UNIVERZITA KARLOVA V PRAZE

1. LÉKAŘSKÁ FAKULTA
ÚSTAV DĚDIČNÝCH METABOLICKÝCH PORUCH

Studijní obor: **Biochemie a patobiochemie**

Autor: **Mgr. Lenka Nosková**

VYUŽITÍ NOVÝCH GENOMICKÝCH TECHNIK VE STUDIU PATOGENEZE VYBRANÝCH VZÁCNÝCH DĚDIČNÝCH ONEMOCNĚNÍ

APPLICATION OF NOVEL GENOMIC TECHNIQUES IN STUDIES OF PATHOGENESIS OF SELECTED RARE INHERITED DISORDERS

Dizertační práce

Vedoucí práce: **Ing. Stanislav Kmoch, CSc.**

Místo a rok vypracování: **Praha, 2013**

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem řádně uvedla a citovala všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 21. 1. 2013

Jméno – Příjmení (hůlkovým písmem)

Podpis

Identifikační záznam:

NOSKOVÁ, Lenka. *Využití nových genomických technik ve studiu patogeneze vybraných vzácných dědičných onemocnění. [Application of novel genomic techniques in studies of pathogenesis of selected rare inherited disorders.]* Praha, 2013. 149 s. Dizertační práce. Univerzita Karlova v Praze, 1. lékařská fakulta, Ústav dědičných metabolických poruch. Vedoucí práce Kmoch, Stanislav.

Abstrakt

Vzácná onemocnění jsou heterogenní skupinou onemocnění, jejichž molekulární podstata je často neznámá. Pro jejich studium nebyly donedávna z důvodů malého počtu pacientů k dispozici vyhovující metodické postupy. Díky novým přístupům ke studiu genomu, především díky technologiím DNA čipů a s rozmachem sekvenování nové generace, je v současné době možno studovat vzácná onemocnění i u jednotlivých rodin nebo u sporadických případů.

Tato dizertační práce se zabývá aplikací nových genomických technik při studiu vzácných dědičných chorob. Popisuje využití technologie DNA čipů pro účely vazebné analýzy, analýzy změn genové exprese, analýzy změn počtu kopií a homozygotního mapování a dále využití sekvenování nové generace. Kombinace těchto metodických postupů byly použity pro studium molekulární podstaty adultní formy neuronální ceroidní lipofuscinózy, Rotorova syndromu, izolovaného defektu ATP syntázy a mukopolysacharidózy typu IIIC.

Klíčová slova

vzácná onemocnění, technologie DNA čipů, exomové sekvenování, neuronální ceroidní lipofuscinóza, Rotorův syndrom, izolovaný defekt ATP syntézy, mukopolysacharidóza typu IIIC

Abstract

Rare diseases are a heterogeneous group of disorders. Knowledge of their molecular basis is poor and till recently there were no appropriate methodical approaches due to a limited number of patients. Novel genomic techniques, especially the DNA array technology and the next generation sequencing emerging in last few years, enabled studies of these diseases even in small families and sporadic cases.

This PhD thesis focuses on application of novel genomic techniques in studies of rare inherited diseases. It describes a use of DNA array technology in linkage analysis, analysis of differential gene expression, analysis of copy number variations and homozygous mapping, and a use of next generation sequencing technology. Combination of these methods was used for identification of molecular basis of adult neuronal ceroid lipofuscinosis, Rotor syndrome, isolated defect of ATP synthase and mucopolysaccharidosis type IIIC.

Key words

rare diseases, DNA array technology, exome sequencing, neuronal ceroid lipofuscinosis, Rotor syndrome, isolated defect of ATP synthase, mucopolysaccharidosis type IIIC

Poděkování

Výsledky prezentované v této práci by nikdy nemohly být výsledkem práce jednoho člověka. Na tomto místě bych chtěla poděkovat celému kolektivu Ústavu dědičných metabolických poruch, zvláště mému školiteli Standovi Kmochovi; díky němu vznikají v ústavu mimořádné projekty a já si velmi vážím toho, že mám to štěstí být součástí jeho týmu. Dále děkuji svým spolupracovníkům – Haně Hartmannové, Viktorovi Stráneckému, Anně Přistoupilové, Lence Piherové, Katce Hodaňové, Petrovi Vyleťalovi, Veronice Barešové, Robertu Ivánkovi, Aleně Čížkové, Heleně Hůlkové, Janě Sovové a dalším, s kterými jsme se pokoušeli a pokoušíme vystopovat příčiny vzácných onemocnění; práce s nimi je po pracovní i osobní stránce nesmírně obohacující. Zásadní roli v této práci hrál profesor Milan Elleder, který založil a léta vedl Ústav dědičných metabolických poruch, vytvořil v něm unikátní tvůrčí prostředí, ve kterém podporoval studenty, vedl je ke kritickému myšlení, dával výzkumným skupinám obrovskou míru tvůrčí svobody, vedl je neformálně a moudře.

Finanční podporu pro projekty zmíněné v této práci poskytly následující grantové agentury a granty: grantová agentura České republiky: 303/03/H065, 303/07/0781, 305/08/H037, grantová agentura Ministerstva zdravotnictví ČR: NR8069-3, NR8069-1, 1A/8239-3, grantová agentura Univerzity Karlovy: 54/20320827/05, 250051 a dále výzkumné projekty Ministerstva školství, mládeže a tělovýchovy MSM0021620806, AV0Z50110509 a 1M6837805002.

Za podporu a motivaci děkuji mému partnerovi Rudovi a mé rodině, hnacím motorem k dopsání této práce bylo také narození mé dcery Žofie. Díky patří mým přátelům a všem karlínským aikidókům, díky kterým se mi více či méně dařilo hledat harmonii a vnitřní klid v práci i v osobním životě.

Obsah

Seznam zkratek

AD	Alzheimer disease, Alzheimerova choroba
ANCL	adultní forma neuronální ceroidní lipofuscinózy
APP	amyloid prekurzor protein
APOA1	apolipoprotein A1
ATP	adenosin trinukleotid fosfát
bp	páry bazí
CAD5	Cath.a diferencované buňky
cDNA	komplementární DNA
CGH	komparativní genomová hybridizace
CNPs	copy number polymorphisms, polymorfismy v počtu kopií
CNVs	copy number variations, varianty v počtu kopií
CSPα	cysteine string protein alfa
CTSD	cathepsin D
Cy3, Cy5	cyanin 3, cyanin 5
DJS	Dubin Johnsonův syndrom
DNA	deoxyribonukleová kyselina
EGFP	enhanced green fluorescent protein
FRET	Fluorescent Resonance Energy Transfer
gDNA	genomová DNA
GROD	granulární osmiofilní depozit
GWAS	genome wide association studies, celogenomové asociační studie
HDL	high density lipoprotein, lipoprotein s vysokou hustotou
Indel	inzerce/delece
LOD	logaritmus poměru pravděpodobnosti rekombinace
Mb	megabáze
MIM	Mendelian Inheritance in Man
miRNA	mikroRNA
MPSIIIC	mukopolysacharidóza typu IIIC
mtDNA	mitochondriální DNA
NCL	neuronální ceroidní lipofuscinózy
NORD	National Organisation of Rare Disorders
OMIM	Online Mendelian Inheritance in Man
PCR	polymerázová řetězová reakce
PPT1	palmitoyl protein thioesteráza
PSEN1,2	presenilin 1,2
RFLP	polymorfismy v délce restrikčních fragmentů
RNA	ribonukleotidová kyselina
RS	Rotorův syndrom
SNPs	single nukleotide polymorfisms, jednonukleotidové záměny
STRs	short tandem repeal, krátké tandemové repetice
TPP1	tripeptidylpeptidáza 1
ÚDMP	Ústav dědičných metabolických poruch
UGT1A1	uridin difosfát glukuronosyl transferáza A1
VFN	Všeobecná fakultní nemocnice
WHO	World Health Organization, Světová zdravotnická organizace
wt	wild type

ČÁST I. Vzácné choroby a přístupy k jejich studiu

Úvod

Tato dizertační práce se zabývá studiem několika vybraných vzácných dědičně podmíněných onemocnění. Pro jejich studium byly použity genomické techniky založené na celogenomové analýze pomocí technologie DNA čipů a na nových sekvenačních metodách. V Ústavu dědičných metabolických poruch, kde tato práce vznikla, jsme prošli několika fázemi vývoje těchto technik. Na počátku jsme sami navrhovali a vyráběli DNA čipy vhodné právě pro námi studované choroby. Později byla dostupná stále se rozšiřující paleta komerčně dostupných čipů, u kterých laboratorní a základní analytická část probíhala v servisních laboratořích. V poslední době je technologie DNA čipů částečně nahrazována celoexomovým sekvenováním pomocí sekvenátorů nové generace.

Ve své dizertační práci bych chtěla prezentovat využitelnost těchto technik na příkladech několika vzácných chorob, kterými jsem se během svého studia zabývala. Zároveň by práce měla ukázat, jak se díky rychlému technologickému vývoji mění přístupy k jejich studiu a jak lze v závislosti na typu dědičnosti a velikosti studovaného souboru využívat kombinace právě těch metod, které vedou nejrychleji a nejefektivněji k cíli – tedy k poznání genetické podstaty studované choroby.

Vzhledem k spíše metodickému zaměření této práce jsou zde prezentovány výsledky studia různých vzácných onemocnění – mitochondriálních, lysozomálních střádavých, neuropsychiatrických a hepatologických, které mají různý typ dědičnosti. Hlavním smyslem však není popis jednotlivých chorob, ale variabilní metodologický přístup závisející na povaze a dostupnosti vzorků pro studium.

Vzácná onemocnění - definice a současný stav poznání

Vzácná onemocnění jsou definována jako taková onemocnění, jejichž prevalence je v Evropě menší než 1:2000 a v USA menší než 1:1250 (Remuzzi and Garattini 2008). Většina onemocnění zahrnutá do této skupiny je však mnohem vzácnější, například Huntingtonova choroba (7 : 100 000), achondroplázie (4,5 : 100 000) nebo Fabryho choroba (1,75 : 100 000).[1] Danou vzácnou chorobou často trpí jen desítky nebo stovky pacientů na celém světě. Například incidence dědičných metabolických poruch coby celé skupiny onemocnění se v naší republice udává mezi 1:1000 až 1:600, u konkrétních chorob se pak pohybuje mezi 1: 10 000 až 1: 1 000 000 (Stastna et al. 2010). Často je na našem území jen jedna či dvě rodiny trpící danou metabolickou poruchou. Jednotlivé vzácné varianty však nemusí být vzácné v izolovaných etnických skupinách. Příkladem může být vyšší prevalence Gaucherovy choroby typu I u aškenázských židů (Vallance and Ford 2003) nebo nesyndromické hluchoty u izolovaných skupin arabských muslimů ze severního Izraele (Zlotogora 2007).

Světová zdravotnická organizace (World Health Organization, WHO) zahrnuje do kategorie vzácných chorob více než 5000 onemocnění, které dohromady postihují asi 30 milionů jedinců v Evropské unii a 25 milionů jedinců v USA (Schieppati et al. 2008, Remuzzi and Garattini 2008).

Vzácná onemocnění jsou velmi heterogenní skupinou nemocí postihující různé orgány a mající nejrůznější klinické projevy. Jejich důsledkem je často výrazné zhoršení kvality života jedinců i jejich rodin. Znalosti o molekulární podstatě jednotlivých chorob - a z nich plynoucí případné možnosti diagnostiky a následné léčby - jsou velmi malé. Pacienti tak získávají jen málo informací o chorobě, kterou trpí, nebo by trpět mohli, a často podstupují řadu náročných vyšetření ve snaze o zpřesnění diagnózy.

Financování studia vzácných onemocnění naráží na mnohé limitace. Hlavními argumenty proti investicím do tohoto typu výzkumu jsou neúměrné náklady a složitost výzkumu ve srovnání s malým ekonomickým ziskem, kterého lze dosáhnout při případném úspěchu s dopadem pouze na velmi malou cílovou skupinu (Senior 1999, Remuzzi and Schieppati 2011). Z těchto důvodů jsou proto upřednostňovány projekty zabývající se spíše populačně častými multifaktoriálními komplexními onemocněními (Antonarakis and Beckmann 2006).

Často jsou vzácná onemocnění označována za sirotky („orphan diseases“) – vědomosti o jejich příčinách a možnostech efektivní terapie jsou prakticky nulové. Přitom je zřejmé, že tato onemocnění postihují funkčně významné esenciální geny propojené do interakčních sítí a jejich budoucí studium přinese řadu praktických aplikací (Zhang et al. 2011).

Lze najít mnoho důvodů, proč studovat vzácná onemocnění.

---

1 http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_decreasing_prevalence_or_cases.pdf

V rovině sociopolitické, kulturní a etické je studium vzácných onemocnění podloženo právem pacientů a jejich rodin na to, profitovat z možností rozvoje biotechnologií a metodologického pokroku, právem na získání vědomostí o své chorobě a právem na diagnózu ([Anonymous] 2008). Společenský tlak na studium vzácných onemocnění byl v USA roku 1983 příčinou vzniku národní organizace pro vzácná onemocnění (National Organisation of Rare Disorders, NORD) zastřešující v současné době více než 2000 organizací sdružujících pacienty s danou vzácnou nemocí. V roce 1986 byla založena Genetické aliance (Genetic Alliance), jejímž cílem je podpora pacientských komunit a podpora spolupráce výzkumných organizací, průmyslu a veřejného sektoru. V Evropě byla po vzoru organizace NORD vytvořena v roce 1997 aliance pacientských sdružení Eurordis (Schieppati et al. 2008, Ayme, Kole and Groft 2008). V Japonsku byla podobná iniciativa založena dokonce již v roce 1972 pod názvem „National Programme on Rare and Intractable Diseases“ (Hayashi and Umeda 2008).

V rovině základního výzkumu přispívá studium vzácných onemocnění významně k pochopení základní biologie. Definuje kauzální geny a může vysvětlit jejich biologické funkce, které dosud u většiny genů nejsou známy. Databáze OMIM, Online Mendelian Inheritance in Man, shromažďující údaje o chorobách s genetickou komponentou a provádějící jejich katalogizaci, popisuje k 18. lednu 2013 pouze 3 683 fenotypů s vysvětlenou molekulární podstatou z celkového počtu 21 600 záznamů (viz tabulka 1). Studium monogenních vzácných onemocnění proto může přinést ještě mnoho informací, které v databázích stále chybí.

Tab. 1 Počet záznamů v databázi OMIM k 18. lednu 2013

	Autozomální	X-vázané	Y-vázané	Mitochondriální	Celkem
Geny se známou sekvencí	13 396	652	48	35	14 131
Geny se známou sekvencí a fenotypem	124	4	0	2	130
Fenotyp se známou molekulární podstatou	3 380	271	4	28	3 683
Fenotyp nebo lokus s neznámou molekulární podstatou	1 628	133	5	0	1 766
Ostatní, především fenotypy s pravděpodobnou mendelovskou dědičností	1 763	125	2	0	1 890
Celkem	20 291	1 185	59	65	21 600

zdroj: http://omim.org/statistics/entry

Vzhledem k tomu, že vzácné nemoci jsou v naprosté většině monogenními onemocněními, jsou pacienti s daným defektem unikátními modely pro pochopení základních fyziologických a patofyziologických mechanismů v lidských buňkách a tkáních. Díky studiu vzácných monogenních onemocnění byly popsány některé nové genetické mechanismy jako například uniparentální disomie

(SPENCE et al. 1988) nebo expanze tripletů (Pearson, Edamura and Cleary 2005). Popsané defekty pak mohou být studovány v izolovaných buněčných modelech, které dále umožní pochopení základních mechanismů. I přes existenci nespočetných myších i jiných zvířecích modelů je člověk sám nejlepším modelem, neboť jeho fenotyp je velmi podrobně prostudován. Ne vždy existují zvířecí ortology lidských genů, klinicky relevantní fenotypy mohou mít také projevy specifické pouze pro člověka (mentální retardace, změny kognitivních funkcí) (Antonarakis and Beckmann 2006).

Mnoho nových poznatků může přinést také studium vzácných monogenních onemocnění u izolovaných populací nebo populací s vysokou mírou příbuzenského křížení, kde mohou segregovat mnohé vzácné recesivní monogenní fenotypy. Příspěvek těchto populací k biologickému poznání může být jistě nemalý, neboť se odhaduje, že podíl příbuzenských sňatků tvoří celosvětově 20 – 50 % (Bittles 2001).

Definice kauzálních genů může dále přispět i ke studiu komplexních onemocnění, neboť přispívá k pochopení širšího kontextu metabolických, signálních a regulačních drah. Například pohled na patogenezi Alzheimerovy choroby (AD) je založen na popisu mutací v genech *APP* (amyloid-prekurzor protein), *PSEN1* (presenilin 1) a *PSEN2* (presenilin 2) u familiárních forem AD, díky kterým byla formulována hypotéza centrální role β-amyloid proteinu. Mutace v uvedených genech vysvětlují jen zhruba 1 % všech případů AD, studium role mutovaných proteinů může však významně přispět k pochopení molekulární patogeneze AD (Bekris et al. 2010). Lepšímu pochopení obecných molekulárních mechanismů neurologických a psychiatrických onemocnění mohou obdobně napomoci výsledky studia vzácných nesyndromických mentálních retardací (Caliskan et al. 2011).

Informace získané studiem vzácných onemocnění mohou dále umožnit popis efektu různých mutací v jednom genu na fenotypovou a klinickou variabilitu onemocnění a umožňuje definovat genetické faktory a modifikující geny. Příkladem je gen *ABCA1* (ATP-binding cassette transporter 1), u kterého byly nalezeny mutace u pacientů s Tangierovou chorobou projevující se nízkou hladinou plazmatického HDL cholesterolu a jeho depozicí (Bodzioch et al. 1999, Brooks-Wilson et al. 1999); varianty v tomto genu modifikují hladinu HDL cholesterolu i v běžné populaci (Frikke-Schmidt et al. 2004). Jiným příkladem je diabetes druhého typu, kde byly asociačními studiemi identifikovány polymorfismy v genech, u nichž byly popsány mutace i u vzácných variant monogenních forem diabetu (Sandhu et al. 2007, Winckler et al. 2007).

Dalším důsledkem studia vzácných onemocnění je možnost využití znalostí vyplývajících z popisu molekulární podstaty vzácných onemocnění při hledání nových způsobů léčby. Identifikace metabolických, regulačních nebo signálních drah ovlivněných danou nemocí umožňuje cíleně vyhledávat inhibitory, regulátory nebo katalyzátory příslušných proteinů.

Studium vzácných onemocnění je cestou k správné diagnóze, cílené léčbě a prevenci jak vzácných chorob, tak i chorob dalších, neboť poznatky na úrovni základního fungovaní dějů v buňce mohou přispět i pochopení podstaty dalších fyziologických i patofyziologických jevů.

Metodické přístupy ke studiu vzácných onemocnění

V posledním desetiletí se díky technologickému pokroku značně rozšířily možnosti studia molekulárních mechanismů a buněčných dějů. Nové technologie umožňují studovat nejen izolované jevy a procesy, ale postihují také komplexitu v rozsahu celého genomu, exomu, transkriptomu, metabolomu… Mezi takové technologie patří především metody DNA čipů a metody celogenomového, respektive celoexomového sekvenování.

Následující kapitola obsahuje přehled a princip těchto metod a specifikuje jejich využitelnost pro studium vzácných onemocnění.

Technologie DNA čipů a její využití pro studium vzácných chorob

DNA čipy (DNA microarrays) jsou analytickým nástrojem umožňujícím simultánní detekci tisíců a desítek tisíc fragmentů DNA najednou. Jejich základním principem je hybridizace fluorescenčně značených komplementárních řetězců fragmentů DNA na imobilizované proby daného čipu, jejich fluorescenční detekce a následná analýza.

Prvopočátky technologie DNA čipů se datují do devadesátých let 20. Století. První praktické využití bylo publikováno v roce 1995 pro studium expresní analýzy 45 genů u *Arabidopsis thaliana* (SCHENA et al. 1995). Rozmach metody přišel s dokončením sekvenace lidského genomu vydaného Human Genome Sequencing Consortium (Consortium 2004, Lander et al. 2001), (http://genome.wellcome.ac.uk), s rozšířením technologických možností výroby DNA čipů a se vstupem komerčních firem zabývajících se vývojem a prodejem těchto čipů na trh.

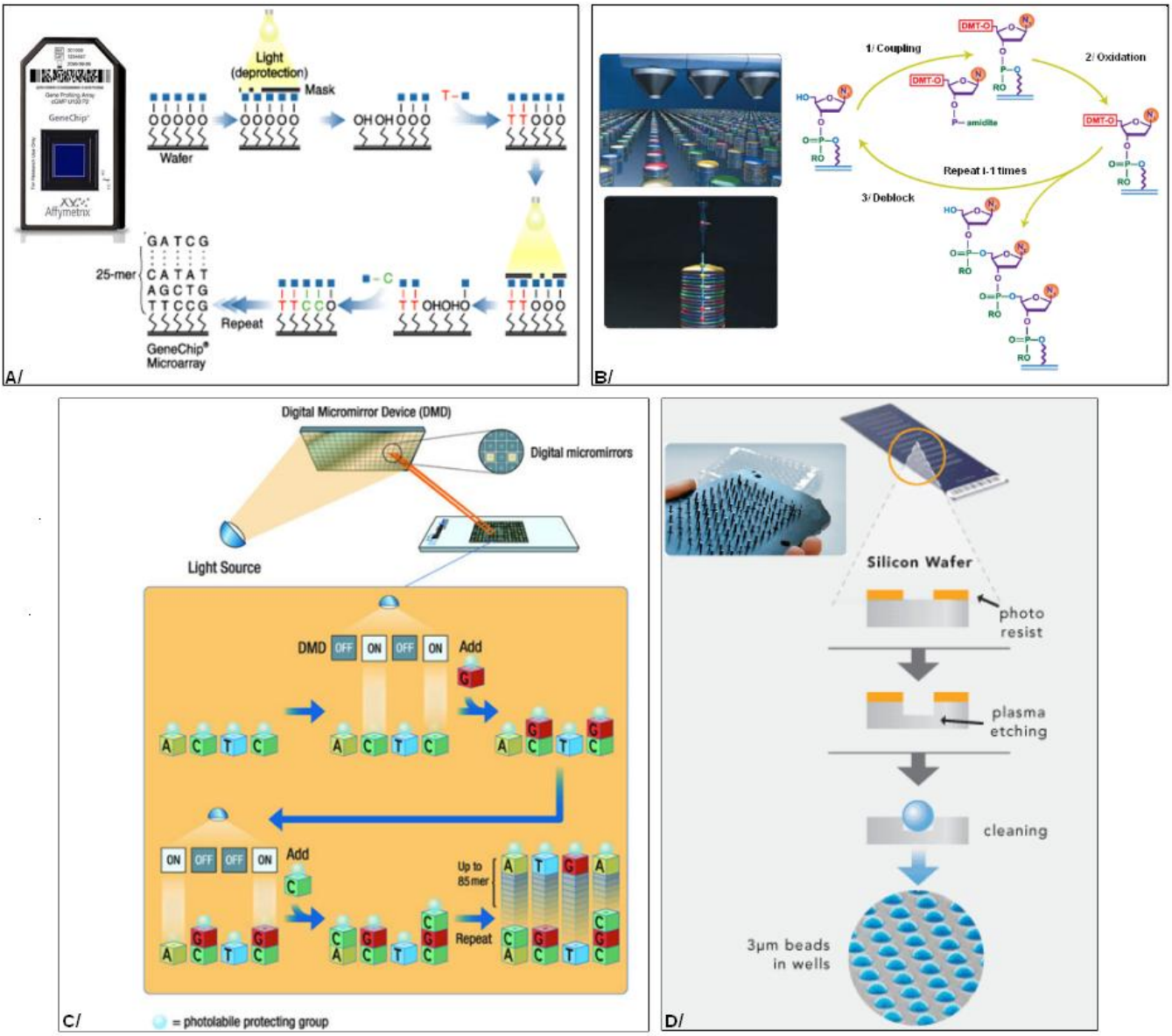
V současné době jsou používány tři technologie výroby DNA čipů a imobilizace prob – *in situ* syntéza, imobilizace na silikátové kuličky a tištěné čipy.

Při *in* situ syntéze jsou na povrchu DNA čipu postupně imobilizovány jednotlivé báze za použití systémů deprotekce funkčních skupin a systémů masek (Affymetrix (obr. 1A)), mikrozrcátek (Roche NimbleGen (obr. 1C)), nebo nekontaktních tiskových hlav (Agilent Technologies (obr.1B)).

Jinou technologií je imobilizace na silikátové kuličky (Illumina (obr. 1D)), při které dochází k nejprve k náhodnému umístění prob a poté k jejich následnému dekódování při sekvenční hybridizaci (Gunderson et al. 2004).

Obě výše zmíněné technologie umožňují výrobu vysokodenzitních čipů obsahujících miliony jednotlivých prob a umožňují tak studium na celogenomové úrovni.

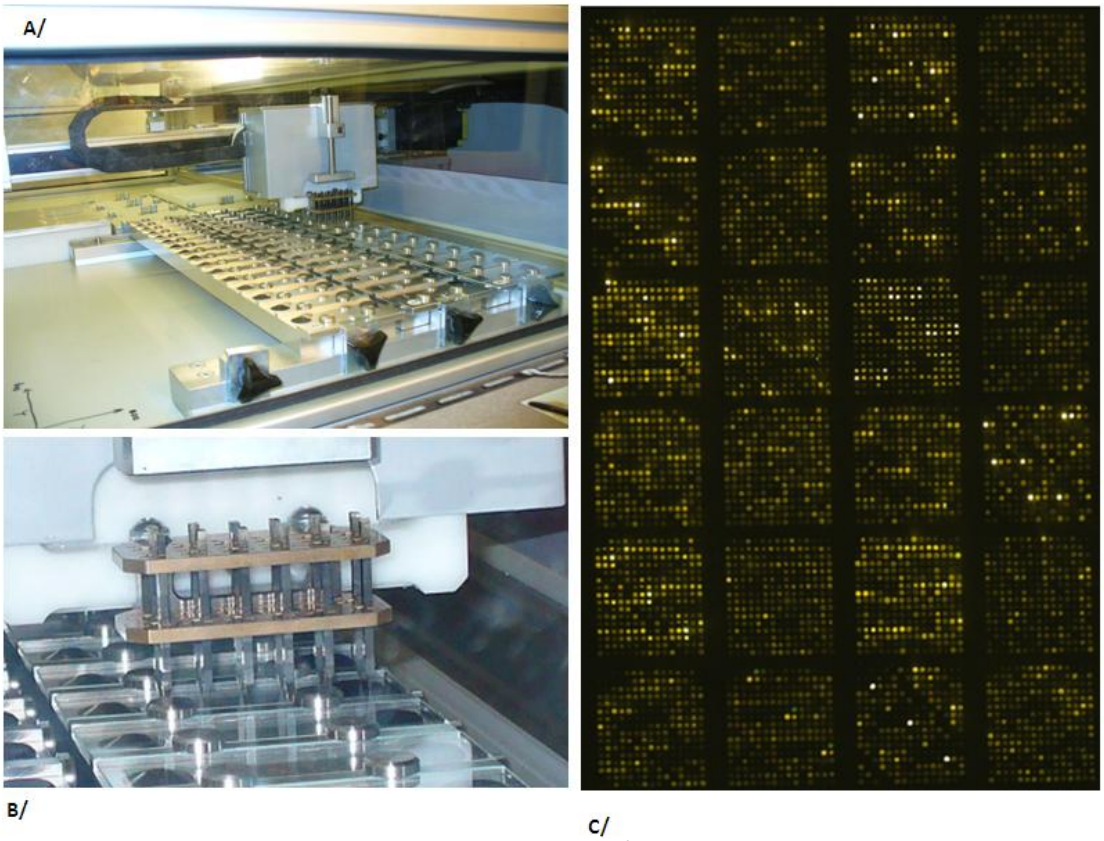
U tištěných čipů jsou proby (oligonukleotidy, dříve cDNA knihovny) nanášeny na povrch podložního skla pomocí robotického stroje. Tento přístup byl v minulosti používán řadou laboratoří při výrobě „custom" DNA čipů - tedy čipů „šitých na míru“ jednotlivým výzkumným projektům. Jejich výhodou byla možnost výběru dané sady genů a nižší finanční náročnost.



Obr. 1 Technologie výroby komerčních DNA čipů
A/ fotolitografická syntéza se světelnou deprotekcí za použití systému masek (Affymetrix) – modifikováno podle www. affymetrix.com
B/ syntéza nekontaktní tiskovou hlavou (Agilent Technologies) – modifikováno podle (http://www.genomics.agilent.com/GenericB.aspx?PageType=Custom&SubPageType=Custom&PageID=2011)
C/ syntéza pomocí systému mikrozrcátek bez použití masek (Roche Nimblegen) – modifikováno podle http://www.nimblegen.com/company/technology/index.html
D/Imobilizace na silikátové kuličky (technologie BeadArray) (Illumina)- modifikováno podle http://www.illumina.com/technology/beadarray_technology.ilmn

V Ústavu dědičných metabolických poruch byl v počátcích rozvoje technologie DNA čipů ve spolupráci s firmou GeneAge Technologies (http://www.geneagetech.com/) navržen, vyroben a testován prototyp „DNA spotteru“, přístroje pro přípravu DNA čipů. Jednalo se o robotické nanášecí zařízení – GeneSurfer, který umožňoval připravit v jednom cyklu až 60 čipů (podložních mikroskopických skel) o maximální hustotě 4900 prob/cm^2, tedy až 80 000 prob na efektivní plochu (viz obrázek 2A,B). Pro přípravu čipů byly optimalizovány metody imobilizace na povrch, metody přípravy fluorescenčně značených vzorků a metody obrazové analýzy a statistického zpracování.

DNA čipy vlastní výroby byly v Ústavu dědičných metabolických poruch využity v rámci několika projektů. Projekt h-MitoArray (viz obrázek 2C) byl určen pro studium exprese genů u mitochondriálních chorob. Pomocí tohoto čipu byla studována genová exprese u pacientů s izolovaným deficitem ATP syntázy, u kterých byla následně identifikována mutace v genu *TMEM70* (Cizkova et al. 2008a, Cizkova et al. 2008b); viz publikace 5 a 6. Dalším čipem připraveným na ÚDMP byl expresní čip navržený pro studium lysozomálního střádavého onemocnění – mukopolysacharidózy typu IIIC (MPSIIIC) (Hrebicek et al. 2006); viz publikace 7. DNA čipy byly na ÚDMP použity i pro metodu komparativní genomové hybridizace, a to při studiu molekulární podstaty Rotorova syndromu (Hrebicek et al. 2007); viz publikace 3.



Obr. 2 Technologie výroby tištěných „custom“ DNA čipů.
A/ Robotické zařízení Genesurfer využívané na UDMP pro výrobu „custom“ DNA čipů
B/ Genesurfer - detail tiskové hlavy
C/ h-Mitoarray využívaná pro studium diferenciální genové exprese u pacientů s mitochondriálními chorobami.

Možnosti aplikace technologie DNA čipů jsou v současné době velmi bohaté. Expresní čipy umožňují detekci poměrného množství mRNA v daném vzorku či vzorcích. U genotypovacích čipů lze odečítat přítomnost jednonukleotidových polymorfismů (SNPs) a variant v počtu kopií (CNVs). Čipy určené pro komparativní genomovou hybridizaci se využívají pro detekci velkých genomových změn. Komerční čipy v sobě často kombinují několik možných typů analýz (detekce SNPs, CNVs i rozsáhlých inzercí a delecí). Dnes jsou dostupné také čipy pro chromatinovou precipitaci, epigenetická studia, analýzu miRNA, promotorovou analýzu a proteinové čipy.

Pro studium vzácných chorob jsou využívány především genotypovací čipy. Díky nim lze efektivně provádět vazebné analýzy v postižených rodinách, detekovat malé inzerce a delece a identifikovat homozygotní oblasti. Využívají se také čipy pro komparativní genomovou hybridizaci (CGH), které umožňují vyšetřit jedince na přítomnost rozsáhlých inzercí a delecí. Expresní čipy mohou hrát pomocnou roli při prioritizaci kandidátních genů z oblastí vymezených vazebnými studiemi.

Vazebná analýza

Principem vazebné analýzy je sledování segregace polymorfních markerů s fenotypem v rodinách se studovanou dědičnou chorobou. Nejprve je provedeno genotypování informativních markerů podél celého genomu u co největšího počtu rodinných příslušníků s cílem nalézt lokusy, které jsou statisticky významně děděny pouze u postižených jedinců. Genotypovací data jsou pak analyzována s ohledem na typ dědičnosti, penetranci onemocnění, frekvenci alely v populaci a předpokládaný vliv vnějšího prostředí. Výsledkem analýzy je identifikace kandidátních oblastí, které by měly obsahovat kauzální gen (y).

V současné době se pro vazebnou analýzu úspěšně využívají genotypovací čipy, které obsahují desetitisíce až statisíce genetických markerů. Jedná se především o jednonukleotidové záměny (SNPs). Tyto genotypovací čipy umožňují zároveň detekovat malé genomové duplikace a delece v řádu několika kilobází.

Genotypovací čipy fungují na principu alelově specifické hybridizace (Affymetrix) nebo jsou založeny na metodě jednobázové extenze (Illumina).

Při alelově specifické hybridizaci se pro každý SNP užívají proby s přesnou sekvencí a se záměnami v sekvenci proby (perfect match a mismatch proby) (Wang et al. 1998). Software pak hodnotí intenzitu signálu u obou těchto typů prob a vypočítává genotypy pro jednotlivé SNP. Pro vazebnou analýzu se často využívají Affymetrix GeneChip® Human Mapping 10K Array; poskytují informace o genotypech přibližně 10 000 jednonukleotidových záměn, což umožňuje definovat vazebné oblasti s přesností na jednotky megabazí. Pro podrobnější genotypování vybraných jedinců je hojně využívána Affymetrix Genome-Wide Human SNP Array 6.0 obsahující více než 900 000 jednonukleotidových záměn, zároveň lze pomocí tohoto čipu detekovat více než 946 000 variant v počtu kopií.

Metoda jednobázové extenze je založena na hybridizaci vzorku na specifickou probu a následné inkorporaci fluorescenčně značeného dideoxynukleotidu pomocí polymerázy. (Steemers et al. 2006). Alternativně je detekována ligace mezi lokus specifickými oligonukleotidy a alelově specifickými oligonukleotidy, kdy ligační produkt slouží jako templát pro následné PCR (Illumina GoldenGate assay)(Fan et al. 2003). Pro vazebnou analýzu jsou v systému Illumina dostupné systémy

HumanLinkage V Panel Set obsahující přes 6 000 SNPs nebo Omni Whole-Genome Arrays pro podrobnější analýzy.

Vazebná analýza je vhodná především pro identifikaci kandidátních oblastí u velkých vícegeneračních rodin s autozomálně dominantní dědičností. Pro takovou analýzu je vhodné genotypovat alespoň 6-12 postižených jedinců a co nejvíce zdravých příbuzných; dostatečný počet jedinců pro vazebnou analýzu se tak stává nejčastější překážkou pro smysluplné využití této metody (Kuhlenbaumer, Hullmann and Appenzeller 2011). I menší počet vyšetřených jedinců a z toho plynoucí větší počet rozsáhlejších kandidátních oblastí však může napomoci při prioritizaci kandidátních genů, pokud je metoda kombinována s dalšími přístupy (analýza genové exprese, exomové sekvenování).

Genotypovací čipy s vyšší hustotou genetických markerů se využívají při populačních studiích a celogenomových asociačních studiích (genome-wide associaton studies, GWAS). Základem využití asociačních studií je představa, že některé běžné polymorfní markery mohou být asociovány s výskytem komplexních chorob. Tento přístup aplikovaný v posledních letech na řadu komplexních onemocnění však dosud objasnil jen malé procento jejich genetické komponenty (Manolio et al. 2009).

Analýza diferenciální genové exprese

Základní myšlenkou při aplikaci metody analýzy genové exprese je porovnání hladiny exprese transkriptů mezi dvěma či více stavy (například poměr exprese genů u pacienta v porovnání s kontrolou, porovnání exprese v tumorové a normální tkáni). Typicky je studovaným materiálem RNA, z které je reverzní transkripcí připravena cDNA. Ta je dále amplifikována a fluorescenčně naznačena. Vzorky, které jsou porovnávány, mohou být hybridizovány na jeden expresní čip; pak jsou značeny dvěma fluorofory - nejčastěji Cy3 a Cy5 (Agilent Technologies, Roche Nimblegen, tištěné čipy), nebo je každý vzorek hybridizován na vlastní expresní čip (Affymetrix, Illumina).

V současné době obsahují komerčně dostupné expresní čipy proby pro transkripty všech známých jaderně i mitochondriálně kódovaných genů, někdy jsou zahrnuty i nekódující RNA. Proby jsou aktualizovány podle informací z databází. Celogenomové expresní čipy obsahují proby pro 24 0000 – 30 000 transkriptů. Jsou využívány především pro popis molekulárních změn spojených se specifickými biologickými podmínkami, umožňují popsat funkční a systémové změny na úrovni buňky.

Expresních čipů lze pro studium vzácných onemocnění využít při prioritizaci kandidátních genů. Předpokládá se přitom, že mutace v daném genu může způsobit snížení množství či naprostou absenci transkriptu. Množství transkriptu může být i zvýšené jako důsledek snahy kompenzovat změněnou či chybějící funkci mutovaného proteinu.

Expresní analýza pomocí DNA čipů je čím dál více nahrazována novými sekvenačními metodami, především sekvenováním RNA. Díky němu lze identifikovat a kvantifikovat i vzácné transkripty bez předchozí znalosti jejich sekvencí a transkripčních variant (Metzker 2010).

Analýza změn počtu kopií a homozygotní mapování

Změny v počtu kopií (copy number variations, CNVs) jsou považovány za nejvýznamnější zdroj genetické variability u člověka. Vykazuje je zhruba 12% lidského genomu. Jedná se o delece, inzerce nových sekvencí nebo mobilních elementů, tandemové či segmentové duplikace, inverze nebo translokace (Alkan, Coe and Eichler 2011). Vznikají *de-novo*, často postihují funkčně neutrální oblasti; ty jsou v databázích uváděny jako populačně časté polymorfismy v počtu kopií (copy number polymorphisms, CNPs). Některé CNVs mohou postihovat i funkčně významné oblasti a být tak příčinou vzácných, ale i populačně častých onemocnění. Geny lokalizované v CNV oblastech pak slouží jako kandidáti pro studium daných onemocnění. Informace o CNVs jsou shromažďovány v databázích genomických variant, jednou z nejpoužívanějších databází je Toronto Database of Genomic Variants (http://projects.tcag.ca/variation/). Ta v současné době obsahuje přes 66 000 CNVs. Pro vyloučení, případně potvrzení významu nalezených CNV je vždy nutná analýza rodičů.

Změny počtu kopií lze detekovat buď na CGH čipech nebo na genotypovacích SNP čipech s vysokým pokrytím. Dostupné CGH čipy obsahují přes dva miliony prob na čip (Agilent SurePrint microarrays, Roche NimbleGen CGH whole-genome tiling arrays), jsou často využívané na klinických pracovištích a jsou schopny nahradit analýzu karyotypu. SNP čipy umožňují detekovat až milion nepolymorfních strukturních změn (Affymetrix Genome-Wide Human SNP Array, Illumina Infinium HD BeadChips).

Data získaná z CGH a SNP čipů lze s úspěchem využívat také pro homozygotní mapování. Tato metoda se s úspěchem využívá především pro autozomálně recesivní onemocnění u rodin s příbuzenským křížením (Collin et al. 2010, Iseri et al. 2010). Principem je hledání homozygotních oblastí v genomu, kdy postižené děti dědí stejnou homozygotní záměnu od obou rodičů.

Sekvenování nové generace a jeho využití pro studium vzácných chorob

Při studiu extrémně vzácných onemocnění se často stává, že je k dispozici pouze malé množství probandů v malých rodinách, či se jedná o nepříbuzné rodiny nebo o sporadické případy. V těchto případech nelze efektivně použít ani metodu vazebné analýzy (nedostatečný počet informativních jedinců v rodině), ani homozygotního mapování (nepříbuznost, nelze předpokládat existenci genového defektu lokalizovaného ve společně sdílené homozygotní oblasti). Genetická podstata takových chorob zůstávala až donedávna neznáma.

V posledních pěti letech se díky biotechnologickému pokroku začaly běžněji využívat také metody genomového, případně exomového sekvenování pomocí sekvenátorů nové generace (next-gen sequencing). Díky tomuto přístupu byly v posledních letech odhaleny kauzální mutace u řady vzácných chorob, například u Kabukiho syndromu (Ng et al. 2010), Fowlerova syndromu (Lalonde et al. 2010) nebo onemocnění Charcot-Marie-Tooth (Lupski et al. 2010).

V principu je možné sekvenovat celý lidský genom a zdá se, že je to směr, kam se budou technologie v příštích letech ubírat. Celogenomové sekvenování totiž umožňuje detekovat i ty strukturní varianty, které zůstávají „neviditelné“ pro běžné celoexomové sekvenování a CGH čipy, tedy inverze, translokace a komplexnější přestavby genomu. Tento přístup je však zatím využíván minoritně, a to jednak z důvodu časové a finanční nákladnosti, jednak kvůli zatím nedostačujícím možnostem bioinformatické analýzy. Vzhledem k tomu, že většina kauzálních mutací byla dosud nalezena v kódujících oblastech DNA a v přilehlých intronových oblastech, využívá se pro hledání kauzálních genů sekvenování lidského exomu. Jeho velikost je přibližně 30 Mb a tvoří tak asi 1% celého genomu. Sekvenování je tedy rychlejší a pro analýzu postačuje mnohem méně sekvenačních dat.

Exomové sekvenování zahrnuje přípravu knihovny genomové DNA, provedení cíleného obohacení a výběru fragmentů obsahujících kódující oblasti, specifickou amplifikaci a samotnou sekvenční reakci.

Příprava DNA knihovny

Při přípravě DNA knihovny je DNA (typicky 2-3 µg) nejprve fragmentována (nebulizací, sonikací). Na konce DNA fragmentů jsou po úpravě jejich konců připojeny specifické adaptory, fragmenty jsou pomocí těchto adaptorů amplifikovány a následně jsou vybrány fragmenty o vhodné délce (podle typu následného cíleného obohacování). Vzorky je možno při amplifikačním kroku označit specifickým kódem v primeru a pracovat s více vzorky najednou. Tento krok významně snižuje pracnost i finanční náročnost přípravy knihovny.

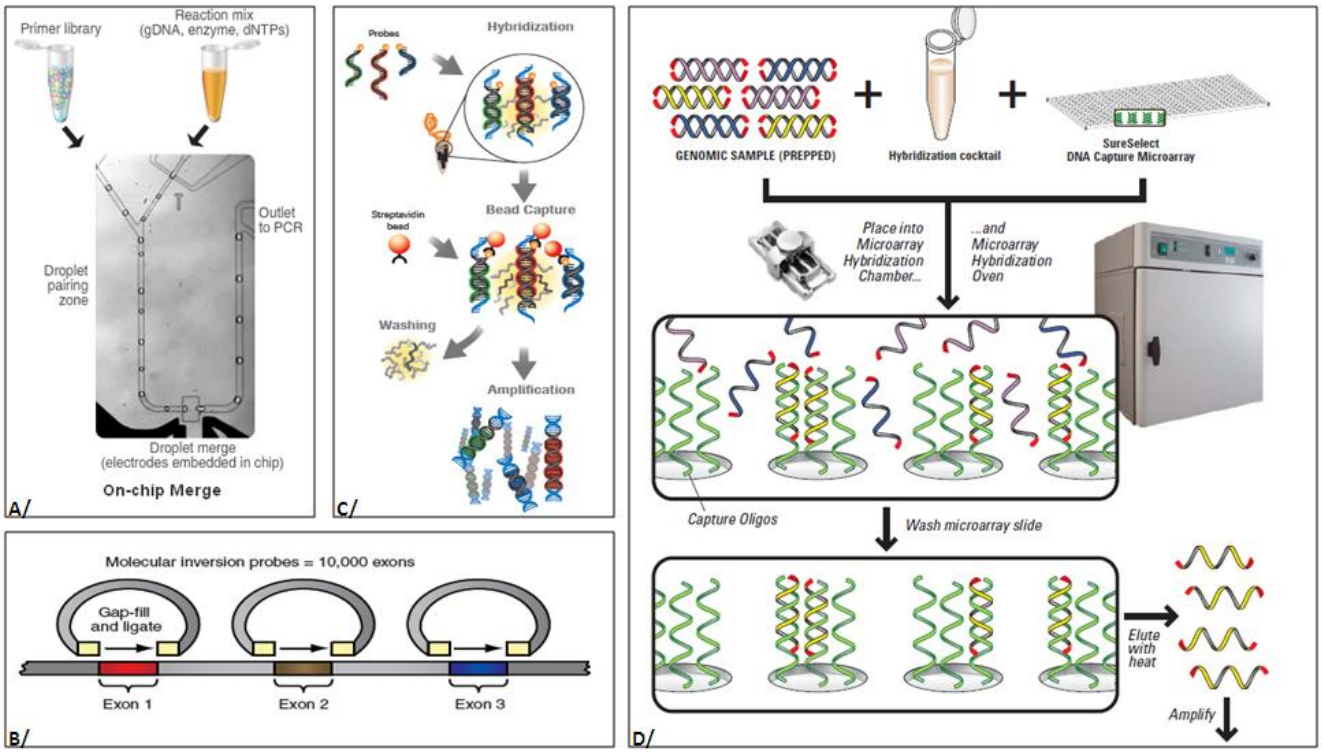
Metody cíleného obohacování

Metody cíleného obohacení umožňují sekvenovat pouze vybrané oblasti genomu, nejčastěji exom či specifické kandidátní oblasti nebo geny. Jsou dostupné díky technologiím vyvinutým některými komerčními firmami, liší se svými principy a jejich efektivita je odlišná. Dnes jsou na trhu dva hlavní způsoby cíleného obohacení – obohacení v roztoku nebo na pevné fázi (Mamanova et al. 2010, Summerer 2009).

Jednou z variant cíleného obohacení v roztoku je využití multiplex PCR využitelné pro následné cílené sekvenování malého množství vybraných PCR produktů. Možností je využití PCR v mikroreaktorech, kdy každý mikroreaktor obsahuje samostatný pár primerů a probíhá v něm samostatná PCR reakce (obr. 3A). Tento přístup nabízí platforma RainDrop™ Digital PCR System (RainDance Technologies), umožňuje cílené obohacení až 20 000 genomických oblastí v jednom vzorku a nabízí několik screeningových panelů (http://www.raindancetech.com/applications/next-generation-sequencing-technology.asp). Dalším přístupem je využití molekulárních inverzních prob (molecular inversion probes, MIP). Principem je využití jednovláknových oligonukleotidů, které obsahují jak univerzální sekvenci, tak sekvenci specifickou pro cílený úsek. Po navázání na specifický úsek je oligonukleotid cirkularizován pomocí ligázy a amplifikován pomocí univerzální sekvence (obr. 3B). Tento přístup je použitelný spíše pro menší množství vybraných cílových sekvencí. Nejvyužívanější metodou cíleného obohacení v roztoku je využití hybridizace DNA knihovny se souborem specifických prob značených biotinem, izolace hybridních molekul afinitní vazbou na streptavidinové magnetické kuličky a následná amplifikace obohacené knihovny pomocí univerzálních adaptorů (obr. 3C). Jako specifické proby slouží buď biotinylované RNA fragmenty (Agilent Technologies) nebo cDNA fragmenty (Roche Nimblegen, Illumina). Dostupné jsou sady pro cílené obohacení celého exomu (Agilent SureSelect All Exon kits, Roche Nimblegen SeqCap EZ Human Exome Library, TruSeq Exome Enrichment Kit), stejně tak je možné navrhnout si pro vlastní potřeby sady specifických prob až do souhrnné velikosti cílené oblasti do 7Mb (Agilent SureSelect Target Enrichment Platform a HaloPlex™ Target Enrichment System, Roche Nimblegen SeqCap EZ Choice Library). Dostupné jsou také validované sety pro cílené obohacení například X chromozomu, kinomu nebo panel pro sekvenování kandidátních genů u některých typů rakovinných onemocnění.

Cílené obohacování na pevné fázi využívá technologii DNA čipů pro imobilizaci specifických prob, následnou hybridizaci komplexního vzorku fragmentované DNA, zachycení specifických sekvencí, odmytí nespecifických vazeb a uvolnění a amplifikace obohacené knihovny (obr. 3D). Stejně jako u předchozí metody jsou k dispozici jak čipy pro obohacení celého exomu (Agilent SureSelect DNA Capture Array, Roche Nimblegen Sequence Capture Human Exome 2.1M Array), stejně tak existuje i možnost navržení vlastního DNA čipu na obohacení sekvencí z oblasti zájmu, maximální souhrnná velikost oblasti se v tomto případě pohybuje kolem 3,5 -5Mb.

Metody cíleného obohacení exomu na principu selektivní hybridizace jsou dnes nejvyužívanější metodou pro celoexomové sekvenování, byly využity například pro projekty 1000 genomů a sekvenování exomů. Limitací je především problematická hybridizace a tedy obohacení GC bohatých úseků, blokování nespecifických vazeb, výběr prob pro repetitivní úseky a hybridizace homologních úseků. Žádná z dostupných metod nepokrývá z těchto důvodů kompletně celý exom. Databáze také dosud pravděpodobně neobsahují všechny existující exony, lidský genom navíc může obsahovat i dosud neidentifikované geny. Žádné nabízené sady nezahrnují regulační oblasti genomu (promotory, enhancery, konzervované oblasti), které mohou obsahovat funkčně významné patogenní varianty.



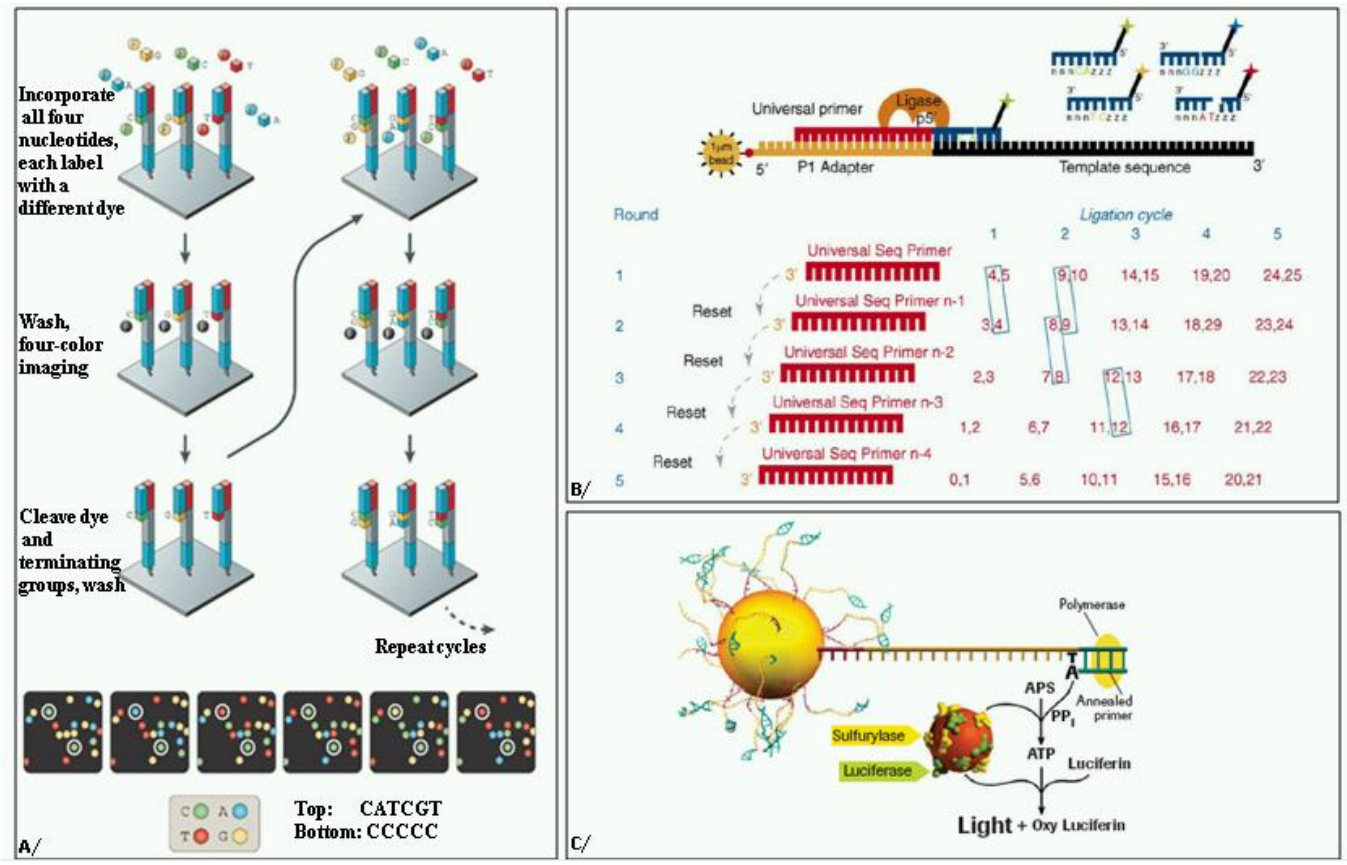
Obr. 3 Metody cíleného obohacení
A/ RainDrop™ Digital PCR System (http://www.raindancetechnologies.com/applications/next-generation-sequencing-technology.asp)
B/ Molekulární inverzní proby - upraveno podle (Mamanova et al. 2010)
C/Cílené obohacení v roztoku pomocí hybridizace specifických prob (http://www.nimblegen.com/products/lit/SeqCapEZChoice_20May2011_LR.pdf)
D/ Cílené obohacení na pevné fázi (http://www.chem.agilent.com/Library/usermanuals/Public/G4458-90000_SureSelect_DNACapture.pdf)

Amplifikace obohacené DNA knihovny

V závislosti na sekvenátoru a sekvenční metodě se využívá několik metod amplifikace obohacené DNA knihovny. Emulzní PCR využívá imobilizaci DNA knihovny na mikrokuličky pomocí univerzálního adaptoru, vytvoření mikroreaktoru s komponentami PCR, které obsahují právě jednu takovou mikrokuličku a následnou PCR reakci (Roche Nimblegen, Life Technologies, Polonator). Amplifikace na pevné fázi (Illumina) využívá metodu můstkové amplifikace, při které se vytváří prostorově oddělené shluky amplifikovaných fragmentů. Některé sekvenační strategie využívají sekvenování jednotlivých molekul DNA templátu (Helicos Biosciences, Pacific Biosciences), amplifikační krok tedy není potřebný.

Sekvenační strategie

Sekvenační strategie závisí na zvolené přípravě DNA knihovny, její amplifikaci a metodách imobilizace. Cyklická reverzibilní terminace (obr. 4A) využívá inkorporace fluorescenčně značeného nukleotidu, jeho detekce a následné odštěpení fluoroforu, přičemž je zajištěna inhibice inkorporace dalších značených nukleotidů pomocí různých typů reverzibilních blokátorů (Illumina, Helicos Biosciences, LaserGen, Pacific Biosciences). Sekvenování pomocí ligace (sequencing by ligation) (obr. 4B) využívá hybridizace speciálně navržených fluorescenčně značených prob, jejich ligace k primeru nebo k předchozí probě a detekce fluorescence podle specifického kódu (kódování pomocí dvou bazí). Po pěti ligačních cyklech jsou proby odštěpeny a pro hybridizaci jsou použity nové primery v pozici (n-1) vůči předchozím, tento cyklus se opakuje pětkrát; celková délka čtení je tedy 25 bp, přítomnost každé báze je ověřena dvěma nezávislými ligacemi (Life Technologies). Při pyrosekvenování (obr. 4C) jsou k amplifikované DNA imobilizované v pikotitračních destičkách přidávány v limitujícím množství jednotlivé dideoxynukleotidy. Uvolněný pyrofosfát je detekován pomocí luciferázové reakce (Roche Nimblegen).



Obr. 4 Sekvenační strategie sekvenátorů nové generace
A/ Cyklická reverzibilní terminace sekvenátoru Illumina, upraveno podle (Metzker 2010)
B/ Sekvenování pomocí ligace (http://www.invitrogen.com)
C/Pyrosekvenování (http://454.com/)

Všechny technologie mají své limity, vývoj a aplikace nových strategií a metod je proto stále hnacím motorem současného biotechnologického výzkumu. Perspektivní metody, které se začínají prosazovat, jsou založeny na sekvenování jedné molekuly DNA. Jejich principy se různí, mohou být založeny na detekci inkorporace nově syntetizované báze pomocí změny pH (Personal Genome Machine, Life Technologies), na metodě FRET (Fluorescent Resonance Energy Transfer), principu optických pastí, nanoporového sekvenování nebo real-time sekvenování (Blow 2008).

Zpracování dat a jejich filtrování

Všechny typy sekvenátorů nové generace jsou založeny na detekci fluorescenčního nebo chemiluminiscentního signálu. Primárními daty je tedy detekce světelných kvant ve formě obrázků, ty jsou softwarem transformovány do jednotlivých bazí a sekvencí, pro každou bázi je vypočtena její kvalita. Další analýza spočívá v mapování krátkých sekvencí na referenční genom a následně určení odlišností sekvenovaného vzorku od referenční sekvence. Všechny tyto kroky lze provádět mnoha komerčními i veřejně dostupnými algoritmy (Dalca and Brudno 2010) a jejich správná volba je úkolem pro zkušeného bioinformatika.

Výstupem bioinformatické analýzy sekvenačních dat je anotovaný seznam odchylek sekvenovaného vzorku od referenčního genomu. Sekvenování exomu jednoho jedince generuje řádově 20 000- 25 000 variant (Singleton 2011). Většinu variant tvoří jednonukleotidové záměny, zatímco inzerce a delece tvoří pouze 1-2% detekovaných variant (Ng et al. 2008). Převážná většina variant je benigních, jedná se o synonymní záměny nebo funkčně neutrální záměny, inzerce či delece.

Prvním krokem ve filtrování exomových dat je přítomnost či nepřítomnost variant ve veřejně dostupných databázích a frekvence jejich výskytu. V současné době jsou volně dostupná data z několika genomových projektů, která toto filtrování umožňují. Databáze dbSNP (Smigielski et al. 2000) (http://www.ncbi.nlm.nih.gov/projects/SNP/) je centrální databází NCBI, obsahuje v současnosti přes 60 000 záznamů polymorfismů v lidské DNA. Je však zatížena poměrně vysokým procentem falešně pozitivních záznamů (Day 2010). Dalším významným zdrojem je projekt 1000 Genomů (Altshuler et al. 2010) (http://www.1000genomes.org), který je zaměřen na identifikaci vzácných sekvenčních variant s frekvencí alespoň 1% ve studovaných populacích nebo mezinárodní projekt HapMap (http://hapmap.ncbi.nlm.nih.gov/), který si klade za cíl vytvořit databázi genetických variant u jednotlivých populací. Využívaným zdrojem je také databáze exomových variant (Exome variant server, http://evs.gs.washington.edu/EVS/), který je výsledkem projektu sekvenování exomů (Tennessen et al. 2012). Kromě veřejně dostupných databází slouží pro efektivní filtrování variant databáze exomových dat jednotlivých laboratoří, díky kterým lze odhalit varianty, které jsou pouze častými sekvenačními chybami danými použitou sekvenační strategií, metodou cíleného obohacení a možnými systematickými chybami dané laboratoře.

Standardně jsou za kandidátní varianty považovány nesynonymní záměny, záměny vnášející nebo rušící terminační kodon, sestřihové varianty a inzerce a delece způsobující posun čtecího rámce. Synonymní záměny, které často tvoří více než polovinu seznamu variant, jsou pro zjednodušení považovány za funkčně nevýznamné.

Dalšími logickými kroky při filtrování seznamu variant jsou filtry na odpovídající typ dědičnosti a na přítomnost variant ve stejném genu i u dalších rodin či sporadických případů. Kandidátní varianta musí také segregovat ve studovaném souboru s daným fenotypem. Jinými užitečnými

nástroji jsou predikce patogenity dané varianty pomocí programů SIFT (Sorting Intolerant From Tolerant, http://sift.jcvi.org), Polyphen (http://genetics.bwh.harvard.edu/pph2/) a SNAP (Screening for Nonacceptable Polymorphisms, http://cubic.bioc.columbia.edu/services/SNAP). Pro další filtrování lze dále hodnotit evoluční konzervovanost varianty nebo přítomnost varianty v genu, který je exprimován v relevantní tkáni. Často mohou velmi napomoci údaje z literatury, popis jiných mutací ve stejném genu a širší biologické souvislosti známých informací o proteinu a jeho role v metabolických či jiných drahách.

Kombinací přístupů k filtrování variant je dnes možné rychle a efektivně získat seznam řádově jednotek, maximálně desítek kandidátních variant, u kterých jsou následně prováděny další funkční studie nebo jsou aplikovány příslušné diagnostické postupy.

Výběr strategie pro studium vzácných chorob podle typu dědičnosti

Prvním krokem při výběru strategie pro studium genetické podstaty konkrétní neznámé choroby musí být stanovení, jedná-li se o familiární či sporadický výskyt. V případě familiárního výskytu je pak nutno získat informace o co největším počtu postižených i nepostižených příbuzných minimálně ve třech generacích. Z rodokmene lze pak odhadnout, o jaký typ dědičnosti by se v daném případě mohlo jednat. Vždy je také výhodné zajistit maximum biologického materiálu, který lze využít pro pozdější molekulární, biochemické a buněčné studie.

Obecně se předpokládá, že většina geneticky podmíněných onemocnění je způsobena bodovými mutacemi, inzercemi a delecemi malého rozsahu nebo variantami v počtu kopií (CNVs). Změny malého rozsahu jsou detekovatelné běžnými sekvenčními metodami i sekvenátory nové generace, detekce CNVs se provádí pomocí genotypovacích čipů. Detekce CNVs je tedy dosud běžně prováděna paralelně s vazebnými analýzami nebo homozygotním mapováním a po prokázání nepřítomnosti velkých delecí či duplikací se přistupuje k celoexomovému sekvenování. Je však snaha upravit algoritmy pro analýzu celoexomových dat a modifikovat celoexomové sekvenování tak, aby bylo možné získat validní a spolehlivé informace o variantách v počtu kopií i jen z celoexomových dat.

U autozomálně dominantních onemocnění je zlatým standardem metoda vazebné analýzy. V závislosti na velikosti studované rodiny, případně na počtu rodin je tak možné získat kandidátní oblasti a analýzou jejich obsahu přímo vytipovat kandidátní geny. Tento přístup sám o sobě již může vést k identifikaci kauzálního genu, pokud jsou kandidátní oblasti dostatečně malé a existují další biologická vodítka, která umožní vybrat „správný“ kandidátní gen (biochemická charakteristika proteinu, jeho lokalizace, předpokládaná funkce). Tímto způsobem byl identifikován kauzální gen například u mukopolysacharidózy typu IIIC (Hrebicek et al. 2006) nebo familiární juvenilní hyperurikemické nefropatie (Zivna et al. 2009). Často však není k dispozici dostatečný počet informativních jedinců pro vymezení dostatečně úzké kandidátní oblasti, vazebná analýza provedená na více rodinách zase může selhat z důvodu heterogenity daného onemocnění. Proto bývá vazebná analýza u malých rodin v současnosti dále doplňována celoexomovým sekvenováním rodičů a jejich dětí. Vhodně filtrovaný seznam vzácných variant je pak propojen s údaji z vazebné analýzy; výsledkem bývají jednotky variant, u nichž je dále kandidátní záměna prioritizována s ohledem na

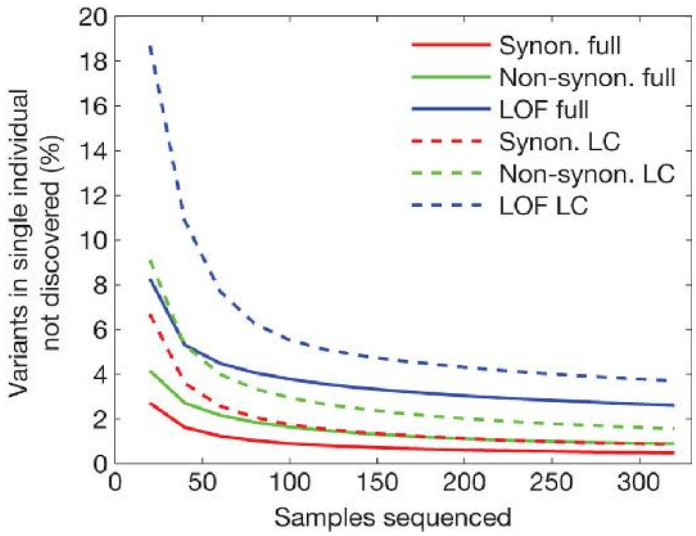
funkci genu, jeho expresi, patogenitu varianty a informace z literatury. Kombinací vazebné analýzy a celoexomového sekvenování byly v nedávné době identifikovány například kauzální geny pro autozomálně dominantní amyotrofní laterální sklerózu (Johnson et al. 2010). U velmi malých rodin se vazebná analýza neprovádí a přistupuje se přímo k exomovému sekvenování; tehdy je vhodné zvolit pro sekvenování jedince, kteří jsou v rodokmenu nejvíce vzdáleni a jsou tedy geneticky nejodlišnější. Validitu vybraných záměn je nutno ověřit Sangerovým sekvenováním a prokázat její segregaci s fenotypem. S ohledem na typ varianty se pak případně provádějí další funkční studie.

Při studiu autozomálně recesivních onemocnění lze opět uplatnit kombinaci genotypování a celoexomového sekvenování se stejnými výhodami i nevýhodami popsanými výše. Filtrováním na recesivní typ dědičnosti (tedy hledání záměn v homozygotním stavu nebo složených heterozygotů v rámci jednoho genu) je většinou získáno mnohem méně kandidátních variant než u dominantního typu dědičnosti. V případě, že se jedná o pacienty z rodin s příbuzenským křížením, je výhodné analyzovat v rámci genotypovací analýzy také ztrátu heterozygocie; takový přístup umožní definovat společné homozygotní oblasti, ve kterých by se měl nacházet kauzální gen. Takto byl definován například kauzální gen pro izolovaný deficit ATP syntázy (Cizkova et al. 2008b) nebo pro jednu z variant spinální muskulární atrofie (Renbaum et al. 2009). Kombinace homozygotního mapování a celoexomového sekvenování umožňuje definovat varianty v homozygotních oblastech u chorob, které nebylo z důvodu malého počtu pacientů možno studovat jinými metodami (Bolze et al. 2010, Ku, Naidoo and Pawitan 2011).

