UNIVERZITA KARLOVA V PRAZE

1. LÉKAŘSKÁ FAKULTA
ÚSTAV DĚDIČNÝCH METABOLICKÝCH PORUCH

Studijní obor: **Molekulární a buněčná biologie, genetika a virologie**

Autor: **Mgr. Hana Hartmannová**

STUDIUM MOLEKULÁRNÍ PODSTATY VYBRANÝCH DĚDIČNĚ PODMÍNĚNÝCH ONEMOCNĚNÍ

MOLECULAR BASIS OF SELECTED INHERITED RARE DISEASES

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, 12.10.2013

Jméno – Příjmení (hůlkovým písmem)

HANA HARTMANNOVÁ

Podpis

Identifikační záznam:

HARTMANNOVÁ, Hana. *Studium molekulární podstaty vybraných dědičně podmíněných onemocnění [Molecular basis of selected inherited rare diseases]* Praha, 2013. 146 s. Dizertační práce. Univerzita Karlova v Praze, 1. lékařská fakulta, Ústav dědičných metabolických poruch. Vedoucí práce Kmoch, Stanislav.

Abstrakt

Vzácná onemocnění jsou klinicky a geneticky heterogenní skupinou onemocnění postihující různé orgány a projevující se v různém věku. Nalézání, charakterizace a studium funkčních dopadů genetických příčin vzácných onemocnění je efektivním nástrojem odhalování funkce lidských genů a genových produktů a otevírá cestu k pochopení molekulárně-biologických mechanismů jednotlivých onemocnění. Znalost příčin a mechanismů vzniku vzácných onemocnění je následně východiskem pro jejich efektivní diagnostiku, cílenou léčbu a prevenci a zároveň poskytuje i poznatky pro pochopení genetických a molekulárních příčin komplexních onemocnění.

Tato dizertační práce dokumentuje základní koncepční a metodický vývoj postupů biochemické genetiky, funkčního klonování, genetického mapování, pozičního klonování, DNA čipů a genomového sekvenování, které jsou dnes základními nástroji efektivního studia všech geneticky podmíněných onemocnění. Rychlý technologický vývoj a praktická využitelnost řady těchto technik je demonstrována na případech studia molekulární podstaty několika vzácných chorob - deficitu adenylosukcinát lyázy, mukopolysacharidózy typu IIIC, Rotorova syndromu, deficitu ATP syntázy, adultní formy neuronální ceroidní lipofuscinózy, GAPO syndromu a X-vázané restriktivní kardiomyopatie, na jejichž objasnění jsem se během svého studia podílela.

Klíčová slova

vzácná onemocnění, technologie DNA čipů, exomové sekvenování, neuronální ceroidní lipofuscinóza, Rotorův syndrom, izolovaný defekt ATP syntázy, mukopolysacharidóza typu IIIC, Gapo syndrom, X-vázaná restriktivní kardiomyopatie

Abstract

Rare diseases represent a clinically and genetically heterogeneous group of diseases affecting various organs and presenting at different ages. Identification and functional characterization of genetic defects causing individual rare diseases represent unique opportunity to understand biological functions of human genes and gene products as well as to basic pathogenetic mechanisms of individual diseases. This knowledge is prerequisite for their effective diagnosis, specific treatment and prevention and it also opens up an avenue for better understanding of complex diseases.

My thesis documents basic conceptual and methodological developments of biochemical genetics, functional cloning, genetic mapping, positional cloning, DNA microarrays and genomic sequencing, which have provided a universal framework for effective characterization of the genetic architecture of almost all human diseases. This conceptual and technological developments are demonstrated on several cases of rare genetic diseases - adenylosuccinate lyase deficiency, mucopolysacharidosis type IIIC, Rotor syndrome, deficiency of ATP synthese, neuronal ceroid lipofuscinosis, GAPO syndrome and X -linked restrictive cardiomyopathy, which genetic and molecular basis I have helped to elucidate.

Key words

rare diseases, DNA array technology, exome sequencing, neuronal ceroid lipofuscinosis, Rotor syndrome, izolated defekt of ATP synthase, mucopolysacharidosis type IIIC, Gapo syndrom, X-linked hypertrophic cardiomyopathy

Poděkování

Ráda bych poděkovala Standovi Kmochovi, jak za odborné vedení nejen při tvorbě dizertační práce, tak i za a všestrannou pomoc. Dále bych chtěla poděkovat kolektivu Ústavu dědičných metabolických poruch za předané zkušenosti, cenné rady a hlavně za vytvoření hezkých pracovních podmínek, díky nimž mohlo vzniknout tolik mimořádných projektů. Poděkování patří zejména Kateřině Hodaňové, Viktorovi Stráneckému, Lence Piherové, Veronice Barešové, Lence Noskové, Petrovi Vyleťalovi, Martině Živné, Marii Zikánové, Anče Přistoupilové, Aleně Čížkové-Vrbacké, Heleně Hůlkové, Janě Sovové, Evě Oliveriusové a dalším.

Nemenší dík patří také mé rodině, zejména Peťkovi, Honzovi, Adélce a mamince.

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, NT13116-4/2012, grantová agentura Univerzity Karlovy: 54/20320827/05, 250051, programy University Karlovy: PRVOUK-P24/LF1/3, UNCE 204011 a SVV2013/266504 a dále výzkumné projekty Ministerstva školství, mládeže a tělovýchovy MSM0021620806, AV0Z50110509 a 1M6837805002.

Obsah

ČÁST I. Vzácné choroby a jejich studium

Úvod

Primární sekvence lidského genomu a její průběžná funkční anotace (ENCODE) (http://www.nature.com/encode) dnes definují více než 20687 protein kódujících genů (GENCODE annotation, V7) (Bernstein B. E. *et al.*, 2012), 18441 nekódujících RNA, 11224 pseudogenů, z nichž je 863 transkripčně aktivních a stovky tisíc regulačních a evolučně konzervovaných oblastí. Biologická funkce většiny těchto genetických elementů je zatím zcela neznámá. Jednou z možností, jak můžeme postupně odhalit biologické funkce těchto genetických elementů, je systematické vyhledávaní jejich genetických variant (mutací) v populaci a korelace těchto variant s fenotypem jejich nositelů.

Unikátní možnosti v tomto směru poskytují vzácná onemocnění. Tato onemocnění jsou ve velké většině monogenní, tj. jsou způsobena mutací jednoho genu nebo jiného funkčně významného genetického elementu. Závažné klinické příznaky přítomné u jednotlivých případů jsou proto přímou demonstrací funkčního významu těchto elementů. Tělní tekutiny, tkáně a tkáňové kultury získané odběrem biologického materiálu jednotlivých případů navíc poskytují nezastupitelný zdroj a nenahraditelné modely pro studium základních biologických a patofyziologických mechanismů u člověka.

Studium jednotlivých případů vzácných onemocnění má též přímý vztah ke studovaným subjektům. Nalezení genetické a molekulární příčiny onemocnění je východiskem pro kvalifikované genetické poradenství v rodinách, umožňuje jejich prevenci formou prenatální, případně preimplantační DNA diagnostiky a v některých případech i cílenou léčbu. Významným výsledkem objasnění základního vztahu mezi genotypem a fenotypem je též možnost nalezení populačně častějších variant kauzálního genu a studium jejich případného podílu na některém z fenotypů původně studovaného onemocnění (např. familiární hypercholesterolémie vs. ischemická choroba srdce).

Teoretická část této dizertační práce dokumentuje základní koncepční a metodický vývoj postupů funkčního klonování, pozičního klonování, DNA čipů a genomového sekvenování, které postupně objasnily více než 3 000 různých typů vzácných onemocnění a jsou dnes základním metodickým nástrojem efektivního studia jak vzácných, tak i populačně častých, geneticky podmíněných onemocnění.

Experimentální část dizertační práce demonstruje rychlý technologický vývoj a praktickou využitelnost řady těchto technik na příkladech několika vzácných chorob - deficitu adenylosukcinát lyázy, mukopolysacharidózy typu IIIC, Rotorova syndromu, deficitu ATP syntázy, neuronální ceroidní lipofuscinózy, GAPO syndromu a X-vázané restriktivní kardiomyopatie, na jejichž výzkumu jsem se během svého studia podílela.

