Infant Behavior and Development 60 (2020) 101409

Contents lists available at ScienceDirect

Infant Behavior and Development

journal homepage: www.elsevier.com/locate/inbede

Leveraging epigenetics to examine differences in developmental trajectories of social attention: A proof-of-principle study of DNA methylation in infants with older siblings with autism

Anna Gui[a,*], Emily J.H. Jones[a], Chloe C.Y. Wong[b], Emma Meaburn[a], Baocong Xia[b], Greg Pasco[c], Sarah Lloyd-Fox[d], Tony Charman[c], Patrick Bolton[b], Mark H. Johnson[d], the BASIS Team[1]

[a] Department of Psychological Sciences, Birkbeck College, University of London, UK
[b] Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[c] Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[d] Department of Psychology, Cambridge University, UK

ARTICLE INFO

Keywords:
Autism
DNA methylation
Developmental trajectory
Social attention
Infant siblings

ABSTRACT

Preliminary evidence suggests that changes in DNA methylation, a widely studied epigenetic mechanism, contribute to the etiology of Autism Spectrum Disorder (ASD). However, data is primarily derived from post-mortem brain samples or peripheral tissue from adults. Deep-phenotyped longitudinal infant cohorts are essential to understand how epigenetic modifications relate to early developmental trajectories and emergence of ASD symptoms. We present a proof-of-principle study designed to evaluate the potential of prospective epigenetic studies of infant siblings of children with ASD.

Illumina genome-wide 450 K DNA methylation data from buccal swabs was generated for 63 male infants at multiple time-points from 8 months to 2 years of age (total N = 107 samples). 11

Abbreviations: 450K array, Illumina Infinium HumanMethylation450 BeadChip; A, adenine; ADHD, Attention deficit and hyperactivity disorder; ADI-R, Autism diagnostic interview – revised; ADOS-2, Autism Diagnostic Observational Schedule – second edition; ADOS-G, Autism Diagnostic Observational Schedule – Generic; AOSI, Autism observation scale for infants; ASD, Autism spectrum disorder; Atyp. Dev., Atypical development; BASIS, British Autism Study of Infant Siblings; C, cytosine; CpG site, Nucleotide where a C is followed by a G in the DNA sequence; CNV, Copy Number Variant; DAWBA, Development and well-being assessment; DNA, Deoxyribonucleic acid; DNAm, DNA methylation; EEG, Electro-encephalography; ERP, Event-related potential; EWAS, Epigenome-wide association study; FDR, False discovery rate; fNIRS, functional near infrared spectroscopy; G, guanine; GWAS, Genome-wide association study; EL, Elevated-likelihood; TL, Typical-likelihood; MSEL, Mullen scales of early learning; mQTL, Methylation Quantitative Trait Locus, i.e., genetic variants that are known to have an effect on the DNA methylation levels at specific methylation sites; ms, milliseconds; N, Number of units; Nc, Negative central; RNA, Ribonucleic Acid; SCQ, Social communication questionnaire; SD, Standard deviation; s.e., Standard error; SFARI, Simons Foundation Autism Research Initiative; SNP, Single nucleotide polymorphism, i.e., variations in nucleotide sequence; SRS, Social responsiveness scale; T, thymine; T1, first visit of the BASIS protocol (~ 8 months of age); T2, second visit of the BASIS protocol (~15 months of age); T3, third visit of the BASIS protocol (~25 months of age); T4, fourth visit of the BASIS protocol, when diagnostic assessment is made (~3 years of age); Typ. Dev., Typical development; VABS, Vineland adaptive behavior scales; WGCNA, Weighted gene co-methylation network analysis

* Corresponding author at: Centre for Brain and Cognitive Development, Birkbeck College, University of London, BMA Tavistock House North (Entrance A), Tavistock Square, London WC1H 9HR, UK.
E-mail address: agui01@mail.bbk.ac.uk (A. Gui).

[1] The BASIS Team members listed in alphabetical order: Baron-Cohen S., Blasi A., Cheung C., Davies K., Elsabbagh M., Fernandes J., Gammer I., Ganea N., Gliga T., Guiraud J., Hendry A., Liersch U., Liew M., Maris H., O'Hara L., Pickles A., Ribeiro H., Salomone E., Taylor, C., Tucker L., Tye C., Wass S.

https://doi.org/10.1016/j.infbeh.2019.101409
Received 28 February 2019; Received in revised form 5 December 2019; Accepted 6 December 2019
0163-6383/ © 2019 Published by Elsevier Inc.

of those infants received a diagnosis of ASD at 3 years. We conducted a series of analyses to characterize DNA methylation signatures associated with categorical outcome and neurocognitive measures from parent-report questionnaire, eye-tracking and electro-encephalography.

Effects observed across the entire genome (epigenome-wide association analyses) suggest that collecting DNA methylation samples within infant-sibling designs allows for the detection of meaningful signals with smaller sample sizes than previously estimated. Mapping networks of co-methylated probes associated with neural correlates of social attention implicated enrichment of pathways involved in brain development. Longitudinal modelling found covariation between phenotypic traits and DNA methylation levels in the proximity of genes previously associated with cognitive development, although larger samples and more complete datasets are needed to obtain generalizable results.

In conclusion, assessment of DNA methylation profiles at multiple time-points in infant-sibling designs is a promising avenue to comprehend developmental origins and mechanisms of ASD.

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by difficulties with social interaction and communication, and the presence of restricted interests, repetitive behaviors and sensory anomalies (American Psychiatric Association, 2013). Behavioral symptoms of ASD emerge gradually over the first few years of life, such that a stable diagnosis can often be made by age 3–5 years (Zablotsky et al., 2017). Identifying the mechanisms that underpin behavioral symptoms is important for understanding the etiology of ASD, and for designing new focused intervention strategies. Since genetic and environmental risk factors for ASD can act prenatally (Lasalle, 2013), to identify the potential mechanisms that lead to symptoms we need to study early development.

One fruitful approach has been the prospective longitudinal study of infants with an older sibling with ASD. Such infants are considered at elevated likelihood of ASD, as they have a higher chance of developing the condition themselves (20 %) compared to the general population (typically ~1 %, Ozonoff et al., 2011). These studies have provided substantial insight into the early behavioral and neurocognitive profiles that precede later symptomatology (Elsabbagh & Johnson, 2010; Szatmari et al., 2016). In parallel, post-mortem genetic and epigenetic studies have discovered that sets of genes implicated in ASD are typically expressed during fetal and infant development (Grayson & Guidotti, 2015). To date these efforts have proceeded largely in parallel, with little attempt to integrate insights from genetics and developmental cognitive neuroscience to provide a full picture of how genetic risk is linked to neurocognitive vulnerability and behavioral symptoms. In this manuscript, we present a proof-of-principle that integrating epigenetics into prospective longitudinal studies of infant siblings of children with ASD holds great potential for closing this gap. These insights will also be valuable to the broader community of researchers interested in the biological foundations of early developmental processes.

1.1. Early neurocognitive pathways to ASD

Although the average age of diagnosis of ASD in the UK is typically around age 5, parents tend to report they first became concerned about their child's behavior before age 2 (Crane, Chester, Goddard, Henry, & Hill, 2016). Prospective longitudinal studies of infants with an older sibling with ASD have examined how behavioral symptoms unfold over developmental time. Such studies follow infants from close to birth to three years, an age at which stable diagnosis of ASD can be made (Ozonoff et al., 2015). Data is analyzed retrospectively based on the child's diagnostic status at the outcome visit. Perhaps surprisingly, 6- to 8-month-old infants who later receive a diagnosis of ASD appear to be relatively typically developing at the behavioral level (Jones, Gliga, Bedford, Charman, & Johnson, 2014). However, in the second postnatal year infants with emerging ASD show gradual declines in social interest and delayed or slower communication development that gradually accumulates (Ozonoff et al., 2014). By 14 months, behavioral signs of ASD begin to show some predictive validity for a later diagnosis (Bussu et al., 2018), and by 24 months a diagnosis is often possible (Szatmari et al., 2016). Thus, the period between 8 and 24 months is particularly critical for identifying the causal processes involved in ASD. The mechanisms that underpin this behavioral trajectory remain unclear. However, one particularly strong candidate that has been suggested to contribute to the emergence of ASD is disruption in how the child's brain engages in focused attentive states during social interaction (Klin, Shultz, & Jones, 2015).

1.1.1. Social attention in the developmental path to ASD

Attention is a domain-general cognitive function which serves to prioritize environmental information for in-depth processing, learning and memory (Amso & Scerif, 2015; Scerif, 2010). A long history of research has identified robust neurocognitive correlates of the state of focused attention that enhances learning in the infant brain. When infants are focally attentive, they tend to show longer epochs of looking with minimal movement (Colombo & Cheatham, 2006; Colombo & Mitchell, 2009). Further, when infants are in an attentive state, averaged profiles of brain activity captured with electroencephalography (EEG) show a larger and faster negative deflection over frontocentral electrodes around 500 ms after a stimulus is presented. This response has been identified as the event-related potential (ERP) Negative Central component or Nc (Richards, Reynolds, & Courage, 2010).

Attention likely plays a critical role in the experience-dependent specialization of the social brain. Basic orienting mechanisms supporting rapid attention to salient, face-like stimuli are present from birth (Senju, Johnson, & Tomalski, 2014) and assist in the tuning up of cortical areas for face processing and social cognition (Johnson, 2005). Evidence from the study of 6- to 12-month-old typically developing infants indicates that there is a progressive selectivity in neural response to social vs non-social stimuli (Jones, Venema, Lowy, Earl, & Webb, 2015). Thus, the first year of life appears to be critical for the infant brain and entails tuning up a wide network of specialized areas that, initially biased towards social cues, gradually learn to select relevant information from these social stimuli (Csibra & Gergely, 2009) and to use this information to actively engage in joint attention (Mundy, 2018).

When look durations and ERPs in response to social and non-social stimuli have been examined, infants with a later diagnosis of ASD show profiles consistent with diminished attention engagement to faces in the first eight months of life (Elsabbagh, Mercure et al., 2012; Jones et al., 2016; Jones & Klin, 2013). Reduced Nc in response to faces with direct gaze has been consistently shown to be a precursor of later social difficulties seen in children with ASD (Dawson et al., 2012; Gui et al., under review; Jones et al., 2016; Webb et al., 2011). Additionally, previous studies using eye-tracking revealed that atypical developmental trajectories of face looking duration are observed after 12 months not only in infants with emerging ASD, but also in infants with a familial history of ASD without a later diagnosis (Elsabbagh, Gliga et al., 2012; Hendry et al., 2018; Jones et al., 2016). Thus, it is possible that genetic or environmental risk factors for ASD impact the brain systems necessary to maintain focused attention to social stimuli. This process may reduce the child's ability to learn from people around them, which gradually makes the social world less comprehensible as it becomes more complicated (Klin et al., 2015). Infants may then gradually withdraw from the people in their environment as an adaptive response, producing subsequent symptoms of ASD (Johnson, 2017).

Evidence supporting this account could be obtained through linking cognitive components like social engagement and other early signs of atypical developmental trajectory in infants with emerging ASD to the biological processes that have been associated with clinical diagnosis in large samples. One way to make progress in this area is through genetics. Indeed, twin studies indicate that social engagement (based on eye-tracking measures of face and eye looking) may be heritable (Constantino et al., 2017), supporting the possibility that early social attention difficulties might serve as *intermediate phenotypes* between ASD risk factors and the emergence of social difficulties. However, just as we need to take a developmental approach to studying the neurocognitive processes underpinning ASD, we need to take a developmental approach to genetics. Traditional behavioral or molecular genetics does not account for the

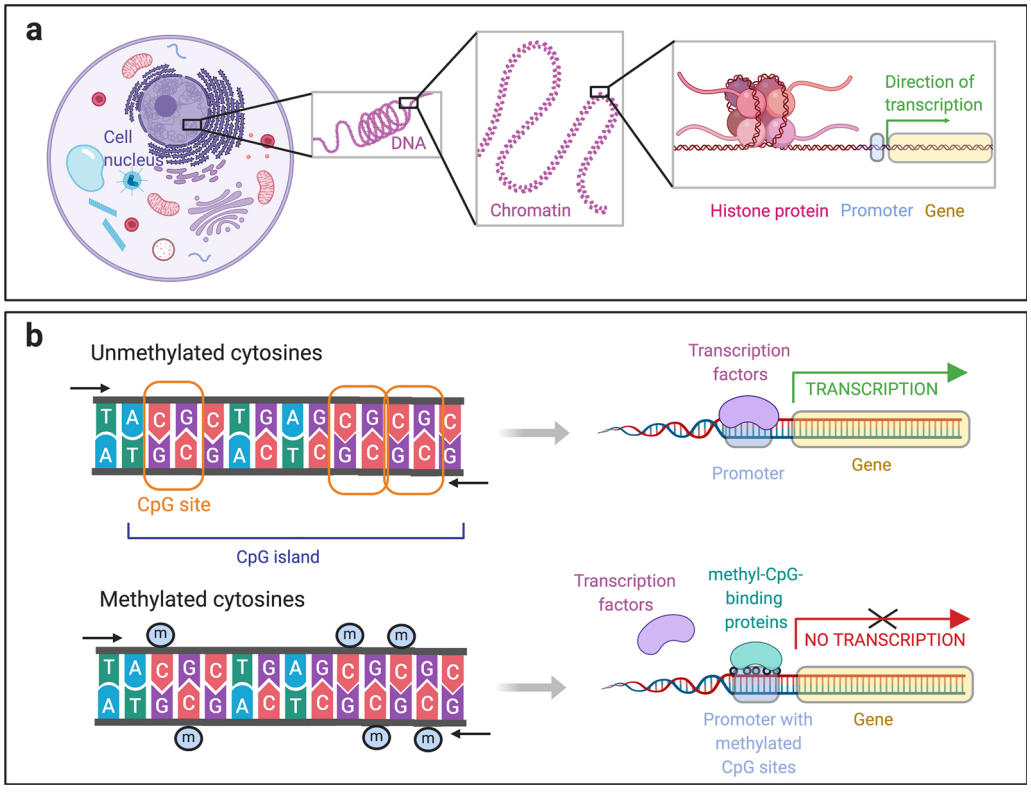

Fig. 1. a) Within the nucleus of each cell, the DNA is packaged into chromatin, a complex formed by long DNA molecules wrapped around histone proteins. Genes are strings of DNA that code for proteins; promoters are regions of DNA that control when and where transcription of a gene is initiated. **b**) CpG islands are strings of DNA where there is a relatively high proportion of Cs whose downstream neighbour is G (CpG sites). When a methyl group (m) attaches to a CpG site in the promoter region of a gene, that gene is less likely to be accessible to transcription factors to be copied into RNA. These diagrams simplify the DNA greatly – promoters and genes can be dozens or even hundreds of base pairs. This figure was created with BioRender.com (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

fact that gene expression patterns are not developmentally static; rather, they change over developmental time and there are substantial individual differences in these changes (Moore, 2015). Understanding the emergence of ASD requires us to study dynamic changes in both neurocognitive systems and genome function in parallel.

1.2. Developmental approaches to genetics: peripheral epigenetics

The term "Epigenetics" literally means "above the genome" and was originally defined as the field that studies the developmental processes connecting genotype and phenotype (Waddington, 1942). The genome is the entire string of DNA (deoxyribonucleic acid) that resides in a cell nucleus. DNA is made up of two intertwined strands composed of nucleotides, which are composed of three elements: a phosphate group, a five-carbon sugar (deoxyribose) and a nitrogenous base attached to it. Nucleotides in DNA can contain one of four different bases, which are referred to by the letters A (adenine), T (thymine), C (cytosine) and G (guanine). These nucleotides are organized into triplets that represent codes for different amino acids. Each 'gene' consists in a long string of triplets. In order to be turned into proteins, gene codes must be copied, or 'transcribed', into RNA (ribonucleic acid) and then 'translated' into strings of amino acids. Amino acids then group together to form proteins, which take up different roles in the cell. The degree to which a particular protein is made within a cell depends on whether the corresponding region of the DNA is available to be transcribed into RNA. Within the nucleus of each cell, DNA is wrapped around histone proteins, forming a complex called chromatin (see Fig. 1.a). Modifications of the chromatin structure determines the accessibility of genes to be transcribed into RNA, while leaving the DNA code intact (Rudenko & Tsai, 2014). Epigenetics involves the study of chemical modifications to chromosome structure which modify gene expression (Wolffe & Matzke, 1999). Critically, epigenetic markers can change across development and between cells, whilst the genome itself is fixed. Studying epigenetics thus gives us an opportunity to probe how genes and environment interact throughout development (Wiers, 2012).

