Large scale differential gene expression analysis identifies genes associated with Bipolar Disorder

Mohamed N. Omar *[1], Mohamed Youssef †[2] and Mohamed Abdellatif ‡[2]

[1]Faculty of Medicine, Mansoura University

[2]Faculty of Medicine, Mansoura University

[2]Faculty of Medicine, Mansoura University

Abstract

Background

Bipolar disorder (BD) remains one of the most common psychiatric disorders with high morbidity and mortality. Several polymorphisms have been found to be implicated in the pathogenesis of BD, however, these loci have small effect sizes that fail to explain the high heritability of the disease. Here, we attempt to better understand the genetic basis of

*mohamed.n.omar1993@gmail.com

†mohamedyou22@gmail.com

‡M.Elmorsy2018@gmail.com

bipolar disorder by identifying differentially expressed genes (DEG) in post-mortem brain tissues of patients with BD versus normal controls.

Methods

Nine publicly available gene expression data sets with 702 samples derived from human post-mortem brain tissues were used for differential expression analysis. 7 data sets were used for the discovery of the gene signature and the other two were left for independent validation. The multi-cohort analysis was performed using a random-effects model utilizing R and MetaIntegrator package.

Results

The analysis resulted in the identification of 13 up-regulated and 9 down-regulated genes in BD. The BD gene signature was further validated in two independent data sets and resulted in an Area Under the ROC Curve (AUC) of 0.74 and 0.73, respectively. Gene set enrichment analysis was performed and resulted in the identification of several biological processes and pathways related to BD including Ca transport, inflammation and DNA damage response.

Conclusion

Our findings support the previous findings that link BD pathogenesis to abnormalities in glial inflammation and calcium transport and also identify several other biological processes not previously reported to be associated with the development of BD. Such findings will improve our understanding of the genetic basis underlying BD and may have future clinical implications.

Keywords— Bipolar Disorder, differential expression, inflammation, calcium transport

---

Background

Bipolar disorder (BD) is a common psychiatric disorder characterized by alternating periods of depression and abnormally elevated mood associated with changes in the activity levels [1]. The lifetime prevalence of Bipolar disorder is approximately 2.4% with average age of onset of 25 years [2].

BD is a highly heritable disease with overall heritability reaching up to 70% based on twin studies [3, 4] with the risk being 10-fold higher in the first degree relatives of those affected with BD when compared with normal population. The biological mechanisms underlying the etiology of bipolar disorder remain unclear with recent evidence pointing toward a significant role of disrupted Ca dynamics and oxidative DNA damage in the pathogenesis of BD [5, 6, 7]. Although several susceptibility loci with small effect sizes have been found to be associated with BD [8, 9], there is no much evidence supporting the role of specific genes in the pathogenesis of the disease [10].

In this study, we attempt to identify the specific genes and biological pathways associated with bipolar disorder. For this purpose, we analyze microarray gene expression data derived from post-mortem brain tissue to identify the deferentially expressed genes (DEG) between BD and normal cases. Detecting those genes and pathways would provide us with a better understanding of the molecular pathogenesis of this disease and may have a future therapeutic implications.

Data collection

We performed an extensive search of the Gene Expression Omnibus (GEO) database [11]looking for microarray gene expression data of bipolar disorder. A total of 9 data sets were identified (Table 1) with 702 human post-mortem brain samples. We used 7 data sets for discovery of the BD gene signature and kept 2 data sets out for independent validation of the gene signature performance.

The meta-analysis

We used R Programming language [12] and "Metaintegrator" package [13] which utilizes a gene expression meta-analysis workflow described by Haynes et al [14]. This approach computes a Hedges effect size [15]for each gene in each data set. These effect sizes are then pooled across all the data sets (see Additional file 1 and 2) using a random-effect model by assuming that results from each study is drawn from a single distribution and that each inter-study difference is just a random effect. The approach computes the log sum of p-values that each gene is up/down-regulated, then combine p-values using Fisher's method and finally performs Benjamini-Hochberg FDR correction across all genes [16].

A gene would be deferentially expressed if the absolute value of its effect size is > 0, false discovery rate is $<= 0.15$ and it has to be significantly up/down regulated in all of the 7 training data sets with a heterogeneity P-value cutoff of 0.05 [17]. All discovery steps were conducted on the training data sets only.

Validation of the resulting gene signature

The performance of the resulting BD gene signature was further validated in the 2 independent cohorts with a total of 156 samples (Table 1). We used the Area Under the Receiver operating characteristic curve (AUC), Area Under the Precision Recall Curve (AUPRC)

and accuracy as evaluation metrics.

Enrichment analysis

Functional Enrichment analysis was performed using Enrichr R package [18, 19] to identify the Gene Ontology (GO) biological processes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in the discovered genes.

Results

The meta-analysis process resulted in the identification of 13 up-regulated and 9 down-regulated genes in the post-mortem brain samples of bipolar disorder cases compared with normal controls (Table 2).

