



RESEARCH ARTICLE

Identification of genomic variants in regulatory elements associated with risk of suicide

Nathalia Cardenas*, Lorena Salazar*

*Department of Biological Sciences, Universidad de los Andes, Bogota, Colombia.

Received: March 23th 2023 / Accepted: April 1st 2023 / Published: June 15th 2023

© The Author(s) 2023. Article published with Open Access.



Abstract

Suicide attempt is the consequence of a complex interaction between social traumatizing relations, environmental and genetics factors that might induce mood disorders. The genetic framework of schizophrenia, depression, bipolarity and anxiety disorders are very similar, with important genes involved in neurotransmission, circadian rhythm, serotonin pathways and more. We perform a non-coding scanning analysis using GWAS datasets of four previous suicidal attempts articles.

P-value significant genetic variants ($p \leq 0.05$) were processed in a novel inferring software (<http://inferno.lisanwanglab.org/index.php>). We map up to 40 regulatory elements near to those genetic variants, that might be interfering in the expression of genes related to excitability of neurons, period and neuronal development, essential for a good mental health. This investigation allows the identification of new loci of interest related with suicidality and emphasizes the importance of including biological context enrichment to post GWAS analysis.

INTRODUCTION

Suicide is the intention of taking one's life. Every year, this public health issue account for approximately 703,000 deaths (WHO, 2021) mostly attributed to men who have a higher reported rate of completed suicide, whereas women have a higher rate of attempted suicide [1]. Unfortunately, suicidality has risen in young people, and it was the third leading cause of death worldwide in 2014 [2]. Temperamental traits of impulsive aggression, neuroticism, childhood abuse and introversion have also been implicated in suicidal risk. However, it is unlikely that these factors, themselves, can explain suicide [1].

Components of suicidal thoughts and behaviors, such as feeling that life is not worth living or contemplating self-harm, are relatively common in the general population, as well as in patients affected by a variety of separate clinical diagnoses [3].

Therefore, diagnosis of suicide tendencies still biases and encompasses a broad spectrum of phenotype and even, genotype traits. In fact, those individuals who do not progress from the preceding steps towards death may be distinct biologically from those that do [4].

Although suicide results from a complex interaction of multiple social and psychological factors, predisposition to suicidality is at least partly genetic [3]. Over 90% of the suicidals had recognizable psychiatric illness at the time of their death [5] and major depressive disorder (MDD), neuroticism, mood instability and schizophrenia have the strongest correlations with suicidality [3]. Genetic framework of mental illness and suicide attempt might be similar; however, the genetic architecture of suicidal behaviors remains poorly understood [6].

First single-nucleotide polymorphism (SNP) studies found statistics supporting relations between some variants and a specific phenotype, but the highly polygenic nature of complex

Corresponding author

Lorena Salazar

Email

il.salazar@uniandes.edu.co

Keywords: suicide, depression, schizophrenia, mood disorders, GWAS, regulatory elements, sleep disorders.

Bioethical aspects: The authors declare that they have no conflicts of interest associated with this manuscript and have obtained informed consent from the patients. This work was approved by the institutional ethics committee.

Funding: The authors declare that they have not received external funding associated with this work.

Data: The anonymized data will be provided upon request to the corresponding author.

DOI: 10.37980/im.journal.ggcl.20232167

traits and common disorders, constitutes an immense challenge for understanding the biological mechanisms linking single variants with phenotypes [7].

Previous genetic studies indicated that suicide attempt has a polygenic architecture [8] and new sequence technology is required for its study. The advent of GWAS contributed to expose unexpected findings that urged to redefine the framework of population genetics [9].

Selected genetic variants in serotonergic system genes have been the major historic focus of suicide related investigations during the past decades, but the polygenetic perspective is nowadays also being addressed by the study of candidate genes in other neurosystems [10].

Recent advances in this area include a GWAS of suicide attempts in approximate 50,000 individuals with and without psychiatric disorders which identified some suggestive loci [3]. These studies have examined individuals with depression or bipolar disorder, comparing attempters with no attempters and testing for genetic variants that might contribute independently to suicide attempt that were underpowered to detect with typical SNP studies [8].

Although, large-scale genetic studies remains difficult due to the necessity of achieve a collection of enough samples to have the power to identify replicable genetic associations, or to directly estimate the proportion of heritability contributed from common genetic variation [6].

