Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Genetic Predisposition of Vulnerable Groups to Schizophrenia and Bipolar I via *AKAP11* Variant Analysis

Peyton Tillie
Sohila Sidhu
Vivian Hu
Ali Mohamad
Elle Vollmer

July 25, 2023

Abstract

Recent exome sequencing analyses have shown that schizophrenia and bipolar I disorder, collectively affecting 2% of the population, may extend from a common genetic origin of the *AKAP11* gene. Though several studies have been put forth to examine the relationship between the pathogenesis of these mental illnesses and *AKAP11* variants, a model has yet to be created that tests this hypothesis by accounting for real–world diagnosis frequencies in addition to a genetic framework. Analogously, no genetic study has been put forth to identify specific vulnerable groups in the development of either schizophrenia or bipolar disorder. To provide additional insight into the pathogenesis of these diseases, we perform ordinal logistic regression on every *AKAP11* variant in the SCHEMA and BipEx data sets. Primarily using the CADD genome annotation, we establish the probability that each variant is deleterious. We use a chi–square goodness of fit test to demonstrate that the amino acids coded for by high–risk variants are similar between schizophrenia and bipolar disorder, suggesting a common underlying genetic origin. Additionally, a sequence of one–sided t tests is run to compare the age and sex frequencies of these high–risk variants to the frequencies of schizophrenia and bipolar I diagnoses among the Canadian population. In total, we find that *AKAP11* exhibits a strong correlation with bipolar I disorder and a moderate correlation with schizophrenia. We also find that males under 30 and females between 50 and 55 are most vulnerable to schizophrenia and bipolar I, respectively.

Introduction

Schizophrenia (SCZ) and bipolar disorder (BPD), widely recognized as highly debilitating hereditary mental illnesses, affect an estimated 2% of the Canadian population (Dobson et al., 2020). Formerly known as manic–depressive illness, BPD is a chronic psychiatric disease that encompasses three differentiable forms: bipolar I disorder (characterized by severe manic or depressive episodes), bipolar II disorder (marked distinct by hypomanic episodes with decreased relative severity), and cyclothymic disorder (recurring hypomanic and depressive episodes). Overarching symptoms include volatility in mood, sleep patterns, and concentration (McCormick et al., 2015). On the contrary, SCZ patients suffer an array of symptoms that are classified as positive, negative, and cognitive. According to Frith et al. (2000), positive symptoms describe any change in behaviour or thought, including hallucinations and delusions; negative symptoms cause patients to become apathetic and lethargic; and cognitive symptoms are marked by impairments in attention, working memory, and executive functions. As a result, SCZ patients experience general motor and cognitive impairment which interrupts thought processes, perceptions, and emotional responsiveness (Chatterton et al., 2008).

Although SCZ and BPD vary comparably in symptomatology, they are similar in that both are strongly associated with an increased prevalence of suicide, diagnosis of clinical depression, crime, and premature mortality (Goeree et al. 2005; Stimmel 2004). The lifetime risk of suicide is 170 and 30 times higher than the general population for BPD and SCZ, respectively (Dome et al. 2019; Zaheer et al. 2020). Pri-

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

mary caregivers, especially in a home environment, experience a tremendous burden in assisting those diagnosed with SCZ or BPD and are recognized as a high–risk group for developing mental disorders themselves (Lasebikan and Ayinde, 2013).

Genetic linkage between SCZ and BPD was once disregarded due to the seemingly unrelated nature of these two mental diseases. However, novel molecular genetic studies have suggested a connection between these disorders via an intermediate phenotype by evidence of certain endophenotypes, namely white matter density (McIntosh et al., 2005). A growing body of research around potential genetic underpinnings supports that variations in the same genes influence how susceptible people are to both conditions (Craddock et al., 2005). For example, the *AKAP11* gene, located on the thirteenth chromosome, has been of particular interest. The AKAP–11 protein, coded for by the *AKAP11* gene, is a part of the protein kinase A family of enzymes, which are tetramers of two catalytic and two regulatory subunits docked at precise intracellular sites (Calejo and Taskén, 2015). AKAP–11 in particular binds to type II regulatory subunits of protein kinase A to assist in various cellular functions, such as liquid metabolism and glycogen regulation (Huang et al., 1997).

