ENVIRONMENTAL-RISK (E-RISK) LONGITUDINAL TWIN STUDY
CONCEPT PAPER FORM

Proposing Author: Matthew Suderman

Author's affiliation, phone, and e-mail address:
Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol
+44 (0) 1173310090
matthew.suderman@bristol.ac.uk

Sponsoring Investigator (if the proposing author is a student, a post-doc or a colleague):
Helen Fisher

Proposed co-authors:
Jean Golding, Marcus Pembrey, Karen Sugden, Helen Fisher

Provisional Paper Title:
AFF2 promoter methylation is associated with FRAXE repeat size and with psychotic experiences in general populations

Date: August 24, 2020

Objective of the study and its significance:

Fragile X syndrome is a genetic disorder characterized by mental impairment that affects males more severely than females. It is caused by silencing of the FMR1 gene on the X chromosome, typically caused by having 200 or more repetitions of CGG triplet within the FRAXA region of FMR1. Multiple lines of evidence indicate that silencing is mediated by DNA methylation of the FMR1 promoter and that FMR1 expression can be rescued by reducing DNA methylation. Fragile XE syndrome is similar but less common, less severe, and typically caused by high numbers of CCG repeats in the FRAXE region near the AFF2 (FMR2) gene.

Although repeat numbers in FRAXA and FRAXE are known to vary in the general population, typically below 40 in FRAXA and below 30 in FRAXE, little is known about the extent of this variation and how it might relate to DNA methylation and mental health.

To answer this question, we have measured FRAXA and FRAXE repeats in nearly 5000 males in the Avon Longitudinal Study of Parents and Children (ALSPAC). Genome-wide DNA methylation has been measured in blood samples collected from a subset of these at birth, age 7, age 15-17 using the Illumina HumanMethylation450 Beadchip and at ages 15-17 and 24 using the Illumina MethylationEPIC Beadchip. Psychotic-like experiences were self-reported by the participants when they were approximately 17.

We observed no evidence of an association of DNA methylation at age 15-17 with FRAXA repeats anywhere in the genome, but did observe extremely strong associations with FRAXE, particularly at two CpG sites cg20321768 and cg25587058 ($p < 1e-26$, n~730). Associations at these two sites are observed at all of the other DNA methylation time-points ($p < 1e-15$, n ~ 450 except for age 24 where n~240).

We observed associations between FRAXE repeats with two psychotic-like experiences: seeing someone no one else could see ($p = 1e-4$, n~1600) and the feeling that someone else's thoughts had been inserted ($p=1.8e-5$, n=1568). DNA methylation at age 15-17 at CpG sites cg20321768 and cg25587058 was associated with the latter ($p=0.027$ and 0.009, respectively, n=531) but not the former ($p>0.2$, n=441) nor with DNA methylation at any other time point.

We would like to determine if associations at CpG sites cg20321768 and cg25587058 with specific adolescent psychotic experiences are replicated in DNA methylation profiles measured in E-Risk participants at age 18.

Statistical analyses:

Associations to be tested using logistic regression with each individual self-reported psychotic experience at age 18 (recoded into binary variables: no/unsure vs yes) in turn as the dependent variable and DNA methylation at cg20321768 and cg25587058 as predictor (models run separately for each CpG site). Models will include covariates to adjust for cell count variation, age at time of interview, sex, BMI, cigarette smoke exposure (using CpG site cg05575921), genetic variation (cis SNP rs138007199) and technical variation in DNA methylation measurements (e.g. plate/batch as necessary). All analyses will account for the non-independence of twin observations. A sensitivity analysis will be undertaken in male study members only as the CpG sites are on the X chromosome and males only have one copy of this chromosome which might affect the results. For comparison with ALSPAC models, summary statistics will also be required for all covariates as well as the correlation between the two CpG sites.