Due to the majority of GWAS hits mapped within non-coding regions (such as intergenic or intronic regions), new functional RNA species (such as lncRNAs or circRNAs) have started to emerge. In recent years, in parallel to an increased knowledge of the non-coding genome, new studies started to characterize disease-associated variants located within non-coding RNA regions [9].

Considering this, we reanalyze four previous genome-wide association studies (GWAS) of suicide attempts, using a novel software to infer the molecular mechanisms of noncoding genetic variants. To our knowledge, this investigation is one of the few that uncover new insights about downstream biological processes related to suicidality.

MATERIALS AND METHODS

Genome-wide association studies (GWAS) data collection

Study subjects were drawn from previous GWAS investigations. We made contact with the correspondence author of each study and they enabled us to use the datasets. First article “Genetic associations with suicide attempt severity and genetic overlap with major depression [4]”, was performed with Yale-Penn sample cohorts of European Americans (EAs, n = 488 affected and 1,951 controls) and African Americans (AAs, n = 643 affected and 3,238 controls), based on severity of suicide attempts.

Second article “GWAS of Suicide Attempt in Psychiatric Disorders and Association With Major Depression Polygenic Risk Scores [8]”, the samples comprised 1,622 suicide attempters and 8,786 no attempters with major depressive disorder; 3,264 attempters and 5,500 no attempters with bipolar disorder; and 1,683 attempters and 2,946 no attempters with schizophrenia. This information resides in the Psychiatric Genomics Consortium and was performed with European individuals.

Third article “Significant shared heritability underlies suicide attempt and clinically predicted probability of attempting suicide [6]” used an online mental health survey taken by a subset of participants (n = 157,366) in the UK Biobank. After quality control, they leveraged a genotyped set of unrelated, white British ancestry participants including 2,433 cases and 334,766 controls. At the same time, they use the leveraged electronic health record (EHR) data from the Vanderbilt University Medical Center (VUMC, 2.8 million patients, 3,250 cases) and machine learning to derive probabilities of attempting suicide in 24,546 genotyped patients.

Lastly, in the investigation “Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation with psychiatric disorders and polygenic association with completed suicide [3]”, they classify samples based on the following criteria: ‘no suicidality’ controls (N = 83,557); ‘thoughts that life was not worth living’ (N = 21,063); ‘ever contemplated self-harm’ (N = 13,038); ‘act of deliberate self-harm in the past’ (N = 2,498); and ‘previous suicide attempt’ (N = 2,666). This information is summarized in Table 1.

Table 1. GWAS dataset summary

Article	Case	Control	Population
Levey D. F. et. al.	1,1131	5,189	European Americans and African-Americans
Mullins N et. al.	6,5569	17,232	European Americans
Ruderfer DM et. al	5,683	334,776	UK Biobank/Europeans
Strawbridge RJ et. al	39,265	83,557	UK Biobank
Total	52,648	440,754	

P- value quality filter

We obtain 31,233,308 variants from the GWAS summary statistics. Based on p-value criteria, we selected 6,492 variants with a significant value between 1×10^{-5} and 1×10^{-8} .

INFERNO: inferring the molecular mechanisms of noncoding genetic variants

Selected variants were entered on INFERNO (<http://inferno.lisanwanglab.org>). First, each top variant is expanded into a set of putative causal variants for further examination to determine LD structure, using a correlation threshold of $r^2 \geq 0.7$ and data collected from 1000 Genomes Project. Users can also select LD structure data for African (AFR), Asian (ASN), and South American (AMR) populations. This LD input block matches with 10,000 random sets of background variants based on the distance to the nearest gene, minor allele frequency, and the number of variants in each tagged LD block [11].

Then, enhancers were annotated with thirty-two tissue and cell type context using FANTOM5, as well as transcription factor binding sites (TFBS) were enriched with the HOMER database. Using the COLOC R package, this software performs the co-localization of the eQTL signals tested in the GTEx database. For each tag region and GTEx tissue, the script identifies all the genes tested for eQTL with the tagging variant. With the same database, lncRNA targets were identified and two correlation measures were performed: Pearson and Spearman [11].

In the final step, INFERNO provides a full summary statistics file including minor frequency allele (MAF) both of input and background variants, r^2 value, the distances to the nearest genes and more useful information. The program prioritizes functional variants within tissue and cell enrichment context and organizes

them depending on its proximity to four types of non-coding regulatory elements: eQTL, miRNA, TFBS and Enhancers.