1.2.1. DNA methylation

The most widely studied mechanism of epigenetic regulation in mammalian genomes is DNA methylation (DNAm). DNAm refers to the addition of a methyl group to a C (cytosine) nucleotide. In humans, methylation of cytosines typically occurs when a C is next to a G in the sequence (i.e., at a CpG dinucleotide or site). When a methyl group attaches to a CpG site, various proteins bind to the methylated C and recruit other proteins at the same location. This then means that the gene that includes that CpG site is less likely to be accessible to being copied into RNA. Often, stretches of DNA where CpG sites occur with higher frequency (CpG islands) are located in 'promoter regions' – regions where key proteins bind to a gene to copy it into RNA. Methylation in CpG islands thus results in silencing of gene expression, because they limit the accessibility of DNA to the key proteins (transcription factors) that are required to transcribe DNA into RNA, critical in later creating proteins (see Fig. 1.b). A complex enzymatic machinery maintains a balance between stability and change in these methylation patterns across the lifespan (Geiman & Muegge, 2010).

1.3. DNA methylation programming during development

DNAm plays a crucial role in brain development. The fastest changes in DNAm occur during the fetal period (Numata et al., 2012). During the early phases of fetal life, two major waves of epigenetic reprogramming take place that remove epigenetic patterns from the sperm and the egg and establish a new epigenetic profile in the developing embryo (Geiman & Muegge, 2010). After the embryo implants in the uterus, changes in DNAm regulate the differentiation of cells into different types (like neurons, somatic cells or glial cells) (Moore, Le, & Fan, 2013). Although *global levels* of methylation do not change drastically after birth (Slieker, Roost, Van Iperen, & Suchiman, 2015), animal work indicates that DNAm *patterns* can change postnatally in interaction with the environment. In mice, extracellular signals and neuronal electrical activity both influence neuronal DNAm levels, inducing long-lasting methylation changes (Guo et al., 2012; Lister et al., 2013). Further, human studies have demonstrated that DNAm levels at CpG sites in the proximity of genes relevant for neurodevelopment can be associated with environmental exposure to adverse events, for example prenatal exposure to toxicants (Kundakovic et al., 2015), parental risks such as substance use or psychopathology (Cecil et al., 2014), malnutrition in infancy (Peter et al., 2015), stress (Essex, Boyce, Hertzman, Lam, & Armstrong, 2013), and bullying victimization (Ouellet-Morin et al., 2013) during childhood. These associations have often been linked to variation in behavioural phenotypes too, suggesting that DNAm may be a mediator between early risk exposure and later psychopathology (Barker, Walton, & Cecil, 2018). Changes in DNAm have also been linked to change in brain activity in adults (e.g., Frodl et al., 2015; Hass et al., 2015; Ursini et al., 2011). Thus, there are preliminary indications of associations between DNAm, brain activity and cognitive functions.

1.4. DNA methylation and ASD

DNAm is essential for typical neurodevelopment and altered methylation patterns have been linked to ASD (Ciernia & Lasalle, 2016). For example, Rett syndrome is a severe neurodevelopmental disorder characterized by brain abnormalities and gradual emergence of autistic features as well as motor and language regression (Neul et al., 2011). Although initially identified through clinical phenotyping, Rett syndrome is now known to be caused by a mutation in the MeCP2 gene. MeCP2 modulates the degree to which other genes are translated into proteins through its ability to bind to CpG sites and recruit other proteins that have the function to switch off genes (i.e., repression factors, Rudenko & Tsai, 2014). This suggests that alterations in DNAm patterns likely produce effects on neurodevelopment similar to those observed when mutations in the MeCP2 gene occur (Rubenstein & Merzenich, 2003). Further, rare *de novo* mutations in the DNMT3a gene, involved in the regulation of embryonic methylation, have been found in people

with ASD (Sanders et al., 2015). Additionally, there are greater age-related DNAm changes in the human cortex during fetal and postnatal development within genes implicated in schizophrenia and ASD relative to other regions of the genome (Jaffe et al., 2016; Numata et al., 2012; Spiers et al., 2015). This suggests that disruptions in the epigenetic machinery could have a particularly significant effect on the formation of brain circuits relevant to these disorders. In line with this hypothesis, a recent large Epigenome-Wide Association Study (EWAS), examining associations of DNAm levels of many individual methylation sites at once with a variable of interest, found that differences in DNAm in newborn blood samples were correlated with increased genetic risk for ASD (Hannon et al., 2018). Infant epigenetic marks, including the location of methylated CpGs, are likely to be traces of developmentally relevant interplays between genes and environment (Lasalle, 2013). Thus, critical period-specific investigations of the association between DNAm levels and developmental neurocognitive phenotypes may illuminate the causal pathway to ASD symptoms.

1.5. Limitations of previous studies that can be addressed in the infant-sibling design

Although highly promising, studying the relation between emergence of ASD and DNAm in humans poses practical, technical and theoretical challenges associated with data collection (Kato & Iwamoto, 2014; Relton & Smith, 2012). Since DNAm is responsible for shaping the differences between neurons, glia and other cell types, epigenomic patterns vary across tissues. Selecting a relevant tissue to study is thus important (Michels et al., 2013). Brain tissue is considered the most relevant tissue for investigating mechanisms involved in neurodevelopmental disorders. Indeed, the study of epigenetic signatures in brain tissue have provided fundamental advances in our knowledge of biological mechanisms underlying the emergence of ASD (Sun, Zang, Shu, & Li, 2014; Wong et al., 2018). However, limited availability of human post-mortem brain tissue is a major issue, especially when studying early development and neurodevelopmental disorders. Moreover, DNAm in post-mortem brains will be affected by how samples are processed, donor characteristics (that are often not available), cause and timing of death (Lim et al., 2014), all of which will introduce noise and hold the possibility to confound findings. Further, to really understand the causal nature of the relationship between DNAm patterns and the emergence of ASD traits across development, multiple measurements from the same individual are required (Martino et al., 2013; Michels et al., 2013). This is clearly impossible when using post-mortem brain tissue.

The importance of tracking individual developmental trajectories of DNAm justifies the use of more available peripheral tissues, such as whole blood and buccal epithelial cells isolated from saliva (Barker et al., 2018). Saliva and buccal swabs are relatively easy to collect, especially from young children and individuals with ASD who might be distressed by a blood draw procedure. Encouragingly, buccal epithelial cells are potentially more closely related to brain methylation patterns than blood, as they come from the same primary germ layer in the embryo, i.e. the ectoderm (Mitchell, Schneper, & Notterman, 2015). Corroborating this, Smith et al. (2015) found that DNAm levels in saliva with a high number of epithelial cells appeared generally more similar to patterns from various brain regions than DNAm levels in blood.

1.6. Aims of the study

Linking DNAm data to neurocognitive measures collected within a prospective longitudinal study of infants at elevated likelihood of ASD (also referred to as 'high-risk infants' in the literature) could enable unique insights into how epigenetics shape brain development. However, the methodological and practical challenges mean this work is highly novel, expensive and risky, and thus the field is in need of proof-of-principle studies to determine its potential value. In this study, we explored the potential of using buccal swab samples to analyze DNAm at multiple time points in a study of infants with emerging ASD. We present four approaches to explore how variations in DNAm signatures associate with behavioural manifestations of atypical development, and with neurocognitive measures of social attention in infancy that are potential intermediate phenotypes between genetic factors and later ASD traits (Constantino et al., 2017; Jones, Venema, Earl, Lowy, & Webb, 2017).

Analyses aimed to identify the association between DNAm profiles and five phenotypes of interest: (*i*) case-control status for ASD and (*ii*) case-control status for atypical development at 3 years of age (categorical outcomes); (*iii*) continuous measure of adaptive skills at 3 years of age (dimensional outcome), and two infant intermediate phenotypes: (*iv*) face looking time and (*v*) neural response to faces, i.e., amplitude of the Nc event-related component. Measures *iv* and *v* have been selected as candidate intermediate phenotypes as they have been previously associated with ASD outcome (Hendry et al., 2018; Jones et al., 2016), and parental variation in ASD-related traits (Jones et al., 2015). We were particularly interested in whether the use of candidate intermediate phenotypes could increase power for identifying relations with epigenetic variation, as has been previously suggested (Flint & Munafò, 2007).

First, in Analysis 1 we explored whether global methylation levels (calculated by averaging all interrogated methylation sites across the microarray, or 'probes') are sensitive to our five phenotypes of interest (*i-v*).

Second, in Analysis 2 we performed EWASes to identify potential significant associations between DNAm levels at individual methylation sites and categorical/dimensional ASD outcomes (*i-iii*), as well as our candidate intermediate phenotypes (*iv,v*). We explored overlap of any identified probes with independent developmental and ASD-related methylation datasets.

Third, in Analysis 3 we used a network-based analysis (Langfelder & Horvath, 2008) to look for associations between the DNAm profiles of groups of related probes and both categorical and dimensional outcomes (*i-iii*) and candidate intermediate phenotypes (iv,v). This method has been previously used to study developmental DNAm changes (Spiers et al., 2015) and relationships with dimensional ASD traits (Ginsberg, Rubin, Falcone, Ting, & Natowicz, 2012). We used this information in association with upstream genetic regulatory information (regions of the genome that are known to have an effect on the DNAm levels of other genes at particular developmental stages – so called 'methylation quantitative traits loci', or mQTL, Gaunt et al., 2016) to ask whether the analysis identified biologically relevant pathways potentially implicated in the early emergence of ASD phenotypes.

Finally, in Analysis 4 we explored the potential for joint longitudinal analysis of both epigenetic change and change in behavior using two dimensional measures available at multiple time-points (adaptive skills from parent-report, *iii*, and looking time to faces, *iv*). Taken together, we aimed to determine whether epigenetics from buccal epithelial cells provides a potentially valuable approach for the field of developmental cognitive neuroscience, and if so what analytic approaches and phenotypic associates are most appropriate to examine.

2. Materials and methods

2.1. Participants

The sample for the present study is derived from the Phase 2 of the British Autism Study of Infant Siblings (http://www.basisnetwork.org), which consists of 143 children (65 females). 116 were younger siblings of children with ASD (elevated-likelihood infants, EL, 52 females) and 27 were infants with no first- or second-degree relatives with ASD (typical-likelihood controls; TL, 13 females). Children took part in research assessments when they were around 8 months (T1), 15 months (T2), 25 months (T3) and 36 months (T4). Research assessments included collection of parent-report questionnaires, experimental tasks with eye-tracking and EEG and standardized behavioral assessments. At the 36-months visit, experienced clinical researchers reviewed all the available measures to determine the clinical outcome.

Genome-wide DNAm profiles from 107 buccal swab DNA samples were available for a sub-set of 63 unrelated male infants from this cohort (49 EL, of whom 11 developed ASD at 3 years of age) across T1, T2 and T3. For 33 participants (21 EL, of whom 5 had ASD), DNA samples were collected at more than one visit and DNAm assessed, resulting in longitudinal DNAm data for these infants. Table 1 provides information on the total number of good-quality DNA samples and the age of the participants when samples were collected.

DNAm was quantified using the Illumina Infinium HumanMethylation450 BeadChip (hereafter, 450 K array), which assesses DNAm at over 450,000 methylation sites (~2 % of all possible sites) throughout the genome at single-nucleotide resolution, including 96 % of the CpG islands and 99% of the known genes (Illumina, 2012). DNA extraction, DNAm array data generation and quality control procedures were performed at King's College London following established guidelines as detailed in Appendix B. For each probe on the BeadChip, relative methylation (β value) was calculated as the ratio of the normalized signal from the methylated probe to the sum of the normalized signals of the methylated and unmethylated probes. Thus, a value ranging from 0 (unmethylated) to 1 (fully methylated) was obtained for each of the 402,971 interrogated probes that passed quality control.

2.2. Phenotypes

2.2.1. Categorical outcome: ASD and atypical development

At three years, children underwent diagnostic assessment for ASD, including standardized and semi-standardized behavioral assessments (the Mullen Scales of Early Learning, MSEL, Mullen, 1995, and Autism Diagnostic Observational Schedule, ADOS, Lord et al., 2000) and parent-report interviews (the Vineland Adaptive Behavior Scales, VABS, Sparrow, Cicchetti, & Balla, 2005, and Autism Diagnostic Interview - Reviewed, ADI-R, Lord, Rutter, & Le Couteur, 1994). Based on these assessments and unstructured observations, experienced clinicians determined whether the EL children met criteria for ASD, and used their determinations to assign the children to outcome groups. Further details on the classification criteria can be found in the Supplementary material (Appendix A). By the outcome visit (T4), none of the TL children had received a community diagnosis of ASD nor met criteria for an ASD research diagnosis (see Table A.1). Following previous work, analyses compared outcome groups based on 1) presence or absence of ASD and 2) presence or absence of atypical development. To evaluate effects of ASD, EL children who received a diagnosis of ASD at T4 (N = 10) were compared with a no-ASD group, which included EL children with no ASD and TL children (N = 41). To evaluate the effects of atypical development, infants with typical developmental outcome at T4 (TL and EL children with no signs of developmental delay, N = 28) were compared with children who received a diagnosis of ASD or showed sub-threshold levels of ASD symptoms and/or more general developmental delays but did not fully meet criteria for ASD (N = 23).

Table 1
Numbers and age of the study participants for the three time-points (visits) and number of unique participants.

Visit	Recruitment groups	N groups	Total N	Mean (SD) age in months	Min – Max age in months
T1	TL	9	51	8.7 (0.8)	8 - 10
	EL	42 (10 ASD)			
T2	TL	9	24	15.3 (1.1)	14 - 18
	EL	15 (6 ASD)			
T3	TL	12	32	25 (1.7)	24 - 32
	EL	20 (3 ASD)			
Unique participants	TL	14	63		
	EL	49 (11 ASD)			

Recruitment groups: TL: typical-likelihood infants; EL: elevated-likelihood infants; ASD: EL infants who received a diagnosis of ASD at age three years. N: number of participants. SD: standard deviation. Min: minimum age. Max: maximum age.

2.2.2. Dimensional outcome: adaptive skills

The standard composite adaptive behavior scores from the second edition of the VABS was used as a measure of dimensional developmental outcome of adaptive behavior at three years of age (Sparrow et al., 2005). VABS is a semi-structured interview measuring adaptive functioning in everyday life. It has been extensively used to study variability in early difficulties in adaptive behavior and to capture differences in developmental trajectories in infants at elevated likelihood of ASD (Bussu et al., 2018; Estes et al., 2015). In the present research, VABS was collected as a parent-report questionnaire at T1 and T2 and as a parent interview at T3 and T4. VABS composite scores at T4 were used as a measure of dimensional adaptive outcome in Analyses 1–3, with lower scores indicating lower adaptive functioning (Sparrow et al., 2005). VABS composite scores of the same scale at T1, T2 and T3 were used in Analysis 4.

2.2.3. Candidate intermediate phenotypes

2.2.3.1. Face looking. Infants completed a face 'pop-out' task at T1, T2 and T3. Data from T1 were used for Analysis 1–3; in Analysis 4 we used data from all three visits. The face pop-out task was created to measure infant's attraction to and preference for faces. The full paradigm has been described previously (Elsabbagh, Gliga et al., 2012; Hendry et al., 2018). At each timepoint, infants were presented with a series of visual arrays composed of five images: face, bird, car, mobile phone and a non-social control stimulus, henceforth "noise", created by randomizing the phase spectra of the face stimulus while keeping the amplitude and color spectra constant (Halit, Csibra, Volein, & Johnson, 2004). An example of the visual array presented as face pop-out stimulus is shown in Fig. 2. Infants watched 10 different arrays, each presented for 15 s. Looking behavior was recorded either with a Tobii 120 eye-tracker at a rate of 60 Hz or a Tobii 1750 at a rate of 50 Hz. Each testing session was preceded by a 5-point calibration sequence.