Variables Needed at Which Ages (names and labels):

Study: E-Risk

Age 5:

FAMILYID	Unique family identifier
ATWINID	Twin A ID (ex chkdig)
BTWINID	Twin B ID (ex chkdig)
RORDERP5	Random Twin Order
ZYGOSITY	Zygosity
SAMPSEX	Sex of Twins: In sample

Age 18:

Relevant psychotic experiences (that match as closely as possible to ALSPAC items):

- FF1E18FIN Thoughts can be read by another - P18 - Elder
- FF3E18FIN Sent messages through radio or TV - P18 - Elder
- FF5E18FIN Being followed or spied on - P18 - Elder
- FF7E18FIN Heard voices others cannot hear - P18 - Elder
- FF9E18FIN Felt under the control of special power - P18 - Elder
- FF11E18FIN Read thoughts of another person - P18 - Elder
- FF13E18FIN See something others cannot see - P18 – Elder

- DNA methylation levels at cg20321768, cg25587058 and cg05575921 from blood with 450k chip
- Cell count variation for samples used to generate DNA methylation profiles
- Technical variation in DNA methylation measurements (e.g. plate/batch)
- Genotypes for SNP rs138007199
- TAGEE18 Age at Interview - P18 – Elder
- bmie18 Body mass index

Data Security Agreement

Provisional Paper Title	AFF2 promoter methylation is associated with FRAXE repeat size and with psychotic experiences in general populations
Proposing Author	Matthew Suderman
Today's Date	July 23, 2020

Please keep one copy for your records
(Please initial your agreement)

__ms__I am familiar with the King's College London research ethics guidelines (https://www.kcl.ac.uk/innovation/research/support/ethics/about/index.aspx) and the MRC good research practice guidelines (https://www.mrc.ac.uk/research/policies-and-guidance-for-researchers/good-research-practice/).

__ms__My project has ethical approval from my institution.

__ms__I am familiar with the EU General Data Protection Regulation (https://mrc.ukri.org/documents/pdf/gdpr-guidance-note-3-consent-in-research-and-confidentiality/), and will use the data in a manner compliant with its requirements.

__ms__My computer is (a) encrypted at the hard drive level, (b) password-protected, (c) configured to lock after 15 minutes of inactivity, AND (d) has an antivirus client which is updated regularly.

__ms__I will treat all data as "restricted" and store in a secure fashion.

__ms__I will not share the data with anyone, including students or other collaborators not specifically listed on this concept paper.

__ms__I will not merge data from different files or sources, except where approval has been given by the PI.

__ms__I will not post data online or submit the data file to a journal for them to post.
Some journals are now requesting the data file as part of the manuscript submission process. The E-Risk Study cannot be shared because the Study Members have not given informed consent for unrestricted open access. Speak to the study PI for strategies for dealing with data sharing requests from Journals.

__ms__Before submitting my paper to a journal, I will submit my draft manuscript and scripts for data checking, and my draft manuscript for co-author mock review, allowing three weeks.

__ms__I will submit analysis scripts and new variable documentation to project data manager after the manuscript gets accepted for publication.

__ms__I will delete the data after the project is complete.

__N/A__ **For projects using location data:** I will ensure geographical location information, including postcodes or geographical coordinates for the E-Risk study member's homes or schools, is <u>never</u> combined or stored with any other E-Risk data (family or twin-level data)

__ms__**For projects using genomic data:** I will only use the SNP and/or 450K data in conjunction with the phenotypes that have been approved for use in this project at the concept paper stage.

Signature: Matt Suderman

CONCEPT PAPER RESPONSE FORM

A. To be completed by the proposing author

Proposing Author: Matthew Suderman

X I have read the E-Risk data-sharing policy guidelines and agree to follow them

Provisional Paper Title: AFF2 promoter methylation is associated with FRAXE repeat size and with psychotic experiences in general populations

Potential co-authors: Jean Golding, Marcus Pembrey, Karen Sugden, Helen Fisher

Potential Journals: TBC

Intended Submission Date (month/year): Oct 2020

Please keep one copy for your records and return one to Louise (louise.arseneault@kcl.ac.uk)

B. To be completed by potential co-authors:

xxx☐ Approved ☐ Not Approved ☐ Let's discuss, I have concerns

Comments:

Please check your contribution(s) for authorship:

ξξ☐ Conceptualizing and designing the longitudinal study

ξξ☐ Conceptualizing and collecting one or more variables

ξξ☐ Data collection

☐ Conceptualizing and designing this specific paper project

☐ Statistical analyses

☐ Writing

☐ Reviewing manuscript drafts

☐ Final approval before submission for publication

ξξξ☐ Acknowledgment only, I will not be a co-author

Signature: .....Temi Moffitt.....