Aggressive behaviour is a common feature of several child and adolescent psychiatric conditions including attention-deficit/hyperactivity disorder (ADHD), conduct disorder (CD) and oppositional defiant disorder (ODD). Several environmental factors are thought to be risk factors for these conditions such as prenatal environmental insults or stressful events during childhood. Recent epigenetic work suggests that variability of DNA methylation levels in neurons could also be implicated. There is growing interest in profiling the epigenetic landscape of different psychiatric conditions. Investigating variation in DNA methylation in different animal models of aggression using bioinformatics pipelines will help us understand the relationship between the moleular aetiology of such child and adolescent conditions and epigenetic factors.

Epigenome-wide association studies (EWAS) aim to identify a relationship between epigenetic variation and a specific phenotype, such as aggression. Epigenetic~~s~~ mechanisms, such as histone and DNA modifications have important roles in regulating gene expression and are independent from DNA sequence. Epigenetic marks regulate processes such as gene silencing, genomic imprinting and X chromosome inactivation. Disregulation of epigenetic marks has been linked to several disorders.

*The focus of this neuroepigenetic study: epigenetic regulation by DNA methylation.*

DNA methylation predominantly occurs on cytosines followed by guanine residues (CpG). CpA, CpT, and CpC (non-CpG) site methylation is less commonly observed. DNA methyltransferases (DNMTs) catalyses the methylation at the fifth carbon of the cytosine, leading to the formation of 5-methylcytosine (5mC). Demethylation of 5mC occurs via a passive or active process. The active process requires oxidation by Ten-Eleven-Translocation (TET) proteins or deamination via activation induced cytidine deaminase (AID)/apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC).

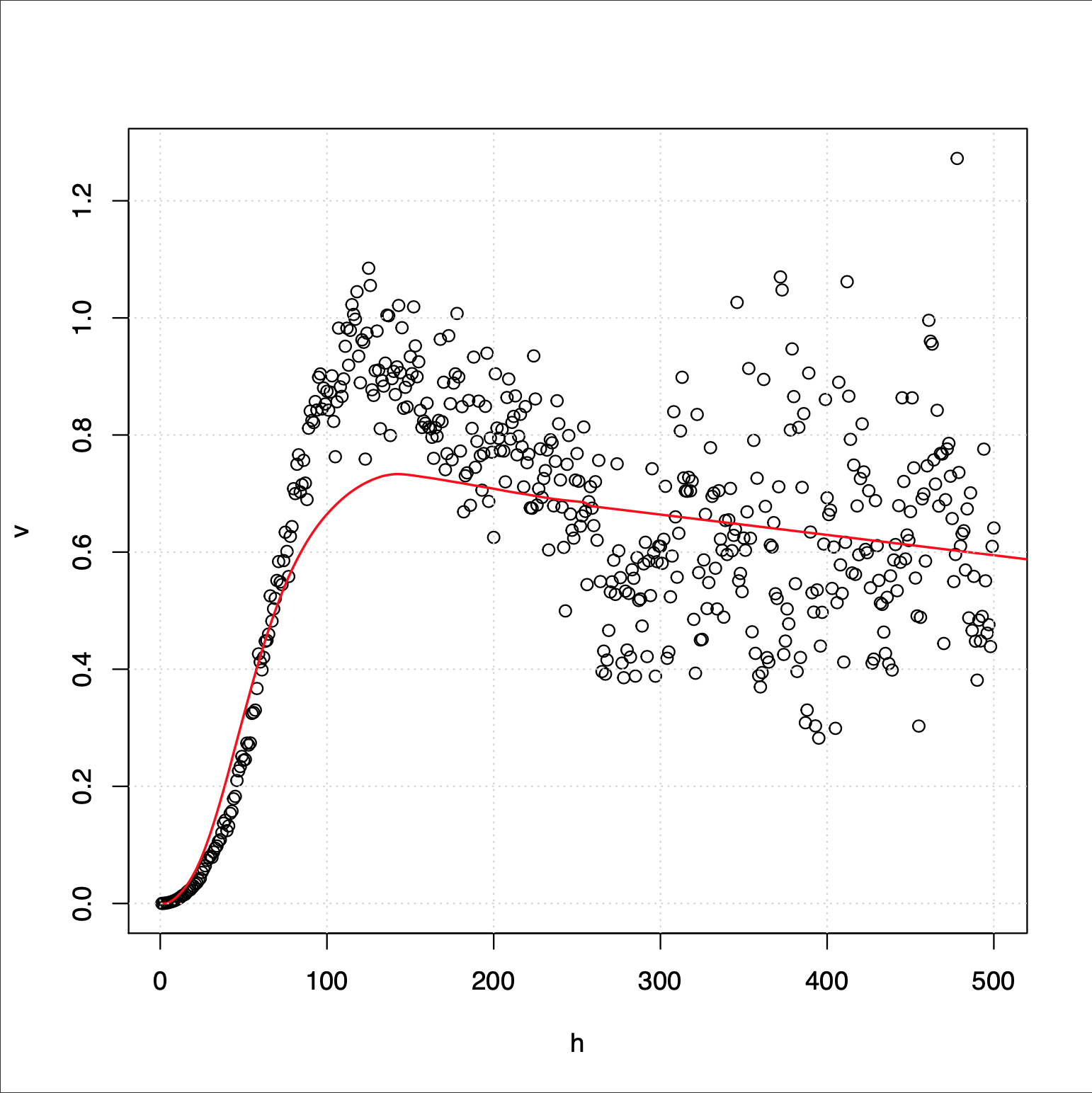
Reduced representation bisulfite sequencing (RRBS) is a sequencing method used to profile DNA methylation levels in the genome. Due to the large CPU power it requires, RRBS is optimally carried out using supercomputers. Luckily, my University had one in handy! I was required to connect my laptop to the University’s supercomputer using a VPN in order to start my project remotely (due to the COVID-19 pandemic).

