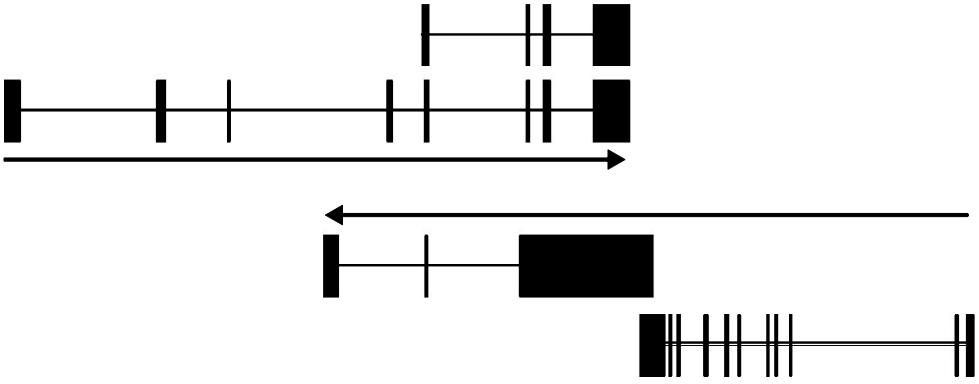
Thesis for doctoral degree (Ph.D.)
2007

Functional Genomics Studies of PINK1



Camilla Scheele

Karolinska Institutet

Thesis for doctoral degree (Ph.D.) 2007
Functional Genomics Studies of PINK1
Camilla Scheele

From the Programme for Genomics and Bioinformatics
Department of Cell and Molecular Biology
Karolinska Institutet, Stockholm, Sweden

FUNCTIONAL GENOMICS STUDIES OF PINK1

Camilla Schéele

Karolinska Institutet

Stockholm 2007

Published by Karolinska Institutet. Printed by Larserics Digital Print AB

ISBN 978-91-7357-376-4

ABSTRACT

Functional genomics has become an important and established research discipline during the last 10 years, mainly as a consequence of the completion of large-scale genome sequencing projects. The human genome is now predicted to transcribe 20,000 – 25,000 protein coding genes, which is only a quarter of the number suggested a few years ago. Instead, a dynamic RNA universe seems to provide diversity to mammalian cells. Around 60-70% of protein coding genes are predicted to generate two or more transcripts by alternative splicing, while non-protein coding RNA are also directly transcribed from the genome. One category of such non-protein coding RNA is *cis*-transcribed natural antisense (NAT). *Cis*-NATs are transcribed from a gene's antisense DNA strand and are suggested to have regulatory functions through direct interaction. The work described in this thesis includes functional genomics studies of the *PINK1* locus.

We undertook a study to discover novel candidate genes associated with physical inactivity, a known risk factor for type 2 diabetes. Gene expression profiling of skeletal muscle from subjects before and after 5 weeks of inactivity by quantitative real-time PCR demonstrated a co-ordinated reduction in mitochondrial gene expression. We thus established a human *in vivo* model for mitochondrial dysfunction. Microarray analyses of the same sample-set suggested that PINK1, a novel mitochondrial kinase, was down regulated during inactivity. PINK1 is transcribed from a complex locus, alternatively spliced and with an annotated *cis*-NAT. Mutations at this locus had also been linked to Parkinson's disease and we thus selected this locus for subsequent functional genomics studies. We utilized human *in vivo* models, gene expression and genomic association analysis and RNA interference (RNAi) to study the regulation of PINK1.

We demonstrated dynamic expression from the *PINK1* locus during modulation of mitochondria *in vivo* in human skeletal muscle. PINK1 was down regulated in our mitochondrial dysfunction model, while a shorter splice variant of PINK1 (svPINK1) and the NAT (naPINK1) were concordantly up regulated. The opposite expression pattern was obtained in a human *in vivo* model for increased mitochondrial activity, suggesting a direct association between svPINK1 and naPINK1. Knockdown of naPINK1 utilizing siRNAs targeting two different sites of naPINK1 reduced the level of svPINK1. This directly supports a role for naPINK1 in promoting the abundance of svPINK1, a novel mechanism for regulation by natural antisense.

In contrast, *all* transcripts from the *PINK1* locus were less abundant in muscle tissue from diabetics, compared to healthy controls. To investigate whether PINK1 transcript levels could affect metabolic fitness or if the lower expression rather was a secondary effect of diabetes, we measured *PINK1* tagging single nucleotide polymorphisms (SNPs). Several SNPs associated with PINK1 transcripts levels. The genotypes associating with higher expression of PINK1 also associated with lower plasma levels of non-esterified fatty acid levels (NEFA) and glucose. Two sets of RNA interference studies provided support for these clinical associations. Firstly, knockdown of PINK1 in human neuroblastoma cells resulted in impaired basal glucose uptake. Secondly, FABP4, a lipid transport protein, was selectively down regulated following PINK1 knockdown in adipocytes. However, mitochondrial genes were not altered when PINK1 expression was ablated, despite the *in vivo* association between such genes. Taken together, our data suggest a role of PINK1 in lipid and glucose metabolism while PINK1 does not appear to be essential for mitochondrial biogenesis in mammalian cells.

LIST OF PUBLICATIONS

The thesis is based on the following original articles, referred to in the text by their Roman numerals I-IV.

I. James A Timmons, Jessica Norrbom, **Camilla Schéele**, Håkan Thonberg, Claes Wahlestedt, Per Tesch. Expression profiling following local muscle inactivity in humans provides a new perspective on diabetes-related genes. Genomics, 2006, 87, 165-172

II. **Camilla Schéele**, Natasa Petrovic, Mohammad A Faghihi, Timo Lassmann, Katarina Fredriksson, Olav Rooyackers, Claes Wahlestedt, Liam Good and James A Timmons.The human PINK1 locus is regulated in vivo by a non-coding natural antisense RNA during modulation of mitochondrial function. BMC Genomics, 2007, 8

III. **Camilla Schéele**, Anders Rinnov Nielsen, Tomas B Waldén, Christian P Fischer, Robert J Brogan, Natasa Petrovic, Ola Larsson, Per A Tesch, Kristian Wennmalm, Dana S Hutchinson, Barbara Cannon, Claes Wahlestedt, Bente K Pedersen and James A Timmons. Altered regulation of the PINK1 locus: a link between Type 2 diabetes
and neurodegeneration?. FASEB Journal, 2007, 21

IV. Paul W Franks*, **Camilla Schéele***, Ruth J Loos, Anders R Nielsen, Francis M Finucane, Claes Wahlestedt, Bente K Pedersen, Nick J Wareham, James A Timmons. Genomic Variants at the *PINK1* Locus are Associated with Transcript Abundance and Biomarkers of Oxidative Energy Metabolism in a Concordant Manner. Submitted

(*These authors contributed equally to this work)

OTHER PUBLICATIONS

Anna Josephson, Alexandra Trifunovski, **Camilla Schéele**, Johan Widenfalk, Claes Wahlestedt, Stefan Brené, Lars Olson, Christian Spenger. Activity-induced and developmental downregulation of the Nogo receptor. Cell and Tissue Research. 2003 Mar;311(3):333-42.

Ola Larsson, **Camilla Schéele**, Zicai Liang, Jürgen Moll, Christina Karlsson and Claes Wahlestedt. Kinetics of Senescence-associated Changes of Gene Expression in an Epithelial, Temperature-sensitive SV40 Large T Antigen Model. Cancer Research. 2004, 64, 482-489.

Håkan Thonberg, **Camilla Schéele**, Cecilia Dahlgren and Claes Wahlestedt. Characterization of RNA interference in rat PC12 cells: requirement of GERp95. Biochemical and Biophysical Research Communications. 2004, 318, 927-934.

CONTENTS

LIST OF ABBREVIATIONS

AGE	Advanced glycation end-products
Ct	Cycle threshold
DDOST	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase
EST	Expressed sequence tag
LD	Linkage disequilibrium
mRNA	Messenger RNA
NAT	Natural antisense
PINK1	PTEN Induced Putative Kinase 1
qPCR	Quantitative real-time PCR
RIN	RNA integrity number
RISC	RNA induced silencing complex
RITS	RNA induced initiation of transcriptional gene silencing
RNAi	RNA interference
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
siRNA	Short interfering RNA
snoRNA	Small nuclear RNA
SNP	Single nucleotide polymorphism
tRNA	Transfer RNA

1 INTRODUCTION

1.1 THE MAMMALIAN TRANSCRIPTOME

RNA molecules have the potential to act both as information carriers and catalysts and could thus catalyze their own synthesis. The "RNA world" hypothesis, which holds that RNA was the first life, is now a well established concept [1,2,3,4]. To date, RNA molecules still have a wide range of functions in the cell. Besides the coding messenger RNA (mRNA), the non coding ribosomal RNA (rRNA) and transfer RNA (tRNA) are crucial components of the protein synthesis machinery. Small nuclear RNA (snoRNA) take part in RNA processing while microRNAs [5,6] has been shown to regulate the abundance of mRNA. Recent efforts in transcriptome analysis indicate that the population of functional RNA in the cell may be even larger. The sequencing of the human genome revealed a surprisingly small number of protein coding genes, around 30,000 [7,8,9], currently revised to 22,701 protein coding genes in humans and 24,118 in mice [10]. In contrast, the mammalian transcriptome was shown to consist of more than 100,000 transcripts [11]. This discrepancy can be partly be explained by alternative splicing, currently predicted to occur for 60-70% of all genes, generating both protein coding and non-coding transcripts [12,13]. However, non-coding RNAs are also transcribed directly from the genome and to a much larger extent than previously realized [8,11,14]. The predicted amount of transcribed non-coding genes in humans is to date 8509, whereas 4613 in mice [10].

The FANTOM consortium has made a substantial contribution in annotating the mouse transcriptome [15,16,17]. By computational prediction, combined with manual annotation of physical poly-adenylated cDNA clones, they identified a wide variety of non-protein coding RNA transcripts in addition to the protein-coding transcripts. The most recent report from the FANTOM consortium describes a total of 103,000 transcripts. Of these, 57,000 were functionally annotated as protein coding while the total number of distinct non-protein coding were determined to 34,000 [17]. The characterization of the mammalian genome is far from completed, thus the numbers given above should be regarded as approximate. Most likely there is much more yet to be discovered. With improved instruments and methods for sequence analysis, we may find out that the transcriptome is even more complex.

