Biochimica et Biophysica Acta xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbadis

Review

Genetics of the neuronal ceroid lipofuscinoses (Batten disease)☆

Sara E. Mole [a,b,*], Susan L. Cotman [c]

[a] MRC Laboratory for Molecular Cell Biology, University College London, Gower Street, London, WC1E 6BT, UK
[b] UCL Institute of Child Health and Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK
[c] Center for Human Genetic Research, Department of Neurology, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA

ARTICLE INFO

Article history:
Received 19 March 2015
Received in revised form 12 May 2015
Accepted 18 May 2015
Available online xxxx

Keywords:
Batten
CLN
Neuronal ceroid lipofuscinosis
NCL, genetics, mutation

ABSTRACT

The neuronal ceroid lipofuscinoses (NCLs) are a group of inherited neurodegenerative disorders that affect children and adults and are grouped together by similar clinical features and the accumulation of autofluorescent storage material. More than a dozen genes containing over 430 mutations underlying human NCLs have been identified. These genes encode lysosomal enzymes (CLN1, CLN2, CLN10, CLN13), a soluble lysosomal protein (CLN5), a protein in the secretory pathway (CLN11), two cytoplasmic proteins that also peripherally associate with membranes (CLN4, CLN14), and many transmembrane proteins with different subcellular locations (CLN3, CLN6, CLN7, CLN8, CLN12). For most NCLs, the function of the causative gene has not been fully defined. Most of the mutations in these genes are associated with a typical disease phenotype, but some result in variable disease onset, severity, and progression, including distinct clinical phenotypes. There remain disease subgroups with unknown molecular genetic backgrounds. This article is part of a Special Issue entitled: "Current Research on the Neuronal Ceroid Lipofuscinoses (Batten Disease)."

The neuronal ceroid lipofuscinoses share common clinical features that include epileptic seizures, progressive psychomotor decline, visual failure, and premature death. NCL disease usually begins in childhood, and most types are inherited in an autosomal recessive manner. However, there are several types with onset in early to late adulthood, including one type with autosomal dominant inheritance. Mutations in more than a dozen genes have been described in families diagnosed with NCL disease (Table 1). Those genes that cause the most prevalent and typical NCL disease with onset in childhood have been identified. There remain families diagnosed with NCL of all ages of onset in which the underlying genetic cause has not been described. Thus far, these cases do not appear to be caused by mutations in other genes that cause NCL-like disease in animals.

The aim of this review is to briefly summarize the genetic basis of NCL and any correlations with disease phenotype. Details on mutations can be found in the NCL mutation database (http://www.ucl.ac.uk/ncl).

NCL disease was first divided into four broad ages of onset in the late 1960s: infantile, late infantile, juvenile, and adult, due to initial simple supposition that there are four genes responsible for NCL disease: *CLN1*, *CLN2*, *CLN3*, and *CLN4*, respectively. The first genes to be identified were in the most common subtypes (*CLN1*, *CLN2*, *CLN3*). However, it was many years, with several further genes identified in less common forms of NCL, before a gene encoding the rare dominant adult type was found (*CLN4*). A variety of experimental approaches, largely reflective of the available technology at the time of identification, were used to identify genes causing NCL (Table 1). The first genes were discovered in 1995 following classic and time-consuming genetic linkage approaches using large numbers of similarly affected families followed by positional cloning of the genes (*CLN1*/*PPT1* and *CLN3*). In contrast, a biochemical approach that detected a missing mannose-6-phosphate tagged lysosomal enzyme in a patient allowed the identification of *CLN2*/*TPP1*. The next wave of gene identification required fewer families to provide sufficient power for genetic linkage analysis, due to the completion of the sequence of the human genome that provided more informative sequence variants (*CLN5*, *CLN6*, *CLN7*/*MFSD8*, *CLN8*). Several of these genes were identified by narrowing the candidate gene region by recognition of the stretches of homozygosity in consanguineous families. Recent improvements in sequencing technology that permit massively parallel sequencing of the whole exome in a relatively short space of time have allowed identification of the disease gene even in single families (*CLN4*/*DNAJC5*, *CLN11*/*GRN*, *CLN12*/*ATP13A2*, *CLN13*/*CTSF*, *CLN14*/*KCTD7*).