A recent exome–sequencing study by Palmer et al. (2022) finds that an excess of *AKAP11* protein–truncating variants (PTVs) is present in both disorders. PTVs are genetic variants predicted to shorten the coding sequence of genes through a premature termination codon, subsequently influencing the development of neuropsychiatric illnesses (Sanders et al., 2019). In this study, we apply a methodology in which *AKAP11* variants are further investigated to establish a genetic background for the diagnosis of both SCZ and BPD. We hypothesize that certain *AKAP11* variants have a statistically significant effect on the diagnosis of both SCZ and BPD. Similarly, we hypothesize that high–risk *AKAP11* variants will demonstrate a correlation with the observed Canadian diagnosis frequencies between age and sex, further testing the validity of the suggested correlation to *AKAP11* using real–world data.

While analyses regarding *AKAP11* variants have been conducted before, particularly with respect to genome sequencing, they have yet to be tested in combination with genome annotation data, which places relative weightings on each variant such that the more deleterious ones can be focused on. Additionally, the current body of literature has yet to compare the results against the true frequencies of SCZ and BPD diagnoses.

Methods

Having demonstrated the basis for related genetic predispositions to schizophrenia and bipolar disease, we turn our attention to a novel methodology of determining the statistical significance of *AKAP11* in the pathogenesis of these two diseases. Specifically, we use *AKAP11* variants to investigate its link to BPD and SCZ. Subsequently, frequencies of severe variants are compared against frequencies of BPD and SCZ across Canada.

Ordinal Logistic Regression

To establish the probability that any given *AKAP11* variant has a statistically significant role in the diagnosis of BPD or SCZ, we run a logistic regression on every variant entered in the SCHEMA (Singh, 2020) and BipEx (Adolfsson, 2021) browsers for SCZ and bipolar I, respectively. Based on the consequence terms of each variant defined by Sequence Ontology (Cunningham et al., 2015), every entry is classified into 4 classifications, j, ranging from statistically unlikely to cause diagnosis to statistically likely to be pathogenic. Any variant without sufficient genome annotation was removed from the analysis (i.e., lack of CADD scores or protein sequence identification), resulting in a variant population of n = 745 and n = 380 for SCZ and BPD, respectively.

With categorized variants, we perform ordinal logistic regression (MacKenzie, 2018) where the probability of a variant being placed into a certain class is assessed via a given set of independent variables. This was performed with the XLSTAT Version 2023.1.2 software package (Lumivero, 2023). Five pertinent values were chosen as independent variables to collectively conclude how likely each variant is to be placed in each class: (i) the CADD functional annotation (Rentzsch et al., 2019), allele number for the (ii) control and (iii) case group, and allele frequency for the (iv) control and (v) case group. 50 iterations were performed to increase the precision of the regression coefficients. Using XLSTAT automatic weightings for each independent variable, we find that there is little significant influence that (ii) through (v) has on the classification of each variant, and thus rely

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

primarily on the CADD functional annotation as a means of determining the probability that any given variant is placed in some class.

