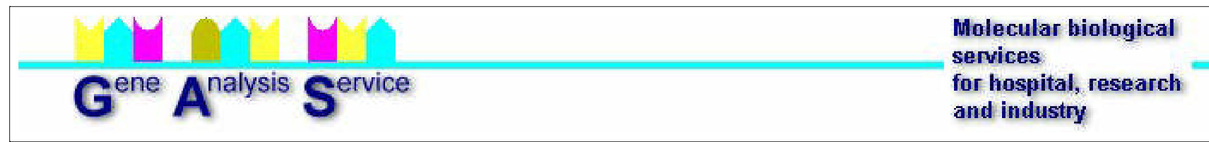
# Molecular biological detection of rare heriditary diseases
# The screening of genes for mutations

**Dr. Alfred Cornelis Looman Ph.D. - Gene Analysis Service GmbH - Berlin - Germany**
+49-(0)30-84707430 - Fax +49-(0)30-84707432 - E-mail: looman@gene-analysis-service.de
Internet: www.gene-analysis-service.de



# 1. GENETIC SCREENING

Since the completion of the working-draft of the human genome in February 2001, nearly every location on the human DNA can be screened for individual changes in the DNA sequence. Base changes, deletions or insertions in protein-coding exons of approximately 30.000 genes, regulatory sequences or splice sites at intron-exon boundaries can be the molecular basis of disease. Up to 6000 diseases are thought to be mono-genetic, that is caused by mutations in one gene. Most other diseases are caused by a combination of mutations in interacting genes and other, non- genetic, factors.

Screening for mutations can help support medical diagnostics in those cases where disease-causing mutations have been identified and can be distinguished from normal polymorphisms. Prenatal screening supports genetic councilling in those cases where lethal mutations can occur. Mutation screening is able to detect disease causing mutations in an early stage before symptoms of a disease are apparent or in heterozygous carriers of a defective allele.

The general principle of a fast genetic analysis by DNA sequencing is illustrated by the screening for mutations in the ATP7B gene. This gene codes for a membrane-bound, copper-transporting protein in the liver and has been recognized as the defective gene in **Wilson's disease**, an inherited autosomal recessive disorder that causes copper accumulation in the liver and the brain. Clinically, patients usually present as older children or young adults with hepatic, neurologic, or psychiatric manifestations, or some combination of these. Wilson's disease occurs in about one of every 40.000 people.

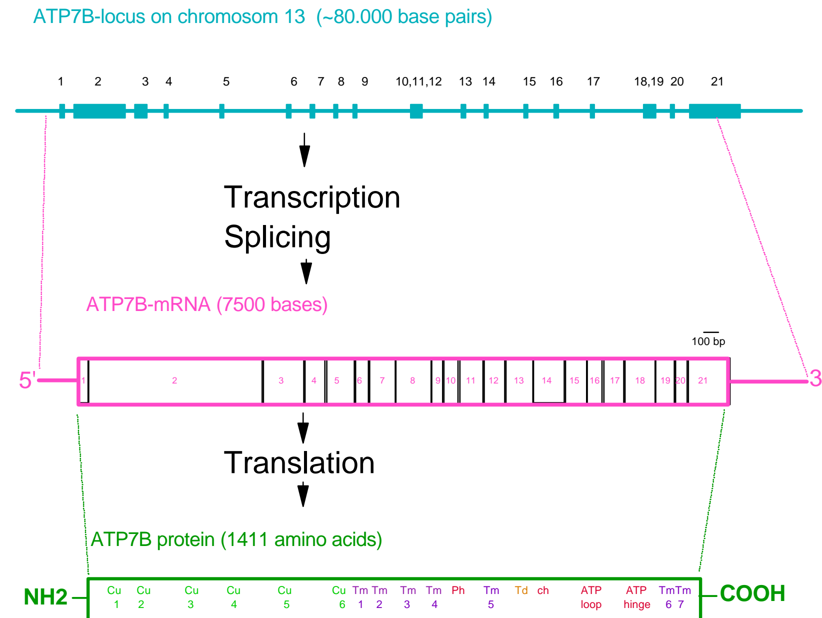
The prevention of severe permanent damage depends upon early recognition and diagnosis. Wilson's disease can be treated very effectively by anticopper treatments like zinc and penicillamine.

Many individual mutations in the protein-coding exons of ATP7B, as well as in intron-sequences, have been reported in patients suffering from the disease. One mutation in exon 14 (His1069Gln) is reported in a majority of cases in Europe.

