Genetic studies of children with mental retardation

Saideh Rajaei

UNIVERSITY OF GOTHENBURG

Department of Medical and Clinical Genetics
Institute of Biomedicine
The Sahlgrenska Academy at the University of Gothenburg,
Gothenburg, Sweden, 2013

Department of Medical and Clinical Genetics
Institute of Biomedicine
The Sahlgrenska Academy at the University of Gothenburg

To my wonderful family

"Nobody can go back and start a new begining but anyone can start today and make a new ending"

Maria Robinson

ABSTRACT

Genetic studies of children with mental retardation

Saideh Rajaei
Department of Medical and Clinical Genetics, Institute of Biomedicine
The Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden, 2013

Mental retardation (MR) is characterised by significant limitations in intellectual function and adaptive behaviour. It is estimated that MR affects up to 3% of the population in Europe. Patients with MR are an aetiologically heterogeneous group. Approximately 25-35% of the patients have a genetic diagnosis. In the last decade, the introduction of molecular karyotyping has proved to add new and important data for MR diagnosis.

In this thesis, we have undertaken extensive genetic analysis for two groups of children with MR. In the first study, a group of fourteen clinically diagnosed early infantile onset Rett Syndrome patients were included. The aim for this patient group was to identify possible pathogenic genetic variations. The second study was population-based and included children (born 1987–1998; 46 000 children) living in the Swedish County of Halland in 2004. 133 patients with SMR were identified and then divided in four categories depending on timing of onset; prenatal, perinatal, postnatal and undetermined timing. 23 patients within the prenatal group (included 82 patients in total), were still undiagnosed. The aims were; firstly to investigate whether the aetiological prevalence and co-morbidity of SMR, as well as the male: female ratio in Scandinavia had changed over time. Secondly, to investigate the impact of new genetic methodology, like molecular karyotyping, on the number of diagnosed children with SMR.
In the early infantile onset RTT patients we found a *MECP2* deletion (1/14) with the initial screening, and molecular karyotyping (SNP array) found three (3/14) copy number variations with uncertain significance.
The SMR study showed the same prevalence as previous Scandinavian studies (2.9 per 1000). The molecular karyotyping resulted in diagnosis of 5/19 patients of the previously undiagnosed patients from the prenatal group, which increased the frequency of diagnosed patients from earlier 4% (using traditional analysis methods) to 22.5% (this study).
Furthermore, we identified *MECP2* duplication syndrome in a female patient with mild to moderate MR and two brothers from the SMR study. These results imply that *MECP2* duplication is a pathogenic CNV in both genders, thus, there are phenotypic differences in females and males. Risk for recurrence is 50% for boys and less for girls because of incomplete penetrance.
In conclusion, this thesis investigates the genetic causes of two specific groups of patients with mental retardation. This follow up is essential for prognosis, management, and genetic counselling which permit prenatal diagnosis and determination of recurrence risk.

List of papers

The thesis is based on the following papers listed in reverse chronological order. They are appended at the end of this thesis and will be referred to by their roman numbers:

I. Rajaei S, Erlandson A, Kyllerman M, Albage M, Lundstrom I, Karrstedt EL, Hagberg B. Early infantile onset "congenital" Rett syndrome variants: Swedish experience through four decades and mutation analysis. *J Child Neurol. 2011 Jan;26(1):65-71.*

II. Lundvall M, Rajaei S, Erlandson A, Kyllerman M. Aetiology of severe mental retardation and further genetic analysis by high-resolution microarray in a population-based series of 6- to 17-year-old children. *Acta Paediatr. 2012 Jan;101(1):85-91.*

III. Rajaei S, Kyllerman M, Albåge M, Erlandson A. Copy-number variations in early infantile onset Rett syndrome variants. *Manuscript in preparation*

IV. Rajaei S, Lundvall M, Hallböök T, Kyllerman M, Stefanova M, Erlandson A. *MECP2* duplication syndrome in three patients with aspects on phenotype and implications for genetic counselling. *Manuscript in preparation*

CONTESTS

ABBREVIATIONS

ASD	Autism spectrum disorders
Bp	Base pair
CDKL5	Cyclin-Dependent Kinase-Like 5 (gene, OMIM#300203)
ChAS	Chromosome Analysis Suite
Chr	Chromosome
CMV	Cytomegalovirus
CNS	Central nervous system
CNV	Copy number variation
COMT	Catechol-O-methyltransferase isoform MB (gene, OMIM#116790)
CT	Computerised tomography
Del	Deletion
DGV	Database of genomic variants(http://projects.tcag.ca/variation)
DNA	Deoxyribonucleic acid
DSM	Diagnostic and statistical manual of mental disorders
Dup	Duplication
EEG	Electroencephalography
FOXG1	Fork head box G1(gene, OMIM#164874)
GLUT1	Glucose transporter 1 (gene, OMIM#606777)
IQ	Intelligence Quotient
IRAK1	Interleukin-1 receptor-associated kinase 1 (gene)
LICAM	L1 cell adhesion molecule (gene, OMIM# 308840)
LOH	Loss of Heterozygosity (LOH)
Mb	Mega base
MEC2	Methyl CpG binding protein 2 (gene, OMIM#300260))
MLPA	Multiplex ligation dependent probe amplification
MR	Mental retardation
MRI	Magnetic resonance imaging
NCBI	National center for biotechnology information (database,http://www.ncbi.nlm.nih.gov/)
Nt	Nucleotide
OMIM	Online mendelian inheritance in man (database, http://www.omim.org/)
PARK2	Parkin, an E3 ubiquiting protein ligase protein ligase (Gene, OMIM#602544)
PubMed	The U.S. national library of medicine (database,http://www.ncbi.nlm.nih.gov/)
RBFOX1	RNA-bindning protein FOXG1, C.elegans, homolog1(gene, OMIM#605104)
RNA	Ribonucleic acid
RTT	Rett syndrome
SIDS	Sudden infant death syndrome
SMR	Severe mental retardation
SNP	Single nucleotide polymorphism
SRO	Shorter region of interest
TBX1	T-box 1 (Gene,OMIM#602054)
UCSC	University of California, santa cruz (Database, http://www.genome.ucsc.edu/)
VOUS	Variants of uncertain clinical significance
XLID	X-Linked Intellectual Disability(Database, http://www.ggc.org/research/moleculargstudies/xlid.html)
XLMR	X-Linked mental retardation

Basic Genetics

All living organisms, from prokaryotic microorganisms such as bacteria to multi-cellular eukaryotic organisms; animals and plants, carry their instruction for how to develop and function in the Deoxyribonucleic acid (DNA) molecules. The inherited information for each organism is coded by its DNA, and the DNA is located in the nucleus in the eukaryotic cells. The DNA is composed of a double-stranded polymer and consists of combination of four bases (nucleotides); adenine (A), guanine (G), cytosine (C) and thymine (T). In each strand, the nucleotides are linked together by covalent phosphodiester bonds that join the 5´ carbon of one deoxyribose group to the 3´ carbon of the next. The two DNA strands form a double helix by the complementary hydrogen bonds between A-T and G-C base pairs. The DNA double helix was first described by Watson and Crick in 1953.[1]

DNA and Genes

According to the official guideline for Human Gene Nomenclature (HGNC), a **gene** is defined as: "A DNA segment that contributes to phenotype/function. In the absence of demonstrated function a gene may be characterized by sequence, transcription or homology".

Genes are spaced at irregular intervals along the DNA sequence and take up only 1-2% of the genome; the complete human genome sequence has been estimated to contain approximately 26,000-40,000 genes, by recent publications supported by the International Human Genome Sequencing Consortium.[2,3] The remaining DNA encodes for non-translated RNA or constitutes regulatory sequences, introns, repetitive sequences, pseudogenes and most regions have unknown function, at least at present. The classical view of a gene structure is that it is composed of **exons** and **introns**. Besides exons and introns, a crucial regulatory element is the **promoter,** which is a short DNA sequence placed upstream (5´) of the gene. The promoter is recognized by transcription factors which initiate the transcription. The exons are the segments that contribute genetic information to the final product by coding for amino acids which build up the protein.

Exons are separated by introns which are noncoding segments that are processed through RNA splicing, which removes intronic RNA segment during transcription. RNA splicing requires recognition of the intronic and exonic sequences (Figure 1). The dinucleotides at the ends of introns are highly conserved, and as a rule the introns start with a GT and end with AG (the *GT-AG rule*).

The exons consist of triplets, called **codons** which are translated to proteins by use of the "universal" genetic codes. Also, the transcription always begins at a specific start codon "ATG" and finishes with stop codons "TAA, TAG and TGA" (Figure 1).
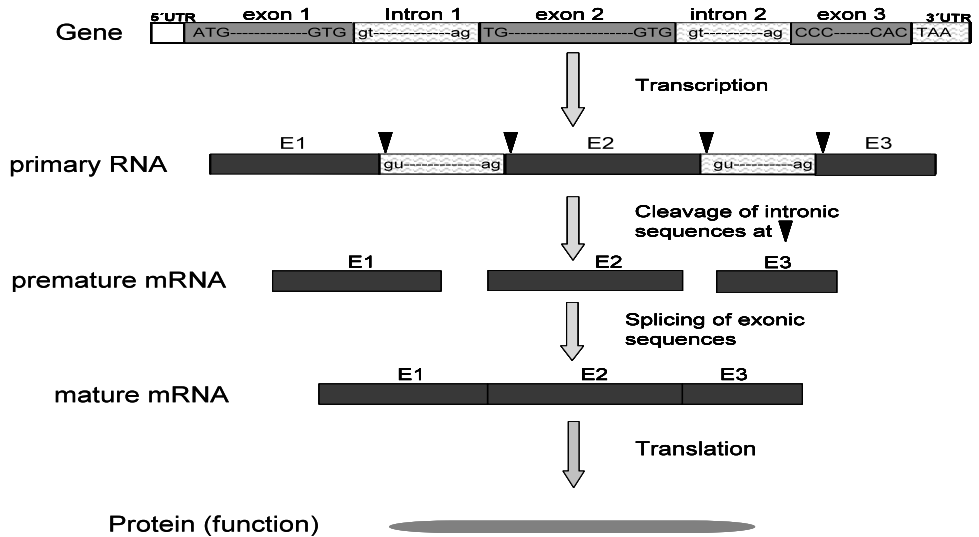
Figure 1   Schematic figure of the process from gene to protein. The gene is transcribed to primary RNA, which is a continuous RNA copy of the gene with both exons and introns. The primary transcript is then cleaved at the exon-intron bounderis. The intronic sequences are snipped out and discarded in the premature mRNA. The maturation of mRNA is complete when the exonic seqences are fused together (spliced) in the same linear order as in the gene. Finally, the mRNA is used as a template for protein which actually carries out the function.