X- vázaná onemocnění je možné efektivně studovat díky možnostem cíleného obohacení a následného sekvenování pouze kódujících oblastí X-chromozomu. Ty jsou dnes běžně nabízeny komerčními firmami (Agilent SureSelect X-Chromosome kit, Roche Nimblegen SeqCap EZ Choice, RainDance XSeq panel). Tento přístup byl již využit například při studiu X-vázaných syndromů (Zhu et al. 2011, Huebner et al. 2011). Díky možnosti specifického kódování vzorků je možné analyzovat cenově efektivně více vzorků najednou. Podobně lze studovat také mitochondriální onemocnění podmíněná mutacemi v mitochondriální DNA, kdy je vstupním materiálem místo celogenomové DNA pouze amplifikovaná mtDNA.

Celoexomové sekvenování umožňuje objasnit také genetickou podstatu sporadických onemocnění a vzácných syndromů. V těchto případech často nelze předem určit typ dědičnosti, může se jednat o homozygotní varianty, složené heterozygoty nebo o *de-novo* mutace (předpokládá se, že v exomu každého jednotlivce je přítomna jedna až dvě *de-novo* vzniklé mutace (Nachman and Crowell 2000)). Často se jedná o jednotky či desítky nepříbuzných pacientů s daným syndromem či onemocněním. Výhodné je proto provádět celoexomové sekvenování rodičů a postiženého jedince, výsledný seznam kandidátních variant filtrovat s ohledem na možné typy dědičnosti a na přítomnost variant nepřítomných u rodičů. V případě, že nejsou pro analýzu k dispozici ani rodiče postiženého jedince, lze provést celoexomové sekvenování a jeho výsledky porovnat s variabilitou v populaci. S rostoucím počtem sekvenovaných lidských genomů roste počet variant v databázích (viz obrázek 5). Odhaduje se, že s výjimkou *de-novo* vzniklých unikátních variant jsou již v databázích přítomny téměř všechny varianty společně s jejich populačními frekvencemi (Altshuler et al. 2010). Při hledání kauzální varianty pro vzácný sporadický případ lze tedy poměrně jednoduše nalézt buď unikátní variantu nepřítomnou v databázi, nebo variantu, jejíž frekvence zhruba odpovídá počtu popsaných případů.

Uvedené kombinace metod mají potenciál i pro studium jiných než vzácných monogenních onemocnění. Studium komplexních onemocnění bylo v posledních deseti letech primárně zaměřeno na identifikaci populačně častých alel, které jsou asociovány s onemocněními, a to pomocí asociačních studií. Hlavními cíli těchto studií byly neuropsychiatrické choroby schizofrenie a autismus, obezita, diabetes nebo choroby srdce. Jejich úspěšnost však nebyla ve srovnání s vynaloženými náklady příliš velká – například u Alzheimerovy choroby, kde se předpokládá vliv genetické komponenty zhruba z 58 – 79%, byla identifikována řada kandidátních variant, ale podstata onemocnění byla dosud vysvětlena zhruba u 1% všech případů Alzheimerovy choroby (Ertekin-Taner 2010). Fenomén „chybějící genetické komponenty“ (Maher 2008) a důvody k němu vedoucí jsou v současnosti velmi diskutovány. Alternativním přístupem pro studium komplexních onemocnění a možností, jak tuto chybějící genetickou komponentu odhalit, se jeví využití sekvenování jednotlivých genomů a detekce CNVs (Cirulli and Goldstein 2010, Manolio et al. 2009, Mardis 2008). Tento přístup by měl být umožněn očekávaným technologickým pokrokem a snížením nákladů na celogenomové sekvenování. Analýza celogenomových dat bude však vyžadovat ještě řadu pokroků především v oblasti bioinformatiky, neboť mapování celogenomových dat na referenční sekvence, především oblasti repetitivních sekvencí a mobilních elementů, je zatím velmi problematické.



Obr. 5 Příspěvek analýzy dalších vzorků k celkové genetické variabilitě populace, The 1000 Genomes Project Consortium *et al. Nature* **467**, 1061-1173 (2010) doi:10.1038/nature09534. S přibývajícím množstvím sekvenovaných vzorků klesá počet nových variant; předpokládá se tedy, že v blízké době budou databáze obsahovat všechny polymorfní oblasti dané populace společně s údaji o jejich frekvenci. Jednotlivé křivky ukazují předpokládaný průběh u nesynonymních, synonymních variant a variant způsobujících ztrátu funkce (LOF) při plném sekvenování a sekvenování s nízkým pokrytím (LC).

ČÁST II. Cíle dizertační práce

Hlavním cílem dizertační práce bylo využití nových genomických technik při studiu vybraných vzácných dědičných onemocnění. Tyto techniky mají vždy za cíl identifikovat kandidátní geny, u kterých je pak nutné dalšími molekulárně-genetickými, biochemickými a buněčnými metodami potvrdit kauzalitu nalezených genomových změn. I tyto následné analýzy jsou proto nedílnou součástí této dizertační práce.

Jednotlivé části této práce mají následující cíle:

1. Využití kombinace vazebné analýzy, analýzy změn genové exprese, analýzy změn v počtu kopií a celoexomového sekvenování při hledání kauzálního genu pro autozomálně dominantní adultní formu neuronální ceroidní lipofuscinózy.

2. Využití homozygotního mapování a analýzy změn v počtu kopií při studiu Rotorova syndromu.

3. Využití homozygotního mapování, analýzy genové exprese a genotypování při studiu autozomálně recesivního izolovaného deficitu ATP syntázy.

4. Studium diferenciální genové exprese u genů z kandidátní oblasti vymezené vazebnou studií pacientů s mukopolysacharidózou typu IIIC (MPSIIIC).

ČÁST III. Studium vybraných vzácných metabolických poruch pomocí kombinace nových genomických technik.

Studium molekulární podstaty autozomálně dominantních neuropsychiatrických onemocnění.

a/ Objasnění molekulární podstaty autozomálně dominantní adultní formy neuronální ceroidní lipofuscinózy (Kufsovy choroby).

Neuronální ceroidní lipofuscinózy (NCL) jsou heterogenní skupinou vzácných geneticky podmíněných neurodegenerativních onemocnění. Jejich incidence se pohybuje mezi 1 - 30 na 100 000 narozených jedinců. Společnými charakteristikami těchto onemocnění je akumulace autofluorescentního materiálu v lysozomech neurálních a periferních tkání a neurodegenerace. Povaha akumulovaného autofluorescentního materiálu není přesně známa, jedná se o směs proteinů, proteolipidů a kovů (Palmer et al. 1997). Na úrovni elektronové mikroskopie se morfologicky rozeznává několik ultrastruktur odlišujících jednotlivé typy NCL – granulární osmiofilní depozity (GRODy), kurvilineární depozity nebo struktury podobné otiskům prstů (Goebel and Wisniewski 2004). Hlavní složku střádavého materiálu tvoří podjednotka c mitochondriální ATP syntázy (PALMER et al. 1992), u některých forem NCL byla popsána také přítomnost saposinů A a D (Seehafer and Pearce 2006).

Neuronální ceroidní lipofuscinózy se projevují progresivním zhoršováním kognitivních i fyzických schopností, poruchami vidění a epileptickými záchvaty. Dělí se tradičně podle věku prvního nástupu nemoci na infantilní, pozdně infantilní, juvenilní a adultní formy. Doposud bylo popsáno více než 400 mutací v devíti genech - *CLN1 (PPT1*, MIM 256730), *CLN2 (TPP1*, MIM 204500), *CLN3*, (MIM 204200), *CLN5* (MIM 256731), *CLN6* (MIM 601780), *CLN7 (MFSD8*, MIM 610951), *CLN8, CLN10 (CTSD*, MIM 610127), *CLN13 (CTSF*), mutace v dalších třech genech byly popsány vždy v jedné rodině – *CLN11 (GRN*)(Smith et al. 2012), *CLN12 (ATP13A2*)(Bras et al. 2012) a *CLN11 (KCTD7*)(Staropoli et al. 2012). Nedávno byla popsána mutace v *CLN6* také u autozomálně recesivní adultní formy (Kufsova choroba, MIM 204300)(Arsov et al. 2011).

Adultní forma NCL byla prvně popsána německým neuropatologem Hugo Kufsem v roce 1926. Jedná se o velmi vzácnou formu a informace o ní jsou ve srovnání s infantilními a juvenilními formami minimální. Sporadické nebo autozomálně recesivní případy jsou popisovány jako Kufsova choroba (CLN4A, MIM 204300), zatímco autozomálně dominantní adultní NCL je nazývána jako Parryho choroba, případně Parryho typ Kufsovy choroby (CLN4B, MIM 162350).

Adultní NCL se objevují ve věku mezi 10 až 50 lety (Nijssen 2011), projevují se epilepsií, ataxií, parkinsonismem, křečemi a zhoršováním kognitivních funkcí. Na rozdíl od ostatních forem NCL nejsou nikdy přítomny poruchy vidění.

Autozomálně dominantní adultní neuronální ceroidní lipofuscinóza (ANCL) byla poprvé popsána v roce 1971 v rodině britského původu z New Jersey („rodina Parry") (Boehme et al. 1971). Další rodiny byly popsány ve Španělsku (FERRER et al. 1980), v USA (Josephson et al. 2001, Burneo et al. 2003) a v Holandsku (Nijssen et al. 2002, Nijssen et al. 2003, Nijssen, Brekelmans and Roos 2009).

V Ústavu dědičných metabolických poruch 1. LF UK a VFN v Praze byla diagnostikována česká rodina s autozomálně dominantní adultní formou NCL. Genomovou DNA členů této rodiny s jasně stanovenou diagnózou jsme získali pro molekulárně genetickou analýzu s cílem odhalit kauzální gen způsobující ANCL v této rodině a přispět tak k odhalení příčiny dosud neznámé autozomálně dominantní ANCL.

Pro studium molekulární podstaty ANCL byla použita kombinace vazebné analýzy, expresní analýzy a exomového sekvenování. Genomová DNA všech dostupných členů rodiny byla nejprve použita pro genotypovaní a vazebnou analýzu. Identifikovali jsme pět kandidátních oblastí na chromozomech 1, 4, 15, 20 a 22 obsahujících zhruba 560 známých genů. U sedmi dostupných pacientů bylo dále provedeno genotypování a analýza změn počtu kopií, nenalezli jsme však žádné potenciálně patogenní rozsáhlejší delece nebo duplikace. Paralelně jsme provedli analýzu genové exprese z leukocytů čtyř pacientů a čtyř kontrol, získali jsme seznam diferenciálně exprimovaných genů, z nichž 65 leželo v oblastech identifikovaných vazebnou analýzou. Funkční anotace a analýza genového obohacení (gene enrichment analysis) ukázaly významnou dysregulaci spliceosomu, upregulaci mnoha složek respiračního řetězce, změněnou expresi genů aktivních u neurodegenerativních chorob - Huntingtonovy choroby, Alzheimerovy choroby a Parkinsonovy choroby a urychlenou proteolýzu. Tyto údaje však stále nebyly dostačující pro nalezení mutace v kauzálním genu. Proto jsme se rozhodli sekvenovat celý exom jednoho pacienta na sekvenátoru SOLiD 4 v institutu CeGaT v Tubingen v Německu. Pro sekvenovaného jedince jsme získali celkem 957 unikátních změn (SNPs a indel záměn).

Po propojení výsledků vazebné analýzy, expresní analýzy a exomového sekvenování se nám podařilo identifikovat kandidátní gen *DNAJC5* ležící v kandidátní oblasti na chromozomu 20q13.33, jehož exprese byla v leukocytech pacienta statisticky významně zvýšena ve srovnání s kontrolami, a který obsahoval unikátní heterozygotní mutaci c.346_348delCTC (p. Leu116del). Segregace této záměny byla ověřena u ostatních členů rodiny Sangerovým sekvenováním.

Díky spolupráci s Rare NCL Gene Consortium shromažďujícím vzorky pacientů s vzácnými typy NCL jsme získali genomovou DNA dalších 23 případů, a to buď familiárních, nebo sporadických. Tyto případy byly diagnostikovány jako Kufsova choroba, klinická a neuropatologická dokumentace však nebyla vždy dostupná. U jednoho dalšího dosud nepublikovaného případu jsme identifikovali totožnou záměnu v genu *DNAJC5* c.346_348delCTC , u dalších tří případů pak byla nalezena heterozygotní mutace v *DNAJC5* c.344T>G (p. Leu115Arg). Klinická data dvou z těchto tří rodin již byla publikována (Josephson et al. 2001, Nijssen et al. 2002). Ve všech případech, kdy byla k dispozici gDNA, byla prokázána segregace mutací v rodinách.

Ani jedna z popsaných záměn v genu *DNAJC5* nebyla popsána v databázích dbSNP a 1000 Genomů, nenalezli jsme ji ani u 200 kontrolních vzorků evropské populace. Abychom ověřili, že mutace vznikly nezávisle, provedli jsme homozygotní mapování genomových oblastí *DNAJC5* a ukázali, že mutace p.Leu116del je přítomna u dvou testovaných rodin na dvou odlišných haplotypech. Mutace tedy pravděpodobně vznikly nezávisle na sobě. Stejně tak mutace p.Leu115Arg byla přítomna na dvou odlišných haplotypech.

DNAJC5 kóduje protein cysteine-string protein alpha (CSPα). Tento protein patří do skupiny J-proteinů, molekulárních chaperonů interagujících s rodinou proteinů Hsp70, které hrají významnou

roli v inhibici neurodegenerativních procesů (Zhao, Braun and Braun 2008). CSPα je membránovým proteinem lokalizovaným v synaptických váčcích, obsahuje na cystein bohatou oblast, jejíž palmitoylace kotví protein do synaptické membrány (Greaves and Chamberlain 2006, Chamberlain and Burgoyne 1998, Greaves et al. 2008). Interaguje s proteinem Hsc70 a SGT a vytváří tak enzymaticky aktivní chaperonový komplex, který ve spolupráci s dalšími chaperony zprostředkovává správnou konformaci řady proteinů účastnících se synaptických dějů. Deplece CSPα vede u myši a u *Drosophila melanogaster* k progresivní neurodegeneraci a redukované délce života (ZINSMAIER et al. 1994, Fernandez-Chacon et al. 2004).

CSPα je vysoce konzervovaný protein a námi popsané mutace ovlivňují evolučně konzervované leuciny lokalizované v cystein bohaté oblasti. Provedli jsme několik funkčních studií, které měly ukázat efekt mutací na funkci proteinu. Pomocí *in silico* analýz jsme prokázali, že obě mutace ovlivňují lokalizaci CSPα v buňce tím, že mění hydrofobicitu a efektivitu palmitoylace. Tato změněná lokalizace byla potvrzena v *in vitro* studiích pomocí transientní exprese *wt* a mutovaných forem EGFP- CSPα konstruktů v CAD5 neuronálních buňkách. Imunohistochemická studia dále ukázala snížené množství proteinu (u p.Leu116del mutace) či naprostou absenci proteinu (u p.Leu115Arg mutace) v šedé hmotě mozkové kůry pacientů. U pacienta s p.Leu116del mutací byla dále metodou Western blot prokázána nepřítomnost solubilního CSPα v mozkovém homogenátu a velké množství agregovaného nerozpustného materiálu obsahujícího CSPα protein. Tyto agregáty pravděpodobně obsahují jak *wt*, tak mutovanou formu proteinu a mohou tak způsobovat depleci CSPα.

Funkčními studiemi jsme prokázali, že mutace v genu *DNAJC5* jsou příčinou adultní formy neuronální ceroidní lipofuscinózy. Zároveň byla prokázána neuroprotektivní funkce CSPα na modelu vzácného onemocnění. Tak se otevřela další cesta pro studium role CSPα u neuronálních ceroidních lipofuscinóz i v dalších neurodegenerativních procesech. V současné době probíhá ve spolupráci s Biotechnologickým institutem v Seville a Univerzitou v Kalifornii studium funkce CSPα na úrovni neuronálních kmenových buněk derivovaných z fibroblastů pacientů s ANCL.

Publikace s názvem „Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis.“ byla publikována v roce 2011 v časopise The American Journal of Human Genetics (IF = 11,680).

Výsledky naší studie byly potvrzeny následnými publikacemi, ve kterých byly popsány tytéž mutace u stejných i dalších rodin a sporadických případů (Benitez et al. 2011, Velinov et al. 2012).

Příspěvek autorky dizertace k této studii:

V rámci této studie jsem se podílela na analýze a interpretaci výsledků vazebné studie, expresních dat a exomového sekvenování, ověřovala jsem segregace mutací, prováděla jsem mutační analýzu u všech dostupných vzorků. Podílela jsem se na *in vitro* studii přípravou konstruktů a koordinací studií, prováděla jsem analýzu buněčných linií a tkáňových lyzátů pomocí metody Western blot. Podílela jsem se na přípravě publikace, na které jsem prvním autorem. Publikace byla v roce 2011 oceněna Bolzanovou cenou v lékařské kategorii udělovanou rektorem Univerzity Karlovy a získala cenu Arnolda Beckmana za rok 2011.

b/ Identifikace mutací v presenilinu 1 u rodiny zařazené do souboru suspektních případů Kufsovy choroby.

Vzhledem k tomu, že identifikací mutací v genu *DNAJC5* byla objasněna zhruba čtvrtina případů, které jsme měli k dispozici, rozhodli jsme se použít podobný přístup, jaký byl popsán v předchozí studii, i u další rodiny, která nám byla zaslána s diagnózou autozomálně dominantní Kufsovy choroby (rodina UCL568, (Noskova et al. 2011)).

I u této rodiny jsme nejprve provedli vazebnou analýzu a identifikovali jsme 14 kandidátních oblastí s pozitivním LOD skóre na chromosomech 3, 4, 8, 9, 10, 13, 14, 16 a 19. Zároveň jsme u jednoho pacienta provedli analýzu změn počtu kopií pomocí Affymetrix GeneChip Mapping 6.0 Array; nenalezli jsme žádné potenciálně patogenní rozsáhlejší delece nebo duplikace. Následně jsme u dvou pacientů a jednoho nepostiženého příbuzného provedli exomové sekvenování na sekvenátoru SOLiD™4 System instalovaném v Ústavu dědičných metabolických za použití exomového obohacení pomocí Agilent SureSelect All Exome Kit. Sekvenováním bylo identifikováno 11 723 jednonukleotidových záměn a 150 inzercí/delecí přítomných v heterozygotním stavu u obou pacientů a nepřítomných u zdravého příbuzného. Z tohoto počtu bylo pouze 65 jednonukleotidových záměn buď nových nebo přítomných v databázích dbSNP, 1000 Genomes, Exome variant server a v místní exomové databázi s frekvencí nižší než 0,001. Po propojení dat exomového sekvenování s informacemi z vazebné analýzy jsme získali seznam sedmi jednonukleotidových záměn, ve kterém jsme jako prioritní kandidátní mutaci označili heterozygotní mutaci c.509C>T (p.Ser170Phe) v genu presenilin 1. Presenilin 1 (*PSEN1*) je jedním ze čtyř genů, u něhož jsou mutace spojovány s familiární Alzheimerovou chorobou s časným nástupem (Bekris et al. 2010). Tato mutace je přítomna v databázi dbSNP pod kódem rs63750577 a je považována za potenciálně patogenní. Je přítomna také v databázi Disease & Frontotemporal Dementia Mutation Database (http://www.molgen.ua.ac.be).

Totožná mutace (*PSEN1* S170F) byla už popsána v jednom případu familiární Alzheimerovy choroby s časným nástupem (Snider et al. 2005) a u dalších tří sporadických případů (Golan et al. 2007, Piccini et al. 2007, Langheinrich et al. 2011). Dosud bylo popsáno 185 patogenních mutací v genu pro presenilin 1, což činí mutace v *PSEN1* nejčastější příčinou u dominantní familiární Alzheimerovy choroby s časným nástupem. PSEN1 je membránový protein, který je součástí γ-sekretázového komplexu (Selkoe and Wolfe 2007). Princip patogeneze není dostatečně znám, předpokládá se, že mutace v *PSEN1* mění specifitu γ-sekretázy vedoucí k zvýšenému množství delších forem amyloidu β, hlavní součásti amyloidních plaků (Citron et al. 1997, Hutton and Hardy 1997, Scheuner et al. 1996).

Díky funkční anotaci variant nalezených exomovým sekvenováním byl zároveň u všech pacientů identifikován známý polymorfismus c.C173>T (rs17571) v genu kathepsin D (*CTSD*). Tento polymorfismus byl nedávno identifikován jako rizikový faktor Alzheimerovy choroby (Schuur et al. 2011). Mutace v *CTSD* byly zároveň popsány u CLN10 - kongenitální a pozdně infantilní varianty NCL (Jalanko and Braulke 2009). Pro další studie se tedy nabízí otázka modifikujícího vlivu tohoto polymorfismu na projevy a průběh choroby.

Tato studie neodhalila další kauzální gen pro adultní formu neuronální ceroidní lipofuscinózy, rodina zařazena do souboru byla nesprávně diagnostikována a byla u ní identifikována již známá mutace podmiňující Alzheimerovu chorobu. Metodický přístup ale potvrdil, že kombinací vazebné

analýzy a exomového sekvenování a díky stále se rozšiřujícímu množství exomových dat v databázích jsme schopni rychle a efektivně odhalit mutaci podmiňující onemocnění. Zároveň ukázala možnost zvažovat vliv dalších polymorfismů jako modifikujících faktorů studovaných chorob.

Publikace s názvem „Cerebellar dysfunction in a family harbouring the *PSEN1* mutation co-segregating with a Cathepsin D variant p.A58V.“ byla přijata v roce 2013 k publikování v časopise Journal of the Neurological Sciences (IF = 2,353).

Příspěvek autorky dizertace k této studii:

V rámci této studie jsem se podílela na přípravě vzorků pro vazebnou analýzu a exomové sekvenování a na jejich analýze, ověřovala jsem segregaci mutací a podílela jsem se na přípravě publikace, na které jsem společným prvním autorem.

Studium molekulární podstaty Rotorova syndromu.

Rotorův syndrom (RS, MIM237450) je vzácné benigní autozomálně recesivní onemocnění. Projevuje se konjugovanou hyperbilirubinémií, koproporfyrinurií a sníženou absorpcí diagnostických sloučenin játry. Svými projevy připomíná další chorobu projevující se konjugovanou hyperbilirubinémií, a to Dubin-Johnsonův syndrom (DJS, MIM 237500). Na rozdíl od DJS nejsou u RS přítomny typické jaterní pigmentové depozity. Také podíl koproporfyrinových izomerů I a III je u RS nižší než u DJS (WOLKOFF et al. 1976). U RS se na rozdíl od DJS objevuje prodloužená clearance nekonjugovaných anionických diagnostických barviv (bromsulfathalein, indocyanová zeleň) a játra ani žlučovody se nezobrazují při cholescintigrafii (FRETZAYAS et al. 1994, BARMEIR et al. 1982). Molekulární podstata RS nebyla dosud známa.

Exkreční dráha bilirubinu v játrech byla dosud popisována jako jednosměrný proces složený ze dvou kroků. Nejprve je nekonjugovaný bilirubin přenesen jaterními přenašeči nebo pasivní transmembránovou difúzí do jaterní buňky, kde je konjugován glukuronovou kyselinou za katalýzy enzymem uridin difosfát glukuronosyl transferázou 1A1 (UGT1A1). Konjugace probíhá na membránách endoplazmatického retikula. Konjugovaný bilirubin je následně secernován do žlučových kanálků pomocí bilirubinové exportní pumpy, proteinu ABCC2.

Jsou známy mutace u obou výše zmíněných proteinů. Mutace v *UGT1A1* byly popsány u Gilbertova syndromu (MIM 143500) projevujícího se nekonjugovanou bilirubinémií (Koiwai et al. 1995). Mutace v *ABCC2* jsou příčinou Dubin-Johnsonova syndromu (Kartenbeck et al. 1996, Paulusma et al. 1997). V případě Dubin-Johnsonova syndromu nedochází k exkreci bilirubinu do žluče, bilirubin je pomocí transportéru ABCC2/MRP3 přesměrován do plazmy.

Rotorův syndrom jsme na ÚDMP začali studovat v roce 2005 ve spolupráci s Institutem klinické a experimentální medicíny, kde byli zachyceni pacienti s tímto onemocněním a kde byl následně díky zahraniční spolupráci shromážděn soubor rodin a sporadických případů s Rotorovým syndromem.

Vzhledem k ne zcela jasnému vymezení Rotorova syndromu a Dubin-Johnsonova syndromu a vzhledem k nespecifickému charakteru rysů obou onemocnění byla nejprve testována hypotéza, že RS a DJS jsou alelickými variantami stejného onemocnění. Proto byla u dvou jedinců s RS provedena histologická a mutační analýza kandidátního genu *ABCC2* . Imunohistochemické nálezy potvrdily normální lokalizaci proteinu ABCC2, sekvenování kódujících a promotorových oblastí genu *ABCC2* nepotvrdily přítomnost patogenní mutace v tomto genu. Pomocí Sangerova sekvenování lze však detekovat pouze nukleotidové záměny, případně inzerce a delece malého rozsahu. Proto jsme se rozhodli provést analýzu případných delecí či inzercí jednotlivých exonů genu pomocí komparativní genomové hybridizace. V této fázi studie nebyly ještě běžně dostupné celogenomové a „custom“ čipy pro komparativní genomovou hybridizaci, proto jsme připravili vlastní CGH čip pro detekci změn počtu kopií jednotlivých exonů několika vybraných genů včetně genu *ABCC2*. Jako proby byly použity oligonukleotidy specifické pro jednotlivé úseky jednotlivých exonů. Čipy byly připraveny na robotickém přístroji Genesurfer metodou vyvinutou v Ústavu dědičných metabolických poruch. Na jeden čip byla vždy hybridizována fluorescenčně značená genomová DNA pacienta a kontroly v dye swap modu (gDNA pacienta značená Cy3 fluoroforem, gDNA kontroly značená Cy5 fluoroforem a

naopak; důvodem je normalizace odlišného chování obou fluoroforů). Pomocí genomové komparativní hybridizace nebyly ani u jednoho ze dvou pacientů nalezeny žádné změny v počtu kopií v žádném z 32 exonů genu *ABCC2*.

Tyto výsledky, společně s dalšími mutačními a imunohistochemickými analýzami, byly publikovány v roce 2007 a potvrdily tak, že Rotorův syndrom a Dubin-Johnsonův syndrom jsou dva odlišné defekty mající odlišnou molekulární podstatu. Publikace s názvem „Rotor-type hyperbilirubinaemia has no defect in the canalicular bilirubin export pump." byla publikována v roce 2007 v časopise Liver International (IF = 2, 559).

Ve studiu molekulární podstaty Rotorova syndromu jsme dále pokračovali genotypováním pacientů z osmi rodin. Homozygotní mapování odhalilo u pacientů ze všech rodin jedinou homozygotní oblast na chromozomu 12. V oblasti byly identifikovány tři odlišné haplotypy segregující s onemocněním, jejich propojením jsme získali kandidátní oblast obsahující pět genů. Paralelně jsme provedli analýzu změn počtu kopií a detekovali u dvou haplotypů homozygotní delece zasahující v jednom případě gen *SLCO1B3*, v druhém případě geny *SLCO1B3*, *SLCO1B1 a LST-3TM12*. Následná sekvenační analýza odhalila patogenní mutace v genech *SLCO1B3* a *SLCO1B1*, u každého haplotypu byly přítomny mutace nebo delece vždy v obou genech. Předpokládaným důsledkem změn byly vždy vážné změny exprese proteinů či jejich úplná absence. Závažnost mutací byla podpořena imunohistochemickými studiemi, které ukázaly v jaterních biopsiích pacientů absenci proteinů OATP1B1 a OATP1B3 kódovanými geny *SLCO1B3* a *SLCO1B3*. U několika zdravých příbuzných jsme v jednotlivých rodinách nalezli heterozygotní i homozygotní mutace či delece samostatně buď v jednom, nebo ve druhém genu, což potvrdilo představu, že k projevům Rotorova syndromu je třeba kompletního defektu obou genů.

OATP1B3 a OATP1B1 proteiny jsou transportéry organických anionických sloučenin. Jsou lokalizovány na sinusoidní membráně hepatocytů a účastní se přenosu řady sloučenin, jako jsou konjugovaný bilirubin, žlučové kyseliny, steroidy, tyroidní hormony, řada léků, toxinů a jejich konjugátů (Hagenbuch and Meier 2004, Hagenbuch and Gui 2008).

Naše výsledky identifikující kompletní defekt obou proteinů OATP1B1 a OATP1B3 jako příčinu Rotorova syndromu byly propojeny s výsledky holandské skupiny z Amsterodamu. Ta se zabývala studiem myšího modelu s deficiencí proteinů Oatp1a/1b, myších homologů lidských proteinů OATP1B1 a OATP1B3. Z výsledků práce na tomto modelu vyplývá, že u Oatp1a/1b - deficientní myši je důležitým faktorem protein Abcc3. U Slco1a/1b$^{-/-}$ myši jsou pozorovány zvýšené hodnoty plazmového bilirubinu, ty jsou významně sníženy u Slco1a/1b; Abcc3$^{-/-}$ myši, přičemž bylo dokázáno, že Abcc3 protein odpovídá z největší části za zvýšené hodnoty bilirubinu v plazmě. Abcc3 secernuje konjugovaný bilirubin z hepatocytů zpět do krve a proteiny Oatp1a a Oatp1b transportují tento bilirubin z krve zpět do hepatocytů. Specifická jaterní exprese lidského OATP1B3 nebo OATP1B1 proteinu v Slco1a/1b$^{-/-}$ myši způsobí normalizaci hladin bilirubinu v krvi i moči, což potvrzuje, že oba lidské proteiny OATP1B3 i OATP1B1 v játrech efektivně absorbují konjugovaný bilirubin z plazmy.

Tyto dvě samostatné studie – studium pacientů s Rotorovým syndromem a studium částečného myšího modelu Rotorova syndromu - Slco1a/1b$^{-/-}$ myši umožnilo formulovat hypotézu, že exkreční dráha bilirubinu není jednoduchým jednosměrným transportem bilirubinu z krve do žluče, ale že zahrnuje smyčku absorpce, sekrece do krve, reabsorpce a sekrece do žluče. Bilirubin vstupuje

z krve do periportálních hepatocytů, kde je glukuronidován. Velká část bilirubinu konjugovaného v hepatocytech je secernována pomocí přenašeče ABCC3 zpět do krve, odkud je zpětně reabsorbována perivenózními hepatocyty pomocí proteinů OATP1B3 a OATP1B1. Následně je transportován do žluče pomocí ABCC2 přenašeče. Tento mechanismus zabraňuje lokální saturaci ABCC2 transportéru v periportálních hepatocytech a umožňuje efektivní exkreci konjugovaného bilirubinu do žluče v perivenózních hepatocytech. Pokud je reabsorpce bilirubinu blokována z důvodu současného defektu obou přenašečů OATP1B3 a OATP1BA, hromadí se konjugovaný bilirubin v krvi a způsobuje projevy odpovídající Rotorovu syndromu.

Absence obou přenašečů vysvětluje diagnostické nálezy u Rotorova syndromu - sníženou clearence anionických diagnostických barev, které jsou substráty právě těchto přenašečů, a zhoršení či absenci vizualizace jater a žlučových kanálků při cholescintigrafii. Také zvýšený poměr koproporfyrinů lze vysvětlit sníženou absorpcí těchto látek hepatocyty.

Význam studie tkví nejen v identifikaci molekulární podstaty Rotorova syndromu, ale má také širší klinický dopad. Ačkoli jsou defekty v obou genech velmi vzácné, výskyt mutací v jednom nebo druhém genu je populačně mnohem častější. Jedinci s mutacemi v proteinu OATP1B1 nebo OATP1B3 mohou vykazovat hypersenzitivitu na látky transportované těmito přenašeči, například na běžně užívané statiny. Souvislost variant v genu *SLCO1B1* se statiny indukovanou myopatií již byla popsána na základě celogenomových populačních studií (Link et al. 2008).

Pomocí kombinace homozygotního mapování a analýzy změn počtu kopií se nám ve spolupráci s další skupinou pracující na myším modelu podařilo vysvětlit genetickou a molekulární podstatu Rotorova syndromu. Vyslovili jsme dále hypotézu, že jaterní exkrece mnoha sloučenin z krve do žluče zahrnuje opětovnou sekreci konjugovaných sloučenin do krve a jejich následnou reabsorpci proteiny OATP1B3 a OATP1B1 v perivenózních hepatocytech, což je mechanismus, který zabraňuje lokální saturaci přenašečů na membráně periportálních hepatocytů.

Publikace s názvem „Complete OATP1B1 and OATP1B3 deficiency causes human Rotor syndrome by interrupting conjugated bilirubin reuptake into the liver.“ byla publikována v roce 2011 v časopise The Journal of Clinical Investigations (IF = 14,152).

Příspěvek autorky dizertace k této studii:

V první části studie jsem se účastnila návrhu designu a přípravy array pro analýzu počtu kopií genu *ABCC2*, provedla hybridizace a účastnila se analýzy dat.

Ve druhé části studie jsem se účastnila genotypování osmi rodin s Rotorovým syndromem, prováděla mutační analýzu, validaci a přesné mapování delecí pomocí real time kvantitativního PCR, long range PCR a Sangerova sekvenování a přispěla k identifikaci jednotlivých haplotypů u rodin s Rotorovým syndromem.

Studium molekulární podstaty izolovaného defektu ATP syntázy.

Mitochondriální onemocnění jsou klinicky, biochemicky a geneticky heterogenní skupinou onemocnění, které vycházejí z defektů mitochondriální biogeneze, defektů komplexů respiračního řetězce nebo defektů jednotlivých mitochondriálních proteinů. Výskyt mitochondriálních onemocnění v populaci se odhaduje na 1: 5000 (Thorburn 2004). Obecnými klinickými projevy jsou poruchy energetického metabolismu související s poklesem produkce ATP. První příznaky se mohou objevit již v novorozeneckém období, ale i v dospělosti a vysokém stáří. V dětském věku jsou nejčastějšími projevy neprospívání, neurologická, kardiologická postižení a svalová hypotonie. Genetickou příčinou mitochondriálních chorob mohou být jak mutace v 37 genech kódovaných mitochondriální DNA, tak mutace v zhruba 1000 jaderně kódovaných mitochondriálních proteinech (http://www.broadinstitute.org/pubs/MitoCarta/human.mitocarta.html).

V Ústavu dědičných metabolických poruch jsou ve spolupráci s laboratoří pro studium mitochondriálních chorob kliniky dětského a dorostového lékařství VFN a Fyziologickým ústavem AV ČR dlouhodobě studovány soubory pacientů s izolovaným defektem ATP syntázy. Mitochondriální ATP syntáza je klíčovým enzymem v mitochondriální energetice. Katalyzuje syntézu ATP v procesu oxidativní fosforylace. Defekty ATP syntázy patří mezi nejzávažnější mitochondriální onemocnění. Mohou být způsobovány mutacemi v mitochondriálně kódovaných genech pro podjednoty *ATP6* nebo *ATP8*, případně mutacemi v jaderně kódovaných genech. Mutace v genech jaderného původu způsobují autozomálně recesivní defekty ATP syntázy, které se obvykle projevují neonatální laktátovou acidózou, encefalokardiomyopatií a 3metylglutakonovou acidurií, v různé míře je postižena centrální nervová soustava (Sperl et al. 2006, Houstek et al. 1999). Biogeneze ATP syntázy je mnohastupňový proces, u savců bylo dosud popsáno jen malé množství homologů chaperonů asemblace ATP syntázy známých u *S. cerevisiae* (Houstek, Kmoch and Zeman 2009, Wang, White and Ackerman 2001).

Vzhledem k tomu, že v době, kdy jsme se začali zabývat studiem pacientů s izolovaným deficitem ATP syntázy, nebyly na komerčních expresních čipech dostupné proby pro mitochondriálně kódované geny, rozhodli jsme se vytvořit vlastní DNA čip pro studium mitochondriálních a lysozomálních poruch. DNA čip nazvaný h-MitoArray obsahoval sadu 1632 genů, z nichž 992 bylo mitochondriálních, 42 lysozomálních, 277 asociovaných s apoptózou a 321 genů účastnících se karcinogeneze, dále 146 housekeeping genů a 10 genů *Arabidopsis thalliana*, které sloužily jako nástroj vnitřní kalibrace a normalizace. Tato h-MitoArray byla nejprve optimalizována a validována a následně využita ke studiu genové exprese u fibroblastů 13 pacientů a 9 kontrol. U všech pacientů byl na biochemické úrovni popsán defekt F_1F_o syntázy, u dvou z nich se jednalo o geneticky charakterizovanou mikrodeleci v genu *ATP6* (Jesina et al. 2004, Cizkova et al. 2008a, Seneca et al. 1996). Na základě porovnání expresních profilů, funkční anotace, metod genového obohacení a analýzy metabolických drah byly vzorky pacientů rozčleněny do tří skupin. První skupina vyčlenila fibroblasty pacientů se známou mutací v mtDNA; expresní profil poukazoval na supresi mitochondriální biogeneze a metabolismu a sníženou expresi genů regulujících G1/S přechod. To podporuje hypotézu o regulaci buněčného cyklu mitochondriemi na transkripční a posttranskripční úrovni (Mandal et al. 2005, Gemin et al. 2005). U druhé skupiny dominovaly v expresním profilu známky aktivované apoptózy a oxidativního stresu, tedy charakteristiky buněčného stárnutí (Shelton

et al. 1999, Stockl et al. 2006). Třetí skupina byla po stránce expresních profilů velmi heterogenní, což odpovídalo i její klinické a biochemické variabilitě. Výsledky expresních studií na h-MitoArray byly verifikovány provedením stejných experimentů na komerční platformě Agilent Human 44k array a byla zjištěna vysoká korelace mezi výsledky obou platforem. Byl tak vytvořen cenově výhodný spolehlivý „custom" DNA čip, který v dané době doplňoval nedostatečné pokrytí mitochondriálních genů u komerčních čipů.

Výsledky studie nazvané „Development of a human mitochondrial oligonucleotide microarray (h-MitoArray) and gene expression analysis of fibroblast cell lines from 13 patients with isolated F(I)F(o) ATP synthase deficiency. " byly publikovány v roce 2008 v časopise BMC Genomics (IF = 3, 926).

Vzhledem k tomu, že analýza genové exprese u pacientů s izolovaným defektem ATP syntázy nepřinesla jasné kandidátní geny pro přímé sekvenování, pokračovali jsme ve studiu pomocí dalších celogenomových technik. U osmi pacientů a jejich zdravých sourozenců a rodičů ze sedmi rodin jsme provedli vazebnou analýzu. Vzhledem k tomu, že všichni pacienti pocházeli z romského etnika s možností efektu zakladatele, bylo provedeno také homozygotní mapování. Tyto analýzy ukázaly jedinou homozygotní oblast na chromozomu 8 sdílenou všemi pacienty. Propojením těchto výsledků s výsledky genové exprese bylo zjištěno, že pouze jediný gen z této oblasti - *TMEM70* měl u všech pacientů sníženou expresi v porovnání s kontrolami. Sekvenační analýzou tohoto genu byla u všech pacientů nalezena homozygotní substituce c.317 -2A>G lokalizovaná v sestřihovém místě exonu 2. Dalšími studiemi bylo prokázáno, žetato mutace vede k abnormálnímu sestřihu a degradaci transkriptu. Pomocí PCR-RFLP analýzy jsme prokázali přítomnost homozygotní mutace u 23 za 25 pacientů. Následné komplementační studie ukázaly obnovení funkce ATP syntázy po vnesení *wt* formy genu *TMEM70* do pacientských fibroblastů; tím byla prokázána kauzalita této mutace. Fylogenetická analýza ukázala přítomnost homologů genu *TMEM70* u vyšších eukaryot, ne však u hub a kvasinek. *TMEM70* byl tak identifikován jako nový faktor účastnící se u vyšších eukaryot biogeneze ATP syntázy. Jeho defekt je zodpovědný za velkou část případů izolovaných defektů ATP syntázy jaderného původu.

Publikace s názvem „*TMEM70* mutations cause isolated ATP synthase deficiency and neonatal mitochondrial encephalocardiomyopathy." byla publikována v roce 2008 v časopise Nature Genetics(IF = 30, 259).

V důsledku publikace této práce byly dosud charakterizovány u čtyř nepříbuzných rodin arabsko – muslimského původu další čtyři další homozygotní mutace v genu *TMEM70* (Spiegel et al. 2011). Naše laboratoř se dále zabývala také charakterizací nově popsaného proteinu (Hejzlarova et al. 2011).

Příspěvek autorky dizertace k této studii:

V rámci první části studie jsem se podílela na optimalizaci výroby a přípravy „custom" DNA čipů pro analýzu genové exprese vybraných genů, podílela se na jejich validaci a byla správcem lokální databázi BASE obsahující informace o celé výrobě a experimentech na tomto druhu čipů. V druhé části studie jsem se prováděla DNA analýzy genu *TMEM70* a připravila konstrukty pro funkční studie.

Studium molekulární podstaty mukopolysacharidózy typu IIIC.

Mukopolysacharidóza typu IIIC (MPSIIIC, Sanfilippo syndrom C, OMIM #252930) je vzácné autozomálně recesivní lysozomální střádavé onemocnění. Patří do širší skupiny mukopolysacharidóz, onemocnění způsobených deficitem enzymů katalyzujících degradaci glykosaminoglykanů. Bylo popsáno celkem jedenáct enzymopatií, skupina mukopolysacharidóz typu III zahrnuje onemocnění způsobená deficitem enzymů degradujících heparan sulfát. U MPSIII jsou klasifikovány čtyři enzymopatie – MPSIIIA (Sanfilippo A, defekt heparan-N-sulfatázy), MPSIIIB (Sanfilippo B, defekt α-N-glukosaminidázy), MPSIIIC (Sanfilippo C, defekt acetyl-koenzymA: α-glukosamin N-acetyltransferázy) a MPSIIID (Sanfilippo D, defekt N-acetylglukosamin 6-sulfatázy).

Acetyl-koenzymA: α-glukosamin N-acetyltransferáza katalyzuje transmembránovou N-acetylaci amino skupiny koncového glukosaminu. Enzym byl částečně purifikován a charakterizován jako transmembránový glykoprotein o velikosti cca 100 kDa, který obsahuje aktivní místo (Ausseil et al. 2006). Prevalence mukopolysacharidózy IIIC byla určena jako 0,07, 0,12 a 0,21 na 100 000 v Austrálii (Meikle et al. 1999), v Portugalsku (Pinto et al. 2004) a v Nizozemí (Poorthuis et al. 1999). Věk nástupu prvních symptomů je mezi čtvrtým a šestým rokem života, mezi symptomy patří změny chování, hyperaktivita a agresivita, následuje postupná mentální retardace, neuropsychiatrické problémy, ztráta sluchu a drobnější viscerální změny. Většina pacientů umírá ještě v dětském věku (BARTSOCAS et al. 1979, KLEIN et al. 1981). MPSIIIC byla poslední enzymopatií skupiny mukopolysacharidóz, jejíž genetická podstata nebyla známa. Vazebná analýza provedená na 44 pacientech a nepostižených příbuzných z 31 rodin vymezila kandidátní oblast o velikosti 8,3 cM na chromozomu 8 (Ausseil et al. 2004).

V Ústavu dědičných metabolických poruch jsme měli možnost studovat celkem pět pacientů ze čtyř nepříbuzných rodin, u nichž bylo onemocnění diagnostikováno na základě biochemického vyšetření aktivity N-acetyltransferázy (VOZNYI et al. 1993). U všech rodin byla provedena vazebná analýza pomocí STR markerů, zároveň bylo na spolupracujícím pracovišti v Montrealu provedeno genotypování 22 mikrosatelitních markerů u 60 pacientů a 44 nepostižených příbuzných. Tyto analýzy zúžily kandidátní oblast na interval 2,6 cM na chromozomu 8 mezi markery D8S1051 a D8S1831. Tento interval obsahoval 32 známých a predikovaných genů.

Díky zavedené metodě přípravy vlastních DNA čipů pomocí robotického přístroje jsme navrhli expresní čip, který obsahoval oligonukleotidové proby genů z kandidátní oblasti. Provedli jsme tak analýzu genové exprese v leukocytech dvou pacientů a čtyř zdravých kontrol. Výsledky ukázaly sníženou expresi transkriptu genu *TMEM76* u obou pacientů ve srovnání se zdravými kontrolami. Gen *TMEM76* byl také vybrán jako kandidátní gen na základě jeho charakteristik lysozomálního transmembránového glykoproteinu o velikosti odpovídající částečně purifikovanému enzymu (Ausseil et al. 2006).

Sekvenováním kandidátního genu byly identifikovány patogenní mutace v genu *TMEM76* nejprve u pěti studovaných českých pacientů, později celkem u třiceti rodin. Jednalo se u 4 nonsense mutace, 14 missense mutací, 3 frameshift mutace a 6 sestřihových mutací. Funkční význam genu *TMEM76* v patogenezi mukopolysacharidózy typu IIIC byl dále potvrzen funkčními expresními studiemi s použitím pacientských fibroblastů.

Nezávisle byly stejné výsledky publikovány další výzkumnou skupinou (Fan et al. 2006), která dospěla k identifikaci kauzálního genu na základě proteomické analýzy. Současné oficiální označení genu *TMEM76* je *HGSNAT* (heparan-α-glucosaminide N-acetyltransferase). Dosud bylo identifikováno více než 50 patogenních mutací v tomto genu (Feldhammer et al. 2009, Canals et al. 2011).

Publikace s názvem „Mutations in *TMEM76** cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome).“ byla publikována v roce 2006 v časopise The American Journal of Human Genetics (IF = 12,629).

Příspěvek autorky dizertace k této studii:

V rámci této studie jsem se podílela na přípravě DNA čipu pro analýzu genové exprese kandidátních genů. Účastnila jsem se optimalizace výroby čipů, optimalizace přípravy vzorků a jejich fluorescenčního značení a provedla všechny hybridizační experimenty. Účastnila jsem se také analýzy výsledků.

ČÁST IV. Souhrn výsledků

Tato dizertační práce představuje využití nových genomických technik, především analýzy pomocí DNA čipů a celoexomového sekvenování pomocí sekvenátoru nové generace. Tyto metody byly aplikovány v případě studia několika vzácných neurologických, hepatologických, mitochondriálních a lysozomálních střádavých chorob s familiárním výskytem.

Souhrnnými výsledky s ohledem na cíle dizertační práce jsou:

1a. Identifikace mutací v genu *DNAJC5* jako kauzální příčiny části případů s adultní formou neuronální ceroidní lipofuscinózy (ANCL) pomocí kombinace vazebné analýzy, analýzy genové exprese, analýzy změn v počtu kopií a celoexomového sekvenování.

1b. Identifikace mutací v genu *PSEN1* a polymorfismu v genu *CTSD* u rodiny původně zařazené do souboru suspektních případů ANCL pomocí metod vazebné analýzy a celoexomového sekvenování.

2a. Vývoj a optimalizace „custom“ DNA čipu pro analýzu změn počtu kopií, jehož využití přispělo k potvrzení odlišné molekulární podstaty dvou vzácných defektů - Rotorova a Dubin-Johnsonova syndromu.

2b. Identifikace delecí a mutací v genech *SLCO1B1* a *SLCO1B3*, které podmiňují vznik Rotorova syndromu pomocí analýzy změn v počtu kopií a homozygotního mapování.

3a. Vývoj a aplikace „custom“ oligonukleotidového DNA čipu pro studium genové exprese u mitochondriálních a lysozomálních střádavých chorob (h-MitoArray) a jeho využití při studiu souboru pacientů s izolovaným defektem ATP syntázy.

3b. Identifikace mutací v genu *TMEM70* jako kauzální příčiny onemocnění u skupiny pacientů s izolovaným deficitem ATP syntázy pomocí vazebné analýzy, analýzy genové exprese a homozygotního mapování.

4. Příspěvek k identifikaci mutací v genu *TMEM76 (HGSNAT)* jako kauzální příčiny mukopolysacharidózy typu IIIC pomocí analýzy genové exprese.

ČÁST V. Praktický význam dosažených výsledků

Část prezentovaných výsledků umožnila také detailnější popis patogeneze některých onemocnění a přispěla k poznání nových fyziologických a patofyziologických mechanismů. Důsledkem některých studií bylo také zavedení nových metod DNA diagnostiky. Význam výsledků jednotlivých studií je shrnut v následujících bodech.

1. Díky popisu kauzálního genu pro adultní formu NCL byly charakterizovány základní patogenetické mechanismy onemocnění. Identifikace genu umožňuje DNA diagnostiku v postižených rodinách a iniciovala další stratifikaci souboru pacientů a rodin s dominantní a sporadickou formou ANCL. Zároveň tento poznatek umožňuje přípravu buněčných a zvířecích modelů pro studium mechanismu neurodegenerace u pacientů a pro studium funkce CSPα obecně.

2. Na základě studia Rotorova syndromu byl popsán nový mechanismus transportu bilirubinu, který zahrnuje opětovnou sekreci konjugovaných sloučenin do krve a jejich následnou reabsorpci proteiny OATP1B3 a OATP1B1. Studie má také širší klinický dopad pro farmakogenomiku v souvislosi s hypersenzitivitou jedinců s mutacemi v popsaných přenašečích na látky jimi transportované. Identifikace mutací také umožňuje diferenciální diagnostiku konjugovaných hyperbilirubinémií.

3. Identifikace mutace v genu *TMEM70* přispěla k poznání biogeneze a regulace funkce ATP syntázy. Umožnila diferenciální diagnostiku izolovaných poruch ATP syntázy a v postižených rodinách nabídla možnost prenatální a postnatální diagnostiky.

4. Na základě identifikace kauzálního genu u mukopolysacharidózy typu IIIC byly připraveny buněčné a zvířecí modely jednak pro studium funkce genu *HGSNAT*, ale i pro širší studium mechanismů neurologického postižení u pacientů. Tyto studie následně vyústily v definici možného terapeutického přístupu pro pacienty s MPSIIIC. I v tomto případě identifikace kauzálního genu umožnila prenatální a postnatální diagnostiku v postižených rodinách.

ČÁST VI. Literatura

Alkan, C., B. Coe & E. Eichler (2011) APPLICATIONS OF NEXT-GENERATION SEQUENCING Genome structural variation discovery and genotyping. *Nature Reviews Genetics,* 12, 363-375.

Altshuler, D., R. Durbin, G. Abecasis, D. Bentley, A. Chakravarti, A. Clark, F. Collins, F. De la Vega, P. Donnelly, M. Egholm, P. Flicek, S. Gabriel, R. Gibbs, B. Knoppers, E. Lander, H. Lehrach, E. Mardis, G. McVean, D. Nickerson, L. Peltonen, A. Schafer, S. Sherry, J. Wang, R. Wilson, D. Deiros, M. Metzker, D. Muzny, J. Reid, D. Wheeler, J. Li, M. Jian, G. Li, R. Li, H. Liang, G. Tian, B. Wang, W. Wang, H. Yang, X. Zhang, H. Zheng, L. Ambrogio, T. Bloom, K. Cibulskis, T. Fennell, D. Jaffe, E. Shefler, C. Sougnez, N. Gormley, S. Humphray, Z. Kingsbury, P. Koko-Gonzales, J. Stone, K. McKernan, G. Costa, J. Ichikawa, C. Lee, R. Sudbrak, T. Borodina, A. Dahl, A. Davydov, P. Marquardt, F. Mertes, W. Nietfeld, P. Rosenstiel, S. Schreiber, A. Soldatov, B. Timmermann, M. Tolzmann, J. Affourtit, D. Ashworth, S. Attiya, M. Bachorski, E. Buglione, A. Burke, A. Caprio, C. Celone, S. Clark, D. Conners, B. Desany, L. Gu, L. Guccione, K. Kao, A. Kebbel, J. Knowlton, M. Labrecque, L. McDade, C. Mealmaker, M. Minderman, A. Nawrocki, F. Niazi, K. Pareja, R. Ramenani, D. Riches, W. Song, C. Turcotte, S. Wang, D. Dooling, L. Fulton, R. Fulton, G. Weinstock, et al. (2010) A map of human genome variation from population-scale sequencing. *Nature,* 467, 1061-1073.

Antonarakis, S. & J. Beckmann (2006) Opinion - Mendelian disorders deserve more attention. *Nature Reviews Genetics,* 7, 277-282.

Arsov, T., K. Smith, J. Damiano, S. Franceschetti, L. Canafoglia, C. Bromhead, E. Andermann, D. Vears, P. Cossette, S. Rajagopalan, A. McDougall, V. Sofia, M. Farrell, U. Aguglia, A. Zini, S. Meletti, M. Morbin, S. Mullen, F. Andermann, S. Mole, M. Bahlo & S. Berkovic (2011) Kufs Disease, the Major Adult Form of Neuronal Ceroid Lipofuscinosis, Caused by Mutations in CLN6. *American Journal of Human Genetics,* 88, 566-573.

Ausseil, J., J. Loredo-Osti, A. Verner, C. Darmond-Zwaig, I. Maire, B. Poorthuis, O. van Diggelen, T. Hudson, T. Fujiwara, K. Morgan & A. Pshezhetsky (2004) Localisation of a gene for mucopolysaccharidosis IIIC to the pericentromeric region of chromosome 8. *Journal of Medical Genetics,* 41, 941-944.

Ausseil, K., K. Landry, V. Seyrantepe, S. Trudel, A. Mazur, F. Lapointe & A. Pshezhetsky (2006) An acetylated 120-kDa lysosomal transmembrane protein is absent from mucopolysaccharidosis IIIC fibroblasts: A candidate molecule for MPS IIIC. *Molecular Genetics and Metabolism,* 87, 22-31.

Ayme, S., A. Kole & S. Groft (2008) Empowerment of patients: lessons from the rare diseases community. *Lancet,* 371, 2048-2051.

BARMEIR, S., J. BARON, U. SELIGSON, F. GOTTESFELD, R. LEVY & T. GILAT (1982) TC-99M-HIDA CHOLESCINTIGRAPHY IN DUBIN-JOHNSON AND ROTOR SYNDROMES. *Radiology,* 142, 743-746.

BARTSOCAS, C., H. GROBE, J. VANDEKAMP, K. FIGURA, H. KRESSE, U. KLEIN & M. GIESBERTS (1979) SANFILIPPO TYPE-C DISEASE - CLINICAL FINDINGS IN 4 PATIENTS WITH A NEW VARIANT OF MUCOPOLYSACCHARIDOSIS-III. *European Journal of Pediatrics,* 130, 251-258.

Bekris, L., C. Yu, T. Bird & D. Tsuang (2010) Genetics of Alzheimer Disease. *Journal of Geriatric Psychiatry and Neurology,* 23, 213-227.

Benitez, B., D. Alvarado, Y. Cai, K. Mayo, S. Chakraverty, J. Norton, J. Morris, M. Sands, A. Goate & C. Cruchaga (2011) Exome-Sequencing Confirms *DNAJC5* Mutations as Cause of Adult Neuronal Ceroid-Lipofuscinosis. *Plos One,* 6.

Bittles, A. (2001) Consanguinity and its relevance to clinical genetics. *Clin Genet,* 60, 89-98.

Blow, N. (2008) DNA sequencing: generation next-next. *Nature Methods,* 5, 267-+.

Bodzioch, M., E. Orso, T. Klucken, T. Langmann, L. Bottcher, W. Diederich, W. Drobnik, S. Barlage, C. Buchler, M. Porsch-Ozcurumez, W. Kaminski, H. Hahmann, K. Oette, G. Rothe, C. Aslanidis, K. Lackner & G. Schmitz (1999) The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. *Nature Genetics,* 22, 347-351.

Boehme, D. H., J. C. Cottrell, S. C. Leonberg & W. Zeman (1971) A dominant form of neuronal ceroid-lipofuscinosis. *Brain,* 94, 745-60.

Bolze, A., M. Byun, D. McDonald, N. Morgan, A. Abhyankar, L. Premkumar, A. Puel, C. Bacon, F. Rieux-Laucat, K. Pang, A. Britland, L. Abel, A. Cant, E. Maher, S. Riedl, S. Hambleton & J. Casanova (2010) Whole-Exome-Sequencing-Based Discovery of Human FADD Deficiency. *American Journal of Human Genetics,* 87, 873-881.

Bras, J., A. Verloes, S. Schneider, S. Mole & R. Guerreiro (2012) Mutation of the parkinsonism gene ATP13A2 causes neuronal ceroid-lipofuscinosis. *Human Molecular Genetics,* 21, 2646-2650.

Brooks-Wilson, A., M. Marcil, S. Clee, L. Zhang, K. Roomp, M. van Dam, L. Yu, C. Brewer, J. Collins, H. Molhuizen, O. Loubser, B. Ouelette, K. Fichter, K. Ashbourne-Excoffon, C. Sensen, S. Scherer, S. Mott, M. Denis, D. Martindale, J. Frohlich, K. Morgan, B. Koop, S. Pimstone, J. Kastelein, J. Genest & M. Hayden (1999) Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nature Genetics,* 22, 336-345.

Burneo, J., T. Arnold, C. Palmer, R. Kuzniecky, S. Oh & E. Faught (2003) Adult-onset neuronal ceroid lipofuscinosis (Kufs disease) with autosomal dominant inheritance in Alabama. *Epilepsia,* 44, 841-846.

Caliskan, M., J. Chong, L. Uricchio, R. Anderson, P. Chen, C. Sougnez, K. Garimella, S. Gabriel, M. DePristo, K. Shakir, D. Matern, S. Das, D. Waggoner, D. Nicolae & C. Ober (2011) Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13. *Human Molecular Genetics,* 20, 1285-1289.

Canals, I., S. Elalaoui, M. Pineda, V. Delgadillo, M. Szlago, I. Jaouad, A. Sefiani, A. Chabas, M. Coll, D. Grinberg & L. Vilageliu (2011) Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles. *Clinical Genetics,* 80, 367-374.

Chamberlain, L. & R. Burgoyne (1998) The cysteine-string domain of the secretory vesicle cysteine-string protein is required for membrane targeting. *Biochemical Journal,* 335, 205-209.

Cirulli, E. & D. Goldstein (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics,* 11, 415-425.

Citron, M., D. Westaway, W. Xia, G. Carlson, T. Diehl, G. Levesque, K. JohnsonWood, M. Lee, P. Seubert, A. Davis, D. Kholodenko, R. Motter, R. Sherrington, B. Perry, H. Yao, R. Strome, I. Lieberburg, J. Rommens, S. Kim, D. Schenk, P. Fraser, P. Hyslop & D. Selkoe (1997) Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. *Nature Medicine,* 3, 67-72.

Cizkova, A., V. Stranecky, R. Ivanek, H. Hartmannova, L. Noskova, L. Piherova, M. Tesarova, H. Hansikova, T. Honzik, J. Zeman, P. Divina, A. Potocka, J. Paul, W. Sperl, J. Mayr, S. Seneca, J. Houstek & S. Kmoch (2008a) Development of a human mitochondrial oligonucleotide microarray (h-MitoArray) and gene expression analysis of fibroblast cell lines from 13 patients with isolated F(I)F(o) ATP synthase deficiency. *Bmc Genomics,* 9.

Cizkova, A., V. Stranecky, J. Mayr, M. Tesarova, V. Havlickova, J. Paul, R. Ivanek, A. Kuss, H. Hansikova, V. Kaplanova, M. Vrbacky, H. Hartmannova, L. Noskova, T. Honzik, Z. Drahota, M. Magner, K. Hejzlarova, W. Sperl, J. Zeman, J. Houstek & S. Kmoch (2008b) *TMEM70* mutations cause isolated ATP synthase deficiency and neonatal mitochondrial encephalocardiomyopathy. *Nature Genetics,* 40, 1288-1290.

Collin, R. W., C. Safieh, K. W. Littink, S. A. Shalev, H. J. Garzozi, L. Rizel, A. H. Abbasi, F. P. Cremers, A. I. den Hollander, B. J. Klevering & T. Ben-Yosef (2010) Mutations in C2ORF71 cause autosomal-recessive retinitis pigmentosa. *Am J Hum Genet,* 86, 783-8.

Consortium, I. H. G. S. (2004) Finishing the euchromatic sequence of the human genome. *Nature,* 431, 931-45.

Dalca, A. & M. Brudno (2010) Genome variation discovery with high-throughput sequencing data. *Briefings in Bioinformatics,* 11, 3-14.

Day, I. (2010) dbSNP in the Detail and Copy Number Complexities. *Human Mutation,* 31, 2-4.

Ertekin-Taner, N. (2010) Genetics of Alzheimer disease in the pre- and post-GWAS era. *Alzheimers Res Ther,* 2, 3.

Fan, J., A. Oliphant, R. Shen, B. Kermani, F. Garcia, K. Gunderson, M. Hansen, F. Steemers, S. Butler, P. Deloukas, L. Galver, S. Hunt, C. McBride, M. Bibikova, T. Rubano, J. Chen, E. Wickham, D. Doucet, W. Chang, D. Campbell, B. Zhang, S. Kruglyak, D. Bentley, J. Haas, P. Rigault, L. Zhou, J. Stuelpnagel & M. Chee (2003) Highly parallel SNP genotyping. *Cold Spring Harbor Symposia on Quantitative Biology,* 68, 69-78.

Fan, X., H. Zhang, S. Zhang, R. Bagshaw, M. Tropak, J. Callahan & D. Mahuran (2006) Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). *American Journal of Human Genetics,* 79, 738-744.

Feldhammer, M., S. Durand, L. Mrazova, R. Boucher, R. Laframboise, R. Steinfeld, J. Wraith, H. Michelakakis, O. van Diggelen, M. Hrebicek, S. Kmoch & A. Pshezhetsky (2009) Sanfilippo Syndrome Type C: Mutation Spectrum in the Heparan Sulfate Acetyl-CoA: alpha-Glucosaminide N-Acetyltransferase (HGSNAT) Gene. *Human Mutation,* 30, 918-925.

Fernandez-Chacon, R., M. Wolfel, H. Nishimune, L. Tabares, F. Schmitz, M. Castellano-Munoz, C. Rosenmund, M. Montesinos, J. Sanes, R. Schneggenburger & T. Sudhof (2004) The synaptic vesicle protein CSP alpha prevents presynaptic degeneration. *Neuron,* 42, 237-251.

FERRER, I., T. ARBIZU, J. PENA & J. SERRA (1980) GOLGI AND ULTRASTRUCTURAL-STUDY OF A DOMINANT FORM OF KUFS DISEASE. *Journal of Neurology,* 222, 183-190.

FRETZAYAS, A., A. GAROUFI, C. MOUTSOURIS & T. KARPATHIOS (1994) CHOLESCINTIGRAPHY IN THE DIAGNOSIS OF ROTOR SYNDROME. *Journal of Nuclear Medicine,* 35, 1048-1050.

Frikke-Schmidt, R., B. Nordestgaard, G. Jensen & A. Tybjaerg-Hansen (2004) Genetic variation in ABC transporter A1 contributes to HDL cholesterol in the general population. *Journal of Clinical Investigation,* 114, 1343-1353.

Gemin, A., S. Sweet, T. Preston & G. Singh (2005) Regulation of the cell cycle in response to inhibition of mitochondrial generated energy. *Biochemical and Biophysical Research Communications,* 332, 1122-1132.

Goebel, H. & K. Wisniewski (2004) Current state of clinical and morphological features in human NCL. *Brain Pathology,* 14, 61-69.

Golan, M., M. Styczynska, K. Jozwiak, J. Walecki, A. Maruszak, J. Pniewski, R. Lugiewicz, S. Filipek, C. Zekanowski & M. Barcikowska (2007) Early-onset Alzheimer's disease with a de novo mutation in the presenilin 1 gene. *Experimental Neurology,* 208, 264-268.

Greaves, J. & L. Chamberlain (2006) Dual role of the cysteine-string domain in membrane binding and palmitoylation-dependent sorting of the molecular chaperone cysteine-string protein. *Molecular Biology of the Cell,* 17, 4748-4759.

Greaves, J., C. Salaun, Y. Fukata, M. Fukata & L. H. Chamberlain (2008) Palmitoylation and membrane interactions of the neuroprotective chaperone cysteine-string protein. *J Biol Chem,* 283, 25014-26.

Gunderson, K., S. Kruglyak, M. Graige, F. Garcia, B. Kermani, C. Zhao, D. Che, T. Dickinson, E. Wickham, J. Bierle, D. Doucet, M. Milewski, R. Yang, C. Siegmund, J. Haas, L. Zhou, A. Oliphant, J. Fan, S. Barnard & M. Chee (2004) Decoding randomly ordered DNA arrays. *Genome Research,* 14, 870-877.

Hagenbuch, B. & C. Gui (2008) Xenobiotic transporters of the human organic anion transporting polypeptides (OATP) family. *Xenobiotica,* 38, 778-801.

Hagenbuch, B. & P. Meier (2004) Organic anion transporting polypeptides of the OATP/SLC21 family: phylogenetic classification as OATP/SLCO superfamily, new nomenclature and molecular/functional properties. *Pflugers Archiv-European Journal of Physiology,* 447, 653-665.

Hayashi, S. & T. Umeda (2008) 35 years of Japanese policy on rare diseases. *Lancet,* 372, 889-890.

Hejzlarova, K., M. Tesarova, A. Vrbacka-Cizkova, M. Vrbacky, H. Hartmannova, V. Kaplanova, L. Noskova, H. Kratochvilova, J. Buzkova, V. Havlickova, J. Zeman, S. Kmoch & J. Houstek (2011) Expression and processing of the *TMEM70* protein. *Biochimica Et Biophysica Acta-Bioenergetics,* 1807, 144-149.

Houstek, J., P. Klement, D. Floryk, H. Antonicka, J. Hermanska, M. Kalous, H. Hansikova, H. Houst'kova, S. Chowdhury, S. Rosipal, S. Kmoch, L. Stratilova & J. Zeman (1999) A novel deficiency of mitochondrial ATPase of nuclear origin. *Human Molecular Genetics,* 8, 1967-1974.

Houstek, J., S. Kmoch & J. Zeman (2009) *TMEM70* protein - A novel ancillary factor of mammalian ATP synthase. *Biochimica Et Biophysica Acta-Bioenergetics,* 1787, 529-532.

Hrebicek, M., T. Jirasek, H. Hartmannova, L. Noskova, V. Stranecky, R. Ivanek, S. Kmoch, D. Cebecauerova, L. Vitek, M. Mikulecky, I. Subhanova, P. Hozak & M. Jirsa (2007) Rotor-type hyperbilirubinaemia has no defect in the canalicular bilirubin export pump. *Liver International,* 27, 485-491.

Hrebicek, M., L. Mrazova, V. Seyrantepe, S. Durand, N. Roslin, L. Noskova, H. Hartmannova, R. Ivanek, A. Cizkova, H. Poupetova, J. Sikora, J. Urinovska, V. Stranecky, J. Zeman, P. Lepage, D. Roquis, A. Verner, J. Ausseil, C. Beesley, I. Maire, B. Poorthuis, J. van de Kamp, O. van Diggelen, R. Wevers, T. Hudson, T. Fujiwara, J. Majewski, K. Morgan, S. Kmoch & A. Pshezhetsky (2006) Mutations in *TMEM76** cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *American Journal of Human Genetics,* 79, 807-819.