Strong performance in the two validation data sets

We measured the discriminatory ability of our gene signature by applying it to the 2 independent data sets to assess its performance on unseen data (Figure 1).

In the first validation data set (GSE5388), the signature achieved an AUC of 0.74, AUPRC of 0.72, accuracy of 74% and sensitivity of 90%. In the second validation data set (GSE35977), the signature achieved an AUC of 0.73, AUPRC of 0.664, accuracy of 68.4% and sensitivity of 89%.

The signature performance in the validation data sets showed great consistency with the performance in the discovery data sets without overfitting (Summary AUC = 0.74, 95% CI = 0.67-0.80) (Figure 2).

Enrichment analysis

Enrichment analysis has revealed that the discovered gene signature is incorporated in many significant KEGG pathways (Figure 3, Figure 4) and GO biological processes (Additional file 3, Additional file 4). Several of these biological processes were previously reported to be associated with bipolar disorder and other psychiatric disease, for example: inflammation, apoptosis, regulation of DNA damage response and Ca transport. We confirm this association and report other pathways and biological processes not previously reported to be associated with BD.

Discussion

Bipolar disorder is a highly heterogeneous disease and the genetic basis underlying its pathophysiology is complex [20]. Several genome-wide association studies (GWAS) have identified several copy number variants (CNVs) and single-nucleotide polymorphisms (SNPs) that seem to play a role in the high heritability of the disease but each of these variants has a small effect [21, 22]. Moreover, these studies failed to identify specific genes that may have a large effect in the pathogenesis of the disease, most likely due to small sample size.

Here we tried to overcome the problem of inadequate sample size by incorporating multiple microarray gene expression studies in a meta-analysis process to achieve high statistical power. Through this large scale gene expression analysis of 702 post-mortem brain samples, we managed to identify 13 up-regulated and 9 down-regulated genes in BD. The resulting gene signature managed to achieve great performance in the two independent validation data sets. In the first validation cohort, the signature achieved an AUC of 0.74 and AUPRC of 0.72. In the second validation cohort, the signature achieved an AUC of 0.73 and AUPRC of 0.664.

The up-regulated genes were enriched in many biological processes [see Additional file 3] like Ca ion transport, cytoskeletal assembly, apoptotic signaling, intracellular signal transduction, cell migration, inflammation and response to DNA damage. Several of these biological processes have been reported to be associated with BD and other psychiatric disorders like Schizophrenia [23, 24, 25, 26]. These same processes are incorporated in the pathogenesis of atherosclerosis which may explain part of the association between BD and other cardiovascular diseases [27, 28, 29].

The 9 down-regulated genes were enriched in important biological processes and pathways [see Additional file 4] like the positive regulation of lymphocyte proliferation, antigen processing/presentation and regulation of macrophages cytokines production supporting the notion that neuronal and glial inflammation may have a role in the pathogenesis of BD.

Conclusion

In this study, we provide more insights into the genetic basis of BD through the identification of differentially expressed genes in post-mortem brain tissues of BD patients compared with normal controls. Our discovered BD gene signature proved to be accurate and consistent as shown by its performance in the independent validation cohorts. The BD signature genes were enriched in a number of biological processes and pathways that were previously reported to have a role in the pathogenesis of BD, most importantly the pathways regulating Ca transport and inflammatory processes. Such findings, together with previous research will provide a more clear understanding of the molecular pathophysiology of BD and may have future therapeutic potentials by targeting these specific genes and pathways.

List of abbreviations

- **BD**: Bipolar Disorder.
- **DEG**: Differentially Expressed Genes.
- **ROC**: Receiver Operating Characteristic.
- **AUC**: Area Under the ROC Curve.
- **GEO**: Gene Expression Omnibus.
- **GO**: Gene Ontology.
- **KEGG**: Kyoto Encyclopedia of Genes and Genomes.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data sets used in this study are publicly available on the Gene Expression Omnibus (GEO) data base under the corresponding accession number. The code for this analysis is available on GitHub and can accessed using the following link: https://github.com/MohamedOmar2020/BD

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was received for this study.

Author's contributions

MO conceived the research question and performed the analysis. MY and MA collected the data sets. All authors contributed equally to the manuscript writing, read and approved the final manuscript.

Acknowledgements

Not applicable

References

[1] Ian M Anderson, Peter M Haddad, and Jan Scott. Bipolar disorder. *BMJ*, 345, 2012.

[2] Lakshmi N Yatham, Sidney H Kennedy, Sagar V Parikh, Ayal Schaffer, David J Bond, Benicio N Frey, Verinder Sharma, Benjamin I Goldstein, Soham Rej, Serge Beaulieu, Martin Alda, Glenda MacQueen, Roumen V Milev, Arun Ravindran, Claire O'Donovan, Diane McIntosh, Raymond W Lam, Gustavo Vazquez, Flavio Kapczinski, Roger S McIntyre, Jan Kozicky, Shigenobu Kanba, Beny Lafer, Trisha Suppes, Joseph R Calabrese, Eduard Vieta, Gin Malhi, Robert M Post, and Michael Berk. Canadian network for mood and anxiety treatments (canmat) and international society for bipolar disorders (isbd) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disord*, 20(2):97–170, Mar 2018.

