panel [Mattingsdal, M. *et al.* "The genetic structure of Norway"]. For phenotype data, 16,029 Norwegian individuals were included and 9,201 had a diagnosis of cancer while 1690 of them had BC.

The imputation was conducted using different pipelines and different reference panels (RPs). The aim was to see the effects of RPs when performing risk stratification of Breast Cancer. We imputed data using Norgene, MoBa [bioRxiv 2022.06.23.496289] and Antegenes pipelines employing Norwegian, HRC and 1kgp RPs respectively.

We chose various available PRS models available for BC in the literature. We calculated PRS using PRSICE 2, and the obtained individual risk scores were standardized to Normal distribution. Next, we evaluated the relationship between BC status and standardized PRS of the dataset with a logistic regression model to estimate the odds ratio per 1 standard deviation of PRS (OR) and Area Under the ROC Curve (AUC). Hazard ratios per 1 unit of standardized PRS (HR) were computed using a right-censored and left-truncated Cox regression survival model. The start of time interval was defined as the age of recruitment; follow-up time was set as the time of diagnosis for cases and at the time of last health data linkage for controls.

Results

We calculated AUC, OR, HR and c-index of PRS obtained from Norwegian Data with different imputation pipelines and compared with the results already obtained from EGC and UKB (Table 1). Norwegian data yields similar AUC and OR values as Antegenes obtained for EGC as well as UKB data. Norwegian RP provides better AUC, OR and HR values compared to other RPs.

Metrics	SNPSET	Antgegenes Pipeline (1kgp)	MoBa (HRC)	Norgene (Norwegian ref. panel)	EGC (with Estonian	UKB

					ref. panel)	
AUC (# of SNPs used)	77 SNPs (PMID: 25855707)	0.582 (66)	0.584 (66)	0.600 (66)	0.591 (73)	0.607
	313 SNPs (PMID: 30554720)	0.605 (232)	0.605 (210)	0.617 (228)	0.604 (257)	0.625
	3102 SNPs (from Ante-genes)	0.616 (2511)	0.617 (2379)	0.618(2474)	0.615(2802)	0.632
	3820 SNPs (PMID: 30554720)	0.621 (2706)	0.620 (2482)	0.625 (2698)	0.611(3081)	0.632
OR (SE)	77 SNPs (PMID: 25855707)	1.460 (0.034)	1.465 (0.034)	1.503 (0.033)	1.37 (0.06)	1.48(0.01)
	313 SNPs (PMID: 30554720)	1.465 (0.034)	1.475 (0.034)	1.534 (0.034)	1.43 (0.06)	1.43(0.01)
	3102 SNPs (from Ante-genes)	1.526 (0.34)	1.538 (0.034)	1.534 (0.034)	1.479 (0.06)	1.616(0.01)
	3820 SNPs (PMID: 30554720)	1.546 (0.034)	1.528 (0.034)	1.567 (0.034)	1.473 (0.06)	1.617(0.01)
HR (%95 confi-dence in-terval)	3820 SNPs	1.462 (1.375-1.554)	1.458 (1.371-1.55)	1.494 (1.406-1.588)	1.66 (1.5 – 1.84)	1.56 (1.53-1.6)
c-index	3820 SNPs	0.602	0.601	0.607	0.654	0.625

		(se=0.01)	(se=0.01)	(se= 0.01)	(se=0.01 5)	(se=0.01)

Table 1. The PRS performance of Norwegian Data (N≈1,487 for cases and N≈8,857 for controls) for BC with different imputation pipelines and their comparison with EGC and UKB.

Conclusions

The obtained results imply that available PRS models can be applicable to Norwegian data. The investigated PRS models show consistent performance with the results from EGC and UKB data. Our next aim is developing absolute risk estimation procedure combining PRS models presented here and available risk models in the literature such as iCare (PMC7001949).

Conflict of interest

None.

[28] IMPROVED RARE AND LOW-FREQUENCY VARIANT IMPUTATION WITH THE NOVEL SISU 4.2 FINNISH POPULATION-SPECIFIC GENOTYPE IMPUTATION REFERENCE PANEL.

Kalle Pärn,[1], Shuang Luo,[2], Samuli Ripatti,[2], Mark Daly,[2], Aarno Palotie,[2], Priit Palta,[2]

[1] University of Helsinki, Fimm/Finngen, Helsinki, Finland, [2] University of Helsinki, Fimm, Finngen, Helsinki, Finland

Introduction

Genotype imputation has become a standard tool in human genetics and is an essential step before carrying out genome-wide association studies (GWAS) and/or meta-analysis. Although there are many publicly accessible imputation reference panels (IRPs) available that allow imputation with satisfactory results, it has been demonstrated that imputation quality can be increased further, especially in the rare and low-frequency variants, by applying a population-specific IRP on the target data.

Hereby we introduce the novel SISu 4.2 IRP which combines the increased sample count of SISu 4.0 with the more relaxed approach to call-rate filtering and the inclusion of LCR and PAR regions to further improve the imputation quality and the overall coverage of the population-specific IRP.

Methods

SISu 4.2 SNV/indel IRP is based on high-coverage (~25X) WGS data for 10,500 Finnish samples, obtained from 5 different cohorts. The initial processing and quality control steps were performed, after which high-quality genotype data with call rate > 0.9 for 8,554 individuals and 41,1M SNV and indel variants remained in the data, including the newly added LCR and chromosome X PAR regions. After the application of AC >=3 filter, the resulting 21,3M variants were phased into the final IRP, increasing the total number of variants by 1,136,488.

For the assessment the imputation quality of the IRP, we masked the WGS data for 1,209 samples to obtain positions data comparable to a QC-d FinnGen Affymetrix genotyping chip. We then used the SISu 4.0 and 4.2 IRPs together with the masked set to perform genotype imputation. Concordance analysis was performed by comparing the masked-imputed set against

the full WGS data containing the same set of samples with variant-wise INFO threshold of >=0.6.

