Universitat de les Illes Balears

BACHELOR'S THESIS

VALIDATION PHASE OF COPY NUMBER VARIANTS (CNVs) ASSOCIATED WITH MAJOR MENTAL DISORDER IDENTIFIED IN FAMILIES IN THE NORTH OF MALLORCA IN ORDER TO DEMONSTRATE THEIR PATHOGENICITY AND THEIR POTENTIAL FOUNDING EFFECT

Alessandro Palou Kötters

Degree in Biology

Faculty of Science

Academic Year 2020-21

VALIDATION PHASE OF COPY NUMBER VARIANTS (CNVs) ASSOCIATED WITH MAJOR MENTAL DISORDER IDENTIFIED IN FAMILIES IN THE NORTH OF MALLORCA IN ORDER TO DEMONSTRATE THEIR PATHOGENICITY AND THEIR POTENTIAL FOUNDING EFFECT

Alessandro Palou Kötters

Bachelor's Thesis

Faculty of Science

University of the Balearic Islands

Academic Year 2020-21

Key words:

Major mental disorder, psychosis, schizophrenia, bipolar disorder, complex diseases, heritability, rare genetic variants, CNVs, SNPs, ddPCR, founder effect

Thesis Supervisor's Name: Cristòfol Vives Bauzà

Acknowledgements

Primer de tot, vull agrair al Dr. Cristòfol Vives per haver-me donat l'oportunitat de fer el Treball de Fi de Grau en el seu equip d'investigació de Neurobiologia a l'Hospital Universitari de Son Espases, així com per tots els coneixements que m'ha transmès, per la seva ajuda i suport que m'ha donat sempre que ho necessitava i per la paciència que ha tingut en mi. També vull donar les gràcies a la Jessica, Aina, Joan, Mónica, Aarne i a tots els altres membres de la Unitat d'Investigació que en un moment o altre m'han ajudat o dels que he tingut l'oportunitat d'aprendre alguna cosa durant la meva estada.

Abstract

Psychotic major mental disorders – schizophrenia, bipolar disorder and schizoaffective disorder – are heritable diseases caused by a complex interplay between genetic and environmental factors and have a pathology consisting, from a genetic point of view, of the combined effect of common and rare variants. In this work, we have checked previously identified rare copy number variants (CNVs), Insertions and Deletions (INDELS) and single nucleotide variants (SNVs) in new recruited patients from the same geographical regions of the two originally previous studied families with high prevalence (Family 1 and Family 2) in order to replicate results that allow us to demonstrate their pathogenicity and to prove founder effects. SNVs and INDELS genotyping were performed by PCR amplification of DNA extracted from peripheral blood, followed by DNA electrophoresis, PCR purification and Sanger sequencing. CNVs were analyzed by droplet digital PCR. As our results show evidence of the presence of the genotyped CNVs in subjects from the same geographical regions as the previous studied families, we conclude that these CNVs are pathogenic and that they have originated a founder effect.

Index

INTRODUCTION

1. Major mental disorders

A major mental disorder is defined as a behavioural, emotional or mental illness that causes severe functional disability, limits or significantly interferes with one or more major life activities (NIMH, National Institute of Mental Health). They are usually characterized by a significant disturbance in terms of behaviour, regulation of emotions and / or cognition in an individual at a clinical level, reflecting a dysfunction in the biological, psychological and / or developmental processes underlying mental functioning[1].

Some major mental disorders, especially schizophrenia and bipolar disorder, include or can include the presence of psychosis, which is known as a domain of symptoms marked by a loss of contact with reality such as delusions (false beliefs) or hallucinations (sensory experiences created by the mind and not based on real stimuli)[1,2]. While an essential characteristic of schizophrenia is the presence of psychotic episodes, bipolar disorder is characterized by disruptive mood swings (from depression to mania) although psychotic symptoms are significant too, particularly in subtype I [3–7].

Since Kraepelin – *The Kraepelinian dichotomy* –, schizophrenia and bipolar disorder have been considered two separate entities in clinical classifications[8–10]. Although, it must be mentioned their substantial phenotypic overlap, as clinical distinction between both disorders may be so confusing that in some cases an initial diagnosis of schizophrenia has ended up with a definitive diagnosis of bipolar disorder or vice versa[11–13]. Due to these unclear borders, Dr. Jacob Kasanin proposed a “mixed” diagnosis category: the schizoaffective disorder, defined as a disorder between bipolar disorder and schizophrenia that requires the presence of psychotic symptoms as well as mood symptoms[14–17].

2. Mental disorders genetics

Genetic research of schizophrenia, schizoaffective disorder and bipolar disorder has tried to answer if their phenotypic overlap translates into a common genetic susceptibility. Evidence collected so far seems to confirm that these disorders share at least part of their genetic susceptibility, and the psychosis domain may play an important role in it[2,7,18–23].

Psychotic disorders are considered complex diseases because of its important genetic contribution and as the inheritance pattern is the consequence of a complex interaction between genetic and environmental factors[24–26].

Genetic contribution is normally quantified by estimating Heritability, which is the phenotypic variation rate in a disease or trait defined by inherited genetic variants[27,28]. Thus, it enables to determine to which extent a disease or trait can be associated with genes[27]. It ranges between 0 (genetics explain a low proportion

of the phenotypic variability) and 1 (genetics explain a high proportion of the phenotypic variability). There are 3 types of heritability: broad heritability (H^2), narrow heritability (h^2) and SNP-based heritability (h^2_{SNP}). While H^2 is described as the quotient between the variance of the trait explained by genetic factors and the total variance of the trait in the population, h^2 divides the variance due to genetic factors into additive, dominant/recessive and gene-environment interactions effects and focuses on additive genetics effects. SNP-based heritability, a subtype of h^2, assesses SNP correlation among unrelated cases and compares it with SNP correlation among unrelated controls. It is calculated directly from genetic data (genotyped SNPs from SNP array)[2,29].

Regarding schizophrenia and bipolar disorder, they are highly heritable disorders, since their heritability has been estimated in multiple studies and the principal conclusion has remained the same. As for schizophrenia, studies point to a Heritability about 80% [30], whereas for bipolar disorder it is around 70% [31]. In both cases, environmental factors contribute to their disease risk[32].

Several factors define the genetic architecture of a complex disorder, describing the relationship between the phenotype and the genotype: the number of risk genetic variants that contribute to the disease, their frequency in the population (common and rare variants), the probability of being affected by the disease given a certain genotype and how the variants interact with each other[25,33,34].

There are two main hypotheses to understand the genetic architecture of complex diseases: the 'Common Disease, Common Variant (CDCV)' hypothesis and the 'Common Disease, Rare Variant (CDRV)' hypothesis[35–37]. The CDCV hypothesis argues that the genetic susceptibility to the disease is the consequence of the combined or cumulative effect of many genetic variations with noticeable frequency, but low 'penetrance' (effect size or the probability that a carrier of the genetic variation will express the disease). As carrying just one or some of these variants with small or modest effect is not enough to develop the disease, it appears when a determinate burden threshold of risk variants is reached or surpassed[38–40]. Alternatively, the CDRV hypothesis argues that the genetic susceptibility to the disease is the result of rare and gene-disrupting DNA sequence variations of recent origin with high penetrance[37,41,42], in which this thesis is focused on.

In both hypotheses, environmental factors are supposed to also play a relevant role[37]. While CDCV has been tested in genome-wide association studies (GWAS)[38,43] and recently with their derived polygenic risk scores (PRS), CDRV was originally implicit in first linkage analyses and has been lately busted by copy number variant (CNV) studies[44,45] and whole-exome/genome-based studies[46–48]. A copy number variation is when the number of copies of a concrete gene differs from one individual to the next, according to the National Human Genome Research Institute (NHGRI). It is a type of structural variation with a variable number of copies of a specific DNA segment longer than 1 kilobase (Kb) and

shorter than 3-5 megabases (Mb)[2]. CNVs may play a role in human pathology via distinct mechanisms. They can influence protein expression multiplying or removing a particular gene or its regulatory factors[49]. Genomic CNV studies have established a role for rare variations in the aetiology of schizophrenia[50]. Evidence confirm that schizophrenia patients have a higher burden of rare and large CNVs (frequency <1% and size >100 kb) in comparison to controls and a higher frequency of *de novo* mutations CNVs[50–54]. Today, the contribution of rare CNVs to complex diseases like schizophrenia is well known[52]. Early CNV studies results propose that rare variations also play a role in bipolar disorder[55–57].

3. Studied Familes

Family-based approaches are a method to analyse rare variants (CNVs, SNPs and INDELs)[50,58–61]. These studies are focused on non-synonymous variants, including frame-shift variants (single point deletion or insertion resulting in a frame-shift), nonsense variants (point mutations which result in a stop codon or altering a splicing site) and missense variants (point mutations resulting in amino acid substitution)[2].

The main idea of family-based approaches relies on the identification and analysis of large extended pedigrees with a high prevalence of the disorder to be studied in order to identify rare highly penetrant variants co-segregating with the disease, which are rare in the population but enriched in families[2,62–64].