## 2. DETECTION OF THE RESPONSIBLE GENE

The molecular basis of Wilson's disease are defects in the copper transporting P-Type-ATPase (ATP7B). The gene coding for the ATP7B-protein was localized on the human chromosome 13 by several research groups between 1985 and 1988. In 1993 the gene was isolated and its structure and sequence completely elucidated [1-3].

The information of the ATP7B-gene lies spread over 80.000 bases of the human genome and is divided into 21 protein-coding exons. The gene is transcribed and spliced into the 7500 base messenger RNA, which is translated into the 1411 amino acid copper transporting ATPase (Fig. 1). The sequencing of the genomic DNA of the protein coding exons and the exon-intron boundaries detects splice site mutations leading to exon skipping, insertions and deletions leading to frame-shift mutations as well as point mutations in the coding sequence of the gene leading to amino acid changes.




### Figure 1. From the gene to an effective protein

The ATP7B-gene was localized on chromosome 13q14.3. It comprises 80.000 base pairs on 21 protein coding exons (thick lines) and 20 non-coding introns (thin lines).

The ATP7B gene is transcribed and processed (*spliced*) into the 7500 bases mRNA in liver, brain, kidney and placenta tissue.

The mRNA is translated into the 1411 amino acid, 159 kDa ATPase. It contains several functionally important domains: for copper binding (Cu 1-6); for correct localization in the cell membrane (Transmem- brane domains Tm 1-7); for the passage of copper through the membrane (Transducing domain Td); and for the energy-dependent translocation of bound copper molecules: ATP-binding, ATP-hinge, ATP-loop, phosphatase , and channel).

## 3. DISCRIMINATING BETWEEN MUTATIONS AND POLYMORPHISMS

Since the complete sequencing of the ATP7B gene in 1993, a molecular biological analysis of Wilson's disease is possible. By comparing the ATP7B sequences of patients with control groups, many variable positions in the gene sequence were found. They divide in such variations that can be found in the control population and are not associated with a malfunction of the protein (*polymorphisms*), and such that are only found in the WD patient group. These mutations impair the function of the ATP7B-protein.

To date, over 200 mutations were found that impair or completely abolish the copper transporting function of the ATP7B-protein[4,5)]. Among them are deletions, insertions, point mutations and splice site mutations in all exons of the gene. The most common mutation is a C to A transversion in codon 1069 of exon 14, causing a histidine to glutamine change, impairing ATP-binding (S). Many WD patients are compound heterozygous, carrying different mutations on both chromosomes. Mutations in functional domains of the ATP7 protein impair or even prevent copper transport. Only the correct interaction of all protein domains allows a regulated copper secretion. The ATP7B-protein embedded in the cell membrane is schematically shown in Fig. 2.

The methods used in a genetic analysis are presented in the following sections.

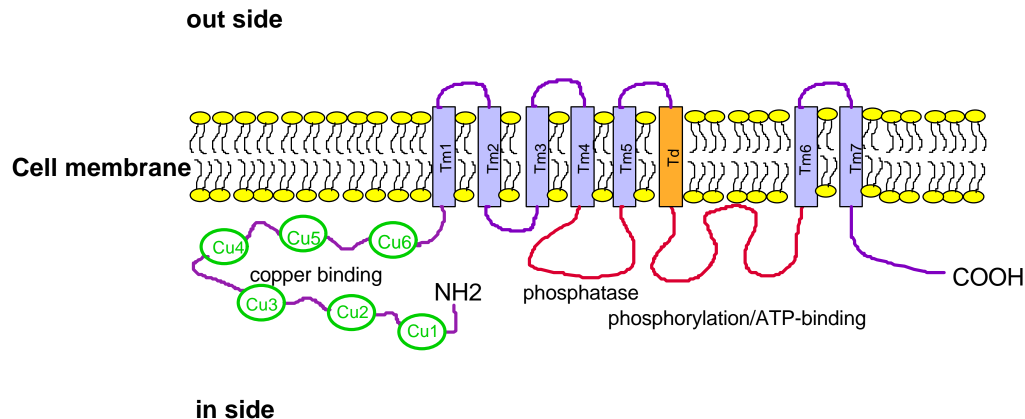



**Fig. 2. The protein at is position in the canilicular membrane**
Positioning of the ATP7B-Protein in the cell membrane[6]. The copper-binding domains **Cu1-6** contain the amino acid motif Gly-Met-Thr-Cys-X-Ser-Cys. The phosphatase domain contains the motif Thr-Gly-Glu-Ala. The transmembrane domain **Td** contains the motiv Cys-Pro-Cys. The phosphorylation domain is characterized by the amino acid sequence Asp-Lys-Thr-Gly-Thr, the ATP-binding site by Ser-Glu-**His**-Pro-Leu, Thr-Gly-Asp-Asn and Gly-Asp-Gly-Val- Asn-Asp.

