

## ORIGINAL ARTICLE

## Epigenetic differences in monozygotic twins discordant for major depressive disorder

K Malki<sup>1</sup>, E Koritskaya<sup>2</sup>, F Harris<sup>1</sup>, K Bryson<sup>2</sup>, M Herbster<sup>2</sup> and MG Tosto<sup>1,3</sup>

Although monozygotic (MZ) twins share the majority of their genetic makeup, they can be phenotypically discordant on several traits and diseases. DNA methylation is an epigenetic mechanism that can be influenced by genetic, environmental and stochastic events and may have an important impact on individual variability. In this study we explored epigenetic differences in peripheral blood samples in three MZ twin studies on major depressive disorder (MDD). Epigenetic data for twin pairs were collected as part of a previous study using 8.1-K-CpG microarrays tagging DNA modification in white blood cells from MZ twins discordant for MDD. Data originated from three geographical regions: UK, Australia and the Netherlands. Ninety-seven MZ pairs (194 individuals) discordant for MDD were included. Different methods to address non independently-and-identically distributed (non-i.i.d.) data were evaluated. Machine-learning methods with feature selection centered on support vector machine and random forest were used to build a classifier to predict cases and controls based on epivariations. The most informative variants were mapped to genes and carried forward for network analysis. A mixture approach using principal component analysis (PCA) and Bayes methods allowed to combine the three studies and to leverage the increased predictive power provided by the larger sample. A machine-learning algorithm with feature reduction classified affected from non-affected twins above chance levels in an independent training-testing design. Network analysis revealed gene networks centered on the *PPAR-γ (NR1C3)* and *C-MYC* gene hubs interacting through the *AP-1 (c-Jun)* transcription factor. *PPAR-γ (NR1C3)* is a drug target for pioglitazone, which has been shown to reduce depression symptoms in patients with MDD. Using a data-driven approach we were able to overcome challenges of non-i.i.d. data when combining epigenetic studies from MZ twins discordant for MDD. Individually, the studies yielded negative results but when combined classification of the disease state from blood epigenome alone was possible. Network analysis revealed genes and gene networks that support the inflammation hypothesis of MDD.

*Translational Psychiatry* (2016) 6, e839; doi:10.1038/tp.2016.101; published online 14 June 2016

## INTRODUCTION

Major depressive disorder (MDD) is a pervasive psychiatric disorder characterized by a number of clinical symptoms including: persistent low mood, anhedonia, insomnia, low energy, feelings of guilt and ideation of death or suicide.<sup>1,2</sup> MDD is also associated with a range of social impairments, including educational and occupational problems and with an increased risk of developing systemic disease, such as cardiovascular disease and Type 2 diabetes.<sup>3</sup> Epidemiological studies have shown links between MDD and increased levels of mortality due to either suicide or resulting diseases.<sup>4,5</sup>

Behavioral genetic research into the etiology of depression reports heritability estimates between 31 and 42% (refs 6–8) but uncovering common sequence variants associated with the pathology has been challenging. To date, there are no common genetic variants of sufficiently high penetrance to account for the pathology that have clinical significance, although variants associated with the disease at the genome-wide level have been recently announced.<sup>9</sup> Environmental factors, such as early- and late-life stressors, are also thought to increase the risk of developing depression; however, the interaction between genetic and environmental factors remains poorly understood.<sup>10</sup>

Studying depression from the epigenetic perspective sheds light on its etiology by uncovering how environmental factors modulate gene expression.<sup>11</sup> Epigenetics is the study of cellular modification and differentiation that is independent of alterations in the DNA sequence. This can be defined simply as the variation in phenotypic expression produced by the up- and down-regulation of genes using DNA methylation and histone modification, rather than changes in the DNA sequence. It has been suggested that through these processes environmental factors can cause lasting changes in gene expression.<sup>12</sup> For example, early-life stress leads to epigenetic changes in neural tissue involved in the stress response in both animal and human studies.<sup>13,14</sup>

Quantitative genetics research typically compares phenotypic concordance rates between monozygotic (MZ) and dizygotic twins. However, discordant MZ twins (who are assumed to share 100% of their genome) can also be used to identify epigenetic factors, such as differing levels of DNA methylation at specific loci that may contribute to explain part of their phenotypic discordance.<sup>15</sup>

The genome (the DNA sequence) is consistent throughout all cell types in the human body: hence, buccal epithelial cells are largely used in genetic research because of their convenience.

<sup>1</sup>King's College London, MRC Social, Genetic and Developmental Psychiatry Centre at the Institute of Psychiatry, Psychology and Neuroscience, London, UK; <sup>2</sup>Department of Computer Science, University College London, London, UK and <sup>3</sup>Laboratory for Cognitive Investigations and Behavioural Genetics Tomsk State University, Tomsk, Russia. Correspondence: Dr K Malki, King's College London, MRC Social, Genetic and Developmental Psychiatry Centre at the Institute of Psychiatry, Psychology and Neuroscience, SGDP Research Centre (PO80), De Crespigny Park, Denmark Hill, London SE5 8AF, UK. E-mail: karim.malki@kcl.ac.uk

Received 6 December 2015; revised 5 April 2016; accepted 20 April 2016

However, epigenetic changes are more tissue-specific. In the field of psychiatry, this presents a problem as, although neural tissues of interest can be obtained post-mortem, in live patients, only cells in peripheral tissues can generally be obtained. A recent study by Oh *et al.*<sup>16</sup> successfully replicated multiple epigenetic features of depression between brain and non-brain tissues, suggesting that peripheral blood cells are adequate for investigating the epigenetics of depression.<sup>16</sup> Studies using peripheral blood cells may also have an advantage in that they may show fewer changes associated with disease-related external factors, including exposures to stress and pharmacological treatments.<sup>17</sup> On the other hand, it is likely that disease-relevant epigenetic variations may be considerably weaker in peripheral blood cells than in the neural tissue of interest or even be undetectable.

The study by Oh *et al.*<sup>16</sup> analyzed epigenetic variations across different tissues in three data sets of discordant MZ twin pairs from three centers in different countries in order to identify epigenetic markers associated with depression.<sup>16</sup> However, when they combined the data sets collected from blood, they were not able to discriminate affected siblings. This study used the same three previously published peripheral blood data sets from MZ twins discordant for MDD, aiming to address some of the statistical challenges of this data set to produce a global classification model. This required several data consolidation issues to be addressed. First, batch effects may lead to the data not being independently and identically distributed (non-i.i.d.); this refers to error introduced by systemic differences in the ways samples are taken and analyzed between different laboratories. The second refers to population stratification and noise variance. This study aims to explore methods to combine these data sets, adjusting for sources of error and then use the increased power of the combined data to classify affected and non-affected individuals according to epigenetic markers from peripheral blood samples of MZ twins discordant for MDD.

## MATERIALS AND METHODS

### Design

Merging raw data often violates independent-and-identically distributed (i.i.d.) assumptions and thus leads to biased model parameter estimations, higher generalization error and lower cross-validation model performance. The study first used an atheoretical, data-driven approach to evaluate methods for batch effect removal and non-i.i.d. data consolidation. In the second part of the study, we used two different machine-learning approaches (linear support vector machines (SVMs) and random forest (RF)) in order to build a model that could classify affected siblings from non-affected ones using the merged data set from the three, independent, cross-country, epigenetic studies described below. Levels of methylation were used as features and affected and non-affected status within each discordant twin pair as outcome labels. Training and testing of the algorithms was performed on independent samples (80:20% split). Feature selection methods for each algorithm were chosen based on cross-validation. Finally, several different thresholds for the number of selected features were evaluated to extract epigenetics markers likely to contribute to variation in the pathology. These features were then carried forward for gene-network analysis.

