

ARTICLE OPEN



Immunoglobulin genes expressed in lymphoblastoid cell lines discern and predict lithium response in bipolar disorder patients

Liron Mizrahi^{1,7}, Ashwani Choudhary^{1,7}, Polina Ofer¹, Gabriela Goldberg², Elena Milanese³, John R. Kelsoe⁴, David Gurwitz⁵, Martin Alda⁶, Fred H. Gage² and Shani Stern¹✉

© The Author(s) 2023

Bipolar disorder (BD) is a neuropsychiatric mood disorder manifested by recurrent episodes of mania and depression. More than half of BD patients are non-responsive to lithium, the first-line treatment drug, complicating BD clinical management. Given its unknown etiology, it is pertinent to understand the genetic signatures that lead to variability in lithium response. We discovered a set of differentially expressed genes (DEGs) from the lymphoblastoid cell lines (LCLs) of 10 controls and 19 BD patients belonging mainly to the immunoglobulin gene family that can be used as potential biomarkers to diagnose and treat BD. Importantly, we trained machine learning algorithms on our datasets that predicted the lithium response of BD subtypes with minimal errors, even when used on a different cohort of 24 BD patients acquired by a different laboratory. This proves the scalability of our methodology for predicting lithium response in BD and for a prompt and suitable decision on therapeutic interventions.

Molecular Psychiatry (2023) 28:4280–4293; <https://doi.org/10.1038/s41380-023-02183-z>

INTRODUCTION

Bipolar disorder (BD) is a chronic, heritable neuropsychiatric disorder with alternating episodes of mania and depression in patients [1–3]. With a prevalence of 1–2 % in the general population, BD is also a notable cause of disability and premature deaths due to suicide or co-morbidities [2, 4]. Clinical management of BD is dependent on lithium and classes of antipsychotic and anticonvulsant drugs [5]. Due to its effectiveness, lithium is used as a first-line treatment for BD patients [3, 6, 7]. However, more than half of BD patients respond to lithium treatment inadequately; therefore, BD patients have been classified and characterized as lithium responders (LRs) and non-responders (NRs) [7, 8]. The presence of subpopulations among BD patients further hints at the confounding genetic complexity and heterogeneity among the patients. Some studies have also pointed toward the neurodevelopmental origin of BD, although the clinical symptoms are only visible in adulthood [8]. Genome-wide association studies (GWAS) in BD patients have identified genomic loci associated with BD but, in the absence of functional validation, the causative genes are obscure [9–13].

Animal models for BD are imperfect mainly due to the unknown genetic component and inability to mimic the extreme mood shifts observed in patients [14]. Disease modeling via induced pluripotent stem cell (iPSC) technology has opened a new possibility for studying psychiatric diseases by allowing researchers to investigate patient-specific neural cells in vitro in 2D and 3D culture systems [12]. Recently, researchers have employed iPSC

technology to study BD patient-specific neurons to understand the cellular pathophysiology and mechanisms underlying BD [15–17]. In 2018, we found that patch clamp recordings of neurons derived from BD patients could be used to predict lithium response with a very low error rate [16]. While these investigations were expensive and well-trained personnel were needed to perform neuronal differentiation and electrophysiological experiments, they showed a proof of concept that the prediction of lithium response was indeed feasible with a low error rate. In addition to iPSCs, LCLs have been used to study drug response in BD. LCLs are immortal lymphoblastoid cell lines derived by the transformation of B lymphocytes by Epstein-Barr virus (EBV) [12, 18]. LCLs can expand indefinitely and are low-cost, patient-specific polyclonal cells as compared to patient-specific iPSCs, which are generally monoclonal and require higher cost and expertise to study [18].

Poor understanding of the etiology and genetics of BD also makes its diagnosis difficult, in turn affecting the therapy received by the patients. Hence, appropriate biomarkers are necessary to diagnose BD correctly as well as predict the patients' response to lithium [19, 20]. Previous studies have found serum and plasma levels of BDNF; the altered response of lymphocytes to glucose deprivation; and oxidative stress markers to be potential biomarkers of BD [21–23]. Gene expression analysis using DNA microarrays from whole blood RNA of BD patients has found some candidate genes that are especially related to myelination or growth factor signaling [24]. While iPSCs offer an excellent system

¹Sagol Department of Neurobiology, Faculty of Natural Sciences, University of Haifa, Haifa 3498838, Israel. ²Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA. ³Victor Babes National Institute of Pathology, Bucharest 050096, Romania. ⁴Department of Psychiatry, University of California, San Diego, La Jolla, CA 92093, USA. ⁵Department of Human Molecular Genetics and Biochemistry, Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel. ⁶Department of Psychiatry, Dalhousie University, Halifax, NS B3H 2E2, Canada. ⁷These authors contributed equally: Liron Mizrahi, Ashwani Choudhary. ✉email: sstern@univ.haifa.ac.il

Received: 13 September 2022 Revised: 3 July 2023 Accepted: 6 July 2023

Published online: 24 July 2023

to study the cellular pathology and molecular mechanisms of the disease, due to their inexpensive nature LCLs are a good source to look for genetic biomarkers, especially in a large sample size including BD patient subtypes [24]. There are hypotheses linking an immune response with psychiatric disorders [25, 26], further suggesting that LCLs will have a predictive value. Earlier studies using LCLs from BD patients have reported altered calcium signaling and lower ER stress response in patients as compared to controls [12, 27, 28]. Subsequently, several studies have reinforced the idea of conducting gene expression analysis in the LCLs of BD patients to identify potential transcriptomic biomarkers [29–32].

Recently, machine-learning (ML) algorithms have found widespread applications in the field of healthcare [16, 33, 34], with the potential to provide effective support in clinical management through disease prediction, detection, and diagnosis [35]. Supervised classification algorithms are a type of ML approach that can be utilized as significant tools in the diagnosis and treatment decisions in complex disorders such as BD [36, 37]. We hypothesized that by analysing transcriptomic datasets with supervised classification algorithms, patterns of gene expression and biomarkers can be identified that are associated with BD and its subtypes. This information can be further used to train models and thus can be utilized for the prediction of treatment response in BD.

In this study, we have performed RNA-seq of LCLs from control and BD patients, including LR and NRs, and thereby report the analysis of DEGs to find appropriate genetic signatures that can be used as potential biomarkers in BD. We have also incorporated the RNA-seq data previously published by Milanese et al. [38]. for BD LR and BD NR patient subtypes and analysed them jointly to add more statistical power to our results. Importantly, we trained supervised classification algorithms on the RNA-seq data, which predicted the lithium response from the RNA-seq data acquired by a different laboratory with a very low error rate (Fig. 1). Such validation is necessary since we are seeking global biomarkers that will work in algorithms that are trained in one laboratory and predictions need to be made based on data that are acquired elsewhere. Our study also found convergence in altered genes with earlier studies, providing insights into the important biological pathways affected in BD.

METHODS

Ethics approval

All participants signed informed consent. The study was approved by the Research Ethics Board of the Nova Scotia Health Authority, Canada. The participants were diagnosed with BD and classified into subtypes according to the lithium response by the psychiatrists as described previously [16].

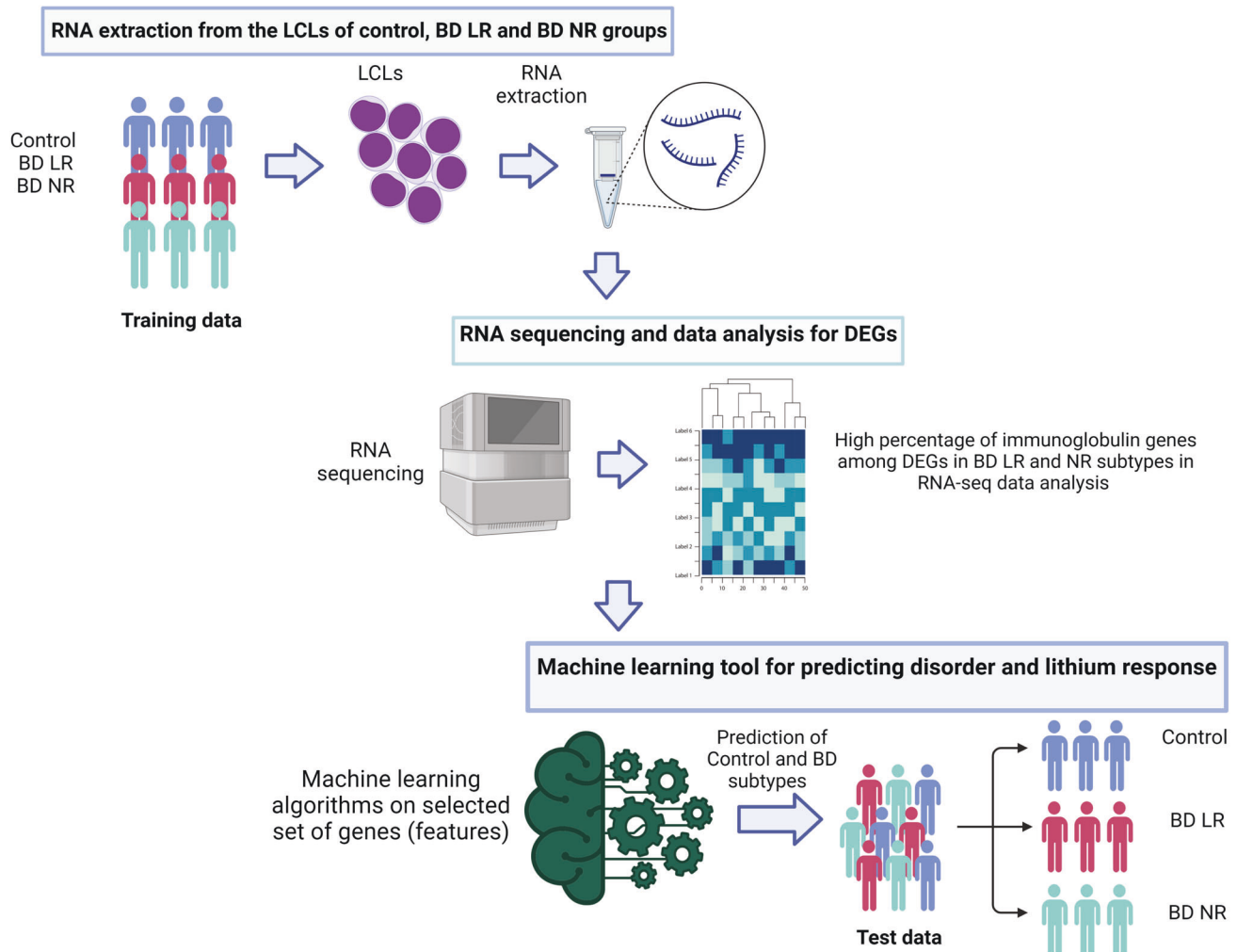


Fig. 1 Graphical representation of the workflow and the experimental design. RNA sequencing was performed from the LCLs of Control and BD subtypes from the train data to find genetic signatures using machine learning algorithms to discern between Control, BD, and its subtypes and used it for classification of test data (Lithium Responsive (LR) and Lithium Non-responsive (NR)). The graphical image was created with BioRender.com.

