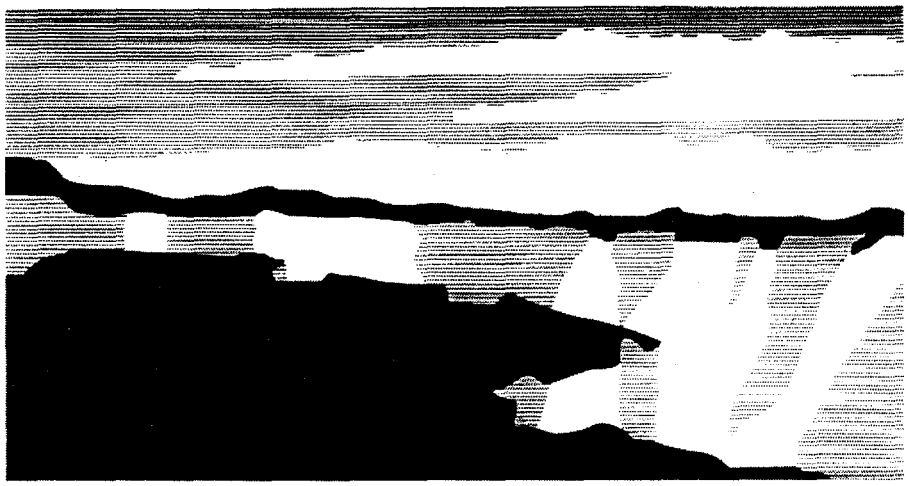
LA-UR- 96-0238

Title: **Positional Cloning of Disease Genes on Chromosome 16**

Author(s): **Norman Doggett, LS-3**
Martin Bruening, Leiden University
David Callen, Adelaide Women's and Children's Hospital
Mark Gardiner, University College, London
Terry Lerner, Massachusetts General Hospital

Submitted to: **DOE Office of Scientific and Technical Information (OSTI)**



Los Alamos
NATIONAL LABORATORY

Form No. 836 R5
ST 2629 10/91

Positional Cloning of Disease Genes on Chromosome 16

Norman Doggett*, Martin Bruening (Leiden University), David Callen (Adelaide Women's and Children's Hospital), Mark Gardiner (University College, London), and Terry Lerner (Massachusetts General Hospital)

Abstract

This is the final report of a three-year, Laboratory-Directed Research and Development (LDRD) project at the Los Alamos National Laboratory (LANL). The project seeks to elucidate the molecular basis of an important genetic disease (Batten's disease) by molecular cloning of the affected gene by utilizing an overlapping clone map of chromosome 16. Batten disease (also known as juvenile neuronal ceroid lipofuscinosis) is a recessively inherited neurodegenerative disorder of childhood characterized by progressive loss of vision, seizures, and psychomoter disturbances. The Batten disease gene was genetically mapped to the chromosome region 16p12.1 in close linkage with the genetic markers D16S299 and D16S298. Exon amplification of a cosmid containing D16S298 yielded a candidate gene that was disrupted by a 1 kb genomic deletion in all patients containing the most common haplotype for the disease. Two separate deletions and a point mutation altering a splice site in three unrelated families have confirmed the gene as the Batten disease gene. The disease gene encodes a novel 438 amino acid membrane binding protein of unknown function.

1. Background and Research Objectives

Positional cloning involves a combination of several molecular genetic strategies to clone disease genes by mapping their position to a limited region of DNA, and identifying mutations specific to affected individuals within these genes [1]. Prior to these new molecular techniques there was no way to identify a disease gene without first knowing what molecular defect occurred at the protein level. When the protein defect was identified, a predicted DNA sequence probe could be synthesized and used to isolate the disease gene. With most inherited

---

*Principal investigator, e-mail: doggett@lanl.gov

diseases however, no information is known about which proteins are involved, and therefore efforts to isolate these disease genes are based solely on determining their position in the genome, "positional cloning."