### Sample and variable description

The data used for this study have been obtained from a recently published study by Oh *et al.*, and are available from the Gene Expression Omnibus (accession number GSE37579). Briefly, methylation profiles were obtained using Illumina 8.1-K CpG island microarrays from white blood cells sampled from MZ twins discordant for MDD. The samples originated from three different regions: the United Kingdom (28 samples), Australia (80 samples) and the Netherlands (86 samples). Levels of methylation were measured for either coding or noncoding DNA regions common across the samples. Following quality control procedures, a total of 8448 features with log-normalized levels of methylation for each sample were obtained that were consistent across the three studies. Depending on the source of information, depressive symptoms were evaluated by structured clinical

interview. A diagnosis of MDD was determined by the results of these questionnaires according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition criteria. Australian participants were drawn from the Australian Twin Registry. Forty pairs were selected for this study, including 31 female pairs. The average age of the twins was 41.2 years, with a s.d. of 11.5 years. Dutch participants were drawn from the Netherlands Twin Register. Forty-three twin pairs were selected for this study, with an average age of 38.4 years and a s.d. of 12.7 years. UK participants were drawn from the St Thomas's Hospital Twin Registry, the Maudsley Hospital (London) Twin Register and from ongoing studies of volunteer twins. Fourteen pairs including two males were selected for this study. The average age of participants was 53.7 years, with a range from 21 to 65 years. This data set has been chosen for several key reasons. First, it is based on MZ discordant twins, which are considered matched for: genetics, age, sex, cohort effects, maternal influences and common environments. By design, MZ studies have more power to detect disease-related epigenetic differences than studies on unrelated individuals. Second, the same Illumina array has been used across the three studies avoiding potential confounds resulting from matching markers on different arrays and artefacts injected by imputation methods. Lastly, Oh *et al.* reported that, using all the twin samples (UK, Australian and Dutch), they were unable to discriminate affected individuals from control subjects from blood. Given their promising results from other tissue, there was good reason to believe that important information could still be gained from this published data set. Further information on the samples, ethical statements and approval data collection and data pre-processing is described elsewhere.<sup>16</sup>

### Statistical analysis

Empirical non-i.i.d. data consolidation methods, including principal component analysis (PCA) and non-parametric Bayes methods (ComBat), have been evaluated in isolation and in combination to batch effect removal, population stratification and other noise variance. Machine-learning methods and feature selection using linear SVMs and RF were used to build a classifier to predict cases and controls for MDD from epigenetic markers and extract features with higher probability of explaining variations in pathology.

### Data normalization and evaluation metric

Two main approaches of batch effect removal were evaluated in isolation and in combination. The PCA method relies on the idea that the direction with higher variance might relate to noise or population stratification rather than disease. The non-parametric ComBat approach is an empirical Bayes method that aims to adjust for unknown, unmodeled or latent sources of noise and systematic bias. MZ twin studies discordant on any phenotype are intrinsically balanced by nature: in each twin pair, we have an affected and non-affected twin. This balance was preserved in the training and test data sets during the resampling procedure via randomly sampling from twin pairs rather than from the whole data set. Although receiver operating characteristic (ROC) returns generally higher values, in this study accuracy was preferred as a more representative, conservative and honest measure of model performance.

### Data consolidation approaches

PCA, ComBat and mixture approaches were evaluated to address issues of non-i.i.d. data and control for potential confounding effects. The PCA method relies on the notion that eigenvectors with higher variance relate to subgroup phenotypes as opposed to disease groups. This approach removes unwanted variance by subtracting a matrix achieved via eigenvector decomposition. Removal of unwanted variance could relate to the removal of batch effect as described by the paper by Nielsen *et al.*,<sup>18</sup> and to control for population stratification as described by Price *et al.*<sup>19</sup> This study considered the approaches above, and unwanted variance was removed from each data set before merging data into one larger set as follows.

For each region  $i$ , data were adjusted as follows:  $Y_i^* = Y_i - U_i W_i^k V_i^T$ , which represents the eigenvector decomposition of the data matrix. Matrices  $U$  and  $V$  contain the top  $k$  eigenvectors corresponding to the top  $k$  eigenvalues given in  $W$ .  $Y$  represents gene expression and  $W_i^k$  is a diagonal matrix corresponding to the top  $k$  eigenvalues. For the purpose of evaluation we subtracted the matrix related to the most informative principal components ranked by eigenvalue. Therefore,  $PCA.1$  subtracts the matrix related to the first principal component where  $PCA.2$ ,  $PCA.3$ ,  $PCA.4$

and *PCA.5* sequentially subtracted the matrix related to principle components 2–5.

The second approach was based on the non-parametric ComBat method implemented in the Surrogate Variable Analysis 'sva' R-package available from Bioconductor.<sup>20</sup> This is an empirical Bayes method aimed to adjust for unknown, unmodeled or latent sources of noise. ComBat adjusts for systematic batch bias common across genes, assuming that batch effect factors often affect many genes in similar ways, similar to increased expression or higher variability. The other benefit of adjusting for systematic bias with ComBat is that it robustly adjusts batch bias for even small batch sizes.<sup>21</sup> ComBat is a three-step empirical Bayes method: (1) standardization of the data is achieved using the formula:

$$Z_{ijg} = \frac{Y_{ijg} - \hat{\alpha}_g - X\hat{\beta}_g}{\hat{\sigma}_g}$$

where  $\hat{\sigma}_g^2 = 1/N \sum_{ij} (Y_{ijg} - \hat{\alpha}_g - X\hat{\beta}_g)^2$  and  $\hat{\alpha}_g, \hat{\beta}_g, \hat{\gamma}_{ig}$  are estimations of parameters  $\alpha_g, \beta_g, \gamma_{ig}$  in a model

$$Y_{ijg} = \alpha_g + X\beta_g + \gamma_{ig} + \delta_{ig}\epsilon_{ijg}$$

where  $N$  is the total number of samples,  $m$  is number of batches,  $n_i$  is number of samples within a batch  $i$  for  $i = 1, \dots, m$ , for genes  $g = 1, \dots, G$ .  $Y_{ijg}$  represents the expression value for gene  $g$  for sample  $j$  from batch  $i$ ,  $\alpha_g$  is the overall gene expression,  $X$  is a design matrix for sample conditions and  $\beta_g$  is the vector of regression coefficients corresponding to  $X$ . The error terms,  $\epsilon_{ijg}$ , can be assumed to follow a Normal distribution with expected value of zero and variance  $\sigma_g^2$ . The  $\gamma_{ig}$  and  $\delta_{ig}$  represent the additive and multiplicative batch effects of batch  $i$  for gene  $g$ , respectively.

(2) Batch effect parameters are estimated batch using empirical priors. Assuming that standardized data  $Z_{ijg} \sim N(Y_{ijg}, \sigma_{ig})$  are normally distributed, the non-parametric estimates are:

$$\hat{\gamma}_{ig}^* = \frac{\sum_{g'} w_{ig'} \hat{\gamma}_{ig'}}{\sum_{g'} w_{ig'}}, \quad \hat{\delta}_{ig}^{2*} = \frac{\sum_{g'} w_{ig'} \hat{\delta}_{ig'}^2}{\sum_{g'} w_{ig'}}$$

where  $\hat{\gamma}_{ig} = 1/n_i \sum_j Z_{ijg}$ ,  $\hat{\delta}_{ig}^2 = 1/(n_i - 1) \sum_j (Z_{ijg} - \hat{\gamma}_{ig})^2$  and  $w_{ig'} = L(Z_{ijg} | \hat{\gamma}_{ig'}, \hat{\delta}_{ig'}^2)$

(3) The method adjusts the data for batch effects

$$Y_{ijg}^* = \frac{Y_{ijg} - \hat{\alpha}_g - X\hat{\beta}_g - Y_{ijg}^*}{\hat{\delta}_g} + \hat{\alpha}_g + X\hat{\beta}_g$$

using estimated batch effects  $\hat{\gamma}_{ig}^* = \hat{\sigma}_{ig} / \hat{\delta}_{ig}^{2*} (Z_{ijg} - \hat{\gamma}_{ig}^*) + \hat{\alpha}_g + X\hat{\beta}_g$

Lastly, we evaluated mixture approaches by combining PCA with empirical Bayes method: first, we removed unwanted variance from irrelevant factors and then tried to identify remaining noise and systematic bias using non-parametric Bayes method. Different classification algorithms were subsequently used on the different combinations.