1.1.1 Natural antisense

Natural antisense transcripts (NATs) are a subset of the non-protein coding transcripts with sequence complementarities to coding sequences in the genome. These NATs can be transcribed from the antisense DNA strand (*cis*-NATs) or from separate loci (*trans*-NATs) and have been suggested to regulate protein-coding RNA, influencing the transcriptional output of the genome [18,19]. Transcription from the antisense DNA strand has been addressed in a genome-wide manner by several groups [14,20,21,22] and the proportion of sense-antisense pairs in the genome has been predicted to be as high as 22-25% of all transcriptional clusters [21,22]. Chen and colleagues hypothesized that co-expression of sense-antisense pairs would indicate a regulatory role and demonstrated that this occurs more frequently than expected by chance. They

further reported that antisense genes tended to be inversely expressed with their sense partners [23]. However, it has been suggested that standard reverse transcription can represent a source of natural antisense artifacts [24]. Genome wide screens based on cDNA data therefore should be followed up and the natural antisense transcripts investigated one by one.

Some complex loci have been characterized in more detail at a functional level. *Cis*-NATs have been suggested to regulate the expression of their target genes in several biological conditions, for example development [25], circadian clock function [26] and hypoxia [27]. *Cis*-NATs have also been shown to direct imprinting [28]. Although the mechanisms for regulation are still largely unknown, *cis*-NATs have been suggested to act through transcriptional interference, RNA masking, double-stranded RNA-dependent mechanisms or chromatin remodeling [19,29] (Fig 1).

The huge amount of information emerging from these genome-wide mapping projects has created a starting point for assigning functions to numerous, previously unknown protein coding and non-coding transcripts. This scientific discipline is called Functional genomics.

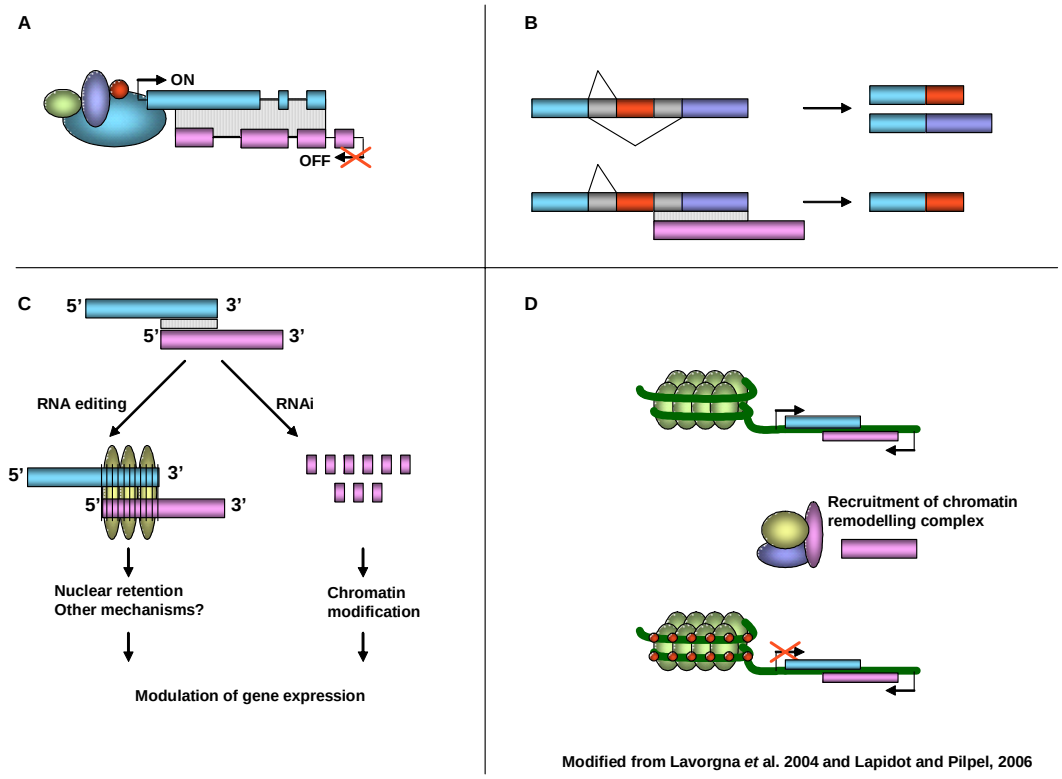


Figure 1. Mechanisms by which natural antisense transcripts may regulate gene expression (A-D). (A) Transcriptional interference. Simultaneous transcription from both strands may be inhibited by physical collision of transcription complexes (B) RNA masking, where the natural antisense masks a splice site and thereby promotes the abundance of one specific splice variant. (C) Double-stranded RNA-dependent mechanisms such as RNA editing and RNA interference require simultaneous presence of sense and antisense transcripts for duplex formation. These mechanisms may therefore explain the numerous observed co-expressed sense-antisense pairs. (D) Chromatin remodelling. Transcription of

non-coding antisense transcripts is involved in monoallelic gene expression, including genomic imprinting, X-inactivation and clonal expression of lymphocyte genes. In these processes, antisense transcripts have been suggested to silence the expression of nearby gene clusters by chromatin remodelling, most likely through the recruitment of histone modifying enzymes.

1.2 FUNCTIONAL GENOMICS

There are several techniques to study gene function. All have advantages and limitations. Given the recently presented information about the frequent existence of complex loci, caution needs to be undertaken when choosing a model system to study gene function. The models used for this thesis studies includes RNA interference in cell lines and primary cultures as well as gene expression and genomic association analysis of human *in vivo* models.

1.2.1 Cell culture models

Cell lines offer a relatively homogenous biological model in which gene expression can be easily manipulated. It is common to choose a cell type emerging from the tissue of interest. However, in many cases it is questionable how much of the original tissue features remain conserved in a cell culture [30]. As cell lines most commonly are derived from tumors, several changes in the phenotype have occurred already prior to cell culture isolation. Primary cultures could be considered as physiologically closer to the parent tissue than cell lines as they are isolated directly from a normal tissue. One limitation of this model is that the cultures can be heterogeneous. Furthermore, these cells cannot divide indefinitely, entering senescence after a certain number of cell divisions [31,32], rendering them less flexible a model. Thus, the appropriate cell model to use varies depending on the particular question.

1.2.2 RNA interference – mechanism

Andrew Fire and Craig Mello discovered that long double-stranded RNA molecules could be used for gene silencing in *C.Elegans* [33]. This phenomenon was termed RNA interference. Unfortunately, long double-stranded RNA induces interferon response as a viral defence in mammalian cells [34]. However, Tuschl and colleagues discovered that by utilizing shorter double-stranded RNAs (~20 nucleotides), called short interfering RNAs (siRNAs) the interferon response in mammalian cells was avoided and gene silencing was obtained [35]. The principle of RNA interference is illustrated in figure 2. RNA molecules complementary to particular mRNAs will bind the mRNA through the RNA induced silencing complex (RISC), leading to cleavage and degradation of the mRNA. The efficiency of a particular siRNA is determined by its sequence and structure, as well as details of the target mRNA. Schwartz and colleagues demonstrated that there is an asymmetry dependent bias for which of the siRNA strands that will get incorporated into the RISC. The strand with the less tightly base-paired 5'-end will usually end up in the RISC, whereas the other strand will be degraded [36]. The RISC mediated gene targeting mechanism is also utilized to regulate gene expression by endogenously expressed small non-coding hairpin RNA molecules, microRNAs, [37]. In yeast, siRNAs has also been shown to act through another

complex – RITS (RNA induced initiation of transcriptional gene silencing). In this case, the siRNAs are used as probes by the protein complex to initiate heterochromatin assembly where there is a sequence match [38]. The advantages of utilizing siRNA in mammalian cell models are that it allows for post-transcriptional silencing, and thus the intervention does not need to integrate into the genome. However, in the light of the increasing knowledge about a functional transcriptome, it is easy to imagine that these siRNAs may have other effects than just knocking down their target gene.

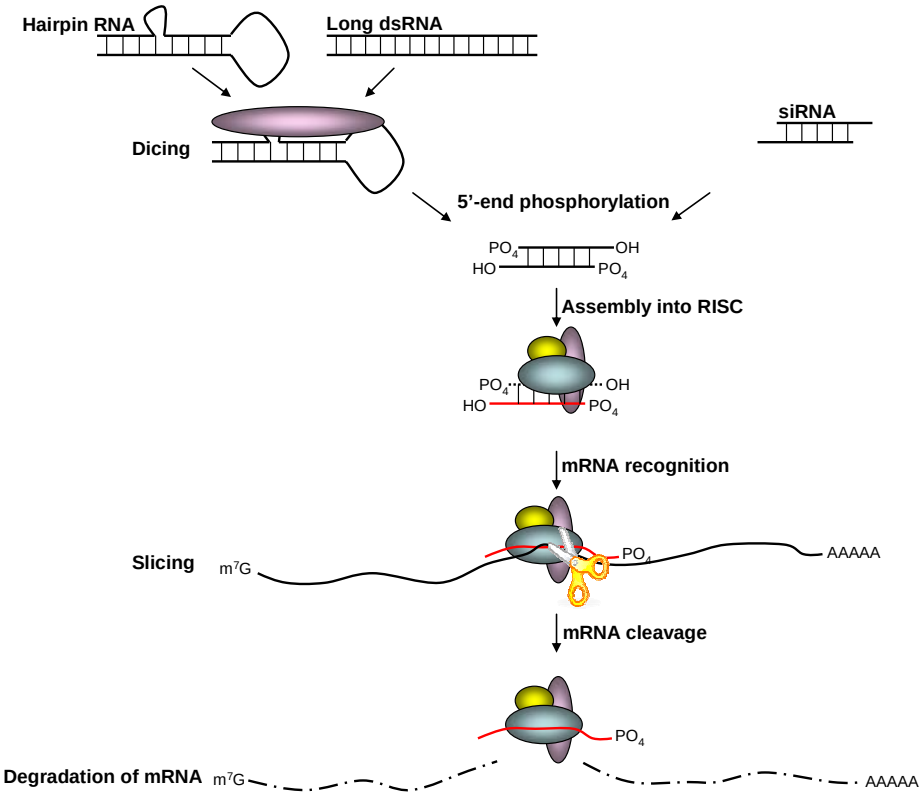


Figure 2. A summary of the RNA interference pathway. Long double-stranded RNA and hairpin RNA gets cleaved into ~20 nucleotide short interfering RNAs (siRNAs) by Dicer. Following 5' end phosphorylation, the siRNAs assembles RISC. The siRNA strand with the less tightly base paired 5' end remains in the RISC complex. Target mRNA is recognized, cleaved and degraded.

