Univerzita Karlova v Praze

1. lékařská fakulta

Charles University in Prague

1st Faculty of Medicine

Autoreferát disertační práce

Autoreview of the Ph.D. thesis

MECHANISMY REGULACE EXPRESE GENŮ PRO ORNITIN TRANSKARBAMYLÁZU A BETA-GLUKOCEREBROSIDÁZU A JEJICH VÝZNAM V DIAGNOSTICE

REGULATORY MECHANISMS OF ORNITHIN TRANSCARBAMYLASE AND BETA-GLUCOCEREBROSIDASE GENE EXPRESSION AND THEIR RELEVANCE TO DIAGNOSTICS

Mgr. Ondřej Lukšan

2014

Doktorské studijní programy v biomedicíně

Univerzita Karlova v Praze a Akademie věd České republiky

Obor: Biochemie a patobiochemie

Předseda oborové rady: Prof. MUDr. Stanislav Štípek, DrSc.

Školicí pracoviště: Institut klinické a experimentální medicíny, Laboratoř experimentální hepatologie, Praha, ČR

Školitel: doc. MUDr. et Mgr. Milan Jirsa, CSc.

Konzultant: RNDr. Lenka Dvořáková, CSc.

Disertační práce bude nejméně pět pracovních dnů před konáním obhajoby zveřejněna k nahlížení veřejnosti v tištěné podobě na Oddělení pro vědeckou činnost a zahraniční styky Děkanátu 1. lékařské fakulty.

Table of contents

Abstrakt

Konečná diagnóza dědičných metabolických poruch je obvykle určena na základě vyšetření enzymové aktivity (což je mnohdy komplikované) a/nebo za pomoci molekulárně genetického vyšetření. Výsledky standardní mutační analýzy však mohou být někdy zavádějící, zejména v případě rozsáhlých reorganizací na úrovni chromozomu či u mutací v regulačních oblastech. Předkládaná práce je zaměřena na charakterizaci komplexních mutací v genu pro ornitin transkarbamylázu (*OTC*), studium regulačních oblastí *OTC* a genu pro β-galaktozidázu (*GBA*).

V rámci studie provedené na souboru čítajícím 37 pacientů s kongenitální hyperamonémií II. typu (OTCD) bylo identifikováno a charakterizováno 14 nových mutací včetně tří rozsáhlých delecí. Byl zaznamenán i případ symptomatické heterozygotky pro hypomorfní mutaci p.R129H, u níž byl fenotyp OTCD zřejmě podmíněn posunem v X-inaktivaci ve prospěch mutované alely. Nález byl podložen sledováním nevýhodné X-inaktivace ve třech různých periferních tkáních.

Za účelem prověření patogenity promotorové variace c.-366A>G v *OTC* nalezené v jedné rodině s mírným fenotypem OTCD byly identifikovány tři alternativní počátky transkripce (PT) a vymezen promotor lidského *OTC*. Transkripční aktivita promotoru, stejně tak jako aktivita nově nalezeného distálního enhanceru, byla ověřena funkční esejí. Výsledky ukázaly klíčový význam interakce promotoru s enhancerem pro tkáňově specifickou expresi *OTC* v játrech. Přítomnost variace c.-366A>G vedla k padesátiprocentním snížení aktivity mutovaného promotoru v interakci s enhancerem.

Detailní charakterizace promotoru lidského *OTC* odhalila dva silné *cis*-aktivní regulační elementy odpovídající vazebným místům pro HNF-4α. V oblasti proximálního promotoru byl nalezen třetí vysoce konzervovaný motiv rozpoznávaný HNF-4α, přičemž všechna tři vazebná místa se nalézají ve vzdálenosti do 35 bp nad PT. Promotor *OTC* přitom postrádá základní elementy nezbytné pro iniciaci transkripce v konvenčních pozicích. Dosažené výsledky silně podporují úlohu HNF-4α v iniciaci transkripce *OTC* u člověka.

Metodický postup zavedený při studiu *OTC* byl následně aplikován na charakterizaci alternativního promotoru *GBA*. V případě alternativního transkriptu nesoucího dva nepřekládané exony -1 a -2 byly nalezeny tři různé PT. Při porovnání s normálním promotorem vykazoval alternativní promotor nižší transkripční aktivitu, expresní profil napříč několika různými tkáněmi však byl srovnatelný. Byla vyslovena hypotéza, že výrazné rozdíly ve fenotypu u pacientů se stejným genotypem *GBA* mohou být dány variabilní expresí mutovaného alternativního transkriptu. Tuto hypotézu se však nepodařilo experimentálně potvrdit na souboru 20 pacientů s Gaucherovou chorobou.

Výsledky výše uvedených studií rozšiřují možností molekulární diagnostiky OTCD a Gaucherovy choroby.

Klíčová slova: Ornitin transkarbamyláza, kongenitální hyperamonémie II. typu, kyselá β-glukocerebrosidáza, Gaucherova choroba, genová exprese, regulace transkripce, promotor, enhancer

Abstract

Definitive diagnosis of inherited metabolic disorders commonly depends on the measurement of enzyme activity (which is often complicated) and/or molecular genetic testing. Yet even the standard mutation analysis can bring false negative results in the case of gross chromosomal rearrangements or incorrect regulation of gene expression due to the mutations in regulatory regions. In the present study I focused on characterization of complex mutations affecting the gene encoding ornithin transcarbamylase (OTC) followed by studies of regulatory regions of *OTC* and *GBA* (the gene encoding β-glucocerebrosidase).

In the first study we identified 14 novel mutations including three large deletions in a cohort of 37 patients with OTC deficiency (OTCD). Subsequently we evaluated clinical significance of all these mutations. We also found a heterozygote carrying a hypomorphic mutation and manifesting OTCD most likely due to unfavorable X-inactivation which was observed independently in three different peripheral tissues.

In order to evaluate the clinical significance of a promoter variation c.-366A>G found in a family with mild OTCD we identified three alternative transcription start sites (TSSs) of human *OTC* and delimited the promoter. We also found a distal enhancer and performed functional analysis of both regulatory regions. Our results indicate that tissue specific expression of *OTC* in the liver depends on the promoter-enhancer interaction. The variation c.-366A>G decreased the promoter-enhancer transcriptional activity by 50%.

A detailed characterization of human *OTC* promoter revealed two positive *cis*-acting regulatory elements corresponding to HNF-4α binding sites. Both sites, similarly as a third HNF-4α recognition motif found in the proximal promoter, are located within 35 bases upstream of the TSSs. Since the *OTC* promoter lacks general core promoter elements such as TATA-box or initiators on standard positions, our results strongly suggest an important role of HNF-4α in the control of *OTC* transcription in human.

A similar approach as in the *OTC* gene studies was used in studies of an upstream promoter of *GBA*. We identified three alternative TSSs and performed function analysis of the alternative promoter. Its transcriptional activity was lower than that of the normal promoter while expression profiles across multiple tissues were comparable. We hypothesized that phenotypic differences in patients with the same genotype may be caused by variable expression of mutant *GBA*; however, our hypothesis was not confirmed experimentally in a group of twenty Gaucher patients.

In conclusion, our findings extend the possibilities of molecular genetic testing for OTCD and Gaucher disease.

Keywords: Ornithine transcarbamylase, ornithine transcarbamylase deficiency, acid β-glucocerebrosidase, Gaucher disease, gene expression, regulation of transcription, promoter, enhancer.

1. Introduction

1.1. OTC deficiency

1.1.1. Nitrogen metabolism and urea cycle

Nitrogen is an essential component of many physiologically active compounds. A central role in nitrogen metabolism of higher organisms plays ammonia, the major byproduct of protein and nucleic acid catabolism. Ammonia is toxic in even small amounts and must be transformed into a neutral molecule to be properly transported to the kidneys and excreted via urine. Quantitatively the most important disposal route for ammonia is the conversion to urea. Urea is formed from NH_3 and CO_2 in the liver during urea cycle - a step wise cyclic process described by Krebs and Henseleit (1932).