From this data, we extracted the average duration of the longest look (peak look) at the face image (Jones et al., 2016). This biometric is considered to reflect endogenous attention allocation from 6 months of age (Colombo & Cheatham, 2006). Peak look durations were calculated as described by Hendry and colleagues (2018). Briefly, the peak look to the face was identified for each trial in which there was a minimum of two looks (> 100 ms). Peak look durations were averaged across the trials (minimum 3/child) to provide a more stable characterization of individual differences.

2.2.3.2. Nc to face. At T1, infants sat on their parent's lap at 60-cm distance from a 40 × 29-cm computer screen while brain activity was measured simultaneously with a 128-channel Hydrocel Sensor Net. The EEG task was as described in Elsabbagh, Mercure et al. (2012). Briefly, infants were presented with a series of 50 trial blocks. Each trial block began with a static colorful fixation stimulus, followed by four to seven color pictures of a female model whose gaze is either direct or averted. Additionally, in approximately one third of the blocks a control-stimuli (the 'noise' stimulus described above) was presented. The trial duration was 800 ms, followed by

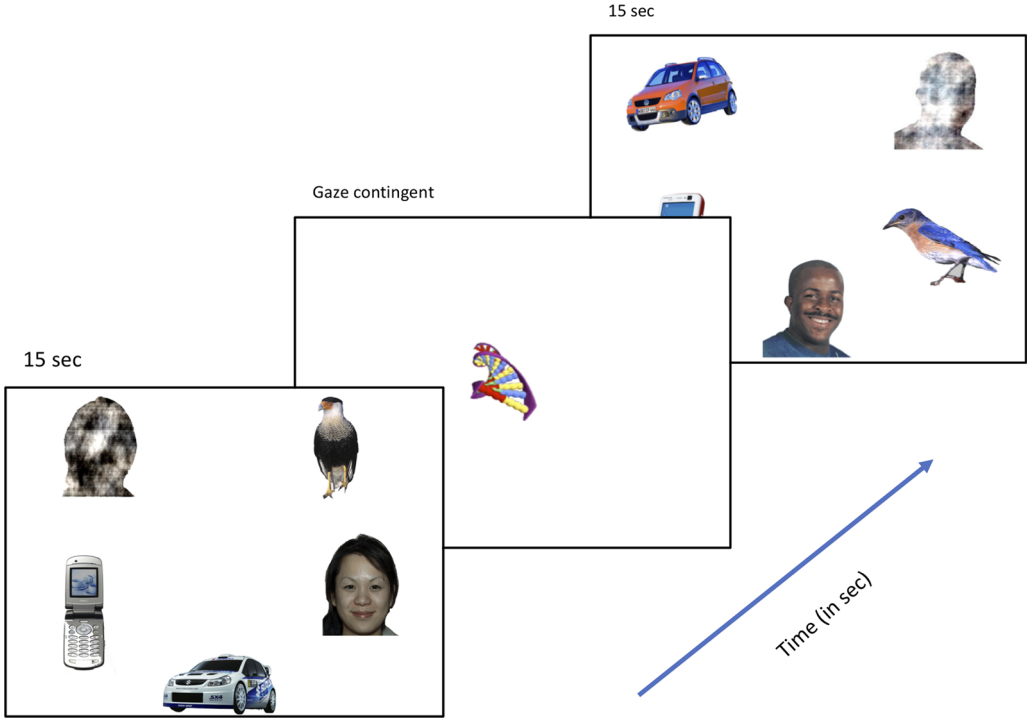

Fig. 2. Examples of the slides of the face pop-out eye-tracking task. The first and third slide display arrays of five static visual stimuli including a face; the second slide presents an animation which served to attract the infant's gaze to the center of the screen before the presentation of the next slide.

a 500-ms interval with no visual stimulus. Fig. 3 illustrates the task stimuli. Gaze was recorded with a video camera and trials were excluded for inattention, blinks or looking away during the 800 ms following stimulus onset.

In NetStation 5, the EEG recording was vertex referenced, band-pass filtered (0.1–100 Hz) and segmented into stimulus-locked epochs (-200 to 800 ms peri-stimulus window) by condition. Segments showing significant artifact were removed through automatic detection and hand-editing. Clean segments were averaged within condition; bad channels were replaced through spine interpolation and then data were average re-referenced and baseline-corrected (195 ms before stimulus onset). Infants were excluded if good quality EEG data was obtained for fewer than 10 trials for the conditions of interest.

Following previous research (Jones et al., 2016; Richards et al., 2010; Webb et al., 2011), the Nc component was used as a neural correlate of attention engagement. Nc amplitude was defined as the mean amplitude of the negative deflection between 300 and 800 ms after stimulus onset across left and right frontal regions. Our phenotypic measure of interest was the Nc amplitude difference score between the face with direct gaze (henceforth face) and noise stimuli, to reflect attention engagement processes specific to salient social content. Previous analyses including the present participants demonstrated that this variable was associated with later individual variability in social adaptive skills, and that this effect was specific to faces with direct gaze (Gui et al., under review). Therefore, EEG data for the face with averted gaze condition were not included in this study.

2.3. Statistical analyses

Four different investigations (Analyses 1–4) were conducted on the present dataset. Table 2 illustrates the number of DNA samples for each Analysis and when they were collected.

2.3.1. Analysis 1: global DNA methylation levels

A global level of methylation was calculated for each individual at each visit (T1, T2, T3) as a mean of methylation levels across all 402,971 probes. All 107 samples were included in this analysis. Multi-level mixed models using the 'lme' function within the nlme R-package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2018) were used to explore the longitudinal trajectory of the change in global methylation level, while accounting for repeated measurements from the same participant with a random-effect term. Fixed effects of group, visit and their interaction were tested by comparing the overall fit of the multilevel mixed-effects models using a chi-square likelihood ratio test. The significant effect of visit was estimated with a linear model with T1 as reference. Finally, linear regressions were used to test associations with dimensional outcome measures (adaptive skills) and candidate intermediate phenotypes (Nc to face, face looking). To control for possible batch effects (i.e., systematic errors occurring when samples are processed in multiple sets), we examined effects of prenatal exposure to smoke and alcohol (Grayson & Guidotti, 2015) and effects of age differences within each visit, Appendix C (Tables C.1-C.5) reports results of all the multilevel mixed-effects models including these variables as covariates.

2.3.2. Analysis 2: Epigenome-wide association analyses (EWAS)

EWAS is a statistical approach that involves systematically testing thousands of probes distributed throughout the genome for association with a phenotypic trait. Typically, linear models are employed where the phenotype is the dependent variable, and a quantitative measure of DNAm (ranging from 0 to 1) is the independent variable. Due to the multiple testing issue, a stringent p-value threshold of $p \leq 2.4 \times 10^{-7}$ has been established as the threshold for statistical significance for studies using the 450 K array, like the present one (Saffari et al., 2018). When no epigenome-wide significant hits are identified, a common approach in the ASD literature is to highlight probes for follow-up analyses that have a more liberal "discovery" p-value threshold, for example $p < 5 \times 10^{-5}$ (see Andrews et al., 2018; Hannon et al., 2018).

In the present study, DNAm collected at 8 months was used for EWAS analyses (N = 51). The choice to restrict this analysis to the first visit was motivated by the importance of narrow age ranges for DNAm analyses (Michels et al., 2013) and the fact that behavioral symptoms begin to emerge after this time-point (Szatmari et al., 2016). Table 3 summarizes the sample sizes for this analysis for each of the phenotypes of interest. Table A.2 provides mean (standard deviation) values for the continuous phenotypes, by

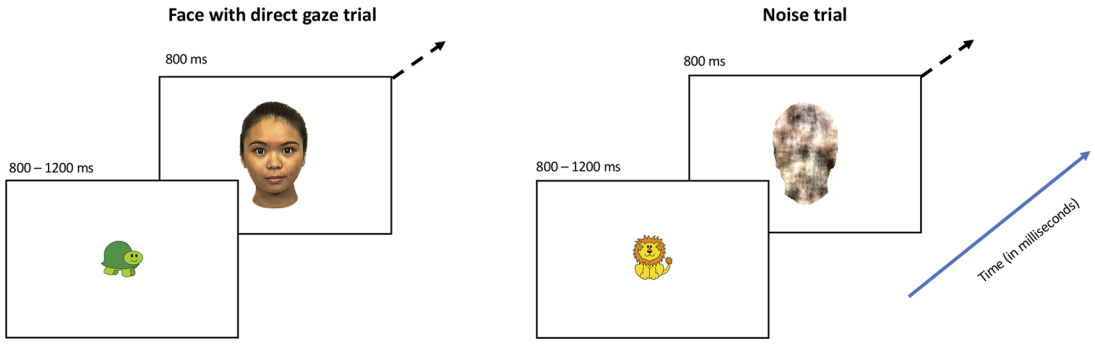

Fig. 3. Examples of the 'face with direct gaze' and 'noise' trials of the EEG task. Blocks started with a figure of a toy to attract the infant's gaze to the center of the screen. Each block presented either images of faces or the 'noise' stimulus.

Table 2
Number of samples included in all the analyses for the present study and specification of the visit when the samples were collected (T1: 8 months, T2: 15 months, T3: 25 months).

	DNA samples for the present study	
	Total N	Visit
Analysis 1: Global DNA methylation levels*	107	T1 (N = 51), T2 (N = 24), T3 (N = 32)
Analysis 2: Epigenome-wide association analysis	51	T1
Analysis 3: Co-methylation network Analysis	51	T1
Analysis 4: Longitudinal analysis*	33 VABS/ 27 Face pop-out	T1,T2,T3

* model accounting for repeated measurements from the same individual.

outcome group. A multiple regression was implemented through the *cpg.assoc* R-package (Barfield, Kilaru, Smith, & Conneely, 2012) to test the association between DNAm levels of individual probes, while controlling for possible confounders due to ancestry differences (Barfield et al., 2014) and batch effect (see Appendix D). Five EWASes were run to test for association with our five phenotypes of interest: three measures of developmental outcome (categorical outcome of ASD, categorical outcome of atypical development and dimensional outcome measured with VABS composite score) at T4, and the two infant phenotypes measured at the same visit when the DNA sample was collected (T1): Nc to faces and face looking.

We then used two approaches to examine whether the discovery-significant probes ($p < 5 \times 10^{-5}$, as in previous research testing the association between peripheral DNAm and later ASD, Hannon et al., 2018) appear to be identifying potentially meaningful signals. Given the focus on ASD, our first approach was to examine the proportion of probes located in high confidence ASD genes in the SFARI Gene database (www.gene.sfari.org, Banerjee-Basu & Packer, 2010, release Nov. 2018). This publicly available database includes all genes which have been previously associated with ASD in published scientific work or are candidate genes for their molecular or functional role (N = 1037 genes). The second approach was to examine the overlap with probes previously related to fetal brain development (Spiers et al., 2015, N = 16,543 probes) given that the study involves early neurocognitive development. To identify the location of each probe with respect to gene regions, EWAS discovery-significant probes were annotated based on UCSC Genome Browser, build hg19 (Karolchik et al., 2004). Chi-square statistics were used to test whether the percentage of discovery-significant genes/probes which could be found in the SFARI and Spiers et al., 2015 datasets, respectively, was higher than it would have been by chance (4.1% of all interrogated probes are located in SFARI genes and 7.1% of them have been associated with fetal brain development in Spiers et al., 2015).

2.3.3. Analysis 3: Co-methylation network analysis

Identification of aberrant DNAm profiles associated with ASD might reflect either the cause or the consequence of atypical development. To investigate a potential causal link, we explored whether differentially methylated probes identified in our sample reflected known ASD-associated biological processes coupled to underlying genetic variations. Recent research has shown that genetic variation influences DNAm levels in a developmental-specific manner (mQTLs, Gaunt et al., 2016), and a relationship between known ASD-associated genetic variants and DNAm levels is observable at birth in peripheral tissue (Hannon et al., 2018). We therefore used available bioinformatic resources to investigate whether biological pathways through which genetic factors contribute to early epigenetic signatures associated with our phenotypes of interest.

First, we ran a Weighted Gene Co-methylation Network Analysis (WGCNA, Langfelder & Horvath, 2008) to identify networks, also called modules, of probes whose DNAm levels were consistently correlated across infants at T1. The module eigenvalues, representing the methylation profiles of the modules, were correlated with the five phenotypes of interest: categorical outcome of ASD and atypical development; dimensional outcome (adaptive skills at T4) and the two candidate intermediate phenotypes (face looking and Nc to face at T1). False Discovery Rate (FDR) method was used to adjust p-values controlling for simultaneous testing of the association of multiple modules with the phenotypes of interest. Module-trait relationships were considered significant at an FDR-adjusted p-value < 0.05.

Subsequently, to examine early genetically driven effects on DNAm levels in the significant module, we used the mQTL database published by Gaunt and colleagues (2016, http://mqtldb.org) to extract a list of variations in nucleotide sequence, or Single Nucleotide Polymorphisms (SNPs), known to influence DNAm of the probes of interest (i.e., mQTLs) during pregnancy and at birth. We then looked for biological pathways in which the mQTLs might be involved. To do this, we used eSNPO (http://bioinfo.hrbmu.edu.cn/esnpo/), which utilizes available resources to provide functional enrichment pathways for sets of SNPs (Li et al., 2016), indicating what biological function each set of SNPs play all together based on their gene expression profiles. Pathways for the two sets of SNPs (mQTLs at birth or during pregnancy) were obtained. FDR correction was applied to control for multiple testing of all possible pathways associated with the input SNPs. The entire procedure is described in detail in Appendix F and in Fig. 6.

2.3.4. Analysis 4: longitudinal analysis

Tracking the stability and change of association between DNAm and behavioral traits in the first two years of life is important to understand which probes will be potentially involved in developmental mechanisms as they show plasticity (i.e. change across

Table 3
Summary of the results of the EWAS for the five phenotypes of interest: ASD, atypical development and adaptive skills at T4, face looking and Nc to face at T1.

Phenotype	Participant N with DNAm data at T1 for EWAS analysis	Normality	λ	Discovery sign. probes	Effect Size >0.01	Max Effect Size	% in SFARI (4.1% by chance)	% fetal brain dev. (7.1% by chance)
ASD	41 no-ASD vs 10 ASD	/	1	32	32	0.22	6.25% (SND1, CACNA2D1)	6.25% (DDR2, CFLAR)
Atyp. Dev.	28 Typ. Dev. vs 23 Atyp. Dev.	/	1.12	33	33	0.07	9% (CHD1, PLXNA3, GIGYF1)	6% (cg01257697, ECE2)
Adaptive Skills	49	W = 0.98, p=0.4	1.01	23	22	0.04	8.97% (PGLYRP2, CDKL5)	8.97% (FOXK1, ADCY10)
Face looking	42	W = 0.91, p = 0.03	1	4	4	0.07	0	25 % (PLEKHH3)
Nc to face	26	W = 0.93, p=0.07	1	9	5	0.08	11.1% (GRID1)	0

Normality: normality of the distribution assessed with Shapiro-Wilk test for the three continuous phenotypes; λ: genomic inflation for each EWAS due to population stratification, where 1 represents no-inflation. Discovery sign. probes: number of probes which were significant at a discovery threshold of $p < 5 \times 10^{-5}$; Effect Size > 0.01 : number of probes with effect size higher than 1%, where the effect size corresponds to the beta estimates of the effect size resulting from the EWAS multiple regression; Max Effect Size: maximum absolute effect size observed among the discovery-significant probes; % in SFARI: percentage of discovery-significant probes in genes enriched for ASD according to the SFARI dataset (Banerjee-Basu & Packer, 2010, last release Nov. 2018); % fetal brain dev: percentage of discovery-significant probes which were also found to be implicated in fetal brain development by Spiers et al. (2015).

development) in relation to the phenotype. In this analysis we aimed to explore if the epigenetic signals suggestively associated with later outcome have a time-specific effect on dimensional traits. We built on the EWAS results to conduct Analysis 4 on two sets of candidate probes: those associated with ASD (N = 32) and those associated with atypical development (N = 33) at a discovery threshold of $p < 5 \times 10^{-5}$. FDR method was used to correct p-values for multiple testing.

