Biotechnology Master Program

Bethlehem University

Faculty of Science

Joint Biotechnology Master Program

Identification of Susceptible Copy Number Variants in a Palestinian Family affected with Intellectual Disability

By
Alaa B. Darwish

……

In Partial Fulfillment of the Requirements for the Degree
Master of Biotechnology

……

December 2014

The undersigned hereby certify that they have read and recommend to the Faculty of Scientific Research and Higher Studies at the Palestine Polytechnic University and the Faculty of Science at Bethlehem University for acceptance a thesis entitled:

"Identification of Susceptible Copy Number Variants in a Palestinian Family affected with Intellectual Disability"

By

Alaa B. Darwish

In partial fulfillment of the requirements for the degree of Master of Science in biotechnology

Graduate Advisory Committee:

Committee Member ,(supervisor)	Date
Committee Member (Internal Examiner)	Date
Committee Member (External Examiner)	Date

Approved for the Faculties

Dean of Faculty of Scientific Research and Higher studies Palestine Polytechnic University	Dean of Faculty of science , Bethlehem university
Date	Date

Abstract

Intellectual disability (ID) is a common problem with major implications for the public health, education and community services. The causes of mental retardation have been found to have a definite etiological basis, which may be biochemical, chromosomal, mendelian genetic disorders, or due to environmental effects. The underlying causes remain unknown in a significant percentage of cases, but genetic causes account for 25-50% of such cases. In the present study, we undertook the challenge of identifying genes involved in ID by performing genome-wide copy number variation analysis of one consanguineous Palestinian family with 3 kids of intellectual disability using cytoscan high density microarrays. We identified copy number variants implicating genes on the chromosome11 *(DLG2)* as well as chromosome X (1ILRB1) as ID candidate genes. Our findings highlight the importance of chromosome 11and X in the etiology of intellectual disability, and demonstrate the power of copy number variation analysis in the identification of disease genes, in particular for complex genetic disorders such as intellectual disability.

ملخص

الإعاقة العقلية (ID) هي مشكلة شائعة مع آثار كبيرة على الخدمات الصحية العامة والتعليم والمجتمع. أسباب هذه الإعاقة مختلفة ومتعددة، والتي قد تكون بيوكيميائية، خلل كروموسومي، الاضطرابات الوراثية المندلية، أو بسبب عوامل بيئية. لا تزال الأسباب الكامنة غير معروفة في نسبة كبيرة من الحالات، ولكن تشكل الأسباب الوراثية ما يقارب 25-50 ٪ من هذه الحالات. في هذه الدراسة، نحاول تحديد الجينات المسؤولة عن هذه الإعاقة عن طريق إجراء تحليل للخلل التركيبي للجينوم في أسرة فلسطينية لها 3 أطفال من المعاقين باستخدام microarray عالية الكثافة والدقة. حددنا خلل وراثي في جين ال DLG2 (chromosome11)، وكذلك جين 1ILRB1 (.X chr) كجينات مرشحة. النتائج التي توصلنا إليها تلقي الضوء على أهمية كروموسوم اكس في مسببات الإعاقة الذهنية، وإظهار قوة تحليل التباين الجيني تحديد جينات المرض، ولا سيما بالنسبة لاضطرابات وراثية معقدة مثل الإعاقة العقلية.

DECLARATION

I declare that the Master Thesis entitled " **Identification of Susceptible Copy Number Variants in a Palestinian family affected with Intellectual disability**" is my own original work, and hereby certify that unless stated, all work contained within this thesis is my own independent research and has not been submitted for the award of any other degree at any institution, except where due acknowledgment is made in the text.

Name and signature: Alaa B Darwish

Date:_____

STATEMENT OF PERMISSION TO USE

Presenting of this thesis is considered as partial fulfillment of the requirements needed to be awarded a Master of Science degree in biotechnology. I agree that the library shall make it available to borrowers under rules of the library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from, reproduction, or publication of this thesis may be granted by my main supervisor or in [his/her] absence, by the Dean of Higher Studies. The proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature: Alaa B Darwish

Date: _____

DEDICATION

I would dedicate this thesis first and foremost to my husband whom without his constant support, love and patience I would have never made it this far. To my kind parents from whom I learned that patience is the key to my success.

Acknowledgment

My deep appreciation to my supervisor Professor Moien Kanaan, to my fellow lab members and best friends whom I spent great days with.
I would to thank all the coordinators of the biotechnology master program from Bethlehem and Polytechnic universities. Finally, I would like to appreciate the contribution of patients and family member, which were collaborative and generous.

List of Abbreviations

ID	Intellectual disability
SID	Syndromic intellectual disability
NS-ID	Non-syndromic intellectual disability
CNS	Central nervous system
MR	Mental retardation
WGSA	whole-genome sampling analysis
SNP	Single nucleotide polymorphism
HD	High density
NGS	Next-generation sequencing
PHA	Phytoheamagglutinin
PCR	Polymerase chain reaction
GTYPE	Genotyping Analysis Software
GCOS	GeneChip Operating Software
SAPE	Streptavidin Phycoerythin
DGV	Database of genomic variation
MAGUK	Membrane-associated guanylate kinase
PSD	post-synaptic density

List of Figures

List of Tables

Table of Contents

CHAPTER 1

Introduction and Literature review

1.1 Intellectual disability: Intellectual disability (ID) is a common form of cognitive impairment and one of the most frequent and disabling neurological abnormalities in children, affecting between 1 and 3% of the population of industrialized countries (1). In developing countries, it seems more frequent, due to environmental factors such as poor health care and malnutrition (2). It is characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills. It is among the more significant public health problems because of its prevalence, the few therapeutic options that are currently available, and the resulting life-long harm to the affected persons, their families, and society as a whole (3). Although there is debate over the definition and classification of ID, psychologists use the diagnostic judgment that mostly based on the measurement of IQ that define ID as an IQ of approximately 70 or below. Also the diagnosis is built upon certain criteria such as: limitations in adaptive behaviour in the following skills: communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health and safety; and onset of the symptoms before 18 years of age (4).

Intellectual disability is divided according to the diagnostic and statistical manual of mental disorders into 5 categories based on IQ: Mild, moderate, severe, profound and unable to classify. However, epidemiological studies often use a simplified classification, grouping their subjects into mild ID (IQ50-70) and severe ID (IQ<50) (5). While the prevalence of severe ID is relatively stable, the prevalence of mild ID is variable and often depends heavily on external environmental factors such as level of maternal education, access to education/opportunity and access to healthcare (6). From a clinical perspective, ID can also be grouped into syndromic intellectual disability (SID) associated with clinical, radiological, metabolic or biological features, and non-syndromic intellectual disability (NS-ID) in which the cognitive impairment represents the only manifestation of the disease (7).

Many environmental such as exposure to certain teratogens, premature birth, viruses or radiation can cause ID, as can severe head trauma or injury causing lack of oxygen to the brain (8). While these factors explain some cases of NS-ID, it is also important to consider genetic factors that can cause ID, including chromosomal abnormalities, and single-gene mutations (9). Genetic causes are present in more than 50% of severely mentally retarded persons (10) identification of these genetic causes considered one of the greater challenges of molecular genetics. Genetic factors include chromosome abnormalities, monogenetic disorders and polygenic factors. Determination of the cause of the ID is of uttermost importance since only if the cause is known, then a prognosis of the disease, a risk assessment and a prenatal diagnosis is possible. Genetic forms of ID are subdivided into two major categories; syndromic ID characterized by clinical, radiological, metabolic or biological features, and non-syndromic characterized only by the cognitive impairment. It is hard to distinguish between both types and in some cases of non-syndromic, it could be recognized as syndromic forms (11).

In general, ID resulting from dysfunction of genetic information that might arise through (i) chromosomal rearrangements, (ii) deregulation of the imprinting of specific genes (iii) and dysfunction of single genes. Determining a specific genetic cause of ID will facilitate the diagnosis and may predict the prognosis of the disease, therefor, the accurate medical care will be proper. The genetic bases of familial ID was not investigated as the sporadic one since its incidence is much lower. Recently, characterization of genetic basis of X-linked and autosomal recessive ID has led to the identification of new genes and mutations (12). However, in Palestine no previous studies has been conducted to identify the underlying genetic causes of familial ID.