The excess of ammonia is channeled to the mitochondria of liver cells and converted into carbamoyl phosphate (CP). The reaction of HCO3- and NH3 is catalyzed by CP synthetase 1 and requires co-operation of the allocteric activator, N-acetyl glutamate (NAG). Still in the mitochondrial matrix, carbamoyl phosphate enters the urea cycle. Ornithine transcarbamylase (OTC) transfers the carbamoyl group to ornithine, producing citrulline. Citrulline moves to the cytoplasm and acquires a second atom of waste nitrogen by condensation with aspartate yielding argininosuccinate. The reaction is catalyzed by argininosuccinate synthetase. Argininosuccinate is cleaved by argininosuccinate lyase into arginine and fumarate which is recycled for use by the citric acid cycle. The final reaction of the urea cycle is the hydrolysis of arginine to yield urea and regenerate ornithine, which re-enters the mitochondrion.

In short-term, regulation of urea synthesis is mediated by three factors: substrate availability, changes in the concentration of N-acetyl glutamate and the acid-base status. Longer-term changes occur in response to alterations in amino acid nitrogen flux. Glucagon, insulin, and glucocorticoids play major roles in mediating changes in activity of the urea cycle enzymes under these conditions.

1.1.2. Urea cycle disorders

Urea cycle disorders (UCDs) are inborn errors in the metabolism of waste nitrogen that can lead to brain damage and death due to increased levels of blood ammonia. They involve deficiencies in one of the enzymes required by the urea cycle and deficiencies of N-acetylglutamate synthase and the mitochondrial ornithine/citrulline transporter ORNT1 (Table 1). Except for OTC deficiency, all listed disorders are inherited in an autosomal recessive manner.

Table 1: Overview of the urea cycle disordes

Disease	**Gene symbol**	**MIM ID**	**Mechanism**
N-acetylglutamate synthase deficiency	*NAGS*	237310	Enzymatic block within urea cycle
Carbamoyl phosphate synthetase 1 deficiency	*CPS1*	237300	Enzymatic block within urea cycle
Ornithine carbamoyltransferase deficiency	*OTC*	311250	Enzymatic block within urea cycle
Argininosuccinate synthetase deficiency	*ASS1*	215700	Enzymatic block within urea cycle
Argininosuccinate lyase deficiency	*ASL*	207900	Enzymatic block within urea cycle
Arginase 1 deficiency	*ARG1*	207800	Enzymatic block within urea cycle
Hyperammonemia-Hyperornithinemia-Homocitrullinuria syndrome	*SLC25A15*	238970	Lack of mitochondrial ornithine as one of the OTC substrates
Citrullinemia type II	*SLC25A13*	603471, 605814	Lack of aspartic acid as one of the ASS substrates

The incidence of UCDs is estimated to be at least 1:30,000 births but partial defects may make the number much higher.

UCDs are commonly characterized by hyperammonemia followed by encephalopathy, and respiratory alkalosis. Newborns with an UCD commonly appear normal but they develop symptoms within 24 hours. The hyperammonemia is less severe in patients with partial enzyme deficiencies. In such cases, the first recognized clinical episode may be delayed for months or years. The most common of these late-onset diseases occurs in female carriers of a mutation in OTC on one of their X chromosomes.

The diagnosis of a urea cycle disorder is based on clinical suspicion and biochemical and molecular genetic testing. Potential presence of an UCD is indicated by elevated plasma ammonia associated with a

normal anion gap and a normal plasma glucose concentration. A definitive diagnosis of a urea cycle defect depends on either results of molecular genetic testing (which is available for all UCDs) or measurement of enzyme activity in a liver biopsy specimen.

1.1.3. The *OTC* gene

Human OTC gene is located on the short arm of chromosome X (Xp2l.1) (Lindgren V, et al., 1984), spans 73 kb and comprises 10 exons encoding a 354-amino acid monomer (Hata A, et al., 1988). The transcription start site of human *OTC* was first assigned to the position c.-135 (GenBank K02100; Horwich AL, et al., 1984). Later published data suggest that transcription of human *OTC* does not initiate at a particular site but within a region of about 70 bases (Brusilow SW and Horwich AL, 2004).

OTC is expressed predominantly in hepatocytes and epithelial cells of the intestinal mucosa (Ryall J, et al., 1985; Hamano Y, et al., 1988). The regulation of *Otc* transcription was studied in rodents. It has been shown that the 0.8 kb 5'-flanking region of the mouse gene (Veres G, et al., 1986) and 1.3 kb region of the rat gene (Murakami T, et al., 1989) contained sufficient information to control *Otc* gene expression. Another positive regulatory element, a distal enhancer ranging 230 bp and situated 11 kb upstream of the transcription start site has been identified in rat (Murakami T, et al., 1990).

Two positive *cis*-acting regulatory elements recognized by HNF4α elements were identified in rat (Kimura A, et al., 1993) and mouse (Inoue Y, et al., 2002). Moreover, HNF4α-null mice exhibited a phenotype closely similar to UCD (Inoue Y, et al., 2002). Another two HNF-4 binding sites and two sites recognized by C/EBP family members were identified within the 110 bp rat minimal enhancer (Nishiyori A, et al., 1994). In co-transfection experiments the authors have shown that both HNF-4 and C/EBPβ are necessary and neither alone is sufficient for activation of the reconstituted enhancer in non-hepatic cells. Combinatorial operation of these two liver-enriched (but not strictly liver-specific) transcription factors leads to more restricted liver-specific transcription of the *Otc* gene (Nishiyori A, et al., 1994).

These data indicate that the tissue-specific *Otc* expression in rodents is based on the promoter-enhancer interaction in the hepatocyte-specific transcription factor milieu. No such regulatory elements have been reported in humans.

1.2. Gaucher disease

1.2.1. Lysosomal storage diseases

Lysosomal storage diseases (LSDs) represent a heterogeneous group of more than 50 rare inherited disorders characterized by accumulation of waste products in the lysosomes. Progressive accumulation of these products leads to cellular dysfunction (Wenger DA, et al., 2003; Wilcox WR, 2004). Widespread cellular destruction eventually causes tissue and organ dysfunction observed in clinical abnormalities. LSDs are commonly linked to the deficiencies of the lysosomal hydrolases but they encompass also deficiencies in other proteins, such as activators, transporters and proteins involved in post-translational modifications of lysosomal enzymes.

Almost all LSDs are inherited as autosomal recessive traits, except for the Fabry, Hunter and Danon diseases which are X-linked (Wilcox WR, 2004; Sugie Ket al., 2002). The individual incidence of UCDs varies between 1 in 50 000 and 1 in 4×10^6, but taken together they are found in about 1 : 8000 births, which makes them a relatively common health problem (Meikle PJ, et al., 1999). Certain populations have even a higher incidence of particular LSD compared with the general population.

The spectrum of clinical phenotypes is very broad including organomegaly and central nervous system dysfunction. The patients suffering from LSDs are mostly born apparently healthy and the symptoms develop progressively. Clinical manifestations are often inconsistent, even within families, and patients with identical genotypes can have significantly different phenotypes (Wilcox WR, 2004).

The diagnosis is definitively confirmed by the detection of enzymatic deficiency at the molecular level. Biochemical testing is often accompanied with molecular genetic testing, which also enables prenatal and postnatal testing and allows the provision of genetic counseling. In addition, molecular genetic testing can clarify the type of genetic variation and its impact on the protein and on the presence of residual enzyme activity. This information is crucial in evaluating treatment options (Marsden D and Levy H, 2010).

For a majority of LSDs, therapeutic management consists of symptomatic care of disease manifestations with no possibility for cure. Supportive care measures are disease specific, depending on the organs involved and degree of physical impairment

1.2.2. Metabolism of glycosphingolipids and Gaucher disease

Glycosphingolipids (GSLs) are ubiquitous constituents of eukaryotic plasma membranes and the major glycans of the vertebrate brain. They are composed of a glycan structure attached to a lipid tail that contains the sphingolipid ceramide.