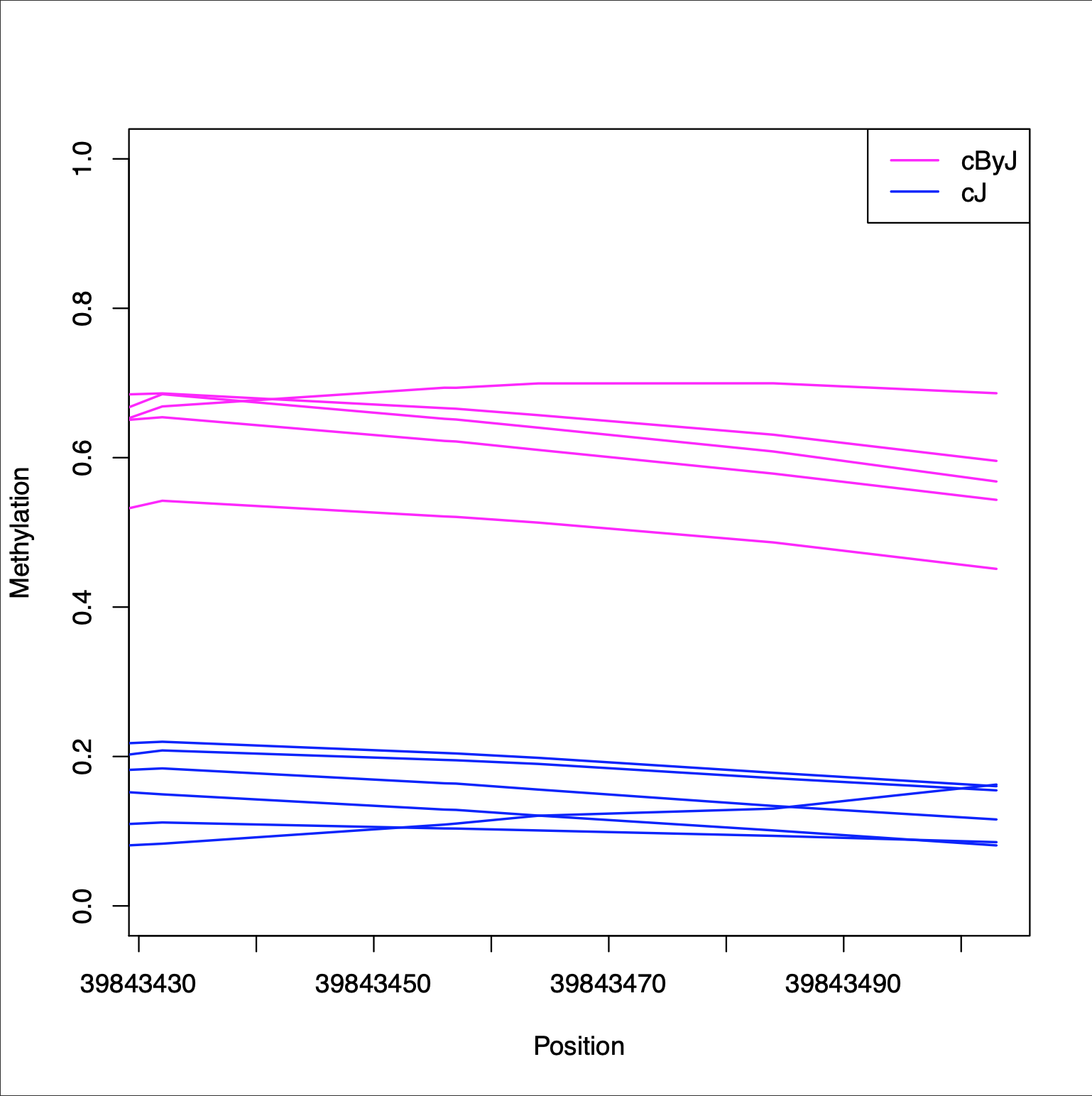
Once connected to the University server, I began investigating DNA methylation levels iand its correlation with aggression. To investigate this we used RBBS data from brain samples of 5 mice with a non-aggressive phenotype (cByJ) and 6 mice with an aggressive phenotype (cJ). Three regions of the brain were analysed; Anterior cingulate cortex (ACC), Midcingulate cortex (MCC) and Ventromedial nucleus (VMH). According to previous studies, the brain regions are involved in the regulation of aggressive behaviour in mice. ACC and MCC are complementary towards each other in their functions: ACC demonstrated an ability to regulate basic threat recognition and MCC mediates approach/avoidance selection during aggressive encounters.

