

PBL project

How genetic influences are responsible for human behavior and mental diseases such as schizophrenia?

Luis Botella Robin, Diego Gómez Guillén, Júlia Alonso Tor, Marta Martínez Lago

Department of 2nd Biomedicine, UIC, Introduction to Bioinformatics subject

Contact us : [\(1\)](#), [\(2\)](#), [\(3\)](#), [\(4\)](#)

KEYWORDS: DISC1, schizophrenia, bioinformatics, behavioural genetics

Abstract

Genetic factors have long been recognized as responsible for variation in behavior both within the normal period and as mental illnesses. There is no specific gene for each of the behaviors, but rather an unknown set of genes that interact with each other, as well as that the same gene acts on different behaviors, depending on its disposition. Genetic studies have established a role of disrupted-in-schizophrenia-1 (DISC1) in chronic mental diseases, such as schizophrenia.

DATA DRIVEN PROJECT

The aim of this project is to collect gene data related to Schizophrenia in order to find out significant differences comparing wild type genes with Schizophrenia mutated genes already discovered, to conclude that these mutated genes are responsible for Schizophrenia and not for irrelevant mutations or other non-behavioural pathologies. We decided to turn it into a data driven project, as our main aim is to browse and explore project data to fight against project obstacles in order to get better results.

INTRODUCTION

Behavioral genetics is a discipline that studies the role of environmental and genetic influences on behavior, with subspecialties focused on the behavioral genetics of humans and animals and the variations between them. [\(1\)](#)

In addition to employing methodologies to understand the nature and origins of individual behavioral differences, the field of behavioral genetics is highly interdisciplinary, contributing knowledge of biology, neuroscience, genetics, epigenetics, ethology, psychology, and statistics. Additionally, behavioral genetics is associated with the "nature and nurture" debate.

Within behavioral genetics there are interrelated subfields, which are psychiatric genetics, epigenetic research in behavior, and genetic research in neuroscience. Genetic factors have long been recognized as responsible for variation in behavior of people with mental illness, among which autism, bipolar disorder and schizophrenia are the most relevant.

Studies have corroborated that there is no specific gene for each of the behaviors, but rather an unknown set of genes that interact with each other, as well as that the same gene acts on different behaviors. And those responsible for these can be both environmental and genetic factors, and what we will do is focus on the last named, to give evidence that these peculiar diseases so little studied can be passed on to the offspring.

The initial objective of our PBL was to associate and work with a large number of different genes relevant to behavioral genetics and associate them with the most well-known mental illnesses cited above. However, given the complexity of the search, and, being such a broad field, we narrowed down the search and decided to focus on schizophrenia. Finally, once the genes with which we were going to work had been decided, we realized that the databases of said disease are not complete enough to carry out a quality work, realizing that, today, schizophrenia, despite its relevance worldwide is very little studied, hence, finally, we have had to focus on the study and work of a single gene and its encoded scaffold protein: DISC1 (Disrupted in schizophrenia 1).

When the word schizophrenia is heard, everyone knows at a minimum what it refers to, however, they are not aware of what it really is. Schizophrenia is defined as a severe, heritable, refractory and a very serious psychiatric disorder. It affects approximately 1% of the world's population. This disease deteriorates the most basic processes of perception, emotion and judgment of the people who suffer from it. Many times, people with this disorder hear or see things that are not present. They also often think that others may read their minds, control their thoughts, or conspire to harm them; hence it terrifies them and they become withdrawn people.

The biology of schizophrenia is complex with multiple hypotheses (dopamine, glutamate, neurodevelopmental) well supported to underlie the disease. Several studies have shown that the disrupted in schizophrenia 1 (*DISC1*) gene is closely associated with schizophrenia by its role in neuronal morphology and synaptic function, so it is one of the most promising candidate genes for major mental behavioural disorders. Indeed, schizophrenia pathways are mainly centered on the risk factor "disrupted in schizophrenia 1" (*DISC1*).

THEORETICAL FRAMEWORK

*"Even though *DISC1* was first linked to major mental illness in 2000, and its importance subsequently confirmed and replicated in numerous independent genetic studies, no full-length or even partial/fragment experimental 3-D structures have been forthcoming. Indeed, to date, biophysical characterization of the full-length protein is still almost totally lacking."* [\(2\)](#)

Nowadays, the structure of *DISC1* protein can only be seen in two species: homo sapiens and mice. In fact, to date, the biophysical characterization of the full-length protein is still almost completely lacking. In the absence of structural information for *DISC1*, experimental work has relied on shorter constructs and domain delineation based on sequence analysis. This lack of structural information leaves us with a large information gap to understand the underlying effects of missense mutations and single nucleotide polymorphisms (SNPs) that have been linked to the disease susceptibility.

Some of the questions we try to answer during this project are:

- Which is the most prevalent gene, that if disrupted, causes phenotype hints of schizophrenia?
- What are the mutations more susceptible to cause schizophrenia disease? Which is the most important one?
- Does sequence variation in *DISC1* impact the structure and function of the expressed protein? If so, how?

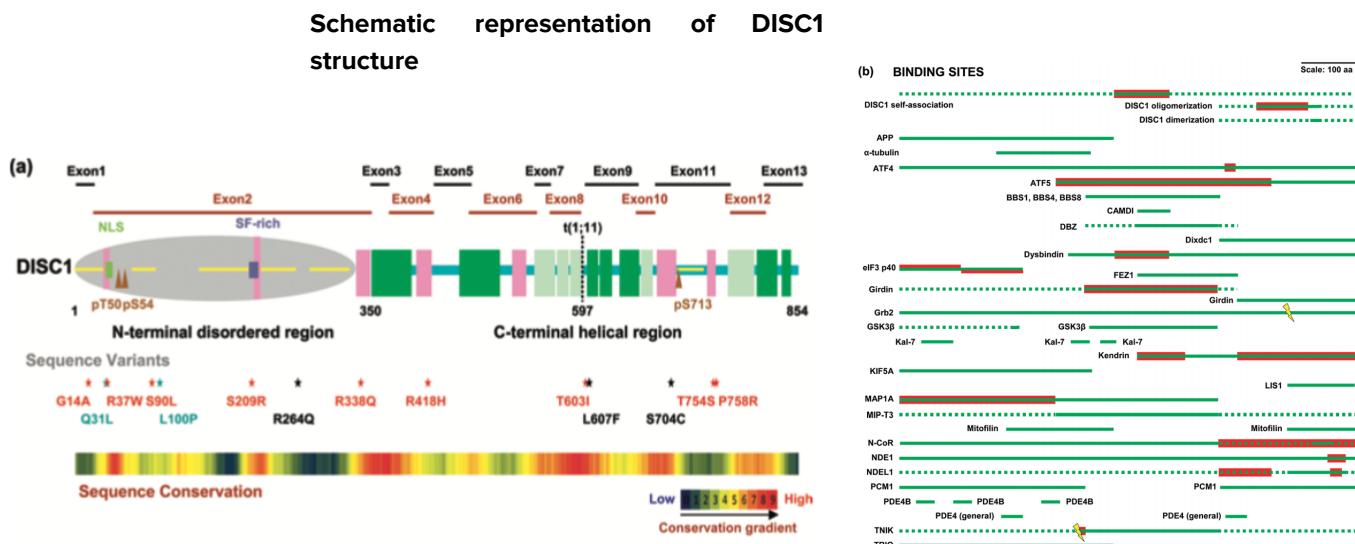
1. STRUCTURAL INFORMATION OF DISC1

First, let's describe the gene: it encodes for a protein which is located in the nucleus, cytoplasm and mitochondria; its transcript is expressed in the ventricular/subventricular zone during embryonic development and in the dentate gyrus of hippocampus in adults, as well in all the brain regions. The protein is involved in neurite outgrowth and cortical development through its interaction with other proteins. (3)

Disrupted in schizophrenia 1 (DISC1) is a protein originally identified from a gene directly disrupted by a translocation in a large Scottish family with a higher burden of serious mental illness. Exist considerable genetic evidence as a risk factor for a wide range of psychiatric illnesses, including schizophrenia, bipolar disorder, depression and possible autism spectrum, in fact, there is a strong evidence of non-neutral, accelerated evolution particularly for exon 2, the only coding region within the schizophrenia-associated haplotype (4) Moreover, DISC1 is present in several neurodevelopmental pathways of which affect mental disease pathology. It acts as a scaffold, binding a number of other proteins, several of which have been shown to be independent risk factors for serious mental illness.