The central dogma of molecular biology

The hybridization of DNA and RNA allows for accurate copying of the genetic material either prior to cell division (***replication***) or for the production of RNA (***transcription***) and protein (***translation***) from specific genes in the DNA (Figure 2). The replication has a semi-conservative approach; the two strands of DNA separate and new strands are synthesized onto the existing strands. The resulting DNA molecules are hybrids of new and old DNA.[1,4] Genetic information generally flows in a one-way direction in the sequence: DNA is decoded to make RNA, and then RNA is used to make polypeptides that are subsequently forms of protein. This flow of genetic information has been described as the central dogma of molecular biology, because it is universal among most of living organisms (Figure 2).
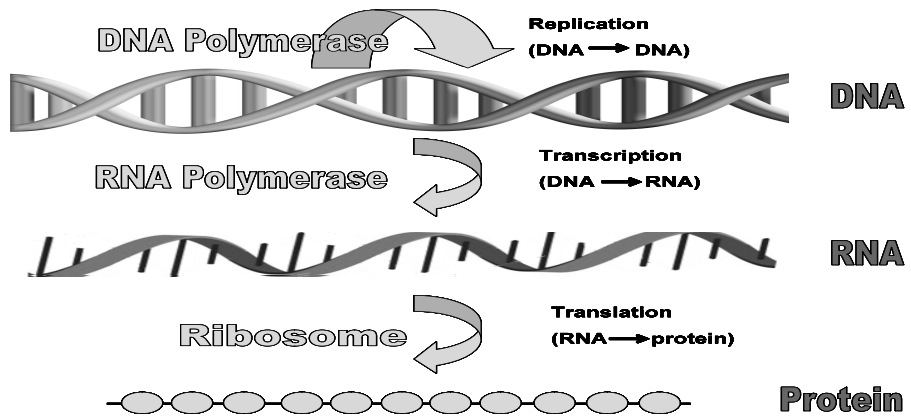
Figure 2   The central dogma. The cell division begins with DNA *replication* when the DNA polymerase in a semi-conservative approach makes two DNA double strands the old DNA strand as a template.The process of genetic information is essential in two steps: 1) *transcription*, which happens in the nucleus (in the eukaryotic cells), where DNA is used as a template for RNA synthesizing by RNA polymerase enzymes. 2) *Translation*, which takes place in ribosomes in the cytoplasm, where the mRNA are decoded to polypeptides.

Human Genetics

 The human genome consists of more than 3 billion base pairs (bp), as the cells enter mitosis the chromatin condenses to chromosomes.[5] The entire genome is then organized in 23 chromosome pairs, in which 22 pairs are autosomes and one pair, consist of the sex chromosomes; X and Y (Figure 3A). There are two copies of each autosome (chromosomes 1-22) in both females and males. The sex chromosomes are different; there are two copies of the X chromosome in females, but males have a single X chromosome paired with a Y chromosome. Chromosomes are generally numbered in order of decreasing size; the largest chromosome is 1, followed by chromosome 2 and 3 etc. The sex chromosomes X and Y and also chromosome 21 which is smaller than chromosome 22 are the exception to this rule (Figure 3A).

The short arm of each chromosome is denoted *p* and the long arm *q*. The specific chromosomal locations have traditionally been referred to the banding pattern obtained by Trypsin/Giesma staining of mitotic chromosomes (Figure 3B) (ISCN; International Standing committee on human cytogentic nomenclature 1995) according to this system the chromosomal bands are numbered from the centromere and outwards along the chromosome arms. Recently, the principle has changed to some extent and the position of a particular gene is now commonly given in mega bases (Mb), as calculated distance from the p-terminal of the chromosome.
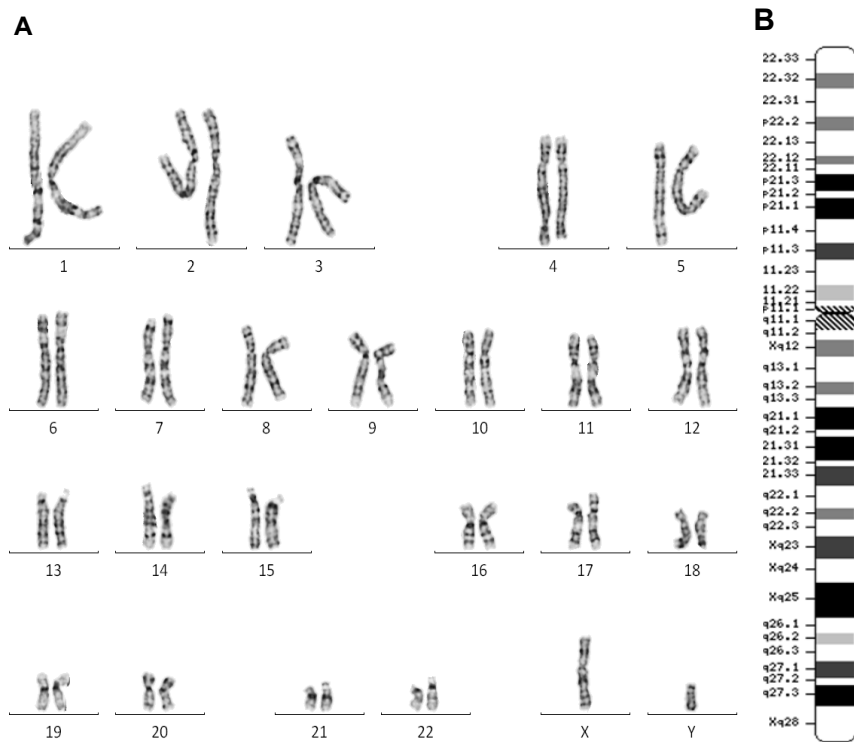
Figure 3 **A)** A Human normal male karyotype showing 46 chromosomes, 23 pairs. Karyotype kindly provided by Ylva Sirenius, Department of Clinical Genetics, SU/Sahlgrenska. B) Schematic picture of the X chromosome, displaying the banding pattern obtained by Trypsin/Giesma staining (NCBI).

Genetic variation and diseases

Generally, there are two main types of aberrations to be assayed for genetic diseases and disorders:

 1. Small scale alteration at DNA level such as basepair substitutions (also refferd to as ***point mutations),*** and small ***insertions*** and ***deletions*** in specific genes. These variations often affect the gene product; the protein. The greatest source of mutations is spontaneous errors in DNA replication and repair (Strachan & Reed 4th edition).[6] New mutations arise in the somatic cells or in the germline of an individual. Germline mutations can be pathogenic (disease causing) and are inheritable, which means that they can be transmitted to offspring of the individual. In contrast, most somatic mutations are neither inheritable nor pathogenic since only single cellclones are affected (with exceptions in cancer). On the other hand, it is common that more than one nucleotide can localize at the same genomic position in the population, this phenomenon is called ***single nucleotide polymorphisms*** or ***SNPs***. However, about one nucleotide of 300 is polymorphic. These SNPs are catalogued in the public dbSNP database; (hTTP://www.ncbi.nlm.nih.gov/projects/SNP).

2. Large scale aberration at a chromosomal level; chromosomal abnormalities are classified in two types: ***Constitutional abnormalities*** which occur very early in development (abnormality in sperm or egg and maybe abnormal event in the very early embryo) resulting in presence of the aberration in all cells of body. The abnormality is *somatic* if it is only present in clones of cells or specific tissues. An individual with a somatic abnormality is called to be ***mosaic*** when there are two populations of cells with different chromosomal constitution; both derived from the same zygote. Chromosomal alterations, whether constitutional or somatic, are mostly divided in two categories; A) ***Numerical abnormalities*** as trisomy; three copies of a chromosome instead of two copies (for example trisomy 21 in Down syndrome; e.g. 47,XX,21) or monosomy, when one of the chromosomes is missing (for example monosomy X in Turner syndrome; 45,X). B) ***Structural abnormalities*** such as ***translocations***, or ***gains*** and ***losses***. These types of aberrations, with the exception of balanced translocations, can also cause ***copy number variation*** (***CNV***) in the affected chromosomal regions and can be identified using a variety of cytogenetic methods or copy number assays (see also page 21).

Genetic aberrations include "epigentic" mechanisms which may affect the phenotype in genetic diseases. Epigenetics is commonly defined as a mitotic and/or meiotic heritable change of gene expression caused by other mechanisms than variations in the underlying DNA sequence. ***DNA methylation*** and ***histone modification*** are the most common epigenetic mechanisms that affect gene expression. The most common epigenetic change is the X chromosome inactivation in females.[6]

Mendelian inheritance

Monogenic inherited traits are caused by single locus variations in the genome, which are transmitted through ***Mendeline inheritance.*** Named after the father of modern genetics; George Mendel who in the 1860 grounded the principle of inheritance pattern. The monogenic inheritance is called ***dominant*** if only one allele of two possible alleles in the locus leads to demonstration of a certain phenotype. But the inheritance is ***recessive*** if demonstration of the phenotype is depending on the presence of the same alleles in both positions in the locus. Nowadays, many thousands of Mendelian traits or diseases are known and information about them have been catalogued in the OMIM (Online Mendelian Inheritance in Man) database; (http://www.omim.org/).

X chromosome inactivation

The human X chromosome carries many essential genes expressed during the development. In contrast, the Y chromosome in males carries very few genes; which are mainly related to male sexual functions. It means that females have twice as many of the X chromosomes genes as males. X chromosome inactivation (XCI) is the modulator for this excess of genes in females. Due to XCI one of the two X chromosomes in all cells in women undergo inactivation. It is a random process which occurs early in embryogenesis (probably in the 10-20 cell stage), in each cell independent to other cells.[6] The XCI is irreversible during lifetime of the cell and the particular inactivated X chromosome remains inactivated in all its daughter cells. Thus, all females are mosaics of clones in which different X chromosomes (either paternal or maternal inherited X) are expressed.

Inactivation is usually ***random*** which means that one of the X chromosomes is inactivated in 50% of a female's cells and the other X chromosome is inactivated in the remaining cells. X chromosome inactivation is considered as ***skewed*** in some conditions if the ratio of active to inactive X is less or equal to 75:25.[7]

This process may explain why female carriers of recessive X-linked conditions (as the *MECP2* duplication syndrome in this thesis) only display minor signs of the phenotype, and also why females, heterozygous for a dominant X-linked condition, usually show milder and more variable phenotype than males with the same X-linked condition.

BACKGROUND

Mental Retardation

Mental retardation (MR) is characterised by significant limitations in intellectual function and adaptive behaviour. It is estimated that MR affects up to 3% of the European population.[8] MR is one of the most important unsolved clinical determinants in healthcare.[8] The ability of cognitive functions is commonly determined by the definition of the **Intelligence Quotient (IQ)**. Assuming a population mean IQ of 100, MR can be subdivided into four degrees of severity: mild (IQ 50–55 to approximately 69), moderate (IQ 35–40 to 50–55), severe (IQ 20–25 to 35–40), and profound (IQ below 20-25)according to World Health Organization (TheICD-10 classification of mental and behavioral disorders.WHO: 2001).

