Title

T27. FEMA-GWAS: FAST AND EFFICIENT MIXED-EFFECTS ALGORITHM FOR DISCOVERY OF GENOME-WIDE AGE-DEPENDENT EFFECTS

Authors

Parekh, Pravesh

Parker, Nadine

Smith, Diana

et al.

Publication Date

2024-10-01

DOI

10.1016/j.euroneuro.2024.08.337

Peer reviewed

therefore confounded by lifelong exposure to the disease, disease-related environmental factors, and use of other medications for comorbid conditions. This project aims to fill the knowledge gap on what is the impact of pharmacological compounds on gene regulation in human neurons and to highlight how this knowledge can be harvested to study MDx, their possible new treatments, treatment outcomes, and possible side effects.
Methods: Human glutamatergic neurons were exposed to two different doses of compounds (CPx) of high interest for MDx (valproic acid, fluoxetine, ketamine, cannabidiol), or cortisol. An additional group of neurons was first treated with cortisol and then by each of the CPx. Genome-wide DNAm levels from all of the exposed and unexposed neurons were quantified with the use of the Infinium MethylationEPIC BeadChip. We subsequently performed epigenome-wide association studies (EWASes) to: i) determine where in the genome CPx and cortisol treatment changes DNAm in human neurons, ii) determine if the CPx- and cortisol-induced DNAm changes vary across doses, and iii) answer if the epigenetic dysregulation caused by exposure to cortisol can be attenuated by CPx treatment. We further investigated if genes that undergo epigenetic changes due to CPx and cortisol exposure are significantly enriched in pathways, tissue expression profiles, and genes associated with MDx and their intermediate phenotypes.
Results: We have successfully established a protocol for a pharmacoepigenomic study of human glutamatergic neurons and will provide recommendations for a successful study design that i) minimizes confounding technical effects and ii) allows for comparison of epigenetic changes between multiple CPx and cortisol treatments in one experimental setup. High quality whole-genome DNAm data is now available for 114 samples and all CPx and cortisol treatment groups are well represented in the data. We will present this newly created pharmacoepigenomic resource, as well as results from the EWASes and pathway enrichment analyses on changes in epigenetic regulation in human neurons occurring upon the abovementioned CPx and cortisol treatments.
Discussion: To provide better and novel treatments for mental illness it is crucial to first understand how do compounds used or considered for treatment of MDx impact gene regulation of a tissue relevant for these conditions. We will present a unique novel neuronal pharmacoepigenomic resource to the research community that can be combined with other omics approaches to study treatment approaches for MDx.

Disclosure: Nothing to disclose.

doi: 10.1016/j.euroneuro.2024.08.336

T27. FEMA-GWAS: FAST AND EFFICIENT MIXED-EFFECTS ALGORITHM FOR DISCOVERY OF GENOME-WIDE AGE-DEPENDENT EFFECTS

Pravesh Parekh [1], Nadine Parker [1], Diana Smith [2], Diliana Pecheva [2], Carolina Makowski [2], Evgeniia Frei [1], Piotr Jaholkowski [1], Viktoria Birkenæs [3], Nora Refsum Bakken [1], Dennis van der Meer [1], Alexey Shadrin [1], Thomas E. Nichols [4], Oleksandr Frei [1], Anders M. Dale [2], Ole Andreassen [1]

[1] *University of Oslo;* [2] *University of California San Diego;* [3] *Oslo University Hospital;* [4] *University of Oxford*