Batten disease (also known as Spielmeyer-Vogt disease, or juvenile onset neuronal ceroid-lipofuscinosis, or CLN-3) is a juvenile onset neurodegenerative disorder characterized by the accumulation of autofluorescent lipopigment in neurons (neuronal ceroid-lipofuscinosis) and other cells. Ultrastructurally, most cell types in affected patients display inclusions that have been termed fingerprint bodies. Patients suffer from pigmentary retinopathy and cerebral atrophy, resulting in progressive loss of vision and progressive spastic paralysis. In the terminal stages of the disease, patients are in a vegetative state, with death occurring when patients are in their twenties and thirties. Batten disease is inherited in an autosomal recessive fashion and is the most common neurodegenerative disorder of childhood with an incidence estimated at one in every twenty-five thousand live births. A variety of metabolic pathways including lipid peroxidation and dolichol-linked oligosaccharide metabolism are affected by the disorder, however the basic biochemical defect remains unknown. The Batten disease gene was localized to chromosome 16 by genetic linkage analysis [2, 3], raising hope that the disease gene may be identified and cloned by molecular genetic methods. A collaborative research agreement has been initiated between Los Alamos and other laboratories to work together to identify and clone the Batten gene.

The research objective of this project was to elucidate the molecular basis of Batten disease by molecular cloning of the affected gene with aid of chromosome 16 maps and resources of the Los Alamos Human Genome Program.

2. Importance to LANL's Science and Technology Base and National R&D Needs

The Human Genome Program at Los Alamos has produced low-resolution yeast artificial chromosome (YAC) map of human chromosome 16 [4]. This low-resolution YAC map is tightly integrated with a higher resolution, 60 percent coverage cosmid map [5] and a genetic map of chromosome 16. The value of this map lies in what can be learned about genome structure and function, and in what medically or scientifically important genes can be identified. The importance of genome structure and function to the genome mapping effort has been emphasized by the DOE. The isolation of Batten and other disease genes on chromosome 16 would clearly demonstrate the power of cloned physical maps for medical genetics. The development of such expertise in the field of positional cloning is of particular importance to the Human Genome Project at LANL.

3. Scientific Approach and Results

Batten disease was assigned to chromosome 16 by genetic linkage studies: DNA samples from families affected with the disease were collected and typed with genetic markers from all chromosomes [2]. Additional linkage studies have placed the Batten gene within an approximate 5 million base pair (Mb) region in the 16p11.2 position [5]. (Chromosome 16 is approximately 100 Mb, and within a 5 Mb region we would expect about 150 genes.) High density cosmid grids with these closest markers are screened to develop new markers from ends of these contigs. Linkage studies of these new markers (in collaboration with Mark Gardiner, University College in London) indicate which contig ends extend into the region and which ends extend outward. Additional cosmid clones are identified from probes synthesized from the inward contig ends and these steps repeated (cosmid walking) to identify the next adjacent cosmid or contig [6]. The cosmid vector, sCOS1, into which the library was constructed is specifically designed for the rapid generation of end probes for high efficiency cosmid walking [7,8].

Pulsed-field gel electrophoresis is used to place all genetically linked probes or markers on a physical map of the region. Several cosmid and YAC clones have now been mapped to the same region by somatic-hybrid-panel mapping and these are placed on the pulsed-field physical map. A pulsed-field map of the Batten region guides us in the use of additional strategies to track down the disease gene. If the region bracketing the Batten defect remains larger than 1 Mb after the first contigs and walks are identified, additional walking is necessary. The isolation of YAC clones is pursued to acquire large DNA fragments from the region. The CEPH MegaYAC library has already been screened extensively with sequence tagged sites (STSs) from the Batten region. These YACs however, have a high probability of containing internal deletions or of being chimeric clones. Thus, two 2X-coverage, chromosome-16-specific YAC libraries arrayed into pools have been screened using the polymerase chain reaction (PCR) with STSs that were developed from cosmid clones. The chromosome-16-specific YAC libraries have been found to have a lower chimeric rate and are believed to have fewer deletions as well.