1.2.3 RNA interference – off-target effects

Algorithms for designing a siRNA sequence that provides efficient knockdown are relatively well established, whereas approaches for avoiding off-target effects remains to be optimized. There are two broad types of off-target definitions: motif-dependent toxic effects and silencing of genes other than the target genes. A specific motif, UGUGU in siRNA sequences, has been shown to induce interferon response *in vivo* in mice and *in vitro* in human peripheral blood mononuclear cells [39]. Another motif, UGGC, was identified to cause a reduction in cell survival when present in the antisense strand of the siRNA [40]. Expression profiling using microarray, revealed that in some cases, 11 bases match between siRNA and mRNA was enough to cause an

off-target effect. [41]. This was further investigated in another study, demonstrating that off-targeting is associated with the hexamer or heptamer seed region (positions 2-7 or 2-8) of the antisense strand of the siRNA perfectly matching with the 3'UTR of several target mRNAs [42]. This manner of silencing, similar to the gene expression regulation by the endogenously expressed miRNAs, was further confirmed in another study [43]. This is of major concern because single miRNAs can regulate large sets of genes.

A chemical modification of the guide (antisense) strand, a 2'-O-methyl ribosyl substitution at position 2, has been shown to reduced the silencing of 3'UTR targeted transcripts while the silencing of perfectly matched mRNA transcripts were unaffected [44]. Clearly, although having a huge potential, siRNAs as a tool to study gene function still needs to be optimized and current data should be interpreted with caution.

1.2.4 Human models

Comparing gene expression between healthy controls and people with a disease is a way of finding candidate genes that may contribute to a disease. However, only limited conclusions can be drawn from these kinds of analyses. For example, individual variations within the control group and disease group respectively, could mask putative candidate genes. In addition, the analysis cannot explain whether a change in gene expression is *causing* the disease or if it rather reflects a *secondary effect*. In particular, this can be an issue in microarray analyses where new hypotheses are formed based on the differential gene expression when testing >1000 of genes. One example of this kind of problem recently occurred within type 2 diabetes research [45]. A closer look at the subjects in this study reveals that they compared non-obese normal glucose tolerant people with a blend of non-obese and obese diabetics. Thus, the difference between their sample sets may be derived from the differences in obesity between the two groups rather than differences causing or being induced by diabetes. However, the way they compared sets of genes involved in the same pathways instead of comparing the genes one by one, was novel and elegant. Human *in vivo* intervention studies offer cleaner models, as material (i.e. blood samples and biopsies) can be collected before and after the intervention. The material collected before the intervention will serve as control for each individual and thus changes in the endogenous gene expression profile can be directly linked to the intervention excluding the background from genetic variations between individuals.

1.2.5 Gene expression analysis – Quantitative real-time PCR

Quantitative real-time PCR (qPCR) is a commonly used method to compare the relative gene expression between samples. The method utilizes fluorescent markers that are linked to the amplification so that each PCR cycle will result in an increase in fluorescence. This works by using a gene specific probe labeled in each end with fluorescent dyes – a reporter and a quencher (quenching the fluorescent spectra of the reporter). The probe binds to the DNA prior to the gene specific primers but is degraded by the polymerase during the amplification. The increased physical distance between the quencher and the reporter dye results in an increase in fluorescence,

detectable by the instrument. Another approach is to use SYBR Green instead of a probe. SYBR Green binds to all double-stranded DNA and the binding will increase its fluorescence. To assure specificity, a heating-dissociation step is added subsequently to the PCR amplification. By measuring the decrease in fluorescence (caused by dissociated PCR amplicons), it is possible to detect whether there is more than one temperature that will cause a decrease in fluorescence and thus whether there is more than one product.

QPCR allows for sensitive comparison of the relative abundance of one gene in different samples. However, the fluorescence levels do not reflect an absolute amount of mRNA copies in each sample, as different PCR amplicons may vary in strength of their fluorescent signals. The raw data obtained from qPCR are cycle threshold (Ct) values. These values reflect the PCR cycle where the signal from a specific sample reaches the threshold set for the particular gene. The same amount of RNA is utilized for all samples in the cDNA synthesis, but to adjust for differences in reverse transcription efficiency, the gene expression of an endogenous control is measured in the same samples. Commonly used controls includes 18S and house-keeping genes such as GAPDH and β-actin, but should be assessed to ensure absence of systematically change by the experimental set-up [46]. Ct values of the endogenous control that varies largely between the samples may indicate problems with RNA integrity or purity. This can be assessed by running the RNA on an agarose gel or by using a bioanalyser, which demands substantially lower amounts of RNA and also provides a RIN-score. The RIN is an RNA integrity number, representing a standardized RNA quality control based on a large collection of microcapillary electrophoretic RNA measurements. An algoritm describing RNA integrity was extracted using multiple RNA separation signal measurements and constructing regression models based on a Bayesian learning technique [47]. In some cases, there can be PCR inhibiting factors remaining from the RNA isolation. This issue can usually be solved by diluting the cDNA (and thus also diluting the inhibitors) or optimizing the RNA isolation.

1.2.6 Gene expression analysis – Microarray

Microarray analysis allows for a genome-wide comparison of differential gene expression between two conditions. Two main approaches for construction of microarrays exist: spotting of pre-synthesized materials and on-chip synthesis of DNA. Spotted materials are cDNA or oligos. The on-chip synthesis method was developed and first introduced by Affymetrix [48] and to date used by several other companies. In a pilot study included in this thesis, we used the *Pegasus* chip, a cDNA chip spotted with druggable gene targets to search for differentially expressed genes in human soleus muscle following inactivity (Paper III). Published microarray data are usually stored in publicly available databases, such as ArrayExpress and Gene Expression Omnibus (GEO), which allows for analyses across several data-sets. This is a new research tool and is a revolution in terms of gene expression profiling.

1.2.7 Single nucleotide polymorphisms

Single nucleotide polymorphisms (SNPs) are 1 nucleotide genetic variations in a population where the minor allele by definition has a frequency of at least 1% in the population. On average, these SNPs occur at about one in every 1,200 bases and are by far the most common type of genetic variation in an otherwise, between humans, rather identical genome. Very recently, the diploid human genome of Craig Venter was presented [49]. Comparison of this genome and the National Center for Biotechnology Information human reference assembly revealed more than 4.1 million DNA variants, encompassing 12.3 Mb. These variants included 3,213,401 SNPs. Groups of SNPs mapped within same chromosomal region and inherited together are known as haplotypes. A haplotype block represents a discrete chromosome region of high linkage disequilibrium (LD), low haplotype diversity and low recombination. To compare these haplotype blocks between individuals, only a small number of representative characteristic "tagging" SNPs is needed [50,51].

The International HapMap Project is a consortium aiming to identifying common haplotypes in four populations from different parts of the world [52,53]. The consortium is characterizing patterns of LD in the human genome and a pilot study showed 'haplotype blocks' in 51 regions scattered throughout the genome. In parallel to this mapping, instruments and software are developed to allow for genome wide association analysis of haplotypes in complex diseases [54,55] to identify candidate genes for potential drug targets. Genome-wide haplotype association studies have for example been made for Parkinson's disease [56], type 2 diabetes [57,58], ischemic stroke [59] and amyotrophic lateral sclerosis [60]. If a particular haplotype occurs more frequently in affected individuals compared with controls, a gene influencing the disease may be located within or near that haplotype. Similar to microarray data, the results are made publicly available for further data mining by other researchers.

1.2.8 Allelic discrimination

The genotypes for a specific SNP of an individual can be determined by utilizing a real-time PCR based assay. This assay includes one set of primers and two probes, each labeled with different fluorescent markers and each recognizing a different genotype. The correct probe for the genotype will bind and get cleaved and the fluorescent signal will increase as described above for the qPCR assay. The melting temperature of the incorrect probe will be lower than the correct probe, due to the one base mismatch. Hence, the incorrect probe will be destabilized and not result in any signal. Thus, if only one probe is producing a signal for a particular sample, the individual is a homozygote for this SNP. If both probes results in a signal, the individual is a heterozygote.

1.3 DISCOVERY OF THE PINK1 GENE

An expressed sequence tag (EST) was induced in cancer cells over-expressing PTEN [61] and was cloned by 5' rapid amplification of cDNA ends (RACE) and named PTEN Induced Putative Kinase 1 (PINK1) [62]. The PINK1 transcript consists of 2600

bp and encodes a 581 amino acid protein with a predicted molecular mass of 62.8 kD [62]. PINK1 gene expression decreased in tumors with mutated PTEN [62]. PINK1 was later identified as a "novel gene" in another study, and was further characterized and given another name, BRPK [63]. In 2004, mutations in PINK1 were linked to hereditary early-onset Parkinson's disease in a large Italian family [64]. This study was a break-through for the identification of the *PARK6* locus. Other studies have later confirmed and expanded this link to other populations and mutations in PINK1 [65,66]. In addition, a parkinsonian patient had a PINK1 whole gene deletion and a 23 bp deletion causing a change in splice sites and several new transcripts was identified in her family [67].

