bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Identifying Changes to RNA Editome in Major Depressive Disorder

Saam Hasan[1, 2], **Shahriyar Mahdi Robbani**[1, 2], **Tamanna Afroze**[2], **Giasuddin Ahsan**[3]

Muhammad Maqsud Hossain[1, 2*]

[1]Department of Biochemistry and Microbiology, North South University, Dhaka, Bangladesh

[2]NSU Genome Research Institute, North South University, Dhaka, Bangladesh

[3]Department of Public Health, North South University, Dhaka, Bangladesh

*Corresponding author:

Saam Hasan

Mailing address: Plot 15, Block B, Bashundhara, Dhaka 1229, Bangladesh

Mobile phone: +88-01829153284

E-mail: muhammad.maqsud@northsouth.edu

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Major Depressive Disorder (MDD) is one of the most significant psychiatric disorders in the world today. Its incidence is widespread in society and its heavy adverse impact on the quality of life is well documented. Previously genetic studies on MDD had identified a hereditary component of the disease as well as crediting RNA editing with a role in its development. The later due to an overexpression of a heavily edited isoform of the Serotonin 2c receptor. Here we used publicly available RNA sequence data from suicide patients diagnosed with MDD as well as controls for identifying RNA editing sites unique to MDD. After variant calling and several steps of filtering, we identified 142 unique RNA editing sites in the MDD patients. These included intronic, downstream, UTR3 and exonic edits. The latter comprising several amino acid changes in the encoded protein. The genes implicated to be uniquely edited in MDD included the aforementioned and previously implicated Serotonin 2c receptor, others involved in functions that play roles in depression and suicide such as Cannabinoid Receptor 1, Frizzled Class 3 Receptor, Neuroligin 3 and others.

Keywords: RNA Editing, Major Depressive Disorder, Suicide, Post Transcriptional

Modification

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Major Depressive Disorder represents one of the most statistically and clinically significant psychiatric disorders in the world. Its adverse impact on the quality of life has already been well documented by research (Kapp, 2001). MDD significantly reduces work capacity through its affect on the willingness to perform and work and through the physical damage it does through body function dysregulations. Its impact also extends to personal life as individuals suffering from MDD tend to struggle at maintaining and keeping up with personal relations. The disorder displays both chronic, where it persists for up to a duration three months or even more, and episodic versions which comprises of separate episodes beginning and ending independently of others. The major observable symptom of the condition is unhealthy depressed state of mind that hampers most activities as well a loss of positive emotions and pleasure towards most things in life (Patten et al., 2009). Diagnosis and identification of individuals suffering from MDD relies on the observation of symptoms in a given time span. This is taken as observing five or more symptoms associated with MDD within a span of five days. The symptoms included as a measure of MDD are as per the *Diagnostic and Statistical Manual of Mental Disorders* (Diagnostic and statistical manual of mental disorders, 2002). Research has in the past identified possible genetic factors predisposing one to the risk of developing MDD. This has identified heritable and familial factors as contributing to the development of the condition. The latter showing that genetic factors acquired through inheritance is indeed a contributing factor to the disease. This is further backed up by twin studies (Lohoff, 2010). Another kind of genetic change that has been speculated as a contributing factor albeit to a lesser degree is RNA editing. RNA editing is a form of post transcriptional change to RNA (Nishikura, 2010). It has been shown to play a role in regulation of several disease as well as increasing complexity of the transcriptome (Park et al., 2017). The only previous connection between MDD and RNA editing was the upregulation of a heavily edited isoform of the HTC2R gene which is a serotonin receptor (Lyddon et al., 2012). Although that remains a clear connection to psychiatric disorders and MDD, no specific variants or editing sites were identified specific to the condition. Here we make the first attempt at characterizing overall changes in the human RNA editome in depression and look to identify editing sites specific to MDD. We used data sequenced by Labonté et al in their study published in 2017 on sex specific transcriptional regulation in Depression (Labonté et al., 2017). We used the RNA sequence data from both MDD and Controls to compare for Editing sites found only in the MDD models. Our major goals of this study were to identify changes in editing patterns in MDD compared to control individuals. To compare the frequency of editing in the two models. And to identify specific editing sites that can serve as potential diagnostic markers for MDD and potentially suicidal behaviour.