Previously studied families in the thesis “Disentangling the genomic architecture of psychosis in families with high prevalence through system genomic approaches” are the following:

Family 1 (See Figure 1A): The family comes originally from two small towns in Mallorca. The proband (SZ3) is a woman affected by schizoaffective disorder, characterized by changing between manic and depressive events together with delusional ideation of paranoid type. She is a mother of 5 children. The symptoms began to arise in 1985 after the birth of her twin daughters. The proband has three sons: SZ7 started to present schizophrenia when he was 19 years old with prevalent negative symptomatology; SZ8 onset of psychotic symptoms at the age of 16 and diagnosed with paranoid schizophrenia. He died in 2014 by suicide. And SZ9 evidence anxiety, behavioural disorders, and somatic symptoms from the age of 14, later diagnosed as schizoaffective disorder. The twin daughters (subjects 10 and 11) do not show any psychiatric symptoms. The proband’s family history includes an alcoholic brother who died of cirrhosis (subject 304), an uncle with mental illness and epilepsy (subject 206), another uncle with psychosis and autism (subject 203), and a cousin with anxiety (subject 6) who, in turn, has a daughter affected by bipolar disorder subtype I (subject BD12) whose father is diagnosed with major depressive disorder (subject 5).

Family 2 (See Figure 1B): The second family is hugely multigenerational and has a high prevalence of mental disorder. It comes from “Las Alpujarras” (Andalusia), a schizophrenia hotspot[65]. Part of the family resides in the Balearic Islands and their family members were recruited by the psychiatry department of the University Hospital Son Espases (HUSE), in Mallorca. Through an informative member of the family, the other individuals of the family were contacted by psychiatrists. A total of 34 individuals were recruited and participated in the study (9 psychotic patients, 10 non-psychotic mental disorder patients, 14 healthy individuals and 1 with an unknown diagnosis).

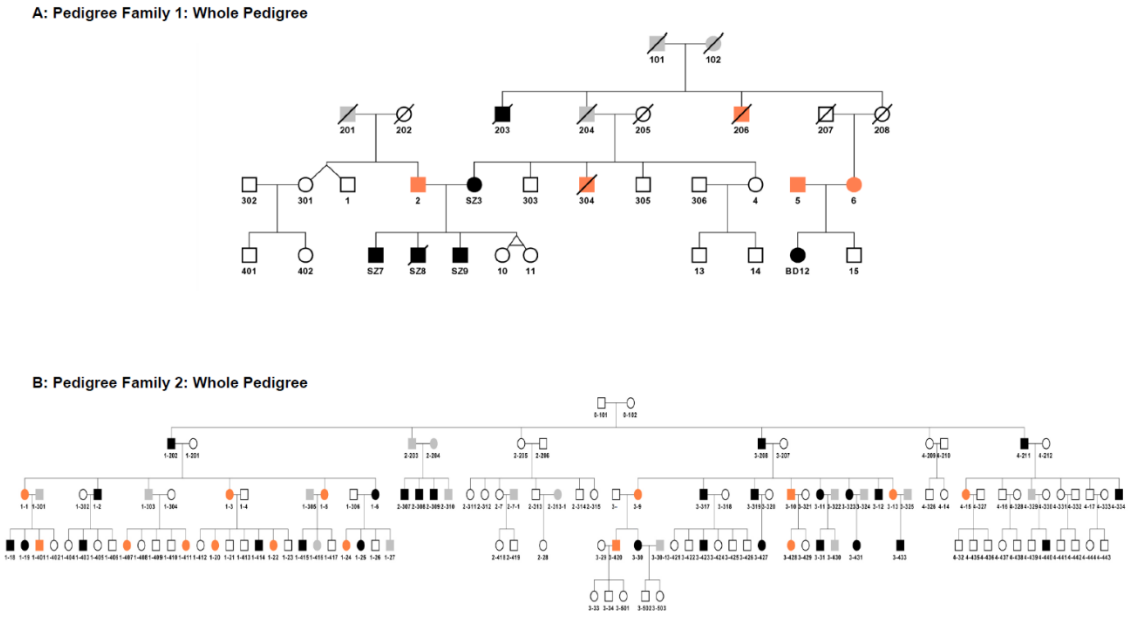

Figure 1. Pedigrees Family 1 (A) and Family 2 (B).

4. Preliminary data

CNV analyses identified a total of 14 CNVs between both families. In Family 1, 10 subjects had at least a CNV (5 narrow phenotype patients, 2 wide phenotype patients and 3 healthy subjects) whereas in Family 2, 19 subjects had at least a CNV (7 narrow phenotype patients, 7 wide phenotype patients, 4 healthy subjects and 1 undefined phenotype subject). As a result, the percentage of subjects that carried at least a CNV affecting a gene was 85.71% for the narrow phenotype, 69.23% for the wide phenotype and 33% for healthy individuals. This corroborates the published evidence that psychotic individuals carry more rare CNVs affecting genes than non-psychotic or healthy individuals[44,45,52,66].

In Family 1, two rare duplications (DUP3p26.3 and DUP16p23.3) were identified.

DUP3p26.3, located on chromosome 3 (from 1159787 to 1781739 bp), affects the *CNTN6* gene which encodes the protein contactin-6. CNVs affecting this gene have been widely identified in subjects with neuropsychiatric disorders, including schizophrenia and bipolar disorder among others, indicating an important role for this type of disorders as well as for neurodevelopmental disorders[67–74]. Besides, the penetrance of this duplication in Family 1 is incomplete and variable since some healthy subjects are carriers. Similar duplications have described at this genomic location showing incomplete penetrance [67,72,74], which may also suggest that CNVs affecting the *CNTN6* gene could be neutral unless they are combined with other predisposing variations[75].

DUP16p23.3 is located on chromosome 16 (from 82180075 to 83664582 bp) containing the cadherin-13 (*CDH13*) gene. CDH13 expression is implicated in inhibitory and excitatory synapses in the hippocampus. It seems to have a relevant role in the equilibrium between excitation and inhibition, so dysfunctions in it may drive to an excitation/inhibition imbalance[76,77]. Cadherins, CDH13's family, have been related with many neuropsychiatric disorders, schizophrenia and bipolar disorder among them[78–82].

In Family 2, three rare SNP/INDEL variants were identified in different genomic regions.

In narrow regions has been identified the variant g. 99006061 G>A (p. H321Y), which is the only rare variant shared among all psychotic subjects, also present in almost all non-psychotic subjects plus in a healthy subject. This variant affects *ARHGAP19*, a gene located on chromosome 10 that encodes a member of the Rho GTPase-Activating protein (RhoGAP) family. Being a RhoGAP, it is a gene with probabilities to be related with schizophrenia, as some RhoGAPs have been associated with it[83–87]. Additionally, *ARHGAP19* is targeted by miR-24, a micro-RNA upregulated throughout neuronal differentiation[88,89], which has been also related with altered expression in the prefrontal cortex of schizophrenia subjects[90].

In non-filtered regions was identified a variant, g. 20230301 G>A (p. R119W), previously associated with schizophrenia[91], which affects *RTN4R*, a gene with solid evidence of involvement in schizophrenia[91–93]. *RTN4R* is located in the region 22q11 and so, it is affected among other genes by the DEL22q11, the main genetic risk factor for schizophrenia[94,95].

Finally, in non-coding regions, a rare small 2 bp deletion was detected in chromosome 9, affecting the gene *LINC00474* (chr9: 118667077) in all non-psychotic subjects.

HYPOTHESIS AND OBJECTIVES

Hypothesis

As linkage and family-based association studies have shown a relation of these CNVs and SNPs with major mental disorders, it is expected to find them in subjects from the same geographical area.

Objectives

The objective is to genotype these CNVs and SNPs in patients from the same geographical region in order to:

1. Demonstrate that they are pathogenic.
2. Demonstrate a founder effect.

METHODOLOGY

1. DNA extraction from peripheral blood

No more than one or two days after the blood collection, DNA was extracted using the DNA Blood Extraction Kit (5Prime®) according to the manufacturer's protocol. After extraction DNA samples were quantified using nanodrop and stored in a -20 ºC freezer.

The following sections were carried out for subjects from the same geographical region as the Family 2 patients.

2. DNA amplification by PCR

50 ng/µl of DNA were used per reaction. Temperature gradients were performed to set up the melting temperature for each pair of primers. PCR reaction was performed in a total volume of 20 µl containing 2 µl buffer 5X, 1.5 µl $MgCl_2$ (25mM), 2 µl dNTPs (2,5mM), 2 µl of each primer (5 µM) and 0.25 µl (5u/µl) of Go Taq® G2 Flexi DNA polymerase. Thermocycler conditions were as follows: Initial denaturing at 95 ºC for 5 minutes followed by 35 cycles of denaturing at 95 ºC for 30 seconds, annealing for 30 seconds at 55 ºC for *ARHGAP19* and 57 ºC for *LINC00494* and extension at 72 ºC for 90 seconds, and a final extension time for 5 minutes at 72 ºC and 4 ºC to finish the process.

3. DNA electrophoresis in agarose gels

PCR products were checked running 3 µl of each sample in a gel made of 2% agarose and TAE buffer (40 mM Tris-acetate, 1 mM EDTA pH 8.0).

4. DNA amplified purification

The resting 17 µl were purified using the Kit DNA Clean & Concentrator™-5 (Zymo Research), a spin column-based DNA purification method.

5. Sanger sequencing

Samples were sent to the biotechnology company *Macrogen* to be sequenced by the Sanger chain-termination method, consisting of using a mix of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) during amplification. The ddNTPs prevent the formation of the phosphodiester bond as they lack a 3'-hydroxyl group, terminating the synthesized chain. Each of the four ddNTPs (ddGTP, ddATP, ddTTP and ddCTP) is labelled with a different fluorescent dye, each of which emit light at different wavelengths.

Droplet Digital PCR (ddPCR) was used for subjects from the same geographical region as the Family 1 patients.

6. ddPCR

Digital PCR was performed using the QX200 Droplet Digital PCR System (Bio-Rad).