Statistical based-on filter

After LD analysis was performed, the blocks of variants with $\text{TAG-MAF} \leq 0.1$ and $r^2 \geq 0.9$ were selected. We searched the regulatory elements and the genes linked to those blocks, using two public databases (UniProt <https://www.uniprot.org>) and NCBI (<https://www.ncbi.nlm.nih.gov>)

RESULTS

It was found a major proportion of TFBS (52%) above Enhancers (36%) and miRNA regulatory elements (12%) near to p-value significant variants, eQTL could not be found (Table 2). We set up 23 TFBS, 18 Enhancer and 6 mi-RNA regulatory elements. After this, we reorganized these regulatory elements in nine groups, based on the similarity of the biological functions of the target gene (Figure 1).

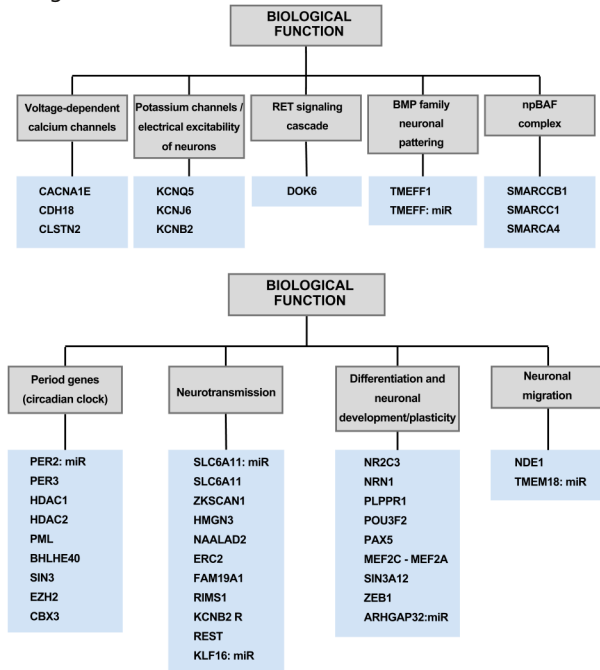
Voltage-dependent calcium channels

Calcium ions are involved in several human cellular processes including transcription, corticogenesis and synaptogenesis

Table 2. List of enhancers, TFBS and mi-RNA.

Article	Enhancer	TFBS	mi-RNA
<i>Levey DF et. al</i>	CACNA1E CDH18	ZKSCAN1	TMEM18: miR PER2: miR
<i>Mullins N et. al.</i>	PER3 POU3F2 PLPR1 NRN1 ERC2 CLSTN2 NAALAD2 FAM191A1	DOC6 SLC6A1 KCNJ6 TMEFF1 KCNQ5 NDE1	ARHGAP32: miR SLC6A11: miR TMEFF1: miR
<i>Ruderfer DM et. al.</i>	KCNB2	SIN3A12 TCF3 NR2C2 FOXP2 REST HMG3 SIN3A HDAC2 ZEB1 SMARCC1 SMARCB1 BHLHE40 SMARCA4	KLF16: miR
<i>Strawbridge RJ et. al</i>	RIMS1	CBX3 EZH2	

Figure 1. Clusters of regulatory elements based on similar biological functions.



[12]. These channels mediate the entry of calcium ions into excitable cells and they are essential in the neurotransmitter release and in a variety of processes including muscle contraction). Dysregulation of these processes and associated alterations in calcium channel activity have been linked to various types of neurological disorders, including epilepsy, migraine, and chronic pain [13].

We identify CACNA1E gene on enhancer, it maintains synaptic plasticity, neuronal survival and fear conditioning and it is expressed on dendrites [12]. The role of calcium channels in mental health is further supported by findings that show that SNPs located in CACNA1C are linked to development of other psychiatric disorders such as bipolar disorder, schizophrenia, and depression [13]. Some schizophrenic patients have an intergenic deletion between classical CDH12 and CDH18 genes [14], suggesting the important role of cadherin genes in neural function. Finally, CLSTN2 gene is predominantly expressed in the brain, and it is involved in learning and memory process.

Potassium channels /electrical excitability of neurons

Voltage-gated potassium (Kv) channels represent the most complex class of voltage-gated ion channels from both functional and structural standpoints. Their diverse functions include regulating neurotransmitter release, heart rate, insulin

secretion, neuronal excitability. Little is known about which ion channels determine the resting electrical properties of presynaptic membranes. However, blockers and activators near the KCNQ5 gene, that we identify on enhancer, could interfere in the depolarization or hyperpolarization of the nerve terminals, altering resting conductance [15]. Moreover, the background conductance set by KCNQ5 channels, in concert with Na⁺ and HCN channels, determined the size and time course of the response to subthreshold stimuli. Thus, ion channels determine the presynaptic resting potential and control synaptic strength [15]. Other genes of interest like KCNJ6 and KCNB2 are also included in this neuronal pathway.