The functions of membrane-associated glycosphingolipids could be divided into two major categories: mediating cell–cell interactions via binding to complementary molecules on opposing plasma membranes (*trans* recognition) and modulating activities of proteins in the same plasma membrane (*cis* regulation). More complex glycosphingolipids thus contribute to the cell–cell recognition and to the regulation of signal transduction.

The breakdown of complex GSLs occurs stepwise under the action of lysosomal hydrolases until they are metabolized to the common components - monosacharides, sphingosine and fatty acids - which are then available for reuse

Gaucher's disease (GD) is the most common glycosphingolipid storage disorder. The rare autosomal recessive disease is caused by the by a functional deficiency of the β-glucocerebrosidase (EC.3.2.1.45) which leads to accumulation of glucocerebroside in the body, predominantly in the liver, spleen, and bone marrow. Rare variant forms of Gaucher disease result from deficiency of the sphingolipid activator protein, saposin C (Beutler E and Grabowski GA, 2001). The GD has been divided into three clinical subgroups reflecting the degree and rate of progression of involvement in the central nervous system (Westbroek W, et al., 2011). Severity and rate of disease progression widely varies, especially in adults, which makes treatment decisions extremely difficult in some patients. The most frequent form is the non-neuronopathic GD type 1 (OMIM#230800) affecting about 90% of patients.
Currently there is no effective treatment for type 2 Gaucher patients. Therapeutic management of type 1 and type 3 GD include enzyme replacement therapy using intravenously applied artificial enzymes or bone marrow transplantation.

1.2.3. The *GBA* gene

The acid β-glucocerebrosidase gene (*GBA*, EC.3.2.1.45, MIM#606463) is located at a gene-rich locus 1q21 (Entrez Gene) or 1q22 (Ensembl, HGNC). It comprises 11 exons and 10 introns, spanning a sequence of 7.8 kb. A highly homologous pseudogene (ps*GBA*), resulting from a tandem duplication event involving ancestral GBA and neighbouring metaxin genes, is located 16 kb downstream (Horowitz M, et al., 1989) and spans 5.5 kb in length, retaining the same organization of exons and introns as the functional gene. The presence of a conserved pseudogene in the same locus is responsible for a subset of Gaucher mutations originating from recombination events between *GBA* and *psGBA* (Eyal N, et al., 1990; Latham TE, et al., 1991).

Beside the major transcript of *GBA* spanning approximately 2.6 kb (Horowitz M, et al., 1989), several alternative mRNA sequences have been identified, including variants with additional non coding exons -1 and -2. Human *GBA* cDNA contains 2 in-frame ATG start codons (Sorge J, et al., 1985) resulting in the synthesis of two signal peptides differing in hydrophobicity. Both precursor variants are sharing the 536-residue sequence of functional glucocerebrosidase monomer.

Promoter regions of genes encoding lysosomal enzymes commonly share characteristic features of housekeeping genes. The *GBA* promoter is unusual in this respect since it contains TATA and CAAT boxes but no Sp1 binding site (Horowitz M, et al., 1989). Moreover, the *GBA* mRNA is expressed differentially and glucocerebrosidase activity varies widely between different cell types, thus indicating that regulation of transcription is an important factor in the control of glucocerebrosidase expression. On the other hand, other regulatory mechanisms should be considered as well as the ratio between *GBA* mRNA and enzyme activity differs between cell lines (Doll RF and Smith FI, 1993).

1.3. Regulation of gene expression

1.3.1. Gene expression and control

Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product. The multi-level process consisting of chromatin remodeling, transcription, RNA processing, translation and post-translational modification of proteins is controlled at different stages.

On the largest scale, the expression of genetic information from chromosomes is performed by changes in chromatin structure. In eukaryotic organisms, chromatin organization plays a critical role in transcriptional regulation. Most of chromatin is present in the condensed form to inhibit access of transcription factors (TFs) to the DNA and prevent non-controlled over-expression of related genes.

In order for transcription to occur, the area around a prospective transcription zone needs to be unwound to make specific DNA sequences accessible. This is a complex process requiring the coordination of histone modifications, TF binding and other chromatin remodeling activities. Activator TFs located in the promoter are responsible for the recruitment of RNA polymerase II (Pol II) and the control of mRNA synthesis.

As most human genes are divided into exons and introns, and only the exons carry information required for protein synthesis, the primary transcripts have to be processed by splicing. Another processes, important for stability and transport of mature mRNA, are modifications on both, 5'- and 3'- termini. Transcript processing provides an additional level of regulation because only mature mRNA can translocate through nuclear membrane to follow the process of gene expression. Moreover, due to the control of RNA splicing, eukaryotic cells can differentially express genes in a variety of alternative transcripts performing various functions in the organism.

Once present in the cytoplasm, the mRNA is finally competent to perform its function - translation of the sequence encoded by the gene into the protein. The regulation at the level of translation is characterized by differential utilization of preexisting mRNAs. Most regulation is exerted at the level of initiation where the pre-initiation complex (PIC) scans the 5'-untranslated region for the initiation codon. Initiation rate changes involve mainly the interaction of *trans*-acting factors (proteins or occasionally RNAs) with *cis*-acting elements located in the 5′ or 3′ untranslated regions of a particular mRNA (Sonenberg N and Hinnebusch AG, 2009; Jackson RJ, et al., 2010).

1.3.2. Regulation of transcription

The control of gene expression at the level of transcription is performed by altering transcription rate. The dynamic process is mediated through interaction of TFs with *cis*-acting regulatory elements. Each TF has a specific DNA binding domain that recognizes a 6-10 base-pair motif in the DNA sequence, as well as an effector domain. All TFs bind at the promoters just upstream of eukaryotic genes; however, they also bind at distal regulatory regions.

The transcription mediated by Pol II initiates at the core promoter in the close proximity of TSS (Orphanides G, et al., 1996; Woychik and Hampsey, 2002), which is considered as the minimal contiguous DNA sequence that is sufficient to direct the transcription. Typically, it encompasses the site of transcription initiation and extends either upstream or downstream for other additional 35-40 nucleotides. The core promoter sequence contains sequence motifs such as the TATA box, initiator element (Inr) and other general core promoter elements, which are responsible for recruitment and stabilization of proteins forming the PIC of Pol II. However, each of these particular elements is found in some but not all core promoters.

In addition to the core promoter, other *cis*-acting DNA sequences that regulate Pol II transcription include the proximal promoter, enhancers, silencers, and boundary/insulator elements (Blackwood EM and Kadonaga JT, 1998; Bulger M and Groudine M, 1999). These elements contain recognition sites for a variety of sequence-specific TFs. Enhancers and silencers can be located hundreds or thousands of nucleotides from the TSS and act either to activate or to repress transcription. Boundary/insulator elements appear to prevent the spreading of the activating effects of enhancers or the repressive effects of silencers or heterochromatin.

The regulation of transcription is a complex phenomenon. Some target genes have a number of different response elements, with each one capable of responding to a particular signal that is sufficient to initiate transcription.

2. Aims of the study

Our main goal was to extend the possibilities of molecular genetic testing for OTCD and GD. We focused on the following specific aims:

1) To perform a detailed molecular characterization of *OTC* in 37 OTCD patients originating from 29 families, including description of novel disease alleles, precise breakpoint mapping of large rearrangements and genotype–phenotype correlation.
2) To define the 5'UTR, identify human *OTC* regulatory regions and investigate the functional impact of a novel unique variation in the 5'-flanking region in a female patient with clinical and biochemical signs of OTCD.
3) To perform detailed characterization of human *OTC* promoter and enhancer and identify key elements responsible for transcriptional regulation of *OTC* expression.
4) To identify transcription start site(s) of human *GBA* non-coding exon -2, investigate the upstream region for potential promoter activity and examine the expression pattern of alternative transcripts originating from the two alternative promoters in a group of 20 Ashkenazi Jewish Gaucher patients homozygous for the common mild mutation p.N370S.

