Genetic aspects of Rett syndrome

Laurent Villard, PhD

Aix Marseille University, Inserm UMR-S 910, Faculté de Médecine de Marseille, France

DNA, genes and proteins: necessary concepts

Rett syndrome is a genetic disease. Genetic diseases are caused by the presence of variations in the sequence of the DNA molecule, contained within each cell. The DNA contains all instructions necessary to build an organism and make it live. Each instruction in the DNA is called a gene. The genes themselves are separated into several smaller consecutive DNA portions called "exons". There are approximately 30,000 genes in the human DNA. Each gene contains the instructions to assemble single amino acids into a chain of amino acids called a protein. These proteins make up our entire body, from our muscles to our skin and from our joints to our brain.

The human DNA is contained within 23 pairs of chromosomes (23x2 = 46 chromosomes). Chromosome pairs 1 to 22 are present in males and females. Chromosomes determining the gender are different in males and females: males have one X chromosome and one Y chromosome (XY) while females have two X chromosomes (XX). In females one of the two X chromosomes is silenced by a process called X chromosome inactivation (see below).

DNA variations are permanent changes in the DNA molecule. These variations can be transmitted from parents to children. When these variations occur in a DNA sequence coding for a protein, they can have several consequences: 1- the replacement of an amino acid by another in the protein, these are called "missense variations"; 2- the replacement of an amino acid by an instruction causing the interruption of the synthesis of the protein, these are called "non-sense variations"; 3- the loss or gain of DNA segments of varying length, called deletions or insertions, respectively. In the case of deletions or insertions, the consequence in the gene is called a "frameshift", usually causing an interruption of protein synthesis; 4- no consequence (these variations are called “polymorphisms”).

Why do variations occur in the DNA?

When a cell is dividing (during the development or in the adult), it needs to duplicate its DNA. This process is particularly complex and it is not completely error-free. These errors introduce variations in the newly synthetized DNA molecule. Geneticists have calculated that every new-born has new random variations that were absent from their parents (the so-called "de novo" variations). In addition, they also carry a specific set of variations inherited from their parents. Across generations, this process introduced an enormous amount of variations in the DNA molecule and hundreds of thousands of variations are known in the human DNA [1]. Most of them do not cause a disease. However, from time to time, a random variation occurs in a DNA segment where it modifies the coding sequence of a gene in such a way that the corresponding protein will lose its normal function (completely or partially). This subset of variations can cause a genetic disease such as Rett syndrome.

Genes whose variations cause Rett syndrome

The major gene involved in Rett syndrome is called MECP2. The acronym stands for methyl CpG binding protein 2. The MECP2 gene was identified for the first time in mice in 1992 by the laboratory of Adrian Bird in Scotland [2]. The corresponding human gene was identified four years later [3]. This work was performed long before mutations in this gene were found to cause a genetic disease, a breakthrough discovery made by the laboratory of Huda Zoghbi in 1999 [4].

The MECP2 gene is located on the X chromosome in humans. It is composed of 4 exons encoding a 486 amino acids protein. The MECP2 protein is a so-called "transcriptional modulator", meaning that it regulates the expression of hundreds of genes in our genome. Its function is particularly important for the brain to function normally [5].

The CDKL5 gene (also called STK9) is located on the X chromosome. In 2004, two studies published simultaneously have identified mutations in the CDKL5 gene in patients sharing clinical signs with Rett syndrome patients [6,7]. The clinical picture of the children carrying a mutation in the CDKL5 gene is only partially overlapping with that of classical Rett syndrome [8]. Several authors question the classification of patients mutated for CDKL5 in the group of «Rett syndrome» patients. Rather, they propose that these children suffer from a different disease [9].

The involvement of the FOXG1 gene in a variant form of Rett syndrome was published in 2008 [10], in children affected by the congenital variant of Rett syndrome. Mutations in CDKL5 or FOXG1 are a rare cause of Rett syndrome and the clinical phenotype of these patients is somewhat different from that of classical Rett syndrome patients.

The sentences "Rett syndrome gene" or "my daughter has the gene for Rett syndrome" are misleading. There is no "Rett syndrome gene". There are variations causing Rett syndrome when they occur in specific genes. The normal copies of genes do not cause a disease. The MECP2, CDKL5 and FOXG1 genes are present in every individual and they encode a protein performing one or several functions.

Why are Rett syndrome patients almost always female?

Variations in the DNA happen by chance when cells divide (see above). Sperm cells are produced by billions during the life of male individuals. This means billions of DNA copies, and billions of possibilities for a disease-causing variation to occur. Hence, the risk of containing a variation causing a genetic disease is much higher in sperm cells than in female eggs. Because fathers transmit their Y chromosome to their son (and not their X) there cannot transmit a MECP2 variation to a son. This probably explains why the majority of cases of Rett syndrome are females. Male patients with disease causing variations in MECP2 variations do exist. However, their disease is quite different from Rett syndrome and ranges from very severe fatal encephalopathy to mild intellectual disability. These cases are extremely rare compared to female cases of Rett syndrome.