This analysis included the subset of infants with repeated measurements for both DNAm and phenotypic measure (Table G.1 presents the number of participants for each visit). Two phenotypic measures were available at multiple visits and were therefore used as dependent variables: VABS composite score as a dimensional measure of adaptive skills (Estes et al., 2015; Robinson et al., 2016, N = 33) and face looking as a behavioral correlate of social visual attention engagement (Colombo & Cheatham, 2006; Hendry et al., 2018; Jones et al., 2016, N = 27). The relation between DNAm level of a probe and continuous phenotypic measure was modelled over time using multilevel mixed-effects linear models (Laird & Ware, 2007). A random intercept model was used to account for individual differences in DNAm at the first visit, thus controlling for possible effects of ancestry differences and other between-individual sources of variability. Multilevel mixed-effects models handle missing time-points in the measurements and have been used previously in longitudinal research testing the relationship between DNAm and continuous phenotype measures (Simpkin et al., 2015). Parameters estimated by the models used in this analysis allowed us to explore changes in the direction of the relation between DNAm and phenotype across toddlerhood, possibly reflecting epigenetic changes going along with symptom emergence.

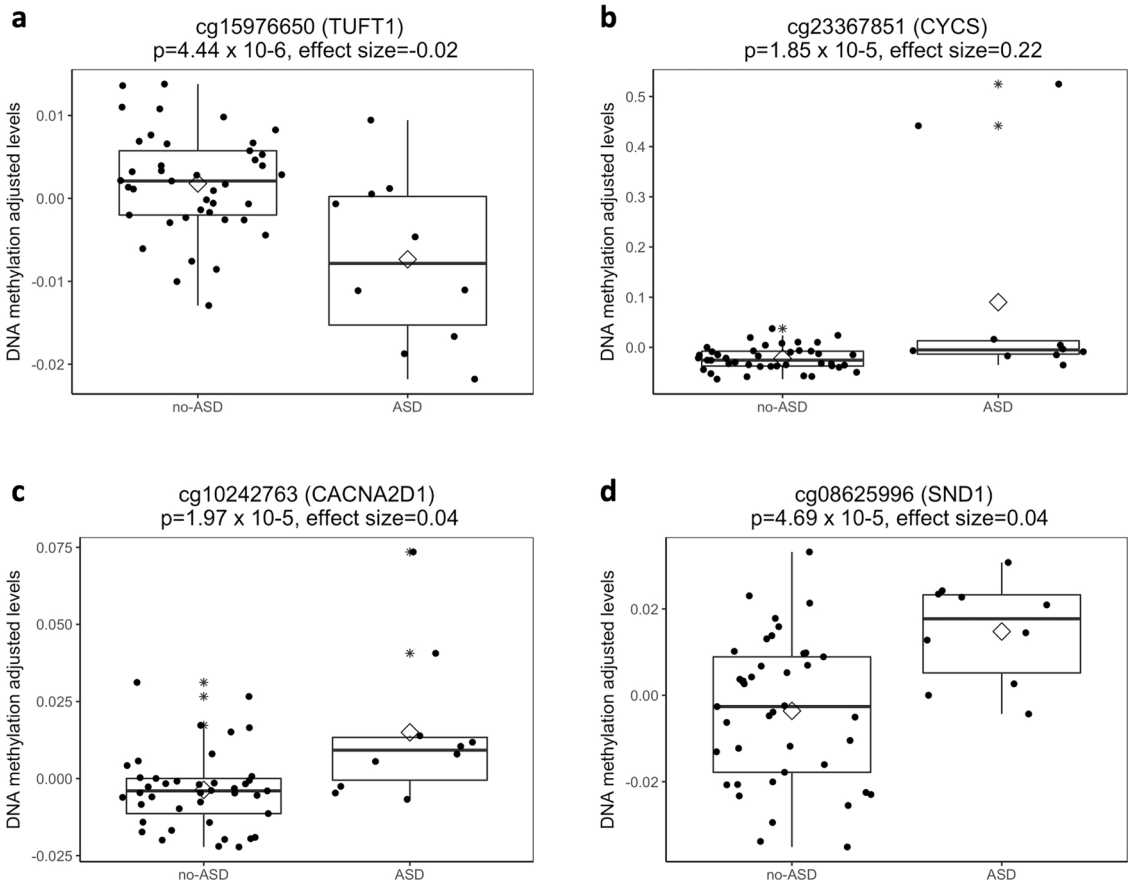

Fig. 4. Boxplots of the pattern of association with four of the top-ranked probes associated with ASD. **a**) The most significant signal (probe cg15976650 in the CpG island of gene TUFT1), **b**) the signal with the highest effect size (probe cg23367851 in the shore of gene CYCS), **c**) probe cg10242763 located in gene CACNA2D1 and **d**) probe cg08625996 located in gene SND1, which have both been previously associated with ASD. DNA methylation levels are adjusted for ancestry principal components and batch effect. In all figures, the lower and upper hinges of the boxplots correspond to the first and third quartiles (the 25th and 75th percentiles). Individual data are represented by points. The upper and lower whiskers extend from the hinges to the largest and smallest values, respectively, no further than 1.5 x inter-quartile range. Data beyond the end of the whiskers are represented by stars. Horizontal lines within boxplots indicate median values, while rhombi represent mean values.

3. Results

3.1. Analysis 1: global DNA methylation levels

3.1.1. Categorical outcome

There was no difference between children with and without ASD in terms of global level of methylation (β = 0.0005, standard error = 0.0007, p = 0.428). Similarly, there was no significant difference between children with typical and atypical development (β = 0.0001, s.e. = 0.0005, p = 0.86). These analyses remained non-significant after correcting for batch, age (in months) at the time of DNA collection (visit) and prenatal maternal smoke and alcohol intake (Tables C.1 and C.2).

Global methylation levels increased with time ($\chi^2(6) = 23.46$, $p < 0.0001$), as they were significantly higher at 25 months than at 8 months (β = 0.002, s.e. = 0.004, $p < 0.0001$). This result was expected based on a previous study reporting general increase in DNAm in buccal cells the first two years of life (Martino et al., 2013). No difference was observed between 8 and 15 months (β = 0.0006, s.e. = 0.0005, p = 0.212). The slope of the change in global methylation between 8 and 25 months did not differ between children with and without ASD ($\chi^2(9) = 0.588$, p = 0.745), nor when comparing children with typical and atypical developmental outcome at 3 years of age ($\chi^2(9) = 0.302$, p = 0.86).

3.1.2. Associations with dimensional phenotypes

Global methylation level was not associated with adaptive behavior at three years (β = -0.0001, s.e. = 0.0003, p = 0.66). Similarly, no variation in global methylation was associated with face looking (β = -0.0004, s.e. = 0.0003, p = 0.27) or Nc to face at 8 months (β = 0.0001, s.e. = 0.0003, p = 0.76). Analyses controlling for possible covariates led to the same results and can be found in Tables C.3-C.5.

3.2. Analysis 2: epigenome-wide association analyses (EWAS)

EWAS analyses were carried out for the three continuous traits in the same way as for categorical outcomes. Results are summarized in Table 3.

3.2.1. Categorical outcome

3.2.1.1. ASD. EWAS was first performed for prediction of categorical ASD outcome comparing no-ASD (N = 41) vs ASD (N = 10) groups for their DNAm levels at T1. No probe was statistically significant at the empirical epigenome-wide significant threshold after controlling for multiple testing ($p < 2.4 \times 10^{-7}$). Figures D.1a and D.1b illustrate the results for the ASD analysis with plots that are commonly used in EWAS to display significant probes. At a discovery threshold of $p < 5 \times 10^{-5}$, 32 probes were significantly associated with ASD, of which 2 probes are in genes that have been previously linked with ASD, according to the SFARI dataset (6.25% of the discovery probes, not significantly higher than that expected by chance, $\chi^2(1) = 0.373$, p = 0.27). Table D.1 reports discovery-level significant probes. Fig. 4 illustrates patterns of group differences for four selected probes: TUFT1, the gene associated with the probe with the most significant (i.e. lowest) p-value; CYCS, the gene associated with the probe with the largest effect size;

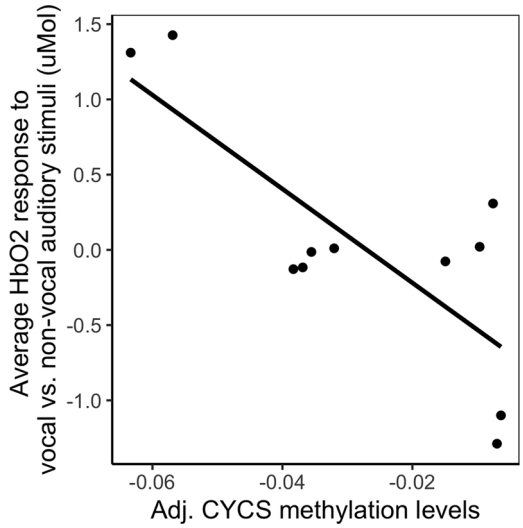

Fig. 5. Scatter plot representing the relationship between DNA methylation of the cg23367851 probe in the CYCS gene at 8 months (x-axis) and difference in oxygenated haemoglobin concentration (HbO2) in the auditory social (vocal) vs non-social (non-vocal) conditions at 5 months (y-axis) for 11 infants who provided both NIRS and DNA methylation data. The black line represents the linear relationship between the two variables. CYCS methylation levels have been adjusted for the effect of ancestry principal components and batch effect.

and the two probes previously associated with ASD: SND1, whose mutation leads to presynaptic myasthenic congenital syndrome 7 (GeneCards, Stelzer et al., 2016) and whose common (Holt et al., 2010) and rare (Iossifov et al., 2012) variations have been previously associated with ASD; and CACNA2D1, which is involved in voltage-gated calcium channels (GeneCards, Stelzer et al., 2016). Missense variation (Iossifov et al., 2014) or deletions (Vergult et al., 2015) in this gene have been found in a patient with ASD and 3 patients with intellectual disability and epilepsy, respectively.

Post-hoc exploratory analyses on the signal with the highest effect size (cg23367851 probe associated with the promoter region of the CYCS gene) were conducted as a proof-of-principle of how phenotypic information collected longitudinally in an infant-sibling design could be leveraged to further investigate possible developmental mechanisms involving epigenetic variation. The CYCS gene plays a crucial role in the electron-transport chain in the mitochondria (GeneCards, Stelzer et al., 2016). Thus, we investigated whether DNAm-driven atypicalities in the electron-transport chain were associated with available hemodynamic data, collected in a functional near infra-red spectroscopy (fNIRS) study from the same cohort at age 4–6 months (details on the study paradigm can be found in Appendix E and in Lloyd-Fox et al., 2013, 2018). We selected this dataset because fNIRS can be used to measure task-related oxygen concentration changes in haemoglobin within the brain, which indirectly reflect mitochondrion-mediated energy metabolism (Siddiqui et al., 2017). Importantly, Lloyd-Fox et al. (2013) reported attenuated socially selective hemodynamic responses – during the presentation of social (human vocal) and non-social (toys, water, bells) sounds – in infants with an older sibling with ASD. In a subset of these children who provided both DNAm and NIRS data (N = 12), we found that DNAm level in the CYCS probe at 8 months (adjusted for principal components and batch effect as in the ASD EWAS conducted in Analysis 2) was negatively correlated with cortical oxygenation changes in response to the vocal vs non-vocal conditions over the temporal lobes (rho = -0.40, t(10) = -1.39, p = 0.097, Figure E.2). This relation held when excluding the data point identified as an outlier for methylation levels (rho = -0.78, t(9) = -3.7, p = 0.002, Fig. 5).

3.2.1.2. Atypical development. The most significant signal for the EWAS comparing infants who later underwent typical (N = 28) vs atypical development (N = 23) was probe cg21973914, located in the shore of gene F10 ($p = 3.34 \times 10^{-7}$, non-significant based on the established significance threshold of 2.4×10^{-7} for EWAS studies, Saffari et al., 2018). Gene F10's primary role is to encode the vitamin K-dependent coagulation factor X of the blood coagulation cascade, and is involved in calcium ion binding and phospholipid binding (GeneCards, Stelzer et al., 2016). The probe with the highest effect size (cg06963664), which was 7.3% more methylated in infants who show signs of atypical development at 3 years, is located in gene PLXNA3. Deleterious variants of this gene have been found in ASD probands from the Simons Simplex Collection, Autism Sequencing Consortium and a Chinese population (Guo et al., 2017; Krumm et al., 2015).

Overall, 33 probes were associated with atypical development at a discovery p-value threshold $< 5 \times 10^{-5}$ (Table D.2, Figure D.2). Three of these probes (9%, slightly higher than that expected by chance: $\chi^2(1) = 1.83$, $p = 0.088$) have been found to be associated with genes within the SFARI dataset. One of these has been previously described as the probe with the highest effect size (cg06963664, gene PLXNA3). Additionally, probe cg21082921 is located in the promoter region of gene CHD1. Missense and *de novo* mutation of this gene have been found in individuals with ASD (Iossifov et al., 2014; Neale et al., 2012) and other neurodevelopmental disabilities such as speech apraxia and developmental delay with and without seizures (Pilarowski et al., 2018). Last, probe cg05922723 is located in the GIGYF1 gene, whose *de novo* mutation has been recognized in ASD probands by three different studies (De Rubeis et al., 2014; Iossifov et al., 2014; Krumm, O'Roak, Shendure, & Eichler, 2014).

3.2.2. Association with dimensional phenotypes

When testing the association between DNAm levels of single probes and dimensional outcome (VABS at 3 years), face looking and Nc to face, no probes reached corrected significance in any analysis. However, all analyses revealed a number of probes that were significant at a discovery threshold of $p < 5 \times 10^{-5}$. Of note, for all measures except face looking, the percentage of probes identified in the vicinity of high-confidence ASD genes (SFARI datadase) was higher than that expected by chance (4.1%), though not significantly. Interestingly, the neural metric (Nc to face) resulted in the highest percentage of probes overlapping with the SFARI database (11%, $\chi^2(1) = 0.97$, $p = 0.16$), consistent with the possibility that imaging measures provide stronger signals for genetic analysis (Nikolova & Hariri, 2015; Wiers, 2012). Of note, signals associated with face looking were not significant in control analyses

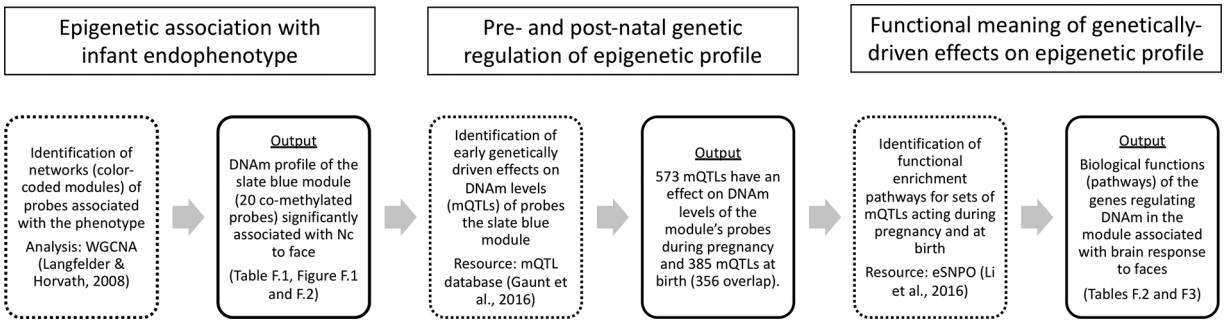

Fig. 6. Diagram illustrating the analyses (dotted boxes) and outputs (solid boxes) of the network analysis (Weighted Gene Co-methylation Network Analysis, WGCNA), methylation quantitative trait loci (mQTLs) analysis and functional enrichment pathway analysis.

using robust regression to account for non-normality of the distribution (Table D.4). Appendix D reports all results (Tables D.3-D.5, Figures D.3-D.5).

