TAMPERE UNIVERSITY OF TECHNOLOGY

PHILIPS, ANJU KARINGADAMPALLIL

EXOME SEQUENCING IN X-LINKED INTELLECTUAL DISABILITY: A NOVEL MUTATION IN THE *CUL4B*-GENE UNDERLYING CABEZAS SYNDROME

Master of Science Thesis

Examiner(s):Docent Irma Järvelä
Professor Matti Karp
Examiner(s) and topic was approved by the Faculty Council of the Faculty of Science and Bioengineering on 07.12.2011.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
International Master Degree Programme in Science and Bioengineering
PHILIPS, ANJU KARINGADAMPALLIL: Exome Sequencing in X-Linked Intellectual Disability: A Novel Mutation in the *CUL4B* gene underlying Cabezas syndrome.
Master of Science Thesis, 66 pages
February 2012
Major subject: Biotechnology
Examiner(s): Docent Irma Järvelä and Professor Matti Karp
Keywords: exome sequencing, gene, mutation, *CUL4B*, Cabezas syndrome

---

In this study we have applied exome sequencing of the X-chromosome in order to identify a mutation in a Finnish family with X–Linked Intellectual Disability (XLID). We identified a novel mutation in the Cullin 4B gene (*CUL4B)* that has previously shown to cause Cabezas syndrome. The mutation was identified using Agilent array that covers 93% of the coding region of chromosome X. The mutation is located in exon 20 resulting in premature stop codon in exon 21 where aspartic acid is changed to a premature stop codon D806X. Here we present a detailed clinical phenotype of the three affected brothers. *CUL4B* is a ubiquitin E3 ligase subunit implicated in the regulation of several biological processes, and *CUL4B* is the first XLID gene that encodes an E3 ubiquitin ligase. Our findings elucidate the functional significance of *CUL4B* in human cognition and in other aspects of human development.

PREFACE

This Master's Thesis has been written as a partial fulfilment for the International Master of Science Degree at Tampere University of Technology. The work has been carried out at the Department of Medical Genetics, University of Helsinki during the period May 2011 to February 2012.

This work was financially supported by the Sigrid Jusélius Foundation.

I would like to express my sincere gratitude to my supervisor associate professor Irma Järvelä for giving me this opportunity to work in this project and examine my thesis. I would like to thank my Professor Matti Karp for his valuable support and advice.

I would like to thank my wonderful colleagues Liisa Ukkola-Vuoti, Katri Kantojärvi, Fatma Doagu, Chakri Kanduri, Minna Ahvenainen and Jaana Oikkonen for always cheering me up and motivating me and creating an inspirational working atmosphere. I am grateful to Fatma Doagu and Minna Ahvenainen for their skillful guidance in several lab methods. I would like to thank Chakri Kanduri for his valuable comments and advices. I thank my loving sister Smitha Ancy Cherian for her constant motivation.

This work is dedicated to my parents Philipose Thomas and Achiamma Philipose, with out them this work would never have been possible.

Finally, all families who have participated in this study are most gratefully acknowledged.

Table of Contents

1. INTRODUCTION 1
2. THEORETICAL BACKGROUND 2
2.1 Definition of Intellectual Disability 2
2.2 Prevalence of Intellectual Disability 4
2.3 X-Linked Intellectual Disability 4
2.4 X Chromosome 5
2.5 Classification of XLID 6
2.6 Etiology of Intellectual Disability 7
2.6.1 Environmental and sociocultural factors 7
2.6.2 Chromosomal aberrations and ID 7
2.6.3 Copy number variants in Intellectual Disability 9
2.6.4 Single base pair mutation 10
2.7 Identification of XLID genes 14
2.7.1 Positional Cloning & Linkage analysis 16
2.7.2 High Throughput genomic DNA sequencing 17
2.7.3 Array Technology 18
2.7.4 Cytogenetic methods 20
2.7.5 Next generation sequencing 20
2.8 Molecular Basis of Intellectual Disability 24
3. AIMS OF THE PRESENT STUDY 27
4. MATERIALS AND METHODS 28
4.1 The family material 28
4.2 DNA isolation 29
4.3 Exome Sequencing 29
4.3.1 Exon enrichment and high throughput sequencing 29
4.3.2 Sequence coverage of targeted exons 29
4.3.3 Calling of Single Nucleotides Polymorphisms 29
4.3.4 Filtering out polymorphisms and selection of disease causing variants 29
4.4 PCR –Sequencing 31
4.5 Transcript Analysis 31
4.5.1 RNA Isolation 31
4.5.2 RNA Integrity Check 32
4.5.3 cDNA synthesis 32
4.5.4 Gel Purification 33
4.6 Bioinformatics Analysis 33

5. RESULTS 34
5.1 Clinical Results 34
5.2 Sequencing results and mutation analysis 36
6. DISCUSSION 42
7. CONCLUSIONS 46
8. REFERENCES 49

ABBREVIATIONS

AAIDD	American Association on Intellectual Disabilities
BAC	Bacterial Artificial Chromosome
CNV	Copy Number Variation
ESE	Exonic splicing enhancers
FISH	Fluorescent In Situ Hybridization
HGMD	Human Gene Mutation Database
IQ	Intelligent Quotient
ISS	Intronic Splicing Enhancer
MB	Megabase
NS-XLID	Non Syndromic X – Linked Intellectual Disability
NGS	Next Generation Sequencing
SNP	Single Nucleotide Polymorphism
S-XLID	Syndromic X – Linked Intellectual Disability
XLID	X-Linked Intellectual Disability
XLMR	X-Linked Mental Retardation
WHO	World Health Organization

1. INTRODUCTION

Intellectual disability (ID) can be defined as a significantly reduced ability to interpret new or complex information, and to learn new skills with a reduced ability to cope independently. It is characterized by significant limitations both in intellectual functioning as well as in adaptive behaviour with onset before the age of 18 years. [Luckasson et al. 2002]. The prevalence of intellectual disability in developed countries is between 1% and 3% and is responsible for 5 to 10 % of health care expenditure in some developed countries [Ropers. 2008].

X-linked Intellectual Disability (XLID) is a clinically complex and heterogeneous disorder, [Gecz et al. 2009] and mutations in more than 90 genes X chromosomal genes have been found to be associated with this disease [Tarpey et al. 2009; Ropers. 2010]. X-linked gene defects are thought to be the important causes of intellectual disability and they account for roughly 10 -15 % of the intellectual disability in males [Hamel et al. 2008]. Intellectual disability can be an outcome of genetic as well as environmental causes that act on the development and functioning of the CNS prenatally, perinatally or postnatally [Chiurazzi et al. 2008]. Genetic causes can be due to large deletions, large duplications or aneuploidies that affect multiple genes. It can also be due to mutations of individual genes leading to autosomal dominant, autosomal recessive or X linked genetic disease. [Inlow et al. 2004].

Among intellectually disabled patients, an excess of males has been observed and is explained by the fact that the presence of many genes responsible for ID are on the X chromosome than on the autosomes [Lehrke R. 1972].

X-linked intellectual disability is divided into syndromic and non syndromic forms. In syndromic forms ID is expressed with a specific pattern of physical, neurological, and/or metabolic abnormalities. In non-syndromic X- linked intellectual disability male patients have no consistent phenotypic manifestations other than intellectual disability. [Renieri et al. 2005].

2. THEORETICAL BACKGROUND

2.1 Definition of Intellectual Disability

Intellectual disability is a disability until recently referred to as mental retardation is characterized by the significant limitations both in intellectual functioning and adaptive behaviour as expressed in conceptual, social and practical adaptive skills. This disability originates before the age of 18. [Schalock et al. 2007].

According to the tenth revision of the WHO (World Health Organization), Intellectual Disability is a disorder defined by the presence of incomplete or arrested mental development, principally characterized by the deterioration of concrete functions at each stage of development that contributes to the overall level of intelligence such as cognitive, language, motor and socialization functions where adaptation to the environment is always affected.

According to the American Association on Intellectual Disabilities (AAIDD), there are five assumptions that are essential for the application of the definition of intellectual disability:

1. Limitations in present functioning must be considered within the context of community environments typical of the individual’s age peers and culture.
2. Valid assessment considers cultural and linguistic diversity as well as differences in communication, sensory, motor, and behavioral factors.
3. Within an individual, limitations often coexist.
4. An important purpose of describing limitations to develop a profile of needed supports.
5. With appropriate personalized supports over a sustained period, the life functioning of the person with intellectual disability generally will improve.

[AAIDD, http://www.aaidd.org/]

In contrast, the WPA Section on Psychiatry of Intellectual Disability considers IDD to be a health condition: a syndromic grouping or meta-syndrome analogous to the construct of dementia, which is characterized by a deficit in cognitive functioning prior to the acquisition of skills through learning. The intensity of the deficit is such that it inter-

feres in a significant way with individual normal functioning as expressed in limitations in activities and restriction in participation (disabilities).[Carulla et al. 2008.]

Table 1: LIMITATIONS ASSOCIATED WITH INTELLECTUAL DISABILITY

1. Communication
2. Personal Care
3. Home Life
4. Social Skills
5. Utilization of the Community
6. Health and Safety
7. Functional academic skills
8. Leisure Time
9. Work

* Based on American Association on Intellectual and Developmental Disabilities (AAIDD).

Clinically ID is described by three basic criteria:
a) an intelligence quotient (IQ) below 70
b) limitation in two or more adaptive behaviors such as communication, self care, social skills, community use, self direction, health and safety
c) evidence that the mental manifestations began before the age of 18.
Four degrees of severity of intellectual impairment can be categorised: mild, moderate, severe, and profound. (Refer Table 2)
Mild: About 50 % of persons fall into this category. People with this level of ID need support on as needed basis, episodic or short term. IQ ranges from 50-55 till 70.
Moderate: This group constitutes about 10 % of the entire population of people with ID. The IQ level usually varies from 35-49 till 50-55. They need support consistent with time but the amount of time is limited.

Severe: This group constitutes of 3% - 4% of individuals with ID. People in this group need regular, consistent and lifetime support. Regular support is needed in at least one such aspect such as school, work or home. The IQ level ranges from 20-25 till 35- 40.
Profound: This group approximately constitutes 1%- 2% of people with intellectual disability. They possess little or no ability to care for their own basic needs and require constant help and supervision. IQ level is under 20-25.

Table 2: Classification of Intellectual Disability

Security Level (Percentage of Individu-als with ID)	Intelligent Quotient Range	Support needed in daily living activities
Mild (50%)	From 50-55 till 70	Intermittent
Moderate (10%)	From 35-49 till 50-55	Limited
Severe (4%)	From 20-25 till 35-40	Extensive
Profound (1%)	Less than 20-25	Pervasive

Based on American Psychiatric Association, DSM-IV TR, 2000; American Association of Intellectual and Developmental Disabilities, 2002.

2.2 Prevalence of Intellectual Disability

The prevalence of ID is estimated to affect approximately 2 % to 3 % of the population. X linked gene defects are considered to be responsible for approximately 10 % of the ID found in males (Ropers et al. 2005). Their prevalence is around 1% in high income countries and 2 % in low and middle income countries [Maulik et al. 2011; Durkin. 2002].

2.3 X-Linked Intellectual Disability

X-linked mental retardation is the proportion of mental retardation showing the distinctive pattern of inheritance associated with the X chromosome. The XLID phenotype in a family can be detected when the transmission follows the general characteristics of X-linked recessive phenotypes:

- only males are affected
- all daughters of affected males will be carriers

- carrier females may have mild expression of the disease
- sons of carrier females have a 50% risk of being affected
- daughters of carrier females have an 50% risk of being carriers
- male to male transmission does not occur
- unaffected males cannot transmit the phenotype

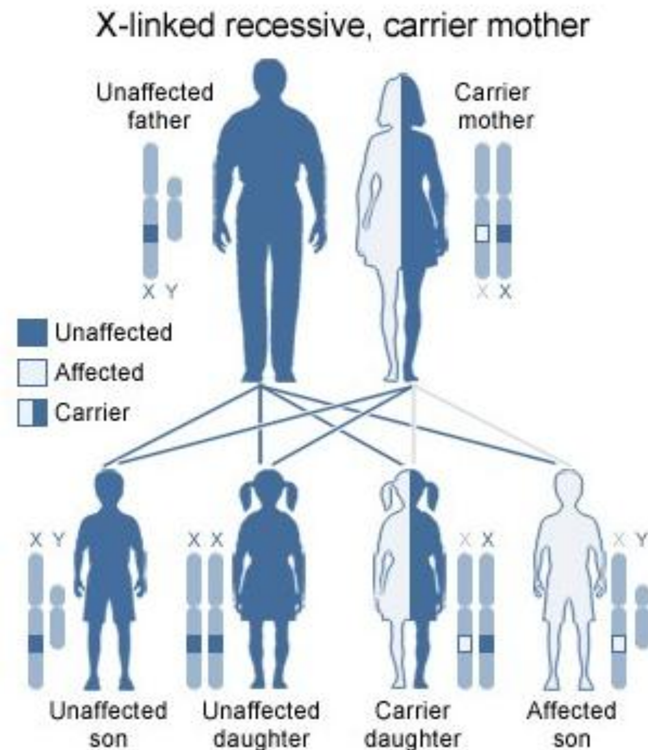


Figure 1. X chromosomal pedigree adapated from Human Molecular Genetics by Stratchan and Andrew P Read, Wiley Liss (1999).

2.4 X Chromosome

X chromosome has many attributes that are exclusive in the human genome. The size of the X chromosome is roughly 155MB and contains a total of 1860 genes (Figure 2). The X chromosome holds a special place in the field of Medical Genetics. Although the X chromosome contains only 4 % of the genes, it is accountable for almost 10 % of the diseases with the Mendelian Inheritance. X chromosome inactivation in mammals attains dosage compensation between males and females for X linked gene products. Females inherit an X chromosome from each parent but males inherit only a single mater-

nal X chromosome. [Ross et al. 2005]. Intellectual disability is one among the common problems in clinical genetics and it affects more males than females. Many genes on the human X chromosome get away from X-inactivation, or at least partly. For most of these, there is no functionally comparable homologue on the Y chromosome, which itself explains the gender –specific distinction accountable for the susceptibility to certain diseases. [Ropers. 2006].

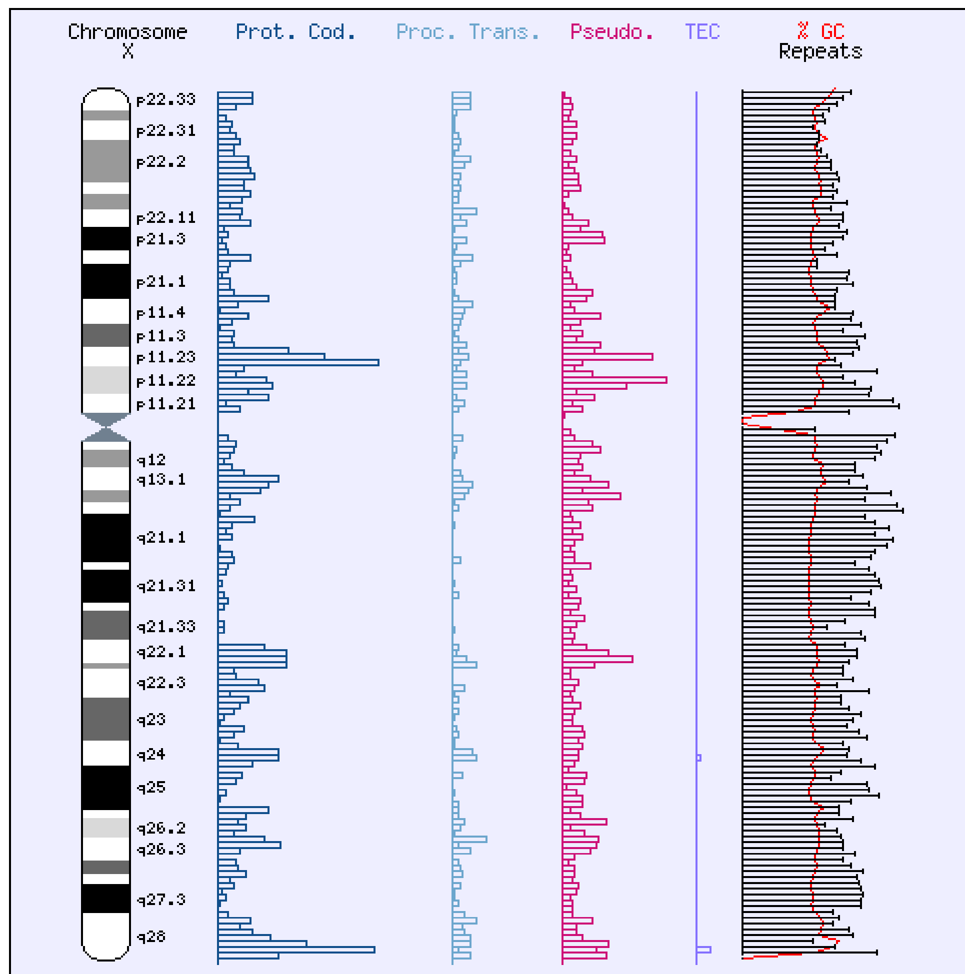


Figure 2. The Composition of X chromosome. Adapted from http://vega.sanger.ac.uk/Homo_sapiens/Location/Chromosome?r=X

2.5 Classification of XLID

XLID is usually categorized into the "syndromic" and "non syndromic" forms. In the syndromic forms, Intellectual disability is present in association with specific pattern of physical, neurological and/or metabolic abnormalities. Until now out of 215 total condi-

tions 98 genes related to syndromic ID have been identified (http://xlmr.interfree.it/home.htm). The term non specific XLID was coined by Kerr et al. in 1991 to illustrate the condition segregating in an X linked manner in which the male patients do not demonstrate any specific phenotypic manifestations other than the Intellectual Disability [Kerr et al. 1991; Mulley et al. 1992]. Till date 66 genes related to non-specific XLID out of the total 215 genes have been identified.
(http://xlmr.interfree.it/home.htm). ID can be an outcome of genetic as well as environmental factors that are responsible for affecting the development and normal functioning of the nervous system prenatally, perinatally and postnatally. It can also be due to various environmental factors mainly malnutrition in pregnancy, pre- and postnatal infections, exposure to neurotoxic compounds, premature birth and peri- and postnatal asphyxia or other trauma. The genetic causes of ID are mainly recognized in severe intellectual disability and they involve chromosome aneusomies, chromosomal structural abnormalities, genomic disorders and monogenic disorders. Genetic factors account for roughly 65 % of moderate to severe ID [Chelly et al. 2006; Chiurazzi et al. 2008; Patel et al. 2010].

2.6 Etiology of Intellectual Disability

2.6.1 Environmental and sociocultural factors

Epidemiological studies have repeatedly revealed that an evident link exists between the poverty and intellectual disability. It is clear that this notable link reflects two distinct processes. The first one indicating that a connection between poverty and exposure to a wide range of environmental and psychosocial factors exist and the second one reveals that families with members who suffer from intellectual disability have an increased risk of catastrophic expenses that considerably affect poverty levels. These factors are considered to be straight forward reasons of intellectual disability in developing countries. [Leonard et al. 2005].

2.6.2 Chromosomal aberrations and ID

Chromosomal abnormalities occur in 6% of all recognized congenital malformations, The most common ones listed in Table 3.