Once Batten disease is confined to a region of a few to several hundred kb, exon trapping and exon amplification methods are applied to isolate genes. Candidate genes that are found are examined for gene expression in the brain, a target organ in Batten disease. Ultimately, it will be necessary to find mutations in the gene that occur in affected individuals but not unaffected ones. Gross level mutations such as large deletions will be immediately evident in affected DNA. Small deletions or point mutations will require a more detailed analysis. Two approaches are pursued to identify such mutations; 1) direct DNA sequencing of

PCR-amplified gene segments and (2) denaturing gradient gel analysis of PCR-amplified gene regions. Cloning of the Batten gene and identification of the genetic mutation(s) involved will elucidate the molecular basis of the disease and open new avenues for diagnosis, treatment, and cures.

During the last year of this project, two phenol sulphotransferase genes were mapped to the Batten disease region of chromosome 16 and were thus considered as candidates for the disease [9-11]. Mutations were not found in either of these genes, however, and they were therefore excluded for further consideration. A detailed physical map through the Batten disease region was completed [12,13], and this map provided the clonal materials for isolating new genes in the region [14]. A Batten disease patient with a microdeletion of chromosome 16 that encompassed the closest flanking genetic marker for the disease was discovered which further pinpointed the likely location for the disease [15]. Exon amplification of a cosmid that encompassed this deletion yielded a candidate gene that also carried a deletion. Mutational analysis of several patients revealed that this gene contains a 1 kb deletion in all Batten disease patients tested with the most common haplotype for the disease. In two unrelated families another deletion and a point mutation confirmed that this gene is the Batten disease gene. The disease gene encodes a novel 438 amino acid membrane binding protein of unknown function [16].

References

[1] Collins, F.S. Genome project can benefit search for disease genes. Human Genome News **2**: 5 (1990).

[2] Gardiner, M., Sandford, A., Deadman, M., Poulton, J., Cookson, W., Reeders, S., Jokiaho, I., Peltonnen, L., Eiberg, H., and Julier, C. Batten disease (Spielmeyer-Vogt disease, juvenile onset neuronal ceroid-lipofuscinosis) gene (CLN3) maps to human chromosome 16. Genomics **8**: 387-390 (1990).

[3] Mitchison, H., Thompson, A.D., Mulley, J.C., Kozman, H.M., Richards, R.I., Callen, D.F., Stallings, R., Doggett, N., Attwood, J., McKay, T.R., Sutherland, G.R., and Gardiner, R.M. Fine Genetic mapping of the Batten disease locus (CLN3) by haplotype analysis and demonstration of allelic association with chromosome 16p microsatellite loci. *Genomics* **16**:455-460 (1993).

[4] Doggett, N.A., Goodwin, L.A., Tesmer, J.G., Meincke, L.J., Bruce, D.C., Clark, L.M., Altherr, M.R., Ford, A.A., Chi, H.-C., Marrone, B.L., Longmire, J.L., Lane, S.A., Whitmore, S.A., Lowenstein, M.G., Sutherland, R.D., Mundt, M.O., Knill, E.H., Bruno, W.J., Macken, C.A., Torney, D.C., Wu, J.R., Griffith, J., Sutherland,

G.R., Deaven, L.L., Callen, D.F., and Moyzis R.K. An Integrated Physical Map of Human Chromosome 16. *Nature* **377:Suppl**:335-365 (1995).

[5] Stallings, R.L., Torney, D.C., Hildebrand, C.E., Longmire, J.L., Deaven, L.L., Jett, J.H., Doggett, N.A., and Moyzis, R.K. Physical mapping of human chromosomes by repetitive sequence fingerprinting. *Proc. Natl. Acad. Sci. USA.* **87**:6218-6222 (1990).