[3] Jack Edvardsen, Svenn Torgersen, Espen Roysamb, Sissel Lygren, Ingunn Skre, Sidsel Onstad, and Per Anders Oien. Heritability of bipolar spectrum disorders. unity or heterogeneity? *J Affect Disord*, 106(3):229–240, Mar 2008.

[4] Nick Craddock and Pamela Sklar. Genetics of bipolar disorder. *Lancet*, 381(9878):1654–1662, May 2013.

[5] Michael J Berridge. Calcium signalling and psychiatric disease: bipolar disorder and schizophrenia. *Cell Tissue Res*, 357(2):477–492, Aug 2014.

[6] Xiaoyun Guo, Dengtang Liu, Tong Wang, and Xingguang Luo. Aetiology of bipolar disorder: contribution of the l-type voltage-gated calcium channels. *Gen Psychiatr*, 32(2):e100009, 2019.

[7] Klaus Munkholm, Henrik Enghusen Poulsen, Lars Vedel Kessing, and Maj Vinberg. Elevated levels of urinary markers of oxidatively generated dna and rna damage in bipolar disorder. *Bipolar Disord*, 17(3):257–268, May 2015.

[8] Seth A Ament, Szabolcs Szelinger, Gustavo Glusman, Justin Ashworth, Liping Hou, Nirmala Akula, Tatyana Shekhtman, Judith A Badner, Mary E Brunkow, Denise E Mauldin, Anna-Barbara Stittrich, Katherine Rouleau, Sevilla D Detera-Wadleigh, John I Jr Nurnberger, Howard J Edenberg, Elliot S Gershon, Nicholas Schork, Nathan D Price, Richard Gelinas, Leroy Hood, David Craig, Francis J McMahon, John R Kelsoe, and Jared C Roach. Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proc Natl Acad Sci U S A*, 112(11):3576–3581, Mar 2015.

[9] M Kataoka, N Matoba, T Sawada, A-A Kazuno, M Ishiwata, K Fujii, K Matsuo, A Takata, and T Kato. Exome sequencing for bipolar disorder points to roles of de novo loss-of-function and protein-altering mutations. *Mol Psychiatry*, 21(7):885–893, Jul 2016.

[10] Michael C O'Donovan and Michael J Owen. The implications of the shared genetics of psychiatric disorders. *Nat Med*, 22(11):1214–1219, Nov 2016.

[11] Tanya Barrett, Stephen E. Wilhite, Pierre Ledoux, Carlos Evangelista, Irene F. Kim, Maxim Tomashevsky, Kimberly A. Marshall, Katherine H. Phillippy, Patti M. Sherman, Michelle Holko, Andrey Yefanov, Hyeseung Lee, Naigong Zhang, Cynthia L. Robertson, Nadezhda Serova, Sean Davis, and Alexandra Soboleva. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research*, 41(D1):D991–D995, 11 2012.

[12] Ross Ihaka and Robert Gentleman. R: A language for data analysis and graphics. 5(3):299–314, 1996.

[13] Purvesh Khatri, Silke Roedder, Naoyuki Kimura, Katrien De Vusser, Alexander A. Morgan, Yongquan Gong, Michael P. Fischbein, Robert C. Robbins, Maarten Naesens, Atul J. Butte, and Minnie M. Sarwal. A common rejection module (crm) for acute rejection across multiple organs identifies novel therapeutics for organ transplantation. *Journal of Experimental Medicine*, 210(11):2205–2221, 2013.

[14] Winston A Haynes, Francesco Vallania, Charles Liu, Erika Bongen, Aurelie Tomczak, Marta Andres-Terrè, Shane Lofgren, Andrew Tam, Cole A Deisseroth, Matthew D Li, Timothy E Sweeney, and Purvesh Khatri. Empowering multi-cohort gene expression analysis to increase reproducibility. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 22:144–153, 2017.

[15] Larry V. Hedges. Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2):107–128, 1981.

[16] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 2019/07/24 1995.

[17] David C. Hoaglin. Misunderstandings about q and 'cochran's q test' in meta-analysis. *Statistics in Medicine*, 35(4):485–495, 2016.

[18] Maxim V Kuleshov, Matthew R Jones, Andrew D Rouillard, Nicolas F Fernandez, Qiaonan Duan, Zichen Wang, Simon Koplev, Sherry L Jenkins, Kathleen M Jagodnik, Alexander Lachmann, Michael G McDermott, Caroline D Monteiro, Gregory W

Gundersen, and Avi Ma'ayan. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*, 44(W1):W90–7, Jul 2016.

[19] Edward Y Chen, Christopher M Tan, Yan Kou, Qiaonan Duan, Zichen Wang, Gabriela Vaz Meirelles, Neil R Clark, and Avi Ma'ayan. Enrichr: interactive and collaborative html5 gene list enrichment analysis tool. *BMC Bioinformatics*, 14:128, Apr 2013.