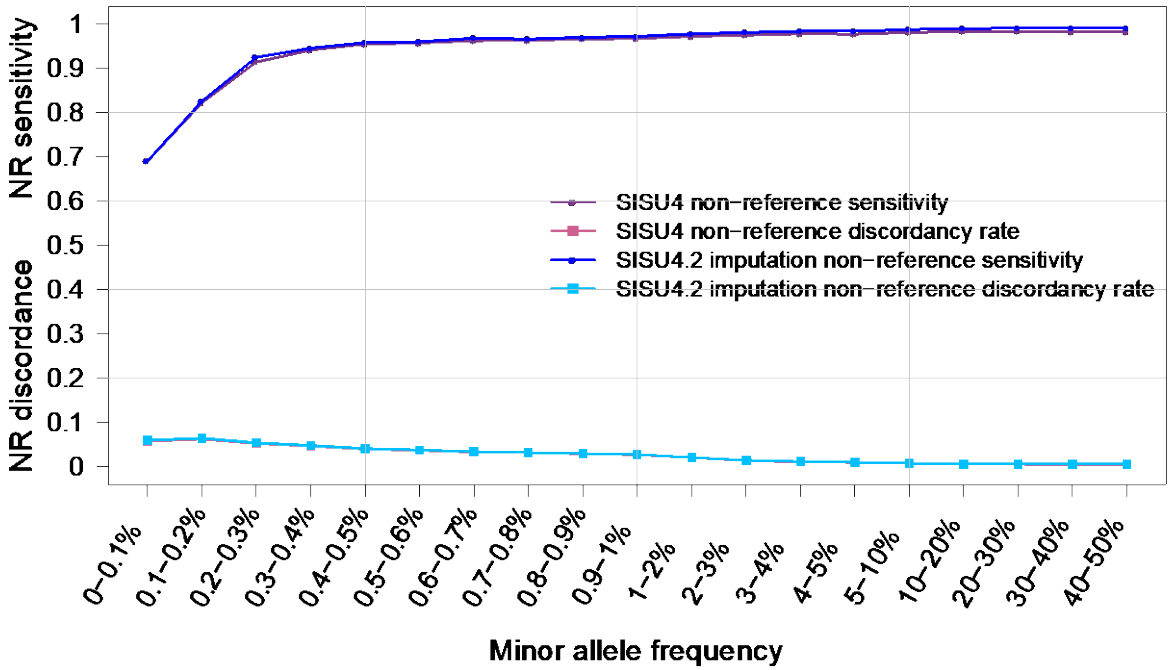
Results

In order to determine the imputation quality of the SISu 4.2 IRP, we compared the results of the concordance analysis against a similar analysis performed with its predecessor, SISu 4.0 (Figure 1). The proportion of recovered variants (NR sensitivity) has increased across all the studied allele frequency bins while the proportion of incorrectly imputed genotypes (NR discrepancy) has remained similar to the values obtained with the previous IRP. Especially so in the 0.2-0.3% frequency bin, where we achieved 1.1% increase in NR sensitivity with only 0.1% increase in NR discrepancy. These results demonstrate that we have managed to further increase the high quality of our Finnish population specific IRP-s, especially considering the rare variants, and that relaxing the call-rate threshold did not result in any adverse effects on the overall quality.

Conclusions

It can be concluded that the SISu 4.2 IRP allows us to recover a greater proportion of variants when compared to the previous iteration, most significantly in the rare allele frequency region, without resulting in increases in the error rate. Rare variant imputation is something that is difficult to reliably achieve with international panels: while in many studies the lower allele frequency threshold has been 1-5%, our IRP allows us to confidently impute well below the 1% limit. This demonstrates the importance of the population-specific IRPs, especially in genetically less diverse regions such as Finland. These results also provide grounds to further look into the quality of imputation with our latest IRP, by focusing more specifically on areas that are difficult to impute (rare variants) or difficult to call (LCR region).

Figure 1. *The proportion of all imputed non-reference genotypes (NR sensitivity) and incorrectly imputed genotypes (NR discordance) across the allele frequency spectrum for the previous (SISu v4.0) and the new (SISu v4.2) imputation reference panel. (Figure on next page).*



Conflict of interest

None.

[29] MENDELIAN RANDOMIZATION STUDY OF BIRTH WEIGHT AND RISK OF PSYCHIATRIC DISORDERS LATER IN LIFE

Xiaoping Wu,[1], Frank Geller,[1], Ipsych Group .,[2], Bjarke Feenstra,[1]

[1] Statens Serum Institute, Epidemiology, Copenhagen, Denmark, [2] .

Introduction

Low birth weight has been associated with a higher risk of psychiatric disorders later in life. It is however not clear what the causal mechanisms are that underlie this relationship. In this study, we investigate whether variation in fetal growth has a direct causal effect on mental health later in life.

Methods

We used birth weight as a proxy measure for fetal growth and first performed logistic regression analyses to assess associations between observed birth weight and later diagnosis of psychiatric disorders in the Danish Initiative for Integrative Psychiatric Research (iPSYCH) study. Next, we leveraged publicly available genome-wide association study (GWAS) summary statistics to construct polygenic scores (PGS) reflecting the effect of fetal genetic variation across the genome on birth weight and test for association with psychiatric disorders. Then, using the individual-level data from the iPSYCH cohort, we performed one-sample Mendelian randomization (MR) analyses with 87 single-nucleotide polymorphisms as instrumental variables by using the two-stage least-squares (2SLS) method. Finally, we performed two-sample MR analyses based on the largest GWAS meta-analyses available for the six psychiatric disorders.

Results

1) ***Observed birth weight associated with psychiatric disorders.*** Higher observed birth weight was associated with lower risk of several psychiatric disorders, including ADHD (OR = 0.89, [95% CI 0.88-0.91], P =3.06×10^{-30}), major depression (OR = 0.86, [95% CI 0.85-0.88], P =1.05×10^{-56}), schizophrenia (OR = 0.83, [95% CI 0.80-0.86], P =8.76×10^{-23}), bipolar disorder (OR = 0.86, [95% CI 0.83-0.90], P =1.04×10^{-10}), and anorexia nervosa (OR = 0.96, [95% CI 0.93-0.998], P =0.04). There was no association between observed birth weight and ASD (OR = 0.99, [95% CI 0.97-1.02], P =0.62).

2) ***Polygenic score of birth weight associated with psychiatric disorders.*** Linear regression of Z-scores of observed birth weight on deciles of PGS for birth weight showed increasing patterns for the controls as well as for the six different psychiatric disorders. Of note, for ADHD, major depression and schizophrenia the observed birth weight for a given PGS decile appeared to be lower compared to the control group. Increasing PGS for birth weight based on fetal genetic associations showed negative association with ADHD (OR = 0.86 per PGS standard deviation, [95% CI 0.79-0.93], P $=1.08\times10^{-4}$) and major depression (OR = 0.92, [95% CI 0.86-0.99], P = 0.02). For ASD there was a positive association (OR = 1.09, [95% CI 1.002-1.18], P =0.04).

3) ***One-sample and two-sample Mendelian Randomization.*** The one-sample MR analyses showed no evidence (all with $p > 0.05$) for a causal relationship of birth weight on any of the six psychiatric disorders. In the two-sample MR analyses, we also found little evidence of a causal effect of birth weight on any of the six psychiatric disorders. A series of sensitivity analysis did not substantially change the association results. The MR-PRESSO global test suggested some degree of horizontal pleiotropy effect for all the six two-sample Mendelian randomization analysis. The outlier test detected different numbers of outliers (0 - 9). However, the distortion test showed that removing those outliers led to no substantial difference in causal estimates.