### Feature reduction and filtering

Several filtering scores were compared in order to select the best-performing method for each classification algorithm. Average classification accuracy and its s.d. were obtained from 200 random, independent resamplings. For each resampling step, the data set was split 80/20 into training and test sets, respectively, and the features with the highest scores on the training set were selected. The model was then built on the training set and performance was evaluated on the test set. The following filtering scores have been compared and the best-performing methods for support vector machine and RF classification algorithms were carried forward for the purposes of analysis:

- **T-statistic:** The absolute value of Welch's *t*-test statistic value. In case of a two-class classification problem, *t*-statistic is equivalent to the difference between two means for each class adjusted for s.d. for both classes.
- **Mean difference:** The absolute value of the difference between average values in each class.
- **P-value for Welch's t-test with degrees of freedom.**
- **Correlation:** Absolute value of Pearson correlation between a variable and the predicting factor.
- The package 'gene-filter' for R, available from Bioconductor, has been used to apply the above filtering criteria.<sup>22</sup>
- **P-value Limma:** The *P*-value calculated for the single-variable regression model using Limma. Bioconductor limma R package has been used for computing this score.<sup>23</sup>

- **RankProd:** RankProd modifies and extends the rank product method proposed by Breitling et al.<sup>24</sup> to integrate multiple studies from different laboratories and/or platforms. In case of pairwise ratios, for each gene it calculates feature rank for each comparison and returns—in this case, a log ratio of methylation levels in twins pair. Rank product is a geometric mean of ranks under different comparisons. The method has been implemented using the 'RankProd' R package.<sup>25</sup>

### Statistical learning

Two machine-learning classification approaches differing in several key aspects, namely linear SVMs and RF, were used to evaluate the effectiveness of the methods chosen to remove batch and population stratification effects and to classify affected from non-affected siblings based on methylation variation.<sup>26,27</sup> SVM and RF have parameters that need to be tuned. The linear SVM was implemented using a soft margin parameter *C*. RF has a number of variables randomly sampled as candidates at each split and used as parameter. The tuning and selection of parameters is required for both classifiers.

Linear SVM is a powerful method with an efficient computation time for a large number of variables and a small number of samples. As many other  $\ell_2$  norm classification methods, SVM is sensitive to noise coming from all variables used. Selecting relevant features and decreasing the number of features aids noise reduction. The linear SVM model used has a soft margin parameter *C*, which is responsible for misclassification penalties. Higher values of parameter *C* forces the SVM to a higher penalty of misclassification on a training set that could lead to a higher generalization error if the final performance is only marginally above chance. Lower values of *C* could lead to lower generalization error. A soft margin parameter of *C* = 0.05 was chosen following cross-validation.

RF has an embedded control for sensitivity to noise, but it is computationally inefficient for a large number of variables. Feature selection aids computation efficiency with additional noise reduction achieved by removal of irrelevant features, and therefore it is particularly important for large data sets using RF. RF has a tuning parameter, which is effectively the ratio of the number of variables randomly sampled as candidates at each split over the total number of variables. Control over the number of variables at each split aims to avoid correlation between trees in a forest. Prediction of a single tree is highly sensitive to noise in a training set, whereas average prediction across many trees is less sensitive if trees are not correlated. Different ratios for the RF model (with top 100 features selected by limma *P*-value) were evaluated and the ratio = 0.1 was chosen as it returned higher average accuracy over 200 resampling.

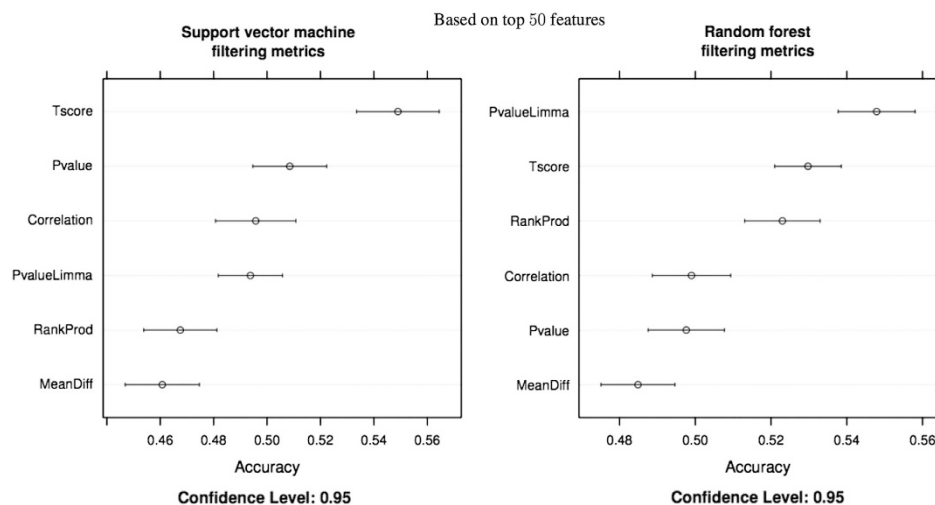
Average classification accuracy as a measure of classification performance and its s.d. was obtained from 200 random independent resamples. At each resampling step, the data set was split into a training and a test set using a 80/20 ratio; features with the highest scores on the training set were selected. A model was built on the training set and performance evaluated on the independent test set. Consolidated data set were compared against the following unadjusted data sets:

- **Australian, the Netherlands and UK**—classification accuracy of the three regions was evaluated independently.
- **RAW data**—raw data from each region were merged without applying non-i.i.d. consolidation methods.
- **NormMean**—the mean normalized raw data before merging (for each data set, the mean normalization subtracts average values for each feature).
- **NormScale**—the mean and s.d. normalized raw data before merging (for each data set, the mean and s.d. normalization subtract average values for each feature and divide by s.d.).

The mean normalization and scaling aims to remove inequality between data sets in cases where for each epigenetic marker relative values between samples remain similar across different data sources, whereas the absolute level could vary together with variance. The above methods provide a baseline to evaluate the increase in performance following correction methods for violation of non-i.i.d. assumptions.

### Network analysis

In order to gain further biological insight into the most predictive features, MetaCore (<https://portal.genego.com/>) was used to explore the intersection between the list of genes mapped to variants uncovered from this



**Figure 1.** Linear support vector machines and random forest performances with features selected using different filtering scores. *t*-statistics is the feature reduction method that yielded higher classification accuracy for support vector machine (SVM), whereas Limma *P*-value showed better performance using random forest (RF). The results are based on top 50 features.

study and known pathway maps and networks. Features most likely to be selected following feature reduction methods and that yielded the highest classification accuracy were mapped to known genes using the UHN Microarray Centre's CpG Island instrument (<https://www.pmgenomics.ca/cpg>). MetaCore scores and prioritizes networks and pathway based on the relevance of the genes uploaded. The gene list returned by the analysis presented may point at genes that show potential interaction, differential expression and may be involved in the pathology. The gene list could also potentially point at possible targets for therapeutic drug discovery. MetaCore evaluates the magnitude of the intersection between the reference gene list and the set of genes corresponding to a network module and returns different metrics including *P*-values and *G*-scores. *P*-values are calculated based on hypergeometric distributions, and these are used to establish whether saturation with the genes of interest is higher than random. When exploring signaling cascades, it is possible to evaluate whether a network contains any fragments of well-understood (canonical) signaling pathways. The *G*-score is another metric used by the software, which effectively modifies the *Z*-score based on the number of the linear canonical pathway fragments contained within the network. A high *G*-score therefore points at a network highly saturated with reference genes and containing several canonical pathways. In this study we have explored the top-ranking networks by *P*-value and by *G*-score.

#### Code availability

All code including R scripts used to generate this analysis are available from the corresponding author on request or can be downloaded online (<http://www.adamlab.org/epigenetics-of-depression/>).