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Methods

The data used was downloaded from NCBI SRA database using the NCBI SRA Toolkit. The prefetch and fastq-dump functions were used for this purpose (Leinonen, Sugawara and Shumway, 2010). We chose RNA sequence data from the Nucleus Acumbens region owing to its functional significance in mood and behaviour regulation (Pavuluri, Volpe and Yuen, 2017). Table 1 shows the accession numbers of the runs we used. In total five samples chosen were individuals diagnosed with MDD and committed suicide and two were controls who died naturally.

Table 1:Detals and SRA Accession Numbers of Data Used

Run Number	Phenotype	Cause of Death	SRA Accession Number
1	MDD	Suicide	SRR5962012
3	MDD	Suicide	SRR5961992
4	Control	Natural	SRR5961986
5	MDD	Suicide	SRR5961999
6	Control	Natural	SRR5961987
12	MDD	Suicide	SRR5961995
13	MDD	Suicide	SRR5961996

We followed a modified version of the GATK best practices pipeline for variant discovery from RNA sequence data. STAR was chosen as the mapping tool. We used two pass STAR which is designed for better mapping splice junctions. We set the mismatch parameter to 3 and overhang to 49 for matching the fastq files which had 50 base pair read lengths (Dobin et al., 2012). Afterwards we used the GATK pipeline for calling variants (DePristo et al., 2011). This included started with adding read groups and marking duplicates with picard. Then we split N Cigar reads using GATK's SplitNCigarReads functionality, recalibrated bases with BaseRecalibrator and finally called variants with HaplotypeCaller. We then split the SNPs and indels, keeping only the former using SelectVariants. Indels do not represent editing sites and hence we didn't consider those variants any further. The variants were then analysed with the Mann-Whitney-Wilcoxon Rank Sum Test using the BaseQRankSum function of GATK's VariantAnnotator. These were subsequently marked for common variants having dbsnp entries using VariantAnnotator again before being filtered for SNP clusters of 3 SNPs within 35 base pair windows using VariantFiltration. The variants were then annotated using Annovar developed by Wang labs. Annovar's built in RefGene database was used for annotating variants (Wang, Li and Hakonarson, 2010). Afterwards we filtered variants for actual editing sites from SNPs in a series of steps. Through STAR's mapping quality filters we had already filtered out most possible wrong base calls during sequencing. Now we first found the common variants between all the data from individuals with MDD who committed suicide. This followed the rationale from Ramaswami et al who established that any rare SNP is unlikely to be common between two separate individuals being studied and any common variants thus are likely to be editing sites

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Ramaswami et al., 2012). Afterwards we removed any editing sites that were also found in either of the controls. Next we used the rationale followed by Zhang and Xiao that allelic linkage should not be observed in editing sites in contrast to genomic SNPs (Zhang and Xiao, 2015). We took the predicted genotype for each variant from HaplotyCaller and compared it to the observed number of alleles of reads harboring that particular variant. Variants where the genotype and allele ratio were in agreement were then discarded as possible genomic SNPs. The rest were retained and subsequently filtered again through those that did not pass the minimum z score value from the BaseQRankSum test, those having a Fisher Strand bias test score of over 30 and finally those variants with a allele depth value of under 2. Finally, variants with dbsnp entries were removed to leave the final list of RNA editing sites specific to depression. Further functional annotation and pathway analysis of these final RNA editing sites was carried using ReactomPA on R available on Bioconductor and the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang, Sherman and Lempicki, 2009, Yu and He, 2016).