X-chromosome inactivation

The phenomenon of X-chromosome inactivation plays an important role in Rett syndrome. X-chromosome inactivation occurs in every female individual in which one of the two X-chromosomes is silenced and does not contribute to the production of proteins from the genes it harbours [11]. The process of X-chromosome inactivation occurs early during the development of the female embryo. It occurs randomly and chromosome X_1 or chromosome X_2 can be silenced by a given cell. On average, one female will have 50% of her cells expressing X_1 and 50% of her cells expressing X_2. She will be a "mosaic" of cells expressing one or the other X chromosome (but never the two simultaneously in a given cell). In Rett syndrome, when a disease-causing variation is present in the MECP2 gene, a single X chromosome carries the variation (the child is said to be "heterozygous"). If the variation is located on the expressed X chromosome, then the cell will be "sick". If the variation is located on the inactive X chromosome, then the cell will be normal. Hence, Rett syndrome females are a mosaic of affected and normal cells.
In the population of healthy females, the percentage of X-chromosome inactivation varies from one female to another, and from expected 50:50 ratio to >90:10 in 7% of females under

25 years of age [12]. The percentage of "normal" versus "sick" cells in a given Rett syndrome female certainly has an impact on the severity of her clinical presentation. The X-inactivation profile is not transmitted from mothers to daughters.

Recurrence risk

The vast majority of Rett syndrome cases are caused by *de novo* variations. This means that the parents of the affected child are not carriers of the disease-causing variation. In this context, the recurrence risk is low and equals the risk of the general population to have a MECP2 disease-causing variation (0.43 to 0.71 per 10,000 females [13]).

There are two extremely rare situations where recurrence can be observed. The first situation is when the variation is present in more than 1 gamete (egg or sperm) of one parent (this is called "germinal mosaicism"). The second instance is when a mother carries a disease-causing variation in MECP2 and simultaneously has a completely skewed X-chromosome inactivation (see above) silencing the X chromosome carrying the variation. This "disease causing" chromosome will be silenced in the unaffected mother but can be passed on to affected children, including males.

These two extremely rare situations can be easily managed by genetic laboratories when a disease-causing variation has been identified in an affected child, and genetic counselling or prenatal diagnosis can be proposed if necessary. It must be mentioned here that the risk of having a second affected child in these two situations in much lower than the risk of losing the foetus following amniocentesis.
A consultation with a geneticist is necessary to discuss these issues because all individual situations and experiences are different and personalized counselling is needed.

Gene therapy

The function of the MECP2 gene is not totally understood. Hence, restoring the normal function of its target pathways to "cure" Rett syndrome might prove difficult. One alternative therapeutic strategy could be to replace the disease-causing version of MECP2 by a normal version. This gene therapy protocol would deliver the normal copy of the gene to the cells of the patient. Because Rett syndrome is primarily a brain disease, brain cells would be the major targets. To this aim, scientists use viruses called adeno-associated viruses (AAVs) modified to achieve a Trojan horse strategy. Viruses are made safe by removing their viral DNA and replacing it by the MECP2 coding sequence. This strategy is being tested in animal models of Rett syndrome, with some success [14-17]. However, a long road is ahead before the same protocol can be used in human patients.

References

[1] Lek, M., et al. (2016). Nature. 536:285-291.
[2] Lewis, J. D. et al. (1992). Cell. 69, 905-914.
[3] D'Esposito, M. et al. (1996). Mamm Genome. 7:533-535.
[4] Amir, R. E. et al. (1999). Nat. Genet. 23:185-188.
[5] Shah, R. R. & Bird, A. P. (2017). Genome Med. 9:17.
[6] Tao, J. et al. (2004). Am. J. Hum. Genet. 75:1149-1154.
[7] Weaving, L. S. et al. Am. J. Hum. Genet. 75:1079-1093.
[8] Bahi-Buisson, N. et al. (2008). Brain. 131:2647-2661.
[9] Fehr, S. et al. (2013). Eur J Hum Genet. 21:266-273.
[10] Ariani, F. et al. (2008). Am J Hum Genet. 83:89-93.
[11] Galupa, R. & Heard, E. (2015). Curr Opin Genet Dev. 31:57-66.
[12] Sharp, A., Robinson, D. & Jacobs, P. (2000). Hum Genet. 107:343-349.
[13] Bienvenu, T., et al. (2006). Pediatr Neurol. 34:372-375.
[14] Gadalla, K. K. et al. (2013). Mol Ther. 21:18-30.
[15] Garg, S. K. et al. (2013). J Neurosci. 33:13612-13620.
[16] Matagne, V. et al. (2017). Neurobiol Dis. 99:1-11.
[17] Sinnett, S. E. et al. (2017). Mol Ther Methods Clin Dev. 5:106-115.