[20] Nick Craddock and Pamela Sklar. Genetics of bipolar disorder. *Lancet*, 381(9878):1654–1662, May 2013.

[21] Liping Hou, Sarah E Bergen, Nirmala Akula, Jie Song, Christina M Hultman, Mikael Landén, Mazda Adli, Martin Alda, Raffaella Ardau, Bárbara Arias, Jean-Michel Aubry, Lena Backlund, Judith A Badner, Thomas B Barrett, Michael Bauer, Bernhard T Baune, Frank Bellivier, Antonio Benabarre, Susanne Bengesser, Wade H Berrettini, Abesh Kumar Bhattacharjee, Joanna M Biernacka, Armin Birner, Cinnamon S Bloss, Clara Brichant-Petitjean, Elise T Bui, William Byerley, Pablo Cervantes, Caterina Chillotti, Sven Cichon, Francesc Colom, William Coryell, David W Craig, Cristiana Cruceanu, Piotr M Czerski, Tony Davis, Alexandre Dayer, Franziska Degenhardt, Maria Del Zompo, J Raymond DePaulo, Howard J Edenberg, Bruno Étain, Peter Falkai, Tatiana Foroud, Andreas J Forstner, Louise Frisén, Mark A Frye, Janice M Fullerton, Sébastien Gard, Julie S Garnham, Elliot S Gershon, Fernando S Goes, Tiffany A Greenwood, Maria Grigoroiu-Serbanescu, Joanna Hauser, Urs Heilbronner, Stefanie Heilmann-Heimbach, Stefan Herms, Maria Hipolito, Shashi Hitturlingappa, Per Hoffmann, Andrea Hofmann, Stephane Jamain, Esther Jiménez, Jean-Pierre Kahn, Layla Kassem, John R Kelsoe, Sarah Kittel-Schneider, Sebastian Kliwicki, Daniel L Koller, Barbara König, Nina Lackner, Gonzalo Laje, Maren Lang,

Catharina Lavebratt, William B Lawson, Marion Leboyer, Susan G Leckband, Chunyu Liu, Anna Maaser, Pamela B Mahon, Wolfgang Maier, Mario Maj, Mirko Manchia, Lina Martinsson, Michael J McCarthy, Susan L McElroy, Melvin G McInnis, Rebecca McKinney, Philip B Mitchell, Marina Mitjans, Francis M Mondimore, Palmiero Monteleone, Thomas W Mühleisen, Caroline M Nievergelt, Markus M Nöthen, Tomas Novák, John I Jr Nurnberger, Evaristus A Nwulia, Urban Ösby, Andrea Pfennig, James B Potash, Peter Propping, Andreas Reif, Eva Reininghaus, John Rice, Marcella Rietschel, Guy A Rouleau, Janusz K Rybakowski, Martin Schalling, William A Scheftner, Peter R Schofield, Nicholas J Schork, Thomas G Schulze, Johannes Schumacher, Barbara W Schweizer, Giovanni Severino, Tatyana Shekhtman, Paul D Shilling, Christian Simhandl, Claire M Slaney, Erin N Smith, Alessio Squassina, Thomas Stamm, Pavla Stopkova, Fabian Streit, Jana Strohmaier, Szabolcs Szelinger, Sarah K Tighe, Alfonso Tortorella, Gustavo Turecki, Eduard Vieta, Julia Volkert, Stephanie H Witt, Adam Wright, Peter P Zandi, Peng Zhang, Sebastian Zollner, and Francis J McMahon. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Hum Mol Genet*, 25(15):3383–3394, Aug 2016.

[22] Consortium Cross-Disorder Group of the Psychiatric Genomics. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*, 381(9875):1371–1379, Apr 2013.

[23] F M Benes, D Matzilevich, R E Burke, and J Walsh. The expression of proapoptosis genes is increased in bipolar disorder, but not in schizophrenia. *Mol Psychiatry*, 11(3):241–251, Mar 2006.

[24] C J Carter. Multiple genes and factors associated with bipolar disorder converge on growth factor and stress activated kinase pathways controlling translation initiation:

implications for oligodendrocyte viability. *Neurochem Int*, 50(3):461–490, Feb 2007.

[25] Christopher J. Carter. EIF2B and Oligodendrocyte Survival: Where Nature and Nurture Meet in Bipolar Disorder and Schizophrenia? *Schizophrenia Bulletin*, 33(6):1343–1353, 02 2007.

[26] Roos C Padmos, Manon H J Hillegers, Esther M Knijff, Ronald Vonk, Anne Bouvy, Frank J T Staal, Dick de Ridder, Ralph W Kupka, Willem A Nolen, and Hemmo A Drexhage. A discriminating messenger rna signature for bipolar disorder formed by an aberrant expression of inflammatory genes in monocytes. *Arch Gen Psychiatry*, 65(4):395–407, Apr 2008.

[27] Marion Leboyer, Isabella Soreca, Jan Scott, Mark Frye, Chantal Henry, Ryad Tamouza, and David J Kupfer. Can bipolar disorder be viewed as a multi-system inflammatory disease? *J Affect Disord*, 141(1):1–10, Dec 2012.