Conclusions

Our study did not find evidence for a direct causal effect of fetal growth (as proxied by birth weight) on the risk of psychiatric disorders later in life. Other explanations for the observed epidemiological associations are necessary. Potential scenarios could involve effects of an adverse intrauterine environment acting independently on both fetal growth and neurodevelopment, or genetic variants in the maternal genome that result in reduced fetal growth and when passed on to the offspring genome pleiotropically increases the risk of psychiatric disorders.

Conflict of interest

None.

[30] MULTI-OMIC PROFILING OF SPONTANEOUS MIGRAINE ATTACKS TREATED WITH A TRIPTAN

Lisette Kogelman,[1], Katrine Falkenberg,[1], Filip Ottosson,[2], Madeleine Ernst,[2], Francesco Russo,[2], Valdemar Stentoft-Hansen,[3], Samuel Demharter,[3], Peer Tfelt-Hansen,[1], Arieh Cohen,[2], Jes Olesen,[4], Thomas Folkmann Hansen,[5]

[1] Danish Headache Center, Rigshospitalet Glostrup, Dept of Neurology, Glostrup, Denmark, [2] Statens Seruminstitut, Denmark, [3] Abzu Aps, Denmark, [4] Glostrup Hospital, University of Copenhagen Department of Neurology, Dept of Neurology, Rigshospitalet Glostrup, Glostrup, Denmark, [5] Copenhagen University Hospital, Danish Headache Center, Rigshospitalet Glostrup, Dept of Neurology, Glostrup, Denmark

Introduction

Migraine is a common, polygenic disorder that is characterized by moderate to severe headache attacks that lasts for 4-72 hours. Migraine attacks are commonly treated with triptans, which are serotonin receptor agonists. However, the mode and site of action of triptans are debated. Furthermore, the therapeutic gain for 2hours pain freedom is relatively low, varying from 13 to 32 per cent. Understanding the molecular mechanisms of triptans, including pharmacokinetics and -dynamics, provides knowledge on migraine mechanisms and may explain why triptans are not working sufficiently in a large part of the migraine patients.

Methods

We collected temporal multi-omics profiles on 24 migraine patients, using blood samples collected at a) migraine attack, b) two hours after treatment with a triptan, c) when headache-free, and d) after a cold-pressor test. Blood samples were used for untargeted metabolomics (LC-MS/MS) and transcriptomics (RNA-sequencing). Differential metabolomic analysis was performed to find metabolites associated with treatment. Their effect was further investigated by integration with transcriptomics using correlation analysis and a machine learning approach.

Results

We found three differential metabolites: cortisol, sumatriptan and glutamine. The change in sumatriptan levels in the blood correlated with a change in *GNAI1* and *VIPR2* gene expression, both known to regulate cyclic AMP (cAMP) levels. This mechanism is not only

suggested to be affected by the serotonin receptor, but also migraine provocation models using e.g., CGRP and PACAP are suggested to act via increasing cAMP. Correlation analysis suggested a role for *ZNF433* as a regulator of gene expression related to cAMP levels. Furthermore, we found several carnitine's to be correlated with sumatriptan and glutamine, suggesting a role for fatty acid oxidation in migraine and/or its treatment with triptans.

Conclusions

Using an integrative approach of untargeted metabolomics and transcriptomics we find evidence for a role of cAMP regulation, glutamine, and fatty acid oxidation in the molecular mechanisms of migraine and/or the effect of triptans.

Conflict of interest

V.S-H and S.D. are employed at Abzu, developers of the QLattice, the machine learning method used in this study. The QLattice is freely available for non-commercial use. The other authors declare that they have no competing interests.

[31] MULTI-PGS ENHANCES POLYGENIC PREDICTION: WEIGHTING 937 POLYGENIC SCORES

Clara Albiñana,[1], Zhihong Zhu,[2], Andrew Schork,[3], Andrés Ingason,[3], Hugues Aschard,[4], Isabell Brikell,[5], Cynthia Bulik,[6], Liselotte Petersen,[7], Esben Agerbo,[7], Jakob Grove,[8], Merete Nordentoft,[9], David Hougaard,[9], Thomas Werge,[9], Anders Børglum,[9], Preben Bo Mortensen,[9], John Mcgrath,[7], Benjamin Neale,[10], Florian Privé,[9], Bjarni Vilhjalmsson,[9]

[1] Aarhus University, Aarhus V, Denmark, [2] Aarhus University, [3] Institute of Biological Psychiatry, Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Copenhagen 2100, Denmark, [4] Department of Computational Biology, Institut Pasteur, Université de Paris, 25-28 Rue du Dr Roux, 75015, Paris, France, [5] Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden, [6] Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, Nc 27514, USA, [7] National Centre for Register-Based Research, Aarhus University, 8210 Aarhus V, Denmark, [8] Department of Biomedicine and Center for Integrative Sequencing, Iseq, Aarhus University, 8000 Aarhus C, Denmark, [9] The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Ipsych, 8210 Aarhus V, Denmark, [10] Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA

Introduction

Psychiatric disorders are prevalent and highly heritable traits with complex genetic architectures. Polygenic scores (PGS) trained on summary statistics from genome-wide association studies (GWAS) for these disorders are generally underpowered and capture only a fraction of the estimated heritable variation. However, many psychiatric and neurological disorders have shared etiology and are genetically correlated with behavioural and cognitive traits that have been studied in large GWAS. PGS from the GWAS based on these correlated phenotypes can be leveraged into a weighted multivariate predictor to increase the prediction accuracy of the outcome disorders (multi-PGS). By using a combination of population health registers and individual-level biobank genetic data (iPSYCH), we generated genetic predictors for a range of ICD10-codes with specific focus on psychiatric disorders, sub-diagnosis, and case-case differentiation.

Methods

We propose a two-step approach to generate multi-PGS for the range of desired outcomes. In

the first step, which was common for all outcomes, we built an internal library of 937 PGS with LDpred2-auto (Privé et al., 2020) from publicly available GWAS summary statistics. Second, we projected the PGS weights into the iPSYCH data genotypes and trained LASSO penalized regression models for a range of psychiatric and neurological disorders and their subtypes based on the Danish Registers, diseases without available GWAS summary statistics and case-case predictions (e.g., ADHD/ASD).

Results

The multi-PGS increased prediction R2 for all major psychiatric disorders with reported GWAS summary statistics in iPSYCH compared to the single PGS for the disorder alone by mean 3-fold and up to 9-fold for attention-deficit/hyperactivity disorder (ADHD). We also computed multi-PGS for a range of ICD10 codes including substance abuse disorders (F1), neurotic, stress-related and somatoform disorders (F4), specific personality disorders (F6), mental retardation (F7), constituting the first PGS for some of the disorder subtypes. By removing the disorder specific PGS from the PGS library pre training the model, we showed how a GWAS for each disorder is not necessary, as the prediction R2 did not change significantly. These results will be further validated using FinnGen data.