Huebner, A. K., M. Gandia, P. Frommolt, A. Maak, E. M. Wicklein, H. Thiele, J. Altmüller, F. Wagner, A. Viñuela, L. A. Aguirre, F. Moreno, H. Maier, I. Rau, S. Giesselmann, G. Nürnberg, A. Gal, P. Nürnberg, C. A. Hübner, I. del Castillo & I. Kurth (2011) Nonsense mutations in SMPX, encoding a protein responsive to physical force, result in X-chromosomal hearing loss. *Am J Hum Genet,* 88, 621-7.

Hutton, M. & J. Hardy (1997) The presenilins and Alzheimer's disease. *Human Molecular Genetics,* 6, 1639-1646.

Iseri, S. U., A. W. Wyatt, G. Nürnberg, C. Kluck, P. Nürnberg, G. E. Holder, E. Blair, A. Salt & N. K. Ragge (2010) Use of genome-wide SNP homozygosity mapping in small pedigrees to identify new mutations in VSX2 causing recessive microphthalmia and a semidominant inner retinal dystrophy. *Hum Genet,* 128, 51-60.

Jalanko, A. & T. Braulke (2009) Neuronal ceroid lipofuscinoses. *Biochimica Et Biophysica Acta-Molecular Cell Research,* 1793, 697-709.

Jesina, P., M. Tesarova, D. Fornuskova, A. Vojtiskova, P. Pecina, H. Hansikova, V. Kaplanova, J. Zeman & J. Houstek (2004) Molecular mechanisms and biochemical consequences of 9205delTA mutation in ATP6 gene. *Biochimica Et Biophysica Acta-Bioenergetics,* 1657, 85-85.

Johnson, J., J. Mandrioli, M. Benatar, Y. Abramzon, V. Van Deerlin, J. Trojanowski, J. Gibbs, M. Brunetti, S. Gronka, J. Wuu, J. Ding, L. McCluskey, M. Martinez-Lage, D. Falcone, D. Hernandez, S. Arepalli, S. Chong, J. Schymick, J. Rothstein, F. Landi, Y. Wang, A. Calvo, G. Mora, M. Sabatelli, M. Monsurro, S. Battistini, F. Salvi, R. Spataro, P. Sola, G. Borghero, G. Galassi, S. Scholz, J. Taylor, G. Restagno, A. Chio, B. Traynor & I. Consortium (2010) Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron,* 68, 857-864.

Josephson, S., R. Schmidt, P. Millsap, D. McManus & J. Morris (2001) Autosomal dominant Kufs' disease: a cause of early onset dementia. *Journal of the Neurological Sciences,* 188, 51-60.

Kartenbeck, J., U. Leuschner, R. Mayer & D. Keppler (1996) Absence of the canalicular isoform of the MRP gene-encoded conjugate export pump from the hepatocytes in Dubin-Johnson syndrome. *Hepatology,* 23, 1061-1066.

KLEIN, U., J. VANDEKAMP, K. VONFIGURA & R. POHLMANN (1981) SANFILIPPO SYNDROME TYPE-C - ASSAY FOR ACETYL-COA - ALPHA-GLUCOSAMINIDE N-ACETYLTRANSFERASE IN LEUKOCYTES FOR DETECTION OF HOMOZYGOUS AND HETEROZYGOUS INDIVIDUALS. *Clinical Genetics,* 20, 55-59.

Koiwai, O., M. Nishizawa, K. Hasada, S. Aono, Y. Adachi, N. Mamiya & H. Sato (1995) Gilbert's syndrome is caused by a heterozygous missense mutation in the gene for bilirubin UDP-glucuronosyltransferase. *Hum Mol Genet,* 4, 1183-6.

Ku, C., N. Naidoo & Y. Pawitan (2011) Revisiting Mendelian disorders through exome sequencing. *Human Genetics,* 129, 351-370.

Kuhlenbaumer, G., J. Hullmann & S. Appenzeller (2011) Novel Genomic Techniques Open New Avenues in the Analysis of Monogenic Disorders. *Human Mutation,* 32, 144-151.

Lalonde, E., S. Albrecht, K. Ha, K. Jacob, N. Bolduc, C. Polychronakos, P. Dechelotte, J. Majewski & N. Jabado (2010) Unexpected Allelic Heterogeneity and Spectrum of Mutations in Fowler Syndrome Revealed by Next-Generation Exome Sequencing. *Human Mutation,* 31, 918-923.

Lander, E., L. Linton, B. Birren, C. Nusbaum, M. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. Waterston, R. Wilson, L. Hillier, J. McPherson, M. Marra, E. Mardis, L. Fulton, A. Chinwalla, K. Pepin, W. Gish, S. Chissoe, M. Wendl, K. Delehaunty, T. Miner, A. Delehaunty, J. Kramer, L. Cook, R. Fulton, D. Johnson, P. Minx, S. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.

Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, et al. (2001) Initial sequencing and analysis of the human genome. *Nature,* 409, 860-921.

Langheinrich, T., C. Romanowski, S. Wharton & M. Hadjivassiliou (2011) PRESENILIN-1 MUTATION ASSOCIATED WITH AMNESIA, ATAXIA, AND MEDIAL TEMPORAL LOBE T2 SIGNAL CHANGES. *Neurology,* 76, 1435-1436.

Link, E., S. Parish, J. Armitage, L. Bowman, S. Heath, F. Matsuda, I. Gut, M. Lathrop, R. Collins & S. C. Grp (2008) *SLCO1B1* variants and statin-induced myopathy - A genomewide study. *New England Journal of Medicine,* 359, 789-799.

Lupski, J. R., J. G. Reid, C. Gonzaga-Jauregui, D. Rio Deiros, D. C. Chen, L. Nazareth, M. Bainbridge, H. Dinh, C. Jing, D. A. Wheeler, A. L. McGuire, F. Zhang, P. Stankiewicz, J. J. Halperin, C. Yang, C. Gehman, D. Guo, R. K. Irikat, W. Tom, N. J. Fantin, D. M. Muzny & R. A. Gibbs (2010) Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med,* 362, 1181-91.

Maher, B. (2008) Personal genomes: The case of the missing heritability. *Nature,* 456, 18-21.

Mamanova, L., A. Coffey, C. Scott, I. Kozarewa, E. Turner, A. Kumar, E. Howard, J. Shendure & D. Turner (2010) Target-enrichment strategies for next-generation sequencing. *Nature Methods,* 7, 111-118.

Mandal, S., P. Guptan, E. Owusu-Ansah & U. Banerjee (2005) Mitochondrial regulation of cell cycle progression during development as by the tenured mutation revealed in Drosophila. *Developmental Cell,* 9, 843-854.

Manolio, T., F. Collins, N. Cox, D. Goldstein, L. Hindorff, D. Hunter, M. McCarthy, E. Ramos, L. Cardon, A. Chakravarti, J. Cho, A. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. Rotimi, M. Slatkin, D. Valle, A. Whittemore, M. Boehnke, A. Clark, E. Eichler, G. Gibson, J. Haines, T. Mackay, S. McCarroll & P. Visscher (2009) Finding the missing heritability of complex diseases. *Nature,* 461, 747-753.

Mardis, E. (2008) The impact of next-generation sequencing technology on genetics. *Trends in Genetics,* 24, 133-141.

Meikle, P., J. Hopwood, A. Clague & W. Carey (1999) Prevalence of lysosomal storage disorders. *Jama-Journal of the American Medical Association,* 281, 249-254.

Metzker, M. (2010) APPLICATIONS OF NEXT-GENERATION SEQUENCING Sequencing technologies - the next generation. *Nature Reviews Genetics,* 11, 31-46.

Nachman, M. & S. Crowell (2000) Estimate of the mutation rate per nucleotide in humans. *Genetics,* 156, 297-304.

Ng, P., S. Levy, J. Huang, T. Stockwell, B. Walenz, K. Li, N. Axelrod, D. Busam, R. Strausberg & J. Venter (2008) Genetic Variation in an Individual Human Exome. *Plos Genetics,* 4.

Ng, S. B., A. W. Bigham, K. J. Buckingham, M. C. Hannibal, M. J. McMillin, H. I. Gildersleeve, A. E. Beck, H. K. Tabor, G. M. Cooper, H. C. Mefford, C. Lee, E. H. Turner, J. D. Smith, M. J. Rieder, K. Yoshiura, N.

Matsumoto, T. Ohta, N. Niikawa, D. A. Nickerson, M. J. Bamshad & J. Shendure (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet,* 42, 790-3.

Nijssen, P., G. Brekelmans & R. Roos (2009) Electroencephalography in autosomal dominant adult neuronal ceroid lipofuscinosis. *Clinical Neurophysiology,* 120, 1782-1786.

Nijssen, P. C., E. Brusse, A. C. Leyten, J. J. Martin, J. L. Teepen & R. A. Roos (2002) Autosomal dominant adult neuronal ceroid lipofuscinosis: parkinsonism due to both striatal and nigral dysfunction. *Mov Disord,* 17, 482-7.

Nijssen, P. C., C. Ceuterick, O. P. van Diggelen, M. Elleder, J. J. Martin, J. L. Teepen, J. Tyynelä & R. A. Roos (2003) Autosomal dominant adult neuronal ceroid lipofuscinosis: a novel form of NCL with granular osmiophilic deposits without palmitoyl protein thioesterase 1 deficiency. *Brain Pathol,* 13, 574-81.

Nijssen, P. C. G. 2011. Autosomal dominant adult neuronal ceroid lipofuscinosis. Faculty of Medicine / Leiden University Medical Center (LUMC), Leiden University.

Noskova, L., V. Stranecky, H. Hartmannova, A. Pristoupilova, V. Baresova, R. Ivanek, H. Hulkova, H. Jahnova, J. van der Zee, J. Staropoli, K. Sims, J. Tyynela, C. Van Broeckhoven, P. Nijssen, S. Mole, M. Elleder & S. Kmoch (2011) Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis (vol 89, pg 241, 2011). *American Journal of Human Genetics,* 89, 589-589.

PALMER, D., I. FEARNLEY, J. WALKER, N. HALL, B. LAKE, L. WOLFE, M. HALTIA, R. MARTINUS & R. JOLLY (1992) MITOCHONDRIAL ATP SYNTHASE SUBUNIT C STORAGE IN THE CEROID-LIPOFUSCINOSES (BATTEN DISEASE). *American Journal of Medical Genetics,* 42, 561-567.

Palmer, D., R. Jolly, H. vanMil, J. Tyynela & V. Westlake (1997) Different patterns of hydrophobic protein storage in different forms of neuronal ceroid lipofuscinosis (NCL, Batten disease). *Neuropediatrics,* 28, 45-48.

Paulusma, C., M. Kool, P. Bosma, G. Scheffer, F. terBorg, R. Scheper, G. Tytgat, P. Borst, F. Baas & P. Elferink (1997) A mutation in the human canalicular multispecific organic anion transporter gene causes the Dubin-Johnson syndrome. *Hepatology,* 25, 1539-1542.

Pearson, C., K. Edamura & J. Cleary (2005) Repeat instability: Mechanisms of dynamic mutations. *Nature Reviews Genetics,* 6, 729-742.

Piccini, A., G. Zanusso, R. Borghi, C. Noviello, S. Monaco, R. Russo, G. Damonte, A. Armirotti, M. Gelati, R. Giordano, P. Zambenedetti, C. Russo, B. Ghetti & M. Tabaton (2007) Association of a presenilin 1 S170F mutation with a novel Alzheimer disease molecular phenotype. *Archives of Neurology,* 64, 738-745.

Pinto, R., C. Caseiro, M. Lemos, L. Lopes, A. Fontes, H. Ribeiro, E. Pinto, E. Silva, S. Rocha, A. Marcao, I. Ribeiro, C. Lacerda, G. Ribeiro, O. Amaral & M. Miranda (2004) Prevalence of lysosomal storage diseases in Portugal. *European Journal of Human Genetics,* 12, 87-92.

Poorthuis, B., R. Wevers, W. Kleijer, J. Groener, J. de Jong, S. van Weely, K. Niezen-Koning & O. van Diggelen (1999) The frequency of lysosomal storage diseases in The Netherlands. *Human Genetics,* 105, 151-156.

Remuzzi, G. & S. Garattini (2008) Rare diseases: what's next? *Lancet,* 371, 1978-1979.

Remuzzi, G. & A. Schieppati (2011) Why rare diseases? *Annali Dell Istituto Superiore Di Sanita,* 47, 94-97.

Renbaum, P., E. Kellerman, R. Jaron, D. Geiger, R. Segel, M. Lee, M. C. King & E. Levy-Lahad (2009) Spinal muscular atrophy with pontocerebellar hypoplasia is caused by a mutation in the VRK1 gene. *Am J Hum Genet,* 85, 281-9.

S., S., M. V., K. E., C. P., L. J. & D. L. 2010. Metabolická příručka 2010. Praha: Všeobecná fakultní nemocnice v Praze.

Sandhu, M., M. Weedon, K. Fawcett, J. Wasson, S. Debenham, A. Daly, H. Lango, T. Frayling, R. Neumann, R. Sherva, I. Blech, P. Pharoah, C. Palmer, C. Kimber, R. Tavendale, A. Morris, M. McCarthy, M. Walker, G. Hitman, B. Glaser, M. Permutt, A. Hattersley, N. Wareham & I. Barroso (2007) Common variants in WFS1 confer risk of type 2 diabetes. *Nature Genetics,* 39, 951-953.

SCHENA, M., D. SHALON, R. DAVIS & P. BROWN (1995) QUANTITATIVE MONITORING OF GENE-EXPRESSION PATTERNS WITH A COMPLEMENTARY-DNA MICROARRAY. *Science,* 270, 467-470.

Scheuner, D., C. Eckman, M. Jensen, X. Song, M. Citron, N. Suzuki, T. Bird, J. Hardy, M. Hutton, W. Kukull, E. Larson, E. LevyLahad, M. Viitanen, E. Peskind, P. Poorkaj, G. Schellenberg, R. Tanzi, W. Wasco, L. Lannfelt, D. Selkoe & S. Younkin (1996) Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nature Medicine,* 2, 864-870.

Schieppati, A., J. Henter, E. Daina & A. Aperia (2008) Why rare diseases are an important medical and social issue. *Lancet,* 371, 2039-2041.

Schuur, M., M. Ikram, J. van Swieten, A. Isaacs, J. Vergeer-Drop, A. Hofman, B. Oostra, M. Breteler & C. van Duijn (2011) Cathepsin D gene and the risk of Alzheimer's disease: A population-based study and meta-analysis. *Neurobiology of Aging,* 32, 1607-1614.

Seehafer, S. & D. Pearce (2006) You say lipofuscin, we say ceroid: Defining autofluorescent storage material. *Neurobiology of Aging,* 27, 576-588.

Selkoe, D. & M. Wolfe (2007) Presenilin: Running with scissors in the membrane. *Cell,* 131, 215-221.

Seneca, S., M. Abramowicz, W. Lissens, M. Muller, E. Vamos & L. deMeirleir (1996) A mitochondrial DNA microdeletion in a newborn girl with transient lactic acidosis. *Journal of Inherited Metabolic Disease,* 19, 115-118.

Senior, K. (1999) Studies on rare genetic diseases: a good use of funds? *Lancet,* 354, 1798-1798.

Shelton, D., E. Chang, P. Whittier, D. Choi & W. Funk (1999) Microarray analysis of replicative senescence. *Current Biology,* 9, 939-945.

Singleton, A. (2011) Exome sequencing: a transformative technology. *Lancet Neurology,* 10, 942-946.

Smigielski, E., K. Sirotkin, M. Ward & S. Sherry (2000) dbSNP: a database of single nucleotide polymorphisms. *Nucleic Acids Research,* 28, 352-355.

Smith, K., J. Damiano, S. Franceschetti, S. Carpenter, L. Canafoglia, M. Morbin, G. Rossi, D. Pareyson, S. Mole, J. Staropoli, K. Sims, J. Lewis, W. Lin, D. Dickson, H. Dahl, M. Bahlo & S. Berkovic (2012) Strikingly Different Clinicopathological Phenotypes Determined by Progranulin-Mutation Dosage. *American Journal of Human Genetics,* 90, 1102-1107.

Snider, B., J. Norton, M. Coats, S. Chakraverty, C. Hou, R. Jervis, C. Lendon, A. Goate, D. McKeel & J. Morris (2005) Novel presenilin 1 mutation (S170F) causing Alzheimer disease with Lewy bodies in the third decade of life. *Archives of Neurology,* 62, 1821-1830.

SPENCE, J., R. PERCIACCANTE, G. GREIG, H. WILLARD, D. LEDBETTER, J. HEJTMANCIK, M. POLLACK, W. OBRIEN & A. BEAUDET (1988) UNIPARENTAL DISOMY AS A MECHANISM FOR HUMAN GENETIC-DISEASE. *American Journal of Human Genetics,* 42, 217-226.

Sperl, W., P. Jesina, J. Zeman, J. Mayr, L. DeMeirleir, R. VanCoster, A. Pickova, H. Hansikova, H. Houst'kova, Z. Krejcik, J. Koch, J. Smet, W. Muss, E. Holme & J. Houstek (2006) Deficiency of mitochondrial ATP synthase of nuclear genetic origin. *Neuromuscular Disorders,* 16, 821-829.

Spiegel, R., M. Khayat, S. Shalev, Y. Horovitz, H. Mandel, E. Hershkovitz, F. Barghuti, A. Shaag, A. Saada, S. Korman, O. Elpeleg & I. Yatsiv (2011) *TMEM70* mutations are a common cause of nuclear encoded ATP synthase assembly defect: further delineation of a new syndrome. *Journal of Medical Genetics,* 48, 177-182.

Staropoli, J., A. Karaa, E. Lim, A. Kirby, N. Elbalalesy, S. Romansky, K. Leydiker, S. Coppel, R. Barone, W. Xin, M. MacDonald, J. Abdenur, M. Daly, K. Sims & S. Cotman (2012) A Homozygous Mutation in KCTD7 Links Neuronal Ceroid Lipofuscinosis to the Ubiquitin-Proteasome System. *American Journal of Human Genetics,* 91, 202-208.

Stastna, S., V. Maskova, E. Kostalova, P. Chrastina, J. Ledvinova & L. Dvořákova. 2010. Metabolická příručka 2010. Praha: Všeobecná fakultní nemocnice v Praze.

Steemers, F., W. Chang, G. Lee, D. Barker, R. Shen & K. Gunderson (2006) Whole-genome genotyping with the single-base extension assay. *Nature Methods,* 3, 31-33.

Stockl, P., E. Hutter, W. Zwerschke & P. Jansen-Durr (2006) Sustained inhibition of oxidative phosphorylation impairs cell proliferation and induces premature senescence in human fibroblasts. *Experimental Gerontology,* 41, 674-682.

Summerer, D. (2009) Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. *Genomics,* 94, 363-368.

Tennessen, J., A. Bigham, T. O'Connor, W. Fu, E. Kenny, S. Gravel, S. McGee, R. Do, X. Liu, G. Jun, H. Kang, D. Jordan, S. Leal, S. Gabriel, M. Rieder, G. Abecasis, D. Altshuler, D. Nickerson, E. Boerwinkle, S. Sunyaev, C. Bustamante, M. Bamshad, J. Akey, B. GO, S. GO, N. E. S. Project, B. GO, S. GO & N. E. S. Project (2012) Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science,* 337, 64-69.

Thorburn, D. (2004) Mitochondrial disorders: Prevalence, myths and advances. *Journal of Inherited Metabolic Disease,* 27, 349-362.

Vallance, H. & J. Ford (2003) Carrier testing for autosomal-recessive disorders. *Critical Reviews in Clinical Laboratory Sciences,* 40, 473-497.

Velinov, M., N. Dolzhanskaya, M. Gonzalez, E. Powell, I. Konidari, W. Hulme, J. F. Staropoli, W. Xin, G. Y. Wen, R. Barone, S. H. Coppel, K. Sims, W. T. Brown & S. Züchner (2012) Mutations in the gene *DNAJC5* cause autosomal dominant Kufs disease in a proportion of cases: study of the Parry family and 8 other families. *PLoS One,* 7, e29729.

VOZNYI, Y., E. KARPOVA, T. DUDUKINA, I. TSVETKOVA, A. BOER, H. JANSE & O. VANDIGGELEN (1993) A FLUOROMETRIC ENZYME ASSAY FOR THE DIAGNOSIS OF SANFILIPPO DISEASE-C (MPS-III-C). *Journal of Inherited Metabolic Disease,* 16, 465-472.

Wang, D., J. Fan, C. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. Hudson, R. Lipshutz, M. Chee & E. Lander (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science,* 280, 1077-1082.

Wang, Z., P. White & S. Ackerman (2001) Atp11p and Atp12p are assembly factors for the F-1-ATPase in human mitochondria. *Journal of Biological Chemistry,* 276, 30773-30778.

Winckler, W., M. Weedon, R. Graham, S. McCarrolll, S. Purcell, P. Almgren, T. Tuomi, D. Gaudet, K. Bostrom, M. Walker, G. Hitman, A. Hattersley, M. McCarthy, K. Ardlie, J. Hirschhorn, M. Daly, T. Frayling, L. Groop & D. Altshuler (2007) Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. *Diabetes,* 56, 685-693.

WOLKOFF, A., E. WOLPERT, F. PASCASIO & I. ARIAS (1976) ROTORS SYNDROME - DISTINCT INHERITABLE PATHOPHYSIOLOGIC ENTITY. *American Journal of Medicine,* 60, 173-179.

Zhang, M., C. Zhu, A. Jacomy, L. Lu & A. Jegga (2011) The Orphan Disease Networks. *American Journal of Human Genetics,* 88, 755-766.

Zhao, X., A. Braun & J. Braun (2008) Biological roles of neural J proteins. *Cellular and Molecular Life Sciences,* 65, 2385-2396.

Zhu, H., D. Shang, M. Sun, S. Choi, Q. Liu, J. Hao, L. E. Figuera, F. Zhang, K. W. Choy, Y. Ao, Y. Liu, X. L. Zhang, F. Yue, M. R. Wang, L. Jin, P. I. Patel, T. Jing & X. Zhang (2011) X-linked congenital

hypertrichosis syndrome is associated with interchromosomal insertions mediated by a human-specific palindrome near SOX3. *Am J Hum Genet,* 88, 819-26.

ZINSMAIER, K., K. EBERLE, E. BUCHNER, N. WALTER & S. BENZER (1994) PARALYSIS AND EARLY DEATH IN CYSTEINE STRING PROTEIN MUTANTS OF DROSOPHILA. *Science,* 263, 977-980.

Zivna, M., H. Hulkova, M. Matignon, K. Hodanova, P. Vylet'al, M. Kalbacova, V. Baresova, J. Sikora, H. Blazkova, J. Zivny, R. Ivanek, V. Stranecky, J. Sovova, K. Claes, E. Lerut, J. Fryns, P. Hart, T. Hart, J. Adams, A. Pawtowski, M. Clemessy, J. Gasc, M. Gubler, C. Antignac, M. Elleder, K. Kapp, P. Grimbert, A. Bleyer & S. Kmoch (2009) Dominant Renin Gene Mutations Associated with Early-Onset Hyperuricemia, Anemia, and Chronic Kidney Failure. *American Journal of Human Genetics,* 85, 204-213.

Zlotogora, J. (2007) Multiple mutations responsible for frequent genetic diseases in isolated populations. *European Journal of Human Genetics,* 15, 272-278.

[Anonymous] (2008) Making rare diseases a public-health and research priority. *Lancet,* 371, 1972-1972.

ČÁST VII. Kopie publikovaných prací tvořících základ dizertační práce

1. Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis.
2. Cerebellar dysfunction in a family harbouring the *PSEN1* mutation co-segregating with a Cathepsin D variant p.A58V
3. Rotor-type hyperbilirubinaemia has no defect in the canalicular bilirubin export pump.
4. Complete OATP1B1 and OATP1B3 deficiency causes human Rotor syndrome by interrupting conjugated bilirubin reuptake into the liver.
5. Development of a human mitochondrial oligonucleotide microarray (h-MitoArray) and gene expression analysis of fibroblast cell lines from 13 patients with isolated F(I)F(o) ATP synthase deficiency.
6. *TMEM70* mutations cause isolated ATP synthase deficiency and neonatal mitochondrial encephalocardiomyopathy.
7. Mutations in *TMEM76** cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome).

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

ARTICLE

Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis

Lenka Nosková,[1,2,9] Viktor Stránecký,[1,2,9] Hana Hartmannová,[1,2] Anna Přistoupilová,[1,2] Veronika Barešová,[1,2] Robert Ivánek,[1,2] Helena Hůlková,[1] Helena Jahnová,[1] Julie van der Zee,[3,4] John F. Staropoli,[5] Katherine B. Sims,[5] Jaana Tyynelä,[6] Christine Van Broeckhoven,[3,4] Peter C.G. Nijssen,[7] Sara E. Mole,[8] Milan Elleder,[1,2] and Stanislav Kmoch[1,2,*]

Autosomal-dominant adult-onset neuronal ceroid lipofuscinosis (ANCL) is characterized by accumulation of autofluorescent storage material in neural tissues and neurodegeneration and has an age of onset in the third decade of life or later. The genetic and molecular basis of the disease has remained unknown for many years. We carried out linkage mapping, gene-expression analysis, exome sequencing, and candidate-gene sequencing in affected individuals from 20 families and/or individuals with simplex cases; we identified in five individuals one of two disease-causing mutations, c.346_348delCTC and c.344T>G, in *DNAJC5* encoding cysteine-string protein alpha (CSPα). These mutations—causing a deletion, p.Leu116del, and an amino acid exchange, p.Leu115Arg, respectively—are located within the cysteine-string domain of the protein and affect both palmitoylation-dependent sorting and the amount of CSPα in neuronal cells. The resulting depletion of functional CSPα might cause in parallel the presynaptic dysfunction and the progressive neurodegeneration observed in affected individuals and lysosomal accumulation of misfolded and proteolysis-resistant proteins in the form of characteristic ceroid deposits in neurons. Our work represents an important step in the genetic dissection of a genetically heterogeneous group of ANCLs. It also confirms a neuroprotective role for CSPα in humans and demonstrates the need for detailed investigation of CSPα in the neuronal ceroid lipofuscinoses and other neurodegenerative diseases presenting with neuronal protein aggregation.

Introduction

The neuronal ceroid lipofuscinoses (NCLs) are a heterogeneous group of inherited neurodegenerative disorders with an incidence of between 1 and 30 per 100,000. Common findings in the NCLs are an accumulation of autofluorescent storage material in neural and peripheral tissues and neurodegeneration. Although mutations in eight genes—*CLN1* (*PPT1* [MIM 256730]), *CLN2* (*TPP1* [MIM 204500]), *CLN3* (MIM 204200), *CLN5* (MIM 256731), *CLN6* (MIM 601780), *CLN7* (*MFSD8* [MIM 610951]), *CLN8* (MIM 600143), and *CLN10* (*CTSD* [MIM 610127])—have been identified in autosomal-recessive childhood and juvenile NCLs[1] and recently also in autosomal-recessive adult-onset NCL (Kufs disease [MIM 204300])[2], the genetic and molecular basis of adult-onset NCL with dominant inheritance (Parry type [MIM 162350]) remains unknown.

Autosomal-dominant adult-onset neuronal ceroid lipofuscinosis (ANCL) was first described in a family of British descent from New Jersey, USA (Parry disease),[3] and in a second family reported in Spain.[4] More recently, a large American family with English ancestry (UCL563 in this study),[5] another family from Alabama, USA (UCL562),[6,7] and a third family from the Netherlands (N1)[8] were presented. Common characteristics of affected individuals included generalized seizures, movement disorders, cognitive deterioration, and progressive dementia; the age of onset varied between 25 and 46 years.

In this work we describe a Czech family (P1) with autosomal-dominant ANCL in whom, by using a combination of linkage mapping, gene-expression analysis, and exome sequencing, we identified a unique heterozygous mutation in *DNAJC5* encoding cysteine-string protein alpha (CSPα [MIM 611203]; information on CSPα is accessible in the National Center for Biotechnology Information [NCBI] Gene Entrez database under GeneID 54968). The same or a second heterozygous *DNAJC5* mutation was found in four additional unrelated ANCL families that, together with altered palmitoylation-dependent sorting of mutant proteins in a cellular model and a reduced amount of CSPα in neuronal cells of affected individuals, confirmed the causality of CSPα mutations in autosomal-dominant ANCL.

[1]Institute for Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague, 120 00 Prague, Czech Republic; [2]Center for Applied Genomics, First Faculty of Medicine, Charles University in Prague, 120 00 Prague, Czech Republic; [3]Neurodegenerative Brain Diseases Group, Department of Molecular Genetics, VIB, B-2610 Antwerp, Belgium; [4]Laboratory of Neurogenetics, Institute Born-Bunge, University of Antwerp, B-2610 Antwerp, Belgium; [5]Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA; [6]Institute of Biomedicine/Biochemistry and Developmental Biology, University of Helsinki, 00014 Helsinki, Finland; [7]Department of Neurology, St. Elisabeth Hospital, 5022 Tilburg, The Netherlands; [8]MRC Laboratory for Molecular Cell Biology, Institute of Child Health and Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK
[9]These authors contributed equally to this work
*Correspondence: skmoch@lf1.cuni.cz
DOI 10.1016/j.ajhg.2011.07.003.

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

Material and Methods

Subjects

The Czech family (P1) was ascertained at the Institute of Inherited Metabolic Disorders in Prague. Some families were described earlier—an American family from USA with English ancestry UCL563,[5] a family from Alabama, USA (UCL562),[6,7] and one from the Netherlands (N1)[8]. Previously unpublished data from families from the USA, France, the Netherlands, Belgium, Poland, Austria, Italy, and Germany were collected under the auspices of the Rare NCL Gene Consortium by Sara Mole. Enzyme assay or analysis of known genes in which mutations lead to NCL had excluded these mutations as the cause in some but not all subjects. Diagnosis of ANCL disease is very challenging, partly because of its rarity but also because for some cases it can only be verified by finding the characteristic pathology in the brain, and not all affected individuals undergo this procedure. The cases included here were diagnosed by clinicians in different countries over two decades. Because full documentation was not always accessible, some medical histories could not be reviewed. However, we chose to test as many likely cases as possible and to fully report negative findings. Investigations were approved by participating centers' institutional review boards and were conducted according to the Declaration of Helsinki principles. Written, informed consent was obtained from all subjects.

Genotyping and Linkage Analysis

Genomic DNA was isolated by standard technology. We genotyped DNA samples by using Affymetrix GeneChip Mapping 10K 2.0 arrays (Affymetrix, Santa Clara, CA) according to the manufacturer's protocol at the microarray core facility of the Institute of Molecular Genetics in Prague. We extracted raw feature intensities from the Affymetrix GeneChip Scanner 3000 7G images by using the GeneChip operating Software (GCOS) 1.4 and generated individual SNP calls by using Affymetrix Genotyping Analysis Software (GTYPE) 4.1.

We carried out multipoint parametric linkage analysis along with a determination of the most likely haplotypes by using affected-only analysis under the assumption of an autosomal-dominant mode of inheritance with a 0.99 constant, age-independent penetrance, 0.01 phenocopy rate, and 0.001 frequency of disease allele; the analysis was performed with version 1.1.2 of Merlin software.[9] The results were visualized in the version 1.032 of the HaploPainter software[10] and in version 2.9.2 of R-project statistical software.

Gene-Expression Analysis

We isolated leucocytes from freshly drawn blood by using a standard erythrocyte lysis protocol and isolated total RNA from freshly isolated cells by using TRIZOL solution (Invitrogen, Carlsbad, CA). RNA concentration was determined spectrophotometrically at A260 nm by NanoDrop (NanoDrop Technologies), and quality was checked on an Agilent 2100 Bioanalyser (Agilent Technologies). Aliquots of isolated RNA were stored at –80°C until analysis. Expression analysis was performed on the Illumina HumanRef-8_V2 BeadChip at the microarray core facility of the Institute of Molecular Genetics in Prague. Hybridized slides were scanned on an Illumina BeadArray Reader, and bead level data were summarized by Illumina BeadStudio Software v3. Bead summary data were imported into R-project statistical software v.2.9.2 and normalized with the quantile method in the Lumi package. Differential gene-expression analysis was performed with the Limma package and the lmFit function. A multiple testing correction was performed with the Benjamini and Hochberg method. Database for Annotation, Visualization and Integrated Discovery version 6.7 (DAVID) was used for functional annotation. Details on the experiment and raw expression data are available at the Gene Expression Omnibus (GEO) repository under accession GSE30369.

Copy-Number Analysis

DNA samples from seven individuals of family P1 (II.2, IV.1, IV.2, IV.3, IV.4, IV.7, and IV.8) were genotyped with Affymetrix GeneChip Mapping 6.0 array (Affymetrix, Santa Clara, CA) at the microarray core facility of the Institute of Molecular Genetics in Prague according to the manufacturer's protocol. Raw feature intensities were extracted from the Affymetrix GeneChip Scanner 3000 7G images with the GeneChip Control Console Software 2.01. We generated individual SNP calls by using Affymetrix Genotyping Console Software 3.02. Copy-number changes were identified in Affymetrix Genotyping Console Software (GTC version 3.02). We used data from both SNP and copy-number probes to identify copy-number aberrations relative to a built-in reference. Only regions larger than 10 Kb and containing at least five probes were reported.

Exome Sequencing

We performed DNA enrichment by using 3 μg of DNA from individual IV.7 and the SureSelect All Exome kit (Agilent, Santa Clara, USA) according to the manufacturer's protocol. DNA sequencing was performed on the captured DNA library with one-quarter of a SOLiD 4 slide (Applied Biosystems, Carlsbad, USA) at CeGaT (Tubingen, Germany). We aligned reads in color space to the reference genome (hg19) by using NovoalignCS version 1.01 (Novocraft, Malaysia) with the default parameters. Sequence variants in the analyzed sample were identified with the SAMtools package (version 0.1.8).[11] The high-confidence variants list (SNP quality $>$ 100 and indel quality $>$ 50) was annotated with the SeattleSeq Annotation server (hg19). Sequence variants that were not annotated in the dbSNP or 1000 Genomes databases were prioritized for further analysis.

DNA Sequencing and Mutation Analysis

All exons and corresponding exon-intron boundaries of *DNAJC5* (NM_025219.2), encoding CSPα, were amplified by PCR from genomic DNA of the probands and sequenced with version 3.1 Dye Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA) with electrophoresis on an ABI 3500XL Avant Genetic Analyzer (Applied Biosystems). Data were analyzed with Sequencing Analysis software. Segregation of the candidate mutations was assessed by PCR and direct sequencing of the corresponding genomic DNA fragments. Primer sequences are available in Table S1, available online.

Homozygosity-Haplotype Analysis

DNAJC5 genomic fragments containing multiple SNPs with high-heterozygosity values were amplified by PCR from genomic DNA of probands and sequenced as described above. Genotypes for individual SNPs were obtained, and homozygous haplotypes were defined as described recently.[12] We compared the resulting homozygous haplotypes across individuals to determine whether

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

the chromosomal segments around the same identified mutations could be identical by descent.

Bioinformatic Analysis of the Cysteine-String Domain

Hydrophobicity of the wild-type and mutant cysteine-string domains were analyzed with a Kyte-Doolittle algorithm available at Expasy server. Potential effects of detected mutations on CSPα palmitoylation were assessed with the prediction program CSS-Palm 2.0. Obtained hydrophobicity values and palmitoylation score values were exported for each of the sequences and plotted with an Excel function. We assessed possible impacts of the p.Leu115Arg substitution on the structure and function of CSPα by using SIFT and PolyPhen-2 servers.

CSPα-Expression Vectors

DNAJC5/CSPα cDNA were amplified by RT-PCR from a control and an affected individuals' leucocytes with primers incorporating a *Bsp*EI site at the 5′ end of PCR products. Resulting PCR products were first cloned into pCR4 TOPO vector (Invitrogen) and, after sequencing verification, these were further subcloned in frame into a pEGFP-C1 vector with *Bsp*EI and *Apa*I restriction sites. The initiating methionine codon was removed from *DNAJC5*/CSPα in all enhanced green fluorescent protein (EGFP)-CSPα constructs.

Transient Expression of EGFP-CSPα

pEGFP-CSPα constructs were transfected into CAD-2A2D5 (CAD5) cells derived from Cath.a-differentiated (CAD) cells (provided by Sukhvir Mahal, The Scripps Research Institute, Jupiter, FL, USA). One day before transfection, 8×10^4 cells/cm^2 were seeded with OptiMEM medium (OptiMEM; Invitrogen) containing 9% BGS (HyClone, Logan, UT), 90 units penicillin/ml, and 90 g of streptomycin/ml. Cells were transfected by either 0.8 μg or 4.5 μg of plasmid constructs with Lipofectamine 2000 (Invitrogen) in serum and antibiotics free OptiMEM medium according to the manufacturer's protocol. Transfection experiments were performed in more than five replicates.

Immunofluorescence Analysis

Cells were fixed 24 hr after transfection with 4% paraformaldehyde, permeabilized in 0.1% TRITON, washed, blocked with 5% bovine serum albumin (BSA), and incubated for 1 hr at 37°C with anti-protein disulfide isomerase (PDI) mouse monoclonal IgG1 (Stressgen, San Diego, CA) for endoplasmic reticulum (ER) localization, anti-GS28 mouse IgG1 (Stressgen, San Diego, CA) for Golgi localization, and anti-GFP rabbit polyclonal IgG (Abcam) for EGFP-CSPα detection. For fluorescence detection, corresponding species-specific secondary antibodies Alexa Fluor 488 and Alexa Fluor 555 (Molecular Probes, Invitrogen, Paisley, UK) were used. Nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI). Prepared slides were mounted in fluorescence mounting medium Immu-Mount (Shandon Lipshaw, Pittsburgh, PA) and analyzed by confocal microscopy.

Image Acquisition and Analysis

XYZ images sampled according to Nyquist criterion were acquired with a TE2000E C1si laser scanning confocal microscope, a Nikon PlanApo objective (40×, N.A.1.30), 488 nm and 543 nm laser lines, and 515 ± 15 nm and 590 ± 15 nm band-pass filters. Images were deconvolved with the classic maximum likelihood restoration algorithm in Huygens Professional software (SVI, Hilversum, The Netherlands).[13] Colocalization maps employing single pixel overlap coefficient values ranging from 0–1[14] were created with Huygens Professional software. The resulting overlap coefficient values are presented as pseudocolor (the scale is shown in the corresponding lookup tables).

Immunoblot Analysis

Transfected CAD5 Cells

Cells were harvested in PBS; centrifuged at 500 *g* for 7 min; and resuspended in 10 mM Tris, 10 mM KCl, 2 mM EDTA, 4% glycerol, 1 mM DTT, and Complete Protease Inhibitor Cocktail (Roche); homogenized by sonication followed by centrifugation at 20,000 *g* for 15 min at 4°C; and assessed for protein content in the supernatant with the Bradford assay.

Brain Homogenates

Frozen autopsy materials were homogenized under liquid nitrogen; dissolved in 10 mM Tris, 10 mM KCl, 2 mM EDTA, 4% glycerol, 1 mM DTT, and Complete Protease Inhibitor Cocktail (Roche); centrifuged at 20,000 *g* for 15 min at 4°C; and assessed for protein content in the supernatant with the Bradford assay. Homogenate aliquots corresponding to 30 μg of total protein in brain homogenates or 20 μg of total protein in CAD5 cells were resolved on 12% SDS-PAGE under nonreducing or reducing conditions and transferred to the polyvinylidene fluoride (PVDF) membrane. Membranes were blocked by 5% BSA and 0.05% Tween 20 in PBS. CSPα or CSPα-EGFP protein was visualized by incubation with rabbit CSP antibody (Stressgen) at 1: 500 in 5% BSA and 0.05% Tween 20 in PBS for 90 min or rabbit GFP antibody (Abcam) at 1:5000 in 5% BSA and 0.05% Tween 20 in PBS for 90 min, followed by incubation with goat anti-rabbit HRP (Pierce) at 1:10000 in 0.05% Tween 20 in PBS for 60 min and detection by SuperSignal West Femto Maximum Sensitivity Substrate (Pierce). For depalmitoylation studies, samples were depalmitoylated prior to SDS-PAGE by treatment with neutral 1 M hydroxylamine or 1 M Tris as a control for 20 hr at room temperature.

Immunohistochemical and Histochemical Studies

Formaldehyde-fixed brain samples were analyzed. Immunodetection of CSPα on paraffin sections was performed with rabbit CSP antibody (Stressgen; diluted 1:750 in 5% BSA) in PBS. Synaptic regions were detected with monoclonal mouse IgG1 synaptobrevin antibody (Sigma, Saint Louis, USA; diluted 1:8000 in 5% BSA) in PBS, which was applied after heat-induced epitope retrieval at pH 6.0. Detection of the bound primary antibody was achieved with Dako EnVision + TM Peroxidase Rabbit kit (Dako, Glostrup, Denmark) with 3,3′-diaminobenzidine as substrate. The specificity of the antigen detection was always ascertained by omitting of the primary antibody-binding step.

Stored ceroid material was best detected because of its prominent autofluorescence via filter block with an excitation wavelength of 400–440 nm (fluorescence microscope Nikon E800, filter block BV-2A).

Results

Clinical Observations and Biochemical Findings

The diagnosis of ANCL in family P1 (Figure 1A) was based on clinical presentation and examination of proband III.6, who presented at age 30 with myoclonic epilepsy, generalized tonic-clonic seizures, and progressive cognitive

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

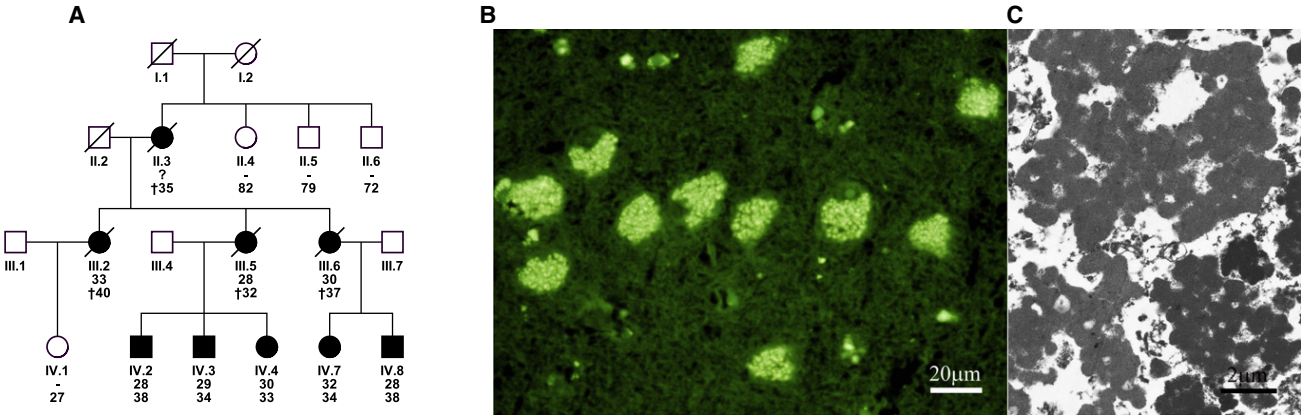


Figure 1. Pedigree and Neuropathology Findings in Family P1
(A) Pedigree of the Czech family. Black symbols denote affected individuals; open symbols denote unaffected individuals. Age of onset is shown above current age or age of death (indicated by †).
(B) Epifluorescence. Hippocampal pyramidal neurons with prominent lysosomal storage of autofluorescent material representing the general neurolysosomal storage pattern in the brain cortex. The autofluorescence was demonstrated with the filter block with an excitation wavelength of 400–440 nm.
(C) Electron micrograph. GROD-type ultrastructure of the storage lysosomes.

deterioration with depression; these symptoms were followed by progressive motor neurological symptoms leading to death at age 37 years. There was normal activity of palmitoyl-protein-thioesterase 1 (PPT1) in leucocytes. Neuropathological examination of postmortem brain tissue showed characteristic neurolysosomal storage of autofluorescent material with ultrastructural appearance corresponding to granular osmiophilic deposits (GRODs) (Figures 1B and 1C). A skin biopsy was free of lysosomal storage at the ultrastructural level. An affected status in other family members was assigned if a very similar clinical course starting with myoclonic and/or generalized tonic-clonic seizures followed after 1–2 years by progressive cognitive deterioration and depressive symptomatology. All affected individuals showed generalized epileptic discharges in electroencephalograms and manifested brainstem and central pyramidal neurological symptomatology in the later period of disease. Other ANCL families analyzed in this study are described in Table 1. Previously unpublished families and cases with mutation in CSPα are described in more detail below.

The proband of family UCL328 was a male of European descent and in good health until his first generalized tonic-clonic seizure at age 34. This was followed by evidence of progressive confusion and dementia as well as more frequent, medically refractory generalized seizures. Long-term electroencephalography showed generalized periodic epileptiform discharges superimposed on a background of diffuse low-amplitude, high-frequency activity consistent with a dementing process. A brain MRI at age 38 showed prominence of cortical sulci and cerebellar folds and mild enlargement of the lateral ventricles consistent with diffuse cerebral and cerebellar atrophy. Concurrent neuropsychiatric testing showed a verbal IQ of 77, a performance IQ of 71, and a full-scale IQ of 73. Regression of gross and fine motor skills began at age 40, and there was ensuing evidence of ataxia and myoclonus. By age 45, the proband was wheelchair-bound and required nursing-home care. Visual function was normal. A frontal lobe brain biopsy revealed numerous neurons containing homogeneous eosinophilic material with a golden-brown hue. The pigmented material stained intensely by the periodic acid-Schiff reaction and was found to be autofluorescent. Ultrastructural examination showed multiple neurons distended by granular osmiophilic deposits. There was no family history of seizures, early-onset dementia, or other neurologic abnormality.

The proband of family UCL519 is one of at least five similarly affected individuals over three generations with apparent autosomal-dominant inheritance. He showed obsessive behavior starting in his mid-20s, and the first seizure occurred when he was in his early 30s. His speech regressed, his short-term memory became impaired, and he had difficulty in walking without an aid. No further details are available.

Identification of CSPα Mutation in Family P1 by a Combination of Linkage Analysis, Copy-Number Analysis, Gene-Expression Analysis, and Exome Sequencing

To map the disease locus, we used Affymetrix GeneChip Mapping 10K v2.0 arrays, genotyped all available and informative family members, and performed linkage analysis. We identified five candidate regions with positive LOD scores on chromosomes 1, 4, 15, 20, and 22 (Figure 2A). In parallel, we used Affymetrix GeneChip Mapping 6.0 array, genotyped seven individuals, and assessed copy-number changes; we found no indication for a potentially disease-causing deletion or duplication.

To identify a mutation that affected the amount of transcript, we compared gene-expression profiles in leucocytes isolated from four affected individuals to those from four

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

Table 1. ANCL Families Analyzed in This Study

Family No.	Mutation in CSPα	Country	Diagnosis	References
P1	p.Leu116del	Czech Republic	ANCL, autosomal dominant	
N1	p.Leu115Arg	The Netherlands	ANCL, autosomal dominant	8,31,32
UCL563	p.Leu115Arg	USA	ANCL, autosomal dominant	5
UCL328	p.Leu115Arg	USA, French-Canadian	Kufs	
UCL519	p.Leu116del	USA	Kufs, autosomal dominant	
UCL417	–	France	Kufs, autosomal dominant	
UCL562	–	USA	Kufs, autosomal dominant	6,7
UCL572	–	USA/Italy	Kufs, autosomal dominant?	
UCL327	–	USA	Kufs, with ALS in extended family	
UCL385	–	Belgium	Kufs Type A or atypical juvenile NCL, autosomal recessive	
UCL403	–	France	Kufs Type B, autosomal recessive	
UCL450	–	Poland	variant juvenile or ANCL, autosomal recessive (heterozygous change in *CLCN6* already known)	33
UCL472	–	Germany	variant juvenile or ANCL	34
UCL482	–	The Netherlands	ANCL	
UCL508	–	USA	Kufs	
UCL520	–	USA	Kufs	
UCL522	–	USA	Kufs	
UCL545	–	Netherlands	Kufs	
UCL568	–	Austria	Kufs	
UCL571	–	Netherlands	Kufs	

Diagnosis is provided as reported by referring clinician. In all cases there was no visual failure, and no distinction was made according the mode of inheritance, if apparent.

age-matched controls by using Illumina HumanRef-8v2 Expression BeadChips. This analysis identified a set of 2131 differentially expressed genes, of which 65 were localized within candidate regions identified by linkage analysis (Figure 2B and Table S2). At the same time, we analyzed gene-expression changes by using gene-enrichment analysis and found that the identified profiles indicated significant dysregulation of spliceosome, upregulation of many components of respiratory chain complexes, altered expression of genes active in pathways involved in neurodegenerative diseases, and accelerated proteolysis (Table 2 and Figures S1–S7).

To directly identify possible disease-causing mutation(s) among the candidate genes defined by this combination of linkage analysis and gene-expression profiling, we performed exome sequencing in individual IV.7. From the sequencing run we obtained 94.7 M sequencing reads, of which we were able to map 50.2 M on the human genome reference sequence. After removing PCR generated duplicate reads (23.6 M), we obtained 26.6 M unique reads, of which 19.5 M (73.3%) mapped on a targeted exome sequence and were 92% covered at least once. When the sequence of the proband was compared to the reference sequence, 22,617 single nucleotide variants (SNP quality > 100) and 2604 indels (indel quality > 50) were revealed in the proband, of which 957 (617 SNPs and 340 indels) were novel (e.g., were not present in the dbSNP and 1000 Genomes databases).

We intersected the results of exome sequencing with the mapping information and the gene-expression changes, and this analysis illuminated a single gene, *DNAJC5*, encoding the protein CSPα, located in the candidate region on chromosomal region 20q13.33, (*DNAJC5* hg19 coordinates chr20:62526518–62565394) and showing a significant increase in transcript levels in affected individuals' leucocytes (Figure 2B), and had a unique heterozygous mutation c.346_348delCTC (p.Leu116del) compatible with autosomal-dominant inheritance of the disease (Table 3).

CSPα Mutations Segregate with ANCL in Additional Families

Through sequence analysis of *DNAJC5* genomic DNA, we found consistent segregation of the c.346_348delCTC mutation with the ANCL phenotype within the Czech family P1 (Figure 2C). Moreover, among 20 additional ANCL families and/or simplex cases tested (Table 1), we identified the same mutation in a previously unreported

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

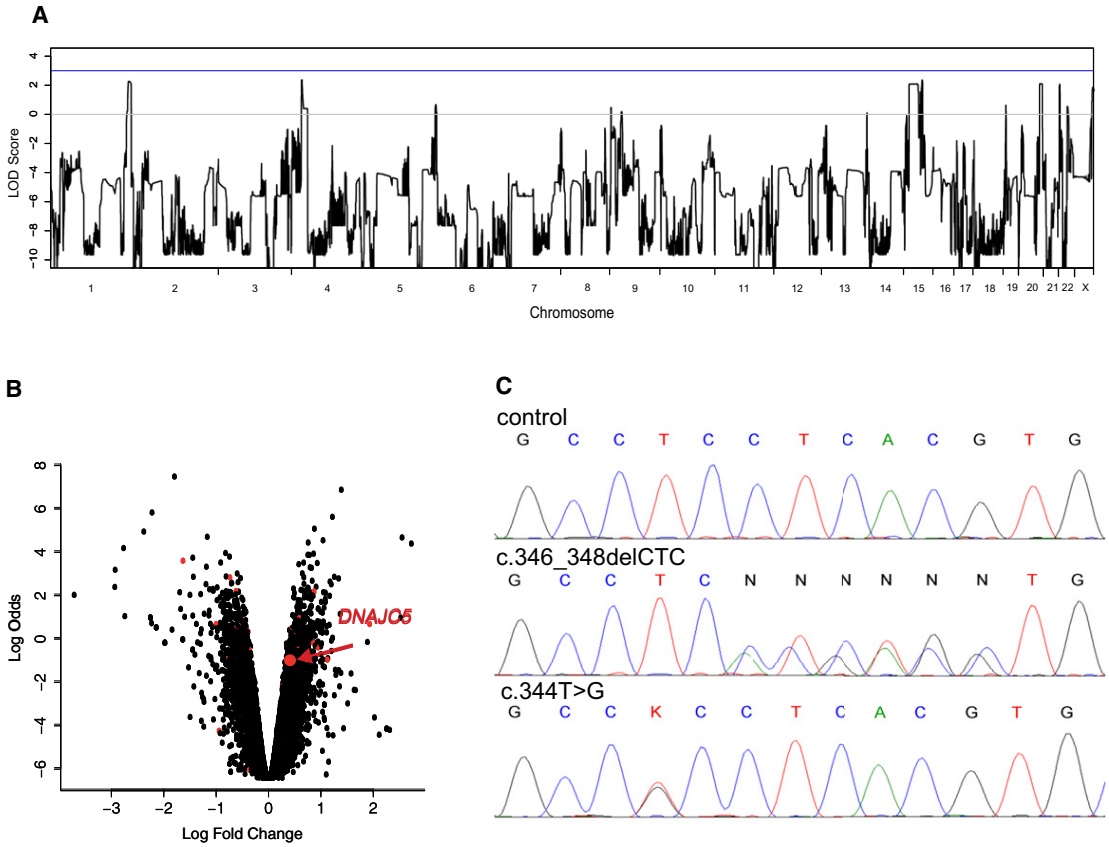


Figure 2. Identification of *DNAJC5* Mutations

(A) A whole-genome parametric linkage analysis showing candidate regions reaching the theoretical maximum LOD scores of 2.1 attainable in this family on chromosomes 1 (1: 233,697,529–249,250,621), 4 (4: 23,561,661–28,920,119), 15 (15: 39,049,915–61,382,423; 65,139,935–67,296,086; 71,515,415–78,819,152), 20 (20:53,448,624–63,025,520), and 22 (22: 1–21,982,248). All coordinates refer to hg19.

(B) Gene-expression changes in leucocytes from four affected individuals compared to those of four controls. The logarithm of the probability that the gene is differentially expressed (log odds) is plotted as a function of the logarithm of the gene-expression fold change (log fold change) between the patient and control samples. Differentially expressed genes located in the candidate regions are shown as red dots, and *DNAJC5* is specifically indicated. The list of differentially expressed genes located within the linked regions is, together with log fold changes and corresponding t test values, p-values and adjusted p-values, provided in Supplemental Data.

(C) Chromatograms of *DNAJC5* genomic DNA sequences showing identified heterozygous mutations. (Upper panel) Sequence of an unaffected individual, (middle panel) sequence showing heterozygous mutation c.346_348delCTC in the proband from family P1, and (lower panel) sequence showing heterozygous mutation c.344T>G in the proband from family N1.

American family, UCL519, and a second heterozygous mutation (c.344T>G [p.Leu115Arg]) (Figure 1D) segregating with the phenotype in the Dutch family N1[8] and the American family UCL563[5] and present in a previously unreported simplex case UCL328. Mutations were found in all 14 affected individuals (five Czech, six Dutch, and one in each of the other pedigrees) across these five families and were absent in all seven unaffected siblings (two Czech, six Dutch, and one from American family UCL563) from whom DNA was available for testing. In addition to this, the identified mutations were absent in 200 control samples of European descent and were not present in the dbSNP or 1000 Genomes databases.

Haplotypes segregating with ANCL phenotype in Czech family P1 and Dutch family N1 were obtained from genotypes generated with Affymetrix GeneChip Mapping 10K v2.0 arrays and are shown in Figures S8 and S9. For simplex cases, phased haplotypes could not be obtained. To reveal whether probands carrying the same mutation might be distantly related and share a mutation-carrying chromosomal segment from a common ancestor, we examined homozygosity haplotypes across the *DNAJC5* genomic region (Table S3). The c.346_348delCTC (p.Leu116del) mutations in families P1 and UCL519 are present on two distinct haplotypes, indicating that these families are probably not related and that the mutations appeared independently. The mutations c.344T>G (p.Leu115Arg) are also present on two distinct haplotypes, one in UCL328 and one shared by family N1 and UCL563. This mutation therefore probably also appeared independently in two different lineages, but it is possible that families N1 and UCL563 are identical by descent.

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

Table 2. Functional Annotation of Gene-Expression Changes and KEGG Pathways Defined by Gene-Enrichment Analysis

Term	Count	%	p Value	Population Hits	Population Total	Fold Enrichment	FDR
hsa03040: spliceosome	41	2.48	$2.3 \times^{-10}$	126	5085	2.92	$2.9 \times^{-7}$
hsa05016: Huntington disease	46	2.78	$7.7 \times^{-8}$	180	5085	2.29	$9.5 \times^{-5}$
hsa05010: Alzheimer disease	43	2.60	$8.5 \times^{-8}$	163	5085	2.37	$1.1 \times^{-4}$
hsa05012: Parkinson disease	35	2.11	$7.0 \times^{-7}$	128	5085	2.45	$8.7 \times^{-4}$
hsa00190: oxidative phosphorylation	35	2.11	$1.0 \times^{-6}$	130	5085	2.41	$1.3 \times^{-3}$
hsa00520: amino sugar and nucleotide sugar metabolism	13	0.79	$2.4 \times^{-3}$	44	5085	2.65	$2.9 \times^{0}$
hsa03050: proteasome	12	0.73	$1.2 \times^{-2}$	47	5085	2.29	$1.4 \times^{1}$
hsa04120: ubiquitin mediated proteolysis	25	1.51	$1.5 \times^{-2}$	137	5085	1.64	$1.7 \times^{1}$
hsa04662: B cell receptor signaling pathway	16	0.97	$1.7 \times^{-2}$	75	5085	1.91	$1.9 \times^{1}$
hsa04621: NOD-like receptor signaling pathway	14	0.85	$1.7 \times^{-2}$	62	5085	2.03	$1.9 \times^{1}$
hsa00052: galactose metabolism	8	0.48	$2.0 \times^{-2}$	26	5085	2.76	$2.2 \times^{1}$
hsa03010: ribosome	17	1.03	$2.9 \times^{-2}$	87	5085	1.75	$3.0 \times^{1}$

Identified Mutations Affect Palmitoylation-Dependent Sorting and the Amount of CSPα in Neuronal Cells

Both identified mutations affect conserved dileucine residues located in the cysteine-string domain implicated in palmitoylation and membrane trafficking of CSPα[16]. Using in silico analysis, we found that p.Leu115Arg is predicted to decrease the hydrophobicity of the cysteine-string domain that is needed for initial binding of CSPα to the endoplasmic reticulum (ER) (Figure 3A), whereas p.Leu116del probably affects the efficiency of palmitoylation of adjacent cysteine residues (Figure 3B). SIFT analysis

Table 3. Exome Sequencing and a List of High-Confidence Novel Coding Variants Revealed by Exome Sequencing

Chromosome	Position	Reference Base	Sample Alleles	Function Genome Variation Server	Amino Acids	Protein Position	Gene List
Single nucleotide variants							
1	235,715,488	C	C/T	missense	ARG.GLN	50/76	*GNG4*
1	236,987,512	C	C/T	synonymous	none	286/1266	*MTR*
1	247,835,885	G	C/G	synonymous	none	153/308	*OR13G1*
15	43,552,700	G	G/T	missense	HIS.ASN	30/721	*TGM5*
15	43,900,153	C	C/T	synonymous	none	1234/1776	*STRC*
15	45,028,847	G	G/T	utr-5	none	NA	*TRIM69*
15	59,500,166	A	A/G	missense	ILE.VAL	343/382	*MYO1E*
15	65,555,518	A	A/G	synonymous	none	220/324	*PARP16*
15	66,857,721	C	C/T	utr-5	none	NA	*LCTL*
15	75,116,809	G	A/G	missense	VAL.MET	481/527	*LMAN1L*
20	60,884,827	G	A/G	synonymous	none	3631/3696	*LAMA5*
22	20,097,643	C	C/T	utr-3	none	NA	*DGCR8*
22	21,138,487	C	C/T	synonymous	none	373/500	*SERPIND1*
Indels							
4	25,678,161	TGC	−TGC	coding	none	NA	*SLC34A2*
20	62,562,227	CTC	−CTC	coding	none	NA	*DNAJC5*

All coordinates refer to hg19. SNP qualilty > 100 and indel quality > 50. Only Variants located within the linkage candidate regions and not present in dbSNP or 1000 Genomes databases are shown.

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

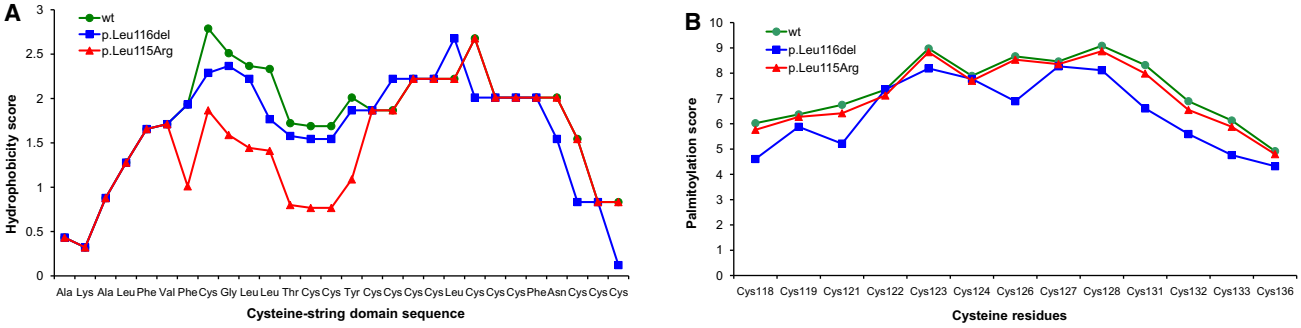


Figure 3. In Silico Analysis of Properties of the Cysteine-String Domain
(A) p.Leu115Arg mutation decreases the hydrophobicity of this domain, which is needed for initial binding to the ER.
(B) The p.Leu116del mutation decreases the palmitoylation score, that is, the confidence that cysteine residues adjacent to Leu116 might be efficiently palmitoylated.

(score = 0.00) predicted that the p.Leu115Arg mutation affects protein function, and analysis with Polyphen (overall score = 0.782; sensitivity = 0.85; and specificity = 0.93) predicted that it is possibly damaging. No such predictions can be obtained for the identified deletion p.Leu116del.

To study an effect of the identified mutations, we transiently expressed wild-type EGFP-tagged CSPα or mutant protein containing either p.Leu115Arg or p.Leu116del in CAD5 neuronal cells. Using immunofluorescence analysis, we found wild-type EGFP-CSPα predominantly at the plasma membrane, whereas both mutated proteins showed diffuse intracellular staining and abnormal colocalization with markers for the ER and Golgi apparatus (Figure 4A). In addition, using immunoblot analysis of transfected cell lysates, we found that both mutated proteins were less efficiently palmitoylated than the wild-type protein (Figure 4B).

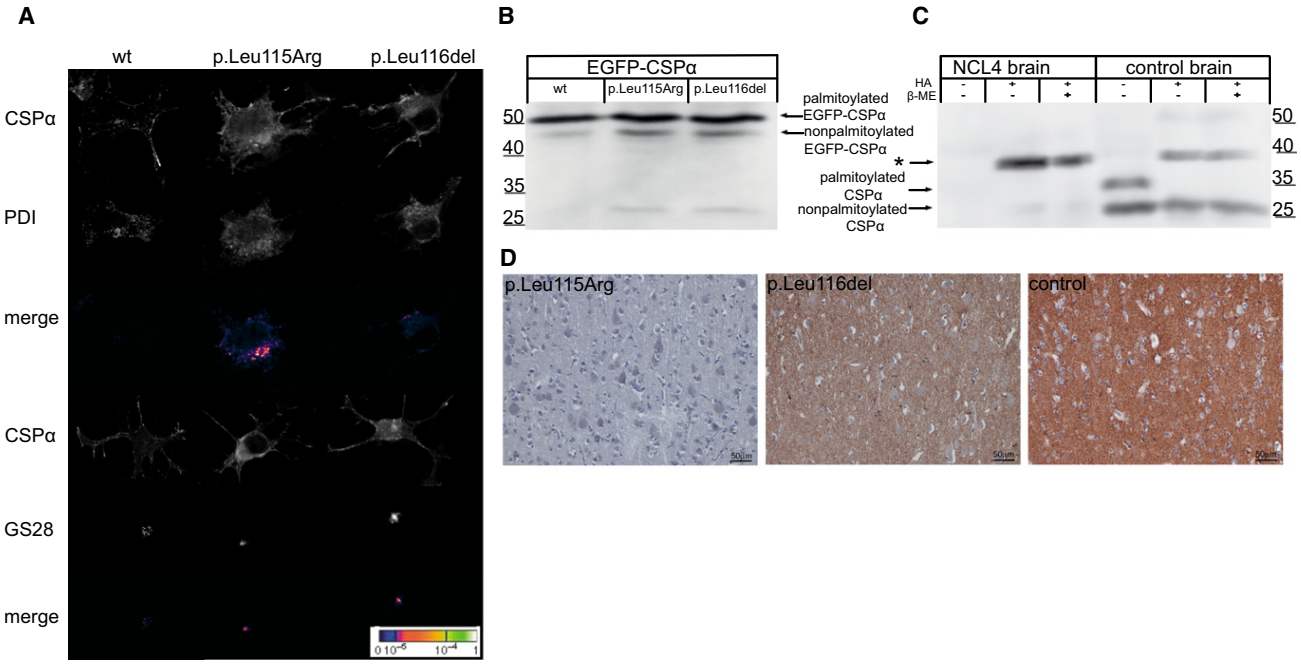


Figure 4. Characterization of Mutated CSPα
(A) Immunofluorescence analysis of transiently expressed EGFP-CSPα proteins in CAD5 cells showing prominent membrane localization of wild-type CSPα compared to the diffuse cytoplasmic staining and marked colocalization of mutated CSPα with endoplasmic reticulum represented by PDI and Golgi apparatus represented by Golgi-SNARE of 28 kDa (GS28).
(B) Immunoblot analysis of transiently expressed EGFP-CSPα proteins showing higher levels of nonpalmitoylated protein precursors for mutant proteins compared the wild-type (wt) protein.
(C) Immunoblot analysis of brain homogenates showing no soluble CSPα and the marked presence of CSPα-containing beta-mercaptoethanol (β-ME)-resistant aggregate (indicated by the asterisk) released upon hydroxylamine (HA) treatment in the affected individual (NCL4) compared to the brain homogenates of the control.
(D) Immunohistochemistry analysis of CSPα in gray matter of the cerebral cortex showing, at a low field, a significant decrease of CSPα in affected individuals compared to the strong CSPα staining in the age-matched control.