Vzácná onemocnění, definice a současný stav poznání

Vzácná onemocnění jsou definována jako onemocnění, jejichž prevalence je v Evropě menší než 1:2000 a v USA menší než 1:1250 (Remuzzi G. and Garattini S., 2008). Odhaduje se, že existuje přes 7 000 vzácných chorob (Heemstra H. E. *et al.*, 2009). Ve velké většině se jedná o onemocnění dědičně podmíněná, způsobená mutacemi jednotlivých genů. Mohou postihovat pacienty již od velmi útlého věku, projevovat se těžkým fyzickým či mentálním postižením výrazně zhoršujícím kvalitu života pacientů a jejich rodin (Dear J. W. *et al.*, 2006). Celkový počet pacientů trpících vzácným onemocněním se v Evropě odhaduje na 30 milionů, v Severní Americe na 25 milionů (Schieppati A. *et al.*, 2008).

Vzácná onemocnění jsou velmi heterogenní skupinou nemocí postihující různé orgány a mající nejrůznější klinické projevy. Vyskytují se jak velmi vzácně, pouze u několika pacientů na světě, tak i častěji. Příkladem častějšího onemocnění je cystická fibróza (1:2700-3800) nebo narkolepsie s kataplexií s incidencí 1:2000 (Dear J. W. *et al.*, 2006). Příkladem méně častého metabolického onemocnění je metylmalonová acidurie s frekvencí výskytu 1:50000 nově narozených dětí.

Znalosti o molekulární podstatě jednotlivých chorob, z nich plynoucí případné možnosti diagnostiky a následné léčby, jsou velmi malé, pacienti tak získávají jen málo informací o chorobě, kterou trpí. Často podstupují ve snaze o zjištění diagnózy náročná vyšetření. V současné době existuje léčba pro pouhých 5% chorob (Rohn J., 2013).

Studium vzácných onemocnění bylo dlouhodobě podceňováno, neboť postihovalo jen malou skupinu pacientů, jejich výzkum se proto jevil jako finančně nevýhodný. Až v roce 1983 byla v USA založena organizace National Organisation of Rare Disorders (NORD), která se začala významně podílet na podpoře výzkumu způsobu léčby vzácných onemocnění. V Evropě pacienty se vzácnými onemocněními zastupuje nevládní organizace EURORDIS (Rare Diseases Europe), která sdružuje 561 organizací z 51 zemí. Tato uskupení se snaží zlepšit kvalitu života pacientů s vzácným onemocněním a jejich blízkých.

Studium vzácných onemocnění významně přispívá k pochopení základní biologie, definuje kauzální geny a může pomoci vysvětlit jejich biologickou funkci, neboť funkce většiny genů je stále neznámá. Databáze OMIM, Online Mendelian Inheritance in Man, shromažďující údaje o chorobách s genetickou komponentou a provádějící jejich katalogizaci, popisuje k 7. říjnu 2013 pouze 3 919 fenotypů s vysvětlenou molekulární podstatou z celkového počtu 7527 (viz tabulka 1). Studium monogenních vzácných onemocnění proto může přinést ještě mnoho informací, které v databázích stále chybí.

	Autozomální	X-vázané	Y-vázané	Mitochondriální	Celkem
Geny se známou sekvencí	13 641	664	48	35	14 388
Geny se známou sekvencí a fenotypem	111	4	0	2	117
Fenotyp se známou molekulární podstatou	3 609	278	4	28	3 919
Fenotyp nebo lokus s neznámou molekulární podstatou	1 595	132	5	0	1 732
Ostatní, především fenotypy s pravděpodobnou mendelovskou dědičností	1 755	119	2	0	1 876
Celkem	20 711	1 197	59	65	22 032

Tabulka 1. Počet záznamů v databázi OMIM k 7. říjnu 2013 (http://www.omim.org/statistics/entry)

Identifikace genů podmiňujících vzácná onemocnění

Velká většina vzácných onemocnění je způsobena mutacemi jednotlivých genů či funkčně významných genetických elementů. Základním východiskem pro efektivní studium těchto onemocnění je proto určení jejich kauzální genetické příčiny.

Historicky byla vzácná onemocnění nejprve spojována s klinickými, biochemickými či histopatologickými charakteristikami pozorovanými u jednotlivých pacientů. Tento přístup vedl k definici řady klinicko-patologických jednotek, systematickému shromažďování a detailní charakterizaci skupin pacientů s podobnými příznaky.

Určení kauzální genetické příčiny v jednotlivých případech bylo umožněno až na základě objasnění principu genetického kódu a rozvoje biochemických a molekulárně

biologických metod. Toto znalostní a metodické vybavení vedlo ke vzniku konceptu funkčního klonování a biochemické genetiky.

Funkční klonování vychází z předběžné znalosti funkce (dysfunkce) příslušného proteinu a z možnosti jeho detekce a izolace. Následné určení aminokyselinové sekvence pomocí Edmanovy degradace či hmotově spektrometrické analýzy a příprava specifických protilátek proti tomuto proteinu (či jeho části) umožňují hybridizační vyhledávání a izolaci příslušných komplementárních DNA (cDNA) nebo genů či jejich fragmentů z cDNA či genomových (gDNA) knihoven. Sekvenční analýza takto získaných cDNA a gDNA odhalí sekvenci a organizaci genu, umožní odhad primární sekvence proteinu a otevírá cestu k analýze příslušného genu ve skupinách pacientů s podezřením na jeho poruchu. Tímto způsobem byly určeny historicky první mutace v genech krevního faktoru VIII (Gitschier J. *et al.*, 1984), fenylalanin hydroxylázy (Kwok S. C. *et al.*, 1985) (Robson K. J. *et al.*, 1984) a následně celé řady dalších protein kódujících genů. Znalost sekvence jednotlivých genů a určování jejich mutací ve vztahu ke studovaným onemocněním umožnilo též homologní klonování jejich paralogů nebo ortologů a studium případného kauzálního efektu mutací v těchto genech u podobně klinicky postižených pacientů.

Postupy funkčního klonování jsou obecně limitovány znalostí biochemické či histopatologické podstaty onemocnění, znalostí funkce příslušného proteinu a možností jeho detekce a izolace. Tento předpoklad však splňuje jen velmi malá část vzácných onemocnění.

Určení kauzální genetické příčiny vzácných onemocnění bez předchozí znalosti funkce genového produktu umožnil koncept genetického mapování a pozičního klonování. (Botstein D. *et al.*, 1980) (Petes T. D. and Botstein D., 1977) Tento postup je založen na studiu segregace polymorfních markerů, historicky nejprve proteinů a následně polymorfních oblastí DNA (délky restrikčních fragmentů (restriction fragment length polymorphism, RFLP), tandemových repetitivních oblastí (short tandem repeat, STR) a jednonukleotidových polymorfismů (single nukleotide polymorphism, SNPs)) v rodinách a skupinách pacientů se studovaným fenotypem. Segregace polymorfních markerů (vazba) se studovaným fenotypem umožňuje poměrně přesně lokalizovat hledaný gen na genetické a fyzické mapě lidského genomu a následně izolovat a sekvenčně charakterizovat jednotlivé geny lokalizované v kandidátní oblasti. Tento postup vedl na konci 80-let k určení chromozomální lokalizace kauzálního genu Huntingtonovy choroby (Gusella J. F. *et al.*, 1983), pozičnímu klonování chronické granulomatózní nemoci (Royer-Pokora B. *et al.*, 1986), cystické fibrózy (Riordan J. R. *et al.*, 1989) a řady dalších onemocnění (Collins F. S., 1995). Postupy genetického

mapování a pozičního klonování byly urychleny díky celé řadě nových technologických a koncepčních přístupů v genotypování (analýza STR markerů pomocí polymerázové řetězové reakce (PCR) a genotypování SNP pomocí DNA čipů), vývoji výpočetních programů pro vazebnou analýzu a určení haplotypové struktury, konstrukci nových generací genetických, fyzických a genových map lidského genomu, metodám klonování rozsáhlých oblastí DNA v podobě artificiálních chromozomů, automatickému sekvenování DNA a existenci veřejně dostupných dat generovaných v rámci projektu sekvenace a funkční charakterizace lidského genomu.

Sekvenování lidského genomu

Projekt Hugo (The Human Genome Project) (Venter J. C. *et al.*, 2001), který si kladl za cíl stanovit úplnou nukleotidovou sekvenci lidského genomu, byl zahájen v roce 1990. První podoba genomu člověka byla zveřejněna v roce 2001 (Consortium T. I. H. G. M., 2001), ale nejnovější data byla publikována až v roce 2006 (Gregory S. G. *et al.*, 2006). Přečtení DNA, která představuje celý lidský genom, bylo netrpělivě očekáváno širokou vědeckou komunitou. Zároveň se předpokládalo, že pomůže porozumět lidské evoluci, příčinám řady nemocí a vztahu k životnímu prostředí.