The diagnosis of MR is established based on three following criteria:[8]

1. Onset of MR symptoms before 18 years of age.
2. Intellectual functions significantly lower than average (IQ < 70).
3. Poor adaptive skills in at least two of the following areas: communication, self-care, social/interpersonal   skills,   self-guidance, school performance, work, leisure, health, and safety.

MR is exceptionally heterogeneous and complex in its aetiology, which is predominantly genetic but can also be environmental (for example fetal alcohol syndrome). Environmental factors such as perinatal hypoxia or infection and teratogenic agents (such as viruses or chemicals that cross the placenta during pregnancy), as well as genetic variations such as chromosomal aberrations, known microdeletion/microduplication syndromes and point mutations in specific gens, all have been implicated to cause MR and dysmorphology.

Genetics of MR

It is estimated that 25–35% of patients with mental retardation might have a genetic background.[8] A search of MR in the online Mendelian Inheritance in Man database (OMIM) has a score of 2726 hit of known genetic condition. Many of these are chromosomal **syndromic** disorders such as Down's syndrome (trisomy 21) and rare syndromes which result from submicroscopic deletions (like 22q11 deletion syndrome) or duplications. These conditions are usually associated with specific congenital abnormalities and dysmorphology.[9] There are also some syndromic disorders which have known and well described inclusion criteria, and are associated to pathogenic variations in one or several known gene/genes, such as Rett syndrome (associated with the *MECP2* gene).

On the other hand, there is a group of **non-syndromic** MR patients who have no specific symptoms, dysmorphology or other congenital features which enable a clinical diagnosis. Although there are state-of-the-art strategies implicated in the genetic testing of MR patients, the majority of these patients still lack aetiological diagnosis.[10]

X-linked mental retardation

The estimated male:female ratio is ranging from 1.3:1–1.8:[1,8,11-14] and the over all male dominance for mental retardation has been well described in several studies. The prevalence of MR in males, up to 30% higher than for females, may be explained by the large amount of **X-Linked mental retardation (XLMR)** genes reported in literature.

More than 200 conditions and defects and 109 XLMR genes are listed on the website of XLID (Greenwood Genetic center, http://www.ggc.org/research/moleculargstudies/xlid.html, updated November 2012). The inheritance pattern of X linked genes in males is different from inheritance of genes located on the autosomes, since there is just one allele for each gene on the X chromosomes. Intellectual disabilities, predominantly affecting males, are commonly due to pathogenic variations in XLMR genes, which probably are partly responsible for the higher prevalence of MR among males compared to females.[15,16] An example is the *MECP2 duplication syndrome* which affects male patients very severely but the same syndrome has variable expressivity in females, from asymptomatic to mild or moderate mental retardation.[7] Accordingly, some X-linked conditions are not compatible with life for males. An example is *Rett syndrome* (*RTT*) which almost exclusively affects only females. However a few male patients with *MECP2* mutation have been reported. The male phenotype range from sever congenital encephalopathy, mild MR and various neurological symptoms.[17,18]

Rett Syndrome

History

Rett Syndrome (RTT) is named after Andreas Rett, an Austrian paediatric neurologist, who first recognised the characteristic features of the syndrome. In 1966, Rett described similar findings in 22 patients, for the first time as a unique clinical entity [Rett,1966, in German]. Meanwhile, in 1960, young female patients in Sweden with quite similar symptoms were observed by Bengt Hagberg. RTT became recognized in the medical community in 1983, when Hagberg and his colleagues reported 35 cases of RTT (from Sweden, France and Portugal) in the English language.[19] RTT with an incidence of 1:10 000 females, is now recognised worldwide as one of the most common causes of mental retardation in girls.[20]

Clinical features of Rett Syndrome

Rett syndrome diagnosis is based on several characteristic clinical criteria.[21] RTT (MIM 312750) is a postnatal progressive neurodevlopmental disorder characterised by normal development up to the age of 6-18 months. Then the development stagnates, this is followed by rapid deterioration of motor development, autistic behavior, loss of purposeful use of the hands, jerky truncal ataxia, microcephaly and epilepsy.[19,20,22,23] Rett syndrome is clinically divided to two subgroups:

1) *Classical Rett syndrome,* which is the most typical form of Rett Syndrome (80-90% of RTT diagnosed patients).[18] The development of classical RTT is divided in four stages:[24]

Stage I) Early onset stagnation: which arise after 6-18 month of normal development, when the girls almost stop to acquire new skills, head growth decelerates with onset of autistic behavior.

Stage II) Rapid developmental regression (onset age 1-4 year) with loss of previously acquired skills.

Stage III) Pseudo stationary period (onset age 4-7 years), in this stage most girls regain some of activities and improve eye contact.

Stage IV) Late motor deterioration (in the following 10-12 years), lower motor neurons impairment becomes prominent resulting in that the most adult RTT patients are restricted to wheelchairs. They are severely mental retarded and develop scoliosis.

2) **Variant Rett syndrome,** deviates from the typical form in different aspects: there are several atypical types of RTT, some are milder as forme fustes (FF), preserved speech variant (PSV) and late regression[24,25] and some are more severe as congenital and early onset infantile variants. [24,26] Beside of RTT variant girls, there are male patients with RTT-like features which attracting an increasing research interest.

Genetic basis of Rett syndrome

In parallel with the worldwide recognition of RTT in the 1980s, For the first time, a connection was established between DNA methylation and heritable effects of gene expression.[27] Identification of the site of almost all DNA methylation in mammalian genomes as the dinucleotide 5´-CG-3´ (known as CpG), was the beginning of exploration of the effects of this modification on gene activity. Two mechanisms have been suggested for methylation-mediated gene repression. The first suggests that methylation of CpG sites (called CpG islands) within gene promoters will inhibit sequence-specific binding of transcription factors. In the second, more prevalent mechanism, the repression is mediated by proteins which specifically bind to methylated CpGs (methyl-CpG binding proteins) and thereby alter the chromatin structure, rendering it inaccessible to the transcription machinery.[28] In 1992, Adrian Bird and coworkers identified a novel mammalian protein that binds methylated CpGs, methyl-CpG binding protein 2(MeCP2).[28]

The gene behind RTT was identified in 1999, by Amir et al. They reported the first cases of RTT patients with mutations in the *MECP2* gene located on chromosome Xq28.[29] They suggested the abnormal epigenetic regulation as the underlying mechanism for the pathogenesis of RTT.[29] However, pathogenic variations in *MECP2* are identified in more than 95% of classical RTT cases. The variations include missense, nonsense, and frameshift mutations, which mostly are *de novo* of paternal orignin and often involve a C to T transition at CpG dinucleotides. [18,30,31] Detection of larger deletions in *MECP2* became possible when the MLPA method was introduced in the *MECP2* analysis.[32] The most common deleterious variations (refered to as "**hotspot**" mutations in the *MECP2* gene) as well as gross deletions (including one or several exons) are presented in the RettBASE: IRSF MECP2 Variation Database (http://mecp2.chw.edu.au/), together with phenotype-genotype correlations.

Early infantile onset Rett Syndrome variant

In this thesis, we have focused on genetic analysis of the "early infantile onset RTT variant" (paper I and III). This variant may be more severe and the disease debut is in very early infancy; before 6 month of age. At older age, the symptoms appear more and more like the classical RTT (Table 1). However, pathogenic variants in *MECP2* are rarely identified in RTT variants.[33]

Other candidate genes have been suggested to be relevant for some of the RTT variants, like congenital RTT variant, namely the Cycling Dependent Kinase-Like 5, (*CDKL5*) gene located at Xp22 and the Fork head box G1 (*FOXG1*) gene located at 14q12.

Table 1   Early onset infantile Rett syndrome variant clinical criteria[1]

Primary criteria, ≥ 3 required	Supportive criteria, ≥ 5 required
-Loss of acquired fine finger skill or never present	-Breath irregularities (hyperventilation and/or breathholding)
-Loss of learned single words/phrase	-Bloating/marked air swallowing
-RTT hand stereotypies, hands together or apart	-Characteristic RTT teeth grinding, Gait dyspraxia/apraxia
-Early deviant disturbed communicative ability	-Neurogenic scoliosis (high kyphosis, ambulant)
-Deceleration of head growth	-Appearing abnormal lower limb neurology
-The RTT disease profile:	-Small blue/cold feet
a regression period (stage II)	-Unmotivated sudden laughing/screaming spells
followed by a come back (stage III)	-Impaired/delayed nociception indicated
contrasting to slow neuromotor regression	-The RTT characteristi eye pointing

NOTE: RTT= Rett syndrome
[1]Unpublished data from paperI

MECP2; one gene, several diseases

The human *MECP2* gene consists of four exons resulting in expression of two protein isoforms due to alternative splicing of exon 2. The splice variants differ only in their N-terminus, and include the more abundant MeCP2-e1 isoform (encoded by *MECP2B*) as well as the MeCP2-e2 isoform (encoded by *MECP2A*).[34] Since recognition of a causative relationship between *MECP2* and RTT, many research laboratories have tried to find out the key molecular signaling pathway of *MECP2* and its expression pattern using different genetic mouse models. The impressive large amount of different types of the *MECP2* manipulated mouse models; including *Mecp2* null mice, models with various defected (mutated) *Mecp2* and also mouse models with over expressed *Mecp2*, have contributed to the knowledge about *MECP2* today. In brief, using genetic mouse models, key molecular signaling pathways that contribute to the deficits in synaptic function, the numbers of synapses and dendrites as well as maturation of neurons have been observed.[35-39] Once identified, these mouse models have also been used to experimentally validate possible therapeutic avenues using genetics, pharmacological, and behavioral approaches.[40] However, the results of these studies suggest the MeCP2 protein function as a genome-wide modulator with global effect on many other genes.[40-42] The MeCP2 protein is an important member in a complex cascade of reactions that regulates other genes as *BDNF, IGF* and *Dlx5/6* by interaction with other factors as Sin 3-histone deacetylase complex and other coactivators. [43,44] Interestingly, recent results suggest that MeCP2-e2 isoform is upregulated in Aβ-treated cortical neurons and promotes neuronal death in post-mitotic neurons, a pathway normally inhibited by fork head protein *FOXG1*.[45] This finding may explain the role of pathogenic variants in *FOXG1* which have been reported in congenital RTT variants.[46-49]

Furthermore, *MECP2* mutations have been reported in patients with diagnosises different from RTT: Autism (1-2%, mostly in the C terminal of gene), mild MR, Angelman syndrome-like (which has overlapping criteria with early infantile onset RTT) and also juvenile onset schizophrenia.