The full-length DISC1 sequence consist of two regions: an **N-terminal** “head” domain containing about ~1–350 amino acid residues with does not share homology with any know folds, and a **C-terminal** spiral region about ~350–854 that shows greater conservation among orthologs than the N-terminus.

The N-Terminal Region of DISC1 has only two regions of conservation that correspond to a nuclear localization signal sequence and a serine-phenylalanine motif highly conserved. The predicted regions of disorder are assigned to areas of least conservation. DISC1 is composed of about 13 exons.



The N-terminal region possesses only two notable regions of conservation that correspond to a nuclear localization signal sequence: motif18 and a serine-phenylalanine-rich (SF-rich) motif. The N-terminus of DISC1 is often referred to as the globular head domain, this region contains numerous low-complexity segments. Delineating the Coiled-Coil Regions within the C-Terminus and Assessing DISC1 Secondary Structure. The chief characteristic of most coiled-coils is a regular seven-residue sequence pattern known as the heptad repeat (alpha helix). These alpha helix are readily identified by sequence analysis

and structurally represent a distorted R-helix where the seven-residue repeat forms two turns (or 3.5 residues per turn) as opposed to a regular R-helix where 3.6 residues form one turn. Despite this apparent simplicity and relatedness at the sequence level, coiled-coils display a considerable degree of structural diversity; the helices may be arranged parallel or antiparallel and may form a variety of oligomeric states.⁽⁵⁾

2. ABOUT ISOFORMS OF DISC1

We believe that once the structure of DISC1 has been explained, it is important to take into account that there are various isoforms for this protein. In this way, we have made an MSA, in Python; with the isoforms registered in Uniprot for this specific gene. Thus, we can study in which these isoforms diverge, and which are the most conserved areas. ([FULL MSA](#)).

As we can see in the displayed alignment, the least conserved region among DISC1 isoforms is the C-terminal region. This makes us think that maybe the N-terminal region conservation leads to higher interaction phenomena among all these isoforms, and the fact that it may have more interactions than C-terminal region could justify the fact that the N-terminal region (corresponding to exon 2 in genomic sequence) is the most pathogenic region of human DISC1.

Alignment with 11 rows and 855 columns

```

MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG...AQA sp|Q9NRI5|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG...AQA sp|Q9NRI5-2|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-3|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-4|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG...LQ- sp|Q9NRI5-5|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-6|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-7|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-8|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-9|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG..... sp|Q9NRI5-10|DISC1_HUMAN
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPG----- sp|Q9NRI5-11|DISC1_HUMAN

```

The conserved position indexes are:

```

[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]
['M', 'P', 'G', 'G', 'G', 'P', 'Q', 'G', 'A', 'P', 'A', 'A', 'A', 'G', 'G', 'G', 'G', 'V', 'S', 'H', 'R']

```

Phylogenetic analysis of isoforms with IQtree ([tree files](#)):

```

***sp_Q9NR15_5_DISC1_HUMAN
| 
+**sp_Q9NR15-2_DISC1_HUMAN

    +**sp_Q9NR15-3_DISC1_HUMAN
    +-+ (0/57)
    |   +-sp_Q9NR15-5_DISC1_HUMAN
    |   +-+ (71.4/47)
    |       +**sp_Q9NR15-9_DISC1_HUMAN
    +-+ (0/33)
    +-+sp_Q9NR15-11_DISC1_HUMAN

+** (0/20)
| 
| +-----+sp_Q9NR15-4_DISC1_HUMAN
| +-----+ (96/98)
| 
| +-+ (0/43)
|   +-+sp_Q9NR15-7_DISC1_HUMAN
| 
+----+ (90.2/70)
|   +-+sp_Q9NR15-6_DISC1_HUMAN
| 
+-+ (86.5/77)
|   +-+sp_Q9NR15-8_DISC1_HUMAN

```

3. CONDUCT AND DESCENDANTS TRANSMISSION

Schizophrenia is not usually transmitted from parents to children, because people who have it do not usually have children; most people with schizophrenia have no family history (approximately 63%). Hence, most of the genetic mutations that occur are new. However, it has been associated with the behaviors that parents develop, in terms of communication to children if they can generate a certain spectrum of schizophrenia. (6)

In contrast inheritance of the translocation is causal and increases risk of developing one of these disorders by ~50-fold in comparison to the general population.

The more relatives who suffer from the disease, the more likely they are to suffer from it. Although the closeness of those relatives is also very important: it is not the same as being a cousin than a father or a brother.

RISK FACTORS	if factor is present, number of times that increases general risk
Family history of schizophrenia monozygotic twins both parents with disease dizygotic twin grandparents with disease 3rd grade relatives with disease	50-70 40-60 9-18 3-6 2-3
Specific gene variability	1.1-1.5
Living on a town	2-3
Being emigrant	2-3
Food infection	2-3
Birth on winter	1.1
Complications in childbirth	2-3
Use of cannabis	2-3
>35 y/o parents	1.5-3
Being male	1.4

Table showing the prevalence to Schizophrenia risk among both genetic (family history) and environmental factors proving that genetic influences exceed the risk index for developing the disease

In fact, Disrupted in schizophrenia 1 (DISC1) is a protein originally identified from a gene directly disrupted by a balanced translocation ((1;11)(q42;q14.3)) in a large Scottish family with a high loading and phenotype of major mental illness.

4. FILOGENIA

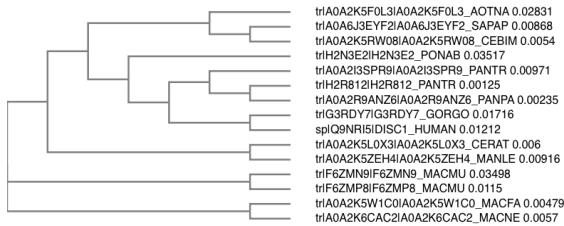
It has been seen that the protein with which we are going to work, DISC1, is very similar to the DISC1 proteins that belong to a series of selected primates; [MSA PRIMATES](#).