Before ddPCR, gDNA of each subject was digested for 1 hour at 37 ºC with the restriction enzyme Hind III (1000 ng gDNA, 1 µl Hind III HF, 2 µl Cut Smart Buffer, and H_2O in a final volume of 20 µl). PCR master mixes for each subject being analyzed were then prepared (ddPCR™ Supermix for Probes (No dUTP) 1X, Primer Target Forward (900 nM), Primer Target Reverse (900 nM), Primer Reference Forward (900 nM), Primer Reference Reverse (900 nM), FAM probe (Target) and VIC probe (Reference) (5 mM) and H_2O in a final volume of 20 µl). Afterward, they were partitioned into drops using the Droplet Generation Oil for Probes (Bio-Rad) and the QX200 AutoDG Droplet Digital PCR System (Bio-Rad). The plate was sealed using the PX1 PCR Plate Sealer (Bio-Rad). The PCR was performed in a C1000 touch thermocycler (Bio-Rad) with the following cycle conditions: Initial denaturalization at 95 ºC for 10 minutes, followed by 39 cycles at 94 ºC for 30 seconds and an extension at 57.1 ºC for 1 minute, with a final denaturing at 98 ºC for 10 minutes. Droplet analysis was performed using the Q200 Droplet Reader (Bio-Rad).

CNVs were calculated using the QuantaSoft software (Bio-Rad), dividing the number of copies of the target by the number of copies of the reference and multiplying it by 2 (the number of copies of the reference locus in the genome).

RESULTS

1. Primers sets optimization

First, optimal annealing temperatures of the pair of primers used to amplify the SNPs studied were determined.

The annealing temperature for rs145032100 (*ARHGAP19*) (55 ºC) was already determined within the context of the PhD thesis: “Disentangling the genomic architecture of psychosis in families with high prevalence through system genomic approaches”, on which the present study is based.

For rs74315508 (*RTN4R*) a PCR was done together with rs145032100 (*ARHGAP19*) (Figure 2).

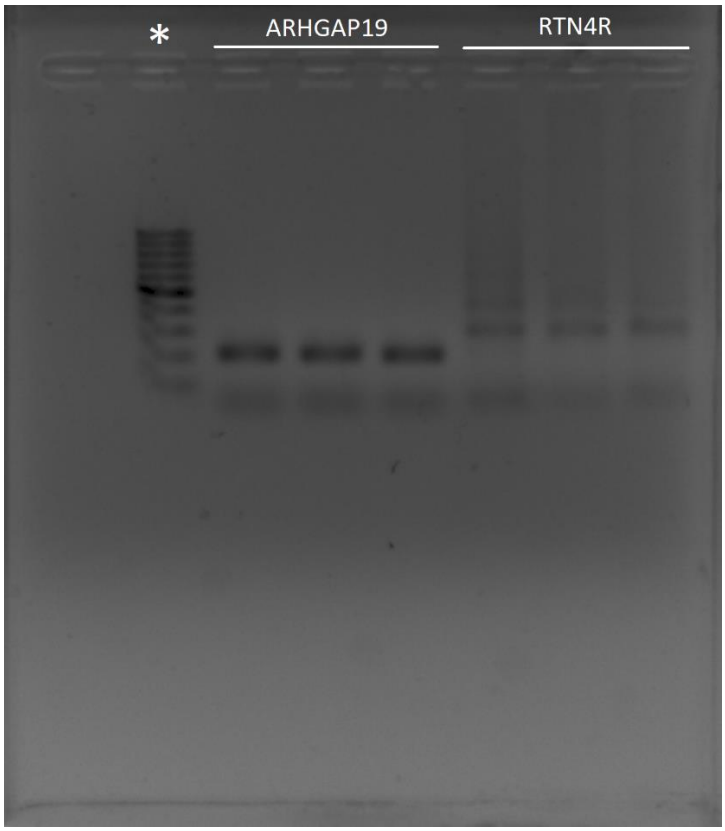

Figure 2. Agarose gel for *ARHGAP19* and *RTN4R*. * = ladder.

As figure 2 shows, rs145032100 (*ARHGAP19*) did amplify well while rs74315508 (*RTN4R*) amplification was too weak to carry on with the same conditions for it. Therefore, and as a test, it was decided to change the double-distilled water in the mix for betaine 0.5% DMSO and annealing temperature from 55 ºC to 57 ºC (Figure 3).

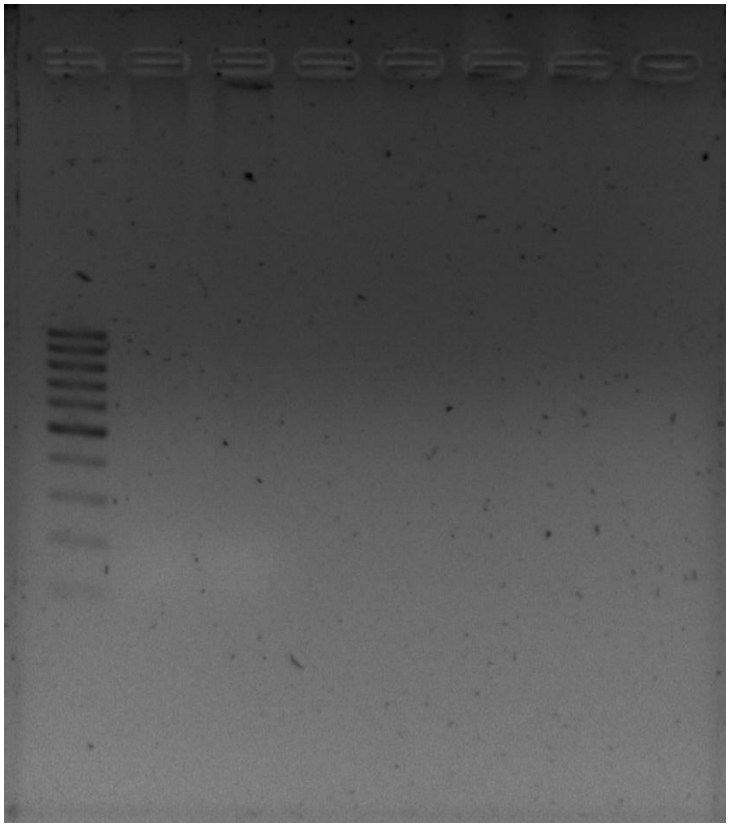

Figure 3. Agarose gel for *RTN4R* with betaine 0.5% DMSO and 57 ºC as annealing temperature. Control DNA has been used as it was a conditions optimization test.

Unfortunately, and as seen in figure 3, the introduced changes did not work. An attempt has been made to amplify the region of interest of rs74315508 (*RTN4R*), rich in C-G content, but given the negative results, it has been decided to change the primers and the corresponding work will be done in the context of another study.

Whereas for *LINC00474* a PCR was made with a temperature gradient (57 ºC, 59 ºC, 60 ºC, 62 ºC), showing 57 ºC as the most optimal annealing temperature (Figure 4).

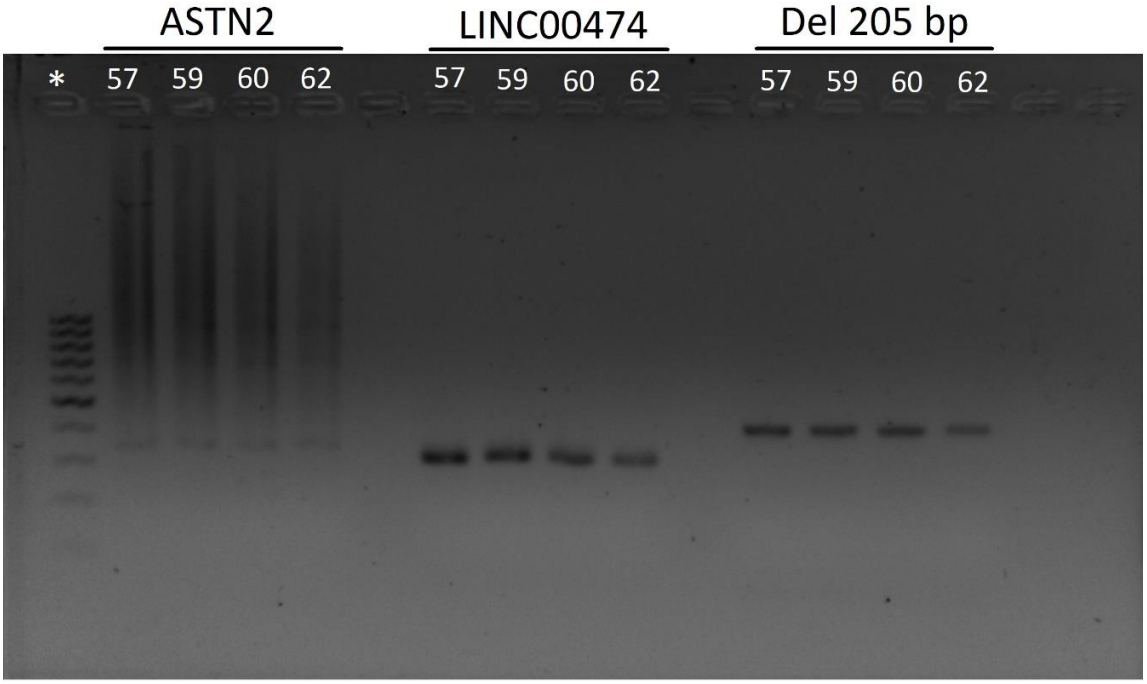

Figure 4. Agarose electrophoresis showing the results obtained after performing the gradient temperature for *ASTN2*, *LINC00474* and Del 205 bp. Samples were loaded as follows from left to right: DNA ladder, *ASTN2* (57 ºC, 59 ºC, 60 ºC, 62 ºC), *LINC00474* (57 ºC, 59 ºC, 60 ºC, 62 ºC) and Del 205 bp (57 ºC, 59 ºC, 60 ºC, 62 ºC). * = ladder.