Cell culture and RNA extraction from LCLs

LCLs were generated from the PBMCs of all the patients and controls (see details about the cohorts below) as described previously [39]. Briefly, PBMCs were isolated from the blood using BD Vacutainer® CPT™ Cell Preparation Tubes with Sodium Citrate (Cat no: 362761) following the manufacturer's protocol. Approximately, 2 million cells were infected with the EBV produced from the B95-8 cell line as mentioned in detail previously [39, 40]. The transformed cells formed aggregates within a week of EBV infection and formed larger aggregates with time. These cells were further passaged 3–4 times to ascertain the establishment of lymphoblastoid lines (Supplementary Fig 1). The LCLs were cultured in T25 and T75 tissue culture flasks in complete RPMI medium containing RPMI 1640 (Biological Industries, Cat no: 01-100-1A), 1X Anti-Anti (ThermoFisher Scientific, Cat no: 15240062) 1% Glutamax (ThermoFisher Scientific, Cat no: 35050061), 1% Sodium pyruvate (ThermoFisher Scientific, 11360070) and 15% heat-inactivated FBS (Sigma, Cat no: F9665) as described previously [40, 41] with media changes on alternative days. To ascertain the confluency, the LCLs were counted using Trypan blue exclusion assay with Bio-Rad TC20™ automated cell counter following the manufacturer's protocol and were passaged regularly when they reached a density of 200,000 cells/ml. All the cell lines used were tested using a Hy-Mycoplasma PCR kit (Hylabs, Cat No. KI 5034I) and found to be free of mycoplasma contamination.

For RNA extraction, approximately 10 million cells were collected in 1 ml of TRIzol™ reagent (ThermoFisher Scientific, Cat no: 15596026) and placed on ice for 5 min before storing at -80°C . Total RNA was isolated using the Zymo Quick RNA kit (Zymo Research, Cat no: R1054) according to the manufacturer's protocol. The RNA was further purified using the RNA clean and concentrator kit (Zymo Research, Cat no: R1013) following the manufacturer's instructions. The RNA quality and integrity were checked using an ND-1000 Nanodrop spectrophotometer (ThermoFisher Scientific) and Tape station 2200 (Agilent). All the RNA samples sent for sequencing had RNA integrity number values > 9 .

First cohort

We performed RNA sequencing from the LCLs of 9 LR (Lithium responsive) BD patients (15 sets of RNA samples), 10 BD NR (Lithium Non-responsive) patients (17 sets of RNA samples), and 10 control individuals (14 sets of RNA samples) grown and sequenced in 3 batches. Some of the samples had replicates and hence a total of 46 sets of RNA were sequenced and analyzed. (Details about the samples in the Github file link <https://github.com/Precision-Disease-Modeling-Lab/Lithium-Respose-Predictor/blob/main/ML/LRvsNR/samples.csv>). 24 samples (5 LR in duplicates, 4 NR in duplicates, 3 control in duplicates) in Batch 1; 10 samples (1 LR in duplicate, 2 NR in duplicate, 2 control in duplicate) in Batch 2; and 12 samples in Batch 3 (3 LR-no replicate, 1 NR in duplicate plus 4 NR, 4 control-no replicate). We have used batch correction to account for the differences that arise due to the batch effect as described below.

Second cohort

The data for the second cohort were obtained from the published work of Milanesi et al. [38], and consisted of 12 BD LR (12 sets of RNA samples—no replicates) and 12 BD NR (12 sets of RNA samples—no replicates) patients for which the RNA-sequencing was performed from LCLs. The LCLs used by Milanesi et al. [38], were also generated using a similar method i.e. by infecting the PBMCs with Epstein Barr Virus produced from the B95-8 cell line. Similarly, a batch correction algorithm was performed as described below.

RNA sequencing, analysis, and batch correction

Libraries for the LCL RNAs from BD and control subjects were prepared using a TruSeq RNA Library Prep Kit v2 (Illumina) following the manufacturer's instructions. Quality control of the raw FASTQ files was performed using FastQC [42] (v0.11.5). Sequencing reads were aligned to the human genome (GRCh38.104) and quantified using STAR [43] (v2.7.9a). The R-based Bioconductor package DESeq2 [44] (v1.34.0) was used to perform differential gene expression analysis. The experimental design was modeled to consider batch and condition (~batch + condition). Batch-effect was accounted for by either including batch as a covariate in the linear model (allowing for correct estimation of degrees of freedom) or by using the remove batch effect function from the limma [45] package in R for Principal Component Analysis (PCA) and correlation analysis. This

function uses linear modeling to estimate and reduce batch effects from gene expression data.

To account for the false positives when identifying DEGs, we performed the false discovery Rate (FDR) analysis. In DESeq2, the FDR is calculated using the Benjamini-Hochberg (BH) procedure, which adjusts the p-values of the hypothesis tested for multiple comparisons [46, 47]. In this study, for each pairwise comparison, genes that reached BH false discovery rate (FDR) ($p\text{-Adjusted}$) < 0.05 and $\log_2\text{fold-change} \geq 1$ were considered significant DEGs.

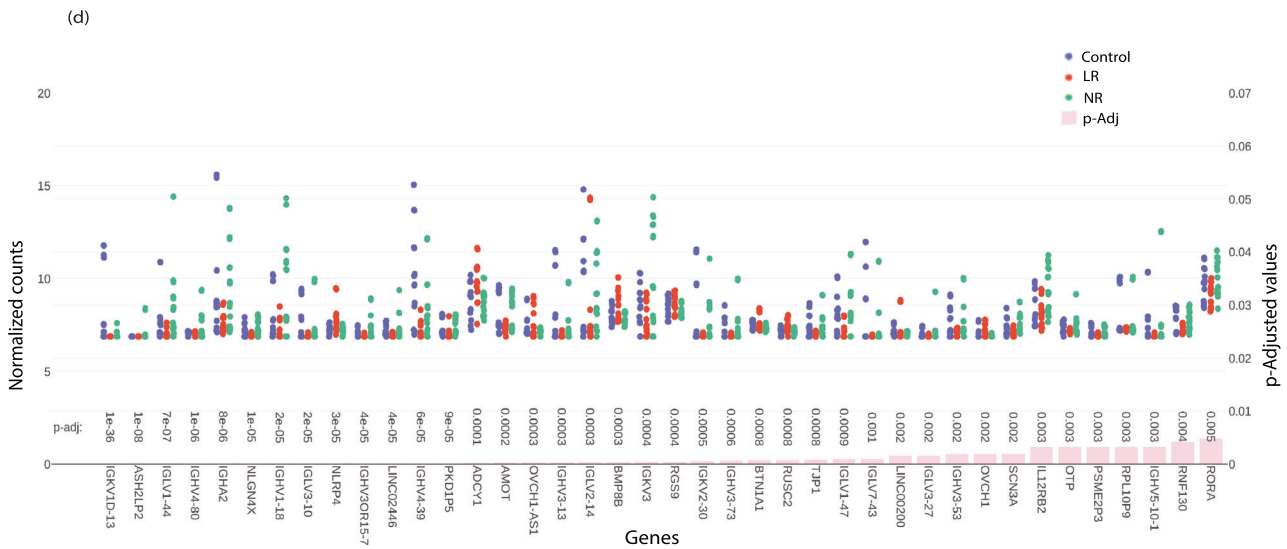
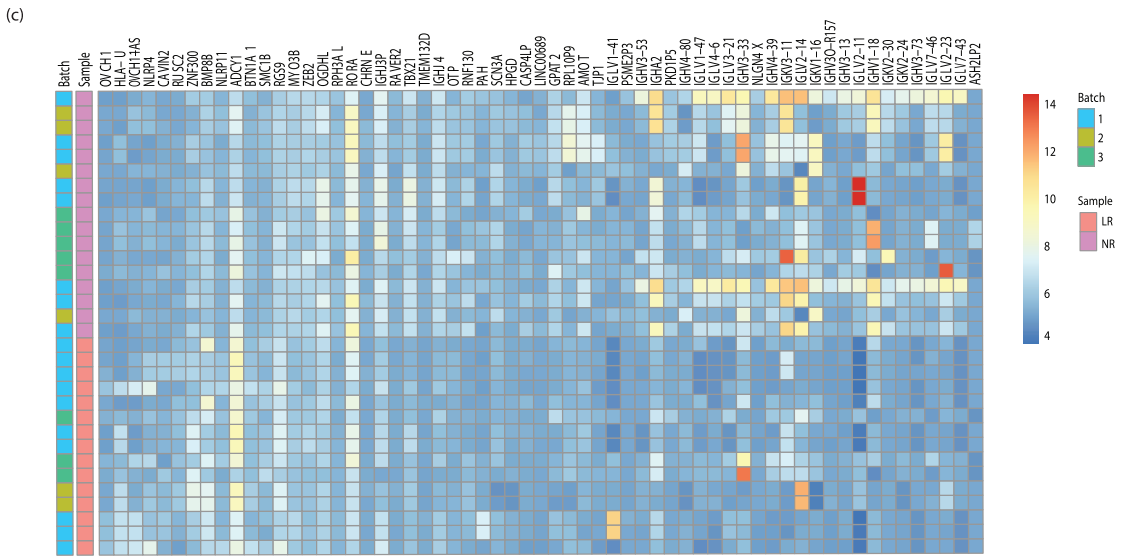
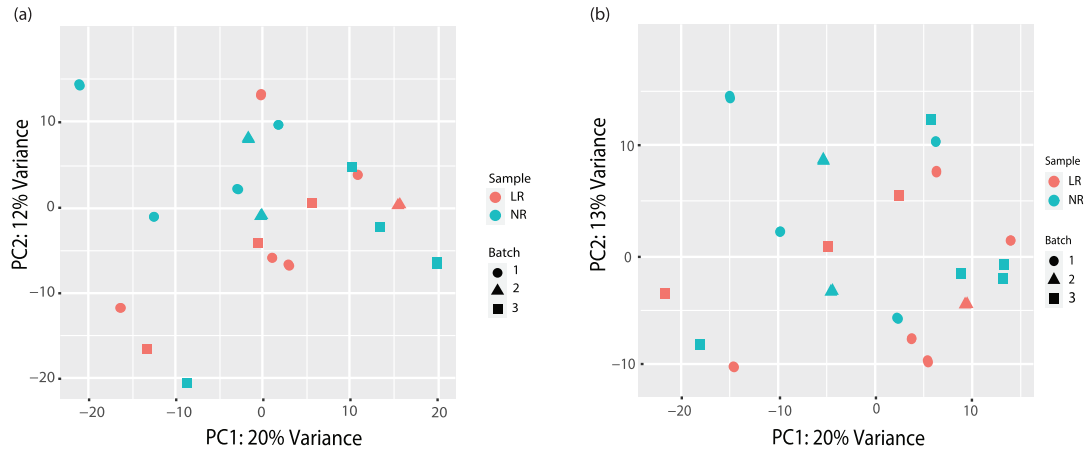
Machine learning predictor analysis

The ML predictor analysis was performed using Python-based packages. For the predictor analysis, we used PyComBat [48] (Python-based package) to correct for the batch effect in the RNA sequencing data (that was different from the R-based package for batch correction used during DEG analysis as mentioned above). PyComBat uses the empirical Bayes method to adjust for the batch effects [48]. The empirical Bayes method is effective even with smaller sample sizes (> 10) [49].