1.4 BIOCHEMISTRY OF PINK1

At the beginning of the work for this thesis nothing was known about the characteristics or function of PINK1 protein. Since *PINK1* was established as the candidate gene for the *PARK6* locus in 2004, several studies have provided a number of new biochemical observations around PINK1 function. PINK1 is a serine threonine kinase that has been shown to be heavily ubiquitinated [68] and exhibit autophosphorylation capacity, assessed using GST-PINK1 fusion proteins [63,69]. Mutations in PINK1 affected the autophosphorylation activity [69]. One of the mutations, W437X, increased the autophosphorylation activity [69], while another, L347P, decreased activity by destabilizing the protein [70]. The increase in phosphorylation activity could be explained by the C-terminal part of the protein having a suppressing role of the phosphorylation activity. This assumption was further supported by an increase in autophosphorylation capacity when using a fusion protein of GST and the PINK1 kinase domain only, compared to the GST-PINK1 whole protein [69]. Another group obtained opposite results, demonstrating that a recombinant PINK1 protein containing both the kinase domain and the C-terminal tail had a higher phosphorylation activity than a recombinant PINK1 protein with the kinase domain only [71]. In addition, C-terminal truncation mutations decreased the phosphorylating activity of PINK1 [71].

The PINK1 protein includes an N-terminal mitochondrial signal sequence. Mitochondrial localization has been experimentally confirmed by subcellular fractionation followed by western blot [69,72] or immunocyto- or histochemistry [64,70,73]. The signal sequence is cleaved off upon mitochondrial import, but the mature protein has also been observed in the cytosol [70]. In the mitochondria, PINK1 localizes to the mitochondrial inner membrane (as shown with sodium carbonate treatment of the mitochondrial fraction followed by ultracentrifugation and western blot) [72]. PINK1 is ubiquitously expressed in most tissues with a higher abundance in mitochondrial rich tissues such as heart, skeletal muscle and testis [62]. No differences between brain regions were observed when comparing expression of the mouse and rat homologoue to PINK1 in cortex, striatum, thalamus, brainstem and cerebellum [74]. PINK1 localized to Lewy bodies in post-mortem sporadic Parkinson's disease patients' brains [72].

1.5 FUNCTIONAL STUDIES OF PINK1

1.5.1 PINK1 and mitochondrial integrity in *Drosophila*

In *Drosophila*, PINK1 homolog mutants demonstrated degenerated mitochondria in the oxidative flight muscle, together with abnormal wing posture [75,76,77]. There was also a higher frequency of enlarged mitochondria [76], fragmentation of mitochondrial cristae [75] and mitochondrial defects in spermatids [75]. In some, but not all of the studies, there was a modest decrease in the amount of dopaminergic [76,77,78] and ommatidia neurons [78]. In addition, a reduction in cranial dopamine content was observed in the aging (18-30 days old) fly, whereas there was no difference in the 2-3 days old fly [76,77]. This indicates that loss of PINK1 causes an accumulative damage to cells.

1.5.2 PINK1 in oxidative stress and apoptosis

PINK1 was identified as a survival kinase in a RNA interference screen in HeLa cells of the 650 known and putative kinases [79]. Stimulation with MPP+, the neurotoxic byproduct of MPTP metabolic oxidation, resulted in a decreased viability of stably transfected L399P mutant PINK1 SH-SY5Y neuroblastoma cells, compared to wild type cells [68]. In line with this, it was recently demonstrated that over expression of several different mutated PINK1 constructs failed to protect HEK293 cells from apoptosis induced by proteasome inhibitor MG132 [80]. This was in comparison to the over expression of wild type PINK1. Further, stable expression of PINK1 in the human neuroblastoma cell line SH-SY5Y, resulted in a decrease in the number of both basal and staurosporine-stimulated TUNEL-positive SH-SY5Y cells and a reduction in cytochrome c translocation. PINK1 over expression also reduced staurosporine-induced caspase-9, -3, and -7 and PARP activation. Transiently over expressing PINK1 decreased basal caspase-3 activity; an effect abolished when over expressing PINK1 constructs that included the PINK1 mutations E240K, L489P and K219M [81].

Human neuroblastoma SH-SY5Y cells treated with MPP+, were sensitized for apoptosis induction by PINK1 knockdown [82]. Importantly, this study only utilized one siRNA sequence and off-target effects can thus not be excluded. PINK1 was further claimed to protect against oxidative stress by phosphorylating and inactivating the mitochondrial chaperone TRAP1 [83]. Using the rat pheocytochroma cell line PC12 as a model, the authors found that PINK1 phosphorylated TRAP1 in response to H_2O_2 stimulation. This event seemed to protect the cells from apoptosis in terms of decreased cytochrome c release. PINK1 has also recently been shown to phosphorylate HtrpA2, a susceptibility factor in Parkinson's disease (*PARK13* locus). PINK1 phosphorylation of HtrpA2 was increased on activation of p38 stress-sensing pathway [84].

Taken together the studies above imply that PINK1 in some way promotes cell survival. However, the majority of this data is generated from cell culture over expression models or, in some cases, *in vitro* phosphorylation assays. Due to supra-physiological expression levels, the models become very rough and it is difficult to judge the accuracy. In reality, the effect of PINK1 ablation is most likely more subtle.

1.5.3 Loss of Pink1 in mice reduces evoked dopamine release

Following the mitochondria linked phenotypes found in Drosophila, and given the link to Parkinson's disease in humans, ablation of Pink1 could be expected to cause death of dopaminergic neurons in mouse. However, silencing of Pink1 by conditional RNA interference had no effect on dopaminergic neuron survival, the levels of neurotransmitter dopamine or its metabolites DOPAC and HV, and neither on motor activity of 6 months old mice [85]. The same conclusions were reached in a study of a Pink1 knockout mouse. This may be due to alternative pathways and compensatory mechanisms existing in higher mammals, as well as variations in metabolic control between the species. However, the Pink1 knockout mouse did have a reduction in the evoked release of dopamine and catecholamine following stimulation. This suggests a role of PINK1 in dopamine release in the nigrostriatal pathway [86].

2 AIMS OF THE THESIS

Mitochondrial dysfunction has a complex relationship with physical inactivity, genetic inherited aerobic capacity and degenerative diseases such as Diabetes and Parkinson's disease. Mitochondrial capacity has been linked to insulin resistance [87] and PGC-1α, a master regulator of mitochondrial gene expression, has been suggested to be a candidate gene for type 2 diabetes [88]. In a search for other candidate genes in a model for mitochondrial dysfunction (characterized in paper I), we identified PINK1, a mitochondrial gene previously linked to Parkinson's disease. Thus, the aims in this thesis were to:

1. Study the genomic locus of PINK1 (Paper II)

2. Characterize the role of PINK1 in metabolic and mitochondrial biogenesis models (Paper III and IV)

3. Investigate whether PINK1 could represent a candidate gene for type 2 diabetes (Paper III and IV).

All human studies were performed according to the declaration of Helsinki and received appropriate local ethics approval.

3 RESULTS

3.1 FINDING PINK1

Physical inactivity is known to contribute to a decline in metabolic fitness, most likely generated by mitochondrial dysfunction. We thus hypothesized that a human *in vivo* model for mitochondrial dysfunction would be able to identify novel candidate genes involved in metabolic disease. A microarray analysis of mRNA levels in skeletal muscle from individuals before and after five weeks of one leg unloading [89], revealed differentially expressed genes (Paper I and III). We performed an expression profiling analysis of this inactivity model, using qPCR, and demonstrated a coordinated down-regulation of mitochondrial genes including PGC-1α but not PGC-1β (Paper I). PGC-1α is a regulator of multiple aspects of mitochondrial gene expression and has been suggested to play a central role in metabolic homeostasis [90]. PINK1 was one of the down-regulated genes in our microarray study, while the Wood research team published a link between mutations in PINK1 and Parkinson's [64], a disease associated with mitochondrial dysfunction. Thus, already implied to produce a mitochondria-linked disease phenotype in humans, and also down-regulated in our mitochondrial dysfunction model, we selected PINK1 for further functional genomics studies.

3.2 CHARACTERIZATION OF THE *PINK1* LOCUS

When we examined the annotation of the *PINK1* gene we made a number of observations. Firstly, there were two transcripts annotated as potentially protein coding. In addition, there appeared to be a *cis*-transcribed natural antisense (*cis*-NAT) that did not appear to code for any protein. Interestingly, this *cis*-NAT was not conserved in the murine NAT databases [14,17] further supporting that it should be a non-coding RNA. We utilized Northern blot and observed bands corresponding to the sizes of PINK1 (2.6 kb), its shorter splice variant that we named svPINK1 (1.6 kb) and the *cis*-NAT, that we named naPINK1 (4.4 kb) (Paper II). SvPINK1 was expressed at a low level, but became cleaner when polyadenylated RNA was isolated and loaded on the gel. This suggest although not prove that svPINK1 is a fully processed transcript and not an artifact. To further verify the existence of svPINK1 we utilized Rapid-amplification of cDNA ends (RACE), and cloned and sequenced svPINK1.

The 5'-end of naPINK1 was annotated to overlap with the 3'UTR of the *Dolichyl-diphosphooligosaccharide-protein glycosyltransferase (DDOST)* gene. *DDOST* encodes the scavenger receptor AGE-R1 enhancing removal of advanced glycation end-products (AGE). AGE accumulate in ageing, diabetes and neurodegeneration [91,92] and the receptor for AGE (RAGE) promotes oxidative-stress dependent NF-κB activation and inflammatory gene expression. AGE-R1 has been shown to negatively regulate RAGE [93] and thus counteract AGE at two levels. Interestingly, type 1 diabetes patients with diabetic complications failed to up-regulate AGE-R1 when challenged with AGE, indicating de-regulation of AGE-R1 in diabetes [94]. Given our finding that *PINK1* was a novel candidate inactivity gene, a condition that strongly

relates to diabetes, and as the adjacent *DDOST* also relates to diabetes, these initial observations were to direct our further investigations of the *PINK1* locus (Fig 3) in metabolic disease.

