The genetics of mental retardation

F. Lucy Raymond1,* and Patrick Tarpey2

1Department of Medical Genetics, Cambridge Institute of Medical Research, University of Cambridge, Addenbrookes Hospital, Cambridge CB2 2XY, UK and 2The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Received July 3, 2006; Revised July 14, 2006; Accepted July 25, 2006

Genetic abnormalities frequently give rise to a mental retardation phenotype. Recent advances in resolution of comparative genomic hybridization and genomic sequence annotation has identified new syndromes at chromosome 3q29 and 9q34. The finding of a significant number of copy number polymorphisms in the genome in the normal population, means that assigning pathogenicity to deletions and duplications in patients with mental retardation can be difficult but has been identified for duplications of MECP2 and L1CAM. Novel autosomal genes that cause mental retardation have been identified recently including CC2D1A identified by homozygosity mapping. Several new genes and pathways have been identified in the field of X-linked mental retardation but many more still await identification. Analysis of families where only a single male is affected reveals that the chance of this being due to a single X-linked gene abnormality is significantly less than would be expected if the excess of males in the population is entirely due to X-linked disease. Recent identification of novel X-linked mental retardation genes has identified components of the post-synaptic density and multiple zinc finger transcription factors as disease causing suggesting new mechanisms of disease causation. The first therapeutic treatments of animal models of mental retardation have been reported, a Drosophila model of Fragile X syndrome has been treated with lithium or metabotropic glutamate receptor (mGluR) antagonists and a mouse model of NF1 has been treated with the HMG-CoA reductase inhibitor lavastatin, which improves the learning and memory skills in these models.

INTRODUCTION

The genetic basis of mental retardation is now a huge field having started with modest beginnings with an initial survey of patients confined to long stay hospital institutions in the 1930s (1). This review has concentrated on the most recent publications that advance the field rather than attempting to be an exhaustive summary. The review begins with recent work on genomic deletions and duplications and progresses to new single gene abnormalities that have been identified. A broad overview of the classes of genes that give rise to mental retardation are discussed and reference is made to the first two papers suggesting possible therapies for deficits in learning and memory in animal model systems of the disease.

LARGE CHROMOSOME ABNORMALITIES

A genetic aetiology for mental retardation has been recognized for many years and since the first identification of Down syndrome as a chromosome abnormality, the focus of mental retardation research has been to identify smaller and smaller chromosome abnormalities associated with disease (2,3). The aim is to understand the molecular basis of intellectual disability and to provide an ever-improved clinical service to this group of patients and their families. In the early days, the identification of gain or loss of genetic material on routine chromosome analysis at a 500 G banding resolution was usually clinically significant as a karyotype was in effect a visual inspection of the whole genome at the resolution of ~5–10 Mb. Gain or loss of 5–10 Mb of DNA represents a large footprint of genes wherever the abnormality is found in the genome and almost inevitably leads to developmental abnormalities during embryogenesis. The most common effect of a chromosome abnormality is cognitive impairment, but it is also frequently associated with defects of heart formation and dysmorphic features. Occasionally, rare variants are described with no clinical significance and are usually associated with variations of euchromatic material (4).
MICRODELETIONS

In the early 1990s, recurrent small microdeletions of the genome not visible by light microscopy were identified associated with characteristic syndromes. These were only detectable by fluorescence in situ hybridization (FISH) or similar techniques and include the common microdeletion syndromes of: Wolf-Hirshhorn, Cri du Chat, Williams, Prader-Willi, Angelman, Rubinstein-Taybi, Miller-Dieker lissencephaly, Smith-Magenis, Alagille, DiGeorge and 22q11 deletion syndrome. Although each syndrome has characteristic distinguishing features all are also associated with varying degrees of mental retardation and in some cases associated with specific areas of intellectual impairment.

In 1995, Flint et al. (5) extended this technique to develop a strategy to screen for the abnormal inheritance of subtelomic DNA polymorphisms in individuals with mental retardation and dysmorphic features. Since this initial observation screening telomeres for deletions has become routine in clinical service for suitably selected patients with a diagnostic yield of 2.5–15% (6–9). More recently, screening individuals with minimal dysmorphic features but mental retardation as the predominant feature has resulted in the delineation of several new microdeletion syndromes including 3q29 microdeletion syndrome (10). Clinically, these conditions are not easily united but the identification of a common microdeletion forms the basis of the classification (11). This reflects a more general trend that classification of mental retardation syndromes associated with fine chromosomal abnormalities is increasingly based on molecular similarities (recurrent similar deletions or duplications) based on grouping of clinical dysmorphology features alone.