As shown in figure 4, *ASTN2* did not work at any temperature, whereas the other two amplicons amplified well at all tested temperatures, although *LINC00474* amplified better at 57 ºC and Del 205 bp (Start 116463148 and End 116463353. It affects an intergenic region) at both 57 ºC and 59 ºC. Thus, it was decided to amplify all samples of these two genes at 57 ºC. On the other hand, for *ASTN2*, despite having evidence that the primers amplified at 55 ºC, the conditions had to be optimized again. To do so, double-distilled water in the master mix was replaced by betaine 1% DMSO (Figure 5).

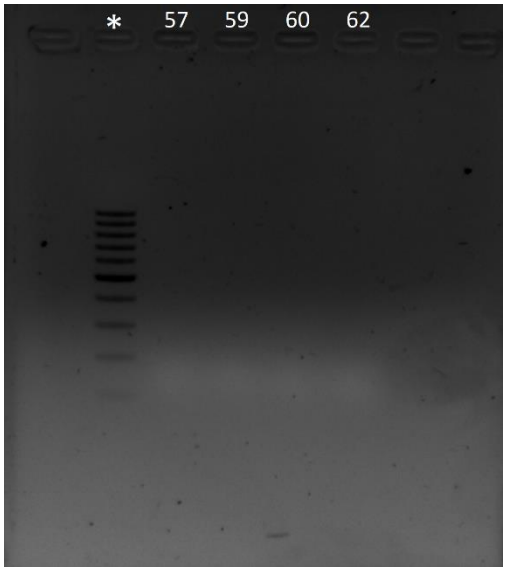

Figure 5. Agarose gel for *ASTN2* with betaine 1% DMSO and temperature gradient (57 ºC, 59 ºC, 60 ºC and 62 ºC).

Given the poor results shown in figure 5, conditions for *ASTN2* still need to be improved to achieve its amplification, work that, as well as for rs74315508 (*RTN4R*), will be done in the context of another study.

2. DNA electrophoresis in agarose gels

Two samples of the amplification checking of the samples to be sequenced are shown below (Figures 6 and 7).

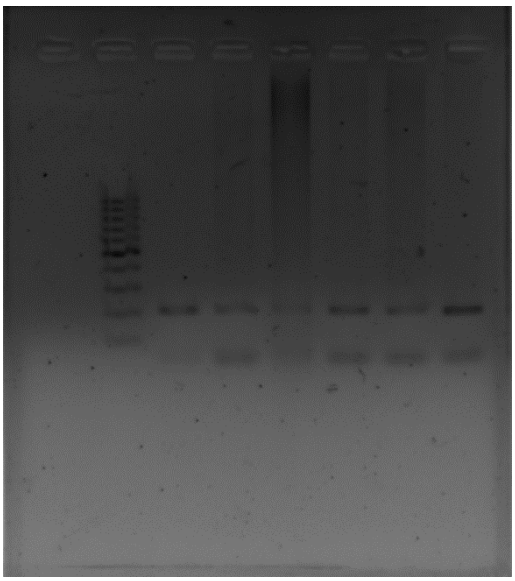

Figure 6. Agarose gel for *ARHGAP19*.

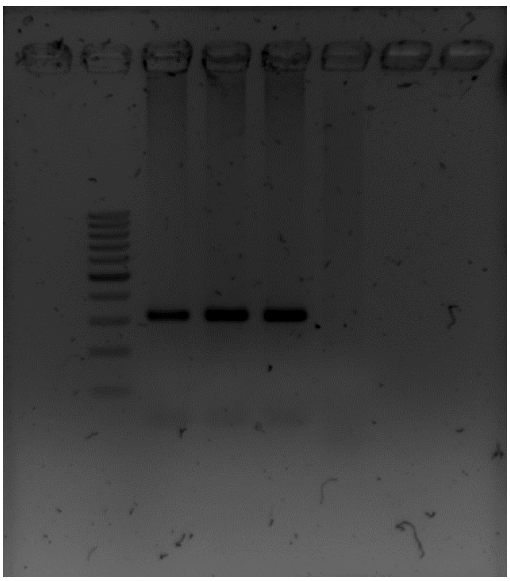

Figure 7. Agarose gel for *LINC00474*.

In total, 27 samples from *ARHGAP19* and 27 from *LINC00474* were amplified and sent to be Sanger sequenced.

Both gels show successful amplification for most samples. Just one sample for each gel did not amplify.

3. Sanger sequencing

PCR products were purified and were sent to the company Macrogen to be sequenced. Unfortunately, the results of the sequencing were not received on time to be included in this thesis and we were not able to conclude whether there is any new patient harboring the mutations tested.

4. Confirmation of non-reported INDELS identified in Family 2

We next checked whether potential novel INDELS, identified through Whole Genome Sequencing (WGS) using the algorithm HaplotypeCaller of GATK (v3.3.0) (https://gatk.broadinstitute.org/), were real. WGS was previously performed in 12 patients from family 2, and two small INDELS were identified in affected members. A deletion of 94 bp in the gene *LPAR* was shared by affected subjects 1-18, 1-19 and 3-13 (pedigree Family 2; Figure 1B). And the second interesting INDEL to validate was a 4 bp insertion in the gene PAPPA, present in subjects 1-24, 1-25, 1-26 and 1-27 of Family 2 (see pedigree Figure 1B).

LPAR1 DEL 94 bp (chr9:113669166-113669260)

The first INDEL analyzed was a 94 bp deletion at chr9. We amplified the genomic region containing the deletion, and the purified PCR products were sequenced by Sanger. As shown in figure 8, we could observe a double electropherogram signal, compatible with the starting point of the deletion at chr9:113669166, coinciding with the coordinates obtained with the algorithm HaplotypeCaller of GATK (v3.3.0), used to identified small INDELS.

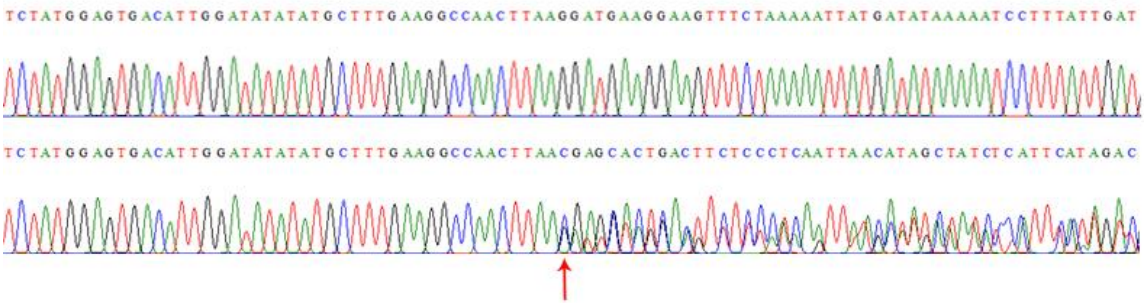

Figure 8. LPAR1 DEL 94 bp (chr9:113669166-113669260). In the 2% agarose gel **(A)**, the different genotypes for this variant for the members of the family are

displayed. In **(B)**, the red arrow indicates the base pair in where the deletion starts, which results in an electropherogram where the two sequences, WT, and DEL overlap.

PAPPA 4 bp INDEL (chr9:119154431)

The second INDEL validated by molecular biology was a 4 bp insertion at *PAPPA* gene (chr9:119154431). In this case, we decided to check whether we could discriminate between WT and inserted sequence by running the product of an amplicon containing the genomic region of interest (Forward 5'-TCGTCTGTGAAGAGAGGA-3' and Reverse 5'-GTTGGTGATGCAGGATTC-3'). For this purpose, the PCR products were run in a 2% agarose gel. Patients 1-24, 1-25, 1-26 and 1-27, harbored the 4 bp insertion, as shown in the electrophoresis (Figure 9a). This INDEL was also validated by Sanger sequence, demonstrating a duplicated electropherogram right at the genomic site predicted by Haplotype Caller (Figure 9b).

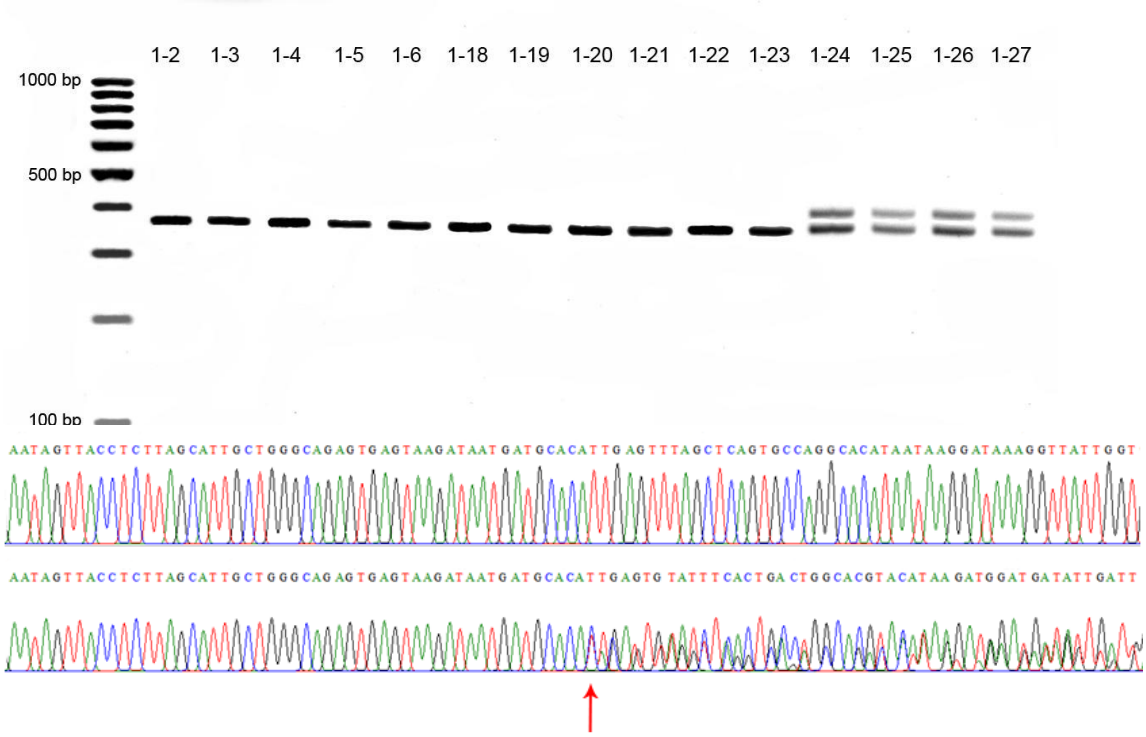

Figure 9. PAPPA 4bp INDEL (chr9:119154431). In the 2% agarose gel **(A)**, the different genotypes for this variant for the members of the family are displayed. In **(B)**, the red arrow indicates the deletion location which results in an electropherogram where the two sequences, WT, and INDEL overlap.