Mutations in the ATP7B-gene lead to a defective mRNA and in consequence to a defective protein, unable to function in copper transport. The frequent mutation **His**1069**Gln** disables ATP-binding.

## 4. ISOLATING GENOMIC DNA FROM BLOOD

The preferred starting material is EDTA-blood (**A**). DNA is isolated from the nucleated white cells, which are lysed enzymatically at $55^{0}$C (**B**). The genomic DNA released by this procedure is further purified by precipitation and centrifugation (**C**).

Thus purified DNA is the source for a genetic analysis. It contains the information for the ~30.000 human genes on 23 chromosome pairs. From 0.2 ml whole blood ca. 6 microgram DNA is extracted, sufficient for more than 10.000 PCR amplifications. Other material (e.g. pre-natal samples, tissue biopsies or paraffin embedded material) can also be used for DNA isolation.

**A:** Starting material blood

**B:** Lysing of the white blood cells

**C:** Centrifugation of the isolated DNA

## 5. AMPLIFICATION OF SELECTED DNA SEQUENCES

The analysis of one gene of 7500 bases like the ATP7B gene amongst the 3 billion bases of the human DNA is like finding a needle in the haystack. To enable this, gene fragments are specifically amplified by means of the polymerase chain reaction (*PCR*).

In a cabinet, shielded from contaminations (**D**), the following components are mixed together: A small amount of purified genomic DNA, DNA building blocks (*nucleotides*), short (20 bp.) synthetic oligonucleotides homologous to the beginning and the end of the target gene fragment (*primer*) and a thermostable enzyme (*Taq DNA-polymerase*), able to copy DNA molecules reliably *in vitro*.

This reaction mixture runs through 25-35 cycles of temperature changes in a thermocycler (**E-F**): 30 sec at 95°C, 25 sec at 50-65°C, 60 sec at 72°C. In every cycle of this reaction, the targeted DNA fragment is doubled. At the end of the reaction $2^{35}$ or ~34.000.000.000 copies of the same gene fragment are present. The relationship of needle and hay-stack now reversed.

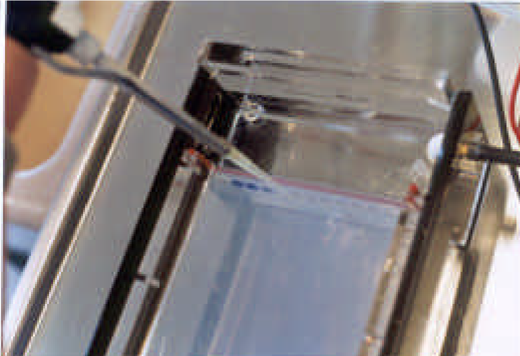
**D:** PCR-working cabinet

**E:** Thermocycler

**F:** Thermocycler-detail

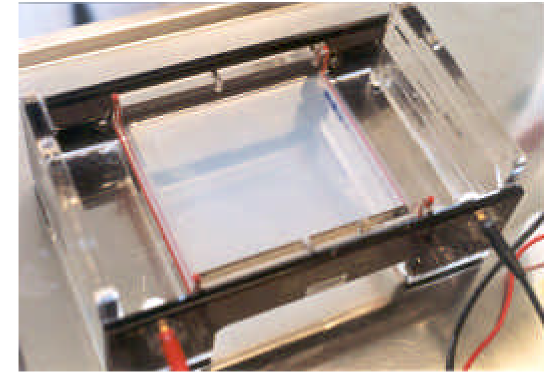
# 6. ISOLATION OF AMPLIFIED GENE FRAGMENTS

The polymerase chain reaction is performed for every fragment of the gene of interest. These amplified fragments are separated according to their size by agarose gel electrophoresis (**G-I**), stained with fluorescent DNA-binding agents, and made visible by UV-light (**J**). Fragments can now be isolated from the gel (**K**) and purified further for applications like restriction fragment analysis or DNA-sequencing.