We evaluated the following five supervised classifiers: Logistic regression (Lr), Random Forest (RF), K-Nearest Neighbors (K-NN), Support Vector Machine (SVM), and Neural Network (NN). Among the classifiers selected for predictor analysis, RF is a well-known ensemble learning algorithm that combines the predictions of multiple decision trees trained on randomly selected subsets of the training data and features [35, 36, 50]. It can achieve high accuracy for a wide range of classification tasks, but may not be the best choice for applications requiring interpretability, scalability, or fast training times [35, 51–53]. For such applications, Lr, SVM, and NN may be better suited depending on the specific requirements [52, 53]. K-NN is a simple and non-parametric supervised learning algorithm that makes predictions based on the proximity and similarity to other data points in the training set [54, 55]. However, it can be computationally expensive and sensitive to the number of neighbor data points (or K parameter) [55]. NN can handle complex problems and a wide range of data types including high dimensional datasets but can also be computationally expensive, difficult to interpret, and prone to overfitting [56, 57]. Lr, another important classifier, is particularly useful for binary classification where interpretability and simplicity are important [35, 57]. However, it may not perform as well as other complex models such as NN or RF for tasks that involve complex nonlinear relationships. SVM can be used for high-dimensional data and nonlinear relationships but can be computationally expensive and sensitive to hyper-parameters required for learning [35, 55].

To avoid overfitting, a feature selection was performed to reduce the number of genes used for training the classifier. To this end, a Mann-Whitney U test was performed between the two groups using the datasets (batches 1–4). (The dataset had 39,000 genes. We dropped the genes with less than 10 counts after which a total of 22,000 genes were kept for analysis). Twenty genes (out of the total 22,000 genes) with the smallest p-value were selected (Supplementary Fig 5,6). An exhaustive search with all gene combinations was then performed, and the subset of genes with the best chance to be used as a feature list in Lr was identified. From all gene combinations, only those with an accuracy score above 0.93 were selected. The 7 potential and promising genes (for LR vs. NR) and 5 promising genes (for BD vs. CTRL), which constantly appeared in the gene combination with high accuracy scores, were chosen as features (see Supp. Fig. 5b & 6b the prediction accuracy as a function of the number of features selected out of the chosen 20 genes). The selected features were used to train the classifiers using another Python-based package, sci-kit-learn [58]. We then performed cross-validation by splitting the data randomly into a 50% train and a 50% test, and the classification was repeated 50 times with these random selections (BD vs. CTRL and LR vs. NR). For BD vs. control, a total of 29 subjects (19 BD + 10 controls) from the original dataset were used. For LR vs. NR, a total of 43 subjects (19 BD subjects from the first cohort and 24 BD subjects from the second cohort were used). In addition to the cross-validation method which uses a shuffle train/test split from the entire dataset, we also performed a predictor evaluation for LR vs. NR by training the model with our in-house (original) datasets (the first 3 batches) and performing the classification on the second cohort (the different dataset from another laboratory) labeled as the batch 4 dataset [38].

To further evaluate the validity of this approach, we tested the performance of all 5 supervised classification algorithms that used the



provided input features (5&7 for disorder and lithium response prediction respectively). For a binary classification problem, the resulting model decisions could fall into 4 categories: true positives (t_p) when the model correctly predicts the positive class, erroneous positive predictions (false positives, f_p), and, analogously, true negatives (t_n) and false negatives (f_n). The accuracy, the Receiver Operator Characteristic (ROC),

and the confusion matrix were evaluated, and a cross-validation method was used to ensure the robustness of the classification. The results were averaged, or aggregated in the case of the confusion matrix, to create a span of results for each iteration. We used the Area under the Curve (AUC) of the ROC and the confusion matrix to assess the performance of the classifiers using the following metrics: Accuracy, Precision, and the

Fig. 2 Differentially expressed genes between LR and NR LCLs are enriched in Immunoglobulin genes. **a** A PCA analysis of LR vs. NR LCLs' RNA in the original (in-house) dataset before batch correction. Please note different colors are assigned for LR vs. NR & different shapes are assigned for different batches 1–3. LR RNA samples: 15; NR RNA samples: 17. 24 samples (5 LR in duplicates, 4 NR in duplicates, 3 control in duplicates) in Batch 1; 10 samples (1 LR in duplicate, 2 NR in duplicate, 2 control in duplicate) in Batch 2; and 12 samples in Batch 3 (3 LR-no replicate, 1 NR in duplicate plus 4 NR, 4 control-no replicate). In this figure, only LR and NR are plotted. **b** Same as (a) but after batch correction. **c** A heat map of the 61 DEGs (out of a total of 117 detected genes with $FDR < 0.05$) that reached a significance of $FDR < 0.01$ and fold-change > 2 between BD LR and NR subtypes in their respective samples in the original dataset. (LR RNA samples: 15; NR RNA samples: 17. RNA samples were sequenced in replicates for some of the subjects as mentioned above). **d** The plot counts of the top ~40 significant DEGs between BD LR and NR samples (as mentioned above) compared to controls arranged according to the ascending order of FDR (p -Adjusted) values < 0.05 with a fold-change > 2 in the original dataset (batches 1–3). The upper part of the graph shows the gene expression counts while the lower part of the graph presents the FDR values. The remaining DEGs have been plotted in Supplementary Fig 2a & b. (LR RNA samples: 15; NR RNA samples: 17. The expression counts have been normalized in DEseq2 in the Log2 scale).

AUC of the ROC:

$$\text{accuracy}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} 1(\hat{y}_i = y_i) \quad (1)$$

$$\text{precision} = \frac{tp}{tp+fp} \quad (2)$$

$$\text{recall} = \frac{tp}{tp+fn} \quad (3)$$

RESULTS

Differentially expressed genes between LR and NR LCLs are enriched in Immunoglobulin genes

To seek biomarkers for lithium response (Fig. 1.), we performed RNA sequencing of the LCLs of 19 BD patients. Nine of them were LRs and ten belonged to the NR subtype of BD. Since the LCLs were cultured and sequenced in three different batches (first cohort, batches 1–3), we examined them for batch effects. Batch effects refer to the unwanted variations in the data that can arise due to technical differences, such as differences in sample preparation, sequencing, or batch processing. We thus performed a sensitivity analysis (PCA) to evaluate the impact of batch effects on the downstream analysis. We found the variability arising in PC1 associated with the batches as visualized in the PCA plot (Fig. 2a). For example, the second batch (triangles) can be seen concentrated near to first quadrant (upper right) of the PCA plot and the third batch (squares) are in the lower part (quadrants 3&4) in the PCA plot (Fig. 2a) regardless of the sample identity (LR or NR). We, therefore, performed a batch correction (see Methods) to account for this technical variability. Figure 2b shows the PCA after the batch correction and reduction of the unwanted variation in the RNA-seq dataset. We then performed differential gene expression analysis (see Methods) between the subtypes (LR vs. NR) of BD subjects. The sixty-one DEGs (out of approx. 117 genes with $FDR < 0.05$) with a fold change above 2 and an $FDR < 0.01$ between the LR and NR subtypes across different samples in different batches (batches 1–3) are presented in Fig. 2c as a heatmap. The extensive list of DEGs (approximately 117 genes with $FDR < 0.05$) with the plot counts has been plotted in Fig. 2d and Supplementary 2a & b in ascending order of FDR (p -Adj) values. Among the top 10 DEGs with the lowest p -Adj values, 70% belonged to the immunoglobulin heavy and light chain genes (Fig. 2d). A volcano plot is also shown in Supplementary Fig 2c. for the highly significant DEGs (p -Adj < 0.05) with a fold change of above 2 between LR vs. NR BD subtypes.

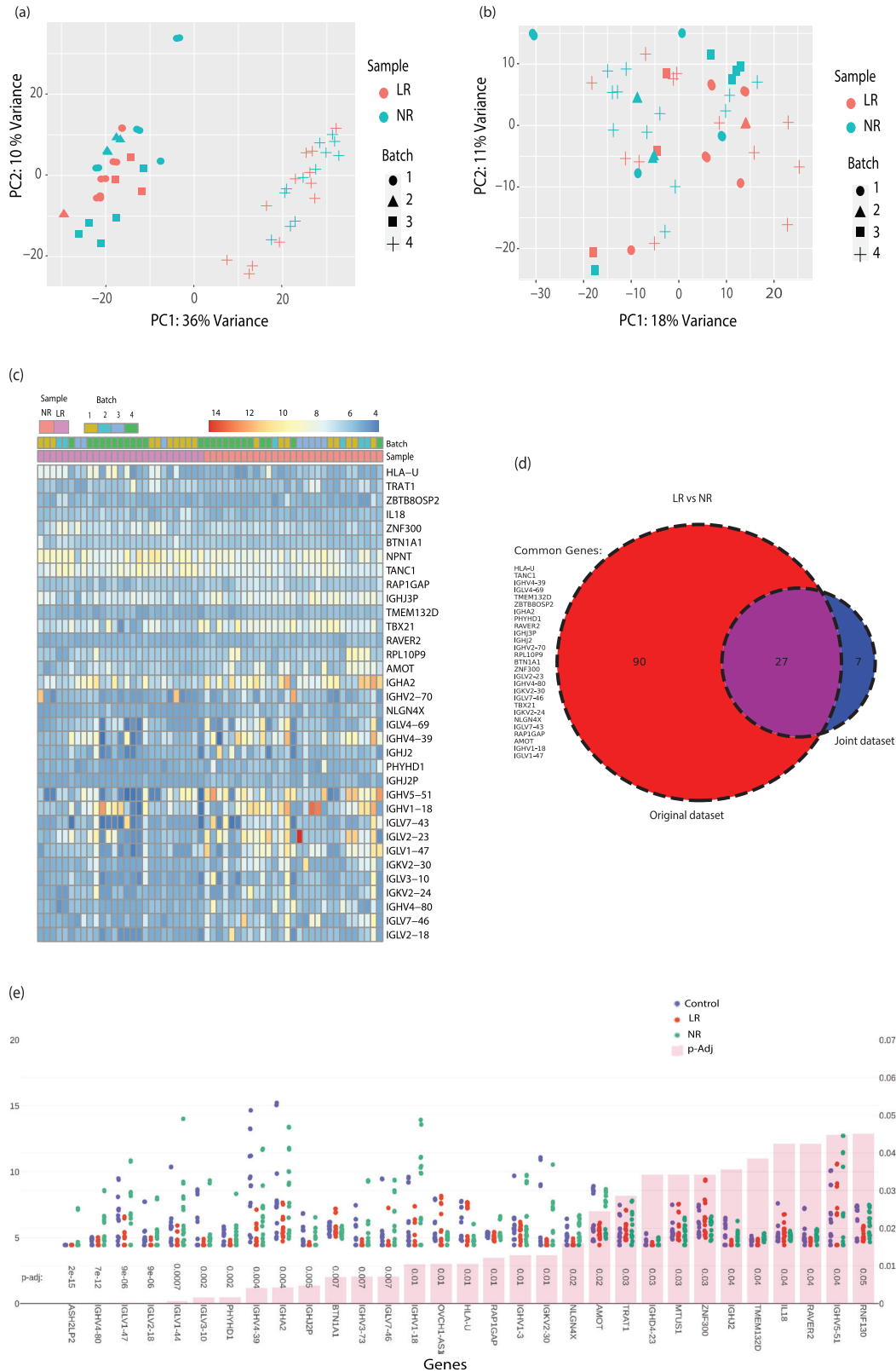
Differentially expressed genes between LR and NR LCLs are highly enriched in Immunoglobulin genes when adding a dataset from another study