3. Methods

3.1. General methods

Nucleic acid and protein isolation from human and animal cells and tissues
Polymerase chain reaction
Quantitative real-time polymerase chain reaction
DNA electrophoresis and restriction fragment length polymorphism analyses
Reverse transcription of total RNA (generation of cDNA)
DNA sequencing using the Sanger method
Molecular cloning
Cell culture techniques

3.2. Specific methods

Mutation analysis

Genomic DNA was isolated from peripheral leukocytes using QIAamp columns (Qiagen, Valencia, CA). Exons of human *OTC*, including the intron-exon boundaries, and four upstream non-coding regions, i.e. promoter, enhancer and two highly conserved regions, were analyzed by direct sequencing using the ABI Prism 3100 – Avant and ABI 3500 Genetic analyzers (Applied Biosystems, Foster City, CA). For details, primer sequences and reaction conditions see Ref I and Ref II.

Copy number analysis

In male patients, the approximate extent the deletions was assessed by inspection of presence/absence of PCR products. In females, the copy number of *OTC* coding exons was determined by the multiplex ligation-probe amplification assay (MLPA, MRC Holland, Amsterdam, The Netherlands) according to the manufacturers instructions as decribed in Ref I and Ref II.

Analysis of large deletions and gross chromosomal rearrangements

The analysis of gross chromosomal rearrangements was performed using the Affymetrix Human SNP 6.0 array (Santa Clara, CA). Data analysis was performed within the Affymetrix Genotyping Console 3.0.2 using CEU HapMap samples as the reference set. Breakpoints were located by sequencing of PCR products overlapping the deletion boundaries. For more details, see Ref I.

Determination of transcription start sites

The 5′-terminal mRNA sequence was analyzed by RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) using the First-Choice™ RACE ready liver kit with total human liver or placenta cDNA (Ambion, Austin, TX). Alternativelly, the GeneRacer™ Kit (Invitrogen) with RNAs isolated from human leukocytes, liver, and placenta was carried out. Both assays were performed according to the manufacturer's instructions as described in Ref II and Ref IV. Major PCR products were gel-purified, cloned into the pCR2.1- or pCR4-TOPO® vector (Invitrogen, Carlsbad, CA) and sequenced.

Construction of reporter plasmids
The fragments with corresponding overhangs were ligated into the *XhoI* or *KpnI* (promoter variants) and *SalI* (enhncer) restriction sites of pGL3 or pGL4. Different combinations of inserts in both sense and anti-sense orientation were prepared in order to confirm the specificity of their regulatory function. Serial deletions of promoter-containing constructs were created by site-directed mutagenesis using the QuickChange XL site-directed mutagenesis kit (Stratagene, La Jolla, USA). All resulting constructs were verified by sequencing. For complete list of constructs and other details see Ref II, Ref III and Ref IV.

Transient transfections and reporter gene assays
Transfection using Tfx-20 and FuGene HD reagents (Promega, Mannheim, Germany) were performed in Opti-MEM media (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. The pGL3- or pGL4-derived experimental constructs were cotransfected with the *Renilla* luciferase containing plasmid pRL-TK in order to normalize reaction yields. Reporter assays were performed using Dual Luciferase Reporter assay (Promega, Mannheim, Germany) according to the manufacturer instructions with few modifications as described in Ref II, Ref III and Ref IV. Emitted light was detected using a Berthold Sirius luminometer (Bundoor, Australia) and Synergy 2 automatic microplate reader (BioTek, Winooski, VT).

Nuclear extract preparation
Nuclear and cytoplasmatic extracts were prepared from Hep-G2 cells using the NE-PER nuclear extraction kit (Thermo Scientific, Waltham, MA) according to the manufacturer's instructions. Extracts were dialyzed according to Dignam, et al. (1983). Protein concentrations were determined using the Pierce BCA Protein Assay Kit (Thermo Scientific, Waltham, MA).

DNase I footprinting
DNase I footprinting was used to analyse the interaction of nuclear proteins with the regulatory regions. The method, based on the separation of fluorescently labeled products of DNase I cleavage using capillary electrophoresis (Wilson DO, et al., 2001; Zianni M, et al., 2006), was modified as described in Ref III. Cleaved DNA fragments were purified by phenol-chloroform extraction followed by ethanol precipitation and separated by capillary electrophoresis on the Applied Biosystems 3130 Genetic Analyzer.

Electromobility shift assays
Double-stranded biotin-labeled probes were prepared by standard PCR or by hybridization of complementary oligonucleotides. The binding reactions with nuclear proteins were incubated for 20 min at room temperature, separated using non-denaturing poly-acrylamide gel electrophoresis and electro-blotted on the Amersham Hybond-N nylon membrane (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The membranes were fixed by baking or UV crosslinking, processed using the Chemiluminiscent Nucleic Acid Detection Module (Thermo Scientific, Rockford, IL) according to the manufacturer´s instructions and analyzed with the FUJI LAS 3000 system (Fujifilm, Tokyo, Japan). A non-labeled DNA probe derived from the consensus HNF-4α recognition sequence was used in the competitor assays; A goat anti-HNF-4α polyclonal antibody (Santa Cruz, Dallas, TX) was used in supershift assays. For detail list of DNA probes, buffer compositions and reaction conditions see Ref III.

Real-time PCR
In case of *GBA*, utilization of the transcription initiation sites was studied by quantitative RT-PCR. Primers and TaqMan probes were designed using Primer Express 3.0 software (Applied Biosystems, Foster City, CA). All reactions were performed using StepOne™ Real-time PCR system (Applied Biosystems). The cDNA was obtained by reverse transcription of Human Placenta Total RNA or FirstChoice® Human Total RNA Survey Panel (both from Ambion) using a High Capacity RNA-to-cDNA Kit (Applied Biosystems). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and β-actin (ACTB) were used as endogenous controls for relative quantification. The Ct values were transformed to relative quantities using the delta Ct method and the expression stability was evaluated using geNorm (version 3.5) and NormFinder software packages.

Analysis of DNA methylation
The X-inactivation ratio in *OTC* patients was determined by analysis of methylation status of the human androgen-receptor locus (HUMARA) as described elsewere (Dobrovolny R, et al., 2005). The methylation status of the CpG island overlapping with P2 promoter of *GBA* gene was studied using bisulphite sequencing. Bisulphite-modified genomic DNA of three control individuals served as a template for

amplification of the CpG-rich sequence in the vicinity of P2 promoter in three fragments. Gel-purified products were sequenced. For more details see Ref I and Ref IV.

In-silico predictions and other bio-informatic analyses

Human *OTC* promoter and enhancer regions were predicted on the basis of orthologous regions in rat and mouse. Homology was assessed using the
NCBI Blast tool http://blast.ncbi.nlm.nih.gov/Blast.cgi) and
UCSC genome browser Multiz Alignment&Conservation utility (http://genome.ucsc.edu).
Prediction of transcription factor binding sites in was performed with
MatInspector (http://www.genomatix.de/shop/evaluation.html),
Match (http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi),
TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) and
Alibaba2 (http://www.gene-regulation.com/pub/programs.html#alibaba2).
All reference DNA and transcript sequences were obtained from
NCBI (http://www.ncbi.nlm.nih.gov/refseq/)
The *GBA* ESTs containing sequence homologous to exons −2 and −1 were identified using BLAST and UCSC Genome Browser. The multiple alignment of exon −2 and P2 was modified from 44-way alignment of assembled genomes available from Galaxy (http://main.g2.bx.psu.edu/).

