Discovery of the *NUBPL* gene and its role in Complex I assembly

Janneke Balk

University of East Anglia and John Innes Centre, Norwich, UK

www.jic.ac.uk/profile/janneke-balk.asp

For many decades now, scientists have been trying to find out what the function is of every one of the 20,000 genes in our genome. This is a very slow process, but it has accelerated in the age of genome sequencing. Often, human genes have similar counterparts in other organisms, such as bacteria, yeast or plants, that are easier to study.

In 2005, Dr Janneke Balk and Dr Katrine Bych, then at the University of Cambridge, decided to study the function of a gene of unknown function in plants. The gene sequence translated into a protein with interesting features, such as possible targeting to mitochondria and binding of iron-sulfur clusters (Figure 1), two major research interests of Dr Balk. Experiments confirmed that the protein was located in mitochondria, and that it bound iron. However, to study the precise function, they needed to disrupt the gene. This was more easily done in a yeast species called *Yarrowia lipolytica*, for which they collaborated with Dr Stephan Kerscher and Professor Ulrich Brandt at the University of Frankfurt. It took 3 months to isolate the yeast mutant lacking the gene, which they called *IND1*. It was immediately clear that the activity of Complex I, a large protein complex in mitochondria, was strongly decreased in this mutant (Bych *et al* 2008). With this knowledge, Dr Alex Sheftel in the laboratory of Professor Roland Lill, University of Marburg, went on to show that depletion of the analogous protein in human cells, called NUBPL, also decreased Complex I activity (Sheftel *et al* 2009).
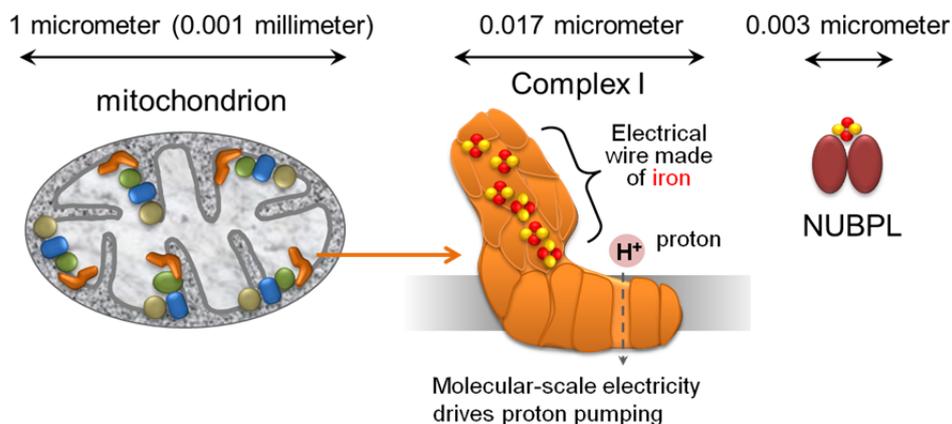

Figure 1: What are mitochondria, Complex I and NUBPL? Our body consists of more than 30 trillion cells. Each cell combusts small amounts of dietary carbon to generate energy. Most of the combustion process takes place in tiny sub-compartments of the cell called mitochondria (plural for mitochondrion). A series of protein complexes ensures the combustion happens step-by-step and in a controlled manner. Complex I is one of those protein complexes. It consists of more than 40 individual proteins. NUBPL is not part of Complex I, but helps specifically with its assembly. Currently we don't know what the precise function of NUBPL is. This is the main research question addressed by Andrew Maclean, a PhD student in Dr Balk's laboratory.