Conclusions

We weighted PGSs for 937 traits to generate multi-PGS for psychiatric disorders that are easily computable for many outcomes of interest in psychiatry without having to pre-select the most genetically correlated traits, and that they consistently increase the prediction accuracy. To obtain comparable accuracies from single or meta-analyzed GWAS summary statistics, hundreds of thousands of cases would need to be genotyped. We benchmark the multi-PGS framework against other PGS re-weighting methods and highlight its potential application to new emerging biobanks or register-based genetic cohorts to generate PGS for every available phecode or defined phenotype in their system.

Conflict of interest

None.

[32] PARENTAL TRANSMITTED AND NON-TRANSMITTED ALLELES RECAPITULATE KNOWN BIOLOGY OF NEONATAL JAUNDICE

Pol Sole-Navais,[1], Julius Juodakis,[1], Karin Ytterberg,[1], Marc Vaudel,[2], Christopher Flatley,[1], Oyvind Helgeland,[2], Pål Njolstad,[2], Louis Muglia,[3], Stefan Johanson,[2], Zhang Ge,[3], Bo Jacobsson,[1]

[1] University of Gothenburg, Department of Obstetrics and Gynaecology, Sahlgrenska Academy, Institute of Clinical Science, Gothenburg, Sweden, [2] University of Bergen, Center for Diabetes Research, Department of Clinical Science, Bergen, Norway, [3] Cincinnati Children's Hospital Medical Center, Division of Human Genetics, Center for the Prevention of Preterm Birth, Perinatal Institute, Cincinnati, United States

Introduction

Jaundice, caused by the accumulation of bilirubin, affects almost all neonates in their first days of life. This condition usually resolves without treatment, but 10% of newborns require phototherapy. The genetic effects on neonatal jaundice have been limited to candidate gene studies using the fetal genome, and our knowledge about it stems from genetic studies of adult bilirubin levels.

Methods

We performed a genome-wide association study of neonatal jaundice (i.e., receiving phototherapy treatment) in 29,765 mothers (n cases = 2,536) and 27,722 neonates (n cases = 1,910).

Results

A common missense variant resulting in the substitution of proline to threonine at position 70C>A of *UGT1A4* was associated with a five-fold lower risk of neonatal jaundice (rs6755571, OR = 0.2, p-value = 3.2× 10-58), effect that was limited to the parental transmitted alleles. This variant is in low LD with the lead variant for adult bilirubin levels (rs35754645, *UGT1A** gene region, R2 = 0.02), suggesting moderate differences in their genetic effects. We further identified an associated locus in the maternal *ABO* gene region. By leveraging data from 23,461 parent-offspring trios, the maternal non-transmitted and paternal transmitted alleles showed an opposite effect direction, consistent with maternal-fetal *ABO* blood group incompatibility.

Conclusions

GWAS of neonatal jaundice recapitulates known biology and a potential role for *UGT1A4* in its etiology.

Conflict of interest

None.

[33] PATTERNS OF GENETIC OVERLAP ACROSS MENTAL DISORDERS AND TRAITS BEYOND GENETIC CORRELATION

Guy Hindley,[1], Oleksandr Frei,[2], Alexey Shadrin,[3], Weiqiu Cheng,[3], Kevin S O'connell,[3], Romain Icick,[4], Nadine Parker,[5], Shahram Bahrami,[3], Naz Karadag,[6], Daniel Roelfs,[5], Børge Holen,[3], Aihua Lin,[5], Chun Chieh Fan,[7], Srdjan Djurovic,[8], Anders M Dale,[7], Olav Smeland,[9], Ole A Andreassen,[10]

[1] University of Oslo, University of Oslo, Norment, Oslo, Norway, [2] Norment, Center for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway, Norway, [3] Norment, University of Oslo, Norment, Institute of Clinical Medicine, Oslo, Norway, [4] Inserm, Université Paris Cité, Umr_s1144, Paris, France, [5] Norment, [6] University of Oslo, Norment, Institute of Clinical Medicine, Oslo, Norway, [7] University of California San Diego, [8] Norment, Oslo University Hospital, [9] Institute of Clinical Medicine, University of Oslo, Norment, Oslo, Norway, [10] Norment, Norway

Introduction

It is well recognized that most mental disorders are moderately to strongly genetically correlated, mirroring their overlapping symptom profiles. However, genetic correlation (rg) does not capture all dimensions of genetic overlap since it is unable to differentiate genetic overlap with a balanced mixture of concordant and discordant effects on two phenotypes from an absence of genetic overlap. We therefore aimed to describe the patterns of genetic overlap across mental disorders, and between mental disorders and cognitive and personality traits when accounting for mixed effect directions.

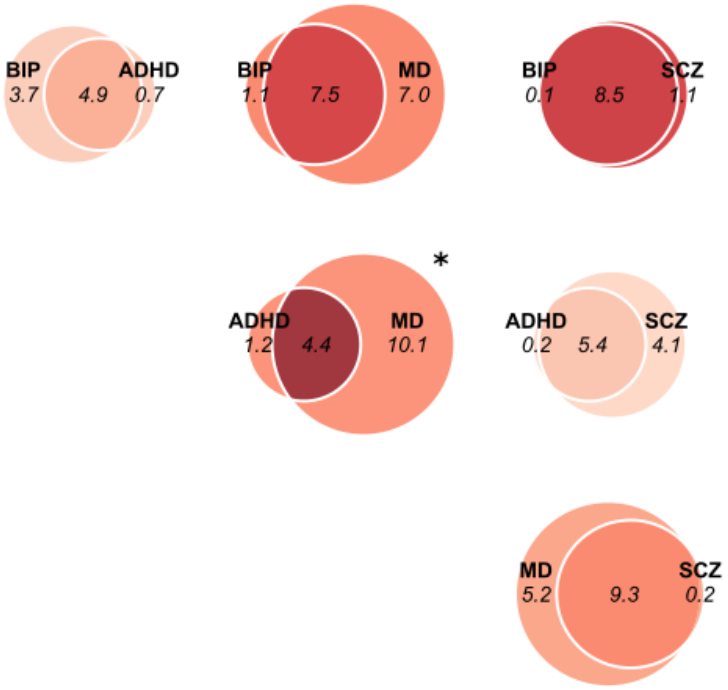
Methods

We applied the bivariate causal mixture model (MiXeR) and local analysis of covariance (LAVA) to genome-wide summary statistics for schizophrenia, bipolar disorder, depression, and ADHD, as well as general intelligence, educational attainment, neuroticism, and subjective well-being. We used height as a polygenic, somatic comparator. Sample sizes ranged from 53,293 to 766,345. MiXeR uses Gaussian mixture models to infer the polygenicity of a single trait, reported as the number of causal variants (single nucleotide polymorphisms, SNPs) required to explain 90% of the total polygenicity (univariate analysis), and the number of shared and unique variants for two traits irrespective of effect direction (bivariate analysis). We used LAVA to compute local genetic correlations within 2,495 semi-independent

genome-wide loci. We compared MiXeR and LAVA estimates of mixed effect directions using Pearson correlation.