Table 3: Chromosomal Abnormalities Associated with Intellectual Disability

SYNDROME	PREVALENCE	CHROMOSOMAL ABNORMALITY
Down syndrome	1 in 1000 live births	Approximately 94 % of the cases are caused by trisomy 21, 3,5% by translocation, & 2,5 % by mosaicism.
Turner's syndrome	1 in 2000 to 1 in 5000 females	Complete or partial absence of one X chromosome (45, XO).
Klinefelter's syndrome	1 in 1000 males	Maternal nondisjunction (47, XXY)
Prader-Willi syndrome	1.2 to 1.3 per 10 000	Microdeletion on chromosome 15, paternal origin at the locus 15q11-13
Angelman's syndrome	1 in 20 000 to 1 in 30 000	Microdeletion on chromosome 15, maternal origin at the locus 15q11-13
Cri-Du chat syndrome	1 in 50 000	deletion 5p
Di-George syndrome	1 in 5000	microdeletion of chromosome 7q11
Smith Magenis syndrome	1 in 25 000	microdeletion 17p11
Rubinstein– Taybi syndrome	1 in 12 500 live births	autosomal dominant, microdeletion 16p 13.3
William's syndrome	1 in 5 500 live births	microdeletion of chromosome 7q11

*Adapted from Ahuja et al. 2005.

These account for 30% to 40% of severe intellectual disability and 10 % of mild intellectual disability. [Raynham et al. 1996; Ahuja et al. 2005]. Chromosomal abnormalities include deletions, microdeletions, translocations, inversions and duplications. Sex chromosome anomalies occur in roughly 1 in 400 live births. These anomalies are commonly due to chromosomal nondisjunction, the risk of which increases with maternal age. The most common include Turner's syndrome (45, XO), Klinefelter's syn-

drome (47, XXY) and the 47, XXX and 47, XYY karyotypes [Ahuja et al. 2005]. It is certainly true that individuals with chromosomal aneuploidy demonstrate some nonspecific features in common such as poor growth, microcephaly, epicanthic folds and unusual palmar creases, in addition to features more specific to the chromosomes involved. Down syndrome (trisomy 21) still remains the most important cause of intellectual disability. Cytogenetically visible chromosomal aberrations are found almost in one out every seven individuals with severe cognitive impairment. [Leonard et al. 2002].

2.6.3 Copy number variants in Intellectual Disability

The term "copy number variation" can referred to as an intermediate-scale genetic change, operationally defined as segments greater than 1,000 base pairs in length but typically less than 5 megabases, which is the cytogenetic level of resolution. CNVs include additional copies of sequences (duplications) and losses of genetic material (deletions). [Eichler et al. 2008]. Unlike the SNPs the CNVs span larger regions which might affect one or more genes by decreasing or increasing the gene dosage, revealing a recessive allele, causing aberrant expression or even might lead to alternative spliced or fusion genes. Rare CNVs have been found to cause neurological disorders such as mental retardation, schizophrenia and autism [Redon et al. 2006; Marshall et al. 2008]. Studies to identify copy number variants (CNVs) on the X-chromosome have revealed novel genes important in the causation of X-linked Intellectual disability (XLID). Froyen and colleagues screened a cohort of 108 subjects with ID by X chromosome array CGH and identified CNVs in 14 subjects (13%). The most common XLID associated chromosomal aberrations reported are the duplications of Xq28 comprising the MECP2 gene, which have been found in more than 100 cognitively impaired individuals with characteristic facial features, hypotonia, seizures, speech delay and recurrent infections [Van Esch et al. 2005; Friez et al. 2006; Madrigal et al. 2007]. Approximately 10-15% of cases in intellectual disability in males have been reported due to copy number variations on the X chromosome [Fab et al. 2007; Koolen et al. 2009]. Until now, very little has been known about the parental origin and the effect of increased parental age on the genesis of genomic copy number variation (CNV), including those which underlie a significant percentage of patients with cognitive disorder. Studies on a specific class of *de novo* CNVs—namely, those associated with recurrent microdeletion and microduplication syndromes—have implied out that there is no significant bias in the parent of

origin. The majority of rare *de novo* CNVs associated with intellectual disability (ID), however, has non-recurrent breakpoints and does not generally involve known syndromes [Koolen et al. 2009]. Numerous of CNVs have been found to be associated with specific phenotypes, and in patients with ID, recurrent interstitial microdeletions and duplications involving at least 19 genomic intervals have been identified [Vissers et al. 2009]. For most of the CNVs associated in intellectual disability, reliable frequency estimates have not been determined to the extreme heterogeneity of ID and relatively limited size of the cohort analyzed so far. In one of the largest study reported so far, Mefford et al. (2008) reported on the frequency of recurrent CNVs at 69 out 130 putative genomic hot spots for NAHR in 1010 children with unexplained ID and also in 2493 previously screened healthy adults. Pathogenic and possibly pathogenic CNVs were identified in 5.4% of the ID cohort [Itsara et al. 2009]. These studies reveal that CNVs are the basic cause of many specific forms of ID and various other disorders. In addition, numerous CNVs (at 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11 and 22q11.2) have been found to the major predisposing factors for wide spectrum of neuropsychiatric disorders, including intellectual disability, autism and epilepsy [Mefford et al. 2008; Stefansson et al. 2008; Weiss et al. 2008].

2.6.4 Single base pair mutation

A single base substitution can be defined where a single nucleotide is replaced by another nucleotide. These single base changes are also referred to as point mutations. They are the most frequent type of alterations in DNA. For example if a purine (a, t) is replaced by a purine and pyrimidine by a pyrimidine (c, g) then this kind of substitution is referred to as transversion. Single base substitutions are broadly categorized into four types mainly – missense mutations, nonsense mutations, silent mutations and splice site mutations.

Missense mutation: In a missense mutation, the base alters the codon which results in a different amino acid being incorporated into the protein chain (Figure 3).

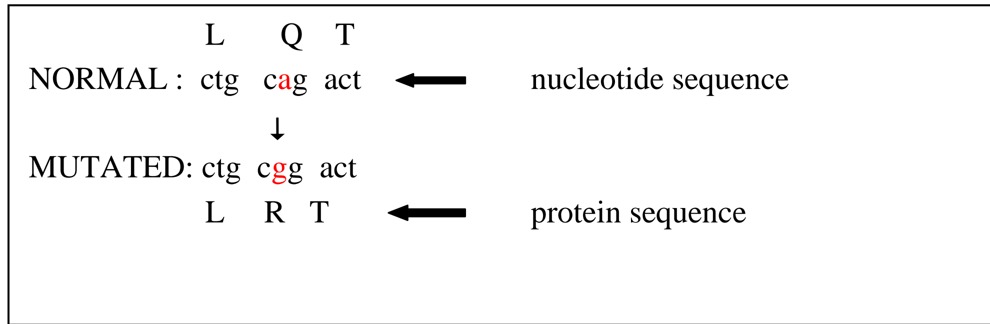


Figure 3. Diagrammatic representation to illustrate a missense mutation.

Lets take an example to illustrate missense mutations. In the above figure, the substitution of “a” (highlighted in red) in the second codon to “g” (highlighted in red) leads to an amino acid substitution of glutamine (Q) to arginine R.

Nonsense mutation: In a nonsense mutation, the new base change in a codon cause one of the stop codons (taa, tag, tga). This will cause translation of mRNA to stop prematurely and a truncated protein is produced (Figure 4).

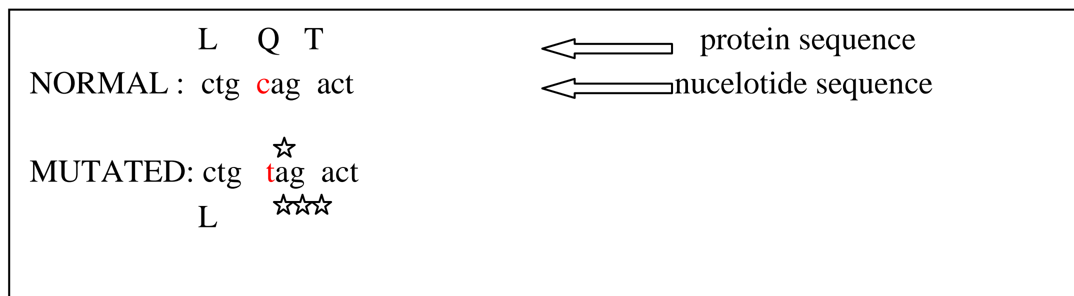


Figure 4. To illustrate nonsense mutations

In the above figure the second codon “c” is changing to “t”. In the normal nucleotide sequence the second codon “cag” codes for glutamine (Q) while in the mutated second codon “tag” codes for STOP leading to a premature termination of the protein.

Silent mutation: Silent mutations are those that do not have any alteration in the final protein product and can only be identified by sequencing the gene. These do not have any deleterious effect as mutated codons will still be coding the same amino acid.

In a recent study, exon screening of 86 known XLID genes yielded probably pathogenic mutations in no more than seven out of 21 families tested [Hao et al. 2010]. Recent studies in a multigenerational German family revealed and confirmed the presence of novel GDI1 mutations with nonsyndromic intellectual disability. Their study revealed a frameshift mutation in GDI1 that co-segregated with the disease [Strobl-Wildemann et al. 2011]. Large scale next generation resequencing of the X chromosome genes have helped in identification of a missense mutation in the CLIC2 gene on Xq28 in a male with X linked Intellectual Disability [Takano et al. 2011]. Mutations in the Jumonji AT –rich interactive domain 1C (JARID1C/SMCX/KDM5C) gene, located at Xp11.22 are emerging as frequent causes of X- linked intellectual disability [Fintelman et al. 2011].

Splice site mutation: Changes in splicing can lead to the disease directly, modify the severity of the disease phenotype or be linked with disease susceptibility. One large scale analysis of human genetic variation revealed that 51 (1.3%) of the SNPs identified within and around the exons of 313 genes studied were found to be within the consensus splice sites [Stephens et al. 2001]. Accurate pre –mRNA splicing needs exon-intron boundaries to be correctly recognized by the nuclear ribonucleoprotein complex known as the splicesome, for the introns to be correctly excised and also for the exons to be joined perfectly to produce mature mRNA [Hastings et al. 2001]. This process required the recognition of a variety of different motifs, including the fairly degenerate consensus sequences flanking the GT and AG dinucleotides at the 5' donor and 3' acceptor splice sites and also the branch site which almost some 15-35 bases upstream of the 3' splice site. In the Human Gene Mutation Database [Stenson et al. 2003], single base pair substitutions within the splice site constitute around 9.5% of all the mutations causing human inherited diseases. Splice site mutations may result in exon skipping, activation of the cryptic splice sites, creation of the pseudo-exon within an intron, or intron retention [Nakai et al. 1994]. Exon skipping is generally considered as the most common effect, and usually thought to be caused by the failure of the normal and mutant splice sites to define an exon. Most of the cryptic mutations initiate splice sites of the same type and are located within the few hundred nucleotides of the natural site [Hawkins 1988]. Also splice site mutations in very short or terminal introns have been found to result in intron retention [Dominski et al. 1991]. Usually mutations that are far away from the natural splice sites create cryptic sites that are activated in the presence of a nearby cryptic splice site of opposite polarity leading to production of a novel non-

coding exon with the intron. Cryptic acceptor splice sites are more frequent in exons than in introns mainly due to depletion of AG dinucleotides upstream of the original acceptor sites [Christensen et al. 2005]. Intron retention is the least studied type of alternative and aberrant splicing. Compared to the other types of alternative splicing which involves the choice between different splice sites, intron retention represents complete absence of splicing. Retained introns have been found to be shorter than the constitutively spliced out ones and also exhibited the tendency to occur in 5' and 3' untranslated regions [Sakabe et al. 2007].

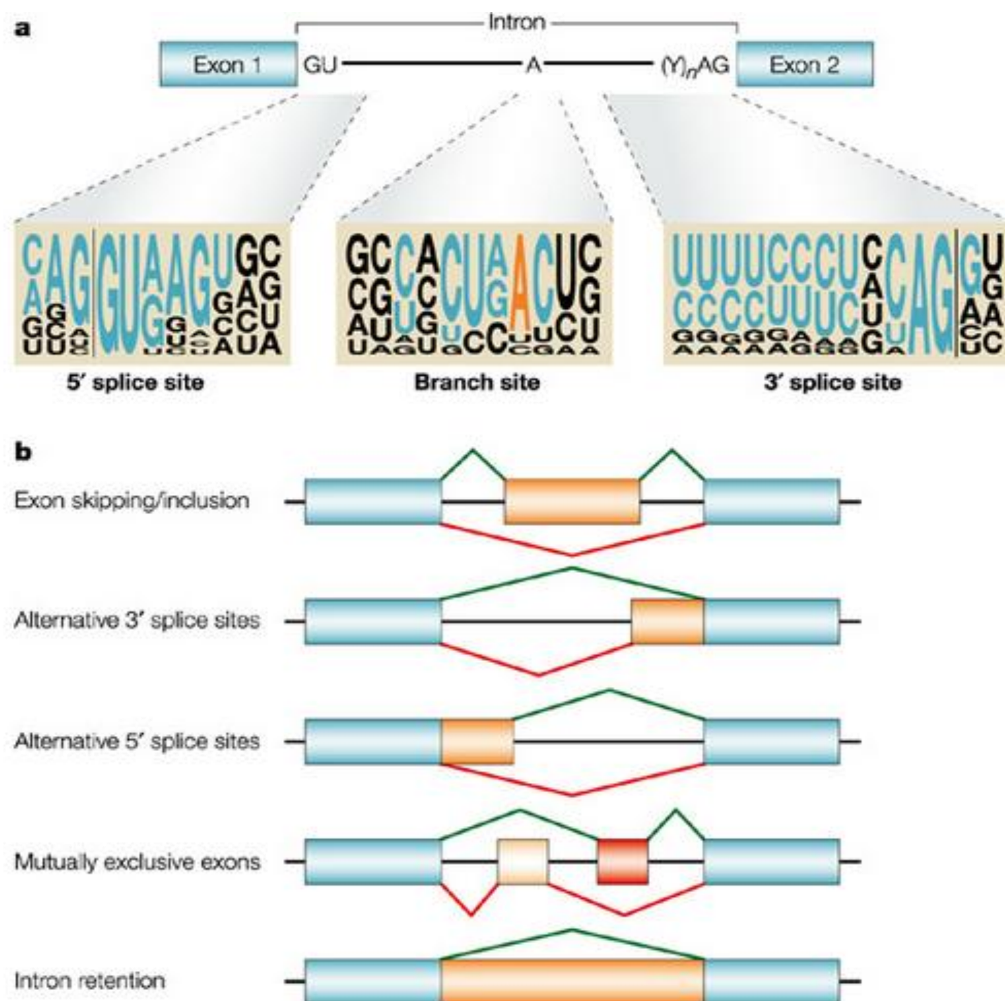


Figure 5. Classical splicing signal and modes of alternative splicing. A) Conserved motifs at or near the intron ends. The nearly invariant GU and AG dinucleotides at the intron ends, the polypyrimidine tract preceding the 3' AG, and the A residue that serves as a branch point. B) Five common modes of alternative splicing. Adapted from Luca Cartegni et al. 2002.

Sequence alterations occurring in exons or introns may affect the correct processivity of the mRNA by disrupting the splice site, exonic splicing enhancers (ESE), intronic

splicing enhancers (ISE) or altering the secondary structure of the mRNA. In case of nonsense mutations suspected of affecting splicing, along with disruption of possible ESE and ESS, two other possible mechanisms are also involved. The first mechanism is referred to as nonsense mediated decay (NMD). In this case when a pretermination codon (PTC) is present the entire mRNA is degraded. Nonsense altered splicing is acknowledged to be more controversial. In this case a translational like machinery scans the reading frame and surveys its integrity before splicing. It is still not clear how frequently sequence variations involving splicing are involved in the onset of disease. In a survey conducted, where the mutations considered consisted only of those directly affecting the standard consensus splice sites, 15% of mutations have been found to result in a human genetic disease through RNA splicing defects.[Krawczak et al. 1992].

2.7 Identification of XLID genes

Searching for genetic effects that are accountable for syndromic forms of XLID is not different from the various gene hunting procedures which we use for any monogenic condition. However finding molecular causes for NS-XLID has been really very challenging due to its high level of genetic heterogeneity, which includes pooling of linkage information from unrelated families and all this greatlymakes it difficult to search for the various mutations.

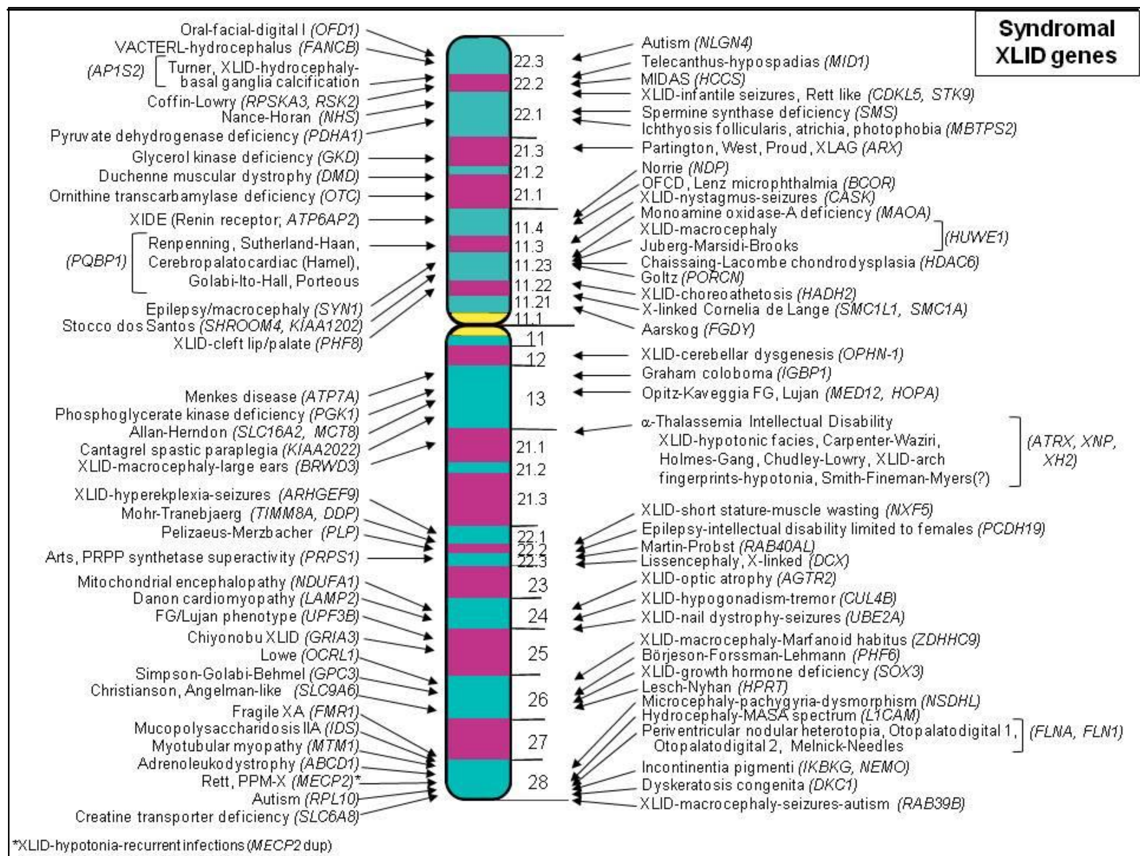


Figure 6. Ideogram of the X chromosome with the position of the syndromal XLID genes. Adapted from http://www.ggc.org/research/molecular-studies/xlid.html .