Please cite this article in press as: Nosková et al., Mutations in *DNAJC5*, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis, The American Journal of Human Genetics (2011), doi:10.1016/j.ajhg.2011.07.003

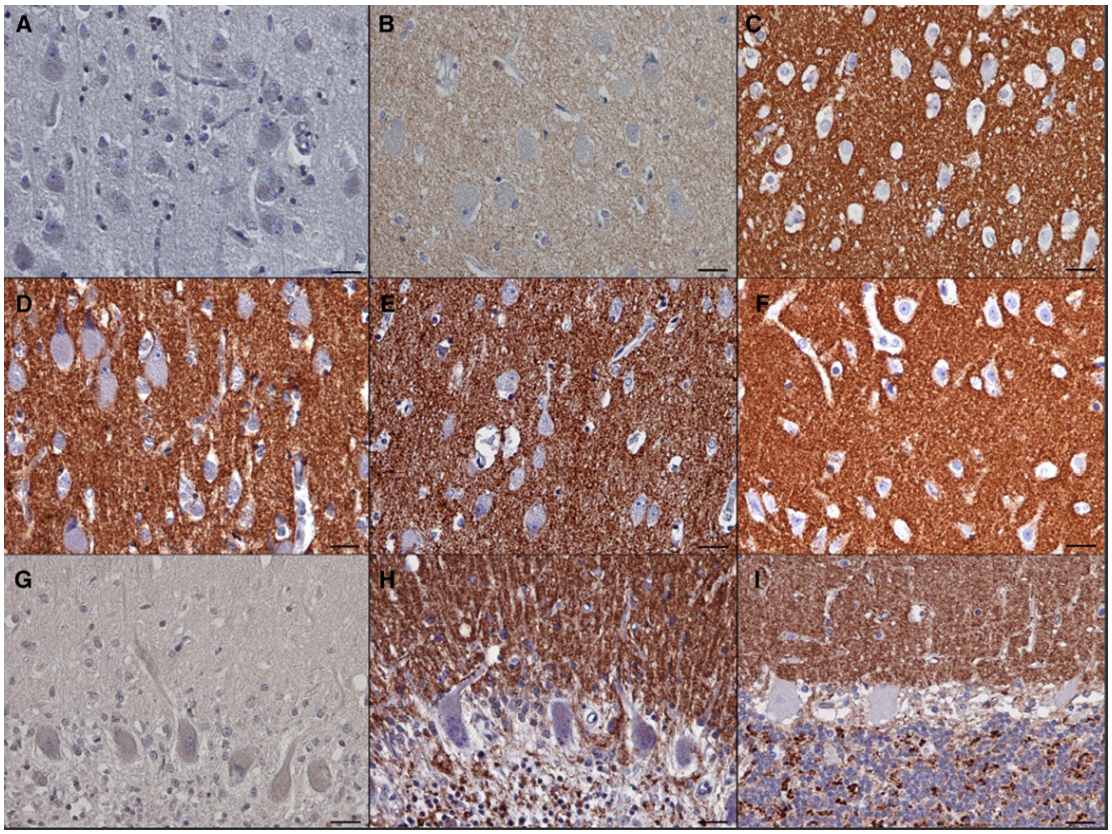


Figure 5. Brain Immunohistochemistry
(A–C) Detail of the CSPα staining in neuropil in the cerebral cortex that is absent in the individual with mutation p.Leu115Arg (A), decreased in the individual with mutation p.Leu116del (B), and strong in the age-matched control (C). Note the prominent neuronal storage, shown by large cell bodies, in both affected individuals.
(D–F) Staining pattern of the synaptic marker synaptobrevin in neuropil in the same regions in the individual with mutation p.Leu115Arg (D), in the individual with mutation p.Leu116del (E), and in the age-matched control (F).
(G and H) Cerebellar cortex of the case with p.Leu115Arg mutation. Similar to that in the cerebral cortex, CSPα staining is absent (G). This contrasts with a strong signal for synaptobrevin in the corresponding area in all three cerebellar cortical layers that are preserved adjacent to areas undergoing neurodegeneration (H).
(I) Strong CSPα staining in a control cerebellum. Note that the CSPα signal in the control (I) as well as the synaptobrevin signal in the individual with mutation p.Leu115Arg (H) are confined to the well defined synaptic regions (i.e., to the dendrites in the molecular layer, to the surface of the Purkinje cells, and to the synaptic glomeruli in the granular cell layer). The scale bars represent 25 μm.

To correlate these observed effects with in vivo, we analyzed post-mortem brain specimens by immunoblotting. Although both palmitoylated and nonpalmitoylated CSPα were present in control brain lysates, we could not detect any CSPα in brain lysates from a Dutch case (family N1) with the p.Leu115Arg mutation. However, after chemical depalmitoylation, we detected a chemiluminescence signal, probably corresponding to an otherwise insoluble CSPα-containing aggregate, which appeared much stronger in brain lysate from the case, than in the control (Figure 4C). Using immunohistochemical staining of CSPα in paraffin-embedded brain sections, we consistently found an absence of CSPα staining in synaptic regions in both the cerebral and the cerebellar cortex of individuals with the p.Leu115Arg mutation and significantly reduced CSPα staining in the cerebral cortex of individuals with the p.Leu116del mutation when we compared these individuals to age-matched controls (Figures 4D and 5).

Discussion

We carried out linkage mapping, gene-expression analysis, exome sequencing, and candidate-gene sequencing in affected individuals from 20 families and/or simplex cases of European descent suffering from autosomal-dominant adult-onset neuronal ceroid lipofuscinosis previously referred to as Parry disease. Using this approach, we identified in five of these families two recurrent mutations, c.346_348delCTC (p.Leu116del) and c.344T>G (p.Leu115Arg), in *DNAJC5* encoding cysteine-string protein alpha (CSPα). To prove their causality, we performed haplotype analysis, which revealed that the mutations had to appear independently in at least four lineages, and by using targeted genotyping of seven unaffected siblings and 200 control individuals as well as searching the 1000 Genome and dbSNP databases, we found that the mutations are exclusively present in 14 affected individuals.

CSPα is a highly conserved protein with no amino acid sequence variant found in humans so far. The identified mutations affect evolutionary conserved dileucine residues located in the cysteine-string domain that is implicated in palmitoylation and membrane targeting of CSPα.[15–17] Functional studies in transfected cell lines proved that these mutations affect palmitoylation and intracellular location of CSPα and thus decrease the level of the CSPα protein in the brain of affected individuals.

The molecular mechanisms underlying the dominant negative effect of the identified mutations on CSPα amounts in neuronal cells are not clear. It is known that CSPα forms detergent-resistant dimers[18] and that the presence of these dimers correlates with an inhibition of synapse formation and synaptic transmission.[19] Immunoblot analysis of brain lysate from affected individuals showed CSPα to be exclusively present in such an aggregate form. It is probable that the presence of mutant protein catalyzes accelerated aggregation and that the resulting aggregates will be composed equally of both mutant and wild-type proteins, and this will result in CSPα depletion. Another explanation of the dominant negative effect—nicely compatible with the observed lysosomal storage—would be a gradual accumulation of nondegradable CSPα aggregates in the lysosomal system. We followed this lead experimentally; however, we failed to identify CSPα in storage lysosomes by using immunohistochemistry analysis of fixed brain samples as well as in storage granules isolated from affected individuals' brains by using immunoblot analysis (data not shown).

CSPα associates with 70 kDa heat-shock cognate protein (Hsc70) and small glutamine-rich tetratricopeptide repeat domain protein (SGT) and forms an enzymatically active chaperone complex that is tethered to synaptic vesicles and ensures, in cooperation with other chaperones such as 40 kDa heat-shock protein (Hsp40),[20] 90 kDa heat-shock protein (Hsp90),[21] Hsc70 interacting protein (HIP)[22] and Hsp70 organizing protein (HOP),[22] correct conformation of many proteins essential for the functionality of synapses. It was shown that CSPα deletion causes progressive neurodegeneration and reduced life span in *Drosophila melanogaster*[23] and knockout mice.[24,25] Depletion of CSPα interferes with SNARE complex formation and has a profound effect on presynaptic vesicle release and synaptic function.[19,24,26–29] Thus, these CSPα mutations might lead to presynaptic dysfunction, explaining some of the neurological symptoms observed in affected individuals. In parallel, dysfunction of the CSPα/Hsc70/SGT chaperone complex might affect the folding quality of many client proteins and make them vulnerable to aggregation and degradation.[30] This could, in the long term, lead to lysosomal accumulation of misfolded and proteolysis-resistant proteins in the form of characteristic ceroid deposits in neurons.

Our finding of neurodegenerative disease caused by mutations in *DNAJC5* thus confirms a neuroprotective role for CSPα in humans and advocates detailed investigation of CSPα in the NCLs and other neurodegenerative diseases presenting with neuronal protein aggregation. It is interesting that there is no visual failure in the cases reported here, in contrast to the rapid loss of vision in mice completely lacking CSPα function.[25]

In this study we were able to explain ~25% of ANCL cases tested, though not all were known to be autosomal-dominant and some could have been misdiagnosed. Those families that do not carry mutations in *DNAJC5* or other known NCL genes provide a resource for identification of further genes whose disruption causes late-onset NCL.

In conclusion, we believe that our work represents an important step in the genetic dissection of a genetically heterogeneous group of ANCLs. From a clinical perspective, and in the absence of specific biochemical markers, our finding, together with the recent identification of *CLN6* mutations in adult-onset recessive Kufs type A disease,[2] provide essential information allowing efficient DNA-based testing in families as well as simplex cases with ANCL presentation.

Supplemental Data

Supplemental Data include nine figures and three tables and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

This work was supported by the Grant Agency of Charles University of Prague (project 299911), the Ministry of Education of the Czech Republic (projects 1M6837805002 and MSM0021620806), Belgian Science Policy Office Interuniversity Attraction Poles program P6/43, Flemish Government Methusalem Excellence grant, Research Foundation Flanders (J.v.d.Z, postdoctoral fellowship), and the Batten Disease Support and Research Association. We thank clinical colleagues and families who contributed samples used in this study, especially John Morris and Joanne Porter (UCL563), David Sleat and the late Krystyna Wisniewski (UCL519), and Aristotle Siakotos (UCL328).

Received: May 3, 2011
Revised: July 4, 2011
Accepted: July 9, 2011
Published online: August 4, 2011

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://www.1000genomes.org/
CSS-Palm 2.0, http://csspalm.biocuckoo.org/online.php
DAVID, Database for Annotation, Visualization and Integrated Discovery version 6.7, http://david.abcc.ncifcrf.gov/
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
ExPASy, http://expasy.org
Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/
GeneReviews, Mole, S.E., and Williams, R.E. (2010). Neuronal Ceroid-Lipofuscinoses, www.ncbi.nlm.nih.gov/books/NBK1428
Online Mendelian Inheritance in man (OMIM), http://www.omim.org
PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/

R-project for Statistical Computing, http://www.r-project.org
SIFT BLink, http://sift.jcvi.org/www/SIFT_BLink_submit.html

Accession Numbers

Gene-expression data are available at the Gene Expression Omnibus (GEO) repository under accession GSE30369.

References

1. Mole, S.E., Williams, R.E., and Goebel, H.H. (2011). The Neuronal Ceroid Lipofuscinoses (Batten Disease) (Oxford: Oxford University Press).
2. Arsov, T., Smith, K.R., Damiano, J., Franceschetti, S., Canafoglia, L., Bromhead, C.J., Andermann, E., Vears, D.F., Cossette, P., Rajagopalan, S., et al. (2011). Kufs disease, the major adult form of neuronal ceroid lipofuscinosis, caused by mutations in CLN6. Am. J. Hum. Genet. *88*, 566–573.
3. Boehme, D.H., Cottrell, J.C., Leonberg, S.C., and Zeman, W. (1971). A dominant form of neuronal ceroid-lipofuscinosis. Brain *94*, 745–760.
4. Ferrer, I., Arbizu, T., Peña, J., and Serra, J.P. (1980). A golgi and ultrastructural study of a dominant form of Kufs' disease. J. Neurol. *222*, 183–190.
5. Josephson, S.A., Schmidt, R.E., Millsap, P., McManus, D.Q., and Morris, J.C. (2001). Autosomal dominant Kufs' disease: A cause of early onset dementia. J. Neurol. Sci. *188*, 51–60.
6. Burneo, J.G., Arnold, T., Palmer, C.A., Kuzniecky, R.I., Oh, S.J., and Faught, E. (2003). Adult-onset neuronal ceroid lipofuscinosis (Kufs disease) with autosomal dominant inheritance in Alabama. Epilepsia *44*, 841–846.
7. Sims, K.B., Cole, A.J., Sherman, J.C., Caruso, P.A., and Snuderl, M. (2011). Case records of the Massachusetts General Hospital. Case 8-2011. A 32-year-old woman with seizures and cognitive decline. N. Engl. J. Med. *364*, 1062–1074.
8. Nijssen, P.C., Brusse, E., Leyten, A.C., Martin, J.J., Teepen, J.L., and Roos, R.A. (2002). Autosomal dominant adult neuronal ceroid lipofuscinosis: Parkinsonism due to both striatal and nigral dysfunction. Mov. Disord. *17*, 482–487.
9. Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2002). Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat. Genet. *30*, 97–101.
10. Thiele, H., and Nürnberg, P. (2005). HaploPainter: A tool for drawing pedigrees with complex haplotypes. Bioinformatics *21*, 1730–1732.
11. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
12. Jiang, H., Orr, A., Guernsey, D.L., Robitaille, J., Asselin, G., Samuels, M.E., and Dubé, M.P. (2009). Application of homozygosity haplotype analysis to genetic mapping with high-density SNP genotype data. PLoS ONE *4*, e5280.
13. Landmann, L. (2002). Deconvolution improves colocalization analysis of multiple fluorochromes in 3D confocal data sets more than filtering techniques. J. Microsc. *208*, 134–147.
14. Manders, E.M.M., Verbeek, F.J., and Aten, J.A. (1993). Measurement of Colocalization of Objects in Dual-Color Confocal Images. Journal of Microscopy *169*, 375–382.
15. Greaves, J., Salaun, C., Fukata, Y., Fukata, M., and Chamberlain, L.H. (2008). Palmitoylation and membrane interactions of the neuroprotective chaperone cysteine-string protein. J. Biol. Chem. *283*, 25014–25026.
16. Greaves, J., and Chamberlain, L.H. (2006). Dual role of the cysteine-string domain in membrane binding and palmitoylation-dependent sorting of the molecular chaperone cysteine-string protein. Mol. Biol. Cell *17*, 4748–4759.
17. Chamberlain, L.H., and Burgoyne, R.D. (1998). The cysteine-string domain of the secretory vesicle cysteine-string protein is required for membrane targeting. Biochem. J. *335*, 205–209.
18. Swayne, L.A., Blattler, C., Kay, J.G., and Braun, J.E. (2003). Oligomerization characteristics of cysteine string protein. Biochem. Biophys. Res. Commun. *300*, 921–926.
19. Xu, F., Proft, J., Gibbs, S., Winkfein, B., Johnson, J.N., Syed, N., and Braun, J.E. (2010). Quercetin targets cysteine string protein (CSPalpha) and impairs synaptic transmission. PLoS ONE *5*, e11045.
20. Gibbs, S.J., Barren, B., Beck, K.E., Proft, J., Zhao, X., Noskova, T., Braun, A.P., Artemyev, N.O., and Braun, J.E. (2009). Hsp40 couples with the CSPalpha chaperone complex upon induction of the heat shock response. PLoS ONE *4*, e4595.
21. Sakisaka, T., Meerlo, T., Matteson, J., Plutner, H., and Balch, W.E. (2002). Rab-alphaGDI activity is regulated by a Hsp90 chaperone complex. EMBO J. *21*, 6125–6135.
22. Rosales-Hernandez, A., Beck, K.E., Zhao, X., Braun, A.P., and Braun, J.E. (2009). RDJ2 (DNAJA2) chaperones neural G protein signaling pathways. Cell Stress Chaperones *14*, 71–82.
23. Zinsmaier, K.E., Eberle, K.K., Buchner, E., Walter, N., and Benzer, S. (1994). Paralysis and early death in cysteine string protein mutants of Drosophila. Science *263*, 977–980.
24. Fernández-Chacón, R., Wölfel, M., Nishimune, H., Tabares, L., Schmitz, F., Castellano-Muñoz, M., Rosenmund, C., Montesinos, M.L., Sanes, J.R., Schneggenburger, R., and Südhof, T.C. (2004). The synaptic vesicle protein CSP alpha prevents presynaptic degeneration. Neuron *42*, 237–251.
25. Schmitz, F., Tabares, L., Khimich, D., Strenzke, N., de la Villa-Polo, P., Castellano-Muñoz, M., Bulankina, A., Moser, T., Fernández-Chacón, R., and Südhof, T.C. (2006). CSPalpha-deficiency causes massive and rapid photoreceptor degeneration. Proc. Natl. Acad. Sci. USA *103*, 2926–2931.
26. Burgoyne, R.D., and Morgan, A. (2011). Chaperoning the SNAREs: A role in preventing neurodegeneration? Nat. Cell Biol. *13*, 8–9.
27. García-Junco-Clemente, P., Cantero, G., Gómez-Sánchez, L., Linares-Clemente, P., Martínez-López, J.A., Luján, R., and Fernández-Chacón, R. (2010). Cysteine string protein-alpha prevents activity-dependent degeneration in GABAergic synapses. J. Neurosci. *30*, 7377–7391.
28. Burré, J., Sharma, M., Tsetsenis, T., Buchman, V., Etherton, M.R., and Südhof, T.C. (2010). Alpha-synuclein promotes SNARE-complex assembly in vivo and in vitro. Science *329*, 1663–1667.
29. Sharma, M., Burré, J., and Südhof, T.C. (2011). CSPα promotes SNARE-complex assembly by chaperoning SNAP-25 during synaptic activity. Nat. Cell Biol. *13*, 30–39.
30. Johnson, J.N., Ahrendt, E., and Braun, J.E. (2010). CSPalpha: The neuroprotective J protein. Biochem. Cell Biol. *88*, 157–165.
31. Nijssen, P.C., Brekelmans, G.J., and Roos, R.A. (2009). Electroencephalography in autosomal dominant adult neuronal ceroid lipofuscinosis. Clin. Neurophysiol. *120*, 1782–1786.
32. Nijssen, P.C., Ceuterick, C., van Diggelen, O.P., Elleder, M., Martin, J.J., Teepen, J.L., Tyynelä, J., and Roos, R.A. (2003). Autosomal dominant adult neuronal ceroid lipofuscinosis: A

novel form of NCL with granular osmiophilic deposits without palmitoyl protein thioesterase 1 deficiency. Brain Pathol. *13*, 574–581.

33. Poët, M., Kornak, U., Schweizer, M., Zdebik, A.A., Scheel, O., Hoelter, S., Wurst, W., Schmitt, A., Fuhrmann, J.C., Planells-Cases, R., et al. (2006). Lysosomal storage disease upon disruption of the neuronal chloride transport protein ClC-6. Proc. Natl. Acad. Sci. USA *103*, 13854–13859.
34. Reif, A., Schneider, M.F., Hoyer, A., Schneider-Gold, C., Fallgatter, A.J., Roggendorf, W., and Pfuhlmann, B. (2003). Neuroleptic malignant syndrome in Kufs' disease. J. Neurol. Neurosurg. Psychiatry *74*, 385–387.

Cerebellar dysfunction in a family harbouring the PSEN1 mutation co-segregating with a Cathepsin D variant p.A58V

Rainer Ehling[a*], Lenka Nosková[b*], Viktor Stránecký[b], Hana Hartmannová[b], Anna Přistoupilová[b], Kateřina Hodaňová[b], Thomas Benke[a], Gabor G. Kovacs[c], Thomas Ströbel[c], Ulrike Niedermüller[d], Michaela Wagner[e], Wolfgang Nachbauer[a], Andreas Janecke[f], Herbert Budka[c], Sylvia Boesch[a°] and Stanislav Kmoch[b°]

[a] *Department of Neurology, Innsbruck Medical University, Innsbruck, Austria*

[b] *Institute for Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague, Czech Republic*

[c] *Institute of Neurology, Medical University Vienna, Austria*

[d] *Department of Neurology, Krankenhaus der Barmherzigen Schwestern Ried, Austria*

[e] *Department of Radiology, Innsbruck Medical University, Innsbruck, Austria*

[f] *Division of Human Genetics, Innsbruck Medical University, Innsbruck, Austria*

* These authors contributed equally to this work.

° These authors share senior authorship.

Correspondence:

Stanislav Kmoch (genetics): Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague, Ke Karlovu 2, 128 00, Czech Republic, phone: +420 22496 7691, fax: +420 22496 7168, e-mail: skmoch@lf1.cuni.cz

Sylvia Boesch (clinics): Department of Neurology, Innsbruck Medical University, Anichstr. 35, 6020 Innsbruck, Austria, phone: 0043 512 504 81815; fax: 0043 512 504 26286 , email: sylvia.boesch@i-med.ac.at

Keywords:

Early Onset Alzheimer Disease, Adult Neuronal Ceroid Lipofuscinosis, Exome Sequencing, Presenilin1, Ataxia, Cathepsin D.

Abstract

Presenile dementia may be caused by a variety of different genetic conditions such as familial Alzheimer's disease, prion disease as well as several hereditary metabolic disorders including adult onset neuronal ceroid lipofuscinosis. We report a multigenerational family with autosomal dominant presenile dementia harboring a cerebellar phenotype. Longitudinal clinical work-up in affected family members revealed ataxia accompanied by progressive cognitive decline, rapid loss of global cognition, memory, visuospatial and frontal-executive functions accompanied by progressive motor deterioration and early death. Linkage analysis and exome sequencing identified the p.S170F mutation of Presenilin 1 in all affected individuals, which is known to be associated with very early onset Alzheimer's disease. Additional search for potentially modifying variants revealed in all affected individuals of the third generation a paternally inherited variant p.A58V (rs17571) of Cathepsin D which is considered as an independent risk factor for Alzheimer's disease. Involvement of cerebellar and brainstem structures leading to functional decortication in addition to rapid progressive presenile dementia in this PSEN1 family may therefore indicate an epistatic effect of the p.A58V Cathepsin D variant on the deleterious course of this disease.

Introduction

Early onset dementia may be caused by a variety of different conditions including familial Alzheimer's disease (AD), prion disease and several hereditary metabolic disorders including leukodystrophies and neuronal ceroid lipofuscinoses (NCLs). AD mutations are among the most frequent cause of inherited dementia. Causative mutations for autosomal dominant AD have been found in the amyloid -β protein precursor (AβPP) as well as in the Presenilin 1 (*PSEN1*) and Presenilin 2 (*PSEN2*) genes (1). Among them, point mutations in the *PSEN1* are most common (www.molgen.ua.ac.be). They have been associated with severe forms of AD characterized by complete penetrance and disease onset as early as 30 years of age. Besides progressive dementia, there is considerable clinical heterogeneity in patients with PSEN1 AD ranging from parkinsonism to spasticity (2).

NCLs are hereditary progressive neurodegenerative diseases with an incidence ranging in different countries from 1.3 to seven per 100.000 live births. Initial signs and symptoms of adult NCL usually appear around age 30 years, with death occurring about ten years later. Clinical profile of adult NCL type 4 (NCL4) shares common features with early onset AD and is characterized by behavior abnormalities and dementia which may be associated with motor dysfunction, ataxia, extra pyramidal and brainstem signs (3).

We report a large Austrian family encompassing five affected individuals in three generations who proved to suffer from autosomal dominant dementia presenting with a predominant cerebellar phenotype. With respect to the consistently severe phenotype and the known heterogeneity of familial presenile dementia, genome-wide linkage analysis and exome sequencing were performed as comprehensive genetic testing and allowed for detection of multiple genetic mutations which traditional genetic approaches would likely have not identified. This together with extensive clinical, neuroradiological evaluation and neuropsychological work up of this family provides for the first time a comprehensive longitudinal insight in clinical signs

and symptoms of a family with a specific PSEN1 mutation associated with an additional AD modifier variant in Cathepsin D (CTSD), a causal gene for NCL type 10.

Materials and methods

Clinical evaluation

Probands were identified from a multigenerational Austrian family exhibiting rapid progressive dementia in combination with an ataxic movement disorder leading to death within few years after onset. Disease developed in 5 individuals out of 3 generations of this family (Figure 1). Three affected individuals in the third generation (III.3, III.5 and III.6) were consecutively referred to the Department of Neurology in Innsbruck, Medical University Innsbruck within several years after clinical onset. Standard neuropsychological tests were used to assess dementia. Clinical information for deceased affected individuals (I.2 and II.2) was obtained by interview of collateral sources and review of available medical records. Investigations were conducted according to the Declaration of Helsinki principles and patients and family members gave written informed consent.

Neuropathological examination

Neuropathological study was performed on formalin-fixed, paraffin-embedded blocks. Sections of frontal cortex, amygdala, anterior segment of the hippocampus with transentorhinal and entorhinal cortex, caudate nucleus, putamen and cerebellar cortex were examined. Sections were stained using hematoxylin and eosin, Klüver-Barrera, and modified Bielschowsky methods. The following monoclonal antibodies were used for immunohistochemistry: anti-tau AT8 (pS202, 1:200, Pierce Biotechnology, Rockford, IL, USA), anti-phospho-TDP-43 (pS409/410, 1:20,000, Cosmo Bio, Tokyo, Japan), anti-α-synuclein (1:10,000, clone 4D6, Signet, Dedham, MA, USA), anti-Aβ (1:100, clone 6F/3D, Dako, Glostrup, Denmark), anti-PrP (1:2,000, 12F10, Cayman

Chemical, Ann Arbor, MI, USA). The DAKO EnVision© detection kit, peroxidase/DAB, rabbit/mouse (Dako) was used for visualization of antibody reactions.

Genotyping and linkage analysis

Genomic DNA was isolated by standard technology. All DNA samples were genotyped using Affymetrix GeneChip® Mapping 10K 2.0 Arrays (Affymetrix, Santa Clara, CA) according to the manufacturer's protocol at the microarray core facility of the Institute of Molecular Genetics in Prague. Raw feature intensities were extracted from the Affymetrix GeneChip Scanner 3000 7G images using the GeneChip operating Software (GCOS) 1.4. Individual SNP calls were generated using Affymetrix Genotyping Analysis Software (GTYPE) 4.1.

Multipoint parametric linkage analysis along with determination of the most likely haplotypes, was carried out using affected-only analysis under the assumption of an autosomal dominant mode of inheritance with a 0.99 constant, age-independent penetrance, 0.001 phenocopy rate, and 0.001 frequency of disease allele, and was performed with version 1.1.2 of Merlin software (4). The results were visualized in the version 1.032 of the HaploPainter software (5) and in version 2.9.2 of R-project statistical software.

Copy-number analysis

DNA sample from individual III.5 was genotyped using Affymetrix GeneChip Mapping 6.0 Array (Affymetrix, Santa Clara, CA) at the microarray core facility of the Institute of Molecular Genetics in Prague according to the manufacturer's protocol. Raw feature intensities were extracted from the Affymetrix GeneChip Scanner 3000 7G images using the GeneChip Control Console Software 2.01. Individual SNP calls were generated using Affymetrix Genotyping Console Software 3.02. Copy number changes were identified in Affymetrix Genotyping Console Software (GTC version 3.02). Data from both SNP and copy number probes were used

to identify copy number aberrations compared to a built-in reference. Only regions larger than 10 Kb containing at least 5 probes were reported.

Exome sequencing

Exome sequencing was performed using 3 µg of DNA from two affected individuals (III.5 and III.6) and one unaffected relative (III.4). For DNA enrichment, SureSelect All Exome Kit (Agilent, Santa Clara, USA) was used according to the manufacturer's protocol. DNA sequencing was performed on the captured barcoded DNA library using SOLiD™ 4 System (Applied Biosystems, Carlsbad, USA) at the Institute for Inherited Metabolic Disorders (Prague, Czech Republic). Reads were aligned in color space to the reference genome (hg19) using NovoalignCS version 1.08 (Novocraft, Malaysia) with default parameters. Sequence variants in the analyzed sample were identified using SAMtools package (version 0.1.8) (6). The high confidence variants list was annotated using ANNOVAR Annotation tool (hg19). Only the sequence variants present in both affected individuals and not found in unaffected relative and having frequency lower than 0.001 in the dbSNP, 1000 Genomes, Exome Variant Server (http://evs.gs.washington.edu/EVS/) and internal exome database were prioritized for further analysis. Candidate variants were visualized in Integrative Genomics Viewer (IGV) - version 1.5.65.

Functional annotation of sequence variants

Single nucleotide variants found by SAMtools package were analyzed using Variant Annotation, Analysis and Search Tool version 1.0.1. (VAAST) (7). Variants present in both affected individuals III.5 and III.6, but not in unaffected relative III.4 were selected using Variant Selection Tool (VST, part of VAAST). Variant analysis was performed under the dominant model of inheritance. Candidate genes were functionally annotated using Database for Annotation, Visualization and Integrated Discovery version 6.7 (DAVID).

DNA sequencing and mutation analysis

Mutation-bearing fragments of *PSEN1* (NM_000021.3) and *CTSD* (NM_001909.4) were PCR amplified from genomic DNA of all available individuals from the family and sequenced using version 3.1 Dye Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA) with electrophoresis on an ABI 3500XL Avant Genetic Analyzer (Applied Biosystems). Data were analyzed using Sequencing Analysis software, and the segregation of the candidate mutations was assessed. Primer sequences are available in Supplemental Data.

Results

Family characteristics, clinical findings, neuropathology and genetics

The three-generation pedigree includes five affected individuals of whom clinical reports are available. The grandmother of the index case (individual I.2) died in her early thirties of a disease characterized by abnormal movements and rapid dementia. A rare disease has been suspected although medical records are not available. One of her daughters (individual II.2) exhibited rapid progressive dementia associated with gait unsteadiness and dysarthria starting at the age of 30 with disease duration of four years. In her neuropathological file (1977) “massive cerebral atrophy with narrowed gyri and widened sulci, but of regular configuration” was described. Unfortunately, histopathological work-up was not accessible.

The offspring of patient II.2 consisted of six children. Three of them were affected by the disease. Extensive work-up including clinical follow-up, genetic confirmation and neuropathological examination has been performed:

Individual III.3

This female patient presented with gradual worsening of gait and stance as well as recurrent falls at the age of 33. Rapid progression of symptoms consisted of a reduction in fine motor skills such as writing and progressive loss of postural stability within the following year. Moreover, a

gradual loss in short term memory and affective lability was noted. She became unable to manage activities of daily life and had to end up her profession as a shop assistant.

After disease duration of two years she was examined at the Department of Neurology, Medical University Innsbruck. The patient showed ataxia with marked intention tremor. Her gait was severely ataxic, unassisted walking was not possible. Dystonic features such as abnormal flexion of the hands during movements were noted. Additionally, trunk dystonia was observed while walking. Pyramidal signs were present, thus spasticity was not the prevailing symptom. Severe cognitive impairment and frontal release signs were noted. Communication was markedly impaired because of massive dysarthria. Seizures and myoclonic jerks were absent. About 2 years later the patient was wheel chair bound due to massive ataxia. Additionally, repetitive flexion of arms and extension of legs during active movements resembling decortication pattern occurred. Cognitive decline was highly advanced making a detailed neuropsychological assessment impossible. The patient died few months later because of aspiration pneumonia at the age of 37.

Neuropsychological investigations during admission to our hospital two years after disease onset revealed severe dementia with an MMSE score of 11 and pronounced impairments of memory, visuospatial and executive functions; Table 1). Magnetic resonance imaging (MRI) revealed moderate parietal and mild global cortical atrophy (Figure 2A) as well as hyperintensities bilaterally (FLAIR) in the mesial temporal lobe (hippocampus, amygdala; Figure 2B). TC-99m Ethylencysteinat-Dimer perfusion SPECT detected bilateral temporal and parietal hypointensities (Figure 2C). Standard cerebrospinal fluid (CSF) examination was unremarkable. Motor evoked potentials were consistent with clinical findings and showed involvement of the pyramidal tract. Neuropathological work-up was performed in this patient (Figure 3).

Individual III.5

This female patient reported problems in short term memory that affected activities of daily and professional life at the age of 28 years. Some time before the patient noted depressive symptoms

and personality changes causing family problems. Approximately two years later gait instability, problems in motor skills and difficulties in pronouncing words were noted. Clinical examination at the age of 32, four years after disease onset, severe cognitive impairment including memory, constructional praxis and frontal executive functions (MMSE 14; Table 1) was present. Additionally, she was suffering from severe hemispheric ataxia with marked dysmetria in directed movements. Her gait was ataxic and postural instability led to falling during pull test. Pyramidal involvement with positive Babinski signs was present. Dystonic features, myoclonus and seizures were absent. The course of disease in this patient was rapidly progressive and she died one year later. Cerebral MRI exhibited mild parietal cortical atrophy (Figure 2D). Routine CSF was unremarkable. CSF amyloid-β (1-42) was markedly reduced, while total TAU was within normal range and phosphorylated TAU was increased. Neuropathological work-up is not available in this patient.

Individual III.6

In this male patient speech impairment was noted as a first symptom at the age of 31 years, followed by gait disturbance and problems in upper limb coordination. Within several months disabling difficulties in short term memory tasks interfering with professional life occurred. At first clinical examination two years later, the patient exhibited marked dysarthria, a reduction in fine motor skills as well as gait ataxia. Additionally, impairment of cognitive functions was present. Clinical follow-up within three years revealed a rapidly progressive cognitive deterioration. Dysmetria and ataxic gait disorder deteriorated less than cognitive functions, he was still able to perform activities of daily living. At the age of 36 and disease duration of five years he had moderate dysarthria, marked gait ataxia and dementia (MMSE 23; deficits of verbal and figural memory, executive functions; Table 1). This patient is followed up until to date.

Neuropsychological evaluation detected prominent impairment of orientation and memory, visuospatial functions, executive functions, and working memory as assessed by standardized neuropsychological tests (Table 1). CSF amyloid-β (1-42) was markedly reduced. Total TAU as

well as phosphorylated TAU was found increased. EEG monitoring lasting one week did not reveal seizures or epileptic discharges but showed low pseudoperiodic sequences. Cerebral MRI exhibited moderate parietal and mild cortical atrophy, corresponding to a decreased metabolism in fluoro-2-deoxy-D-glucose PET (Figure 2E and F).

Neuropathological findings in patient III.3

Neuropathology was available for patient III.3. Major microscopic features included spongiosis of the superficial layers in the frontal and entorhinal cortices and neuronal loss accompanied by reactive astrogliosis predominating in frontal and entorhinal cortices, and amygdala. In addition, abundant amyloid plaques were noted together with dystrophic neurites and congophilic amyloid angiopathy. Bielschowsky silver staining revealed numerous neuritic plaques and neurofibrillary tangles in these regions (Figure 3A). There was no spongiform change of the neuropil, and immunostaining for prion protein excluded disease-associated deposits. In addition, we did not observe prominent accumulation of lipofuscin in neurons as seen in NCL patients.

Immunostaining for amyloid-β (Figure 3B) revealed abundant cored and diffuse plaques and vascular, perivascular and subpial amyloid-β deposits in the frontal cortex. Amyloid-β deposition was also prominent in the amygdala and transentorhinal/entorhinal cortex, here also in the form of lake-like granulofibrillar deposits. The latter were observed in the basal ganglia as well. Occasional small cored but more diffuse deposits in the molecular layer following the arborization of the Purkinje cells arranged perpendicular to the surface were seen in the cerebellum.

Identification of sequence variants in two AD associated genes

SNP array-based linkage analysis identified 14 candidate regions with positive LOD scores on chromosomes 3, 4, 8, 9, 10, 13, 14, 16 and 19 (Figure 4A); a high-resolution SNP array copy

number analysis in one patient (III.5), found no indication for a potentially disease-causing deletion or duplication.

To directly identify possible disease-causing mutation(s) among the candidate genes defined by linkage analysis, we performed exome sequencing in genomic DNA of two affected individuals (III.5 and III.6), and one unaffected relative (III.4). From the sequencing run we obtained from these samples 105, 114 and 109M of sequencing reads, respectively, of which we were able to map on average 69 % on the human genome reference sequence. After removing PCR-generated duplicate reads, we obtained for the samples 56, 60 and 59M of unique reads, respectively. We identified 11726 single nucleotide variants (SNVs) and 150 indels that were present in heterozygous form in both affected probands, but were absent in the unaffected relative (candidate variants). From these candidate variants, 65 SNVs and none of the indels were either novel or present at frequencies lower than 0.001 in the dbSNP, 1000 Genomes, Exome Variant Server (http://evs.gs.washington.edu/EVS/) and internal exome database.

Filtering the candidate variants for linkage regions left 7 potentially disease-causing variants (Table 2), of which the heterozygous mutation c.509C>T (p.Ser170Phe) in *PSEN1* (Figure 4B) emerged as causative. This mutation is present in the dbSNP database with accession number rs63750577 and is referred to as pathogenic. It is acknowledged as a disease-associated mutation in the Alzheimer Disease & Frontotemporal Dementia Mutation Database (http://www.molgen.ua.ac.be). Our sequence analysis of the mutation-bearing fragment showed its segregation with the phenotype in the family (Figure 4C).

To identify potentially disease-modifying variants we functionally annotated the candidate variants identified by exome sequencing. This analysis revealed a heterozygous mutation c.C173>T (p.A58V) in *CTSD*, and sequence analysis showed that all three affected individuals of the last generation bearing PSEN1 p.S170F mutation are heterozygous for this mutation, which they inherited from their unaffected father (not shown).

Discussion

We report a large multigenerational family segregating a dominantly inherited phenotype characterized by cerebellar ataxia and premature dementia segregating with a PSEN1 p.S170F mutation.

PSEN1 is a membrane protein that is part of the γ-secretase complex (8). To date 185 mutations in PSEN1 have been described leading to autosomal dominant AD. A number of PSEN1 missense mutations alter the specificity of γ-secretase cleavage leading to increased proportions of longer forms of amyloid-β peptide (Aβ), which are prone to oligomerization and aggregation in amyloid plaques (9, 10). The p.S170F mutation affects a conserved residue located in the transmembrane domain of the protein (11). The mutation is known to be associated with the early onset of Alzheimer disease probably due to a largely increased secretion Aβ(1–42) and accelerated Aβ plaque deposition (12). PSEN1 is also required for autophagy and its mutations may lead to impairment of autophagosome maturation as a result of a selective impairment of autolysosome acidification and cathepsin activation (13). Presence of anbnormal autophagic activity is frequently observed in common neurodegenerative disorders and its role in accummulation of protein aggregates is frequently discussed (14).

Phenotype comparison with another four patients harbouring the p.S170F mutation who have been reported in the literature until to date, (11, 15-17), revealed that in two reported cases the motor symptoms encompassed a prominent cerebellar phenotype early in the disease which is comparable to our cases : i) Piccini and co-workers reported a 28 years old male with delusions and lower limb jerks accompanied by intentional myoclonus and ataxia (16), and ii) another single PSEN1 p.S170F case initially exhibited head titubation, truncal and mild upper limb as well as gait ataxia (15). In contrast, another two individual PSEN1 p.S170F cases exhibited a phenotype characterized by myoclonic jerks, parkinsonism and progressive cognitive decline (11, 17). Epileptic seizures were reported in three out of four cases with a PSEN1 p.S170F mutation (11, 15, 17) but were not observed in our family. Rapid deterioration of motor skills in

our PSEN1 family led to bilateral dystonic flexion – extension pattern accompanied by severe dysphagia. Similar symptoms in late stage disease have also been observed by Piccini and colleagues (16). The above-mentioned clinical motor pattern together with pathological crying indicates functional decerebration in late stage disease.

Detailed neuropsychological assessment of three PSEN1 family members revealed progressive impairment of orientation and memory, visuospatial and executive functions in early disease. Verbal fluency as well as set shifting abilities and working memory functions such as naming, higher motor functions and psychomotor speed were impaired as often seen in late-onset AD cases. However, cognitive decline occurred even earlier in this family than in previous series of AD with PSEN1 mutations (18). Extensive neuropsychological assessment revealed emotional blunting, depression, anxiousness and apathy as well as pseudobulbar crying. Behavioral abnormalities as seen in early disease stage in our series are frequently observed in patients with cerebellar lesions (19). Moreover, together with the pattern of cognitive decline they are suggestive of a cerebellar cognitive affective syndrome as described by Schmahmann and co-workers (20-23). Later on, cognitive deterioration was associated with parietal, temporal and occipital perfusion deficits on functional neuroimaging, indistinguishable from classical 'cortical' AD. Global cortical atrophy emerged in two out of three patients while predominant cerebellar atrophy was not present. Of note, one patient showed FLAIR hyperintensities in the medial temporal region resembling limbic encephalitis. This feature was previously described in a sporadic PSEN1 p.S170F case (15).

Neuropathological evaluation revealed abundant amounts of amyloid plaques accompanied by dystrophic axons and neurofibrillary degeneration as typically found in AD (24). Immunostaining revealed wide-spread amyloid-β deposits. Diffuse amyloid plaques in the cerebellar molecular layer following the arborization of the Purkinje cells were additionally found. This has previously been shown in another individual harboring the S170F mutation (16).

Some features in our patients are reminiscent of adult onset NCL with an autosomal dominantly inherited, neurodegenerative, lysosomal storage disorders characterized by autofluorescent storage material in neuronal tissues, progressive intellectual and motor deterioration, seizures, and early death (Parry type). The here reported family was tested negative for *DNAJC5* mutation - family UCL568, (25), and it is of interest that all affected individuals of the third generation carry in addition to the PSEN1 p.S170F mutation descending from their mother, also a heterozygous variant p.A58V (rs17571) of CTSD, which they inherited from their father. CTSD is likely involved in AβPP processing (26) and recessive mutations of CTSD cause NCL10 (27). In addition, the role of CTSD variants has been studied in AD (28) as well as in other neurodegenerative diseases (29-31), and specifically the heterozygous CTSD variant p.A58V has been associated with an increased risk for AD (32). The presence of the p.A58V variant may thus have an additional impact on the extremely early onset and course of disease in our PSEN1 S170F family.

In conclusion, our work demonstrates effectiveness of genome-wide linkage analysis and exome sequencing based genetic testing in single families. It also endorses previous reports on a cerebellar phenotype in p.S170F patients, which even may precede cognitive decline and AD type dementia. The course of disease suggests a consecutive involvement of cerebellar and brainstem structures leading to functional decortication and death. Some features in our patients are reminiscent of adult onset NCLs with autosomal dominant inheritance, and the presence of the p.A58V variant of Cathepsin D is suggestive of an epistatic effect of this NCL variant on the deleterious course of this disease.

Acknowledgement:

This work was supported by institutional programs of Charles University in Prague PRVOUK-P24/LF1/3, UNCE 204011 and SVV2012/ 2645 and by grants from the Grant Agency of Charles University of Prague (project 299911) and from the Ministry of Health of the Czech Republic (project NT13116-4/2012).

Conflicts of interest and Disclosure:

None.

References

1. Guerreiro RJ, Baquero M, Blesa R, Boada M, Bras JM, Bullido MJ, et al. Genetic screening of Alzheimer's disease genes in Iberian and African samples yields novel mutations in presenilins and APP. Neurobiology of aging. 2010 May;31(5):725-31. PubMed PMID: 18667258. Pubmed Central PMCID: 2850052. Epub 2008/08/01. eng.

2. Larner AJ, Doran M. Genotype-phenotype relationships of presenilin-1 mutations in Alzheimer's disease: an update. Journal of Alzheimer's disease : JAD. 2009;17(2):259-65. PubMed PMID: 19221408. Epub 2009/02/18. eng.

3. Kousi M, Lehesjoki AE, Mole SE. Update of the mutation spectrum and clinical correlations of over 360 mutations in eight genes that underlie the neuronal ceroid lipofuscinoses. Human mutation. 2012 Jan;33(1):42-63. PubMed PMID: 21990111. Epub 2011/10/13. eng.

4. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nature genetics. 2002 Jan;30(1):97-101. PubMed PMID: 11731797. Epub 2001/12/04. eng.

5. Thiele H, Nurnberg P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics. 2005 Apr 15;21(8):1730-2. PubMed PMID: 15377505. Epub 2004/09/21. eng.

6. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078-9. PubMed PMID: 19505943.

7. Yandell M, Huff C, Hu H, Singleton M, Moore B, Xing J, et al. A probabilistic disease-gene finder for personal genomes. Genome Research. 2011 SEP 2011;21(9):1529-42. PubMed PMID: WOS:000294477000013. English.

8. Selkoe DJ, Wolfe MS. Presenilin: running with scissors in the membrane. Cell. 2007 Oct 19;131(2):215-21. PubMed PMID: 17956719. Epub 2007/10/25. eng.

9. Citron M, Westaway D, Xia W, Carlson G, Diehl T, Levesque G, et al. Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. Nature medicine. 1997 Jan;3(1):67-72. PubMed PMID: 8986743. Epub 1997/01/01. eng.

10. Scheuner D, Eckman C, Jensen M, Song X, Citron M, Suzuki N, et al. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. Nature medicine. 1996 Aug;2(8):864-70. PubMed PMID: 8705854. Epub 1996/08/01. eng.

11. Golan MP, Styczynska M, Jozwiak K, Walecki J, Maruszak A, Pniewski J, et al. Early-onset Alzheimer's disease with a de novo mutation in the presenilin 1 gene. Experimental neurology. 2007 Dec;208(2):264-8. PubMed PMID: 17931627. Epub 2007/10/13. eng.

12. Boyle JP, Hettiarachchi NT, Wilkinson JA, Pearson HA, Scragg JL, Lendon C, et al. Cellular consequences of the expression of Alzheimer's disease-causing presenilin 1 mutations in human neuroblastoma (SH-SY5Y) cells. Brain research. 2012 Mar 14;1443:75-88. PubMed PMID: 22297172. Epub 2012/02/03. eng.

13. Lee JH, Yu WH, Kumar A, Lee S, Mohan PS, Peterhoff CM, et al. Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations. Cell. 2010 Jun;141(7):1146-58. PubMed PMID: 20541250. eng.

14. Son JH, Shim JH, Kim KH, Ha JY, Han JY. Neuronal autophagy and neurodegenerative diseases. Exp Mol Med. 2012 Feb;44(2):89-98. PubMed PMID: 22257884. Pubmed Central PMCID: PMC3296817. eng.

15. Langheinrich TC, Romanowski CA, Wharton S, Hadjivassiliou M. Presenilin-1 mutation associated with amnesia, ataxia, and medial temporal lobe T2 signal changes. Neurology. 2011 Apr 19;76(16):1435-6. PubMed PMID: 21502605. Pubmed Central PMCID: 3087402. Epub 2011/04/20. eng.

16. Piccini A, Zanusso G, Borghi R, Noviello C, Monaco S, Russo R, et al. Association of a presenilin 1 S170F mutation with a novel Alzheimer disease molecular phenotype. Archives of neurology. 2007 May;64(5):738-45. PubMed PMID: 17502474. Epub 2007/05/16. eng.

17. Snider BJ, Norton J, Coats MA, Chakraverty S, Hou CE, Jervis R, et al. Novel presenilin 1 mutation (S170F) causing Alzheimer disease with Lewy bodies in the third decade of life. Archives of neurology. 2005 Dec;62(12):1821-30. PubMed PMID: 16344340. Epub 2005/12/14. eng.

18. Verkkoniemi A, Ylikoski R, Rinne JO, Somer M, Hietaharju A, Erkinjuntti T, et al. Neuropsychological functions in variant Alzheimer's disease with spastic paraparesis. Journal of the neurological sciences. 2004 Mar 15;218(1-2):29-37. PubMed PMID: 14759630. Epub 2004/02/05. eng.

19. Parvizi J, Anderson SW, Martin CO, Damasio H, Damasio AR. Pathological laughter and crying: a link to the cerebellum. Brain : a journal of neurology. 2001 Sep;124(Pt 9):1708-19. PubMed PMID: 11522574. Epub 2001/08/28. eng.

20. Bellebaum C, Daum I. Cerebellar involvement in executive control. Cerebellum. 2007;6(3):184-92. PubMed PMID: 17786814. Epub 2007/09/06. eng.

21. Molinari M, Petrosini L, Misciagna S, Leggio MG. Visuospatial abilities in cerebellar disorders. Journal of neurology, neurosurgery, and psychiatry. 2004 Feb;75(2):235-40. PubMed PMID: 14742596. Pubmed Central PMCID: 1738892. Epub 2004/01/27. eng.

22. Ravizza SM, McCormick CA, Schlerf JE, Justus T, Ivry RB, Fiez JA. Cerebellar damage produces selective deficits in verbal working memory. Brain : a journal of neurology. 2006 Feb;129(Pt 2):306-20. PubMed PMID: 16317024. Epub 2005/12/01. eng.

23. Schmahmann JD, Sherman JC. The cerebellar cognitive affective syndrome. Brain : a journal of neurology. 1998 Apr;121 (Pt 4):561-79. PubMed PMID: 9577385. Epub 1998/05/13. eng.

24. Duyckaerts C, Delatour B, Potier MC. Classification and basic pathology of Alzheimer disease. Acta neuropathologica. 2009 Jul;118(1):5-36. PubMed PMID: 19381658. Epub 2009/04/22. eng.

25. Noskova L, Stranecky V, Hartmannova H, Pristoupilova A, Baresova V, Ivanek R, et al. Mutations in DNAJC5, encoding cysteine-string protein alpha, cause autosomal-dominant adult-onset neuronal ceroid lipofuscinosis. American journal of human genetics. 2011 Aug 12;89(2):241-52. PubMed PMID: 21820099. Pubmed Central PMCID: 3155175. Epub 2011/08/09. eng.

26. Dreyer RN, Bausch KM, Fracasso P, Hammond LJ, Wunderlich D, Wirak DO, et al. Processing of the pre-beta-amyloid protein by cathepsin D is enhanced by a familial Alzheimer's disease mutation. European journal of biochemistry / FEBS. 1994 Sep 1;224(2):265-71. PubMed PMID: 7523115. Epub 1994/09/01. eng.

27. Steinfeld R, Reinhardt K, Schreiber K, Hillebrand M, Kraetzner R, Bruck W, et al. Cathepsin D deficiency is associated with a human neurodegenerative disorder. American journal of human genetics. 2006 Jun;78(6):988-98. PubMed PMID: 16685649. Pubmed Central PMCID: 1474096. Epub 2006/05/11. eng.

28. Davidson Y, Gibbons L, Pritchard A, Hardicre J, Wren J, Tian J, et al. Genetic associations between cathepsin D exon 2 C-->T polymorphism and Alzheimer's disease, and pathological correlations with genotype. Journal of neurology, neurosurgery, and psychiatry. 2006 Apr;77(4):515-7. PubMed PMID: 16543533. Pubmed Central PMCID: 2077521. Epub 2006/03/18. eng.

29. Jeong BH, Lee KH, Lee YJ, Yun J, Park YJ, Kim YH, et al. Genetic polymorphism in exon 2 of cathepsin D is not associated with vascular dementia. Acta neurologica Scandinavica. 2011 Jun;123(6):419-23. PubMed PMID: 20597865. Epub 2010/07/06. eng.

30. Kovacs GG, Sanchez-Juan P, Strobel T, Schuur M, Poleggi A, Nocentini S, et al. Cathepsin D (C224T) polymorphism in sporadic and genetic Creutzfeldt-Jakob disease.

Alzheimer disease and associated disorders. 2010 Jan-Mar;24(1):104-7. PubMed PMID: 19571726. Epub 2009/07/03. eng.

31. Schulte T, Bohringer S, Schols L, Muller T, Fischer C, Riess O, et al. Modulation of disease risk according to a cathepsin D / apolipoprotein E genotype in Parkinson's disease. J Neural Transm. 2003 Jul;110(7):749-55. PubMed PMID: 12811635. Epub 2003/06/18. eng.

32. Schuur M, Ikram MA, van Swieten JC, Isaacs A, Vergeer-Drop JM, Hofman A, et al. Cathepsin D gene and the risk of Alzheimer's disease: a population-based study and meta-analysis. Neurobiology of aging. 2011 Sep;32(9):1607-14. PubMed PMID: 19926167. Epub 2009/11/21. eng.

Legends to Figures

Figure 1 Pedigree. Squares = males; circles = females; filled symbols = affected subjects; diagonal line = deceased subjects. Age of onset is shown above current age or age of death (indicated by +). * indicates subjects heterozygous for c.C173>T (pAla58Val) mutation in cathepsin D.

Figure 2 Cerebral MRI of patient III.3 revealed moderate parietal and mild global cortical atrophy (**A**) as well as hyperintensities bilaterally (FLAIR) in the mesial temporal lobe (hippocampus, amygdala; **B**). TC-99m Ethylencysteinat-Dimer perfusion SPECT of patient III.3 detected bilateral temporal and parietal hypointensities (**C**). Cerebral MRI of patient III.5 exhibited mild parietal cortical atrophy (**D**). Cerebral MRI of patient III.6 exhibited moderate parietal and mild cortical atrophy, corresponding to a decreased metabolism in fluoro-2-deoxy-D-glucose PET (**E** and **F**).

Figure 3 Hematoxylin and eosin (H&E) and Bielschowsky silver staining in the frontal cortex (Cx) and amygdala. Note the angiopathy and gliosis in the frontal cortex together with high density of argyrophilic neuritic plaques (two images at left). Prominent astrogliosis and large neuritic plaques were observed in the amygdala (two images at right). Bar graph in the left image represent 100 μm (**A**). Immunostaining for phospho-Tau (AT8) and amyloid-β in the frontal and tranentorhinal cortices, putamen, and cerebellum. Note the abundance of neuropil threads and neurofibrillary tangles in the frontal and transentorhinal cortices, less in the putamen, and lacking in the cerebellar cortex. In contrast, amyloid-β deposits were seen in all examined regions but with distinct morphology: cored and diffuse plaques in the forntal cortex, granulofibrillar diffuse deposits in the transentorhinal cortex, putamen, and cerebellum together with amyloid angiopathy. Note the stripe-like pattern in the molecular layer of the cerebellar cortex following the arborization of Purkinje cells (**B**).

Figure 4 A whole-genome parametric linkage analysis showing candidate regions reaching the theoretical maximum LOD scores of 1.8 attainable in this family on chromosomes 3, 8, 10, 13,

14, 16 and 19 (**A**). Integrative genome viewer (IGV) display showing heterozygous mutation c.509C>T of *PSEN1* in the proband (**B**). Chromatograms of *PSEN1* genomic DNA sequences showing identified heterozygous mutation. (Upper panel) Sequence showing heterozygous mutation c.509C>T in the proband, and (lower panel) sequence of an unaffected individual (C).

Figure 1

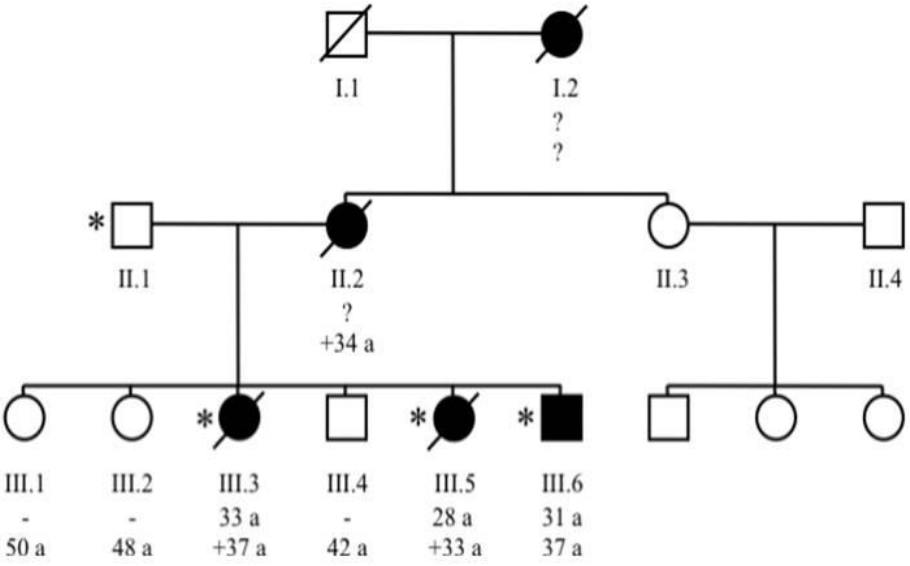

I.1
I.2
?
?
*
II.1
II.2
?
+34 a
II.3
II.4
*
*
*
III.1
-
50 a
III.2
-
48 a
III.3
33 a
+37 a
III.4
-
42 a
III.5
28 a
+33 a
III.6
31 a
37 a

Figure 2

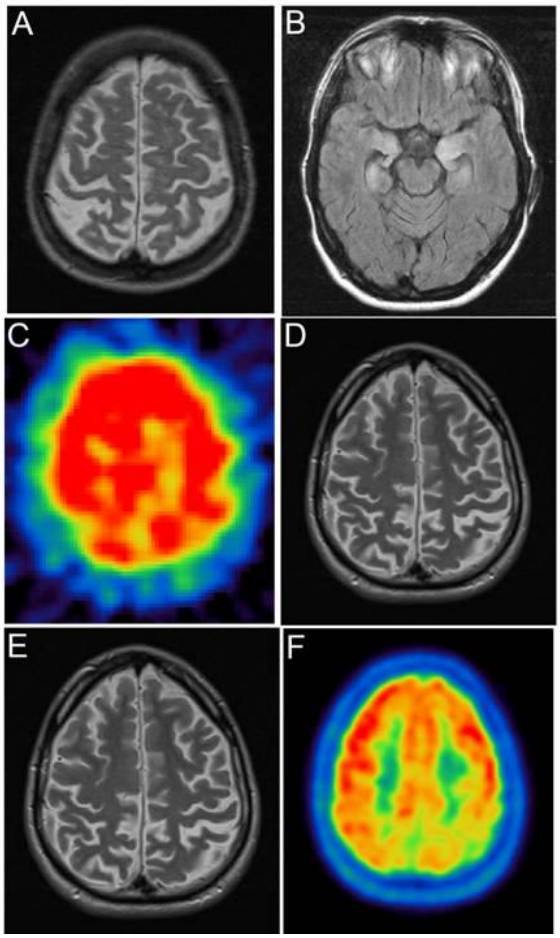


Figure 3

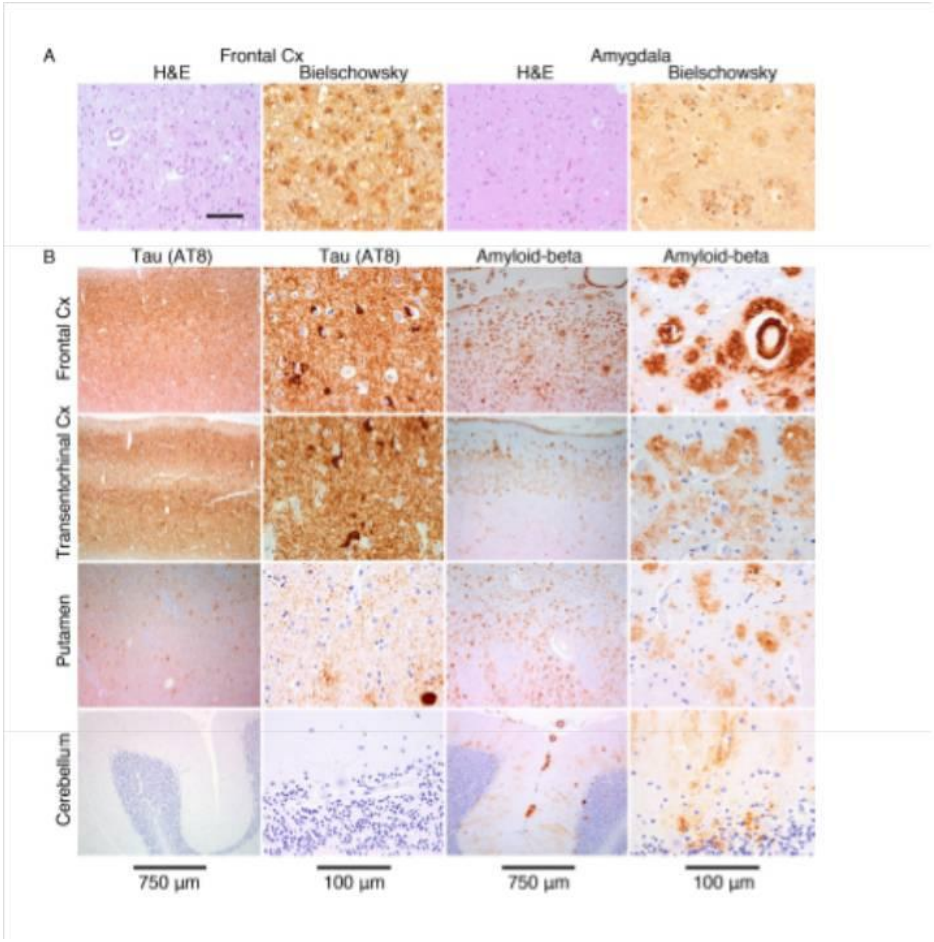

A
Frontal Cx
H&E
Bielschowsky
Amygdala
H&E
Bielschowsky
B
Tau (AT8)
Tau (AT8)
Amyloid-beta
Amyloid-beta
Frontal Cx
Transentorhinal Cx
Putamen
Cerebellum
750 µm
100 µm
750 µm
100 µm

Figure 4

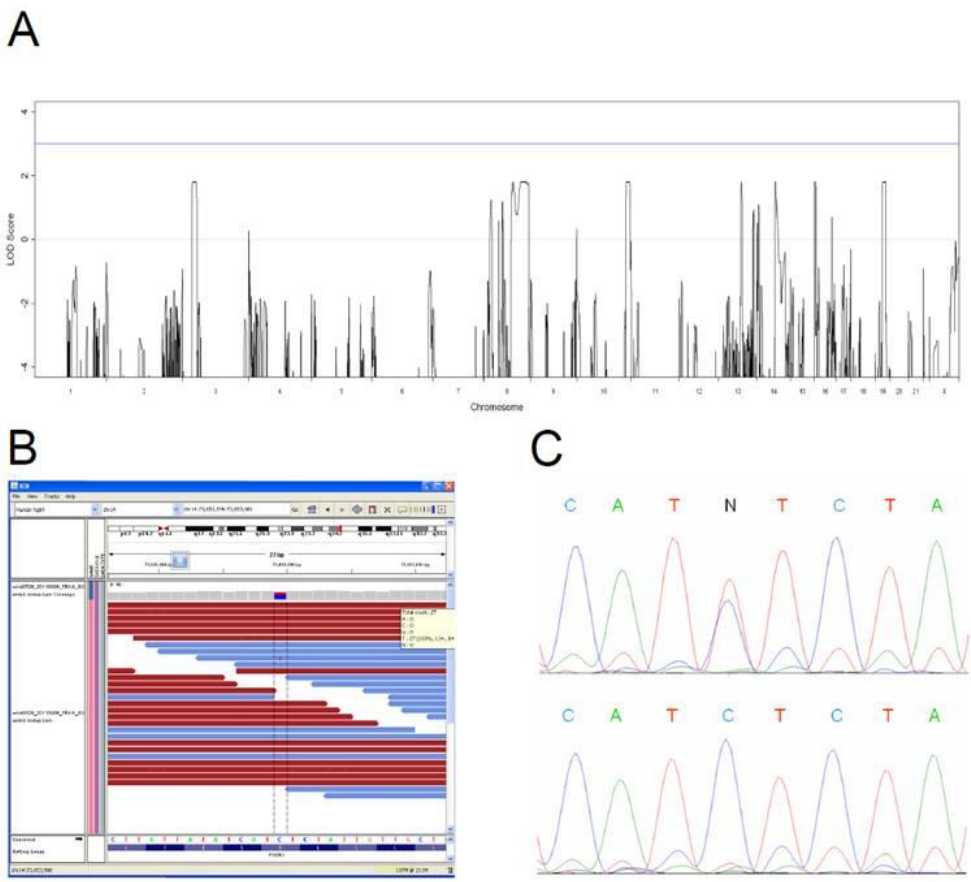


Table 1. Neuropsychological testing in affected individuals of generation III.

affected individual	III.3	III.5	III.6
age at disease onset (years)	33	28	31
disease duration (years)	4+	5+	5
presenting symptom	gait ataxia	memory loss	dysarthria
further symptoms	memory loss	dysarthria, ataxia	ataxia, memory loss
neuropsychological testing (max. possible score)			
age at testing	36	32	35
MMSE (30)	11	14	23
Word list memory (30)	9 (3-3-3)	11 (3-3-5)	9 (2-3-4)
Word list recall (10)	0	0	0
Word list recognition (10)	10	7	9
WLM, false positives / intrusions	0 / 2	0 / 6	0 / 0
Constructional praxis (11)	5	5	10
Constructional praxis, recall (11)	2	5	0
Verbal fluency (animals)	12	15	23
Phonological fluency (letter s)	not performed	4	17
Boston naming test, short version (15)	8	13	14
Clock drawing task (15)	2	8	14
Trail making test A	Aborted	179 sec	94 sec
Trail making test B	Aborted	aborted	aborted
Digit span forward / backward	3 / 2	3 / 2	4 / 3
Other cognitive impairments (bedside testing)	limb apraxia, verbal conceptualization	limb apraxia, language comprehension	limb apraxia, mental calculation, interference naming
Motor speech	severe dysarthria	moderate dysarthria	moderate dysarthria
Psychiatric symptoms	anxiousness and depression	anxiousness and depression,	anxiousness and depression

	pathological crying and laughing	emotional lability	
cerebral MRI	moderate parietal and mild global cortical atrophy hyperintensities on FLAIR bilaterally in mesial temporal lobe	mild parietal cortical atrophy	moderate parietal and mild cortical atrophy
SPECT / PET	bilateral temporal and parietal hypointensities	-	bilateral parietal hypometabolism

28

Table 2. Novel or rare coding variants revealed by exome sequencing and localized in linkage intervals.

Chr	Start	End	Gene	Func	ExonicFunc	AAChange	Other	1000g	dbSNP130	SIFT	PolyPhen2
3	25833079	25833079	OXSM	exonic	nonsynonymous SNV	NM_001145391:c.G568T:p.G190C	het			0	1
8	98827607	98827607	LAPTM4B	exonic	nonsynonymous SNV	NM_018407:c.T536A:p.M179K	het			0	1
10	124091986	124091986	BTBD16	exonic	nonsynonymous SNV	NM_144587:c.T1122A:p.D374E	het			0.06	0.95
10	124812607	124812607	ACADSB	exonic	nonsynonymous SNV	NM_001609:c.G1159A:p.E387K	het			0.1	0.06
14	73653589	73653589	PSEN1	exonic	nonsynonymous SNV	NM_000021:c.C509T:p.S170F	het		rs63750577	0	0.99
14	79181137	79181137	NRXN3	exonic	nonsynonymous SNV	NM_004796:c.G580A:p.A194T	het			0.5	0.89
16	4312335	4312335	TFAP4	exonic	nonsynonymous SNV	NM_003223:c.G344A:p.R115H	het			0	0.75

Individual columns describe localization of the mutation on chromosome (Chr), genomic coordinates of the mutation (Start and End), gene symbol (Gene), localization of the mutation within the gene (Func), type of the mutation (ExonicFunct), cDNA accession number: cDNA mutation: and its predicted effect on amino acid sequence (AAChange), heterozygous or homozygous status of the mutation (Other), absence/presence of the mutation in 1000 genomes (1000g), absence/accession number of the mutation in database of single nucleotide polymorphisms (dbSNP130), and predicted effect of the mutation on protein function (SIFT and PolyPhen2).