To improve the statistical power, we next combined our in-house (original) data with the RNA-seq data from another cohort of BD

LR and NR subtypes that was previously published [38] (also see Methods). As explained above, we repeated the analysis for batch effect. Since the data was acquired from a different laboratory, a higher magnitude of the batch effect was observed as visualized with the separate clustering of the first cohort (batches 1–3) in the left side of the PCA plot (quadrant 3) and the datasets from the second cohort (batch 4) at the right side of the PCA plot (quadrant 4) (Fig. 3a). This observed variability was reduced using our batch correction method. The PCA plots before and after batch corrections are presented in Fig. 3a & 3b respectively to show the effectiveness of the method we employed. A heat map showing the dysregulated genes with a fold change of over 2 between the LR and NR LCLs across different samples in different batches (batches 1–4) with an $FDR < 0.05$ is plotted in Fig. 3c. A Venn diagram showing the total of the shared 27 DEGs between the two datasets (original and joint) is shown in Fig. 3d. Figure 3e represents the plot counts of 31 DEGs for the joint dataset arranged according to the ascending p -Adj (FDR) values. We again found 80% of the top 10 DEGs belonging to the immunoglobulin family in this combined dataset coming from different laboratories and diagnosed by different psychiatrists (Fig. 3e). A volcano plot is presented in Supplementary Fig 3 for highly significant genes with a fold change of above 2 and $FDR < 0.05$ between LR vs. NR BD subtypes in the joint datasets.

Shared genome-wide association studies (GWAS) genes

We were interested to see whether any of the DEGs in the LCLs in our original dataset also appeared in published GWAS of BD patients. We searched the GWAS catalog [59] for "bipolar/lithium" and extracted studies that reported significant genetic variants. We found overall 61 different studies but only 5 studies had one or more genetic variants in common with our DEGs from the original (batches 1–3) datasets after analysis (see the GitHub link for the code used). These genes along with the respective studies have been summarized in Table 1. Among the common genes, RIMS1 and BCL11B were notable because they appeared in 4 GWAS publications and appeared together in 2 of the publications. RIMS1 is a protein involved in neurotransmission because it is required for synaptic vesicle exocytosis [60]. RIMS1 has previously been found to have altered gene expression in cortical brain samples from SCZ and autism patients [61]. BCL11B (also known as CTIP2) is involved in both neuronal and immunological functions. BCL11B has been implicated in Alzheimer's disease, Huntington's disease, Neuro-HIV, learning, and memory, and its cellular role in cortical GABAergic neurons, medium spiny neurons, and vomeronasal sensory neurons have been studied [62]. In its non-neuronal role, BCL11B is crucial for T-cell differentiation and VDJ recombination in immunoglobulin proteins [62, 63]. ADCY1 and NPTX1 were also found to be associated with BD in our analysis as well as other GWAS studies [64, 65]. ADCY1 plays a potential role in learning and memory [66, 67], whereas NPTX1 is required for neural cell specification and is also known to be involved in synaptic plasticity [68, 69]. The other non-overlapping genetic variants from 56 genome-wide studies on BD subjects have been listed in Supplementary Table 1.



Differentially expressed genes between BD and control LCLs are enriched in immunoglobulin genes

We next proceeded to identify genes that are differentially expressed between LCLs of BD patients compared to control individuals in the original dataset. We performed a similar analysis

and compared the samples between the two groups- Control and BD (including LR and NR subtypes). The PCA plots before batch correction are presented in Fig. 4a. We again reduced the ambiguous technical factors by employing batch correction methods which is plotted in Fig. 4b. A heatmap of the top DEGs

Fig. 3 Differentially expressed genes between LR and NR LCLs are highly enriched in Immunoglobulin genes when adding a dataset from another study. **a** A PCA analysis of LR vs. NR LCLs' RNA in the joint dataset (with an additional dataset from another lab, batch 4) before batch correction. Please note different colors are assigned for LR vs. NR samples & different shapes are assigned for different batches 1–4. LR RNA samples: 27; NR RNA samples: 29. 24 samples (5 LR in duplicates, 4 NR in duplicates, 3 control in duplicates) in Batch 1; 10 samples (1 LR in duplicate, 2 NR in duplicate, 2 control in duplicate) in Batch 2; and 12 samples in Batch 3 (3 LR-no replicate, 1 NR in duplicate plus 4 NR, 4 control-no replicate); Batch 4–24 samples (12 LR-no replicates and 12 NR-no replicates). In this figure, only LR and NR are plotted. **b** Same as (a) but after batch correction. **c** A heat map of the DEGs that reached a significance of $FDR < 0.05$ and fold-change > 2 between BD LR and NR subtypes in their respective samples in the joint dataset. (LR RNA samples: 27; NR RNA samples: 29. RNA samples were sequenced in replicates for some of the subjects as mentioned above). **d** A Venn diagram showing 27 common DEGs between the original dataset and joint dataset of BD LR compared to BD NR samples, with a threshold of $FDR < 0.05$ and fold-change > 2 . **e** The plot counts of all significant DEGs in BD LR and BD NR samples compared to control samples arranged according to the ascending order of FDR (p -Adjusted) values < 0.05 with a fold-change > 2 (joint dataset). The upper part of the graph shows the gene expression counts while the lower part of the graph presents the FDR values. (LR RNA samples: 27; NR RNA samples: 29. RNA samples were sequenced in replicates for some of the subjects. The expression counts have been normalized in DEseq2 in the Log2 scale).

with a fold change ≥ 2 and $FDR < 0.05$ is plotted from the original dataset (Batch 1, 2 & 3) (Fig. 4c). In Fig. 4d, we have presented the plot counts of the DEGs in ascending order of p -Adjusted values. A volcano plot is shown in Supplementary Fig 4a for the highly significant DEGs in BD compared to control groups.

We also analyzed to determine the fraction of immunoglobulin genes in the DEG groups vs. the total and substantially expressed immunoglobulin genes detected by RNA sequencing. First, we checked the percentage of immunoglobulin genes when observing just the highly expressed genes with two different thresholds of over a count of 100 (in the sequencing) or a count of 1000 (in the sequencing). As can be observed in Supplementary Fig. 4b (i–iii), the percentage of highly expressed immunoglobulin genes is ~ 1.1 – 1.2% in the LCLs. However, we found the immunoglobulin genes to be highly enriched in the BD vs control group (17.4%). The percentage of these genes was even higher in DEGs of LR vs. NR subtype comparisons in the original (29.1%) and in the joint datasets (52%). The graphs representing this data are presented in Supplementary Fig 4b (iv–vi). This evidence suggests the deregulation of the peripheral immune system in BD patients including the LR and NR subtypes compared to control.

Predicting BD from the RNA of LCLs

We started this study to develop biomarkers that will be easy and fast to implement in the clinic. Batch effects are known to often drive the main differences between datasets of similar groups acquired in different technical conditions. It was therefore important for us to develop a robust protocol that will allow the prediction of BD (Fig. 5) and lithium response (Fig. 6) from the RNA-seq datasets. To examine this, we used five types of supervised classification algorithms: Logistic Regression (Lr), Random Forest (RF), K- Nearest Neighbors (K-NN), Support Vector Machine (SVM), and Neural Network (NN). As described in the Methods, to prevent overfitting we first performed batch correction on RNA-seq datasets using *Pycombat* and implemented a feature selection (see Methods for detail). We partitioned our data for cross-validation (see Methods). The expression values and the p -values of the subset of the 20 genes used for feature selection have been shown in Fig. 5a & Supplementary Fig 5a respectively. We selected a subset of the genes and ran the prediction algorithm (Lr) to see the dependence of the accuracy on the number of features (Supplementary Fig 5b). Notably, 5 features were enough to give an accuracy of about 93%. According to the algorithm, we selected a subset of 5 genes presented in (Fig. 5b) that were highly predictive for the disorder. Supplementary Fig 5c presents the confusion matrix (see Methods) of the 5 classifiers after repeating the split train /test approach 50 times. Figure 5c presents the performance of the respective 5 classifiers in the form of the area under the ROC curve. The Lr, SVM, and NN classifiers gave a successful prediction with a very low prediction error ($AUC = 0.99 \pm 0.00$) (Fig. 5c).

Additionally, other classifiers also gave a small error in the prediction: RF ($AUC = 0.97 \pm 0.04$) and K-NN ($AUC = 0.99 \pm 0.01$) (Fig. 5c). Following randomly splitting the data into a train and test dataset and performing the classification 50 times, the accuracy of the prediction was also plotted in Fig. 5d. The Lr (0.996 ± 0.021), SVM (0.996 ± 0.022), NN (0.973 ± 0.058), and K-NN (0.956 ± 0.106) classifiers were more accurate than RF ($0.888 \pm 0.087\%$).

Predicting BD patients' response to lithium treatment

We next wondered whether the gene expression analysis from the LCLs obtained from the BD (LR & NR subtypes) and control subjects would be sufficient to predict lithium response and classify them into respective groups and subtypes. As described above, to reduce the noise and variation, we performed the batch correction and selected 7 genes to be utilized as features from the list of 20 genes with the lowest p -values (Fig. 6a & b; Supplementary Fig 6a & b) in a similar manner to what was described for selecting 5 genes to predict the disorder. Further, to test the robustness of our algorithms, we trained the above-mentioned five classifiers on the selected subset of 7 genes (Fig. 6b) using a 50-50 split train/test approach repeated 50 times (also see Methods). Supplementary Fig 6c presents the confusion matrix for 5 different classifiers. Figure 6c presents the ROC for the 5 classifiers (50% train, 50%test repeated 50 times for the datasets from batches 1–4). Among the predictors, Lr ($AUC = 0.99 \pm 0.01$), SVM ($AUC = 0.98 \pm 0.03$) & NN ($AUC = 0.98 \pm 0.03$) showed a highly effective prediction with a lower error rate, but also RF ($AUC = 0.96 \pm 0.03$) and K-NN ($AUC = 0.95 \pm 0.03$) gave good results. Similarly, the accuracy of the classifiers has been plotted (Fig. 6d). The Lr (0.965 ± 0.038), NN 0.940 ± 0.052 , and SVM (0.925 ± 0.063) had better accuracy compared to RF ($0.874 \pm 0.071\%$) and K-NN (0.86 ± 0.074) classifiers.

Additionally, for lithium response prediction, we trained the model on our dataset (9 LR and 10 NR) and used it to classify a completely different dataset that was acquired in another laboratory (batch 4) and on different patients [38]. This prediction was 95.8% accurate with SVM & NN and 91.7% accurate with the other three classifiers (RF, Lr, and K-NN).

DISCUSSION

The current understanding of the pathophysiology and progression of BD is inadequate. For better management of BD, biomarkers are necessary for diagnosis as well as for the selection of suitable therapeutic interventions [19, 24, 38]. Differential response to lithium treatment by more than half of BD patients is a conundrum in the field of BD research [8]. Some studies have provided information about a genetic link to non-response to lithium treatment, as family members of NR BD patients usually also do not respond to lithium [8, 70–72]. Nevertheless, mood episodes in BD are also known to be impacted by environmental factors like stress and traumatic experience [8].