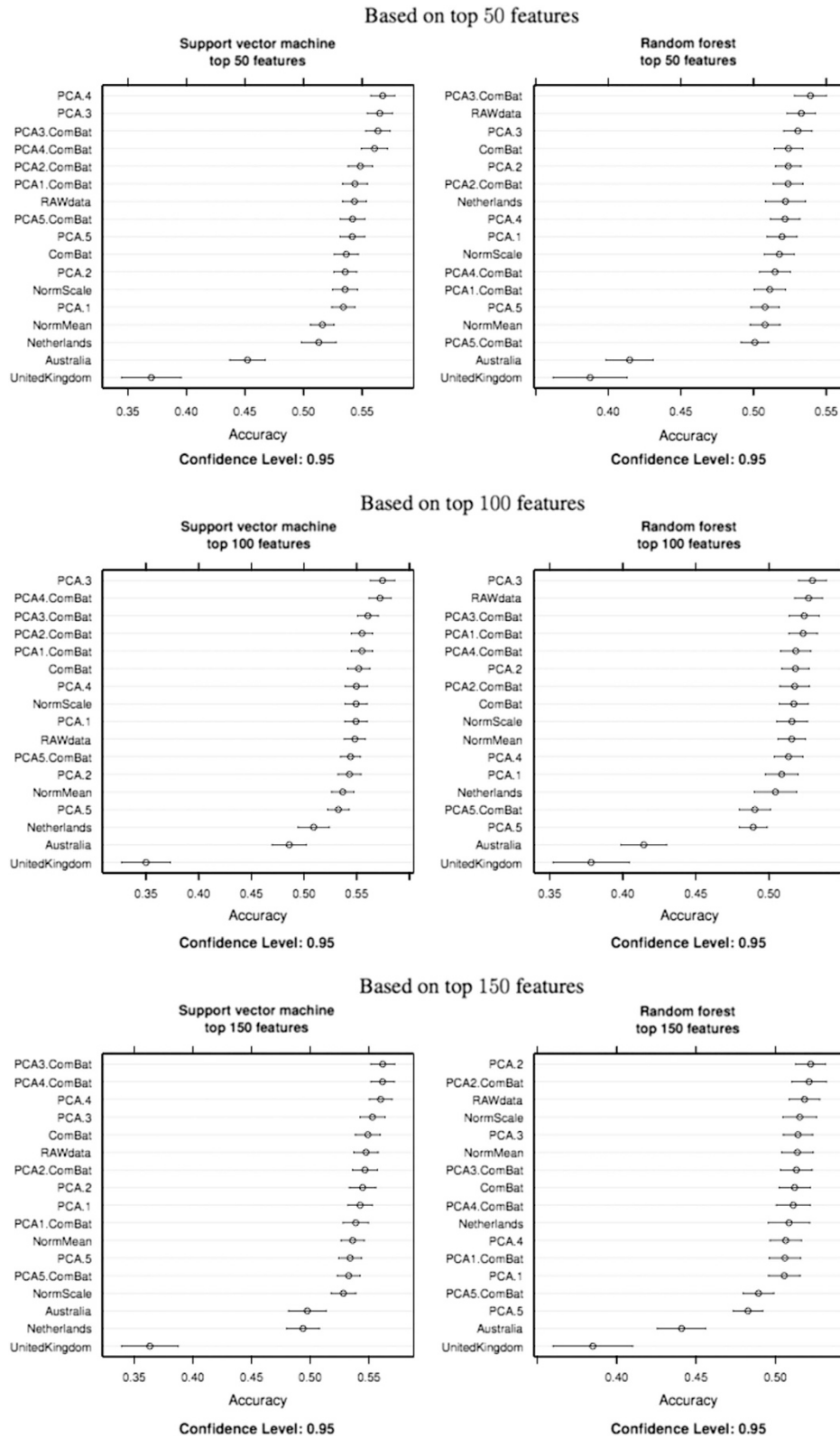
## RESULTS

Data resulting from high-throughput technologies tend to be high-dimensional with many more variables than cases. Methods that can inform on the most relevant features and reduce the dimensionality of the data by identifying variants with larger effect size are used to offer a more favorable signal/noise ratio and to reduce model over-fitting. As each data set has unique structures and substructures, we first empirically evaluated different methods for feature selections for each of the two classification methods used. The results of our evaluation showed that the *t*-test statistics scoring metric was the filtering method that gave the best performance when using an SVM while single-variable linear model *P*-value yielded better-accuracy performance when using RF. Other scores such as correlation, *P*-value and mean difference showed comparatively worse performances (Figure 1). These two methods were carried forward for classification analyses.

Analysis using SVM and RF in each of the three studies independently (UK, Australia and the Netherlands), with and without feature selection, returned overall poor classification accuracy. It was not possible to classify cases above chance level across any of the studies. The worse model performance was obtained in the UK study likely because of the smaller sample size compared with Dutch and Australian studies. The Dutch and Australian studies performed better in comparison with the UK study but were also underpowered and returned accuracy scores around chance levels (Figure 2). We then repeated the analyses using the three samples combined together. However, without correction to overcome issues of non-i.i.d. and removal of potential confounding effects, the combined sample produced a classifier that still did not perform better than chance, even when feature reduction was used (Figure 2). This is consistent with results reported by Oh *et al.* Feature reduction was an important step to improve the performance of the two algorithms but only after various correction methods applied to control for the non-i.i.d. nature of the data. With dimensionality reduction and correction methods, accuracy improved with the highest-accuracy performance achieved with a set of 100 features.

When evaluating different methods to control for confounding effects, *PCA.3* and mixture models using *PCA3.ComBat* showed a steady improvement of performance over raw data set for both linear SVM and RF models with different numbers of selected features. Methods with one more or one less number of removed principal components such as *PCA.4* and *PCA.2* with or without ComBat methods returned better accuracy for some classifiers, but were not steady across different numbers of selected features and different models using independent training and testing sets (Figure 2).

Mixture models using *PCA.3* together with ComBat and *PCA.4* together with ComBat were the best-performing methods when using a SVM, whereas mixture methods using *PCA.2*, ComBat and *PCA.2* returned the best performance using RF. The best overall classification was achieved using SVM and either removing the first three principle components or with a mixture approach using *PCA.3* and ComBat in conjunction with feature elimination. A classification of 58% was achieved from epigenetic variations in blood with 100 features selected. The results are significantly above chance level and remained above 57% even when the number of features was reduced to 50. The top-ranking features selected across the three different studies following resampling are summarized in Table 1. The signal that can be detected from