Liver International ISSN 1478-3223

CLINICAL STUDIES

Rotor-type hyperbilirubinaemia has no defect in the canalicular bilirubin export pump

Martin Hřebíček[1], Tomáš Jirásek[2], Hana Hartmannová[1], Lenka Nosková[1], Viktor Stránecký[1,3], Robert Ivánek[4,1], Stanislav Kmoch[1,3], Dita Cebecauerová[5], Libor Vítek[6], Miroslav Mikulecký[7], Iva Subhanová[6], Pavel Hozák[4] and Milan Jirsa[5,6]

1 Institute of Inherited Metabolic Diseases, Charles University 1st Faculty of Medicine, Prague, Czech Republic
2 Department of Pathology, Charles University 3rd Faculty of Medicine, Prague, Czech Republic
3 Center for Applied Genomics, Charles University 1st Faculty of Medicine, Prague, Czech Republic
4 Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
5 Institute for Clinical and Experimental Medicine, Prague, Czech Republic
6 Institute of Clinical Biochemistry and Laboratory Diagnostics, Charles University 1st Faculty of Medicine, Prague, Czech Republic
7 1st Medical Clinic, Teaching Hospital, Comenius University, Bratislava, Slovak Republic

Keywords
ABCC2 – Dubin-Johnson syndrome – hereditary jaundice – MRP2 – Rotor syndrome – UGT1A1

Correspondence
Milan Jirsa, MD, PhD, Laboratory of Experimental Hepatology, Institute for Clinical and Experimental Medicine, Vídeňská 1958/9, 140 21 Praha 4–Krč, Czech Republic.
Tel: +420 261 362 773
Fax: +420 241 721 666
e-mail: milan.jirsa@medicon.cz

Received 28 August 2006
accepted 22 December 2006

DOI:10.1111/j.1478-3231.2007.01446.x

Abstract

Background: The cause of Rotor syndrome (RS), a rare-familial conjugated hyperbilirubinaemia with normal liver histology, is unclear. We hypothesized that RS can be an allelic variant of Dubin–Johnson syndrome, caused by mutation in *ABCC2*, and investigated *ABCC2* (gene) and ABCC2 (protein) in two patients with RS. *Methods:* A 57-year-old male presented with a 5-year history of predominantly conjugated hyperbilirubinaemia (170 μmol/l). Urinary porphyrin excretion was increased; cholescintigraphy revealed no chromoexcretion. A 68-year-old male presented with lifelong conjugated hyperbilirubinaemia (85 μmol/l). Bromosulfophthalein elimination was typical for RS. Both patients had histologically normal liver, without pigment. ABCC2 expression was investigated by confocal fluorescence microscopy. *ABCC2* was sequenced from genomic DNA and cDNA, and exon deletions/duplications were sought by comparative genomic hybridization on a custom micro-array. *Results:* In both patients, ABCC2 was expressed unremarkably at the apical membrane of hepatocytes and no sequence alterations were found in 32 exons, adjacent intronic regions and the promoter region of *ABCC2*. *Conclusions:* Rotor-type hyperbilirubinaemia is not an allelic variant of *ABCC2* deficiency.

The bilirubin excretory pathway consists of two steps: conjugation of unconjugated bilirubin with glucuronic acid, catalyzed by uridine diphosphate-glucuronosyl transferase 1A1 (UGT1A1), and secretion of conjugated bilirubin into bile via ABCC2, the canalicular bilirubin export pump. Mutations in *ABCC2*, encoding ABCC2, are known to cause Dubin–Johnson syndrome (DJS, OMIM No. 237500) (1), a rare benign predominantly conjugated hyperbilirubinaemia with typical deposits of melanin-like pigment within hepatocyte lysosomes. Rotor syndrome (RS, OMIM No. 237450) represents another form of hereditary jaundice with predominantly conjugated hyperbilirubinaemia. Unlike most patients with DJS, patients with RS have no abnormal hepatic pigmentation. Total porphyrin excretion in urine is increased and the ratio of coproporphyrin isomers I:III is lower than in DJS (2). Unlike patients with DJS, patients with RS exhibit marked retention of bromosulfophthalein (BSP) after injection (3). Neither the liver nor the biliary tree is visualized by cholescintigraphy in RS (4, 5). The molecular basis of RS is unknown.

The definition of RS as a pathophysiological entity distinct from DJS is based on the differences in BSP clearance and total urinary coproporphyrin output. Other features of both disorders such as liver pigmentation and visualization of the gallbladder by cholescintigraphy are less specific (6), making the diagnosis of RS difficult to establish. We hypothesized that the phenotypic differences between RS and DJS do not exclude the possibility that RS and DJS are allelic variants. In our study, we investigated the potential role of *ABCC2* as a candidate gene responsible for Rotor-type hyperbilirubinaemia in two affected subjects.

Patients and methods

Case 1

A 57-year-old male had scleral icterus since birth. Generalized jaundice appeared for the first time at age 7 years and was diagnosed as posthepatitic. At age 18 years, he was judged unfit for compulsory military service because of jaundice. He was hospitalized, aged 52 years, owing to an acute occlusion of the central retinal artery. His total serum bilirubin concentration was 170 µmol/l, with a direct-reacting fraction of 120 µmol/l. Blood counts were within the normal ranges. Serum concentrations of aspartate aminotransferase (AST), alanine transaminase (ALT) and γ-glutamyl transferase activities and of α_1-antitrypsin, copper, ceruloplasmin and bile salts were normal. No serologic evidence of infection with hepatitis A, B and C viruses or with other hepatotropic viruses was found. Urinary coproporphyrin output ranged between 80 and 500 µg/24 h (normal < 200 µg/24 h); isomer I represented 57% of total coproporphyrin. Except for multiple gallstones, which were asymptomatic, no pathological changes of the liver and biliary tree were observed on ultrasonography and endoscopic retrograde cholangiopancreatography (ERCP). In contrast, ^{99m}Tc-HIDA cholescintigraphy revealed no uptake of the radionuclide by the liver, and the bile ducts and gallbladder were not visualized (Fig. 1). No abnormality was found on microscopy of a percutaneous needle liver biopsy specimen. Jaundice has persisted, with biochemical documentation, for the subsequent 5 years, with serum concentrations of total and direct bilirubin oscillating around 170 and 120 µmol/l respectively.

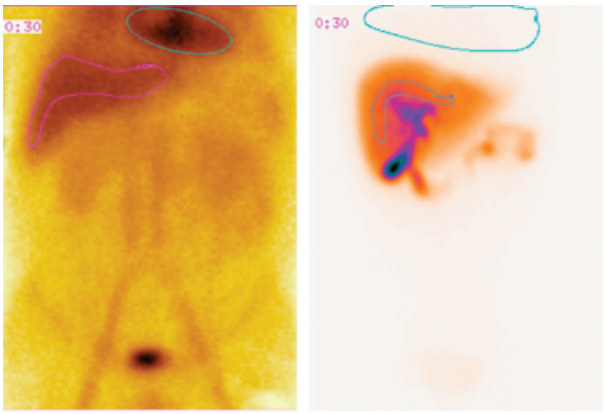


Fig. 1. Cholescintigraphy of patient 1 (left). In contrast to the normal situation (right), the radionuclide was retained in the circulation and no uptake of ^{99m}Tc-HIDA by the liver and visualization of bile ducts and gallbladder was observed.

The family history of the patient was of potential interest because his daughter and one of his four sisters had Gilbert syndrome.

Case 2

Our second patient (7), a 68-year-old male with no family history of liver disease, had scleral jaundice since childhood. Clinical and laboratory investigations on military conscription aged 19 years revealed predominantly direct hyperbilirubinaemia with no other clinical-biochemistry evidence of hepatobiliary injury and without serologic evidence of viral hepatitis. One year later, the patient was re-examined in the Central Military Hospital (Bratislava, Slovak Republic) for fatigue, lack of appetite, weight loss, abdominal pain, dark urine and jaundice. Aside from jaundice, chronic tonsillitis and skin rash, the physical examination was unremarkable. Total serum bilirubin concentrations ranged from 41 to 121 µmol/l; direct-reacting bilirubin represented 53–72%. Total urinary coproporphyrin output was repeatedly increased, but coproporphyrin isomers were not quantitatively differentiated. Other laboratory values were within the normal ranges. BSP retention 30 min after administration of 2 mg/kg was 77.5% (normal $< 10\%$). The gallbladder was repeatedly not visualized by oral cholecystography. Exploratory laparotomy revealed a normal appearance of the liver, bile ducts and gallbladder. Normal architecture without any pigment or signs of cholestasis was found on microscopy of a biopsy specimen of liver. RS was diagnosed (7). During the subsequent 48 years, the patient's liver was assessed twice at surgery as normal in contour and color (ureterotomy for ureterolithiasis, age 44 years; cholecystectomy for gallstones, age 56 years). Persistent jaundice with predominantly conjugated hyperbilirubinaemia (total serum bilirubin ranging between 42 and 170 µmol/l with seasonal fluctuations) has been repeatedly documented.

Histology

Liver from Patient 1 was fixed in formalin; liver from Patient 2 was fixed in formalin or in Carnoy's fixative. All specimens were embedded in paraffin. Sections were stained with haematoxylin/eosin and with periodic acid–Schiff technique.

Antibodies

Mouse anti-ABCC2 monoclonal antibody (clone M_2III-6) was purchased from Kamiya (Seattle, WA, USA). Rabbit polyclonal anti-human carcinoembryonic antigen antibody, which recognizes

Liver International (2007)

carcinoembryonic antigen-related adhesion molecule 1 (CEACAM1) on bile canaliculi of the human liver, was purchased from DAKO (Glostrup, Denmark) together with the EnVision Peroxidase Kit and LSAB+Kit. Fluorescein isothiocyanate-conjugated donkey anti-rabbit antibody and Cy5-conjugated goat anti-mouse antibody were obtained from Jackson (West Grove, PA, USA).

Confocal laser scanning microscopy (CLSM)

For immunohistochemical procedures, 5 μm sections cut from formalin-fixed (Patient 1) and Carnoy's-fixed (Patient 2), paraffin-embedded tissue samples were deparaffinized and treated with 2.7% hydrogen peroxide and 0.1% sodium azide. Double immunolabelling was performed as described (8). The slides were observed in a CLSM Leica TCS SP (Leica Microsystems, Wetzlar, Germany). Simultaneous excitation with an argon–krypton laser (wavelength 488 nm) for fluorescein isothiocyanate and a helium–neon laser (wavelength 633 nm) for Cy5 was used. Sections incubated without primary antibodies were used as negative controls. Sections of a liver-biopsy specimen assessed as exhibiting minimal changes, obtained from an adult patient, were used as positive controls. Sections of a liver-biopsy specimen from a patient with proven ABCC2 deficiency (8) served as a negative control.

Mutation analysis

Written informed consents were obtained from the patients and family members before genetic investigation and skin biopsy. *ABCC2* was analyzed by direct sequencing of polymerase chain reaction (PCR) products amplified from genomic DNA extracted from peripheral blood leukocytes. All exons and the 1500-bp-long promoter region were amplified by PCR using reported intronic oligonucleotide primers (8). All amplicons were gel-purified, extracted with QIA quick spin columns (Qiagen, Hilden, Germany) and sequenced on an automated fluorescent DNA sequencer (AlfExpress, Amersham-Pharmacia, Uppsala, Sweden).

Analysis of copy number changes caused by exon deletions or duplications was performed by comparative hybridization of the patient's and control male genomic DNAs. As probes, the micro-array contained PCR amplified products representing each *ABCC2* exon and 5′-aminomodified 40-mer oligonucleotides corresponding to specific regions of *ABCC2* exons. Oligonucleotide sequences were designed using Oligopicker software and purchased from Illumina (San Diego, CA, USA). The PCR products (100 ng/μl) and oligonucleotides (20 μM) in 3 × SSC were printed in triplicate on aminosilane-modified slides and immobilized by standard techniques that combined baking and UV cross-linking. The slides were pretreated by baking at 80 °C; after UV cross-linkage, they were then washed twice in 0.1% sodium dodecyl sulfate (SDS) for 2 min, twice in 0.2 × SSC for 2 min and four times in MilliQ water, followed by denaturation in boiling water for 2 min. Prehybridization was performed in prehybridization buffer (6 × SSC, 0.5% SDS, 1% bovine serum albumin). Genomic DNA was extracted by the phenol/chloroform method from peripheral blood leukocytes, fragmented with *Mbo*I restriction endonuclease (New England Biolabs, Ipswich, MA, USA) and labelled using Cy3-AP3-dUTP or Cy5-AP3-dUTP (Amersham Biosciences, Piscataway, NJ, USA). Patient samples, control samples and 5 μg of human Cot-1 (Invitrogen, Carlsbad, CA, USA) were combined and dissolved in hybridization buffer (50% formamide, 6 × SSC, 0.5% SDS, 5 × Denhardt's). Hybridizations were performed at 37 °C in an ArrayIt Hybridization Cassette chamber (TeleChem International, Sunnyvale, CA, USA). Patient and control samples were analyzed in a dye swap mode with two replicates of each mode. The hybridized slides were scanned using a GenePix 4200A scanner (Axon Instrument, Union City, CA, USA) with photomultiplier gains adjusted to obtain highest-intensity unsaturated images. Data analysis was performed in the R statistical environment (version 2.2.1) using the Linear models for Microarray data package LIMMA 2.2.0, which is part of the Bioconductor project (www.bioconductor.org) (9). Raw data were processed using lowess normalization and movingmin. Correlation between three duplicate spots per gene in each array was used to increase robustness. The linear model was fitted for each exon given from a series of arrays using lmFit function. The empirical Bayes method (10) was used to rank differential expression signal-fold changes of individual gene exons using eBayes function.

ABCC2 mRNA was isolated from cultured skin fibroblasts and subjected to reverse transcription. Overlapping ≈ 800 bp fragments were amplified by nested PCR from cDNA (see Table 1 for primer sequences), gel purified and sequenced.

Mutations in *UGT1A1* known to be associated with Gilbert syndrome in Caucasians were detected as described (8).

Results

Histology

Routine light microscopy of haematoxylin/eosin-stained sections of liver from both patients found no

Table 1. Sequences of PCR primer pairs used to amplify the overlapping fragments of *ABCC2* from cDNA

Pair number	Forward primer	Reverse primer
1	5′-TAGAAGAGTCTTCGTTCCAGACGCAG-3′	5′-AGTGCCCGCCTGGCTTTC-3′
2	5′-TTCTGAAAGGCTACAAGCGTCCTC-3′	5′-ATTGGGATTACAAGCACCATCACC-3′
3	5′-AACTTCATGCACATGCTGTGGTC-3′	5′-CCTTTATGGTGCCATTCTGAATCC-3′
4	5′-AATCCTCCTTGATATCAGCCATGC-3′	5′-TCAAGGAGTTTCTCAGGGACTTCAG-3′
5	5′-ACAGCTTTCGTCGAACACTTAGCC-3′	5′-GGATAACTGGCAAACCTGATACGG-3′
6	5′-ACCATCATCGTCATTCCTCTTGG-3′	5′-TGTTGAAAGGGTCGAGATTCATCC-3′
7	5′-ATATTGCTTCCATTGGGCTCCAC-3′	5′-TGGGTAGTAGGTTCATGGGTGTTC-3′

abnormalities (Fig. 2). Pigment accumulation, the characteristic histomorphological feature of DJS, was not detected. No autofluorescence was observed on fluorescence microscopy.

ABCC2 protein expression

Immunohistochemical staining showed linear marking for CEACAM1 in the canalicular membrane of hepatocytes in both our patients, as well as in the positive control. CLSM with double immunofluorescence staining confirmed the localization of ABCC2 in the canalicular membrane of hepatocytes (Fig. 3). In contrast, no ABCC2 immunostaining was observed in sections of a liver-biopsy specimen from a patient with proven ABCC2 deficiency (negative control).

DNA analysis

In Patient 1, sequence analysis of *ABCC2* disclosed heterozygosity for the known synonymous polymorphism 3972C/T (I1324I, GenBank dbSNP rs3740066) in exon 28 (11) and for the polymorphism −1023G/A (GenBank dbSNP rs7910642). The sequence of cDNA corresponded with the genomic sequence. Heterozygosity for the 3972C/T polymorphism indicated that both alleles of the gene were expressed at the RNA level in cultured skin fibroblasts. A heterozygous polymorphism $A(TA)_7TAA$ of the TATAA-box and a heterozygous polymorphism −3263T/G in the phenobarbital-responsive enhancer module were detected in the promoter region of the *UGT1A1* gene. Family analysis revealed homozygosity for both variants, $A(TA)_7TAA$ and −3263G, in the proband's daughter, indicating that both heterozygous polymorphisms are located on the same chromosome in the proband, Patient 1. Finally, Patient 1 was homozygous for the wild-type 211G allele.

Sequence analysis of *ABCC2* in Patient 2 revealed homozygosity for two known polymorphisms: 3972C/T (I1324I) in the protein-coding region and −24C/T (GenBank dbSNP rs717620) in the 5′-untranslated region (11). As in Patient 1, the sequence of cDNA corresponded with the genomic sequence. In contrast to Patient 1, the expression of individual alleles could not be addressed because no heterozygous sequence variation was present in the transcribed mRNA. The patient was found to be homozygous for the wild-type

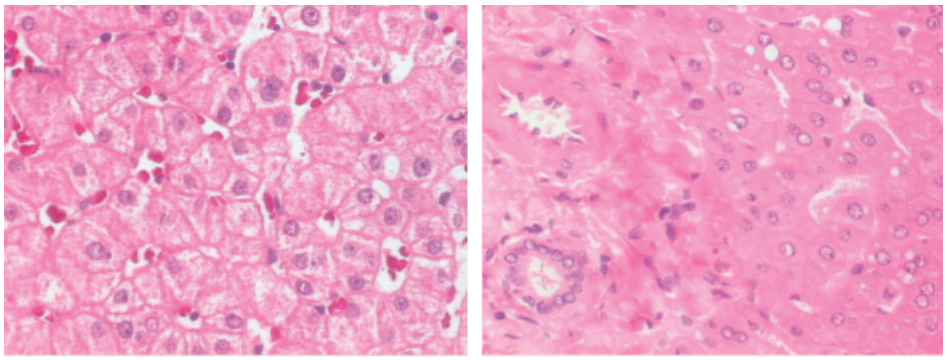


Fig. 2. Liver histology of Patient 1 (centrizonal area, needle biopsy taken at the age of 52 years) and Patient 2 (periportal area, liver excision performed in 1958 at the age of 20 years). Normal histology and cytology with no liver pigment was found in both specimens. The absence of lipopigment in hepatocytes in Patient 1 at the age of 52 is uncommon. Haematoxylin&eosin, original magnification × 400.

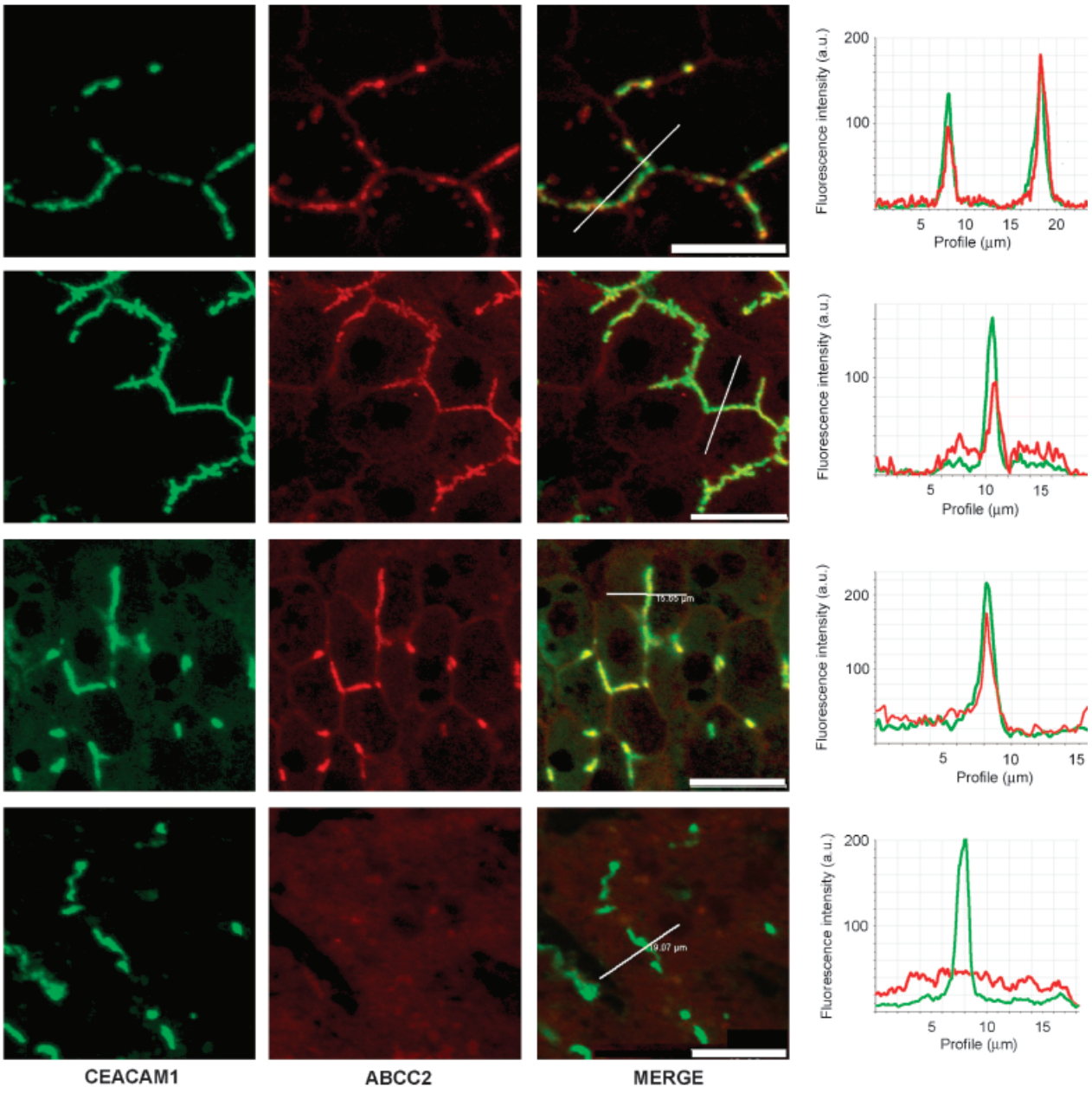


Fig. 3. Confocal laser scanning micrographs with double immunofluorescence staining for CEACAM1 (left column) and ABCC2 (middle column). In the control liver as well as in both samples from the patients with RS, ABCC2 and CEACAM1 colocalize with the canalicular membrane of hepatocytes (right column – yellow color, and graphs). In the liver of a patient with proven ABCC2 deficiency (bottom slides), ABCC2 protein is absent both at canalicular membranes and in the cytoplasm of hepatocytes, whereas expression of CEACAM1 is not affected. Bar = 20 μm, original magnification × 400.

alleles A(TA)$_6$TAA and −3263T of the *UGT1A1* gene promoter and for the wild-type 211G allele in the first exon of *UGT1A1*.

Comparative genomic hybridization to a custom micro-array revealed no significant copy number changes in any of 32 exons of *ABCC2* in either Patient 1 or 2.

Discussion

We attempted to investigate the role of *ABCC2* in two subjects with Rotor-type hyperbilirubinaemia; to our knowledge, ours is the first such attempt. Normal expression and localization of ABCC2 on the canalicular membrane of hepatocytes ruled out the most

common findings seen in DJS, when mutations in both alleles of *ABCC2* abolish ABCC2 expression. Normal localization of ABCC2 excluded potential defects in *RDX*, encoding radixin, a cytoskeletal protein essential in anchoring ABCC2 to the canalicular membrane (12). The immunohistologic findings, however, did not exclude two possibilities. Firstly, a mutation-impairing function but not expression of ABCC2 might be present in at least one allele of *ABCC2*. Secondly, a mutation in the regulatory region of *ABCC2* might decrease but not abolish ABCC2 expression. Both possibilities were checked by sequence analysis of the coding and promoter regions of *ABCC2*. In addition to conventional mutational screening, a thorough search for less common types of mutations – exon deletions and duplications – was performed; none was found. Finally, a contribution of Gilbert syndrome to the unconjugated fraction of elevated serum bilirubin was excluded by *UGT1A1* genotyping.

The rationale for investigation of *ABCC2* and ABCC2 in RS, which is considered a disorder of hepatic bilirubin storage (13), can be questioned. Significant reduction in the actual or apparent capacity of hepatocytes to store unconjugated bilirubin, unconjugated BSP and indocyanine green (which does not undergo conjugation) can result from decreased concentrations of 'ligandin' in cytosol of hepatocytes; from decreased hepatocellular uptake of unconjugated bilirubin, BSP and indocyanine green; and from occupation of binding sites of intracellular 'ligandin' – proteins belonging mainly to the α-class of the glutathione-*S*-transferase family (GST-α).

The hypothesis that RS is a primary disorder of hepatic bilirubin storage is supported by the kinetics of selected anionic dyes and by the immunohistological findings of Abei M et al. (14). To the best of our knowledge, these findings have not been confirmed by quantitative analysis of GST-α isoenzymes or by mutational analysis of the corresponding *GSTA1-5* genes. A compensatory upregulation of GST-α family proteins in the liver has been detected in *Gsta4* null mice; no jaundice has been observed in these animals (15). If the same is true for single gene defects in *GSTA1* or *GSTA2* (less likely in *GSTA3* and *GSTA5* because their expression is low), such defects cannot be expected to reduce substantially the total concentration of GST-α family proteins in the liver.

Defective uptake of unconjugated bilirubin and other organic anions is not compatible with predominantly conjugated hyperbilirubinaemia in RS.

Thirdly, reduction of hepatic bilirubin storage capacity can be caused by occupation of the binding sites of 'ligandin'. Retention of 'ligandin'-binding substrates in cytosol may arise from changes in affinity of mutated ABCC2 for a subgroup of ABCC2 substrates. ABCC2 can be mutated in such a way that retained nonbilirubin substrates do not upregulate expression of ABCC3 in compensation for impaired function of ABCC2. Mutations affecting the affinity of ABCC2 for various substrates have been documented by Ito et al. (16–18). Our presented results exclude this possibility in these patients.

Alternatively, 'ligandin'-binding substrates may be retained in cytosol of hepatocytes owing to deficient transport of polar conjugates from the endoplasmic reticulum (ER) to the cytoplasm. Conjugation takes place on the luminal aspect of ER (19, 20). Transport of highly polar conjugates is likely mediated by an ATP-independent permease specific for conjugates of bilirubin (and possibly other substrates) with glucuronic acid (19, 21). Production of bilirubin glucuronides in the ER lumen and activity of the canalicular bilirubin export pump ABCC2, which keeps cytoplasmic concentrations of bilirubin glucuronides low, constitute the driving force for translocation of conjugated bilirubin across the ER membrane. In the case of impaired export, conjugated bilirubin should be retained in the ER lumen and secreted into the plasma via exocytosis. Increased intra-ER concentrations of conjugated bilirubin (and perhaps other glucuronides) may decrease the rates of conjugation of the corresponding substrates. Unconjugated substrates retained in cytosol (owing to deficient export of conjugates from the ER to the cytoplasm) bind to GST-α and decrease hepatic storage capacity for unconjugated bilirubin, unconjugated BSP and indocyanine green as well as the transport maximum for BSP.

In conclusion, we have shown that RS is not an allelic variant of DJS. The discovery of the molecular background of RS would be helpful in differentiating among forms of hereditary conjugated jaundice.

Acknowledgements

This study was supported by grant MZO 00023001 from the Institute for Clinical and Experimental Medicine. The authors from Institute of Inherited Metabolic Diseases were supported by grant 1A/8239-3 from the Grant Agency of the Ministry of Health of the Czech Republic; their institutional support was provided by grant MSM 0021620806 from the Ministry of Education of the Czech Republic. The authors thank Lucie Budisova for technical assistance, MVDr. Zorka Novotna for the culture of primary fibroblasts and A.S. Knisely, M.D., and Prof. MUDr. Milan Elleder, Dr.Sc., for their comments and help with the preparation of the manuscript.

References

1. Paulusma CC, Kool M, Bosma PJ, *et al*. A mutation in the human canalicular multispecific organic anion transporter gene causes the Dubin–Johnson syndrome. *Hepatology* 1997; **25**: 1539–42.
2. Wolkoff AW, Wolpert E, Pascasio FN, Arias IM. Rotor's syndrome. A distinct inheritable pathophysiologic entity. *Am J Med* 1976; **60**: 173–9.
3. Wolpert E, Pascasio FM, Wolkoff AW, Arias IM. Abnormal sulfobromophthalein metabolism in Rotor's syndrome and obligate heterozygotes. *N Engl J Med* 1977; **296**: 1099–101.
4. Bar-Meir S, Baron J, Seligson U, Gottesfeld F, Levy R, Gilat T. 99mTc-HIDA cholescintigraphy in Dubin–Johnson and Rotor syndromes. *Radiology* 1982; **142**: 743–6.
5. Fretzayas AM, Garoufi AI, Moutsouris CX, Karpathios TE. Cholescintigraphy in the diagnosis of Rotor syndrome. *J Nucl Med* 1994; **35**: 1048–50.
6. Fretzayas AM, Stavrinadis CS, Koukoutsakis PM, Sinaniotis CA. Diagnostic approach of Rotor syndrome with cholescintigraphy. *Clin Nucl Med* 1997; **22**: 635–6.
7. Mikulecky M. Das atypische Dubin–Johnsonsche syndrom. *Gastroenterologia* 1960; **94**: 201–26.
8. Cebecauerova D, Jirasek T, Budisova L, *et al*. Dual hereditary jaundice: simultaneous occurrence of mutations causing Gilbert's and Dubin-Johnson syndrome. *Gastroenterology* 2005; **129**: 315–20.
9. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, eds. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer, 2005; 397–420.
10. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3.
11. Itoda M, Saito Y, Soyama A, *et al*. Polymorphisms in the ABCC2 (cMOAT/MRP2) gene found in 72 established cell lines derived from Japanese individuals: an association between single nucleotide polymorphisms in the 5′-untranslated region and exon 28. *Drug Metab Dispos* 2002; **30**: 363–4.
12. Kikuchi S, Hata M, Fukumoto K, *et al*. Radixin deficiency causes conjugated hyperbilirubinaemia with loss of Mrp2 from bile canalicular membranes. *Nat Genet* 2002; **31**: 320–5.
13. Chowdhury NR, Arias IM, Wolkoff AW, Chowdhury JR. Disorders of bilirubin metabolism. In: Arias IM, ed. *The Liver. Biology and Pathobiology*. Philadelphia: Lippincott Williams & Wilkins, 2001; 290–309.
14. Abei M, Matsuzaki Y, Tanaka N, Osuga T, Adachi Y. Defective hepatic glutathione S-transferase in Rotor's syndrome. *Am J Gastroenterol* 1995; **90**: 681–2.
15. Engle MR, Singh SP, Czernik PJ, *et al*. Physiological role of mGSTA4-4, a glutathione S-transferase metabolizing 4-hydroxynonenal: generation and analysis of mGsta4 null mouse. *Toxicol Appl Pharmacol* 2004; **194**: 296–308.
16. Ito K, Suzuki H, Sugiyama Y. Charged amino acids in the transmembrane domains are involved in the determination of the substrate specificity of rat Mrp2. *Mol Pharmacol* 2001; **59**: 1077–85.
17. Ito K, Suzuki H, Sugiyama Y. Single amino acid substitution of rat MRP2 results in acquired transport activity for taurocholate. *Am J Physiol Gastrointest Liver Physiol* 2001; **281**: G1034–43.
18. Ito K, Oleschuk CJ, Westlake C, Vasa MZ, Deeley RG, Cole SP. Mutation of Trp1254 in the multispecific organic anion transporter, multidrug resistance protein 2 (MRP2) (ABCC2), alters substrate specificity and results in loss of methotrexate transport activity. *J Biol Chem* 2001; **276**: 38108–14.
19. Jansen PL, Mulder GJ, Burchell B, Bock KW. New developments in glucuronidation research: report of a workshop on "glucuronidation, its role in health and disease". *Hepatology* 1992; **15**: 532–44.
20. Radominska-Pandya A, Czernik PJ, Little JM, Battaglia E, Mackenzie PI. Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev* 1999; **31**: 817–99.
21. Csala M, Staines AG, Banhegyi G, Mandl J, Coughtrie MW, Burchell B. Evidence for multiple glucuronide transporters in rat liver microsomes. *Biochem Pharmacol* 2004; **68**: 1353–62.

Research article

Complete OATP1B1 and OATP1B3 deficiency causes human Rotor syndrome by interrupting conjugated bilirubin reuptake into the liver

Evita van de Steeg,[1] Viktor Stránecký,[2,3] Hana Hartmannová,[2,3] Lenka Nosková,[3] Martin Hřebíček,[3] Els Wagenaar,[1] Anita van Esch,[1] Dirk R. de Waart,[4] Ronald P.J. Oude Elferink,[4] Kathryn E. Kenworthy,[5] Eva Sticová,[6] Mohammad al-Edreesi,[7] A.S. Knisely,[8] Stanislav Kmoch,[2,3] Milan Jirsa,[6] and Alfred H. Schinkel[1]

[1]Division of Molecular Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands. [2]Center for Applied Genomics and [3]Institute of Inherited Metabolic Diseases, Charles University of Prague, First Faculty of Medicine, Prague, Czech Republic. [4]Tytgat Institute for Liver and Intestinal Research, Academic Medical Center, Amsterdam, The Netherlands. [5]Department of Drug Metabolism and Pharmacokinetics, GlaxoSmithKline, Ware, United Kingdom. [6]Institute for Clinical and Experimental Medicine, Prague, Czech Republic. [7]Department of Pediatrics, Saudi Aramco Dhahran Health Centre, Dhahran, Saudi Arabia. [8]Institute of Liver Studies, King's College Hospital, London, United Kingdom.

Bilirubin, a breakdown product of heme, is normally glucuronidated and excreted by the liver into bile. Failure of this system can lead to a buildup of conjugated bilirubin in the blood, resulting in jaundice. The mechanistic basis of bilirubin excretion and hyperbilirubinemia syndromes is largely understood, but that of Rotor syndrome, an autosomal recessive disorder characterized by conjugated hyperbilirubinemia, coproporphyrinuria, and near-absent hepatic uptake of anionic diagnostics, has remained enigmatic. Here, we analyzed 8 Rotor-syndrome families and found that Rotor syndrome was linked to mutations predicted to cause complete and simultaneous deficiencies of the organic anion transporting polypeptides OATP1B1 and OATP1B3. These important detoxification-limiting proteins mediate uptake and clearance of countless drugs and drug conjugates across the sinusoidal hepatocyte membrane. OATP1B1 polymorphisms have previously been linked to drug hypersensitivities. Using mice deficient in Oatp1a/1b and in the multispecific sinusoidal export pump Abcc3, we found that Abcc3 secretes bilirubin conjugates into the blood, while Oatp1a/1b transporters mediate their hepatic reuptake. Transgenic expression of human OATP1B1 or OATP1B3 restored the function of this detoxification-enhancing liver-blood shuttle in Oatp1a/1b-deficient mice. Within liver lobules, this shuttle may allow flexible transfer of bilirubin conjugates (and probably also drug conjugates) formed in upstream hepatocytes to downstream hepatocytes, thereby preventing local saturation of further detoxification processes and hepatocyte toxic injury. Thus, disruption of hepatic reuptake of bilirubin glucuronide due to coexisting OATP1B1 and OATP1B3 deficiencies explains Rotor-type hyperbilirubinemia. Moreover, OATP1B1 and OATP1B3 null mutations may confer substantial drug toxicity risks.

Introduction

Rotor syndrome (RS; OMIM %237450) is a rare, benign hereditary conjugated hyperbilirubinemia, also featuring coproporphyrinuria and strongly reduced liver uptake of many diagnostic compounds, including cholescintigraphic tracers (1–6). RS is an autosomal recessive disorder that clinically resembles another conjugated hyperbilirubinemia, the Dubin-Johnson syndrome (DJS; OMIM #237500) (7, 8). In both RS and DJS, mild jaundice begins shortly after birth or in childhood. There are no signs of hemolysis, and routine hematologic and clinical-biochemistry test results are normal, aside from the primarily conjugated hyperbilirubinemia. RS is, however, distinguishable from DJS by several criteria (1, 2, 9, 10): (a) it lacks the hepatocyte pigment deposits typical of DJS; (b) in RS, but not DJS, there is delayed plasma clearance of unconjugated bromsulphthalein (BSP), an anionic diagnostic dye, and no conjugated BSP appears in plasma (4); (c) the liver in RS is scarcely visualized on ^{99m}Tc-N[2,6-dimethylphenyl-carbamoylmethyl] iminodiacetic acid (^{99m}Tc-HIDA) cholescintigraphy, with slow liver uptake, persistent visualization of the cardiac blood pool, and prominent kidney excretion (5); and (d) total urinary excretion of coproporphyrins is greatly increased in RS, with coproporphyrin I being the predominant isomer (11).

DJS is caused by mutations affecting ABCC2/MRP2, a canalicular bilirubin glucuronide and xenobiotic export pump, thus disrupting bilirubin glucuronide excretion into bile (7, 8). Excretion of bilirubin glucuronides is then redirected into plasma by the action of ABCC3/MRP3, a homolog of ABCC2 that is present in the sinusoidal membrane and is upregulated in DJS (12, 13). The molecular mechanism of DJS is in line with the generally accepted paradigm of normal hepatic bilirubin excretion, according to which a unidirectional elimination pathway is postulated: first, uptake of unconjugated bilirubin (UCB) from blood into hepatocytes; subse-

Authorship note: Evita van de Steeg and Viktor Stránecký, and Milan Jirsa and Alfred H. Schinkel contributed equally to this work.

Conflict of interest: The research group of Alfred H. Schinkel receives revenues from commercial distribution of some of the mouse strains used in this study.

Citation for this article: *J Clin Invest* doi:10.1172/JCI59526.

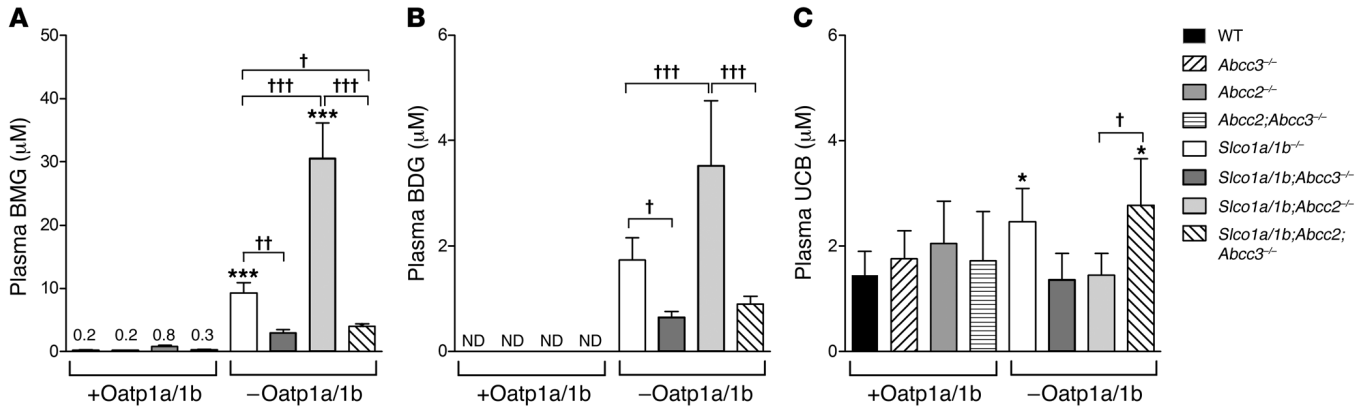


Figure 1
Increased plasma bilirubin glucuronide in *Slco1a/1b*$^{-/-}$ mice is in part dependent on Abcc3. (**A**) BMG, (**B**) BDG, and (**C**) UCB levels in plasma of male wild-type, *Abcc3*$^{-/-}$, *Abcc2*$^{-/-}$, *Abcc2*$^{-/-}$*Abcc3*$^{-/-}$, *Slco1a/1b*$^{-/-}$, *Slco1a/1b;Abcc3*$^{-/-}$, *Slco1a/1b;Abcc2*$^{-/-}$, and *Slco1a/1b;Abcc2;Abcc3*$^{-/-}$ mice (n = 4–7). +Oatp1a/1b denotes strains possessing Oatp1a/1b proteins, and –Oatp1a/1b denotes strains lacking Oatp1a/1b proteins. Data are mean ± SD. $^{*}P < 0.05$, $^{***}P < 0.001$ compared with wild-type mice. Bracketed comparisons: $^{\dagger}P < 0.05$, $^{\dagger\dagger}P < 0.01$, $^{\dagger\dagger\dagger}P < 0.001$. ND, not detectable; detection limit was 0.1 μM.

quent glucuronidation; and finally, secretion of bilirubin glucuronide into bile via ABCC2. Individuals with RS, however, lack *ABCC2* mutations (14), and the mechanistic basis of RS is unknown.

Organic anion transporting polypeptides (OATPs, genes: *SLCO*s) contain 12 plasma membrane–spanning domains and mediate sodium-independent cellular uptake of highly diverse compounds, including bilirubin glucuronide, bile acids, steroid and thyroid hormones, and numerous drugs, toxins, and their conjugates (15, 16). Human OATP1B1 and OATP1B3 localize to the sinusoidal membrane of hepatocytes and mediate the liver uptake of, among other compounds, many drugs (15–19). Various SNPs in *SLCO1B1* cause reduced transport activity and altered plasma and tissue levels of statins, methotrexate, and irinotecan in patients, potentially resulting in life-threatening toxicities (20–24).

In a *Slco1a/1b*$^{-/-}$ mouse model recently generated by our group, the importance of Oatp1a/1b proteins in hepatic uptake and clearance of drugs was confirmed, but the mice also displayed marked conjugated hyperbilirubinemia (25). We therefore hypothesized that sinusoidal Oatps in the normal, healthy mouse liver function in tandem with the sinusoidal efflux transporter Abcc3 to mediate substantial hepatic secretion and reuptake of bilirubin glucuronides and other conjugated compounds (25).

Here we describe how a combination of functional studies in mice to address this hypothesis and independent genetic studies in humans has resulted in elucidation of the genetic and mechanistic basis of Rotor syndrome.

Results

To test our hypothesis regarding the involvement of Abcc3 in the sinusoidal cycling of bilirubin glucuronides, and to assess a possible interplay with Abcc2, we generated *Slco1a/1b*$^{-/-}$*Abcc3*$^{-/-}$ (*Slco1a/1b;Abcc3*$^{-/-}$), *Slco1a/1b*$^{-/-}$*Abcc2*$^{-/-}$ (*Slco1a/1b;Abcc2*$^{-/-}$), and *Slco1a/1b*$^{-/-}$*Abcc2*$^{-/-}$*Abcc3*$^{-/-}$ (*Slco1a/1b;Abcc2;Abcc3*$^{-/-}$) mice by crossbreeding of existing strains. All strains were fertile, with normal life spans and body weights. As previously found for *Abcc2*$^{-/-}$ and *Abcc2*$^{-/-}$*Abcc3*$^{-/-}$ mice (26, 27), liver weights of *Slco1a/1b;Abcc2*$^{-/-}$ and *Slco1a/1b;Abcc2;Abcc3*$^{-/-}$ mice were significantly increased (~30% and ~50%, respectively) compared with wild-type mice (data not shown). Quantitative RT-PCR analysis of functionally relevant uptake and efflux transporters in liver, kidney, and intestine of the single and combination knockout strains revealed only some modest expression changes (Supplemental Table 1 and Supplemental Results; supplemental material available online with this article; doi:10.1172/JCI59526DS1). Hepatic UDP-glucuronosyltransferase 1a1 (*Ugt1a1*) expression was not significantly altered in any of the strains.

Importantly, the markedly increased plasma bilirubin monoglucuronide (BMG) and bilirubin diglucuronide (BDG) levels observed in *Slco1a/1b*$^{-/-}$ mice were substantially reduced in *Slco1a/1b;Abcc3*$^{-/-}$ mice, demonstrating that Abcc3 is necessary for most of this increase (Figure 1, A and B). Plasma BMG levels in *Slco1a/1b;Abcc2*$^{-/-}$ mice, even further increased owing to strongly reduced biliary BMG excretion (Figure 2, A and B), were similarly decreased in *Slco1a/1b;Abcc2;Abcc3*$^{-/-}$ mice (Figure 1, A and B). Thus, Abcc3 secretes bilirubin glucuronides back into blood, and Oatp1a/1b proteins mediate their efficient hepatic reuptake, thereby together establishing a sinusoidal liver-blood shuttling loop. The incomplete reversion of plasma bilirubin glucuronide levels in the Oatp1a/1b/Abcc3-deficient strains (Figure 1, A and B) suggests that additional sinusoidal exporter(s), e.g., Abcc4 (28), can partly take over the sinusoidal bilirubin glucuronide extrusion function of Abcc3.

The biliary output of bilirubin glucuronides in the single and combination knockout mice showed that, as long as Oatp1a/1b was functional, Abcc3 improved the efficiency of biliary bilirubin glucuronide excretion, even though it transports its substrates initially from liver to blood, not bile (Figure 2, A and B, strains +Oatp1a/1b). This suggests that, within liver lobules, the bilirubin glucuronide extruded by Abcc3 in upstream hepatocytes is efficiently taken up in downstream hepatocytes via Oatp1a/1b and then excreted into bile. The resulting relief of possible saturation of (or competition for) biliary excretion in the upstream hepatocytes may explain why the overall biliary excretion is enhanced by this transfer to downstream hepatocytes. However, when Oatp1a/1b was absent, Abcc3 instead decreased biliary bilirubin glucuronide excretion (Figure 2, strains –Oatp1a/1b) and

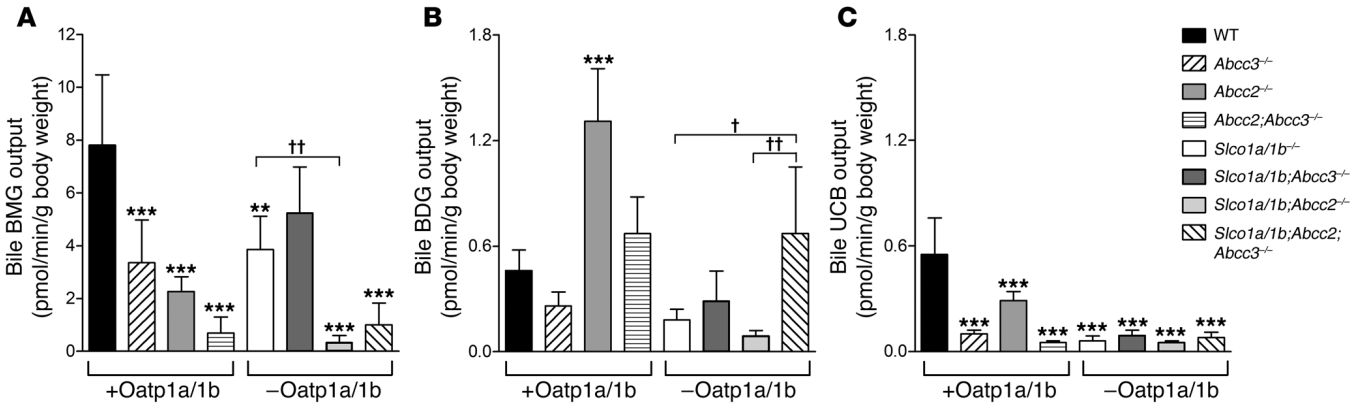


Figure 2

In the presence of Oatp1a/1b, but not in its absence, Abcc3 enhances biliary excretion of bilirubin glucuronides. (**A**) BMG, (**B**) BDG, and (**C**) UCB output in bile of male wild-type, *Abcc3*$^{-/-}$, *Abcc2*$^{-/-}$, *Abcc2*$^{-/-}$*Abcc3*$^{-/-}$, *Slco1a/1b*$^{-/-}$, *Slco1a/1b;Abcc3*$^{-/-}$, *Slco1a/1b;Abcc2*$^{-/-}$, and *Slco1a/1b;Abcc2;Abcc3*$^{-/-}$ mice. Bile collected during the first 15 minutes after gall bladder cannulation was analyzed. +Oatp1a/1b denotes strains possessing Oatp1a/1b proteins, and –Oatp1a/1b denotes strains lacking Oatp1a/1b proteins. Data are shown as mean ± SD (n = 4–7). $^{**}P < 0.01$, $^{***}P < 0.001$ compared with wild-type mice. Bracketed comparisons: $^{\dagger}P < 0.05$, $^{\dagger\dagger}P < 0.01$.

redirected excretion toward urine via the increased plasma bilirubin glucuronide levels (Supplemental Figure 1). Obviously, in the absence of Oatp1a/1b-mediated hepatic reuptake, Abcc3 activity can only decrease hepatocyte levels of bilirubin glucuronide in upstream and downstream hepatocytes alike, and will therefore reduce overall biliary excretion. Thus, both components of the Abcc3 and Oatp1a/1b shuttling loop are necessary to improve hepatobiliary excretion efficiency.

Human hepatocytes express only two OATP1A/1B proteins at the sinusoidal membrane, OATP1B1 and OATP1B3 (15). To test whether these could mediate the identified Oatp1a/1b functions, and in a liver-specific manner, we generated *Slco1a/1b*$^{-/-}$ mice with liver-specific expression of either human OATP1B1 or OATP1B3. Liver-specific expression was obtained using an apoE promoter (29). These strains were viable and fertile, and displayed normal life spans and body weights. Liver levels of transgenic OATP1B1 and OATP1B3 proteins were similar to those seen in pooled human liver samples (data not shown). Both of the transgenic rescue strains displayed a virtually complete reversal of the increases in plasma and urine levels of BMG and BDG seen in *Slco1a/1b*$^{-/-}$ mice (Figure 3, A and B, and Supplemental Figure 2). This indicates that both human OATP1B1 and OATP1B3 effectively reabsorb bilirubin glucuronides from plasma into the liver, in line with their demonstrated in vitro role in bilirubin glucuronide uptake (30). The modest (~1.8-fold) increase in plasma UCB in *Slco1a/1b*$^{-/-}$ mice was also reduced in the rescue strains (Figure 3C), suggesting an ancillary role of these proteins in hepatic UCB uptake.

These findings collectively raised the question as to whether humans with a severe deficiency in OATP1B1 and OATP1B3, possibly leading to a conjugated hyperbilirubinemia, might exist. A literature search suggested RS as a candidate inborn metabolic disorder. A search for RS subjects by part of the present group led to collaboration with another team already working on mapping of the RS gene(s).

In an unbiased approach, scanning the whole genome, we mapped the genomic candidate intervals for RS in 11 RS index subjects from 8 different families, 4 Central European (CE1–CE4), 3 Saudi-Arabian (A1–A3), and 1 Filipino (P1) (Figure 4A and Supplemental Table 2). Homozygosity mapping identified a single genomic region on chromosome 12 for which 8 tested index subjects and no healthy siblings or parents were homozygous (Figure 4B), suggesting inheritance of both alleles from a common ancestor. Three distinct homozygous haplotypes (R1–R3) segregated with RS: R1 in families CE1, CE2, and CE4; R2 in families CE3, A1, A2, and A3; and R3 in family P1 (Figure 4B; for genotyping details, see Methods). Intersection of these haplotypes defined a candidate genomic region spanning the *SLCO1C1*, *SLCO1B3*, *SLCO1B1*, *SLCO1A2*, and *IAPP* genes (Figure 4B). A parallel genome-wide copy number analysis detected a homozygous deletion within the *SLCO1B3* gene in the R1 haplotype and a homozygous approximately 405-kb deletion encompassing *SLCO1B3* and *SLCO1B1* and the *LST-3TM12* pseudogene in the R2 haplotype (Figure 4B and Supplemental Figure 3).

Sequence analysis revealed predictably pathogenic mutations affecting both *SLCO1B3* and *SLCO1B1* in each of the haplotypes (Figure 4, B–D, Table 1, Supplemental Figure 3, and Supplemental Table 3). In the R1 haplotype, a 7.2-kb deletion removes exon 12 of *SLCO1B3*, encoding amino acids 500–560 of OATP1B3 (702 aa long) and introduces a frameshift and premature stop codon, thus removing the C-terminal 3 transmembrane domains. Furthermore, a nonsense mutation in exon 13, c.1738C→T, introduces a premature stop codon (p.R580X) in R1-linked OATP1B1 (691 aa long), removing the C-terminal one-and-a-half transmembrane domains. The 405-kb R2 deletion encompasses exons 3–15 of *SLCO1B3* (sparing only a small N-terminal region) and the whole of *SLCO1B1*, but not *SLCO1A2*. The R3 haplotype harbors a splice donor site mutation, c.1747+1G→A, in intron 13 of *SLCO1B3*. If *SLCO1B3* is still yielding functional mRNA, this would truncate OATP1B3 after amino acid 582, deleting the C-terminal one-and-a-half transmembrane domains. A nonsense mutation, c.757C→T, in exon 8 of R3-linked *SLCO1B1* introduces a premature stop (p.R253X), truncating OATP1B1 before the C-terminal 7 transmembrane domains. All of these mutations would severely disrupt or annihilate proper protein expression and function. Moreover, they all showed consistent autosomal recessive segregation with the RS phenotype in the investigated families (Table 1). No

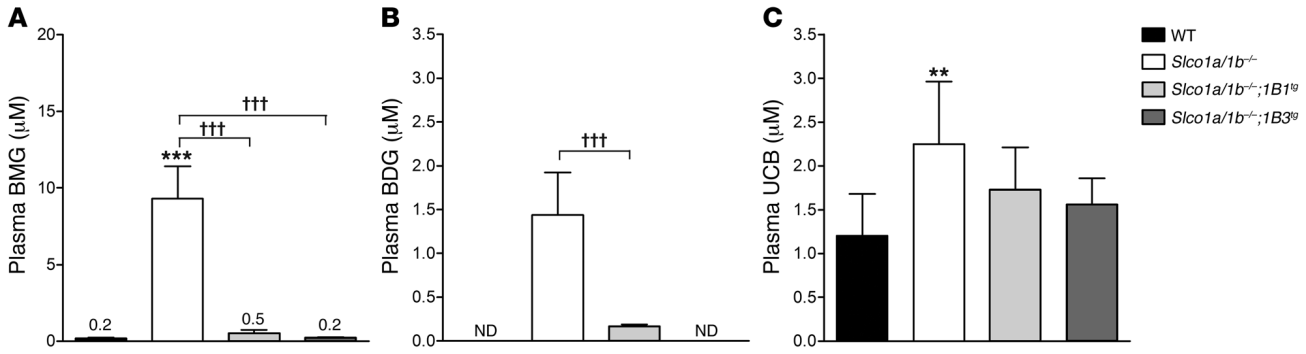


Figure 3
Increased plasma bilirubin glucuronide in *Slco1a/1b*$^{-/-}$ mice is reversed by human OATP1B1 and OATP1B3. (**A**) BMG, (**B**) BDG, and (**C**) UCB levels in plasma of male wild-type and *Slco1a/1b*$^{-/-}$ mice, and of the derived OATP1B1- and OATP1B3-transgenic strains (*Slco1a/1b*$^{-/-}$*;1B1*^{tg} and *Slco1a/1b*$^{-/-}$*1B3*^{tg}, respectively) (n = 5–8). Data are mean ± SD. $^{**}P < 0.01$, $^{***}P < 0.001$ compared with wild-type mice. Bracketed comparisons: $^{†††}P < 0.001$. Detection limit was 0.1 μM.

SLCO1A2 sequence variation was found in probands representing the 3 haplotypes, rendering involvement of OATP1A2 in RS unlikely. The severity of the identified mutations affecting *SLCO1B3* and *SLCO1B1* and their strict cosegregation with the RS phenotype indicate that RS is caused by co-inherited complete functional deficiencies in both OATP1B3 and OATP1B1.

The severity of the mutations was independently supported by immunohistochemical studies of the sparse RS liver biopsy material available. Given their sparseness, immunostaining of these liver biopsies was performed using one antibody recognizing the N terminus of both OATP1B1 and OATP1B3 (31). This revealed absence of detectable staining in probands representing each haplotype (Figure 5). In controls, basolateral membranes of centrilobular hepatocytes stained crisply, as previously reported (31). Thus, the *SLCO1B1* and *SLCO1B3* mutations in each haplotype result in absence of a detectable signal for OATP1B protein in the liver.

In family A2, a heterozygous splice donor site mutation, c.481+1G→T, in intron 5 of *SLCO1B1* would result in dysfunctional RNA or protein. Its co-occurrence with the 405-kb R2 deletion in two asymptomatic family members (Table 1) indicates that a single functional *SLCO1B3* allele can prevent RS.

A search for copy number variations (CNVs) in existing databases and CNV genotyping of more than 2,300 individuals from various populations (see Supplemental Results) revealed additional heterozygous small and large deletions predicted to disrupt *SLCO1B1* or *SLCO1B3* function, including several approximately 400-kb deletions similar or identical to the R2 haplotype–linked deletion. One individual without jaundice, heterozygous for the R1 haplotype–associated c.1738C→T (p.R580X) mutation in *SLCO1B1*, was also homozygous for the R1 haplotype–associated deletion in *SLCO1B3*. Thus, a single functional *SLCO1B1* allele can also prevent RS. Combined with the findings in family A2 described above, this demonstrates that only a complete deficiency of both alleles of *SLCO1B1* and *SLCO1B3* will result in RS.

Discussion

We demonstrate here that RS is an obligate two-gene disorder, caused by a complete deficiency of the major hepatic drug uptake transporters OATP1B1 and OATP1B3. We further identified individuals with a complete deficiency of either OATP1B1 or OATP1B3, which was not recognizable by obvious jaundice. In spite of the documented important functions of especially OATP1B1 in drug detoxification, apparently such deficiencies are compatible with relatively normal life.

Using Oatp1a/1b-knockout mice, which can, retrospectively, be considered to be a partial model for RS, we showed that Abcc3 is an important factor for the RS-like conjugated hyperbilirubinemia. Our data imply that in the normal human liver ABCC3, OATP1B1, and OATP1B3 may form a liver-blood shuttling loop for bilirubin glucuronide, similar to that driven by Oatp1a/1b and Abcc3 in the mouse (Figure 6). A substantial fraction of bilirubin conjugated in hepatocytes is secreted back into the blood by ABCC3 and subsequently reabsorbed in downstream hepatocytes by OATP1B1 and OATP1B3. In RS this reuptake is hampered, causing increased plasma bilirubin glucuronide levels and jaundice. The flexible "hepatocyte hopping" afforded by this loop facilitates efficient detoxification, presumably by circumventing saturation of further detoxification processes in upstream hepatocytes, including excretion into bile. Indeed, we could show that, counterintuitively, but in accordance with the hepatocyte hopping model, loss of Abcc3 in mice resulted in decreased biliary excretion of bilirubin glucuronide, as long as Oatp1a/1b was present (Figure 2). This process likely also enhances hepatic detoxification of numerous drugs and drug conjugates (e.g., glucuronide, sulfate, and glutathione conjugates) transported by OATP1B1/3 and ABCC3. Moreover, this principle may also apply to other saturable hepatocyte detoxifying processes, such as phase I and phase II metabolism, as long as the substrate compounds involved are transported by ABCC3 and OATP1B proteins. Additional sinusoidal efflux and uptake transporters (e.g., ABCC4, OATP2B1, NTCP) will further widen the scope of compounds affected by this hepatocyte hopping process. Results obtained with the *Slco1a/1b;Abcc3*-knockout mice indeed show that in addition to Abcc3 there must be other sinusoidal efflux processes for bilirubin glucuronides. Preventing accumulation of drug glucuronides may be particularly important, since protein adduction by acyl-glucuronides is a well-established cause of drug (hepato)toxicity (32).

One should exercise caution when extrapolating mouse data to humans, and the individual Oatp1a/1b proteins are not straightforward orthologs of human OATP1B1 and OATP1B3. However, there is a strong analogy between the bilirubin phenotypes of Oatp1a/1b-knockout mice and human Rotor subjects. Moreover, the hepatic transgenic expression of human

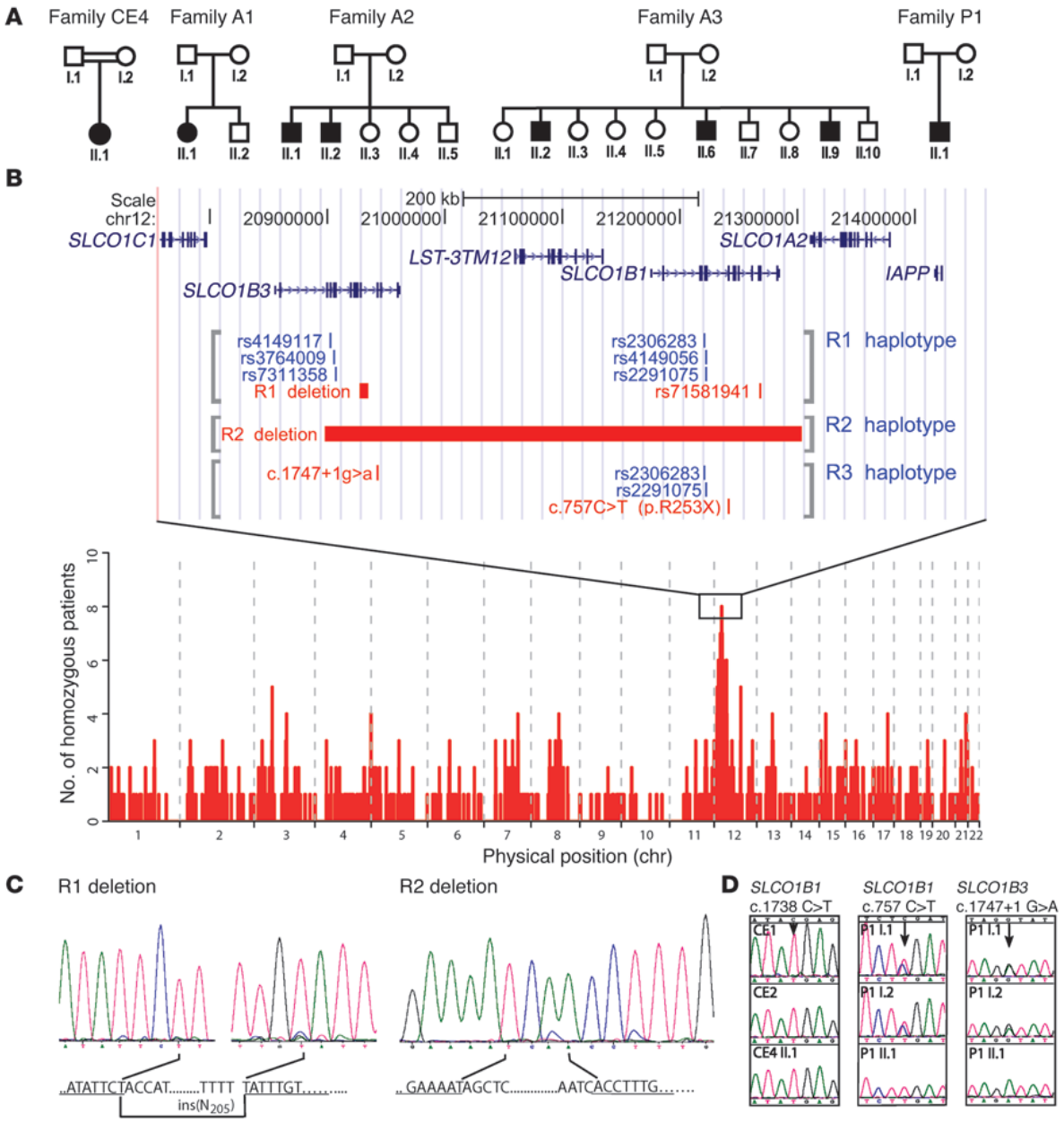


Figure 4

RS families display deficiencies in *SLCO1B1* and *SLCO1B3*. (**A**) Pedigrees of the investigated families. Black symbols denote RS index subjects. Parents in family CE4 had a documented common ancestor. Families CE1–CE3 (only single individuals analyzed) are not shown. (**B**) Homozygosity regions in 8 RS index subjects and overview of detected mutations and polymorphisms. The genome map shows number and location of overlapping homozygosity regions in RS index subjects, gene content of the top candidate region on chromosome 12, and the genotypes forming all 3 identified RS haplotypes. Mutations crucial for RS are shown in red. chr, chromosome. (**C**) Sequences and electropherograms of the R1 and R2 deletion breakpoints. (**D**) Pathogenic point mutations in R1 and R3 haplotypes. Electropherograms indicate the c.1738C→T (p.R580X) mutation in *SLCO1B1* in probands CE1, CE2, and CE4 II.1 and the c.757C→T (p.R253X) and c.1747+1G→A mutations in *SLCO1B1* and *SLCO1B3*, respectively, in family P1.

OATP1B1 or OATP1B3 resulted in virtually complete rescue of the Oatp1a/1b-knockout phenotype for bilirubin handling (Figure 3). This strongly supports that the principles governing bilirubin handling by Oatp1a/1b in mouse liver also apply to OATP1B1 and OATP1B3 in human liver.

Analogous to the mouse data for Oatp1a/1b (25), the extensive glucuronidation of bilirubin in Rotor subjects suggests that OATP1B1 and/or OATP1B3 are not strictly essential for uptake of UCB into the liver. Passive transmembrane diffusion is one likely candidate to take over this process, in hepatocytes and probably many other cell types as well (e.g., ref. 33), but we do not exclude that additional uptake transporters (perhaps OATP2B1) can also contribute to UCB uptake. However, OATP1B1 and/or OATP1B3 probably do contribute to hepatic UCB uptake, since in RS subjects a significant increase in plasma UCB is usually observed and reduced clearance of UCB has been reported (34, 35). Moreover, polymorphisms in *SLCO1B1*

Table 1
Mutations in *SLCO1B* genes detected in RS subjects and their family members

Subject	Family status	Haplotype R1-linked mutations		Haplotype R2-linked mutations		Haplotype R3-linked mutations	
		SLCO1B3 7.2-kb deletion	*SLCO1B1* c.1738C→T (p.R580X) rs71581941	*SLCO1B* locus 405-kb deletion	*SLCO1B1* c.481+1G→T splice site mutation	*SLCO1B3* c.1747+1G→A splice site mutation	*SLCO1B1* c.757C→T (p.R253X)
CE1	**Proband**	del/del	T/T				
CE2	**Proband**	del/del	T/T				
CE4 I.1	Father	del/WT	T/C				
CE4 I.2	Mother	del/WT	T/C				
CE4 II.1	**Proband**	del/del	T/T				
CE3	**Proband**			del/del	–/–		
A1 I.1	Father			del/WT	–/G		
A1 I.2	Mother			del/WT	–/G		
A1 II.1	**Proband**			del/del	–/–		
A1 II.2	Brother			WT/WT	G/G		
A2 I.1	Father			del/WT	–/G		
A2 I.2	Mother			del/WT	–/T		
A2 II.1	**Brother**			del/del	–/–		
A2 II.2	**Proband**			del/del	–/–		
A2 II.3	Sister			del/WT	–/T		
A2 II.4	Sister			del/WT	–/G		
A2 II.5	Brother			WT/WT	G/T		
A3 I.1	Father			del/WT	–/G		
A3 I.2	Mother			del/WT	–/G		
A3 II.1	Sister			WT/WT	G/G		
A3 II.2	**Proband**			del/del	–/–		
A3 II.3	Sister			del/WT	–/G		
A3 II.4	Sister			del/WT	–/G		
A3 II.5	Sister			del/WT	–/G		
A3 II.6	**Brother**			del/del	–/–		
A3 II.7	Brother			del/WT	–/G		
A3 II.8	Sister			del/WT	–/G		
A3 II.9	**Brother**			del/del	–/–		
A3 II.10	Brother			WT/WT	G/G		
P1 I.1	Father					G/A	C/T
P1 I.2	Mother					G/A	C/T
P1 II.1	**Proband**					A/A	T/T

Boldface indicates index subjects with RS (n = 11; 8 probands, 3 affected siblings); 405-kb deletion (assembly NCBI36/hg18) — g.(20898911)_(21303509)del(CA)ins; 7.2-kb deletion (assembly NCBI36/hg18) — g.(20927077)_(20934292)del(N205)ins. WT, wild-type sequence, i.e., sequence from which all exons of *SLCO1B1* and *SLCO1B3* could be amplified. Genotypes for all empty entries were wild-type in sequence and/or heterozygous or homozygous for the large haplotype R2–linked deletion as predicted.

and *SLCO1B3* have been associated with increased serum UCB levels (36, 37). There was also a significant, nearly 2-fold increase in plasma UCB in the *Slco1a/1b*$^{-/-}$ mice, and this was partially reversed by both human OATP1B1 and OATP1B3 expression (Figure 3C).