Results

Univariate MiXeR revealed substantial differences in the polygenicity of mental disorders: ADHD was the least polygenic (5.6K causal variants), followed by bipolar disorder (8.6K), schizophrenia (9.6K), and depression (14.5K). Pairwise bivariate analyses demonstrated extensive genetic overlap across mental disorders and between mental disorders and cognitive and personality traits. This was observed in the presence of strong positive genetic correlation, for example between bipolar disorder and schizophrenia (no. shared variants = 8.5K, rg = 0.78) and weak positive genetic correlation, for example ADHD and schizophrenia (no. shared variants = 5.4K, rg = 0.19). The latter is indicative of mixed effect directions whereby a mixture of shared variants with concordant and discordant effects cancel each other out. In contrast, there was minimal genetic overlap between mental disorders and height (0.7-1.1K). LAVA identified a mixture of loci with positive and negative local correlations, consistent with MiXeR estimates of mixed effect directions (Pearson r = 0.88, p<0.001).



Conclusions

These findings indicate that the SNP variant component of the genetic risk for mental disorders is predominantly driven by a set of highly pleiotropic genetic variants. It is therefore largely the disorder-specific distribution of effect directions and effect sizes which attribute specificity for a phenotype rather than a set of disorder specific variants. This extends to cognitive and personality traits, supporting dimensional conceptualizations of mental disorders. Nevertheless, differences in polygenicity across mental disorders may participate in some of the differences in phenotypic heterogeneity and may provide a means for testing the biological specificity of sub-phenotyping procedures. These findings therefore provide a conceptual advance on our understanding of the genetic underpinnings of mental disorders and may contribute to future attempts to reconceptualize psychiatric nosology.

Conflict of interest

Dr. Dale is a founder of and holds equity in Cortechs.ai and serves on its scientific advisory board; he is a member of the scientific advisory boards of HealthLytix and the Mohn Medical Imaging and Visualization Center (Bergen, Norway); and he receives funding through a research agreement between General Electric Healthcare and UCSD. Prof. Andreassen has received speaking honoraria from Lundbeck and has served as a consultant for HealthLytix. The other authors report no financial relationships with commercial interest.

[34] POLYGENIC HEALTH INDEX, GENERAL HEALTH, AND PLEIOTROPY: SIBLING ANALYSIS AND DISEASE RISK REDUCTION

Erik Widén,[1], Laurent Tellier,[2], Louis Lello,[3], Timothy Raben,[4], Steve Hsu,[5]

[1] Genomic Prediction, Inc., Michigan State University, Department of Physics and Astronomy, North Brunsiwck, United States, [2] Michigan State University, Genomic Prediction, Department of Physics and Astronomy, North Brunswick, United States, [3] Genomic Prediction, Inc., Michigan State University, Department of Physics and Astronomy, North Brunswick, Nj, United States, [4] Michigan State University, Department of Physics and Astronomy, East Lansing, United States, [5] Genomic Prediction, Inc., Michigan State University, Department of Physics and Astronomy, North Brunswick, United States

Introduction

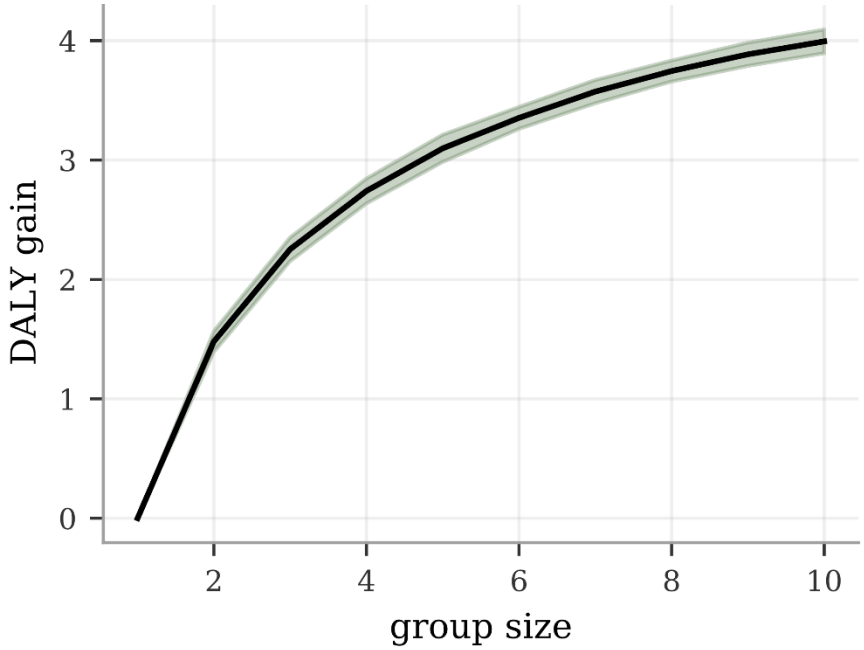
As biobank data sets have grown larger, so have the performances and applicability of PRS. Typically, PRS are trained on and applied to a single disease, but with many such risk predictions available, it is natural to ask whether they could be combined into a general health index — a single number to describe the overall health of an individual, using the DALY: Disability Adjusted Life Years.

Methods

We construct a polygenic health index as a weighted sum of polygenic risk scores for 20 major disease conditions, including, e.g., coronary artery disease, type 1 and 2 diabetes, schizophrenia, etc. Individual weights are determined by population-level estimates of impact on life expectancy and quality. We validate this index in odds ratios and selection experiments using unrelated individuals and siblings (pairs and trios) from the UK Biobank.

Results

Individuals with higher index scores have decreased disease risk across almost all 20 diseases (no significant risk increases), and lower phenotypic disease incidence. We find no statistical evidence for antagonistic trade-offs in risk reduction across these diseases. When Disability Adjusted Life Years (DALYs) due to the 20 diseases were used as the performance metric, the gain from genetic selection (highest index score vs average) among 10 individuals was found to be roughly 4 DALYs, and among 5 individuals was found to be 3 DALYs.



Discussion

These results have important implications for public health and also for fundamental issues such as pleiotropy and genetic architecture of human disease conditions. It is commonly believed that genetic factors influence overall health and longevity. With modern genomic methods we can test the scientific veracity of this hypothesis. By combining Polygenic Risk Scores (PRS) across the most impactful disease conditions, we can build a composite predictor as part of an individual's overall health.

Correlations between the disease risks are found to be mostly positive, and generally mild. This supports the folk notion of a general factor which characterizes overall health, sometimes described as synergistic pleiotropy. These results have important implications for public health and also for fundamental biological questions such as genetic architecture of human disease conditions.