However, these species are not capable of developing schizophrenia; Though psychotic animals may exist, schz-psychosis has never been observed outside of our own species.[\(7\)](#)

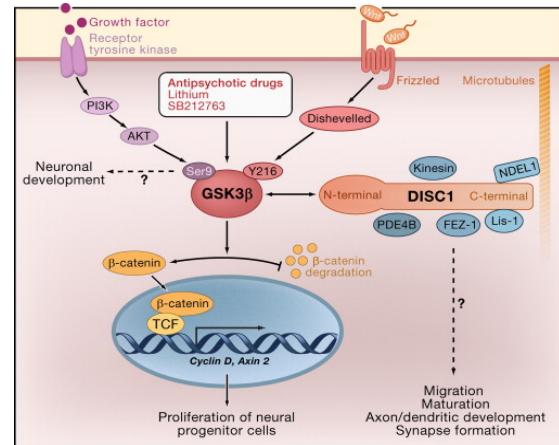
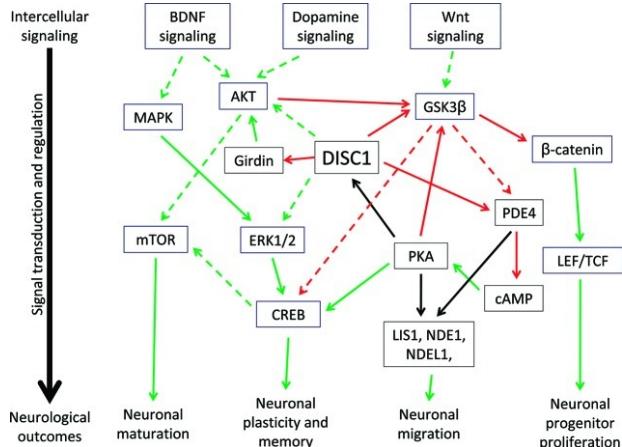
The appearance of human speech and language expression and behaviour is related to the genetics of schizophrenia, since, in other animals, it is not a relevant function. In fact, language dysfunction is a feature of schizophrenia.

Therefore, despite its structural similarity, schizophrenia is a disorder exclusively responsible for modifying human behavior, and not that of other animals

[Phylogenetic tree](#) using the 14th most similar sequences to the human DISC1:



5. DISC 1 PROTEIN INTERACTIONS



Although DISC1 SNPs can confirmly lead to a high risk of schizophrenia, this protein makes huge complexes of interaction leading to its function. It has an important role in the Dopamine and Glutamate signalling pathway, but here we are going to focus on some important interactions with particular proteins: GSK3 β , PDE4 and NDEL1.

GSK3 β interaction: this is a “virtually ubiquitous enzyme present throughout hippocampal neurons, in dendritic spines, and present in both the cytosolic and synaptosomal subcellular fractions. The enzyme has a large array of substrates and multiple well demonstrated roles which include regulation of cell

proliferation and synaptic plasticity with a peak in expression between embryonic day 18 and postnatal day 10, which corresponds to the major period of synapse formation.” [\(8\)](#)

The interaction DISC1-GSK3 β takes place in the N-terminal region of DISC1 as we can see in the picture above, concretely GSK3 β binds directly to fragments 1–120 and 356–595 of DISC1, and these fragments inhibit GSK3 β autophosphorylation. This makes us think that any mutation at the N-terminal region in DISC1 may be a danger for this interaction performance, and thus, leading to a missensed period synapse formation, neuronal proliferation, plasticity... and a potential risk of schizophrenia phenotype. This is what actually happens, “A83V, R264Q and L607F variants cannot stimulate cell division/proliferation compared to WT-DISC1 or the S704C variant”, meaning that these variants have a reduced interaction with GSK3 β and impact directly on brain development.

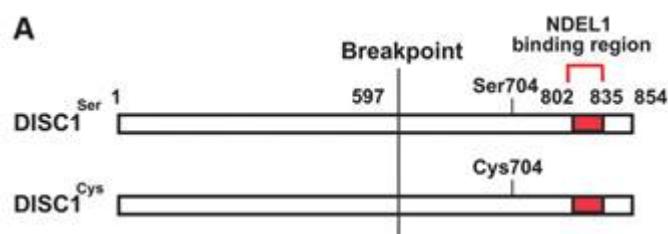


PDE4 interaction: what usually phosphodiesterases do is inactivate cAMP to regulate its spatial and temporal fluxes. *“More recent work has suggested that PDE4 activity is also intrinsically linked to the outcome of GSK3-DISC1 interactions,* but there is no evidence of those 3 proteins forming an interaction complex, making it difficult targeting this pathway with therapeutic compounds for this disorder among others. [\(9\)](#)

As we can see, DISC1-PDE4 occurs at the ending fragment of the N-terminal fragment. This interaction is isoform dependent (DISC1 has may and PDE4 has 4). A key player here is cAMP in the outcome of the interaction DISC1-PDE4-GSK3, so; any disruption in DISC1 at the ending regions of the N-terminal fragment could riskily alter the pathway track and result in a disrupted in a neuronal migration and neuronal plasticity.

NDEL1 interaction: unlike the previous interactions, this one happens in the DISC1 C-terminal fragment, and it is an essential interaction for neurite outgrowth.

“We addressed whether the Ser704Cys variation of DISC1 (resulting in the proteins DISC1Ser and DISC1Cys, respectively) might influence the protein interaction of DISC1 and NDEL1, using the same strategies as above (Fig. 5A). In yeast two-hybrid assays, the DISC1Cys fragment (597–854 amino acids) displays a slightly stronger interaction with NDEL1 when compared with that of the DISC1Ser fragment, although the differences remain in a suggestive range.” [\(10\)](#)



Hashimoto and his colleagues have reported that Cys allele is associated with an increased risk for major depressive disorder and with a reduced gray matter volume in the cingulate cortex. Surprisingly, DISC1Cys has a stronger affinity for NDEL1 compared to DISC1Ser; “such subtle but consistent changes may increase a risk of major mental disorders by providing a chronic disturbance of DISC1–NDEL1 function.” There is a key binding region in DISC1 for NDEL1, a coiled coil leucine zipper in the C-terminal (amino acids 807-828) crucial for the interaction to take place. This interaction is required for neurite growth; and NDEL1 has been reported to associate with cytoskeletal components, playing a role in cytoskeletal stabilization, cell mitosis, membrane trafficking and neuronal migration. Then, if this enzymatic activity is disturbed, potentially neurite growth, neuronal proliferation and migration would be affected too, leading to the known and repeated schizophrenia disease.

METHODS

We needed a reference protein sequence, which is **human DISC1 L isoform**, that corresponds to the entire protein sequence (854 aa).

To find and analyze possible DISC1 (**prot.**) mutations related to schizophrenia, we first looked for mutations of interest in [Uniprot](#) and [Szdb](#) (because in the variant table of [Ensembl](#) not appear any variant provided with clinical significance), and once found; secondly, we proceeded to analyze them. In these databases we collect sequences, papers, annotations ... that served us to follow a protocol to know and discuss these mutations more efficiently.

Chiefly, we decided to study the degree of dissimilarity generated (PID) by the mutation in reference to the L isoform of DISC1 (the most abundant isoform of that gene in the organism); using [Python](#) to make a pairwise sequence alignment (PSA).

Chimera Procedures: The goal at using this tool was at first to have a look at the wild type Human DISC1 protein sequence (854 aa); but unfortunately, there is no structural information or any identifier for Protein Data Bank (PDB). We blasted this sequence to find similar sequences among other species. We selected the PDB as the search set and excluded the human organism option. From this blast we obtained similar sequences to DISC1 at a structural level, and the only ones registered were *mus musculus* DISC1 fragments; these sequences have a 68,89% percentage of identity (not compared to the full human DISC1 protein, but with the structural information about it found in the database). This is our handicap, regardless of it, we managed to use *mus musculus* DISC1 protein sequence, which is 145 aa long.

The procedures went like this:

- Check where the mutation falls in the human sequences and find if the mutation site exists in the mouse sequence.
- If it exists, call the function “tools” → sequence → sequence, and select in green in the sequence viewer the amino acids of interest to see where they fall in the structure.