Since 2000, the rate of identification of novel genes that cause intellectual disability has adavanced rapidly due to various technological advancements. Prior to this, most of the gene identification relied on linkage and positional cloning methods. In each large family, the candidate interval for disease causing gene was found by identifying which parts of the X chromosome was shared by all the affected males and also not shared by unaffected males in the family. Progress was later achieved by improved DNA sequencing methods and the coordination of two large international studies to identify the causative genes, EURO-MRX and IGOLD. Hence to overcome these hurdles and to study the molecular basis of NS-XLID, the European XLMR Consortium was founded in 1995 and since then various members and the associated groups have made important contributions to the identification of about 90 XLID genes identified so far (http://xlmr.interfree.it) (Figure 6 and 7). The search for XLID genes is still an ongoing process. There are various strategies for the identification of the various XLID genes.

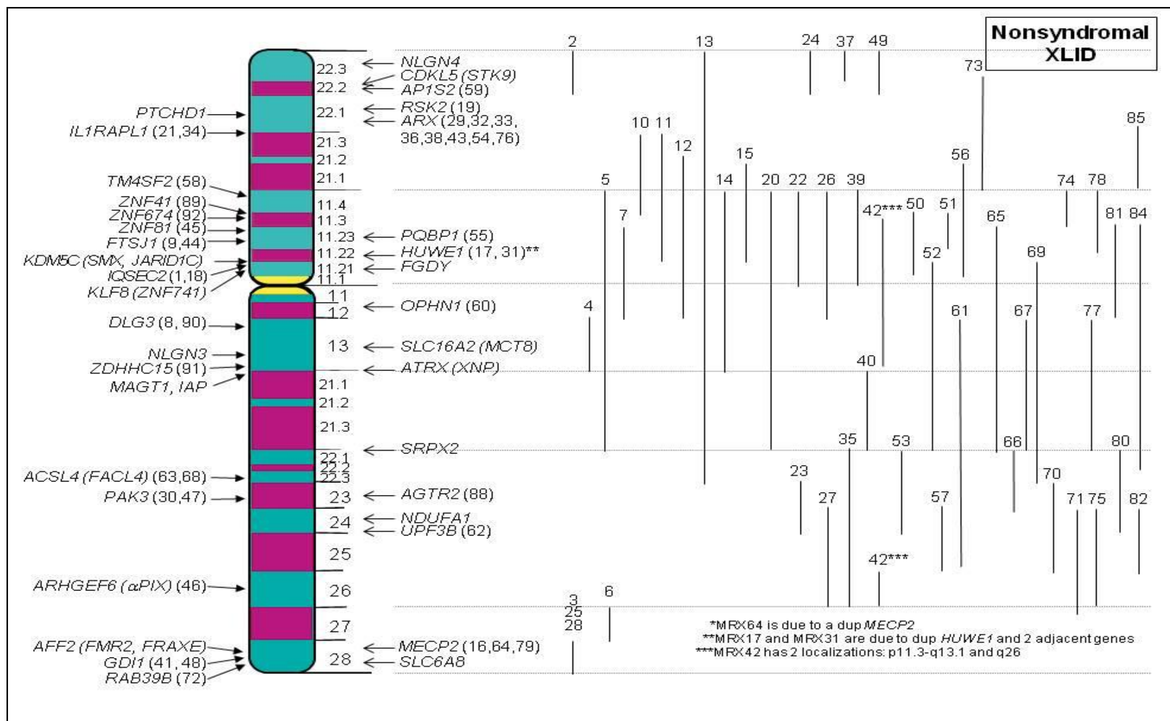


Figure 7. Ideogram of the X chromosome with the position of the nonsyndromal XLID genes. Adapted from http://www.ggc.org/research/molecular-studies/xlid.html .

2.7.1 Positional Cloning & Linkage analysis

Positional candidate gene studies are usually used to identify the disease genes for many human genetic diseases, and these studies involve genome wide linkage analysis to identify the approximate chromosomal location of a disease gene, fine structure genetic mapping to delineate and narrow the chromosomal interval in which the disease gene might be located, and also physical mapping and gene identification in the genetically defined interval to clone the disease gene.[Boehnke M.1994].

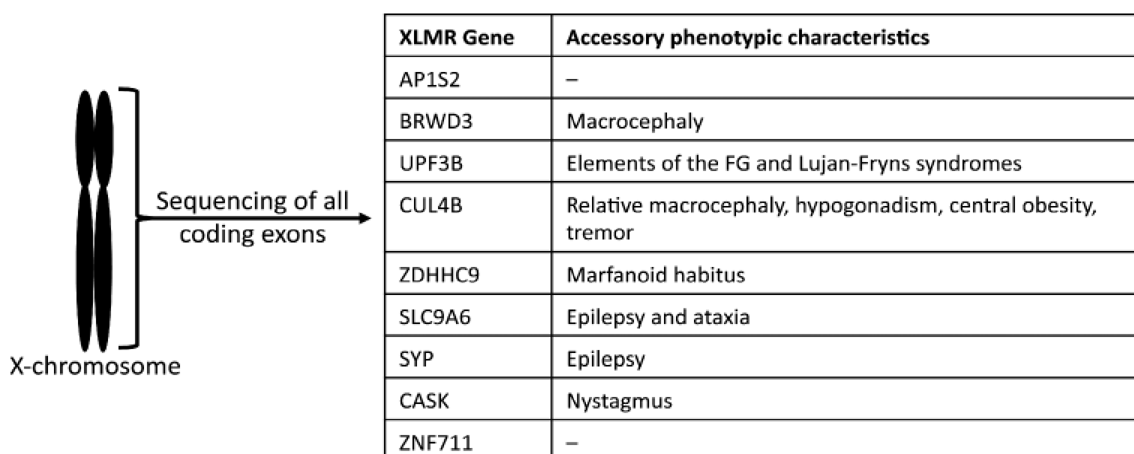
Positional cloning is based on the chromosomal localization of the gene. There is no previous knowledge of the biological function of the gene product. The overall strategy of positional cloning is to map the location of a human disease gene by linkage analysis and then to use the mapped location on the chromosome to clone the gene [Ballabio 1993; Pierce 2003]. Positional cloning of genes responsible for XLID has been based on the investigation of X; autosome balanced translocations, deletion mapping or molecular screening of candidate genes [Castellvi-Bel and Milà 2001]. The disease genes isolated by positional cloning usually fall into two main categories: (i) genes for relatively

common diseases with a large availability of pedigrees and samples for genetic mapping and that are the subject of intense research; (ii) genes representing “easy” targets because of the presence of patients with readily visible cytogenetic abnormalities [Ballabio. 1993]. The identification of XLID genes is greatly facilitated by the presence of cytogenetic aberrations, like translocations, inversions and deletions. Translocations in mentally retarded patients provide an excellent opportunity to clone the X-chromosomal gene disrupted by the translocation. Males with microdeletions on the X chromosome are also informative and are often detected by the presence of a contiguous gene syndrome, in which X-linked disease phenotype is combined with intellectual disability [Chelly and Mandel 2001].

For example *Oligophrenin-1 (OPHN1)* was identified after characterizing an X;12 balanced translocation in a female patient with mild intellectual disability [Billuart et al. 1998]. Also *TM4SF2* was identified by positional cloning after characterizing an X;2 balanced translocation in a female patient with intellectual disability [Zemni et al. 2000]. In addition *ARHGEF6* was identified by a balanced X;21 translocation [Kutsche et al. 2000]. *IL1RAPL* gene was discovered with overlapping microdeletions in two XL-ID families [Carrié et al. 1999].

2.7.2 High Throughput genomic DNA sequencing

The most extensive approach would be to perform a high throughput genomic DNA resequencing of all the genes on the X chromosome (Figure 8).



XLMR Gene	Accessory phenotypic characteristics
AP1S2	–
BRWD3	Macrocephaly
UPF3B	Elements of the FG and Lujan-Fryns syndromes
CUL4B	Relative macrocephaly, hypogonadism, central obesity, tremor
ZDHHC9	Marfanoid habitus
SLC9A6	Epilepsy and ataxia
SYP	Epilepsy
CASK	Nystagmus
ZNF711	–

Figure 8. Diagram to illustrate systematic resequencing of X chromosome coding exons as an approach to identify novel X linked intellectual disability genes. (Figure adapted from Sanders, 2010)

A direct systematic sequencing has been a new approach whereby all coding exons of the X chromosome are fully sequenced in the male XLID or in a female obligate carrier. And from this screen a catalogue of X chromosome sequence variants are generated, some of which will be causative.

Various variants are further analyzed to determine whether they are disease causing. Large-scale systematic resequencing has been anticipated as the future strategy for the discovery of rare, disease-causing sequence variants across the range of various human complex diseases. Tarpey et al (2009) have sequenced the coding exons of 718 X chromosome genes in 208 families with X-linked Intellectual Disability, which has been the largest direct screen which is reported so far. The average coverage of the genes reported here was 75%. 1,858 different coding sequence variants were detected, around 1,769 from the X specific and 89 from the pseudoautosomal X chromosome regions. Also 1,814 single nucleotide changes were reported out of which 980 were causing missense amino acid substitutions, 22 were responsible for causing nonsense termination codons, 13 were abnormalities at highly conserved bases at splice acceptor and donor sites and 799 were causing silent changes. Three variants were missense double nucleotides substitutions, and 41 variants detected small insertions and deletions of which 26 were in frame and 15 were causing translational frameshifts. This dataset has helped in discovering nine genes implicated in XLID, namely *AP1S2*, *BRWD3*, *UPF3B*, *CUL4B*, *ZDHHC9*, *SLC9A6,* SYP, *ZNF711* and *CASK* reported. Also this study emphasized the challenges which are faced in whole genome sequencing screens, specially the loss of function of 1 % or more of X chromosome genes.

2.7.3 Array Technology

Comparative genomic hybridization, CGH is a molecular cytogenetic method for the detection of chromosomal imbalances, which does not depend on the availability of chromosome spreads and is not limited to the analysis of growing cells [Solinas et al. 1997; Pinkel et al. 1998]. In conventional CGH, two differentially labeled genomes, a test and control, are competitively hybridized to metaphase chromosomes. Regions of gain or loss of DNA sequences, such as deletions, duplications, or amplifications, are seen as changes in the ratio of the intensities of the two fluorochromes along the target chromosomes [Kallioniemi et al. 1992]. However, since the DNA in metaphase chromosomes is condensed, the resolution is limited to 5-10 Mb. Thus, there is little reason

to apply CGH to chromosomes of metaphase cells in routine diagnostic settings [Lichter et al. 2000].

Recently, the drawbacks of conventional chromosomal CGH have been overcome by the introduction of array CGH. Here differentially labeled test and reference DNA are co-hybridized onto microarrays. Metaphase chromosomes are replaced with slides arrayed with complementary DNA (cDNA) oligonucleotides or genomic BAC clones as the targets for analysis. Chromosomal microarrays allow the genome wide identification of submicroscopic chromosomal abnormalities at a very high resolution. The first array-CGH experiments in patients with ID relied on the use of homemade bacterial artificial chromosome (PAC) arrays and had a diagnostic yield of about 10% [Vissers et al. 2003]. It does not require an expert clinician to suspect a specific diagnosis and it may cover the entire genome or targets known pathologic loci in a unique test. This new technique has been successful in revealing submicroscopic chromosome aberrations in patients with intellectual disability with normal results from prior cytogenetic analysis with detection rate up to 5- 20% [Liang et al. 2008]. The recently developed array-CGH technique combines the property of a complete genome scan of CGH along with the hybridization on sorted genomic DNA fragments from the microarray technology. In addition to that array CGH has been able to detect and quantify segmental aneuplody with a resolution comparable to that of FISH. Array-CGH, using DNA from uncultured cells, has been efficient in detection of low-level mosaicism, which could be missed by conventional cytogenetic analysis [Ballif et al. 2006]. Large scale array CGH based studies using the Affymetrix 500K oligonuleotide array and the tiling path bacterial artificial chromosomes (BAC) array have identified copy number variations in 12% of the human genome [Zhang et al. 2007]. Froyen and collegues have developed a full coverage X chromosome tiling array CGH for the detection of copy number alterations in patients with suspected XLID [Froyen et al. 2007]. Complementary DNA microarray technology has been developed in recent years to perform large scale quantitative analysis of gene expression at the transcript level. Zhang and collegues have recently managed to use a custom human X chromosome cDNA microarray to identify the candidate genes responsible for XLID [Zhang et al. 2007]. However this approach is limited by requiring expression of potential candidate genes in accessible tissues and also by its inability to identify candidate genes with mutations that do not alter transcript levels. Microarray technology was first used first time in a microarray – based copy number analysis of all human telomeres in patients with Intellectual Disability [Veltman et al.

2002]. The power of microarray technology is being fully utilized for unbiased whole-genome copy number analysis [Koolen et al. 2008].

2.7.4 Cytogenetic methods

Karyotype analysis, biopsy of amniocytes or chorionic villi, is the most descriptive pre-natal diagnostic tool to detect chromosome abnormalities. However, this method is time consuming. To reduce the level of distress among parents, efforts have been made to develop a very quick diagnostic test for specific chromosome abnormalities. Trisomies of chromosomes 13, 18 or 21 (Patau, Edwards or Down syndrome respectively) and sex chromosome aneuploidies (e.g. Turner syndrome) are the most characteristic chromosome abnormalities detected at amniocentesis [Ryan et al. 2005].

2.7.5 Next generation sequencing

Recent advances in the next generation sequencing technologies have helped the researchers to explore both rare and common disorders. While the whole genome sequencing still remains quite costly for most of the applications, exome sequencing is a technique which is the limelight and focuses on the protein coding portion of the genome only. There are currently three major exome enrichment platforms: Agilent Sure Select Human All Exon 50Mb, Roche/Nimblegen's SeqCap EZ Exome Library v2.0 and Illumina's TreSeq Exome Enrichment. [Michael et al. 2011]. The exome consists of all the exons of a genome that are transcribed into a mature RNA. Protein coding genes constitute about 1% of the human genome but harbour 85% of the mutations with affect on disease related traits .[Bamshad et al. 2011]. Exome sequencing is used synchronously with two sampling strategies: family –based phenotypes (to exploit parent child transmission patterns) and extreme phenotypes. In families where multiple individuals are affected with a common trait, one methodology will be to sequence the most distally related individuals.

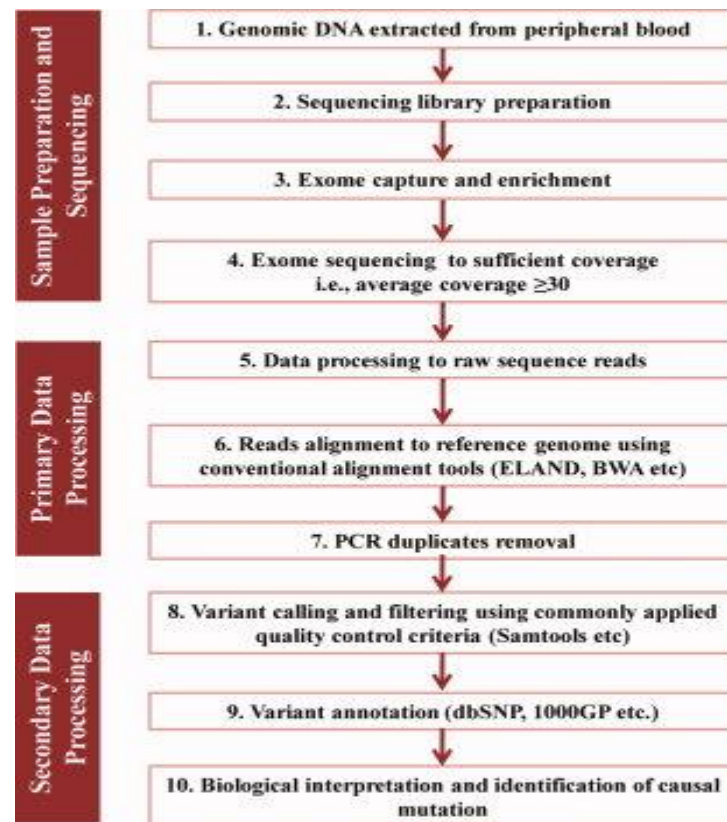


Figure 9. To Illustrate the Workflow of Exome sequencing from genomic DNA extraction to biological interpretation and the identification of a casual mutation. The workflow is categorized into 3 main steps a) Sample Preparation and Sequencing, b) Primary Data Processing, c) Secondary Data Processing. Figure directly adapted from Chee- Seng Ku et al. 2012.

Another alternative would be the family based approach which is useful in identifying *de novo* variants, it involves sequencing parent-offspring trios in which only the offspring is affected. Exome sequencing of a parent –child trios have been found to be an effective concept for identifying de novo coding mutations. This approach is applicable to gene discovery in disorders where most of the cases are sporadic and also when a dominant mode of inheritance is speculated [Hoischen et al. 2010]. Most of the Mendelian disorders are caused by the exonic mutations or splice site mutations that change the amino acids sequence of the affected gene. Exome sequencing is a very powerful tool for the discovery of Mendelian diseases in situations where the conventional approaches have failed [Biesecker et al. 2010]. Exome sequencing for the first time has been successful for the discovery of a novel mutation underlying an autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13 [Caliskann et al. 2010]. A notable accomplishment of the whole exome sequencing has been in

discovering that *de novo* single nucleotide variants contributes to the intellectual disability [Vissers et al. 2010] and to children with sporadic autism [O'Roak et al. 2011]. It has been affirmed that the power of whole exome sequencing plays a very substantial role in identifying the genetic basis of various human diseases. In a recent study the largest published so far deep sequencing has revealed disease causing variants in 50 novel genes for recessive cognitive disorders, it has also assisted in revealing additional mutations in 23 genes earlier implicated in intellectual disability or connected neurological disorders. [Najmabadi et al. 2011].

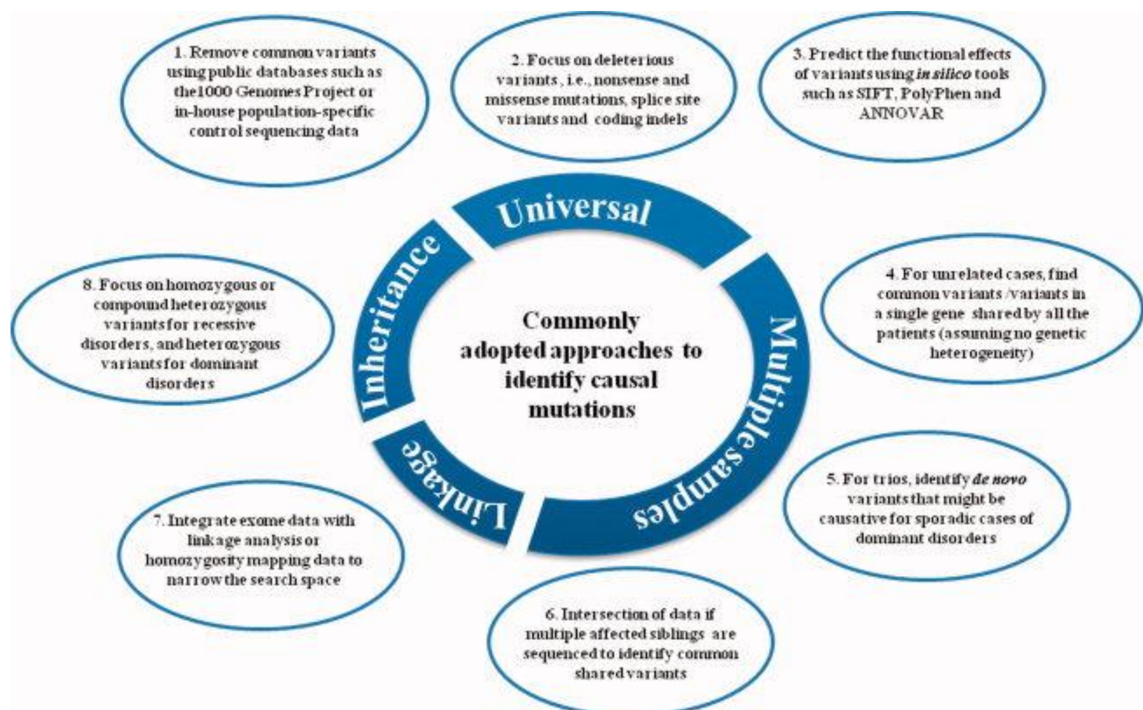


Figure10. Commonly adopted approaches to identify causal mutations. The 3 main criterias which are used to filter the less likely causal variants are 1) removing common variants, 2) focusing on deleterious variants, and 3) predicting and retaining variants with functional effects. Adapted from Chee- Seng Ku et al. 2012.