The ordinal logistic regression model is a form of logistic regression wherein the probability of a categorical event occuring is assessed via a given set of independent variables. Within the context of this study, we take the above listed independent variables (with the primary weighting given to the CADD annotation) to find the probability that a given variant is classified under some j.

$$\text{logit}(\theta_i) = \sum_{n=0}^{u} \beta_n x_{\text{in}}, \tag{1}$$

wherein θ_i is the cumulative probability that the variant is placed in the class $1 \leq j \leq i$; x_{i1}, x_{i2},..., x_{iu} are the independent variables that have some effect on the classification probability, beginning at x_{i1}; and β_1, β_2, ..., β_u are regression coefficients, which define the relative weightings placed on each x value up to u independent variables. Thus, eq. 1 may be expanded to the more pragmatic model:

$$\theta_i = \frac{\exp\left(\beta_{\text{o}} + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_u x_{iu}\right)}{1 + \exp\left(\beta_{\text{o}} + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_u x_{iu}\right)}. \tag{2}$$

Determining the probability α_j that any variant is placed under one non-cumulative class j is modelled by subtracting the previous class probabilities from the specific class. This is shown by the piecewise function:

$$\alpha_j = \begin{cases} \theta_i, & \text{if } j = 1 \\ \theta_{i-1}, & \text{if } 1 \leq j \leq i-1 \\ 1 - \theta_{i-1}, & \text{if } j = 1 \end{cases}. \tag{3}$$

Using these independent variables for each variant of *AKAP11* in SCHEMA and BipEx, we take the frequencies of the amino acids coded for by each of the variants with the highest 10% probabilities of being placed in $j = 4$ (i.e., most likely to cause diagnosis of SCZ and BPD). If there were no correlation with *AKAP11* , we would assume a null hypothesis wherein there is no difference between the amino acid distribution of the highest risk *AKAP11* variants and the natural distribution of amino acids in humans, demonstrating the inability of the gene to be damaging to non–random amino acids. A chi–square goodness of fit test (χ^2) is used to determine if there is any statistical deviation from the amino acid frequencies observed against natural amino acid frequencies, given that the data adhere to the standard assumptions of χ^2. Particularly, variables are ordinal, mutually exclusive, and absolute (McHugh, 2013).

Equivalence Testing

Subsequently, we further test our hypothesis by comparing the frequencies of the variants with the highest $\mathbb{P}(j = 4)$ against the frequencies of schizophrenia and bipolar disorder across Canada. We manually extract demographic distributions for each of the top 10% of variants for SCZ and BPD with the highest probability of being placed in $j = 4$ from the Genome Aggregation Database (gnomAD) browser v2.1.1 and v3.1.2 (Abreu, 2014). Any variant that was not recorded in either version of gnomAD was excluded from the study, resulting in a variant population size of n = 66 and n = 37 for SCZ and BPD, respectively. Additionally, the age and sex frequency distributions of SCZ and BPD crude prevalence in Canada from the 2019–2020 fiscal year were recorded from the Canadian Chronic Disease Surveillance System (CCDSS; of Canada, 2019). Any individual outside of >30, 50–55, or <80 age brackets in addition to any individual who could not be classified as male or female at birth were excluded from analysis.

A two one–sided test (TOST) was used to find a possible correlation between the predicted frequency sets (based on *AKAP11* variants) and observed frequency sets (based on CCDSS). Lower and upper equivalence bounds are ±1.1 for SCZ and±2.3 for BPD, and a failure to show that each data set violates these bounds using t tests is sufficient evidence for equivalence (Lakens, 2017). Similarly, we perform the Shapiro–Wilk and F–Test to test for equal variance and normal distribution, ensuring appropriate use of the t test.

Results

The goal of the first suite of tests was to determine if *AKAP11* has some effect on the diagnosis of SCZ or BPD using the amino acid distributions of its variants. We use the results from the logistic ordinal regression to quantify the probability that each listed

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

variant is statistically likely to be pathogenic. The relative probabilities for these *AKAP11* variants for both SCZ and BPD are shown in Fig. 1, where a larger line indicates a larger probability. In Fig. 1a, the relative probabilities of *AKAP11* variants on schizophrenia are shown. Here, the probabilities of each variant are relatively concentrated with a few minor extremes; however, in Fig. 1b which demonstrates the effect on bipolar I diagnoses, the probability variance is much more dramatic and spread out along the chromosome.