Enter Query Sequence

BLASTP programs search protein databases using a protein query. [more...](#)

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#)

NP_061132.2

Query subrange [?](#)

From To

Or, upload file Seleccionar archivo Ningún archivo seleccionado [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

Align two or more sequences [?](#)

Sequences producing significant alignments									Download	Select columns	Show	100	?	
<input checked="" type="checkbox"/> select all 2 sequences selected				GenPept	Graphics	Distance tree of results	Multiple alignment	MSA Viewer						
	Description			Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession			
	<input checked="" type="checkbox"/> Solution Structure of the DISC1/Ndel1 complex [Mus musculus]			Mus musculus	112	112	10%	2e-28	68.89%	145	5YI4_A			
	<input checked="" type="checkbox"/> Solution structure of DISC1/ATF4 complex [Mus musculus]			Mus musculus	110	110	10%	4e-28	68.89%	133	6IRR_A			

- After getting the visual information, we will try to explain the significance of the mutation and make some conclusion on its structural effect and interaction capacity.

Due to the difficulties we had studying the changes that occurred in the protein structure of DISC1 and due to lack of prior art, we decided to proceed to look for another tool that could allow us this type of analysis in sequences with no PDB ID. For this, we made use of a predictor of secondary structure, an online tool called [SOPMA](#), that predicts in percentages the secondary structure of a specific sequence. SOPMA is an improved SOPM method, correctly predicting 69.5% of amino acids for a three-state description of secondary structure (alpha helix, beta sheet, and spiral) in a comprehensive database containing 126 chains of (less than 25% identity) proteins.

Analyzing with SOPMA, the changes in the structure percentages between the isoform L (reference percentages) and the mutated sequences of interest. In this way, depending on where the variance of percentages occurred, we can suppose where the mutation falls (with the help of reading the papers).

As we already know, the structure is highly linked to the function of a protein, therefore, it was finally convenient to study what types of changes in the function of the protein varied due to the structural change. Thus, analyze the changes produced in the pathway and their role in the development of schizophrenia.

RESULTS

1st mutation:

First, we analyze the most abundant type of DISC1 mutation in relation to the development of schizophrenia: **L → P** at position 815 and **L → P** at position 822, they are linked. To study the mutation we are going to proceed according to the protocol discussed in the methods section. First of all we extract the mutated sequence:

- MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPGYMRSSGPGIGFLSPAVGTLFRFPGGVSG
EESHSESRARQCGLDSRGLLVRSPVSKSAAAPTVSRGTSAGFGIQLRGTRLPDRLSWPCGPGSAGWQQE
FAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPCGPEVPPTPPGSHSAFTSSFSFIRLSLGSAG
ERGEAECPSPREAESHCQSPQEMGAKAASLDGPHEPDPRCLSRPFSLLATRVSAADLAQAARNSSRPERDMHSL
PDMDPGSSSSLDPAGCGGDGSSSGDAHSWDTLLRKWEPVLRDCLLRNRRQMEVISLRLKLQLQDAVEN
DDYDKAETLQRLEDLEQEKFISLHFQLPSRQPALSSFLGHЛАQVQAALRRGATQQASGDDTHTPLRMEPRLEP
TAQDSLHVSITRDWLLQEKKQQLKEIEALQARMFVLEAKDQQLRREIEEQEQQLQWQGCDLTPVGQLSLGQL
QEVSALKQDTLASAGQIPFHAEPPETIRSLQERIKSLSNLKIEITTVCMSKFCSTLRKKVNDIETQLPALLEAKMH
AISGNHFWTAKDLTEEIRSLTSEREGLEGLLSKLLVLSRNVKKLGSVKEDYNRLRREVEHQETAYETSVKENTMKY
METLKNKLCCKCPLLGKWEADLEACRLLIQSQLQEARGLSVEDERQMDDLEGAAPPRLHSEDKRKTPL
KVLEEKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGKKLLYLEDQLHTAIHSHDEDLIQSLRREPQMVKETPQ
AMILQLQPAKEAGEREAAASCMTAGVHEAQ

Pairwise alignment & PID:

```
Pretty print first alignment:
MPGGPGQPAAGGGVGHSRAGSDCLPPAACFRRRRLARRPGYMRSSGPGICGLFLSPAAGVTLFRPGGVGSEESHHSESARQCGLDSRGLLVRSPVSKSAAAPTVTSVRGTSAH
FGIQLGRGGTRPLRSWCPGSGAWQGFEFAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNCSPGCPEVPPTPGSHSAFTSSFSFIRLSLGSAGERGEAEGCCPSSRE
AESHCQSPQEMGAAKASLDGPHEPDPRCLSLRFSLLATRVSADLAQAARNSSRPERDMHSLPDMDPGSSSSLDPSLAGCGGDGSSSGDAHSDTLLRKWEPVLRDCLLRNRQRMEVI
SLRLKLQLQKLODAEVDNDYKAETLQRQLLEDLEQEKKISLHFQLPSRQALPSSFLGHIAAQVQLARRGATQOASGDDHTHFLRMPERFLPTEAQLDSLHSWITSTRDWLQLQQKOLKE
IEALQARMFVLEAKDQQLRRIEEQEQQLWQGQCDLTPLVGQLSLQLEQVFSKALQDTLASAGQIPFHAEFPETIRSLQERIKISLNSLKEITTVCMSKEFCSTLRKVNDIETQL
PALLEAKMHAIISGNHFWTAKDLTEEIRSLTISEREGLEGLLSKLLVLSRSSRNVKLGSVKEDYNNLRREVEHQETAYETSVKENTMKYMETLCKNKLCSCKCPLLGKVWEADLEACRLLI
QKQQLQEARQGSLSVEDMDLEGAAPPTRPLHSEDESKRTPKLVLEEWKTHLISLHAGGEQKEEYSILSAELGEKCEDIGKLLYLEDQLHTAIHSHDEDLIQLSRREL-QMV
KETI_-QAMTLO_OLPRAEAFGRPAASCMTCAGVHFOA
```

MPGGGPGQAPAAAAGGGGVSHRAGSRDCLPAAACFRRRRLARRPGYMRSSGTGPGICFLSPAVGTLFRPGVGVSSEESHHSESRAQRQCGLDSRGLLVRSVPKSAAAPTVTSVRGTSASH
FGIQLRGGTRLPDRLSWPCCPGSAGWQQEFAAMDSSETLDAWSAEACSDGARRVRAAGLSPAEELSSNSCSPGCCPEVPPTPGHSRAFTSSFSFIRLSLGSAGERGEAEGCPPSRE
AESHCQSPQEMGAKAAASLDGPHEPDCKLRSFPSSLATRVSADLAQARNNSSPERDMHSLPDPMDGSSSLDPSLGAZCGGDGSSGSGDGAWSDHSTDLRKLHRWPVLRDCLRNRRQMEVI
SLRKLQLQEDADYDQKATLQLRDELEQEKISLHLQPLSPRQALPSFLHLQAQVQALARRGATQASQGDTHTFLPRMELPRLTEQADLSHVSIITRDRWLQEQKQLQKE
IEALQARMFVLEAKDQQLRREIEEQEQLWQOGCDLTPLVGOLSLGQLEQEVSKALQDTLASAQOIPFHAEPPETIRSLSQERIKLSNLSLKEITTKVCMSEFKFCSTLRKVNDIETQL
PALLEAKHMIAISGNHWFTAKDLTTEEIRLSTSREREGLEGLLSKLVLSSRNVKLGSVKEDYNRNLREVEHQETAYETSVKENTMKYMTFLKNKLCSCCKPILLGVWEADLEACRLLI
OSLQLQEARGLSLSVEDERQMDLEGAAPTPRPLHSEDKRTPLKVLEEWKTHLIPSLHCGAGEQKESYIISLAELEGKCEDIGKLLYLEDQLHTAIHSHDEDLIQSRLRE-PQMV
KET-PQAMILQLQPKAEAGEREEAAASCMTAGVHEAQAA