Results

In each of the five samples from suicide victims previously diagnosed with MDD, we initially found approximately 606000, 401000, 539000, 545000, 703000 variants respectively for the runs labelled 1. 3, 5, 12, 13, including variants with dbsnp entries. The controls, labelled Run 4 and 6, contained approximately 270000 and 530000 variants respectively, including dbsnp variants. After removing common variants with dbsnp entries, the numbers decreased significantly for all the runs. Table 2 shows the exact number of variants for both with and without dbsnp entries.

Table 2:Number of Variants Identified in Each Run With or Without Known Variants (dbsnp)

Run	Run1	Run3	Run4	Run5	Run6	Run12	Run13
Total	606938	401223	270214	599935	530667	545738	703541
Total Without known variants (dbsnp)	374468	214035	137447	329086	293650	295502	382241

Out of these we found the common positions where variants were discovered in all the MDD models. The same was done for the two control runs. We then subtracted the variants that were common to the controls from the variants common to the MDD runs as well removing known variants with dbsnp IDs. This left us with a total of 716 variants. At this stage we applied the concept for distinguishing RNA editing sites and genomic SNPs utilized by Zhang and Xiao as explained in methods. A total of 250 variants were filtered out leaving 466 RNA editing sites. These were then filtered for BaseQRankSum score, Fisher strand bias score and allele depth (methods). This narrowed down potential editing sites to 349. Finally we removed any site that occurred in either of the controls individually. The final list of RNA editing sites contained 143 editing sites. Table 3 lists these sites as well as their relevant functional information.

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 3:Final list of Annotated RNA Editing Sites

Chromosome	Position	Ref	Alt	Location	Gene
chr1	43623269	T	C	UTR3	PTPRF
chr1	66037514	T	C	intronic	PDE4B
chr1	1.47E+08	G	A	UTR3	PRKAB2
chr1	1.71E+08	A	G	intronic	PRRC2C
chr1	2.13E+08	T	C	intronic	TATDN3
chr1	2.34E+08	A	G	intronic	SLC35F3
chr1	2.37E+08	A	G	intronic	RYR2
chr1	2.37E+08	T	G	intronic	RYR2
chr1	2.41E+08	A	G	intronic	RGS7
chr10	72520901	A	G	intronic	MICU1
chr10	1.02E+08	T	C	intronic	MGEA5
chr11	47418793	A	T	UTR3	PSMC3
chr11	88961899	A	T	intronic	GRM5
chr12	22464034	A	G	intronic	C2CD5
chr12	40771604	T	C	intronic	CNTN1
chr12	50930684	A	G	UTR3	METTL7A
chr12	98888847	A	T	intronic	ANKS1B
chr12	99028271	A	G	intronic	ANKS1B
chr13	42045203	A	G	intronic	DGKH
chr13	66785667	C	T	intronic	PCDH9
chr13	67090206	A	G	intronic	PCDH9
chr13	67155829	T	A	intronic	PCDH9
chr13	1.1E+08	A	G	UTR3	IRS2
chr14	33000200	G	A	intronic	NPAS3
chr14	50985284	C	T	intronic	TRIM9
chr14	67001143	T	C	intronic	GPHN
chr14	87933820	A	G	UTR3	GALC
chr15	25391077	T	G	ncRNA_intronic	SNHG14
chr15	36931744	A	T	intronic	MEIS2
chr15	44088342	T	C	intronic	FRMD5
chr15	44170171	T	C	intronic	FRMD5
chr15	44717871	T	C	UTR3	B2M
chr15	63324149	A	G	UTR3	CA12
chr15	74837212	A	T	UTR3	ULK3
chr15	80399677	A	T	intergenic	LINC00927;ARNT2
chr16	23067786	A	G	UTR3	USP31
chr16	69085208	A	G	downstream	TANGO6