Table 1. Overlapping genes between GWAS data and DEGs in BD LR vs. NR LCLs.

Study Accession	Title	Overlapping genes	Total reported genes
GCST005081	Association of Polygenic Score for Schizophrenia and HLA Antigen and Inflammation Genes With Response to Lithium in Bipolar Affective Disorder: A Genome-Wide Association Study [101]	ADCY1	8
GCST008103	Genome-wide association study identifies 30 loci associated with bipolar disorder [102]	RIMS1, BCL11B, NPTX1	146
GCST009600	Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders [103]	RIMS1, BCL11B	111
GCST011102	Novel Risk Loci Associated With Genetic Risk for Bipolar Disorder Among Han Chinese Individuals: A Genome-Wide Association Study and Meta-analysis [104]	RIMS1	23
GCST012465	Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology [105]	BCL11B	63
56 more studies that did not have any overlap with our set of DEGs are mentioned in Supplementary Table 1		Genes with no overlap with DEGS HOMER2, MIR52, NF1A, PTGFR, FAT, MHC, RP11-252P19.1,ADM,NFIA, DGKH NCAN, DPY19L3, CGNL1, SPATS2L, RND1, DDX23, CANCNB3, GRIK5, ST8SIA2, C15orf32, SLITRK1, SLITRK6, SNAP91, PRSS35, SDCCAG8, ANK3, CADM3, ENSG00000258081, POU3F2, MIR2113 FER1L5, LMAN2L, CNNM4, NF1A, ST6GALNAC3, ADCY2, ODZ4, TRANK1, ROR2, NFIL3, AUH, MIR3910, APOB, RPRD2, AC096669.1, LOC440300, LOC388152, LOC642423, GOLGA6L4, DNM1P41, GOLGA6L5, UBE2Q2P1,LOC100506874, ZSCAN2, SCAND2P, WDR73, NMB, SEC11A, ZNF592, ALPK3, SLC28A1OR4F16, OR4F29, LOC729467, LOC100133331, LOC100132287, FGGY, TMEM108, FAM178B, MKLN1, ETV5, FGGY, STAG1	

(Reference database: GWAS catalog [59]; see Methods) and the significant DEGs from the original dataset of BD LR and BD NR samples. The other studies that did not overlap with our set of DEGs are mentioned in Supplementary Fig. 1.

In our study, we compared LCLs from 3 groups, Control, BD LRs, and BD NRs, by performing RNA sequencing to search for DEGs and associated biological pathways between the groups. The advantage of our method is its high-throughput, hypothesis-free approach, and the use of accessible patient-specific LCLs. LCLs have a huge potential for use in the discovery of biomarkers, especially in psychiatric disorders where the brain tissues are not accessible for molecular investigations from live patients [29, 38]. We have also analyzed and compared our in-house (original) dataset of BD LR and NR patient subtypes with a previously published study of 24 BD patients [38] (12 LR and 12 NR) as a joint dataset and found shared DEGs between the two cohorts. Additionally, in our study, we discovered significant DEGs between the BD and control groups.

When comparing the subtypes within the BD group, i.e., the LR vs. the NR patients, a set of genes belonging to the immunoglobulin heavy and light chains variable region were found to be significantly upregulated in lithium NRs as compared to LRs in the joint datasets. Increased expression of the immunoglobulin family genes specifically in lithium NRs entails the deregulation of the peripheral immune system. Apart from the immunoglobulin genes, we found the HLA-U, ZNF300, and TRAT1 genes to be significantly downregulated in NR subtypes. Interestingly, previous GWAS studies have reported the association of these genes with BD and other psychiatric disorders [73–75]. HLA-U is a pseudogene found on the MHC complex [76]. GWAS have revealed the MHC complex as an important risk gene in both BD and Schizophrenia (SCZ) [77–81]. Apart from its vital immunological role, the HLA locus has been proposed to play an important role during neurodevelopment [80, 81]. In our study, we found IL-18 to

be significantly downregulated in NRs. IL-18 is a proinflammatory cytokine reported to be expressed in different regions of the brain and has a known role in neuroinflammation [82]. Some studies have reported high levels of circulating IL-18 in BD and SCZ patients [83–85]. The altered expression of these genes specifically in NRs compared to LRs is intriguing, especially because these genes overlap with previous reports in BD.

Further, when we analyzed the original and joint datasets, 27 genes were found to be common. Apart from HLA-U and IGV genes, TANC1 and TMEM 132D were among the top 5 most common genes. The TANC1 protein has been shown to interact with PSD95 as well as other synaptic proteins such as glutamate receptors [86]. A TANC1 mutation, along with mutations in NRXN1 and RBMS1 genes, were implicated in psychomotor delayed development in a case study of chromosomal inversion [87]. GWAS have linked the TMEM 132D polymorphism to anxiety and panic disorder [88, 89]. TMEM 132D has also been reported to be dysregulated at the mRNA level in brain regions associated with anxiety disorders [89]. The differential expression of these critical genes is even more intriguing since we integrated datasets from another group that was conducted in a different cohort.

We also compared gene expression between BD and control LCLs. We found a downregulation of immunoglobulin kappa chain variable genes in BD patients when compared with control groups. ADAM23, IGHEP1, GNAQ, FRG1EP, and PPP4R4 were upregulated in BD patients, including LRs and NRs as subsets. GNAQ and ADAM23 have earlier been reported to be associated with BD [90, 91]. ADAM23 was previously reported in a microarray-based gene expression study of BD postmortem brain samples [91]. ADAM23 is a membrane protein belonging to the family of

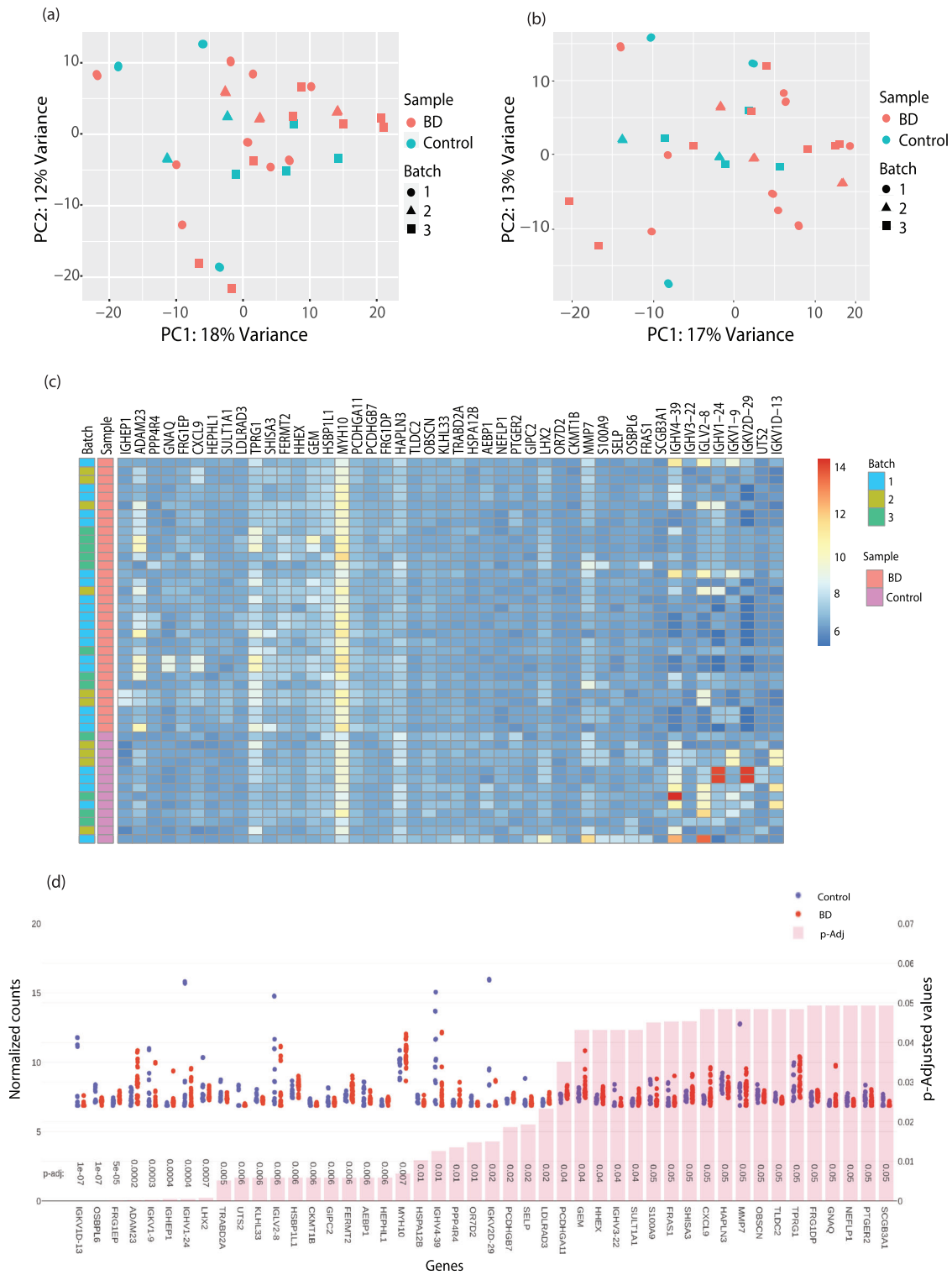


Fig. 4 Differentially expressed genes between BD and control LCLs are enriched in Immunoglobulin genes. **a** PCA analysis of BD vs. control LCLs’ RNA in the original dataset before batch correction. (Please note different colors are assigned for BD vs. Ctrl samples & shapes for different batches 1–3. RNA samples: 32; Control RNA samples: 14. RNA samples were sequenced in replicates for some of the subjects as mentioned in Fig. 2a). **b** Same as (a) but after batch correction. **c** A heat map of the DEGs that reached a significance of FDR < 0.05 and fold-change > 2 between BD and control groups in their respective samples in the original dataset (BD RNA samples: 32; Control RNA samples: 14. RNA samples were sequenced in replicates for some of the subjects). **d** The plot counts of all significant DEGs in BD samples compared to control samples arranged according to the ascending order of FDR (*p*-Adjusted) values < 0.05 with a fold-change > 2. (Original dataset). The upper part of the graph shows the gene expression counts while the lower part of the graph presents the FDR values. (RNA samples: 32; Control RNA samples: 14. The expression counts have been normalized in DEseq2 in the Log2 scale).