Once a significant number of screens of telomeres were performed in individuals with mental retardation, rare small familial polymorphic deletions or duplications were noted making the clinical significance of a deletion in an individual with mental retardation occasionally difficult to interpret but at this stage these were relatively rare (8,9,12).

Having established the principle that small deletions or duplications of chromosome material can lead to mental retardation, the technology has been further developed to identify yet smaller deletions and duplications in the whole genome in a single or few procedures. The use of multiple probes simultaneously is now possible using probes of known location on the genome 1 Mb apart (13). Recent publications have established that a further 10% of patients with mental retardation carry deletions or duplications (14,15). Initially, those patients selected for analysis were both dysmorphic and had significant mental retardation. The deletions that have been detected have been large and generally incorporating several genes.

COPY NUMBER CHANGES

With further development and refinement of array CGH technology leading to more extensive coverage of the genome, single gene abnormalities that result in mental retardation can now be detected. Initially, the development has been on the X chromosome, as this chromosome is well annotated and a large number of X-linked mental retardation genes map to the X chromosome (16). Further developments to generate a novel BAC tiling resolution genome-wide microarray has been successful and the use of this in a study of 100 patients revealed that 10% had de novo abnormalities in a cohort of mentally retarded individuals (17). In this study, however, 97% of patients had a reproducible copy number change and on average each patient had three detected changes throughout the genome. Most of these changes were either familial and therefore thought to be polymorphisms or were unproven as both parents were not available to confirm the origin of the copy number change. Although the genomic coverage of this form of tiling path is a great improvement on the 1 Mb array, the issue of deciding whether or not a rare variant is disease-causing is now a major issue. Where the phenotype is unusual in the population and likely to be due to a single gene or a few single genes within the genome distinguishing copy number changes that are pathological or polymorphic can be relatively easily resolved, but this is less easy when the phenotype is less distinct such as mental retardation and where the number of single genes that cause this phenotype is at minimum 100 and probably very many more.

It is emerging that the human genome is highly variable from one individual to another and large-scale copy number duplications or deletions within the genome are commonly polymorphic (18,19). Furthermore, the extent of inversions and complex rearrangements are far more common than was previously identified (20–23) and the extent to which these predispose to disease is now under close scrutiny (24,25).

COPY NUMBER CHANGES IN SINGLE GENES

Although there appears to be a high degree of selection bias against copy number changes in regions of the genome where Mendelian disorders map, there are many single gene disorders where deletions and duplications are common causes of disease e.g. Duchenne Muscular dystrophy, Pelizaeus Merzbacher and Charcot Marie Tooth disease (26,27).

Recently, a pathological duplication of MECP2 and LICAM has been described in males with severe mental retardation and progressive neurological symptoms but other duplications on the X chromosome at Xp22.3, Xq22.3 and Xq26.3 have also been described with no pathological significance (28,29). Without previous identification of point mutations in LICAM and MECP2 this assignment of a duplication in Xq28 to pathological significance would not have been possible and other duplications of genes in Xq28 have not necessarily been associated with disease (30).

The criterion for identifying whether or not a gene is pathological is by identifying coding sequence abnormalities that are not found in appropriate control populations. For microdeletions associated with dysmorphic feature and mental retardation, the task is to identify which features are associated with a large deletion and which are only due to haploinsufficiency of a single gene. For some microdeletion syndromes e.g. Williams and 22q11 deletion syndrome it still has to be fully resolved which gene in the common deleted region causes which of the phenotypic features. For the recently described 9q34 deletion syndrome, a single gene within the commonly deleted region that causes the disease has been elegantly identified by collecting a series of patients with smaller
and smaller deletions and identifying an apparently balanced translocation with the phenotype (31). Demonstrating the disruption of a gene at a balanced translocation breakpoint is still insufficient to be confident that a particular gene causes disease as subtle gain or loss of genetic material in individuals with translocations and no phenotype have been described (32). The final identification of a point mutation in a single gene within 9q34 in euchromatic histone methyl transferase 1 (EHMT1) has assigned this phenotype to a specific disease-causing gene (33). This gene now adds to the list of diseased genes where mental retardation is the predominant feature with or without additional clinical or dysmorphic features.

**CODING ABNORMALITIES IN SINGLE GENES**

The identification of genes that cause disease have usually relied on the characterization of familial cases. Where the mental retardation is severe in the family, it is usually effectively a reproductive lethal mutation and therefore autosomal dominant pedigrees are unlikely to appear in the population. The prevalence of diseased genes is therefore dependent of the new mutation rate, for autosomal dominant forms. Where the phenotype is common i.e. mental retardation without syndromic features, the identification of these individuals is clinically difficult and thus the number of characterized autosomal dominant genes that give rise to mental retardation is low. The identification of novel genes has relied on the collecting of individuals with mental retardation and distinct dysmorphic features in order to analyse and group together. Conditions like Smith-Magenis syndrome and Rubinstein-Taybi have been identified this way and the single genes that are sufficient to cause the phenotype have been identified by systematic deletion mapping and point mutation analysis (34,35).