5. Validation of DUP16p23.3 and DUP3p26.3 by ddPCR

By using CNVnator, in Family 1 two duplications were identified, DUP16.p23.3 and DUP3p26.3. To validate those duplications in the different family members and in psychotic patients from the same geographic region, ddPCR analyses were performed. Regarding the presence of DUP16p23.3, containing the *CDH13* (cadherin-13) gene, patients 9, 10, 11 and 12 harbored the DUP, as evidenced for the presence of 3 copies instead of the normal 2 DNA copies (Figure 10). As for the presence of DUP3p26.3, which contains the *CNTN6* (contactin-6) gene, the same patients (9, 10, 11 and 12) harbored the duplication, presenting 3 copies in contrast to the 2 copies of the other subjects (Figure 11).

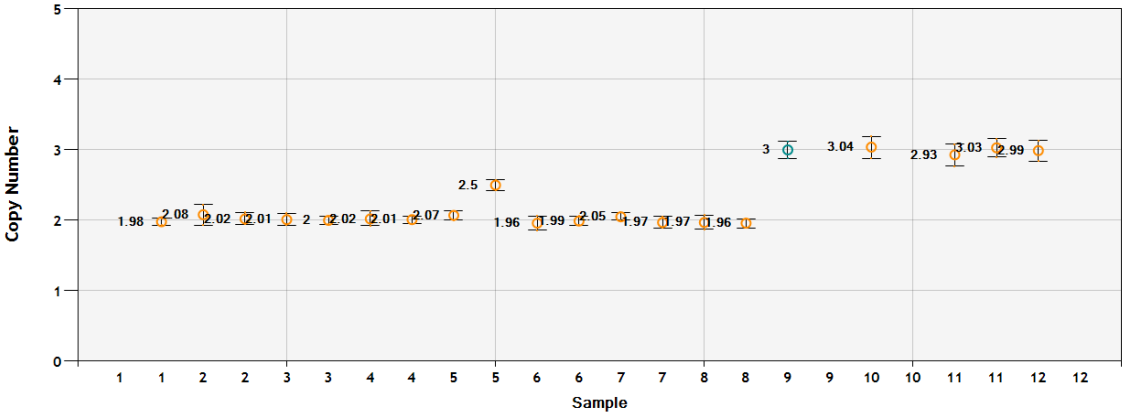

Figure 10. Validation of DUP16p23.3 CNV in Family 1 population subjects.

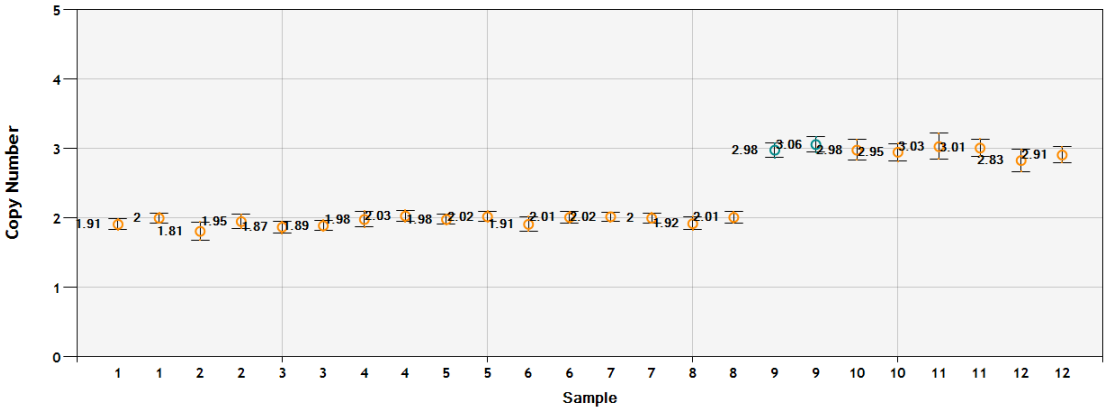

Figure 11. Validation of DUP3p26.3 CNV in Family 1 population subjects.

In both cases, samples 9 and 10 were from Family 1 patients, while the rest of the tested patients were psychotic subjects outside the family but from the same geographic region.

DISCUSSION

Even the study limitations we have had to face, as the little sample size and the lack of Sanger sequencing results, most samples did amplify, and we have purified the amplification products and sent them to be sequenced.

Nevertheless, we have been able to demonstrate the presence of two small INDELs. The first one, a 94bp deletion affecting the open-reading frame of the gene Lysophosphatidic acid receptor 1 (*LPAR1*), related to neurodevelopmental disorders and central nervous system diseases, such as neuropsychiatric disorders[96]. And the second, a 4 bp insertion in a non-coding region of Pappalysin-1 (*PAPPA*), a gene affected in major mental disorder patients[97], both previously identified by algorithms developed by WGS, proving that they are real INDELS, using classical cell biology techniques for their identification. The pathogenicity of both INDELs will have to be explored in further investigations.

Regarding the duplications validated by ddPCR (DUP16p23.3 and DUP3p26.3), the presence of both duplications was evident in patients 9, 10, 11 and 12, as they have 3 DNA copies instead of the usual 2 (Figures 10 and 11). But we also could notice that for certain patient's results were not that clear (Figure 10. Sample 5), as the ratio between the set gene and the normalizing gene was 2.5. (Figure 10B). These results should be repeated in order to clarify whether this patient harbor the DUP. Other samples also evidenced some technical problems, as low droplets were generated. The ddPCR states that at least 10.000 droplets need to be generated in order to fully trust the results (Figures B and D).

The objective of this study was to demonstrate the pathogenicity and a founder effect of the genotyped rare CNVs and SNVs in subjects from the same geographical region of the previous studied families.

A founder effect, also called founder mutation or founder variant, is a genetic alteration observed in a community that is or was culturally and/or geographically isolated, in which a single ancestor was the carrier of the altered gene that has proliferated in the group (NIH).

Various cases of founder effect have been detected until today. In 2007 a published study reported genetic founder effects in hospital-treated cases of psychosis, self-harm and suicide between years 1877 and 2005 in cohorts of a small rural community in Norway[98]. There are also several cases of founder effect

in schizophrenia[98,99], one of them was found in a geographical isolated French island in the Indian ocean[100]. Other cases of founder effect in mental related diseases were detected in Alzheimer's Disease[101] and Fragile X Syndrome[102].

Novel studies including higher number of patients will need to be performed in order to finally demonstrate a founder effect of any of the rare SNVs identified in the family 2 of the isolated region of Las Alpujarras. The same approach will need to be further investigated to elucidate whether the duplications identified in Family 1, which origin is two small villages from the north of Majorca, may have arisen as a founder effect.

CONCLUSIONS

1. We have set up the experimental conditions to validate by Sanger sequence and agarose electrophoresis three rare INDELS (*LINC00474* INS_2bp (chr9: 118667077); Del 205 bp (Start 116463148 and End 116463353. It affects an intergenic region); *LPAR1* DEL_94 bp (chr9:113669166-113669260); *PAPPA* INS_4bp (chr9:119154431).

2. We have studied in a new cohort of 27 patients the potential founder effect of the new rare SNVs rs145032100 (*ARHGAP19*) and rs74315508 (*RTN4R*), by amplifying by PCR the genomic region containing the mutation and Sanger sequencing the purified PCR product.

3. We have validated by ddPCR the presence of two CNVs (DUP16p23.3 and DUP3p26.3) in Family 1 and we have studied its potential founder effect in a new cohort of 5 patients from the same towns.