Sekvenování lidského genomu a genomů modelových organismů bylo zpočátku významně limitováno kapacitou přístrojů a cenou sekvenování. Základní metodický přístup proto vycházel z přípravy knihoven genomové DNA. Jednotlivé klony byly charakterizovány pomocí určení jejich RFLP profilu nebo sekvenace jejich konců (sequenced tag sites, STS). Na základě těchto informací byly v knihovnách vyhledávány gDNA klony vykazující částečně identický RFLP profil nebo obsahující nalezená STS. Vzájemně se překrývající klony umožnily zpětné sestavení přesně pozičně charakterizované fyzické mapy zajišťující pokud možno úplné pokrytí původní sekvence DNA. Vybrané klony gDNA pokrývající genom s minimálním přesahem byly následně subklonovány do sekvenačních vektorů a sekvenovány. Získané sekvence byly v příslušném formátu ukládány v databázi GenBank (http://www.ncbi.nlm.nih.gov). K sekvenční informaci byla přidávána zdrojová data (identifikátory sekvenovaných klonů), mapovací informace (obsah STS) a postupně i známé funkční anotace (genový obsah, sekvence RNA apod.).

Během veřejně podporovaného projektu sekvenování lidského genomu sílila ve Spojených státech amerických snaha o patentování funkčně významných oblastí lidského

genomu (zejména genů). Tato snaha vedla ke vzniku řady firem, např. Human Genome Sciences nebo Incyte, které patentovaly tisíce částečných sekvencí cDNA klonů (expressed sequence tags, EST). V roce 1998 byla založena firma Celera Genomics, která usilovala o prioritní určení a patentování sekvence lidského genomu. Celera Genomics založila svůj přístup na masivním sekvenování náhodně vybraných gDNA klonů (shot-gun sekvenování) a zpětném poskládání získaných sekvencí. Tento přístup vedl k vývoji a výrobě kapilárních sekvenátorů DNA, odstartoval revoluci v automatizaci sekvenování DNA a vedl ke vzniku vysoce výkonných výpočetních a bioinformatických nástrojů umožňujících analýzu a uchovávání velkých objemů sekvenačních dat.

Soutěž mezi veřejně a privátně podporovaným projektem sekvenování lidského genomu měla rozsáhlé odborné, politické i etické důsledky. Ve svém výsledku poskytla základní, veřejně dostupnou informaci o sekvenci lidského genomu a otevřela cestu k následné funkční charakterizaci genomu, studiu genetické variability člověka a efektivnímu výběru kandidátních genů v oblastech definovaných vazebnou analýzou.

Technologický a koncepční vývoj v průběhu finální fáze sekvenace lidského genomu zásadním způsobem snížil cenu a dramaticky zvýšil kapacitu sekvenace DNA. Zkušenosti získané v souvislosti s přípravou a manipulací s rozsáhlými knihovnami cDNA a gDNA klonů vedly k vývoji zcela nových technologických prvků. Byly připraveny první cDNA arraye (Schena M. *et al.*, 1995) a litograficky vyráběné oligonukleotidové čipy (Fodor S. P. A. *et al.*, 1993; Pease A. C. *et al.*, 1994). Sekvenování lidských genomů začalo též postupně odkrývat genetickou variabilitu člověka na úrovni jednotlivých nukleotidů. Toto vedlo k definici skupiny jednonukleotidových polymorfismů, SNPs a k jejich katalogizaci v podobě dbSNP databáze (Smigielski E. M. *et al.*, 2000). Oligonukleotidové čipy umožňující genotypování desetitisíců SNPs (Fan J. B. *et al.*, 2000) umožnily konstrukci nového typu genetických map (Matise T. C. *et al.*, 2003) a kombinovaných genetických a fyzických map (Kong X. *et al.*, 2004). Tyto přístupy stimulovaly vývoj nových bioinformatických nástrojů umožňujících současné zpracování desetitisíců genotypů pro účely vazebné analýzy a rekonstrukce haplotypů (Abecasis G. R. *et al.*, 2002) a vedly k postupnému budování integrovaných databází a prohlížečů jako Map Viewer (Wheeler D. L. *et al.*, 2006), ENSEMBL (Hubbard T. *et al.*, 2002) nebo UCSC genome browser (Kent W. J. *et al.*, 2002), které poskytují možnosti on-line vyhledávání, získávání a analýzy genomových sekvencí a jejich příslušných strukturních a funkčních anotací.

DNA čipy

Metodický rozvoj spojený s přípravou gDNA a cDNA knihoven, a zejména s manipulací a charakterizací statisíců klonů vedl k vývoji nových konceptů a nástrojů funkční genomiky, především k vývoji technologie DNA čipů.

DNA čipy byly vyvinuty na základě hybridizačního vyhledávání cDNA nebo gDNA klonů. Zásadní inovací byla myšlenka adresného robotického rozmístění jednotlivých klonů DNA a porovnávací (diferenciální) hybridizace různě fluorescenčně značených cDNA. Elementárním sledovaným parametrem byl zpočátku poměr intenzit fluorescenčního signálu. Toto jednoduché kritérium bylo následně nahrazeno širokým spektrem statistických nástrojů umožňujících vzájemné porovnávání experimentů, přísný statistický výběr odlišně exprimovaných genů a funkční anotaci a interpretaci získaných expresních dat.

Rozvoj robotiky postupně umožnil přípravu prvních celogenomových čipů. Problémy s přípravou sond, nové metody přípravy oligonukleotidů a zejména možnosti jejich *in situ* syntézy vedly ke vzniku nových generací oligonukleotidových čipů, které postupně umožnily efektivní kvalitativní i kvantitativní analýzu nukleových kyselin – analýzu genové dávky (copy number variation (CNV)), analýzu genové exprese, genotypování i sekvenování DNA. DNA čipy výrazně urychlily proces genotypování a mapování kandidátních oblastí. V řadě případů též přímo napomohly k určení kandidátních či kauzálních genů a umožnily funkční studium a odhalení základních patogenetických mechanismů řady onemocnění. DNA čipy přispěly významným způsobem i k objevu a charakterizaci CNV u člověka a umožnily celogenomové asociační studie u stovek komplexních onemocnění.

Nové metody sekvenování

Sekvenování DNA je dnes jednou z hlavních metod molekulární biologie a genetiky. Znalost sekvence a její analýza umožňují předpovědět řadu základních vlastností kódovaných studovaným fragmentem (předpověď rozpoznávacích sekvencí restrikčních a DNA modifikujících enzymů, sekvencí DNA vazebných proteinů, detekce mutací, předpověď struktury genu a sekvence příslušného transkriptu a proteinu, původ a příbuznost studovaného fragmentu apod.). Obdobně znalost sekvence celého genomu umožňuje předpovědět řadu základních vlastností celého organismu.

Význam znalosti sekvence DNA byl zřejmý ihned po objasnění struktury DNA: určení základních mechanismů uchování a dědičnosti genetické informace, definice centrálního

dogmatu molekulární biologie a pochopení principu genetického kódu v padesátých a šedesátých letech minulého století.

Historie sekvenování

V roce 1953 James D. Watson a Francis H. C. Crick sestavili molekulární strukturu nukleové kyseliny. První sekvenování krátké sekvence DNA provedl v letech 1968 až 1971 R. Wu a A.D. Kaiser (Wu R. and Kaiser A. D., 1968). Tato sekvence se skládala z pouhých dvanácti nukleotidů a pocházela z 3´ konce genomu fága lambda.

Průlom nastal na konci sedmdesátých let minulého století, kdy v roce 1977 angličtí vědci F. Sanger, S. Nicklen a A. R. Coulson vyvinuli enzymatickou sekvenační dideoxy metodu. V témže roce vyvinuli američtí vědci A. Maxam a W. Gilbert chemickou metodu sekvenování molekul DNA. Oba vědecké týmy získaly v roce 1980 Nobelovu cenu za chemii.

Prvním genomem se známou sekvencí byl genom bakteriofága PhiX 174, o velikosti 5375 bazí (Sanger F. *et al.*, 1977). V dnešní době jsou k dispozici sekvence genomů mnoha dalších organizmů.

Sangerova metoda

V roce 1973 Sanger využil pro sekvenování DNA princip syntézy enzymem DNA polymeráza I (Sanger F. *et al.*, 1973). V reakci byla syntéza iniciovaná krátkým syntetickým oligonukleotidem. Nahrazením Mg^{2+} za Mn^{2+} v reakci umožnilo DNA polymeráze I inkorporovat do nově syntetizovaného vlákna kromě deoxyribonukleotidů (dNTP) i ribonukleotidy (NTP). Jestliže reakce obsahovala jen jeden NTP spolu s dNTP bylo možné nově syntetizované vlákno štěpit pomocí RNázy A nebo alkalickou hydrolýzou v místech náhodné inkorporace NTP. Po uspořádání vzniklých fragmentů podle velikosti bylo možné stanovit místa použitého NTP v sekvenci. Pro určení celé sekvence bylo nutné použít reakce se čtyřmi různými NTP.