1.2 Epidemiology of ID

Worldwide prevalence is 2-85 per 1000 individuals and the incidence in some countries is estimated to be around 1-3% (13). The prevalence in the United States is about 14.9 per 1000 (14) while, in India, the national prevalence rate is currently 1.4-25.3 per 1000 (15). This wide range of variation could be due to the different classification and survey systems used in this regard (16). In the developing countries, incidence of ID is much higher compared to the developed countries (17); this could be due to higher infant

mortality in the developing countries (18). In addition, epidemiological studies have shown a higher prevalence among the male population compared to female (ratio 1.9) (18). Low birth weight in and higher mortality rates in males suggest that gender variation could be influence the variation and the epidemiology of ID.

No statistical and updated data regarding the prevalence of intellectual disability is available in Palestine. In general, no published work has been found that discusses the treatment approach to ID in Palestine. ID patients seems to be a neglected group of patients who have not been well studied. Among Palestinians, the rate of consanguinity is very high and some 44.3% of the marriages are between relatives (22.6% of them between first cousins). The importance of genetic factors in various congenital malformations, such as ID was confirmed by the observation of a significantly higher consanguinity rate in Palestinian parents of these patients than is observed in the general population (19).

1.3 Pathophysiology of ID

ID has been defined as a manifestation of a group of disorders of the central nervous system controlled by cortical structures (20). Most individuals with cognitive impairment have no structural abnormalities of the brain (20). Visual CNS malformations such as neural tube defects, hydranencephaly, and microcephaly are only diagnosed in 10-15% of cases (20). In many cases, ID is a complex syndrome characterized of developmental brain abnormalities that lacks the normal brain organization and function. (21). In this regard, two major groups of genes can be categorized; genes involved in ID with brain developmental abnormalities, and those with no specific brain abnormality (21). This subdivision provides the understanding of the biological mechanisms underlying ID, despite that it have many weaknesses. However, understanding the biology of ID is complicated by extremely heterogeneity of genetic causes associated with it (22). Genetic abnormalities can lead to defects in synaptic structure or function and deregulation of specific brain process and pathways (21). This suggest that deficient proteins in postnatal stages of active learning periods can be prevented or improved if early postnatal diagnosis and subsequent intervention are implemented like behavioral and cognitive therapies can help ID patients to reach their maximum potential (23).

1.4 Chromosomal abnormalities

15-30 % of all ID persons harbor chromosomal abnormalities (24). These include numerical abnormalities, partial abnormalities and microdeletions. The frequent numerical abnormalities in ID include trisomy 13 (Patau's syndrome), 18 (Edwards' syndrome) and 21 (Down syndrome) that is the most reported cause of ID (25).

Partial abnormalities include deletions, insertions, inversions, and translocations. Examples of visible deletions include (25) cri-du-chat syndrome, however, not all deletions are cytogenetically visible such as the submicroscopic deletion on chromosome 7q11.23 that causes Williams-Beuren syndrome and the Smith-Magenis syndrome caused by a deletion of chromosome 17p11.2. Also the velocardiofacial syndrome, caused by a 22q11 deletion. Specific emerging deletions in the telomeric region significantly cause the familial ID (26). These regions are gene-rich which explains why small deletions can cause ID such as in the Miller-Dieker syndrome (deletion in 17p telomere) and the a-thalassemia/ mental retardation syndrome (deletion in 16p telomere) (18). Table 1 summarizes some examples of syndromes and chromosomal alterations associated with intellectual disability (27).

Table 1: Syndromes and chromosomal alterations associated with ID.

Syndrome	Chromosomal abnormality	OMIM	Clinical hallmarks
Down	Trisomy 21	190685	Intellectual disability and characteristic faces, congenital malformations of the heart, significant hearing loss, and early onset Alzheimer disease
Prader willi	Del15q11-q13	176270	Diminished fetal activity, obesity, muscular hypotonia, intellectual disability, short stature, and small hands and feet
Angelman	Del15q11-q13	105830	Intellectual disability, movement disorder, abnormal behavior, and severe limitation in speech and language.
Smith-magenis	Del17p11.2	182290	Speech delay with or without hearing loss, growth retardation, Brachycephaly, and midface hypoplasia

Miller-dieker	Del17p13.3	247200	Lissencephaly, microcephaly, cardiac malformation, and intellectual disability with seizures
DiGeorge	Del22q11.2	18840	Neonatal hypocalcemia, infection susceptibility, cardiac malformations, low set ears, small mouth, short stature, and learning difficult.
Edwards	Trisomy 18	-	Kidney malformations, structural heart defects, microcephaly, intellectual disability, cleft palate, narrow eyelid folds, and micrognathia
Williams-beuren	Del17q11.23	194050	Intellectual disability and distinctive facial features

The recent introduction of genome wide techniques such as array CGH and SNP arrays are now used to identify submicroscopic copy number changes. These techniques have been applied to patients with ID (38, 39). Identification of specific copy-number changes in affected patients as compared with control has led to a rapid discovery of novel microdeletion and microduplication syndromes associated with ID (40) as shown in table 2.

Table 2: Novel Recurrent Copy-Number Changes Associated with Intellectual Disability and Related Disorders.

Chromosome region	Coordinates in Mb	Deletion or duplication associated with disorder
1q21.1	**Chr. 1: 145.0-146.35**	**Deletion: intellectual disability, schizophrenia, multiple congenital anomalies, Duplication: intellectual disability, autism**
3q29	**Chr. 3: 197.4-198.9**	**Deletion: intellectual disability, schizophrenia Duplication: intellectual disability.**
10q22-q23	**Chr. 10: 81.12-89.07**	**Deletion: intellectual disability**
15q11.2	**Chr. 15: 20.3-20.7**	**Deletion: intellectual disability, schizophrenia, epilepsy**
15q13.3	**Chr. 15: 28.7-30.2**	**Deletion: intellectual disability, schizophrenia, epilepsy, autism.**
15q24	**Chr. 15: 72.2-73.8**	**Deletion: intellectual disability, autism.**
16p11.2 (a)	**Chr. 16: 29.5-30.1**	**Deletion: intellectual disability, autism, obesity. Duplication: schizophrenia.**
16p11.2 (b)	**Chr. 16: 28.7-29.0**	**Deletion: intellectual disability, obesity.**
16p12	**Chr. 16: 21.8-22.4**	**Deletion: intellectual disability.**
16p13.11	**Chr. 16: 15.4-16.4**	**Deletion: intellectual disability, autism, epilepsy. Duplication: intellectual disability, autism, ADHD**
17q12	**Chr. 17: 31.8-33.3**	**Deletion: intellectual disability, schizophrenia, autism.**
17q21.3	**Chr. 17: 41.0-41.7**	**Deletion: intellectual disability**

Several novel microdeletions and duplication have been identified in patients who have a wide range of phenotypic features and severity. Heterozygous deletions of 17q21.31 are associated with moderate-to-severe intellectual disability (37, 38). All 17q21.31 deletions has never been seen in healthy control subjects. Its prevalence estimated to be 1 in 16,000 persons (41). Deletions of 15q24 is another deletion identified in patients with ID syndrome with features include developmental delay and intellectual disability that is usually moderate to severe; prolonged speech delay or the absence of speech; and dysmorphic features (42). Deletions of 1q21.1 have been associated with variable degrees of ID, and some patients have congenital anomalies, including cataracts and congenital heart disease (43). Deletions of this region have also been associated with schizophrenia (44). Duplications in the same region are also associated with mild ID and autistic features in some patients (43). Another example deletion with variable outcomes is the 16p11.2 deletion. It is present in up to 1% of those with autism spectrum disorders and associated with ID without autistic features (44).

1.5 Monogenic causes of ID:

Recently, Inlow and Restifo (45) identified about 282 autosomal and X-linked monogenic causes of ID, out of those genes, 16% reside on the X chromosome (45). Ropers HH and Chiurazzi P (46) identified 60 X-linked genes involved in ID disorders. The apparent excess of X-linked genes suggest that the human X chromosome contains a high density of genes influencing cognitive abilities (47).

In many cases, ID is part of a complex syndrome comprising developmental brain abnormalities required for normal development of the CNS. Accordingly, two major groups of genes could be distinguished: (i) genes involved in ID disorders with brain developmental abnormalities; (ii) genes involved in ID disorders with no specific brain abnormality, this subdivision allows to highlight genes implicated in potential common genetic and functional pathways and provides bases for understanding pathological mechanisms underlying ID. Most of the genes that are involved in ID with no apparent brain function are X-linked, although the number of autosomal genes involved in autosomal recessive ID is progressively increasing (table 2). Functional analysis of

characterized ID genes fall into distinct functional subclasses such as transcription and chromatin-remodeling factors, transmembrane proteins, microtubule- and actin-associated proteins, regulators and/ or effectors of RhoGTPase pathways.