BIBLIOGRAPHY

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM 5)*.; 2013. https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596.
2. Pol Fuster J. Disentangling the genomic architecture of psychosis in families with high prevalence through system genomic approaches. 2021.
3. Dunayevich E, Keck PE. Prevalence and description of psychotic features in bipolar mania. *Curr Psychiatry Rep*. 2000;2(4):286-290. doi:10.1007/s11920-000-0069-4
4. Gaudiano BA, Zimmerman M. Prevalence of attenuated psychotic symptoms and their relationship with DSM-IV diagnoses in a general psychiatric outpatient clinic. *J Clin Psychiatry*. 2013;74(2):149-155. doi:10.4088/JCP.12m07788
5. Gozdzik-Zelazny A, Borecki L, Pokorski M. Depressive symptoms in schizophrenic patients. *Eur J Med Res*. 2011;16(12):549-552. doi:10.1186/2047-783x-16-12-549
6. Reininghaus U, Böhnke JR, Chavez-Baldini UY, et al. Transdiagnostic dimensions of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *World Psychiatry*. 2019;18(1):67-76. doi:10.1002/wps.20607
7. Bramon E, Sham PC. The common genetic liability between schizophrenia and bipolar disorder: a review. *Curr Psychiatry Rep*. 2001;3(4):332-337. doi:10.1007/s11920-001-0030-1
8. Craddock N, Owen MJ. The Kraepelinian dichotomy - Going, going... but still not gone. *Br J Psychiatry*. 2010;196(2):92-95. doi:10.1192/bjp.bp.109.073429
9. Fischer BA, Carpenter WT. Will the Kraepelinian Dichotomy Survive DSM-V? *Neuropsychopharmacology*. 2009;34(9):2081-2087. doi:10.1038/npp.2009.32
10. Reininghaus U, Böhnke JR, Hosang G, et al. Evaluation of the validity and utility of a transdiagnostic psychosis dimension encompassing schizophrenia and bipolar disorder. *Br J Psychiatry*. 2016;209(2):107-113. doi:10.1192/bjp.bp.115.167882
11. Mukherjee S, Shukla S, Woodle J, Rosen AM, Olarte S. Misdiagnosis of schizophrenia in bipolar patients: A multiethnic comparison. *Am J Psychiatry*. 1983;140(12):1571-1574. doi:10.1176/ajp.140.12.1571
12. Heslin M, Lomas B, Lappin JM, et al. Diagnostic change 10 years after a first episode of psychosis. *Psychol Med*. 2015;45(13):2757-2769. doi:10.1017/S0033291715000720
13. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, et al. Diagnostic stability and evolution of bipolar disorder in clinical practice: A prospective cohort study. *Acta Psychiatr Scand*. 2007;115(6):473-480. doi:10.1111/j.1600-0447.2006.00984.x

14. Maier W. Do schizoaffective disorders exist at all? *Acta Psychiatr Scand*. 2006;113(5):369-371. doi:10.1111/j.1600-0447.2006.00763.x
15. Lake CR, Hurwitz N. Schizoaffective disorders are psychotic mood disorders; there are no schizoaffective disorders. *Psychiatry Res*. 2006;143(2-3):255-287. doi:10.1016/j.psychres.2005.08.012
16. Cheniaux E, Landeira-Fernandez J, Lessa Telles L, et al. Does schizoaffective disorder really exist? A systematic review of the studies that compared schizoaffective disorder with schizophrenia or mood disorders. *J Affect Disord*. 2008;106(3):209-217. doi:10.1016/j.jad.2007.07.009
17. Wilson JE, Nian H, Heckers S. The schizoaffective disorder diagnosis: A conundrum in the clinical setting. *Eur Arch Psychiatry Clin Neurosci*. 2014;264(1):29-34. doi:10.1007/s00406-013-0410-7
18. Craddock N, O'Donovan MC, Owen MJ. Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology. *Schizophr Bull*. 2006;32(1):9-16. doi:10.1093/schbul/sbj033
19. Potash JB. Carving chaos: Genetics and the classification of mood and psychotic syndromes. *Harv Rev Psychiatry*. 2006;14(2):47-63. doi:10.1080/10673220600655780
20. Park N, Juo SH, Cheng R, et al. Linkage analysis of psychosis in bipolar pedigrees suggests novel putative loci for bipolar disorder and shared susceptibility with schizophrenia. *Mol Psychiatry*. 2004;9(12):1091-1099. doi:10.1038/sj.mp.4001541
21. Potash JB, Chiu YF, MacKinnon DF, et al. Familial aggregation of psychotic symptoms in a replication set of 69 bipolar disorder pedigrees. *Am J Med Genet - Neuropsychiatr Genet*. 2003;116 B(1):90-97. doi:10.1002/ajmg.b.10761
22. Ruderfer DM, Fanous AH, Ripke S, et al. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry*. 2014;19(9):1017-1024. doi:10.1038/mp.2013.138
23. Labbe A, Bureau A, Moreau I, et al. Symptom dimensions as alternative phenotypes to address genetic heterogeneity in schizophrenia and bipolar disorder. *Eur J Hum Genet*. 2012;20(11):1182-1188. doi:10.1038/ejhg.2012.67
24. Mitchell KJ. What is complex about complex disorders? *Genome Biol*. 2012;13(1):1-11. doi:10.1186/gb-2012-13-1-237
25. Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci*. 2014;17(6):782-790. doi:10.1038/nn.3708
26. Sklar P. LINKAGE ANALYSIS IN PSYCHIATRIC DISORDERS: The Emerging Picture. *Annu Rev Genomics Hum Genet*. 2002;3:371-413. doi:10.1146/annurev.genom.3.022502.103141
27. Tenesa A, Haley CS. The heritability of human disease: Estimation, uses and abuses. *Nat Rev Genet*. 2013;14(2):139-149. doi:10.1038/nrg3377

28. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era - Concepts and misconceptions. *Nat Rev Genet*. 2008;9(4):255-266. doi:10.1038/nrg2322
29. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010;42(7):565-569. doi:10.1038/ng.608
30. Gejman P V., Sanders AR, Duan J. The role of genetics in the etiology of schizophrenia. *Psychiatr Clin North Am*. 2010;33(1):35-66. doi:10.1016/j.psc.2009.12.003
31. Gordovez FJA, McMahon FJ. The genetics of bipolar disorder. *Mol Psychiatry*. 2020;25(3):544-559. doi:10.1038/s41380-019-0634-7
32. Stilo SA, Murray RM. Non-Genetic Factors in Schizophrenia. *Curr Psychiatry Rep*. 2019;21(10):1-10. doi:10.1007/s11920-019-1091-3
33. Rodriguez-Murillo L, Xu B, Roos JL, Abecasis GR, Gogos JA, Karayiorgou M. Fine mapping on chromosome 13q32-34 and brain expression analysis implicates MYO16 in schizophrenia. *Neuropsychopharmacology*. 2014;39(4):934-943. doi:10.1038/npp.2013.293
34. Visscher PM, Goddard ME, Derks EM, Wray NR. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol Psychiatry*. 2012;17(5):474-485. doi:10.1038/mp.2011.65
35. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev*. 2009;19(3):212-219. doi:10.1016/j.gde.2009.04.010
36. Rodriguez-Murillo L, Gogos JA, Karayiorgou M. The genetic architecture of schizophrenia: New mutations and emerging paradigms. *Annu Rev Med*. 2012;63:63-80. doi:10.1146/annurev-med-072010-091100
37. Mitchell KJ, Porteous DJ. Rethinking the genetic architecture of schizophrenia. *Psychol Med*. 2011;41(1):19-32. doi:10.1017/S003329171000070X
38. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science (80-)*. 1996;273(5281):1516-1517. doi:10.1126/science.273.5281.1516
39. Gottesman II, Shields J. A polygenic theory of schizophrenia. *Proc Natl Acad Sci U S A*. 1967;58(1):199-205. doi:10.1073/pnas.58.1.199
40. Lander ES. The new genomics: Global views of biology. *Science (80-)*. 1996;274(5287):536-539. doi:10.1126/science.274.5287.536
41. McClellan JM, Susser E, King MC. Schizophrenia: A common disease caused by multiple rare alleles. *Br J Psychiatry*. 2007;190(MAR.):194-199. doi:10.1192/bjp.bp.106.025585
42. McClellan J, King MC. Genetic heterogeneity in human disease. *Cell*. 2010;141(2):210-217. doi:10.1016/j.cell.2010.03.032
43. Kruglyak L. The road to genome-wide association studies. *Nat Rev Genet*. 2008;9(4):314-318. doi:10.1038/nrg2316

44. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science (80-)*. 2008;320(5875):539-543. doi:10.1126/science.1155174
45. Stone JL, O'Donovan MC, Gurling H, et al. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*. 2008;455(7210):237-241. doi:10.1038/nature07239
46. Girard SL, Gauthier J, Noreau A, et al. Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet*. 2011;43(9):860-863. doi:10.1038/ng.886
47. O ' Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 2012;485(7397):246-250. doi:10.1038/nature10989
48. Xu B, Ionita-Laza I, Roos JL, et al. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet*. 2012;44(12):1365-1369. doi:10.1038/ng.2446
49. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med*. 2010;61:437-455. doi:10.1146/annurev-med-100708-204735
50. Malhotra D, Sebat J. CNVs: Harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 2012;148(6):1223-1241. doi:10.1016/j.cell.2012.02.039
51. Sebat J, Levy DL, McCarthy SE. Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet*. 2009;25(12):528-535. doi:10.1016/j.tig.2009.10.004
52. Marshall CR, Howrigan DP, Merico D, et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet*. 2017;49(1):27-35. doi:10.1038/ng.3725
53. Szatkiewicz JP, Fromer M, Nonneman RJ, et al. Characterization of Single Gene Copy Number Variants in Schizophrenia. *Biol Psychiatry*. 2020;87(8):736-744. doi:10.1016/j.biopsych.2019.09.023
54. Kirov G, Pocklington AJ, Holmans P, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry*. 2012;17(2):142-153. doi:10.1038/mp.2011.154
55. Malhotra D, McCarthy S, Michaelson JJ, et al. High frequencies of de novo cnvs in bipolar disorder and schizophrenia. *Neuron*. 2011;72(6):951-963. doi:10.1016/j.neuron.2011.11.007
56. Priebe L, Degenhardt FA, Herms S, et al. Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. *Mol Psychiatry*. 2012;17(4):421-432. doi:10.1038/mp.2011.8
57. Zhang D, Cheng L, Qian Y, et al. Singleton deletions throughout the genome increase risk of bipolar disorder. *Mol Psychiatry*. 2009;14(4):376-380. doi:10.1038/mp.2008.144