V roce 1975 Sanger a Coulson vypracovali “plus a minus” metodu (Sanger F. and Coulson A. R., 1975), při které byl během syntézy přítomen jeden dNTP ve výrazně nižší koncentraci. V důsledku nedostatku daného dNTP docházelo k předčasné terminaci syntetizovaného vlákna v místech, kde měl být tento dNTP inkorporovaný. To umožňovalo

stanovit sekvenci DNA. V roce 1977 byla tato metoda zdokonalena použitím chemických terminátorů syntézy DNA- dideoxynukleotidů (ddNTP) (Sanger F. *et al.*, 1977).

Další modifikace Sangerovy metody umožnily zvýšení její efektivity a souvisely s použitím v automatických sekvenátorech. První podstatnou změnou bylo nahrazení radioaktivního značení fluorescenčně značenými primery nebo terminátory (ddNTP), použití různých fluoroforů umožnilo sekvenování pouze v jedné reakci. Podstatným zlepšením bylo také nahrazení separace v polyakrylamidovém gelu kapilární elektroforézou s tekutým polymerem. Asymetrická amplifikace pomocí termostabilních polymeráz podstatně snížila množství templátu potřebného k jedné reakci.

V současné době je tato metoda nejpoužívanější a nejpřesnější sekvenační metodou, s nepatrnými modifikacemi se používá i v automatických sekvenátorech. DNA je obvykle připravena pomocí PCR amplifikace se specifickými primery pro oblast zájmu. DNA je poté cyklicky sekvenována. Délka čtených fragmentů dosahuje velikosti 700-1000 bazí s vysokou přesností a cena se pohybuje okolo 500$ za Mb (Tucker T. *et al.*, 2009).

Maxam-Gilbertova metoda

Metoda vyvinutá Maxamem a Gilbertem využívá pro separaci fragmentů také polyakrylamidový gel (Maxam A. M. and Gilbert W., 1977), příprava produktů se však od Sangerovy metody značně liší. Fragmenty dsDNA jsou na jednom konci značeny radioaktivním fosforem. Poté je vzorek rozdělen do čtyř reakcí, ve kterých je DNA specificky štěpena v daných pozicích. Délky fragmentů se stanovovaly elektroforetickou separací a autoradiograficky (Gilbert W. and Maxam A., 1977).

Touto metodou bylo možné získat sekvenci o délce 200-400 bazí.

Chemická metoda byla zpočátku velmi populární, neboť umožňovala přímé sekvenování DNA. Dnes již byla téměř nahrazena Sangerovou metodou, jednak díky přípravě syntetických oligonukleotidů a dalším důvodem bylo použití toxických chemikálií a problémy s jejich použitím v automatických sekvenátorech.

Sekvenační technologie nové generace

Sangerova metoda sekvenování sehrála důležitou roli v genomových projektech, včetně projektu lidského genomu. Prudký rozvoj funkční a komparativní genomiky, rozšíření palety

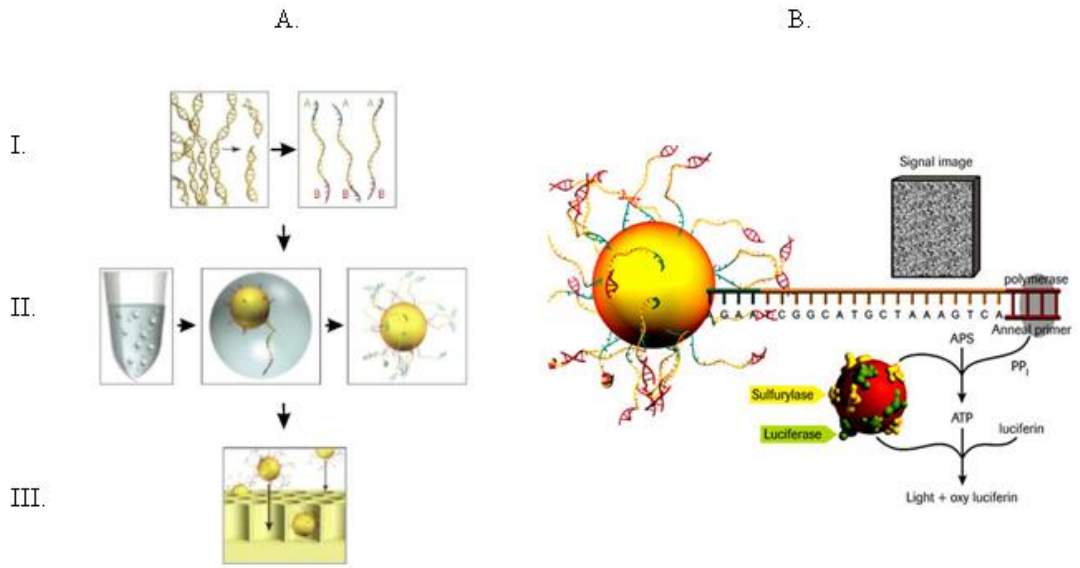
aplikací v diagnostice a perspektivy personalizované medicíny podnítily snahy o vývoj alternativních sekvenačních technologií, jejichž cílem je získávání dat podstatně rychlejším a levnějším způsobem. Nová generace sekvenačních technologií (next generation sequencing, NGS) využívá metody založené na různých principech - ligace oligonukleotidů (sequencing by ligation), syntézy DNA (sequencing by synthesis), hybridizace nukleových kyselin (sequencing by hybridization) nebo sekvenování pomocí nanopórů (nanopore sequencing). Všechny tyto technologie jsou v současné době dostupné.

Pyrosekvenování

Pyrosekvenace je jednou z prvních sekvenačních metod nové generace, která se začala rozvíjet jako alternativa ke klasické Sangerově metodě. V roce 1985 P. Nyrén navrhl, jak lze monitorovat aktivitu DNA polymerázy pomocí bioluminiscence (Nyren P. and Lundin A., 1985). Pyrosekvenováním se označuje série enzymatických reakcí, během kterých se zaznamenává viditelné záření vzniklé začleněním dNTP do syntetizovaného řetězce – inkorporace kteréhokoliv dNTP do syntetizovaného řetězce vede k uvolnění pyrofosfátu (PP), který je dále převeden na ATP pomocí enzymu ATP sulfurylázy. Vzniklé ATP je následně štěpeno luciferázou, která luciferin přemění na oxyluciferin. V reakci se uvolní kvantum viditelného světla, které je úměrné množství PP vzniklého při přeměně dNTP na dNMP. Světlo je zaznamenáno pomocí CCD kamery. Nespotřebované nukleotidy jsou v následujícím kroku odstraněny pomocí enzymu apyráza a sekvenační reakce pokračuje přidáním dalšího typu dNTP.

454 GenomeSequencer FLX instrument (Roche Applied Science)

Trvalo téměř jedenáct let, než byla pyrosekvenace sama o sobě uvedena do praxe. V roce 2000 byla představena technologie FLX Genome Sequencer (FLX GS) firmou 454 Life Sciences (http://454.com). Byla první komerčně dostupným zařízením pro masivní paralelní sekvenování. V roce 2011 přichází na trh další FLX GS série Titanium, která umožnila zlepšení výkonu prodloužením čtecího rámce, který je srovnatelný s kapilární Sangerovou metodou.



Obrázek 1. A. Schematická.příprava knihovny pro sekvenátor 454 GenomeSequencer FLX, .I. Ligace adaptorů na fragmenty DNA, II. navázání fragmentů na kuličky a emulzní PCR, III. nasedání amplifikovaných kuliček na PicoTiter Plate.
B. . Schematický postup pyrosekvenování. (http://454.com/products/technology.asp)

DNA fragmenty jsou ligovány se specifickými adaptéry, díky kterým je na každou mikrokuličku o velikosti 28 μm navázán právě jeden fragment. Pomocí emulzního PCR je DNA amplifikována ve vodních kapkách v oleji obsahujících jedinou kuličku s fragmentem DNA, polymerázou, dNTPs. Amplifikace je nutná pro získání dostatečné intenzity signálu. Dále jsou molekuly DNA na mikrokuličkách ukládány do pikolitrových jamek čipu PicoTiterPlate^{TM} (Jarvie T., 2005) tak, že ve většině jamek je pouze jedna kulička (obr. 1).

K vytvoření individuálních sekvenačních reaktorů jsou do jamek přidávány směsi menších kuliček, které obsahují nadbytek DNA polymerázy, ATP sulfurylázy a luciferázy. Tyto směsi jsou nezbytné k vytvoření světla z volného pyrofosfátu (Margulies M. *et al.*, 2005).