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

chr16	78121913	A	G	intronic	WWOX
chr16	81389940	C	A	intergenic	MIR4720;CMIP
chr17	33922758	T	C	intronic	ASIC2
chr17_GL000205v2_random	42474	T	C	intergenic	NONE;MGC70870
chr18	31624468	A	G	UTR3	B4GALT6
chr18	39871212	T	C	intergenic	MIR924HG;LINC01477
chr18	46825460	T	A	UTR3	PIAS2
chr18	61416859	A	T	intronic	CDH20
chr2	11273650	A	G	intronic	ROCK2
chr2	29158897	A	G	intronic	CLIP4
chr2	50854340	A	T	intronic	NRXN1
chr2	55016758	A	G	intronic	RTN4
chr2	77516387	A	G	intronic	LRRTM4
chr2	98585704	T	C	intronic	INPP4A
chr2	1.65E+08	T	C	intronic	SCN3A
chr2	1.92E+08	A	G	intronic	TMEFF2
chr2	2.07E+08	T	C	intronic	ADAM23
chr2	2.3E+08	A	G	UTR3	TRIP12
chr20	8870795	T	G	intronic	PLCB1
chr20	16443939	A	G	intronic	KIF16B
chr21	28957258	A	C	intronic	LTN1
chr3	25598304	C	T	UTR3	TOP2B
chr3	64176818	A	G	intronic	PRICKLE2
chr3	76313702	T	C	intergenic	ZNF717;ROBO2
chr3	85267463	T	C	intronic	CADM2
chr3	1.18E+08	A	G	intergenic	LINC00901;LINC02024
chr3	1.18E+08	A	G	intergenic	LINC00901;LINC02024
chr3	1.49E+08	A	G	UTR3	HLTF
chr3	1.68E+08	A	G	UTR3	GOLIM4
chr3	1.74E+08	T	C	intronic	NLGN1
chr4	21333789	A	C	intronic	KCNIP4
chr4	21508213	A	G	intronic	KCNIP4
chr4	1.13E+08	T	C	intronic	ANK2
chr4	1.55E+08	A	T	UTR3	MAP9
chr4	1.56E+08	T	C	intergenic	MAP9;GUCY1A3
chr4	1.57E+08	A	G	intronic	GLRB
chr4	1.59E+08	A	C	intronic	RAPGEF2
chr4	1.67E+08	A	G	intronic	SPOCK3
chr4	1.72E+08	G	A	intronic	GALNTL6
chr4	1.76E+08	T	C	UTR3	WDR17
chr5	95078226	A	G	intronic	MCTP1

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

chr5	1.38E+08	A	G	intronic	KLHL3
chr5	1.47E+08	A	G	intronic	PPP2R2B
chr5	1.65E+08	T	C	intergenic	LOC102546299;LINC01947
chr6	12800740	T	A	intronic	PHACTR1
chr6	12804719	A	G	intronic	PHACTR1
chr6	13023635	G	A	intronic	PHACTR1
chr6	56805187	A	G	intronic	DST
chr6	63707499	T	C	intronic	PHF3
chr6	88140380	A	G	UTR3	CNR1
chr6	1.1E+08	A	G	intronic	WASF1
chr6	1.36E+08	A	C	intronic	BCLAF1
chr6	1.38E+08	T	C	UTR3	ARFGEF3
chr6	1.45E+08	T	A	intronic	UTRN
chr6	1.65E+08	A	G	intronic	PDE10A
chr7	11067825	A	C	intronic	PHF14
chr7	77042032	T	C	ncRNA_exonic	PMS2P9
chr7	87162186	A	G	intronic	DMTF1
chr7	1.12E+08	T	C	intronic	DOCK4
chr7	1.21E+08	T	G	intronic	KCND2
chr7	1.21E+08	T	C	intronic	KCND2
chr7	1.22E+08	T	C	intronic	PTPRZ1
chr7	1.52E+08	A	G	intronic	RHEB
chr8	4962072	C	T	intronic	CSMD1
chr8	11153025	A	G	intronic	XKR6
chr8	28571071	T	C	UTR3	FZD3
chr8	50232742	A	T	intronic	SNTG1
chr8	1.04E+08	T	C	UTR3	RIMS2
chr8	1.15E+08	A	C	UTR3	TRPS1
chr9	3492210	A	C	intronic	RFX3
chr9	14677334	A	T	intronic	ZDHHC21
chr9	35385160	T	A	intronic	UNC13B
chr9	38558640	T	C	intergenic	FAM95C;ANKRD18A
chr9	41058879	A	T	ncRNA_intronic	PGM5P2
chr9	1.05E+08	A	G	intronic	ABCA1
chr9	1.09E+08	A	G	UTR3	TMEM245
chr9	1.33E+08	C	T	UTR3	SURF4
chrX	20238266	G	A	intronic	RPS6KA3
chrX	23985387	A	G	UTR3	KLHL15
chrX	73957424	A	G	ncRNA_intronic	JPX
chrX	74116756	A	G	ncRNA_intronic	FTX
chrX	1.15E+08	T	C	intronic	HTR2C