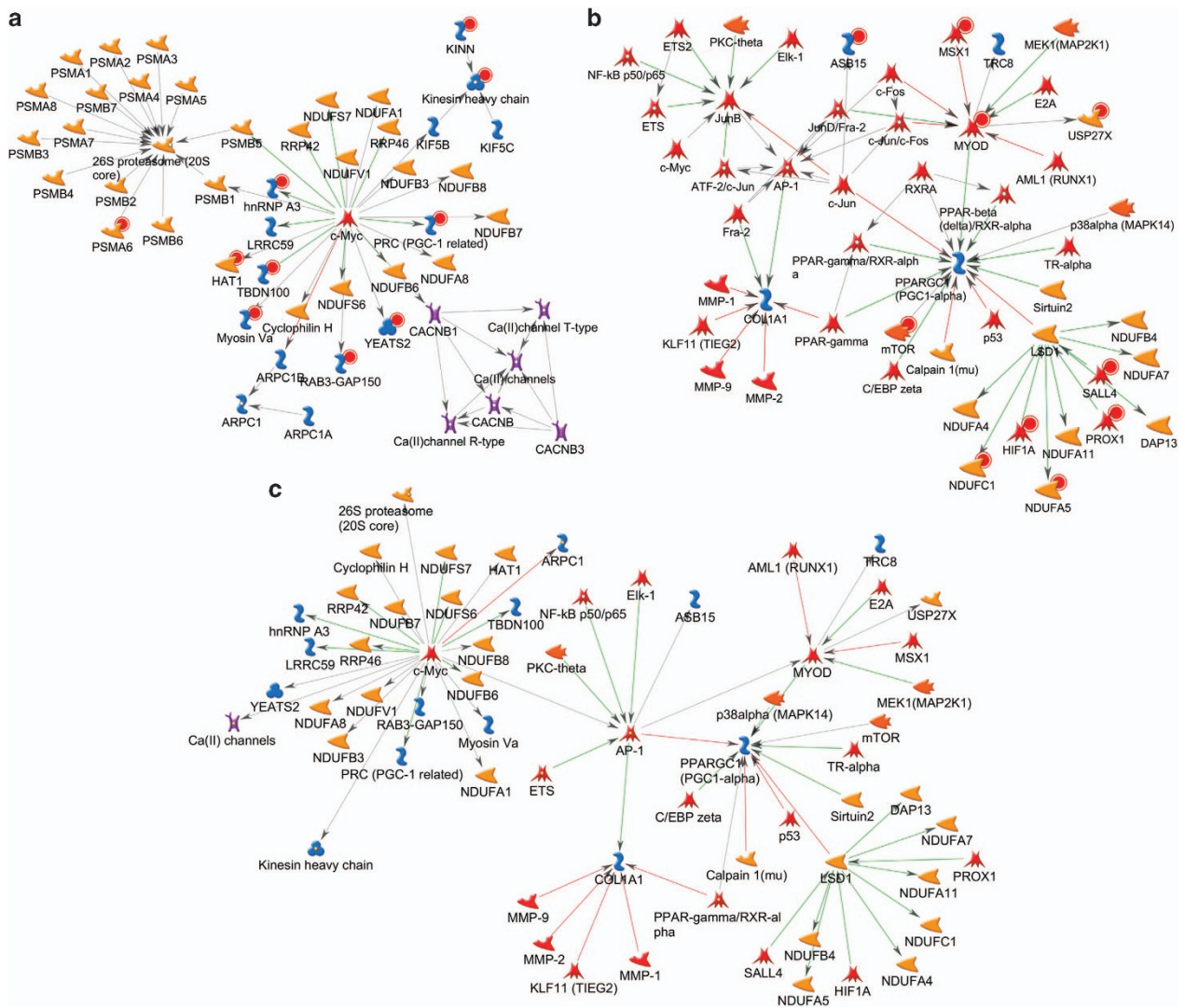


**Figure 2.** Summary of results for two classification methods (support vector machines and random forest) across different methods to control for non-independent-and-identically distributed (non-i.i.d.) data and different depths of feature selection. The mean accuracy percentage scores are reported on the *x* axis. It is not possible to classify cases and controls by analyzing each study individually, even after feature selection. The UK sample, the smallest, performs considerably worse compared with the other two, which show around chance levels. By leveraging the increased power offered by the combined samples and removing potential batch and noise variance it is possible to detect a weak but stable and significant signal. The highest classification accuracy (58%) is achieved with 100 features.

**Table 1.** The table reports the array probes that have higher probability of being selected following cross-validation

UHNID	Genome location	Within strand	Gene symbol	Upstream strand	Gene symbol	Downstream strand	Distance	Gene symbol
UHNhscpg00013724	chrX:49643159-49643620	-	hCG_1993022	+	PAGE4	+	849	USP27X
UHNhscpg00014299	chr1:220700559-220701079	-	TMEM87A	-	RAB3GAP2	+	488	KIAA1477
UHNhscpg00013480	chr15:42564937-42565851	-	TMEM87A	-	VPS39	+	4	GANC
UHNhscpg00000046	chr10:44405717-44405979	+	PMAIP1	-	HNRPA3	+	28471	AX747950
UHNhscpg00004659	chr18:57566852-57567954	-	FAM155A	-	CCBE1	+	470609	MC4R
UHNhscpg00014315	chr13:108325128-108325573	+	ARHGAP15	+	ARGLU1	-	114136	BC043519
UHNhscpg00014319	chr2:143972533-143972830	+	SPTBN4	+	KYNU	+	220332	ARHGAP15
UHNhscpg00003893	chr19:41073929-41074621	+	PTPN9	+	SPTBN4	+	8135	SKBP1
UHNhscpg00000830	chr20:11309478-11309784	-	AK301806	+	BC043370	+	561692	BTBD3
UHNhscpg00005193	chr15:75871291-75871420	+	IQCB1	+	LENG6	-	19005	SNUPN
UHNhscpg00014339	chr7:63774121-63774927	+	Y44	+	ZNF679	+	38441	AK123141
UHNhscpg00014256	chr9:109514015-109514290	+	DIS3	+	AK093363	+	111087	ZNF462
UHNhscpg0000120	chr17:8029524-8030292	-	NDUFA5	-	HES7	-	13496	hPer
UHNhscpg00014546	chr3:121553705-121554108	+	MAP6D1	+	GOLGB1	+	59062	SLC15A2
UHNhscpg00013060	chr6:133561543-133562818	+	SCN2A	+	LOC285735	-	193397	AK093513
UHNhscpg0002422	chr11:131557833-131558280	+	SCN2A	+	AK128059	+	221219	NTM
UHNhscpg00013741	chr13:73355948-73356559	-	DEF6	-	C13orf34	+	272554	KLF5
UHNhscpg00014277	chr7:123197380-123198004	-	FRMD4A	-	IQUB	+	43916	ASB15
UHNhscpg00014284	chr15:52860388-52860692	+	FRMD4A	+	MYO5A	-	12830	DKFZp666F111
UHNhscpg00017254	chr4:56221111-56221294	+	LOX	-	KDR	-	8845	BC047403
UHNhscpg00016077	chr19:57350030-57350493	-	LOX	-	ZIM2	+	1776	MIMT1
UHNhscpg0000829	chr5:121412422-121413979	-	LOX	-	LOX	+	51235	ZNF474
UHNhscpg00014173	chr11:17694766-17694956	+	MAP6D1	+	OTOG	+	46153	MYOD1
UHNhscpg00004403	chr10:103880275-103880427	+	SCN2A	+	LDB1	+	12359	PPRC1
UHNhscpg00013749	chr3:183542146-183542612	+	SCN2A	+	YEATS2	-	4561	PARL
UHNhscpg00014254	chr2:166103020-166103587	+	SCN2A	+	SCN3A	+	46753	SCN2A
UHNhscpg0006419	chr1:24969946-24970475	+	SRRM1	+	C1orf130	+	4776	SRRM1
UHNhscpg0003218	chr14:35761163-35761287	+	KIAA0391	+	KIAA0391	+	286	PSMA6
UHNhscpg00007062	chr6:35265208-35266730	+	DEF6	+	ZNF76	-	41098	BC016143
UHNhscpg0003949	chr10:14287820-14288200	-	FRMD4A	-	FRMD4A	-	79973	KIAA1294
UHNhscpg00013413	chr2:172778466-172778887	-	FRMD4A	-	SLC25A12	+	47	HAT1
UHNhscpg00015148	chr3:61236979-61237436	-	FHIT	-	BRMS1L	-	25873	MBIP
UHNhscpg00013509	chr4:140223250-140223792	+	NARG1	+	NPCR	+	309806	PTPRG
UHNhscpg0005075	chr22:43010321-43011102	-	POLDIP3	-	NDUFC1	+	151168	RAB33B
UHNhscpg00014262	chr1:43683519-43684423	+	ABCG1	+	KIAA1649	+	3711	CYB5R3
UHNhscpg00014278	chr11:64216985-64217570	-	ZFP64	-	UMODL1	+	21566	HOAT4
UHNhscpg00014576	chr20:50719871-50720391	-	ZFP64	-	BC038767	+	105527	ZFP64
UHNhscpg0002409	chrX:128912982-128913773	-	SUGTIP	-	SALL4	-	47428	ZFP64
UHNhscpg00015607	chr9:33473226-33474250	-	MTOR	-	XPNPEP2	+	118	SASH3
UHNhscpg00016477	chr1:1322016-1323073	-	MTOR	-	NOL6	+	74151	BC009440
UHNhscpg00011113	chr4:4859613-4859967	-	KIF5A	+	ANGPTL7	+	10181	UBIAD1
UHNhscpg00015879	chr12:48299099-48299618	+	KIF5A	+	AK056081	+	1424	MSX1
UHNhscpg00015606	chr12:57957633-57958195	+	SNAPC1	+	NR111	+	57711	TMEM106C
UHNhscpg00014519	chr1:214159322-214159791	+	SNAPC1	+	DCTN2	+	20493	BC033961
UHNhscpg00016016	chr14:62228816-62229592	+	TIAM1	+	HIF1A	+	1494	PROX1
UHNhscpg0002174	chr12:45444911-45445553	-	TIAM1	-	DBX2	-	224210	syf14r
UHNhscpg0005059	chr21:32929750-32930539	+	ZNF518A	+	TIAM1	+	10849	FKSG42
UHNhscpg0001248	chr10:97890139-97890702	-	ID11	-	ZNF518A	-	1018	BC014150
UHNhscpg00000005	chr10:1094361-1094422	-	ID11	-	ID11	+	60760	BLNK
							574	BC046483

Column one shows array probe ID. Column two shows genome location including chromosome number and distance. Columns 3, 5 and 7 show the directionality of within, upstream and downstream strand. Columns 4, 6 and 9 show the gene symbols mapped to the array probes.