RET signaling cascade

Rearranged during transfection (RET) is the tyrosine kinase receptor that under normal circumstances interacts with ligands at the cell surface, and mediates essential roles in a variety of cellular processes such as proliferation, differentiation, survival, migration, and metabolism [16]. RET plays a pivotal role in the development of both peripheral and central nervous systems and remains expressed throughout all life stages. However, the physiological ligand-dependent activation of RET receptors is important for the survival and maintenance of several neuronal populations, appetite, and weight gain control [16]. In other words, RET is the receptor-tyrosine kinase for the glial cell line-derived neurotrophic factor (GDNF) family ligands and its influence results in the recruitment and assembly of adaptor protein complexes involved in signal transduction downstream of the receptor like DOK6 that in our data play a role on enhancer.

Expression analysis reveals that DOK6 is highly expressed in the development of the central nervous system and is co-expressed with RET in several locations, including sympathetic, sensory, and parasympathetic ganglia [16]. Taken together, these data identify DOK-6 as a novel DOK-4/5-related adaptor molecule that in vivo may be involved in the transduction of signals that regulate RET-mediated processes such as axonal projection [17].

BMP family neuronal patterning

This is one of the most interesting data that we got. At the beginning, we could not find relevant correlations between TMEFF1 and psychiatric disorders because previous reports linked this gene and SNP related variants with cancer develop-

ment. Back on the analysis data resource, we figure out that this gene was postulated by similarity, it means that INFERNO used bioinformatics tools and determine TMEFF1 as a candidate gene that may inhibit NODAL and BMP signaling during neural patterning, due to the similarity with another gene already found in this pathway in another organism. This could be a new approach to postulate TMEFF1 as a gene of interest, but further investigation is needed.

npBAF complex

SMARCCB1, SMARCC1 and SMARCA4 genes are all subunits of the BAF complex. During the development of the vertebrate nervous system, neural progenitors (npBAF) divide and migrate to create specific morphologies and synaptic connections [18].

Period genes (circadian clock)

Sleep and circadian rhythm disruption (SCRD) are a common feature of schizophrenia and is associated with symptom severity and patient quality of life [19]. Genetic association studies have demonstrated that mutations in several circadian clock-related genes are associated with a higher risk of schizophrenia as well as alterations in dopamine signaling. Thereby, dopamine is closely linked to sleep and the circadian system [19].

Previous GWAS studies have associated mutations in circadian clock genes with an increased risk of schizophrenia. Among them are the single nucleotide polymorphisms (SNPs) in CLOCK, PER2, PER3, RORB, and TIMELESS. Respecting HDAC1, HDAC2 and PML genes, may be involved in the transcriptional repression of circadian target genes, such as PER1 and PER2, mediated by CRY1 through histone deacetylation [20]. Moreover, BHLHE40, a transcriptional repressor is involved in the regulation of the circadian rhythm by negatively regulating the activity of the circadian clock genes and circadian clock-controlled genes [21]. TFBS elements are involved in neuronal differentiation (SIN3A), catalysis of important circadian complex (EZH2), and recruitment of the methyltransferase SUV39H1 to the E-box elements of the circadian target genes (like CBX3).

Differentiation and neuronal development/plasticity

This biological process already has an elevated genetic burden. Due to this, we wanted to focus this finding on a few facts. The MEF2-TFBS complex reduces the expression of MEF2 gene in

hippocampal neurons, suggesting that activation of this factor restricts synapse number [22].

According to Fatjó-Vilas, M., 2016, Neuritin 1 gene (NRN1) is involved in neurodevelopment processes and synaptic plasticity and its expression is regulated by brain-derived neurotrophic factor (BDNF). Results suggest that NRN1 variability is a shared risk factor for both schizophrenia-spectrum disorder and bipolar disorders and that NRN1 may have selective impact on age at onset and intelligence in Somatic Symptom disorder SSD.