3.3. Analysis 3: Co-methylation network analysis

Differently from Analysis 2 (EWAS), where ~400,000 tests were performed simultaneously to test associations at a single-probe level, Analysis 3 (network analysis) tested the association between methylation profiles in 23 networks or modules of co-methylated probes identified in the 8-month-old DNAm data and the five phenotypes of interest. We found that one module (to which the color *slate blue* was automatically assigned; hereafter, the slate blue module) was significantly associated with Nc to face with a p-value = 4×10^{-7} (significant after controlling for multiple testing, FDR < 0.0001). No other association reached corrected significance (see Figure F.1). Probes in the significant module showed higher DNAm in relation to enhanced neural activation in response to noise over face (Figure F.2). This result was driven by a case that showed extreme, though plausible (see Figure F.3) values of Nc amplitude difference between conditions, with greater Nc in response to the noise than to the face stimulus. This participant also had an extreme methylation profile (i.e., module eigenvalue) for the slate blue module, though he did not show atypical DNAm genome-wide (see Figure F.4). This result must be treated with caution given that it is significant only in the presence of one influential case. However, as there was no *a priori* reason to exclude the participant from the analysis, we explored this finding with additional analyses aimed to understand whether it could reflect plausible biological mechanisms that, if altered due to abnormal methylation, could be related to neural atypicalities.

To identify genes that exert a direct influence over DNAm levels in the co-methylated probes of the slate blue module (listed in Table F.1), we identified mQTLs during pregnancy and at birth (Gaunt et al., 2016). We found 573 mQTLs acting during pregnancy and 385 mQTLs at birth. 356 probes were shared signals, indicating that they influenced infants' DNAm levels at both developmental points. FDR-significant pathways indicating the biological function of the set of SNPs regulating methylation levels of the slate blue module are reported in Tables F.2 and F.3. Fig. 6 summarizes the results of this set of analyses.

3.4. Analysis 4: longitudinal analyses

One potentially valuable feature of longitudinal analysis is the ability to examine probes that changed the direction of their association with key phenotypes over time, providing insights into the timing of the relationship between DNAm and the phenotypic trait. Longitudinal analyses focused on the probes that Analysis 2 identified for suggestive significance by the EWAS for ASD (N = 32, Table D.1) and atypical development (N = 33, Table D.2) in Analysis 2. For probes associated with ASD, DNAm levels of one probe (cg21348771) showed an age-specific change of association with face looking between T1 and T2 (β = -0.008, s.e. = 0.003, p = 0.003, FDR = 0.08, Fig. 7). Interestingly, this probe is in the CpG island of gene GFOD1, which has been associated with ADHD (GeneCards, Stelzer et al., 2016). No other signals were significant after controlling for multiple testing (all FDRs > 0.1, Tables G.2 and G.3). When analyzing the probes associated with atypical development, DNAm levels of one probe (cg04089240 in gene TRPM5, Fig. 8) showed significant change in direction of the association between DNAm levels and adaptive skills at 15 months (β = -0.003,

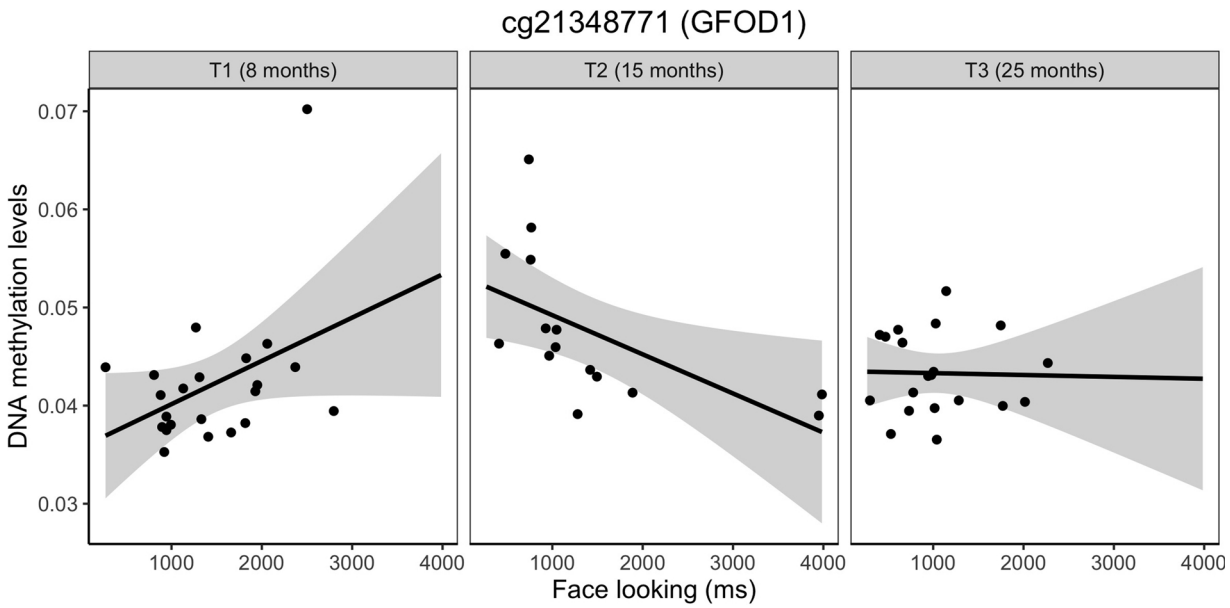

Fig. 7. Scatterplots representing the relationship between face looking (x-axis) and DNA methylation levels of the cg21348771 probe located in the promoter region of the GFOD1 gene (y-axis) for the three visits. Regression lines are depicted, with shaded grey areas representing standard errors.

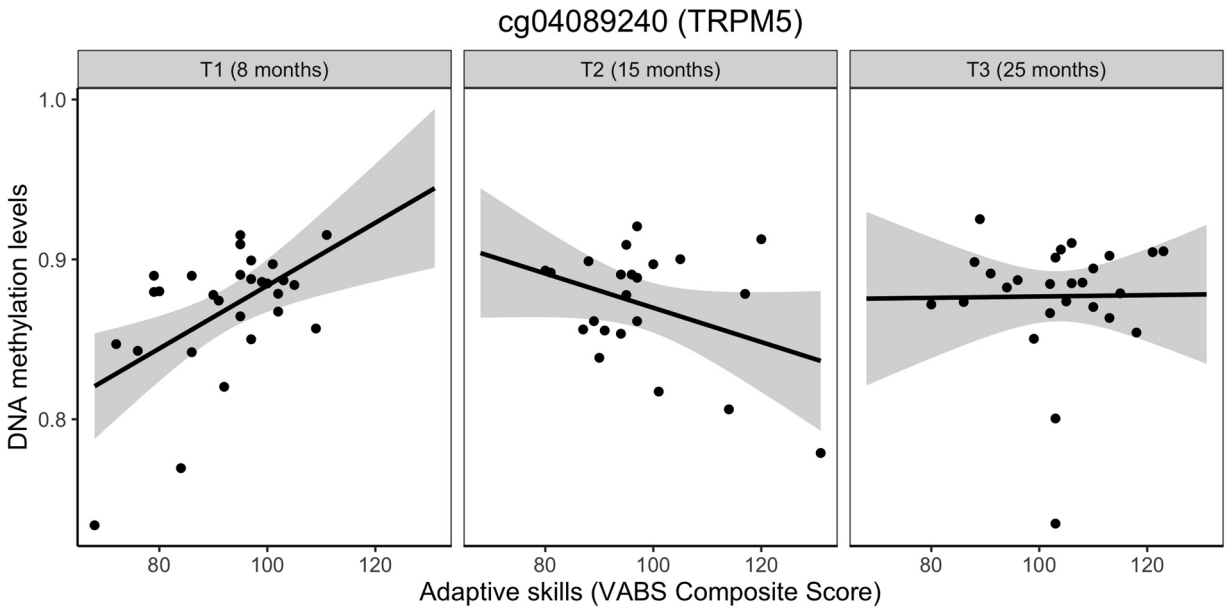

Fig. 8. Scatterplots representing the relationship between parent-reported adaptive skills measured with the Vineland Adaptive Behavior Scales (VABS) composite score (x-axis) and DNA methylation levels of the cg04089240 probe located in the shore of gene TRPM5 (y-axis) for the three visits. Regression lines are depicted, with shaded grey areas representing standard errors.

s.e. = 0.0009, p = 0.002, FDR = 0.03). Of note, *de novo* loss of function mutation in the TRPM5 gene has been found in one child with ASD (Neale et al., 2012). No other probe showed a significant association with adaptive skills (Table G.4) or face looking (Table G.5) after correcting p-values for multiple testing.

4. Discussion

Incorporating epigenetic analysis into prospective longitudinal studies of infants with later ASD could help delineate the mechanisms that underpin ASD emergence. First, we did not find differences in whole-genome DNAm associated with later ASD, in line with some other reports in larger samples (Hannon et al., 2018). Second, we showed that although EWAS did not produce epigenome-wide significant hits in a sample of this size (as expected), we identified altered DNAm profiles in a number of probes that are located in close proximity to high-confidence ASD genes at our discovery-level threshold. Third, we showed that a network-based analysis coupled with neuroimaging phenotypes did produce significant associations after correction for multiple comparisons. Further, biological pathways identified through examination of mQTLs linked to probes in the significant network were relevant to brain development. This supports other evidence that neuroimaging metrics may be highly promising intermediate phenotypes appropriate for analysis of genetic effects on neurodevelopmental disorders (Geschwind & Konopka, 2009; Nikolova & Hariri, 2015; Wiers, 2012). Finally, our longitudinal analysis revealed dynamic changes in the relation between DNAm of candidate probes and ASD-relevant behaviors between the first and second years of life. Since this is the time at which behavioral markers of ASD typically emerge, such longitudinal analyses may hold promise for understanding underpinning mechanisms of change (Gliga, Jones, Bedford, Charman, & Johnson, 2014). Together, our analyses indicate that epigenetic analysis may be a promising strategy for the integration of neurobiology and neurocognitive insights into ASD emergence.

4.1. Developmental changes in whole-genome methylation

As expected based on previous studies of buccal swab samples collected during infancy (Martino et al., 2013), an increase of global methylation was observed in the genome for the entire sample between 8 and 25 months. This confirms the importance of using narrow age ranges in DNAm studies, especially when investigating mechanisms of development (Michels et al., 2013). However, we did not find significant differences in global methylation between infants with and without later ASD or with and without later atypical development. Effect sizes were < 0.0005, suggesting that if any significant difference between groups existed it would have only been identified with extremely large samples. In accordance with other studies (see Dall'Aglio et al., 2018), we did not replicate previous claims about a general disruption of DNAm in individuals with ASD (James, Shpyleva, Melnyk, Pavliv, & Pogribny, 2013; Tsang et al., 2016). Of note, the 450 K array used in the present and other studies, interrogates around 450,000 individual CpG sites genome-wide. These are estimated to be less than 2% of all possible methylated cytosines (Dedeurwaerder, Defrance, & Calonne, 2011), and differences between studies might be due to discrepancies in the interrogated sites. Importantly, previous studies reporting differences in global methylation in ASD have studied children over 2 years (James et al., 2013; Tsang et al., 2016). It is

possible that such differences emerge later in life, as environmental influences on the epigenome become more consistent (Fraga et al., 2005). However, our data gave no indications of different developmental pathways for children with typical and atypical development up to two years of age.

4.2. Exploring associations with phenotypes at a methylation site-level

No individual probes reached a strict epigenome-wide significance threshold in the EWAS analyses. This was as expected since the sample size was relatively small (though comparable to many previous studies on DNAm and ASD; see Dall'Aglio et al., 2018 for a review). However, most effect sizes resulting from our categorical analyses are comparable with previous case-control DNAm studies, with the most highly associated probes having an effect size (β coefficients of regression) ranging from 0.005 to 0.1, indicating a difference of 0.5–10% in DNAm level between individuals with and without ASD (Dall'Aglio et al., 2018). For example, the strongest signal in our sample was a probe with 2% average difference between children with and without ASD. Previous power computations for EWAS signals have suggested sample sizes of 110–120 cases and controls are necessary for epigenome-wide significance of signals for this range of effect sizes (Saffari et al., 2018; Tsai & Bell, 2015). Importantly, these simulation analyses have indicated that more powerful signals are found when variance in DNAm is smaller, and thus more homogenous phenotypic samples may give larger effect sizes (Tsai & Bell, 2015). Accordingly, characteristics in the infant-sibling designs might lead to higher power for EWAS analyses. For example, our study examined DNAm levels of males only, ranging between 7 and 10 months of age, all receiving accurate diagnostic assessment by the same clinical team at the same age. Results of an exploratory power analysis conducted on the observed results confirm that EWAS in infant-sibling designs may require substantially smaller sample sizes than previously estimated (see Appendix D and Table D.6), although these results should be considered with caution given that estimates of the examined effect sizes derived from underpowered analyses and might not be reliable.

As in similar research, we also examined the probes that were associated to each phenotype at a discovery threshold to further probe the potential promise of this approach for the study of early neurodevelopmental trajectories. We identified probes in high confidence ASD genes according to the SFARI Gene resource (Banerjee-Basu & Packer, 2010), and probes previously related to fetal brain development (Spiers et al., 2015). With exception of the face looking metric, the proportion of probes in the discovery set found in the SFARI dataset was higher, although not significantly, than that expected by chance. Associations with the EEG phenotype produced the strongest enrichment for high-confidence ASD genes from SFARI (11% vs 4.1% by chance), due to a single probe located in the CpG island of gene GRID1). GRID1 is the Glutamate D1 Receptor gene (GRID1). In this analysis, an infant with particularly enhanced amplitude of the neural response to noise vs face had 10% higher DNAm levels in this probe. This association is significant also using robust regression (Table D.5), suggesting that this relationship was shared with others in the sample. Of course, this result requires replication. However, GRID1 deletions have been strongly linked to ASD (Glessner et al., 2009) and related conditions like schizophrenia (e.g., Guo et al., 2017; Nenadic et al., 2012); one study looking at rare variations of large effect across the entire genome (copy number variants, CNVs) showed that mutations in this gene were over-represented in ASD patients ($N = 14$) compared to controls ($N = 3$) (odds ratio = 5.412, $p = 0.0031$) (Glessner et al., 2009). Deletion of GRID1 in mouse models has been associated with deficits in emotional and social behaviors (Yadav et al., 2012). Since the imbalance between glutamatergic and GABAergic neurons in the brain is a leading hypothesis for mechanisms underpinning core ASD symptomatology and sensory atypicalities, this may indicate promising avenues for future work (Robertson & Baron-Cohen, 2017; Rubenstein & Merzenich, 2003). Another interesting signal emerging from the EWAS testing association with the neural phenotype is that one of the discovery-significant probes is located in the CpG island of gene FAM167 (see Appendix D and Table D.5), which is involved in auto-immune disease (Mentlein et al., 2018). This signal replicates an epigenome-wide significant finding of a recent EWAS with ~1200 newborns with and without later ASD, which reported an association between DNAm in this gene and ASD polygenic score (Hannon et al., 2018).

With regards to probes previously implicated in fetal brain development, the proportion of shared signals was not higher than that expected by chance for all phenotypes except face looking (Table 3). In fact, DNAm levels of the top probe associated with this behavioral phenotype significantly increase during fetal brain development (Spiers et al., 2015). Further analyses revealed that this result is not robust to influential cases (see Appendix D and Table D.4 showing results of robust regression). This might indicate that possible effects of epigenetic variations are observed only in relation to values at the extreme ends of the phenotypic distribution, although no reliable conclusions can be drawn from this exploratory analysis on a very small sample (Sonuga-Barke & Fearon et al., 2018). Associations with face looking produced no enrichment for genes previously associated with ASD, perhaps reducing the likelihood that this is a meaningful intermediate phenotype for ASD. Indeed, its association with later ASD before the first year of age in previous studies has been mixed (Hendry et al., 2018; Jones et al., 2016), with stronger potential links to executive functioning (Hendry et al., 2018).