Background: Genome-wide association studies (GWAS) on longitudinal phenotypes can be used to discover age-dependent genetic effects. These studies can be useful for discovery of genetic factors driving neurodevelopment, brain maturation, and neurodegeneration, which can then be linked to neuropsychiatric diseases. Linear mixed-effects (LME) is one way to perform these analyses while accounting for the dependencies between observations (repeated measurements, genetic relatedness, others). However, fitting single nucleotide polymorphisms (SNP)-level LMEs for large-scale datasets is computationally intensive and almost impractical for situations involving a large number of outcome variables (for example, neuroimaging variables). Here, we present a novel, fast, and computationally efficient mixed-effects framework for performing longitudinal GWAS, enabling discovery of linear and non-linear age-dependent effects of genetic variants on phenotypes.
Methods: Given the longitudinal nature of the data and the presence of other dependencies between observations, we use mixed-effects models, where SNPs are the fixed effects, modeled individually, in addition to other variables. We use our recently developed fast and efficient mixed-effects algorithm (FEMA) framework for fitting computationally efficient LMEs. The FEMA-GWAS approach is: 1) using FEMA, estimate the effect of the fixed effects (without SNPs) on the phenotype(s) and the variance parameters for the random effects; 2) use a two-stage estimation approach (equivalent to the conventional approach by the Frisch-Waugh theorem) where we pre-residualize the genotype and phenotypes for the effects of other fixed effects. This simplifies the estimation problem to estimating the effect of the residualized SNPs on the residualized phenotype. 3) For non-linear age-dependent effects, we model the interactions of the SNPs with splines (natural cubic splines or B-splines with user-defined knots) – these basis functions are used for expanding the (main) effect of SNP. 4) Finally, we use generalized least squares (re-using the estimated variance parameters from the first step) to estimate the coefficients for this expanded set of predictors. Additionally, we allow these expanded set of predictors to interact with dummy-coded sex variable, thereby separately estimating the effects of the SNP-by-age interactions for males and females.
Results: We have developed FEMA-GWAS, a novel MATLAB-based software that allows computationally efficient discovery of SNP-by-age linear and non-linear interactions across the entire genome for a large number of phenotypes. Using simulations, we show that FEMA-GWAS has i) well-controlled type I error; and provides: ii) equivalent estimates to standard LMEs, and iii) accurate parameter recovery for various interaction terms. As a practical application, we present results for non-linear interaction of SNP and age for cortical thickness using samples from the longitudinal ABCD Study.
Discussion: We have developed a software that enables users to perform longitudinal GWAS and discover non-linear age-dependent effects of genetic variants. Our approach is

designed for large-sample situations and efficiently scales to large number of outcome variables (like neuroimaging variables). Our work will enable users to perform these analyses and discover novel genetic variants associated with trajectories of health and its deviations. FEMA is available at: https://github.com/cmig-research-group/cmig_tools.

Abstract Previously Published

Disclosure: Nothing to disclose.

doi: 10.1016/j.euroneuro.2024.08.337

T28. HUMAN ENDOGENOUS RETROVIRUSES AND LONG INTERSPERSED NUCLEAR ELEMENTS-1 LINKED TO MAJOR PSYCHIATRIC CONDITIONS

Rodrigo RR Duarte [1], Oliver Pain [1], Aitana Padilla Requerey [1], Nicholas J Bray [2], Deepak P Srivastava [1], Douglas F Nixon [3], Timothy R Powell [1]

[1] *King's College London;* [2] *Cardiff University;* [3] *Weill Cornell Medicine*