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

chr10	1.33E+08	T	C	exonic	CALY
chr11	62530656	A	G	exonic	AHNAK
chr13	49551341	A	G	exonic	RCBTB1
chr13	52397324	A	G	exonic	THSD1
chr15	52376380	A	G	exonic	MYO5A
chr18	32268128	C	T	exonic	GAREM1
chr18	70201926	A	G	exonic	RTTN
chr3	58430139	A	T	exonic	PDHB
chr5	1.8E+08	A	G	exonic	TBC1D9B
chr9	1.21E+08	A	C	exonic	FBXW2
chr9	1.32E+08	T	A	exonic	SETX
chrX	53193462	A	T	exonic	KDM5C
chr10	68788322	A	T	exonic	CCAR1
chr1	1.83E+08	T	C	exonic	DHX9
chr13	28712777	A	G	exonic	SLC46A3
chr2	32433746	T	A	exonic	BIRC6
chr3	57575590	T	C	exonic	ARF4
chr3	1.53E+08	T	C	exonic	P2RY1
chr5	37169127	C	T	exonic	C5orf42
chr5	1.38E+08	A	G	exonic	FAM13B
chr6	46881525	A	G	exonic	ADGRF5
chr8	67238774	A	G	exonic	ARFGEF1
chr9	92848175	A	G	exonic	ZNF484

In terms of the type of editing event, as expected A to I (A to G in the data) and C to U (C to T in the data) dominated. A to G edits ranged from 26 to 28% of all edits in all but Run 5. Run 5 being a bit of an anomaly with only about 13% of all edits being A to G. C to U edits represented 9 to 10% of edits in all the runs. Run 5 again had a lowest percentage of C to T edits at about 9%. Figure 1 displays this information.