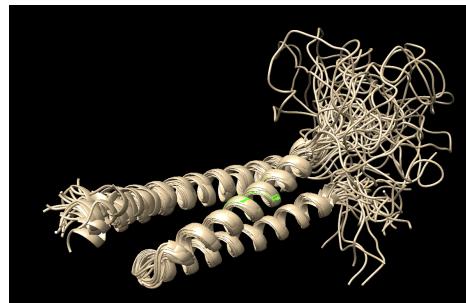
Alignment 0 PID:
0 9976580796252927

Second, we analyze the changes in the protein structure due to the commented mutation. To carry out this analysis we used Chimera, so that we marked the mutated amino acids to know where they fell. As we discussed earlier, we did not obtain PDB ID for DIS1 human protein, that is why we used the protein structure of DISC1 mouse. In humans, the mutated region where Leucines (L) are changed by Prolines (P) are in the last chain of the human DISC1 protein (positions 815 and 822). This region in mouse DISC1 protein is conserved but with a mutation in 1 position (Leucines are conserved through):

- Human region → LQMVKETL
 - Mouse region → LQTVKETL

Now we have to look at the location of this region in chimera on the mouse protein representation:

- Mutation region position:
 - Chain A location 0.20.
 - Positions 815 and 822.
 - C-terminal end.
- Prolines in alpha helices, cause a kink in the helix. This kink is caused by proline being unable to complete the H-bonding chain of the helix and steric or rotamer effects that keep proline from adapting the preferred helical geometry. The kink is 'away' from the proline residue, which is frequently on the hydrophilic side of the helix. [\(11\)](#)



SOPMA :

Alpha helix	(Hh)	:	431	is	50.47%
3 ₁₀ helix	(Gg)	:	0	is	0.00%
Pi helix	(Ii)	:	0	is	0.00%
Beta bridge	(Bb)	:	0	is	0.00%
Extended strand	(Ee)	:	34	is	3.98%
Beta turn	(Tt)	:	29	is	3.40%
Bend region	(Ss)	:	0	is	0.00%
Random coil	(Cc)	:	360	is	42.15%
Ambiguous states (?)	:	:	0	is	0.00%
Other states	:	:	0	is	0.00%

SOPMA :

Alpha helix	(Hh)	:	425	is	49.77%
3 ₁₀ helix	(Gg)	:	0	is	0.00%
Pi helix	(Ii)	:	0	is	0.00%
Beta bridge	(Bb)	:	0	is	0.00%
Extended strand	(Ee)	:	35	is	4.10%
Beta turn	(Tt)	:	29	is	3.40%
Bend region	(Ss)	:	0	is	0.00%
Random coil	(Cc)	:	365	is	42.74%
Ambiguous states (?)	:	:	0	is	0.00%
Other states	:	:	0	is	0.00%

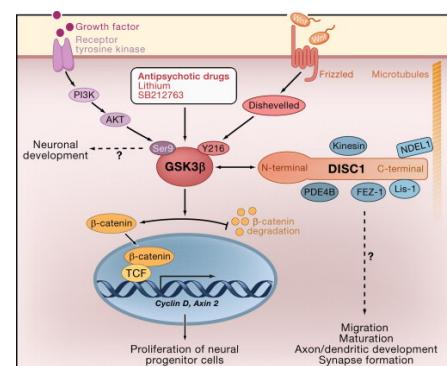
Isoform L

x2 Proline mutation

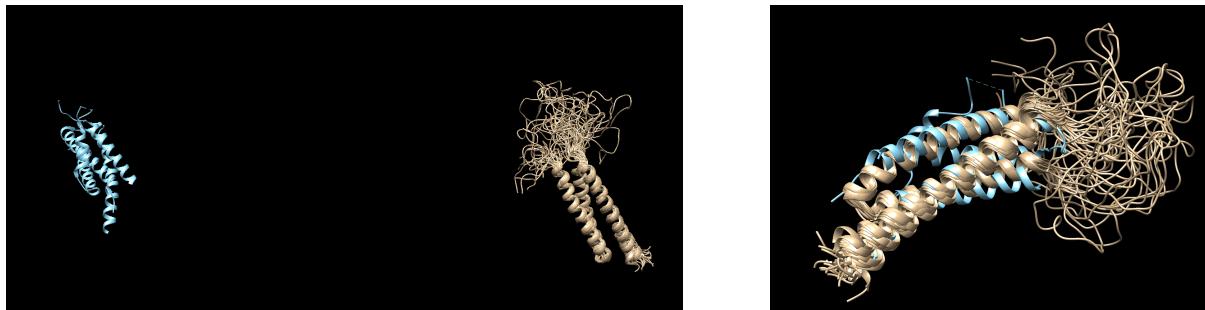
SOPMA secondary structure predictor doesn't seem to be that significant regarding the secondary structure. Nevertheless, it turns out that alpha helix formation % is lower; extended strand formation % is higher, beta turn % remains the same and random coil % is a little higher as well. These results tell us that the mutated protein is less likely to fold as the wt one does, corresponding to the "kink" tendency of Proline amino acids.

Finally, time to discuss what would be the mutation effects on the protein function itself and study in which the interactions with other important pathway proteins would be affected.

- In the linked paper, further confirmation that the NDEL1 interaction site with DISC1 is dependent on the C-terminus of DISC1 was obtained by performing GST extraction assays using extracts of HEK293 cells expressing DISC1 full-length GFP (isoform L). [\(11\)](#)
- DISC1 and NDEL1 interact in the developmental stage of the brain in the activation of GSK3b in neural development.
 - Disruption of a specific leucine zipper / spiral domain in DISC1 inhibits interaction with Nudel



Visual interaction between wt mouse DISC1 and NDEL1:



2nd mutation:

Another SNP; the most prevalent one among schizophrenia affecteds; that leads to a high risk schizophrenia phenotype is the **Ser704Cys mutation**; where a Serine in position 704 is substituted missensely by a Cysteine. According to Epstein coefficient of difference, Ser and Cys differ in 0.6 (60%) on polarity and size. According to the BLOSUM matrix, there is a mismatch between Ser and Cys with a score of -1. This may tell us that there is no benefit on this amino acid replacement. We can see in red where the mutation takes place:

- MPGGGPQGAPAAAGGGGVSHAGSRDCLPPAACFRRRRLARRPGYMRSSSTGPGIGFLSPAVGTLFRPGVGSG
EESHHSERARQCGLDSRGLLVRSPVSKSAAAPTVTSGRTSAHFGIQLRGGTRLPDRLSWPCPGSAGWQQE
FAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPGCGPEVPPTPPGSHSAFTSSFSFIRLSLGSAG
ERGEAEGPCPSREAESHQCSQPQEMGAKAASLDGHEDPRCLSRPFSSLATRVSAADLAQAARNSSRPERDMHSL
PDMDPGSSSLDPSLAGCGGDGSSSGSGDAHSWDTLLRKWEPVLRDCLLRNRRQMEVISRLKLQKLQEDA
DDYDKAETLQRLEDLEQEKEKISLHFQLPSRQPALSSFLGHЛАAQVQAALRRGATQQASGDDTHPLRMEPRLL
TAQDSLHVSITRRDWLLQEKKQLQKEIEALQARMFVLEAKDQQLRREIEEQEQQQLQWQGCDLPLVGQLSLGQL
QEVSQALQDTLASAGQIPFHAEPPETIRSLQERIKSLNLSLKEITTKVCMSEKFCSTLRKKVNDIETQLPAL
AISGNHFWTAKDLTEEIRSLTSEREGLEGLLSKLLVSSRNVKLGSVKEDYNRLRREVEHQETAYETSVKENTMKY
METLKNKLCSCCKCPLLGVWEADLEACRLLIQ**C**LQLQEARGLSVEDERQMDDLEGAAPPPIPRLHSEDKRKTPL
KVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGKKLLYLEDQLHTAIHSHDEDLIQLRRELQMVKETLQA
MILQLQPAKEAGEREAAASCMTAGVHEAQA