By targeting the specific regions of interest selective DNA enrichment improves the overall cost and efficiency of the NGS [Rehman et al. 2010; Volpi et al. 2010]. Targeted sequencing focuses on all protein coding subsequences (the functional exome), which requires roughly 5% as much sequencing compared to that what is needed for the entire human genomes [Pussegoda 2010; Senapathy et al. 2010]. This present scheme helps in reducing the overall cost for the sequencing a single individual. Commercially available products for targeted sequence –enrichment includes Agilent's SureSelect and Nimble-

gen's ASeqCap, Illumina's TruSeq. However the parallel short –read strategy of NGS opens to many hurdles for the bioinformatics to comprehend the short reads and the genetic alterations in human genomes [Myers et al. 2008]. The assets of the NGS can only be utilized when bioinformaticians are able to decipher and make use of the short read sequences, including alignment, assembly [Salzberg et al. 2008]. NGS technologies will certainly allow us to identify all the causative variants including the "rare variants" within the individual. It is also strongly anticipated that the whole genome sequencing or the exome sequencing will make influential impact to our understanding of the genetic etiologies that contribute to the complicated human disease as well as the genetic basis of genomics.

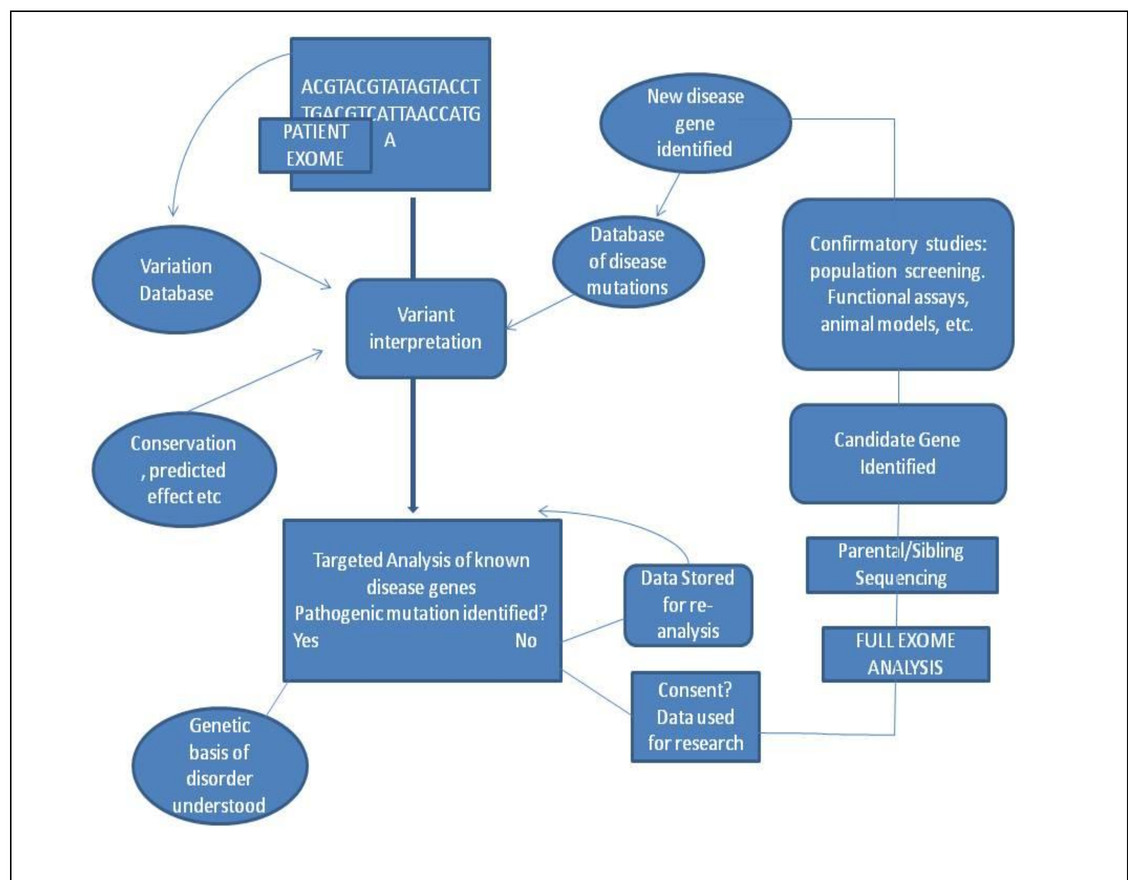


Figure 11. Integrated diagnostic/research information workflow. A patient presents with an idiopathic intellectual disability and her exome is sequenced in diagnostic laboratory. Variants are automatically annotated with respect to population frequency, evolutionary conservation, predicted effect on transcript expression or splicing, effect on protein function, and are checked against the databases of proven disease causing mutations. To make this process faster, all the variant data are also deposited into a local database of genetic variation. The laboratory then undertakes a targeted analysis of the set of genes known to be potentially causative of her condition, and information about the mutation in those genes are reported back again to the clinicians. In cases where no causative mutation is found, her sequence data are stored

and periodically reanalyzed. Full analysis of a patient's exome along with the sequencing of parents or siblings' exomes may result in the identification of new potential candidate genes. (Figure modified from Topper et al. 2011)

2.8 Molecular Basis of Intellectual Disability

The intricacy of the genes accountable for intellectual disability can be understood in terms of modules of several genes acting together in a single pathway or complex, resulting in an equivalent phenotypes when mutated. Functional correlations have helped in identification of approximately 450 genes that have been found to be associated in ID [Inlow et al. 2004]. Several general molecular and cellular mechanisms underlying the pathophysiology of intellectual disability can be identified, including neurogenesis, neuronal migration, synaptic functions, and transcription and translation [Chelly et al. 2006]. Disturbed neurogenesis is commonly noticed in intellectual disability disorders comprising microcephaly, and these conditions seem to have a common origin in defective centrosome function and DNA repair response pathways [Kiaindl et al. 2010]. Larger groups of intellectual disorders refequently involve sharing biological functions in synapse formation and plasticity, cellular signaling, and transcriptional regulation. During the development and until adolescence, the brain is vulnerable to broad structuring of neuronal connectivity by the formation and elimination of synapses. The dynamic regulation of synaptic connectivity is very critical for the various aspects of learning, memory, and cognitive functions in the adult brain. Synapses and spines are highly active in their outlook and can undergo rapid structural changes in response to stimuli. This property is attributed to as the synaptic plasticity, and is believed to be involved in learning and memory [Lamprecht et al. 2004]. Functional analysis of the normal and the disrupted synaptic functions of intellectual disability associated proteins have aided us in learning about these processes. Several different mechanisms involving ID genes care responsible for the impaired synaptic plasticity and therefore they are known to affect cognitive function. Local regulation of protein levels in the PSD is a substantial mechanism in the control of synaptic plasticity. Another level of regulation of the PSD proteins is through the ubiquitin –mediated protein turnover. A large number of ID proteins have been found to be precisely involved in UPS- mediated protein degradation, including UBE3A, UBE2A, HUWE1, CUL4B, and UBR1. Thus it reveals that tight regulation of postsynaptic protein levels is crucial for normal learning and memory processes. Dendritic spine morphology is highly plastic in nature and changes of shape

or size of spines can change within seconds. The dynamic morphology of spines is a result of continous polymerization and breakdown of actin filaments and dynamic microtubules in the spine [Jaworski et al. 2009].

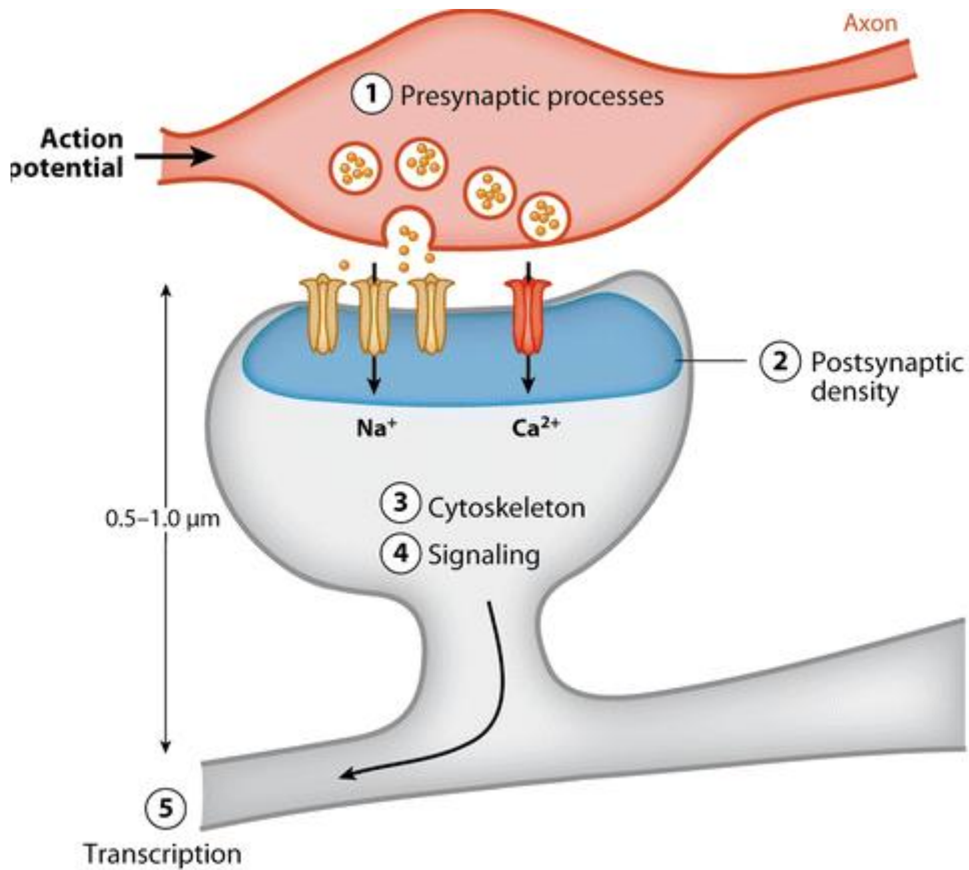


Figure 12: Mechanism how intellectual disability proteins act on the synaptic membranes. The neurotransmitter content is relased after fusion into the synaptic cleft activating neurotransmitter receptors at the postsynaptic cell membrane and leading to the opening of ion channels followed by generation of a postsynaptic potential (Figure adapted from van Bokhoven at al. 2011.)

There is increasing evidence that breakdown of signaling pathways both in excitatory glutamatergic neurons and in inhibitory GABAergic neurons contribute to the cognitive impairment and behavioral anaomalies in ID and ASD [Malinow et al. 2002]

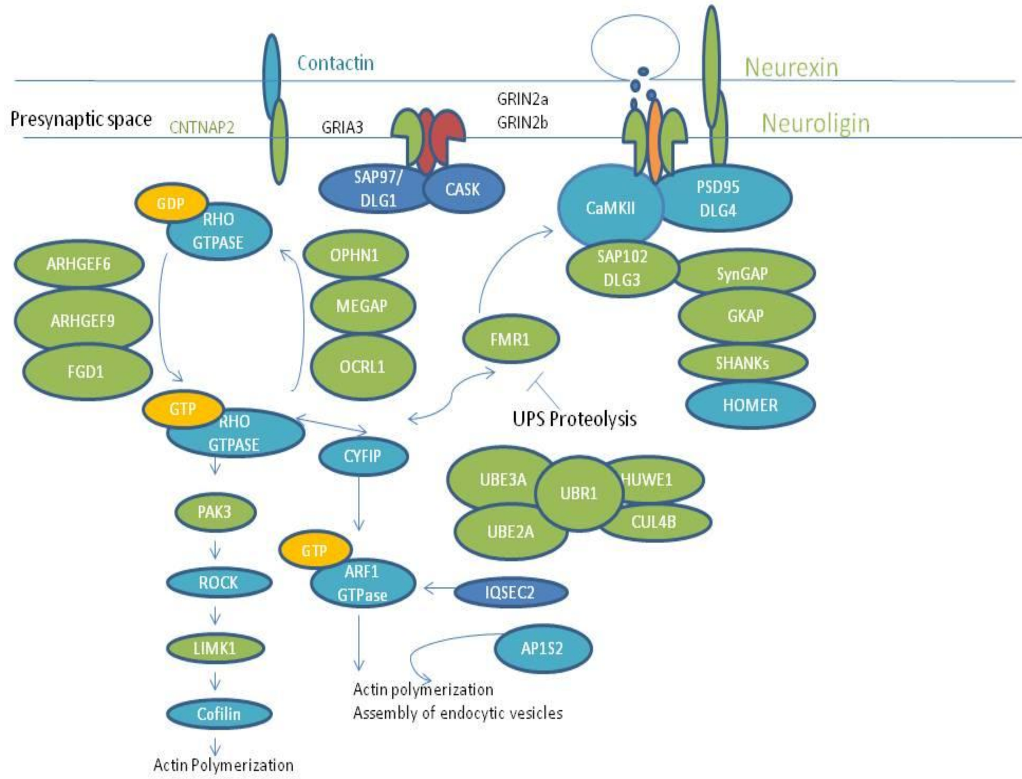


Figure13. Postsynaptic protein networks and pathways involving intellectual disability (ID) proteins. Figure modified from van Bokhoven et al. 2011.

3. AIMS OF THE PRESENT STUDY

1. To identify a mutation in the Finnish family with intellectual disability of unknown cause.
2. To investigate the functional effect of the mutation.

4. MATERIALS AND METHODS

4.1 The family material

The family pedigree is shown in Figure 14. The family belongs to the material consisting of seven families with at least three affected males whose diagnosis has remained unknown since 1980s (data not shown). In 2004 a total of 88 multiallelic microsatellite markers on the X-chromosome were analyzed in these families. In one family a novel mutation in *PAK3* was identified [Peippo M et al., 2007]. Using X-chromosomal CGH-array a duplication in hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 genes on Xp11.2 was identified in another family [Froyen G et al. 2008]. Due to long common haplotypes that contained numerous genes in the remaining five families the identification of the causative gene was too expensive and time-consuming and was not done. The Ethics Committee of the Helsinki university hospital approved the study. Informed written consent was obtained from all the participating individuals or their parents.

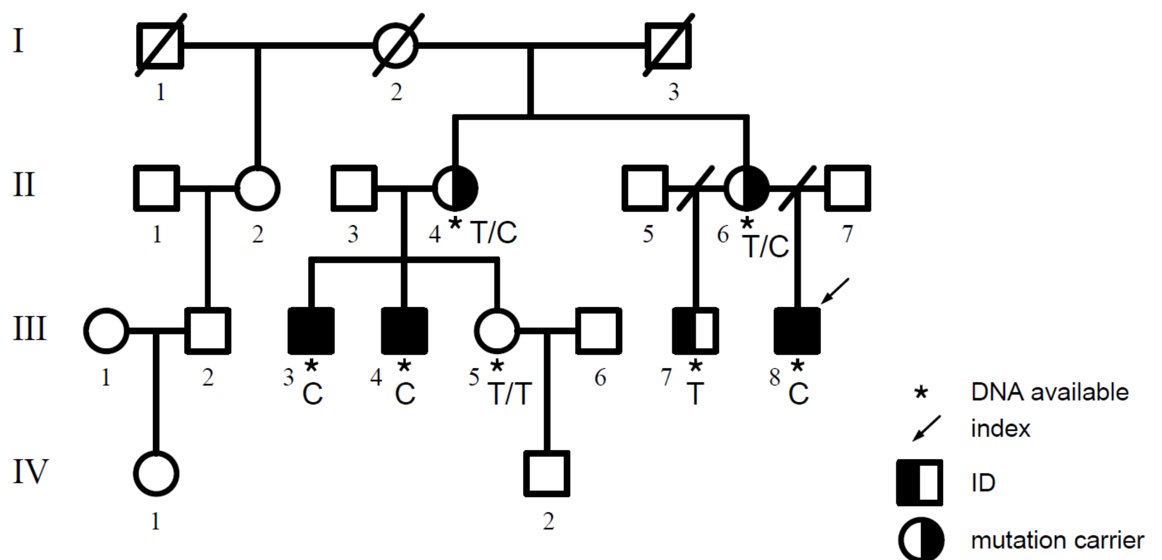


Figure 14. The family pedigree under study.

4.2 DNA isolation

10 ml whole blood was collected in the EDTA vials from the individuals participating in the study. Lymphocytes in the samples are lysed using Igepal CA-630 Sigma to release the nucleus from the cell and a DNA stabilizing agent was added to the sample. The samples were then centrifuged to pellet the nuclei after which the pellet was washed using TKM1 buffer. SDS-detergent was added to dissolve lipids and release the DNA from the nuclei into the solution. The remaining proteins are precipitated and removed using a high concentration salt solution after which the DNA is precipitated using absolute ethanol. The DNA was finally dissolved and stored in TRIS-EDTA and the concentration and the purity of the sample was measured using the NanoDrop.

4.3 Exome Sequencing

4.3.1 Exon enrichment and high throughput sequencing

Exons from the homozygous intervals were enriched with custom-made Agilent Sure Select DNA capture arrays including an average of 60 bp of flanking sequence on either side of the exon and sequenced on an Illumina Genome Analyzer II yielding 76 bp single reads.

4.3.2 Sequence coverage of targeted exons

More than 98% of the targeted exons were covered by at least four non redundant sequence reads, each with a PHRED like quality score of 20 or above.

4.3.3 Calling of Single Nucleotides Polymorphisms

To detect the single nucleotide polymorphisms high quality reads were aligned to the human reference genome by SOAP2.20 with default settings.

4.3.4 Filtering out polymorphisms and selection of disease causing variants

Also to eliminate previously reported, non pathogenic changes, all sequence variants were filtered against dbSNP. In addition to it the OMIM catalogue /http://www.ncbi.nlm.nih.gov/omim) and the Human Gene Mutation Database (HGMD,

http://www.hgmd.org/) were used as filter to identify all previously described pathogenic changes.