[28] Miriam Weiner, Lois Warren, and Jess G Fiedorowicz. Cardiovascular morbidity and mortality in bipolar disorder. *Ann Clin Psychiatry*, 23(1):40–47, Feb 2011.

[29] Giuseppe Maina, Virginio Salvi, Alberto Vitalucci, Virginia D'Ambrosio, and Filippo Bogetto. Prevalence and correlates of overweight in drug-naive patients with bipolar disorder. *J Affect Disord*, 110(1-2):149–155, Sep 2008.

[30] Hamid M Abdolmaleky, Adam C Gower, Chen-Khuan Wong, Jiayi W Cox, Xiaoling Zhang, Arunthathi Thiagalingam, Rahim Shafa, Vadivelu Sivaraman, Jin-Rong Zhou, and Sam Thiagalingam. Aberrant transcriptomes and dna methylomes define pathways that drive pathogenesis and loss of brain laterality/asymmetry in schizophrenia and bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet*, 180(2):138–149, Mar 2019.

[31] Dominique Arion, Zhiguang Huo, John F Enwright, John P Corradi, George Tseng, and David A Lewis. Transcriptome alterations in prefrontal pyramidal cells distinguish schizophrenia from bipolar and major depressive disorders. *Biol Psychiatry*, 82(8):594–600, Oct 2017.

[32] K Iwamoto, C Kakiuchi, M Bundo, K Ikeda, and T Kato. Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol Psychiatry*, 9(4):406–416, Apr 2004.

[33] Angelica de Baumont, Mariana Maschietto, Leandro Lima, Dirce Maria Carraro, Eloisa Helena Olivieri, Alex Fiorini, Luiz Andre Nardin Barreta, Joana Almeida Palha, Paulo Belmonte-de Abreu, Carlos Alberto Moreira Filho, and Helena Brentani. Innate immune response is differentially dysregulated between bipolar disease and schizophrenia. *Schizophr Res*, 161(2-3):215–221, Feb 2015.

[34] Megan Hastings Hagenauer, Anton Schulmann, Jun Z Li, Marquis P Vawter, David M Walsh, Robert C Thompson, Cortney A Turner, William E Bunney, Richard M Myers, Jack D Barchas, Alan F Schatzberg, Stanley J Watson, and Huda Akil. Inference of cell type content from human brain transcriptomic datasets illuminates the effects of age, manner of death, dissection, and psychiatric diagnosis. *PLoS One*, 13(7):e0200003, 2018.

[35] M M Ryan, H E Lockstone, S J Huffaker, M T Wayland, M J Webster, and S Bahn. Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes. *Mol Psychiatry*, 11(10):965–978, Oct 2006.

[36] C Chen, L Cheng, K Grennan, F Pibiri, C Zhang, J A Badner, E S Gershon, and C Liu. Two gene co-expression modules differentiate psychotics and controls. *Mol Psychiatry*, 18(12):1308–1314, Dec 2013.

[37] Chao Chen, Qingtuan Meng, Yan Xia, Chaodong Ding, Le Wang, Rujia Dai, Lijun Cheng, Preethi Gunaratne, Richard A Gibbs, Shishi Min, Cristian Coarfa, Jeffrey G Reid, Chunling Zhang, Chuan Jiao, Yi Jiang, Gina Giase, Amber Thomas, Dominic Fitzgerald, Tonya Brunetti, Annie Shieh, Cuihua Xia, Yongjun Wang, Yunpeng Wang, Judith A Badner, Elliot S Gershon, Kevin P White, and Chunyu Liu. The transcription factor pou3f2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. *Sci Transl Med*, 10(472), Dec 2018.