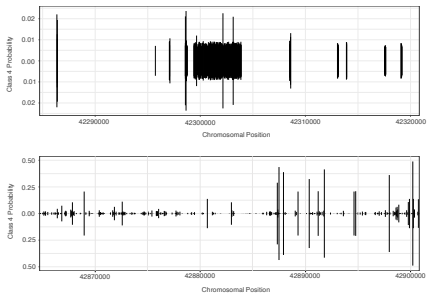

Figure 1 Variant Probability of Causing Diagnosis of Schizophrenia and Bipolar I Along Chromosome 13

These amino acid frequencies of the top 10% of variants with the largest $\mathbb{P}(j = 4)$ are compared against natural amino acid frequencies to test for discrepancies using χ^2. Table 1 shows the recorded amino acid frequencies with their expected frequencies taken from Moura et al. (2013). Interestingly, when taking the top 3 amino acids with the highest adjusted differences between observed and expected frequencies for BPD and SCZ, we find that both disorders are most affected by identical amino acids (Leu, Ser, and Val), indicating a strong genetic linkage.

Running a chi–square goodness of fit test to compare observed frequencies with expected frequencies, we find statistically significant divergence. A χ^2 value of ˜80.82 and ˜109.21 is calculated for BPD and SCZ, respectively, both of which indicate a p value of <0.01. Therefore, we reject the null hypothesis in favour of the alternative hypothesis: variants of *AKAP11* show statistically significant influence over the diagnosis of SCZ and BPD.

Furthermore, we run a second suite of tests to compare the highest risk variants ($j = 4$) used in the first section against the true frequencies of SCZ and BPD in Canada. We assume that these variants have a 100% certainty of causing a SCZ or BPD diagnosis

Bipolar I			Schizop	
Amino Acid	Observed Occurrences	Expected Occurrences	Amino Acid	Observ Occurre
Alanine	3	4.35	Arganine	10
Glutamine + Glutamic Acid	3	6.44	Leucine	9
Isoleucine	2	1.01	Lysine	13
Leucine	5	1.67	Serine	8
Lysine	2	1.35	Threonine	10
Methionine	1	0.21	Tyrosine	11
Proline	3	0.15	Valine	13
Serine	7	2.51		
Threonine	4	2.04		
Tyrosine	3	2.14		
Valine	4	1.85		

Table 1
Amino Acid Occurrences of the Most Statistically Pathogenic Variants.

if present in an individual's genome. The age and sex distributions of SCZ and BPD classified by gnomAD are noted, constrained by ages <30, 50–55, and >80. Fig. 2 demonstrates the age and sex brackets in Canada and those of gnomAD for BPD.

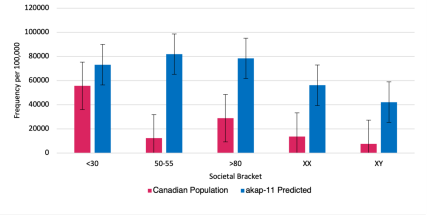

Figure 2 Bipolar Frequencies from *AKAP11* Variants and the Canadian Population

Additionally, Fig. 3 demonstrates the age and sex brackets in Canada and those of gnomAD for SCZ.

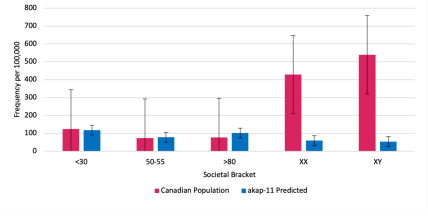

Figure 3 Schizophrenia Frequencies from *AKAP11* Variants and the Canadian Population

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Using the *AKAP11* predicted diagnoses, we find that those most vulnerable to SCZ are males under 30 and those most vulnerable to BPD are females between 50 and 55, as can be seen from the largest frequencies in Fig. 2 and 3.