58. McCarroll SA, Feng G, Hyman SE. Genome-scale neurogenetics: Methodology and meaning. *Nat Neurosci*. 2014;17(6):756-763. doi:10.1038/nn.3716
59. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010;11(6):415-425. doi:10.1038/nrg2779
60. Sanders SJ, Neale BM, Huang H, et al. Whole genome sequencing in psychiatric disorders: The WGSPD consortium. *Nat Neurosci*. 2017;20(12):1661-1668. doi:10.1038/s41593-017-0017-9
61. Sleiman PMA, Hakonarson H. Sequencing Approaches to Map Genes Linked to Schizophrenia. In: *The Neurobiology of Schizophrenia*. Elsevier Inc.; 2016:51-59. doi:10.1016/B978-0-12-801829-3.00012-4
62. Wijsman EM. The role of large pedigrees in an era of high-throughput sequencing. *Hum Genet*. 2012;131(10):1555-1563. doi:10.1007/s00439-012-1190-2
63. Simpson CL, Justice CM, Krishnan M, et al. Old lessons learned anew: family-based methods for detecting genes responsible for quantitative and qualitative traits in the Genetic Analysis Workshop 17 mini-exome sequence data. *BMC Proc*. 2011;5(S9):1-6. doi:10.1186/1753-6561-5-s9-s83
64. Glahn DC, Nimgaonkar VL, Raventós H, et al. Rediscovering the value of families for psychiatric genetics research. *Mol Psychiatry*. 2019;24(4):523-535. doi:10.1038/s41380-018-0073-x
65. Moreno B, García-Alonso CR, Negín Hernández MA, Torres-González F, Salvador-Carulla L. Spatial analysis to identify hotspots of prevalence of schizophrenia. *Soc Psychiatry Psychiatr Epidemiol*. 2008;43(10):782-791. doi:10.1007/s00127-008-0368-3
66. Rees E, Kendall K, Pardiñas AF, et al. Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA Psychiatry*. 2016;73(9):963-969. doi:10.1001/jamapsychiatry.2016.1831
67. Kashevarova AA, Nazarenko LP, Schultz-Pedersen S, et al. Single gene microdeletions and microduplication of 3p26.3 in three unrelated families: CNTN6 as a new candidate gene for intellectual disability. *Mol Cytogenet*. 2014;7(1):1-10. doi:10.1186/s13039-014-0097-0
68. Juan-Perez C, Farrand S, Velakoulis D. Schizophrenia and epilepsy as a result of maternally inherited CNTN6 copy number variant. *Schizophr Res*. 2018;202:111-112. doi:10.1016/j.schres.2018.06.062
69. Huang AY, Yu D, Davis LK, et al. Rare Copy Number Variants in NRXN1 and CNTN6 Increase Risk for Tourette Syndrome. *Neuron*. 2017;94(6):1101-1111.e7. doi:10.1016/j.neuron.2017.06.010
70. Mercati O, Huguet G, Danckaert A, et al. CNTN6 mutations are risk factors for abnormal auditory sensory perception in autism spectrum disorders. *Mol Psychiatry*. 2017;22(4):625-633. doi:10.1038/mp.2016.61
71. Te Weehi L, Maikoo R, Mc Cormack A, et al. Microduplication of 3p26.3

Implicated in Cognitive Development. *Case Rep Genet*. 2014;2014:1-6. doi:10.1155/2014/295359
72. Hu J, Liao J, Sathanoori M, et al. CNTN6 copy number variations in 14 patients: A possible candidate gene for neurodevelopmental and neuropsychiatric disorders. *J Neurodev Disord*. 2015;7(1):1-9. doi:10.1186/s11689-015-9122-9
73. Tassano E, Uccella S, Giacomini T, et al. Clinical and Molecular Characterization of Two Patients with CNTN6 Copy Number Variations. *Cytogenet Genome Res*. 2019;156(3):144-149. doi:10.1159/000494152
74. Noor A, Lionel AC, Cohen-Woods S, et al. Copy number variant study of bipolar disorder in Canadian and UK populations implicates synaptic genes. *Am J Med Genet Part B Neuropsychiatr Genet*. 2014;165(4):303-313. doi:10.1002/ajmg.b.32232
75. Repnikova EA, Lyalin DA, McDonald K, et al. CNTN6 copy number variations: Uncertain clinical significance in individuals with neurodevelopmental disorders. *Eur J Med Genet*. 2020;63(1):103636. doi:10.1016/j.ejmg.2019.02.008
76. Paradis S, Harrar DB, Lin Y, et al. An RNAi-Based Approach Identifies Molecules Required for Glutamatergic and GABAergic Synapse Development. *Neuron*. 2007;53(2):217-232. doi:10.1016/j.neuron.2006.12.012
77. Rivero O, Selten MM, Sich S, et al. Cadherin-13, a risk gene for ADHD and comorbid disorders, impacts GABAergic function in hippocampus and cognition. *Transl Psychiatry*. 2015;5(10):655. doi:10.1038/tp.2015.152
78. Redies C, Hertel N, Hübner CA. Cadherins and neuropsychiatric disorders. *Brain Res*. 2012;1470:130-144. doi:10.1016/j.brainres.2012.06.020
79. Hawi Z, Tong J, Dark C, Yates H, Johnson B, Bellgrove MA. The role of cadherin genes in five major psychiatric disorders: A literature update. *Am J Med Genet Part B Neuropsychiatr Genet*. 2018;177(2):168-180. doi:10.1002/ajmg.b.32592
80. Børglum AD, Demontis D, Grove J, et al. Genome-wide study of association and interaction with maternal cytomegalovirus infection suggests new schizophrenia loci. *Mol Psychiatry*. 2014;19(3):325-333. doi:10.1038/mp.2013.2
81. Otsuka I, Watanabe Y, Hishimoto A, et al. Association analysis of the Cadherin13 gene with schizophrenia in the Japanese population. *Neuropsychiatr Dis Treat*. 2015;11:1381-1393. doi:10.2147/NDT.S84736
82. Xu W, Cohen-Woods S, Chen Q, et al. Genome-wide association study of bipolar disorder in Canadian and UK populations corroborates disease loci including SYNE1 and CSMD1. *BMC Med Genet*. 2014;15(1):1-13. doi:10.1186/1471-2350-15-2
83. Moon SY, Zheng Y. Rho GTPase-activating proteins in cell regulation. *Trends Cell Biol*. 2003;13(1):13-22. doi:10.1016/S0962-8924(02)00004-1

84. Huang GH, Sun ZL, Li HJ, Feng DF. Rho GTPase-activating proteins: Regulators of Rho GTPase activity in neuronal development and CNS diseases. *Mol Cell Neurosci*. 2017;80:18-31. doi:10.1016/j.mcn.2017.01.007
85. Tolias KF, Duman JG, Um K. Control of synapse development and plasticity by Rho GTPase regulatory proteins. *Prog Neurobiol*. 2011;94(2):133-148. doi:10.1016/j.pneurobio.2011.04.011
86. Xu Z, Chen Y, Chen Y. Spatiotemporal Regulation of Rho GTPases in Neuronal Migration. *Cells*. 2019;8(6):568. doi:10.3390/cells8060568
87. Duman JG, Mulherkar S, Tu YK, X. Cheng J, Tolias KF. Mechanisms for spatiotemporal regulation of Rho-GTPase signaling at synapses. *Neurosci Lett*. 2015;601:4-10. doi:10.1016/j.neulet.2015.05.034
88. Fukuda Y, Kawasaki H, Taira K. Exploration of human miRNA target genes in neuronal differentiation. *Nucleic Acids Symp Ser (Oxf)*. 2005;49(49):341-342. doi:10.1093/nass/49.1.341
89. Amelio I, Lena AM, Viticchiè G, et al. Mir-24 triggers epidermal differentiation by controlling actin adhesion and cell migration. *J Cell Biol*. 2012;199(2):347-363. doi:10.1083/jcb.201203134
90. Perkins DO, Jeffries CD, Jarskog LF, et al. microRNA expression in the prefrontal cortex of individuals with schizophrenia and schizoaffective disorder. *Genome Biol*. 2007;8(2):27. doi:10.1186/gb-2007-8-2-r27
91. Sinibaldi L, De Luca A, Bellacchio E, et al. Mutations of the nogo-66 receptor (Rtn4r) gene in schizophrenia. *Hum Mutat*. 2004;24(6):534-535. doi:10.1002/humu.9292
92. Thomas RA, Ambalavanan A, Rouleau GA, Barker PA. Identification of genetic variants of LGI 1 and RTN 4R (NgR1) linked to schizophrenia that are defective in NgR1– LGI 1 signaling . *Mol Genet Genomic Med*. 2016;4(4):447-456. doi:10.1002/mgg3.215
93. Hsu R, Woodroffe A, Lai WS, et al. Nogo receptor 1 (RTN4R) as a candidate gene for schizophrenia: Analysis using human and mouse genetic approaches. *PLoS One*. 2007;2(11):1234. doi:10.1371/journal.pone.0001234
94. Bassett AS, Chow EWC. Schizophrenia and 22q11.2 deletion syndrome. *Curr Psychiatry Rep*. 2008;10(2):148-157. doi:10.1007/s11920-008-0026-1
95. Karayiorgou M, Simon TJ, Gogos JA. 22q11.2 microdeletions: Linking DNA structural variation to brain dysfunction and schizophrenia. *Nat Rev Neurosci*. 2010;11(6):402-416. doi:10.1038/nrn2841
96. Xiao D, Su X, Gao H, Li X, Qu Y. The Roles of Lpar1 in Central Nervous System Disorders and Diseases. *Front Neurosci*. 2021;0:950. doi:10.3389/FNINS.2021.710473
97. Pol-Fuster J, Cañellas F, Ruiz-Guerra L, et al. The conserved ASTN2/BRINP1 locus at 9q33.1–33.2 is associated with major psychiatric disorders in a large pedigree from Southern Spain. *Sci Reports 2021 111*.