For autosomal recessive causes of mental retardation, the use of homozygosity mapping in highly consanguineous families with affected siblings has great potential power. Although there are likely to be a significant number of genes on autosomes that cause mental retardation, the identification of these genes has been relatively slow. Three genes, PRSS12 on chromosome 4q26, CRBN on chromosome 3p26 and most recently CC2D1A on chromosome 19p13.12 have been identified using autozygosity mapping of highly consanguineous families (36–38). This method is powerful for detecting autosomal recessive genes that cause disease but require the identification of rare families with the condition and the enrichment of a population by a founder mutation that will have occurred by chance. The anticipation is that many new genes will emerge using this technique in the next 5–10 years, but it is unlikely to reveal all autosomal recessive genes as it cannot provide a systematic method of surveying the genome.

**X-LINKED MENTAL RETARDATION**

The majority of single genes that have been identified, which give rise to mental retardation are on the X chromosome. There is a male excess of affected individuals compared with females (ratio1: ~1.3) which has been assumed to be due to X-linked inheritance both due to the retention of X-linked genes in the population by the maintenance of reproductive fitness in females contributing to the prevalence in the population and the large number of genes on the X chromosome that gives rise to the condition (1). The first gene to be identified was FMR1 that causes fragile X syndrome and still remains the commonest single gene abnormality to be identified (39–42). Since 1990, a series of genes have been identified either by positional cloning or translocation breakpoint mapping methodologies. Some are only associated with mental retardation and are not reported to be associated with dysmorphic or other neurological symptoms and others are more syndromic although the distinction between these are gradually becoming less clear as syndromic and non-syndromic phenotypes are described for several of the genes (43,44). Currently, genes that are classified as non-syndromic mental retardation genes are: IL1RAPL1 (45), TM4SF2 (46), ZNF41 (47), FTSJ1 (48), DLG3 (49), FACL4 (50), PAK3 (51), ARHGGEF6 (52), FMR2 (53,54), GDI (55), ZNF81 (56) and ZNF674 (57), whereas the genes where syndromic forms of mental retardation is described are NLGN4 (58), RPS6KA3 (RSK2) (59), OPHN1 (60,61), ATRX (62), SLC6A8 (63), ARX (64,65), SYN1 (66), AGTR2 (67), MECP2 (68–70), POBP1 (56,71), FGDI (72,73), SMCX (74) and SLC16A2 (75–77).

In the past 18 months, the number of newly identified genes has not increased exponentially despite technological advances suggesting that either nearly all the genes on the X chromosome have been found or much more likely that the remaining number of genes that are to be found and now increasingly difficult to identify because of their rarity. Support for the later view is the presence of a large number of published and unpublished families with three or four generations affected with moderate to severe mental retardation and a clear X-linked inheritance patterns that have not had the causative mutation identified despite extensive screening of the known genes suggesting that yet more genes need identification. A systematic but limited search of brain expressed genes within Xp11 region that included ~50 genes in this region revealed only three new genes, and a detailed screen of 70 candidate genes based on sequence homology and functional similarity to those that have already been identified in a panel of 300 X-linked families only identified one novel gene where three missense variants were identified (Raymond et al., unpublished) (78). In order to identify the remaining genes on the X chromosome that cause mental retardation, a novel systematic approach is needed. The Genetics of Learning Disability (GOLD) study was established by the authors in Cambridge, UK, to identify novel genes that cause X-linked mental retardation. The group is using high throughput genomic DNA sequencing of all coding exons of all genes on the X chromosome in a cohort of 200 X-linked mental retardation families (49). The current estimate of genes to be screened is 854 (http://vega.sanger.ac.uk), and the number of primer pairs required for coverage is ~7200. The criteria for selecting the family for sequencing is the presence of at least two affected male individuals, a normal karyotype and no known diagnosis. Families with the largest pedigrees are prioritized as disease in these is most likely to be due to a single gene defect. Although this major project has not reported its completed findings yet, it is emerging that there are large numbers of unique rare variants being identified on the X chromosome in the cohort of 200
families and not all of them are likely to be disease-causing. We have evidence that truncating mutations are not infrequently polymorphic and not necessarily associated with the disease phenotype (Fig. 1). This has important implications for the molecular genetics community and emphasizes the need to have samples from distant relatives to track mutations and to be aware that de novo sequence variants are not necessarily pathogenic if the phenotype under scrutiny is caused by a large number of different possible genes. Nevertheless, the identification of many further genes that cause X-linked mental retardation hopes to be a significant contribution to the field.