Background: Human Endogenous Retroviruses (HERVs) and Long Interspersed Nuclear Elements-1 (LINE-1s) are abundant ancient sequences in the genome that have long been considered "junk DNA". Recent advances in expression quantification methods for these retrotransposable elements facilitated research implicating them in health conditions. For instance, our group developed a novel approach called a retrotranscriptome-wide association study (rTWAS), which is a transcriptome-wide association study (TWAS) that incorporates global HERV expression, a component of the 'retrotranscriptome'. We identified neurological HERV expression signatures associated with genetic susceptibility for complex traits, which are independent from the expression of canonical protein coding genes (Duarte et al., 2024). Here, we expand the rTWAS approach to include the expression of LINE-1s and HERVs in the adult and second-trimester foetal brain.
Methods: We analysed RNA-sequencing and whole-genome genotype data from 563 adult dorsolateral prefrontal cortex (DLPFC) samples and 77 second trimester fetal brain samples of European ancestry. We constructed expression weights after assessing the expression of canonical genes using either kallisto or STAR, and the expression of HERVs and LINE-1s using Telescope. We used FOCUS (fine-mapping of causal gene sets) to identify HERV and LINE-1 expression signatures associated with genetic susceptibility, while accounting for the expression of canonical genes in the two expression panels and controlling for expression correlation induced by linkage disequilibrium. We report only HERV and LINE-1 signatures associated with genome-wide significant variants ($P < 5e\text{-}8$) and a posterior inclusion probability above 0.5, which are signatures more likely to be causal. We analysed European GWAS of schizophrenia (Trubetskoy et al., 2022), bipolar disorder (Mullins et al., 2021), depression (Howard et al., 2019), attention deficit hyperactivity disorder (Demontis et al., 2019), and autism spectrum conditions (Grove et al., 2019).
Results: For schizophrenia, our fine-mapping approach revealed 12 LINE-1 and 10 HERV expression signatures in the DLPFC, and seven LINE-1 and one HERV expression signatures in the foetal brain associated with genetic risk. For bipolar disorder, we identified one LINE-1 and one HERV expression signature in the DLPFC, and three LINE-1 and four HERV expression signatures in the foetal brain. For depression, we identified five LINE-1 and four HERV expression signatures in the DLPFC, and two LINE-1 and two HERV expression signatures in the foetal brain. For attention deficit hyperactivity disorder, we observed one LINE-1 expression signature in the DLPFC and one HERV expression signature in the foetal brain. For autism spectrum conditions, we did not observe HERV or LINE-1 expression signatures associated with genetic risk. Most signatures detected were specific to a single disorder, but some were shared, including the expression of MER4_20q13.13 (schizophrenia and bipolar disorder) and LINE-1FLnI_1p31.1n (depression and schizophrenia) in the DLFPC.
Discussion: Building upon our previous work, this study suggests that the expression of LINE-1s, in addition to HERVs, contributes to psychiatric disorder aetiology. Overall, these findings expand our understanding of the genomic basis of psychiatric disorders and suggest that further exploration of these retrotransposable elements in the adult and foetal brain may reveal novel risk mechanisms.

Disclosure: Nothing to disclose.

doi: 10.1016/j.euroneuro.2024.08.338

T29. SPATIAL-MOLECULAR SEX DIFFERENCES IN THE HUMAN VENTROMEDIAL AND ARCUATE HYPOTHALAMUS

Bernard Mulvey [1], Yi Wang [2], Heena R. Divecha [1], Svitlana Bach [1], Kelsey D. Montgomery [1], Sophia Cinquemani [1], Yufeng Du [1], Ryan Miller [1], Joel E. Kleinman [1], Stephanie Cerceo Page [1], Thomas M. Hyde [1], Stephanie C. Hicks [2], Kasper D. Hansen [2], Keri Martinowich [1], Kristen R. Maynard [1]

[1] *Lieber Institute for Brain Development;* [2] *Johns Hopkins University*

Background: The hypothalamus (HYP) coordinates core physiology and behaviors, such as growth, mating, parenting, aggression, appetite, and metabolism. Given these roles, it is unsurprising that HYP is a nexus of behavioral, functional, and molecular sex differences in model species. The ventromedial HYP (VMH) and arcuate nucleus (ARC) are illustrative: VMH neurons containing estrogen receptor (Esr1) are sex-differentially activated by socialization, and neurons of the ARC regulate sex hormone production. Though rodent HYP is well-profiled, little is known about human VMH and ARC at the molecular level. We performed spatial transcriptomics (ST) of post-mortem human VMH and ARC using Visium (10x Genomics), followed by in situ sequencing (ISS) of 366 genes at cellular resolution using Xenium (10x Genomics). We present marker genes, spatial ex-