Figures

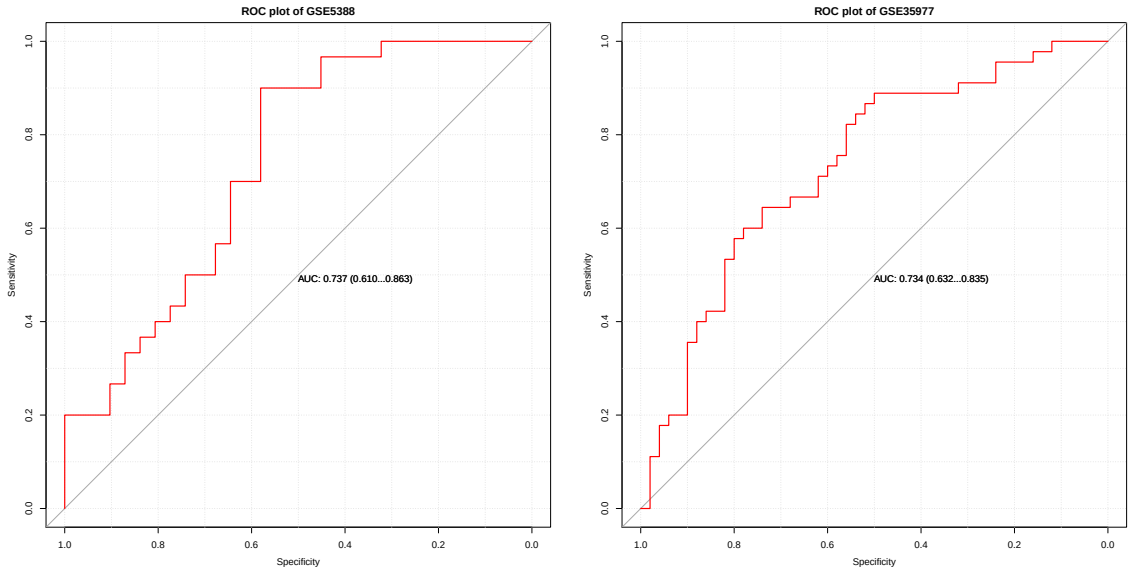

Figure 1: **ROC curves of the two validation data sets.** Strong classification performance (AUC = 0.74 and 0.73 respectively).

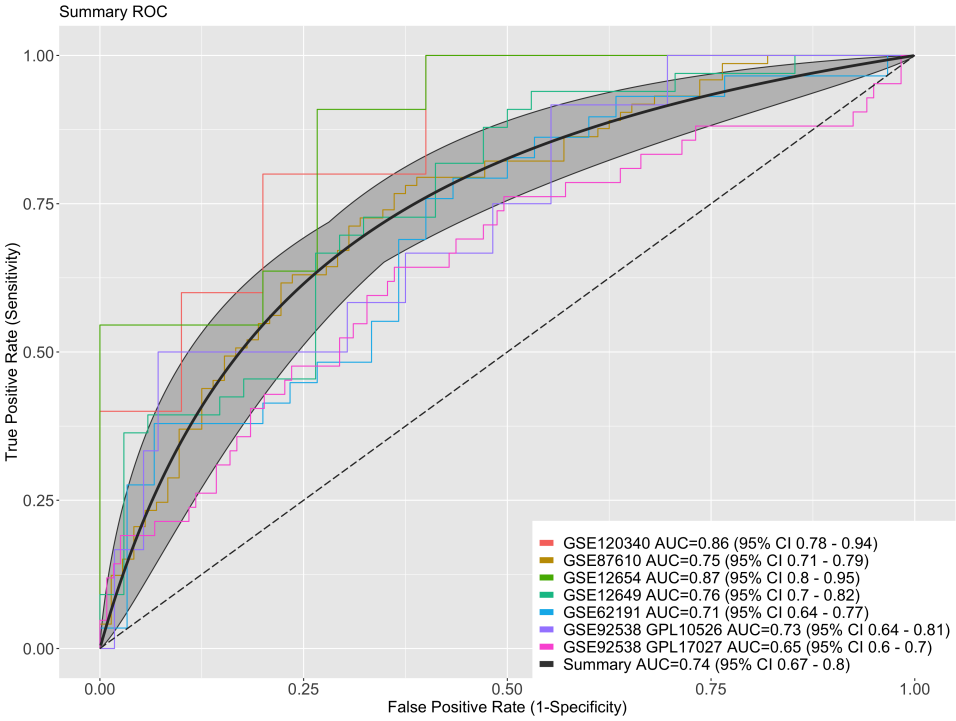

Figure 2: **ROC curves of the 7 discovery data sets.** Notice the summary AUC of 0.74 which is identical to the AUC in each of the validation data sets (no overfitting)