Statistical analysis

Descriptive statistics, parametric and non-parametric tests of statistical hypotheses, analyses of variance (ANOVA) and multiple comparisons were performed using SigmaPlot software (Systat Software, Chicago, IL). All values are given as mean ± SD of replicates. The results were evaluated with Kruskal-Wallis's one way ANOVA on ranks, multiple comparisons were set up using the Dunn's method. Differences at $p \leq 0.01$ were considered as statistically significant.

4. Results and discussion

4.1. Ornithine carbamoyltransferase deficiency: molecular characterization of 29 families

In the first article we report 25 different mutations in 37 probands originating from 29 families. Fourteen of identified genomic variants were novel and three of them, classified as gross chromosomal rearrangements, would be overlooked using standard sequencing methods. A novel 3 bp in-frame deletion p.Leu43del, eliminating leucine 43 residue from a highly conserved domain important for correct folding of the enzyme was identified in one patient. None of the three novel missense mutations p.Ala102Pro, p.Pro158Ser and p.Lys210Glu was indicated to affect splicing or to be directly involved either in the enzyme active site or in the intersubunit interactions (Shi D, et al., 2001). Nevertheless, the mutated residues show critical intramolecular interactions with other spatially neighboring and highly conserved residues. Sequencing of cDNA obtained from the liver specimens of four patients did not reveal any abnormally spliced transcripts caused by a deletion c.126_128delTCT or substitutions c.304G>C and c.829C>T, but uncovered an aberrant splicing in the case of silent transition c.867G>A resulting in the usage of an alternative exonic splicing site followed by an in-frame deletion of 12 nucleotides.

Large deletions were identified in two males and one female. In patient 1, a large deletion of 444 kb encompassed 15 out of 19 exons of *RPGR* and the entire *OTC* and *TSPAN7* genes. An insertion of a 63 bp fragment identical with the region g.41,647–41,704 coinciding with the *AluJb* retroelement was identified between the breakpoints. The complex mutation was likely originated through a two-step insertion/deletion process (Chuzhanova NA, et al., 2003). The insertion may have occured through misalignment between two inexact inverted repeats that occur upstream of the 5'-breakpoint and the intermediate breakpoint and consist of *AluJb* repeats in opposite orientation with subsequent strand switching via interchromosomal serial replication slippage (Chen JM, et al., 2005). A recombinogenic motif CCCCACCCC and its complement found in close proximity to the intermediate breakpoint (position g.41,555) and the 3' breakpoint (position g.445,313), respectively, may have been responsible for DSBs and non-homologous end joining (NHEJ) resulting in deletion.

The other two large deletions affected only parts of *OTC*: the 10 kb-deletion encompassing exons 5 and 6 was identified in patient 8, whereas patient 36 was heterozygous for the 24.5 kb-deletion removing exons 9 and 10. Unlike other large deletions described in the *OTC* gene (Engel K, et al., 2008; Quental R, et al.,2009), no repetitive elements were found in the vicinity of the deletion breakpoints in our case. However,

recombinogenic motifs CCTCCCT and CCTCCTT, as same as their complements, were found in close proximity to the mutation breakpoints in both patients respectively. The deletions have most likely been mediated by DSBs and subsequent NHEJ (Lieber MR, et al., 2003). Small insertions, as seen in our patients, are common for this type of recombination. Such a mechanism has not been previously described in association with *OTC* gene deletions.

In addition to clearly pathogenic loss-of function mutations we found in our cohort a set of hypomorphic mutations. In accordance with the published data (Matsuda I, et al., 1996), we observed that such mutations may cause a broad spectrum of phenotypes even among the male members of the same family. We examined the X-inactivation pattern in the peripheral blood cells, urinary cells and saliva cells of a manifesting carrier with a hypomorphic mutation. The patient skewed significant X-inactivation in favor of the mutant allele, while her asymptomatic heterozygous mother showed the opposite pattern favoring the wild type allele. It has been shown that trends in X-inactivation patterns are comparable among different tissues of the same individual and analysis of X-inactivation accessible tissues (e.g. blood) may be useful to predict the status in inaccessible tissues (Bittel, et al., 2008). In line with these findings our data indicate that manifestation of symptoms caused by hypomorphic mutation might depend on highly skewed X-inactivation in the liver, which is reflected by X-inactivation status in peripheral tissues.

4.2. Disruption of OTC promoter-enhancer interaction in a patient with symptoms of ornithine carbamoyltransferase deficiency

The second article follows a case excluded from the cohort mentioned above and published separately. No predictably pathogenic mutation was found in the coding region of *OTC* of female patient with mild form of OTCD. The only detected sequence variation was a heterozygous single nucleotide substitution NG_008471.1:g.4849A>G (c.-366A>G). The variation was inherited from the patient's father who had no obvious complaints except for eczema but clinical testing revealed mild orotic aciduria and mild elevation of plasma glutamine and glutamic acid and abnormal allopurinol test. The substituted nucleotide c.-366A is conserved in mammals including rat and mouse and the variation has not been published in SNP databases. Therefore we considered the variation a suspect mutation.

Because the 5'-flanking region of *OTC* gene was not studied yet, we had to determine the TSS and decide whereas the variation is located in the 5'-UTR or in the upstream, potentially promoter region. We identified three alternative TSSs located 95, 119 and 169 bp upstream of TIS. Although the human *OTC* TSS was first assigned to the position c.-135 (Horwich AL, et al., 1984), later published data suggested that transcription of human *OTC* does initiate within a region of about 70 bases, mostly at preferred positions c.-95, c.-120, c.-150, c.-161 and c.-166 (Brusilow SW and Horwich AL, 2004). These findings as well as the records from the ESTs library containing spliced clones starting from positions c.-92, c.-120, c.-151 and c.-169 bp upstream of TIS correspond well with our results. Only one TSS was found in mouse and rat, located 136 bp (Veres G, et al., 1986) and 87 bp (Takiguchi M, et al., 1987) upstream of the TSS respectively. These results show that the novel variation c.-366A>G is located upstream of the *OTC* 5'UTR.

A 793 bp fragment encompassing the highly conserved region (72% homology with both, rat and mouse) and ranging from -838 to -46 bp upstream of the TIS was chosen for functional testing of promoter properties. The fragment exhibited significantly increased transcriptional activity in hepatocellular lines Hep-G2, HuH-7 and PCL/PRF/5. Interestingly, similar activity was observed also in HEK293 cells (control renal line). Studies in transgenic mice revealed that introduction of the rat *Otc* controlled by the 1.3 kb rat promoter (Murakami T, et al., 1989) and the human *OTC* controlled by the 0.7 kb mouse promoter (Jones SN, et al., 1990) can conduct small intestine-selective gene expression in mouse at levels comparable with those of the endogenous *Otc*. However, the mRNA levels of the introduced genes were low in the liver. These findings together indicate the presence of some other liver-specific *cis*-acting element(s).

Such an element, a distal enhancer ranging 230 bp and situated 11 kb upstream of the TIS, was indeed identified in rat (Murakami T, et al., 1990). We identified a sequence homologous in 83% with rat enhancer in a conserved region 9 kb upstream of human *OTC* and tested a fragment of 465 bp for its enhancer activity. Subcloning of the enhancer into the reporter construct led to a significant increase in transcriptional activity in Hep-G2 but not in HEK293, thus indicating the role of promoter-enhancer interaction in liver-specific expression of *OTC*.

When comparing the transcriptional activity of normal promoter and promoter carrying the variation c.-366A>G, no significant difference was observed in constructs with promoter alone, but the situation

dramatically changed when enhancer was subcloned. Variant promoter in that case reached only about 50% of the transcriptional activity observed with normal promoter. The results of functional studies thus strongly suggest molecular pathogenicity of the c.-366A>G promoter variation.