The concept of pleiotropy was formulated before the notion of high dimensional spaces of genetic variation became familiar. The conventional logic is that, because a single gene can affect many different complex traits, it must be the case that different complex traits, such as disease risks, are themselves correlated, perhaps antagonistically (e.g., due to balancing selection, or for some deeper biochemical reason). This would entail specific trade-offs, hypothetically: an individual with low diabetes risk might necessarily have higher cancer risk, etc. However, results from the modern era of GWAS and machine learning on large data sets

show that the number of genetic loci which control a specific complex trait is typically in the thousands. The fact that most of the variance can be disjoint across different complex traits is a manifestation of high dimensionality. In this work we focus on sparse algorithms applied to array data which leaves open the possibility that there could be underlying causal loci that could still be correlated. However, the relatively small genetic correlations observed here leave this as an unlikely scenario.

Conflict of interest

I am an employee of Genomic Prediction.

[35] POPULATION-LEVEL STUDY OF CNVS AND THEIR ASSOCIATED RISK OF PSYCHIATRIC DISORDERS IN A DANISH CASE-COHORT

Morteza Vaez,[1], Simone Montalbano,[1], Xabier Calle Sanchez,[1], Ipsych Collaborators,[2], Armin Raznahan,[3], Wesley Kurt Thompson,[4], Thomas Werge,[5], Dorte Helenius Mikkelsen,[6], Andres Ingason,[7]

[1] Institute of Biological Psychiatry, Mental Health Services Capital Region of Denmark, Roskilde, Denmark, [2] The Lundbeck Foundation Initiative for Integrative Psychiatric Research (Ipsych), Copenhagen and Arhus, Denmark, [3] Section on Developmental Neurogenomics, Human Genetics Branch, National Institute of Mental Health Intramural Research Program, Bethesda, Bethesda, United States, [4] Herbert Wertheim School of Public Health and Human Longevity, University of California, San Diego, United States, [5] Copenhagen University Hospital, Institute of Biological Psychiatry, Copenhagen, Denmark, [6] Institute of Biological Psychiatry, Roskilde, Denmark, [7] Institute of Biological Psychiatry, Region Hovedstadens Psykiatri, Roskilde, Denmark

Background/Objectives

Recurrent copy number variants (CNVs) have been shown to increase the risk of neuropsychiatric disorders in clinical case-control studies. Yet, little is known about their pathogenic impact at the level of an individual or an entire population. We sought to estimate the true population prevalence of recurrent CNVs and their associated risk of psychiatric disorders.

Methods

We applied the iPSYCH2015 case-cohort, including all individuals born in Denmark in 1981-2008 and diagnosed with a major psychiatric disorder (Autism, ADHD, affective disorder, Schizophrenia and anorexia) (n= 82626) by 2015, and a random comparison sample from the same birth cohort (n=43346). Samples were genotyped and recurrent CNVs at 30 loci called with PennCNV and verified by visual inspection. Population-representative hazard ratios (HR) were derived using a Cox proportional hazard model with inverse probability of sampling (IPS) weights.

Results

The population prevalence of several CNVs was higher than in the UK Biobank (11 loci including 9 deletions and 2 duplications). Overall, the carrier rate in cases was higher than in

the comparison sample with HR varying widely across disorders and CNVs. The prevalence of duplication was higher than deletions in most of the examined loci in the random cohort, of which 15q11.2 duplication was the most prevalent 0.44% (95% CI, 0.37% – 0.50%) (1:228). In total, we found 47 significant association results, where 16p13.11 deletion, 16p13.11 duplication and 15q13.3 deletion were associated with most of assigned diagnosis.

Discussion

The population-based iPSYCH study enables estimating of true prevalence and associated risk conferred by CNVs of ascertained psychiatric disorders, thus paving the way for implementing genetic predictions into clinical practice.

Conflict of interest

None.

[36] PUBLICLY AVAILABLE PRIVACY-PRESERVING BENCHMARKS FOR POLYGENIC PREDICTION

Menno Witteveen,[1], Emil Pedersen,[1], Joeri Meijsen,[1], Michael Andersen,[2], Florian Prive,[1], Doug Speed,[1], Bjarni Vilhjálmsson,[1]

[1] Aarhus University, [2] Technical University of Denmark

Introduction

In recent years, many different approaches for creating polygenic scores have been proposed and this trend shows no sign of abating. Most of these new polygenic score methods claim to outperform previously proposed methods, making it confusing for users to choose an appropriate PRS approach.

This apparent performance paradox is due to challenges in properly determining what approaches are superior because different datasets, quality control, and pre-processing steps are used to determine performance. As shown previously, these steps can have a significant impact on the overall prediction accuracy for polygenic score methods, and thus lead to an incomplete comparison of PRS methods.

Methods

In the related field of Machine Learning this problem has been addressed by the introduction of publicly available benchmark datasets, which allow for fair comparison between approaches, which has been vital to the advancement of the field. However, the creation of such publicly available benchmark datasets in genetics has, until now, proved challenging because of privacy concerns.

As a solution, we present a privacy-preserving and publicly available benchmark for polygenic prediction, which allows researchers to both *train* and *test* polygenic prediction methods using only linkage disequilibrium (LD) information and summary statistics from genome-wide association studies, thus preserving privacy.

Results

Using UK biobank data and a diverse set of 8 external summary statistics, we show that with our approach we estimate the squared correlation prediction accuracy almost perfectly. The

method only requires linkage disequilibrium (LD) information and genome-wide association summary statistics, and no individual level data. We further used the benchmark data to compare a variety of polygenic scoring methods including PRS-CS, LDPred2, and SBayesR and observe a remarkable concordance, with perfect recovery of model rankings and almost perfect correlation of model performance measures.

Conclusions

This new benchmark for PRS model development and comparison will be made publicly available for researchers to download. We believe this benchmark can be used as a clear and unbiased standard for future polygenic score methods to compare against.

Conflict of interest

None.

[37] QUANTIFYING THE ASSORTATIVE MATING OF PSYCHIATRIC DISORDERS AND ITS CONSEQUENCES IN DENMARK

Saeid Rasekhi Dehkordi,[1], Joeri Meijsen,[2], Dorte Helenius,[2], Richard Zetterberg,[2], John Shorter,[2], Alfonso Buil,[2], Thomas Werge,[3]

[1] Mental Health Services, Copenhagen, University of Copenhagen, Institute of Biological Psychiatry (Ibp), Roskilde, Denmark, [2] Institute of Biological Psychiatry (Ibp), Mental Health Services, Sct. Hans, [3] Institute of Biological Psychiatry (Ibp), Mental Health Services, Sct. Hans, University of Copenhagen

Background

Similarity (positive) or dissimilarity (negative) of couples in phenotype characteristics which deviate from expectation is defined as assortative mating (AM). Positive assortative mating between patients with psychiatric disorders has been described before. However, the consequences of AM in psychiatric disorders are not described nor quantified. This study aims to quantify AM in psychiatric disorders and its consequences in terms of prevalence, heritability, and fecundity in Danish register data.