We perform an F-test and Shapiro–Wilk Test for both diseases to ensure appropriate use of TOST. Table 2 summarizes the p values obtained from these tests. Since both data sets pass the Shapiro–Wilk Test, we reject the null hypothesis that the Canadian and *AKAP11* data sets for SCZ and BPD are not normally distributed, which makes them both a suitable subject for t testing. The bipolar data set does not pass the F-test, so we accept the null hypothesis that it has an unequal variance between the Canadian population and *AKAP11* predicted results. By extension, the bipolar data set requires an unequal variance t test and the schizophrenia data set requires a standard student's t test.

	Bipolar I	Schizophrenia
F-Test	0.38	0.0007
Shapiro-Wilk Test	<0.01	<0.01

Table 2
Results from Variance and Distrbution Testing.

Consequently, we run the appropriate TOST test for each data set based on the rejection or acceptance of variance and distribution from the p values in Table 2. We derive several conclusions from each respective TOST test on each data set.

- We find that BPD falls within the set bounds with a p value of 0.0076. Assuming a $\propto = 0.05$ cutoff, this is well beneath the required value, suggesting that this correlation is not due to random chance.
- However, the SCZ data set, tested with equal variance, obtains a p value of 0.13, which is greater than the predefined $\propto$. This suggests that such a result may be due to random chance.

Therefore, we find that *AKAP11* has a strong correlation with the diagnosis of BPD and a slight correlation to SCZ. This is because we find a correlation for both SCZ and BPD when the variants are analyzed genetically using amino acids, but a statistically insignificant effect on SCZ when tested against the Canadian population. In both cases, we reject the null hypothesis that *AKAP11* has no effect on the diagnosis of SCZ and BPD.

Discussion

We demonstrate that both SCZ and BPD have a theoretical basis for a diagnosis in response to the presence of certain *AKAP11* variants; however, when applied practically to observed diagnosis frequencies, the possibility that SCZ has a correlation with these variants is weak. One potential reason for this discrepancy is disagreements in the eukaryogenesis hypothesis (Koonin, 2006). Current literature presents contradictions in the evolution of modern eukaryotic cells: some suggest that the ability for eukaryotic mRNA splicing is a consequence of intron interruption within genes, while others suggest that intron lineage had no effect on eukaryotic evolution (Belshaw and Bensasson, 2006). Previous studies on the *AKAP11* correlation with mental disease focused solely on protein–coding exons and their role in disease pathogenesis. On the contrary, our study chose to adopt a more outward approach which focused on the gene variants as a whole rather than solely exon interaction. These inherent differences in approach represent differing results from eukaryogenesis hypotheses in that we considered introns and exons equally, though others have excluded introns from analysis. Subsequent research must therefore be conducted to further investigate genetic modelling to provide a more comprehensive understanding of the underlying assumptions made in both models.

However, we obtained perfect overlap between SCZ and BPD when we apply high–risk variants to their expected amino acid frequencies, indicating of a strong correlation between *AKAP11* and both of these diseases. This suggests that the altered bonding of *AKAP11* variants to group A–kinase anchor proteins results in a higher probability of SCZ and BPD diagnosis. Overall, we find that the correlation to *AKAP11* is robust with BPD and moderate with SCZ.

The reason for a less significant effect when SCZ is applied to the Canadian population may be due to limitations in our model. For example, variables in the genome annotations of each variant likely influenced our findings. Namely, the results of the regression are based primarily on the consequence genome annotation as defined by Sequence Ontology to classify each variant. While it is an accepted annotation tool, improvements are still being suggested to this annotation at the current date (Eilbeck et al., 2005). As seen in Fig. 1, there is a vast discrepancy between

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

the probabilities that each variant will be placed in Class 4 between SCZ and BPD (as evidenced by the extreme scaling differences of the relative probabilities of both diseases). Such a large variance in the class probabilities between SCZ and BPD is likely indicative that there may be differences in genome annotation, which is required for both theoretical and practical statistical testing.