Pairwise alignment & PID:

```
Pretty print first alignment:
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPGMRSSTPGIGFLSPAVGTLFRFPGVSGEESHSESRARQCGLDSRGLLVRSPVSKAAAPTVSVRGTS
FGIQLRGTRLPDRSWPCPGSAGWQQFEAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPCGPPEVPTPPGSHSFTSSFSFIRSLSGSAGERGEAEGC
PSRE
AESHCQSPQEMGAKAASLDGPHEPDPRCLSRPFSLLATRVSADLAQAARNSSRPERDMHSLPDMDPGSSSLDPSLAGCGGDGSSGSDAHSDTLLRKWEPVLRDCL
LRRNRRQMEVI
SLRLKLQLQEDAVENTDDYDKAETLQQRLEDLEQEKIISLHFQLPSPRQPALSSLGHIAQVQAALRRGATQQASGDDHTPLRMEPRLLLEPTAQDSLHSITR
RDWLQEKQQLQKE
IEALQARMFVLEAKDQQLRREIEEQEQQQLQWQGCDLTPLVGQLSLQOLQEVSKALQDTLASAQIPFHAEPPETIRSLQERIKSLNLSLKEITTKVCMSEK
CSTLRKKVNDIETQL
PALLEAKMHAISGNHFWTAKDLTEIRSLTSEREGLEGLLSKLLVLSSRNVKLGSVKEDYNRLRREVEHQETAYETSVKENTMKYMETLKNKLCSCCKP
LLGKVWEADLEACRLLI
QS-LQLQEARGSLSVEDERQMDLEGAAPPPIPRLHSEDKRKTPLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGKLLYLEDQLHTAIHSHDE
LIQSLRRELQMV
KETLQAMILQLQPAKEAGEREAAASCMTAGVHEAQA
Score=853
```

Alignment 0 PID:
0.9988290398126464

Second, we analyze the changes in the protein structure due to the commented mutation. To carry out this analysis we used SOPMA, because it turns out that the mutagen region for Ser704Cys does not exist in mus musculus protein sequence. With the intention of studying where the mutations fall, we have compared the prediction percentages of the protein structure of Isoform L of DISC1 and that of DISC1 Ser704Cys, provided by SOPMA

- Isoform L:

SOPMA :	
Alpha helix (Hh) :	431 is 50.47%
β_10 helix (Gg) :	0 is 0.00%
Pi helix (Ii) :	0 is 0.00%
Beta bridge (Bb) :	0 is 0.00%
Extended strand (Ee) :	34 is 3.98%
Beta turn (Tt) :	29 is 3.40%
Bend region (Ss) :	0 is 0.00%
Random coil (Cc) :	360 is 42.15%
Ambiguous states (?) :	0 is 0.00%
Other states :	0 is 0.00%

- Ser704Cysr:

SOPMA :	
Alpha helix (Hh) :	438 is 51.29%
β_10 helix (Gg) :	0 is 0.00%
Pi helix (Ii) :	0 is 0.00%
Beta bridge (Bb) :	0 is 0.00%
Extended strand (Ee) :	31 is 3.63%
Beta turn (Tt) :	24 is 2.81%
Bend region (Ss) :	0 is 0.00%
Random coil (Cc) :	361 is 42.27%
Ambiguous states (?) :	0 is 0.00%
Other states :	0 is 0.00%

Changes are mainly observed in the alpha helix, of the N-terminal. We corroborate the conclusions with the linked paper(12). These results may be indicators of a higher probability of stabilizing the alpha helix (where the change % is more significant), altering the protein interactions.

So, regarding cysteine properties compared to serine, we know that cysteine is a much more acidic amino acid that tends to make disulfide bonds that play a key role in stabilizing proteins. But, as we know, DISC1 has a lot of different isoforms that interact with other proteins, and miss-interactions are a clear cause/risk of schizophrenia.

This makes us think that this particular mutation may “block” in some way interactive regions of some DISC1 isoforms (by wrongly stabilizing the mutated region), leading to a bad interaction with GSK3 β or PDE4 (if this pathway is disrupted, schizophrenia or other mental illnesses may take place).

3rd mutation:

This time, we will study a particular Isoform, **Isoform 11**. This is a type of sequence that at higher than normal concentrations induces schizophrenia. We got the sequence:

- GSRDCLPPAACFRRRLARRPGYMRSSTGPGIGFLSPAVGTLFRPGGVSGEESHHSESRARQCGLDSRGLLVRS
PVSKSAAAPTVSRGTSAHFGIQLRGTRLPDRLSWPCPGSAGWQQFAAMDSSETLDASWEAACSDGAR
RVRAAGSLPSAELSSNCSPCGPEVPPTPPGSHSAFTSSFSFIRSLSGSAGERGEAEGCPPSREAESHQCSPQE
MGAKAASLDGPHEDPRCLSRPFSLLATRVSADLAQAARNSSRPERDMHSLPDMDPGSSSSLDPLSLAGCGGDGS
SGSGDAHSWDTLLRKWEPLVRDCLLRNRRQMEVISLRLKLQKLQEDAVENTDDYDK

Pairwise alignment & PID:

```
Pretty print first alignment:
MPGGGPGQAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPGYMRSSGTGPGLFLSPAVGTLFRFPGGVSGEESHHSSESRARQCGLSDRSGLLVRSPVSKSAAAPTVTSVRGTSAA
FGIQLRGRTPLPDRSLPQGSAGWQFEEAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPCGPEVPPTPPGVSGHSAFTSFISFIERTLSLGASAGEREAEGCPTT
AESHCSPQEMGAAKASLDGPHEDPRCLSLRPLSLLATRVSADLAQARNRSSPERDHMSLPDMPGCSSSLDPLSACGGDGSGSSGDAHSWDTLLRKWEVPLDRCLLNRRQMEVI
SLRLKLQLKLOQDAEVDENDYDKAFTLQLRLEDEQEKKISHLQPLSPQALPSLFLCHLAAQOVALRRGATQOASGDDHTPLTRMLPRLPLETAQDLSVSTTWRDLWLQEOKQLOKE
IEALQARMFVLEAKDQQLRREAQEQLQWQGCDLPTLVLQGSLLSQGLQEVFSKVALQDTLASAGQIFPHEAPEPTIRSLQPERIKSNLJSLEITTKVCMSEFKCSTLRKVNJLETQ
PALLEAKMHAISGNHFWTAKDLTEERISLTSEREGLEGLLSKLVLVLSRNVKLKSVEKDYNLRLRKEVHEHQETAYETSVKENTMKYMETLKNLKLCSCPKLGLWVWEADLECARLII
QLSLOQLEARGSLSVEDERQMDDLEGAQPPIPRLHSEDKDRTPLKVLEEWKTHLIPSLSHCAGGEQEKEYSIIISLAEGLKECDEDIGKLLYLLEDQLHTAIHSHDEDLQLSRLRELQMVK
ETLQMLIOLOPAOEAKERFAAACSATGTVFHQAOA
```

```

-----GSRDCLPPAACFRRRLARRPGYMRSSTGPGLSPAVGTLFRFPGGVSGEESHHSRARQCLDSRGLLVRSPVSKSAAAPTVTSVRGTSAH
FGIQLRGGTRLPDRLSWPAGPCPGSAGWQQFAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPCGPEVPPTPPGSHSAFTSSFSIRL
LSGSAGERGEAEGCPSPREAESHCQSP0EMGAKAAASLDGPHEPDPRCLSRPFSLLATVRSAIDLQAQARNSSRPERDMHSLPDMDPGSSSSLPDSLACGGDGS
SSGSGDAHSDWTLLRKWEPVLRDCLLRRRQM-----
-----E-----V-----I-----S-----
-----L-----R-----L-----K-----LQ-----K-----L-----Q-----
-----L-----K-----ED-----A-----V-----E-----N-----
-----D-----D-----D-----Y-----D-----
-----K-----
Score=350