4.3. HNF-4α regulates expression of human ornithin transcarbamylase through interaction with two positive cis-acting regulatory elements located in the proximal promoter

Our goal in the third publication was to characterize *OTC* promoter and enhancer and identify elements responsible for the control of *OTC* transcription. In a set of reporter assays, using a series of 5'-deleted promoter fragments, we defined the 223 bp core promoter located -46..-268 bp upstream of the TIS and containing all three TSS. Within the minimal promoter we identified two positive *cis*-acting elements: Cis1+ located within the first 81 base-pairs, reaching 83% of the full promoter activity and Cis2+ located within the upstream 52 bp sequence. Using a combination of bioinformatic predictive methods, DNase I footprinting and EMSA we identified both, Cis1+ and Cis2+ as HNF-4α binding motifs situated at positions c.-105..c.-121 and c.-136..c.-152 respectively. Another HNF-4α binding site was recognized 187 bp upstream of the TIS. The sequence is highly conserved in vertebrates and the affinity of HNF-4 to this region has been clearly demonstrated in rat (Kimura A, et al., 1993) and mouse (Inoue Y, et al., 2002). All three HNF-4α binding sites share conserved matrix sequence and could be considered as at least middle affinity binders to HNF-4α (Fang B, et al. 2012).

Activation of the enhancer in rat is associated with the interaction of promoter-bound HNF-4 with the enhancer-bound C/EBPβ (Nishiyori A, et al., 1994). A similar mechanism involving HNF-4 and C/EBPα has been suggested for the promoter of human apolipoprotein B gene (Metzger S, et al., 1993). We identified a binding site for C/EBPβ in the *OTC* core promoter between positions c.-221 and c.-234 but deletion of the region did not change the level of reporter signal in luciferase assays.

On the other hand, *in silico* analysis of the proximal 212 bp part of the enhancer region revealed a set of six repeated C/EBPβ binding motifs. Accordingly, the DNase I cleavage of the proximal enhancer region was decreased and DNA-protein interactions were observed even with EMSA. Interaction of the liver enriched HNF-4 factors bound to the *OTC* core promoter with the set of periodically occurring liver specific C/EBP-like proteins thus may contribute to the tissue specificity of *OTC* expression in the liver, as was seen in the rat model (Nishiyori A, et al., 1994).

The distribution of HNF-4α recognition sites within the *OTC* core promoter region exhibits an important feature: All three HNF-4α binding motifs are located within 35 bases upstream of the TSSs. Moreover, the sequence of 223 bp proximal promoter lacks any general core promoter elements such as Inr, DPE, BRE, MTE, downstream core element (DCE) and the X core promoter element (XCPE) located at standard positions upstream of the TSSs, thereby suggesting the importance of HNF-4α in the initiation of *OTC* transcription.

Few indices suggesting the role of HNF-4α in the transcription from TATA- less and Inr-less promoters can be found in the literature. The HNF-4α was shown to activate transcription *in vitro* by facilitating assembly of the pre-initiation complex through direct physical interaction with TFIIB (Malik S and Karathanasis SK, 1996). Similar mechanism was observed later in estrogen receptor, an another member of the nuclear hormone receptor superfamily (Sabbah M, et al., 1998). In a recent study, a systematic proteomic characterization of the HNF4α-TFIID interactions revealed that HNF4α can directly target TBP subunit of TFIID to promoters containing HNF4α-binding sites (Takahashi H, et al., 2009). The findings suggest that regulatory mechanism of *OTC* transcription may involve physical interaction of HNF4α with general transcription factors forming the pre-initiation complex of Pol II. Such a mechanism would explain characteristic features observed within *OTC* minimal promoter: existence of multiple TSS, site-oriented initiation of transcription in the absence of general core promoter elements and a partial ability to compensate elimination of one *cis*-active element by the function of others.

4.4. Glucocerebrosidase gene has an alternative upstream promoter, which has features and expression characteristic of housekeeping genes

The methodological approach introduced within the study of *OTC* regulatory regions was applied on the glucocerebrosidase gene. The *GBA* promoter (P1) was assigned to the region located immediately upstream of the coding-exon 1 (Horowitz M, et al., 1989). Surprisingly, four of the five known glucocerebrosidase

transcripts contain alternative non-coding exon −2 or both exons −2 and −1. We identified three alternative TSS located at positions c.-347, c.-380 and c.-413 uptream of the TIS and inspected transcriptional activity of a 1156 bp fragment (P2) ranging from c.-353 to c.-1509 upstream of the non-coding exon -2. The P2 promoter was found to be active, but the reached levels of reporter gene signal were about six times lower than those of P1. Since we used both fragments of similar size, we might miss some important *cis*-active element and therefore the reporter gene expression from P1 and P2 constructs may not be proportional to the strength of the native promoters.

It was shown that lysosomal genes exhibit coordinated transcriptional behavior and are regulated by the transcription factor EB (TFEB) (Sardiello M, et al., 2009). The coordinated lysosomal expression and regulation motif (CLEAR) recognized by TFEB is widely enriched in promoters of lysosomal genes. While two such motifs were previously recognized and confirmed in the P1, we identified another two CLEAR-like sequences in P2 at positions 764 bp and 1238 bp upstream of the TIS. We found that in contrary to the P1, the P2 lacks a TATA-box but it carries a multiple Sp-1 binding sites and an unmethylated CpG island thus presenting common properties with the majority of housekeeping promoters. However, the results from the quantitative RT-PCR in RNA isolated from normal adult human tissues did not show significant differences in the expression from both promoters. Moreover, the expression stability of both, the P1 and P2, appears to be higher than that of the commonly used reference housekeeping genes GAPDH and ACTB.

In line with these findings we hypothesized that differences in the manifestation of GD in patients presenting the same genotype may be caused by variable expression of different GBA transcripts due to variations in the 5’-flanking region. We therefore examined the promoter sequences in a group of 20 Ashkenazi Jewish Gaucher patients homozygous for the common mild mutation p.N370S to inspect if the variable phenotype is not caused by aberrant transcription. Unfortunately, we did not find any sequence variation either in the P1 or in the P2 except for a common polymorphism rs10908459.

5. Conclusions

In the present work we described the molecular basis of OTCD in probands coming from 29 families. We identified and characterized 14 novel mutations affecting the protein-coding region including three large deletions and a pathogenic variation located in the 5'-flanking region. We found and functionally characterized human *OTC* promoter and distal enhancer and suggested mechanism of their involvement in the control of transcription. We found and functionally characterized an upstream *GBA* promoter driving transcription of non-coding exons -2 and -1.

The results, presented in four publications, can be summed up as follows:

1) Fourteen out of twenty five mutations altering the coding sequence of *OTC* gene were found novel in DNA samples from 29 OTCD families. Three of them are missense mutations affecting residues of highly conserved OTC architecture. another three are large deletions. A manifesting heterozygote carrying a hypomorphic mutation with unfavorable X-inactivation was identified and unfavorable mosaic skewing was observed in three peripheral tissues.
2) The sequence spanning 793 bp in the 5'-flanking region of human *OTC* acts as a promoter. The transcriptional activity of the promoter depends on a newly discovered upstream enhancer located 9 kb upstream of the translation initiation site. The promoter – enhancer interaction contributes to tissue specific expression of *OTC* in liver-derived cell lines. Single nucleotide substitutions in the regulatory regions such as the c.-366A>G variation may cause OTC deficiency. Therefore, analysis of the regulatory regions of *OTC* should be considered in patients with OTCD and negative finding in the coding region of the gene.
3) HNF-4α plays an essential role in regulation of human *OTC* expression. Three HNF-4α binding motifs are located within 35 bases upstream each of the transcription start sites in the minimal promoter. Two of these sites were identified as *cis*-active regulatory elements responsible for the full transcriptional activity of the promoter and both were shown to cooperate with distal enhancer in the synergistic manner resulting in significantly increased liver-specific expression of *OTC*. Moreover, DNA-bound HNF-4α seems to act as a direct initiator of *OTC* transcription through recruitment of general transcription factors TFIIB and TDFIID and stabilization of the pre-initiation complex to the promoter at particular sites lacking TATA-box or initiator.
4) The region 1 kb upstream of the *GBA* non-coding exon −2 acts as an alternative promoter. The transcriptional activity of the alternative promoter P2 seems to be lower than that of the promoter P1 driving the *GBA* expression from exon 1. The promoter P2 lacks TATA-box but contains two CLEAR-like, multiple Sp-1 binding sites and an unmethylated CpG island. Expression of the transcripts originating from both promoters was stable among multiple tissues. The DNA analysis of the P1 and P2 promoters in a group of twenty Ashkenazi Jewish Gaucher patients with pleiotropic phenotype homozygous for the common mild mutation p.N370S did not reveal any sequence variations responsible for the phenotypic differences.