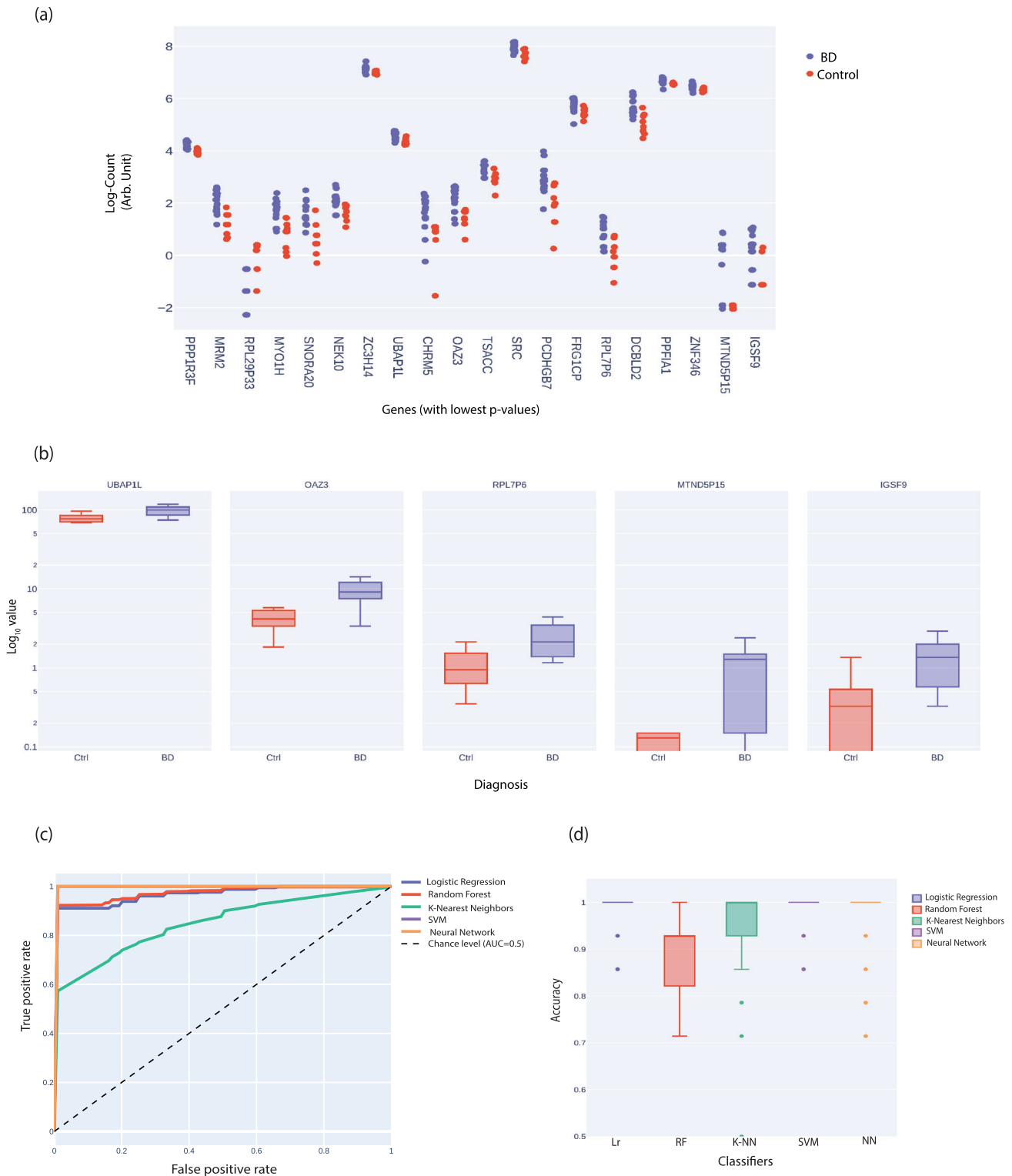
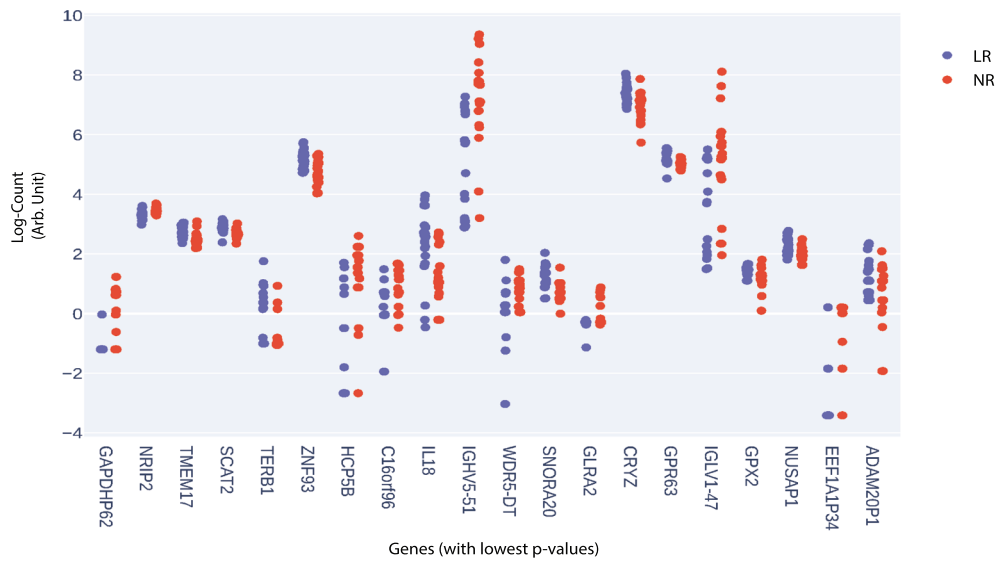
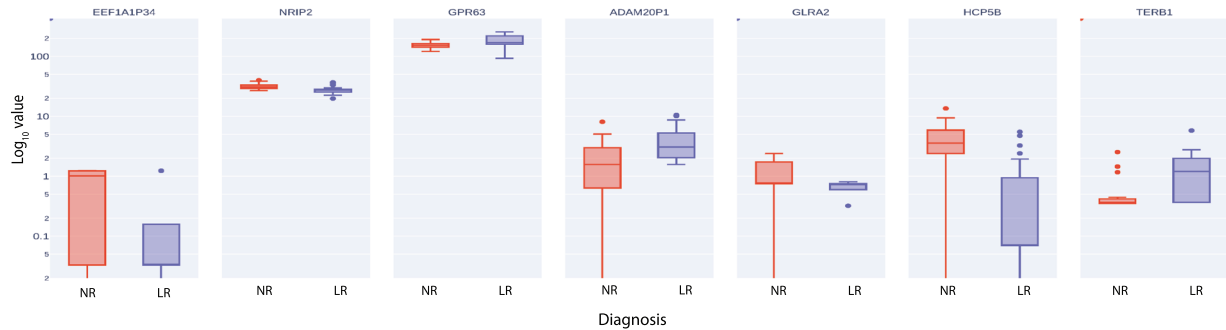


Fig. 5 Predicting BD from the RNA of LCLs. Using 50%-50% cross-validation schemes, we utilized five different supervised classification algorithms to predict which of the LCLs' RNA was extracted from a BD patient or control individual. Approximately 50% of the data (15 subjects) were used as a training set and ~50% (14 subjects) of the dataset were used as a test (out of the dataset that consisted of 19 BD and 10 controls) from cohort 1. The 50%-50% train/test approach was iterated 50 times. The five algorithms used were Logistic Regression (Lr), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), and Neural Network (NN). **a** The gene expression counts of 20 genes in BD and control subjects with the lowest p-values (Mann Whitney U test) were used for the feature selection for the different classifiers. **b** Box plots of the expression of the five genes in BD and control groups that were used for the prediction. **c** A joint ROC for the prediction of BD vs. Control for the five classifiers. **d** The accuracy scores for BD vs. Control prediction for the five classifiers after splitting the data into test/train and repeating 50 times.

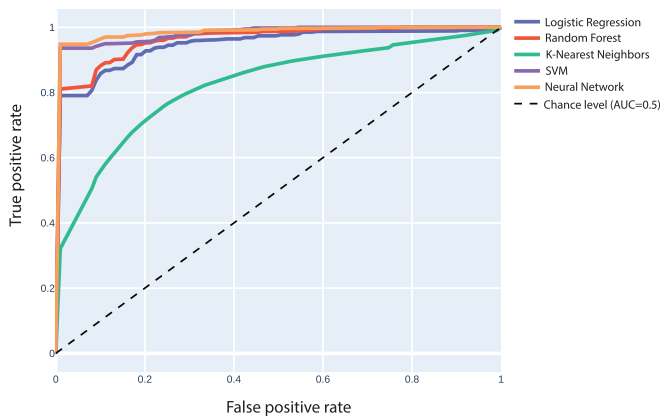
(a)



(b)



(c)



(d)

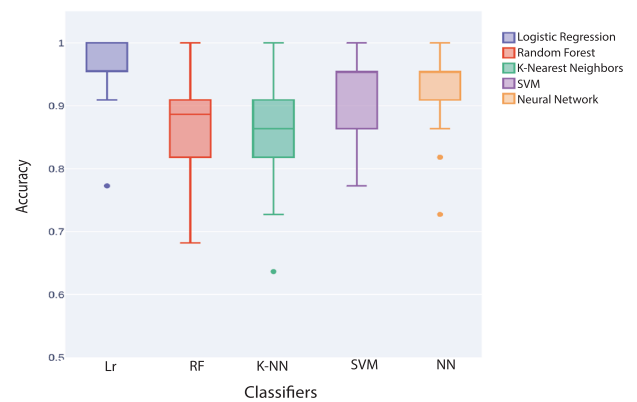


Fig. 6 Predicting BD patients' response to lithium treatment. Similarly as Fig. 5; Approximately 50% of the data (21 subjects) were used as a training set and ~50% (22 subjects) of the datasets were used as a test (out of the datasets that consisted of 19 BD subjects including LR & NR from cohort 1 & 24 BD subjects including LR & NR from cohort 2). **a** The plot counts of 20 genes in LR and NR subjects with the lowest p -values (Mann–Whitney U test) were used for the feature selection for the five classifiers. **b** Box plots of the expression of the seven genes in BD LR and NR subtypes that were used for the prediction. **c** A joint ROC for the prediction of LR vs. NR for the five classifiers. **d** Similarly, the accuracy scores of the classifiers for BD LR vs. NR prediction using a 50–50% cross-validation scheme repeated 50 times. Additionally, training was done on our dataset and the trained model was used to predict the lithium response of the BD patients in the Cohort 2 dataset (Milanesi et al. [38]).

ADAM proteins and is part of a presynaptic complex interacting with the Lgl receptor [92]. While FRG1EP has a reported function in cancer, the function of the other above-mentioned genes is not well understood, especially in the context of psychiatric diseases [93].

Overall, our findings suggest significant dysregulation of genes with immunological function as well as alteration of genes involved in synaptic pathways in BD LCLs. Overwhelming reports have highlighted the importance of immune dysregulation in BD and SCZ [94–97]. We also found genes implicated in neuroinflammation and anxiety disorder altered in BD LCLs. Furthermore, between BD patient subtypes, we found cytokines and immunological genes specifically altered in NR compared to LR patients.

We used five diverse supervised classification algorithms – Lr, RF, K-NN, SVM, and NN to predict BD and lithium treatment response. Each classifier has its own set of pros and cons as described briefly above (in Methods) and reviewed previously [36, 55, 57, 98]. One typical drawback with supervised classification algorithms is that their performance decreases with high-dimensional datasets (like RNA-seq data) with a large number of features leading to overfitting [98, 99]. To prevent overfitting, we reduced the number of genes (features) used for training the classifier to only 5 & 7 (for BD prediction and lithium response respectively) and thus reduced the complexity of the model by selecting only the statistically informative and robust genes for prediction. Further, our model used a random sampling cross-validation scheme to split the data into training and test datasets. This gave robust statistical results with a very low error rate especially using the Lr, SVM, and NN classifiers for lithium response prediction. The error rate was very low (less than 5% with SVM and NN; approx. 8% for Lr, RF, and K-NN) even when trained on our in-house dataset and tested on the data from another laboratory. The BD samples could also be predicted with high conformity and distinguished from the control samples with minimal errors using the Lr, SVM, and NN classifiers.