2021;11(1):1-16. doi:10.1038/s41598-021-93555-4
98. Andersen JE, Hynnekleiv T. Hospital-treated psychosis and suicide in a rural community (1877–2005). Part 2: Genetic founder effects. *Acta Psychiatr Scand*. 2007;116(SUPPL. 436):20-32. doi:10.1111/J.1600-0447.2007.01076.X
99. Maziade M, Paccalet T. A protective-compensatory model may reconcile the genetic and the developmental findings in schizophrenia. *Schizophr Res*. 2013;144(1-3):9-15. doi:10.1016/J.SCHRES.2012.12.022
100. Jay M, Gorwood P, Feingold J, Leboyer M. A one year prevalence study of schizophrenia on Reunion Island. *Eur Psychiatry*. 1997;12(6):284-288. doi:10.1016/S0924-9338(97)84787-6
101. Bird TD, Lampe TH, Nemens EJ, Miner GW, Sumi SM, Schellenberg GD. Familial alzheimer's disease in american descendants of the volga germans: Probable genetic founder effect. *Ann Neurol*. 1988;23(1):25-31. doi:10.1002/ANA.410230106
102. Oudet C, Koskull H von, Nordström AM, Peippo M, Mandel JL. Striking Founder Effect for the Fragile X Syndrome in Finland. *Eur J Hum Genet*. 1993;1(3):181-189. doi:10.1159/000472412

ANNEX

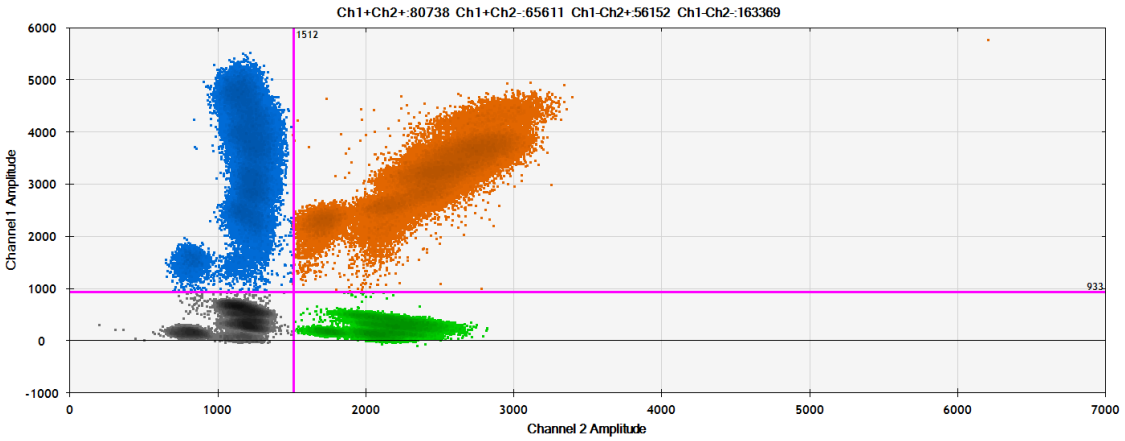

Figure A. DUP16p23.3 CNV validation threshold settle.

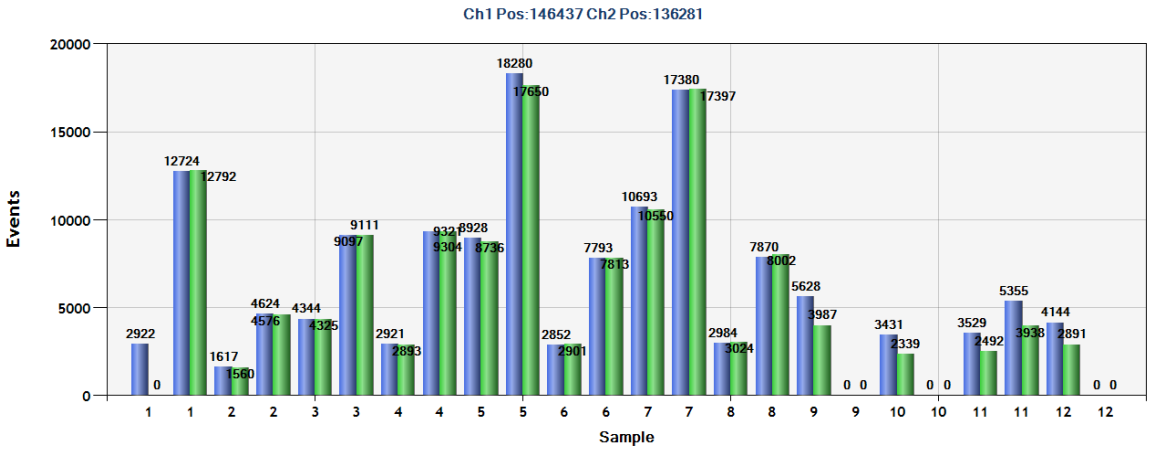

Figure B. DUP16p23.3 CNV validation events.

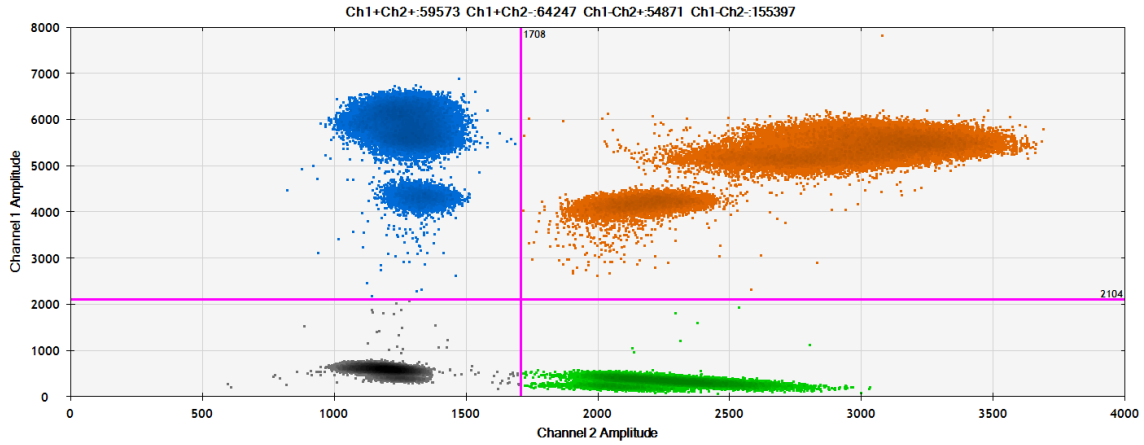

Figure C. DUP3p26.3 CNV validation threshold settle.

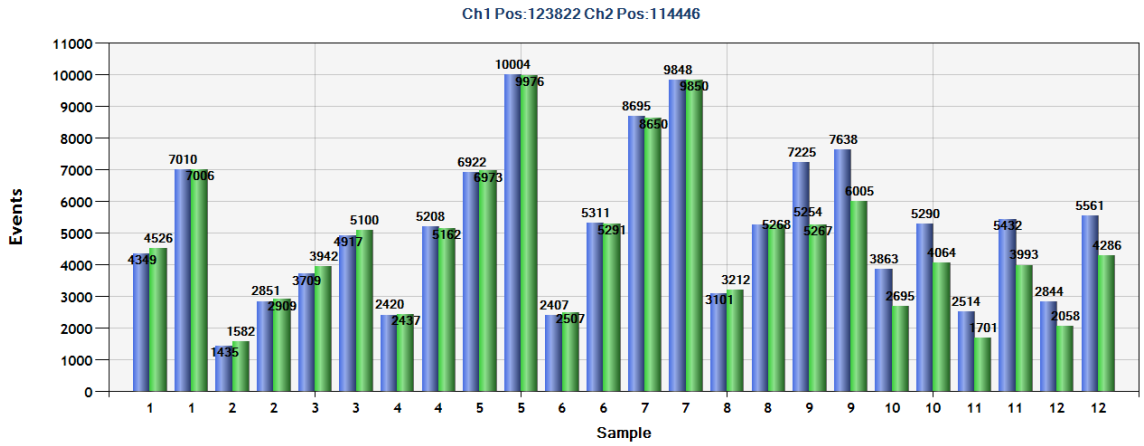

Figure D. DUP16p23.3 CNV validation events.

Table: Set of Primers Used to Validate SNPs, CNVs and INDELs

Name	Forward	Reverse	Probe	PCR
ARHGAP19 (rs145032100)	5'- GACTCCACAAATCTTAATGC-3'	5'- GAGAACCACTGTCCTTGC-3'		PCR
DUP16p23.3	5'-TTGGTGTTTGACCCTGTGAA-3'	5'-TGAGCTAGGGCTCCCACTTA-3'	5'-FAM-TTTGGATTGCTTTGCCTACC-BHQ1-3'	Digital PCR
DUP3p26.3	5'-TCAGTGAAGTGCCTGGTTTG-3'	5'-GGCTGTTCCATGAGGAATGT-3'	5'-FAM-CTAGGCTGGGCTCACTTGTC-BHQ1-3'	Digital PCR