Our data extend the possibilities of genetic testing for OTC and GBA deficiency.

6. References

1. Beutler E and Grabowski GA. Gaucher disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D. The Metabolic and Molecular Basis of Inherited Disease. 8th ed. New York, McGraw-Hill, Inc, 2001; p.3635–68.
2. Bittel DC, Theodoro MF, Kibiryeva N, Fischer W, Talebizadeh Z, Butler MG. Comparison of X-chromosome inactivation patterns in multiple tissues from human females. J Med Genet. 2008; 45:309–13.
3. Blackwood EM, Kadonaga JT. Going the distance: a current view of enhancer action. Science. 1998; 281(5373):60-3.
4. Brusilow SW and Horwich AL. Urea Cycle Enzymes. In: Scriver CR, Beaudet AL, Sly WS, Valle D. Metabolic and Molecular Bases of Inherited Diseases - OMMBID. www.ommbid.com. New York, McGraw Hill, Inc, 2004; Chap. 85.
5. Bulger M and Groudine M. Looping versus linking: toward a model for long-distance gene activation. Genes Dev. 1999; 13(19):2465-77.
6. Chen JM, Chuzhanova N, Stenson PD, Krawczak M, Ferec C, Cooper DN. Intrachromosomal serial replication slippage in trans gives rise to diverse genomic rearrangements involving inversions. Hum Mutat. 2005; 26:362–73.
7. Chuzhanova NA, Anassis EJ, Ball EV, Krawczak M, Cooper DN. Meta-analysis of indels causing human genetic disease: mechanisms of mutagenesis and the role of local DNA sequence complexity. Hum Mutat. 2003; 21:28–44.
8. Dignam JD, Lebovitz RM, Roeder RG. Accurate Transcription Initiation by Rna Polymerase-Ii in a Soluble Extract from Isolated Mammalian Nuclei. Nucleic Acids Res. 1983; 11(5):1475-89.
9. Dobrovolny R, Dvorakova L, Ledvinova J, Magage S, Bultas J, Lubanda JC, Elleder M, Karetova D, Pavlikova M, Hrebicek M. Relationship between X-inactivation and clinical involvement in Fabry heterozygotes. Eleven novel mutations in the alpha-galactosidase A gene in the Czech and Slovak population. J Mol Med. 2005; 83:647–54.
10. Doll RF and Smith FI. Regulation of expression of the gene encoding human acid beta-glucosidase in different cell types. Gene. 1993; 127:255–60.
11. Engel K, Nuoffer JM, Mühlhausen C, Klaus V, Largiadèr CR, Tsiakas K, Santer R, Wermuth B, Häberle J. Analysis of mRNA transcripts improves the success rate of molecular genetic testing in OTC deficiency. Mol Genet Metab. 2008; 94:292–7.
12. Eyal N, Wilder S, Horowitz M. Prevalent and rare mutations among Gaucher patients. Gene. 1990; 96:277–83.
13. Fang B, Mane-Padros D, Bolotin E, Jiang T, Sladek FM. Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. Nucleic Acids Res. 2012; 40:5343-56.
14. Hamano Y, Kodama H, Yanagisawa M, Haraguchi Y, Mori M, Yokota S. Immunocytochemical localization of ornithine transcarbamylase in rat intestinal mucosa. Light and electron microscopic study. J Histochem Cytochem. 1988; 36(1):29-35.
15. Hata A, Tsuzuki T, Shimada K, Takiguchi M, Mori M, Matsuda I. Structure of the human ornithine transcarbamylase gene. J Biochem. 1988; 103(2):302-8.
16. Horowitz M, Wilder S, Horowitz Z, Reiner O, Gelbart T, Beutler E. The human glucocerebrosidase gene and pseudogene: structure and evolution. Genomics. 1989; 4:87–96.
17. Horwich AL, Fenton WA, Williams KR, Kalousek F, Kraus JP, Doolittle RF, Konigsberg W, Rosenberg LE. Structure and expression of a complementary DNA for the nuclear coded precursor of human mitochondrial ornithine transcarbamylase. Science. 1984; 224(4653):1068-74.
18. Inoue Y, Hayhurst GP, Inoue J, Mori M, Gonzalez FJ. Defective ureagenesis in mice carrying a liver-specific disruption of hepatocyte nuclear factor 4alpha (HNF4alpha). J Biol Chem. 2002; 277(28):25257-65.
19. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010; 11(2):113-27.
20. Jones SN, Grompe M, Munir MI, Veres G, Craigen WJ, Caskey CT. Ectopic Correction of Ornithine Transcarbamylase Deficiency in Sparse Fur Mice. J Biol Chem. 1990; 265(24):14684-90.
21. Kimura A, Nishiyori A, Murakami T, Tsukamoto T, Hata S, Osumi T, Okamura R, Mori M, Takiguchi M. Chicken Ovalbumin Upstream Promoter-Transcription Factor (Coup-Tf) Represses Transcription from the Promoter of the Gene for Ornithine Transcarbamylase in a Manner Antagonistic to Hepatocyte Nuclear Factor-Iv (Hnf-4). J Biol Chem. 1993; 268(15):11125-33.
22. Krebs HA and Henseleit K. Untersuchungen uber die harnstoffbildung im tierkorper. Hoppe-Seyler's Z Physiol Chem. 1932; 210:325–32.
23. Latham TE, Theophilus BD, Grabowski GA, Smith FI. Heterogeneity of mutations in the acid beta-glucosidase gene of Gaucher disease patients. DNA Cell Biol. 1991; 10:15–21.
24. Lieber MR, Ma Y, Pannicke U, Schwarz K. Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol. 2003; 4:712–20.
25. Lindgren V, de Martinville B, Horwich AL, Rosenberg LE, Francke U. Human ornithine transcarbamylase locus mapped to band Xp21.1 near the Duchenne muscular dystrophy locus. Science. 1984; 226(4675):698-700.
26. Malik S, Karathanasis SK. TFIIB-directed transcriptional activation by the orphan nuclear receptor hepatocyte nuclear factor 4. Mol Cell Biol. 1996; 16:1824-31.
27. Marsden D and Levy H. Newborn Screening of Lysosomal Storage Disorders. Clin Chem. 2010, 56(7):1071-9
28. Matsuda I, Matsuura T, Nishiyori A, Komaki S, Hoshide R, Matsumoto T, Funakoshi M, Kiwaki K, Endo F, Hata A, Shimadzu M, Yoshino M.Matsuda I, Matsuura T, Nishiyori A. Phenotypic variability in male patients carrying the mutant ornithine transcarbamylase (OTC) allele, Arg40His, ranging from a child with an unfavourable prognosis to an asymptomatic older adult. J Med Genet. 1996; 33:645–8.
29. Meikle PJ, Hopwood JJ, Clague AE, Carey WF. Prevalence of lysosomal storage disorders. JAMA. 1999; 281:249-54