The dosage of MeCP2 has phenotypic impact as well, a mouse model with over-expression of a trans-gene containing the human *MECP2* gene, that showed a near twofold MeCP2 expression, demonstrated severe progressive neurological effects.[50] The effect of *MECP2* over-expression has also been observed in human, where a double dosage of *MECP2* causes severe developmental delay and mental retardation.[51-54]

MECP2 duplication syndrome

In 1999, Lubs et al. presented a family with severe XLMR, hypotonia and a mild myopathy.[51]Three of the five affected males in this family died of secondary complications before the age of 10 years and none have survived past the age of 10. These complications included swallowing dysfunction and gastroesophageal reflux with secondary recurrent respiratory infections. He also described specific characteristic features in this group, including downslanting palpebral fissures, hypertelorism, and short nose with a low nasal bridge.[51] Interestingly, three obligate carriers in Lubs study had an IQ less than 80. The suggested localization for the causal gene was distal to DXS8103 in Xq28.[51] Along with the progress of molecular genetic technologies, several male cases of Xq28 duplications including *MECP2* have been detected (by for example microarray).[52-54] The duplicated regions differ in size and are usually inherited from unaffected mothers. Genotype-phenotype correlation studies suggest that the minimal duplicated region required for the syndromic specific phenotype in males include the entire *MECP2* coding sequence and the adjacent *IRAK1* gene (see also figure1, paper IV). The *MECP2* gene has been suggested to be the dosage sensitive gene that mediate neurological outcome.[53,55] The prevalence of *MECP2* duplication syndrome (OMIM 300260) is estimated to 1% of patients with unexplained XLMR[56-58], and 2% of male patients with severe encephalopathy.[59] The clinical manifestation in male patients is now updated and includes a variety of moderate to severe mental retardation, hypotonia, delayed or absent speech, epilepsy, late onset spasticity as well as feeding difficulties and recurrent respiratory infections.[60-62] Nevertheless, female carriers with almost completely skewed XCI in peripheral blood have been considered as asymptomatic.[61,63,64] Later on, affected females with X:autosome translocation resulting in *MECP2* duplication have been reported.[65] Eventually, other cases of females with *de novo MECP2* duplications were observed.[66-68] The phenotype in females differs from the male phenotype, in females it is broad and ranges from asymptomatic to moderate MR, with learning disabilities, anxiety, slightly dysmorphism and autistic features. The minimal duplicated region required for appearance of phenotype is smaller in females compared to males and includes only *MECP2* (paper IV).[66-68]
However, further studies of this syndrome are needed to clarify the inclusion criteria, especially in females. In female cases other factors as X chromosome inactivation and incomplete penetrance may also be important phenotypic modifiers.

Microdeletion and microduplication syndromes

Genotype, phenotype, penetrance and variable expression

As discussed earlier in chromosomal aberrations (see Page 15), all individuals carry copy number variations (CNVs), also refered to as microdeletions or microduplications, in the genome (refers to genotype). Most common CNVs are normal variations. The effect on the

phenotype depends on several factors; for example the size of the variation, the gene content, or other functional elements located in the region of interest. International guidelines for interpretation of CNVs have been published by Miller et al., 2010.[69] SNP microarrays and DNA sequencing technologies have increased the resolution, thus made it far easier to identify smaller CNVs. Some well known variants can cause very specific phenotypes, as most microdeletion syndromes, while others, often microduplication syndromes, may be associated to variable expressivity.[70] Only genes which are dosage dependent are influence the phenotype (for example haploinsufficient genes).

The probability to display a specific phenotype due to a genetic variation is what is usually referred to as the ***penetrance***. On the other hand, variable expressivity has been described for many syndromes such as 22q11 duplication syndrome, 16q11 duplication syndrome etc. These syndromes present with very broad phenotypic spectra, which even include carriers with very mild symptoms, which are difficult to identify in a clinical setting. It is not unusual that the carriers with mild symptoms only are identified after the diagnosis has been established for a family member with more evident symptoms.

An example is the 22q11.2 deletion that causes a multi-systemic disorder, the 22q11.2 deletion syndrome (also known as velocardiofacial syndrome or DiGeorge syndrome). Approximately 93 % of the patients have a *de novo* deletion, and 7 % have inherited the deletion from a parent, sometimes mildly affected. Many patients with the deletion have learning disorder, congenital heart defects, malformed palate, mild facial abnormalities, neuropsychiatric illness, and sensitivity to infections.[71] The phenotype is variable and may include facial dysmorphism, cardiovascular abnormalities, short stature, cognitive and behavioural impairments and a high risk of schizophrenia.[71,72] Causative genes in the 22q11.2 region include *COMT* (catechol-O-methyltransferase isoform MB- COMT)[73] and T box1; *TBX1* gene.[74]

Duplication of 22q11.2 also has variable clinical presentation, including developmental delay, dysmorphic facial features, autism, and cognitive and behavioural impairments[75], but here the ratio of inherited duplications is almost 70 %. The duplication syndrome has incomplete penetrance, and a parent who has 22q11.2 duplication can have a normal or near-normal phenotype. Because of incomplete penetrance both parents need to be tested to distinguish from *de novo* cases.

In conclusion, interpretation of CNVs is an important stage in diagnostic genetics. Access to the phenotype and liable and strong evidence for correlation between the genotype of interest is crucial. To date, many useful tools have been provided by different international communities with genetic expertise to enable the interpretation. However, there are still variants which are difficult to interpret as data may be inconclusive or missing.

The Database of Genomic Variants (DGV) from Centre for Applied Genomics ; (http://projects.tcag.ca/variation), is the platform where many of the normal variants from the populations and the frequency of them have been collected together. DECIPHER database (DatabasE of Chromosome Imbalance and Phenotype in Humans using Ensembl Resources; https://decipher.sanger.ac.uk) from Welcome Trust Sanger Institute, is another databases that is essential for comparison of detected variations with reported patients. DECIPHER is also a phenotypic database, it is a comprehensive database which incorporates data from several

databases, the information includes, reported variations together with related phenotype, common copy-number changes in healthy populations are also displayed and genes of recognized clinical importance are highlighted. These databases and also searching in OMIM (Online Mendelian Inheritance in Man) home pages and PubMed (the U.S. National Library of Medicine; http://www.ncbi.nlm.nih.gov/), are essential for medical validation.

Other databases as *Unique* database (http://www.rarechromo.co.uk/) are available with description of both old and new microdeletion/microduplication syndromes and the related criteria.

MATERIALS AND METHODS

Patients and controls

In this thesis, we have studied two groups of patients with genetic analysis. The first group consists of 14 patients with early infantile onset Rett Syndrome variant which have resulted in paper I and III. The patients in this study with early infantile onset RTT variant phenotypes (young girls as well as adult women) have been examined and collected over 40 years by professor emeritus Bengt Hagberg. Most of the women have been systematically examined and diagnosed according to the previously published clinical diagnostic programme.[25]

The other group consists of 133 patients with severe mental retardation (SMR), these children have been epidemiologically collected and categorised in different categories. Twenty-three patients in this group were categorised in "the prenatal unknown aetiology" subgroup. We have analysed 19/23 of these patients with microarray technique in paper II. Two brothers in this group had the same maternally inherited X-linked aberration; the brothers have been further studied and compared with a female patient with a partly overlapping genetic aberration in paper IV.

In addition, a reference set of 45 in-house controls with common CNVs in the Western Swedish population was used to exclude common benign CNVs in paper II-IV.

Paper I and III

In paper I we have studied 14 cases with early infantile onset variant of Rett syndrome. The women's phenotypes were collected from medical history and examination of the patients. For inclusion the patients had to have shown definitely abnormal signs or deviant developmental profiles before the age of six months. All fulfilled three or more Rett variant criteria and five or more supportive criteria.[22] Four had documented epilepsy, five had no seizures and five had missing information on seizures.

12 patients from study I remained without causative genetic aberration, in paper III we have further studied these 12 patients for copy number variations.

Paper II and IV

In paper II, children with SMR in the Swedish county of Halland area were investigated. 133 children with SMR between 6-17 years old, born in 1987-1998 living in the county at December 31st, 2004 and registered at the habilitation centres, paediatric clinic and school health services were identified. The children were examined with CT, MRI, metabolic screening, karyotyping and further genetic analyses when indicated at the clinical examination. Classification of aetiologies was made similar to those used in previous Scandinavian studies.[12,13,76] Patients were categorised in 4 main aetiologies; ***prenatal*** in 82 (62 %), ***perinatal*** in 14 (10 %) and ***postnatal*** in 8 (6%) children in relationship of birth time. In 29 children (22 %) it was not possible to relate the condition to the time of birth, and was categorised as ***undetermined timing***, 24 of these children were males with autistic disorders. Each category was then divided in sub group according to the diagnosis of SMR. The prenatal aetiology is divided to ***genetic, acquired*** and ***unknown*** aetiology subgroups. The prenatal genetic group were based on the results of the clinical observations as well as

genetic analysis and includes 57 patients in which 34 with **chromosomal abnormalities** such as Down syndrome and 6 patients diagnosed with other known chromosomal aberration detected with traditional karyotyping; (marker chromosome 15q, partial trisomy 11, partial monosomy 2q, mosaic trisomy 8, Pallister-Killian syndrome and deletion 3p).

The other group encompasses 14 patients with **monogenic** *or* **presumed monogenic disorders**: (tuberous sclerosis, Angelman syndrome, Cornelia deLange syndrome, Laurence-Moon-Biedl, infantile neuroaxonal dystrophy and Rett syndrome). Three girls were diagnosed with **metabolic disorders** and five cases had **malformation syndromes** as diagnosis. The **Acquired** group comprised two children, both with spastic tetraplegic. Microarray analysis was implicated for propose of determination of aetiology in 23 children with unknown prenatal aetiology, and 19/23 patients were subsequently analysed by microarray in the paper II; (for 4 patients in this group, blood samples were not available). In paper IV, we have analysed a female patient with mild to moderate mental retardation who had a duplication at chromosome band Xq28, which the two brothers in paper II also shared and we have here described the female phenotype in comparison to male phenotype.

Molecular Genetic Methods

Mutation detection methods

In order to detect small scale genetic aberration in patients in this thesis, we have used three polymerase chain reactions (PCR) based methods in paper I and III. PCR is an *in vitro* method for amplification of DNA sequences using oligonucleotide primers.

DNA sequencing

DNA sequencing (developed first by Fredrick Sanger, 1977) is the most widely-used analysis method to analyze the base composition of stretches of DNA. We used terminator chemistry DNA sequencing in paper I and III. In order to ensure high quality of sequencing results, the PCR products were purified with for example enzymes or spin columns. Then, the purified PCR products were added to the sequencing reaction, which requires a mix consisting of buffers, dNTP:s, ddNTP:s, sequencing primers (one primer at a time, either forward primer (5´→3´) or reverse primer (3´→5´)) and enzyme for elongation of the sequencing product. The ddNTP:s (A, C, G or T) are labeled with a fluorescent molecule which results in one specific color for each one of them. The labeled ddNTP:s were also modified without a hydroxyl group on the 3´carbon which interrupts the elongation procedure. As both dNTP:s and ddNTP:s were randomly involved in the sequencing process, it results in all possible lengths of sequence products separated by one base in the final mixture. These products were separated according to the product length through electrophoresis, the negatively charged DNA fragments migrate to the positive anode in the electric field and the camera recognized the different florescence dyes and the software processed the results into electropherograms. In this work we used the ABI 3100 instrument (Applied Biosystems) and the Seq Scape collection program (Applied Biosystems) to final analysis of obtained DNA sequences.