1.6 Genes that cause ID

Defects in synaptogenesis and synaptic activities as well as their plasticity, especially in postnatal stage during learning and acquisition of intellectual performances and emotional behavior, are shown to cause cognitive impairment resulting from mutations in ID related genes (48) including FMRP, OPHN1, NLGN4, DLG3, RabGDI, Neurotrypsin and probably PAK3. For ILRAPL. FMRP protein, whose mutation causes fragile X syndrome, is an RNA-binding protein. Koekkoek et al (49) showed that knockout FMR cells exhibit abnormal morphology of dendritic spines with synaptic dysfunction. Interestingly, Schenck et al (50) showed that FMRP function is regulated through Rac1, a member of the RhoGTPase subfamily, dependent signaling pathway (50). The importance of RhoGTPase signaling pathways in the development of ID is also supported by the evidence that ID genes such as OPHN1, PAK3, ARHGEF6 and FGD1 encode regulators and/or effectors of RhoGTPases, and also implicated in the regulation of dendritic spine morphology and synaptic activity (51).

Potential roles in the regulation of synaptic activity have also been proposed for another group of genes involved in ID which encode for transmembrane proteins such as NLGN4, (52) DLG350 and IL1RAPL and for the soluble protein GDI1(53, 54). Primary functions of these proteins are diverse. However, they have in common the subcellular localization at the pre or postsynaptic compartments. The importance of the regulation of synaptic activity in ID development is also supported by the implication of another synaptic protein encoded by the DLG3 gene (55). DLG3 encodes the synapse-associated protein 102 (SAP102), a member of the membrane-associated guanylate kinase protein family. Mutations identified in DLG3 are shown to be associated with MR and predicted to impair the interaction of SAP102 with N-methyl-D-aspartate (NMDA) receptors that involved in NMDA receptor signaling pathways (55).

Altogether, these genetic data in combination with functional studies suggest that deregulation of synaptic activity and plasticity could be regarded as one of the cellular bases that contributes to the pathogeny of ID.

1.7 Genome-wide homozygosity mapping: It is well known that sites of genetic variation could be used as markers to identify disease segregation patterns among families. This approach successfully led to the identification of a number of genes involved in different disorders (56). The strategy of homozygosity mapping depend on the fact that in a child of a consanguineous marriage affected with a homozygous genetic disease mutation, a large region spanning the disease locus is probably also homozygous. Thus, searching for regions that are homozygous by descent in all affected individuals would probably provide a strategy for mapping the recessive gene mutation. Single nucleotide polymorphism are the most common type of sequence variation and are powerful markers due to their abundance (over 7 million), stability, and relative ease of scoring.

The genechip mapping assay, also known whole-genome sampling analysis (WGSA) are now used for SNP genotyping (57, 58). In contrast to traditional genotyping techniques that use multiple oligonucleotides per SNP, the chip assay amplifies a segment of the DNA through a single primer amplification reaction using restriction enzyme digestion and adaptor ligation. The GeneChip® Human Mapping 500K Array is used for genotyping of 500,000 SNPs, this high number of SNPs that can be analyzed in one assay will increase the amount of informative data that can be used for association studies and will offer an alternative tool for the traditional linkage analysis.

1.8 CytoScan Array Cytogenetics is evolving with new techniques to improve test resolution, sensitivity, specificity, and accuracy. Nowadays, genome-wide microarrays are recommended as first choice for evaluating patients with clinical syndrome caused probably by genetic alterations (59) in which conventional karyotyping could not detect any abnormalities. The array analysis successfully identified abnormalities on many cases, thus contributing significantly to their disease prognosis. DNA microarrays offer the latest technological innovations for genome wide analysis, genotyping and diagnosis

(60). Generally, the array is composed of DNA probes that are bound to a solid substrate such as glass. Each spot in the array is composed of identical probes that are complementary to the sequence of interest. During hybridization, specific DNA sequences that are complement to the specific probe sequence will anneal and form a DNA duplex. Those duplex can then be detected by reporter molecule systems. DNA microarrays can be classified according to the type of probes on the array that can be a product of polymerase chain reaction or oligonucleotide DNA (61). Recently, high-density genotyping arrays such as Affymetrix CytoScan HD arrays have been used for copy number variation analysis as well as SNP analysis. They also can provide higher precision and resolution than traditional techniques (62). Affymetrix CytoScan HD arrays have many advantages over other arrays since it provides extensive coverage and highest resolution of molecular cytogenetic analysis tools with its 2,690,000 marker content (750,000 polymorphic (SNP) and 1,900,000 non-polymorphic markers) (59).

As a Brief description for the basic principles of the cytoscan, the total genomic DNA is digested with Nsp I restriction enzymes and ligated to adaptors that recognize the cohesive overhangs produced by Nsp I. All fragments resulting from digestion are ligated. A generic primer that recognizes the adaptor sequence is used to amplify adaptor-ligated DNA fragments. PCR conditions have been optimized to amplify fragments in the 200 to 1,100 bp size range. The amplified DNA is purified, then fragmented, labeled and hybridized to the array.

1.9 Next Generation Sequencing

Inexpensive and fast production of large volumes of sequence data is a great demand for innovative technologies that directed the recent development of next-generation sequencing (NGS) technologies. Sanger sequencing has been the standard method for DNA sequencing. However, despite several improvements throughout the years, its basic process is not adequate for fast and complete sequencing of one or multiple genomes. While Sanger sequencing gives an output of 120,000 bp in 24-hours for the cost of $4,000 per Mb sequenced, a single run in NGS yields more than 30 Gb in the same time for less than 2$ per Mb. Moreover, NGS method can sequence a 3.2 Gb of a single human genome in 1 day at a reasonably low cost, in comparison to Sanger sequencing

that would take up to 73 years at a cost of $200,000. In principle, the concept behind NGS technology is similar to capillary electrophoresis; the bases of small fragment of DNA are identified from signals emitted as each fragment is resynthesized form DNA template strand. NGS extend this process across millions of reactions in a massively parallel fashion, rather than being limited to a single DNA fragments. This enables rapid sequencing of large stretches of DNA in the entire genomes. In brief the process works as follow; the genomic DNA is first fragmented into small segments that can be sequenced in millions of parallel reactions, then the new strings of bases, called reads, are reassembled using a known reference genome (resequencing) or in the absence of a reference genome (de novo sequencing). The final set of aligned reads reveals the entire sequence of the genomic DNA sample.

Diagnosis of ID is challenging due to the broad spectrum of phenotypic presentations, as patients with ID often have congenital anomalies and/or autism spectrum disorders. With the advent of NGS, there has been expectation of an increased detection of the genetic causes of ID. In particular, NGS has been identified as an effective tool for discovery of new disease genes implicating in the development of ID. Different studies have shown that NGS is particularly suited for clinical diagnostic testing for genetically heterogeneous conditions where there are a large number of known candidate genes (63, 64). The NGS for our family member have been done at genetics lab at Washington University.

1.10 Problem statements and objectives

Intellectual disability, whether acquired or genetic, is a major health concern worldwide and even a more serious problem among the Palestinian population where the rate of consanguinity is very high and some 44.3% of the marriages are among relatives. ID is one of the complex disorders known since it is genetically heterogeneous, with many different genetic defects giving rise to clinically indistinguishable phenotypes. More than 90 different gene defects have been identified for X-chromosome-linked intellectual disability alone and over two hundred eighty genes have been linked to ID up to date. Although it is well established that genetic factors play an important role in the etiology of ID, conventional methods such as linkage analysis and association studies have shown a limited success in identification of ID genes. We carried out this study in hereditary research lab at Bethlehem University, we have performed homozygosity mapping, copy number variation analysis, and next-generation sequencing in one consanguineous Palestinian family with intellectual disability

More specifically the study aims to achieve the following goals:

1. Identify the causative mutation in a consanguineous Palestinian family by using different molecular approaches.
2. Identify ID candidate genes or louci that may experience CNV utilizing ship array.