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

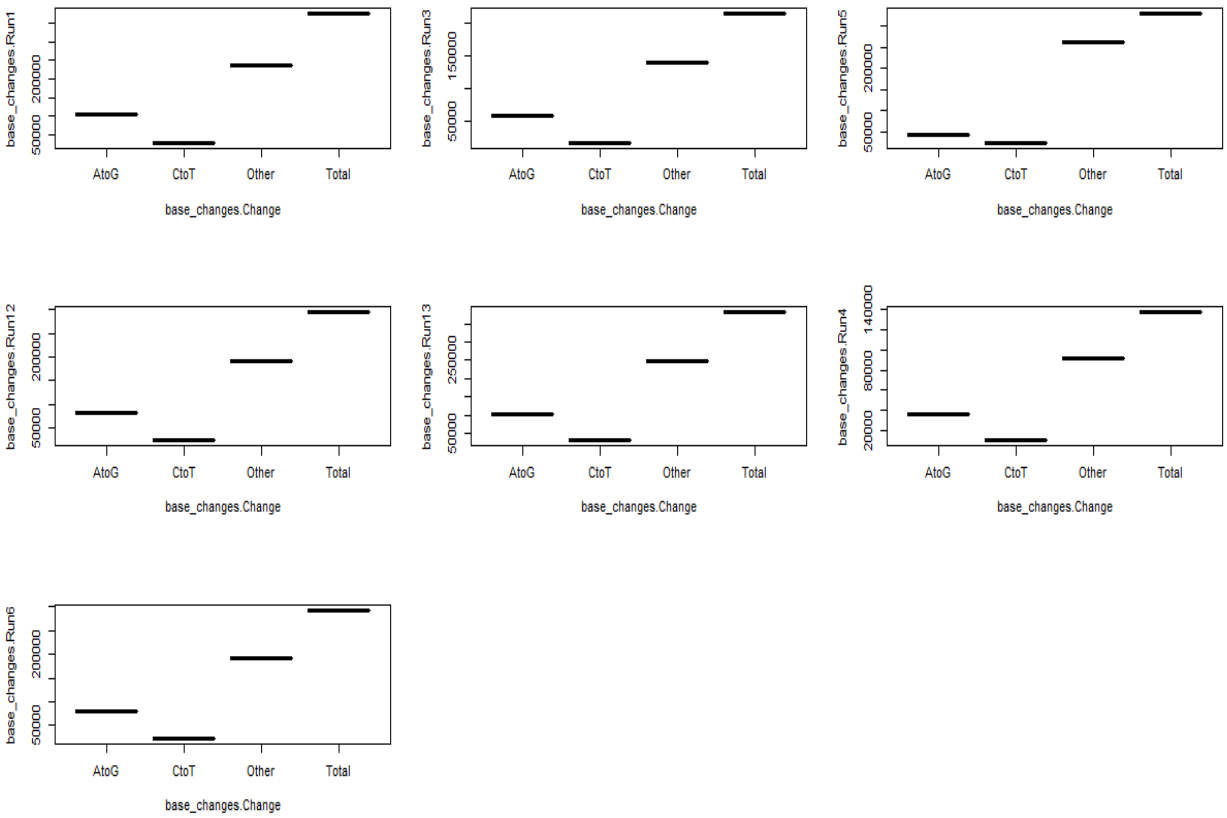

Figure 1:Types of Base Changes for Each Run (After Removing Known Variants)

In terms of exonic changes, which are usually more difficult to identify for RNA editing sites, we found 23 editing sites within exons in our final list of MDD specific editing sites. Among these 12 induced nonsynonymous amino acid changes in the encoded protein. Out of the amino acid changes, four were Methionine to Threonine changes. We performed a pathway enrichment analysis of the list of genes using the ReacomPA package available on Bioconductor on R. This subsequently gave us several relevant functional pathways that maybe dysregulated in MDD as a result of these editing events. Figure 2 shows a dotplot for the pathway enrichment analysis.