```

Alignment 0 PID:

1.0

Second, we analyze the changes in the protein structure due to the commented mutation. In this case, we obtained that it lacks the C-terminal, c-terminal missing

Short transcripts may be associated with schizophrenia in a dominant-negative fashion, because the putative truncated proteins lack the C terminus predicted to mediate protein-protein interactions with a large number of proteins, including nuclear distribution E-like (NDEL1, NUDEL), lissencephaly 1 (LIS1), and elongation protein zeta-1 (FEZ1). Indeed, there is ample evidence from cell and animal models that overexpression of C-terminally truncated DISC1 leads to aberrant phenotypes. Transfection of cells with truncated DISC1 isoforms results in aberrant intracellular localization of DISC1, disruption of interactions with other proteins, including NDEL1, LIS1, dyactin, dynein, and PDE4B, aberrant mitochondrial function, disrupted microtubular network, and inhibition of neurite outgrowth. Mice, in which C-terminally truncated forms of DISC1 were overexpressed on a background of endogenous mouse DISC1, show brain morphological changes, cytoarchitectural anomalies in the hippocampus and cortex, diminished neurite outgrowth, and various behavioral abnormalities reminiscent of aspects of human mental illness.

4th mutation:

Looking at some **natural variants** for DISC1 in the NCBI, we found an interesting mutation associated with susceptibility to schizoaffective disorder which is **L → F** (p.Leu607Phe). We extract the sequence:

- MPGGGPQGAPAAAGGGGVSHAGSRDCLPPAACFRRRLARRPGYMRSSSTGPGIGFLSPAVGTLFRPGGVSG
EESHHSERSRARQCGLDSRGLLVRSPVSKSAAAPTVTSGRTSAHFGIQLRGTRLPDRLSWPCPGSAGWQQE
FAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPCGPEVPPTPPGSHTSAFTSSFSFIRLSLGSAG
ERGEAEGCPPSREAESHQSPQEMGAKAASLDGPHEDPRCLSRRPSLLATRVSADLAQAARNSSRPERDMHSL
PDMDPGSSSSLPSLAGCGGDGSSSGDAHSWDTLLRKWEPVLRDCLLRNRRQMEVISRLKLQKLQEDA
DDYDKAETLQRLEDLEQEKEISLHFQLPSRQPALSSFLGHЛАAQVQAALRRGATQQASGDDTHPLRMEPRLLEP
TAQDSLHVSITRRDWLLQEKKQLQKEIEALQARMFVLEAKDQQLRREIEEQEQQLQWQGCDLTLVQQLSLGQL
QEVSQALQDTLASAGQIPFHAEPPETIRSLQERIKSLNLSLKEITTKVCMSEKFCSTLRKKVNDIETQLPAL
AISGNHFWTAKD**F**TEEIRSLTSEREGLEGLLSKLLVSSRNVKLGSKEDYNRLRREVEHQETAYETSVKENTMKY
METLKNKLCSCCKCPLLGVWEADLEACRLLIQSQLQEARGLSVEDERQMDDLEGAAPPRLHSEDKRKTP
KVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGKKLLYLEDQLHTAIHSHDEDLIQLSRRELQMVKETLQA
MILQLQPAKEAGEREAAASCMTAGVHEAQA

Pairwise alignment & PID:

```
Pretty print first alignment:
MPGGGPQGAPAAAGGGGVSHRAGSRDCLPPAACFRRRLARRPGYMRSSSTGPFIGFLSPAVGTLFRFPGVSGEESHSESARQCGLDSRGLLVRSPVSKSAAAPTVTSVRGTS
FGIQLRGCTRLPDRLSWPCPGSGAGWQEQFAAMDSSETLDASWEAACSDGARRVAAGSLPSAELSNSCSPCGPEVPPTPGSHSAFTSFSFIRSLSGSAGERGEAEGC
PSPREAESHCQSPQEMGAKAASLDGPHEPDPRCLSRPFSSLATRVSDALQAQARNSSPERDMHSLPDMDPGSSSLDPSLAGCGDGSSGSDAHSDTLLRKWEPVLRDCL
LNRRQMEVISLRLKLQLQEDAVENDYDKAETLQQRLDEQEKEKISLHQLPSQPALSSFLGHIAAQVQALRGCATQASGDDHTPLRMEPRLLPTAQDSLHV
SITTRRDWLLOERQQLQKIEALQARMFVLEAKDQQLRREIEEQEQQQLWQGCCDTPLVQGQLSLQOLQEVS
KALQDTLASAGQIIPFHAEPPETIRSLQERIKSLNLSLKEITTKVCMSEFKCSTLRKKVNDIETQLPALLEAKNHAISGNHFWTAKD-TEEIRS
LTSEREGLEGLLSKLVLVLSRNVKLGSVKEDYNRLRREVEHQETAYTSVKE
NTMVKYMETLKNKLCSCCKCP
LLGVWEADLEACRLLIQSLQLQEARGLS
VEDERQMDLEGAA
PP
I
PPLHS
EDKRKTPLKVLEEWKTHLIPSLHCA
GGEQKEESYILS
AELGKCEDIGKLLYLEDQLHTAIHSHDEDLIQSLRRELQMV
KETLQAMILQLQPAKEAGEREEAACMTAGVHEAQ
Score=853
```

Alignment 0 PID:
0.9988290398126464

In this case, at the variant position 607 of the amino acid change, a Leucine residue (L) is changed to a Phenylalanine (F) one. The change implicated is from a medium size hydrophobic residue to a large size and aromatic one.

This is a **missense mutation** (rs6675281), a genetic alteration in which a single base pair substitution alters the genetic code in a way that produces an amino acid that is different from the usual amino acid at that position. Some missense variants (or mutations) will alter the function of the protein.

We analyse the changes in the protein structure of this mutation to see where the mutation falls, with the percentages provided by SOPMA in addition to the literature about it:

Comparing this results of the percentages with the normal sequence (protein structure of L isoform) we can see that in this mutation we have a 50.47% of alpha helix, identical to the wt sequence. This may mean that the pathogenic trace occurs at the interaction moment with other proteins (Lis-1 or NDEL1) that bind DISC1 C-terminal region rather than at a structural level. As well as other mutations, this would affect neuronal migration.

SOPMA :	
Alpha helix	(Hh) : 431 is 50.47%
Beta helix	(Gg) : 0 is 0.00%
Pi helix	(Ii) : 0 is 0.00%
Beta bridge	(Bb) : 0 is 0.00%
Extended strand	(Ee) : 34 is 3.98%
Beta turn	(Tt) : 29 is 3.40%
Bend region	(Ss) : 0 is 0.00%
Random coil	(Cc) : 360 is 42.15%
Ambiguous states (?)	: 0 is 0.00%
Other states	: 0 is 0.00%

In addition to mutations that affect protein function, any mutation or polymorphism that affects the rate at which the DISC1 gene is transcribed might also be expected to predispose an individual to disease.