Similarly, the study is limited by the fact that only the *AKAP11* variants found in SCHEMA and BipEx were included in the analysis. There were many differences between which variants were described in each data set. Without identical data sets, it is impossible to get a complete analysis of the entire *AKAP11* gene; as a result, both diseases have different data inputs for the logistic regression. This therefore reduces the accuracy of comparing the results from two different regression tests to a single data set of the Canadian population.

The results of this statistical model have far–reaching implications. Individuals with a diagnosis of either disorder would require highly individualized treatment (Goldberg, 2019; Vieta, 2010) rather than a universalized system which assumes complete similarity between SCZ and BPD. With respect to legislative consequences, it would be beneficial to invest in a more genetic approach to combating mental diseases. For example, the genome editing CRISPR–Cas–9 system has several applications, including mental illnesses such as schizophrenia and bipolar I disorder. Using such a method, precise and highly targeted edits of potentially threatening genetic alleles (Michael and Brennand, 2021), specifically *AKAP11*. It is also necessary to better subsidize treatment facilities which specifically target at–risk demographics for these diseases. We find that males under 30 and females between 50 and 55 may be particularly vulnerable, and this could aid in creating an improved system of finding and treating individuals affected with SCZ and BPD, respectively. Lastly, the efficiency of our methods could be improved for future research through a computer code (see Appendix) that compares data sets for high–risk AKAP11 variants and demographic diagnoses. This would allow for a more practical application of data for the global population rather than just the Canadian population.

Since our results suggest a strong correlation between *AKAP11* and BPD and a slight correlation with SCZ, further research aimed at the genetic predispositions of mental disease will allow for more confident applications of genetic patterns as it relates to SCZ and BPD.

Conclusion

Our study concludes that variants of the *AKAP11* gene have a strong effect on the diagnosis of bipolar disorder and a moderate effect on the diagnosis of schizophrenia. Furthermore, our novel processes for variant classification based on the likelihood of diagnosis and comparison against nationwide data can offer future direction in identifying the genetic predisposition of certain vulnerable demographics to SCZ and BPD. Further research should be conducted to validate these observed higher probabilities of *AKAP11* variants in the diagnosis of both schizophrenia and bipolar I.

Appendix: A Guide to Future Studies

To define a preliminary script to aid in future studies wherein larger data sets can be made of use and applied to the necessary context, a listing is provided below. This shows a preliminary Java code for automatic comparison of large demographic data sets, which has the versatility to be applied for any size of data analysis. It should be noted, however, that certain attributes of this code may need to be adjusted for the specifics of the study to improve efficiency.

```
//import all libraries, including file readers to deal wit
import java.io.BufferedReader;
import java.io.FileReader;
import java.io.IOException;
import java.util.HashSet;
import java.util.Set;

public class Main {
    public static void main(String[] args) {

    //declare variables
    String file1Path = "pathname2file1.txt";
    String file2Path = "pathname2file2.txt";

    try {
        Set<String> file1Lines = new HashSet<>();
```

```
        //Read lines from file 1 and store them in a set (more efficient for large text file)
        Set<String> file1Lines = new HashSet<>();
        BufferedReader reader1 = new BufferedReader(new FileReader(file1Path));
        String line1;
        while ((line1 = reader1.readLine()) != null) {
            file1Lines.add(line1);
        }
        reader1.close();

        //Compare lines from file 2 with the lines stored in the set
        BufferedReader reader2 = new BufferedReader(new FileReader(file2Path));
        String line2;
        while ((line2 = reader2.readLine()) != null) {
            if (file1Lines.contains(line2)) {
                //Found a matching line
                System.out.println("Matching line found: " + line2);
            }
        }
        reader2.close();
    } catch (IOException e) {
        e.printStackTrace();
    }
    }
}
```

References

Aguilar Salinas C. A. Ahmad T. Albert-C. M. Allred N. Altshuler D. Ardissino D. Atzmon G. Barnard J. Beaugerie L. Beecham G. Benjamin E. J. Boehnke M. Bonnycastle-L. Bottinger E. Bowden D. Bown M. Campos H. Chambers J. ... Xavier R. Abreu, M. Genome Aggregation Database (gnomAD). 2014. URL https://gnomad.broadinstitute.org/.