Method

The study was conducted on two generations of Danish residents born between January 1969 and December 2015. Odds ratio on 5:1 case-control and tetrachoric correlation on the population scale were used to quantify AM within several psychiatric diagnosis, including Alcohol abuse, Schizophrenia, single and recurrent depressive disorder, obsessive-compulsive disorder, anxiety disorders, Borderline type, Autism spectrum, attention-deficit/hyperactivity disorder, and Bipolar. We have evaluated the consequences of non-random mating on prevalence, heritability, and fecundity by randomizing the pattern of mating in the population.

Results

In 571534 unique couples residing in Denmark, substantial similarities between partners' diagnostic statuses for psychiatric disorders were observed. The odds of psychiatric cases, males and females probands on average having an affected partner were substantially higher than matched controls ranging from 2.52 for the single and recurrent depressive disorder to 14.97 for schizophrenia which is in accordance with diagnostic status correlations of partners in the whole population, 0.20 and 0.39 respectively. We have evaluated the consequences of AM for the phenotype 'any mental disorder' on 822209 offspring of those couples.

Randomizing partners suggests that on average (with 1000 replication in each gender) population prevalence of any mental disorder would be 2.25 % less if couples had mated randomly. Moreover, the elimination of spousal phenotypic resemblance in the population results in 30 % decrease in the heritability of offspring generation. Finally, the predictive model for fecundity calculation made by randomized partners reveals that the effect of AM of psychiatric disorders in fecundity is small though significant (a decrease of 7 for every 10000 births).

Conclusion

Assortative mating of psychiatric disorders increased heritability and population prevalence significantly. Also, it decreases the average number of children per family.

Conflict of interest

None.

[38] SWISS NATIONAL SARS-COV-2 GENOME AND VARIANT SURVEILLANCE PROGRAM

Marc Friedli,[1], Samuel Cordey,[2], Erik Boehm,[3], Pauline Vetter,[3], Ioannis Xenarios,[4], Keith Harshman,[5], Lorenzo Cerutti,[5], Cedric Howald,[5], Aitana Neves,[6], Christian Althaus,[7], Tanja Stadler,[8], Emma Hodcroft,[9], Cornelius Roemer,[10], Richard Neher,[10], Laurent Kaiser,[3]

[1] Epfl, Health 2030, Switzerland, [2] University Hospital Geneva (Hug), Switzerland, [3] Hug, Geneva, Switzerland, [4] University of Lausanne, Health 2030 Genome Center, Geneva, Switzerland, Lausanne, Switzerland, [5] Health 2030 Genome Center, Geneva, Switzerland, [6] Sib, Switzerland, [7] University of Bern, Switzerland, [8] Eth Zurich, Switzerland, [9] Hug, Geneva, Switzerland, University of Bern, Switzerland, [10] University of Basel, Switzerland

Introduction

Switzerland initiated a national genomic surveillance program for SARS-CoV-2 in March 2021. The program is funded by the Federal Office of Public Health (FOPH) and will run at least until the end of 2023. The overall goal of the program is to monitor epidemiological trends and to highlight meaningful observations to which policies can be adapted. Because greater transmissibility and/or immune escape potential of the different variants of concern (VOCs) and variants of interest (VOIs) can result in new surges in COVID-19 numbers despite the vaccination campaign, this program aims to closely monitor each variant displaying mutations known to be linked to either increased transmissibility or immune escape potential. In order to cover all geographical regions of Switzerland, 15 diagnostic laboratories have participated to the program, including university hospital centres in Switzerland, private laboratories, cantonal-based laboratories, and 3 high-throughput sequencing platforms (Health 2030 Genome Centre in Geneva, Genomics Facility Basel, Functional Genomics Centre Zurich). Processed sequencing data are shared openly within 14 days from positive PCR result through the GISAID platform (https://www.gisaid.org) and eventually through the Swiss Pathogen Surveillance Platform (SPSP). The analysis of this National Surveillance is centralized, and variants of concern are counted, analysed and all sequences scanned for new variants with potential changes in antibody-spike interactions

(https://nextstrain.org/groups/swiss, https://covariants.org/per-country, https://cov-spectrum.ethz.ch). This program was initiated and developed in close collaboration with the Swiss

National COVID-19 Science Task Force and the Swiss Institute of Bioinformatics (SIB). As of October 2022, Switzerland contributed significantly to the international effort with > 150'000 SARS-CoV-2 genome sequences shared, representing ~ 3.7% of reported Swiss cases or 17.6 SARS-CoV-2 WGS per million in the Swiss population. The program swiftly identified appearance of all SARS-CoV-2 variants of concern - alpha, beta, gamma, delta - and is now tracking the sub-lineages of omicron.

Conflict of interest

None.

[39] THE ROLE OF DEPRESSION AND ITS TREATMENT IN ANTIHYPERTENSIVE MEDICATION ADHERENCE AND PERSISTENCE

Hanna Kariis,[1], Silva Kasela,[1], Tuuli Jürgenson,[1], Aet Saar,[2], Jana Lass,[3], Kristi Krebs,[1], Urmo Võsa,[1], Elis Haan,[1], Lili Milani,[4], Kelli Lehto,[1]

[1] Institute of Genomics, University of Tartu, Estonian Genome Centre, Tartu, Estonia, [2] North Estonia Medical Centre, [3] Tartu University Hospital, [4] Institute of Genomics, Estonian Genome Centre, Tartu, Estonia

Introduction

Patients with major depression (MD) have higher rates of comorbid cardiovascular diseases and its risk factors, including hypertension (HTN) compared to the general population. MD has been associated with lower antihypertensive medication (AMH) adherence and persistence but the research is inconclusive. Therefore, we aim to investigate the role of MD diagnosis and genetic predisposition to MD in AHM adherence and persistence. We will also explore the relationship between antidepressant initiation and intraindividual change in AHM adherence.

Methods

An observational retrospective cohort study was carried out in the Estonian Biobank (N=200,000) using pharmacy refill data from the National Health Insurance database. Adherence and persistence to AHMs were determined for newly diagnosed hypertension cases (during 2009-2017) for a 3-year follow-up period.

Multivariable linear and logistic regressions were used to explore the associations between depression, polygenic risk for depression (MD PRS), and adherence or persistence. To assess the role of depression on adherence or persistence we adjusted for the following sociodemographic and genetic risk factors: age at AHM initiation, sex, educational attainment, body mass index (BMI), MD PRS, systolic blood pressure (SBP) PRS, diastolic blood pressure (DBP) PRS, and five principal components (PCs). To assess the role of genetic predisposition for depression on adherence or persistence we adjusted for age at AHM initiation, sex, educational attainment, BMI, SBP PRS, DBP PRS, and five PCs.