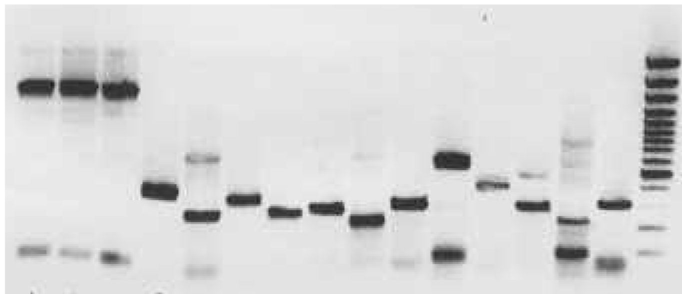
**G:** Loading samples on agarose gel

**H:** Adjusting electrophoresis current



**I:** Separation of DNA-fragments

exon: 2 2 2 3 4 8 9 15 16 17 18-19 5 7 20 14 marker



**J:** ATP7B fragments stained with ethidium bromide

**K:** Isolation of separated fragments

## 7. DNA SEQUENCING OF GENE FRAGMENTS

With a similar method as was used in their amplification (*cycle-sequencing*), the gene fragments are prepared for the analysis of their base sequence. In this reaction, fluorescence labeled nucleotides (**G**, **A**, **T** and **C**) are incorporated into newly synthesized single stranded molecules, using the purified gene fragments as a template. Following the sequencing reaction, these molecules are separated by capillary electrophoresis in an automatic DNA-sequencer (**L,M**) and visualized by laser light. Fluorescence signals are turned into bases by computing (**N**) and in result, the sequence of the fragment can be read.

**L:** DNA-sequencing aparatus

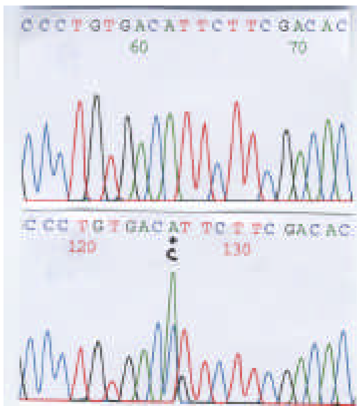
**M:** In side view of sequencer

**N:** Turning fluorescence signals into DNA sequence

# 8. ANALYSIS OF THE RESULTS

The established DNA sequences of gene fragments are compared with the known DNA sequence. Deviations are divided into normal variations (*polymorphisms*) and changes with consequences for the functioning of the gene product (*mutations*). Mutations and polymorphisms can be present on only one chromosome (*heterozygote*), both chromosomes can be affected by the same mutation (*homozygote),* or by different mutations *(compound heterozygote).* In figures (**O-S**) some mutations and polymorphisms in the ATP7B-gene are shown:

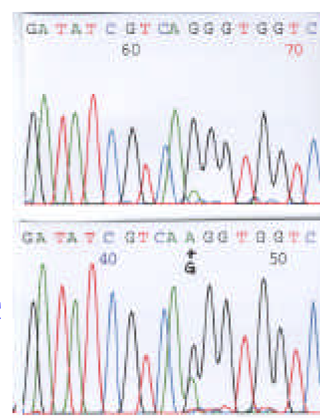
**O**: Polymorphism in Exon 8: Thr$^{762}$Thr



ACA
homozygote

ACA and ACC
Heterozygote

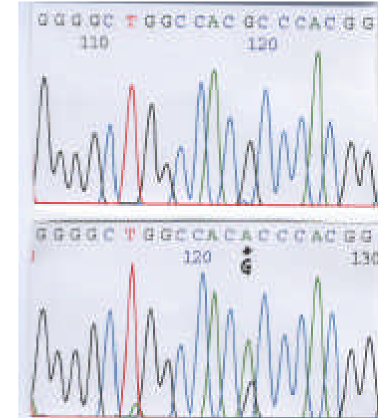
**P**: Polymorphism in Exon 10: Arg$^{832}$Lys