CHAPTER 2

Materials and Methods

2.1 Family data

Figure 1 shows the pedigree of family AUT. Family members participated in the study after obtaining informed consent in accordance with the guidelines of Bethlehem University IRB Committee. The parents are cousins, they have six children; 2 females and 4 males, the affected are the older female (AUT-Z3) and 2 males, (AUT-Z4, and 5) who were scored for ID by Caritas Baby Hospital. The most obvious phenotype of are reduced ability to learn, compared to their normal peers of the same age coupled with delays in language skills, also they characterized by having a retarded movement and balance, poor memory and attention, irresponsibility, emotional imbalance, and excessive movement. It is worth to mention that the female have the most regressive phenotype. DNA was extracted from 5ml of peripheral blood of the family members using the salt method (reference needed here)

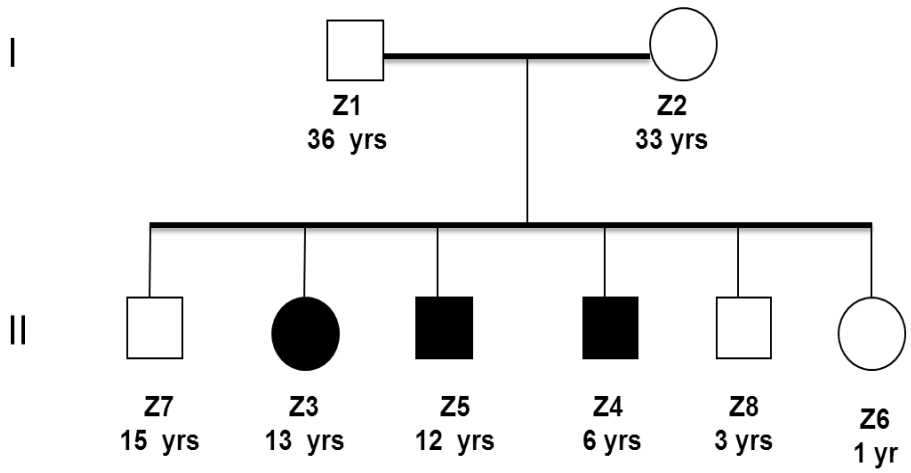

Figure 1: Pedigree of family AUT with intellectual disability syndrome.

2.2 Chromosomal analysis

2.2.1 Harvesting:

0.5 ml of blood sample was cultured into 5 ml PHA mitogen (phytoheamagglutinin) media, inverted and incubated for three days at 37C. Then 50 µl colcemed (10µg/ml) was added to the cultured-sample for 1 hour. Colcemed acts by binding with microtubules to suppress its action, hence, stops the cells in metaphase stage. After that the sample was centrifuged for 10 min at 1000 rpm. The supernatant was removed, the pellet was resususpended in 8-10 ml of hypotonic solution (0.56% KCl) and then incubated for 20 min at 37 C. 2 ml of fresh fixative solution (1:3 glacial acetic acid: absolute methanol) was added and tubes were inverted to mix. Then The sample was centrifuged for 10 min at 1000 rpm. The supernatant was discarded, the pellet was gently suspended in the last few drops, and 8 ml of fixative solution was added and centrifuged for 10 min at 1000 rpm. The last step was repeated for at least 2 times until pellet is clear white. Finally, the pellet was suspended in a small volume of fixative solution.

2.2.2 Slide preparation:

Slides were cleaned by alcohol and rinsed with distilled water. 2 to 3 drops of cell suspension were dropped on wet slides held at 45 degree angle, and placed at a humid chamber at 37- 45 C until dried. Then the slides were placed at the hot plate at 90 C for 1 hour and stained as follow: First, in pH 7 buffered solution which prepared by dissolving 1Gurr buffer tablet in 1 litter 0.9% NaCl. Then, in Trypsin solution for 30 seconds, which prepared by adding 1 ml of 0.25% trypsin in 40 ml pH 7 buffer. After trypsin, slides were rinsed in two jars of pH 7 buffer. Finally, slides were stained using giemsa solution for 2 min, which prepared by placing 1.5 ml of stain with pH 6.8 buffered solution. PH 6.8 buffered solution was prepared by dissolving 1 Gurr tab in 1 litter of DW, and adjusted by pH meter to 6.8 by NaOH or HCl. The slides were visualized using the light microscope and 20 metaphases were analyzed for every sample.

2.3 Testing for Fragile X Syndrome (FXS)

FXS is an X-linked semi-dominant condition with reduced penetrance. It is caused by an expansion of an unstable CGG trinucleotide repeat in the fragile X gene (FMR-1). In addition to being associated with characteristic physical and behavioral features, causes intellectual disability ranging from mild to severe. It is the most common cause of inherited intellectual disability and is second only to Down's syndrome as the most common genetic cause of intellectual disability (98).Normal individuals carry between 5 and 54 copies of CGG repeat. In normal carriers, the number of CGG repeats (the permutation) is between 55 and 200. In individuals clinically manifesting the syndrome (full mutation), the CGG repeats increase to 200–2000 or more. Such a large mutation is usually accompanied by hypermethylation of the DNA sequence, whereby methyl groups attach to the CGG triplets that renders the FMR-1 gene transcriptionally inactive (97).

2.3.1 DNA extraction

DNA was obtained from venous blood samples on EDTA tube. DNA was extracted using the salt base method protocol according to miller et al (99) as following;
Five to ten ml of blood was collected in a sterile EDTA vacutainer tube. Red Blood Cell lysis buffer was added 4 times the volume of the blood and mixed gently. Tubes were kept on ice for 30 minutes and then centrifuged at 2000 rounds per minute (rpm) for 10 minutes at 4°C. The supernatant was carefully removed and the pellet was re-suspended in 3 ml Red Blood Cell lysis buffer and centrifugation was repeated. After breaking the pellet, it was suspended in a mix of 3ml of 1X Lysis buffer, 100ul of 20% *SDS* (Amresoc-Cat#:M112) and 100ul of 5mg/ml Proteinase K (Amresco-Cat#E115), followed by incubation at 37°C overnight. After incubation, 1ml of 6M NaCl was added to the lysate and vigorously vortexed. Then it was centrifuged at 3000 rpm for 20 minutes at room temperature. The supernatant (the clear upper phase) was transferred gently into a 15ml-tube, avoiding the salt protein deposit. Absolute cold ethanol (EtOH) was added in twice the volume of the supernatant and gently mixed by inverting the tube. DNA was removed with a glass Pasteur pipette and washed in 70% EtOH. After air-drying for a few minutes on the Pasteur pipette, DNA was dissolved in 0.02% Sodium Azide (Sigma-

Aldrich- Cat#: S2002) (200-1000ml depending on the amount of DNA) and left at room temperature overnight.

2.3.2 PCR assay

PCR amplifications were performed using fail safe PCR pre-mix selection kit in a total volume of 25 µl containing 1.0 µl of genomic DNA, 0.5 µl of each primer (forward and reverse from 10pmol), 0.25 µl Taq DNA polymerase (60 unit), 10.25 µl H2O, and 12.5 µl buffer J that give the best result among the kit 12 buffers. The PCR cycling profile was as follows: initial denaturation at 98°C for 2 minutes, 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute, and a final extension at 72°C for 10 minutes. Allele sizes were determined using 2 % agarose gel electrophoresis that was prepared by dissolving 2gm of agarose powder in 100ml 1X TAE buffer.

FMR1 Primers:

Forward primer: (5'-GCTCAGCTCCGTTTCGGTTTCACTTCCGGT-3') and

Reverse primer: (5'-AGCCCCGCACTTCCACCACCAGCTCCTCCA-3')

2.4 Genome-wide homozygosity mapping

Whole Genome scan was performed in hereditary research lab by using Genechip Mapping 500K array NspI chip (Affymetrix) for all family members and the data was analyzed using GeneChip Genotyping Analysis Software (GTYPE) and GeneChip Operating Software (GCOS) according to manufacture.

2.5 Cytoscan microarray

2.5.1 DNA extraction

DNA was isolated from whole blood. Blood was drawn using EDTA tubes, DNA was extracted using the FlexiGene DNA Kit.

2.5.2 CytoScan microarray procedure:

DNA samples were processed according to the instructions provided in the Affymetrix CytoScan™ Assay Manual. The protocol is presented in the following detailed stages:

Dilute the Genomic DNA

1. Each DNA sample was diluted to 50 ng/µL using Low EDTA TE buffer.

2. 5 µL of the diluted DNA sample and 5 µL Low EDTA TE buffer (negative control) were transferred to well of the digest/ligate plate.