CCD kamera zachycuje světlo, které je uvolňováno ze sekvenační reakce. Mikrokulička obsahující 10 miliónů kopií templátu přitom emituje takové množství světla, že je CCD kamerou zachycováno přibližně 10 000 fotonů na jeden nukleotid začleněný do DNA řetězce. Každým obrazem snímaným CCD kamerou jsou získána data o velikosti až 30 MB (Jarvie T., 2005).

Nově uvedený přístroj 454 FLX Titanium zpětinásobil množství výstupních dat až na 500 MB a nový pikolitrový čip používá kuličky o průměru 1 μm.

Illumina (Solexa) Genome Analyzer

Sekvenační přístroj firmy Solexa byl představen v roce 2006 a o rok později jej odkoupila firma Illumina. Stejně jako tomu bylo u metody 454/Roche, je princip metody založen na sekvenaci syntézou (Ansorge W. J., 2009) a je podobný Sangerovu sekvenování.

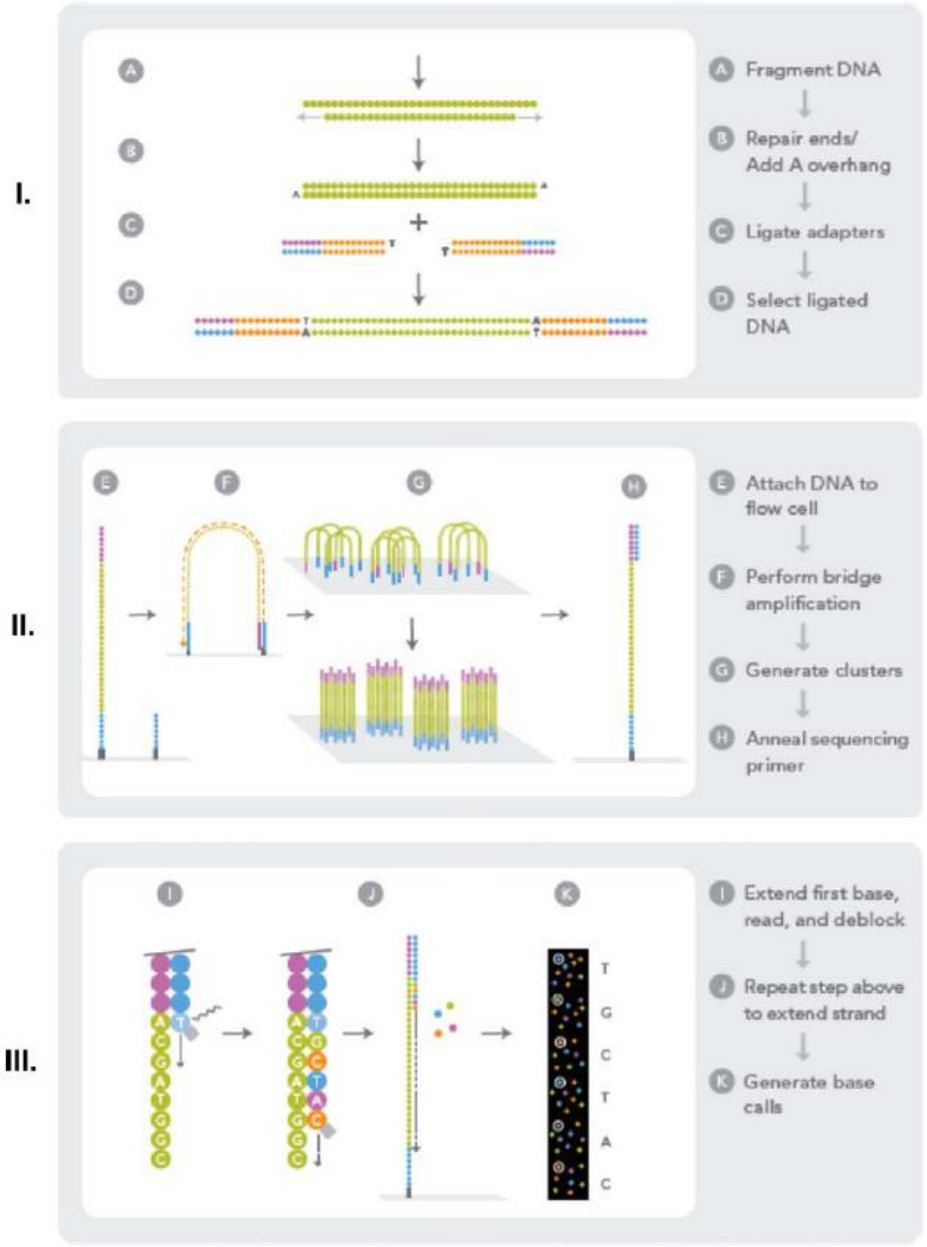
DNA fragmenty jsou po naligování specifických adapterů na oba konce a následné denaturaci imobilizovány jedním koncem na pevný podklad, který je hustě pokryt oligonukleotidy adapterů a jejich komplementárních řetězců. Fragmenty jsou následně amplifikovány v osmnácti cyklech PCR s použitím fúzní DNA polymerázy (Croucher N. J. *et al.*, 2009). Při amplifikaci, která probíhá na amplifikační destičce (flow cell), dochází po přidání enzymů k ohnutí fragmentu, vzniku tzv. můstku (bridge). Dochází tak ke vzniku shluků (cluster) každého fragmentu knihovny a výsledkem této izotermické amplifikace je sto milionů jedinečných svazků. V závěru můstkové amplifikace dochází na konci každého úseku DNA k navázání sekvenčního primeru.

Samotná sekvenace všech amplifikovaných částí DNA probíhá najednou v přístroji Illumina Genome Analyzer™ (obr. 2). Templáty DNA jsou sekvenovány báze po bázi pomocí čtyř odlišných fluorescenčních barviv. Postupně dochází k navázání jednotlivých fluorescenčně značených bází na templát. Po každém cyklu syntézy jsou shluky detekovány laserem přístroje a zaznamenávány CCD kamerou.

V roce 2010 Illumina představila nový sekvenátor HiSeq 2000, princip sekvenování je také založen na sekvenování syntézou pomocí reverzibilních terminátorů.

V roce 2007 bylo jedním sekvenčním během produkováno okolo 1Gb dat, v současnosti je produkováno téměř 1000x více sekvenčních dat a je možno sekvenovat více než 5 lidských genomů v jediném běhu trvajícím přibližně týden za cenu méně než 5000 USD za genom (Liu L. *et al.*, 2012).

Pro srovnání první lidský genom trvalo osekvenovat téměř 10 let a stálo necelé 3 miliardy amerických dolarů.



Obrázek 2. Princip sekvenátoru Illumina Genome Analyzer. (Ansorge W. J., 2009)
příprava knihovny- I. fragmentace, oprava konců a ligace adaptorů
II. nanesení fragmentů na povrch, amplifikace pomocí můstků, vytvoření shluků, hybridizace sekvenačních primerů
III. sekvenování syntézou

Applied Biosystems (AB) SOLiD™ System

Od října roku 2007 je možné sekvenování pomocí technologie AB SOLiD™ (Sequencing by Oligo Ligation and Detection). Zde je využito sekvenačního procesu katalyzovaného DNA ligázou. Modely firmy AB (5500 a 5500xl SOLiDTM System) slibují sekvenaci s přesností nad 99,99% (http://www.lifetechnologies.com/cz/en/home/life-

science/sequencing/next-generation-sequencing/solid-next-generation-sequencing.html). Sekvenační přístroj technologie AB SOLiD přečte řetězce o délce 25-35 bp. Přečtění dat o velikosti 3-4 GB obvykle trvá pět dní (Mardis E. R., 2008; Mardis E. R., 2008). Systém je také schopen přečíst sekvenci až 300 milionů fragmentů o délce dvakrát 50 bází (mate-pairs), v jednom běhu trvajícím 12-14 dní tak generuje až 30 GB dat (Pospíšilová Š. *et al.*, 2009).