Identification of pathogenic mutations in *NUBPL* (2010 – 2012)

Thanks to the internet and unknown to Dr Balk, the 2008 publication on the yeast gene was spotted by Dr Elena Tucker on the other side of the world, in the Murdoch Childrens Research Institute in Melbourne. She was part of an international consortium that were about to perform sequence analysis on 100 patients with Complex I deficiency. As a last minute decision, the human *NUBPL* gene was added to the list of genes for sequencing (Elena Tucker, *personal communication*). Note that these researchers were looking at a subset of genes only, because in 2009 it was still technically difficult and costly to sequence a whole genome. Moreover, only the protein-coding sequence was determined, known as exome sequencing. In one of the patients, they found a point mutation in the paternal copy of *NUBPL* (c.166G>A), while the maternal copy had large sections deleted. The Complex I levels in cells from the patient could be restored to normal with a healthy copy of the *NUBPL* gene (Calvo *et al* 2010).

Dr Tucker tried to show that the c.166G>A mutation was indeed pathogenic, but it actually did not cause Complex I deficiency (Tucker *et al* 2012). She then sequenced the whole *NUBPL* gene, including the introns which are normally removed from the mature transcript before it is translated into a protein (Figure 2). This revealed a c.815-27T>C mutation, namely a mutation 27 nucleotides into an intron, counting back from nucleotide 815 of the coding sequence. Dr Tucker and colleagues showed that the mutation can cause 'exon skipping', namely exon 10 is inadvertently removed and exon 9 is directly fused to exon 11. This results in a shorter protein with a faulty amino acid sequence towards the end. We call this a splice-site or branch-site mutation. Some normal transcript is also produced in the patient's fibroblast cells, but only about 15% compared to normal levels. (Tucker *et al* 2012). The c.815-27T>C mutation occurs at a frequency of 1.2% in the human population, meaning that 1 out of every 120 people carries it. Therefore we expect more NUBPL patients to be identified.
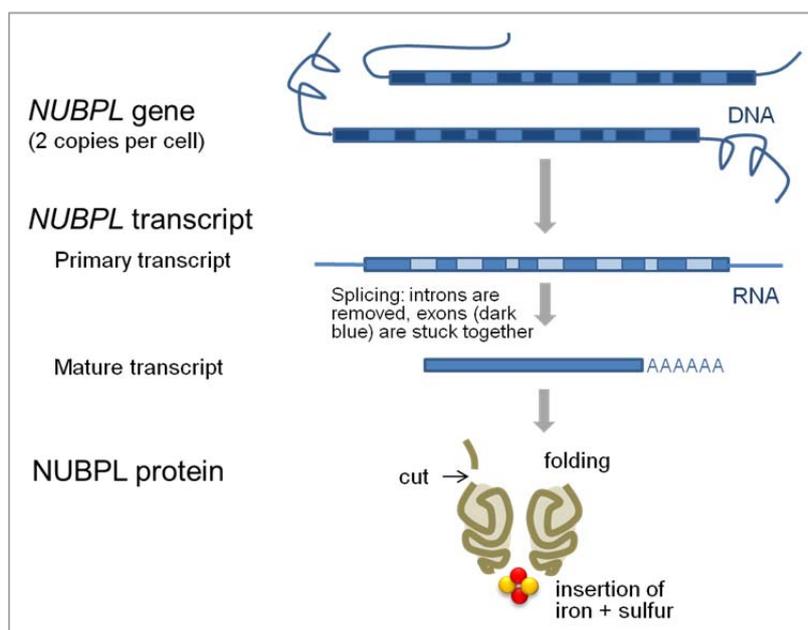

Figure 2: From gene to protein:

1. Each cell has two copies of the *NUBPL* gene in its DNA. One copy is derived from the mother, one from the father.

2. Genes are the hardware of the cell. They are used as a template to make multiple transcripts (RNA).

3. The primary transcript is further processed. So-called introns are removed by a mechanism called splicing. The ends are trimmed and a poly-A tail is attached.

4. The mature transcript is translated to a protein: each combination of 3 nucleotides (A, C, G or T) codes for 1 amino acid. So the *NUBPL* coding sequence of 957 nucleotides makes a protein of 319 amino acids.

5. Like most newly produced proteins, NUBPL is moved to the part of the cell where it does its job. In the case of NUBPL, it travels to the mitochondria where it is trimmed, then folded into a globular structure. Two globules combine, binding iron and sulfur atoms between them (Bych *et al* 2008).