It should be noted that UGT1A1-mediated glucuronidation may also occur in extrahepatic tissues, for instance, colon (38), and we cannot exclude that some of the bilirubin glucuronide observed in RS plasma has resulted from such extrahepatic glucuronidation, possibly enhanced by the increased plasma UCB levels. It seems unlikely, however, that all bilirubin glucuronide in RS subjects would derive from extrahepatic glucuronidation. This would require a complete block of hepatic UCB uptake (due to the OATP1B1 and OATP1B3 deficiency), but at the same time require efficient uptake of UCB into UGT1A1-containing extrahepatic cells (e.g., colonocytes) that do not normally express OATP1B1 and OATP1B3, and certainly not in Rotor subjects. If UCB transmembrane diffusion can do this efficiently, it is hard to see why this would not mediate substantial uptake into the liver as well. Only if hepatic diffusion uptake is negligible (which seems physically unlikely) and an unknown efficient UCB uptake system would function in colonocytes (and not in liver), could one envisage such a situation. On balance, this seems rather implausible.

Elucidation of OATP1B1 and OATP1B3 deficiency as the cause of RS can also readily explain the other diagnostic traits of the disorder. Absence of OATP1B1/3-mediated liver uptake would cause the decreased plasma clearance of anionic diagnostic dyes such as indocyanine green and BSP, an excellent substrate of OATP1B1 and OATP1B3 (15), and the greatly reduced or delayed visualization of the liver by anionic cholescintigraphic radiotracers such as ^{99m}Tc-HIDA and ^{99m}Tc-mebrofenin (5, 6). ^{99m}Tc-mebrofenin, for instance, is efficiently transported by both OATP1B1 and OATP1B3 (39).

The markedly increased urinary excretion of coproporphyrins, and the increased preponderance of isomer I over III in urine of RS subjects, could be simply explained by reduced (re)uptake of

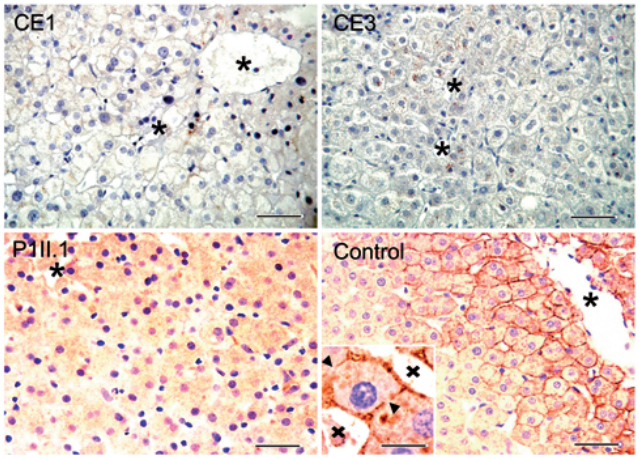


Figure 5
Liver expression of OATP1B proteins in RS subjects and control. With an anti-OATP1B1/3 antibody, basolateral membrane immunostaining of hepatocytes in centrilobular areas was intense in control. Asterisks indicate central veins, arrowheads bile canaliculi, and crosses sinusoids. OATP1B proteins were not detectable in RS subjects CE1 (haplotype R1), CE3 (haplotype R2), and P1 II.1 (haplotype R3). Scale bars: 25 μm (original magnification of CE1 and CE3, ×400; original magnification of P1 II.1 and control, ×200); inset: 5 μm (original magnification, ×1,000).

these compounds into the liver, partly shifting the excretion route from hepatobiliary/fecal to urinary, especially for isomer I. Coproporphyrin I and III thus most likely are transported substrates of OATP1B1 and OATP1B3. Indeed, interaction of several porphyrins with OATP1B1 has recently been demonstrated (40).

Phenotypic abnormalities in RS subjects are surprisingly moderate. Perhaps OATP1B1 and OATP1B3 functions are partly taken over by other sinusoidal uptake transporters, such as OATP2B1. Nevertheless, since even reduced-activity OATP1B1 polymorphisms can result in life-threatening drug toxicities (20–24, 41, 42), such risks are likely increased substantially in RS subjects. Their evident jaundice, however, may have been a warning sign for physicians to prescribe drugs with caution.

The obligatory deficiency in two different, medium-sized genes explains the rarity of RS, with a roughly estimated frequency of about 1 in 10^6, although it might be several-fold lower or higher in different populations. Complete deficiency of either OATP1B1 or OATP1B3 alone will occur much more frequently but will not cause jaundice. For instance, the p.R580X mutation in OATP1B1 occurred at an allele frequency of 0.008 (3 of 354) in a Japanese population (43), suggesting that about 1 in 14,000 individuals in this population would be homozygous for this full-deficiency mutant. Such individuals might demonstrate idiosyncratic hypersensitivity to OATP1B1 substrate drugs, including statins or irinotecan. Similarly, in the present study we identified a non-jaundiced individual homozygously deficient for *SLCO1B3* in our CNV screening of approximately 2,300 individuals, in line with a non-negligible incidence of fully OATP1B3-deficient individuals.

Some drugs, such as high-dose cyclosporine A, can transiently increase plasma levels of conjugated bilirubin without evoking other markers for liver damage (44, 45). Until now, such increases were thought to be primarily mediated by inhibition of ABCC2 as the main biliary excretion factor for bilirubin glucuronide. However, given the insights from the present study, direct inhibition of OATP1B1 and/or OATP1B3 by the applied drug may be an additional or even the main cause of such drug-induced conjugated hyperbilirubinemias. This might for instance apply to cyclosporine A, rifampin, rifamycin SV, or other drugs that are established inhibitors of OATP1B proteins (23). Moreover, heterozygous carriers of the various full-deficiency mutations in OATP1B1/3 might be more susceptible to such inhibitory effects. This also applies to drug-drug interactions mediated through OATP1B1/3 inhibition.

The molecular mechanism we identified in RS may also underlie a similar disorder called hepatic uptake and storage syndrome, or conjugated hyperbilirubinemia type III (OMIM %237550) (46). This hypothesis can now be tested by mutational analysis of OATP1B1 and OATP1B3 in the only reported family to date. Furthermore, a mutant strain of Southdown sheep has also been described as displaying a similar hepatic uptake and storage syndrome (46), and it would not surprise us if these animals would likewise have a deficiency of one or more hepatic sinusoidal OATPs. The observation that mutant Southdown sheep, like the *Slco1a/1b*$^{-/-}$ mice (25), also display strongly reduced clearance of (unconjugated) cholic acid, but not of (conjugated) taurocholic acid (47), further supports this idea.

Collectively, our findings explain the genetic and molecular basis of RS. The demonstration of an Abcc3-, OATP1B1-, and OATP1B3-driven detoxification-enhancing liver-blood shuttling loop in mice and, by implication, most likely also in humans challenges the view of one-way excretion from blood through liver to bile of bilirubin and drugs detoxified by conjugation. Furthermore, the identified full-deficiency alleles of *SLCO1B1* and *SLCO1B3* may contribute to various "idiosyncratic" drug hypersensitivities.

Methods

Mouse strains and conditions. Mice were housed and handled according to institutional guidelines complying with Dutch legislation. *Slco1a/1b*$^{-/-}$, *Abcc2*$^{-/-}$, *Abcc3*$^{-/-}$, and *Abcc2*$^{-/-}$*Abcc3*$^{-/-}$ mice have been described (25–27, 48). Human OATP1B1 transgenic mice have been described (29), and human OATP1B3 transgenic mice were generated in an analogous manner, using an apoE promoter to obtain liver-specific expression of the transgene. Each transgene was crossed back into an *Slco1a/1b*$^{-/-}$ background to obtain the corresponding humanized rescue strains. Routine mouse conditions and analyses of mouse samples are described in Supplemental Methods.

Western blot analysis. Isolation of crude membrane fractions from mouse liver, kidney, and small intestine and Western blotting were as described previously (29). For detection of Abcc2 and Abcc3 primary antibodies, M_2III-5 (dilution 1:1,000) and M_3-18 (dilution 1:25) were used, respectively. For detection of transgenic OATP1B1 and OATP1B3 in mouse liver, the rabbit polyclonal antibodies ESL and SKT, provided by D. Keppler (Deutsches Krebsforschungszentrum, Heidelberg, Germany) were used (17, 18).

RNA isolation, cDNA synthesis, and RT-PCR. RNA isolation from mouse liver, kidney, and small intestine and subsequent cDNA synthesis and RT-PCR were as described previously (49). Specific primers (QIAGEN) were used to detect expression levels of *Slco1a1*, *Slco1a4*, *Slco1a6*, *Slco1b2*, *Slco2b1*, *Slc10a1*, *Slc10a2*, *Abcc2–4*, *Abcb1a*, *Abcb1b*, *Abcb11*, *Abcg2*, *Osta*, *Ostb*, and *Ugt1a1*.

Analysis of bilirubin in mouse plasma, bile, and urine. Gallbladder cannulations and collection of bile and urine in male mice of the various strains (n = 4–7) as well as bilirubin detection were as described (25, 50, 51). For details, see Supplemental Methods.

RS families. We examined 11 RS index subjects (8 probands, 3 siblings of probands) of 8 families and 21 clinically healthy members of 5 of these 8 families. Family members of 3 probands (CE1–CE3) were not available.

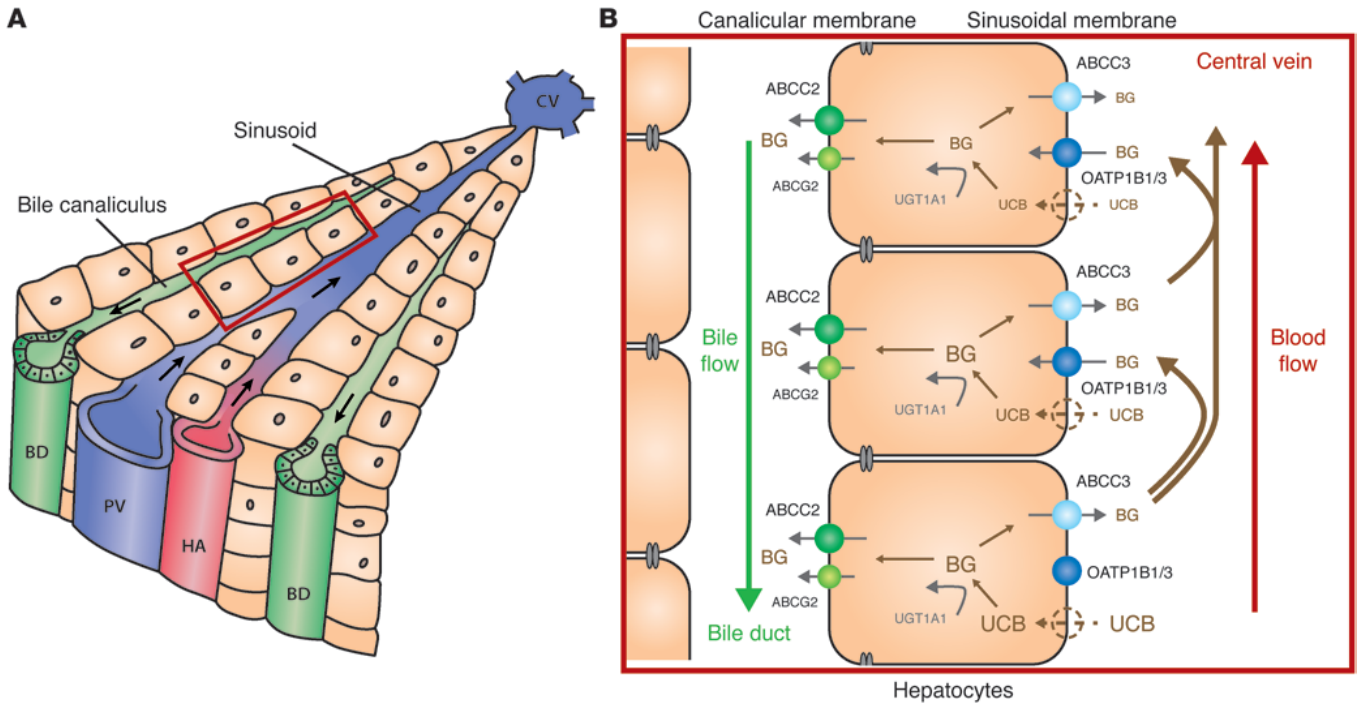


Figure 6
Hepatocyte hopping distributes the biliary excretion load of bilirubin glucuronides across the liver lobule. (**A**) Schematic of liver lobule. Hepatocytes are organized around portal tracts, with branches of the portal vein (PV), hepatic artery (HA), and bile ducts (BD). The PV and HA deliver nutrient- and oxygen-rich blood, respectively, which flows through the sinusoids toward the central vein (CV). Basolateral (sinusoidal) membranes of hepatocytes are flushed with perisinusoidal plasma. Bile flows in the opposite direction toward bile ducts through canaliculi lined by canalicular membranes of hepatocytes. (**B**) Hepatocyte hopping cycle. UCB enters the hepatocytes via passive diffusion and/or transporters, which may include OATP1B1 and/or OATP1B3 in non-Rotor subjects. Conjugation with glucuronic acid by UGT1A1 to bilirubin glucuronides (BG) takes place in endoplasmic reticulum. BG is secreted into bile mainly by ABCC2. ABCG2 also can contribute to this process. Even under physiological conditions, a substantial fraction of the intracellular BG is rerouted by ABCC3 to the blood, from which it can be taken up by downstream hepatocytes via OATP1B1/3 transporters. This flexible off-loading of BG to downstream hepatocytes prevents saturation of biliary excretion capacity in upstream hepatocytes. Relative type sizes of UCB and BG represent local concentrations. Schematic modified, with permission, from ref. 54.

Families CE1–CE4 are of mixed Central European descent by family report. Three families (A1–A3) are Saudi Arabs, and one family (P1) is from the Philippines. Central European families were ascertained at the Institute for Clinical and Experimental Medicine, Prague, and Saudi Arab and Filipino families at the Saudi Aramco Dhahran Health Center. Medical histories were obtained by referring consultants. Subjects CE1 and CE2 were reported as case 1 and case 2, respectively (14).

ABCC2 mutation screening. *ABCC2* mutation screening was performed in 8 probands representing all studied families as described previously (14).

Genotyping. Genotyping was performed using Affymetrix GeneChip Mapping 6.0 Arrays (Affymetrix) according to the manufacturer's protocol. Raw feature intensities were extracted from Affymetrix GeneChip Scanner 3000 7G images using GeneChip Control Console Software 2.01. Individual SNP calls were generated using Affymetrix Genotyping Console Software 3.02. Details of the experiment and individual genotyping data are available at the GEO repository (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE33733.

Multipoint nonparametric and parametric linkage analysis. Multipoint nonparametric and parametric linkage analysis along with determination of the most likely haplotypes was performed with version 1.1.2 of Merlin software (52). Parametric linkage was carried out assuming an autosomal recessive mode of inheritance with a 1.00 constant, age-independent penetrance, 0.00 phenocopy rate, and 0.0001 frequency of disease allele. Results were visualized in version 1.032 of HaploPainter software (53) and in version 2.9.2 of R-project statistical software (http://www.r-project.org/).

Homozygosity mapping. Extended homozygosity regions were identified in Affymetrix Genotyping Console Software version 3.02 using the algorithm comparing values from the user's sample set and SNP-specific distributions derived from a reference set of 200 ethnically diverse individuals. Distribution of extended homozygosity regions in affected and healthy individuals was analyzed and visualized using custom R-script.

Copy number changes. Copy number changes were identified in Affymetrix Genotyping Console Software version 3.02. Data from both SNP and copy number probes were used to identify copy number aberrations compared with built-in reference. Only regions larger than 10 kb containing at least 5 probes were reported.

Quantitative PCR. Quantitative PCR was carried out in duplicate on a LightCycler 480 System (Roche Applied Science). Data were analyzed by LightCycler 480 Software, release 1.5.0. Absolute quantification was used to determine copy number status of a given fragment in analyzed samples. Genomic positions of the analyzed fragments and control genes, corresponding primer sequences, and Universal ProbeLibrary probes used for amplification and quantitation are provided in Supplemental Table 4.

Mutation analysis. Long-range PCR products encompassing the genomic regions of deletion breakpoint boundaries were gel-purified and sequenced using a primer walking approach. DNA sequencing of PCR products and

genomic fragments covering 1 kb of the promoter regions and all of the exons, with their corresponding exon-intron boundaries, of *SLCO1B1*, *SLCO1B3*, and *SLCO1A2* was performed. For details, see Supplemental Methods. Confirmation and segregation of both identified copy number changes and missense mutations in the families, as well as frequency of the mutations in a control population of mixed European descent, were assessed by PCR, PCR-RFLP, and direct sequencing of corresponding genomic DNA fragments. For primer sequences, see Supplemental Table 4.

Histology and immunohistochemistry. Archival liver biopsy specimens were available from 5 unrelated RS index subjects (probands, families CE1, CE2, CE3, and P1; brother [A3 II.9] of proband from family A3). Sections of paraffin-embedded material (formalin or Carnoy solution fixative; 4–6 μm thick) were routinely stained with hematoxylin and eosin and periodic acid–Schiff techniques. For OATP1B1, OATP1B3, and ABCC2 immunostaining, routine techniques were applied (see Supplemental Methods). OATP1B1 and OATP1B3 detection was performed with a primary mouse anti-OATP1B antibody (clone mMDQ, GeneTex; recognizing the N terminus of both OATP1B1 and OATP1B3), 1:100 dilution, overnight at 4°C (31).

Statistics. One-way ANOVA followed by Tukey's multiple comparison test was used to assess statistical significance of differences between data sets. Results are presented as mean ± SD. Differences were considered statistically significant when P was less than 0.05.

Study approval. All mouse studies were ethically reviewed and carried out in accordance with European directive 86/609/EEC and Dutch legislation and the GlaxoSmithKline policy on the Care, Welfare and Treatment of Laboratory Animals. Experiments were approved by the Animal Experimentation Committee (DEC) of the Netherlands Cancer Institute. Investigations involving humans were approved by the Institutional Review Boards of the Institute for Clinical and Experimental Medicine, Prague, Czech Republic, and the Saudi Aramco Dhahran Health Centre, with written informed consent received from participants or their guardians, and conducted according to Declaration of Helsinki principles.

Acknowledgments

The human study was supported in part by the Ministry of Education of the Czech Republic (projects MSM0021620806 and 1M6837805002) and by the Institute for Clinical and Experimental Medicine (MZO 00023001). The mouse work was supported in part by grant S2918 from GlaxoSmithKline and grant NKI 2007-3764 from the Dutch Cancer Society. The authors thank L. Budišová and M. Boučková for technical assistance and L. Vítek, M. Mikulecký, J. Horák, and A. Šuláková for referring patients CE1–CE4.

Received for publication June 16, 2011, and accepted in revised form November 30, 2011.

Address correspondence to: Alfred H. Schinkel, Division of Molecular Biology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. Phone: 31.20.5122046; Fax: 31.20.6961383; E-mail: a.schinkel@nki.nl. Or to: Milan Jirsa, Department of Experimental Medicine, Institute for Clinical and Experimental Medicine, Vídenská 1958/9, 140 00 Prague 4 — Krč, Czech Republic. Phone: 420.261362773; Fax: 420.241721666; E-mail: miji@ikem.cz.

1. Chowdhury JR, Chowdhury NR, Jansen PLM. Bilirubin metabolism and its disorders. In: Boyer TD, Wright TL, Manns MP, Zakim D, eds. *Zakim and Boyer's Hepatology. A Textbook of Liver Diseases*. Vol. 2. Philadelphia, Pennsylvania, USA: Saunders Elsevier; 2006:1449–1474.
2. Chowdhury JR, Wolkoff AW, Chowdhury NR, Arias IM. Hereditary jaundice and disorders of bilirubin metabolism. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. Vol. 2. New York, New York, USA: McGraw Hill; 2001:3063–3101.
3. Rotor AB, Manahan L, Florentin A. Familial nonhemolytic jaundice with direct van den Bergh reaction. *Acta Med Phil*. 1948;5:37–49.
4. Wolpert E, Pascasio FM, Wolkoff AW, Arias IM. Abnormal sulfobromophthalein metabolism in Rotor's syndrome and obligate heterozygotes. *N Engl J Med*. 1977;296(19):1099–1101.
5. Bar-Meir S, Baron J, Seligson U, Gottesfeld F, Levy R, Gilat T. 99mTc-HIDA cholescintigraphy in Dubin-Johnson and Rotor syndromes. *Radiology*. 1982; 142(3):743–746.
6. LeBouthillier G, Morais J, Picard M, Picard D, Chartrand R, Pommier G. Scintigraphic aspect of Rotor's disease with Technetium-99m-mebrofenin. *J Nucl Med*. 1992;33(8):1550–1551.
7. Kartenbeck J, Leuschner U, Mayer R, Keppler D. Absence of the canalicular isoform of the MRP gene-encoded conjugate export pump from the hepatocytes in Dubin-Johnson syndrome. *Hepatology*. 1996;23(5):1061–1066.
8. Paulusma CC, et al. A mutation in the human canalicular multispecific organic anion transporter gene causes the Dubin-Johnson syndrome. *Hepatology*. 1997;25(6):1539–1542.
9. Nowicki MJ, Poley JR. The hereditary hyperbilirubinaemias. *Baillieres Clin Gastroenterol*. 1998; 12(2):355–367.
10. Strassburg CP. Hyperbilirubinemia syndromes (Gilbert-Meulengracht, Crigler-Najjar, Dubin-Johnson, and Rotor syndrome). *Best Pract Res Clin Gastroenterol*. 2010;24(5):555–571.
11. Wolkoff AW, Wolpert E, Pascasio FN, Arias IM. Rotor's syndrome. A distinct inheritable pathophysiologic entity. *Am J Med*. 1976;60(2):173–179.
12. König J, Rost D, Cui Y, Keppler D. Characterization of the human multidrug resistance protein isoform MRP3 localized to the basolateral hepatocyte membrane. *Hepatology*. 1999;29(4):1156–1163.
13. Lee YM, et al. Identification and functional characterization of the natural variant MRP3-Arg1297His of human multidrug resistance protein 3 (MRP3/ ABCC3). *Pharmacogenetics*. 2004;14(4):213–223.
14. Hrebicek M, et al. Rotor-type hyperbilirubinaemia has no defect in the canalicular bilirubin export pump. *Liver Int*. 2007;27(4):485–491.
15. Hagenbuch B, Gui C. Xenobiotic transporters of the human organic anion transporting polypeptide (OATP) family. *Xenobiotica*. 2008;38(7–8):778–801.
16. Hagenbuch B, Meier PJ. Organic anion transporting polypeptides of the OATP/ SLC21 family: phylogenetic classification as OATP/ SLCO superfamily, new nomenclature and molecular/functional properties. *Pflügers Arch*. 2004;447(5):653–665.
17. König J, Cui Y, Nies AT, Keppler D. A novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am J Physiol Gastrointest Liver Physiol*. 2000;278(1):G156–G164.
18. König J, Cui Y, Nies AT, Keppler D. Localization and genomic organization of a new hepatocellular organic anion transporting polypeptide. *J Biol Chem*. 2000;275(30):23161–23168.
19. Abe T, et al. LST-2, a human liver-specific organic anion transporter, determines methotrexate sensitivity in gastrointestinal cancers. *Gastroenterology*. 2001;120(7):1689–1699.
20. Takane H, et al. Life-threatening toxicities in a patient with UGT1A1*6/*28 and SLCO1B1*15/ *15 genotypes after irinotecan-based chemotherapy. *Cancer Chemother Pharmacol*. 2009; 63(6):1165–1169.
21. Link E, et al. SLCO1B1 variants and statin-induced myopathy — a genomewide study. *N Engl J Med*. 2008; 359(8):789–799.
22. Treviño LR, et al. Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. *J Clin Oncol*. 2009;27(35):5972–5978.
23. Kalliokoski A, Niemi M. Impact of OATP transporters on pharmacokinetics. *Br J Pharmacol*. 2009; 158(3):693–705.
24. König J, Seithel A, Gradhand U, Fromm MF. Pharmacogenomics of human OATP transporters. *Naunyn Schmiedebergs Arch Pharmacol*. 2006; 372(6):432–443.
25. Van de Steeg E, et al. Organic anion transporting polypeptide 1a/1b-knockout mice provide insights into hepatic handling of bilirubin, bile acids and drugs. *J Clin Invest*. 2010;120(8):2942–2952.
26. Vlaming ML, et al. Carcinogen and anticancer drug transport by Mrp2 in vivo: studies using Mrp2 (Abcc2) knockout mice. *J Pharmacol Exp Ther*. 2006; 318(1):319–327.
27. Vlaming ML, et al. Impact of Abcc2 (Mrp2) and Abcc3 (Mrp3) on the in vivo elimination of methotrexate and its main toxic metabolite 7-hydroxymethotrexate. *Clin Cancer Res*. 2008;14(24):8152–8160.
28. Rius M, Nies AT, Hummel-Eisenbeiss J, Jedlitschky G, Keppler D. Cotransport of reduced glutathione with bile salts by MRP4 (ABCC4) localized to the basolateral hepatocyte membrane. *Hepatology*. 2003; 38(2):374–384.
29. Van de Steeg E, et al. Methotrexate pharmacokinetics in transgenic mice with liver-specific expression of human OATP1B1 (SLCO1B1). *Drug Metab Dispos*. 2009;37(1):1–5.
30. Cui Y, König J, Leier I, Buchholz U, Keppler D. Hepatic uptake of bilirubin and its conjugates by the human organic anion transporter SLC21A6. *J Biol Chem*. 2001;276(13):9626–9630.
31. Cui Y, et al. Detection of the human organic anion transporters SLC21A6 (OATP2) and SLC21A8 (OATP8) in liver and hepatocellular carcinoma. *Lab Invest*. 2003;83(4):527–538.
32. Zhou S, Chan E, Duan W, Huang M, Chen YZ.

Drug bioactivation, covalent binding to target proteins and toxicity relevance. *Drug Metab Rev*. 2005; 37(1):41–213.
33. Zucker SD, Goessling W, Hoppin AG. Unconjugated bilirubin exhibits spontaneous diffusion through model lipid bilayers and native hepatocyte membranes. *J Biol Chem*. 1999;274(16):10852–10862.
34. Kawasaki H, Kimura N, Irisa T, Hirayama C. Dye clearance studies in Rotor's syndrome. *Am J Gastroenterol*. 1979;71(4):380–388.
35. Fedeli G, et al. Impaired clearance of cholephilic anions in Rotor syndrome. *Z Gastroenterol*. 1983; 21(5):228–233.
36. Zhang W, et al. OATP1B1 polymorphism is a major determinant of serum bilirubin level but not associated with rifampicin-mediated bilirubin elevation. *Clin Exp Pharmacol Physiol*. 2007;34(12):1240–1244.
37. Sanna S, et al. Common variants in the SLCO1B3 locus are associated with bilirubin levels and unconjugated hyperbilirubinemia. *Hum Mol Genet*. 2009; 18(14):2711–2718.
38. Strassburg CP, Manns MP, Tukey RH. Expression of the UDP-glucuronosyltransferase 1A locus in human colon. Identification and characterization of the novel extrahepatic UGT1A8. *J Biol Chem*. 1998; 273(15):8719–8726.
39. Ghibellini G, Leslie EM, Pollack GM, Brouwer KL. Use of tc-99m mebrofenin as a clinical probe to assess altered hatobiliary transport: integration of in vitro, pharmacokinetic modeling, and simulation studies. *Pharm Res*. 2008;25(8):1851–1860.
40. Campbell SD, Lau WF, Xu JJ. Interaction of porphyrins with human organic anion transporting polypeptide 1B1. *Chem Biol Interact*. 2009;182(1):45–51.
41. Morimoto K, Oishi T, Ueda S, Ueda M, Hosokawa M, Chiba K. A novel variant allele of OATP-C (SLCO1B1) found in a Japanese patient with pravastatin-induced myopathy. *Drug Metab Pharmacokinet*. 2004;19(6):453–455.
42. Tirona RG, Kim RB. Pharmacogenomics of organic anion-transporting polypeptides (OATP). *Adv Drug Deliv Rev*. 2002;54(10):1343–1352.
43. Kim SR, et al. Genetic variations and frequencies of major haplotypes in SLCO1B1 encoding the transporter OATP1B1 in Japanese subjects: SLCO1B1*17 is more prevalent than *15. *Drug Metab Pharmacokinet*. 2007;22(6):456–461.
44. Yahanda AM, et al. Phase I trial of etoposide with cyclosporine as a modulator of multidrug resistance. *J Clin Oncol*. 1992;10(10):1624–1634.
45. List AF, et al. Phase I/II trial of cyclosporine as a chemotherapy-resistance modifier in acute leukemia. *J Clin Oncol*. 1993;11(9):1652–1660.
46. Dhumeaux D, Berthelot P. Chronic hyperbilirubinemia associated with hepatic uptake and storage impairment. A new syndrome resembling that of mutant southdown sheep. *Gastroenterology*. 1975; 69(4):988–993.
47. Engelking LR, Gronwall R. Bile acid clearance in sheep with hereditary hyperbilirubinemia. *Am J Vet Res*. 1979;40(9):1277–1280.
48. Zelcer N, et al. Mice lacking Mrp3 (Abcc3) have normal bile salt transport, but altered hepatic transport of endogenous glucuronides. *J Hepatol*. 2006; 44(4):768–775.
49. Van Waterschoot RA, et al. Midazolam metabolism in cytochrome P450 3A knockout mice can be attributed to up-regulated CYP2C enzymes. *Mol Pharmacol*. 2008;73(3):1029–1036.
50. Van Herwaarden AE, et al. The breast cancer resistance protein (Bcrp1/Abcg2) restricts exposure to the dietary carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine. *Cancer Res*. 2003; 63(19):6447–6452.
51. Spivak W, Carey MC. Reverse-phase h.p.l.c. separation, quantification and preparation of bilirubin and its conjugates from native bile. Quantitative analysis of the intact tetrapyrroles based on h.p.l.c. of their ethyl anthranilate azo derivatives. *Biochem J*. 1985;225(3):787–805.
52. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin — rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet*. 2002;30(1):97–101.
53. Thiele H, Nurnberg P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics*. 2005;21(8):1730–1732.
54. Van de Steeg E, Iusuf D, Schinkel AH. Physiological and pharmacological functions of OATP1A/1B transporters: insights from knockout and transgenic mice. In: Van de Steeg E. *Physiological and pharmacological functions of OATP1A/1B transporters. PhD thesis, University of Utrecht.* Enschede, the Netherlands: Gildeprint drukkerijen; 2010:9–37.

BMC Genomics

BioMed Central

Research article

Open Access

Development of a human mitochondrial oligonucleotide microarray (h-MitoArray) and gene expression analysis of fibroblast cell lines from 13 patients with isolated F_1F_o ATP synthase deficiency

Alena Žížková[1,2,5], Viktor Stránecký[1,2], Robert Ivánek[1,2,4], Hana Hartmannová[1,2], Lenka Nosková[2], Lenka Piherová[1,2], Markéta Tesařová[1,3], Hana Hansíková[1,3], Tomáš Honzík[3], Jiří Zeman[1,3], Petr Divina[4], Andrea Potocká[1,5], Jan Paul[1,5], Wolfgang Sperl[6], Johannes A Mayr[6], Sara Seneca[7], Josef Houštěk[1,5] and Stanislav Kmoch*[1,2]

Address: [1]Center for Applied Genomics, 1st Faculty of Medicine, Charles University, Prague, Czech Republic, [2]Institute of Inherited Metabolic Disorders, 1st Faculty of Medicine, Charles University, Prague, Czech Republic, [3]Department of Pediatrics, 1st Faculty of Medicine, Charles University, Prague, Czech Republic, [4]Institute of Molecular Genetics, Academy of Science of the Czech Republic, Prague, Czech Republic, [5]Department of Bioenergetics, Institute of Physiology, Academy of Science of the Czech Republic, Prague, Czech Republic, [6]Department of Pediatrics, Paracelsus Medical University, Salzburg, Austria and [7]Center of Medical Genetics, Free University Brussels, Brussels, Belgium

Email: Alena Žížková - acizk@LF1.cuni.cz; Viktor Stránecký - vstra@LF1.cuni.cz; Robert Ivánek - ivanek@img.cas.cz; Hana Hartmannová - hhart@LF1.cuni.cz; Lenka Nosková - lnosk@LF1.cuni.cz; Lenka Piherová - Lenka.Piherova@LF1.cuni.cz; Markéta Tesařová - Marketa.Tesarova@LF1.cuni.cz; Hana Hansíková - HHansikova@seznam.cz; Tomáš Honzík - HonzikT@seznam.cz; Jiří Zeman - jzem@LF1.cuni.cz; Petr Divina - divina@img.cas.cz; Andrea Potocká - potockaa@biomed.cas.cz; Jan Paul - paulj@biomed.cas.cz; Wolfgang Sperl - w.sperl@salk.at; Johannes A Mayr - h.mayr@salk.at; Sara Seneca - sara.seneca@az.vub.ac.be; Josef Houštěk - houstek@biomed.cas.cz; Stanislav Kmoch* - skmoch@LF1.cuni.cz

* Corresponding author

Published: 25 January 2008

BMC Genomics 2008, **9**:38 doi:10.1186/1471-2164-9-38

Received: 6 September 2007
Accepted: 25 January 2008

This article is available from: http://www.biomedcentral.com/1471-2164/9/38

Abstract

Background: To strengthen research and differential diagnostics of mitochondrial disorders, we constructed and validated an oligonucleotide microarray (h-MitoArray) allowing expression analysis of 1632 human genes involved in mitochondrial biology, cell cycle regulation, signal transduction and apoptosis. Using h-MitoArray we analyzed gene expression profiles in 9 control and 13 fibroblast cell lines from patients with F_1F_o ATP synthase deficiency consisting of 2 patients with mt9205ΔTA microdeletion and a genetically heterogeneous group of 11 patients with not yet characterized nuclear defects. Analysing gene expression profiles, we attempted to classify patients into expected defect specific subgroups, and subsequently reveal group specific compensatory changes, identify potential phenotype causing pathways and define candidate disease causing genes.

Results: Molecular studies, in combination with unsupervised clustering methods, defined three subgroups of patient cell lines – M group with mtDNA mutation and N1 and N2 groups with nuclear defect. Comparison of expression profiles and functional annotation, gene enrichment and pathway analyses of differentially expressed genes revealed in the M group a transcription profile suggestive of synchronized suppression of mitochondrial biogenesis and G1/S arrest. The N1 group showed elevated expression of complex I and reduced expression of complexes III, V, and V-type

ATP synthase subunit genes, reduced expression of genes involved in phosphorylation dependent signaling along MAPK, Jak-STAT, JNK, and p38 MAP kinase pathways, signs of activated apoptosis and oxidative stress resembling phenotype of premature senescent fibroblasts. No specific functionally meaningful changes, except of signs of activated apoptosis, were detected in the N2 group. Evaluation of individual gene expression profiles confirmed already known *ATP6*/*ATP8* defect in patients from the M group and indicated several candidate disease causing genes for nuclear defects.

Conclusion: Our analysis showed that deficiency in the ATP synthase protein complex amount is generally accompanied by only minor changes in expression of ATP synthase related genes. It also suggested that the site (mtDNA vs nuclear DNA) and the severity (ATP synthase content) of the underlying defect have diverse effects on cellular gene expression phenotypes, which warrants further investigation of cell cycle regulatory and signal transduction pathways in other OXPHOS disorders and related pharmacological models.

Background

Mitochondria generate most of the cellular energy in the form of ATP, regulate cellular redox state, cytosolic concentration of Ca^{2+}, are a source of endogenous reactive oxygen species, and integrate many of the signals for initiating apoptosis. By means of retrograde signaling mitochondria communicate all these events to the nucleus and thus modulate nuclear gene expression and cell cycle.

In humans, mitochondrial dysfunction leads to a vast array of pathologies, and hundreds of diseases result from various defects of mitochondrial biogenesis and maintenance, respiratory chain complexes, or individual mitochondrial proteins [1].

The most frequent group of mitochondrial diseases results from genetic defects of the oxidative phosphorylation system (OXPHOS) [2]. OXPHOS defects form a highly diverse group of diseases that affect primarily energy demanding tissues, such as the central nervous system, heart, and skeletal muscles. Their prevalence is estimated as at least 1:5000 [3]. About half of the OXPHOS defects result from mtDNA mutations [4]. Diseases resulting from mtDNA mutations usually show maternal mode of inheritance and variable penetrance of the disease phenotype, reflecting levels of mtDNA heteroplasmy and threshold effects in affected tissues. Remaining OXPHOS defects result from mutations in genes encoded in nuclear DNA. The majority of the nuclear encoded diseases are inherited as autosomal recessive traits and produce severe and usually fatal phenotypes in infants [5]. Up to now, mutations in approximately 50 nuclear genes have been identified, but most of nuclear genetic defects remain unknown and can involve any of approximately 1000 mitochondria related genes [6]. These genes play an essential role in the assembly or maintenance of individual OXPHOS complexes, in maintenance of mtDNA integrity, and mitochondrial biogenesis.

Diagnostic process of OXPHOS defects requires a combination of biochemical, enzymatic, immunohistochemical and molecular biology methods. To distinguish between isolated and combined OXPHOS deficiencies, the diagnostic process starts with measurements of selected mitochondrial enzyme activities and activities of individual OXPHOS complexes. The diagnostic procedure continues with analysis of OXPHOS complex protein composition. The origin of the molecular defect (mtDNA vs ncDNA) is often apparent from clinical presentation and family history. If not, it can be determined by using transmitochondrial cybrid cell analysis. Final steps in the diagnosis represent mutation analysis either in mtDNA or in nuclear encoded candidate genes in accordance with observed clinical and biochemical phenotypes. The diagnostic process is experimentally demanding and time-consuming and in majority of cases leads only to biochemical diagnosis. The molecular basis of the disease, especially in nuclear encoded defects, mostly remains unknown.

Identification of nuclear gene defects in OXPHOS deficiencies requires combination of positional cloning, functional complementation, and candidate gene analysis. Application of these "standard" procedures is however greatly hampered by limited number of affected patients, complexity and overlap of observed diseases phenotypes, difficulties in measurement of biochemical phenotypes *in vitro*, and by the existence of many candidate nuclear genes [7].

Another method having potential to contribute to differential diagnosis and research of OXPHOS defects relies on gene expression profiling. This type of analysis has a potential to provide information on putative diseases subtypes [8], suggest candidate disease causing genes [7,9,10], reveal pathogenic mechanism of the disease [11] and define specific gene expression profiles usable in future disease class prediction [12].

One of the possibilities for long term studies selectively targeted to mitochondrial gene expression analysis involves development and application of a focused microarray interrogating set of all known and hypothetical human mitochondrial genes, and several human mitochondria focused microarrays were prepared recently [13-16]. All these microarray platforms were based on PCR amplified probes prepared from selected IMAGE consortium cDNA clones. This approach however poses a number of technical obstacles. High rate of miss-annotation and contamination in the commercially distributed subset of the IMAGE Consortium cDNA clone collection [17] requires resequencing of individual clone inserts, and subsequent PCR preparation of individual probes is laborious and time consuming. Given these difficulties it has become very attractive to use sets of oligonucleotide probes that obviate much of the probe preparation work. Since the yield of long oligonucleotides has improved and cost has fallen recently, the current trend in preparation of low density, tailor-made microarrays favours oligonucleotide microarrays [18].

In this paper, we describe development and validation of a focused oligonucleotide microarray for expression profiling of human mitochondria related genes – "h-MitoArray" and report gene expression analysis of fibroblast cell lines from 9 controls and 13 patients with isolated deficiency of F_1F_o ATP synthase caused either by microdeletion of mtDNA encoded ATP6 gene [19,20] or by mutation of unknown nuclear genes [21,22].

Results

Microarray design and preparation

For microarray preparation we selected genes coding for known or predicted mitochondrial proteins, genes known to be involved in cell cycle growth and regulation, and genes involved in apoptosis and free radical metabolism.

The final set contained 1632 genes, of which 992 are "mitochondrial" genes, 42 lysosomal genes, 277 genes are associated with apoptosis, and 321 are "oncogenes". For normalization and background correction we included 146 human "housekeeping" genes, 10 *Arabidopsis* genes and 32 blanks. Full list of selected genes with corresponding symbols, accession and LocusLink codes is provided [see Additional file 1]. Functional annotation of selected genes and comparison of the gene content against whole human genome set is provided [see Additional file 2].

Microarray validation

Hybridization properties and performance of designed oligonucleotide probes and control features placed on h-MitoArray were tested by hybridization of fluorescently labeled panomers and fluorescently labeled cDNA prepared from a pool of total RNA isolated from several cell lines (test RNA). Gene expression signal was detected in > 77% of 1820 elements when fluorescently labeled cDNA pool was used.

Following comparison of various labeling strategies and optimization of hybridization conditions a series of self-to-self experiments was performed using test RNA. Data analysis showed acceptable reproducibility with Pearson correlation coefficient ranging 0.987 – 0.991.

Gene expression analysis in ATP synthase deficient fibroblasts

Fluorescent cDNA probes labeled with Cy5 were prepared from 13 patient and 9 control cell lines and were hybridized to common reference cDNA probe labeled with Cy3 in two technical replicates for each sample. Following data acquisition, transformation, normalization and replicate averaging, gene expression signals were obtained for 1264 genes. Ratios of Log2 sample gene intensities against Log2 common reference gene intensities (M) were calculated and are provided [see Additional file 3]. Calculated ratios of individual patient Log2 gene intensities against the Log2 of average of controls gene intensities (M) are provided [see Additional file 4].

Principal component analysis

To assess overall data quality and visualize relations between analyzed samples, we removed from the original data set 47 genes showing low expression variability (based on criteria $|Mmin; Mmax| \leq 0.58$, less than 1-fold change across all the samples) and subjected resulting data set to principal components analysis. Visual inspection of resulting plots showed no gross differences among the individual samples but suggested that several samples from nuclear defect patients group might be distinct from the others (Figure 1B).

Hierarchical clustering

To reveal gene expression changes, survey variation in patient samples, and better interpret the results of principal component analysis (PCA), gene expression signals from individual patient samples were compared to average of gene expression signals from all controls. Hierarchical clustering of all gene ratios across all patient samples was performed using Euclidean distance metrics and average linkage clustering algorithm. Resulting expression map (not shown) and sample dendrogram shown in Figure 1A defined, in agreement with previous PCA, two distinct subgroups of patients with nuclear defect, (N1 and N2 group) which were considered in subsequent gene expression comparisons and functional evaluations.

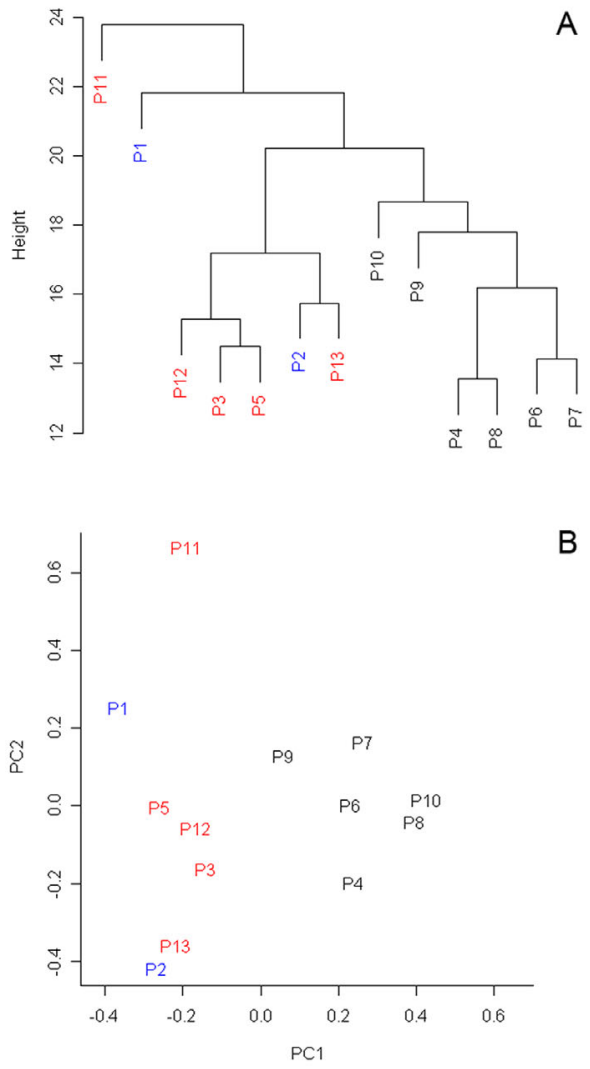


Figure 1
Results of unsupervised clustering methods. **A)** Dendrogram resulting from two-dimensional hierarchical clustering of all genes across all patient samples performed using Euclidean distance metrics and average linkage clustering algorithm. **B)** Two-dimensional PCA plot of all expression data showing the separation of samples forming N1 group. Patients from M, N1 and N2 groups are shown in blue, black and red, respectively.

Overall gene expression changes triggered by ATP synthase deficiency

Comparison of gene expression patterns between ATP synthase deficient and control fibroblast cell lines was performed in R statistical environment as described in methods. This analysis revealed 78 genes to be differentially expressed at adjusted P < 0.01 significance level [see Additional file 5]. Detailed inspection of expression map and evaluation of individual gene expression profiles showed, that although defined as significant, majority of the identified genes was not uniformly altered across all the patient samples.

Identification of subgroup specific gene expression profiles

To identify the subgroup specific gene expression changes, the subgroups of patients defined by a mutation of the *MTATP6* gene of the mtDNA (M group), PCA and hierarchical clustering (N1 and N2 groups) were compared. ANOVA analysis performed in MeV software revealed 97 genes to be differentially expressed at unadjusted P < 0.01 (Figure 2), [see Additional file 6].

Inspection of resulting data showed that the M group was specifically characterized by reduced expression of mitochondria encoded ATP synthase subunit genes *MTATP6*, *MTATP8*, nuclear encoded ATP synthase assembly factor *ATPAF1*, cytochrome *c* oxidase subunit II gene *MTCO2*, mitochondrial transcription factors *TFAM* and *TFB1M*, peroxisome proliferator-activated receptor alpha (*PPARA*), regulatory genes *H2AFX*, *CCNB1*, *C11orf13 (RASSF7)*, *TPR* and *ACO2*. This was accompanied by induction of *NRF1*.

The N1 group was characterized by reduced expression of genes involved in cell growth, differentiation and transduction pathways (*FOS*, *NOV*, *MAGED1*, *IL15RA*, *RARRES3*, *CTSK*, *UPLC1*, *PIM1*), mitochondrial proteosynthesis (*MRPS5*), lysosomal metabolism and function (cathepsins *S*, *K* and *D*, *GBA*, *PPGB*, *NPC*, *CLN2*, *FUCA1*, *HEXB*), protein transport (*AP2A1*), protein phosphorylation (*CDK5*, *PPAP2A*), hydrolase activity (*LIPA*, *LYPLA3*), reactive oxygen species metabolism (*GPX4*) and membrane transport (*SLC17A5*, *CTNS*). This was accompanied by elevated expression of several cell cycle regulatory genes such *WNT5A*, *IL3*, *CSNK1A1*, *BID*, *EIF4A1*, and *ACO2*.

The N2 group showed reduced expression of *WNT5A*, *EMP2*, *ADK*, *MDH2*, *SMAC* and elevated expression of *PPARG* and *GLS*. Extent and range of detected changes were much less than that observed in M and N1 groups.

Following ANOVA analysis, which revealed only intergroup specific differences, a list of group specific gene expression changes was obtained by comparison between defined patient subgroups and controls in R statistical environment as described in Methods. The analysis revealed 61, 215, and 54 genes to be differentially expressed at adjusted P < 0.01 in the M, N1 and N2 groups, respectively. In addition to the above mentioned genes revealed by ANOVA, we found in the M group elevated expression of *mitofusin* and coordinately reduced

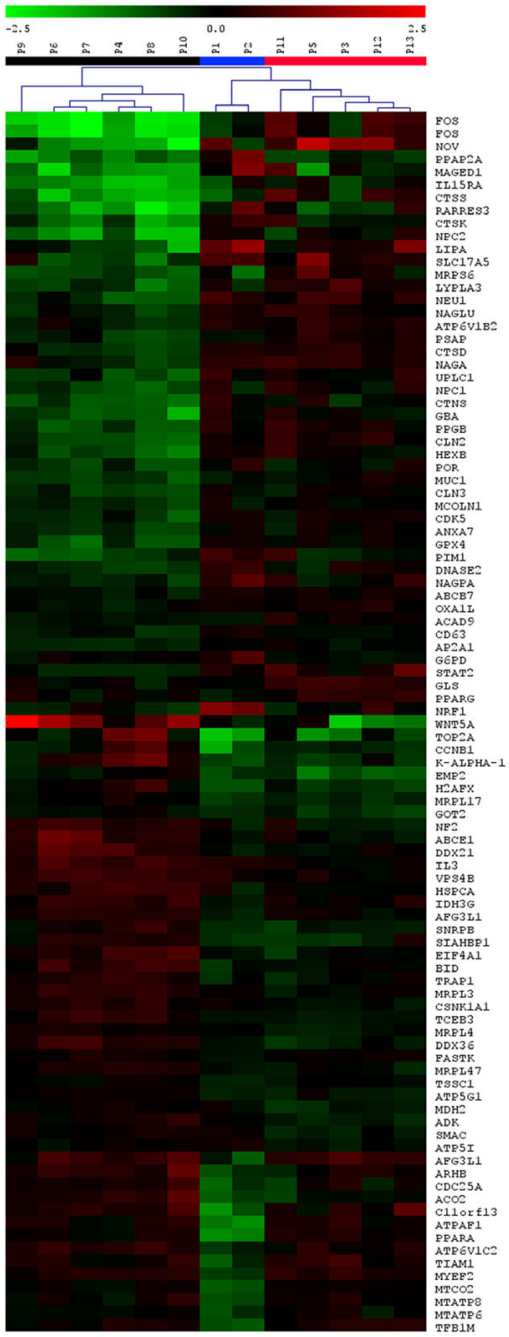


Figure 2
Differentially expressed genes defined by ANOVA analysis. Heatmap of genes detected as differentially expressed between defined patient groups using ANOVA analysis and unadjusted P < 0.01 significance level. The results are shown as Log2 ratio of relative gene expression signal in each patient sample to average of this of control samples. Ratio values are represented as the pseudo-color whose scale is shown in corresponding lookup picture.

expression of genes regulating G1/S phase transitions (*E2F1*, *MYC*, *CDC2*, *GAS1*, *CCNA2*, *CCNB*, *CDK2*, *CDC25A*, *PCNA*), thymidine metabolism (*TK*, *TYMS*) and DNA topology (*H2AFX*, *TOP2*, *LMN2*). In the N1 group, we observed reduced expression of genes regulating cell growth and signaling (*JUNB*, *MAPK3*, *WT1*, *CEBPA*, *CEBPB*) and lysosomal metabolism. We found elevated expression in genes involved in apoptosis (*FAS*, *CYTC*, *SMAC*, *IGFBP3*). In the N2 group, we found signs of started apoptosis (*SMAC*, *CASP8*). Group specific gene lists with expression values and corresponding P-statistics are provided [see Additional file 7, 8, 9].

Biological consequences of identified gene expression changes

To reveal biological consequences and to identify pathways potentially involved in the pathogenesis of the studied defects, we extracted from original expression data for each of the three defined groups all genes found to be differentially expressed at unadjusted $P < 0.05$ and showing expression change $|M| > 0.2$. Resulting expression datasets were uploaded into the DAVID database [23] and gene enrichment analysis was performed against h-MitoArray gene list. Results are provided in Table 1.

As the enrichment analysis suggested group specific dysregulation of several metabolic and signaling pathways, we further uploaded identical datasets into KEGGArray software (KEGG pathway databases – Kyoto Encyclopedia of Genes and Genomes) and inspected gene expression changes in all the indicated pathways.

In the M group, generally reduced expression was observed in cell cycle regulation (Figure 3), Krebs cycle (*OGDH*, *IDH1*, *ACO2*) and gluconeogenesis (*ALDOA*, *LDHA*, *PGAM1*) pathways. With an exception of *MTATP6*, *MTATP8* and *MTCOX2*, no multiple changes in OXPHOS system, valine, leucine, isoleucine, lysine, β-oxidation and MAP kinase pathway were observed. Reduced expression of *CytC* and *NFκB* and elevated expression of *FAS* were detected in the apoptotic pathway. In contrast to the N1 group, elevated expression of genes involved in N-glycan and heparan sulfate was detected.

In the N1 group, the analysis revealed elevated expression of several complex I subunit genes (*ND1*, *ND2*, *ND4*, *ND4L*, *Ndufs1*, *Ndufv2*, *Nufa9*, *Ndufb9 and Ndufa10*) and generally reduced expression of complex IV (*COX4*, *COX5A*, *COX6A*, *COX6B*, *COX6C* and *COX15*) and complex V subunit genes (*ATPAF1*, *ATP5G2*) in OXPHOS system. Generally reduced expression of V-type ATP synthase subunit genes was observed. Elevated transcription activity was found along valine, leucine, isoleucine, lysine and fatty acid β-oxidation pathways. Elevated expression of *FGF*, *FGFR*, *Ras* and *PKC* and reduced expression of *Raf1*,

Table 1: Functional annotation of defined patient subgroups.

M			**N1**			**N2**		
category	n	p	category	n	p	category	n	p
DAVID IDs								
	258			383			238	
Biological processes								
	230			344			203	
DNA replication	13	5E-3	endodome transport	7	1E-3	development	38	9E-3
taxis	9	7E-3	vacuole organization and biogenesis	7	9E-3	reactive oxygen species metabolism	6	1E-2
carbohydrate metabolism	25	9E-3	response to chemical stimuli	23	1E-2	response to oxidative stress	5	3E-2
negative regulation of biological processes	26	1E-2	regulation of enzyme activity	20	3E-2	dephosphorylation	6	2E-2
nucleic acid metabolism	69	2E-2	vesicle mediated transport	18	3E-2	intracellular protein transport	18	3E-2
Molecular function								
	238			347			211	
DNA binding	39	2E-2	protein dimerization activity	14	2E-2	protein domain specific binding	6	3E-2
protein dimerization activity	10	5E-2	hydrolase activity on glycosyl bonds	12	5E-2	GTPase activity	7	5E-2
nucleic acid binding	54	5E-2						
Cellular component								
	285			342			195	
chromosome	11	2E-3	vacuole	44	2E-8	chromosome	8	4E-2
chromatin	7	9E-3	lytic vacuole	39	2E-7			
nucleus	70	6E-3	lysosome	39	1E-7	lytic vacuole	17	4E-2
lytic vacuole	20	2E-3	extracellular region	32	2E-2	lysosome	17	4E-2
lysosome	20	2E-3	endosome	8	4E-2	non-membrane bound organelle	28	4E-2
KEGG pathway								
	122			185			125	
N-glycan degradation	5	2E-2	antigen processing	9	2E-3	Toll-like receptor signaling	10	2E-2
hematopoetic cell lineage	8	3E-2	glycosphingolipid metabolism	7	2E-2	glycosylaminoglycan degradation	6	2E-2
			hematopoetic cell lineage	10	4E-2			
Biocarta pathway								
	68			92			57	
cyclins and cell cycle regulation	9	2E-2	role of ERB2 in signal transduction	9	5E-3	activation of Src	4	3E-2
			IL 3 signaling pathway	7	1E-2	phospholipid signaling intermediates	5	4E-2
			IL 6 signaling pathway	8	1E-2			
			Erk and PI-3 kinase pathway	7	2E-2			
			signaling pathway from G-protein families	7	3E-2			

"n", number of genes involved in the corresponding annotation category; p, modified Fisher exact p-value of the gene enrichment for each category.

MEF1, *ERK*, *Elk1* and *FOS* were found in classical MAP kinase pathway (Figure 4A). Reduced expression of *IL1*, *IL1R*, *AKT*, *Elk1*, *GADD153* and *JunD* with elevated expression of *p53*, *p38* and *Evi1* were found in JNK and p38 MAP kinase pathways (Figure 4A). Reduced expression of *STAT*, *CPB*, *Pim-1*, *AKT* and *BclXL* and elevated expression of *IL2/3* and *IL3R* were found in Jak-STAT signaling pathway. Elevated expression of *Bid* and *CytC* with reduced expression of *Bcl-2/XL* and *CASP9* were detected in the apoptotic pathway. General decrease in expression of genes involved in N-glycan, glycosylaminoglycan, and ganglioside degradation was found. In conjunction with 3-meth-

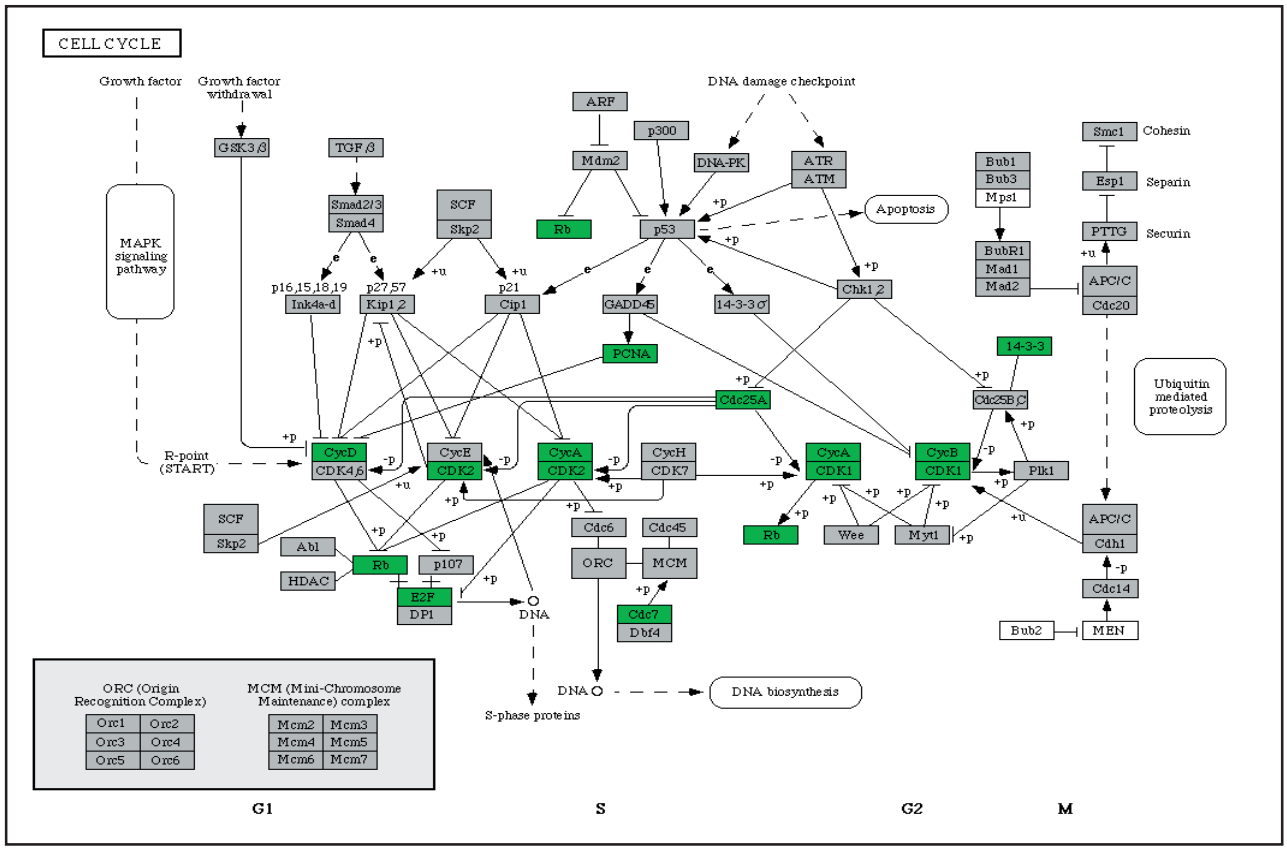


Figure 3
Gene expression changes detected in selected pathways in M group. General changes in cell cycle pathway detected in patients with mtDNA mutation (M group) using KEGGArray software.

ylglutaconic aciduria, which is a characteristic biochemical feature of the patients from this group, inspection of leucine degradation pathway showed moderately reduced expression of 3-methylglutaconyl-CoA hydratase gene, *AUH*, (Figure 4B). Although the extent of the *AUH* expression changes neither directly implicates the deficiency of 3-methylglutaconyl-CoA hydratase nor explains 3-methylglutaconic aciduria present in these patients, it is possible that such changes might be much more pronounced and have functional effects during metabolic stress and/or in metabolically active tissues. In the N2 group, reduced expression of *GRB2*, *RAS* and *ERK* and elevated expression of *FOS*, *JUND* and *Evi1* was found in MAP kinase pathway. This was accompanied by elevated expression of genes involved in N-glycan, glycosylaminoglycan and ganglioside degradation. No multiple changes in the apoptotic and valine, leucine, isoleucine, lysine, β-oxidation degradation pathways were found. All mentioned pathways and gene expression changes identified by KEGGArray software are provided [see Additional file 10 and 11].

Identification of patient specific gene expression profiles and definition of candidate disease causing genes

To get specific information on patient mitochondrial genome expression, we extracted and clustered gene expression data for all 37 mtDNA genes. Resulting mitochondrial genome expression map (Figure 5A) reflects relative mitochondrial DNA amount with generally elevated expression in P11, P3, P10 and P6 and generally reduced expression in P2, P4 and P8. Specific gene expression changes were detected in P1 and P2, where the expression map revealed reduced amount of *MTATP6*, *MTATP8* and *MTCOXII* transcript reflecting disease causing microdeletion of *MTATP6*, and in P12 with specifically reduced expression of *tRNAGly*.

To obtain the information on patient specific ATP synthase complex expression, we extracted and clustered gene expression data for all of its structural genes and assembly factors. Resulting expression map is provided in Figure 5B. In P1 and P2, it shows reduced expression of mitochondrial subunits *MTATP6*, *MTATP8* and also of *ATPAF1*.

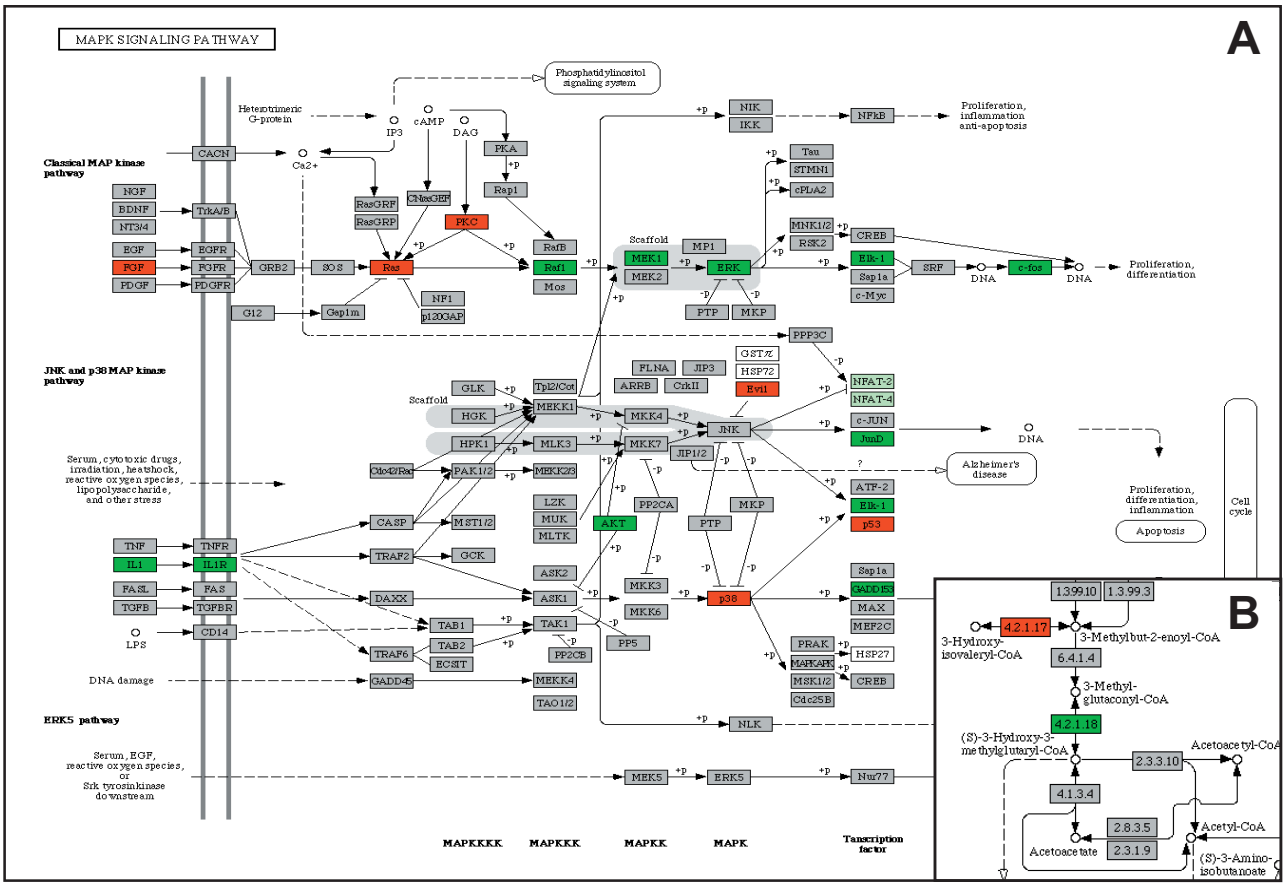


Figure 4
Gene expression changes detected in selected pathways in N1 group. A) Changes in MAPK, JNK and p38 MAP kinase pathways, **B)** reduced expression of AUH, 3-methylglutaconyl-CoA hydratase gene in leucine degradation pathway, detected in patients with nuclear defect (N1 group) using KEGGArray software.