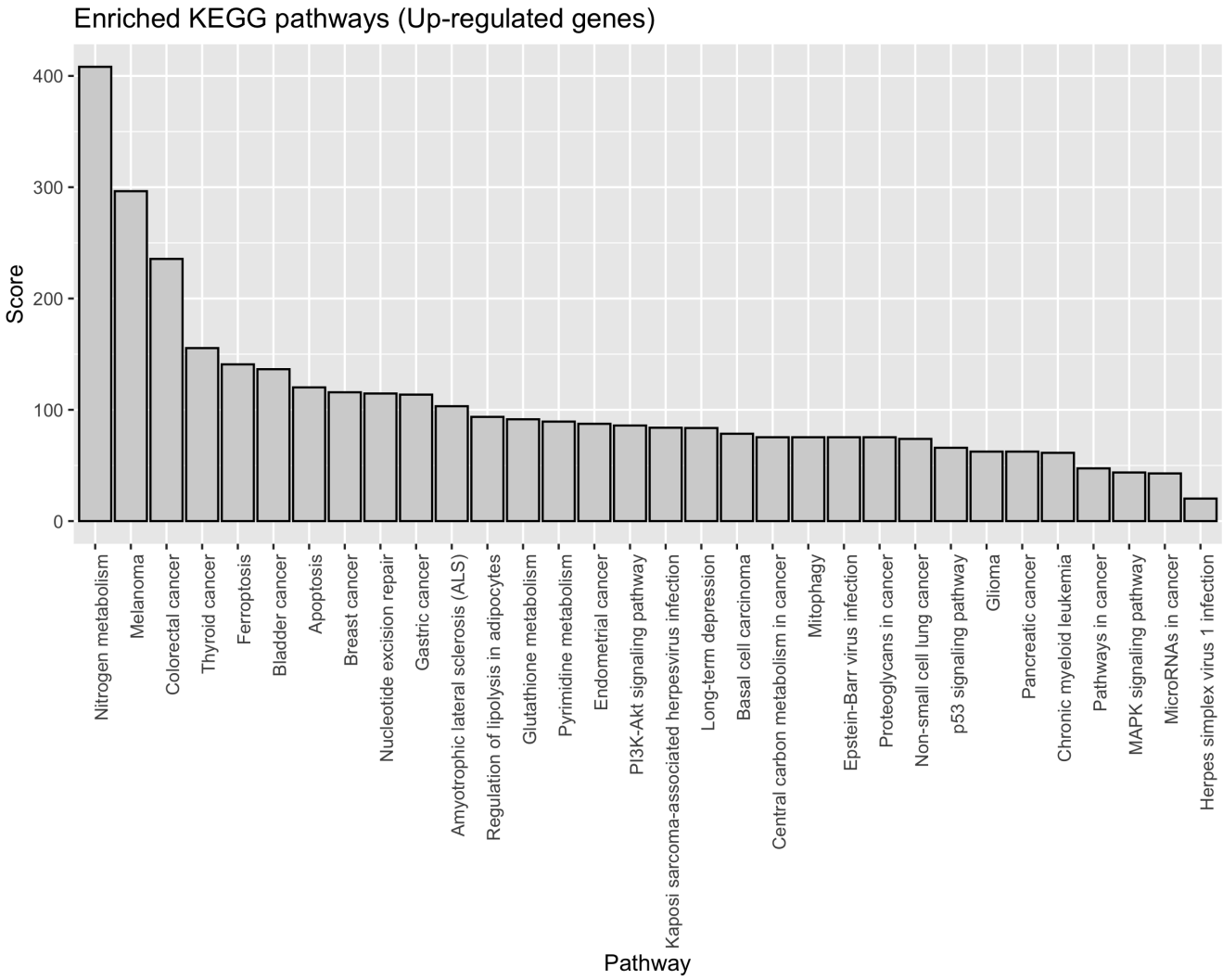

Figure 3: KEGG pathways of the 13 up-regulated genes.