Stage 1: Restriction Enzyme Digestion

The genomic DNA was digested by the Nsp I restriction enzyme. After preparation of the Nsp I Digestion Master Mix as in the below table, 14.75 µl of Master Mix was added to 5 µl of each DNA sample. The samples included genomic DNA sample and one positive and one negative control.

Table 3: Digestion Master Mix

Reagent	**Sample**
Nuclease free water	11.55 µL
10X Nsp I buffer	2.00 µL
100X BSA	0.20 µL
Nsp I	1.00 µL
Total	**14.75 µL**

The plate was placed onto a thermal cycler and run by the CytoScan Digest program (table 3)

Table 4: Digestion PCR Program

Temperature	**Time**
37 °C	**2 hours**
65 °C	**20 minutes**
4 °C	**Hold**

Stage 2: Ligation

The digested samples were ligated simultaneously using Nsp Adaptor. After preparing Nsp Ligation Master Mix (table 4), 5.25 µL of the mix was added to each digested sample. The samples were placed onto a thermal cycler and run by the CytoScan Ligate program (table 5). After finishing the program, the ligated samples were diluted with 75 µl of Chilled Affymetrix® Nuclease-Free water.

Table 5: Ligation Master Mix

Reagent	Sample
10X T4 DNA Ligase Buffer	2.50 µL
50 µM Adaptor, Nsp I	0.75 µL
T4 DNA Ligase	2.00 µL
Total	**5.25 µL**

Table 6: Ligate PCR Program

Temperature	Time
16ºC	**3 hours**
70ºC	**20 minutes**
4ºC	**Hold**

Stage 3: PCR

Ligated sample were transferred into plate. Nsp PCR Master Mix was prepared (table 6) and 90 µL of the Mix was added to 10 µL of ligated sample and each plate was placed on a thermal cycler and CytoScan PCR Program was run (table 7).

Table 7: PCR Master Mix

Reagent	1 Sample
Nuclease-Free Water	**39.5 µL**
10X TITANIUM™ Taq PCR Buffer	**10.0 µL**
GC-Melt Reagent	**20.0 µL**
dNTP Mixture (2.5 mM each)	**14.0 µL**
PCR Primer, 002	**4.5 µL**
50X TITANIUM™ Taq DNA Polymerase	**2.0 µL**
Total	**90.0 µL**

Table 8: Nsp PCR program

Temperature	Time	Cycles
94ºC	3 minutes	1X
94ºC 60ºC 68ºC	30 seconds 45 seconds 15 seconds	} 30X
68ºC	7 minutes	1X
4ºC	HOLD	
Volume: 100 µL		

After finishing the PCR Program, PCR products were run on 2% TBE agarose gel at 140 V for 1 hour for verifying that the PCR product distribution is between ~150 bp to 2000 bp.

Stage 4: PCR Product Purification

Nsp PCR reactions were pooled to a 1.5 mL Eppendorf Safe-Lock tube and 720 mL of the magnetic beads were added to each pool, mixed well by inverting and incubated at room temperature for 10 min. The tubes were centrifuged for 3 min at maximum speed (16,100 rcf) and Placed on a magnetic stand (MagnaRackTM). The supernatant were discarded without disturbing the bead pellet and 1.0 mL Purification Wash Buffer were added to each tube. After that, tubes were loaded into the foam adapter, and vortex at maximum setting for 2 min, then centrifuged for 3 min at 16,100 rcf. Then, placed back on the magnetic stand. The supernatant were discarded. Tubes were spin for 30 sec at 16,100 rcf, then placed back on the magnetic stand and 52 mL of Elution Buffer were added to each tube. Tubes were loaded into the foam adapter, and vortex at maximum power for 10 min to resuspend the beads. Centrifuged for 3 min at maximum speed 16,100 rcf. Tubes were placed on the magnetic stand for 10 min until all beads are pulled to the side. Finally, 47 mL of eluted sample were transferred to the appropriate well. Vortex at high speed for 1 sec and spin down at 2000 rpm for 1 min.

Stage 5: Quantitation

The concentration of PCR products after purification with magnetic beads was measured using a spectrophotometer (NanoDrop®). The OD of each PCR product was measured at 260 nm.

Stage 6: Fragmentation

PCR products were fragmented using Fragmentation Reagent. After preparing the fragmentation Master Mix table 8) the 10 µl of Master Mix was added to the purified PCR product, placed the plate onto a thermal cycler and the CytoScan Fragment Program (table 9) was run. When the program was completed the results of this stage was checked by running 4 µl of each reaction on a 4% agarose gel to ensure that majority of fragment distribution between 25 to 125 bp.

Table 9: Fragmentation Master Mix

Reagent	**Reagent concentration (enzyme: 2.5 U/µL)**
Chilled AffymetrixR Nuclease-FreeWater	**123.8 µL**
10X Fragmentation Buffer	**158.4 µL**
Fragmentation Reagent	**5.8 µL**
Total	**288.0 µL**

Table 10: Fragmentation PCR Program

Temperature	**Time**
37ºC	**35 minutes**
95ºC	**15 minutes**
4ºC	**Hold**

Stage 7: Labeling

The fragmented samples were labeled using the DNA Labeling Reagent. After preparing the Labeling Master Mix (table 10) the 19.5 µl of Master Mix was added to each sample. The samples were placed onto a thermal cycler and the CytoScan Label Program (table 11) was run.

Table 11: Labeling Master Mix

Reagent	1 Sample
5X TdT buffer	14.0 µL
30 mM DNA labeling reagent	2.0 µL
TdT	3.5 µL
Total	19.5 µL

Table 12: labeling PCR program

Temperature	Time
37 ºC	4 hr
95 ºC	15 minutes
4 ºC	Hold

Stage 8: Hybridization

Arrays were allowed to warm to room temperature by leaving them on the table for 15 minutes. A 200 µl pipette tip was inserted into the upper right septum of the each array. The Hybridization Master Mix was prepared according to table 12.

Table 13: Hybridization master mix

Reagent	1 Sample
Hyb buffer 1	165.0 µL
Hyb buffer 2	15.0 µL
Hyb buffer 3	7.0 µL
Hyb buffer 4	1.0 µL
Oligo control reagent 0100	2.0 µL
Total	190 µL

After preparation of Hybridization Master Mix, 190 µl of Master Mix was added to each sample on the Label Plate. After vortexing and spinning the plate to assure that the sample was mixed properly, the plate was placed onto the thermal cycler and the CytoScan Hyb Program (Table 13).

Table 14: Hyb PCR Program

Temperature	Time
95 ºC	10 minutes
49 ºC	Hold

Stage 9: Load the Samples onto Arrays

200 µL of each sample were removed and immediately injected into an array. The septa were covered on the array with the 1/2" Tough-Spots that were previously placed on the top edge of the array. The loaded samples were placed in the hybridization oven and allowed to rotate at 50 °C, 60 rpm for 16 to 18 hr.

Stage 10: Washing and staining arrays

Washing and staining steps were carried out by using GeneChip® Fluidics Station (Figure 6) and it was operated using. The Affymetrix staining protocol for mapping arrays was a three stage process which consisted of 1) Streptavidin Phycoerythin (SAPE) stain, 2) an antibody amplification step, and 3) a final stain with Streptavidin Phycoerythin (SAPE). All three solutions were placed on the Fluidics Stations into the sample holders 1, 2 and 3 in vials (500 µL, 500 µL, 800 µL of Stain Buffer 1, Stain Buffer 2, and Array Holding Buffer respectively). The arrays were inserted into the designated modules of the fluidics station while the cartridge lever is in the Down or Eject position. Before starting the program, the fluidics station was primed to ensure that the lines were filled with the appropriate buffers and the fluidics station was ready to run fluidics station protocols.

Scanning arrays

Scanning was performed by using the GeneChip® Scanner 3000 7G. The glass surface of the array was cleaned with the non-abrasive towel before scanning and both septa on the array were carefully covered with Tough Spots.

After placing the arrays in the scanner and selecting the right experiment, the Start button was clicked and scanning was started. The scanning of one array lasted approximately 15 minutes.