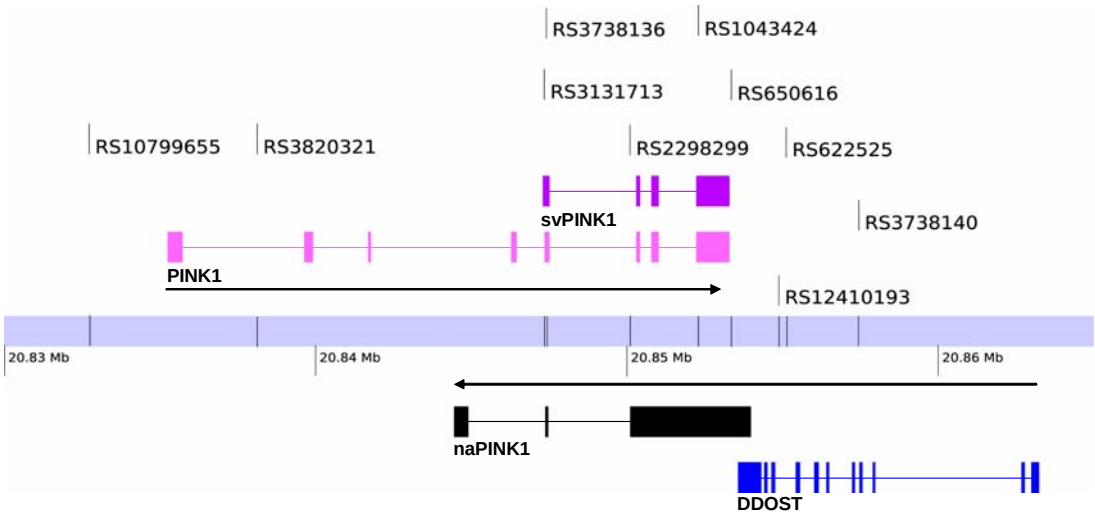


Figure 3. The *PINK1* locus includes full-length PINK1 and svPINK1 (a transcript predicted to be protein coding). On the opposite DNA strand, a natural antisense, naPINK1 is transcribed, partly complementary to PINK1 and svPINK1 and over-lapping with the 3'UTR of the neighboring gene DDOST. The 10 tagging SNPs analyzed in Paper IV are annotated.

3.3 THE *PINK1* LOCUS IS REGULATED BY A *CIS*-TRANSCRIBED NAT

We utilized two different human *in vivo* models for physiological changes of mitochondrial content/activity. Using qPCR, we found an opposing expression pattern in skeletal muscle of the full-length PINK1 as compared to svPINK1 and naPINK1 (Paper II and III). Whereas PINK1 was down-regulated following 5 weeks of inactivity and up-regulated following 6 weeks of endurance training, svPINK1 and naPINK1 demonstrated the opposite net expression levels. Linear regression models demonstrated a strong concordant correlation between svPINK1 and naPINK1 expression in these two models. In contrast, the expression of PINK1 did not correlate with either svPINK1 or naPINK1. We observed the same relationships in a cell model, where siRNA knockdown of naPINK1 resulted in a down regulation of svPINK1, indicating that naPINK1 was somehow promoting the abundance of svPINK1. PINK1 was not significantly changed by this manipulation. The abundance of DDOST was not altered by inactivity, activity or knockdown of naPINK. This demonstrates that under physiological conditions these genes do not interact.

3.4 PINK1 HAS A ROLE IN OXIDATIVE METABOLISM IN HUMANS

We extended our investigation of the *PINK1* locus to a sample-set of non-obese and obese type 2 diabetics and normal glucose tolerant controls (Paper III). Here we observed a lower expression in skeletal muscle of PINK1, naPINK1 and DDOST in type 2 diabetes compared to normal glucose tolerant controls. There was also a discordant non-linear correlation between fasting glucose levels and HbA1c and PINK1 RNA levels in skeletal muscle. In contrast, PGC-1α did not correlate with HbA1c or fasting glucose (Paper III).

To investigate whether genetic variation at the *PINK*1 locus could contribute to the progression of diabetes, we used single nucleotide polymorphism (SNP) analysis. We mapped two different cohorts for 10 non-coding and coding tagging SNPs dispersed over the *PINK1* gene (Fig 3). PINK1 mRNA levels associated with 6 of the 10 SNPs (located both in the intergenic and the coding regions) in the smaller cohort (208 subjects) (Paper IV). These SNPs further associated to biomarkers of oxidative metabolism in a larger cohort (1701 subjects). In other words, an allele linked to a higher PINK1 gene expression had in general lower levels of plasma glucose and non-esterified fatty acids (NEFA). This indicates that low levels of PINK1 could contribute to the progression of type 2 diabetes. An observation confirmed by examination of the publicly available data from the Diabetes Genetics Initiative cohort (cases: n=1464, controls: n=1467) [58]. In addition, our data indicate that the *PINK1* locus deregulation associates with obesity.

To further examine the role of PINK1 in metabolism, we knocked down PINK1 in cell models using siRNA. In cultures of brown pre-adipocytes, the lipid transporter FABP4 was subsequently down regulated, while basal glucose uptake was impaired in neuroblastoma cell lines (Paper II and IV). Furthermore, PINK1 increased during mitochondrial biogenesis in differentiating brown adipocytes, yet ablation of PINK1 had no effect on the expression of mitochondria related genes, such as mtnd5, citrate synthase, tfam, ucp1 or parkin. Neither did it have any apparent importance for cell viability (Paper III and IV). Taken together, our experiments support our human *in vivo* findings that PINK1 may play an important role in oxidative metabolism. In contrast, PINK1 does not seem to have a role in mammalian mitochondrial biogenesis.

4 DISCUSSION

4.1 A NATURAL ANTISENSE REGULATES THE *PINK1* LOCUS

The mechanisms for the actions of NATs are largely unclear. We demonstrate that a NAT regulates the *PINK1* locus (Paper II and III), and we suggest that it functions by promoting alternative splicing of PINK1. NATs have previously been suggested to control alternative splicing. *Cis*-transcribed natural antisense RNA are predicted to be more commonly expressed at loci where there are also splice variants [18]. RevErb is an example of a natural antisense thought to be involved in splicing. RevErb has been suggested to regulate the ratio of ErbA isoforms by masking interaction sites for splicing regulatory factors. ErbA encodes the thyroid receptor and the mRNA is spliced into ErbAα1 and ErbAα2 which are translated into proteins with opposite functions [95,96].

In our study, the shorter PINK1 splice variant, svPINK1, has a concordant expression with naPINK1 during physiological conditions in humans *in vivo* or during manipulation of naPINK1 expression in cell culture systems. This indicates that rather than masking a splice site, naPINK1 may promote the splicing of PINK1 pre-mRNA into svPINK1 (Fig 4), a mechanism for natural antisense action not previously elucidated. One suggested but not yet proven function of natural antisense is to guide the RNA editing machinery [29]. RNA editing and splicing have been shown to act coordinately and editing was shown to create alternative splice sites [97,98,99]. Thus, one possibility is that naPINK1 act as a guiding strand to accomplish RNA editing and promote alternative splicing of PINK1, resulting in a higher abundance of svPINK1. Another possibility is that naPINK1 is not involved in the splicing but rather physically stabilizes svPINK1. The exact mechanism of how naPINK1 regulates svPINK1 in a concordant manner remains to be investigated.

Although PINK1 had an opposing expression pattern to naPINK1 and svPINK1, we did not see a negative correlation between PINK1 and naPINK1 in the human *in vivo* models. Neither was PINK1 significantly up regulated when naPINK1 was knocked down in cell culture using siRNA. This could indicate that the expression of PINK1 is regulated at several levels and that the natural antisense regulation only represents a fine-tuning mechanism for alternative splicing. Alternatively, it could reflect the difficulties of detecting relative changes in the higher abundant PINK1 versus the lower abundant svPINK1.

Over expression of aHIF, a NAT suggested to destabilize HIF1α, a splice variant of HIF1, has previously been linked to disease [100]. We therefore preliminary hypothesized that naPINK1 and svPINK1 would be up regulated in metabolic disease. However, the dynamic expression pattern of the *PINK1* locus observed during physiological inactivity or activity was absent in obesity and type 2 diabetes samples (Paper III). Indeed, *all* transcripts from the locus were suppressed. As both DNA strands are affected, we speculate that this reflects epigenetic changes suppressing the *PINK1* locus, generated by chronically impaired metabolic homeostasis. It may be that

DDOST, encoding AGER1, represents the primarily target for gene silencing due to deregulation induced by chronic exposure of AGEs.