Příprava DNA knihovny a amplifikace

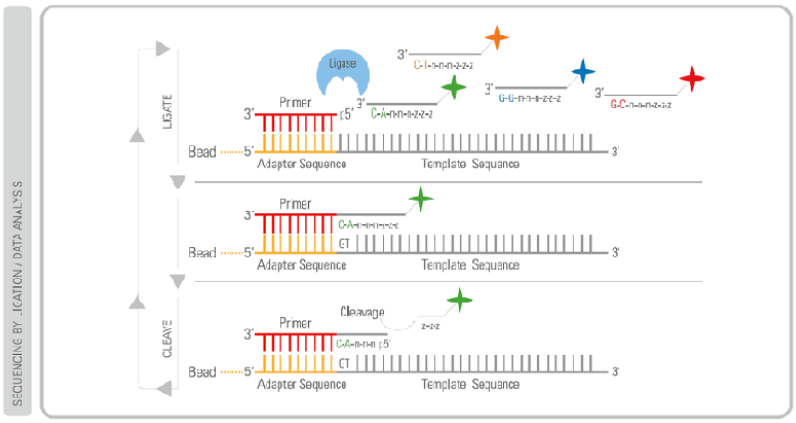
U metody SOLiD může příprava DNA knihovny probíhat několika způsoby. V současné době v Ústavu dědičných metabolických poruch používáme tzv. fragmentovanou DNA knihovnu. Po izolaci je DNA rozštěpena na fragmenty o délce 150-250 bp, na jejichž 5‘-konec se naváže 25 bp dlouhý adaptér P1 a na 3‘-konec se naváže stejně dlouhý adaptér P2 (Ansorge W. J., 2009).

K amplifikaci DNA fragmentů dochází obdobně, jako tomu bylo u metody 454/Roche. DNA úseky se naváží na 1μm mikrokuličky, které jsou následně podrobeny emulznímu PCR. Následně dochází k denaturaci šablon a kuličky s amplifikovanými fragmenty jsou odděleny od nežádoucích. Vybrané mikrokuličky jsou dále na 3’konci modifikovány, tak aby mohly vytvořit kovalentní vazbu s podkladem (Mardis E. R., 2008).

Postup sekvenace s metodou AB SOLiD™ System

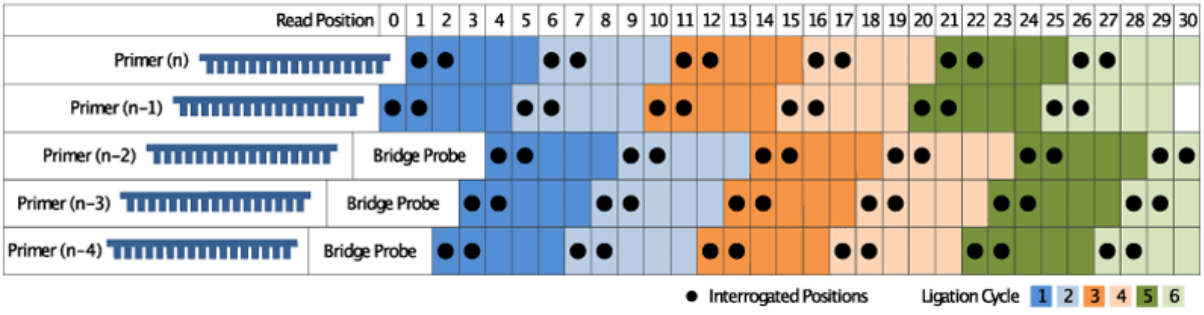
Sekvenování pomocí technologie SOLiD™ se od předchozích dvou metod liší, neboť využívá hybridizaci krátkých fluorescenčně značených sond. Značené sondy jsou definovány prvními dvěma dNTPs (Pospíšilová Š. *et al.*, 2009). V prvním kroku je primer hybridizován s adaptérem, a poté dochází k navázání směsi oligonukleotidových oktamerů. V těchto oktamerech jsou obsaženy dvojice čtvrtých a pátých bází, které jsou charakteristické jednou ze čtyř fluorescenčních barviv na konci oktameru (Ansorge W. J., 2009). Sada fluorescenčně značených dvoufázových sond soutěží o ligaci k sekvenačnímu primeru. Po fluorescenční detekci označených bazí jsou oktamery štěpeny za pátou bází, fluorescence je odstraněna, poté může být proces ligace opakován. Množství cyklů ligace, detekce fluorescence a štěpení ligovaných sond závisí na požadované délce přečtené sekvence (obr. 3). K přečtení kompletní sekvence je zapotřebí proces hybridizace a následné ligace zopakovat nejméně pětkrát. Tím, že je každá pozice v sekvenci charakterizována dvěma fluorescenčními signály, dochází k zajištění vyšší spolehlivosti určení báze. Po dosažení požadované délky je nově vzniklý

řetězec odstraněn a celý proces se opakuje s novým sekvenčním primerem komplementárním k n -1 pozici předchozího primeru (obr. 4).



Obrázek 3. Princip sekvenování ligací.
http://www.lifetechnologies.com/cz/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing/solid-next-generation-sequencing-systems-reagents-accessories/solid-next-generation-sequencing-chemistry.html

Jedinečná vlastnost sekvenace založené na ligaci a označování oktamerů je tzv. dvoubázové dekódování (two base encoding), pomocí něhož dokážeme v oktamerech identifikovat dNTP rozdíly a chyby v průběhu analýzy dat (Mardis E. R., 2008). Prostřednictvím procesu vyměňování primerů je detekována každá báze dvěma různými primery ve dvou nezávislých ligacích. Dvojí detekce je zásadní pro velkou přesnost, kterou se vyznačuje AB SOLiD™ Systém.

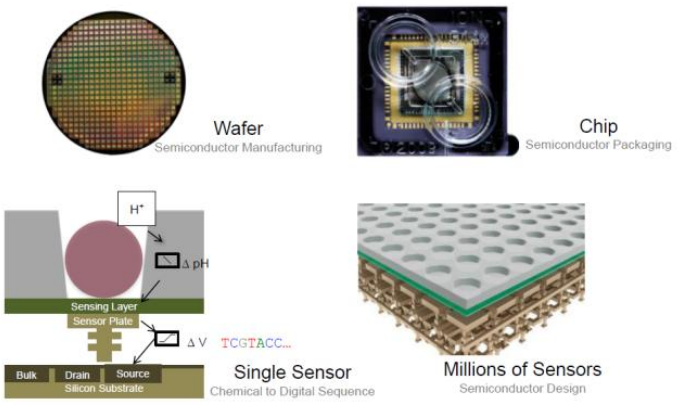


Obrázek 4. Postupná ligace primerů při sekvenování ligací.
http://www.lifetechnologies.com/cz/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing/solid-next-generation-sequencing-systems-reagents-accessories/solid-next-generation-sequencing-chemistry.html

Ion Torrent™ sekvenátory

Ion Personal Genome Machine (PGM) byla představena koncem roku 2010. PGM využívá principu polovodičového sekvenování (obr. 5). Při inkorporaci nukleotidu do nově vznikajícího řetězce pomocí polymerázy je uvolněn proton, který způsobí změnu pH roztoku,

která je zaznamenána ion senzorem umístěným na dně mikročipu a ihned převedena do digitálního formátu. PGM je první komerčně dostupná technologie, která nevyužívá flourescenci ani skenování CCD kamerou, což má za následek vysokou rychlost, nízké náklady a menší nároky na velikost sekvenčního přístroje. V současné době je schopna přečíst 200 nukleotidů za necelé 2 hodiny.



Obrázek 5. Princip sekvenátoru Ion Torrent (http://www.genomics.cn/en/navigation/show_navigation?nid=4147)

V porovnání všech systémů sekvenování nové generace má Illumina HiSeq největší množství výstupních dat a nejnižší cenu chemikálií, Solid systém je nejpřesnější a Roche 454 má nejdelší přečtené délky fragmentů (Liu L. *et al.*, 2012). Detaily jsou uvedeny v tabulce 2.

a)

Sekvenátor	454 GS FLX	Hiseq 2000	SOLiD4	Sanger 3730xl
Princip sekvenátoru	pyrosekvenování	Sekvenování syntézou	Sekvenování ligací a dvoufázové dekódování	Dideoxy terminace
Délka čtení	700 bp	50SE, 50PE, 101PE	50+35 bp nebo 50+50 bp	400-900 bp
Přesnost	99,9%	98% (100PE)	99,94% raw data	99,999%
Počet čtení	10^6	$3x10^9$	$1,2x10^9$-$1,4x10^9$	-
Výstupní data/běh	0,7 Gb	600 Gb	120 Gb	1,9 – 84 Kb
Doba běhu	24 h	3-10 dní	7 dní SE, 14 dní PE	20 min- 3 h
Výhody	Délka čtení, rychlost	Vysoká výkonnost	přesnost	Vysoká kvalita, dlouhé délky čtení
Nevýhody	Vysoká chybovost, vysoká cena, nízká výkonnost	Krátká čtení	Krátká čtení	Vysoká cena, nízká vykonnost

b)