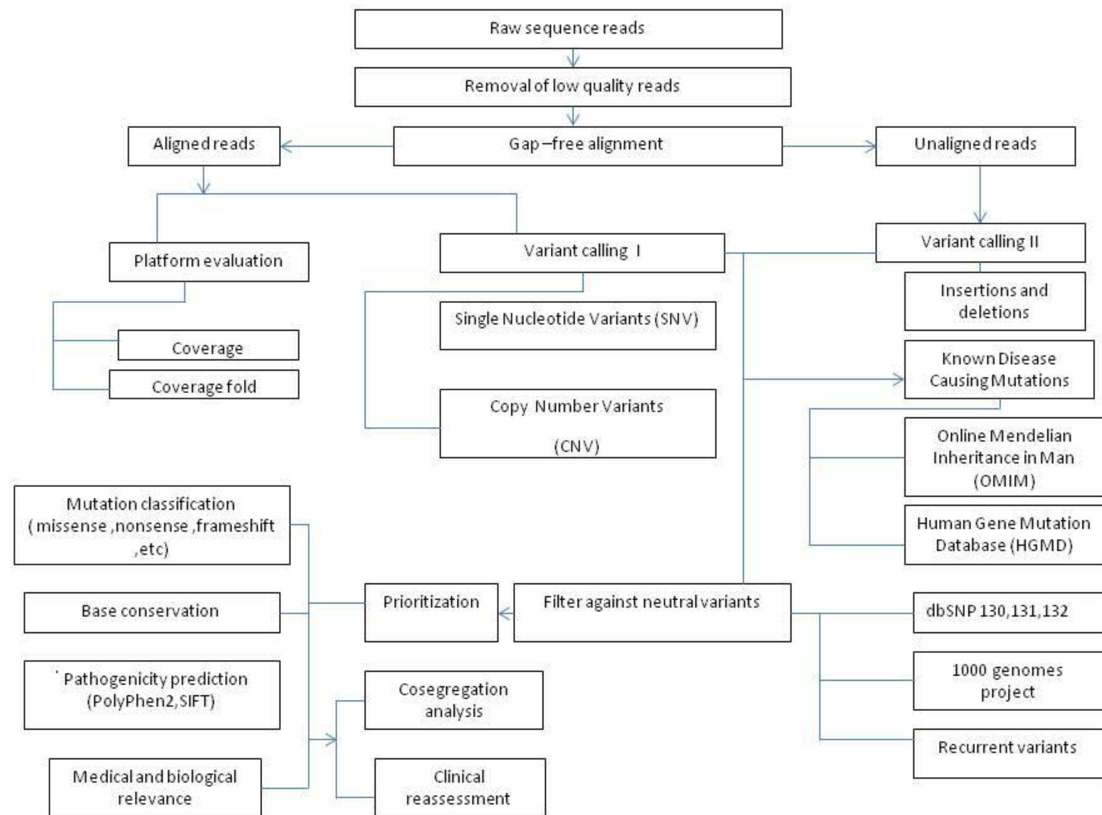


Figure 15. Variant calling, filtering and prioritization scheme adapted from Najambadi et al. 2011.

Several different measures were used to rank and screen the plausible disease causing missense changes or small in-frame deletions. Other criterias included: 1) the presence of a single non-polymorphic variant in the family; 2) the evolutionary conservation of the relevant nucleotide, as defined by the PhyloP score; 3) the pathogenicity of these variants, as predicted by PolyPhen2, SIFT etc and 4) the available biological and medical evidence supporting a role of this gene in the brain function, including disease links in humans and animal models and a range of other functional clues. Variants were validated by Sanger Sequencing. [Najmabadi et al. 2011].

4.4 PCR –Sequencing

PCR for sequencing reactions was performed in a 15 µl reaction volume. The reaction volume contained 220pmol of both primers and 10nmol of each nucleotide. The polymerase enzyme used in this reaction was Dynazyme II (Finnzymes).

The DNA of the study subjects were amplified by the polymerase chain reaction (PCR) using 2720 Thermal Cycler (Applied Biosystems). The polymerase chain reaction condition were as follows: 1 min at 95 °C followed by 35 cycles of the denaturation step of 30 sec at 95 °C, annealing step of 20 sec at 55 °C, the elongation step of 1 min at 72 °C and the final extension for 10 min at 72 °C.

PCR products were separated by electrophoresis on 1.5 % agarose gel with ethidum bromide to verify the success of the PCR reaction. A 100 bp size standard (O'Range ruler 100bp DNA ladder, Fermentas) was included to detect the size of the PCR products. Purification of the PCR products was done using ExoSAP IT (Affymetrix). It consisted of 15 min incubation at 37°C followed by enzyme activation at 80°C for a further 15 min. The sequencing was performed in a volume of 10µl with BigDye 3.1 terminator (Applied Biosystems) according to manufacturer's instructions. The sequencing reaction was performed in 25 cycles for 1 min with initial denaturation at 96 °C, 10 s denaturation at 96 °C, 5s primer annealing at primer specific temperature (55 °C), 4 min extension at 60 °C. PCR and sequencing primers are presented. Sequenced products were electrophoresed on an ABI 3730 DNA Analyzer (Applied Biosystems) and base calling was performed using Seqencing Analysis 5.2 software (Applied Biosystems). Reference sequence was obtained from the UCSC Human Genome Browser and sequence analysis was performed using Sequencer 4.8 (Gene Codes, USA).

4.5 Transcript Analysis

Patients with mutations leading to abnormal mRNA splicing were further studied through transcript analysis.

4.5.1 RNA Isolation

Blood samples were collected in PAX-gene RNA-tubes (Qiagen). RNA was isolated using the PAXgene Blood miRNA kit (Qiagen). The total RNA was purified from the

stabilized blood samples using the PAXgene silica membrane technology. PAXgene Blood RNA tubes were first centrifuged to pellet the samples, then washed with water and resuspended in Buffer BM1. Followed by the digestion in Buffer BM2 with proteinase K, the samples were homogenized by centrifugation through PAXgene shredder spin columns. Isopropanol was then added to the samples to optimize the binding conditions and the samples were then centrifuged through PAXgene RNA spin column whereby the total RNA binds to the PAXgene silica membrane. The bound RNA was subjected to DNase digestion to remove genomic DNA contamination and washed with buffer BM3 followed by Buffer BM4. Pure RNA was then eluted in buffer BR5.

4.5.2 RNA Integrity Check

RNA integrity was checked with the help of the Agilent 2100 Bioanalyzer. The chip used to measure the samples was first primed using a gel-dye mix and the chip priming station. 5 µl of the RNA 6000 Nano marker was then loaded to the ladder well and each sample well in the chip, followed by the loading of 1 µl of the ladder and each sample into the chip. The chip was then briefly mixed for 1 minute and then loaded into the Bioanalyzer. The run was started using the software connected to the computer. The software automatically allocates each sample a RIN, which usually ranges from 1 to 20.

4.5.3 cDNA synthesis

5 µg of the total RNA was reverse transcribed using the SuperScript III First Strand synthesis (Invitrogen) to generate the complementary cDNA synthesis. cDNA synthesis was performed in the first step using 5 µg total RNA with 50 µM oligo(dT) primers along with 10mM dNTP mix and DEPC treated water in a 10 µl reaction volume. The reaction mix was incubated at 65 °C for 5 minutes and then placed on ice for 1 minute. cDNA synthesis mix comprising of the following components 10X RT buffer, 25mM MgCl, 0.1M DTT, RNaseOUT and Superscript III RT was prepared and 10 µl of cDNA synthesis mix was added to each of the RNA/primer mixture and incubated for 50 minutes at 50 °C, reaction was terminated at 85 °C for 5 minutes. Finally 1 µl of RNase H was added and incubated for 20 minutes at 37°C. In the second step PCR was performed in separate tubes using 10X High Fidelity buffer, 10 Mm dNTP, 50 Mm

$MgSO_4$, forward and the reverse primers (20 µM gene specific primers), Platinum Taq DNA polymerase and autoclaved water to a 50 µl reaction volume. PCR reaction was carried out using 2 µl of cDNA. All PCR reactions were performed with a 2720 Thermal Cycler (Applied Biosystems). Initial denaturation at 94 °C for 1 minute followed by denaturation at 94 °C for 30 seconds, DNA amplification was carried out in a 35 cycles of 30 seconds annealing at temperature 55 °C, extension at 72 °C and the final extension at 72 °C for 7 minutes.

4.5.4 Gel Purification

Gel purification was done with the help of Pure Link Quick Gel Extraction kit. The desired band was excised using a clean sharp razor blade. The gel slice containing the DNA fragment was weighed and then placed into 1.5 ml polypropylene microcentrifuge tube. Then 3 volumes of Gel solubilization Buffer were added to 1 volume of gel. The tube containing the gel slice and the gel solubilization buffer was placed into a 50^0C water bath and incubated for 10 minutes. When the gel slice seemed to be dissolved, the tube was incubated for additional 5 minutes. The dissolved gel piece was then placed onto a Quick Gel Extraction Column containing silica membrane. The DNA was bound to the membrane either by centrifugation or vacuum extraction. The membrane was then washed with the 500 µl Wash buffer containing ethanol to remove the impurities and the purified DNA was then eluted into a recovery tube using 50 µl Elution Buffer (10Mm Tris-Hcl, pH 8.5).

4.6 Bioinformatics Analysis

Primers were designed by Primer3 and sequence comparison was done with the help of NCBI-BLAST. Clustal-W and LALIGN programs were used to compare the protein and the DNA sequences. The pathogenicity of the sequence variants was analyzed with the help of NNsplice (http://www.fruitfly.org/seq_tools/splice.html), PolyPhen2 (http://genetics.bwh.harvard.edu/pph/).

5. RESULTS

5.1 Clinical Results

The index patient is a 30-year-old man with a severe learning disability. He was born from a pregnancy complicated by pre-eclampsia. A cesarean section was performed at term due to a prolonged delivery. His birth weight was 2920 grams and length 49 centimeters. Apgar scores were 7, 8, and 9 after 1, 5, and 10 minutes. After birth dysmorphic features were noted consisting of abnormal auricles, an asymmetric thorax, and a small penis. At 10 months spasticity was observed in the hands. At 18 months hypotonia and ataxia were detected, and also, epilepsy was diagnosed. He was able to walk independently at 2 years. Brain CT, ERG, and VEP were normal at 22 months. Karyotype and cytogenetic fragile-X analysis gave normal results. At a clinical geneticist´s consultation at 2 years a prominent forehead, up-slanting palpebral fissures, and hyperextensible fingers were seen. At 30 years he had no speech. He walks in a slightly forward-flexed walking posture with small steps, and does not move hands. He is constantly wandering around and has aggressive bursts.

The older one of index patient´s maternal male cousins is a 45-year-old man with a moderate learning disability. This man was born after an uneventful pregnancy and delivery at term with a birth weight of 2650 grams, and a length of 46 centimeters. Apgar score was 10 at 1 minute. After birth he had feeding difficulties. The psychomotor development was delayed from birth. He learned to walk at 26 months. At 6 years he used only two words. At that time tremor intentionalis was diagnosed. Strabismus was seen in the eye investigation at 9 years. Also, a generalized slowing and disorganization of the EEG was seen. No seizures have occurred. Bone age was delayed. Karyotype was normal. He was referred to the Department of Medical Genetics, The Family Welfare Federation, Helsinki, at 17 years. Dysmorphic features were noted consisting of a high forehead, deep set eyes, hyperplastic supra-orbital ridges, up-slanting of palpebral fissures, abnormal and low-set ears, and a broad base to nose. He also had hypotonia and pes planus. At 45 years he uses two-word sentences, and walks. He has a pleasant personality

Table 4. The phenotypes observed in patients of the CUL4B family with XLID.

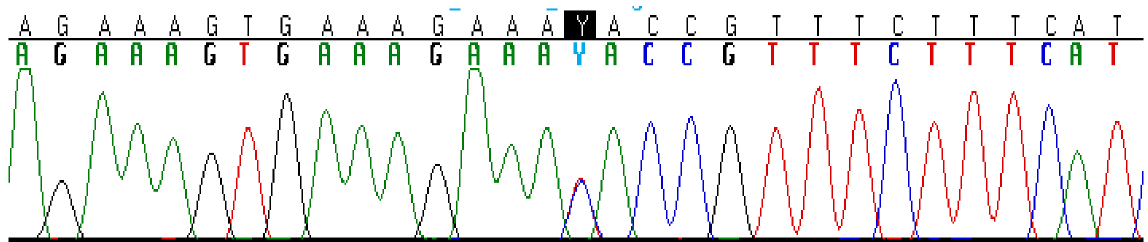
Measure	III-8	III-4	III-3
Age	30	45	39
Birth Weight(gm)	2920	2650	3080
Height(cm)	49	46	49
ID(level)	Severe	Moderate	Moderate
Motor Delay	Yes	Yes	Yes
Speech Delay	Yes	Yes	Yes
Tremor	Yes	Yes	Yes
Seizures	Yes	No	No
Gait Abnormality	Yes	No	No
Malformed low set ears	Yes	Yes	Yes
Gynaecomastia	Yes	Yes	Yes
Short feet	No	Yes	Yes
Prominent lower lip	No	No	Yes
Small Testes	Yes	Yes	Yes
Kyphosis	Yes	No	No
Behavioural problems	Yes	No	No
Sandal Gap	Yes	Yes	Yes
Narrow palpebral fissures	No	Yes	Yes
Syndactyly	Yes	Yes	Yes

The younger brother of index patient´s maternal cousins is a 39-year-old man with a moderate learning disability He was born after an uneventful pregnancy and delivery at 37 weeks of gestation with a birth weight of 3080 grams and a length of 49 centimeters. Apgar score was 9 at 1 minute. He learned to walk at 13 months. At 2 years speech delay was diagnosed. The audiological study was normal. Strabismus was seen in the eye examination. The EEG analysis revealed a generalized slowing and disorganization, but no signs of paroxysmal discharges were seen. Karyotype and cytogenetic fragile-X analysis gave normal results. He was referred to the Department of Medical Genetics, The Family Welfare Federation, Helsinki, at 12 years. Dysmorphic features were detected consisting of a high forehead, down-slanting palpebral fissures, a broad base to

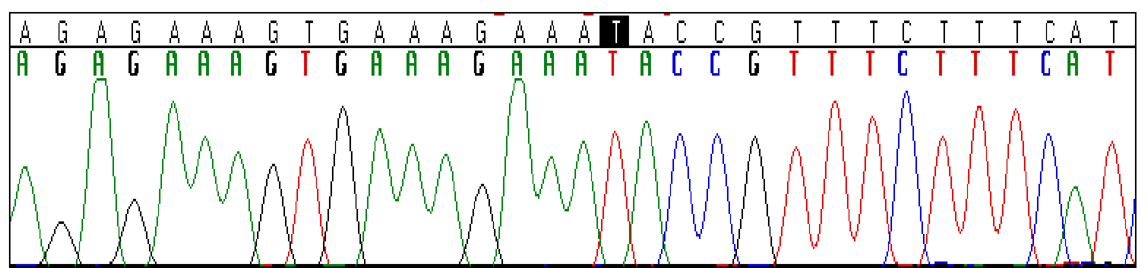
nose, and abnormal and low-set ears. Also brachydactyly, II-III syndactyly, pes planus and hypotonia were detected. At 39 years he uses single words, and walks. He has an easy-going personality.

5.2 Sequencing results and mutation analysis

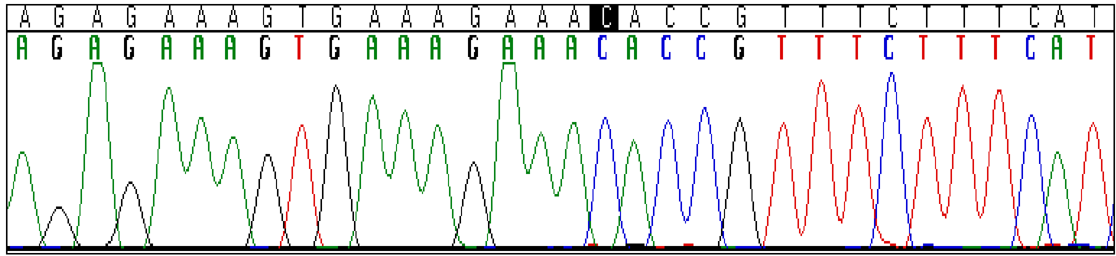
Exome sequencing performed in Max Planck institute, Berlin, revealed single base substitution T > C in intron 20 at position Chr X:-119,666,274 (GRCh 37/hg19) in index patient (III-8). Sequencing of genomic DNA of four affected male patients in Finland showed that the mutation was found in three of the four affected patients. The mothers (II-4 and II-6) were carriers of the mutation. To analyze the frequency of the mutation in the Finnish population we screened 200 anonymous blood donors and did not find any mutation



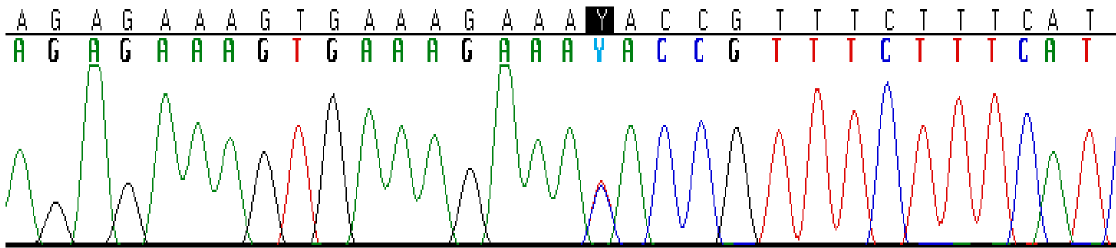
A) PATIENT II-6



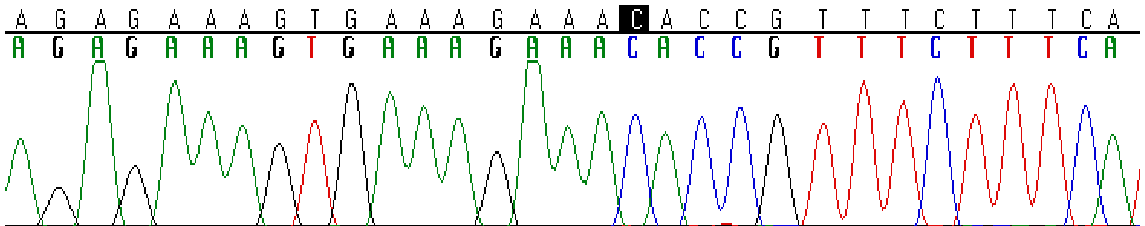
B) PATIENT III-7



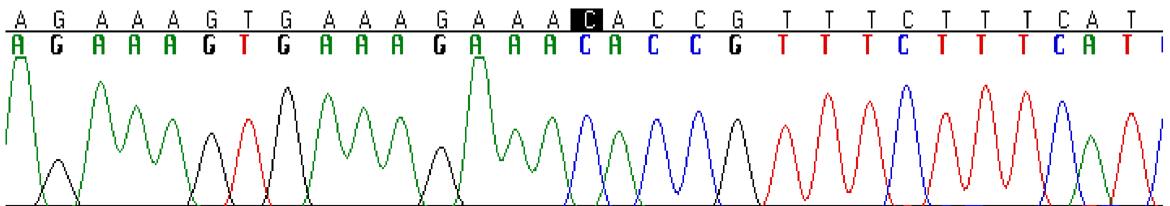
C) INDEX PATIENT III-8



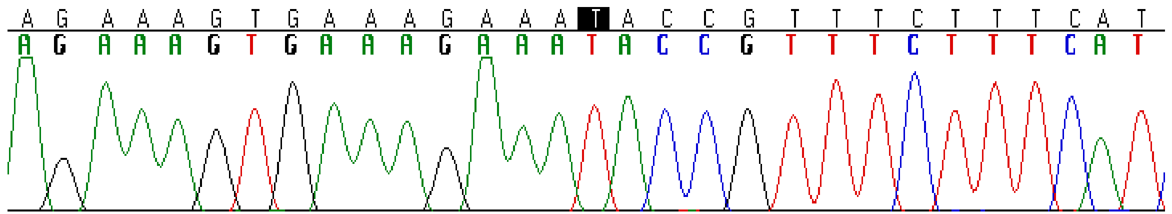
D) PATIENT II-4



E) PATIENT III-4

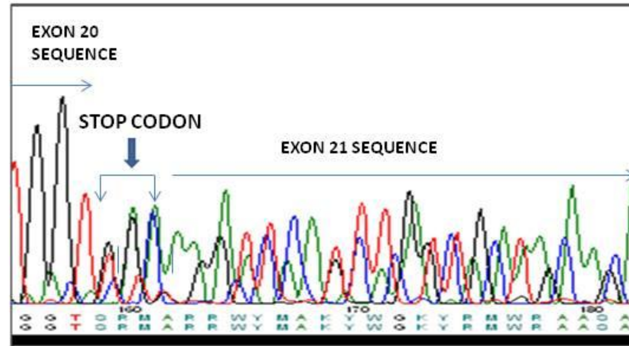


F) PATIENT III-3

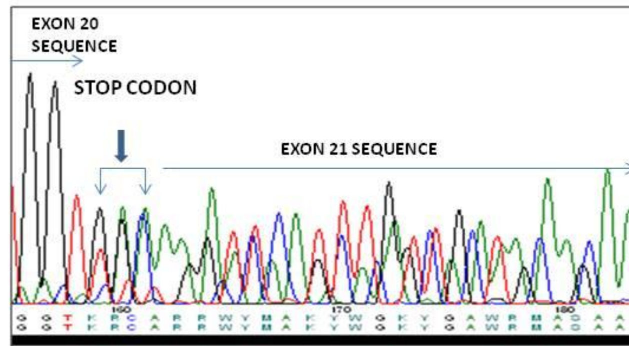


G) PATIENT III-7

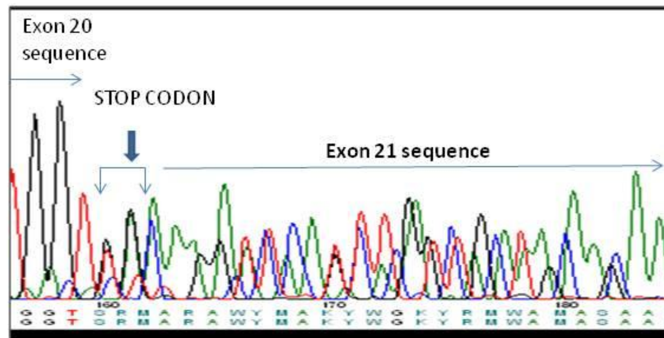
Further we performed Direct Sequencing of the RT-PCR products, which revealed the deletion of 78bp from exon 20 and 2bp deletion from exon 21, therefore leading to a total deletion of 80 bp nucleotides resulting into a premature stop codon at D806X.