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

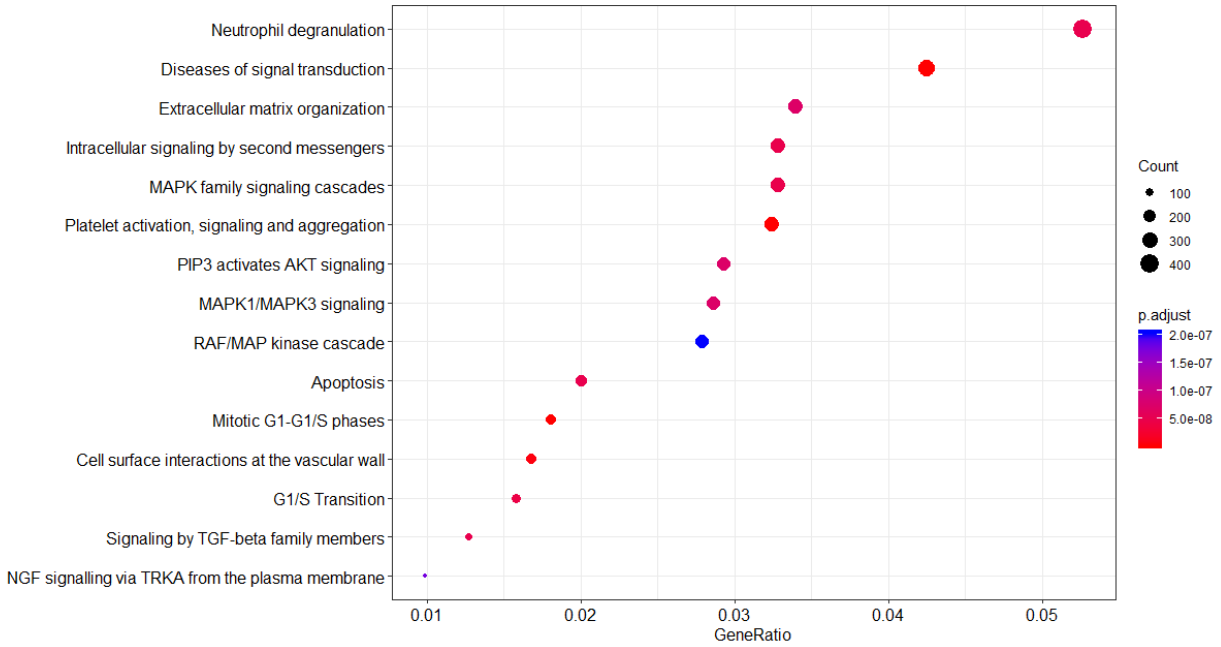

Figure 2:Results of Pathway Enrichment Analysis of Final List of Editing Sites. This shows the 15 most significantly affected functional pathways by the Genes Harbouring Said Editing Sites

Discussion

We attempted to identify new RNA editing sites in individuals diagnosed with Major Depressive Disorder and who subsequently were suicide victims. We also wished to characterize any changes in RNA editome that may be specific to the disease. While we did not find any alterations to overall RNA editing pattern, we identified a number of new editing sites that were unique to the disease models. A significant number of these were exonic and even caused nonsynonymous codon changes leading to amino acid substitutions. Among these, our findings included the Serotonin 2c Receptor. However while our initial results included a large number of editing sites in this gene from the MDD models, after the filtering steps we were only left with one intronic editing site in the gene in our final list. This could possibly imply a lesser role for the edited isoform of the gene's transcript in MDD than previously speculated. Several of our other genes however offer potential new markers for the disease. The Cannabinoid Receptor 1 has functions implicated in depression associated disorders and may offer a potential marker through the editing site we found here (an A to I edit in its UTR3). The frizzled class 3 receptor is another gene previously implicated in depression. This gene harbored a T to C edit in its UTR3. Neuroligin 3 was another gene we found to have MDD specific editing that is involved in depression associated pathways. This contained an intronic T to C editing site. Lastly

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

phosphodiesterase 4B was our fifth gene to have a direction functional relationship with depression. This contained an A to I intronic edit. In addition to these, several of the genes we identified have functions associated with bi polar disorder such as multiple C2 and transmembrane domain containing 1, Phosphodiesterase 10A and aryl hydrocarbon receptor nuclear translocator 2. We also found several genes implicated in schizophrenia that contained unique RNA editing sites. This included protein phosphatase 2 regulatory subunit B'beta, protein tyrosine phosphatase, receptor type Z1, reticulon 4 and others. Other genes we identified were relevant in context of mental retardation, attention deficiency, autism, smoking disorder and alcoholism (UniProt Consortium, 2018). All of which in terms of their negative impact on quality of life can have potential contributions to depression and suicidal behavior. At the time of submission of this paper, we are still furthering this research through running analysis on more RNA sequence samples from MDD diagnosed suicide patients, controls and non MDD suicide patients. Here we present our initial findings that may serve as potential markers for the identification and early diagnosis of MDD and potential suicidal tendencies.