Sekvenátor	454 GS FLX	Hiseq 2000	SOLiD4	Sanger 3730xl
Cena přístroje	Přístroj 500.000 USD, 7.000 USD běh	Přístroj 690.000 USD, 6.000 USD /(30x) lidský genom	Přístroj 495.000 USD, 15.000 USD/100 Gb	Přístroj 95.000USD, 4 USD/800 bp reakce
CPU	2x Intel Xeon X5675	2x Intel Xeon X5560	8x procesor 2,0 GHz	Pentium IV 3,0 GHz
Paměť	48 GB	48 GB	16 GB	1 GB
Hard disk	1,1 TB	3 TB	10 TB	280 GB
Automatická příprava knihovny	ano	ano	ano	ne
Přídavné přístroje	REM e systém	cBot systém	EZ bead systém	ne
Cena/milión bazí	10 USD	0,07 USD	0,13 USD	2400 USD

c)

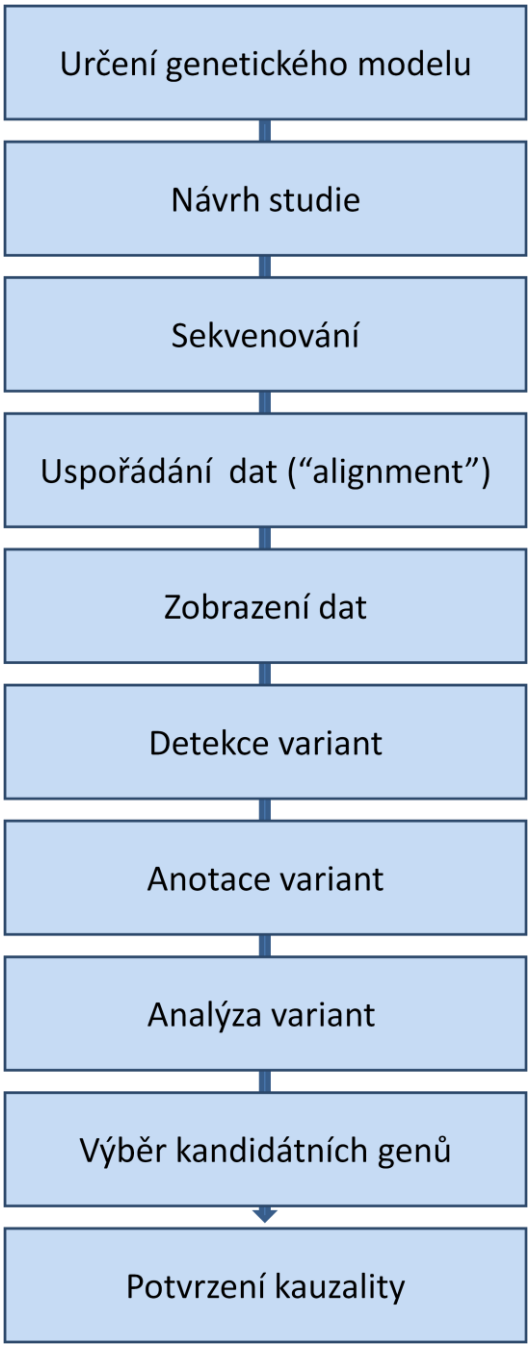
Sekvenátor	454 GS FLX	Hiseq 2000	SOLiD4	Sanger 3730xl
Resekvenování		ano	ano	
De novo sekvenování	ano	ano		ano
Panel vybraných genů	ano	ano	ano	
Vzorek GC bohatý	ano	ano	ano	
Velký genom	ano	ano		
Detekce mutací	ano	ano	ano	ano

Tabulka 2. a) Mechanismus a výhody sekvenátorů, b) Komponenty a ceny sekvenátorů, c) Využití sekvenátorů (Liu L. *et al.*, 2012)

Komplexní analýza genomů jako univerzální nástroj identifikace kauzálních genů a mutací u vzácných onemocnění

Nové metody sekvenování DNA umožnily analýzu desetitisíců genomů, případně jejich protein-kódujících oblastí. Tento vývoj odhalil základní strukturu genetické variability člověka a významným způsobem zjednodušil vyhledávání vzácných genetických mutací způsobujících vzácná onemocnění.

Současný přístup určení genetických příčin vzácných onemocnění vychází z univerzálního konceptu porovnávání sekvence genomu postižených jedinců s genetickou variabilitou různých populací (obr. 6).



Obrázek 6. Schematický přehled postupů vedoucích k identifikaci kauzálních genů

U vzácných onemocnění je výsledkem takového porovnání určení populačně vzácných či privátních genetických variant studovaných jedinců. Tyto varianty jsou následně hodnoceny na základě obecného genetického modelu onemocnění a segregace jednotlivých variant se studovaným onemocněním v rodině. Významným faktorem je opakovaný výskyt

vzácných variant ve vybraném kandidátním genu u nepříbuzných pacientů. Další možnosti poskytuje též analýza mezidruhové sekvenční konzervovanosti, předpokládaného efektu nalezených variant a funkční anotace genů s nalezenými variantami. Vhodným doplňkem kvalifikované analýzy a výběru kandidátních genů je též poziční informace definovaná např. paralelně provedeným genetickým mapováním, analýza RNA v postižených tkáních či tkáňových kulturách.

Výsledkem takového postupu je určení omezeného počtu kandidátních genů a variant, jejichž kauzalita musí být ve většině případů ještě následně studována a charakterizována pomocí vhodných biochemických, molekulárně biologických či histopatologických metod.

Automatizace sekvenace DNA a rychlý nárůst produkce sekvencí DNA vyžadovaly nové nástroje pro jejich analýzu a uchovávání. Tento požadavek vedl k založení The National Center for Biotechnology Information (NCBI), postupnému budování a udržování databáze literatury (MEDLINE), sekvencí DNA (GenBank) a vývoji nových bioinformatických nástrojů umožňujících přenos (FASTA), vzájemné porovnávání (BLAST 3, CLUSTAL V 16), interpretaci a veřejné sdílení sekvenčních dat DNA a proteinů (ENTREZ).

Identifikaci kauzálních genů umožnila též dostupnost panelů somatických hybridů, cDNA a gDNA knihoven a postupy funkční komplementace buněčných kultur pacientů se studovanými onemocněními (Strathdee C. A. *et al.*, 1992), případně vhodných buněčných modelů (zejména mutantních kvasinek a křeččích ovariálních buněk) nebo modelových organismů.

Základním předpokladem pro úspěšné použití všech výše uvedených přístupů byla znalost příslušného biochemického, biologického, případně histopatologického fenotypu a možnost jeho relativně snadného měření či sledování. Avšak tento předpoklad splňovala (a splňuje) pouze velmi malá část studovaných fenotypů.

ČÁST II. Identifikace kauzálních genů a mutací a studium molekulární podstaty vybraných vzácných onemocnění v Ústavu dědičných metabolických poruch v letech 1996 – 2013

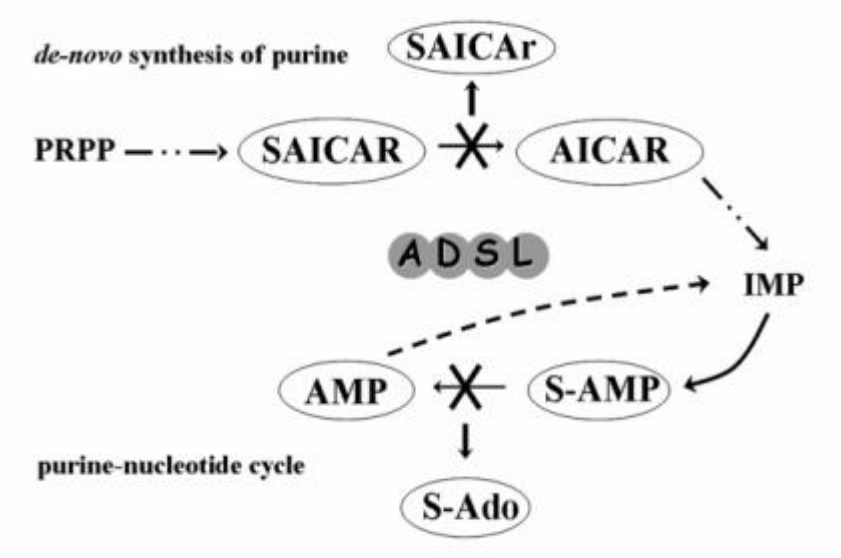
Cíle práce

1. Rozvoj technologie DNA čipu a její využití při studiu vzácných onemocnění.
2. Zavedení a využití postupu genetického mapování a pozičního klonování při studiu vzácných onemocnění.
3. Zavedení a využití celoexomového sekvenování.
4. Identifikace kauzálních genů a mutací a studium molekulární podstaty vybraných vzácných onemocnění (deficit ADSL, deficit ATP syntázy, MPSIIIC, Rotorův syndrom, ANCL, GAPO syndrom, X-vázaná restriktivní hypertrofická kardiomyopatie).