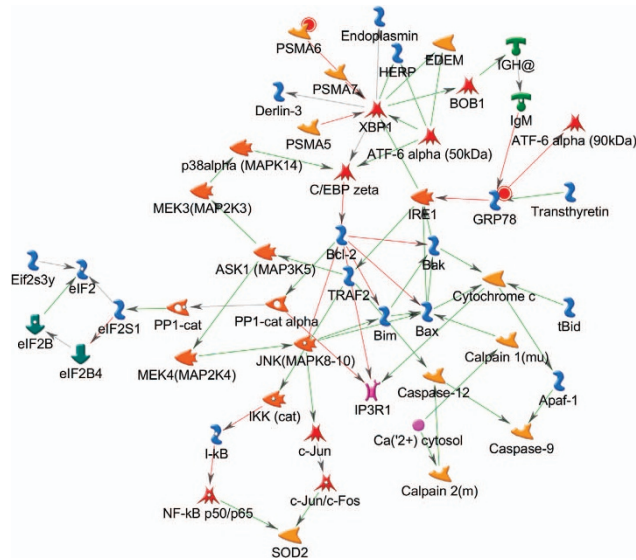


**Figure 3.** Top ranking gene networks. (a) The first score (by *P*-value) network from the study. Reference genes are marked with red circles. The network is centered on the *c-Myc* gene hub. *c-myc* is known as a proto-oncogene and is associated with cell proliferation and as a pro-apoptotic molecule. (b) The second scoring (by *P*-value) network from the study. The network includes 10 reference genes that are marked with red circles. The network is centered on the *PPARG* (human accelerated region 20 (HAR20)). Peroxisome proliferator-activated receptor-gamma (*PPAR-γ*) is a transcriptional coactivator involved in the regulation of energy metabolism and in mitochondrial biogenesis. Importantly, the gene has a role in the regulation of cAMP response element-binding protein (CREB). The transcription factor CREB has been implicated in signaling pathways relevant for pathogenesis and therapy of depression, which includes the *c-Jun* N-terminal kinases (JNKs). (c) Merging of the two top-ranking networks ranked by *P*-value reveal that the two central gene hubs (*c-MYC* and *PPAR-γ*) are linked by the activator protein (AP)-1-binding hub. The AP-1 is associated with dimeric transcription factors composed of Jun and Fos subunit. The AP-1 target is particularly relevant for depression as it regulates gene expression in response to different stimuli including cytokines. Together with the nuclear factor (NF)-kappaB, AP-1 controls T-cell activation, followed by binding of foreign antigens to the T-cell receptor leading to cytokine secretion. AP-1 therefore has a key role in the initiation inflammatory response by activating immune cells through expression and secretion of chemokines and cytokines. Inflammation is an event that has been associated with increased risk of major depressive episodes. Higher levels of peripheral inflammatory markers taken from blood samples have been found in depressed patients.

blood clearly comes from a very small subset of the total number of features. However, this may be expected, given that the data was obtained from a peripheral tissue as opposed to a more disease-related one.

A network analysis using MetaCore was performed in order to gain additional insight into the potential relationship between the feature-set uncovered from the above analysis. The top-ranking genes that could be mapped to probes on the 8.1-K array with higher probability of being selected across the different resampling were uploaded as reference molecule to MetaCore's database. We first explored the top two ranking networks

returned by *P*-value. The first ranking network ( $P < 1.29 \times 10^{-21}$ , *G*-score = 48.28) was centered on the *c-MYC* gene hub (Figure 3a). *c-MYC* is known as a proto-oncogene and is associated with cell proliferation and as a pro-apoptotic molecule. The second network by *P*-value ( $P < 1.62 \times 10^{-21}$ , *G*-score = 47.79) includes 10 reference genes and is centered on the *PPARGC1A* (human accelerated region 20 (HAR20)) gene hub (Figure 2b). The gene is a transcriptional coactivator involved in the regulation of energy metabolism and in mitochondrial biogenesis. Importantly, the gene has a role in the regulation of cAMP response element-binding protein. The transcription factor cAMP response element-



**Figure 4.** The highest-ranking network by G-score. The network reveals a number of potential interaction that may be relevant to the disease and have been previously associated with inflammation. The bottom of the network shows a cluster of genes including *c-Jun*, *Fos*, *AP-1* (together forming the APP complex), *JNK* (MAPK) and *NF- $\kappa$ B*. On the upper part, the network is clustered around the *XPB-1* gene hub, which is involved in modulation of the cellular response to endoplasmic reticulum (ER) stress brought about by stressful stimuli.

binding protein has been implicated in signaling pathways relevant for pathogenesis and is associated with the *c-Jun* N-terminal kinases. The network points at potential mechanisms involved in inflammation, which have been extensively associated with MDD. The two networks were subsequently merged in order to gain further insight into their potential relationship and interactions (Figure 3c). Interestingly, the gene hubs are one interaction away and modulated through a pathway centered on the AP-1 hub (*C-JUN*). *c-Jun*, in combination with *c-Fos*, forms the AP-1 early-response transcription factor. Activation is dependent on double phosphorylation by the *c-Jun* N-terminal kinase pathway, which has an important role in initiating inflammatory cellular responses.<sup>28</sup>

Lastly, we explored the top-ranking network by G-score (Figure 4). The network reveals a number of interesting genes including *c-Jun* (*AP-1*), with interactions between *c-Jun* N-terminal kinase (MAPK) and the X-box-binding protein 1 (*XPB-1*) gene. The *XPB1* gene encodes a key transcription factor in the unfolded protein stress response and has been implicated in the pathophysiology of MDD.<sup>29</sup>

## DISCUSSION

Landscapes of DNA methylation can show great variation between different tissues, but accessibility to brain tissues in human studies is limited to post-mortem and surgical resection. Therefore, peripheral tissues, including blood, can be an important source of information for the identification of biomarkers and mechanisms associated with the pathology that are assumed to also be manifested in the brain. However, disease-associated epivariations in blood may be less pronounced, more sparsely distributed and complicated to detect because of noise variance compared to disease-affected brain tissues. In this study we combined three non-i.i.d. data sets using a data-driven approach to uncover a suitable method to control for both batch effect and noise variance in order to build a global classifier that could classify affected siblings in MZ pairs discordant for MDD from epivariation in peripheral blood.

Of the range of methods evaluated for non-i.i.d. data consolidation, PCA with the top three eigenvector methods and mixture approaches using PCA and ComBat showed steady positive results on our data sets. The resulting model could classify the disease state above chance levels in an independent testing set using differences in methylation levels. Out of several methods considered for classification, the linear support vector machine with the top 100 selected features returned the highest classification accuracy. The accuracy prediction above chance level can be considered important, given the complex molecular architecture underpinning the pathology. Accuracy measures were reported instead of receiver operating characteristic scores, even if these can often be higher. Indeed, high sensitivity and specificity can be achieved even in the absence of high accuracy. Being prudent in quantifying the ability to discriminate discordant twins for MDD from epigenetic markers in peripheral blood alone was preferred to reporting higher ROC scores. This is particularly important in the context of disease classification where it is likely that accumulation of epigenetic changes may explain only a small fraction of individual differences in the pathology. Enrichment analysis of the top-ranking features points at an epigenetic signature of MDD that can be detected in peripheral blood and may be used to inform candidate gene selection in future molecular studies of MDD.

The results of the analysis identified a number of key genes, which have previously been shown to be involved in the pathophysiology of MDD. These genes include *PPAR- $\gamma$*  (peroxisome proliferator-activated receptor-gamma), *AP-1*, *XPB-1* and *NF- $\kappa$ B*. One of the gene maps uncovered from this study focused around *PPAR- $\gamma$*  as the center gene hub (Figure 3b). Merging the two top-ranking networks showed that *AP-1* appears to be a connection molecule mediating between the two gene hubs of the two networks; these two top-ranking networks by *P*-value appear to be one interaction away. It is possible that DNA methylation may modulate a number of regulatory processes affecting these networks. The precise mechanisms by which these epigenetic changes may affect the pathophysiology of depression or whether these can be used as potential biomarkers from peripheral fluids needs further exploration.