Enrichment for ASD genes was greater for the comparison between typical and atypical outcome (around 9%) than for ASD vs no ASD (6%). This may be due to the fact that groups were more balanced in terms of participant numbers in the former analysis (28 vs 23) than in the latter (41 vs 10). However, the ASD vs no ASD contrast did reveal an intriguingly strong signal for a particular probe (cg23367851). This probe was very highly methylated in two individuals (see Fig. 4.b), although non-parametric analysis still gave nominally significant results ($W = 99$, $p = 0.011$). To check that this result was not an artifact of a genetic variant being present in the probe location, we looked for the presence of SNPs in the probe sequence and flanking region. No SNPs lies within the probe binding region, although one SNP (rs58704610) lies at 10 base pairs (i.e. 10 nucleotides) distance from it. Given that genotype data was not available for the infants it is not possible to exclude the possibility that the DNAm signal is confounded by local DNA sequence variation. However, the pattern of DNAm levels did not correspond to three qualitative groups of methylation levels that would have indicated an effect of genotype (high and low DNAm levels representing homozygosity for each of the two alleles

respectively, and mid-level between the other two groups representing heterozygosity).

Probe cg23367851 is associated with the promoter region of the CYCS gene, which is a central component of the electron transport chain in mitochondria (GeneCards, Stelzer et al., 2016). As mitochondrial dysfunctions have been found previously in individuals with ASD (Siddiqui, Elwell, & Johnson, 2016), we leveraged the availability of a range of phenotypic measures collected in the first three years for the infant siblings to conduct post-hoc exploratory verification of the plausibility of such mechanism. Examination of phenotypic data from these two individuals revealed that parents reported lower motor skills at T1 compared to the other adaptive domains (Figure E.1a). Data from standardized behavioral assessments confirmed that at T1 these infants showed lower gross and fine motor skills, respectively, with respect to their peers (Figures E.1b and E.1c), in line with findings reporting motor delay and fatigability in individuals with mitochondrial dysfunction and ASD (Weissman et al., 2008). Longitudinal data available revealed that this effect was not persistent, suggesting that, if our result reflects a true pathophysiological mechanism, this represents a transient effect, rather than a stable long-term pathological dysfunction. Disruptions in the electron transport chain might lead to early inefficient oxygen metabolism in the brain, having downstream effects on the hemodynamic response to social stimuli which can be observed with functional Near Infrared Spectroscopy (fNIRS, Siddiqui et al., 2017). Significant correlations between CYCS methylation levels and brain oxygenation responses measured at 4 to 6 months with fNIRS (Lloyd-Fox et al., 2013) seem to support the hypothesis that increased DNAm in this gene might be linked with reduced responses to social vs non-social auditory stimuli (see Fig. 5 and E.2), which is a pattern observed in infants who go on to develop ASD (Lloyd-Fox et al., 2018). Of note, this probe is promoter-associated and not tissue-specific (Table D.1), thus a systemic epigenetic effect is not implausible. Moreover, Goldenthal et al. (2015) reported extensive mitochondrial enzyme deficiencies in the activity of the respiratory complex in buccal swabs of 39 (42%) children with ASD compared with controls. A deletion in the region of the genome where the CYCS gene is located was reported in one individual with a developmental delay in the Deciphering Developmental Disorders study (Fitzgerald et al., 2014), and expression levels have been reported to be downregulated in a mouse model of schizophrenia (Sowers et al., 2019). Given increasing drives to stratify participants with ASD into meaningful subgroups (Jeste & Geschwind, 2014), genetic and epigenetic signals related to mitochondrial function may be a meaningful area for future investigation. Such findings can generate hypotheses for biologically possible mechanisms through which epigenetic markers could reflect or contribute to the gradual emergence of more severe phenotypes.

4.3. Functional networks and genetic influences

Network-based analysis allowed us to take into account the interrelation between different epigenetic probes when associating their DNAm levels with particular phenotypes. This approach may thus increase power to detect effects in small samples. Indeed, in the present study we found a significant association between one network of co-methylated probes, the slate blue model, and our putative neural intermediate phenotype between genetic risk and emerging ASD (Nc to face, see also Jones et al., 2016). Combined with the greater enrichment of ASD-relevant genes in the EWAS incorporating this marker, this result further suggests that neuroimaging measures may be more powerful in detecting relations to epigenetics than categorical outcome. Of note, these signals were driven by an influential case who showed extremely (though not implausible, see Appendix D and Figure F.3) reduced neural activation to social than to non-social stimuli. Although we recommend caution in the interpretation of the observed signals, it is possible that variation in DNAm profiles might have an impact on neurodevelopment only when extreme patterns emerge. The influential case in this analysis showed subthreshold ASD symptoms at T4, thus early intermediate phenotype patterns might indeed be indications of an atypical developmental trajectory.

Understanding the pathophysiological pathway underlying this association and ASD is complicated, especially when observing epigenetic signal in a tissue different from the brain. One approach to dealing with the issue of tissue specificity of the DNAm signal is to move away from the interpretation of the individual effects of significant probes and examine the functional effect of genetic variants that are known to exert upstream influences on DNAm levels of the probes in the significant network (mQTLs). These genetic variants are of course themselves present across tissue types. We can then examine the biological processes controlled by these genetic variants to see whether some are plausibly involved in neurodevelopment. This approach uses DNAm data as a signature readout of a combination of genetic and epigenetic differences that may impact development (Ciernia & Lasalle, 2016; Geschwind, 2008). We found that 385 genetic variants influence DNAm of the probes in the significant co-methylation network at birth and 573 SNPs during pregnancy. Importantly, functional enrichment analysis revealed that genes in our significant network are significantly involved in several brain-related pathways including neural cells migration and differentiation, synaptic transmission and composition of the myelin sheath (Tables F.2 and F.3). The identified mQTLs also have other biological functions that are potentially associated with ASD, such as: response to oxygen limitation (ischemia) which is one potential source of vulnerability for the autistic brain (Johnson, 2017), sleep, which is atypical in people with ASD (Won, Feldman, & Huffman, 2019) and inhibitory processes in the brain (calcineurin complex and calcineurin-NFAT signaling cascade, involved in activity-dependent GABAergic interneuron tuning, Bannai et al., 2009). Various pathways are also involved in mitochondrial generation function, intriguing in the light of the EWAS results for later ASD discussed above and given that mitochondrial dysfunctions are a proposed mechanism for ASD (Essa et al., 2013; Siddiqui et al., 2016). This approach thus produced encouraging signals that could guide future structural genomic work in similar cohorts, and could ultimately help us to generate hypotheses for potential biomarkers that are understood from biology to behavior.

4.4. Sensitive periods for epigenetic effects

Although no significant change in global methylation was observed between 8 and 15 months of age, our longitudinal analysis

revealed significant association between probe-specific methylation profiles and behavioral phenotypes across times. Altered brain development caused by genetic or environmental perturbations may affect DNAm in a system- and period-specific way, as the brain attempts to compensate for suboptimal conditions (Ciernia & Lasalle, 2016). Changes in DNAm occurring at specific loci in sensitive developmental windows are strictly regulated and highly important for neurodevelopment, as demonstrated in studies using fetal brain tissue (Numata et al., 2012). One intriguing feature of the longitudinal analysis was that significant changes in the association between DNAm and the behavioral phenotypes was observed between 8 and 15 months.

While very preliminary due to the small sample size, these results showing changes in the relationship between phenotype and DNAm during the second year are interesting because the time period between 8 and 15 months has been associated with emerging atypicality related to later ASD in both looking behavior (Elsabbagh, Gliga et al., 2012; Hendry et al., 2018) and behavioral measures including the adaptive skills measured with the VABS (Bussu et al., 2018; Estes et al., 2015) in previous work. In fact, the age of 12–15 months appears to be particularly significant. For example, differences in adaptive skills between EL infant siblings with and without ASD are significant at 12 but not 6 months (Estes et al., 2015); difficulties in attention shifting emerge between 6 and 12–14 months (Elsabbagh et al., 2013; Elsabbagh, Gliga et al., 2012); the behavioral composite symptom measure from the Autism Observation Scale for Infants (AOSI) is elevated in infants with later ASD at 14 but not 8 months (Gammer et al., 2015); and prediction of later ASD from a range of cognitive and adaptive skills is possible with reasonable sensitivity and specificity at 14 but not 8 months (Bussu et al., 2018). The present study indicates there may be a change in the degree of relation between the methylation of particular epigenetic probes and behavioral phenotypes within this same window. Future studies with larger samples could fruitfully focus on this developmental window.

Longitudinal modelling revealed that, among the probes whose methylation levels were highly associated with ASD at 3 years, one probe located in gene GFOD1 showed an age-specific change of association with face looking between T1 and T2. Encouragingly, DNAm levels of this probe in saliva and brain have been shown to be correlated within individuals (rho = 0.64, p = 0.04, Braun et al., 2019). Genome-wide association analyses revealed that GFOD1 is significantly associated with ADHD at the population level (Lasky-Su et al., 2008). Moreover, this gene was found to be associated with differences in IQ in a large sample of individuals with ASD (N = 1,590, Wang, Qin, Guo, Samuels, & Shugart, 2013). In the present study, decreased DNAm in this gene was associated with shorter looking time at 8 months but longer looking time at the face at 15 months. Across this period, look duration becomes increasingly under the influence of the executive attention system, which has been identified as a domain of vulnerability for a significant minority of children with ASD (Hendry et al., 2018). Thus, it is possible that variation in DNAm levels of the GFOD1 gene contributes to this developmental mechanism. In fact, ADHD and ASD often co-occur in the same individuals and shared genetic and familial influences have been observed on ADHD and ASD (Rommelse, Franke, Geurts, Hartman, & Buitelaar, 2010). Causal modelling investigating the co-occurrence of ADHD and ASD in family members of people with a neurodevelopmental condition revealed a strong causal link between impulsivity and social difficulties (Sokolova et al., 2017). This finding supports the idea that shared biological mechanisms might lead to the emergence of common behavioral traits at the end of the first year of life in children with elevated likelihood of neurodevelopmental disorders (Johnson, Gliga, Jones, & Charman, 2014).

5. Conclusions

In a proof-of-principle study, we explored the potential of imaging-behavior-epigenetics studies in longitudinal cohorts of individuals with a familial history of neurodevelopmental disorders. Given the exploratory nature of the present study, an extension, as well as replication, of these findings is warranted. Our sample size, although similar to many previous studies, is too small to provide high confidence in the results (Sonuga-Barke & Fearon et al., 2018). Undoubtedly, our EWAS analyses were underpowered to provide epigenome-wide significant results (see Table D.6). Larger samples can help to identify weaker signals but are counterbalanced by important limitations such as increasing heterogeneity and less precise and less informative phenotypes. We only included male infants to reduce heterogeneity, but obviously this means our results will not necessarily generalize to females. As noted, the Illumina 450 K array used in the current study covers only 2% of the CpG sites; future studies should consider obtaining better genome coverage using EPIC BeadChip (which interrogates > 850,000 CpG sites) or whole-genome bisulfite sequencing (Pidsley et al., 2016), though this has cost implications. In practice there is often a trade-off between sample size and density of coverage that can be achieved. Future researchers will need to carefully evaluate means to conduct well-powered and biologically informative studies.

Promising signals were observed at the single-probe level, replicating the contribution of several genes implicated in biological mechanisms for ASD. We showed how infant-siblings designs offer the possibility to investigate the effects of suggestive signals on experimental data collected as part of the longitudinal study, to test specific hypotheses of biological mechanisms underlying individual differences. Mapping networks of co-methylated probes associated with neural correlates of social attention showed several candidate pathways potentially involved in brain development. This is consistent with previous evidence that neuroimaging phenotypes may produce higher effect sizes because they are closer to the neurobiological level (Gliga et al., 2014; Nikolova & Hariri, 2015). Longitudinal modelling indicated promising signals associated with developmental changes in looking behavior during social attention and adaptive skills, although larger samples and complete datasets are needed to obtain generalizable results. We suggest that incorporation of epigenetic analysis into prospective longitudinal designs is a potentially fruitful approach to illuminating the neurobiological mechanisms underlying ASD.

Funding

The project was funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-

Curie [grant number 642996] and Wellcome Trust Institutional Strategic Support Fund to A.G. The Simons Foundation Autism Research Initiative (SFARI) Pilot Award [grant number 511504] supported E.J.H.J and E.M. in conceiving the general approach. Data collection was funded by MRC Programme [grant numbers G0701484 and MR/K021389/1], the BASIS funding consortium led by Autistica (www.basisnetwork.org), EU-AIMS (the Innovative Medicines Initiative joint undertaking grant agreement number 115300, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme [grant number FP7/2007-2013] and EFPIA companies' in-kind contribution) and AIMS-2-TRIALS (Innovative Medicines Initiative 2 Joint Undertaking under grant agreement number 777394. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI). Epigenetic data generation was funded by MRC Centenary Award to C.C.Y.W.

Ethics approval and consent to participate

The research was conducted in accordance with APA ethical standards in the treatment of the study sample. Ethical approval for this study was obtained from NHS Health Research Authority (06/MRE02/73). Written informed consent was obtained from all participants' parents.

CRediT authorship contribution statement

Anna Gui: Conceptualization, Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Emily J.H. Jones:** Conceptualization, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Chloe C.Y. Wong:** Conceptualization, Resources, Data curation, Writing - review & editing, Supervision, Funding acquisition. **Emma Meaburn:** Conceptualization, Writing - review & editing, Supervision. **Baocong Xia:** Data curation. **Greg Pasco:** Investigation, Data curation, Project administration. **Sarah Lloyd-Fox:** Investigation, Data curation, Writing - review & editing. **Tony Charman:** Writing - review & editing, Project administration, Funding acquisition. **Patrick Bolton:** Conceptualization, Writing - review & editing. **Mark H. Johnson:** Conceptualization, Supervision, Writing - review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgements

We heartfully thank all the families who took part in our research. We also thank Ayden Saffari and Pasco Fearon for helpful advice on the analyses, and Laurel Fish and Eleanor Braithwaite for precious comments on the final version of the manuscript. The BASIS Team members listed here in alphabetical order contributed to the design and/or data collection of the longitudinal BASIS study: Baron-Cohen S., Blasi A., Cheung C., Davies K., Elsabbagh M., Fernandes J., Gammer I., Ganea N., Gliga T., Guiraud J., Hendry A., Liersch U., Liew M., Maris H., O'Hara L., Pickles A., Ribeiro H., Salomone E., Taylor, C., Tucker L., Tye C., Wass S. In particular, we would like to acknowledge Charlotte Tye and Alexandra Hendry who pre-processed the EEG and eye-tracking data, respectively, used in the present study.

Appendix A. Supplementary data

Supplementary material (Appendices A-G) related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.infbeh.2019.101409.

References

American Psychiatric Association (2013).). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Authorhttps://doi.org/10.1176/appi.books.9780890425596.744053.

Amso, D., & Scerif, G. (2015). The attentive brain : insights from developmental cognitive neuroscience. *Nature Reviews, 16*(10), 606–619. https://doi.org/10.1038/nrn4025.

Andrews, S. V., Sheppard, B., Windham, G. C., Schieve, L. A., Schendel, D. E., Croen, L. A., ... Ladd-Acosta, C. (2018). Case-control meta-analysis of blood DNA methylation and autism spectrum disorder. *Molecular Autism, 9*(1), 1–11. https://doi.org/10.1186/s13229-018-0224-6.

Banerjee-Basu, S., & Packer, A. (2010). SFARI Gene: An evolving database for the autism research community. *Disease Models & Mechanisms, 3*(3), 133–135. https://doi.org/10.1242/dmm.005439 Retrieved from.

Bannai, H., Lévi, S., Schweizer, C., Inoue, T., Launey, T., Racine, V., ... Triller, A. (2009). Activity-dependent tuning of inhibitory neurotransmission based on GABAAR diffusion dynamics. *Neuron, 62*(5), 670–682. https://doi.org/10.1016/j.neuron.2009.04.023.

Barfield, R. T., Almli, L. M., Kilaru, V., Smith, A. K., Mercer, K. B., Duncan, R., ... Conneely, K. N. (2014). Accounting for population stratification in DNA methylation studies. *Genetic Epidemiology, 38*(3), 231–241. https://doi.org/10.1002/gepi.21789.Accounting.

Barfield, R. T., Kilaru, V., Smith, A. K., & Conneely, K. N. (2012). CpGassoc: An R function for analysis of DNA methylation microarray data. *Bioinformatics, 28*(9), 1280–1281. https://doi.org/10.1093/bioinformatics/bts124.