Some studies have used supervised machine learning algorithms for Diabetes, Cardiovascular diseases, Cancers, Alzheimer's disease, and Parkinson's disease using clinical parameters and neuroimaging datasets [100]. Ours is the first study to use transcriptomic datasets from LCLs of BD patients to predict the disorder as well as the responsiveness to lithium. The method is cost-effective and scalable and can be easily implemented in a psychiatric clinic. While one of the limitations of the study is the insufficient clinical sample size given the complexity and polygenic nature of BD, we showed the robustness of our algorithm by testing it on a new dataset from another lab. Further, the use of transformed B cell lines (LCLs) may not be the ideal tissue choice for a neuropsychiatric disorder like BD but the fact that we could find statistically significant DEGs and were able to predict and classify BD subtypes using the classifiers provides proof about the existing specific genetic signatures between BD subtypes.

In conclusion, using RNA-seq, we have found a set of DEGs from LCLs of BD patients that can be used as potential biomarkers to diagnose as well as classify BD patient subtypes. Functional studies of these genes in model systems should also aid in elucidating the cellular and molecular processes underlying BD pathophysiology. Importantly, we hope that our study and our developed algorithm will serve as an easy and ready-to-use protocol for deciding on effective treatment in the clinic within days of diagnosis.

DATA AND CODE AVAILABILITY

The data including raw files are available in the following Github link below. All the codes used in this study for RNA-seq data analysis, Machine learning predictor analysis, and GWAS analysis are also available in the following link. <https://github.com/Precision-Disease-Modeling-Lab/Lithium-Response-Predictor>.

REFERENCES

- Vieta E, Berk M, Schulze TG, Carvalho AF, Suppes T, Calabrese JR, et al. Bipolar disorders. *Nat Rev Dis Primers*. 2018;4:18008.
- Soreca I, Fagiolini A, Frank E, Houck PR, Thompson WK, Kupfer DJ. Relationship of general medical burden, duration of illness and age in patients with bipolar I disorder. *J Psychiatr Res*. 2008;42:956–61.
- Edvardsen J, Torgersen S, Roysamb E, Lygren S, Skre I, Onstad S, et al. Heritability of bipolar spectrum disorders. Unity or heterogeneity? *J Affect Disord*. 2008;106:229–40.
- Belmaker RH. Bipolar disorder. *N Engl J Med*. 2004;351:476–86.
- Baldessarini RJ, Tondo L, Vazquez GH. Pharmacological treatment of adult bipolar disorder. *Mol Psychiatry*. 2019;24:198–217.
- Girardi P, Brugnoli R, Manfredi G, Sani G. Lithium in bipolar disorder: optimizing therapy using prolonged-release formulations. *Drugs R D*. 2016;16:293–302.
- Geddes JR, Burgess S, Hawton K, Jamison K, Goodwin GM. Long-term lithium therapy for bipolar disorder: systematic review and meta-analysis of randomized controlled trials. *Am J Psychiatry*. 2004;161:217–22.
- Cruceanu C, Alda M, Turecki G. Lithium: a key to the genetics of bipolar disorder. *Genome Med*. 2009;1:79.
- Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Molecular psychiatry*. 2008;13:197–207.
- Dieset I, Djurovic S, Tesli M, Hope S, Mattingsdal M, Michelsen A, et al. Up-regulation of NOTCH4 gene expression in bipolar disorder. *Am J Psychiatry*. 2012;169:1292–1300.
- Lee KW, Woon PS, Teo YY, Sim K. Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt? *Neurosci Biobehav Rev*. 2012;36:556–71.
- Nayak R, Rosh I, Kustanovich I, Stern S. Mood stabilizers in psychiatric disorders and mechanisms learnt from in vitro model systems. *Int J Mol Sci*. 2021;22:17.
- Psychiatric GCBWDWG. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet*. 2011;43:977–83.
- Nestler EJ, Hyman SE. Animal models of neuropsychiatric disorders. *Nat Neurosci*. 2010;13:1161–9.
- Santos R, Linker SB, Stern S, Mendes APD, Shokhirev MN, Erikson G, et al. Deficient LEF1 expression is associated with lithium resistance and hyperexcitability in neurons derived from bipolar disorder patients. *Mol Psychiatry*. 2021;26:2440–56.
- Stern S, Santos R, Marchetto MC, Mendes APD, Rouleau GA, Biesmans S, et al. Neurons derived from patients with bipolar disorder divide into intrinsically different sub-populations of neurons, predicting the patients' responsiveness to lithium. *Mol Psychiatry*. 2018;23:1453–65.
- Stern S, Sarkar A, Galor D, Stern T, Mei A, Stern Y, et al. A physiological instability displayed in hippocampal neurons derived from lithium-nonresponsive bipolar disorder patients. *Biol Psychiatry*. 2020;88:150–8.
- Gurwitz D. Human iPSC-derived neurons and lymphoblastoid cells for personalized medicine research in neuropsychiatric disorders. *Dialogues Clin Neurosci*. 2016;18:267–76.
- Stern S, Linker S, Vadodaria KC, Marchetto MC, Gage FH. Prediction of response to drug therapy in psychiatric disorders. *Focus (Am Psychiatr Publ)*. 2019;17:294–307.
- Sigitova E, Fisar Z, Hroudova J, Cikankova T, Raboch J. Biological hypotheses and biomarkers of bipolar disorder. *Psychiatry Clin Neurosci*. 2017;71:77–103.
- Hashimoto K. Brain-derived neurotrophic factor as a biomarker for mood disorders: an historical overview and future directions. *Psychiatry Clin Neurosci*. 2010;64:341–57.
- Emamghoreishi M, Schlichter L, Li PP, Parikh S, Sen J, Kamble A, et al. High intracellular calcium concentrations in transformed lymphoblasts from subjects with bipolar I disorder. *Am J Psychiatry*. 1997;154:976–82.
- Naydenov AV, MacDonald ML, Ongur D, Konradi C. Differences in lymphocyte electron transport gene expression levels between subjects with bipolar disorder and normal controls in response to glucose deprivation stress. *Arch Gen Psychiatry*. 2007;64:555–64.
- Le-Niculescu H, Kurian SM, Yehyawi N, Dike C, Patel SD, Edenberg HJ, et al. Identifying blood biomarkers for mood disorders using convergent functional genomics. *Mol Psychiatry*. 2009;14:156–74.
- Misiak B, Frydecka D, Stanczykiewicz B, Samochowiec J. Editorial: peripheral markers of immune response in major psychiatric disorders: where are we now and where do we want to be? *Front Psychiatry*. 2019;10:5.
- Bennett FC, Molofsky AV. The immune system and psychiatric disease: a basic science perspective. *Clin Exp Immunol*. 2019;197:294–307.
- So J, Warsh JJ, Li PP. Impaired endoplasmic reticulum stress response in B-lymphoblasts from patients with bipolar-I disorder. *Biol Psychiatry*. 2007;62:141–7.