All NCL genes lie on autosomes, and in most cases, disease is inherited in a recessive manner, where deleterious mutations are present in both disease gene alleles. There are two notable exceptions: (1) adult onset NCL caused by mutations in *CLN4*/*DNAJC5* is dominantly inherited in all families described [1]; (2) there is one published report of uniparental disomy in the NCLs in which a patient with complete

Abbreviations: NCL, Neuronal ceroid lipofuscinosis; EPMR, Epilepsy with progressive mental retardation; PME, Progressive myoclonic epilepsy

☆ This article is part of a Special Issue entitled: "Current Research on the Neuronal Ceroid Lipofuscinoses (Batten Disease)."

* Corresponding author at: MRC Laboratory for Molecular Cell Biology, University College London, Gower Street, London WC1E 6BT, UK. Tel.: +44 207 679 7257; fax: +44 207 679 7805.
E-mail address: s.mole@ucl.ac.uk (S.E. Mole).

http://dx.doi.org/10.1016/j.bbadis.2015.05.011
0925-4439/

Table 1
Summary of the identification of genes that cause NCL.

Gene	Year of identification	Main approaches used for locus identification	Other approaches used	Reference
CLN1/PPT1	1995	Linkage	Linkage disequilibrium	[31]
CLN2/TPP1	1997	Biochemical	Linkage	[32]
CLN3	1995	Linkage	Linkage disequilibrium	[33]
CLN4/DNAJC5	2011	Linkage, exome sequencing	Gene expression	[1]
CLN5	1998	Linkage	Linkage disequilibrium	[34]
CLN6	2002	Linkage	Homozygosity mapping	[35]
CLN7/MFSD8	2007	Linkage	Homozygosity mapping	[36]
CLN8	1999	Linkage, animal model	Homozygosity mapping	[8]
CLN9	Not known			
CLN10/CTSD	2006	Animal model		[37,38]
CLN11/GRN	2012	Linkage, exome sequencing		[10]
CLN12/ATP13A2	2012	Exome sequencing		[21]
CLN13/CTSF	2012	Linkage, exome sequencing		[39]
CLN14/KCTD7	2012	Exome sequencing		[17]

isodisomy of chromosome 8, leading to homozygosity of a maternally inherited deletion in *CLN8*, is described [2].

The majority of NCL genes encode proteins that reside in the secretory and/or endo/lysosomal pathways. In most cases, these are lysosomal proteins, which include enzymes and a soluble protein (encoded by *CLN1/PPT1*, *CLN2/TPP1*, *CLN5*, *CLN10/CTSD*, *CLN13/CTSF*) and transmembrane proteins (encoded by *CLN3*, *CLN6*, *CLN7/MFSD8*, *CLN12/ATP13A2*). However, the transmembrane proteins CLN6 and CLN8 both localize to the endoplasmic reticulum (ER), and progranulin, encoded by *CLN11/GRN*, is demonstrated to reside in compartments in the secretory pathway [3]. Two other NCL proteins, encoded by *CLN4/DNAJC5* and *CLN14/KCTD7*, are both cytoplasmic and peripherally associated with cellular membranes. The *in vivo* substrates for the soluble lysosomal enzymes are incompletely defined, and knowledge regarding the primary functions of the membrane proteins in these compartments largely remains unresolved.

For the enzymatic deficiencies, enzyme replacement and gene therapy are promising approaches that are actively undergoing clinical development [4]. However, the unresolved functions for the other NCL proteins, which are less compatible with gene/protein replacement approaches, remain problematic for targeted therapy development. It is unclear which cellular defects due to their loss of function are primary versus secondary, and which of these defects play a central role in the cellular dysfunction and death that ensues. Despite these gaps in understanding of NCL protein function, there is wide evidence to indicate that cellular processes involving lipid and protein trafficking in the endocytic pathway and the regulation of endosomal and lysosomal transport are points of convergence for a number of the NCL protein deficiencies. A detailed overview of the current understanding of function for each of the NCL-encoded proteins can be found in this special issue [5].