In the secondary analysis we used a linear mixed-effects model to understand if antidepressant therapy initiation has an impact on intraindividual change in AHM adherence, while adjusting for age at AHM initiation and sex. We compared AHM adherence 6 months before and 6 months after the first antidepressant purchase, while excluding the first month after AD purchase due to a delay in the onset of antidepressant therapeutic effects.

Results

We identified 20,724 individuals with newly diagnosed HTN (6294 cases with MD and 14,430 controls). Controlling for sociodemographic and genetic variables, we found that MD was associated with 2% lower adherence (beta=-0.015, CI95=-0.025–(-0.004)) and 16% lower odds of persistence (OR=0.845, CI95=0.778–0.916). One standard deviation increase in MD PRS was associated with 0.8% lower adherence (beta=-0.008, CI95=-0.012–(-0.003)) and 5% lower persistence (OR=0.953, 95%CI=0.919–0.988). Starting antidepressant therapy increased AHM adherence by 8% compared to six months before and six months after initiating antidepressant therapy (beta=0.078; 95%CI=0.025–0.131).

Conclusions

We found that a lifetime MD diagnosis and genetic liability to depression were associated with lower AHM adherence and persistence, after controlling for sociodemographic and genetic factors. This may suggest that patients with depression may need more support integrating their doses to their daily routine and re-motivation to continue the treatment. To our knowledge, this is the first study to investigate the effect of genetic predisposition for depression on AHM medication adherence and persistence. The association between depression PRS and medication adherence or persistence may suggest a shared contribution of common genetic variants to depression and AHM adherence and persistence. In addition, antidepressant therapy may help to improve adherence to hypertensive medications in patients with depression.

Conflict of interest

None.

[40] TIME-VARYING EFFECTS ARE COMMON IN GENETIC CONTROL OF GESTATIONAL DURATION

Karin Ytterberg,[1], Julius Juodakis,[2], Christopher Flatley,[2], Pol Sole-Navais,[2], Bo Jacobsson,[3]

[1] Clinical Science, Obstetrics and Gynecology, Gothenburg, Sweden, [2] University of Gothenburg, Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Gothenburg, Sweden, [3] University of Gothenburg, and Norwegian Institute of Public Health, Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Gothenburg, Sweden

Introduction

Preterm birth is a major burden to neonatal health worldwide, determined in part by maternal genetics. Recently, large studies have discovered a number of genes associated with this trait or its continuous equivalent – gestational duration. However, the effect timing and clinical relevance of the loci identified so far are still unclear.

Methods

Here, we use genotyping data from a cohort of 31,000 births from Norway to investigate different models of the genetic pregnancy "clock". We conduct genome-wide association studies using gestational duration or preterm birth. We calculate relative power differences between the settings. We then re-analyse the individual variants in piecewise-exponential additive models that allow identifying genetic effects that vary over time.

Results

In the association studies, we replicate known maternal associated variants and finding one new foetal variant. We illustrate how the interpretation of the results for dichotomized and continuous phenotypes is complicated by the loss of power when dichotomizing. Using flexible survival models, we are able to resolve this complexity, and find that many of the known loci have time-varying effects, often stronger early in pregnancy. Overall the polygenic control of birth timing appears to be shared in the term and preterm, but not very preterm periods, and we present exploratory results suggesting involvement of the major histocompatibility complex genes in the latter.

Conclusions

These findings show the clinical relevance of the known gestational duration loci, and should provide the next step towards designing further experimental studies of these loci. The models introduced here can be applied to time-to-onset analysis of late-adulthood diseases as well.

Conflict of interest

None.

[41] ULTRALOW-COST GENOMIC POPULATION SCREENING

Ulrik Stoltze,[1], Christian Hagen,[2], Thomas van Overeem Hansen,[3], Anna Byrjalsen,[4], Anne-Marie Gerdes,[5], Viktor Yakimow,[2], Simon Rasmussen,[6], Marie Bækvad-Hansen,[2], David Michael Hougaard,[2], Kjeld Schmiegelow,[7], Henrik Hjalgrim,[8], Karin Wadt,[3], Jonas Bybjerg-Grauholm,[9]

[1] Rigshospitalet, Department of Pediatrics and Adolescent Medicine, Copenhagen, Denmark, [2] Ssi, Denmark, [3] Rigshospitalet, Denmark, [4] Rigshospitalet, Department of Clinical Genetics, Copenhagen East, Denmark, [5] Rigshospitalet, ., Denmark, [6] University of Copenhagen, Nnf Center for Protein Research,, København N, Denmark, [7] Rigshospitalet, University of Copenhagen, Department of Paediatrics and Adolescent Medicine, Copenhagen, Denmark, [8] Department of Epidemiology Research, Statens Serum Institut, Copenhagen Denmark, Danish Cancer Society Research Center, Danish Cancer Society, Copenhagen, Denmark, Denmark, [9] Ssi, Copenhagen, Denmark

Introduction

Next-generation sequencing (NGS) based population screening holds great promise for disease prevention and earlier diagnosis, but the costs associated with screening millions of humans remain prohibitive. New methods for population genetic testing that lower the costs of NGS without compromising diagnostic power are needed.

Methods

We developed Double Batched Sequencing where DNA samples are batch-sequenced twice – directly pinpointing individuals with rare variants. We sequenced batches of at-birth blood spot DNA using a commercial 113-gene panel in an explorative (n = 100) and a validation (n = 100) cohort of children who went on to develop pediatric cancers. All results were benchmarked against individual whole genome sequencing data.

Results

We demonstrated fully replicable detection of cancer-causing germline variants, with positive and negative predictive values of 100% (95%CI, 0.91–1.00 & 95%CI, 0.98–1.00, respectively). Pathogenic and clinically actionable variants were detected in RB1, TP53, BRCA2, APC, and 19 other genes. Analyses of larger batches indicated that our approach is highly scalable, yielding more than 95% cost reduction or less than 3 cents per gene screened for

rare disease-causing mutations. We also show that Double Batched Sequencing can cost-effectively prevent childhood cancer deaths through broad genomic testing.

Conclusions

Our ultracheap genetic diagnostic method, which uses existing sequencing hardware and standard newborn blood spots, should readily open up opportunities for population-based presymptomatic genetic screening across many fields of clinical genetics and genomics.

Conflict of interest

None.

Thank you to our sponsors and exhibitors

Silver Sponsors

illumina®

lifebit

MGI

ThermoFisher SCIENTIFIC

Bronze Sponsors

BD

BGI

BioXpedia

CeGaT

eurofins | Genomics

IDT INTEGRATED DNA TECHNOLOGIES

Olink®

Oxford NANOPORE Technologies

PacBio

Promega

Supporter

Steno Diabetes Center Aarhus

midt regionmidtjylland