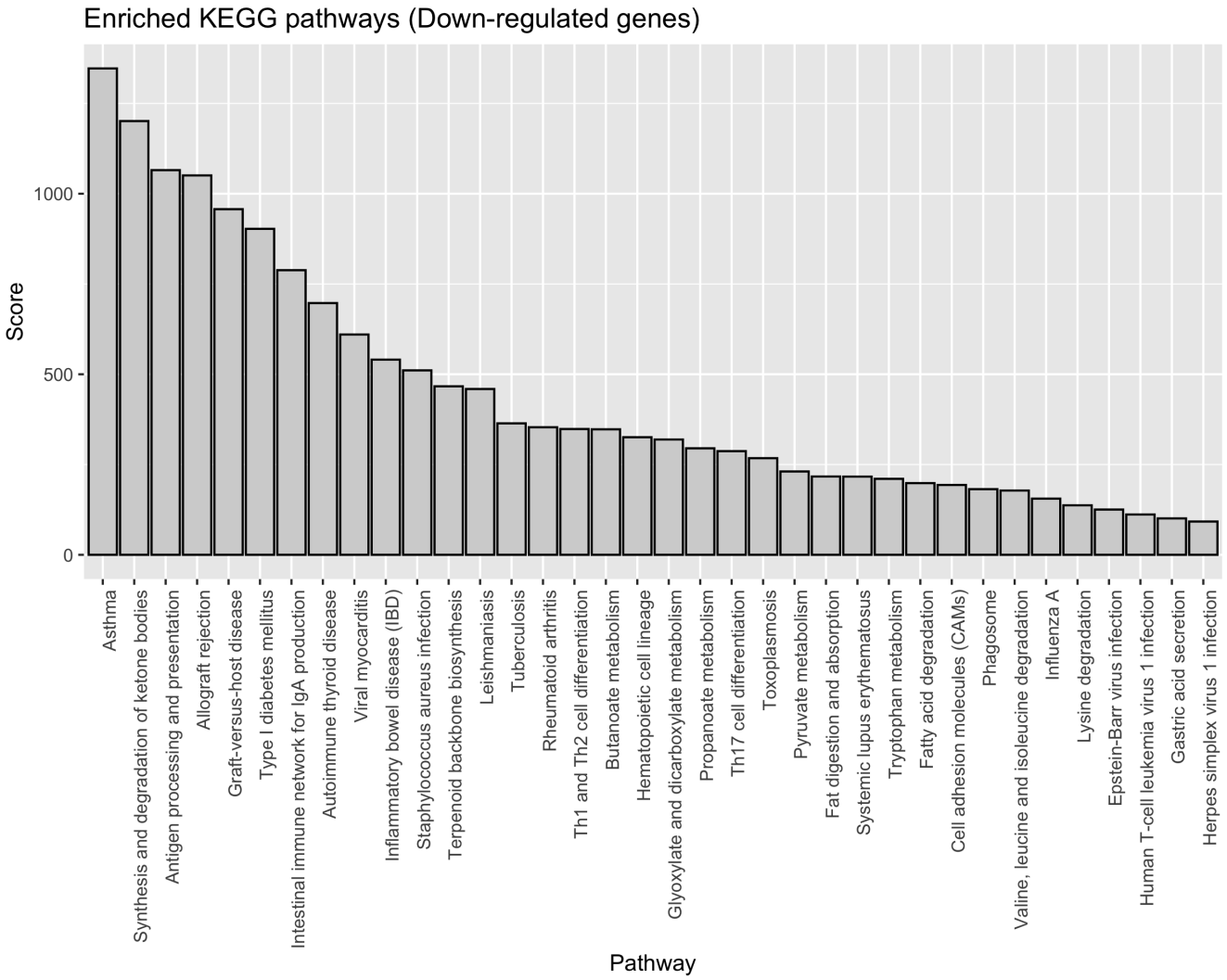

Figure 4: KEGG pathways of the 9 down-regulated genes.

Tables

Table 1: Collected data sets showing the number of samples with the number of bipolar disease (BD) cases. 7 data sets were used for discovery of the BD gene signature and 2 data sets were used for independent validation

GEO accession	Type of experiment	Number of samples (BD cases)	Tissue source
		Discovery	
GSE120340 [30]	Expression profiling by array	20 (10)	Dorsolateral prefrontal cortex
GSE87610 [31]	Expression profiling by array	145 (73)	Dorsolateral prefrontal cortex
GSE12654 [32]	Expression profiling by array	26 (11)	Anterior prefrontal cortex
GSE12649 [32]	Expression profiling by array	67 (33)	Dorsolateral prefrontal cortex
GSE62191 [33]	Expression profiling by array	59 (29)	Frontal cortex
GSE92538-GPL10526 [34]	Expression profiling by array	68 (12)	Dorsolateral prefrontal cortex
GSE92538-GPL17027 [34]	Expression profiling by array	161 (42)	Dorsolateral prefrontal cortex
		Validation	

Table 1: Collected data sets showing the number of samples with the number of bipolar disease (BD) cases. 7 data sets were used for discovery of the BD gene signature and 2 data sets were used for independent validation

GEO accession	**Type of experiment**	**Number of samples (BD cases)**	**Tissue source**
GSE5388 [35]	Expression profiling by array	61 (30)	Dorsolateral prefrontal cortex
GSE35977 [36, 37]	Expression profiling by array	95 (45)	Parietal cortex

Table 2: BD signature genes. 13 up-regulated and 9 down-regulated genes

Gene Symbol	**Effect size**	**Effect size p-value**
	Up-regulated genes	
MYO6	0.233803528781553	3.92088529079236e-05
ERCC8	0.223683432497003	5.06105158365377e-05
SLC12A4	0.153840135678242	0.000140523441674623
CA12	0.249039212602444	0.000169533685049415
ZNF273	0.287473913205269	0.000202310305620391
TP53	0.166740750366376	0.000301337862141146
MYRIP	0.240157780766432	0.000306767694381059

Table 2: BD signature genes. 13 up-regulated and 9 down-regulated genes

Gene Symbol	**Effect size**	**Effect size p-value**
FGF2	0.230057248064757	0.000314051978487104
BCL2L11	0.14089492142396	0.000369520305484846
COL11A1	0.279394249891102	0.000763160058191395
PRKG1	0.194638754941504	0.000863840829673692
RRM1	0.197841881310803	0.000948698091142044
PAQR3	0.18648054393229	0.00097136180697817
	Down-regulated genes	
CD74	-0.314560020628663	3.71498955514197e-07
CNP	-0.231525261417171	4.41001367269943e-06
HLA-DPB1	-0.183506935227843	1.17960248475937e-05
DDX17	-0.318235420127414	2.63152971729961e-05
HLA-DPA1	-0.268023572622133	5.76742550284089e-05
MRPL33	-0.260414913702845	0.000125094044339841
CCKBR	-0.327943659949979	0.000258927502648349
ACAT2	-0.330687000537771	0.000359707958498527
CD53	-0.359439343217732	0.000366805911760364

Additional Files

Additional File 1— EffectSizes-UpRegulated

PDF file showing the effect sizes of each of the 13 up-regulated genes.

Additional File 2— EffectSizes-DownRegulated

PDF file showing the effect sizes of each of the 9 down-regulated genes.

Additional File 3 — GO-BP-UpRegulated

CSV table showing the GO biological processes of the 13 up-regulated genes.

Additional File 4— GO-BP-DownRegulated

CSV table showing the GO biological processes of the 9 down-regulated genes.