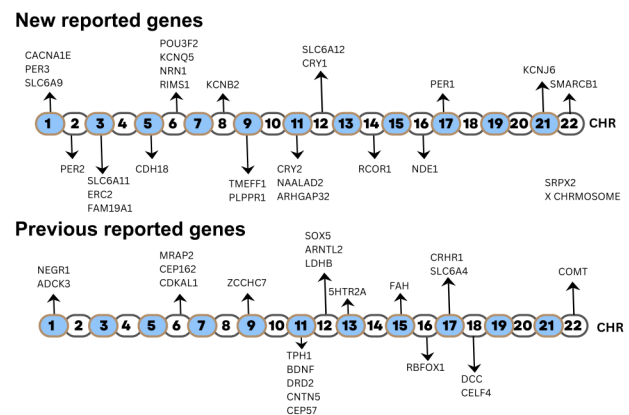
Finally, we compare the distribution of some previous reported genes against the new findings (Figure 2). Suggestive evidence has been reported mostly in chromosome 11, 6 and 12 for serotonin transporter SLC6A4, serotonin receptor 2A 5HT2A, Tryptophan hydroxylase TPH1, TPH2, catechol-o-methyltransferase COMT, brain derived neurotrophic factor BDNF, dopamine receptor D2 DRD2, and Cortico-trophin-releasing hormone 1 CRHR1 gene [1].

Another review reported genes related with voltage-gated potassium (Kv) channel KCNIP4, neuron channel activation PRKCE, neurotrophic factor receptor GFRA1, neural cell adhesion NCAM1 and more biological process related with PAPLN, IL28RA, GDA, ELP3, CTNNA3 and more related genes [22].

CONCLUSION

Of all the psychiatric phenotypes, suicidality remains especially challenging to assess due to the highly polygenic condition.

Figure 2. Candidate genes and distribution in chromosomes.



There is an urgent need to better understand its etiology. As seen in the GWAS of other psychiatric disorders, the number of genetic associations is expected to increase with large sample size. This investigation proposes a different way of approaching the understanding of suicide, highlighting variants affecting regulatory elements that might interfere with the normal transcription of important genes related to behavior and synaptic functions. Exponential use of bioinformatics tools and big data would lead further research in non-coding regions that can provide invaluable biological insights.

This research is the first approach to understanding the functionality and potential of INFERNO. The volume of the output data was significantly high, approximately 100 MB or more for each of the four GWAS data sets. We have processing tools limitation, so we had to prioritize the information that would give us better results, and it was the block of variants with a minor MAF and good linear correlation value that was expressed in Brain tissue. However, we recognize that other variants of a medium-high frequency could be potentially associated with suicidal behaviors and they must be examined in future studies. We obtained GWAS from Americans and United Kingdom people, whereas Latin American, Asian and other European countries were sub-represented.

Data availability

Derived data supporting the findings of this study are available from the corresponding author on request.

Acknowledgements

We specially thank Daniel F. Levey, Niamh Mullins, Douglas M. Ruderfer and Rona J Strawbridge, the correspondence author of each article who gave us the GWAS data. Also to SIGEN laboratory at Los Andes university for the guidance in this project.