### *AP-1* and *NF- $\kappa$ B*

The results from this study support previous human and animal studies on MDD that have uncovered genes centered around a stress-response cascade involving the activator protein 1 (*AP-1*) and nuclear factor kappa-light-chain-enhancer of activated B cells (*NF- $\kappa$ B*).<sup>30</sup> *AP-1* is a gene that codes for a transcription factor that regulates gene expression in response to cytokines as well as environmental stress and bacterial and viral infections. *AP-1* downregulation has been implicated as part of the mechanism by which administration of IFN-alpha therapy induces depression symptoms.<sup>31</sup> *NF- $\kappa$ B* is a transcription regulator that has a role in peripheral inflammation with both pro- and anti-inflammatory effects.<sup>32,33</sup>

### *PPAR- $\gamma$*

One of the loci uncovered by the network pathway analysis is the *PPAR- $\gamma$*  gene that codes for the glitazone receptor (NR1C3). Activation of the *PPAR- $\gamma$*  system in the central nervous system is thought to decrease parainflammation, endoplasmic reticulum (ER) stress, formation of reactive oxygen species and glutamate toxicity while increasing neurogenesis and neuroplasticity.<sup>34</sup> It has been shown that activation of *PPAR- $\gamma$*  leads to an increased neurogenesis as well as antidepressant effects in rodent models.<sup>35,36</sup> While the mechanisms by which this leads to an antidepressant effect are unclear, it has been shown that exposure to stress is associated with decreased hippocampal neurogenesis. Induction of neurogenesis has been shown to be a mechanism of



action of several antidepressant medications in animal models.<sup>37–39</sup>

Interestingly, several clinical cross-over and randomized control trials have evaluated the efficacy of the insulin-sensitizing PPAR- $\gamma$  agonists thiazolidinediones (troglitazone, pioglitazone and rosiglitazone) for the treatment of patients with concomitant MDD and metabolic syndrome or diabetes, as adjunctive therapy in patients with moderate-to-severe MDD in the absence of other metabolic disorders and as monotherapy.<sup>40–43</sup> In addition to their action as insulin-sensitizing agents, these drugs also have anti-inflammatory, neuroprotective and anti-excitotoxic properties.<sup>44</sup> Activation of PPAR- $\gamma$  receptors by their natural (15d-PGJ2) or synthetic ligands has been shown to support neuronal glucose and glutamate metabolism following exposure to stress and to increased levels of neurotrophic factors.<sup>45,46</sup> These processes are believed to be dysregulated in MDD, thus making PPAR- $\gamma$  activation a drug target of interest.

#### XPB-1

The top-ranking network by G-score revealed a network with the *XPB-1* as a principal hub (Figure 4). *XPB-1* is involved in the cellular response to ER stress brought about by stressful stimuli. ER stress occurs when demands on the ER to fold and process proteins are increased beyond its capacity, leading to the production of unfolded proteins.<sup>47</sup> This initiates the ER-stress response, which promotes protein folding and secretion, as well as unfolded protein degradation. Failure of this response leads to calcium ATPase (Ca2+) release from the ER, leading to cell apoptosis. XPB-1 activates unfolded protein degradation as well as production of chaperones for protein secretion.<sup>34</sup> A polymorphism leading to decreased *XPB-1* expression has also been implicated in bipolar disorder.<sup>48</sup> Sodium valproate, an anticonvulsant with mood-stabilizing properties, affects the ER-stress response, reducing the likelihood of cell apoptosis. One mechanism of action for this is the activation of *ATF6*, the gene that activates *XPB-1* in the hippocampus and the cerebral cortex.<sup>49,50</sup>

#### Strengths and limitations

DNA methylation is known to exhibit tissue specificity, but human studies are limited to either post-mortem, surgical resection or often surrogate tissues, which include blood cells. The extent to which epivariations in surrogate tissues resemble those in central tissues of interest is unknown. Given the complexity of the pathology, it is likely that many epigenetic variants involved in a number of mechanisms underpinning MDD cannot be detected in blood. On the other hand, blood cells may show fewer epigenetic changes associated with factors associated with the disorder (such as in response to pharmacological treatment) but not involved in etiology. According to the peripheral inflammation hypothesis of depression, it is systemic inflammation that could potentially explain part of the molecular etiology of the disorder. Many of the inflammatory molecules implicated by this hypothesis are detectable in peripheral blood cells. The most predictive epigenetic markers uncovered by this study were mapped to genes and gene networks with mechanisms associated to inflammation, which is reasonable to expect from blood. It is, however, unclear whether peripheral inflammation is involved in the etiology of depression or is purely driven in response to it.

This study was able to control for the non-i.i.d. nature of the data, allowing the integration of three separate data sets, each from a different center and country, overcoming one of the limitations from a previous published study using the same three data sets. In isolation, the studies were too underpowered to allow classification of affected and non-affected siblings, but leveraging the increased power of the larger sample allowed detection of a clear signal. Independent testing and training and careful optimization of parameters to avoid model over-fitting increased

the probability of true-positive findings. Although classification was possible above chance levels from peripheral blood, classification accuracy remains generally low. On the other hand it is likely that only a small portion of the phenotypic variance is explained by epigenetic variations. However, this approach allowed us to explore the etiology of MDD beyond gene + environment and gene  $\times$  environment etiological paradigms.

This method of modeling allows the identification of epigenetic loci that account for a portion of the variance in phenotype between MZ twin pairs discordant for depression. Whereas these loci can be identified, this methodology does not give any information as to their role. The construction of networks of epigenetic loci identified from our analysis revealed which epigenes are linked, suggesting pathways and relationships between the proteins these genes code for. The genes implicated by the model can then be corroborated with the literature, shedding light on their role in MDD pathophysiology, although this is inevitably subjected to bias. Furthermore, the majority of the literature does not look at the role of specific genes from the epigenetic perspective; therefore, there is potential for further research involving genes both implicated in the pathophysiology of depression, such as the peripheral inflammation hypothesis, and the epigenetics of depression.

Many of the genes identified in the networks of this analysis are yet to be included in any literature regarding the pathophysiology of depression. *C-Myc* is a good example of this (Figure 3a). It is an oncogene implicated in the pathophysiology of a number of cancers, but it has not yet been researched extensively in the field of psychiatry.

This study employs a cross-sectional design, meaning that it is not possible to infer causality. Accessibility of peripheral blood samples from the same MZ twin pairs over time could allow to identify changes in DNA methylation in relation to time of onset of depression and potentially identify a causal link.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGMENTS

Dr Malki is supported by an MRC grant (G9817803).

#### REFERENCES

- 1 Belmaker RH, Agam G. Major depressive disorder. *N Engl J Med* 2008; **358**: 55–68.
- 2 Mullins N, Perroud N, Uher R, Butler AW, Cohen-Woods S, Rivera M et al. Genetic relationships between suicide attempts, suicidal ideation and major psychiatric disorders: a genome-wide association and polygenic scoring study. *Am J Med Genet B* 2014; **165**: 428–437.
- 3 Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. *Diabetes Care* 2008; **31**: 2383–2390.
- 4 Perroud N, Aitchison KJ, Uher R, Smith R, Huezio-Diaz P, Marusic A et al. Genetic predictors of increase in suicidal ideation during antidepressant treatment in the gendep project. *Neuropsychopharmacology* 2009; **34**: 2517–2528.
- 5 Van Praag HM. Depression, suicide and the metabolism of serotonin in the brain. *J Affect Disord* 1982; **4**: 275–290.
- 6 Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2014; **157**: 1552–1562.
- 7 Kendler KS, Gardner CO, Neale MC, Prescott CA. Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? *Psychol Med* 2001; **31**: 605–616.
- 8 McGuffin P, Katz R, Watkins S, Rutherford J. A hospital-based twin register of the heritability of dsm-iv unipolar depression. *Arch Gen Psychiatry* 1996; **53**: 129–136.
- 9 Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.
- 10 Malki K, Keers R, Tosto MG, Lourdasamy A, Carboni L, Domenici E et al. The endogenous and reactive depression subtypes revisited: integrative animal and