With the exception of reduced expression of *ATP5G2* in P5 and maybe also *ATP5C1* in P3 no additional subgroup and/or patient specific profile were found.

To define potential candidate disease genes, we finally compared gene expression data of individual patients with a group of controls in R statistical environment as described in methods, and searched for genes showing significantly reduced expression and having known function either in ATP synthase biogenesis, mitochondrial protein trafficking or mitochondrial biogenesis. In P1 and P2, we detected reduced expression of ATP synthase structural subunits *MTATP6*, *MTATP8* and also of *ATPAF1*. In P3 we detected reduced expression of *ATP5C1* and *ATP5O*. In P4 and P8 we detected reduced expression of *TOM 7*. Mitochondrial carrier homolog 1 (*C. elegans*) (*MTCH1*) transcript was reduced in P6, P10, P11 and P12. In P10 we detected reduced expression of mitochondrial elongation factor *EFG1* and *TOM22*. In P11 we found reduced expression of *TIM23*, *TIM8* and *TOM34* homologs, *ATP5H* and ATP5E. In P12 we found reduced expression of *mitofusin* and *ATP5H*. The lists of all the differentially expressed genes are shown [see Additional file 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24].

Confirmation of the hybridization results

To rule-out platform specific bias, we re-analyzed all RNA samples from the N1 and control groups using the same common reference RNA on Agilent 44 k arrays. We used available annotations and extracted from the Agilent data gene expression values for the genes identified as significantly ($P < 0.05$), differentially expressed in the N1 group on our platform. Correlation coefficient of expression values of 102 identified genes was 0.925.

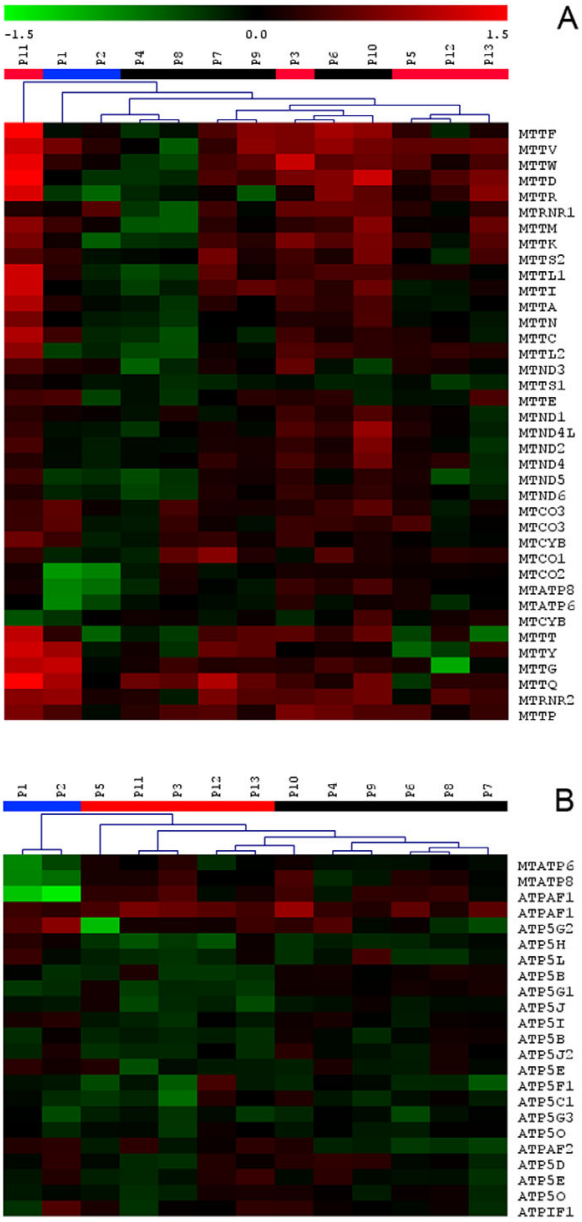


Figure 5
Two-dimensional hierarchical clustering of patient samples. A) Expression matrix of all 37 mtDNA encoded genes. **B)** Expression matrix of structural and assembly factor genes involved in ATP synthase complex biogenesis. Selected genes were clustered across all patient samples using Euclidean distance metrics and average linkage clustering algorithm.

Correlation of expression data with available RT-PCR and Western blot results

ATP synthase deficiency of nuclear genetic origin is characterized at the protein level by pronounced decrease of the individual subunits and the mature ATP synthase protein complex amounts (Table 2). However, our data in patient cell lines, in agreement with previous Q-PCR analyses, did not show pronounced alterations in ATP synthase subunits or of ATP synthase-specific assembly factors mRNA levels that could explain it easily. Only in the M group, the data showed decrease of *MTATP6* and *MTATP8* mRNA levels which correspond with previously performed Northern blot and Q-PCR analysis showing that this mutation affects processing of *ATP8/ATP6/COX-III* polycistronic transcript and results in decreased levels and/or stability of mature *ATP8/ATP6* mRNA [19,24]. Many of mitochondrial diseases are associated with compensatory changes in the cellular content of mitochondria and/or the content of one or more OXPHOS complexes. Western blot analysis of fibroblasts with ATP synthase deficiency has previously shown increased mitochondrial content of complex I and complex III [25]. In agreement with this observation, our data showed elevated expression of complex I subunit genes in N1 group. Expression of complex III subunit genes was however decreased. Parallel analyses of the fibroblasts with nuclear ATP synthase defects used in this study revealed variable changes in fibroblast COX and/or SDH specific content (Table 2). These changes were not associated with generally elevated expression of COX and SDH subunit genes. Detailed inspection of individual gene expression profiles [see Additional file 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24] however suggested that elevated expression of *COX7A2L* and *SDHA* may correlate with this observation (P3, P5, P6, P8, P10).

Discussion

Platform selection and evaluation

In our work, we attempted to set up an experimental platform which will allow in a cost-effective way prospective gene expression analysis of cell lines and tissues from patients with various genetically determined OXPHOS defects. We considered availability of biological materials for the analysis, estimated the number of informative genes, evaluated gene content of commercially available microarrays and took into account instrumentation availability and platform related running costs. Since cultured skin fibroblasts are the most accessible, relatively well standardized, and multiple analysis amendable source for gene expression analysis especially in nuclear encoded OXPHOS defects, we estimated (based on Gene Expression Omnibus database data), that in fibroblasts, reliable expression signal may be obtained for approximately 6000 (HG-U95 array) to 10000 (HG-U133 Plus 2.0 Array) genes, of which only part may be meaningful to detect and understand anticipated changes in mitochondrial biology and related basic cell responses. In addition, we also evaluated representation of mitochondria encoded genes on available whole genome arrays. We found that

Table 2: Clinical, biochemical and molecular description of patients (P1 – P13).

Patient (group)	Phenotype	Biochemical data	Genetic defect	ATPase (% of C)	SDH (% of C)	COX (% of C)	Ref.
P1 (M)	PMR, encephalomyopathy, spastic quadruparesis, microcephalia,	lactate: 1.0–3.4 3 MGA: <15	mt9205ΔTA	*80–120	120–200	80–120	[19]
P2 (M)	transient lactic acidosis, nystagmus, GR	lactate: 3.9–10	mt9205ΔTA	*80–120	80–120	80–120	[20]
P3 (N2)	PMR, HCMP, hypotonia, peripheral neuropathy,	lactate: 1.4–10 3 MGA: 133–281	ncDNA, unknown	<30	120–200	120–200	[21]
P4 (N1)	Fatal lactic acidosis, HCMP	lactate: 30–36	ncDNA, unknown	<30	120–200	80–120	[82]
P5 (N2)	PMR, HCMP, hypotonia, dysmorphy, microcephaly	lactate: 1.6–8 3 MGA: 22–225	ncDNA, unknown	<30	>200	>200	[21]
P6 (N1)	PMR, HCMP, hypotonia, dysmorphy, microcephaly	lactate: 3.6–4.5 3 MGA: 28–260	ncDNA, unknown	<30	>200	>200	NR
P7 (N1)	PMR, HCMP, hypotonia, dysmorphy, microcephaly, epilepsy	lactate: 2.2–6.0 3 MGA: 28–161	ncDNA, unknown	<30	80–120	120–200	NR
P8 (N1)	PMR, hypotonia, dysmorphy, microcephaly	lactate: 3.6–6.7 3 MGA: 56–252	ncDNA, unknown	<30	120–200	>200	NR
P9 (N1)	PMR, hypotonia, dysmorphy, microcephaly	lactate: 2.2–10 3 MGA: 62–150	ncDNA, unknown	<30	>200	>200	[21]
P10 (N1)	PMR, hypotonia, dysmorphy, microcephaly	lactate: 1.4–4.6 3 MGA: 64–270	ncDNA, unknown	<30	120–200	120–200	NR
P11 (N2)	PMR, hypotonia, GR, HCMP dysmorphy, microcephaly	lactate: 1.5–8.2 3 MGA: 34–254	ncDNA, unknown	<10	80–120	80–120	[21]
P12 (N2)	PMR, hypotonia, HCMP	lactate: 2–6.0 3 MGA: 115–460	ncDNA, unknown	<10	80–120	80–120	[25]
P13 (N2)	PMR, GR, microcephaly, mild spasticity, hepatopathy	lactate: 1.2–3.9 3 MGA: 37–132	ncDNA, unknown	<30	120–200	120–200	[21]

Patient assignment to groups is based on DNA sequencing data (M) and results of PCA and hierarchical clustering (N1, N2). PMR – psychomotor retardation, HCMP – hypertrophic cardiomyopathy, GR – growth retardation, lactate – blood lactate (mmol/l), 3 MGA – 3-methylglutaconic aciduria (mg/g creatinine). ATPase (complex V), SDH (complex II) and COX (complex IV) represent enzyme protein content in fibroblast homogenates quantified by SDS PAGE/WB as in [19], using specific primary antibodies (MitoSciences, OR), Alexa Fluor® 680-labeled secondary antibodies and an Odyssey® Infrared Imaging System (LI-COR Biotechnology, Lincoln, NE). Data are presented as % of control values. * Decreased content of subunit a (ATP6). NR means not reported.

(at the time of project planning) no complete coverage of mitochondrial tRNA, rRNA and OXPHOS structural subunits have been available on() Affymetrix HG_U95Av2 Array (contained just *TRNC, TRNY* and *TRNS1*), HG-U133 Plus 2.0 Array (no tRNAs, rRNAs and *ND1*, *ND4L*, *CYTB*) and Agilent 44 k Array (no tRNAs, rRNAs and *ND4L*). Considering this data and also available instrumentation, we then decided to construct focused an oligonucleotide microarray and employ competitive two-color hybridization approach with common reference experimental design.

Selected gene content allows gene expression analysis of the entire mitochondrial genome and almost all of "mitochondria" related genes in context of key DNA synthesis, growth response, regulatory and apoptotic genes. Hybridization signal was obtained from 78% of the designed oligonucleotides. Vast majority of the oligonucleotides giving no hybridization signal were designed to detect regulatory genes and transcription factor transcripts probably not transcribed in the analyzed materials. Interestingly, we detected hybridization signals for almost all mitochondrial tRNA and rRNA probes which is, in respect to oligo-dT labeling strategy, suggestive that all those transcript are also at least partially polyadenylated [26].

Gene expression analysis in patients with defect of F_1F_o ATP synthase

In the work presented herein, we analyzed and compared gene expression profiles in fibroblast cell lines from 9 control individuals and 13 patients with biochemically proven but genetically heterogeneous F_1F_o ATP synthase deficiency. We aimed to identify gene expression changes indicating how affected cells react to and compensate for the common biochemical defect, use gene expression data to assign patients into already defined and/or putative disease subgroups, identify candidate disease causing genes, and define potential pathogenetic mechanisms associated with the disease.

The magnitude of observed expression changes was moderate with only several dozens of genes exceeding 2-fold changes. Comparing all the patient cell lines with all control cell lines, we have not identified any common and meaningful gene expression changes attributable to ATP synthase deficiency *per se*. It has been suggested recently that the degree and compartmentalization of ATP deple-

tion may be defect specific and may thus have also specific biological consequences [27]. Our data support this view.

Cell lines with mtDNA mutation (M group) showed gene expression changes suggestive of suppressed mitochondrial biogenesis and metabolism characterized by down regulation of *TFAM* and *TFB1M*, master regulators of mitochondrial transcription, accompanied by reduced expression of other mitochondria encoded transcripts (*MTCO2*, *MTATP6*, *MTATP8*, and *MTND6*), reduced expression of *ATPAF1*, *E2F1*, *ACO2* (component of mitochondria to nucleus retrograde pathway) and *PPARA*. This "mitochondria silencing" activity seems to be sensed and counterbalanced by elevated expression of *NRF1*, which is however not accompanied by expression changes of any NRF-1 target and/or coactivator genes [28,29]. Inhibition of mitochondrial biogenesis is synchronized with reduced expression of genes regulating the G1/S phase transition (*E2F1*, *MYC*, *Rb*, *CycA*, *CycD*, *CDK2*, *Cdc7*, *Cdc25A*, *PCNA*) and associated thymidine metabolism (*TK*, *TYMS*) [30]. We interpret this gene expression pattern as an ATP depletion mediated G1/S arrest [31] associated with synchronized replication arrest of mitochondrial genome [32] and repression of NRF-1 activity [33]. Our observations are quite similar to that made in *Drosophila* mutants, in which low ATP levels lead to arrest in the G1 phase without affecting cellular differentiation and cell viability [34,35]. Furthermore, our observation conforms to the view that mitochondria co-regulate cell cycle progression and that this regulation is executed not only at posttranscriptional [34] but also at transcriptional level.

The N1 group differed from M group in that it showed very minor signs of mitochondrial response suggested only by slightly elevated expression of *PPGC-1*, *TFAM*, *TFB2M* and *ACO2*. More significant and distinct changes were however observed in signal transduction pathways regulating mitochondrial oxidative phosphorylation [36]. The gene expression portrait, reduced expression of many transcription factors and cytokines regulating cell growth and differentiation (*FOS* [37], *JUNB* and *MAPK3* [38], *CEBPA* and *CEBPB*, *CXCL1* and *CXCL2* [39]), elevated expression of *IGFBP3* [40] and *CAV2* [41], together with activated apoptosis (*BCL2L1*, *SMAC*, *CYCS*, *FAF1*), signs of oxidative stress (*TR2*) [42] and general decrease in lysosomal activities [43], resemble characteristic signs of senescent fibroblasts [44,45]. However all the cell lines from the N1 group have originated from very young donors, all but one were in their early passages and all showed the same passage frequency of 5–6 days, (Table 3). It has been shown that inhibition of oxidative phosphorylation may play an active role in the process of cellular senescence in human fibroblasts [46], and that changes in transcription activity may be governed by changes in protein phosphorylation [47]. We therefore interpret the observed gene expression pattern as accelerated stress induced premature senescence phenotype resulting from impaired oxidative phosphorylation and profoundly reduced ATP availability for critical energy-dependent cellular processes. Our explanation of N1 cellular phenotype is the following. Mitochondrial ATP synthesis is markedly decreased in fibroblasts derived from patients with nuclear DNA-related disorders but only variably so in patients with mtDNA mutations [48]. ATP depletion is sensed by AMP-activated protein kinase which acts as a metabolic sensor or "fuel gauge" that monitors cellular AMP and ATP levels [49]. Once activated, the enzyme switches off ATP-consuming anabolic pathways and switches on ATP-producing catabolic pathways [50], such as fatty acid oxidation (elevated expression of *ECH1*, *ECHS1*, *ETFDH*, *CABC1*) and amino acid catabolism. Despite this compensatory effort, mitochondrial ATP depletion persists due to intrinsic ATP synthase defect, activation of AMPK persist and leads to accelerated p53-dependent cellular senescence [51]. AMPK activity also leads to decrease of HuR cytosolic translocation, which influences the mRNA-stabilizing function of HuR [52] and diminishes the expression and half-lives of HuR target transcripts, such as *FOS* [53] or *CDKN1A* [54] which also leads to the premature senescence phenotype [55]. ATP availability probably modulates cytoplasmic translocation and recruitment of other RNA-binding proteins stabilizing various mRNAs [56]. In this context it is interesting that we have detected reduced expression (or transcript abundance) of two RNA-binding protein genes *CUGBP1* and *AUH*. CUGBP1 affects translation of *CDKN1A* [57] and *CEBPB* [58], and our data show decrease in those two transcripts as well. AUH stabilizes *FOS* and other immediate early mRNA's [59], and its deficiency is also causing methylglutaconic aciduria [60], a characteristic biochemical phenotype observed specifically in this group of nuclear encoded ATP synthase deficient patients [21] (Table 2).

Table 3: Growth characteristics of the fibroblast cell lines.

	patients									controls										
	1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3	4	5	6	7	8	9
passage number	17	19	15	4	20	9	6	6	28	12	28	17	12	22	22	14	27	16	17	11	16	13
passage frequency (days)	5	5	6	6	6	6	6	6	5	6	6	7	7	9	3	3	4	4	5	4	4	7

Resulting transcriptional silencing and other ATP depletion mediated disturbances of intracellular signal transduction cascades lead thus to premature senescence phenotype, reduced proteasome activity and accumulation of oxidized proteins, which may explain observed discrepancies between gene expression and Western blot data. Patients forming this group are of common ethnic origin, and this is suggestive that common genetic defect may underlie this specific gene expression profile.

The N2 group showed neither signs of mitochondria response observed in the M group, nor signs of premature senescence observed in the N1 group. Expression profile is suggestive of partly activated apoptosis (*SMAC*) and disturbances of intracellular signaling transduction cascades (down regulation of several cytokines, early genes, and regulatory proteins). However, all these changes were not uniformly present in all cell lines, which together with variability in clinical and biochemical data is suggestive of further genetic heterogeneity within this group of patients.

Selection of candidate disease causing genes

As gene expression changes may be used for selection of candidate disease causing genes [9,61], we evaluated group specific and individual gene expression profiles. This approach was successful in both patients from the M group, in whom detected alterations clearly indicated the involvement of ATP6/ATP8/COXIII transcript. In other patients we first focused on expression of ATP synthase subunits. Inspection of this expression profile (Figure 5B) suggested involvement of ATP synthase assembly factor ATPAF2 in several patients from N1 group. Mutation of *ATPAF2* has been found in the case with ATP synthase deficiency [62] and this warrant sequence analysis of this gene in this group of patients. Other candidate genes may be *ATP5G2*, the expression of which is decreased in P5 and possibly also *ATP5C1* found lowered in P3. From other genes, no clear candidates for immediate sequence analysis may be defined yet. However, more focused interpretation will be possible once candidate disease genomic intervals are defined by ongoing linkage studies.

Conclusion

We designed, produced, and validated an oligonucleotide microarray focused on expression profiling of human mitochondria related genes, and searched for gene expression changes in genetically heterogeneous group of 13 patients with F_1F_o ATP synthase deficiency. The analysis classified patients into three distinct groups and suggested that site (mtDNA vs nucleus) and severity (residual content of ATP synthase) of underlying biochemical defect have diverse effects on cell gene expression phenotype. Comparisons with controls, between defined groups and among individual patient cell lines did not show any uniform transcription changes explaining pronounced decrease in ATP synthase content and alterations of the other OXPHOS complexes observed at the protein level. The analysis nevertheless confirmed the already known and indicated candidate disease causing genes, and suggested that defects in ATP synthesis lead to deregulation of signal transduction pathways and affect mitochondrial and nuclear DNA replication. These may be important pathogenic mechanisms involved not only in F_1F_o ATP synthase deficiency but also in other OXPHOS defects. Observed gene expression changes therefore warrant further investigation of major cell cycle regulatory and signal transduction pathways in other OXPHOS disorders and pharmacological models. Full potential of the constructed h-MitoArray platform will be further revealed in ongoing positional cloning studies in herein analyzed patients and in gene expression studies in other groups of OXPHOS deficient cell lines.

Methods

Database of human mitochondrial genes

Lists of "mitochondrial" and "mitochondria related" genes were extracted and merged from various public databases such as Mitomap [63], Mitop [64], Migenes [65], Mitoproteom [66], Molecular Signature Database [67], OMIM, RefSeq and Unigene sections at NCBI [68], Gene Ontology database [69] and UniProt resource [70]. Full annotation of selected genes has been obtained and deposited in a locally installed database BASE [71].

Microarray preparation

For each of the selected 1632 genes, a single 5'-aminomodified 40-mer oligonucleotide was designed using Oligopicker software [72]. Blast searches were performed with each candidate probe to exclude possibility of cross hybridization with homologous genes prior to the synthesis of oligonucleotide probes. Synthesized oligonucleotides, Generi Biotech (Czech Republic) and Illumina (San Diego, CA), were resuspended at 20 μM concentration in 3 × SSC, printed in triplicates on aminosilane modified slides, and immobilised by standard technique using combination of baking and UV cross-link as previously described [61]. Qualities of arrays from individual printing series were assessed using fluorescently labelled panomers (Invitrogen, Carlsbad, CA).

Mixed RNA for microarray validation

As a standard for microarray optimisation, standardization and validation total RNA was isolated from HeLa G, ECV 304, 293, U 937, JURKAT and A 301 cell lines using the TRIZOL solution (Invitrogen, Carlsbad, CA). Isolated RNA samples were pooled, and aliquots were stored at -80°C until the analysis.

Reference RNA preparation

As a common reference RNA for gene expression studies, total RNA from cultured HeLa cells was chosen. Total RNA was extracted as above. Concentration was determined spectrophotometrically at A 260 by NanoDrop (NanoDrop Technologies, Wilmington, DE) and quality was checked on Agilent 2100 bioanalyser – RNA Lab-On-a-Chip (Agilent Technologies, Santa Clara, CA). Aliquots of isolated RNA were stored at -80°C until the analysis.

Control group

Selected control fibroblasts cell lines were used repeatedly in previous diagnostic biochemical tests and showed no signs of any mitochondrial or other metabolic defect.

Patients

Fibroblast cell lines from 13 patients were used in this study. All the patients showed major clinical symptoms associated with OXPHOS defect. Biochemical diagnosis of ATP synthase deficiency was based on absence or significant decrease of mature ATP synthase complex and of its subunits in electrophoretic analysis of OXPHOS complexes in cultured fibroblasts and other available tissues [73]. Mitochondrial genome sequencing performed in all patients revealed disease causing mitochondrial DNA mutations in two patients (P1, P2, M group) [19]. Molecular basis of defect in the other patients has not yet been defined. Relevant clinical, biochemical and molecular data and references on individual patients included in this study are provided in Table 2.

Cell culturing

Growth characteristics of the cell lines used in this study are provided in Table 3. Skin fibroblasts were cultured in the Dulbecco's modified Eagle's medium supplemented by 10% fetal calf serum, 20 mM HEPES pH 7.5, 0.2% $NaHCO_3$ and gentamycin 0.02 mg/ml at 37°C in a 5% CO_2 humidified atmosphere. For experiments, confluent cell were harvested using 0.05% trypsine and 0.02% EDTA. Detached cells were diluted in ice-cold culture medium, sedimented by centrifugation (600 g) and washed twice in phosphate buffered saline (140 mM NaCl, 5.4 mM KCl, 8 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.2).

RNA preparation, cDNA labeling and hybridization

Total RNA was extracted from cultured cells and QC controlled as described above.

Five μg of total RNA was reverse transcribed and labeled by Array 900 Expression Detection Kit (Genisphere, Hatfield, PA) according to the manufacturer protocol. The slides were pretreated by baking at 80°C, UV cross-linked and washed twice in 0.1% SDS for 2 minutes, twice in 0.2 × SSC for 2 min, four times in MilliQ water, followed by denaturation in boiling water for 2 minutes. Prehybridization was performed using hybridization buffer (Genisphere, Hatfield, PA) according to the manufacturer protocol. All hybridizations were performed in humid hybridization chamber, ArrayIt Hybridization Cassette chamber (TeleChem International, Sunnyvale, CA).

Microarray scanning

The hybridized slides were scanned with GenePix 4200A scanner (Axon Instruments, Union City, CA) with PMT gains adjusted to obtain highest intensity unsaturated images. GenePix Pro software (Axon Instruments, Union City, CA) was used for image analysis of the TIFF files, as generated by the scanner.

Experimental setup and data normalization

All 13 patient samples and 9 controls were hybridized to common reference (HeLa cell lines) in two replicates of each sample. All arrays were hybridized with a Cy5-labeled sample cDNA and a Cy3-labeled reference cDNA.

Expression data were obtained using GenPix Pro software. Comparative microarray analysis was performed according to MIAME guidelines [74]. Normalization was performed in R statistical environment [75] using Limma package [76] which is part of the Bioconductor project [77]. Raw data from individual arrays were processed using Loess normalization and normexp background correction. Gquantile function was used for normalization between arrays. The correlation between 3 replicate spots per gene on each array was used to increase the robustness. Linear model was fitted for each gene given a series of arrays using lmFit function. The empirical Bayes method was used to rank differential expression of genes using eBayes function. Multiple testing correction was performed using Benjamini & Hochberg method [78].

Quality control

Variation among feature replicates on the array was calculated by conversion of raw data to log-ratios. Data were further normalized using Loess function. Features with less than double background intensity (A < 8.5) were removed. For each feature on the array the deviation from the mean computed as the difference between the ratio of the feature and the mean of the set of feature replicates was calculated. Standard deviation of the error distribution using all of the replicates was calculated and converted to coefficient of variability using equation.

$$CV = \sqrt{\exp[(\ln 2 * SD2)^2] - 1}$$

The variability between the duplicate spots ranged from 8.1% to 27.5%. Arrays with variability higher then 18% were removed from the analysis.

Statistical analysis

Principal component analysis, hierarchical clustering, ANOVA and SAM analyses were performed in TIGR Multiexperiment Viewer (MeV), version 4.0 [79], available [80]. Significant gene expression changes between defined subgroups were identified using t-test in R statistical environment [75]. Applied parameters are provided in corresponding result sections.

Functional annotation

Functional annotation and pathway enrichment analysis was performed in DAVID (The Database for Annotation, Visualization and Integrated Discovery [23]). Visualization of gene expression changes along affected pathways was performed in KEGGArray software (KEGG pathway databases – Kyoto Encyclopedia of Genes and Genomes) [81].

Data accession

Description of h-MitoArray platform and gene expression data reported in this study are stored and available in Gene Expression Omnibus repository under accessions GPL5150 and GSE8648.

Ethics

The project was approved by the Scientific Ethics Committee of the 1st Faculty of Medicine of Charles University of Prague under reference NR/8069-3. Patient participation in the project was made on a voluntary basis after oral and written information and consent according to the Helsinki V Declaration.

Authors' contributions

A.Ж. tested, compared and optimized labeling and hybridization condition, performed all the RNA sample isolations, cDNA labeling, microarray hybridizations, data acquisition, and contributed to data analysis and interpretation. V.S performed all data analysis in R and MeV environments. R.I. designed oligonucleotide probes, updated functional annotation of selected gene set, and participated in data analysis. H.H., L.N. a L.P. set and kept BASE database, optimized methods for microarray manufacturing and prepared all microarrays used in this study. M.T, H. Han., T.H., J.Z., J.A.M., W.S., A.P., J.P. and J.H. performed all relevant clinical, biochemical, and molecular investigations in studied patients, and together with S.S. provided patient and control cell lines. S.K. conceived and coordinated the study, was involved together with J.H. in data analysis, result interpretation and manuscript preparation.

Additional material

Additional file 1

Annotation of h-MitoArray. List of selected genes with corresponding symbols, accession and LocusLink codes. Probes showing hybridization signal with fluorescently labeled panomers and fluorescently labeled cDNA are presented as "signal detected".
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S1.xls]

Additional file 2

Functional annotation of the h-MitoArray. Functional annotation of selected genes and comparison of the h-MitoArray gene content against whole human genome reference set.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S2.xls]

Additional file 3

Expression matrix – samples compared to common reference. Ratios of Log2 sample gene intensities against Log2 gene intensities of common reference.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S3.xls]

Additional file 4

Expression matrix – patients compared to controls. Ratios of individual patient Log2 gene intensities against the average of the Log2 controls intensities.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S4.xls]

Additional file 5

Differentially expressed genes between all patients and controls. List of genes detected as differentially expressed between all studied ATP synthase deficient and control fibroblast cell lines at adjusted $P < 0.01$ significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S5.xls]

Additional file 6

Characterization of defined patient groups using ANOVA analysis. List of genes detected as differentially expressed between defined patient groups using ANOVA analysis and unadjusted $P < 0.01$ significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S6.xls]

Additional file 7

Lists of differentially expressed genes in M group. Lists of genes detected as differentially expressed in M group, when compared to controls at adjusted $P < 0.01$ significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S7.xls]

Additional file 8
Lists of differentially expressed genes in N1 group. Lists of genes detected as differentially expressed in N1 group, when compared to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S8.xls]

Additional file 9
Lists of differentially expressed genes in N2 group. Lists of genes detected as differentially expressed in N2 group, when compared to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S9.xls]

Additional file 10
Pathway analysis in M group. Pathways and gene expression changes identified by KEGGArray software in M group.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S10.PDF]

Additional file 11
Pathway analysis in N1 group. Pathways and gene expression changes identified by KEGGArray software in N1 group.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S11.PDF]

Additional file 12
Differentially expressed genes in patient P1. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S12.xls]

Additional file 13
Differentially expressed genes in patient P2. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S13.xls]

Additional file 14
Differentially expressed genes in patient P3. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S14.xls]

Additional file 15
Differentially expressed genes in patient P4. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S15.xls]

Additional file 16
Differentially expressed genes in patient P5. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S16.xls]

Additional file 17
Differentially expressed genes in patient P6. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S17.xls]

Additional file 18
Differentially expressed genes in patient P7. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S18.xls]

Additional file 19
Differentially expressed genes in patient P8. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S19.xls]

Additional file 20
Differentially expressed genes in patient P9. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S20.xls]

Additional file 21
Differentially expressed genes in patient P10. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S21.xls]

Additional file 22
Differentially expressed genes in patient P11. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S22.xls]

Additional file 23

Differentially expressed genes in patient P12. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S23.xls]

Additional file 24

Differentially expressed genes in patient P13. Lists of genes detected as differentially expressed in individual patients comparing to controls at adjusted P < 0.01 significance level.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-9-38-S24.xls]

Acknowledgements

This work was supported by grant NR8069-3 from the Grant Agency of the Ministry of Health of the Czech Republic. Further support was provided by grants 303/03/H065 and 303/07/0781 from the Grant Agency of the Czech Republic, 54/203208 27/05 from the Grant Agency of the Charles University of Prague and the Czech-Austrian Bilateral Cooperation Project (Kontakt 2006/3). Institutional support was provided by Ministry of Education of Czech Republic grants 1M6837805002, AVOZ 50110509 and MSM0021620806.

References

1. McFarland R, Taylor RW, Turnbull DM: **Mitochondrial disease – its impact, etiology, and pathology.** *Curr Top Dev Biol* 2007, **77:**113-155.
2. DiMauro S, Schon EA: **Mitochondrial respiratory-chain diseases.** *N Engl J Med* 2003, **348:**2656-2668.
3. DiMauro S: **Mitochondrial DNA medicine.** *Biosci Rep* 2007, **27:**5-9.
4. Chinnery PF: **Searching for nuclear-mitochondrial genes.** *Trends Genet* 2003, **19:**60-62.
5. Shoubridge EA: **Nuclear gene defects in respiratory chain disorders.** *Semin Neurol* 2001, **21:**261-267.
6. Calvo S, Jain M, Xie X, Sheth SA, Chang B, Goldberger OA, Spinazzola A, Zeviani M, Carr SA, Mootha VK: **Systematic identification of human mitochondrial disease genes through integrative genomics.** *Nat Genet* 2006, **38:**576-582.
7. Thorburn DR, Sugiana C, Salemi R, Kirby DM, Worgan L, Ohtake A, Ryan MT: **Biochemical and molecular diagnosis of mitochondrial respiratory chain disorders.** *Biochim Biophys Acta* 2004, **1659:**121-128.
8. Slonim DK: **From patterns to pathways: gene expression data analysis comes of age.** *Nat Genet* 2002, **32(Suppl):**502-508.
9. Mootha VK, Lepage P, Miller K, Bunkenborg J, Reich M, Hjerrild M, Delmonte T, Villeneuve A, Sladek R, Xu F, Mitchell GA, Morin C, Mann M, Hudson TJ, Robinson B, Rioux JD, Lander ES: **Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics.** *Proc Natl Acad Sci USA* 2003, **100:**605-610.
10. Kirby DM, Salemi R, Sugiana C, Ohtake A, Parry L, Bell KM, Kirk EP, Boneh A, Taylor RW, Dahl HH, Ryan MT, Thorburn DR: **NDUFS6 mutations are a novel cause of lethal neonatal mitochondrial complex I deficiency.** *J Clin Invest* 2004, **114:**837-845.
11. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC: **PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.** *Nat Genet* 2003, **34:**267-273.
12. Chen JJ: **Key aspects of analyzing microarray gene-expression data.** *Pharmacogenomics* 2007, **8:**473-482.
13. Alesci S, Manoli I, Michopoulos VJ, Brouwers FM, Le H, Gold PW, Blackman MR, Rennert OM, Su YA, Chrousos GP: **Development of a human mitochondria-focused cDNA microarray (hMitChip) and validation in skeletal muscle cells: implications for pharmaco- and mitogenomics.** *Pharmacogenomics J* 2006, **6:**333-342.
14. Kerstann KW, Procaccio VF, Yen HC, Hosseini SH, Golik PZ, Wallace DC: **Microarray Analysis of Human Mitochondrial Disease Patients.** *Am J Hum Genet* 2000, **67:**271.
15. Van Der Westhuizen FH, Van Den Heuvel LP, Smeets R, Veltman JA, Pfundt R, Van Kessel AG, Ursing BM, Smeitink JA: **Human mitochondrial complex I deficiency: investigating transcriptional responses by microarray.** *Neuropediatrics* 2003, **34:**14-22.
16. Bai X, Wu J, Zhang Q, Alesci S, Manoli I, Blackman MR, Chrousos GP, Goldstein AL, Rennert OM, Su YA: **Third-generation human mitochondria-focused cDNA microarray and its bioinformatic tools for analysis of gene expression.** *Biotechniques* 2007, **42:**365-375.
17. Halgren RG, Fielden MR, Fong CJ, Zacharewski TR: **Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones.** *Nucleic Acids Res* 2001, **29:**582-588.
18. Holloway AJ, van Laar RK, Tothill RW, Bowtell DD: **Options available – from start to finish – for obtaining data from DNA microarrays II.** *Nat Genet* 2002, **32(Suppl):**481-489.
19. Jesina P, Tesarova M, Fornuskova D, Vojtiskova A, Pecina P, Kaplanova V, Hansikova H, Zeman J, Houstek J: **Diminished synthesis of subunit a (ATP6) and altered function of ATP synthase and cytochrome c oxidase due to the mtDNA 2 bp microdeletion of TA at positions 9205 and 9206.** *Biochem J* 2004, **383:**561-571.
20. Seneca S, Abramowicz M, Lissens W, Muller MF, Vamos E, de Meirleir L: **A mitochondrial DNA microdeletion in a newborn girl with transient lactic acidosis.** *J Inherit Metab Dis* 1996, **19:**115-118.
21. Sperl W, Jesina P, Zeman J, Mayr JA, Demeirleir L, VanCoster R, Pickova A, Hansikova H, Houst'kova H, Krejcik Z, Koch J, Smet J, Muss W, Holme E, Houstek J: **Deficiency of mitochondrial ATP synthase of nuclear genetic origin.** *Neuromuscul Disord* 2006, **16:**821-829.
22. Houstek J, Pickova A, Vojtiskova A, Mracek T, Pecina P, Jesina P: **Mitochondrial diseases and genetic defects of ATP synthase.** *Biochim Biophys Acta* 2006, **1757:**1400-1405.
23. **The Database for Annotation, Visualization and Integrated Discovery (DAVID) 2007** [http://david.abcc.ncifcrf.gov]
24. Chrzanowska-Lightowlers ZM, Temperley RJ, Smith PM, Seneca SH, Lightowlers RN: **Functional polypeptides can be synthesized from human mitochondrial transcripts lacking termination codons.** *Biochem J* 2004, **377:**725-731.
25. Mayr JA, Paul J, Pecina P, Kurnik P, Forster H, Fotschl U, Sperl W, Houstek J: **Reduced respiratory control with ADP and changed pattern of respiratory chain enzymes as a result of selective deficiency of the mitochondrial ATP synthase.** *Pediatr Res* 2004, **55:**988-994.
26. Slomovic S, Laufer D, Geiger D, Schuster G: **Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark.** *Mol Cell Biol* 2005, **25:**6427-6435.
27. Gajewski CD, Yang L, Schon EA, Manfredi G: **New insights into the bioenergetics of mitochondrial disorders using intracellular ATP reporters.** *Mol Biol Cell* 2003, **14:**3628-3635.
28. Kelly DP, Scarpulla RC: **Transcriptional regulatory circuits controlling mitochondrial biogenesis and function.** *Genes Dev* 2004, **18:**357-368.
29. Scarpulla RC: **Nuclear control of respiratory gene expression in mammalian cells.** *J Cell Biochem* 2006, **97:**673-683.
30. Hu CM, Chang ZF: **Mitotic control of dTTP pool: a necessity or coincidence?** *J Biomed Sci* 2007.
31. Gemin A, Sweet S, Preston TJ, Singh G: **Regulation of the cell cycle in response to inhibition of mitochondrial generated energy.** *Biochem Biophys Res Commun* 2005, **332:**1122-1132.
32. Martinez-Diez M, Santamaria G, Ortega AD, Cuezva JM: **Biogenesis and Dynamics of Mitochondria during the Cell Cycle: Significance of 3'UTRs.** *PLoS ONE* 2006, **1:**e107.
33. Wang C, Li Z, Lu Y, Du R, Katiyar S, Yang J, Fu M, Leader JE, Quong A, Novikoff PM, Pestell RG: **Cyclin D1 repression of nuclear res-**

piratory factor 1 integrates nuclear DNA synthesis and mitochondrial function. *Proc Natl Acad Sci USA* 2006, **103:**11567-11572.
34. Mandal S, Guptan P, Owusu-Ansah E, Banerjee U: **Mitochondrial regulation of cell cycle progression during development as revealed by the tenured mutation in Drosophila.** *Dev Cell* 2005, **9:**843-854.
35. Liao TS, Call GB, Guptan P, Cespedes A, Marshall J, Yackle K, Owusu-Ansah E, Mandal S, Fang QA, Goodstein GL, Kim W, Banerjee U: **An efficient genetic screen in Drosophila to identify nuclear-encoded genes with mitochondrial function.** *Genetics* 2006, **174:**525-533.
36. Boneh A: **Regulation of mitochondrial oxidative phosphorylation by second messenger-mediated signal transduction mechanisms.** *Cell Mol Life Sci* 2006, **63:**1236-1248.
37. Seshadri T, Campisi J: **Repression of c-fos transcription and an altered genetic program in senescent human fibroblasts.** *Science* 1990, **247:**205-209.
38. Chalmers CJ, Gilley R, March HN, Balmanno K, Cook SJ: **The duration of ERK1/2 activity determines the activation of c-Fos and Fra-1 and the composition and quantitative transcriptional output of AP-1.** *Cell Signal* 2007, **19:**695-704.
39. Limatola C, Mileo AM, Giovannelli A, Vacca F, Ciotti MT, Mercanti D, Santoni A, Eusebi F: **The growth-related gene product beta induces sphingomyelin hydrolysis and activation of c-Jun N-terminal kinase in rat cerebellar granule neurones.** *J Biol Chem* 1999, **274:**36537-36543.
40. Moerman EJ, Thweatt R, Moerman AM, Jones RA, Goldstein S: **Insulin-like growth factor binding protein-3 is overexpressed in senescent and quiescent human fibroblasts.** *Exp Gerontol* 1993, **28:**361-370.
41. Park WY, Park JS, Cho KA, Kim DI, Ko YG, Seo JS, Park SC: **Up-regulation of caveolin attenuates epidermal growth factor signaling in senescent cells.** *J Biol Chem* 2000, **275:**20847-20852.
42. Lee SR, Kim JR, Kwon KS, Yoon HW, Levine RL, Ginsburg A, Rhee SG: **Molecular cloning and characterization of a mitochondrial selenocysteine-containing thioredoxin reductase from rat liver.** *J Biol Chem* 1999, **274:**4722-4734.
43. Cuervo AM, Dice JF: **How do intracellular proteolytic systems change with age?** *Front Biosci* 1998, **3:**d25-43.
44. Cristofalo VJ, Lorenzini A, Allen RG, Torres C, Tresini M: **Replicative senescence: a critical review.** *Mech Ageing Dev* 2004, **125:**827-848.
45. Shelton DN, Chang E, Whittier PS, Choi D, Funk WD: **Microarray analysis of replicative senescence.** *Curr Biol* 1999, **9:**939-945.
46. Stockl P, Hutter E, Zwerschke W, Jansen-Durr P: **Sustained inhibition of oxidative phosphorylation impairs cell proliferation and induces premature senescence in human fibroblasts.** *Exp Gerontol* 2006, **41:**674-682.
47. Whitmarsh AJ, Davis RJ: **Regulation of transcription factor function by phosphorylation.** *Cell Mol Life Sci* 2000, **57:**1172-1183.
48. Shepherd RK, Checcarelli N, Naini A, De Vivo DC, DiMauro S, Sue CM: **Measurement of ATP production in mitochondrial disorders.** *J Inherit Metab Dis* 2006, **29:**86-91.
49. Hardie DG, Hawley SA, Scott JW: **AMP-activated protein kinase – development of the energy sensor concept.** *J Physiol* 2006, **574:**7-15.
50. Zong H, Ren JM, Young LH, Pypaert M, Mu J, Birnbaum MJ, Shulman GI: **AMP kinase is required for mitochondrial biogenesis in skeletal muscle in response to chronic energy deprivation.** *Proc Natl Acad Sci USA* 2002, **99:**15983-15987.
51. Jones RG, Plas DR, Kubek S, Buzzai M, Mu J, Xu Y, Birnbaum MJ, Thompson CB: **AMP-activated protein kinase induces a p53-dependent metabolic checkpoint.** *Mol Cell* 2005, **18:**283-293.
52. Wang W, Fan J, Yang X, Furer-Galban S, Lopez de Silanes I, von Kobbe C, Guo J, Georas SN, Foufelle F, Hardie DG, Carling D, Gorospe M: **AMP-activated kinase regulates cytoplasmic HuR.** *Mol Cell Biol* 2002, **22:**3425-3436.
53. Wang W, Yang X, Cristofalo VJ, Holbrook NJ, Gorospe M: **Loss of HuR is linked to reduced expression of proliferative genes during replicative senescence.** *Mol Cell Biol* 2001, **21:**5889-5898.
54. Wang W, Furneaux H, Cheng H, Caldwell MC, Hutter D, Liu Y, Holbrook N, Gorospe M: **HuR regulates p21 mRNA stabilization by UV light.** *Mol Cell Biol* 2000, **20:**760-769.
55. Wang W, Yang X, Lopez de Silanes I, Carling D, Gorospe M: **Increased AMP:ATP ratio and AMP-activated protein kinase activity during cellular senescence linked to reduced HuR function.** *J Biol Chem* 2003, **278:**27016-27023.
56. Chen HH, Xu J, Safarpour F, Stewart AF: **LMO4 mRNA stability is regulated by extracellular ATP in F11 cells.** *Biochem Biophys Res Commun* 2007, **357:**56-61.
57. Iakova P, Wang GL, Timchenko L, Michalak M, Pereira-Smith OM, Smith JR, Timchenko NA: **Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate.** *Embo J* 2004, **23:**406-417.
58. Timchenko LT, Salisbury E, Wang GL, Nguyen H, Albrecht JH, Hershey JW, Timchenko NA: **Age-specific CUGBP1-eIF2 complex increases translation of CCAAT/enhancer-binding protein beta in old liver.** *J Biol Chem* 2006, **281:**32806-32819.
59. Nakagawa J, Waldner H, Meyer-Monard S, Hofsteenge J, Jeno P, Moroni C: **AUH, a gene encoding an AU-specific RNA binding protein with intrinsic enoyl-CoA hydratase activity.** *Proc Natl Acad Sci USA* 1995, **92:**2051-2055.
60. L I, Loupatty FJ, Ruiter JP, Duran M, Lehnert W, Wanders RJ: **3-Methylglutaconic aciduria type I is caused by mutations in AUH.** *Am J Hum Genet* 2002, **71:**1463-1466.
61. Hrebicek M, Mrazova L, Seyrantepe V, Durand S, Roslin NM, Noskova L, Hartmannova H, Ivanek R, Cizkova A, Poupetova H, Sikora J, Urinovska J, Stranecky V, Zeman J, Lepage P, Roquis D, Verner A, Ausseil J, Beesley CE, Maire I, Poorthuis BJ, van de Kamp J, van Diggelen OP, Wevers RA, Hudson TJ, Fujiwara TM, Majewski J, Morgan K, Kmoch S, Pshezhetsky AV: **Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome).** *Am J Hum Genet* 2006, **79:**807-819.
62. De Meirleir L, Seneca S, Lissens W, De Clercq I, Eyskens F, Gerlo E, Smet J, Van Coster R: **Respiratory chain complex V deficiency due to a mutation in the assembly gene ATP12.** *J Med Genet* 2004, **41:**120-124.
63. **MITOMAP – A human mitochondrial genome database** [http://www.mitomap.org]
64. **Mitochondrial Proteome, Database for mitochondria-related genes, proteins and diseases** [http://www.mitop2.de]
65. **MiGenes Database** [http://www.pharm.stonybrook.edu/migenes]
66. **MitoProteome Database** [http://www.mitoproteome.org]
67. **Molecular Signatures Database** [http://www.broad.mit.edu/gsea/msigdb/msigdb_index.html]
68. **National Center for Biotechnology Information** [http://www.ncbi.nlm.nih.gov]
69. **Gene Ontology** [http://www.godatabase.org/dev/database]
70. **UniProt – the universal protein resource** [http://www.expasy.uniprot.org]
71. **BASE – BioArray Software Environment** [http://base.thep.lu.se]. http://base.img.cas.cz
72. **OligoPicker** [http://pga.mgh.harvard.edu/oligopicker]
73. Klement P, Nijtmans LG, Van den Bogert C, Houstek J: **Analysis of oxidative phosphorylation complexes in cultured human fibroblasts and amniocytes by blue-native-electrophoresis using mitoplasts isolated with the help of digitonin.** *Anal Biochem* 1995, **231:**218-224.
74. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M: **Minimum information about a microarray experiment (MIAME)-toward standards for microarray data.** *Nat Genet* 2001, **29:**365-371.
75. **The R Project for Statistical Computing** [http://www.r-project.org]
76. Smyth GK: *Limma: linear models for microarray data* New York: Springer; 2005.
77. **BIOCONDUCTOR – open source software for bioinformatics** [http://www.bioconductor.org]
78. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing.** *Journal of the Royal Statistical Society* 1995, **B, 57:**289-300.
79. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J: **TM4: a free, open-source system for microarray data management and analysis.** *Biotechniques* 2003, **34:**374-378.
80. **TM4 – microarray software suite** [http://www.tm4.org]

81. **KEGG: Kyoto Encyclopedia of Genes and Genomes** [http://www.genome.jp/kegg/]
82. Houstek J, Klement P, Floryk D, Antonicka H, Hermanska J, Kalous M, Hansikova H, Hout'kova H, Chowdhury SK, Rosipal T, Kmoch S, Stratilova L, Zeman J: **A novel deficiency of mitochondrial ATPase of nuclear origin.** *Hum Mol Genet* 1999, **8:**1967-1974.

nature genetics

TMEM70 mutations cause isolated ATP synthase deficiency and neonatal mitochondrial encephalocardiomyopathy

Alena Čížková[1,2], Viktor Stránecký[1], Johannes A Mayr[3], Markéta Tesařová[4], Vendula Havlíčková[2], Jan Paul[2], Robert Ivánek[1], Andreas W Kuss[5], Hana Hansíková[4], Vilma Kaplanová[2], Marek Vrbacký[2], Hana Hartmannová[1], Lenka Nosková[1], Tomáš Honzík[4], Zdeněk Drahota[2], Martin Magner[4], Kateřina Hejzlarová[2], Wolfgang Sperl[3], Jiří Zeman[4], Josef Houštěk[2] & Stanislav Kmoch[1]

We carried out whole-genome homozygosity mapping, gene expression analysis and DNA sequencing in individuals with isolated mitochondrial ATP synthase deficiency and identified disease-causing mutations in *TMEM70*. Complementation of the cell lines of these individuals with wild-type *TMEM70* restored biogenesis and metabolic function of the enzyme complex. Our results show that TMEM70 is involved in mitochondrial ATP synthase biogenesis in higher eukaryotes.

Mitochondrial ATP synthase, a key enzyme of mitochondrial energy provision, catalyzes synthesis of ATP during oxidative phosphorylation. ATP synthase is a 650-kDa protein complex composed of 16 types of subunits; 6 form the globular F_1 catalytic part and 10 form the transmembraneous F_o part with two connecting stalks[1]. Two mammalian ATP synthase subunits, ATP6 and ATP8, are encoded by mtDNA; all the others are encoded by nuclear DNA. Biogenesis of ATP synthase is a stepwise process requiring a concerted action of assembly factors. Several of these factors have been described in yeast (for example, ATP10, ATP11, ATP12, ATP22, ATP23 and FMC1)[2], but only three have been found in mammals—homologs of F_1-specific factors ATP11 and ATP12 (refs. 2–4) essential for assembly of F_1 subunits α and β, and a homolog of the F_o-related ATP23 with unclear function in mammals[5].

Inherited disorders of ATP synthase belong to most deleterious mitochondrial diseases, which typically affect the pediatric population[6]. Maternally transmitted ATP synthase disorders are caused by heteroplasmic mutations of *MT-ATP6* (ref. 7) and rarely of *MT-ATP8* (ref. 8). These defects impair the energetic function of the F_o proton channel and thus prevent ATP synthesis, although the rate of ATP hydrolysis and the concentration of the enzyme complex remain largely unchanged. In contrast, ATP synthase defects of nuclear genetic origin (MIM604273) are characterized by selective decrease of ATP synthase concentrations (to <30%) and a profound loss of both synthetic and hydrolytic activities[9]. Most affected individuals show neonatal lactic acidosis, hypertrophic cardiomyopathy and/or variable central nervous system involvement and 3-methylglutaconic aciduria. The disease outcome is severe, and half of affected individuals die in early childhood[10]. During the last decade, an increasing number of affected individuals, mostly of Roma (Gypsy) ethnic origin, have been reported[10–13], but a mutation affecting the F_1-specific factor ATP12 was only found in one case[11]. To identify the genetic defect in the other affected individuals with isolated deficiency of ATP synthase we used Affymetrix GeneChip Mapping 250K arrays and genotyped eight index affected individuals, their healthy siblings and parents from six families (**Supplementary Methods** and **Supplementary Fig. 1** online) and performed linkage analysis (**Supplementary Fig. 2** online) and homozygosity mapping (**Fig. 1a** and **Supplementary Fig. 3** online). To prioritize candidate genes, we intersected the mapping information with Agilent 44K array gene expression data[13]. This analysis illuminated a single gene, *TMEM70*, as it has previously been localized in a top-candidate region on chromosome 8 (**Fig. 1a**), showed reduced transcript amount in fibroblast cell lines from affected individuals (**Fig. 1b**,**c**,**d**) and encodes what has been characterized as a mitochondrial protein[14]. Through sequence analysis of genomic DNA (**Supplementary Table 1** online), we identified in affected individuals a homozygous substitution, 317-2A>G, located in the splice site of intron 2 of *TMEM70* (NM-017866; **Fig. 1e**), which leads to aberrant splicing and loss of *TMEM70* transcript (**Fig. 1b**,**d**). We carried out PCR-RFLP analysis in investigated families and proved autosomal recessive segregation of the mutation, as all the affected individuals were homozygous, all parents were heterozygous and unaffected siblings showed either the wild-type or heterozygous genotype. We screened for the 317-2A>G mutation among 25 individuals with low ATP synthase content being studied in our institutions, and found 23 who were homozygous for the mutation (**Supplementary Table 2** online). In an additional single heterozygous individual, P27, we identified on the second allele the frameshift mutation 118_119insGT (**Supplementary Fig. 4** online), which encodes a truncated TMEM70 protein, Ser40CysfsX11. We did not find any mutation in affected individual P3, in whom *TMEM70* transcript amount was also unchanged (**Fig. 1c**). We did not find any of the identified mutations in 100 control individuals.

To prove that TMEM70 is necessary for the biogenesis of the ATP synthase, we carried out RT-PCR analysis of several human tissues (**Fig. 1d**) and found no evidence of distinct *TMEM70* splicing variants

[1]Institute of Inherited Metabolic Disorders, Charles University of Prague, First Faculty of Medicine, Prague 12808, Czech Republic. [2]Department of Bioenergetics, Institute of Physiology, Academy of Science of the Czech Republic, Prague 14220, Czech Republic. [3]Department of Pediatrics, Paracelsus Medical University, Salzburg A5020, Austria. [4]Department of Pediatrics, Charles University of Prague, First Faculty of Medicine, Prague 12808, Czech Republic. [5]Max Planck Institute for Molecular Genetics, Berlin 14195, Germany. Correspondence should be addressed to J.H. (houstek@biomed.cas.cz) or S.K (skmoch@lf1.cuni.cz).

Received 3 April; accepted 28 August; published online 26 October 2008; doi:10.1038/ng.246

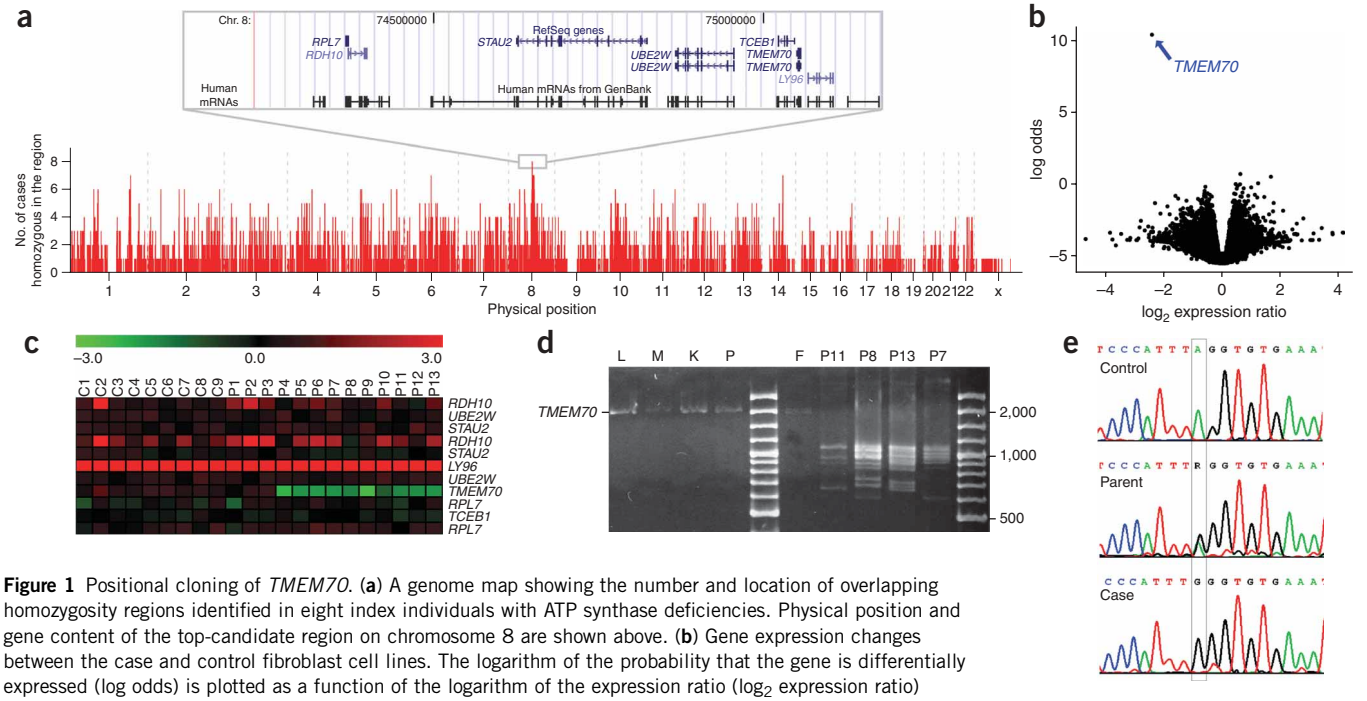


Figure 1 Positional cloning of *TMEM70*. (**a**) A genome map showing the number and location of overlapping homozygosity regions identified in eight index individuals with ATP synthase deficiencies. Physical position and gene content of the top-candidate region on chromosome 8 are shown above. (**b**) Gene expression changes between the case and control fibroblast cell lines. The logarithm of the probability that the gene is differentially expressed (log odds) is plotted as a function of the logarithm of the expression ratio ($\log_2$ expression ratio) between the case and control samples. (**c**) Expression matrix of the genes located in the candidate region showing reduced *TMEM70* transcript amount in all but one case (P3) with a nuclear defect. Normal *TMEM70* transcript amount is present in cases P1 and P2, as these individuals had the mt9205ΔTA microdeletion and not the *TMEM70* mutation. The results are shown as $\log_2$ ratios of gene expression signal in each sample to that of a common reference sample. (**d**) *TMEM70* cDNA analysis. L, M, K, P and F lanes show presence of a single RT-PCR product in control human liver, muscle, kidney, pancreas and fibroblasts, respectively. Abnormal RT-PCR products were observed in fibroblasts from affected individulas (lanes P11, P8, P13 and P7). (**e**) Chromatograms of *TMEM70* genomic DNA sequence showing 317-2A>G substitution.

reported in genomic databases. We cloned *TMEM70* cDNA into the pEF-DEST51 expression vector and transfected skin fibroblast cell lines of several affected individuals (**Fig. 2a**). We found that transfected cells increased the amount of both F_1 and F_o structural subunits of ATP synthase (**Fig. 2b**) and produced normal concentrations of the full size, assembled ATP synthase complex (**Fig. 2c**). Consequently,

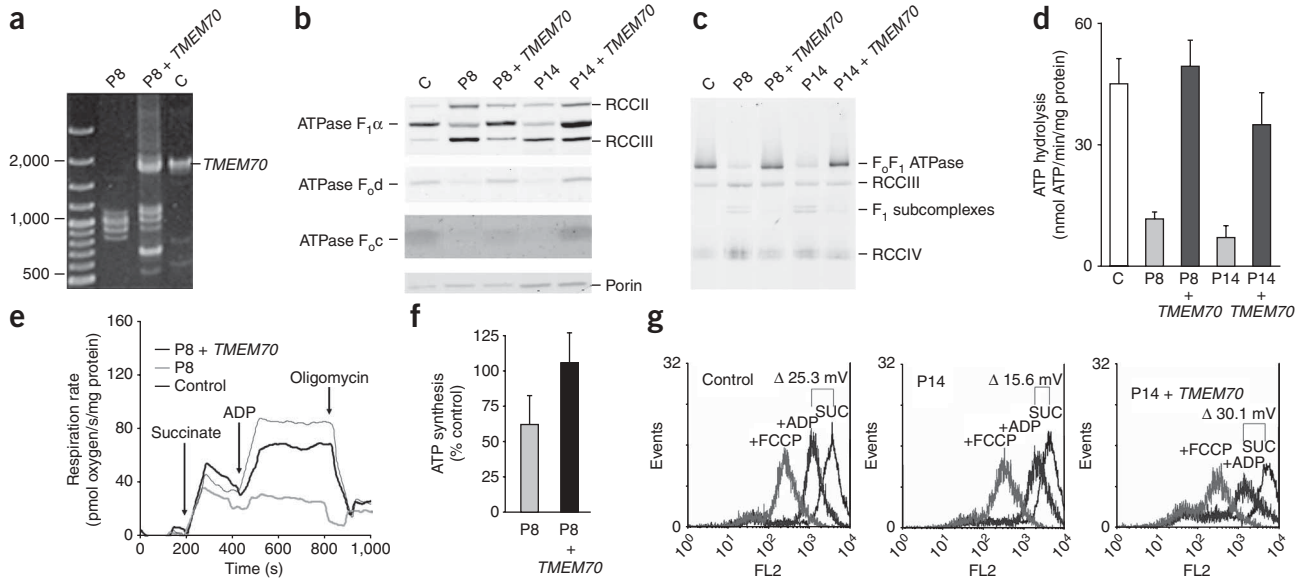


Figure 2 *TMEM70* complementation of ATP synthase deficiency. (**a**) *TMEM70* cDNA is present after transfection. (**b**) SDS-PAGE protein blot of fibroblasts shows a specific increase of the content of ATP synthase subunits relative to the respiratory chain complexes and porin. (**c**) BN-PAGE protein blot of mitochondria shows increase of the full-size assembled ATP synthase 650-kDa complex relative to respiratory chain complexes. (**d**) Oligomycin-sensitive ATP synthase hydrolytic activity is restored. (**e**) ADP stimulation is enhanced in digitonin-permeabilized cells. (**f**) Analysis of ATP formation shows restoration of mitochondrial ATP synthesis. (**g**) TMRM cytofluorometric measurements in permeabilized cells show restoration of the ADP-induced drop of mitochondrial membrane potential at state 4. Data in **d** and **f** are shown as mean ± s.d.; $n = 3$.

the vector restored oligomycin-sensitive ATP hydrolysis (**Fig. 2d**), ADP-stimulated respiration (**Fig. 2e**), mitochondrial ATP synthesis (**Fig. 2f**) and ADP-induced decrease of mitochondrial membrane potential (**Fig. 2g**).

TMEM70 contains the conserved domain DUF1301 and two putative transmembrane regions. Using phylogenetic analysis, we found *TMEM70* homologs in genomes of multicellular eukaryotes and plants, but not in yeast and fungi (**Supplementary Fig. 5** online). This indicates that the evolution of TMEM70 may be an important factor accounting for differences in the ATP synthase assembly process in higher eukaryotes, yeast and bacteria[2,3].

We have identified TMEM70 as a protein involved in the biogenesis of the ATP synthase in higher eukaryotes and shown that its defect is relatively frequent among individuals, particularly Romanies, with mitochondrial energy provision disorders. Existence of the prevalent mutation and co-occurrence of cases with severe and milder phenotypes, probably representing varying quality and functionality of individual nonsense-mediated RNA decay systems, open a way for investigation of translational bypass therapy in this group of individuals.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This study was supported by grants from Ministry of Education of Czech Republic (1M6837805002, AV0Z 50110509, MSM0021620806, Kontakt 14/2006), GAČR (305/08/H037), OeNB 12568, Päd. Forschungsverein and PMU Salzburg (06/04/022). We thank R.Gallyová, Š.Rosipal, V.Smolka, A.Hlavatá, P.Freisinger, M.Huemer and O.Bodamer, who provided samples from affected individuals for this study, and D. Seelow for bioinformatic support.

AUTHOR CONTRIBUTIONS

A.C., H.Hartmannová and L.N. carried out DNA and gene expression analysis and TMEM70 cloning. V.S. and R.I. were responsible for genotyping, gene expression analysis and bioinformatics. J.A.M. carried out biochemical diagnosis and DNA analysis. A.W.K. did genotyping and homozygosity mapping. M.T. and H.Hansíková carried out biochemical diagnosis, cell culturing and transfections. V.H., J.P. and V.K. carried out transfections, complementation studies, ELFO/WB analysis and bioinformatics. M.V., Z.D. and K.H. were responsible for functional studies. T.H. and M.M. were responsible for family ascertainment and sample collection, and J.Z. and W.S. handled diagnosis and clinical characterization. S.K. and J.H. initiated and coordinated the study and wrote the manuscript.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Collinson, I.R., Skehel, J.M., Fearnley, I.M., Runswick, M.J. & Walker, J.E. *Biochemistry* **35**, 12640–12646 (1996).
2. Ackerman, S.H. & Tzagoloff, A. *Prog. Nucleic Acid Res. Mol. Biol.* **80**, 95–133 (2005).
3. Pickova, A., Potocky, M. & Houstek, J. *Proteins* **59**, 393–402 (2005).
4. Wang, Z.G., White, P.S. & Ackerman, S.H. *J. Biol. Chem.* **276**, 30773–30778 (2001).
5. Zeng, X., Neupert, W. & Tzagoloff, A. *Mol. Biol. Cell* **18**, 617–626 (2007).
6. Houstek, J. *et al. Biochim. Biophys. Acta* **1757**, 1400–1405 (2006).
7. Schon, E.A., Santra, S., Pallotti, F. & Girvin, M.E. *Semin. Cell Dev. Biol.* **12**, 441–448 (2001).
8. Jonckheere, A. *et al. J. Med. Genet.* **45**, 129–133 (2007).
9. Houstek, J. *et al. Hum. Mol. Genet.* **8**, 1967–1974 (1999).
10. Sperl, W. *et al. Neuromuscul. Disord.* **16**, 821–829 (2006).
11. De Meirleir, L. *et al. J. Med. Genet.* **41**, 120–124 (2004).
12. Mayr, J.A. *et al. Pediatr. Res.* **55**, 988–994 (2004).
13. Cizkova, A. *et al. BMC Genomics* **9**, 38 (2008).
14. Calvo, S. *et al. Nat. Genet.* **38**, 576–582 (2006).

Mutations in *TMEM76** Cause Mucopolysaccharidosis IIIC (Sanfilippo C Syndrome)

Martin Hřebíček, Lenka Mrázová, Volkan Seyrantepe, Stéphanie Durand, Nicole M. Roslin, Lenka Nosková, Hana Hartmannová, Robert Ivánek, Alena Čížková, Helena Poupětová, Jakub Sikora, Jana Uřinovská, Viktor Stránecký, Jiří Zeman, Pierre Lepage, David Roquis, Andrei Verner, Jérôme Ausseil, Clare E. Beesley, Irène Maire, Ben J. H. M. Poorthuis, Jiddeke van de Kamp, Otto P. van Diggelen, Ron A. Wevers, Thomas J. Hudson, T. Mary Fujiwara, Jacek Majewski, Kenneth Morgan, Stanislav Kmoch,† and Alexey V. Pshezhetsky

Mucopolysaccharidosis IIIC (MPS IIIC, or Sanfilippo C syndrome) is a lysosomal storage disorder caused by the inherited deficiency of the lysosomal membrane enzyme acetyl–coenzyme A:α-glucosaminide *N*-acetyltransferase (*N*-acetyltransferase), which leads to impaired degradation of heparan sulfate. We report the narrowing of the candidate region to a 2.6-cM interval between *D8S1051* and *D8S1831* and the identification of the transmembrane protein 76 gene (*TMEM76*), which encodes a 73-kDa protein with predicted multiple transmembrane domains and glycosylation sites, as the gene that causes MPS IIIC when it is mutated. Four nonsense mutations, 3 frameshift mutations due to deletions or a duplication, 6 splice-site mutations, and 14 missense mutations were identified among 30 probands with MPS IIIC. Functional expression of human *TMEM76* and the mouse ortholog demonstrates that it is the gene that encodes the lysosomal *N*-acetyltransferase and suggests that this enzyme belongs to a new structural class of proteins that transport the activated acetyl residues across the cell membrane.