For most NCL genes, there is a typical disease phenotype associated with complete loss of function. However, for many, there are also phenotypes recognized as NCL that are more protracted or have a later age of onset (Table 2). These arise from mutations that have 'milder' effects on gene function, and these phenotypes can vary quite considerably. For example, classic CLN6 disease begins in early childhood, but disease onset can be delayed until adulthood, with no associated visual failure [6]. There is also clinically similar NCL disease arising from mutations in more than one gene (e.g. late infantile variant NCL can be caused by mutations in *CLN5*, *CLN6*, *CLN7*, or *CLN8*). This led to the development of a new classification system that is gene based and takes into account these marked phenotypic consequences [7].

There are an increasing number of examples of different mutations in a single gene giving rise to quite different diseases (Table 2): (1) A recessive missense mutation in *CLN8* [8] causes progressive epilepsy with mental retardation (EPMR). This disease is an intellectual developmental disorder that presents with seizures in the juvenile age range that cease at adulthood and was the first genetic disease to be recognized for *CLN8*. Mutations that cause a more typical NCL were described later. (2) A particular missense mutation in CLN2/TPP1 is associated with spinocerebellar ataxia SCAR7, a slowly progressing but not life-limiting disease with no ophthalmologic abnormalities or epilepsy, and absence of typical storage [9]; A homozygous recessive mutation in *GRN* that is associated with rectilinear profiles, as found in progranulin-deficient mice [10], leads to CLN11 disease, whereas heterozygous mutations are a major cause of frontotemporal lobar degeneration with TDP-43 inclusions (FTLD-TDP), the second most common type of early-onset dementia. In this NCL family, it seems that carriers of this mutation are at risk of developing dementia with increasing age, which had been masked by the early death of many older members on both sides of the family. The age-at-onset and neuropathology of FTLD-TDP and NCL are markedly different and exemplify genetic links that have been emerging between rare diseases and common neurological disorders, such as Niemann–Pick C disease with Alzheimer's disease [11] and type 1 Gaucher disease with Parkinson's disease [12]. (3) Evidence suggests that the most common and very widespread mutation in *CLN3*, a 1-kb deletion, does not completely abolish CLN3 function, suggesting that disease caused by complete loss of function has not yet been recognized or may be lethal [13]; other recently described distinct phenotypes associated with *CLN3* mutations include retinitis pigmentosa without other clinical symptoms, even in mid-late adulthood [14] and a distinct disease described as autophagic myopathy associated with heart failure [15], considerably extending the phenotype of *CLN3*-associated disease. (4) There are also reports of other families with mutations in NCL genes that have predominantly visual problems [16]. (5) The only autosomal dominant type of NCL is CLN4 disease. Disease caused by complete loss of CLN4 function is not known, although animal models would predict very severe and early-onset disease. (6) Mutations in *CLN14/KCTD7* have now been reported to cause three different diseases, including NCL-like CLN14 disease [17–20]. (7) One family has been diagnosed with CLN12 disease, an atypical NCL [21], whereas all other mutations in *ATP13A2* cause Kufor–Rakeb syndrome. (8) One mutation in *SGSH* was described in a case diagnosed with adult onset NCL. Mutations in this gene usually underlie late infantile onset disease mucopolysaccharidosis type IIIA (MPSIIIA) [22].

Mutations in NCL genes range from those that are described in one or only a few families, to those that are more common in certain populations due to local founder effects. Several NCL genes have widespread distribution across several continents due to ancient founder effects (Table 2). The common mutation causing juvenile CLN3 disease is the best example of this, though it is unclear whether this global spreading provided a genetic advantage in the past. Diagnostic testing can be appropriately targeted to these common mutations.

There are reports of patients carrying changes in more than one NCL gene. One that was later found to be compound heterozygous for mutations in *CLN5* also carries a single mutation in the *CLCN6* gene. Another family has been described in which a single mutation in *CLCN6* is the

Please cite this article as: S.E. Mole, S.L. Cotman, Genetics of the neuronal ceroid lipofuscinoses (Batten disease), Biochim. Biophys. Acta (2015), http://dx.doi.org/10.1016/j.bbadis.2015.05.011