[6] Hadfield, C. Chromosome Walking. Focus **5**: 1-5 (1983).

[7] Evans, G.A., & Wahl, G.M. Cosmid vectors for genome walking and rapid restriction mapping. *Methods in Enzymology* **152**: 604-610 (1987).

[8] Evans, G.A., & Lewis, K.A. Physical mapping of colmplex genomes by cosmid multiplex analysis. *Proc. Natl. Acad. Sci. USA* **86**:5030-5034 (1989).

[9] Dooley, T.P., Obermoeller, R.D., Probst, P., Mole, S., Doggett, N.A., Callen, D.F., and Siciliano, M. Phenol Sulfotransferases: Candidate genes for Batten disease. *American Journal of Medical Genetics* **57**:327-332 (1995).

[10] Dooley, T.P., Mitchison, H.M., Munroe, P.B., Probst, P., Neal, M., Siciliano, M.J., Deng, Z., Doggett, N.A., Callen, D.F., Gardiner, R.M., and Mole, S.E. Mapping of two phenol sulphotransferase genes, STP and STM, to chromosome 16p: Candidate genes for Batten Disease. *Biochem. Biophys. Res. Comm.* **205**:482-489 (1994).

[11] Dooley, T.P., Probst, P., Munroe, P.B., Mole, S.E., Liu, Z, and Doggett, N.A. Genomic organization of the human catecholamine-sulfating phenol sulfotransferase gene (STM). *Biochem. Biophys. Res. Comm.* **205**:1325-1332 (1994).

[12] Järvelä, I.E., Mitchison, H.M., Callen, D.F., Lerner, T.J., Doggett, N.A., Taschner, P.E.M., Gardiner, R.M., and Mole, S.E. Physical map of the region containing the gene for Batten disease (*CLN3*). *American Journal of Medical Genetics* **57**:316-319 (1995).

[13] Järvelä, I.E., Mitchison, H.M., O'Rawe, A.M., Taschner, P.E.M., de Vos, N., Lerner, T.J., D'Arigo, K.L., Callen, D.F., Thompson, A.D., Knight, M., Marrone, B.L., Mundt, M.O., Meincke, L., Breuning, M.H., Gardiner, R.M., Doggett, N.A., and Mole, S.E. YAC and cosmid contigs spanning the Batten disease region (CLN3) at 16p12.1-p11.2. *Genomics* **29**:478-489 (1995).

[14] Lerner, T.J., D'Arigo, K.L., Haines, J.L., Doggett, N.A., Taschner, P.E.M., Buckler, A.J., and the Batten Disease Consortium. Isolation of genes from the Batten candidate region using exon amplification. *American Journal of Medical Genetics* **57**:320-323 (1995).

[15] Taschner, P.E.M., De Vos, N., Thompson, A.D., Callen, D.F., Mole, S.E., Dooley, T.P., Doggett, N.A., Barth, P.G., and Breuning, M.H. Chromosome 16 microdeletion

in a patient with juvenile neuronal ceroid lipofuscinosis (Batten desease). *American Journal of Human Genetics* **56**:663-668 (1995).

[16] The International Batten Disease Consortium: Lerner, T.J., Boustany, R.-M.N., Anderson, J.W., D'Arigo, K.L., Schlumpf, K., Buckler, A.J., Gusella, J.F., and Haines, J.L.; Kremmidiotis, G., Lensink, I.L., Sutherland, G.R., and Callen, D.F.; Taschner, P.E.M., de Vos, N., and Breuning M.H.; Doggett, N.A., Meincke, L.J., Liu, Z.-Y., Goodwin, L.A., and Tesmer, J.G.; Mitchison, H.M., O'Rawe, A., Munroe, P.B., Järvelä, I.E., Gardiner, R.M., and Mole, S.E. The major Batten disease (CLN3) mutation is a 1 kb deletion in a novel gene. *Cell* **82**:949-957 (1995).

DISCLAIMER