In addition to the difficulty of finding mutations in large X-linked families, the observations of Mandel and Chelly that the prevalence of mutations in ARX in X-linked pedigrees was high at 6.6% but was only 0.13% in singletons, challenges the assumption that all male mental retardation is due to high penetrant abnormalities on the X chromosome (79). This suggests that in males there is a general predisposition to mental retardation akin to that of autism where genetic predisposing alleles need to be identified and are not necessarily located on the X chromosome, but may account for the increased prevalence of disease in male sib pairs quite separate from the rarer X-linked family pedigrees. This polygenic model

Figure 1. Mutation analysis of families with X-linked mental retardation. Probands are indicated by black arrows and individuals where mutation tracking was inconsistent are indicated by red arrows. For the family with a mutation in GPR119 individuals I1, I2, II1, II2, II4, II5, III1 and III2 were tested. For the family with a mutation in TEX13B individuals I1, I2, II3, II4, II6, IV1 and IV2 were tested. For the family with a mutation in ATXN3L individuals I2, II1, II3, II4, III1 and III2 were tested. Sequence traces of genomic DNA from the probands are shown with the mutation indicated. The location of the mutations and the subsequent protein alterations are based on the following RefSeq sequences: GPR119 NM_178471, NP_848566; TEX13B NM_031273, NP_112563; ATXN3L XN_045705, XP_045705.
of disease would fit with the empirical recurrence risks given, which are largely based on multiple observations of sib-pair recurrence risks between 1971 and 1987 and are 2–14% or 8.4% in a most recent population-based study in Atlanta, USA, where recurrences in children born between 1981 and 1991 for isolated mental retardation were noted (80).

**TYPES OF GENES THAT CAUSE MENTAL RETARDATION**

It is emerging that mental retardation can result from a wide range of protein abnormalities. The genes identified earliest were frequently signalling molecules in the RhoGTPase pathway (GDI, PAK3, ARHGEF6) or associated with chromatin remodelling (RPS6KA3, ATRX). Most recently, components of the synaptic vesicle or components necessary for its formation, have been identified as defective (SYN1, SLC6A8, NLGN4 and DLG3) and a number of novel transcription factors (ZNF41, 81 and 674) have been found. The former is especially exciting as the identification of mutations in these genes brings together the earlier work on rodent models of learning and memory defects which identified molecules in the post-synaptic density as key mediators of long-term potentiation and depression (LDP and LTP) (81).

The number of genes identified that cause mental retardation suggests that a mental retardation phenotype can emerge as the final common pathway of many different types of abnormal cellular processing and that no one overriding mechanism is likely to be the cause of mental retardation. This is both encouraging to us as humans who like to think that our intellectual processing is both sophisticated and complex but is somewhat daunting to the mental retardation biologist who in the early days naively thought that the identification of the genes responsible for the disease would give us immediate insights into the mechanism of the disease processes.

Despite the complexity and number of genes that have been identified, two recent papers have suggested that some defects of learning and memory may be amenable to therapeutic agents. McBride et al. (82) demonstrate that synaptic plasticity and courtship behaviour can be restored in a model of Fragile X syndrome by treatment with metabotropic glutamate receptors (mGluR) antagonists or lithium. Also, the learning and attention deficits in a mouse model of neurofibromatosis type 1 can be restored using an HMG-CoA reductase inhibitor lavastatin (83). Although these observations are a far way from suggesting therapeutic effects in humans, the possibility of offering drug treatments for some aspects of learning and memory disabilities in children with idiopathic mental retardation is promising. The next challenge is to provide cheap, accurate and effective sequence variant analysis in families with mental retardation. Together with this we need to develop rapid and effective functional assessments of pathogenicity of variants for all genes. Currently, we are only able to offer this for SLC16A2 where free T3 levels are abnormality high and for SLC6A8 where urine and plasma creatine:creatinine levels are abnormal.

We next need to create and properly assess the validity of animal models of disease in order to understand the pathways involved and increase the possibility of therapeutic targets for learning and memory disabilities in patients in the future.

**ACKNOWLEDGEMENTS**

This work was funded by the Wellcome Trust, UK. We wish to acknowledge the help and support from the families in our study, clinicians and key collaborators Gillian Turner, Jozef Gecz, Charles Schwartz, Roger Stevenson and the many members of the Sanger Institute Cancer Genome team lead by Mike Stratton, Andy Futreal and Richard Wooster.

Conflict of Interest statement. None declared.

**REFERENCES**