Blackwood D. Chapman S. B. Chen C.-Y. Churchhouse C. Corvin A. P. Craddock N. Di Florio A. Goes F. S. Howrigan D. P. Landén M. McIntosh A. McQuillin-A. Morris D. W. Neale B. M. O'Donovan M. C. Ophoff R. A. Owen M. J. Palmer D. S. ... Curtis D. Adolfsson, R. BipEx: Bipolar Exomes. 2021. URL https://bipex.broadinstitute.org/.

R Belshaw and D Bensasson. The rise and falls of introns. *Heredity (Edinb)*, 96:208–13, Mar 2006.

AI Calejo and K Taskén. Targeting protein-protein interactions in complexes organized by A kinase anchoring proteins. *Front Pharmacol*, 6:192, 2015.

MJ Chaston, X Kong, DE Lewis, K Rajagopalan, and A Lazarus. Impact of bipolar disorder on temporal utilization of health care resources. *P T*, 33:15–34, Jan 2008.

N Craddock, MC O'Donovan, and MJ Owen. The genetics of schizophrenia and bipolar disorder: dissecting psychosis. *J Med Genet*, 42:193–204, Mar 2005.

F Cunningham, B Moore, N Ruiz-Schultz, GR Ritchie, and K Eilbeck. Improving the Sequence Ontology terminology for genomic variant annotation. *J Biomed Semantics*, 6:32, 2015.

KC Dobson, SN Vigod, C Mustard, and PM Smith. Trends in the prevalence of depression and anxiety disorders among Canadian working-age adults between 2000 and 2016. *Health Rep*, 31:12–23, Dec 2020.

P Dome, Z Rihmer, and X Gonda. Suicide Risk in Bipolar Disorder: A Brief Review. *Medicina (Kaunas)*, 55, Jul 2019.

K Eilbeck, SE Lewis, CJ Mungall, M Yandell, L Stein, R Durbin, and M Ashburner. The Sequence Ontology: a tool for the unification of genome annotations. *Genome Biol*, 6:R44, 2005.

CD Frith, S Blakemore, and DM Wolpert. Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res Brain Res Rev*, 31:357–63, Mar 2000.

R Goeree, F Farahati, N Burke, G Blackhouse, D O'Reilly, J Pyne, and JE Tarride. The economic burden of schizophrenia in Canada in 2004. *Curr Med Res Opin*, 21:2017–28, Dec 2005.

JF Goldberg. Personalized Pharmacotherapy for Bipolar Disorder: How to Tailor Findings From Randomized Trials to Individual Patient-Level Outcomes. *Focus (Am Psychiatr Publ)*, 17:206–217, Jul 2019.

LJ Huang, K Durick, JA Weiner, J Chun, and SS Taylor. Identification of a novel protein kinase A anchoring protein that binds both type I and type II regulatory subunits. *J Biol Chem*, 272:8057–64, Mar 1997.

EV Koonin. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on 25 Jul 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.169029169.95176267/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

early versus introns-late debate? *Biol Direct*, 1:22, Aug 2006.

D Lakens. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Soc Psychol Personal Sci*, 8:355–362, May 2017.

VO Lasebikan and OO Ayinde. Family Burden in Caregivers of Schizophrenia Patients: Prevalence and Socio-demographic Correlates. *Indian J Psychol Med*, 35:60–6, Jan 2013.

Lumivero. XLSTAT (Version 2023.1.2) [Computer software]. 2023. URL https://www.xlstat.com/.

Nichols J. D. Royle J. A.-Pollock K. H. Bailey L. L. & Hines J. E. MacKenzie, D. I. *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence*. Academic Press, Elsevier., 2 edition, 2018.