REFERENCES

- [1] Mandelli, L., Carli, V., Serretti, A., & Sarchiapone, M. (2011). Suicidal behaviour: genes, environmental stress and temperamental traits. *Suicidology*, 16(2).
- [2] Stein, M. B., Ware, E. B., Mitchell, C., Chen, C. Y., Borja, S., Cai, T., ... & Jain, S. (2017). Genomewide association studies of suicide attempts in US soldiers. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 174(8), 786-797.
- [3] Strawbridge, R. J., Ward, J., Ferguson, A., Graham, N., Shaw, R. J., Cullen, B., ... & Pell, J. P. (2019). Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation with psychiatric disorders and polygenic association with completed suicide. *EBioMedicine*, 41, 517-525.
- [4] Levey, D. F., Polimanti, R., Cheng, Z., Zhou, H., Nuñez, Y. Z., Jain, S., ... & Smoller, J. W. (2019). Genetic associations with suicide attempt severity and genetic overlap with major depression. *Translational psychiatry*, 9(1), 1-12.
- [5] Tanney, B. L. (2000). Psychiatric diagnoses and suicidal acts. In *Comprehensive Textbook of Suicidology* (pp. 311-341). New York: Guilford. Tsai, S. J., Wang, Y. C., Hong, C. J., & Chiu, H. J. (2003). Association study of oestrogen receptor alpha gene polymorphism and suicidal behaviours in major depressive disorder. *Psychiatric Genetics*, 13(1), 19-22.
- [6] Ruderfer, D. M., Walsh, C. G., Aguirre, M. W., Tanigawa, Y., Ribeiro, J. D., Franklin, J. C., & Rivas, M. A. (2020). Significant shared heritability underlies suicide attempt and clinically predicted probability of attempting suicide. *Molecular psychiatry*, 25(10), 2422-2430.
- [7] Krapohl, E., Patel, H., Newhouse, S., Curtis, C. J., von Stumm, S., Dale, P. S., ... & Plomin, R. (2018). Multi-polygenic score approach to trait prediction. *Molecular psychiatry*, 23(5), 1368-1374.
- [8] Mullins, N., Bigdeli, T. B., Børglum, A. D., Coleman, J. R., Demontis, D., Mehta, D., ... & Anjorin, A. (2019). GWAS of suicide attempt in psychiatric disorders and association with major depression polygenic risk scores. *American journal of psychiatry*, 176(8), 651-660.
- [9] Giral, H., Landmesser, U., & Kratzer, A. (2018). Into the wild: GWAS exploration of non-coding RNAs. *Frontiers in cardiovascular medicine*, 5, 181.
- [10] Wasserman, D., & Wasserman, C. (2009). *Oxford Textbook of Suicidology and Suicide Prevention: A Global Perspective*. New York: Oxford University Press.
- [11] Alexandre Amlie-Wolf, Mitchell Tang, Elisabeth E Mlynarski, Pavel P Kuksa, Otto Valladares, Zivadin Katanic, Debby Tsuang, Christopher D Brown, Gerard D Schellenberg, Li-San Wang; INFERNO: inferring the molecular mechanisms of noncoding genetic variants, *Nucleic Acids Research*, 2018, gky686, <https://doi.org/10.1093/nar/gky686>

- [12] Kessi, M., Chen, B., Peng, J., Yan, F., Yang, L., & Yin, F. (2021). Calcium channelopathies and intellectual disability: a systematic review. *Orphanet Journal of Rare Diseases*, 16(1), 1-30
- [13] Simms, B. A., & Zamponi, G. W. (2014). Neuronal voltage-gated calcium channels: structure, function, and dysfunction. *Neuron*, 82(1), 24-45.
- [14] Seong, E., Yuan, L., & Arikath, J. (2015). Cadherins and catenins in dendrite and synapse morphogenesis. *Cell adhesion & migration*, 9(3), 202-213.
- [15] Huang, H., & Trussell, L. O. (2011). KCNQ5 channels control resting properties and release probability of a synapse. *Nature neuroscience*, 14(7), 840-847.
- [16] Mahato, A. K., & Sidorova, Y. A. (2020). RET Receptor Tyrosine Kinase: Role in Neurodegeneration, Obesity, and Cancer. *International Journal of Molecular Sciences*, 21(19), 7108.
- [17] Crowder, R. J., Enomoto, H., Yang, M., Johnson, E. M., & Milbrandt, J. (2004). Dok-6, a Novel p62 Dok family member, promotes Ret-mediated neurite outgrowth. *Journal of Biological Chemistry*, 279(40), 42072-42081.
- [18] Staahl, B. T., Tang, J., Wu, W., Sun, A., Gitler, A. D., Yoo, A. S., & Crabtree, G. R. (2013). Kinetic analysis of npBAF to nBAF switching reveals exchange of SS18 with CREST and integration with neural developmental pathways. *Journal of Neuroscience*, 33(25), 10348-10361.
- [19] Ashton, A., & Jagannath, A. (2020). Disrupted sleep and circadian rhythms in schizophrenia and their interaction with dopamine signaling. *Frontiers in Neuroscience*, 14, 636. Clapham, D. E. (2007). Calcium signaling. *Cell*, 131(6), 1047-1058.
- [20] Hirayama, J., Sahar, S., Grimaldi, B., Tamaru, T., Takamatsu, K., Nakahata, Y., & Sassone-Corsi, P. (2007). CLOCK-mediated acetylation of BMAL1 controls circadian function. *Nature*, 450(7172), 1086-1090.
- [21] Honma, S., Kawamoto, T., Takagi, Y., Fujimoto, K., Sato, F., Noshiro, M., ... & Honma, K. I. (2002). Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature*, 419(6909), 841-844.
- [22] Flavell, S. W., & Greenberg, M. E. (2008). Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. *Annu. Rev. Neurosci.*, 31, 563-590.
- [23] Sokolowski, M., Wasserman, J., & Wasserman, D. (2014). Genome-wide association studies of suicidal behaviors: a review. *European Neuropsychopharmacology*, 24(10), 1567-1577.