CHAPTER 3

Results

3.1 Results of karyotype

Karyotype showed normal set of 46 chromosomes among all family members including the affected and non-affected individuals as shown in the figure 2 below

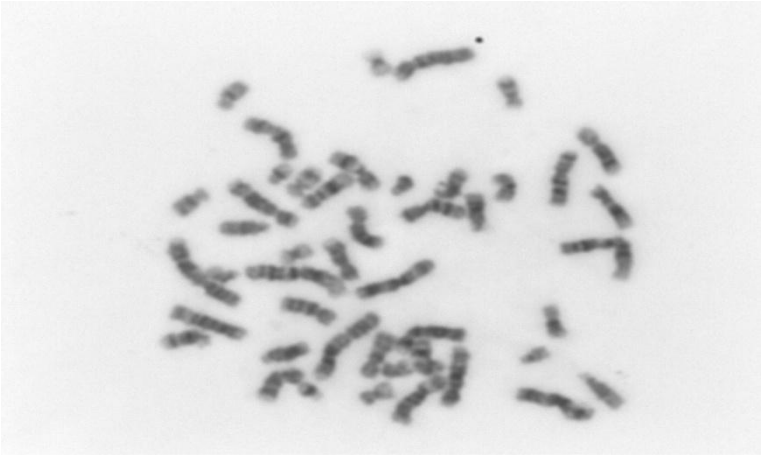

Figure 2: Karyotype analysis, 46, XY for affected case (Z5)

3.2 Results of fragile X PCR:

To determine the number of the CGG repeats, we used multiple primers around the expansion region of the FMR1 gene. We found that the PCR product size within the normal range (less than 45 CGG repeats) as shown in figure # 3, thus FMR1 repeat expansion cannot explain the ID features in the affected kids and other genetic causes should be explored.

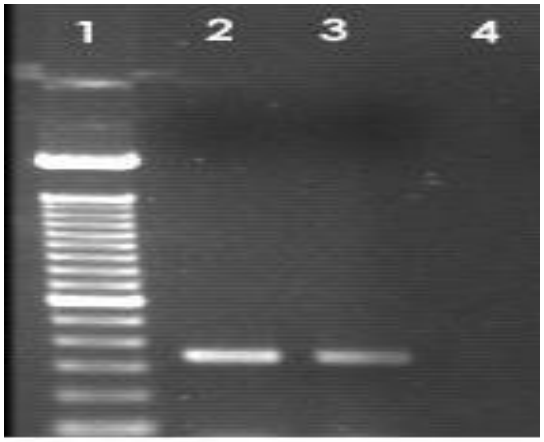

Figure 3: Agarose gel analysis for fragile X PCR product. Lanes 1, 2, 3 and 4 represent DNA molecular size marker (100-bp DNA ladder), AUT-Z3, Wild type DNA and non template control (pure water) respectively.

3.3 Results of Homozygosity mapping:

Eight homozygous regions were identified at chromosomes 1, 5, 7, 8, 10, 11, 18, 2 and were analyzed for any candidate genes as shown in the table below. One region at chromosome 5 have 2 candidate genes that are involved in the development of ID; those are HOMER1 and DHFR genes. Further analysis using whole exome sequencing exclude those genes and also did not identify any candidate gene that can be associated with the development of ID in the family.

Table 15: Homozygous regions identified using the GeneChip® Mapping 500K Assay

Chromosome	Block name	StartPos	EndPos	No. of SNPs	Size (bp)
5	chr5.4	76,400,618	84,004,601	635	7,603,984
7	chr7.2	14,305,196	19,342,398	583	5,037,203
21	chr21.1	9,764,385	14,775,729	7	5,011,345
18	chr18.3	15,079,294	18,611,243	4	3,531,950
8	chr8.3	123,470,712	126,649,545	236	3,178,834
10	chr10.2	45,868,889	48,367,621	25	2,498,733
1	chr1.5	147,814,694	150,023,407	12	2,208,714
11	chr11.4	71,286,539	73,413,121	63	2,126,583

5.4 Results of CytoScan microarray:

We used the Affymetrix® CytoScan™ Arrays (NspI Chips) to scan the genomes of the three ID cases to assess genomic imbalances that may be associated with the disease. A total of 122 genomic chromosomal CNVs were detected among the samples. Among these, 74 CNVs were gain (duplications) and 48 were loss (deletions). These CNVs were further analyzed for overlap with polymorphic CNVs in controls, and CNVs annotated in the database of genomic variation DGV. It was observed that among these variants, 2 CNVs were specific to ID cases, as they were not identified in any of the controls. The CNVs identified range from a few kilobases (Kb) to several hundred Kb in size, and some of them encompass functionally important candidate genes such as *IL1RAPL1*and *DLG2*. However, it is possible that these regions may contain some novel and uncharacterized genes or regulatory regions that may play a role in ID susceptibly. Two ID specific CNVs, a 14 Kb deletion at chromosome Xp21.3 and a 59.595 Kb deletion at chromosome 11q14.1 that involve 2 candidate genes. The microarray for AUT-Z3 case showed a significant deletion at chromosome X in the intronic region of IL1RAPL1 gene at the P arm location 21.3, in the position between 29,160,655-29,174,711 bp as showed in figure # 4. As well as a second deletion in the intronic region of DLG2 gene at chromosome 11 q arm location 14.1 in the region between 84,488,144-84,547,739 bp as showed in figure #5.

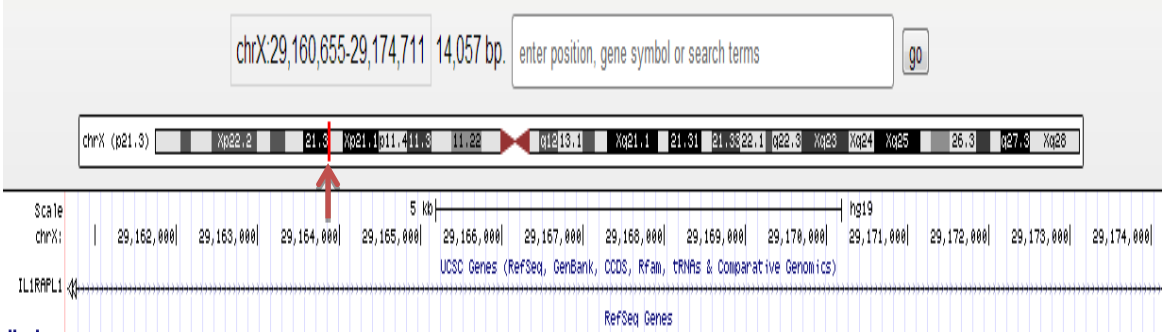

Figure 4: Genomic region showing a 14 Kb deletion at Xp21.3 which involves intron of the IL1RAPL1 **gene.**

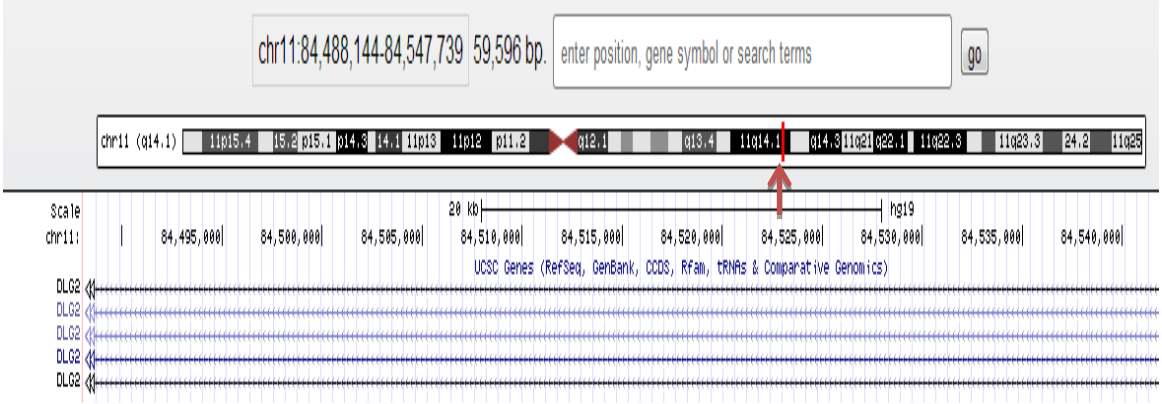

Figure 5: Genomic region showing a 59.595 Kb deletion at chr. 11q14.1 which involves intron 6 of the DLG2 **gene.**

3.5 Result of next generation sequencing

The next generation sequencing did not reveal any significant genes that can be identified as a candidate in the development of ID syndrome among our family.