28. Warsh JJ, Andreopoulos S, Li PP. Role of intracellular calcium signaling in the pathophysiology and pharmacotherapy of bipolar disorder: current status. *Clin Neurosci Res.* 2004;4:201–13.
29. Kato T, Hayashi-Takagi A, Toyota T, Yoshikawa T, Iwamoto K. Gene expression analysis in lymphoblastoid cells as a potential biomarker of bipolar disorder. *J Hum Genet.* 2011;56:779–83.
30. Yoon IS, Li PP, Siu KP, Kennedy JL, Macchiardi F, Cooke RG, et al. Altered TRPC7 gene expression in bipolar-I disorder. *Biol Psychiatry.* 2001;50:620–6.
31. Washizuka S, Kakiuchi C, Mori K, Tajima O, Akiyama T, Kato T. Expression of mitochondria-related genes in lymphoblastoid cells from patients with bipolar disorder. *Bipolar Disord.* 2005;7:146–52.
32. Iwamoto K, Bundo M, Washizuka S, Kakiuchi C, Kato T. Expression of HSPF1 and LIM in the lymphoblastoid cells derived from patients with bipolar disorder and schizophrenia. *J Hum Genet.* 2004;49:227–31.
33. Tripathi U, Mizrahi L, Alda M, Falkovich G, Stern S. Information theory characteristics improve the prediction of lithium response in bipolar disorder patients using a support vector machine classifier. *Bipolar Disord.* 2023;25:110–27.
34. Mizrahi L, Shekhdem HA, Stern S. Age separation dramatically reduces COVID-19 mortality rate in a computational model of a large population. *Open Biol.* 2020;10:200213.
35. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak.* 2019;19:281.
36. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet.* 2015;16:321–32.
37. Tarca AL, Carey VJ, Chen X-W, Romero R, Drăghici S. Machine learning and its applications to biology. *PLoS Comput Biol.* 2007;3:e116.
38. Milanesi E, Voinsky I, Hadar A, Srouji A, Maj C, Shekhtman T, et al. RNA sequencing of bipolar disorder lymphoblastoid cell lines implicates the neurotrophic factor HRP-3 in lithium's clinical efficacy. *World J Biol Psychiatry.* 2019;20:449–61.
39. Darlington GJ. Epstein-Barr virus transformation of lymphoblasts. 2006; 2006: pdb. prot4481.
40. Neitzel H. A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum Genet.* 1986;73:320–6.
41. Scheinfeldt LB, Hodges K, Pevsner J, Berlin D, Turan N, Gerry NP. Genetic and genomic stability across lymphoblastoid cell line expansions. *BMC Res Notes.* 2018;11:558.
42. S. A. FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinformatics* 2010.
43. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15–21.
44. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
45. 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.
46. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B (Methodological).* 1995;57:289–300.
47. Haynes W. Benjamini-Hochberg Method. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds). *Encyclopedia Sys Biol.* Springer New York: New York, NY, 2013, pp 78–78.
48. Behdenna A, Haziza J, Azencott C-A, Nordor AJB. pyComBat, a Python tool for batch effects correction in high-throughput molecular data using empirical Bayes methods. 2020: 2020.2003. 2017.995431. <https://doi.org/10.1101/2020.03.17.995431>.
49. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2006;8:118–27.
50. Sarker IH. Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci.* 2021;2:160.
51. Chen X, Ishwaran H. Random forests for genomic data analysis. *Genomics.* 2012;99:323–9.
52. Han J, Liu Y, Sun XJ. CoSE, Science S. A scalable random forest algorithm based on MapReduce. 2013:849–52.
53. Han J, Pei J, Tong H. Data mining: concepts and techniques. 2022. Morgan kaufmann. https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:11082.
54. Yao Z, Ruzzo WL. A Regression-based K nearest neighbor algorithm for gene function prediction from heterogeneous data. *BMC Bioinform.* 2006;7:511.
55. Bansal M, Goyal A, Choudhary A. A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics J.* 2022;3:100071.
56. Shen X, Jiang C, Wen Y, Li C, Lu Q. A brief review on deep learning applications in genomic studies. *Front Syst Biol.* 2022;2. <https://doi.org/10.3389/fsysb.2022.87717>.
57. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol.* 1996;49:1225–31.
58. Fabian Pedregosa GV, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Duchesnay E. Scikit-learn: Machine Learning in Python. *J Mach Learning Res.* 2011;12:2825–30.
59. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* 2023;51:D977–d985.
60. Kiyonaka S, Wakamori M, Miki T, Uriu Y, Nonaka M, Bito H, et al. RIM1 confers sustained activity and neurotransmitter vesicle anchoring to presynaptic Ca²⁺ channels. *Nat Neurosci.* 2007;10:691–701.
61. Guan J, Cai JJ, Ji G, Sham PC. Commonality in dysregulated expression of gene sets in cortical brains of individuals with autism, schizophrenia, and bipolar disorder. *Trans Psychiatry.* 2019;9:152.
62. Lennon MJ, Jones SP, Lovelace MD, Guillemin GJ, Brew BJ. Bcl11b-A critical neurodevelopmental transcription factor-roles in health and disease. *Front Cell Neurosci.* 2017;11:89.
63. Wakabayashi Y, Watanabe H, Inoue J, Takeda N, Sakata J, Mishima Y, et al. Bcl11b is required for differentiation and survival of αβ T lymphocytes. *Nat Immunol.* 2003;4:533–9.
64. Genetics CoL. Association of polygenic score for schizophrenia and HLA antigen and inflammation genes with response to lithium in bipolar affective disorder: a genome-wide association study. *JAMA Psychiatry.* 2018;75:65–74.
65. Rajkumar AP, Christensen JH, Mattheisen M, Jacobsen I, Bache I, Pallesen J et al. Analysis of t(9;17)(q33.2;q25.3) chromosomal breakpoint regions and genetic association reveals novel candidate genes for bipolar disorder. 2015;17: 205–11.
66. Wang H, Ferguson GD, Pineda VV, Cundiff PE, Storm DR. Overexpression of type-1 adenylyl cyclase in mouse forebrain enhances recognition memory and LTP. *Nature neuroscience.* 2004;7:635–42.
67. Devasani K, Yao Y. Expression and functions of adenylyl cyclases in the CNS. *Fluids and Barriers of the CNS.* 2022;19:23.
68. Boles NC, Hirsch SE, Le S, Corneo B, Najm F, Minotti AP, et al. NPTX1 regulates neural lineage specification from human pluripotent stem cells. *Cell Rep.* 2014;6:724–36.
69. Gómez de San José N, Massa F, Halbgebauer S, Oeckl P, Steinacker P, Otto M. Neuronal pentraxins as biomarkers of synaptic activity: from physiological functions to pathological changes in neurodegeneration. *J Neural Transmission.* 2022;129:207–30.
70. Alda M, Grof E, Cavazzoni P, Duffy A, Martin R, Ravindran L, et al. Autosomal recessive inheritance of affective disorders in families of responders to lithium prophylaxis? *J Affect Disord.* 1997;44:153–7.
71. Alda M, Grof P, Grof E, Zvolsky P, Walsh M. Mode of inheritance in families of patients with lithium-responsive affective disorders. *Acta Psychiatr Scand.* 1994;90:304–10.
72. Hou L, Heilbronner U, Degenhardt F, Adli M, Akiyama K, Akula N, et al. Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *Lancet.* 2016;387:1085–93.
73. Tamouza R, Krishnamoorthy R, Leboyer M. Understanding the genetic contribution of the human leukocyte antigen system to common major psychiatric disorders in a world pandemic context. *Brain Behav Immun.* 2021;91:731–9.
74. Gamazon ER, Badner JA, Cheng L, Zhang C, Zhang D, Cox NJ, et al. Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Mol Psychiatry.* 2013;18:340–6.
75. Kuo CY, Chen TY, Kao PH, Huang W, Cho CR, Lai YS, et al. Genetic pathways and functional subnetworks for the complex nature of bipolar disorder in genome-wide association study. *Front Mol Neurosci.* 2021;14:772584.
76. Geraghty DE, Koller BH, Hansen JA, Orr HT. The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. 1992;149: 1934–46.
77. Choudhary A, Peles D, Nayak R, Mizrahi L, Stern S. Current progress in understanding schizophrenia using genomics and pluripotent stem cells: A meta-analytical overview. *Schizophrenia Res.* 2022. <https://doi.org/10.1016/j.schres.2022.11.001>.
78. Stern S, Zhang L, Wang M, Wright R, Cordeiro D, Peles D et al. Monozygotic twins discordant for schizophrenia differ in maturation and synaptic transmission. 2022: 2022.2005.2013.491776.
79. Bergen SE, O'Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry.* 2012;17:880–6.
80. McAllister AK. Major histocompatibility complex I in brain development and schizophrenia. *Biol Psychiatry.* 2014;75:262–8.

81. Elmer BM, McAllister AK. Major histocompatibility complex class I proteins in brain development and plasticity. *Trends Neurosci.* 2012;35:660–70.
82. Alboni S, Cervia D, Sugama S, Conti B. Interleukin 18 in the CNS. *J Neuroinflammation.* 2010;7:9.
83. Munkholm K, Weikop P, Kessing LV, Vinberg M. Elevated levels of IL-6 and IL-18 in manic and hypomanic states in rapid cycling bipolar disorder patients. *Brain Behav Immun.* 2015;43:205–13.
84. Luo Y, He H, Zhang M, Huang X, Fan N. Altered serum levels of TNF-alpha, IL-6 and IL-18 in manic, depressive, mixed state of bipolar disorder patients. *Psychiatry Res.* 2016;244:19–23.
85. Szabo A, O'Connell KS, Ueland T, Sheikh MA, Agartz I, Andreou D, et al. Increased circulating IL-18 levels in severe mental disorders indicate systemic inflammatory activation. *Brain Behav Immun.* 2022;99:299–306.
86. Gasparini A, Tosatto SCE, Murgia A, Leonardi E. Dynamic scaffolds for neuronal signaling: in silico analysis of the TANC protein family. *Sci Rep.* 2017;7:6829.
87. Granot-Hershkovitz E, Raas-Rothschild A, Frumkin A, Granot D, Silverstein S, Abeliovich D. Complex chromosomal rearrangement in a girl with psychomotor-retardation and a de novo inversion: inv(2)(p15;q24.2). *Am J Med Genet Part A.* 2011;155a:1825–32.
88. Erhardt A, Czibere L, Roeske D, Lucae S, Unschuld PG, Ripke S, et al. TMEM132D, a new candidate for anxiety phenotypes: evidence from human and mouse studies. *Mol Psychiatry.* 2011;16:647–63.
89. Naik RR, Sotnikov SV, Diepold RP, Iurato S, Markt PO, Bultmann A, et al. Polymorphism in *Tmem132d* regulates expression and anxiety-related behavior through binding of RNA polymerase II complex. *Trans Psychiatry.* 2018;8:1.
90. Greenwood TA, Bipolar Genome Study C, Kelsoe JR. Genome-wide association study of irritable vs. elated mania suggests genetic differences between clinical subtypes of bipolar disorder. *PLoS One.* 2013;8:e53804.
91. Chen H, Wang N, Zhao X, Ross CA, O'Shea KS, McClinnis MG. Gene expression alterations in bipolar disorder postmortem brains. *Bipolar Disord.* 2013;15:177–87.
92. Fukata Y, Lovero KL, Iwanaga T, Watanabe A, Yokoi N, Tabuchi K, et al. Disruption of *LGI1*-linked synaptic complex causes abnormal synaptic transmission and epilepsy. *Proc Natl Acad Sci USA.* 2010;107:3799–804.
93. Parrish RL, Gibson GC, Epstein MP, Yang J. TIGAR-V2: Efficient TWAS tool with nonparametric Bayesian eQTL weights of 49 tissue types from GTEx V8. *HGG Adv.* 2022;3:100068.
94. Rosenblat JD, McIntyre RS. Bipolar disorder and immune dysfunction: epidemiological findings, proposed pathophysiology and clinical implications. *Brain Sci* 2017;7:144. <https://doi.org/10.3390/brainsci7110144>.
95. Barbosa IG, Machado-Vieira R, Soares JC, Teixeira AL. The immunology of bipolar disorder. *Neuroimmunomodulation.* 2014;21:117–22.
96. Rosenblat JD. Targeting the immune system in the treatment of bipolar disorder. *Psychopharmacology (Berl).* 2019;236:2909–21.
97. Hamdani N, Bengoufa D, Godin O, Doukhan R, Le Guen E, Daban-Huard C, et al. Immunoglobulin sub-class distribution in bipolar disorder and schizophrenia: potential relationship with latent *Toxoplasma Gondii* infection. *BMC Psychiatry.* 2018;18:239.
98. Nasteski VJHb. An overview of the supervised machine learning methods. 2017;4:51–62.
99. Mahendran N, Durai Raj Vincent PM, Srinivasan K, Chang C-Y. Machine learning based computational gene selection models: A survey, performance evaluation, open issues, and future research directions. *Front Genet.* 2020;11:603808. <https://doi.org/10.3389/fgene.2020.603808>.
100. Ahsan MM, Luna SA, Siddique Z. Machine-learning-based disease diagnosis: a comprehensive review. *Healthcare (Basel, Switzerland)* 2022;10:541. <https://doi.org/10.3390/healthcare10030541>.
101. International Consortium on Lithium G, Amare AT, Schubert KO, Hou L, Clark SR, Papiol S, et al. Association of polygenic score for schizophrenia and HLA antigen and inflammation genes with response to lithium in bipolar affective disorder: a genome-wide association study. *JAMA Psychiatry.* 2018;75:65–74.
102. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet.* 2019;51:793–803.
103. Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address pmhe, cross-disorder group of the psychiatric genomics C. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell.* 2019;179:1469–82.e1411.
104. Li HJ, Zhang C, Hui L, Zhou DS, Li Y, Zhang CY, et al. Novel risk loci associated with genetic risk for bipolar disorder among han chinese individuals: a genome-wide association study and meta-analysis. *JAMA Psychiatry.* 2021;78:320–30.
105. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet.* 2021;53:817–29.

ACKNOWLEDGEMENTS

JPB Foundation, Dolby Family Foundation, NIH U19 MH106434 for Prof. Fred H. Gage. The graphical image for this publication was created with [Biorender.com](https://biorender.com) and the figures were assembled in Adobe Illustrator.

AUTHOR CONTRIBUTIONS

LM and AC (co-first authors) contributed equally to this work. LM performed the analysis of data; machine learning predictor analysis, and helped in drafting the manuscript. AC performed experiments & analysis of the data; drafted and compiled the manuscript. PO and GG performed the experiments. EM and DG provided data for the second cohort used in the study. MA provided the clinical samples for the first cohort study. JRK provided resources for the study. FHG contributed to the study design and provided resources for the study. SS (corresponding author) conceptualized and supervised the study as well as analyzed data and drafted the manuscript. All authors reviewed the manuscript.

FUNDING

The Zuckerman STEM leadership program and Israel Science Foundation grants - 1994/21 and 3252 /21 for Dr. Shani Stern.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41380-023-02183-z>.

Correspondence and requests for materials should be addressed to Shani Stern.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023