Carriers of the F607 allele have been shown to have a reduction in gray matter in the cortex area, compared to homozygous L607, in addition to suffering more severe and representative symptoms of schizophrenia variant F607 is also associated with reduced norepinephrine release, lower levels of the PCM1 protein that interacts with DISC1 in the centrosome, as well as mitochondrial trafficking defects.

5th mutation:

As we have mentioned before, there are certain isoforms of DISC1 that, due to the imbalance in its concentrations, could be a direct cause of the development of schizophrenia. In this case, we believe that it is important to comment that **Isoform 10** is one of the protagonists. We obtained its sequence:

- LQKLQEDAVENDDYDKAETLQQRLEDLEQEKFISLHFQLPSRQPALSSFLGHAAQVQAALRRGATQQASGDDTHTPLRMEPRLLPETAQDSLHVSITRRDWLLQEKKQLQEIEALQARMFVLEAKDQQLRREIEEQEQQLQWQGCDLTPLVGQLSLGQLQEVSALKQDTLASAGQIPFHAEPPETIRSLQERIKSLSNLKEITTCKVMSEKFCSTLRKVKNDIETQLPALLEAKMHAISGNHFWTAKDLTEEIRSLTSEREGLEGLLSKLLVLSRNVKKLSVKEDYNRLRREVEHQETAYETSVKENTMKYMETLKNKLCSCCKPCLLGKVWEADLEACRLLIQSQLQLQEARGSLSVEDERQMDDLEGAAPPIPPRLHSEDKRKTPLKVLEEWKTHLIPSLLCAGGEQEKEYSILSAELEGEKCEDIGKKLLYLEDQLHTAIHSDEDLIQLSLRRELQMVKETLQAMILQLQPAKEAGEREAAASCMTAGVHEAQA

Pairwise alignment & PID:

Alignment 0 PID:

1.0

After analyzing these results, we reached the conclusion that:

- N-terminal missing.
 - 350-356: VISLRLK → LRRYNKD

We could associate the pathogenesis of the Isoform 10 due to the lack of N-terminal. In consequence, if this isoform is expressed in too high concentrations, there will be a lack of interactions with GSK3 β or PDE4, and a lack in neuronal migration and neuronal plasticity and memory.

DISCUSSION

Taking into account all the results, we can assume that they have agreement. It is interesting to note that most of these myths that trigger schizophrenia actually only diverge from very few amino acids, and can trigger a great effect in terms of phenotype.

As we have been able to see and demonstrate in our ongoing work, schizophrenia is a disease caused not by a single specific factor, as it can occur in other diseases, but by a myriad of different causes and mutations. And not only that, since, to this day, the lack of information about it, means that there are still endless discoveries to be found. We have focused on a single protein, DISC1, but even so, it has been shown that there are others such as DISC2, CACNA1, CACNA1C, CACNB2, ATXN7, SETD8, whose relevance is also important in the development of said behavioral disease, all and that , there is still a long way to go into the research of other triggers of it.

All mutations affect the pathway, generating some type of modification that will ultimately affect the interaction with gsk3b, responsible for neuronal development.

Another thing to keep in mind is the fact that, during our review, we have found ourselves in front of contradictory papers, which determined different exons when acting as causative agents of schizophrenia. With our worked results, we have been able to reach a coherent conclusion explained in the next point.

CONCLUSION

We have looked at the major mutations in the DISC1 protein that can trigger schizophrenia, but there are thousands that could, less frequently. With the analyzes that we have carried out we can affirm that schizophrenia can occur due to purely genetic mutations. We believe that it is very important to clarify that just as genetic mutations are a precursor of this disease, but... it must also be taken into account that environmental factors, outside of genetics, are very important when developing a mental illness such as schizophrenia.

From the methods we used to understand the different isoforms or mutations for the gene DISC1 we can finally conclude that the N-terminal region (codified by exon 2) is the most pathogenic one because it is the more conserved region comparing through other species like primates and as we know, the more conserved, the more restricted function causing pathogenesis. Although, mutations in the C-terminal are also likely to cause this disease. So, both regions mutations are susceptible to cause schizophrenia, suggesting that probably the N-terminal has more severe symptoms and this is why it is the more mutated one.

Despite this, in our work, we have decided to represent mutations in both terminal regions to corroborate that it can occur due to many types of alterations and not especially due to exon 2, as stated by several papers used.

Findings confirm earlier conclusions that deficits in family communication processes are especially reliable correlates of schizophrenia and high risk for schizophrenia in offspring. This finding is particularly true for sons, although comparisons by sex of offspring are rare. Findings also corroborate that genetic factors trigger disease in a more prevailing way, unlike environmental factors.

To conclude, we have demonstrated that Schizophrenia is a human exclusive disease, despite the fact that at the protein level, it is similar to many other species; this suggests that the development of this behavioral disease has been carried out thanks to evolution, and as a consequence of developing essential behaviors, such as exercising speech.

REFERENCES

Papers of interest from which we have collected information about the pathology:

<https://pubmed.ncbi.nlm.nih.gov/14962739/>

[genes-environment-and-behavior](#)

http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0016-38132006000300014

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-59#Abs1>

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-290>

<https://onlinelibrary.wiley.com/doi/full/10.1111/j.1601-5223.2009.02138.x>

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-3508-8>

<https://annals-general-psychiatry.biomedcentral.com/articles/10.1186/1744-859X-5-S1-S77>

<https://annals-general-psychiatry.biomedcentral.com/articles/10.1186/1744-859X-7-S1-S89>

<https://pubmed.ncbi.nlm.nih.gov/17336946/>

<https://aacijournal.biomedcentral.com/articles/10.1186/1710-1492-4-1-37>

<https://healthitanalytics.com/news/whole-genome-sequencing-reveals-genetic-risks-of-schizophrenia>

<https://news.unchealthcare.org/2020/04/whole-genome-sequencing-reveals-genetic-structural-secrets-of-schizophrenia/>

<https://www.nih.gov/news-events/nih-research-matters/common-genetic-factors-found-5-mental-disorders>

https://www.ncbi.nlm.nih.gov/nuccore/NC_000001.11?report=fasta&from=231626790&to=23204127

<https://www.ncbi.nlm.nih.gov/search/>

https://www.ncbi.nlm.nih.gov/datasets/tables/genes/?table_type=transcripts&key=761a12aaa786d09cb101c80c8ac13f9f

<https://www.uniprot.org/blast/uniprot/B2021041663E7E78CFC6242B71761763234FF46DC0033125>

<https://www.nature.com/articles/4002079>

<https://www.scientificamerican.com/article/why-don-t-animals-get-schizophrenia-and-how-come-we-do/>

Evolutionary behavioral genetics = <nihms638687.pdf>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4288764/>

Tool to induce artificial mutations on DISC 1 protein sequence :

(https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl)

<https://pubs.acs.org/doi/full/10.1021/cn200062k>

Schizophrenia database : <http://www.szdb.org/SZDB/cnv1.php>

Ensembl phylogenetic trees :

http://www.ensembl.org/Homo_sapiens/Gene/Compara_Ortholog?db=core;g=ENSG00000162946;r=1:231626790-232041272

http://www.ensembl.org/Homo_sapiens/Gene/SpeciesTree?db=core;g=ENSG00000162946;r=1:231626790-232041272

http://www.ensembl.org/Homo_sapiens/Location/View?db=core;g=ENSG00000162946;r=1:231626790-232041272