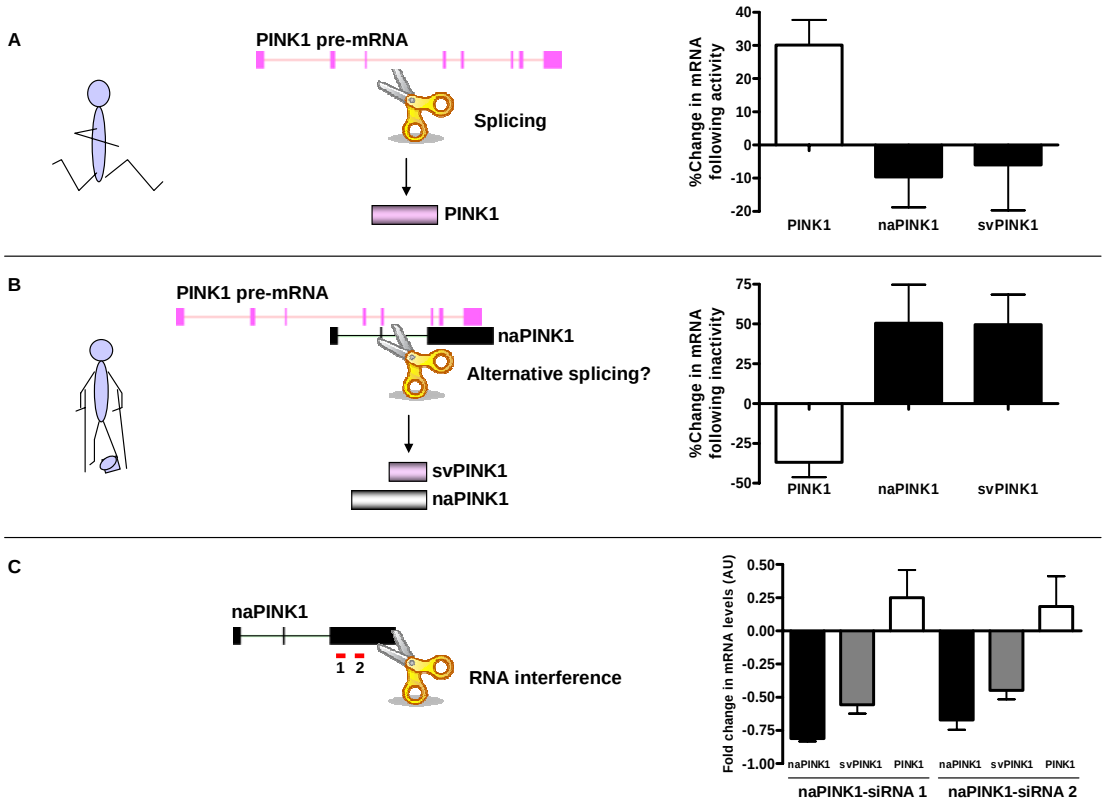


Figure 4. The *PINK1* locus is regulated by a natural antisense (naPINK1) following physical activity and inactivity. A possible mechanism for this is that naPINK1 in some way directs alternative splicing of the PINK1 pre-mRNA (A-C). (A-B) Differences in gene expression before and after activity/inactivity were tested by paired t-tests comparing to the expression of the particular gene before the intervention. (A) PINK1 is increased ($P<0.001$) whereas naPINK1 ($P<0.05$) and svPINK1 ($P<0.02$) are decreased following physical activity (n=23-24). The graph is modified from Paper II. (B) PINK1 is decreased ($P<0.03$) whereas naPINK1 and svPINK1 are increased ($P<0.08$ and $P<0.05$, respectively) following physical inactivity, possibly by primarily increased levels of naPINK1, directing alternative splicing of the PINK1 pre-mRNA (n=6). The graph is modified from Paper III. (C) Knockdown of naPINK1 in SH-SY5Y cells using two siRNAs (red) targeting different sites on naPINK1, results in a significant down regulation of svPINK1 (n=5-6). 18S adjusted mRNA levels of siRNA transfected cells were compared to the mRNA levels in control siRNA transfected cells. The graph is modified from Paper II. Statistical significance was assessed by Anova one-way analysis followed by Tukey post-hoc tests generating P-values. For naPINK1 and svPINK1 expression and both siRNAs the *P*-value was <0.001, while the expression of PINK1 was not significantly changed compared to the control siRNA (not shown).

4.2 PINK1 HAS A ROLE IN MITOCHONDRIAL OXIDATIVE METABOLISM

During the time of our studies other research groups, mainly focusing on untangling the link between PINK1 and Parkinson's disease, published several reports on functional studies of PINK1. In these studies, PINK1 is suggested to protect against apoptosis

induced by oxidative stress. Numerous studies have assessed the effects of challenging mutated PINK1 over expressing cells with apoptotic stimuli [68,80,81]. One such study identified TRAP1, a mitochondrial chaperone as a substrate for PINK1 [83]. Another study suggested that PINK1 was phosphorylating the mitochondrial protease HtrA2 [84]. Both these studies describe an involvement in stress-sensing pathways. This could be interpreted as a PINK1 inhibition of stress-induced apoptotic response. However, it could also reflect that a cell with impaired energy homeostasis (i.e. with a defect PINK1) becomes sensitized to exogenous stress stimulation. It is plausible that PINK1 can influence cell viability at multiple levels. For example one study in *Drosophila* suggested that PINK1 protected cells from oxidative stress. Here, *Pink1* mutants were rescued from degeneration of ommiditial and dopaminergic neurons by treatment with the antioxidants SOD and vitamin E. [78].

Our results provide a novel perspective on PINK1. The human *in vivo* intervention and epidemiological studies together with the cell culture RNA interference studies imply a role for PINK1 in the regulation of fatty acid and glucose metabolism. Intracellular accumulation of fatty acid metabolites can contribute to the production of ROS. Thus, we hypothesize that PINK1 dysfunction may increase the production of ROS due to impaired metabolic homeostasis, rather than by directly protecting cells from ROS. This idea is supported by the changes in peroxidation and glutathione metabolism in cultured fibroblasts obtained from skin samples of hetero- or homozygotes for the PINK1 mutation G309D compared to healthy controls [73]. Levels of malondialdehyde (MDA), an indicator of lipid oxidation was significantly higher in the homozygotes, compared to heterozygotes and controls. In addition, MnSOD protein was significantly higher expressed in the homozygotes compared to controls and both glutathione reductase and glutathione-S-transferase activities were higher in fibroblasts from homozygotes compared to controls. This indicates that PINK1 ablation leads to an increased lipid oxidation and production of reactive oxygen species with a subsequent increase in antioxidant activity.

4.3 PINK1 AND PARKIN MAY ACT THROUGH PARALLEL PATHWAYS

Several *Drosophila* reports have suggested that Parkin, also associated with Parkinson's, acts downstream to Pink1 in a common pathway. The basis for this assumption is that the *Pink1*-mutant *Drosophila* demonstrate a similar mitochondrial degenerative phenotype to the *Parkin*-mutant *Drosophila* and transgene flies over-expressing human Parkin was rescued from the *Pink*- mutant phenotype in muscle [75,76,77], dopaminergic [76,77] and spermatid [75]. However, there is no description on whether endogenous levels of Parkin are affected by Pink1 ablation in *Drosophila*. Another possible scenario is that Parkin is acting through a parallel pathway and that the high levels obtained by exogenous over expression are able to compensate for the loss of Pink1. Parkin has been demonstrated to enhance mitochondrial biogenesis in mammalian cell culture [101] and thus we assessed the role of PINK1 in mitochondrial biogenesis to evaluate whether there is a linear link between PINK1 and Parkin with regard to mitochondrial status. We demonstrated that both PINK1 and Parkin gene expression were increased in our model. However, siRNA knockdown of Pink1 in primary cultures of the mitochondrial rich murine brown adipocytes did not affect

Parkin gene or protein expression (Paper III and IV). This suggests that in mammals, PINK1 and Parkin act through parallel pathways. This is despite of both playing important roles in mitochondrial metabolism and Parkinson's disease [102]. One can argue that if Parkin was acting through a parallel pathway it should have been up regulated in our cell model, to compensate for the loss of PINK1. This may happen when the cells are challenged with any mitochondrial stress agent, but remains to be investigated.

4.4 IS PINK1 A CANDIDATE GENE FOR MITOCHONDRIAL DISEASE?

PINK1 clearly has an important role in cell function and especially in tissues with high-energy turnover. Diabetes and Parkinson's are complex diseases more common in older people and both are related to mitochondrial dysfunction and oxidative stress. This implies pathophysiological mechanisms that include accumulating damage and suggests that diabetes and Parkinson's disease should demonstrate some degree of epidemiological association. Indeed, type 2 diabetes appear as risk factors for Parkinsons's disease [103]. In addition, up to 80% of Parkinson's patients have been suggested to suffer from glucose intolerance [104]. Reactive oxygen species (ROS) have been suggested to cause an accumulating damage to cells. The highest production of endogenous ROS occurs in the mitochondria as a by-product of energy production. Previous functional genomics studies focused on explaining the link between PINK1 and Parkinson's disease. The present investigation extends this perspective to include other conditions with impaired mitochondrial function.

Collectively, our data implies that a higher expression of PINK1 results in a higher metabolic fitness. It is hard to imagine that this would occur only by neutralising ROS and preventing the cell from entering apoptosis [83,84]. Rather, a regulatory role in energy utilization or uptake would be able to explain the PINK1 link between Parkinson's [64], the subtle phenotype in knockout/knockdown mice [85,86] and our data on metabolic fitness and diabetes in humans and in cell models (Paper II, III and IV). On one hand, our data indicates that mutations in PINK1 could be a risk factor not only for developing Parkinson's disease, but also for type 2 diabetes. On the other hand, our results emphasize and provide molecular support for the action of exercise to inhibit disease and promote health.

5 ACKNOWLEDGEMENTS

A number of people have helped, inspired and encouraged me during the progression of this thesis. In particular I would like to thank:

My supervisors:
Jamie Timmons, Thank you for everything you taught me about science. For a demanding supervision with constant questioning, but also for inspiring discussions, continuous support and for all the fruitful collaborations you created. Thanks for spotting some potential in me and spending so much time and energy to develop it. I am truly grateful for all the things I learned.
Claes Wahlestedt, for creating a functional genomics-bioinformatics research center with an outstanding potential. Thank you for letting me to be a part of such inspiring environment and for encouraging and generous support whenever I asked for it.

Friends, collaborators and colleagues at Karolinska Institutet
Liam Good, for general support and for always being ready to answer or discuss a question in a very precise way. Special thanks for your helpful comments on this thesis. **Per Tesch,** for giving me the opportunity to study gene expression in your human *in vivo* models. **Ola Larsson,** for everything you taught me when we were working together, and for your continuous encouragement later on. **Katarina Fredriksson,** it started with Trizol and it grew into friendship, thanks for all discussions about science and life. **Elin Enervald,** for sharing my passion for cell culture, for midnight caspase assays, very ambitious meze and for all our chats about everything. Also special thanks for your help with my thesis application. **Anna Birgersdotter,** for inspiring discussions and support and for all the exploding laughs in front of the computer when chatting with you. **Therese Andersson** and **Kairi Tammoja** thanks for your support and for all the fun we had! Special thanks to Therese for all your help during the thesis process. **Timo Lassman,** for a great collaboration and for creating the very useful *PINK1* locus figure. **Joakim Elmén,** you were a support at KI, thanks for also helping me adapt to my new life in Copenhagen and for your comments on the RNAi part of this thesis. **Alistair Chalk,** for the siRNA-tool, for all the parties and for useful comments on this thesis. **Håkan Thonberg,** for answering all my questions about molecular biology, wine and the geography of Söder. **Joel Zupicich,** for your help with cloning and for nice chats. **Robbie Brogan,** for being a brilliant summer student, thanks for the help. **Cecilia Dahlgren,** for all the fun we had as room mates and for help with computer issues numerous of times. **Alexandra Trifunovski**, for collaboration and friendship. **Kristian Wennmalm,** for making me laugh until I almost lost my breath. **Fredrik S, Ruben, Sussie, Jakob**, **Shan, Omid, Abbas, Mohammad, Geert, Carolina Å, Ola H, Boris, Olav R, Gerry, Bruce, Emily, Shane, Marcela, Eva B, Jessica N, Abhiman, Vivan, Raju, David F, Mia B, Erik A, Hagit, Ellen, Elizabeth, Rikard D, Anna H, Daniel L, Margareta F, Elsebrit, Weilin, Peter S, Jona, Iréne, Zdravko, Brittis, Gitt, Elisabeth, Bent** and anyone I may have forgotten, thank you for all those years!