U McCormick, B Murray, and B McNew. Diagnosis and treatment of patients with bipolar disorder: A review for advanced practice nurses. *J Am Assoc Nurse Pract*, 27:530–42, Sep 2015.

ML McHugh. The chi-square test of independence. *Biochem Med (Zagreb)*, 23:143–9, 2013.

AM McIntosh, DE Job, TW Moorhead, LK Harrison, SM Lawrie, and EC Johnstone. White matter density in patients with schizophrenia, bipolar disorder and their unaffected relatives. *Biol Psychiatry*, 58: 254–7, Aug 2005.

Deans PJ Michael and KJ Brennand. Applying stem cells and CRISPR engineering to uncover the etiology of schizophrenia. *Curr Opin Neurobiol*, 69: 193–201, Aug 2021.

A Moura, MA Savageau, and R Alves. Relative amino acid composition signatures of organisms and environments. *PLoS One*, 8:e77319, 2013.

Public Health Agency of Canada. Canadian Chronic Disease Surveillance System (CCDSS) [Data set]. 2019. URL https://health-infobase.canada.ca/ccdss/.

DS Palmer, DP Howrigan, SB Chapman, R Adolfsson, N Bass, D Blackwood, MPM Boks, CY Chen, C Churchhouse, AP Corvin, N Craddock, D Curtis, Florio A Di, F Dickerson, NB Freimer, FS Goes, X Jia, I Jones, L Jones, L Jonsson, RS Kahn, M Landén, AE Locke, AM McIntosh, A McQuillin, DW Morris, MC O'Donovan, RA Ophoff, MJ Owen, NL Pedersen, D Posthuma, A Reif, N Risch, C Schaefer, L Scott, T Singh, JW Smoller, M Solomonson, DS Clair, EA Stahl, A Vreeker, JTR Walters, W Wang, NA Watts, R Yolken, PP Zandi, and BM Neale. Exome sequencing in bipolar disorder identifies AKAP11 as a risk gene shared with schizophrenia. *Nat Genet*, 54:541–547, May 2022.

P Rentzsch, D Witten, GM Cooper, J Shendure, and M Kircher. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*, 47:D886–D894, Jan 2019.

SJ Sanders, M Sahin, J Hostyk, A Thurm, S Jacquemont, P Avillach, E Douard, CL Martin, ME Modi, A Moreno-De-Luca, A Raznahan, A Anticevic, R Dolmetsch, G Feng, DH Geschwind, DC Glahn, DB Goldstein, DH Ledbetter, JG Mulle, SP Pasca, R Samaco, J Sebat, A Pariser, T Lehner, RE Gur, and CE Bearden. A framework for the investigation of rare genetic disorders in neuropsychiatry. *Nat Med*, 25:1477–1487, Oct 2019.

Poterba T. Curtis D. Akil H. Al Eissa M. Barchas J. D. Bass N. Bigdeli T. B. Breen G. Bromet E. J. Buckley P. F. Bunney W. E. Bybjerg-Grauholm J. Byerley W. F. Chapman S. B. Chen W. J. Churchhouse C. Craddock N. Cusick C. M. ... Daly M. J. Singh, T. SCHEMA: Schizophrenia exome meta-analysis consortium. 2020. URL https://schema.broadinstitute.org/.

GL Stimmel. Economic grand rounds: the economic burden of bipolar disorder. *Psychiatr Serv*, 55:117–8, Feb 2004.

E Vieta. Developing an individualized treatment plan for patients with schizoaffective disorder: from pharmacotherapy to psychoeducation. *J Clin Psychiatry*, 71 Suppl 2:14–9, 2010.

J Zaheer, M Olfson, E Mallia, JSH Lam, Oliveira C de, D Rudoler, AF Carvalho, BJ Jacob, A Juda, and P Kurdyak. Predictors of suicide at time of diagnosis in schizophrenia spectrum disorder: A 20-year total population study in Ontario, Canada. *Schizophr Res*, 222:382–388, Aug 2020.