Denaturing High Performance Liquid Chromatography (DHPLC)

In summary**,** DHPLC is a mutation screening method based on liquid chromatography for detection of heteroduplexes in PCR amplified fragments. We applied this method in paper I for analysis of *CDKL5* gene. Melting profiles and DHPLC run conditions can be determined by Navigator Software (Transgenomic). PCR products are denatured and slowly renatured, which create perfectly matching homoduplexes (wild type) and heteroduplexes (mismatched bp due to mutation), the hetroduplexes have lower melting temperature than homoduplexes which results in different peak-patterns from wild-type to mutant in the chromatogram.

Multiplex Ligation-depended Probe Amplification (MLPA)

MLPA is a multiplex PCR method which can be used for the relative quantification of DNA (or RNA).[77] It can be used to detect moderate to large insertions, deletions, duplications and copy number variations in targeted genes or regions. The MLPA reactions comprises of 5 steps: 1) Denaturation of DNA and hybridization of specific unique MLPA probes (each probe is designed in two parts). 2) Ligation using Ligase enzyme to fill the gap between the two parts of each hybridized probe. 3) Multiplex PCR reaction using universal primers which amplifies all amplicons in one reaction. 4) Separation of amplification products according to their unique length in each MLPA kit by electrophoresis. 5) Data analysis. We used the Gene Scan Analysis Software v3.7 (Applied Biosystems) and the Sequence Pilot version 1.2 software (JSI medical system). The *MECP2* MLPA kit P015 from MRC Hollad was used in paper I to detect deletion in *MECP2* and also to verify the result of *MECP2* duplication syndrome in paper IV. In addition MLPA was used to verify some of our findings in paper II (22q11 deletion).

SNP array

Single nucleotide polymorphisms (SNPs) are the most common source of genetic variation in the human genome, and refer to genomic positions where two or more bases are found in populations. The frequency of about 10 million, evenly dispersed across the genome, makes SNPs a suitable marker to copy number analysis.[78] Early studies with chromosomal karyotyping (with low resolution) and fluorescent in situ hybridization (FISH) were limited in their ability to detect only the largest CNVs due to large probes, but increased resolution of SNP microarrays have made smaller CNVs far easier to identify.[70]
In paper II-IV, we have used Affymetrix Genome-Wide Human SNP 6.0 Array and Affymetrix cytogentic Whole Genome 2.7M array to investigate and identify CNVs in our patients. Affymetrix GeneChip™ technology applies a combination of photolithography and combinatorial chemistry with light- directed in situ synthesis oligonucleotides on a glass surface. The Genome-wide Human SNP Array 6.0, has 1.8 million markers which are 25 bases long oligonucleotides probes with 1-5 kb median marker spacing and consist of SNP probes (906,600 in SNP 6.0 array) and Copy number probes (CNP; 946,000 probes in SNP 6.0).

We used Affymetrix cytogenetic Whole Genome 2.7M array in paper II which provides whole genome coverage with a high density of 2.7 million oligonuclotide markers, of which 400 103 are SNP probes. The principle of the SNP array method is shown in Figure 4. Data from Genome Wide Human SNP 6.0 Array was analysed with theAffymetrix Genotyping Consol v2.1 software and then Affymetrix Chromosome Analysis Suite (ChAS) 1.0.1 software and the NetAffx build 32 was used to visualise the patient's copy number variation. ChAS Software is designed for cytogenetic researchers. The provided *in-silico* controls in the software and also in-house controls can be used to excludes the most common CNVs in population. In additions, there is direct linking to useful databases as OMIM, DVG and UCSC integrated in the software that allows researchers to compare and interpret the region of interest with these databases.
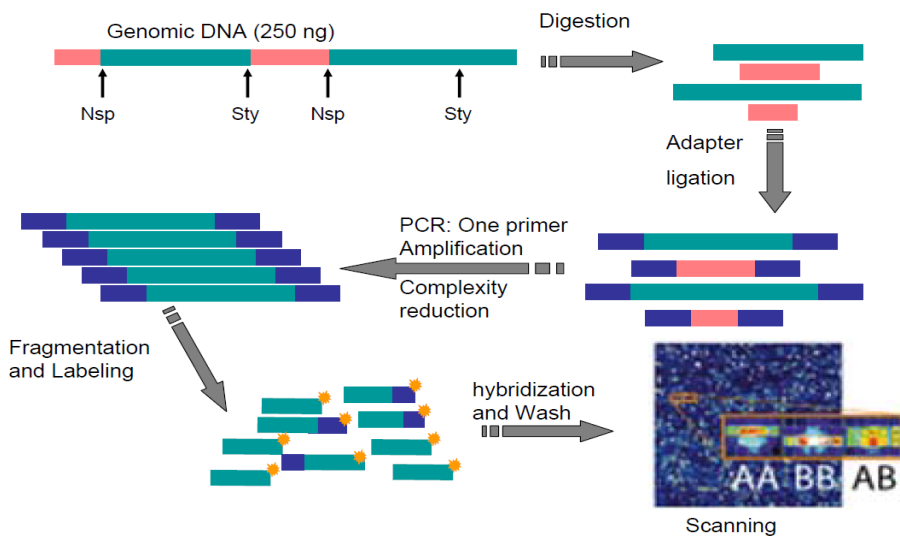
Figure 4   The SNP array reactions step by step are as follows; 1) Digestion of 250 ng of genomic DNA with each restriction enzymes Nsp1 and Sty1(total 500ng DNA). 2) Ligation to adaptors. 3) Generic primers recognizing the enzyme specific adaptor sequences amplify adaptor-ligated DNA. 4) PCR purification with magnetic beads. 5) Fragmentation and labelling of PCR products. 6) Hybridization in the Affymetrix GeneChip® Hybridization Oven 640. 7) Wash and stain in the Affymetrix GeneChip® Fluidics Station 450. 8)Array scan with the Affymetrix GeneChip® Scanner 3000 7G.

RESULTS

Early onset infantile-congenital-Rett syndrome (Paper I and III)

In paper I, we studied all 14 clinically diagnosed early onset infantile RTT patients for aberration in *MECP2* and *CDKL5* genes. One of the patients had a pathogenic partial deletion including 2 exons of the *MECP2* gene, detected by MLPA. Patient 2 in this study had a *de novo* deletion of chromosome 3p, del(3)(pter→3p25.1~25.2), which was detected and described in pervious studies.[79,80] We suggest that the 3p deletion has more likely caused her phenotype, since we have not detected any pathogenic mutations in the *MECP2* or *CDKL5* genes. In addition, the microarray analysis in our following study (paper III) confirmed that the 3p deletion is the only pathogenic CNV in her genome. This deletion contains 12 CNS expressed genes such as *ATG7* and *SLC6A1* which have showed relationship with mental retardation and epilepsy.[79,80]

We did not find any pathogenic mutation in the *MECP2* or *CDKL5* genes in other patients in this series (Table 2).

Table 2   **Mutations analysis and microarray analysis results in early onset infantile Rett syndrome patients in summary: paper I and paper III**

Patient	*MECP2*		*CDKL5*		*FOXG1*	CNV	
	Mutation	**NV**	**Mutation**	**NV**		**Type**	**size**
1	ND	ND	ND	c.*131GA>AT	ND	-	-
2	ND	ND	ND	c.*131GA>AT	ND	Distal 3p deletion[b]	11 Mb
3	ND	ND	ND	ND	ND	-	
4	ND	ND	ND	ND	ND	-	-
5	ND	ND	ND	c.*131GA>AT	ND	6q26 gain[c]	309 Kb
6	ND	c.[1035A>G (+) 1233C>T]	ND	ND	ND	-	-
7	ND	992A>G	ND	c.*131GA>AT	ND	-	-
8	ND	ND	ND	ND	ND	-	-
9	ND	ND	ND	ND	ND	16p13.2 deletion[c]	102 Kb
10	ND	ND	ND	ND	ND	6q26 gain[c]	191 Kb
11	ND	ND	ND	ND	ND	-	-
12	ND	ND	ND	ND	ND	-	-
13	ND	ND	ND	ND	ND	-	-
14[a]	Del exon 1-2	ND	-	-	-	-	-

NOTE: NV = normal variant; ND = not detected; Del = deletion

[a]*CDKL5* analysis has not been performed for this patient since we identified a large deletion in *MECP2*, we hypothesis that this deletion might cause her phenotype.

[b]distal 3p deletion in pervious study by FISH.

[c]Intragenic CNV

Normal sequence variants (SNPs) were detected in 6 patients, 3 SNPs in the *MECP2* (2 silent SNPs in patient 6 and another SNP in patient 7 inherited from her normal father). In *CDKL5,* we found the same intronic SNP in 4 patients. Angelman Syndrome was excluded for the cohort, by both clinical criteria and methylation analysis. In paper III, we began with mutation analysis for the *FOXG1* gene in all 12 remaining patients from paper I (excluding patient 2 and patient 14) with DNA sequencing and we did not detect any pathogenic mutations in *FOXG1* in the 12 patients. Then, we have further analysed the patients using affymetrix SNP6.0 array. We found 3 patients with variants of uncertain clinical significance (VOUS): two different duplications located on chromosome band 6q26 in the *PARK2* gene, and a deletion at chromosome 16p13.2 in the *RBFOX1* gene (Table 2). We also identified many regions with copy neutral Loss of Heterozygosity (LOH) in patient 4; including 7q11.22-q22.1, 10p15.1-p15.3, 14q11.2-q13.3, 21q11.2-q21.1, Xp11.4-p22.33 and Xq23-q25.

Severe mental retardation (paper II)

The main objective of this study was to investigate the prevalence, co-morbidities and aetiologies of severe mental retardation in a cohort of 133 Swedish children (figure 5) and to further penetrate aetiologies in a group with ***unknown*** causes by application of updated clinical-genetic methods to estimate how todays methods as microarray affects the rate of genetic diagnosis in this group. Assignation to SMR defined by IQ < 50 was based on psychological testing and/or careful evaluation of the cognitive developmental level (see also page 17). All children with Down syndrome were included.
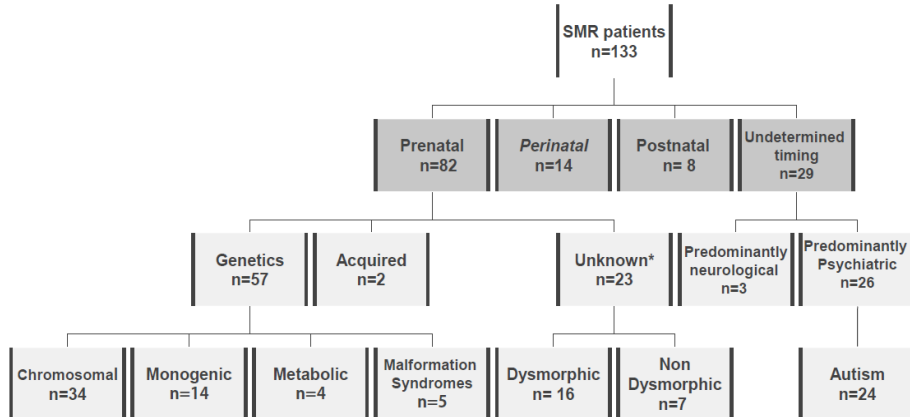

Figure 5   Flow-chart showing the process of inclusion of aetiology categorization of SMR patients, and the number of patients in each group.
*The number of patients in "Unkown aetiology" decreased to 18 after SNP array analysis.