To investigate the functionality of the shorter protein, Dr Mateusz Wydro in the laboratory of Dr Balk recreated the mis-spliced transcript in Yarrowia yeast. Dr Wydro found that it could not compensate for the lack of the gene in the deletion strain (Wydro and Balk, 2013). He also noticed that the shorter protein was not stable, i.e. most was degraded. This was also observed in patient cells by Dr Tucker. These data demonstrate the pathogenicity of the c.815-27T>C mutation, although a contribution from the c.166G>A mutation can as yet not be ruled out.

Other mutations in *NUBPL* (2013 -)

Once mutations in *NUBPL* were linked to Complex I deficiency, this helped with finding other pathogenic mutations and provided a genetic diagnosis for a number of patients (Kevelam *et al* 2013; Kimonis *et al*, unpublished). Clinical symptoms manifest themselves at a young age; therefore patients come in through paediatrics departments.

All 12 NUBPL patients found so far have the c.815-27T>C branch-site mutation. The c.166G>A is usually also present in the same gene copy, but not in a recent case found by Dr Kimonis. The branch-site mutation can be inherited paternally or maternally. None of the patients found so far have two copies of the gene with the branch-site mutation. Instead, the other copy of the *NUBPL* gene always carries different mutations, which is termed 'compound heterozygous' by geneticists. These mutations are predicted to severely impair the function of NUBPL, either because most of the sequence is deleted (patient DT35 in Calvo *et al* 2010), because the protein terminates early, or because a very important amino acid is changed into the 'wrong' amino acid. In all the different types of *NUBPL* mutation, clinical symptoms are very similar. In particular, patterns observed in MRI scans of the brain are rather specific for NUBPL defiency (Kevelam *et al* 2013).

KEY RESEARCH QUESTIONS

- What is the molecular function of NUBPL? How does NUBPL influence Complex I assembly?
- What is the best treatment for NUBPL patients? Is there a specific treatment or is a generic treatment for mitochondrial disease better?

Dr Virginia Kimonis, Dr Janneke Balk and Andrew Maclean in the laboratory in Norwich, UK

References

Bych, K., Kerscher, S., Netz, D. J., Pierik, A. J., Zwicker, K., Huynen, M. A., Lill, R., Brandt, U. and Balk, J. (2008). The iron-sulphur protein Ind1 is required for effective complex I assembly. *EMBO J*. **27**, 1736-1746.

Calvo, S. E., Tucker, E. J., Compton, A. G., Kirby, D. M., Crawford, G., Burtt, N. P., Rivas, M., Guiducci, C., Bruno, D. L., Goldberger, O. A. et al. (2010). Highthroughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nat. Genet.* **42**, 851-858.

Kevelam, S. H., Rodenburg, R. J., Wolf, N. I., Ferreira, P., Lunsing, R. J., Nijtmans, L. G., Mitchell, A., Arroyo, H. A., Rating, D., Vanderver, A. et al. (2013). NUBPL mutations in patients with complex I deficiency and a distinct MRI pattern. *Neurology* **80**, 1577-1583.

Sheftel, A. D., Stehling, O., Pierik, A. J., Netz, D. J., Kerscher, S., Elsässer, H. P., Wittig, I., Balk, J., Brandt, U. and Lill, R. (2009). Human ind1, an iron-sulfur cluster assembly factor for respiratory complex I. *Mol. Cell. Biol*. **29**, 6059-6073.

Tucker, E. J., Mimaki, M., Compton, A. G., McKenzie, M., Ryan, M. T. and Thorburn, D. R. (2012). Next-generation sequencing in molecular diagnosis: NUBPL mutations highlight the challenges of variant detection and interpretation. *Hum. Mutat.* **33**, 411-418.

Wydro, M. M. and Balk, J. (2013). Insights into the pathogenic character of a common *NUBPL* branch-site mutation associated with mitochondrial disease and complex I deficiency using a yeast model. *Dis. Models Mech.* **6**, 1279-1284.