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

DePristo, M., Banks, E., Poplin, R., Garimella, K., Maguire, J., Hartl, C., Philippakis, A., del Angel, G., Rivas, M., Hanna, M., McKenna, A., Fennell, T., Kernytsky, A., Sivachenko, A., Cibulskis, K., Gabriel, S., Altshuler, D. and Daly, M. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5), pp.491-498.

Diagnostic and statistical manual of mental disorders. (2002). Washington, DC: American Psychiatric Association.

Dobin, A., Davis, C., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. (2012). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp.15-21.

Huang, D., Sherman, B. and Lempicki, R. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), pp.44-57.

Kapp, C. (2001). WHO report aims to increase understanding of mental health. *The Lancet*, 358(9289), p.1248.

Labonté, B., Engmann, O., Purushothaman, I., Menard, C., Wang, J., Tan, C., Scarpa, J., Moy, G., Loh, Y., Cahill, M., Lorsch, Z., Hamilton, P., Calipari, E., Hodes, G., Issler, O., Kronman, H., Pfau, M., Obradovic, A., Dong, Y., Neve, R., Russo, S., Kazarskis, A., Tamminga, C., Mechawar, N., Turecki, G., Zhang, B., Shen, L. and Nestler, E. (2017). Sex-specific transcriptional signatures in human depression. *Nature Medicine*, 23(9), pp.1102-1111.

Leinonen, R., Sugawara, H. and Shumway, M. (2010). The Sequence Read Archive. *Nucleic Acids Research*, 39(Database), pp.D19-D21.

Lohoff, F. (2010). Overview of the Genetics of Major Depressive Disorder. *Current Psychiatry Reports*, 12(6), pp.539-546.

Lyddon, R., Dwork, A., Keddache, M., Siever, L. and Dracheva, S. (2012). Serotonin 2c receptor RNA editing in major depression and suicide. *The World Journal of Biological Psychiatry*, 14(8), pp.590-601.

Nishikura, K. (2010). Functions and Regulation of RNA Editing by ADAR Deaminases. *Annual Review of Biochemistry*, 79(1), pp.321-349.

Park, E., Guo, J., Shen, S., Demirdjian, L., Wu, Y., Lin, L. and Xing, Y. (2017). Population and allelic variation of A-to-I RNA editing in human transcriptomes. *Genome Biology*, 18(1).

Patten, S., Kennedy, S., Lam, R., O'Donovan, C., Filteau, M., Parikh, S. and Ravindran, A. (2009). Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical Guidelines

bioRxiv preprint doi: https://doi.org/10.1101/419309; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

for the Management of Major Depressive Disorder in Adults. I. Classification, Burden and Principles of Management. *Journal of Affective Disorders*, 117, pp.S5-S14.

Pavuluri, M., Volpe, K. and Yuen, A. (2017). Nucleus Accumbens and Its Role in Reward and Emotional Circuitry: A Potential Hot Mess in Substance Use and Emotional Disorders. *AIMS Neuroscience*, 4(1), pp.52-70.

Ramaswami, G., Lin, W., Piskol, R., Tan, M., Davis, C. and Li, J. (2012). Accurate identification of human Alu and non-Alu RNA editing sites. *Nature Methods*, 9(6), pp.579-581.

UniProt Consortium, T. (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, 46(5), pp.2699-2699.

Wang, K., Li, M. and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), pp.e164-e164.

Yu, G. and He, Q. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Molecular BioSystems*, 12(2), pp.477-479.

Zhang, Q. and Xiao, X. (2015). Genome sequence–independent identification of RNA editing sites. *Nature Methods*, 12(4), pp.347-350.