CHAPTER 4

Discussion

In present study, we have provided an evidence for the existent of the CNVs on the 11 and X chromosome that might be associated with intellectual disability. The X chromosome is of particular interest for mapping genes involved in the etiology of ID. Using high density arrays, we identified one CNV in chromosome 11 and one in chromosome X in 3 ID affected siblings

After comparison with the controls, the CNVs regions were filtered out based on the following criteria; data base genomic variation, regions with pseudogenes, normal variations found among healthy Palestinian. We attempted to identify ID candidate genes by characterizing the disease associated CNV regions. One of our compelling findings was the identification of 2 CNVs directly involving previously known ID loci. At the beginning, we identified a 14 kb intergenic deletion of *IL1RAPL1*, a member of interleukin 1 receptor family, previously implicated in non-syndromic ID (65). The gene members of this interleukin family have been previously proposed in the development of autism and ID (66). Later studies further reporting *IL1RAPL1* mutations in ID patients. A study by Bhat et al reported a pericentromeric inversion resulting in disruption of *IL1RAPL1* gene in a patient with ID (67).

The deletion in this region were not validated since it need a long series of primer pairs that covers different region across the deletion knowing that the exact boundaries of the deletion were not identified.

We also identified a 59.595 Kb deletion spanning intron 6 of *DLG2* gene. The protein encoded by this gene is known as Disks large homolog 2 with alternative names known as Channel-associated protein of synapse-110 (Chapsyn-110) and Postsynaptic density protein (PSD-93). It is a member of the membrane-associated guanylate kinase (MAGUK) family. Membrane-associated guanylate kinases originally referred to a family of scaffold proteins highly concentrated at the sites of cell-cell junctions (68, 69). This family play essential roles in diverse cellular processes, including cell-cell

communication, cell polarity establishment and maintenance, and cellular signal transduction (70). MAGUK, now encompass a large family with very diverse biological functions and can be divided into several different classes (70). For example, DLG subfamily proteins are widely expressed in the brain and are essential for the formation and plasticity of glutamatergic synapses (71). DLG proteins consist of several modular protein interacting domains. The PDZ, Src homolog 3 (SH3), and guanylate kinase (GK) domains which are the characteristic protein domains shared in all MAGUK proteins.

DLG2 have been localized at central nervous system glutamatergic synapses (72) as well as at the cholinergic synapses (73) and is highly enriched in the post-synaptic density (PSD) (74). Indeed, mammalian brain is the tissue expressing the greatest diversity of MAGUK proteins. It has often been hypothesized that the most likely function of this subfamily of MAGUKs is as central organizers of vertebrate Central nervous system synapses. MAGUK protein family are expressed at different stages of postnatal development, DLG2 is highly expressed at later stages (75), thus PSD-93 are suggested to play roles in postsynaptic development in mammalian brain.

The encoded protein forms a heterodimer with a related family member that may interact at postsynaptic sites to form a multimeric scaffold for the clustering of receptors, ion channels, and associated signaling proteins. Multiple transcript variants encoding different isoforms have been found for this gene (genecards ID: GC11M083166).
Functional alteration at synaptic sites appears to be among the earliest events in the initiation of the cognitive decline that characterizes ID and can be considered the best pathological indicator of cognitive impairment in this pathology.

Recent studies analyzed the mechanisms involved in synaptic dysfunction in transgenic mice, a mouse model that appears to mimic early cognitive impairment in the early phases in ID (76). Interestingly, in these animals the reduced synaptic activity is strictly correlated to changes in composition of the DLG protein members of the MAGUK family. A similar modification in the synaptic levels of DLG proteins was found also in

another animal model for AD (77) suggesting a role for PSD-MAGUKs proteins in the earliest alterations of synaptic composition and activity in ID as well as AD.

DLG2 gene has been previously implicated in neorodevelopmental disorders such as schizophrenia. A study conducted by Trilochan et al. found that certain deletions in DLG2 is considered as candidate gene for neurodevelopmenta disorders (78). Soumya et al. similarly reported that rare deletions in DLG2 are involved in neuro-developmental diseases (79). Moreover, Walsh et al. highlighted this gene among other three genes as potentially pathogenic based on neurodevelopmental pathway analysis (80).

The *DLG2* gene is longest annotated transcript; (Chr11:82843701-85015962, NCBI Build 36.1 hg18) and is one of the genes that are among the largest in the human genome, with transcripts extending from 550 kb to 2.2 Mb of genome. All of these genes have been identified to be involved in the development of neurodevelopment diseases.

Thus, *DLG2* is another gene originally implicated in ID, and in our study we found a disease associated CNV also involving this gene. Based our findings, we proposed *DLG2* as an ID candidate gene. Further studies are needed to explore the phenotypic expression of such intronic deletion and its potential pathway that is involved in the development of the intellectual disability syndrome.

References

1. Leonard, H., and Wen, X., 2002. The epidemiology of mental retardation: challenges and opportunities in the new millennium. Mental Retardation and Developmental Disabilities Research Reviews, 2002; 8(3), p.117-134.

2. Durkin, M. The epidemiology of developmental disabilities in low-income countries. Mental Retardation and Developmental Disabilities Research Reviews, 2002; 8(3), p.206-211.

3. Leonard H,Wen X: The epidemiology of mental retardation: challenges and opportunities in the new millennium. Ment Retard Dev Disabil Res Rev 2002; 8: 117–34.

4. Salvador-Carulla, L., Rodríguez-Blázquez, C., and Martorell,A.. Intellectual disability: an approach from the health sciences perspective. Salud Pública de México, 2008; 50, p.s142-s150.

5. Ropers HH, Hamel BC. X-linked mental retardation. Nat Rev Genet. 2005; 6:46–57.

7. Ropers HH. X-linked mental retardation: many genes for a complex disorder. Curr Opin Genet Dev. 2006; 16:260–9.

8. Rauch A, Hoyer J, Guth S, et al. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. Am J Med Genet A. 2006;140:2063–74.

9. Kinsbourne,M., andW.D. Graf. Disorders of mental development. 2000; pp. 1155–1211 in *Child Neurology*.

10. Chelly J, Khelfaoui M, Francis F et al.: Genetics and pathophysiology of mental retardation. Eur J Hum Genet 2006; 14: 701–13.

11. Frints SG, Froyen G, Marynen P, Fryns JP: X-linked mental retardation: vanishing boundaries between non-specific (MRX) and syndromic (MRXS) forms. Clin Genet 2002; 62: 423– 432.

12. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, et al: Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 2008; 478:57-63.

13. McLaren J. and Bryson S.E. *American Journal of Mental Retardation*. 1987; 92, 243-254.

14. Larson S.A., Lakin K.C., Anderson L., Kwak N., Lee J.H., An-derson D. *American Journal of Mental Retardation*. 2001; 106, 231-52.

15. Murali M.S. *Indian Journal of Community Medicine*, 2001; 26, 198-200.

16. Leonard H. and Wen X. *Mental Retardation and Developmental Disability Research Reviews*, 2002; 8, 117-134.

17. Emerson E. *Mental Retardation and Developmental Disability Research Reviews*, 2007; 13, 107-113.

18. Knight, S.J.L., Regan, R., Nicod, A., Horsley, S.W., Kearney, L., Homfray, T., Winter, R.M., Bolton, P. and Flint, J. Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet*, 1999; 354, 1676-1681.

19. Genetic disorders among Palestinian Arabs: Effects of consanguinity. Zlotogora J. Am J Med Genet. 1997 Feb 11;68 (4):472-5

20. Zeldin A.S. *Mental Retardation: Medicine Neurology*. 2010.

21. Chelly J., Khelfaoui M., Francis F., Chérif B., Bienvenu T. *European Journal of Human Genetics*, 2006; 14, 701-713.

22. Francis F., Meyer G., Fallet-Bianco C., Moreno S., Kappeler C., Socorro A.C., Tuy F.P., Beldjord C., Chelly J. *European Journal of Neuroscience*, 2006; 23, 877-893.

23. Butler F.M., Miller S.P., Lee K.H., Pierce T. *Mental Retardation*, 2001; 39, 20-31.

24. Chiurazzi P, Oostra BA: Genetics of mental retardation. Curr Opin Pediatr 2000; 12: 529–35.

25. Curry, C.J., Stevenson, R.E., et al.. Evaluation of Mental Retardation: Recommendations of a Consensus Conference. *American Journal of Medical Genetics* 1997, 72, 468-477.

26. Holinski-Feder E., et al. Familiar mental retardation syndrome ATR-16 due to an inherited cryptic subtelomeric translocation t(3;16)(q29;p13.3). *America Journal of Human Genetics*, 2000; 66, 16-25.

27. Ropers, H.H. Genetics of early onset cognitive impairment. Annual Review of Genomics and Human Genetics, 2010; 11, p.161-187.