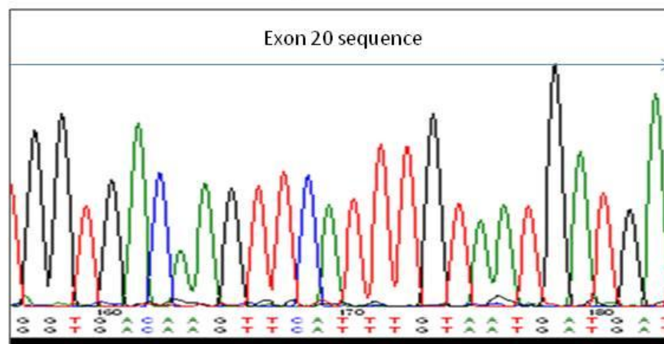
a) PATIENT III-3



b) PATIENT III-4



c) PATIENT III-8



d) PATIENT III-7

We also did a pairwise alignment to strengthen our results. With the help of EMBOSS Needle alignment program, we aligned our patient's cDNA sequence with the RefSeQ sequence of the CUL4B gene. Patient III-7 ‘ s cDNA sequence aligned completely with the reference sequence confirming that he did not have the CUL4B mutation and exon 20 and exon 21 remain intact. Instead, in patients III-4, III-3 & III-8, a deletion of 78 bp of exon 20 and 2 bp of exon 21 were found to be deleted that resulted in premature stop codon 806 where aspartic acid in changed to a stop codon D806X the protein sequence. (p.806D>X) and thereby affecting the translation and the protein function.

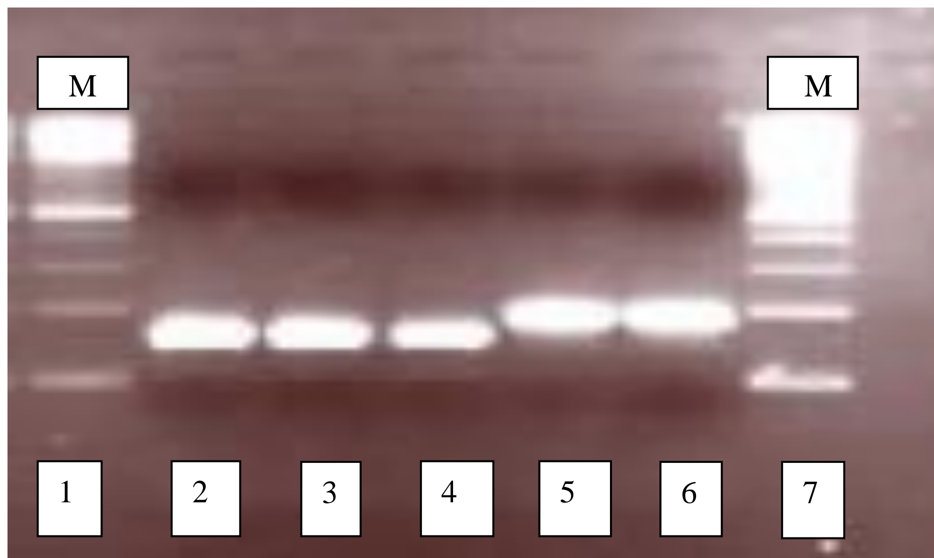


Figure 16. Lanes 1 & 7 represent the 200bp ladder. Lane 2 is Patient III-3, lane 3 is Patient III-4, Lane 4 is Patient III-8. Lane 5 is Patient who doesn't have the mutation and Lane 6 is the control sample.

cDNA analysis in patients depicts a band with lower molecular weight than the control, demonstrating the deletion of a total of 80 bp nucleotides.

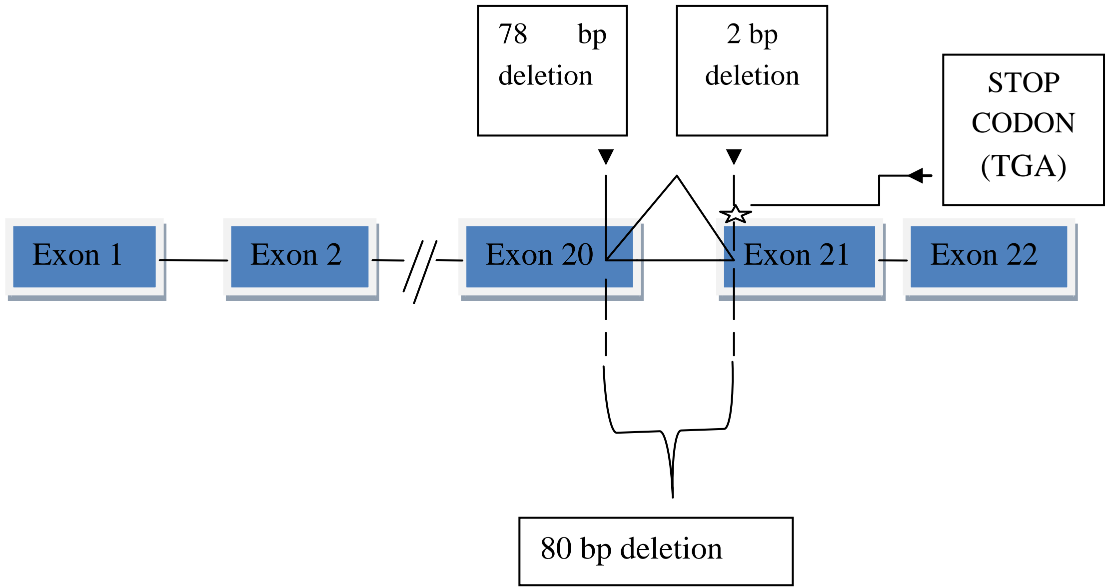


Figure 17: To illustrate the *CUL4B* mutation

6. DISCUSSION

In this study we report a novel mutation in *CUL4B* gene in three out of four male patients with X-linked intellectual disability. The mutation follows the rule of X linked inheritance in the family. The mothers of the affected children are the mutation carriers. The mutation is rare since it has not been found in 200 anonymous blood donors.

Previously mutations in *CUL4B* have been shown to cause Cabezas syndrome [Cabezas et al. 2000; Tarpey et al. 2007; Zou et al. 2007]. Cabezas syndrome (MIM 300354) is a syndromic form of the X linked intellectual disability caused by mutations in the *CUL4B* gene. The main clinical features of the syndrome were first described in a single family [Cabezas at al., 2000]. The clinical features of Cabezas syndrome are severe mental retardation, speech impairment, hyperactivity, seizures, intention tremor, inguinal hernia, small feet, syndactyly of the second and third toes and skin manifestations (hyperhydrosis and keratosis pilaris) and craniofacial dysmorphic features [Cabezas et al. 2009]. The clinical features in our patients were consistent with the Cabezas syndrome. A total of twelve families with *CUL4B* mutations have been reported till date [Badura-Stronka et al. 2010; Isidor et al. 2010; Tarpey et al. 2007; Zou et al. 2007; Ravn et al. 2011]. Wei et al. (1993) initially reported a large X- linked pedigree in which five males were affected, Zou et al. 2007 had limited the candidate interval on Xq25 and identified a nonsense mutation in the *CUL4B* gene.

Table 5. Identified *CUL4B* variants.

CUL4B VARIANT	*MUTATION CLASS*	*NUMBER OF AFFECTED INDIVIDUALS*	*PROTEIN CHANGE*	*REFERENCE*
c.638C>T exon 4	Missense	5	T213I	Tarpey et al., 2007 ☆
c.901-2A>G exon 7	Splice	3	not known	Tarpey et al., 2007 ☆
c.1007_1011delTTAT exon 8	Deletion	7	not known	Tarpey et al ., 2007 ☆
c.1162C>T exon 9	Nonsense	3	R388X	Tarpey et al., 2007 ☆
c.1714C>T exon 14	Missense	6	R572C	Tarpey et al., 2007 ☆
c.2107A>T Exon 18	Nonsense	3	K703X	Badura et al., 2009 ☆
c.2243T>C Exon 19	Missense	4	V745A	Tarpey et al., 2007 ☆
c.2493G>A Exon 20	Splice	3	T831T	Tarpey et al ., 2007 ☆
c. 2413T>C Exon 20	Nonsense	3	D806X	Our study*
c.2566C>T Exon 21	Nonsense	8	R856X	Tarpey et al., 2007 ☆

Later on Cabezas reported an unrelated X linked family with six males with intellectual disability, small testes, muscle wasting in legs, abnormal gait, tremor and prominent lower lip [Cabezas et al. 2000]. In 2007 Tarpey identified mutations in the *CUL4B* gene in eight independent families. In parallel, a nonsense mutation in *CUL4B* was identified in the family [Zou et al. 2007] initially reported by Wei and coworkers.

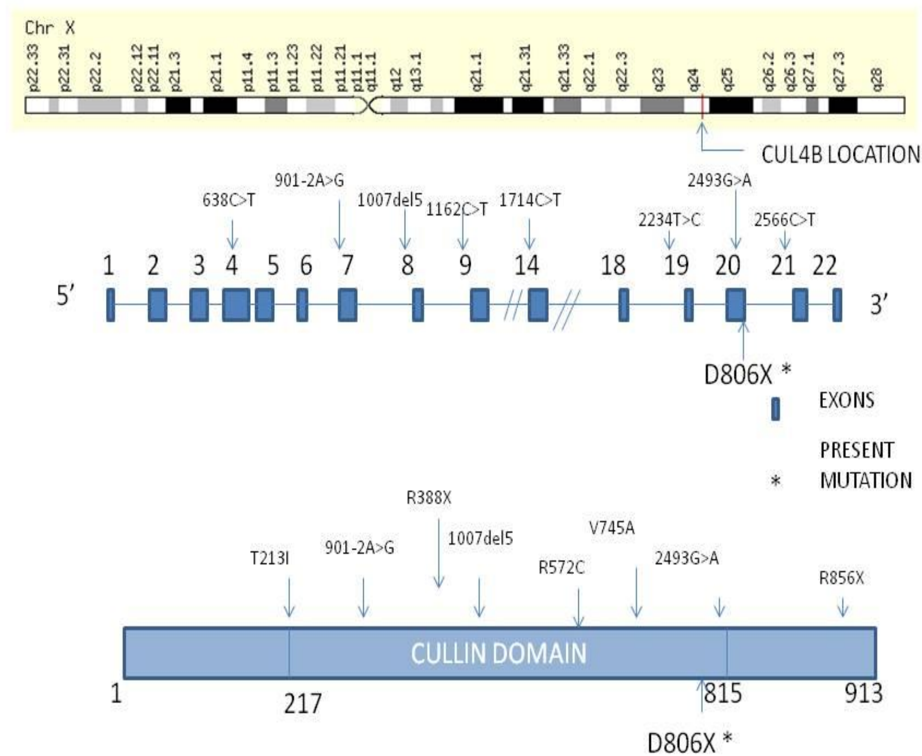


Figure 18. Schematic representation of the genomic structure of mutations found in eight families along with the present mutation. Also schematic representation of the protein sequence.

The *CUL4B* gene is composed of 22 exons and it encodes a protein of 913 amino acids (GenBank accession number NM_003588). CUL4B is a member of the family of cullin proteins that function primarily as scaffold proteins for a series of ubiquitin- protein ligase complexes that are responsible for regulating the degradation of cellular proteins [Hershko et al. 1992]. CUL4B has been found to be highly expressed in brain, testis, prostrate, colon and leukocytes. Cullins are a family of proteins that are characterized by the presence of distinct globular C terminal domain (cullin-homology domain) and a series of N terminal repeats of a five- helix bundle .[Tarpey et al. 2007]. Cullin –RING complexes comprise the largest known class of ubiquitin ligases. Cullins are involved in several biological processes which include cell cycle regulation, signal transduction, oxygen regulation and DNA repair. The cullin family consists of at least seven members

in mammals. [Petroski et al. 2005]. Although CUL4A and CUL4B are 80 % identical in their protein sequences, CUL4B has a unique N terminus that is 149 amino acids longer than CUL4A. The N terminus of CUL4B assembles a specific ubiquitin ligase complex that targets the estrogen receptor alpha for degradation and this function has been found to play in role in this XLID. [Zou Y et al. 2009]. *CUL4B* is a component of the ubiquitin system. Ubiquitin-dependent proteolysis is an integral cellular mechanism for regulating protein activity. It is found to be implicated in diverse set of biochemical processes, including signal transduction, transcription receptor down-regulation, and endocytosis. It has also found to be involved in regulation of the cell cycle, immune response, development, and programmed cell death. The proteolytic effects of ubiquitination have been widely studied, but it has been evident from the studies that it can influence activities of proteins through processes other than the degradation. [Petroski et al. 2005]. CUL4 is one of three founding cullins evolutionarily conserved from yeast to humans. Genetic analyses in various organisms have revealed a wide range of cellular and organismal functions mainly chromosome condensation, heterochromatin formation and DNA replication and repair [Harper et al. 2007]. *CUL4* is present as a single gene in yeast, plants and invertebrates, but the vertebrates express two closely related paralogs *CUL4A* and *CUL4B* [Higa et al. 2006]. Deletion of *CUL4B* in mice has resulted in embryonic lethality and defects in nervous system and heart development [Cox et al. 2010] whereas in humans loss of function mutations in *CUL4B* have been identified in patients with X linked Intellectual Disability [Badura-Stronka et al. 2010; Isidor et al. 2010; Tarpey et al. 2007; Zou et al. 2007]. *CUL4B* mutations including missense, frame-shifts and primary truncations appear to be distributed throughout the gene and in most cases result in significant reduced levels of CUL4B protein expression [Kerzendorfer et al. 2010]. Tarpey et al.(2007) findings of eight families with XLID, approximately 3% of the 250 families screened that have *CUL4B* mutations clearly indicates that this might be one of the most frequently mutated genes underlying XLID.

7. CONCLUSIONS

Although the traditional gene mapping approaches like karyotyping, homozygosity mapping , linkage analysis have helped in understanding Mendelian diseases over the past few decades, they are still unable to detect all forms of structural variation [Vissers et al. 2004; Lander et al. 1987; Kerem et al. 1989](Refer Table 6). Exome sequencing of a DNA sample from a single individual will generally reveal about 25,000 variants; the challenge then lies not in finding variants, but in identifying the particular mutation accountable for disease. In a single experiment it is possible that nearly all the coding content of the genome can be analyzed.

It has been found that whole exome sequencing has the potential to identify the causative mutations in diseases with genetic and phenotypic heterogeneity [Ng et al. 2010; Gilissen et al. 2010]. Exome sequencing definitely has the potential to identify rare variants. Apart for the success in finding mutations that causes rare, familial forms of disease it has been also successful in diseases caused by de novo mutations [Vissers et al. 2010].

Table 6: Mendelian Disease Gene Identification approaches

Approach	Applicable to	Pros	Cons
Candidate gene	Any disease	Easy to perform for one or two genes; requires no mapping, can directly identify the causative variant/mutation	Relies heavily on current biological knowledge; success rate very low
Genetic mapping by karyotyping	Any disease	Easy to perform; no familial cases required; can detect (large) balanced events	Low resolution, only detects large chromosomal aberrations; mutation detection requires second step
Genetic mapping by linkage analysis	Inherited disease	Easy to perform	Requires large families, often identifies large loci; mutation detection requires second step
Genetic mapping by homozygosity mapping	Recessive monogenic diseases	Small families can be used	Most useful for consanguineous families; often identifies large loci; mutation detection requires second step
Genetic mapping by CNV analysis	Monogenic/monolocus disease	High resolution CNV screening; no familial cases required; can potentially identify small loci	Only investigates CNVs; cannot detect balanced events, no base-pair resolution; mutation detection requires second step
Whole exome sequencing (WES)	Any disease	Base-pair resolution exome-wide; detects most types of genomic variation; can directly identify the causative variant/mutation	Unable to detect non-coding variants; limited resolution for CNVs and other structural variation; coverage variability due to enrichment process; relatively expensive
Whole genome sequencing (WGS)	Any disease	Base-pair resolution genome-wide; detects all types of genomic variation; can directly identify the causative variant/mutation	Data analysis complex; even more expensive than exome sequencing

Adapted from Christian Gilissen et al. 2011

Exome sequencing has already affirmed its worth in rapid genetic diagnosis and screening for many neurological diseases. (Singleton et al. 2011). However there are still various restraints to this technique, the exome data created by the current technologies are partial. It skips completely the structural variation. It has been also found that it also skips a certain set of exons and in case if the casual variants do not lie inside these exons, then they are not at all targeted. It is due to the partiality in the capture, sequencing and the various alignment processes. Whole genome sequencing is also being increasingly used. It has additional edge of capturing all of the exome and also it can afford to provide information on the structural variation present around the genome. Nevertheless, whole exome sequencing provides most of the advantages of the whole genome sequencing but with the lower costs in terms of sequencing as well as the storage and analysis of the data. [Majewski et al. 2011]. We anticipate that the whole exome sequencing in the coming years will drastically improve to elucidate the molecular basis of most remaining Mendelian disorders.

8. REFERENCES

Ahuja, A.S., Thapar, A. & Owen, M.J. 2005. Genetics of mental retardation. Indian J Med Sci 59, 9, pp. 407-417.

Badura-Stronka, M., Jamsheer, A., Materna-Kiryluk, A., Sowińska, A., Kiryluk, K., Budny, B. & Latos-Bieleńska, A. 2010. A novel nonsense mutation in CUL4B gene in three brothers with X linked mental retardation syndrome. Clin Genet 77, pp.141-144.