30. Metzger S, Halaas JL, Breslow JL, Sladek FM. Orphan Receptor Hnf-4 and Bzip Protein C/Ebp-Alpha Bind to Overlapping Regions of the Apolipoprotein-B Gene Promoter and Synergistically Activate Transcription. J Biol Chem. 1993; 268:16831-8.
31. Murakami T, Nishiyori A, Takiguchi M, Mori M. Promoter and 11-kilobase upstream enhancer elements responsible for hepatoma cell-specific expression of the rat ornithine transcarbamylase gene. Mol Cell Biol. 1990; 10(3):1180-91.
32. Murakami T, Takiguchi M, Inomoto T, Yamamura K, Mori M. Tissue- and developmental stage-specific expression of the rat ornithine carbamoyltransferase gene in transgenic mice. Dev Genet. 1989; 10(5):393-401.
33. Nishiyori A, Tashiro H, Kimura A, Akagi K, Yamamura K, Mori M, Takiguchi M. Determination of tissue specificity of the enhancer by combinatorial operation of tissue-enriched transcription factors. Both HNF-4 and C/EBP beta are required for liver-specific activity of the ornithine transcarbamylase enhancer. J Biol Chem. 1994; 269(2):1323-31.
34. Orphanides G, Lagrange T, Reinberg D. The general transcription factors of RNA polymerase II. Genes Dev. 1996; 10(21):2657-83.
35. Quental R, Azevedo L, Rubio V, Diogo L, Amorim A. Molecular mechanisms underlying large genomic deletions in ornithine transcarbamylase (OTC) gene. Clin Genet. 2009; 75:457–64.
36. Ryall J, Nguyen M, Bendayan M, Shore GC. Expression of nuclear genes encoding the urea cycle enzymes, carbamoyl-phosphate synthetase I and ornithine carbamoyl transferase, in rat liver and intestinal mucosa. Eur J Biochem. 1985; 152(2): 287-92.
37. Sabbah M, Kang KI, Tora L, Redeuilh G. Oestrogen receptor facilitates the formation of preinitiation complex assembly: Involvement of the general transcription factor TFIIB. Biochem J. 1998; 336:639-46.
38. Sardiello M, Palmieri M, di Ronza A, Medina DL, Valenza M, Gennarino VA, Di Malta C, Donaudy F, Embrione V, Polishchuk RS, Banfi S, Parenti G, Cattaneo E, Ballabio A. A gene network regulating lysosomal biogenesis and function. Science. 2009; 325(5939):473-7
39. Shi D, Morizono H, Yu X, Tong L, Allewell NM, Tuchman M. Human ornithine transcarbamylase: crystallographic insights into substrate recognition and conformational changes. Biochem J. 2001; 354:501-9.
40. Sonenberg N and Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009; 136(4):731-45.
41. Sorge J, West C, Westwood B, Beutler E. Molecular cloning and nucleotide sequence of human glucocerebrosidase cDNA. Proc Natl Acad Sci. 1985; 82:7289-93
42. Sugie K, Yamamoto A, Murayama K, Oh SJ, Takahashi M, Mora M, Riggs JE, Colomer J, Iturriaga C, Meloni A, Lamperti C, Saitoh S, Byrne E, DiMauro S, Nonaka I, Hirano M, Nishino I. Clinicopathological features of genetically confirmed Danon disease. Neurology. 2002; 58:1773-8
43. Takahashi H, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC. Proteomics Reveals a Physical and Functional Link between Hepatocyte Nuclear Factor 4 alpha and Transcription Factor IID. J Biol Chem. 2009; 284:32405-12.
44. Takiguchi M, Murakami T, Miura S, Mori M. Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-linked gene with an atypical promoter. Proc Natl Acad Sci. 1987; 84(17):6136-40.
45. Veres G, Craigen WJ, Caskey CT. The 5'-flanking region of the ornithine transcarbamylase gene contains DNA sequences regulating tissue-specific expression. J Biol Chem. 1986; 261(17):7588-91.
46. Wenger DA, Coppola S, Liu SL. Insights into the diagnosis and treatment of lysosomal storage diseases. Arch Neurol. 2003; 60:322-8.
47. Westbroek W, Gustafson AM, Sidransky E. Exploring the link between glucocerebrosidase mutations and parkinsonism. Trends Mol Med. 2011;17:485–93.
48. Wilcox WR. Lysosomal storage disorders: the need for better pediatric recognition and comprehensive care. J Pediatr. 2004; 144:S3-S14
49. Wilson DO, Johnson P, McCord BR. Nonradiochemical DNase I footprinting by capillary electrophoresis. Electrophoresis. 2001; 22(10):1979-86.
50. Woychik NA and Hampsey M. The RNA polymerase II machinery: structure illuminates function. Cell. 2002; 108(4):453-63.
51. Zianni M, Tessanne K, Merighi M, Laguna R, Tabita FR. Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. J Biomol Tech. 2006; 17(2):103-13.

7. List of publications

7.1. Publications related to this thesis

Ref I:

Storkanova G, Vlaskova H, Chuzhanova N, Zeman J, Stranecky V, Majer F, Peskova K, Luksan O, Jirsa M, Hrebicek M, Dvorakova L. Ornithine carbamoyltransferase deficiency: molecular characterization of 29 families. Clin Genet. 2013; 84(6):552-9. *IF(2012)=4.247*

Ref II:

Lukšan O, Jirsa M, Eberová J, Minks J, Trešlová H, Boučková M, Štorkánová G, Vlášková H, Hřebíček M, Dvořáková L. Disruption of OTC promoter-enhancer interaction in a patient with symptoms of ornithine carbamoyltransferase deficiency. Hum Mutat. 2010; 31(4):E1294-E1303. *IF=5.956*

Ref III:

Lukšan O, Dvořáková L, Jirsa M. HNF-4α regulates expression of human ornithin transcarbamylase through interaction with two positive *cis*-acting regulatory elements located in the proximal promoter. Fol Biol. 2014; 60:133-143. *IF(2012)=1.219*

Ref IV:

Svobodová E, Mrázová L, Lukšan O, Elstein D, Zimran A, Stolnaya L, Minks J, Eberová J, Dvořáková L, Jirsa M, Hřebíček M. Glucocerebrosidase gene has an alternative upstream promoter, which has features and expression characteristic of housekeeping genes. Blood Cells Mol Dis. 2011; 46(3):239-45. *IF=2.351*

7.2. Publications not related to this thesis

1. Neřoldová M, Fraňková S, Stránecký V, Honsová E, Lukšan O, Beneš M, Michalová K, Kmoch S, Jirsa M. Hereditary haemochromatosis caused by homozygous HJV mutation evolved through paternal disomy. Clin Genet. 2014 Feb 12. doi: 10.1111/cge.12346. *IF(2012)=4.247*
2. Subhanova I, Muchova L, Lenicek M, Vreman HJ, Luksan O, Kubickova K, Kreidlova M, Zima T, Vitek L, Urbanek P. Expression of Biliverdin Reductase A in peripheral blood leukocytes is associated with treatment response in HCV-infected patients. PLoS One. 2013; 8(3):e57555. *IF(2012)=3.730*
3. Sticova E, Elleder M, Hulkova H, Luksan O, Sauer M, Wunschova-Moudra I, Novotny J, Jirsa M. Dubin-Johnson syndrome coinciding with colon cancer and atherosclerosis. World J Gastroenterol. 2013; 19(6):946-50. *IF(2012)=2.547*
4. Fafilek B, Krausova M, Vojtechova M, Pospichalova V, Tumova L, Sloncova E, Huranova M, Stancikova J, Hlavata A, Svec J, Sedlacek R, Luksan O, Oliverius M, Voska L, Jirsa M, Paces J, Kolar M, Krivjanska M, Klimesova K, Tlaskalova-Hogenova H, Korinek V. Troy, a Tumor Necrosis Factor Receptor Family Member, Interacts With Lgr5 to Inhibit Wnt Signaling in Intestinal Stem Cells. Gastroenterology. 2012; 144(2):381-91. *IF=12.821*
5. Treepongkaruna S, Gaensan A, Pienvichit P, Luksan O, Knisely AS, Sornmayura P,Jirsa M. Novel ABCB11 mutations in a Thai infant with progressive familial intrahepatic cholestasis. World J Gastroenterol. 2009; 15(34):4339-42. *IF=2.092*