AGG (Arg)
homozygote

AGG and AAG
Arg and Lys
heterozygote

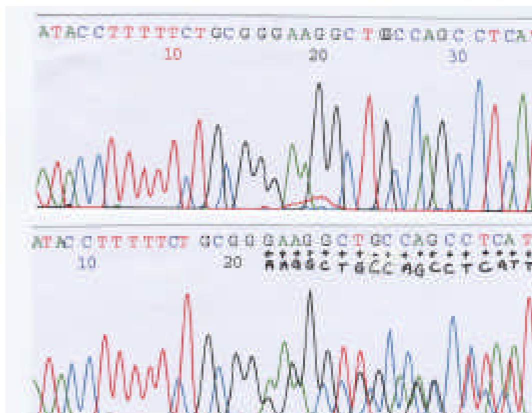
**Q**: Polymorphism in Exon 13: Thr$^{991}$Thr



ACG
homozygote

ACG and ACA
heterozygote

**R**: Mutation in Exon 15: 3400delC

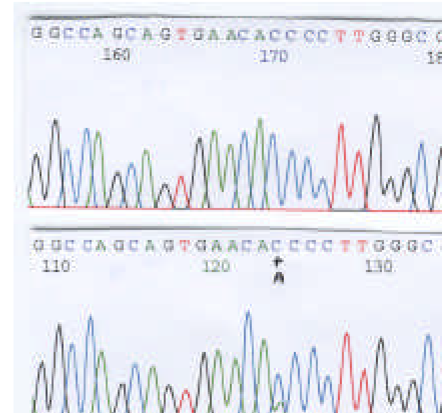


Normal base sequence

Homozygote

On one chromosome a C is
deleted at position 3400
heterozygote
After amino acid Pro$^{1134}$ the
protein is truncated.

**S**: Mutation in Exon 14: His$^{1069}$Gln



Normal base sequence: CAC (His)
homozygote

CAC (His) and CAA (Gln)
heterozygote
From one chromosome a
defective protein is expressed.

# 9. CONCLUSIONS

As a result of the developments of the last decades, it has become possible to track down the molecular basis, and with it, the cause of many inherited diseases. The publication of the results of the Human Genome Project only marks the beginning of a new era of medical diagnosis and therapy. By establishing the individual gene sequences and detection of mutations and polymorphisms, **predictions about the occurrence** and the severity of a disease will become possible. Individual and optimal therapy becomes feasible.

Gene analysis supports **the clinical diagnose** in cases of doubt or when clear symptoms of a disease are not present yet.

By gene analysis, **carriers of a genetic disease** among relatives of patients who show no symptoms themselves, but can inherit the disease further, can be discovered. **Prenatal gene analysis** can help consulting parents that carry genetic diseases or are affected by a serious genetic disease.

In theory, it is possible to transport and integrate functional gene sequences into cells and organs to „repair" gene defects (*gene therapy*). A still unsolved problem is the permanent, efficient transfection of all cells of target organs or cell populations.

## LITERATURE REFERENCES

[1] Bull, P.C. *et al*. (1993) *The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene*. Nature Genetics **5**:327-337.

[2] Tanzi, R.E. *et al*. (1993) *The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene*. Nature Genetics **5**:344-350.

[3] Petrukhin, K. *et al*. (1993) *Mapping, cloning and genetic characterization of the region containing the Wilson disease gene*. Nature Genetics **5**:338-343.

[4] Kenney, S. and Cox, D.W. (March 2001) *Wilson Disease mutation database*. http://www.medgen.med.ualberta.ca/database.html

[5] Bonne-Tamir, B. and Kreisberg-Zakarin (Feb. 2001) *Hum. Genet. Disease Database-Wilson's Disease.* http://life2.tau.ac.il/GeneDis/Tables/Wilson/wilson.html

[6] DiDonato, M. and Sarkar, B. (1997) *Copper transport and its alterations in Menkes and Wilson diseases*. Biochim. Biophys. Acta **1360**:3-16.

## CREDITS

We thank Dr. E. Heyl, Dr. W. Parr and Dr. J. Ruprecht of Heyl Chem.-Pharm. Fabrik GmbH & Co. KG, Berlin, for initiating Wilson's disease molecular diagnostics. Our special gratitude goes to Dr. H.-J. Kühn, Neurological Clinic, University of Leipzig, for providing us with patient material. We are obliged to Mr. Greg Morel Bsc (Hons) Msc -Trinity, Jersey, United Kingdom, for correction of the manuscript.

**Gene Analysis Service GmbH - Goerzallee 253 - D 14167 Berlin - Germany**

Tel. +(49)-30-84707430 - Fax +(49)-30-84707432 - E-mail: looman@gene-analysis-service.de - Internet: www.gene-analysis-service.de