Ballabio, A. 1993. The rise and fall of positional cloning. Nat Genet 3(4), pp.277-9.

Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., EMOND, M.J., Nickerson, D.A. & Shendure, J. 2011. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12(11), pp.745-55.

Bisecker, L.G. 2010. Exome sequencing makes medical genomics a reality. Nat Genet 42(1), pp.30-35.

Billuart, P., Bienvenu, T., Ronce, N., des Portes, V., Vinet, M.C., Zemni, R., Carrié ,A., Beldjord, C., Kahn, A., Moraine, C. & Chelly J. 1998. Oligophrenin 1 encodes a rho-GAP protein involved in X-linked mental retardation. Pathol Bio 46(9), pp.678.

Boehnke, M. 1994. Limits of resolution of genetic linkage studies: implications for the positional cloning of human disease genes. Am J Hum Genet 55(2), pp.379-390.

Cabezas, DA., Slaugh, R., Abidi, F., Arena, J.F., Stevenson, RE., Schwartz, C.E. & Lubs H.A. 2000. A new X linked mental retardation (XLMR) syndrome with short stature, small testes, muscle wasting, and tremor localises to Xq24-q25. J Med Genet 37, pp. 663-668.

Çalışkan, M., Chong, J.X., Uricchio, L., Anderson, R., Chen, P., Sougnez, C., Garimella, K., Gabriel, S.B., dePristo, M.A., Shakir, K., Matern, D., Das, S., Waggoner, D., Nicolae, D.L. & Ober, C. 2011. Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13. Hum Mol Genet 20(7), pp.1285-1289.

Carrié ,A., Jun, L., Bienvenu, T., Vinet, M.C., McDonell, N., Couvert, P., Zemni, R., Cardona, A., Van Buggenhout, G., Frints, S., Hamel, B., Moraine, C., Ropers, H.H., Strom, T., Howell, G.R., Whittaker, A., Ross, M.T., Kahn, A., Fryns, J.P., Beldjord, C., Marynen, P. & Chelly, J. 1999. A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation. Nat Genet 23(1), pp. 25-31.

Cartegni, L., Chew, S.L. & Krainer, A.R. 2002. Listening to Silence and Understanding Nonsense : Exonic Mutations That Affect Splicing. Nature Reviews Genetics 3, pp. 285-298.

Castellvi-Bel, S. & Milà, M. 2001. Genes responsible for nonspecific mental retardation. Mol Genet Metab 72(2), pp.104-108.

Chee- Seng, Ku., David N, Cooper., Constantin, Polychronakos., Nasheen, Naidoo., Mengchu, Wu. & Richie Soong. 2012. Ann Neurol 71, pp.5-14.

Chelly, J. & Mandel, J.L. 2001. Monogenic causes of X-linked mental retardation. Nat Rev Genet 2(9), pp.669-80.

Chelly, J., Khelfaoui, M., Francis, F., Cherif, B. & Bienvenu, T. 2006. Genetics and pathophysiology of mental retardation. Eur J Hum Genet 14(6), pp.701-13

Chiurazzi, P., Tabolacci, E. & Neri J. 2004. X-linked mental retardation (XLMR): from clinical conditions to cloned genes. Crit Rev Clin Lab Sci 41(2), pp.117- 58.

Chiurazzi, P., Schwartz, C.E., Gecz, J. & Neri G. 2008. XLMR genes: update 2007. Eur J Hum Genet 16(4), pp.422-34.

Chiurazzi, P. 2011. Mental Retardation: Is naming the Real Issue? Am J Med Genet Part A 155, pp.974-975.

Christensen, MB., Královicová, J. & Vorechovský, I. 2005. Biased exon/intron distri-

bution of cryptic and de novo 3' splice sites. Nucleic Acids Res 33, pp.4882-4898.

De Brouwer A.P., Yntema H.G., Kleefstra, T., Lugtenberg, D., Oudakker, A,R., de Vries, B.B., van Bokhoven, H., Van Esch, H., Frints, S.G., Froyen, G., Gryns, J.P., Raynaud, M., Moiard, M.P., Ronce, N., Bensalem, A., Moraine, C., Poirier, K., Castelnau, L., Saillour, ., Bienvenu, T., Beldjord, C., des Portes, V., Chelly, J., Turner, G., Fullston, T., Gecz, J., Kuss, A.W., Tzschach, A., Jensen, L.R., Lenzner, S., Kalscheuer, V.M., Ropers, H.H. & Hamel B.C. 2007. Mutation frequencies of X linked Mental Retardation genes in families from the EUROMRX consortium. Hum Mut 28, pp.207-208.

Dominski, Z. & Kole, R. 1991. Selection of splice sites in pre-mRNAs with short internal exons. Mol Cell Biol 11, pp.6075–6083.

Durkin M. 2002. The epidemiology of developmental disabilities in low income countries. Ment Retard Dev Disabil Res Rev 8 (3), pp.206-211.

Eichler, E.E. 2009. Population Analysis of Large Copy Number Variants and Hotspots of Human Genetic Disease. Am J Hum Genet 84, pp.148-161.

Fab, Y.S., Jayakar, P., Zhu, H., Barbouth ,D., Sacharow, S., Morales, A., Carver, V., Benke, P., Mundy, P. & Elsas, L.J. 2007. Detection of pathogenic gene copy number variations in patients with mental retardation by genomewide oligonucleotide array comparative genomic hybridization. Hum Mutat 28, pp.1124-1132.

Friez, M.J., Jones, J.R., Clarkson, K., Lubs, H., Abuelo, D., Bier, J.A., Pai, S., Simensen, R., Williams, C., Giampietro, P.F., Schwartz, C.E. & Stevenson, R.E. 2006. Recurrent infections, hypotonia, and mental retardation caused by duplication of MECP2 and adjacent region in Xq28. Pediatrics 118, pp.1687e95.

Froyen, G., Van Esch, H., Bauters, M., Hollanders, K., Frints, S.G., Vermeesch, J.R., Devriendt, K., Fryns, J.P. & Marynen, P. 2007. Detection of genomic copy number changes in patients with idiopathic mental retardation by high resolution X- array – CGH: important role for increased gene dosage of XLMR genes. Hum Mut 28(10), pp.1034-1042.

Froyen, G., Corbett, M., Vandewalle, J., Jarvela, I., Lawrence, O., Meldrum, C., Bauters ,M., Govaerts, K., Vandeleur, L., Van Esch, H., Chelly, J., Sanlaville, D., van Bokhoven, H., Ropers, H.H., Laumonnier, F., Ranieri, E., Schwartz, C.E., Abidi, F., Tarpey. P.S., Futreal, P.A., Whibley, A., Raymond, F.L., Stratton, M.R, Fryns, J.P, Scott, R., Peippo, M., Sipponen, M., Partington, M., Mowat, D., Field, M., Hackett, A., Marynen. P., Turner, G. & Gécz, J. 2008. Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are associated with mental retardation. Am J Hum Genet 82(2), pp.432-443.

Gecz, J., Shouldbridge, C. & Corbett, M. 2009. The genetic landscape of intellectual disability arising from chromosome X. Trends Genet 25, pp.308-316.

Hastings, M.L. & Krainer, A.R. 2001. Pre-mRNA splicing in the new millennium. Curr Opin Cell Biol 13(3), pp.302-309.

Hawkins, J.D. 1988. A survey on intron and exon lengths. Nucleic Acids Res 16(21), pp.9893-9908.

Hershko, A. & Ciechanover, A. 1992. The ubiquitin system for protein degradation. Ann Rev Biochem 61, pp.761-807.

Higa, L.A., Yang, X., Zheng, J., Banks, D., Wu, M., Ghosh, P., Sun, H. & Zhang H. 2006. Involvement of CUL4 ubiquitin E3 ligases in regulating CDK inhibitors Dacapo/p27Kip1 and cyclin E degradation. Cell Cycle 5(1), pp.71-77.

Hoischen, A., Gilissen, C., Arts, H.H., Spruijt, L., Mans, D.A., Arts, P., van Lier, B.,Steehouwer, M., van Reeuwijk, J., Kant, S.G., Roepman, R., Knoers, N.V., Veltman, J.A. & Brunner, H.G. 2010. Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. Am J Hum Genet 87(3), pp.418-423.

Hori, T., Chiba, T., Hori, T., Osaka, F., Miyamoto, C., Okabayashi, K., Shimbara, N., Kato, S., & Tanaka, K. 1999. Covalent modifications of all members of human cullin family proteins by NEDD8. Oncogene 18, pp.6829-6834.

Hu, H., Wrogemann, K., Kalscheuer, V., Tzschach, A., Richard, H., et al. 2010. Mutation screening in 86 known X-linked mental retardation genes by droplet-based multiplex PCR and massive parallel sequencing. The HUGO J. 11568-010-9137-y.

Inlow, J.K. & Restifo, L.L 2004. Molecular and Comparitive Genetics of Mental Retardation. Genetics 166, pp.835-881.

Itsara, A., Gregory, M., Cooper, B.C., Girirajan, S., Li, J., Absher, D., Krauss, R.M., Myers, R.M., Ridker, P.M., Chasman, D.I., Mefford, H., Ying, P., Nickerson, D.A. & Kerr 1991

Jaworski, J., Kapitein, L.C., Gouveia, S.M,, Dortland, B.R., Wulf, P.S., Grigoriev. I., Camera, P., Spangler, S.A., Di Stefano, P., Demmers, J., Krugers, H., Defilippi, P. & Akhmanova, A. 2009. Neuron 61, pp.85–100.

Kahrizi, K., Hu, C.H., Garshasbi, M., Abedini, S.S., Ghadami, S., Kariminejad, R., Ullmann, R., Chen, W., Ropers, H.H., Kuss, A.W., Najmabadi, H. & Tzschach A. 2011. Next generation sequencing in a family with autosomal recessive Kahrizi syndrome reveals a homozygous frameshift mutation in SRD5A3. Eur J Hum Genet 19, pp.115-117.

Kallioniemi, A. & Kallioniemi,O.P.(1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258(5083), pp.818-21.

Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M. & Tsui, L.C. 1989. Identification of the cystic fibrosis gene: genetic analysis. Science 245, pp.1073–1080.

Kerr, B., Turner, G., Mulley, J., Gedeon, A. & Partington, M. 1991. Non-specific X linked mental retardation. J Med Genet 28(6), pp.378-382.

Kerzendorfer, C., Hart, L., Colnaghi, R., Carpenter, G., Alcantara, D., Outwin, E., Carr, AM. & O'Driscoll M. 2011. CUL4B-deficiency in humans: understanding the clinical consequences of impaired Cullin 4-RING E3 ubiquitin ligase function. Mech Ageing Dev. 132(8-9), pp.366-73.

Kaindl, A.M., Passemard, S., Kumar, P., Kraemer, N. & Issa, L. 2010. Many roads lead to primary autosomal recessive microcephaly. Prog Neurobiol 90, pp.363–83

Koolen, D.A., Sistermans, E.A., Nilessen, W., Knight, S.J., Regan, R., Liu, Y.T., Kooy, R.F., Rooms, L., Romano, C., Fichera, M., Schinzel, A., Baumer, A., Anderlid, B.M., Schoumans, J., van Kessel, A.G., Nordenskjold, M. & de Vries, B.B. 2008. Identification of non-recurrent submicroscopic genome imbalances: the advantage of genome-wide microarrays over targeted approaches. Eur Hum J Genet 16(3), pp.395-400.

Koolen, D.A., Pfundt, R., de Leeuw, N., Hehir-Kwa, J.Y., Nillesen, W.M., Neefs, I., Scheltinga, I., Sistermans, E., Smeets, D., Brunner, H.G., van Kessel ,A.G., Veltman, J.A. & de Vries, B.B. 2009. Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. Hum Mut 30, pp.283-92

Kramer JM, van Bokhoven H. 2009. Genetic and epigenetic defects in mental retardation. Int J Biochem Cell Biol 41, pp. 96-107.

Krawczak, M., Reiss, J. & Cooper D.N. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: Causes and consequences. Hum Genet 90, pp.41–54.

Krawczak, M., Thomas, N.S., Hundrieser, B., Mort, M., Wittig, M., Hampe, J. & Cooper D.N. 2007. Single Base-Pair Substitutions in Exon-Intron Junctions of Human Gene: Nature, Distribution and Consequences for mRNA splicing. Hum Mut 28 (2), pp.150-158.

Kurmangaliyev, Y.Z. & Gelfand, M.S. 2008. Computational analysis of splicing errors and mutations in human transcripts. BMC Genomics 9, pp.13.

Kuss, A.W., Santos-Rebouças, C.B., Fintelman-Rodrigues, N., Jensen, L.R., Ribeiro, M.G., Campos, M Jr., Santos, J.M. & Pimentel, M.M. 2011. A novel nonsense mutation in KDM5C/ JARID1C gene causing intellectual disability, short stature and speech delay. Neuroscience 498, pp.67-71.

Kutsche, K., Yntema, H., Brandt, A., Jantke, I., Nothwang, H.G., Orth, U., Boavida, M.G., David, D., Chelly, J., Fryns, J.P., Moraine, C., Ropers, H.H., Hamel, B.C., van Bokhoven, H. & Gal, A. 2000. Mutations in ARHGEF6, encoding a guanine nucleotide exchange factor for Rho GTPases, in patients with X-linked mental retardation. Nat Genet 26(2), pp.247-250.

Lander, E.S. & Botstein, D. 1987. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 236(4808), pp.1567-1570.

Lamprecht, R. & LeDoux, J. 2004. Structural plasticity and memory. Nat Rev Neurosci 5, pp.45–54.

Lehrke, R. 1972. A theory of X linkage of major intellectual traits. Am J Ment Defic 76, pp.611- 619.

Leonard, H. & Wen X. 2002. The epidemiology of mental retardation: challenges and opportunities in the new millenium. Mental retardation and developmental disabilities 8, pp. 117- 134.

Lichter, P., Joos, S., Bentz, M. & Lampel, S. 2000. Comparative genomic hybridization: uses and limitations. Semin Hematol 37(4), pp.348- 357.

Liang, J.S., Shimojima, K. & Yamamoto, T. 2008. Application of array based comparative genome hybridization in children with developmental delay or mental retardation. Pediatr Neonatol 49(6), pp.213- 217.

Luckasson, R., Borthwick-Duffy, S. & Buntinx, WHE. 2002. Mental Retardation: Definition, Classification, and Systems of Supports 10th edn. American Association on Mental Retardation: Washington, DC.

Madrigal, I., Rodrı ´guez-Revenga, L., Armengol, L., Gonza´lez, E., Rodriguez, B., Badenas, C., Sa´nchez, A., Martı ´nez, F., Guitart, M., Ferna´ndez, I., Arranz, J.A., Tejada, M., Pe´rez-Jurado, L.A., Estivill, X. & Mila`, M. 2007. X-chromosome tiling

path array detection of copy number variants in patients with chromosome X-linked mental retardation. BMC Genomics 8, pp.443.

Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, Alexandre. & Jabado, N. 2011. What can exome sequencing do for you? J Med Genet 48(9), pp.580- 589.

Malinow, R. & Malenka, R.C. 2002. AMPA receptor trafficking and synaptic plasticity. Annu Rev Neurosci 25, pp.103-126.

Maulik, P.K., Yasamy, M.T., Tomlinson, M., Lund, C., Van Ommeren, M. & Saxena, S. 2011. Responsible governance for mental health research in low resource countries. Plos Med 8(11), pp.e10001126.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren. Y., Thiruvahindrapduram. B., Fiebig. A., Schreiber, S., Friedman, J., Ketelaars, C.E., Vos, Y.J., Ficicioglu, C., Kirkpatrick, S., Nicolson, R., Sloman, L., Summers, A., Gibbons, C.A., Teebi, A., Chitayat, D., Weksberg, R., Thompson, A., Vardy, C., Crosbie, V., Luscombe, S., Baatjes, R., Zwaigenbaum, L., Roberts, W., Fernandez, B., Szatmari, P. & Scherer, S.W. 2008. Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 82(2), pp.477-488.

Mefford, H.C., Sharp, A.J., Baker, C., Itsara, A., & Jiang, Z. 2008. Recurrent rearrangements of chromosome1q21.1 and variable pediatric phenotypes. N Engl J Med 359, pp.1685–99

Miller, D,T., Adam, M.P., Aradhya, S., Biesecker, L.G., Brothman, A.R., Carter, N.P., Church, D.M., Crolla, J.A., Eichler, E.E., Epstein, C.J., Faucett, W.A., Feuk, L., Friedman, J.M., Hamosh, A., Jackson, L., Kaminsky, E.B., Kok, K., Krantz, I.D., Kuhn ,R.M., Lee, C., Ostell, J,M., Rosenberg, C., Scherer, S.W., Spinner, N.B., Stavropoulos, D.J., Tepperberg, J.H., Thorland, E.C., Vermeesch, J.R., Waggoner, D.J., Watson, M.S., Martin, C.L, & Ledbetter DH. 2010. Consensus Statement: Chromosomal Microarray Is a First-Tier Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. Am J Hum Genet 86, pp.749-764.

Moeschler, J.B. & Nisbeft, J. 2011. Invited comment on terminology. Am J Med Genet Part A 155, pp.972- 973.

Mulley, J.C., Kerr, B., Stevenson, R. & Lubs, H. 1992. Nomenclature guidelines for X-linked mental retardation. Am J Med Genet 43(1-2), pp.383-391.

Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., Behjati, F., Haas, S., Jamali, P., Zecha, A., Mohseni, M., Puttmann, L., Vahid, L.N., Jensen, C., Moheb, L.A., Bienek, M., Larti, F., Mueller, I., Weissmann, R., Darvish, H., Wrogemann, K., Hadavi, V., Lipkowitz, B., Esmaeeli –Nieh., Wieczorek, D., Kariminejad, R., Firouzabadi, S.G., Cohen, M., Fattahi, Z., Rost, I., Mojahedi, F., Hertzberg, C., Dehghan, A., Rajab, A., Banavandi, M.J., Hoffer, J., Falah, M., Musante, L., Kalscheuer, V., Ullmann, R., Kuss, A.W., Tzschach, A., Kahrizi, K. & Ropers HH. 2011. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 478(7367), pp.57-63.

Nakai, K. & Sakamoto, H. 1994. Construction of a novel database containing aberrant splicing mutations of mammalian genes. Gene. 141(2), pp.171-177.

Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K,, Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., Shendure, J. & Bamshad, M.J. 2010. Exome sequencing identifies the cause of a Mendelian disorder. Nat Genet 42, pp.30-35.

Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., Bamshad, M., Nickerson, D.A. & Shendure, J. 2009 . Targeted Capture and Massively Sequencing of Twelve Human Exomes. Nature 461(7261), pp.271-276.

O'Roak, B.J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J.J., Girirajan, S., Karakoc, E., Mackenzie, A.P., Ng, S.B., Baker, C., Rieder, M.J., Nickerson, D.A., Bernier, R., Fisher, S.E., Shendure, J. & Eichler, E.E. 2011. Exome Sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat Genet 43, pp.585-589.

Patel, D.R., Greydanus, D.E., Calles, J.L. Jr. & Pratt, H.D. 2010. Developmental disabilities across the lifespan. Dis Mon 56(6), pp.304-97.

Petroski, M.D. & Deshaies, R.J. 2005. Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. Cell 123(6), pp.1107-1120.

Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W.L., Chen, C., Zhai, Y., Dairkee, S.H., Ljung, B.M., Gray, J.W. & Albertson, D.G. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20(2), pp.207-11.

Pussegoda, K.A. 2010. Exome sequencing: locating causative genes in rare disorders. Clin Genet 78(1), pp. 32-33.