37. Lejeune J, Gautier M, Turpin R. CR Hebd Seances Acad Sci 1959 ;248:1721-2.

38. Koolen DA, Vissers LE, Pfundt R, et al. A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. Nat Genet 2006;38:999-1001.

39. Sagoo GS, Butterworth AS, Sanderson S, Shaw-Smith C, Higgins JP, Burton H. Array CGH in patients with learning disability and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects. Genet Med 2009;11:139-46.

40. Mefford HC, Eichler EE. Duplication hotspots, rare genomic disorders, and common disease. Curr Opin Genet Dev 2009; 19:196-204.

41. Koolen DA, Sharp AJ, Hurst JA, et al. Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. J Med Genet 2008; 45:710-20.

42. Mefford HC, Rosenfeld JA, Shur N, et al. Further clinical and molecular delineation of the 15q24 microdeletion syndrome. J Med Genet 2012; 49:110-8.

43. Mefford HC, Sharp AJ, Baker C, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. N Engl J Med 2008; 359:1685-99.

44. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. Nature 2008; 455: 232-6.

44. Shinawi M, Liu P, Kang SH, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet 2010; 47:332-41.

45. Inlow JK, Restifo LL: Molecular and comparative genetics of mental retardation. Genetics 2004; 166: 835– 881.

46. Ropers HH, Hamel BC: X-linked mental retardation. Nat Rev Genet 2005; 6: 46– 57.

47. Skuse DH: X-linked genes and mental functioning. Hum Mol Genet 2005; 14: R27–R32.

48. Bahi N, Friocourt G, Carrie A et al: IL1 receptor accessory protein like, a protein involved in X-linked mental retardation, interacts with Neuronal Calcium Sensor-1 and regulates exocytosis. Hum Mol Genet 2003; 12: 1415– 1425.

49. Koekkoek SK, Yamaguchi K, Milojkovic BA et al: Deletion of FMR1 in Purkinje cells enhances parallel fiber LTD, enlarges spines, and attenuates cerebellar eyelid conditioning in fragile X syndrome. Neuron 2005; 47: 339– 352.

50. Schenck A, Bardoni B, Langmann C, Harden N, Mandel JL, Giangrande A: CYFIP/Sra-1 controls neuronal connectivity in Drosophila and links the Rac1 GTPase pathway to the fragile X protein. Neuron 2003; 38: 887– 898.

51. Meng J, Meng Y, Hanna A, Janus C, Jia Z: Abnormal long-lasting synaptic plasticity and cognition in mice lacking the mental retardation gene Pak3. J Neurosci 2005; 25: 6641–6650.

52. Laumonnier F, Bonnet-Brilhault F, Gomot M et al: X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. Am J Hum Genet 2004; 74: 552– 557.

53. D'Adamo P, Menegon A, Lo Nigro C et al: Mutations in GDI1 are responsible for X-linked non-specific mental retardation. Nat Genet 1998; 19: 134– 139.

54. Bienvenu T, des Portes V, Saint Martin A et al: Non-specific Xlinked semidominant mental retardation by mutations in a Rab GDP-dissociation inhibitor. Hum Mol Genet 1998; 7: 1311– 1315.

55. Tarpey P, Parnau J, Blow M et al: Mutations in the DLG3 gene cause nonsyndromic X-linked mental retardation. Am J Hum Genet 2004; 75: 318– 324.

56. Roeleveld, N., G. A. Zielhuis and F. Gabreels,. The prevalence of mental retardation: a critical review of recent literature. Dev. Med. Child Neurol. 1997; 39: 125–132.

57. Larson S.A., Lakin K.C., Anderson L., Kwak N., Lee J.H., An-derson D. *American Journal of Mental Retardation*, 2001; 106, 231-52.

58. Murali M.S. *Indian Journal of Community Medicine*, 2001; 26, 198-200.

59. Chelly J, Khelfaoui M, Francis F et al.: Genetics and pathophysiology of mental retardation. Eur J Hum Genet 2006; 14: 701–13.

60. Kinsbourne,M., andW.D. Graf, Disorders ofmental development, 2000. 1155–1211 in *Child Neurology*.

61. Leonard H,Wen X: The epidemiology of mental retardation: challenges and opportunities in the new millennium. Ment Retard Dev Disabil Res Rev 2002; 8: 117–34.

62. Wang k, Inlow JK, Restifo LL: Molecular and comparative genetics of mental retardation. Genetics 2004; 166: 835– 881.

63. Meder B, Haas J, Keller A, Heid C, Just S, et al. Targeted next generation sequencing for the molecular genetic diagnostics of cardiomyopathies.
Circ Cardiovasc Genet 2011; 4: 110–122.

64. Neveling K, Collin RW, Gilissen C, van Huet RA, Visser L, et al. Next generation genetic testing for retinitis pigmentosa. Hum Mutat 2012; 33: 963–972

65. Tabolacci, E., Pomponi, M. G., Pietrobono, R., Terracciano, A., Chiurazzi, P., Neri, G. A truncating mutation in the IL1RAPL1 gene is responsible for X-linked mental retardation in the MRX21 family. Am. J. Med. Genet. 2006; 140A: 482-487.

66. Marshall et al., Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet. 2008 Feb;82(2):477-88.

67. Bhat SS, Ladd S, Grass F, Spence JE, Brasington CK, Simensen RJ, Schwartz CE, Dupont BR, Stevenson RE, Srivastava AK. Disruption of the IL1RAPL1 gene associated with a pericentromeric inversion of the X chromosome in a patient with mental retardation and autism. Clin Genet. 2008 Jan;73(1):94-6.

68. Woods D F, Bryant P J. ZO-1, DlgA and PSD-95/SAP90: homologous proteins in tight, septate and synaptic cell junctions. Mech Dev, 1993; 44(2-3): 85–89

69. Funke L, Dakoji S, Bredt D S. Membrane-associated guanylate kinases regulate adhesion and plasticity at cell junctions. Annu Rev Biochem, 2005; 74(1): 219–245

70. de Mendoza A, Suga H, Ruiz-Trillo I. Evolution of the MAGUK protein gene family in premetazoan lineages. BMC Evol Biol, 2010; 10(1): 93

71. Montgomery J M, Zamorano P L, Garner C C. MAGUKs in synapse assembly and function: an emerging view. Cell Mol Life Sci, 2004; 61(7–8): 911–929

72. Aoki C, Miko I, Oviedo H, Mikeladze-Dvali T, Alexandre L, Sweeney N, Bredt DS Electron microscopic immunocytochemical detection of PSD-95, PSD-93, SAP-102, and SAP-97 at postsynaptic, presynaptic, and nonsynaptic sites of adult and neonatal rat visual cortex. Synapse 2001; 40:239–257.

73. Conroy WG, Liu Z, Nai Q, Coggan JS, Berg DK PDZ-containing proteins provide a functional postsynaptic scaffold for nicotinic receptors in neurons. Neuron 2003; 38:759–771.

74. El-Husseini AE, Topinka JR, Lehrer-Graiwer JE, Firestein BL, Craven SE, Aoki C, Bredt DS Ion channel clustering by membraneassociated guanylate kinases. Differential regulation by N-terminal lipid and metal binding motifs. J Biol Chem 2000; 275:23904–23910.

75. Kim E, Sheng M PDZ domain proteins of synapses. Nat Rev Neurosci 2004; 5:771–781.

76. Dewachter I, Filipkowski RK, Priller C, Ris L, Neyton J, Croes S, Terwel D, Gysemans M, Devijver H, Borghgraef P, Godaux E, Kaczmarek L, Herms J, Van Leuven F. Deregulation of NMDA-receptor function and down-stream signaling in APP[V717I] transgenic mice. Neurobiol Aging. 2007: Jul 26; doi:1016

77. Almeida CG, Tampellini D, Takahashi RH, Greengard P, Lin MT, Snyder EM, Gouras GK Beta-amyloid accumulation in APP mutant neurons reduces PSD-95 and GluR1 in synapses. Neurobiol Dis 2005; 20:187–198.

78. Trilochan Sahooet al. Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. Genetics IN Medicine. 2011 Volume 13, Number 10.

79. Soumya Raychaudhuri. Accurately Assessing the Risk of Schizophrenia Conferred by Rare Copy-Number Variation Affecting Genes with Brain Function. PLoS Genetics. 2010,Volume 6, Issue 9.

80. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 2008, 320: 539–543.