*What does RRBS involve?*

As a preliminary step, the DNA samples undergo bisulfite conversion in the laboratory. Unmethylated Cs are converted to Us in this process, which after PCR amplification are converted to Ts. Methylated Cs are protected against bisulfite and are not converted, allowing clear distinction between methylated and unmethylated CpG sites.

After this step, I could begin the bioinformatics protocols on the sequencing reads. A variety of software was used, for example, fastQC, used to check for low quality raw sequencing reads. Adapters which were ligated to the DNA in the preliminary steps and low-quality sequence ends were trimmed using TrimGalore. The sequencing reads from each sample were aligned to a reference genome, and cDNA methylation at CpG sites was estimated. Finally, Bismark is used to map bisulfite-treated sequencing reads to a genome of interest and extract methylation information.





BiSeq is an R package used for statistical analysis of RRBS data. By producing R scripts which plotted the data, I identified significant patterns. Figure 1 demonstrates spatial correlation between pairs of CpG sites in the MCC brain region as a variogram. It plots distance between pairs (x-axis) against their semi variance (variation between methylation levels, y-axis). Results show a typical semi variance curve, suggesting that CpG sites with a similar methylation level are close together. As the distance increases between pairs, the variation becomes higher, until it stabilises.

Scatter plots were developed for each differentially methylated region (DMR) in the region. Figure 2 represents one DMR from the MCC brain region and compares the methylation levels between the cJ and cByJ phenotypes in that region of the genome. The variation of methylation between the two groups indicates that the DMR contains a gene associated with an increase in aggression; suppressed by methylation in the cByJ group and expressed in the cJ group.

*What would be the next step?*

By researching the role of epigenetic variation in the aetiology of psychiatric disorders scientists can progress their understanding of disease characteristics. The data shows some very interesting results which can be investigated further with other datasets. Pinpointing key DMRs and their positions in the genome would allow us to identify genes associated with aggressive traits.

Luckily, the COVID-19 pandemic did not hinder the completion of my project, which was initially part of the Institute of Metabolic and Systems Research Summer School (IMSR). Although many exciting seminars provided by the IMSR were unfortunately cancelled, I am grateful for the opportunity to explore neuroepigenetics and bioinformatics from my own office (also known as my bed)! Working from home was a lovely experience, thanks to my supervisor, Dr Joana Viana, who allowed me the freedom to self-direct my learning towards my own interests. Despite these trying times, I could experience this opportunity thanks to the financial aid from the Genes and Development Summer Studentship, offered by the Genetics Society. Learning about RRBS, R and neuroepigenetics? I couldn’t have asked for a more intriguing lockdown challenge!