Barker, E. D., Walton, E., & Cecil, C. A. M. (2018). Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 59*(4), 303–322. https://doi.org/10.1111/jcpp.12782.

Braun, P. R., Han, S., Hing, B., Nagahama, Y., Gaul, L. N., Heinzman, J. T., ... Shinozaki, G. (2019). Genome-wide DNA methylation comparison between live human

brain and peripheral tissues within individuals. *Translational Psychiatry, 9*(1), 47. https://doi.org/10.1038/s41398-019-0376-y.
Bussu, G., Jones, E. J. H., Charman, T., Johnson, M. H., Buitelaar, J. K., Baron-Cohen, S., ... Volein, A. (2018). Prediction of autism at 3 years from behavioural and developmental measures in high-risk infants: A longitudinal cross-domain classifier analysis. *Journal of Autism and Developmental Disorders, 48*(7), 2418–2433. https://doi.org/10.1007/s10803-018-3509-x.
Cecil, C. A. M., Lysenko, L. J., Jaffee, S. R., Pingault, J. B., Smith, R. G., Relton, C. L., ... Barker, E. D. (2014). Environmental risk, Oxytocin Receptor Gene (OXTR) methylation and youth callous-unemotional traits: a 13-year longitudinal study. *Molecular Psychiatry, 19*(10), 1071–1077. https://doi.org/10.1038/mp.2014.95.
Ciernia, A. V., & Lasalle, J. M. (2016). The landscape of DNA methylation amid a perfect storm of autism aetiologies. *Nature Reviews Neuroscience, 17*(7), 411–423. https://doi.org/10.1038/nrn.2016.41.The.
Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior,* 283–322.
Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory, 92*(2), 225–234. https://doi.org/10.1016/j.nlm.2008.06.002.Infant.
Constantino, J. N., Kennon-McGill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., ... Jones, W. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature, 547*(7663), 340–344. https://doi.org/10.1038/nature22999.
Crane, L., Chester, J. W., Goddard, L., Henry, L. A., & Hill, E. (2016). Experiences of autism diagnosis: A survey of over 1000 parents in the United Kingdom. *Autism, 20*(2), 153–162. https://doi.org/10.1177/1362361315573636.
Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005.
Dall'Aglio, L., Muka, T., Cecil, C. A. M., Bramer, W. M., Verbiest, M. M. P. J., Nano, J., ... Tiemeier, H. (2018). The role of epigenetic modifications in neurodevelopmental disorders: A systematic review. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/J.NEUBIOREV.2018.07.011.
Dawson, G., Jones, E. J. H., Merkle, K., Venema, K., Lowy, R., Faja, S., ... Webb, S. J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(11), 1150–1159. https://doi.org/10.1016/j.jaac.2012.08.018.
De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature, 515*(7526), 209–215. https://doi.org/10.1038/nature13772.
Dedeurwaerder, S., Defrance, M., & Calonne, E. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics, 3*(6), 771–784.
Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry, 74*(3), 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030.
Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., & Johnson, M. H. (2012). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research, 251*, 147–154. https://doi.org/10.1016/j.bbr.2012.07.030.
Elsabbagh, M., & Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences, 14*(2), 81–87. https://doi.org/10.1016/j.tics.2009.12.005.
Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., ... Johnson, M. H. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology, 22*(4), 338–342. https://doi.org/10.1016/j.cub.2011.12.056.
Essa, M. M., Subash, S., Braidy, N., Al-Adawi, S., Lim, C. K., Manivasagam, T., & Guillemin, G. J. (2013). Role of NAD+, oxidative stress, and tryptophan metabolism in autism spectrum disorders. *International Journal of Tryptophan Research, 6*(SUPPL.1), 15–28. https://doi.org/10.4137/IJTR.S11355.
Essex, M. J., Boyce, W. T., Hertzman, C., Lam, L. L., & Armstrong, J. M. (2013). Epigenetic vestiges of early developmental adversity: Childhood stress exposure and DNA methylation in adolescence marilyn. *Child Development, 84*(1), 58–75. https://doi.org/10.1111/j.1467-8624.2011.01641.x.Epigenetic.
Estes, A. M., Zwaigenbaum, L., Gu, H., St. John, T., Paterson, S., Elison, J. T., ... Piven, J. (2015). Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *Journal of Neurodevelopmental Disorders, 7*(1), 1–10. https://doi.org/10.1186/s11689-015-9117-6.
Fitzgerald, T. W., The Deciphering Developmental Disorders, Gerety, S. S., Jones, W. D., van Kogelenberg, M., King, D. A., ... Hurles, M. E. (2014). Large-scale discovery of novel genetic causes of developmental disorders. *Nature, 519*, 223. https://doi.org/10.1038/nature14135 Retrieved from.
Flint, J., & Munafò, M. R. (2007). The endophenotype concept in psychiatric genetics. *Psychological Medicine, 37*(2), 163–180. https://doi.org/10.1017/S0033291706008750.
Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., ... Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences, 102*(30), 10604–10609. https://doi.org/10.1097/00005721-200605000-00016.
Frodl, T., Szyf, M., Carballedo, A., Ly, V., Dymov, S., Vaisheva, F., ... Booij, L. (2015). DNA methylation of the serotonin transporter gene (SLC6A4) is associated with brain function involved in processing emotional stimuli. *Journal of Psychiatry and Neuroscience, 40*(5), 296–305. https://doi.org/10.1503/jpn.140180.
Gammer, I., Bedford, R., Elsabbagh, M., Garwood, H., Pasco, G., Tucker, L., ... Pickles, A. (2015). Behavioural markers for autism in infancy: Scores on the Autism Observational Scale for infants in a prospective study of at-risk siblings. *Infant Behavior & Development, 38*, 107–115. https://doi.org/10.1016/j.infbeh.2014.12.017.
Gaunt, T. R., Shihab, H. A., Hemani, G., Min, J. L., Woodward, G., Lyttleton, O., ... Relton, C. L. (2016). Systematic identification of genetic influences on methylation across the human life course. *Genome Biology, 17*(61), 1–14. https://doi.org/10.1186/s13059-016-0926-z.
Geiman, T. M., & Muegge, K. (2010). DNA methylation in early development. *Molecular Reproduction and Development, 77*(2), 105–113. https://doi.org/10.1002/mrd.21118.
Geschwind, D. H. (2008). Autism: Many genes, common pathways? *Cell, 135*(3), 391–395. https://doi.org/10.1016/j.cell.2008.10.016.
Geschwind, D. H., & Konopka, G. (2009). Neuroscience in the era of functional genomics and systems biology. *Nature, 461*(7266), 908–915. https://doi.org/10.1038/nature08537.
Ginsberg, M. R., Rubin, R. A., Falcone, T., Ting, A. H., & Natowicz, M. R. (2012). Brain transcriptional and epigenetic associations with autism. *PloS One, 7*(9), e44736. https://doi.org/10.1371/journal.pone.0044736.
Glessner, J. T., Wang, K., Cai, G., Korvatska, O., Kim, C. E., Wood, S., ... Hakonarson, H. (2009). Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature, 459*(7246), 569–573. https://doi.org/10.1038/nature07953.
Gliga, T., Jones, E. J. H., Bedford, R., Charman, T., & Johnson, M. H. (2014). From early markers to neuro-developmental mechanisms of autism. *Developmental Review, 34*(3), 189–207. https://doi.org/10.1016/j.dr.2014.05.003.
Goldenthal, M. J., Damle, S., Sheth, S., Shah, N., Melvin, J., Jethva, R., ... Legido, A. (2015). Mitochondrial enzyme dysfunction in autism spectrum disorders; A novel biomarker revealed from buccal swab analysis. *Biomarkers in Medicine, 9*(10), 957–965. https://doi.org/10.2217/bmm.15.72.
Grayson, D. R., & Guidotti, A. (2015). Merging data from genetic and epigenetic approaches to better understand autistic spectrum disorder. *Epigenomics, 8*. https://doi.org/10.2217/epi.15.92 epi.15.92.
Gui, A., Bussu, G., Tye, C., Elsabbagh, M., Pasco, G., Charman, T., ... Jones, E. J. H. (2020). *Diminished engagement of attentive brain states to faces predicts later autism* (under review).
Guo, H., Peng, Y., Hu, Z., Li, Y., Xun, G., Ou, J., ... Xia, K. (2017). Genome-wide copy number variation analysis in a Chinese autism spectrum disorder cohort. *Scientific Reports, 7*(November 2016), 1–9. https://doi.org/10.1038/srep44155.
Guo, J. U., Ma, D. K., Mo, H., Ball, M. P., Jang, M., Bonaguidi, A., ... Zhang, K. (2012). Neuronal activity modifies DNA methylation landscape in the adult brain. *Nature Neuroscience, 14*(10), 1345–1351. https://doi.org/10.1038/nn.2900.Neuronal.
Halit, H., Csibra, G., Volein, A., & Johnson, M. H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 45*(7), 1228–1234. https://doi.org/10.1111/j.1469-7610.2004.00321.x.
Hannon, E., Schendel, D., Ladd-acosta, C., Grove, J., Hansen, C. S., Andrews, S. V., ... Reichenberg, A. (2018). Elevated polygenic burden for autism is associated with differential DNA methylation at birth. *Genome Medicine, 10*(19), 1–13. https://doi.org/10.1111/eva.12494.
Hass, J., Walton, E., Wright, C., Beyer, A., Scholz, M., Turner, J., ... Ehrlich, S. (2015). Progress in neuro-psychopharmacology & biological psychiatry associations between DNA methylation and schizophrenia-related intermediate phenotypes — A gene set enrichment analysis. *Progress in Neuro-psychopharmacology &*

Biological Psychiatry, 59, 31–39. https://doi.org/10.1016/j.pnpbp.2015.01.006.
Hendry, A., Jones, E. J. H., Bedford, R., Gliga, T., Charman, T., Johnson, M. H., ... Wass, S. V. (2018). Developmental change in look durations predicts later effortful control in toddlers at familial risk for ASD. *Journal of Neurodevelopmental Disorders, 10*(1), https://doi.org/10.1186/s11689-017-9219-4.
Holt, R., Barnby, G., Maestrini, E., Bacchelli, E., Brocklebank, D., Sousa, I., ... Monaco, A. P. (2010). Linkage and candidate gene studies of autism spectrum disorders in European populations. *European Journal of Human Genetics, 18*(9), 1013–1020. https://doi.org/10.1038/ejhg.2010.69.
Illumina, Inc (2012). *Infinium HumanMethylation450 BeadChip data sheet.* Retrieved fromhttps://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_humanmethylation450.pdf.
Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., ... Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature, 515*(7526), 216–221. https://doi.org/10.1038/nature13908.
Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., ... Wigler, M. (2012). De novo gene disruptions in children on the autistic spectrum. *Neuron, 74*(2), 285–299. https://doi.org/10.1016/j.neuron.2012.04.009.De.
Jaffe, A. E., Gao, Y., Deep-Soboslay, A., Tao, R., Hyde, T., Weinberger, D. R., & Kleinman, J. E. (2016). Mapping DNA methylation across development, genotype, and schizophrenia in the human frontal cortex. *Nature Neuroscience, 19*(1), 40–47. https://doi.org/10.1038/nn.4181.Mapping.
James, S. J., Shpyleva, S., Melnyk, S., Pavliv, O., & Pogribny, I. P. (2013). Complex epigenetic regulation of Engrailed-2 (EN-2) homeobox gene in the autism cerebellum. *Translational Psychiatry, 3*(2), e232–8. https://doi.org/10.1038/tp.2013.8.
Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nature Reviews Neurology, 10*(2), 74–81. https://doi.org/10.1038/nrneurol.2013.278.
Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience, 6*(10), 766–774. https://doi.org/10.1038/nrn1766.
Johnson, M. H. (2017). Autism as an adaptive common variant pathway for human brain development. *Developmental Cognitive Neuroscience, 25*, 5–11. https://doi.org/10.1016/j.dcn.2017.02.004.
Johnson, M. H., Gliga, T., Jones, E. J. H., & Charman, T. (2014). Annual Research Review: Infant development, autism, and ADHD - early pathways to emerging disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 3*, 1–20. https://doi.org/10.1111/jcpp.12328.
Jones, E. J. H., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience and Biobehavioral Reviews, 39*, 1–33. https://doi.org/10.1016/j.neubiorev.2013.12.001.
Jones, E. J. H., Venema, K., Earl, R., Lowy, R., Barnes, K., Estes, A. M., ... Webb, S. J. (2016). Reduced engagement with social stimuli in 6-month-old infants with later autism spectrum disorder: A longitudinal prospective study of infants at high familial risk. *Journal of Neurodevelopmental Disorders, 8*, 7. https://doi.org/10.1186/s11689-016-9139-8.
Jones, E. J. H., Venema, K., Earl, R., Lowy, R., & Webb, S. J. (2017). Infant social attention: an endophenotype of ASD-related traits? *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 58*(3), 270–281. https://doi.org/10.1111/jcpp.12650.
Jones, E. J. H., Venema, K., Lowy, R., Earl, R., & Webb, S. J. (2015). Developmental changes in infant brain activity during naturalistic social experiences. *Developmental Psychobiology, 2*(74), 842–853. https://doi.org/10.1126/scisignal.274pe36.Insulin.
Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6 month-olds later diagnosed with autism. *Nature, 504*(7480), 427–431. https://doi.org/10.1038/nature12715.Attention.
Karolchik, D., Hinrichs, A. S., Furey, T. S., Roskin, K. M., Sugnet, C. W., Haussler, D., & James Kent, W. (2004). The UCSC table browser data retrieval tool. *Nucleic Acids Research, 32*(Database issue), D493–D496. https://doi.org/10.1093/nar/gkh103.
Kato, T., & Iwamoto, K. (2014). Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders. *Neuropharmacology, 80*, 133–139. https://doi.org/10.1016/j.neuropharm.2013.12.019.
Klin, A., Shultz, S., & Jones, W. (2015). Social visual engagement in infants and toddlers with autism: Early developmental transitions and a model of pathogenesis. *Neuroscience and Biobehavioral Reviews, 50*, 189–203. https://doi.org/10.1016/j.neubiorev.2014.10.006.
Krumm, N., O'Roak, B. J., Shendure, J., & Eichler, E. E. (2014). A de novo convergence of autism genetics and molecular neuroscience. *Trends in Neurosciences, 37*(2), 95–105. https://doi.org/10.1016/j.tins.2013.11.005.
Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K. T., ... Eichler, E. E. (2015). Excess of rare, inherited truncating mutations in autism. *Nature Genetics, 47*(6), 582–588. https://doi.org/10.1038/ng.3303.
Kundakovic, M., Gudsnuk, K., Herbstman, J. B., Tang, D., Perera, F. P., & Champagne, F. A. (2015). DNA methylation of BDNF as a biomarker of early-life adversity. *Proceedings of the National Academy of Sciences of the United States of America, 112*(22), 6807–6813. https://doi.org/10.1073/pnas.1408355111.
Laird, N. M., & Ware, J. H. (2007). Random-effects models for longitudinal data. *Science, 38*(4), 963–974.
Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics, 9*. https://doi.org/10.1186/1471-2105-9-559.
Lasalle, J. M. (2013). Epigenomic strategies at the interface of genetic and environmental risk factors for autism. *Journal of Human Genetics, 58*(7), 396–401. https://doi.org/10.1038/jhg.2013.49.Epigenomic.
Lasky-Su, J., Neale, B. M., Franke, B., Anney, R. J. L., Zhou, K., Maller, J. B., ... Faraone, S. V. (2008). Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics, 147B*(8), 1345–1354. https://doi.org/10.1002/ajmg.b.30867.
Li, J., Wang, L., Jiang, T., Wang, J., Li, X., Liu, X., ... Guo, M. (2016). ESNPO: An eQTL-based SNP Ontology and SNP functional enrichment analysis platform. *Scientific Reports, 6*, 6–11. https://doi.org/10.1038/srep30595.
Lim, A. S. P., Srivastava, G. P., Yu, L., Chibnik, L. B., Xu, J., Buchman, A. S., ... De Jager, P. L. (2014). 24-hour rhythms of DNA methylation and their relation with rhythms of RNA expression in the human dorsolateral prefrontal cortex. *PLoS Genetics, 10*(11), https://doi.org/10.1371/journal.pgen.1004792.
Lister, R., Mukamel, E. A., Nery, J. R., Urich, M. A., Puddifoot, C. A., Nicholas, D., ... Ecker, J. R. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science, 341*, 629–643. https://doi.org/10.1126/science.1237905.
Lloyd-Fox, S., Blasi, A., Elwell, C. E., Charman, T., Murphy, D., & Johnson, M. H. (2013). Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proceedings. Biological Sciences / The Royal Society, 280*(1758), 20123026. https://doi.org/10.1098/rspb.2012.3026.
Lloyd-Fox, S., Blasi, A., Pasco, G., Gliga, T., Jones, E. J. H., Murphy, D. G. M., ... Yemane, F. (2018). Cortical responses before 6 months of life associate with later autism. *The European Journal of Neuroscience, 47*(6), 736–749. https://doi.org/10.1111/ejn.13757.
Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... Rutter, M. (2000). The autism diagnostic schedule – generic: A standard measures of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223. https://doi.org/10.1023/A:1005592401947.
Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders, 24*, 659–685.
Martino, D., Loke, Y. J., Gordon, L., Ollikainen, M., Cruickshank, M. N., Saffery, R., & Craig, J. M. (2013). Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biology, 14*(5), R42. https://doi.org/10.1186/gb-2013-14-5-r42.
Mentlein, L., Thorlacius, G. E., Meneghel, L., Ramirez Sepulveda, J. I., Brauner, S., Espinosa, A., & Wahren-Herlenius, M. (2018). FAM167A/BLK is a susceptibility locus in autoimmune diseases: Characterisation of the FAM167 gene family. *Annals of the Rheumatic Diseases, 77*(Suppl 1), https://doi.org/10.1136/annrheumdis-2018-EWRR2018.39 A19-A19.
Michels, K. B., Binder, A. M., Dedeurwaerder, S., Epstein, C. B., Greally, J. M., Gut, I., ... Irizarry, R. A. (2013). Recommendations for the design and analysis of epigenome-wide association studies. *Nature Methods, 10*(10), 949–955. https://doi.org/10.1038/nmeth.2632.
Mitchell, C., Schneper, L. M., & Notterman, D. A. (2015). DNA methylation, early life environment, and health outcomes. *Pediatric Research,* (November), 1–8. https://doi.org/10.1038/pr.2015.193.
Moore, D. S. (2015). *The developing genome* (1st edition). New York: Oxford University Press.
Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology, 38*(1), 23–38. https://doi.org/10.1038/npp.2012.112.