Heparan sulfate is a polysaccharide found in proteoglycans associated with the cell membrane in nearly all cells. The lysosomal membrane enzyme, acetyl–coenzyme A (CoA):α-glucosaminide *N*-acetyltransferase (*N*-acetyltransferase) is required to *N*-acetylate the terminal glucosamine residues of heparan sulfate before hydrolysis by the α-*N*-acetyl glucosaminidase. Since the acetyl-CoA substrate would be rapidly degraded in the lysosome,[1] *N*-acetyltransferase employs a unique mechanism, acting both as an enzyme and a membrane channel, and catalyzes the transmembrane acetylation of heparan sulfate.[2] The mechanism by which this is achieved has been the topic of considerable investigation, but, for many years, the isolation and cloning of *N*-acetyltransferase has been hampered by its low tissue content, instability, and hydrophobic nature.[3–5]

Genetic deficiency of *N*-acetyltransferase causes mucopolysaccharidosis IIIC (MPS IIIC [MIM 252930], or Sanfilippo syndrome C), a rare autosomal recessive lysosomal disorder of mucopolysaccharide catabolism.[6–8] MPS IIIC is clinically similar to other subtypes of Sanfilippo syndrome.[9] Patients manifest symptoms during childhood with progressive and severe neurological deterioration causing hyperactivity, sleep disorders, and loss of speech accompanied by behavioral abnormalities, neuropsychiatric problems, mental retardation, hearing loss, and relatively minor visceral manifestations, such as mild hepatomegaly, mild dwarfism with joint stiffness and biconvex dorsolumbar vertebral bodies, mild coarse faces, and hypertrichosis.[7] Most patients die before adulthood, but some survive to the 4th decade and show progressive dementia and retinitis pigmentosa. Soon after the first 3 patients with MPS IIIC were described by Kresse et al.,[6] Klein et al.[8,10] reported a similar deficiency in 11 patients who had received the diagnosis of Sanfilippo syndrome, therefore suggesting that the disease is a relatively frequent subtype. The birth prevalence of MPS IIIC in Australia,[11] Portugal,[12] and the Netherlands[13] has been estimated to be 0.07, 0.12, and 0.21 per 100,000, respectively.

The putative chromosomal locus of the MPS IIIC gene was first reported in 1992. By studying two siblings who received the diagnosis of MPS IIIC and had an apparently

From the Institute for Inherited Metabolic Disorders (M.H.; L.M.; L.N.; H.H.; R.I.; A.Č.; H.P.; J.S.; J.U.; V. Stránecký; J.Z.; S.K.) and Center for Applied Genomics (R.I.; A.Č.; V. Stránecký; J.Z.; S.K.), Charles University 1st School of Medicine, and Institute of Molecular Genetics, Academy of Sciences of the Czech Republic (R.I.), Prague; Hôpital Sainte-Justine and Départements de Pédiatrie (V. Seyrantepe; S.D.; J.A.; A.V.P) and Biochimie (A.V.P), Université de Montréal, and Research Institute of the McGill University Health Centre (N.M.R.; T.J.H.; T.M.F.; K.M.), McGill University and Genome Quebec Innovation Centre (P.L.; D.R.; A.V.; T.J.H.; J.M.), and Departments of Human Genetics (T.J.H.; T.M.F.; J.M.; K.M.), Medicine (T.J.H.; T.M.F.; K.M.), and Anatomy and Cell Biology (A.V.P.), McGill University, Montreal; Biochemistry, Endocrinology & Metabolism Unit, UCL Institute of Child Health, London (C.E.B.); Hôpital Debrousse, Lyon, France (I.M.); Department of Medical Biochemistry, Academic Medical Center UVA (B.J.H.M.P.), and Department of Clinical Genetics, VU University Medical Center (J.v.d.K.), Amsterdam; Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands (O.P.v.D.); and Laboratory of Pediatrics and Neurology, University Medical Center, Nijmegen, The Netherlands (R.A.W.)

Received June 8, 2006; accepted for publication August 8, 2006; electronically published September 8, 2006.

Address for correspondence and reprints: Dr. Alexey V. Pshezhetsky, Service de Génétique Médicale, Hôpital Sainte-Justine, 3175 Côte Sainte-Catherine, Montreal, Quebec H3T 1C5, Canada. E-mail: alexei.pchejetski@umontreal.ca

* Footnote added in proof: the gene name has been changed to *HGSNAT*.

† S.K. has led the Prague team.

Am. J. Hum. Genet. 2006;79:807–819. 0002-9297/2006/7905-0004$15.00

balanced Robertsonian translocation, Zaremba et al.[14] suggested that the mutant gene may be located in the pericentric region of either chromosome 14 or chromosome 21, but no further confirmation of this finding was provided. Previously, we performed a genomewide scan on 27 patients with MPS IIIC and 17 unaffected family members, using 392 highly informative microsatellite markers with an average interspacing of 10 cM. For chromosome 8, the scan showed an apparent excess of homozygosity in patients compared with their unaffected relatives.[15] Additional genotyping of 38 patients with MPS IIIC for 22 markers on chromosome 8 identified 15 consecutive markers (from *D8S1051* to *D8S2332*) in an 8.3-cM interval for which the genotypes of affected siblings were identical in state. A maximum multipoint LOD score of 10.6 was found at marker *D8S519,* suggesting that this region includes the locus for MPS IIIC.[15] Recently, localization of the MPS IIIC causative gene on chromosome 8 was confirmed by microcell-mediated chromosome transfer in cultured skin fibroblasts of patients with MPS IIIC.[16]

Here, we report the results of linkage analyses that narrowed the candidate region for MPC IIIC to a 2.6-cM interval between *D8S1051* and *D8S1831* and the identification of the *TMEM76* gene, located within the candidate region, as the gene that codes for the lysosomal *N*-acetyltransferase and, when mutated, is responsible for MPS IIIC.

Material and Methods

Families

In Montreal, 33 affected individuals and 35 unaffected relatives comprising 15 families informative for linkage were genotyped. The families came from Europe, North Africa, and North America. An additional 27 affected individuals and 9 unaffected relatives in uninformative pedigrees, as well as 40 controls, were also genotyped. Eleven of these families and the controls have been reported elsewhere.[15] In addition, 54 individuals from four MPS IIIC–affected families from the Czech Republic were studied in Prague (fig. 1). One family had two affected brothers, whereas the remaining three families each had one affected individual. The families came from various regions of the Czech Republic and were not related within the four most-recent generations. The diagnosis for affected individuals was confirmed by the measurement of *N*-acetyltransferase activity in cultured skin fibroblasts or white blood cells.

Genotyping

The samples in Montreal were genotyped for 22 microsatellite markers in the pericentromeric region of chromosome 8 spanning 8.9 cM on the Rutgers map, version 2.0.[17] The genotyping was performed as described by Mira et al.[18] at the McGill University and Genome Quebec Innovation Centre on an ABI 3730xl DNA Analyzer platform (Applied Biosystems). Alleles were assigned using Genotyper, version 3.6 (Applied Biosystems). The random-error model of SimWalk2, version 2.91,[19,20] was used to detect potential genotyping errors, with an overall error rate of 0.025. Nine genotypes for which the posterior probability of being incorrect was >0.5 were removed before subsequent analyses. In addition, nine genotypes for one marker in one family were removed because of a suspected microsatellite mutation. The samples from the Czech Republic were genotyped in Prague for 18 microsatellite markers in an 18.7-cM region that includes the 8.9-cM region mentioned above. The genotyping was performed on an LI-COR IR2 sequencer by use of Saga genotyping software (Li-Cor) as described elsewhere.[21] Genotypes were screened for errors by use of the PedCheck program.[22]

Linkage Analysis

For the families genotyped in Montreal, multipoint linkage analysis was performed using the Markov chain–Monte Carlo (MCMC) method implemented in SimWalk2, version 2.91,[19] since one pedigree was too large to be analyzed by exact computation. A fully penetrant autosomal recessive parametric model was used with a disease-allele frequency of 0.0045. Marker-allele frequencies were estimated by counting alleles in the available parents of patients with MPS IIIC and in control individuals. To check the consistency of the results, the MCMC analysis was repeated four times.

N-acetyltransferase activity was measured in all participants of the four families from the Czech Republic.[23] Individuals were classified as affected, carriers, or unaffected on the basis of the results of this assay. Mean affected and carrier activities were determined from the five affected individuals and their seven obligate heterozygote parents, respectively, whereas the mean control activity was determined from a sample of 89 unrelated individuals. Four individuals were unable to be classified because their values were within 2 SDs of the means of both the control and carrier groups. Multipoint linkage analysis was performed using a codominant model with a penetrance of 0.99 and a phenocopy rate of 0.01, to account for the possibility of misclassification or genotyping errors. The same disease-allele frequency of 0.0045 was used. Marker-allele frequencies were estimated by counting all genotyped individuals. Exact multipoint linkage analysis was run on 18 microsatellite markers by use of Allegro 1.2c,[24] which was also used to infer haplotypes.

Gene-Expression Analysis

For each of 32 genes located in the candidate interval, a single 5′-amino–modified 40-mer oligonucleotide probe (Illumina) was spotted in quadruplicate on aminosilane-modified microscopic slides and was immobilized using a combination of baking and UV cross-linking. Total RNA (250–1,000 ng) from white blood cells of two patients with MPS IIIC (patients AIV.8 and BIII.5) and four healthy individuals were amplified using the SenseAmp plus RNA Amplification Kit (Genisphere) and were reverse transcribed using 300 ng of poly(A)-tailed mRNA. Reverse transcription and microarray detection were done using the Array 900 Expression Detection Kit (Genisphere) according to the manufacturer's protocol. The two patient samples and four control samples were analyzed in dye-swap mode, in two replicates of each mode. The hybridized slides were scanned with a GenePix 4200A scanner (Molecular Devices), with photomultiplier gains adjusted to obtain the highest-intensity unsaturated images. Data analysis was performed in the R statistical environment (The R Project for Statistical Computing, version 2.2.1) by use of the Linear Models for Microarray Data package (Limma, version 2.2.0).[25] Raw data were processed using loess normalization and a moving minimum background correction on individual arrays and quantile

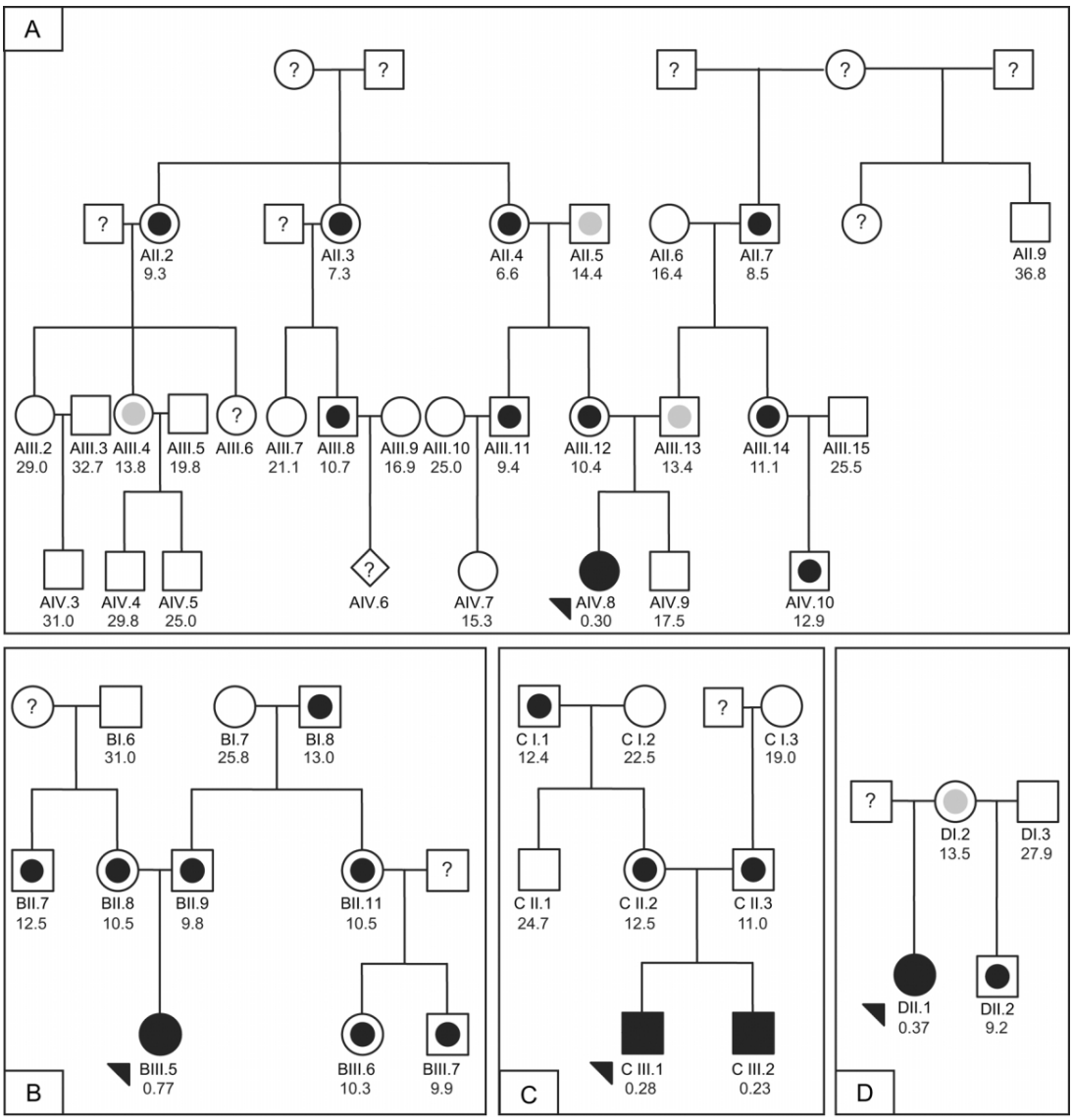


Figure 1. Four families from the Czech Republic used in the linkage and mutation analyses. Fully blackened symbols indicate individuals with MPS IIIC; arrowheads indicate probands. Measurements in seven obligate heterozygotes from these pedigrees (mean ± SD 11.6 ± 1.5 nmol/h/mg) and 89 controls not known to be related to members of the pedigree (mean ± SD 24.4 ± 5.7 nmol/h/mg) were used to establish *N*-acetyltransferase activity ranges for heterozygotes (symbols with blackened inner circle) and normal homozygotes (open symbols). An individual was assigned to a class if his or her enzyme activity was within 2 SDs of the class, unless the value was within the overlap of the upper end of the obligate heterozygotes and the lower end of the controls. Individuals with values within the open interval 13.0–14.6 nmol/h/mg were classified as unknown (symbols with gray inner circle). A symbol with a question mark (?) indicates that no material was available for the enzyme assay. DNA was available for individuals with ID numbers, and *N*-acetyltransferase activity measurements in white blood cells are shown below the ID numbers.

normalization between arrays. The correlation between four duplicate spots per gene on each array was used to increase the robustness. A linear model was fitted for each gene given a series of arrays by use of the lmFit function. The empirical Bayes method[26] was used to rank the differential expression of genes by use of the eBayes function. Correction for multiple testing was performed using the Benjamini and Hochberg false-discovery–rate method.[27] We considered genes to be differentially expressed if the adjusted *P* value was <.01.

DNA and RNA Isolation and Sequencing

Cultured skin fibroblasts from patients with MPS IIIC and normal controls were obtained from cell depositories (Hôpital Debrousse,

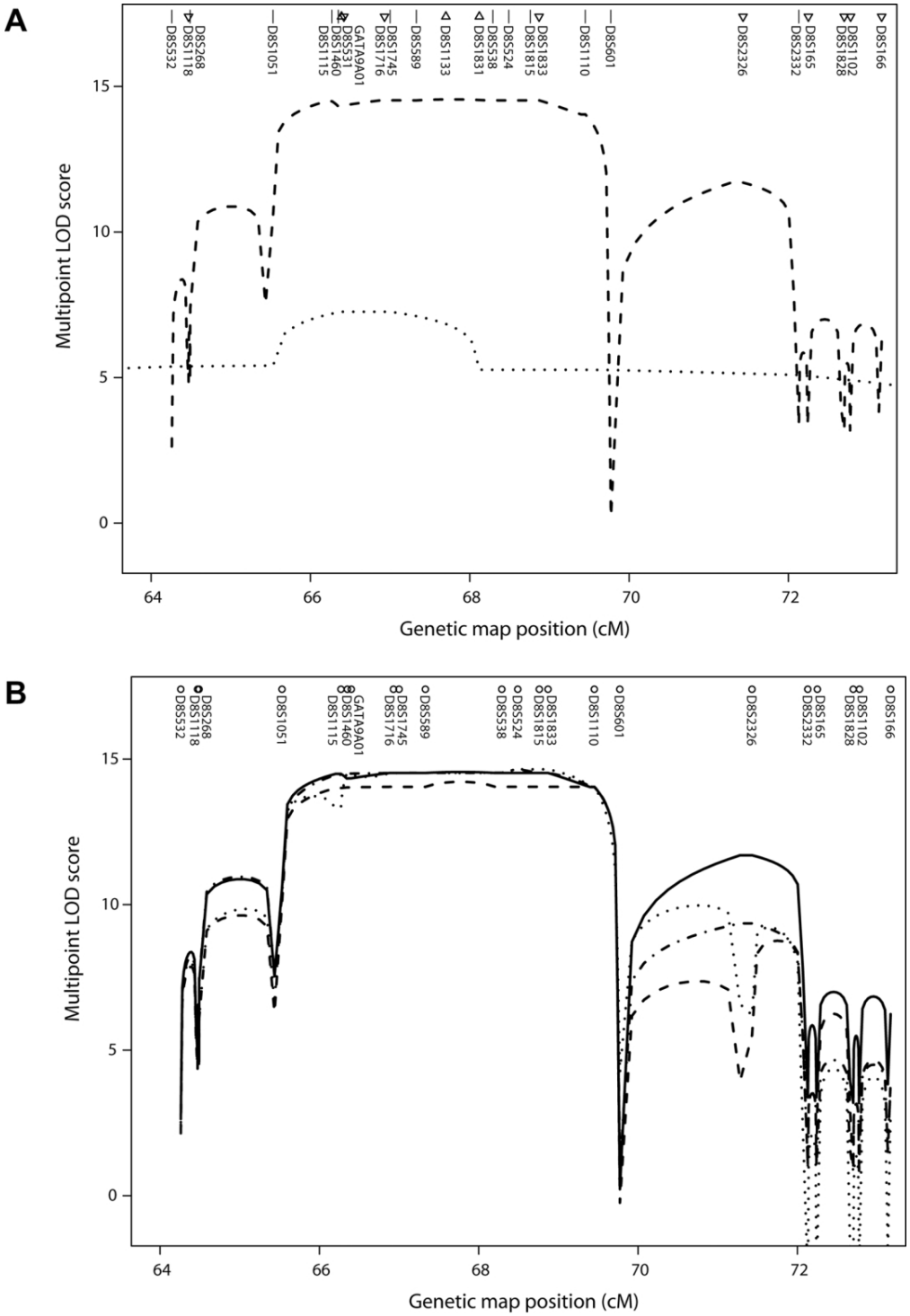


Figure 2. Multipoint linkage analysis of MPS IIIC on chromosome 8. *A,* Multipoint LOD scores in an 8.9-cM interval from two sets of families. Symbols above the marker names indicate the map position. Marker names are listed in the correct order but may be displaced from the symbols for visibility. The dashed line is based on families genotyped in Montreal, and the dotted line on families genotyped in Prague. Straight lines next to marker names indicate that the markers were typed in both data sets. Triangles pointing down indicate markers typed only in the Montreal data set, and triangles pointing up indicate markers typed only in the Prague data set. For the Montreal data, the SimWalk2 run with the highest likelihood is shown. *TMEM76* lies between *D8S1115* and *D8S1460,* and, according to the March 2006 freeze of the human genome sequence from the University of California–Santa Cruz Genome Browser,[30] the order is *D8S1115*–(500 kb)–*TMEM76*–(800 kb)–centromere–(200 kb)–*D8S1460*. *B,* Multipoint LOD scores from the Montreal data from four runs of SimWalk2, version 2.91,[19] showing the variation between runs.

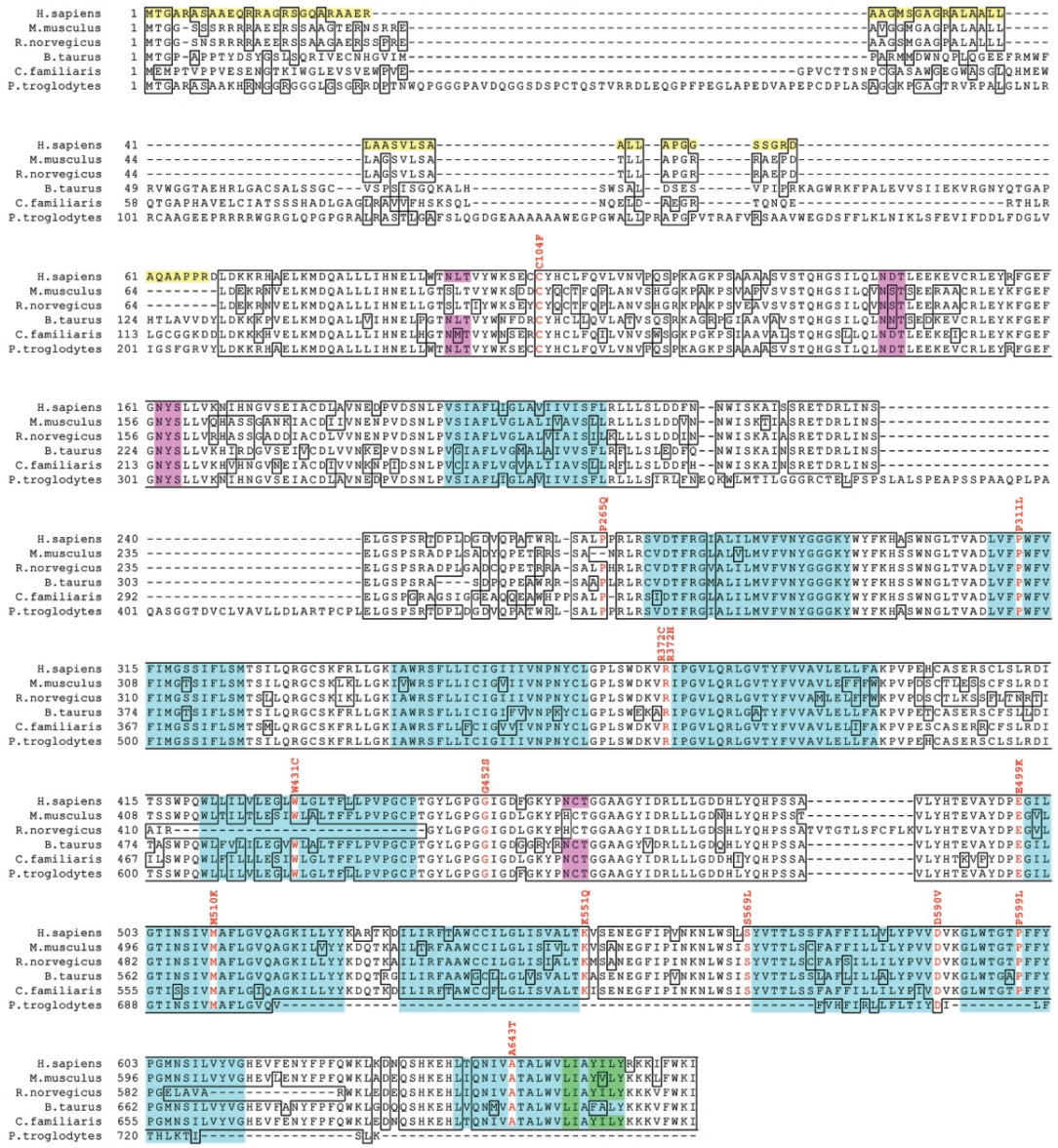


Figure 3. Predicted amino acid sequence of the TMEM76 protein. Amino acid sequence alignment of *Homo sapiens* TMEM76 with orthologs from *Mus musculus* (cloned sequence), *Canis familiaris* (GenBank accession number XP_539948.2), *Bos taurus* (XP_588978.2), *Rattus norvegicus* (XP_341451.2), and *Pan troglodytes* (XP_519741.1) by use of BLAST. All cDNA sequences are predicted except the sequence for *M. musculus*. The identical residues are boxed, the residues with missense mutations in patients with MPS IIIC are shown in red, and the amino acid changes are indicated above the sequence. The first 67 aa of the human sequence shown as black on yellow comprise the predicted signal peptide. The predicted transmembrane domains in the human sequence are shown as black on turquoise. The topology model[5–7] strongly predicts that the N-terminus is inside the lysosome and the C-terminus is outside. Four predicted *N*-glycosylation sites are shown as black on pink, and the predicted motifs for the lysosomal targeting, as black on green.

France; NIGMS Human Genetic Mutant Cell Repository; Montreal Children's Hospital, Canada; and Department of Clinical Genetics, Erasmus Medical Center, The Netherlands). Blood samples from patients with MPS IIIC, their relatives, and controls were collected with ethics approval from the appropriate institutional review boards. DNA from blood or cultured skin fibroblasts was extracted using the PureGene kit (Gentra Systems). Total RNA from cultured skin fibroblasts and pooled tissues (spleen, liver, kidney, heart, lung, and brain) of a C57BL/6J mouse was isolated using Trizol (Invitrogen), and first-strand cDNA synthesis was prepared with SuperScript II (Invitrogen). DNA fragments containing *TMEM76* exons and adjacent regions (~40 bp from each side; primer sequences are shown in appendix A) were amplified by PCR from genomic DNA and were purified with Montage PCR96

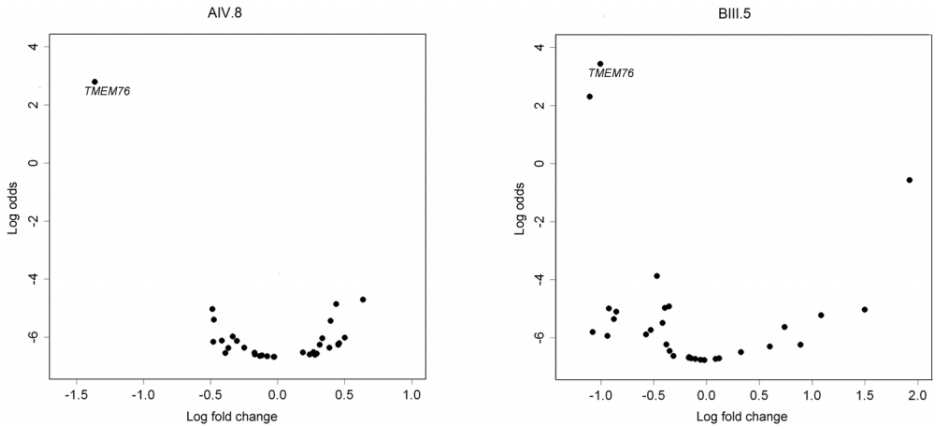


Figure 4. Volcano plot of genes located within the MPS IIIC candidate region, showing significantly reduced expression of the *TMEM76* gene in white blood cells of two patients with MPS IIIC: AIV.8 and BIII.5. The natural logarithm of the probability that the gene is differentially expressed (Log odds) is plotted as a function of the logarithm of the gene-expression $\log_2$ fold change (Log fold change) between the patient and control samples.

filter plates (Millipore). Each sequencing reaction contained 2 μl of purified PCR product, 5.25 μl of H_2O, 1.75 μl of 5 × sequencing buffer, 0.5 μl of 20 μM primer, and 0.5 μl of Big Dye Terminator v3.1 (all from Applied Biosystems). In Montreal, PCR products were analyzed using an ABI 3730xl DNA Analyzer (Applied Biosystems). In Prague, PCR products were analyzed on an ALFexpress DNA sequencer (Pharmacia), as described elsewhere.[28] Included in the sequencing analysis were 30 probands with MPS IIIC who were considered unrelated and 105 controls. The controls were unrelated CEPH individuals, and amplified DNAs were combined in pools of two before sequencing.

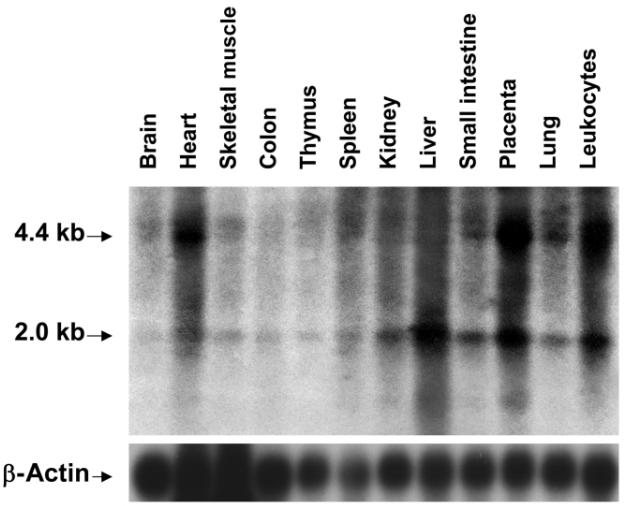


Figure 5. Northern-blot analysis of *TMEM76* mRNA in human tissues. A 12-lane blot containing 1 μg of poly A+ RNA per lane from various adult human tissues was hybridized with a [^{32}P]-labeled 220-bp cDNA fragment corresponding to exons 8–10 of the *TMEM76* gene or β-actin, as described in the Material and Methods section.

Northern Blotting

A 12-lane multiple-tissue northern blot containing 1 μg of poly A+ RNA per lane from various human tissues (BD Biosciences Clontech) was hybridized with the 220-bp cDNA fragment corresponding to exons 8–10 of the human *TMEM76* gene or the entire cDNA of human β-actin labeled with [^{32}P]-dCTP by random priming with the MegaPrime labeling kit (Amersham). Prehybridization of the blot was performed at 68°C for 30 min in ExpressHyb (Clontech). The denatured probes were added directly to the prehybridization solution and were incubated at 68°C for 1 h. The blots were washed twice for 30 min at room temperature with 2 × sodium chloride–sodium citrate (SSC) solution and 0.05% SDS and once for 40 min at 50°C with 0.1 × SSC and 0.1% SDS and were exposed to a BioMax film for 48 h.

Mouse and Human TMEM76 *cDNA Cloning*

Mouse coding sequence was amplified by PCR (forward primer 5′-GAATTCATGACGGGCGGGTCGAGC-3′; reverse primer 5′-ATATGTCGACGATTTTCCAAAACAGCTTC-3′) and was cloned into pCMV-Script, pCMV-Tag4A (Stratagene), and pEGFP-N3 (BD Biosciences Clontech) vectors by use of the *Eco*RI and *Sal*I restriction sites of the primers. The cloned sequence was identical to GenBank accession number AK152926.1, except that an "AT" was needed to complete an alternate ATG initiation codon. GenBank accession number AK149883.1 provides what we consider to be the complete clone and encodes a 656-aa protein. The GenBank sequences differ by 1 aa and three silent substitutions.

A 1,907-bp fragment of the human *TMEM76* cDNA (nt +75 to +1992) was amplified using Platinum High Fidelity *Taq* DNA polymerase (Invitrogen), a sense primer with an *Hin*dIII site (5′-AAGCTTGGCGGCGGGCATGAG-3′), and an antisense primer with an *Sal*I site (5′-GTCGACCTCAGTGGGAGCCATCAGATTTT-3′) and was cloned into pCMV-Script expression vector (Stratagene). Since high GC content (85%) of the 5′ region of human *TMEM76* cDNA prevented its amplification by PCR, a synthetic 186-bp codon-optimized double-stranded oligonucleotide fragment (5′-AAGCTTATGACCGGAGCGAGGGCAAGCGCCGCCG-

AACAAAGAAGAGCCGGACGGTCCGGCCAGGCTAGGGCCGC-AGAGCGAGCTGCTGGCATGTCAGGTGCAGGGCGCGCACTTG-CCGCCTTGCTGCTCGCCGCGAGTGTGCTGAGCGCTGCCCTC-CTGGCTCCCGGAGGCTCTTCCGGGCGGGAC-3′) corresponding to nt +1 to +186 of human *TMEM76* cDNA was purchased from BioS&T. A 177-bp 5′ fragment was combined with rest of the cDNA by use of *Hin*dIII and *Sap*I sites. The cloned sequence is identical to GenBank accession number XM_372038.4 from nt 131 to nt 1946, except for the presence of SNP *rs1126058*.

Cell Culture and Transfection

Skin fibroblasts and COS-7 cells were cultured in Eagle's minimal essential medium (Invitrogen) supplemented with 10% (v/v) fetal bovine serum (Invitrogen) and were transfected with the full-size mouse *Tmem76* (*Hgsnat*) coding sequence subcloned into pCMV-Script, pCMV-Tag4A, and pEGFP-N3 vectors or with the full-size human *TMEM76* coding sequence subcloned into pCMV-Script vector by use of Lipofectamine Plus (Invitrogen) according to the manufacturer's protocol. The cells were harvested 48 h after transfection, and *N*-acetyltransferase activity was measured in the homogenates of *TMEM76*-transfected and mock-transfected cells (i.e., transfected with only the cloning vector).

Enzyme Assay

N-acetyltransferase enzymatic activity was measured using the fluorogenic substrate 4-methylumbelliferyl β-D-glucosaminide (Moscerdam) as described elsewhere.[23] Protein concentration was measured according to the method of Bradford.[29] This assay was used for the activity measurements in cultured skin fibroblasts or white blood cells from patients and all participating members of the Czech families and for the functional expression experiments.

Confocal Microscopy

To establish colocalization of the tagged protein with the lysosomal compartment, the skin fibroblasts expressing mouse TMEM76-EGFP were treated with 50 nM LysoTracker Red DND-99 dye (Molecular Probes), were washed twice with ice-cold PBS, and were fixed with 4% paraformaldehyde in PBS for 30 min. Slides were studied on an LMS 510 Meta inverted confocal microscope (Zeiss).

Results

Linkage Analysis

Previously, we performed a genomewide linkage study that indicated that the locus for MPS IIIC is mapped to an 8.3-cM interval in the pericentromeric region of chromosome 8.[15] To reduce this interval, we genotyped the families from that study as well as newly obtained MPS IIIC–affected families for 22 microsatellite markers (Montreal data). Linkage analysis under an autosomal recessive model resulted in LOD scores >14 in the 4.2-cM region spanning *D8S1051* to *D8S601,* which included the centromere (fig. 2). The results of multiple MCMC runs showed consistent trends. Linkage was also performed in four families from the Czech Republic by use of an autosomal codominant model (Prague data). For these data, linkage analysis produced a maximum LOD score of 7.8 at 66.4 cM at *D8S531* and reduced the linked region for the Montreal data to a 2.6-cM interval between *D8S1051* and *D8S1831.* This region was defined by inferred recombinants at *D8S1051* in one family in each of the Montreal and Prague data sets, and a recombinant at *D8S1831* in an additional family in the Prague data set. This interval contains 32 known or predicted genes and ORFs.

Identification of a Candidate Gene

On the basis of our previous studies that defined the molecular properties of the lysosomal *N*-acetyltransferase,[31] we searched the candidate region for a gene encoding a protein with multiple transmembrane domains and a molecular weight of ~100 kDa, which allowed us to exclude the majority of the genes in the region. In contrast, the predicted protein product of the *TMEM76* gene has multiple putative transmembrane domains. The predicted coding region in GenBank accession number XM_372038.4 was extended by 28 residues at the 5′ end on the basis of the transcript in GenBank accession number DR000652.1 (which includes 14 of the 28 residues), examination of the genomic sequence in NT_007995.14, and comparison with mouse sequence AK149883.1. We predict that the modified *TMEM76* contains 18 exons, corresponding to an ORF of 1,992 bp, and codes for a 73-kDa protein. A comparison of human TMEM76 with five vertebrate orthologs is shown in figure 3. Furthermore, of all the genes present in the candidate interval, only *TMEM76* showed a statistically significant reduction of the transcript level in the cells of two patients with MPS IIIC (AIV.8 and BIII.5; adjusted P values < .001) in the custom oligonucleotide-based microarray assay (fig. 4). Further, we showed that both patients carried nonsense mutations presumably causing mRNA decay (R534X and L349X; see table 2).

Analysis of the TMEM76 *Transcript by Northern Blot and RT-PCR*

Northern-blot analysis identified two major *TMEM76* transcripts of 4.5 and 2.1 kb ubiquitously expressed in various human tissues (fig. 5). The highest expression was detected in leukocytes, heart, lung, placenta, and liver, whereas the gene was expressed at a much lower level in the thymus, colon, and brain, which is consistent with the expression patterns of lysosomal proteins. Consistent with the northern-blot results, a full-length 4.5-kb cDNA containing 1,992 bp of coding sequence and two polyadenylation signals as well as two shorter transcripts were amplified by RT-PCR from the total RNA of normal human skin fibroblasts, white blood cells, and skeletal muscle. In one transcript, exons 9 and 10 were spliced out, leading to an in-frame deletion of 64 aa, which contains the predicted transmembrane domains III and IV. Most likely, this transcript does not encode an active enzyme, since it was also detected in the RNA of two patients with MPS IIIC (patients CIII.1 and CIII.2) who had almost complete loss of

N-acetyltransferase activity. Another transcript lacked exons 3, 9, and 10.

The deduced amino acid sequence predicts 11 transmembrane domains and four potential *N*-glycosylation sites (fig. 3), consistent with the molecular properties of lysosomal *N*-acetyltransferase.[31] The first 67 aa may comprise the signal peptide, with length and composition resembling those of lysosomal proteins. According to the predictions made by empirical computer algorithms,[32–34] the C-terminus of the TMEM76 protein is exposed to the cytoplasm and contains conserved Tyr-X-X-Θ and Leu-Leu sequence motifs involved in the interaction with the adaptor proteins responsible for the lysosomal targeting of membrane proteins.[35]

Mutation Analysis

We identified 27 *TMEM76* mutations in the DNA of 30 MPS IIIC–affected families (table 1) that were not found in DNA from 105 controls. Among the identified mutations, there were 4 nonsense mutations, 14 missense mutations, 3 predicted frameshift mutations due to deletions or duplications, and 6 splice-site mutations. All the missense mutations occur at residues conserved among five species with the most homologous TMEM76 sequences (fig. 3), except for P265Q, which is not conserved in the mouse, and W431C, which is not conserved in the rat. There were three instances of two mutations on the same allele that were found in patients who were homozygous, and these are designated as complex mutations in table 1. cDNA sequencing of one of the patients homozygous for the splice-site mutation in intron 2 and a missense mutation (P265Q) demonstrated that the splice-site mutation disrupts the consensus splice-site sequence between exon 2 and intron 2 and causes exon 2 skipping and a frameshift (not shown).

Consanguinity was reported in 4 of the 13 families in which the patients were homozygous for *TMEM76* mutations: the two Moroccan families, the French family with two missense mutations (W431C and A643T), and the Turkish family with the splice-site mutation in intron

Table 1. Mutations in *TMEM76* Identified in Patients from 30 Families with MPS IIIC

Mutation Group and Mutation[a]	Predicted Effect on Protein	No. of Alleles	Location in *TMEM76*
Nonsense mutations:			
c.1031G→A	p.W344X	2	Exon 10
c.1046T→G	p.L349X	1	Exon 10
c.1234C→T	p.R412X	8	Exon 12
c.1600C→T	p.R534X	1	Exon 15
Missense mutations:			
c.311G→T	p.C104F	1	Exon 2
c.932C→T	p.P311L	3	Exon 9
c.1114C→T	p.R372C	3	Exon 11
c.1115G→A	p.R372H	1	Exon 11
c.1354G→A	p.G452S	2	Exon 13
c.1495G→A	p.E499K	3	Exon 14
c.1529T→A	p.M510K	1	Exon 14
c.1706C→T	p.S569L	4	Exon 17
c.1769A→T	p.D590V	1	Exon 17
c.1796C→T	p.P599L	1	Exon 17
Frameshift mutations:			
c.1118_1133del	p.I373SfsX3	1	Exon 11
c.1420_1456dup	p.V488GfsX22	1	Exon 13
c.1834delG	p.V612SfsX16	1	Exon 18
Splice-site mutations:			
c.202+1G→A	p.L69EfsX32[b]	1	Intron 1
c.577+1G→A	p.P193HfsX20[b]	1	Intron 4
c.935+5G→A	p.F313X	1	Intron 9
c.1334+1G→A	p.G446X[b]	1	Intron 12
c.1810+1G→A	p.S567NfsX14	2	Intron 17
Complex mutations:			
c.[318+1G→A; 794C→A]	p.[D68VfsX19; P265Q]	6	Intron 2; exon 7
c.[577+1G→A; 1650A→C]	p.[P193HfsX20; K551Q]	2	Intron 4; exon 16
c.[1293G→T; 1927G→A]	p.[W431C; A643T]	2	Exon 12; exon 18

[a] Mutation names were assigned according to the guidelines of the Human Genome Variation Society and on the basis of the cDNA sequence from GenBank accession number NT_007995.14, except that the first exon includes 84 nt 5′ of the stated ATG initiation codon. Thus, +1 corresponds to the A of the ATG at nt 13315945 (instead of nt 13316029).

[b] The mutations were named under the assumption that no exon skipping takes place; cDNA sequencing was not done.

17. The two Moroccan families were not known to be related to each other or to the Spanish patient homozygous for the same mutations (table 2). The parents of the French patient are second cousins in two ways (see family F1 in the work of Ausseil et al.[15]).

The splice-site mutation in the above-mentioned Turkish family disrupts the consensus splice-site sequence between exon 17 and intron 17 and causes exon 17 skipping and a frameshift in all transcripts, as detected by sequencing of multiple RT-PCR clones (not shown). The two affected siblings in this family (family F8 in the work of Ausseil et al.[15]) had a severe form of MPS IIIC and showed almost complete loss of *N*-acetyltransferase activity in cultured skin fibroblasts. Among other severely affected patients with MPS IIIC, a patient of French origin was homozygous for a nonsense mutation (W344X) in exon 10, which may result in the synthesis of a truncated protein or RNA decay. A patient of Polish origin was a compound heterozygote for a 37-bp duplication in exon 13 and a missense mutation (S569L) in exon 17 (table 2). The duplication results in a frameshift, whereas the substitution of a strictly conserved small polar Ser for a bulky hydrophobic Leu may have a significant structural impact (fig. 3).

The five patients from four Czech families are all compound heterozygotes for eight different mutations (table 2). Five of the eight mutations are predicted to result in truncated products (three nonsense mutations, one 16-bp deletion, and one splice-site mutation leading to the inclusion of 89 bases from the 5′ end of intron 9 and the splicing out of exon 10 in the transcript, and the remaining three are missense mutations affecting residues conserved among multiple species and located either in the predicted transmembrane regions (fig. 3) or in their close vicinity, suggesting that they may have a serious structural impact. In the Czech families, the mutations completely segregated with reduced enzyme activity. That is, all individuals assigned to be heterozygotes on the basis of the enzyme assay as well as the four individuals who were within 2 SD of the lower end of the controls (symbols with gray inner circle in fig. 1) were found to carry *TMEM76* mutations.

Functional Expression Studies

The fibroblast cell line from a patient homozygous for a splice-site mutation in intron 17 with negligible *N*-acetyltransferase activity was transfected with plasmids containing human *TMEM76* cDNA or cDNA of the mouse ortholog of TMEM76 carrying a FLAG tag on the C-terminus or of a fusion protein of mouse TMEM76 with enhanced green fluorescent protein (EGFP). All constructs increased the *N*-acetyltransferase activity in the mutant fibroblast cells to approximately normal level (fig. 6*A*). Significant increase in activity was also observed in transfected COS-7 cells, confirming that the TMEM76 protein

Table 2. TMEM76 Predicted Mutations in Probands from 30 Families with MPS IIIC

Patient Group and Mutation 1	Mutation 2	No. of Patients	Geographic Origin of Patient(s)
Patients from Czech families:			
p.I373SfsX3	p.R534X	1	Czech Republic
p.L349X	p.M510K	1	Czech Republic
p.F313X	p.R412X	1	Czech Republic
p.R372H	p.P599L	1	Czech Republic
Patients homozygous for *TMEM76* mutations:			
p.[D68VfsX19; P265Q]	p.[D68VfsX19; P265Q]	3	Morocco, Morocco, and Spain
p.[P193HfsX20; K551Q]	p.[P193HfsX20; K551Q]	1	France
p.P311L	p.P311L	1	United Kingdom
p.W344X	p.W344X	1	France
p.R372C	p.R372C	1	United Kingdom
p.R412X	p.R412X	2	Turkey and Poland
p.[W431C; A643T]	p.[W431C; A643T]	1	France
p.G452S	p.G452S	1	Canada
p.E499K	p.E499K	1	Canada
p.S567NfsX14	p.S567NfsX14	1	Turkey
Patients compound heterozygous for *TMEM76* mutations:			
p.C104F	…	1	Belarus
p.E499K	p.D590V	1	France
p.P193HfsX20	p.R412X	1	Canada
p.P311L	p.R372C	1	France
p.R412X	…	1	Poland
p.R412X	p.G446X	1	Poland
p.S569L	…	2	France and Portugal
p.S569L	p.L69EfsX32	1	United States
p.V488GfsX22	p.S569L	1	Poland
p.V612SfsX16	…	1	Finland
Families with no mutations identified to date	…	2	North Africa and Portugal

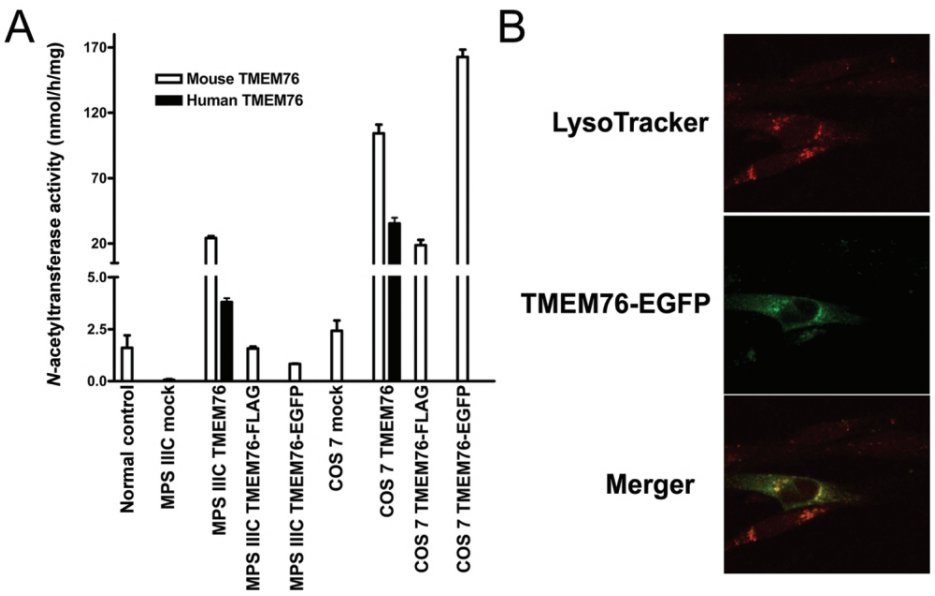


Figure 6. Functional expression of human and mouse TMEM76 protein. *A,* The full-size human and mouse *TMEM76* coding sequences subcloned into pCMV-Script, pCMV-Tag4A, and pEGFP-N3 vectors were expressed in COS-7 cells and in cultured skin fibroblasts from a patient with MPS IIIC. The cells were harvested 48 h after transfection, and *N*-acetyltransferase activity was measured in the homogenates of TMEM76-transfected and mock-transfected fibroblast or COS-7 cells by use of the artificial fluorometric substrate 4-methylumbelliferyl-β-D-glucosaminide.[23] Values represent means $\pm$ SD of four independent experiments. *B,* The intracellular localization of TMEM76 was studied by expressing the fusion protein of the mouse TMEM76 with EGFP. Before fixation, the cells were treated for 45 min with 50 nM lysosomal marker, LysoTracker Red DND-99 dye. Slides were analyzed on an LMS 510 Meta confocal microscope (Zeiss). Magnification $\times$1000. The image was randomly selected from 30 studied panels, all of which showed a similar localization of TMEM76-EGFP. The fluorescence of EGFP was not quenched as it would have been if the fluorophore had been exposed to the acidic lysosomal microenvironment, confirming that the C-terminus of TMEM76 faces the cytoplasmic side of the lysosomal membrane.

by itself has *N*-acetyltransferase activity. Confocal fluorescent microscopy shows that TMEM76-EGFP (fig. 6*B*) or TMEM-FLAG (not shown) peptides are targeted in human fibroblasts to cytoplasmic organelles, colocalizing with the lysosomal-endosomal marker LysoTracker Red.

Discussion

Degradation of heparan sulfate occurs within the lysosomes by the concerted action of a group of at least eight enzymes: four sulfatases, three exo-glycosydases, and one *N*-acetyltransferase, which work sequentially at the terminus of heparan sulfate chains, producing free sulfate and monosaccharides. The inherited deficiencies of four enzymes involved in the degradation of heparan sulfate cause four subtypes of MPS III: MPS IIIA (heparan *N*-sulfatase deficiency [MIM 252900]), MPS IIIB (α-*N*-acetylglucosaminidase deficiency [MIM 252920]), MPS IIIC (acetyl-CoA:α-glucosaminide acetyltransferase deficiency), and MPS IIID (*N*-acetylglucosamine 6-sulfatase deficiency [MIM 252940]). Since the clinical phenotypes of all these disorders are similar, precise diagnosis relies on the determination of enzymatic activities in patients' cultured skin fibroblasts or leukocytes. The biochemical defect in MPS IIIC was identified 30 years ago as a deficiency of an enzyme that transfers an acetyl group from cytoplasmically derived acetyl-CoA to terminal α-glucosamine residues of heparan sulfate within the lysosomes, resulting in the accumulation of heparan sulfate. Therefore, for identification of the molecular basis of this disorder, we used two complementary approaches. First, we performed a partial purification of human and mouse lysosomal *N*-acetyltransferase, which suggested that the enzyme has properties of an oligomeric transmembrane glycoprotein, with an ~100-kDa polypeptide containing the enzyme active site.[31] Second, by linkage analysis, we narrowed the locus for MPC IIIC to a 2.6 cM-interval (*D8S1051–D8S1831*) and, third, compared the level of transcripts of the genes present in the candidate region between normal control cells and those from patients with MPS IIIC. Thus, an integrated bioinformatic search and gene-expression analysis both pinpointed a single gene, *TMEM76,* which encodes a 73-kDa protein with predicted multiple transmembrane domains and glycosylation sites. DNA mutation analysis showed that patients with MPS IIIC harbor *TMEM76* mutations incompatible with the normal function of the predicted protein, whereas expression of human *TMEM76* and the mouse ortholog proved that the protein has *N*-acetyltransferase activity and lysosomal lo-

calization, providing evidence that *TMEM76* is the gene that codes for the lysosomal *N*-acetyltransferase.

The TMEM76 protein does not show a structural similarity to any known prokaryotic or eukaryotic *N*-acetyltransferases or to other lysosomal proteins, on the basis of sequence homology searches. Thus, we think that it belongs to a new structural class of proteins capable of transporting the activated acetyl residues across the cell membrane. Moreover, TMEM76 shares homology with a conserved family of bacterial proteins COG4299 (uncharacterized protein conserved in bacteria) (Entrez Gene GeneID 138050). All 146 members of this family are predicted proteins from diverse bacterial species, including Proteobacteria, Cyanobacteria, and Deinococci. Since many of these bacteria are capable of synthesizing heparan sulfate and other structurally related glycosaminoglycans and perform reactions of transmembrane acetylation, it is tempting to speculate that this activity may also be performed by the proteins of the COG4299 family. Previous studies suggested two contradictory mechanisms of transmembrane acetylation. Bame and Rome[2,36,37] proposed that it is performed via a ping-pong mechanism. First, the acetyl group of acetyl-CoA is transferred to an His residue in the active site of the enzyme. This induces a conformational change that results in the translocation of the protein domain containing the acetylated residue to the lysosome, where the acetyl residue is transferred to the glucosamine residue of heparan sulfate. In contrast, Meikle et al.[38] were unable to demonstrate any specific acetylation of the lysosomal membranes and proposed an alternative mechanism that involved the formation of a tertiary complex of the enzyme, acetyl-CoA, and heparan sufate. Identification of *N*-acetyltransferase as a 73-kDa protein with multiple transmembrane domains, together with our previous data that showed that *N*-acetyltransferase is acetylated by [^{14}C]acetyl-CoA in the absence of glucosamine,[31] strongly supports the ping-pong mechanism of transmembrane acetylation.

For 23 of the 30 probands included in this study for mutation analysis, *TMEM76* mutations were identified in both alleles. Five probands were heterozygous for a missense mutation, with a second mutation yet to be identified. In two probands from North Africa and Portugal, we did not identify any mutations in the coding regions or immediate flanking regions. These patients are homozygous for the microsatellite markers throughout the entire MPS IIIC locus and may be homozygous for a yet-to-be-identified *TMEM76* mutation; however, we cannot formally exclude defects in other genes. Additional studies have been initiated to search for mutations in the introns and promoter regions. The patients with MPS IIIC with the identified frameshift and nonsense mutations all have a clinically severe early-onset form. The almost complete deficiency of *N*-acetyltransferase activity in cultured skin fibroblasts from these patients is consistent with the predicted protein truncations and/or nonsense-mediated mRNA decay. Further expression studies are necessary to confirm the impact of the identified substitutions of the conserved amino acids on enzyme activity. Nevertheless, the identification of the lysosomal *N*-acetyltransferase gene which, when mutated, accounts for the molecular defect in patients with MPS IIIC sets the stage for DNA-based diagnosis and genotype-phenotype correlation studies and marks the end of the gene-discovery phase for lysosomal genetic enzymopathies.

Acknowledgments

We thank the patients, their families, and the Czech Society for Mucopolysaccharidosis, for participating in our study, and members of the sequencing and genotyping facilities at the McGill University and Genome Quebec Innovation Centre, for their technical support. We also acknowledge Nina Gusina, Joe Clarke, and Tony Rupar, for providing cell lines from patients with MPS IIIC; Mila Ashmarina, Milan Elleder, J. Loredo-Osti, and Johanna Rommens, for helpful discussions; Karine Landry, for technical support; and Maryssa Canuel, for help with confocal microscopy. The Montreal study was supported by operating grants from the Sanfilippo Children's Research Foundation (to A.V.P.) and by the Canadian Networks of Centres of Excellence Program—the Mathematics of Information Technology and Complex System network (to K.M.). The Prague study was supported by grants NR8069-1 and 1A/8239-3 from the Grant Agency of the Ministry of Health of the Czech Republic. Institutional support was provided by Ministry of Education of the Czech Republic grant MSM0021620806. A.V.P. is a National Investigator of the Fonds de la Recherche en Santé du Québec.

Appendix A

Table A1. Exon-Flanking Primers Used for PCR Amplification and Sequencing of the Exons in the Human *TMEM76* Gene

Primer	Sequence (5′→3′)
TMEM76_Exon1_F	CTCCCGAAGACAAACACTCC
TMEM76_Exon1_R	GCGAAGTCGCAGCAACAGC
TMEM76_Exon2_F	AAGCTTTTGAGAAGCACTACTGG
TMEM76_Exon2_R	GAAGGGCTTTAGACATGAGAGC
TMEM76_Exon3_F	GGAAAAGTCATGTCAGGATCTCC
TMEM76_Exon3_R	GAATAATACATGTTCCTGGGTACG
TMEM76_Exon4_F	TTATTCTGCCTCCATGATATTAGC
TMEM76_Exon4_R	CTACAGAAAGCGTCATGGACTGC
TMEM76_Exon5_F	GGAAATTCAGCATGAGAATATAGG
TMEM76_Exon5_R	GCCACTTGAGGGTGACAGC
TMEM76_Exon6_F	GAATATGAGCTTTAATTTTATTTCC
TMEM76_Exon6_R	TTAGGAATACGGGAGCTACAACC
TMEM76_Exon7_F	CAAAATGAAATTTACCCCTTAGC
TMEM76_Exon7_R	ACATCCAAGAAATCCTTCCTAGC
TMEM76_Exon8_F	CCTTCCTTTTCACATAGCAAACC
TMEM76_Exon8_R	GCTCTGTGAAGGACGTATATAAGC
TMEM76_Exon9_F	CCCCTGGGTTTACTTTCTATACC
TMEM76_Exon9_R	CCAGCATCATCTGAAAAACAGG
TMEM76_Exon10_F	GGGGCTATATTCTGAACTCTTCC
TMEM76_Exon10_R	ACCTGAGATGGAGGAATTGC
TMEM76_Exon11_F	CTGGGATGAGAGGAGAAGTCC
TMEM76_Exon11_R	ACTTGAAGCCAGGAGTGAGG
TMEM76_Exon12_F	CCTTCTATTTGCATTTAGTTCACC
TMEM76_Exon12_R	GAGAATTCCTCTGACTCGAGACC
TMEM76_Exon13_F	TTTTATTCTTGTCCCTCTGTTCG
TMEM76_Exon13_R	CACTTCTGAAAGCCTGAGTTCC
TMEM76_Exon14_F	TTGGTCTAGGAGCTGTTTGTACG
TMEM76_Exon14_R	CCATAGCACAAGAGAGAATATGC
TMEM76_Exon15_F	TCTTTGTCAGGTAGTTAAGACAGTGG
TMEM76_Exon15_R	GTGAAGGAAAGGAATTTTAGC
TMEM76_Exon16_F	ACAAGTTTCAGCCCTCTCTACG
TMEM76_Exon16_R	GTGGAGGAGACGTTTCAGTGC
TMEM76_Exon17_F	ATGCTGAAATTGGATTTGTTCC
TMEM76_Exon17_R	ACCAAGGATGCTCCAGAGG
TMEM76_Exon18_F	AGTAGCCAACAATGGAAGTGC
TMEM76_Exon18_R	GAGCCGTGTCACAGTTAACC

NOTE.—For bidirectional sequencing on the ALFexpress DNA sequencer, all primers have the universal overhang synthesized on the 5′ end (AATACGACTCACTATAG for forward [F] primers and CAGGAAACAGCTATGAC for reverse [R] primers).

Web Resources

Accession numbers and URLs for data presented herein are as follows:

BLAST, http://www.ncbi.nlm.nih.gov/blast/ (used to identify ortholog protein sequences)

Entrez Gene, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene (for GeneID 138050)

GenBank, http://www.ncbi.nih.gov/Genbank/ (for accession numbers AK152926.1, AK149883.1, DR000652.1, XM_372038.4, NT_007995.14, XP_539948.2, XP_588978.2, XP_341451.2, and XP_519741.1)

Human Genome Variation Society, http://www.hgvs.org/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for MPS IIIA, IIIB, IIIC, and IID)

The R Project for Statistical Computing, http://www.r-project.org/

References

1. Rome LH, Hill DF, Bame KJ, Crain LR (1983) Utilization of exogenously added acetyl coenzyme A by intact isolated lysosomes. J Biol Chem 258:3006–3011
2. Bame KJ, Rome LH (1985) Acetyl-coenzyme A:α-glucosaminide *N*-acetyltransferase: evidence for a transmembrane acetylation mechanism. J Biol Chem 260:11293–11299
3. Pohlmann R, Klein U, Fromme HG, von Figura K (1981) Localisation of acetyl-CoA: α-glucosaminide *N*-acetyltransferase in microsomes and lysosomes of rat liver. Hoppe Seylers Z Physiol Chem 362:1199–1207
4. Hopwood JJ, Freeman C, Clements PR, Stein R, Miller AL (1983) Cellular location of N-acetyltransferase activities toward glucosamine and glucosamine-6-phosphate in cultured human skin fibroblasts. Biochem Int 6:823–830
5. Meikle PJ, Whittle AM, Hopwood JJ (1995) Human acetyl-coenzyme A:α-glucosaminide *N*-acetyltransferase: kinetic characterization and mechanistic interpretation. Biochem J 308:327–333
6. Kresse H, von Figura K, Bartsocas C (1976) Clinical and biochemical findings in a family with Sanfilippo disease, type C. Clin Genet 10:364
7. Bartsocas C, Grobe H, van de Kamp JJ, von Figura K, Kresse H, Klein U, Giesberts MA (1979) Sanfilippo type C disease: clinical findings in four patients with a new variant of mucopolysaccharidosis III. Eur J Pediatr 130:251–258
8. Klein U, Kresse H, von Figura K (1978) Sanfilippo syndrome type C: deficiency of acetyl-CoA:α-glucosaminide *N*-acetyltransferase in skin fibroblasts. Proc Natl Acad Sci USA 75: 5185–5189
9. Sanfilippo SJ, Podosin R, Langer LO Jr, Good RA (1963) Mental retardation associated with acid mucopolysaccharidurie (heparitin sulfate type). J Pediatr 63:837–838
10. Klein U, van de Kamp JJP, von Figura K, Pohlmann R (1981) Sanfilippo syndrome type C: assay for acetyl-CoA:α-glucosaminide *N*-acetyltransferase in leukocytes for detection of homozygous and heterozygous individuals. Clin Genet 20:55–59
11. Meikle PJ, Hopwood JJ, Clague AE, Carey WF (1999) Prevalence of lysosomal storage disorders. JAMA 281:249–254
12. Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, Ribeiro H, Pinto E, Silva E, Rocha S, Marcao A, Ribeiro I, Lacerda L, Ribeiro G, Amaral O, Sa Miranda MC (2004) Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet 12: 87–92
13. Poorthuis BJ, Wevers RA, Kleijer WJ, Groener JE, de Jong JG, van Weely S, Niezen-Koning KE, van Diggelen OP (1999) The frequency of lysosomal storage diseases in The Netherlands. Hum Genet 105:151–156
14. Zaremba J, Kleijer WJ, Juijmans JG, Poorthuis B, Fidzianska E, Glogowska I (1992) Chromosomes 14 and 21 as possible candidates for mapping the gene for Sanfilippo disease type IIIC. J Med Genet 29:514
15. Ausseil J, Loredo-Osti JC, Verner A, Darmond-Zwaig C, Maire I, Poorthuis B, van Diggelen OP, Hudson TJ, Fujiwara TM, Morgan K, Pshezhetsky AV (2004) Localization of a gene for mucopolysaccharidosis IIIC to chromosome region 8p11-8q11. J Med Genet 41:941–945

16. Seyrantepe V, Tihy F, Pshezhetsky AV (2006) The microcell-mediated transfer of human chromosome 8 restores the deficient *N*-acetylytransferase activity in skin fibroblasts of mucopolysaccharidosis type IIIC patients. Hum Genet 120:293–296
17. Kong X, Murphy K, Raj T, He C, White PS, Matise TC (2004) A combined linkage-physical map of the human genome. Am J Hum Genet 75:1143–1148
18. Mira MT, Alcais A, Nguyen VT, Moraes MO, Di Flumeri C, Vu HT, Mai CP, Nguyen TH, Nguyen NB, Pham XK, Sarno EN, Alter A, Montpetit A, Moraes ME, Moraes JR, Dore C, Gallant CJ, Lepage P, Verner A, Van De Vosse E, Hudson TJ, Abel L, Schurr E (2004) Susceptibility to leprosy is associated with PARK2 and PACRG. Nature 427:636–640
19. Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. Am J Hum Genet 58:1323–1337
20. Sobel E, Papp JC, Lange K (2002) Detection and integration of genotyping errors in statistical genetics. Am J Hum Genet 70:496–508
21. Hodanova K, Majewski J, Kublova M, Vyletal P, Kalbacova M, Stiburkova B, Hulkova H, Chagnon YC, Lanouette CM, Marinaki A, Fryns JP, Venkat-Raman G, Kmoch S (2005) Mapping of a new candidate locus for uromodulin-associated kidney disease (UAKD) to chromosome 1q41. Kidney Int 68:1472–1482
22. O'Connell JR, Weeks DE (1998) PedCheck: a program for identifying genotype incompatibilities in linkage analysis. Am J Hum Genet 63:259–266
23. Voznyi YV, Karpova EA, Dudukina TV, Tsvetkova IV, Boer AM, Janse HC, van Diggelen OP (1993). A fluorimetric enzyme assay for the diagnosis of Sanfilippo disease C (MPS III C). J Inher Metab Dis 16:465–472
24. Gudbjartsson DF, Jonasson K, Frigge M, Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. Nat Genet 25:12–13
25. Smyth GK (2005) Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (eds) Bioinformatics and computational biology solutions using R and Bioconductor. Springer, New York, pp 397–420
26. Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3:article 3
27. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300
28. Kmoch S, Hartmannova H, Stiburkova B, Krijt J, Zikanova M, Sebesta I (2000) Human adenylosuccinate lyase (ADSL), cloning and characterization of full-length cDNA and its isoform, gene structure and molecular basis for ADSL deficiency in six patients. Hum Mol Genet 9:1501–1513
29. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254
30. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, et al (2006) The UCSC Genome Browser Database: update 2006. Nucleic Acids Res 34:D590–D598
31. Ausseil J, Landry K, Seyrantepe V, Trudel S, Mazur A, Lapointe F, Pshezhetsky AV (2006) An acetylated 120-kDa lysosomal transmembrane protein is absent from mucopolysaccharidosis IIIC fibroblasts: a candidate molecule for MPS IIIC. Mol Genet Metab 87:22–31
32. Kahsay RY, Gao G, Liao L (2005) An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics 21:1853–1858
33. Jensen LJ, Gupta R, Blom N, Devos D, Tamames J, Kesmir C, Nielsen H, Staerfeldt HH, Rapacki K, Workman C, Andersen CA, Knudsen S, Krogh A, Valencia A, Brunak S (2002) Prediction of human protein function from post-translational modifications and localization features. J Mol Biol 319:1257–1265
34. Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4:1633–1649
35. Bonifacino JS, Traub LM (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. Annu Rev Biochem 72:395–447
36. Bame KJ, Rome LH (1986a) Acetyl-coenzyme A:α-glucosaminide *N*-acetyltransferase: evidence for an active site histidine residue. J Biol Chem 261:10127–10132
37. Bame KJ, Rome LH (1986b) Genetic evidence for transmembrane acetylation by lysosomes. Science 233:1087–1089
38. Meikle PJ, Whittle AM, Hopwood JJ (1995) Human acetyl-coenzyme A:α-glucosaminide *N*-acetyltransferase: kinetic characterization and mechanistic interpretation. Biochem J 308:327–333