The **prenatal unknown group** included 23 children with indications of prenatal but unclassifiable aetiology. Although some of these patients had unspecific dysmorphic features, brain anomalies, epilepsy and additional dysfunctions, the clinical diagnosis were undetermined. The chromosome microarray method was implicated to find out if there are some pathogenic CNVs presented in this group, the aim was to investigate the impact of genetic methods choices on number of aetiologically diagnosed SMR patients. We identified pathogenic CNVs in 5/23 children, or 4/22, if the two brothers are calculated as one occurrence (a *distal 10q deletion syndrome*, a *5q14.3-q15 microdeletion syndrome*, a *22q11 deletion syndrome* and *Xq28/MECP2 duplication syndrome* in 2 brothers; Table 3).

Table 3   **Overview of detected pathogenic CNVs in prenatal unknown SMR patients by Cytogenetics Whole_Genome 2.7 M array and the Genome wide SNP6.0 array (paper II)**

Sex	Chr	CNV type	Locus	Start position[1](bp)	End position[1](bp)	Size (Kb)	Genes	inheritance
Female	10	Loss	10q26.12-q26.3	122703669	133745133	12527	98	*de novo*
Male	5	Loss	5q14.3	83132102	89994111	6862	17	*de Novo*
Male	22	Loss	22q11.1-q11.21	17369300	19790008	2421	67	*de Novo*
Male*	X	Gain	Xq28	152768158	153201808	433	18	Inherited (Mat)

[1]UCSC Build 36
ND: not determind
*Two brother s with same maternally inherited Xq28 duplication including *MECP2* gene, the xq28/*MECP2* duplication was detected by microarray analysis in one brother and the other brother was analysed using *MECP2* specific MLPA assay.
Mat: maternally

MECP2 duplication syndrome (paper IV)

The all over aim of this study was to further investigate and compare the *MECP2* duplication syndrome phenotype in female versus male patients. As the most of *MECP2* duplication syndrome in male patients were inherited from healthy mothers, the duplication was considered as asymptomatic in females. The X chromosome inactivation (XCI) mechanism in female had been suggested to rescue the phenotype in carrier females. In paper IV; we identified a 101 kb tandem *de novo MECP2* duplication, containing the *MECP2* gene and exon 1-7 of the *IRAK1* gene, in a female patient. XCI status in leucocytes was random and she had slightly dysmorphic features which were in accordance with earlier reported phenotypes in affected females; high palate, pointed teeth and delayed developmental milestones.[7]
The comparison between the SRO (shorter region of interest) in earlier reported cases in both gender and our cases in this study confirms that the *MECP2* is the gene dosage sensitive gene (paper IV, figure 1).

The two brothers with *MECP2* duplication syndrome (from prenatal unknown aetiology in study II) and their carrier mother and sister were also included in this study for further clinical observations. Although both the mother and sister were asymptomatic carrier the XCI pattern in blood was random for the mother and skewed (15:85) for the sister's normal allele. However, the brothers showed different phenotype and responded to epilepsy treatment differently despite the same inherited (most likely identical) duplication. The younger brother died suddenly at 13 years age in coincidence with an unexpected response to respiratory infection.

The results of this study underlines that the *MECP2* duplications syndrome in males is more severe and may even lead to sudden death, this duplication in females has incomplete penetrance and variable expressivity, thus the symptoms might display milder and extend from asymptomatic to moderate MR. Female carrier status may be diagnostically important as some symptoms, such as learning difficulties, are common for females with *MECP2* duplication syndrome.

DISCUSSION

In this thesis, we have undertaken extensive genetic analysis for the two groups of children with mental retardation. As discussed earlier the early infantile onset RTT patients have already had the clinical diagnosis and the main aim for this group was to find out the possible genetic variations associated to this diagnosis. The approach of genetic analysis of the other group (SMR with unknown aetiology) was the other way around; to identify a probable genetic variation as an implement for diagnosis.

The findings in the unknown aetiology group were comprehensible with other available data and the interpretations of all four detected CNVs with microarray were easier since all of these findings were already described as syndromes (paper II). The clinical evaluation was affected by the lack of some symptoms in accordance with the syndromes, together with other additional symptoms in our patients, which directed the patients to the unknown aetiology group. Since this study is a based on epidemiologically collected material, one aim was to determine whether the prevalence of SMR in Scandinavia have changed compared to pervious studies, the hypothesis was that the implication of updated new diagnostic methods as well as the progress in genetic methodology may affect the number of diagnosed children with SMR. However, the outcome of this study implies that the aetiological prevalence of SMR is unchanged and is in accordance with those previously reported.[12,13,81-83] In contrast, the high throughput microarray analysis elevated the rate of diagnosis from earlier 4 % to remarkable 22.5 % in unclear cases. At the initiation of the study in the year 2005, there were no coherent data supportive of association of CNVs with autism, which was why we excluded twenty-six children with autistic disorders without dysmorphic signs from the genetic analysis, these patients were classified as having undetermined aetiology. After that, many research groups have used a similar approach in autism studies and association of CNVs including both gains and losses has been reported.[84-88] These results implies that microarray analysis in the excluded autistics group (n=26) might change our results somewhat and probably would have increased the number of detected pathogenic CNVs.

On the other hand, we found various genetic abnormalities in the early infantile onset RTT patients. One explanation may be the phenotypic heterogeneity in this group, although all the girls included in this group full field the criteria that are required for classification of RTT variants, but within the group, the patients full field different criteria. Lack of *MECP2* aberration in the most of early infantile onset RTT variant patients (13/14 patients in paper I and III), motivated the microarray analysis approach. The MeCP2 protein has a repressor function, with many other proteins involved in the pathway. Mutation in these genes might lead to a phenotypic presentation similar to the phenotype of RTT.
However, the microarray analysis results in this group include 3 different autosomal intra-genic CNVs (2 in *PARK2* and 1 in *RBFOX1*). Although CNVs including both of these genes have showed association with autism and also other neurological disorders as schizophrenia, juvenile Parkinson disease and Alzimers disease.[70,87] However, these CNVs have uncertain significance and should be interpreted with caution since the CNVs are inherited from asymtomatic parents and are represented in the normal population (DVG database) as well.

Many factors have to be considered in these cases:

Firstly, incomplete penetrance and variable expressivity for inherited CNVs, which have been reported for other microduplications/microdeletions syndromes as 22q11 duplications syndrome.

Secondly, the possibility of additional modifying factors as variations in combination, epigenetic factors or environmental elements may play the second-hit   role   in affected patients.

Thirdly, MECP2 mutations have also been reported in patients with other neurological disorders as autism and schizophrenia. This may implicate pathways with both *MECP2* and for example *PARK2* gene.

In this thesis, we have analyzed early infantile onset "congenital" Rett syndrome variant patients with mutation analysis for candidate genes and further copy number variation analysis using SNP array. The children with SMR were epidemiologically investigated and categorised in our other study and the unknown aetiology group within the prenatal SMR group were analyzed with SNP array. The findings of our studies can be summarised as below:

- The *MECP2* gene aberrations may leads to earlier onst of Rett syndrome since we found a *MECP2* deletion in 1/14 early infantile onset RTT patients. Therefore, we conclude that mutations screening of *MECP2* should be recommended as the first step in genetic analysis considering RTT variant patients.

- The causative genetic mechanisms of early infantile onset RTT is complex and may be different as compared with classical Rett syndrome. This variant consists of a phenotypically heterogeneous group of patients which has been reflected in genetic heterogeneity; from mutations in candidate genes to CNVs. These results suggest that copy number variant analysis with the microarray technique might be useful for genetic analysis of this particular group as the second step in genetic analysis in *MECP2* negative patients.

- The prevalence and co-morbidity of severe mental retardation in Sweden is unchanged and similar to previous Scandinavian studies (2.9 per 1000). The high-resolution chromosomal microarrays are valuable diagnostic tools of patients with unexplained SMR and increased the number diagnosed patients from earlier 4% (using traditional analysis methods) to todays 22.5%.

- *MECP2* duplication is a pathogenic CNV in both genders with recurrence risk for both females and males; risk for recurrence is 50% for boys and less for girls because of incomplete penetrance. This has to be recognised in genetic counselling for carrier females. Follow up of carriers can be recommended because of the performance of late-onset neurological features in female carriers.

Medical relevance

In conclusion, this thesis makes an effort to investigate the cause of two specific groups of patients with mental retardation. This follow up is essential for prognosis, management, and genetic counselling.
Identification of underlying variations for MR has crucial implications for prevention and treatment. It is also important to provide inexpensive, accurate, and effective molecular analysis for families with mental retardation. The molecular diagnosis has major impact on the possibility to provide prognosis, correct treatment, and also to allow networking with patient organisations. Family studies might identify carriers of pathogenic variations and genetic counselling to determine risk of recurrence in future pregnancies.

FUTURE PROSPECT

Recent genetic studies in the field of mental retardation have contributed to progress of different areas from genetic counselling to treatment. The advent of new technologies as microarray in the last decade has been important for diagnostic methodology in MR. In the years to come, I believe research and genetic analysis concerning patients with SMR unknown aetiology and RTT variants will highly increase. Nowadays, families are more aware of the possibility to use genetic counselling and prenatal diagnosis becomes more and more included in the healthcare of pregnant women.

For many years screening of the *MECP2* gene for diagnostic proposes (for classical RTT) has been a well established method. Rett mouse models indicate an effective treatment strategy which suggesting new applications for possible human gene therapy. These results motivate us to further study RTT variants. Both the early infantile onset RTT variant and male patients with Rett-like features can contribute to find the disease causing genes, CNVs as well as other genetic elements. Exome sequencing analysis or whole genome sequencing analysis of these patients may identify other key elements. This combination of methods will provide an opportunity to find out both small and gross variations in the genome.

The final goal will be to introduce the causative genetic vatiations for prenatal and screening of newborns. If this happens, it will be of assistance and improve the public healthcare system of service to patients and their families, and hopefully an aid to find an eventual cure (especially for the RTT patients).