Rauch, A., Hoyer, J., Guth, S., Zweier, C., Kraus, C., Becker, C., Zenker, M., Hüffmeier, U., Thiel, C., Rüschendorf, F., Nürnberg, P., Reis, A. & Trautmann, U. 2006. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. Am J Med Genet A 140, pp.2063-2074.

Ravn, K., Lindquist, S., Nielsen, K., Dahm, T. & Tümer, Z. 2011. Deletion of CUL4B leads to concordant phenotype in a monozygotic twin pair. Clin Genet

Raynham, H., Gibbons, R., Flint, J. & Higgs, D. 1996. The genetic basis for mental retardation. QJM 89(3), pp.169- 75.

Raymond, F.L., Cox, J.J., Holden, S.T., Dee, S &, Burbridge, J.I. 2003. Identification of a 650 kb duplication at the X chromosome breakpoint in a patient with 46,X,t(X;8)(q28;q12) and non –syndromic mental retardation. J Med Genet 40(3), pp.169- 174.

Raymond, F.L. 2006. X linked mental retardation: a clincal guide. J Med Genet 43, pp.193- 200.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., González, J.R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., Marshall, C.R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M.J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D.F., Estivill, X., Tyler-Smith,C., Carter, N.P., Aburatani, H., Lee, C., Jones, K., Scherer, S.W. & Hurles, M.E. 2006. Global variation in copy number in the human genome. Nature 444(7118), pp.444-454.

Renieri, A., Pescucci, C., Longo, I., Ariani, F., Mari, F. & Meloni, I. 2005. Non-syndromic X-linked mental retardation: from a molecular to a clinical point of view. J Cell Physiol 204(1), pp.8- 20.

Rogan, P.K., Faux, B.M. & Schneider, T.D. 1998. Information Analysis of Human Splice Site Mutations. Hum Mut 12, pp.153-171.

Ross, M.T., Grafham, D.V., Coffey, A.J., Scherer, S., McLay, K., Muzny, D, Platzer, M., Howell, G.R., Burrows, C., Bird, C.P., Frankish, A., Lovell, F.L., Howe, K.L., Ashurst, J.L., Fulton, R.S., Sudbrak, R., Wen, G., Jones, M.C., Hurles, M.E., Andrews, T.D., Scott, C.E., Searle, S., Ramser, J., Whittaker, A., Deadman, R., Carter, N.P., Hunt, S.E., Chen, R., Cree, A., Gunaratne, P., Havlak, P., Hodgson, A., Metzker, M.L., Richards, S., Scott, G., Steffen, D., Sodergren, E., Wheeler, D.A., Worley, K.C., Ainscough, R., Ambrose, K.D., Ansari-Lari, M.A., Aradhya, S., Ashwell, R.I., Babbage, A.K., Bagguley, C.L., Ballabio, A., Banerjee, R., Barker, G.E., Barlow, K.F., Barrett, I.P., Bates, K.N., Beare, D.M., Beasley, H., Beasley, O., Beck, A., Bethel, G., Blechschmidt, K., Brady, N., Bray-Allen, S., Bridgeman, A.M., Brown, A.J., Brown, M.J., Bonnin, D., Bruford, E.A., Buhay, C., Burch, P., Burford, D., Burgess, J., Burrill, W., Burton, J., Bye, J.M., Carder, C., Carrel, L., Chako, J., Chapman, J.C., Chavez, D., Chen, E., Chen, G., Chen, Y., Chen, Z., Chinault, C., Ciccodicola, A., Clark, S.Y., Clarke, G., Clee, C.M., Clegg, S., Clerc-Blankenburg, K., Clifford, K., Cobley, V., Cole, C.G., Conquer, J.S., Corby, N., Connor, R.E., David ,R., Davies, J., Davis, C., Davis, J., Delgado, O., Deshazo, D,, Dhami, P., Ding, Y., Dinh, H., Dodsworth, S., Draper, H., Dugan-Rocha, S., Dunham, A., Dunn, M., Durbin, K.J., Dutta, I., Eades, T., Ellwood, M., Emery-Cohen, A., Errington, H., Evans, K.L., Faulkner, L., Francis, F.,

Frankland, J., Fraser, A.E., Galgoczy, P., Gilbert, J., Gill, R., Glo¨ckner, G., Gregory, S.G., Gribble, S., Griffiths, C., Grocock, R., Gu, Y., Gwilliam, R., Hamilton, C., Hart, E.A., Hawes, A., Heath, P.D., Heitmann, K., Hennig, S., Hernandez, J., Hinzmann, B., Ho, S., Hoffs, M., Howden, P.J., Huckle, E.J., Hume, J., Hunt, P.J., Hunt, A.R., Isherwood, J., Jacob, L., Johnson, D., Jones, S., de Jong, P.J., Joseph, S.S., Keenan, S., Kelly, S., Kershaw, J.K., Khan, Z., Kioschis, P., Klages, S., Knights, A.J., Kosiura, A., Kovar-Smith, C., Laird, G.K., Langford, C., Lawlor, S., Leversha, M., Lewis, L., Liu, W., Lloyd, C., Lloyd, D.M., Loulseged, H., Loveland, J.E., Lovell, J.D., Lozado, R., Lu, J., Lyne, R., Ma, J., Maheshwari, M., Matthews, L.H., McDowall, J., McLaren, S., McMurray, A., Meidl, P., Meitinger, T., Milne, S., Miner, G., Mistry, S.L., Morgan, M., Morris, S., Mu¨ller, I., Mullikin, J.C., Nguyen, N., Nordsiek, G., Nyakatura, G., O'Dell, C.N., Okwuonu, G., Palmer, S., Pandian, R., Parker, .D, Parrish ,J., Pasternak, S., Patel, D., Pearce, A.V., Pearson, D.M., Pelan, S.E., Perez, L., Porter, K.M., Ramsey, Y., Reichwald, K., Rhodes, S., Ridler, K.A., Schlessinger, D., Schueler, M.G., Sehra, H.K., Shaw-Smith, C., Shen, H., Sheridan, E.M., Shownkeen, R., Skuce, C.D., Smith, M.L., Sotheran, E.C., Steingruber, H.E., Steward, C.A., Storey, R., Swann, R.M., Swarbreck, D., Tabor, P.E., Taudien, S., Taylor, T., Teague, B., Thomas, K., Thorpe, A., Timms, K., Tracey, A., Trevanion, S., Tromans, A.C., d'Urso, M., Verduzco, D., Villasana, D., Waldron, L., Wall, M., Wang ,Q., Warren ,J., Warry, G.L., Wei, X., West, A., Whitehead, S.L., Whiteley, M.N., Wilkinson, J.E., Willey, D.L., Williams, G., Williams, L., Williamson, A., Williamson, H., Wilming, L., Woodmansey, R.L., Wray, P.W., Yen, J., Zhang, J., Zhou, J., Zoghbi, H., Zorilla, S., Buck, D., Reinhardt, R., Poustka, A., Rosenthal, A., Lehrach, H., Meindl, A., Minx, P.J., Hillier, L.W., Willard, H.F., Wilson, R.K., Waterston, R.H., Rice, C.M., Vaudin, M., Coulson, A., Nelson, D.L., Weinstock, G., Sulston, J.E., Durbin, R., Hubbard, T., Gibbs, R.A., Beck, S., Rogers, J. & Bentley DR. 2005. The DNA sequence of the human X chromosome. Nature 434, pp.325-337.

Ropers, H.H., Hoeltzenbein, M., Kalscheuer, V., Yntema, H., Hamel, B., Fryns, J.P., Chelly, J., Partington, M., Gecz, J. & Moraine, C. 2003. Nonsyndromic X-linked mental retardation: where are the missing mutations? Trends Genet 19(6), pp.316-20.

Ropers, H.H. & Hamel, B.C. 2005. X-linked mental retardation. Nat Rev Genet 6(1), pp. 46-57.

Ropers, H.H. 2008. Genetics of Intellectual Disability. Current Opinion in Genetics & Development 18, pp.241-250.

Ropers H.H. 2010. Genetics of early onset cognitive impairment. Annu Rev Genomics Hum Genet Vol 11, pp.161-187.

Ryan, M., Diack, J., Watson, V. & Smith, N. 2005. Rapid prenatal diagnostic testing for Down syndrome only or longer wait for full karyotype: the views of pregnant women. Prenat Diagn 25(13), pp.1206-1211

Salzberg, S.L. & Pop, M. 2008. Bioinformatics challenges of new sequencing technology. Trends Genet. 24(3), pp.142-149.

Schalock, R.L., Luckasson, R.A., Shogren, K.A., Borthwick-Duffy, S., Bradley, V., Buntinx, W.H., Coulter, D.L., Craig, E.M., Gomez, S.C., Lachapelle, Y., Reeve, A., Snell, M.E., Spreat, S., Tassé, M.J., Thompson, J.R., Verdugo, M.A., Wehmeyer, M.L. & Yeager, M.H. 2007. The renaming of mental retardation: understanding the change to th term intellectual disability. Intellect Dev Disabil 45, pp.116-24.

Senapathy, P., Bhasi, A., Mattox, J., Dhandapany, P.S. & Sadayappan, S. 2010. Targeted genome-wide enrichment of functional regions. Plos One 5(6), pp.e11138.

Singleton, A.B. 2011. Exome sequencing: a transformative technology. Lancet Neurol 10(10), pp.942-946.

Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nickolenko, J., Benner, A., Döhner, H., Cremer, T. & Lichter, P. 1997. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer 20(4), pp.399-407.

Stefansson, H., Rujescu, D., Cichon. S., Pietilainen, O.P. & Ingason A. 2008. Large recurrent microdeletions associated with schizophrenia. Nature 455, pp.232–236

Stenson, P.D., Ball, E.V., Mort, M., Phillips, A.D., Shiel, J.A., Thomas, N.S., Abeysinghe, S., Krawczak, M. & Cooper, D.N. 2003. Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 21(6), pp.577-581.

Stephens, J. C., Schneider, J.A., Tanguay, D.A., Choi, J., Acharya, T., Stanley, S.E., Jiang, R., Messer, C.J., Chew, A., Han, J.H., Duan, J., Carr, J.L., Lee ,M.S,, Koshy, B., Kumar, A.M., Zhang, G., Newell, W.R., Windemuth, A., Xu, C., Kalbfleisch, T.S., Shaner ,S.L., Arnold, K., Schulz, V., Drysdale, C.M., Nandabalan ,K., Judson, R.S., Ruano ,G. & Vovis, G.F. 2001. Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 293, pp.489–493

Stevenson, R.E. 2005. Advances in X linked mental retardation. Curr Opin Pediatr 17, pp.720-724.

Strobl-Wildemann, G., Kalscheuer,V.M., Hu, H., Wrogemann, K., Ropers, H.H. & Tzschach A. 2011. Novel GDI1 mutation in a large family with Non- syndromic Intellectual Disability. Am J Med Genet Part A 155, pp.3067-3070.

Tabor, H.K., Berkman, B.E., Hull, S.C. & Bamshad, M.J. 2011. Really gets personal:How Exome and Whole Genome Sequencing Challenge the Ethical Framework of Human Genetics Research. Am J Hum Genet Part A 155, pp.2916-2924.

Takano, K., Schwartz, C., Witham, S. & Alexov, E. 2011. A missense mutation in CLIC2 associated with intellectual disability is predicted by in silico modeling to affect protein stability and dynamics. Proteins 79, pp.2444- 2454.

Tarpey, P.S., Raymond, F.L., O'Meara, S., Edkins, S., Teague, J., Butler, A., Dicks, E., Stevens, C., Tofts, C., Avis, T., Barthorpe, S., Buck, G., Cole, J., Gray, K., Halliday, K., Harrison, R., Hills, K., Jenkinson, A., Jones, D., Menzies, A., Mironenko, T., Perry, J., Raine, K., Richardson, D., Shepherd, R., Small, A., Varian, J., West, S., Widaa, S., Mallya, U., Moon, J., Luo, Y., Holder, S., Smithson, S.F., Hurst, J.A., Clayton-Smith, J., Kerr, B., Boyle, J., Shaw, M., Vandeleur, L., Rodriguez, J., Slaugh, R., Easton, D.F., Wooster, R., Bobrow, M., Srivastava, A.K., Stevenson, R.E., Schwartz, C.E., Turner, G., Gecz, J., Futreal P.A., Stratton M.R. & Partington, M. 2007. Mutations in CUL4B

which encodes a Ubiquitin E3 Ligase Subunit, Cause an X –linked Mental Retardation Syndrome Associated with Aggressive Outbursts, Sezures, Relative Macrocephaly, Central Obesity, Hypogonadism, Pes Cavus, and Tremor. Am J Hum Genet 80, pp.345-352.

Tarpey, P.S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy, C., O'Meara, S., Latimer, C., Dicks, E., Menzies, A., Stephens, P., Blow, M., Greenman, C., Xue, Y., Tyler-Smith, C., Thompson, D., Gray, K., Andrews, J., Barthorpe, S., Buck, G., Cole, J., Dunmore, R., Jones, D., Maddison, M., Mironenko, T., Turner, R., Turrell, K., Varian, J., West, S., Widaa, S., Wray, P., Teague, J., Butler, A., Jenkinson, A., Jia, M., Richardson, D., Shepherd, R., Wooster, R., Tejada, M.I., Martinez, F., Carvill, G., Goliath, R., de Brouwer, A.P,, van Bokhoven, H., Van Esch, H., Chelly, J., Raynaud, M., Ropers, H.H., Abidi, F.E., Srivastava, A.K., Cox, J., Luo, Y., Mallya, U., Moon, J., Parnau, J., Mohammed, S., Tolmie, J.L., Shoubridge, C., Corbett, M., Gardner, A., Haan, E., Rujirabanjerd, S., Shaw, M., Vandeleur, L., Fullston, T., Easton, D.F., Boyle, J., Partington, M., Hackett, A., Field, M., Skinner, C., Stevenson, R.E., Bobrow, M., Turner, G., Schwartz, C.E., Gecz, J., Raymond, F.L., Futreal, P.A. & Stratton, M.R. 2009. A systematic large scale resequencing screen of X chromosome coding exons in mental retardation. Nat Genet 41, pp.535- 543.

Teer, J.K. & Mullikin, J.C. 2010. Exome sequencing: the sweet spot before the whole genomes. Hum Mol Genet 19:R145-R151.

ten Bosch, J.R. & Grody W.W. 2008. Massively Parallel Sequencing in Clinical Diagnostics. Journal of Molecular Diagnostics 10, pp.484-492.

Tucker, T., Marra, M. & Friedman, J.M. 2009. Massively parallel sequencing: the next big thing in genetic medicine. Am J Hum Genet 85, pp.142-154.

van Bokhoven H. 2011. Genetic and epigenetic networks in intellectual disabilities. Ann Rev Genet 45, pp.81– 104.

Van Esch, H., Bauters, M., Ignatius, J., Jansen, M., Raynaud, M., Hollanders, K., Lugtenberg , D., Bienvenu, T., Jensen, L.R., Gecz, J., Moraine, C., Marynen, P., Fryns, J.P.

& Froyen, G.2005. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males.Am J Hum Genet 77(3), pp.442-453.

Veltman, J.A., Schoenmakers, E.F., Eussen, B.H., Janssen, I., Merkx, G., van Cleef, B., van Ravenswaaij, C.M., Brunner, H.G., Smeets, D., & Geurts van Kessel, A. 2002. High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization. Am J Hum Genet 70, pp.1269–76.

Vissers, L.E., de Vries, B.B., Osoegawa, K., Janssen, I.M., Feuth, T., Choy, C.O., Straatman, H., van der Vliet, W., Huys, E.H., van Rijk, A., Smeets, D., van Ravenswaaij-Arts ,C.M., Knoers, N.V., van der Burgt, I., de Jong, P.J., Brunner, H.G., van Kessel, A.G. & Schoenmakers, E.F. 2003. Array –based comparative genomic hybridization for the genomewide detetion of submicrosopic chromosomal abnormalities. Am J Hum Genet 73, pp.1261-1270.

Vissers, L.E., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., van Bon, B.W., Hoischen, A., de Vries, B.B., Brunner, H.G. & Veltman, J.A. 2010. A de novo paradigm for mental retardation. Nat Genet 42, pp.1109-1112.

Vissers, L.E., de Vries, B.B. & Veltman, J.A. 2010. Genomic microarrays in mental retardation from copy number variation to gene, from research to diagnosis. J Med Genet 47, pp.289- 297.

Volpi, L., Roversi, G., Colombo, E.A., Leijsten, N., Concolino, D., Calabria, A., Mencarelli, M.A., Fimiani., Macciardi, F., Pfundt, R., Schoenmakers, E.F. & Larizza, L. 2010. Targeted next-generation sequencing appoints c16orf57 as clericuzio-type poikiloderma with neutropenia gene. Am J Hum Genet 86(1), pp.72-76.

Wei, J., Chen, B., Jiang, Y., Yang, Y., Guo, Y. 1993. Smith-Fineman-Myers syndrome: report on a large family. Am J Med Genet 47, pp.307- 311.

Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E. & Miller, D.T. 2008. Association be-

tween microdeletion and microduplication at 16p11.2 and autism. N Engl J Med 358, pp.667–675

Wong, K.K., deLeeuw, R.J., Dosanjh, N.S., Kimm, L.R., Cheng, Z., Horsman, D.E. & MacAulay C, Ng. 2007. A comprehensive analysis of common copy- number variations in the human genome. Am J Hum Genet 80, pp.91-104.

Zemni, R., Bienvenu, T. & Vinet, M.C. 2000. A new gene involved in X-linked mental retardation identified by analysis of an X;2 balanced translocation. Nat Genet 24, pp.167–170.

Zhang, Z.F., Ruivenkamp, C., Staaf, J., Zhu, H., Barbaro, M., Petillo, D., Khoo, S.K., Borg, A., Fan ,Y.S. & Schoumans, J. 2008. Detection of submicroscopic constitutional chromosome aberrations in clinical diagnostics: a validation of the practical performance of different array platforms. Eur J Hum Genet 16(7), pp.786-92.

Zhang, J., Chiodini, R., Badr, A. & Zhang, G. 2011. The impact of next generation sequencing on genomics. J Genet Genomics 38(3), pp.95- 109.

Zou Y, Liu Q, Chen B, Zhang X, Guo C, Zhou H, Li J, Gao G, Guo Y, Yan C, Wei J, Shao C, Gong Y. 2007. Mutation in CUL4B, which encodes a member of cullin-RING ubiquitin ligase complex, causes X-linked mental retardation. Am J Hum Genet 80, pp. 561-6.

Zou Y, Mi J, Cui J, Lu D, Zhang X, Guo C, Gao G, Liu Q, Chen B, Shao C, Gong Y. 2009. Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression. J Biol Chem 284, pp. 33320- 32.

Electronic-Database Information

UCSC genome browser, http://genome.ucsc.edu/

XLMR genes update website, http://xlmr.interfree.it/home.htm

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM

American Association on Mental Retardation (AAMR), http://www.aamr.org/About_AAMR/index.shtml

PolyPhen program, http://tux.embl-heidelberg.de/ramensky/polyphen.cgi

Webgene Splice View, http://l25.itba.mi.cnr.it/~webgene/wwwspliceview.html

Gene Splicer Web Interface, http://www.tigr.org/tdb/GeneSplicer/gene_spl.html

NatGen2 Server, http://www.cbs.dtu.dk/services/NetGene2/

Universal Protein Resource, http://www.pir.uniprot.org/

Simple Modular Architecture Research Tool, http://smart.embl-heidelberg.de/

dbSNP database , http ://www.ncbi.nlm.nih.gov/projects/SNP/