- human studies implicate multiple distinct molecular mechanisms underlying major depressive disorder. *BMC Med* 2014; **12**: 73.
- 11 Schroeder M, Krebs MO, Bleich S, Frieling H. Epigenetics and depression: current challenges and new therapeutic options. *Curr Opin Psychiatry* 2010; **23**: 588–592.
  - 12 Dalton VS, Kolshus E, McLoughlin DM. Epigenetics and depression: return of the repressed. *J Affect Disord* 2014; **155**: 1–12.
  - 13 Murgatroyd C, Patchev AV, Wu Y, Micale V, Bockmühl Y, Fischer D et al. Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nat Neurosci* 2009; **12**: 1559–1566.
  - 14 McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonté B, Szyf M et al. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci* 2009; **12**: 342–348.
  - 15 Castillo-Fernandez JE, Spector TD, Bell JT. Epigenetics of discordant monozygotic twins: implications for disease. *Genome Med* 2014; **6**: 60.
  - 16 Oh G, Wang S-C, Pal M, Chen ZF, Khare T, Tochigi M et al. Dna modification study of major depressive disorder: beyond locus-by-locus comparisons. *Biol Psychiatry* 2015; **77**: 246–255.
  - 17 Mill J, Petronis A. Molecular studies of major depressive disorder: the epigenetic perspective. *Mol Psychiatry* 2007; **12**: 799–814.
  - 18 Nielsen TO, West RB, Linn SC, Alter O, Knowling MA, O'Connell JX et al. Molecular characterisation of soft tissue tumours: a gene expression study. *Lancet* 2000; **359**: 1301–1307.
  - 19 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–909.
  - 20 Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012; **28**: 882–883.
  - 21 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics* 2007; **8**: 118–127.
  - 22 Gentleman R, Carey V, Huber W, Hahne F. Genefilter: Genefilter: Methods for Filtering Genes from Microarray Experiments. *R package version, 1* 2007. <http://bioconductor.org/packages/release/bioc/html/genefilter.html>.
  - 23 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res* 2015; **43**: e47.
  - 24 Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 2004; **573**: 83–92.
  - 25 Hong F, Wittner B, Breitling R, Smith C, Battke F et al. Rankprod: Rank Product Method for Identifying Differentially Expressed Genes with Application in Metaanalysis. *R version, 2* 2009. <https://bioconductor.org/packages/release/bioc/html/RankProd.html>.
  - 26 Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; **20**: 273–297.
  - 27 Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell* 1998; **20**: 832–844.
  - 28 Raison CL, Capuron L, Miller AH. Cytokines sing the blues: inflammation and the pathogenesis of depression. *Trends Immunol* 2006; **27**: 24–31.
  - 29 Grunebaum MF, Galfalvy HC, Y-y Huang TB, Cooper AK, Burke M, Agnello MA et al. Association of x-box binding protein 1 (xbp1) genotype with morning cortisol and 1-year clinical course after a major depressive episode. *Int J Neuropsychopharmacol* 2009; **12**: 281–283.
  - 30 Malki K, Pain O, Tosto MG, Du Rietz E, Carboni L, Schalkwyk LC. Identification of genes and gene pathways associated with major depressive disorder by integrative brain analysis of rat and human prefrontal cortex transcriptomes. *Transl Psychiatry* 2015; **5**: e519.
  - 31 Felger JC, Cole SW, Pace TWW, Hu F, Woolwine BJ, Doho GH et al. Molecular signatures of peripheral blood mononuclear cells during chronic interferon- $\alpha$  treatment: relationship with depression and fatigue. *Psychol Med* 2012; **42**: 1591–1603.
  - 32 Felger JC, Lotrich FE. Inflammatory cytokines in depression: neurobiological mechanisms and therapeutic implications. *Neuroscience* 2013; **246**: 199–229.
  - 33 Lawrence T. The nuclear factor  $\kappa$ -b pathway in inflammation. *Cold Spring Harb Perspect Biol* 2009; **1**: a001651.
  - 34 Gold PW, Licinio J, Pavlatou MG. Pathological parainflammation and endoplasmic reticulum stress in depression: potential translational targets through the cns insulin, klotho and ppar- $\gamma$  systems. *Mol Psychiatry* 2013; **18**: 154–165.
  - 35 Sadaghiani MS, Javadi-Paydar M, Gharedaghi MH, Fard YY, Dehpour AR. Antidepressant-like effect of pioglitazone in the forced swimming test in mice: the role of ppar-gamma receptor and nitric oxide pathway. *Behav Brain Res* 2011; **224**: 336–343.
  - 36 Cimini AM, Cerù MP. Emerging roles of peroxisome proliferator-activated receptors (ppars) in the regulation of neural stem cells proliferation and differentiation. *Stem Cell Rev* 2008; **4**: 293–303.
  - 37 Malberg JE, Eisch AJ, Nestler EJ, Duman RS. Chronic antidepressant treatment increases neurogenesis in adult rat hippocampus. *J Neurosci* 2000; **20**: 9104–9110.
  - 38 Malki K, Lourdasamy A, Binder E, Paya-Cano J, Sluyter F, Craig I et al. Antidepressant-dependent mrna changes in mouse associated with hippocampal neurogenesis in a mouse model of depression. *Pharmacogenet Genomics* 2012; **22**: 765–776.
  - 39 Surget A, Tanti A, Leonardo ED, Laugeray A, Rainer Q, Touma C et al. Antidepressants recruit new neurons to improve stress response regulation. *Mol Psychiatry* 2011; **16**: 1177–1188.
  - 40 Sepanjnia K, Modabbernia A, Ashrafi M, Modabbernia M-J, Akhondzadeh S. Pioglitazone adjunctive therapy for moderate-to-severe major depressive disorder: randomized double-blind placebo-controlled trial. *Neuropsychopharmacology* 2012; **37**: 2093–2100.
  - 41 Lin KW, Wroolie TE, Robakis T, Rasgon NL. Adjuvant pioglitazone for unremitted depression: Clinical correlates of treatment response. *Psychiatry Res* 2015; **230**: 846–852.
  - 42 Kashani L, Omidvar T, Farazmand B, Modabbernia A, Ramzanzadeh F, Tehrani-nejad ES et al. Does pioglitazone improve depression through insulin-sensitization? results of a randomized double-blind metformin-controlled trial in patients with polycystic ovarian syndrome and comorbid depression. *Psychoneuroendocrinology* 2013; **38**: 767–776.
  - 43 Kemp DE, Ismail-Beigi F, Ganocy SJ, Conroy C, Gao K, Obral S et al. Use of insulin sensitizers for the treatment of major depressive disorder: a pilot study of pioglitazone for major depression accompanied by abdominal obesity. *J Affect Disord* 2012; **136**: 1164–1173.
  - 44 Delerive P, Fruchart J-C, Staels B. Peroxisome proliferator-activated receptors in inflammation control. *J Endocrinol* 2001; **169**: 453–459.
  - 45 Garca-Bueno B, Caso JR, Pérez-Nievas BG, Lorenzo P, Leza JC. Effects of peroxisome proliferator-activated receptor gamma agonists on brain glucose and glutamate transporters after stress in rats. *Neuropsychopharmacology* 2007; **32**: 1251–1260.
  - 46 Toyomoto M, Ohta M, Okumura K, Yano H, Matsumoto K, Inoue S et al. Prostaglandins are powerful inducers of ngf and bdnf production in mouse astrocyte cultures. *FEBS Lett* 2004; **562**: 211–215.
  - 47 Winnay JN, Boucher J, Mori MA, Ueki K, Kahn CR. A regulatory subunit of phosphoinositide 3-kinase increases the nuclear accumulation of x-box-binding protein-1 to modulate the unfolded protein response. *Nat Med* 2010; **16**: 438–445.
  - 48 Kakiuchi C, Iwamoto K, Ishiwata M, Bundo M, Kasahara T, Kusumi I et al. Impaired feedback regulation of xbp1 as a genetic risk factor for bipolar disorder. *Nat Genet* 2003; **35**: 171–175.
  - 49 Chen B, Wang JF, Young LT. Chronic valproate treatment increases expression of endoplasmic reticulum stress proteins in the rat cerebral cortex and hippocampus. *Biol Psychiatry* 2000; **48**: 658–664.
  - 50 Bown CD, Wang J-F, Chen B, Young LT. Regulation of er stress proteins by valproate: therapeutic implications. *Bipolar Disord* 2002; **4**: 145–151.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016