POPULÄRVETENSKAPLIG SAMMANFATTNING

Utvecklingsstörning definieras som ett tillstånd där en människas kognitiva förmåga inte utvecklats normalt. Det medicinska uttrycket för utvecklingsstörning är ***mental retardation (MR).*** Förekomsten av mental retardation är beräknad till 3 % i Europa**.**
MR orsakas av genetiska förändringar hos individen i 25-35 % av alla fall, vilka antingen kan vara nedärvda från föräldrarna eller orsakas av nya förändringar i genomet, s.k. *de novo* mutationer. Därför är genetiska utredningar viktiga för att diagnostisera patienter med MR. Delarbetena i denna avhandling baseras på genetiska studier av två patientgrupper med MR.

Den första studien avser genetiska studier av barn med en variant av **Rett syndrom** (RTT) med tidig debut; "early infantile onset Rett syndrome variant" (delarbete I och III). Rett syndrom är en neurologisk sjukdom som främst drabbar flickor med en incidens av 1:10 000. RTT kan förekomma hos pojkar men det är mycket ovanligt.
Barnen med RTT utvecklas normalt initialt och symptomen uppträder efter 6-18 månader. Vid early infantile onset RTT har patienterna en diagnostisk debut före 6 månaders ålder vilket gör att de första månaderna av normal utveckling uteblir. RTT orsakas främst av genförändringar i *MECP2*, lokaliserad på X-kromosomen (Xq28). På senare år har mutationer i två andra gener *CDKL5* (Xp22) och *FOXG1* (14q12) rapporterats i varianter av RTT. Därför har vi gjort en genetisk utredning av vår patientgrupp om 14 barn genom mutationsscreening av ovanstående 3 gener samt kompletterat med kopietalsanalys (Copy number variation, CNV). Kopietalsanalys är en relativt ny analysmetod som kan upptäcka betydligt mindre genetiska förändringar i arvsmassan på kromosomnivå, än en traditionell kromosomanalys. I delarbete I upptäckte vi en sjukdomsorsakande deletion (förlust av genetiskt material) i *MECP2* genen hos en patient medan en annan patient uppvisade en stor deletion av den korta armen av kromosom 3. Vidareutredning med kopietalsanalys av resterande 12 patienter (delarbete III) visade tre olika CNVs. Två patienter hade en duplikation (extramaterial) på kromosomband 6q26 där *PARK2* är lokaliserad, medan en patient hade deletion på kromosomband 16p13.2 i *RBFOXG1*.
Dessa CNVs har föreslagits som potentiella sjukdomsorsakande avvikelser i litteraturen och kan inte uteslutas som bidragande orsak till patienternas symptom.

Delarbete II och IV i avhandlingen är en epidemiologisk studie av barn med **svår mental retardation** (SMR) och genetisk uppföljning av dessa. Denna studie genomfördes i samarbete med en barnneurolog i Hallands län där 133 barn med SMR identifierats. Patienterna är indelade i fyra kategorier i relation till beräknad tid för sjukdomens uppkomst; prenatal (innan födelse), perinatal (under förlossning), postnatal (efter födelse) eller vid obestämd debut. Den prenatala gruppen utgör den största andelen (82 patienter) varav 23 patienter är utan känd diagnos. Syftet med dessa studier var att undersöka den genomiska profilen hos dessa 23 patienter med hjälp av kopietalsanalys. Patogena CNVs upptäcktes hos 5 patienter i delarbete II och dessa är sedan tidigare beskrivna som kända syndrom: Distal 10q deletion syndrome, 5q14.3-q15 microdeletion syndrome samt 22q11deletion syndrome fanns hos vardera en patient medan Xq28/*MECP2* duplication syndrome sågs hos två bröder.

I delarbete IV jämfördes sedan fenotypen mellan bröderna från arbete II med en kvinna med *MECP2* duplikation för att se könsspecifika skillnader.

ÖVERGRIPANDE SLUTSATSER:

- En patient med early onset infantile Rett syndrom med deletion i *MECP2* visar att förändringar av *MECP2* kan ge upphov till en mycket tidigare sjukdomsdebut än tidigare känt.

- Kopietalsanalys (SNP array) är ett effektivt genetiskt analysredskap för diagnostisering av mental retardation. Metoden har lett till en ökning av antalet diagnostiserade SMR patienter, från 4 % till 22,5 %.

- *MECP2* duplikation är sjukdomsorsakande med 50 % upprepningsrisk för män och något lägre för kvinnor, vilket bör beaktas vid genetisk vägledning.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to everyone who has been involved in this thesis:

PhD **Anna Erlandson**, my supervisor, for giving me the opportunity to work with this exciting project. For sharing your knowledge and experiences in the broad field of genetic analysis of mental retardation. For guiding me through this journey and for help in writing papers and this thesis.

Associated professor **Mårten Kyllerman**, my co-supervisor, for fruitful meetings and discussions and for sharing your clinical knowledge. I am truly grateful for your fast response to all my questions and your help with writing this thesis and your encouraging attitude.

Professor emeritus **Jan Wahlström**, my co-supervisor, for giving me the opportunity to work with the Rett syndrome project and for supporting me during work with this project.

I have to express my gratitude to *Professor emeritus* **Bengt Hagberg**, for your generosity sharing your great experiences of over 40 years of Rett syndrome research, it has been a privilege knowing you!

To all our collaborators and co-authors, especially *Dr* **Mikael Lundvall**, for excellence collaboration and sharing the SMR study with me. For your enthusiasm about the project and always answering my clinical questions.

All the present and former researcher at Department of Clinical Genetics, especially Professor **Afrouz Behboudi** for fruitful discussions and your advises in every thing from research to family. **Rose-Marie Sjöberg** for your willingness and encouragement, for sharing your experience in the lab and amusing discussions about various topics. **Niloofar Javanmardi** for being a good friend and being there when I needed you. **Åsa Naluai, Susanne Fransson, Anna Djos, Tara Stanne, Frida Abel** and **Kittichate Visuttijai** for making the floor 3 a pleasant working place.

To everyone on the fourth and fifth floors of the **Department of Clinical Gentics** for creating a nice and warm working atmosphere. For all the friendly conversations in the lunch room.

Especial thanks to **Maria Yhr** for helping me with the array analysis in end of my project. For being a true friend and good listener, I especially appreciate your positive spirit and endless support. **Camilla Friberg**, my dear friend for all conversations about small and big things in life or lab.

I would also express my gratitude to my fantastic **parents,** my mentors in life**,** for support and encouragement. For loving me and always showing interest in what I do.

My *sisters* and **brother**, for being there even you are miles away, thank you for your support.

Especial thanks to my wonderful family; **Said** my love in life and soul mate! For loving me and always supporting me.
Arian, **Nicki** and **Armin**, my lovely children for your unconditional love, for always making me smile and for reminding me what life is about, you are the sunshine in my life!

I would like also to express my deepest gratitude to all **the families and children** that have participated in our studies. This work would never been possible without your contribution.

"Just as opportunities also contains problems,
every problem brings an opportunity."

DAN MILLMAN
"The life you were born to live"

REFERENCES

1. Watson, J.D. & Crick, F.H. The structure of DNA. *Cold Spring Harb Symp Quant Biol* **18**, 123-31 (1953).
2. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
3. Venter, J.C. *et al.* The sequence of the human genome. *Science* **291**, 1304-51 (2001).
4. Crick, F.H. The Complementary Structure of DNA. *Proc Natl Acad Sci U S A* **40**, 756-8 (1954).
5. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421**, 448-53 (2003).
6. Strachan &Read, . Human molecular genetics. *Book, 4th edition*, (Garland Science, 2011).
7. Bijlsma, E.K. *et al.* Xq28 duplications including MECP2 in five females: Expanding the phenotype to severe mental retardation. *Eur J Med Genet* **55**, 404-13 (2012).
8. Lisik, M.Z. & Sieron, A.L. X-linked mental retardation. *Med Sci Monit* **14**, RA221-9 (2008).
9. Grayton, H.M., Fernandes, C., Rujescu, D. & Collier, D.A. Copy number variations in neurodevelopmental disorders. *Prog Neurobiol* **99**, 81-91 (2012).
10. Rauch, A. *et al.* Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet A* **140**, 2063-74 (2006).
11. Blomquist, H.K., Gustavson, K.H. & Holmgren, G. Mild mental retardation in children in a northern Swedish county. *J Ment Defic Res* **25**, 169-86 (1981).
12. Matilainen, R., Airaksinen, E., Mononen, T., Launiala, K. & Kaariainen, R. A population-based study on the causes of mild and severe mental retardation. *Acta Paediatr* **84**, 261-6 (1995).
13. Stromme, P. Aetiology in severe and mild mental retardation: a population-based study of Norwegian children. *Dev Med Child Neurol* **42**, 76-86 (2000).
14. Wellesley, D., Hockey, A. & Stanley, F. The aetiology of intellectual disability in Western Australia: a community-based study. *Dev Med Child Neurol* **33**, 963-73 (1991).
15. Chiurazzi, P., Schwartz, C.E., Gecz, J. & Neri, G. XLMR genes: update 2007. *Eur J Hum Genet* **16**, 422-34 (2008).
16. Lundvall, M., Rajaei, S., Erlandson, A. & Kyllerman, M. Aetiology of severe mental retardation and further genetic analysis by high-resolution microarray in a population-based series of 6- to 17-year-old children. *Acta Paediatr* **101**, 85-91 (2012).
17. Moog, U. *et al.* Neurodevelopmental disorders in males related to the gene causing Rett syndrome in females (MECP2). *Eur J Paediatr Neurol* **7**, 5-12 (2003).
18. Wan, M. *et al.* Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots. *Am J Hum Genet* **65**, 1520-9 (1999).
19. Hagberg, B., Aicardi, J., Dias, K. & Ramos, O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* **14**, 471-9 (1983).
20. Hagberg, B. Clinical manifestations and stages of Rett syndrome. *Ment Retard Dev Disabil Res Rev* **8**, 61-5 (2002).

21. Hagberg, B., Goutieres, F., Hanefeld, F., Rett, A. & Wilson, J. Rett syndrome: criteria for inclusion and exclusion. *Brain Dev* **7**, 372-3 (1985).

22. Hagberg, B. Rett syndrome: the diagnostic criteria and the clinical variants. Rev Neurol *Separata* **38**, 26-28 (2004).

23. Hagberg, B. Rett syndrome: long-term clinical follow-up experiences over four decades. *J Child Neurol* **20**, 722-7 (2005).

24. Hagberg, B. & Witt-Engerstrom, I. Rett syndrome: a suggested staging system for describing impairment profile with increasing age towards adolescence. *Am J Med Genet Suppl* **1**, 47-59 (1986).

25. Hagberg, B.A. & Skjeldal, O.H. Rett variants: a suggested model for inclusion criteria. *Pediatr Neurol* **11**, 5-11 (1994).

26. Hagberg, B., Hanefeld, F., Percy, A. & Skjeldal, O. An update on clinically applicable diagnostic criteria in Rett syndrome. Comments to Rett Syndrome Clinical Criteria Consensus Panel Satellite to European Paediatric Neurology Society Meeting, Baden Baden, Germany, 11 September 2001. *Eur J Paediatr Neurol* **6**, 293-7 (2002).