Mullen, E. M. (1995). Mullen scales of early learning. *Circles Pines, MN:AGS.*.
Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *The European Journal of Neuroscience, 47*(6), 497–514. https://doi.org/10.1111/ejn.13720.
Neale, B. M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K. E., Sabo, A., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature, 485*(7397), 242–246. https://doi.org/10.1038/nature11011.
Nenadic, I., Maitra, R., Scherpiet, S., Gaser, C., Schultz, C. C., Schachtzabel, C., ... Schlösser, R. G. M. (2012). Glutamate receptor delta 1 (GRID1) genetic variation and brain structure in schizophrenia. *Journal of Psychiatric Research, 46*(12), 1531–1539. https://doi.org/10.1016/j.jpsychires.2012.08.026.
Neul, J. L., Kaufmann, W. E., Glaze, D. G., Clarke, A. J., Leonard, H., Bailey, M. E. S., ... Renieri, A. (2011). Rett syndrom: Revised diagnostic criteria and nomenclature. *Annals of Neurology, 68*(6), 944–950. https://doi.org/10.1002/ana.22124.Rett.
Nikolova, Y. S., & Hariri, A. R. (2015). Can we observe epigenetic effects on human brain function? *Trends in Cognitive Sciences, 6*(7), 1–8. https://doi.org/10.1016/j.tics.2015.05.003.
Numata, S., Ye, T., Hyde, T. M., Guitart-Navarro, X., Tao, R., Wininger, M., ... Lipska, B. K. (2012). DNA methylation signatures in development and aging of the human prefrontal cortex. *American Journal of Human Genetics, 90*(2), 260–272. https://doi.org/10.1016/j.ajhg.2011.12.020.
Ouellet-Morin, I., Wong, C. C. Y., Danese, A., Pariante, C. M., Papadopoulos, A. S., Mill, J., & Arseneault, L. (2013). Increased serotonin transporter gene (SERT) DNA methylation is associated with bullying victimization and blunted cortisol response to stress in childhood: A longitudinal study of discordant monozygotic twins. *Psychological Medicine, 43*(9), 1813–1823. https://doi.org/10.1017/S0033291712002784.Increased.
Ozonoff, S., Young, G. S., Belding, A., Hill, A., Hutman, T., Johnson, S. P., ... Iosif, A. (2014). The Broader autism phenotype in infancy: When does it emerge? *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(4), 398–407. https://doi.org/10.1016/j.jaac.2013.12.020.
Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., ... Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics, 128*(3), 488–495. https://doi.org/10.1542/peds.2010-2825.
Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., ... Iosif, A. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: a baby siblings research consortium study. *Journal of Child Psychology and Psychiatry, 56*(9), 988–998. https://doi.org/10.1111/jcpp.12421.Diagnostic.
Peter, C. J., Fischer, L. K., Kundakovic, M., Garg, P., Jakovcevski, M., Dincer, A., ... Akbarian, S. (2015). DNA methylation signatures of early childhood malnutrition associated with impairments in attention and cognition. *Biological Psychiatry, 80*(10), 765–774. https://doi.org/10.1016/j.biopsych.2016.03.2100.
Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., ... Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology, 17*(1), 1–17. https://doi.org/10.1186/s13059-016-1066-1.
Pilarowski, G. O., Vernon, H. J., Applegate, C. D., Boukas, L., Cho, M. T., Gurnett, C. A., ... Henderson, L. B. (2018). Missense variants in the chromatin remodeler CHD1 are associated with neurodevelopmental disability. *Journal of Medical Genetics, 55*(8), 561–566. https://doi.org/10.1136/jmedgenet-2017-104759.Missense.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team, A. (2018). *Nlme: Linear and nonlinear mixed effects models.* R package version 3.1-137. Retrieved fromhttps://cran.r-project.org/package=nlme.
Relton, C. L., & Smith, D. G. (2012). Is epidemiology ready for epigenetics? *International Journal of Epidemiology, 41*(1), 5–9. https://doi.org/10.1093/ije/dys006.
Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science, 19*(1), 41–46. https://doi.org/10.1177/0963721409360003.The.
Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience, 18*(11), 671–684. https://doi.org/10.1038/nrn.2017.112.
Robinson, E. B., Anttila, V., Kosmicki, J. A., Bulik-sullivan, B., Grove, J., Maller, J., ... Daly, M. J. (2016). Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nature Genetics, 48*(5), 552–555. https://doi.org/10.1038/ng.3529.Genetic.
Rommelse, N. N. J., Franke, B., Geurts, H. M., Hartman, C. A., & Buitelaar, J. K. (2010). Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *European Child & Adolescent Psychiatry, 19*(3), 281–295. https://doi.org/10.1007/s00787-010-0092-x.
Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain, and Behavior, 2*(5), 255–267. https://doi.org/10.1046/j.1601-183X.2003.00037.x.
Rudenko, A., & Tsai, L.-H. (2014). Epigenetic modifications in the nervous system and their impact upon cognitive impairments. *Neuropharmacology, 80*, 70–82. https://doi.org/10.1016/j.neuropharm.2014.01.043.
Saffari, A., Silver, M. J., Zavattari, P., Moi, L., Columbano, A., Meaburn, E., & Dudbridge, F. (2018). Estimation of a significance threshold for epigenome-wide association studies. *Genetic Epidemiology, 42*(1), 20–33. https://doi.org/10.1002/gepi.22086.
Sanders, S. J., He, X., Willsey, A. J., Ercan-Sencicek, G. A., Samocha, K. E., Cicek, E. A., et al. (2015). Insights into autism Spectrum disorder genomic architecture and biology from 71. *Risk Loci. Neuron, 87*(6), 1215–1233. https://doi.org/10.1007/s12026-008-8082-5.Thymus.
Scerif, G. (2010). Attention trajectories, mechanisms and outcomes: At the interface between developing cognition and environment. *Developmental Science, 13*, 805–812. https://doi.org/10.1111/j.1467-7687.2010.01013.x.
Senju, A., Johnson, M. H., & Tomalski, P. (2014). The two-process theory of face processing : Modifications based on two decades of data from infants and adults. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2014.10.009.
Siddiqui, M. F., Elwell, C. E., & Johnson, M. H. (2016). Mitochondrial dysfunction in autism Spectrum disorders. *Autism-open Access, 6*(5), https://doi.org/10.4172/2165-7890.1000190.Mitochondrial.
Siddiqui, M. F., Lloyd-Fox, S., Kaynezhad, P., Tachtsidis, I., Johnson, M. H., & Elwell, C. E. (2017). Non-invasive measurement of a metabolic marker of infant brain function. *Scientific Reports, 7*(1), 1–6. https://doi.org/10.1038/s41598-017-01394-z.
Simpkin, A. J., Suderman, M., Gaunt, T. R., Lyttleton, O., McArdle, W. L., Ring, S. M., ... Relton, C. L. (2015). Longitudinal analysis of DNA methylation associated with birth weight and gestational age. *Human Molecular Genetics, 24*(13), 3752–3763. https://doi.org/10.1093/hmg/ddv119.
Slieker, R. C., Roost, M. S., Van Iperen, L., & Suchiman, H. E. D. (2015). DNA methylation landscapes of human fetal development. *PLoS Genetics, 11*(10), 1–19. https://doi.org/10.1371/journal.pgen.1005583.
Smith, A. K., Kilaru, V., Klengel, T., Mercer, K. B., Bradley, B., Conneely, K. N., ... Binder, E. B. (2015). DNA extracted from saliva for methylation studies of psychiatric traits: Evidence tissue specificity and relatedness to brain. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics, 1*(0), 36–44. https://doi.org/10.1002/ajmg.b.32278.DNA.
Sokolova, E., Oerlemans, A. M., Rommelse, N. N. J., Groot, P., Hartman, C. A., Glennon, J. C., ... Buitelaar, J. K. (2017). A causal and mediation analysis of the comorbidity between Attention Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum disorder (ASD). *Journal of Autism and Developmental Disorders, 47*(6), 1–10. https://doi.org/10.1007/s10803-017-3083-7.
Sonuga-Barke, E. J. S., & Fearon, P. (2018). Commentary: Whither the epigenetics of child psychopathology? Some reflections provoked by Barker et al. (2018). *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 59*(4), 323–326. https://doi.org/10.1111/jcpp.12906.
Sowers, M. L., Di Re, J., Wadsworth, P. A., Shavkunov, A. S., Lichti, C., Zhang, K., & Laezza, F. (2019). Sex-specific proteomic changes induced by genetic deletion of fibroblast growth factor 14 (FGF14), a regulator of neuronal ion channels. *Proteomes (Preprint), 7*(1), 5. https://doi.org/10.3390/proteomes7010005.
Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales* (second edition). (Vineland-2). (MN: Pearson Assessments, Ed.). Livonia.
Spiers, H., Hannon, E., Schalkwyk, L. C., Smith, R., Wong, C. C. Y., Donovan, M. C. O., ... Mill, J. (2015). Methylomic trajectories across human fetal brain development. *Genome Research,* 1–16. https://doi.org/10.1101/gr.180273.114.Freely.
Stelzer, G., Rosen, R., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., ... Lancet, D. (2016). *GeneCards - the human gene database.* https://doi.org/10.1002/cpbi.5.
Sun, W., Zang, L., Shu, Q., & Li, X. (2014). From development to diseases: The role of 5hmC in brain. *Genomics, 104*(5), 347–351. https://doi.org/10.1016/j.ygeno.2014.08.021.
Szatmari, P., Chawarska, K., Dawson, G., Georgiades, S., Landa, R., Lord, C., ... Halladay, A. (2016). Prospective longitudinal studies of infant siblings of children with autism: Lessons learned and future directions. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(3), 179–187. https://doi.org/10.1007/s10741-014-9462-7.Natural.

Tsai, P. C., & Bell, J. T. (2015). Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *International Journal of Epidemiology, 44*(4), 1429–1441. https://doi.org/10.1093/ije/dyv041.

Tsang, S. Y., Ahmad, T., Mat, F. W. K., Zhao, C., Xiao, S., Xia, K., & Xue, H. (2016). Variation of global DNA methylation levels with age and in autistic children. *Human Genomics, 10*(1), 1–6. https://doi.org/10.1186/s40246-016-0086-y.

Ursini, G., Bollati, V., Fazio, L., Porcelli, A., Iacovelli, L., Catalani, A., ... Bertolino, A. (2011). Stress-related methylation of the Catechol-O-Methyltransferase Val158 allele predicts human prefrontal cognition and activity. *Journal of Neuroscience, 31*(18), 6692–6698. https://doi.org/10.1523/JNEUROSCI.6631-10.2011.

Vergult, S., Dheedene, A., Meurs, A., Faes, F., Isidor, B., Janssens, S., ... Menten, B. (2015). Genomic aberrations of the CACNA2D1 gene in three patients with epilepsy and intellectual disability. *European Journal of Human Genetics, 23*(5), 628–632. https://doi.org/10.1038/ejhg.2014.141.

Waddington, C. H. (1942). The epigenotype. *Endeavour,* 18–20. https://doi.org/10.1093/ije/dyr184.

Wang, H. Z., Qin, H., Guo, W., Samuels, J., & Shugart, Y. Y. (2013). New insights into the genetic mechanism of IQ in autism spectrum disorders. *Frontiers in Genetics, 4*(October), 195. https://doi.org/10.3389/fgene.2013.00195.

Webb, S. J., Jones, E. J. H., Merkle, V. K., Greenson, J., Murias, M., & Dawson, G. (2011). Developmental change in the ERP responses to familiar faces in toddlers with Autism Spectrum disorders versus typical development. *Child Development, 82*(6), 1868–1886. https://doi.org/10.1007/s12020-009-9266-z.A.

Weissman, J. R., Kelley, R. I., Bauman, M. L., Cohen, B. H., Murray, K. F., Mitchell, R. L., ... Natowicz, M. R. (2008). Mitochondrial disease in autism spectrum disorder patients: A cohort analysis. *PloS One, 3*(11), 1–6. https://doi.org/10.1371/journal.pone.0003815.

Wiers, C. E. (2012). Methylation and the human brain: Towards a new discipline of imaging epigenetics. *European Archives of Psychiatry and Clinical Neuroscience, 262*(3), 271–273. https://doi.org/10.1007/s00406-011-0261-z.

Wolffe, A. P., & Matzke, M. A. (1999). Epigenetics: Regulation through repression. *Science, 286*, 481–486. https://doi.org/10.1126/science.286.5439.481.

Won, D. C., Feldman, H. M., & Huffman, L. C. (2019). Sleep problem detection and documentation in children with autism spectrum disorder and attention-deficit/hyperactivity disorder by developmental-behavioral pediatricians - A DBPNet study. *Journal of Developmental & Behavioral Pediatrics, 40*(1), 20–31.

Wong, C. Y., Smith, R. G., Hannon, E., Ramaswami, G., Parikshak, N. N., Assary, E., ... Mill, J. (2018). Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post- mortem human brain tissue. *BioRxiv*.

Yadav, R., Gupta, S. C., Hillman, B. G., Bhatt, J. M., Stairs, D. J., & Dravid, S. M. (2012). Deletion of glutamate delta-1 receptor in mouse leads to aberrant emotional and social behaviors. *PloS One, 7*(3), https://doi.org/10.1371/journal.pone.0032969.

Zablotsky, B., Colpe, L. J., Pringle, B. A., Kogan, M. D., Rice, C., & Blumberg, S. J. (2017). Age of parental concern, diagnosis, and service initiation among children with autism spectrum disorder. *Am J Intellet Dev Disabil, 122*(1), 49–61. https://doi.org/10.1352/1944-7558-122.1.49.Age.