Stockholm University
Barbara Cannon and Jan Nedergaard, thank you for collaboration and for adopting me for a few months and allowing me to be a part of your group – I learned a lot from you during my time at Stockholm University! **Natasa Petrovic,** for all your kind help and brilliant ideas. **Tomas Waldén** for our collaboration and for all the fun we had. Also many thanks to the rest of the zoofys family for making me feel so welcome and almost like I was one of you.

Copenhagen University
Bente Klarlund Pedersen, thank you for exiting collaborations and for giving me the opportunity to work with such valuable human material. For inviting me to work in your group and for being very understanding and letting me pursue this work in parallel to my present project. **Anders Rinnov Nielsen**, for great collaborations and for collecting "the biggest diabetes cohort in the world".

Umeå unversitet/Cambridge
Paul Franks, thank you for an interesting collaboration and for introducing me to genotyping.

My beloved family
My sister **Susanne**, for always listening, always supporting and for all your advices. My parents **Monika and Håkan**, for believing in me, for always being ready to help me with anything and for showing such interest in my research. My grandmother **Majken,** for all your care.

Søren, for your solid support and encouragement, for telling me to "lugna ner" and for making me smile every day.

6 REFERENCES

1. Szathmáry E (1993) Coding coenzyme handles: a hypothesis for the origin of the genetic code. Proceedings of the National Academy of Sciences of the United States of America 90: 9916-9920.
2. Szathmáry E (1999) The origin of the genetic code: amino acids as cofactors in an RNA world. Trends in genetics : TIG 15: 223-229.
3. Ma W, Yu C (2006) Intramolecular RNA replicase: possibly the first self-replicating molecule in the RNA world. Origins of life and evolution of the biosphere : the journal of the International Society for the Study of the Origin of Life 36: 413-420.
4. Ma W, Yu C, Zhang W, Hu J (2007) Nucleotide synthetase ribozymes may have emerged first in the RNA world. RNA.
5. Lee R, Ambros V (2001) An extensive class of small RNAs in Caenorhabditis elegans. Science (New York, NY) 294: 862-864.
6. Lau N, Lim L, Weinstein E, Bartel D (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science (New York, NY) 294: 858-862.
7. Venter J, Adams M, Myers E, Li P, Mural R, et al. (2001) The sequence of the human genome. Science (New York, NY) 291: 1304-1351.
8. Lander E, Linton L, Birren B, Nusbaum C, Zody M, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860-921.
9. Waterston R, Lindblad-Toh K, Birney E, Rogers J, Abril J, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520-562.
10. Hubbard T, Aken B, Beal K, Ballester B, Caccamo M, et al. (2007) Ensembl 2007. Nucleic Acids Research.
11. Carninci P, Kasukawa T, Katayama S, Gough J, Frith M, et al. (2005) The Transcriptional Landscape of the Mammalian Genome. Science 309: 1559-1563.
12. Johnson J, Castle J, Garrett-Engele P, Kan Z, Loerch P, et al. (2003) Genome-Wide Survey of Human Alternative Pre-mRNA Splicing with Exon Junction Microarrays. Science 302: 2141-2144.
13. Sharov A, Dudekula D, Ko M (2005) Genome-wide assembly and analysis of alternative transcripts in mouse. Genome Research 15: 748-754.
14. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, et al. (2005) Antisense transcription in the mammalian transcriptome. Science 309: 1564-1566.
15. Kawai J, Shinagawa A, Shibata K, Yoshino M, Itoh M, et al. (2001) Functional annotation of a full-length mouse cDNA collection. Nature 409: 685-690.
16. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, et al. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420: 563-573.
17. Maeda N, Kasukawa T, Oyama R, Gough J, Frith M, et al. (2006) Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. PLoS genetics 2: 62-62.
18. Li Y-Y, Qin L, Guo Z-M, Liu L, Xu H, et al. (2006) In silico discovery of human natural antisense transcripts. BMC Bioinformatics 7: 18-18.
19. Lapidot M, Pilpel Y (2006) Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. EMBO reports 7: 1216-1222.
20. Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O, et al. (2003) Widespread occurrence of antisense transcription in the human genome. Nat Biotechnol 21: 379-386.

21. Chen J, Sun M, Kent W, Huang X, Xie H, et al. (2004) Over 20% of human transcripts might form sense-antisense pairs. Nucleic acids research 32: 4812-4820.
22. Engström P, Suzuki H, Ninomiya N, Akalin A, Sessa L, et al. (2006) Complex Loci in Human and Mouse Genomes. PLoS Genetics 2: 47-47.
23. Chen J, Sun M, Hurst L, Carmichael G, Rowley J (2005) Genome-wide analysis of coordinate expression and evolution of human cis-encoded sense-antisense transcripts. Trends in genetics : TIG 21: 326-329.
24. Perocchi F, Xu Z, Clauder-Münster S, Steinmetz L (2007) Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. Nucleic Acids Research.
25. Coudert A, Pibouin L, Vi-Fane B, Thomas B, MacDougall M, et al. (2005) Expression and regulation of the Msx1 natural antisense transcript during development. Nucleic acids research 33: 5208-5218.
26. Kramer C, Loros J, Dunlap J, Crosthwaite S (2003) Role for antisense RNA in regulating circadian clock function in Neurospora crassa. Nature 421: 948-952.
27. Uchida T, Rossignol F, Matthay M, Mounier R, Couette S, et al. (2004) Prolonged hypoxia differentially regulates hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha expression in lung epithelial cells: implication of natural antisense HIF-1alpha. The Journal of biological chemistry 279: 14871-14878.
28. Sleutels F, Zwart R, Barlow DP (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature 415: 810-813.
29. Lavorgna G, Dahary D, Lehner B, Sorek R, Sanderson CM, et al. (2004) In search of antisense. Trends Biochem Sci 29: 88-94.
30. Sandberg R, Ernberg I (2005) Assessment of tumor characteristic gene expression in cell lines using a tissue similarity index (TSI). Proc Natl Acad Sci U S A 102: 2052-2057.
31. Hayflick L, Moorhead PS (1961) The serial cultivation of human diploid cell strains. Exp Cell Res 25: 585-621.
32. Hayflick L (1965) The Limited in Vitro Lifetime of Human Diploid Cell Strains. Exp Cell Res 37: 614-636.
33. Fire A, Xu S, Montgomery M, Kostas S, Driver S, et al. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391: 806-811.
34. Stark G, Kerr I, Williams B, Silverman R, Schreiber R (1998) How cells respond to interferons 67(1):227. Annual Review of Biochemistry 67: 227-264.
35. Elbashir S, Harborth J, Lendeckel W, Yalcin A, Weber K, et al. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411: 494-498.
36. Schwarz D, Hutvágner G, Du T, Xu Z, Aronin N, et al. (2003) Asymmetry in the assembly of the RNAi enzyme complex. Cell 115: 199-208.
37. Valencia-Sanchez M, Liu J, Hannon G, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. Genes & development 20: 515-524.
38. Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, et al. (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. Science (New York, NY) 303: 672-676.
39. Judge AD, Sood V, Shaw JR, Fang D, McClintock K, et al. (2005) Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. Nat Biotechnol 23: 457-462.
40. Fedorov Y, Anderson E, Birmingham A, Reynolds A, Karpilow J, et al. (2006) Off-target effects by siRNA can induce toxic phenotype. RNA (New York, NY) 12: 1188-1196.
41. Jackson A, Bartz S, Schelter J, Kobayashi S, Burchard J, et al. (2003) Expression profiling reveals off-target gene regulation by RNAi. Nat Biotech 21: 635-637.
42. Birmingham A, Anderson E, Reynolds A, Ilsley-Tyree D, Leake D, et al. (2006) 3[prime] UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nat Meth 3: 199-204.

43. Jackson A, Burchard J, Schelter J, Chau B, Cleary M, et al. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. RNA (New York, NY) 12: 1179-1187.
44. Jackson A, Burchard J, Leake D, Reynolds A, Schelter J, et al. (2006) Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. RNA (New York, NY) 12: 1197-1205.
45. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, et al. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34: 267-273.
46. Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, et al. (1999) Housekeeping genes as internal standards: use and limits. J Biotechnol 8: 291-295.
47. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, et al. (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC molecular biology 7: 3.
48. Lockhart D, Dong H, Byrne M, Follettie M, Gallo M, et al. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. Nature biotechnology 14: 1675-1680.
49. Levy S, Sutton G, Ng P, Feuk L, Halpern A, et al. (2007) The Diploid Genome Sequence of an Individual Human. PLoS Biology 5: 254-254.
50. Patil N, Berno A, Hinds D, Barrett W, Doshi J, et al. (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science (New York, NY) 294: 1719-1723.
51. Johnson G, Esposito L, Barratt B, Smith A, Heward J, et al. (2001) Haplotype tagging for the identification of common disease genes. Nature genetics 29: 233-237.
52. Consortium IH (2003) The International HapMap Project. Nature 426: 789-796.
53. Consortium IH (2005) A haplotype map of the human genome. Nature 437: 1299-1320.
54. Gibbs J, Singleton A (2006) Application of genome-wide single nucleotide polymorphism typing: simple association and beyond. PLoS genetics 2: 150-150.
55. Cardon L, Abecasis G (2003) Using haplotype blocks to map human complex trait loci. Trends in genetics : TIG 19: 135-140.
56. Fung H-C, Scholz S, Matarin M, Simón-Sánchez J, Hernandez D, et al. (2006) Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. Lancet neurology 5: 911-916.
57. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445: 881-885.
58. Saxena R, Voight B, Lyssenko V, Burtt N, de Bakker P, et al. (2007) Genome-Wide Association Analysis Identifies Loci for Type 2 Diabetes and Triglyceride Levels. Science 316: 1331-1336.
59. Matarín M, Brown W, Scholz S, Simón-Sánchez J, Fung H-C, et al. (2007) A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. Lancet neurology 6: 414-420.
60. Schymick J, Scholz S, Fung H-C, Britton A, Arepalli S, et al. (2007) Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet neurology 6: 322-328.
61. Matsushima-Nishiu M, Unoki M, Ono K, Tsunoda T, Minaguchi T, et al. (2001) Growth and gene expression profile analyses of endometrial cancer cells expressing exogenous PTEN. Cancer Res 61: 3741-3749.
62. Unoki M, Nakamura Y (2001) Growth-suppressive effects of BPOZ and EGR2, two genes involved in the PTEN signaling pathway. Oncogene 20: 4457-4465.
63. Nakajima A, Kataoka K, Hong M, Sakaguchi M, Huh N (2003) BRPK, a novel protein kinase showing increased expression in mouse cancer cell lines with higher metastatic potential. Cancer Lett 201: 195-201.

64. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, et al. (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 304: 1158-1160.
65. Valente EM, Salvi S, Ialongo T, Marongiu R, Elia AE, et al. (2004) PINK1 mutations are associated with sporadic early-onset parkinsonism. Ann Neurol 56: 336-341.
66. Wang F, Feng X, Ma J, Zou H, Chan P (2006) A common A340T variant in PINK1 gene associated with late-onset Parkinson's disease in Chinese. Neurosci Lett 410: 121-125.
67. Marongiu R, Brancati F, Antonini A, Ialongo T, Ceccarini C, et al. (2006) Whole gene deletion and splicing mutations expand the *PINK1* genotypic spectrum. Human mutation 943.
68. Tang B, Xiong H, Sun P, Zhang Y, Wang D, et al. (2006) Association of PINK1 and DJ-1 confers digenic inheritance of early-onset Parkinson's disease. Hum Mol Genet 15: 1816-1825.
69. Silvestri L, Caputo V, Bellacchio E, Atorino L, Dallapiccola B, et al. (2005) Mitochondrial import and enzymatic activity of PINK1 mutants associated to recessive parkinsonism. Hum Mol Genet 14: 3477-3492.
70. Beilina A, Van Der Brug M, Ahmad R, Kesavapany S, Miller DW, et al. (2005) Mutations in PTEN-induced putative kinase 1 associated with recessive parkinsonism have differential effects on protein stability. Proc Natl Acad Sci U S A 102: 5703-5708.
71. Sim C, Lio D, Mok S, Masters C, Hill A, et al. (2006) C-terminal truncation and Parkinson's disease-associated mutations down-regulate the protein serine/threonine kinase activity of PTEN-induced kinase-1. Hum Mol Genet 15: 3251-3262.
72. Gandhi S, Muqit MM, Stanyer L, Healy DG, Abou-Sleiman PM, et al. (2006) PINK1 protein in normal human brain and Parkinson's disease. Brain 129: 1720-1731.
73. Hoepken HH, Gispert S, Morales B, Wingerter O, Del Turco D, et al. (2007) Mitochondrial dysfunction, peroxidation damage and changes in glutathione metabolism in PARK6. Neurobiol Dis 25: 401-411.
74. Taymans JM, Van den Haute C, Baekelandt V (2006) Distribution of PINK1 and LRRK2 in rat and mouse brain. J Neurochem.
75. Clark IE, Dodson MW, Jiang C, Cao JH, Huh JR, et al. (2006) Drosophila pink1 is required for mitochondrial function and interacts genetically with parkin. Nature 441: 1162-1166.
76. Park J, Lee SB, Lee S, Kim Y, Song S, et al. (2006) Mitochondrial dysfunction in Drosophila PINK1 mutants is complemented by parkin. Nature.
77. Yang Y, Gehrke S, Imai Y, Huang Z, Ouyang Y, et al. (2006) Mitochondrial pathology and muscle and dopaminergic neuron degeneration caused by inactivation of Drosophila Pink1 is rescued by Parkin. Proc Natl Acad Sci U S A 103: 10793-10798.
78. Wang D, Qian L, Xiong H, Liu J, Neckameyer WS, et al. (2006) Antioxidants protect PINK1-dependent dopaminergic neurons in Drosophila. Proc Natl Acad Sci U S A 103: 13520-13525.
79. MacKeigan JP, Murphy LO, Blenis J (2005) Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. Nat Cell Biol 7: 591-600.
80. Wang H, Chou A, Yeh T, Li A, Chen Y, et al. (2007) PINK1 mutants associated with recessive Parkinson's disease are defective in inhibiting mitochondrial release of cytochrome c.
81. Petit A, Kawarai T, Paitel E, Sanjo N, Maj M, et al. (2005) Wild-type PINK1 prevents basal and induced neuronal apoptosis, a protective effect abrogated by Parkinson disease-related mutations. J Biol Chem 280: 34025-34032.
82. Deng H, Jankovic J, Guo Y, Xie W, Le W (2005) Small interfering RNA targeting the PINK1 induces apoptosis in dopaminergic cells SH-SY5Y. Biochem Biophys Res Commun 337: 1133-1138.
83. Pridgeon JW, Olzmann JA, Chin LS LL (2007) PINK1 Protects against Oxidative Stress by Phosphorylating Mitochondrial Chaperone TRAP1. PLoS Biol 5.

84. Plun-Favreau H, Klupsch K, Moisoi N, Gandhi S, Kjaer S, et al. (2007) The mitochondrial protease HtrA2 is regulated by Parkinson's disease-associated kinase PINK1. Nat Cell Biol.
85. Zhou H, Falkenburger BH, Schulz JB, Tieu K, Xu Z, et al. (2007) Silencing of the Pink1 gene expression by conditional RNAi does not induce dopaminergic neuron death in mice. Int J Biol Sci 3: 242-250.
86. Kitada T, Pisani A, Porter DR, Yamaguchi H, Tscherter A, et al. (2007) Impaired dopamine release and synaptic plasticity in the striatum of PINK1-deficient mice. Proc Natl Acad Sci U S A 104: 11441-11446.
87. Petersen KF, Dufour S, Befroy D, Garcia R, Shulman GI (2004) Impaired mitochondrial activity in the insulin-resistant offspring of patients with type 2 diabetes. N Engl J Med 350: 664-671.
88. Patti ME, Butte AJ, Crunkhorn S, Cusi K, Berria R, et al. (2003) Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. Proc Natl Acad Sci U S A 100: 8466-8471.
89. Tesch PA, Trieschmann JT, Ekberg A (2004) Hypertrophy of chronically unloaded muscle subjected to resistance exercise. J Appl Physiol 96: 1451-1458.
90. Puigserver P, Spiegelman BM (2003) Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (PGC-1 alpha): transcriptional coactivator and metabolic regulator. Endocr Rev 24: 78-90.
91. Yamagishi S-I, Matsui T, Ueda S-I, Nakamura K, Imaizumi T (2007) Advanced glycation end products (AGEs) and cardiovascular disease (CVD) in diabetes. Cardiovascular & hematological agents in medicinal chemistry 5: 236-240.
92. Dalfó E, Portero-Otín M, Ayala V, Martínez A, Pamplona R, et al. (2005) Evidence of oxidative stress in the neocortex in incidental Lewy body disease. Journal of neuropathology and experimental neurology 64: 816-830.
93. Lu C, He J, Cai W, Liu H, Zhu L, et al. (2004) Advanced glycation endproduct (AGE) receptor 1 is a negative regulator of the inflammatory response to AGE in mesangial cells. Proceedings of the National Academy of Sciences 101: 11767-11772.
94. He C, Koschinsky T, Buenting C, Vlassara H (2001) Presence of diabetic complications in type 1 diabetic patients correlates with low expression of mononuclear cell AGE-receptor-1 and elevated serum AGE. Molecular medicine (Cambridge, Mass) 7: 159-168.
95. Hastings M, Ingle H, Lazar M, Munroe S (2000) Post-transcriptional regulation of thyroid hormone receptor expression by cis-acting sequences and a naturally occurring antisense RNA. The Journal of biological chemistry 275: 11507-11513.
96. Hastings M, Milcarek C, Martincic K, Peterson M, Munroe S (1997) Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. Nucleic acids research 25: 4296-4300.
97. Laurencikiene J, Källman A, Fong N, Bentley D, Öhman M (2006) RNA editing and alternative splicing: the importance of co-transcriptional coordination. EMBO Rep 7: 303-307.
98. Rueter S, Dawson T, Emeson R (1999) Regulation of alternative splicing by RNA editing. Nature 399: 75-80.
99. Bratt E, Öhman M (2003) Coordination of editing and splicing of glutamate receptor pre-mRNA. RNA 9: 309-318.
100. Thrash-Bingham C, Tartof K (1999) aHIF: a natural antisense transcript overexpressed in human renal cancer and during hypoxia. Journal of the National Cancer Institute 91: 143-151.
101. Kuroda Y, Mitsui T, Kunishige M, Shono M, Akaike M, et al. (2006) Parkin enhances mitochondrial biogenesis in proliferating cells. Hum Mol Genet 15: 883-895.
102. Gandhi S, Wood NW (2005) Molecular pathogenesis of Parkinson's disease. Hum Mol Genet 14 Spec No. 2: 2749-2755.
103. Hu G, Jousilahti P, Bidel S, Antikainen R, Tuomilehto J (2007) Type 2 diabetes and the risk of Parkinson's disease. Diabetes care 30: 842-847.

104. Ristow M (2004) Neurodegenerative disorders associated with diabetes mellitus. Journal of molecular medicine (Berlin, Germany) 82: 510-529.