27. Chahrour, M. & Zoghbi, H.Y. The story of Rett syndrome: from clinic to neurobiology. *Neuron* **56**, 422-37 (2007).

28. Lewis, J.D. *et al.* Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905-14 (1992).

29. Amir, R.E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23**, 185-8 (1999).

30. Erlandson, A., Hallberg, B., Hagberg, B., Wahlstrom, J. & Martinsson, T. MECP2 mutation screening in Swedish classical Rett syndrome females. *Eur Child Adolesc Psychiatry* **10**, 117-21 (2001).

31. Trappe, R. *et al.* MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am J Hum Genet* **68**, 1093-101 (2001).

32. Erlandson, A. *et al.* Multiplex ligation-dependent probe amplification (MLPA) detects large deletions in the MECP2 gene of Swedish Rett syndrome patients. *Genet Test* **7**, 329-32 (2003).

33. Rajaei, S. *et al.* Early infantile onset "congenital" Rett syndrome variants: Swedish experience through four decades and mutation analysis. *J Child Neurol* **26**, 65-71 (2011).

34. Mnatzakanian, G.N. *et al.* A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat Genet* **36**, 339-41 (2004).

35. Adachi, M., Keefer, E.W. & Jones, F.S. A segment of the Mecp2 promoter is sufficient to drive expression in neurons. *Hum Mol Genet* **14**, 3709-22 (2005).

36. Alvarez-Saavedra, M., Saez, M.A., Kang, D., Zoghbi, H.Y. & Young, J.I. Cell-specific expression of wild-type MeCP2 in mouse models of Rett syndrome yields insight about pathogenesis. *Hum Mol Genet* **16**, 2315-25 (2007).

37. Blackman, M.P., Djukic, B., Nelson, S.B. & Turrigiano, G.G. A critical and cell-autonomous role for MeCP2 in synaptic scaling up. *J Neurosci* **32**, 13529-36 (2012).

38. Giacometti, E., Luikenhuis, S., Beard, C. & Jaenisch, R. Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proc Natl Acad Sci U S A* **104**, 1931-6 (2007).

39. Nelson, E.D., Kavalali, E.T. & Monteggia, L.M. MeCP2-dependent transcriptional repression regulates excitatory neurotransmission. *Curr Biol* **16**, 710-6 (2006).

40. Banerjee, A., Castro, J. & Sur, M. Rett syndrome: genes, synapses, circuits, and therapeutics. *Front Psychiatry* **3**, 34 (2012).

41. Chahrour, M. *et al.* MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**, 1224-9 (2008).

42. Della Ragione, F., Filosa, S., Scalabri, F. & D'Esposito, M. MeCP2 as a genome-wide modulator: the renewal of an old story. *Front Genet* **3**, 181 (2012).

43. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-9 (1998).

44. Kokura, K. *et al.* The Ski protein family is required for MeCP2-mediated transcriptional repression. *J Biol Chem* **276**, 34115-21 (2001).

45. Dastidar, S.G. *et al.* Isoform-specific toxicity of Mecp2 in postmitotic neurons: suppression of neurotoxicity by FoxG1. *J Neurosci* **32**, 2846-55 (2012).

46. Ariani, F. *et al.* FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet* **83**, 89-93 (2008).

47. Bahi-Buisson, N. *et al.* Revisiting the phenotype associated with FOXG1 mutations: two novel cases of congenital Rett variant. *Neurogenetics* (2009).

48. Jacob, F.D., Ramaswamy, V., Andersen, J. & Bolduc, F.V. Atypical Rett syndrome with selective FOXG1 deletion detected by comparative genomic hybridization: case report and review of literature. *Eur J Hum Genet* (2009).

49. Yeung, A. *et al.* 4.45 Mb microduplication in chromosome band 14q12 including FOXG1 in a girl with refractory epilepsy and intellectual impairment. *Eur J Med Genet* (2009).

50. Collins, A.L. *et al.* Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum Mol Genet* **13**, 2679-89 (2004).

51. Lubs, H. *et al.* XLMR syndrome characterized by multiple respiratory infections, hypertelorism, severe CNS deterioration and early death localizes to distal Xq28. *Am J Med Genet* **85**, 243-8 (1999).

52. Bartsch, O. *et al.* Four unrelated patients with Lubs X-linked mental retardation syndrome and different Xq28 duplications. *Am J Med Genet A* **152A**, 305-12.

53. Friez, M.J. *et al.* Recurrent infections, hypotonia, and mental retardation caused by duplication of MECP2 and adjacent region in Xq28. *Pediatrics* **118**, e1687-95 (2006).

54. del Gaudio, D. *et al.* Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genet Med* **8**, 784-92 (2006).

55. Echenne, B. *et al.* Neurologic aspects of MECP2 gene duplication in male patients. *Pediatr Neurol* **41**, 187-91 (2009).

56. Honda, S. *et al.* The incidence of hypoplasia of the corpus callosum in patients with dup (X)(q28) involving MECP2 is associated with the location of distal breakpoints. *Am J Med Genet A* **158A**, 1292-303 (2012).

57. Honda, S. *et al.* Concomitant microduplications of MECP2 and ATRX in male patients with severe mental retardation. *J Hum Genet* **57**, 73-7 (2012).

58. Lugtenberg, D. *et al.* Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. *Eur J Hum Genet* **17**, 444-53 (2009).

59. Honda, S. *et al.* Copy-number variations on the X chromosome in Japanese patients with mental retardation detected by array-based comparative genomic hybridization analysis. *J Hum Genet* **55**, 590-9 (2010).

60. Hanchard, N.A. *et al.* A partial MECP2 duplication in a mildly affected adult male: a putative role for the 3' untranslated region in the MECP2 duplication phenotype. *BMC Med Genet* **13**, 71 (2012).

61. Ramocki, M.B., Tavyev, Y.J. & Peters, S.U. The MECP2 duplication syndrome. *Am J Med Genet A* **152A**, 1079-88 (2010).

62. Sanmann, J.N. *et al.* Characterization of six novel patients with MECP2 duplications due to unbalanced rearrangements of the X chromosome. *Am J Med Genet A* **158A**, 1285-91 (2012).

63. Bauters, M. *et al.* Nonrecurrent MECP2 duplications mediated by genomic architecture-driven DNA breaks and break-induced replication repair. *Genome Res* **18**, 847-58 (2008).

64. Breman, A.M. *et al.* MECP2 duplications in six patients with complex sex chromosome rearrangements. *Eur J Hum Genet* **19**, 409-15 (2011).

65. Auber, B. *et al.* An unbalanced translocation resulting in a duplication of Xq28 causes a Rett syndrome-like phenotype in a female patient. *Clin Genet* **77**, 593-7 (2010).

66. Grasshoff, U. *et al.* De novo MECP2 duplication in two females with random X-inactivation and moderate mental retardation. *Eur J Hum Genet* **19**, 507-12 (2011).

67. Makrythanasis, P. *et al.* De novo duplication of MECP2 in a girl with mental retardation and no obvious dysmorphic features. *Clin Genet* **78**, 175-80 (2010).

68. Mayo, S. *et al.* De novo interstitial triplication of MECP2 in a girl with neurodevelopmental disorder and random X chromosome inactivation. *Cytogenet Genome Res* **135**, 93-101 (2011).

69. Miller, D.T. et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 86, 749-64 (2010).

70. Glessner, J.T., Connolly, J.J. & Hakonarson, H. Rare Genomic Deletions and Duplications and their Role in Neurodevelopmental Disorders. *Curr Top Behav Neurosci* **12**, 345-60 (2012).

71. Drew, L.J. *et al.* The 22q11.2 microdeletion: fifteen years of insights into the genetic and neural complexity of psychiatric disorders. *Int J Dev Neurosci* **29**, 259-81 (2011).

72. Karayiorgou, M., Simon, T.J. & Gogos, J.A. 22q11.2 microdeletions: linking DNA structural variation to brain dysfunction and schizophrenia. *Nat Rev Neurosci* **11**, 402-16 (2010).

73. Shashi, V. *et al.* Cognitive correlates of a functional COMT polymorphism in children with 22q11.2 deletion syndrome. *Clin Genet* **69**, 234-8 (2006).

74. Paylor, R. *et al.* Tbx1 haploinsufficiency is linked to behavioral disorders in mice and humans: implications for 22q11 deletion syndrome. *Proc Natl Acad Sci U S A* **103**, 7729-34 (2006).

75. Portnoi, M.F. Microduplication 22q11.2: a new chromosomal syndrome. *Eur J Med Genet* **52**, 88-93 (2009).

76. Fernell, E. Aetiological factors and prevalence of severe mental retardation in children in a Swedish municipality: the possible role of consanguinity. *Dev Med Child Neurol* **40**, 608-11 (1998).

77. Schouten, J.P. *et al.* Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* **30**, e57 (2002).

78. Bignell, G.R. *et al.* High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res* **14**, 287-95 (2004).

79. Wahlstrom, J. *et al.* Congenital variant Rett syndrome in a girl with terminal deletion of chromosome 3p. *J Med Genet* **36**, 343-5 (1999).

80. Malmgren, H., Sahlen, S., Wide, K., Lundvall, M. & Blennow, E. Distal 3p deletion syndrome: Detailed molecular cytogenetic and clinical characterization of three small distal deletions and review. *Am J Med Genet A* **143**, 2143-9 (2007).

81. Gustavson, K.H., Hagberg, B., Hagberg, G. & Sars, K. Severe mental retardation in a Swedish county. I. Epidemiology, gestational age, birth weight and associated CNS handicaps in children born 1959--70. *Acta Paediatr Scand* **66**, 373-9 (1977).

82.   Gustavson, K.H., Hagberg, B., Hagberg, G. & Sars, K. Severe mental retardation in a Swedish county. II. Etiologic and pathogenetic aspects of children born 1959-1970. Neuropadiatrie 8, 293-304 (1977).

83.   Gustavson, K.H., Holmgren, G., Jonsell, R. & Son Blomquist, H.K. Severe mental retardation in children in a northern Swedish county. *J Ment Defic Res* **21**, 161-80 (1977).

84.   Bremer, A. *et al.* Copy number variation characteristics in subpopulations of patients with autism spectrum disorders. *Am J Med Genet B Neuropsychiatr Genet* **156**, 115-24 (2011).

85.   Bremer, A. *et al.* Screening for copy number alterations in loci associated with autism spectrum disorders by two-color multiplex ligation-dependent probe amplification. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 280-5 (2010).

86.   Casey, J.P. *et al.* A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Hum Genet* **131**, 565-79 (2012).

87.   Glessner, J.T. *et al.* Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569-73 (2009).

88.   Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368-72 (2010).

other sources:
ISCN; International Standing committee on human cytogentic nomenclature 1995

ICD-10 classification of mental and behavioral disorders.WHO: 2001

Human Gene Nomenclature (HGNC), http://www.genenames.org/