Studium molekulární podstaty deficitu adenylosukcinát lyázy

Deficit adenylosukcinát lyázy (ADSL) (OMIM 103050) je metabolická porucha *de novo* syntézy purinů a purin-nukleotidového cyklu popsaná v roce 1984 (Jaeken J. and Vandenberghe G., 1984). ADSL katalyzuje dva kroky syntézy purinových nukleotidů. Prvním je přeměna sukcinylaminoimidazolkarboxamidribotidu (SAICAR) na aminoimidazolkarboxamidribotid (AICAR) v *de novo* syntéze purinů a druhým vznik adenosinmonofosfátu (AMP) z sukcinyladenosinmonofosfátu (SAMP) v purin-nukleotidovém cyklu (obr.7).

Patogenetické mechanismy, které vedou k rozvoji individuálních klinických symptomů a způsobují fenotypovou heterogenitu, zůstávají nejasné. Hlavní patogenetický efekt je přisuzován hromadění toxických defosforylovaných substrátů ADSL – sukcinyladenosinu (S-Ado) a sukcinylaminoimidazolkarboxamidribosidu (SAICAr) v mozkomíšním moku a moči (Stone T. W. *et al.*, 1998). Přestože celkové koncentrace S-Ado a SAICAr nekorelují s tíží onemocnění, bylo zjištěno, že hodnoty poměru koncentrací S-Ado a SAICAr v mozkomíšním moku (S-Ado/SAICAr ratio) odpovídají třem základním fenotypickým skupinám deficitu ADSL (Mouchegh K. *et al.*, 2007; Van den Bergh F. *et al.*, 1993).



Obrázek 2. Schéma funkce ADSL v de novo syntéze purinů a v purin nukleotidovém cyklu. Příslušné enzymatické defekty jsou znázorněny křížky a zakresleny jsou alternativní metabolické cesty vedoucí ke vzniku SAICAr a S-Ado. PRPP - fosforibosyl pyrofosfát; AICAR - aminoimidazolkarboxamid ribotide; IMP- inosinmonofosfát (Zikanova M. *et al.*, 2010)

Typickými příznaky deficitu ADSL jsou vážné neurologické symptomy jako psychomotorická retardace, epilepsie, hypotonie a autismus. Stupeň postižení je u jednotlivých pacientů různý a klinicky klasifikovány jsou tři typy tohoto onemocnění: neonatální forma (prenatální hyperkineze, plicní hypoplázie a prenatální zastavení růstu, následovaná fatální neonatální encefalopatií) (Mouchegh K. *et al.*, 2007), těžká forma – typ I (těžká psychomotorická retardace, často brzká úmrtí) (Jaeken J. *et al.*, 1988; Maaswinkel-Mooij P. D. *et al.*, 1997), střední až mírná forma – typ II (psychomotorická retardace, hypotonie a autismus) (Jaeken J. *et al.*, 1988; Jaeken J. *et al.*, 1992; Valik D. *et al.*, 1997). Do roku 1995 bylo na světě popsáno pouhých 14 pacientů (Salerno C. *et al.*, 1995).

V předkládané práci se nám metodou CapFinder podařilo identifikovat úplnou sekvenci genu *ADSL* včetně jeho izoforem (GenBank, AF067853). Pro charakterizaci genomické struktury *ADSL* bylo pro predikci konzervovaných sestřihových míst v cDNA využito známé uspořádání myšího Adsl genu. Byly izolovány a sekvenovány genomické fragmenty pokrývající *ADSL* včetně promotoru, všech exonů a přilehlých částí intronů. Bylo zjištěno, že gen pro *ADSL* se skládá ze 13 exonů. Genomická struktura *ADSL* vykazuje identitu s myším genem. Provedli jsme první podrobnou studii 7 pacientů s deficitem ADSL zahrnující podrobný klinický popis a biochemickou charakterizaci pacientů a identifikovali jsme mutace podmiňující deficit ADSL. Klonováním a expresí rekombinantních mutantních i wild-type proteinů jsme posuzovali a charakterizovali vlastnosti jednotlivých proteinů ve vztahu k fenotypu jednotlivých pacientů (příloha 1). V souvislosti s uvedenou studií jsme

zároveň identifikovali a určili sukcinyladenosin jako normální složku likvoru (Krijt J. *et al.*, 1999), naznačili možnou roli sukcinylpurinů v patogenesi postižení centrální nervové soustavy, jak u deficitu ADSL, tak u deficitu fumarázy (Zeman J. *et al.*, 2000) a využili dostupnosti rekombinantního lidského enzymu k přípravě substrátů a jejich defosforylovaných produktů nezbytných pro studium patogeneze deficitu ADSL (Zikanova M. *et al.*, 2005).

Díky nově vyvinutému diagnostickému přístupu (Krijt J. *et al.*, 1999) se nám podařilo shromáždit největší světový souboru pacientů (http://www1.lf1.cuni.cz/udmp/adsl). Tento soubor zahrnuje pacienty s různou tíží klinického postižení a je zdrojem pro studium patogenetických mechanismů vzniku tohoto onemocnění a pro studium základních biologických mechanismů vzniku a funkce purinozómu.

Rozvoj technologie DNA čipů a její využití při studiu vzácných chorob.

V Ústavu dědičných metabolických poruch 1.LF UK a VFN v Praze byl v roce 1998 ve spolupráci s firmou GeneAge Technologies navržen, vyroben a testován přístroj pro tisk DNA čipů, GeneSurfer. Tento přístroj a jeho vlastní softwarové ovládání umožňovalo přesný 3D pohyb tiskové hlavy po pracovní ploše, na níž byly rozmístěny komponenty nutné k přípravě DNA čipů - zásoba sond v mikrotitračních destičkách, mikroskopická skla, sonikační stanice pro oplach a stanice pro sušení jehel.

Zavedli jsme technologii přípravy skel a sond tak, že jsme byli schopni připravit v jednom cyklu až 60 čipů a nanést sondy o maximální hustotě 4900 sond/cm^2. Podařilo se nám zoptimalizovat metodu imobilizace sond na povrch, metodu 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 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*.

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. H-MitoArray byl využit ke studiu genové exprese u fibroblastů 13 pacientů s popsaným defektem F1Fo syntázy a 9 kontrol. 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 mitochondriální DNA; expresní profil poukazoval na supresi mitochondriální biogeneze a metabolismu a sníženou expresi genů regulujících přechod z G1 do S fáze. Tyto výsledky podpořily hypotézu o regulaci buněčného cyklu mitochondriemi na transkripční a posttranskripční úrovni (Gemin A. *et al.*, 2005; Mandal S. *et al.*, 2005). U druhé skupiny dominovaly v expresním profilu známky aktivované apoptózy a oxidativního stresu, tedy charakteristiky buněčného stárnutí (Shelton D. N. *et al.*, 1999; Stockl P. *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 (příloha 2).

V ÚDMP jsme měli také možnost studovat pacienty s mukopolysacharidózou typu IIIC (MPSIIIC, Sanfilippo syndrom C, OMIM #252930). Jedná se o vzácné autozomálně recesivní lysozomální střádavé onemocnění. Měli jsme k dispozici celkem pět pacientů ze čtyř nepříbuzných rodin, u nichž bylo onemocnění diagnostikováno na základě biochemického vyšetření aktivity N-acetyltransferázy (Voznyi Ya V. *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. 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 sondy 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 J. *et al.*, 2004).

Sekvenováním kandidátního genu byly pak identifikovány patogenní mutace v genu *TMEM76* nejprve u pěti studovaných českých pacientů, později celkem u třiceti rodin. Jednalo se o 4 nonsense mutace, 14 missense mutací, 3 mutace způsobující posun čtecího rámce a 6 sestřihových mutací. Funkční význam genu *TMEM76* v patogenezi mukopolysacharidózy typu IIIC byl dále potvrzen funkčními expresními studiemi s použitím fibroblastů pacientů (Hrebicek M. *et al.*, 2006).

DNA čipy byly v ÚDMP použity i pro metodu komparativní genomové hybridizace, a to při studiu molekulární podstaty Rotorova syndromu (Hrebicek M. *et al.*, 2007), kterým jsme se zabývali i v následujícím roce (viz dále).

Expresní čipy připravené na ÚDMP našly využití také v zemědělství. Ve spolupráci s Jihočeskou Universitou v Českých Budějovicích jsme vytvořili oligonukleotidový čip schopný efektivně a spolehlivě detekovat viry v listech brambor (Sip M. *et al.*, 2010).

Využití postupů genetického mapování a pozičního klonování

Vzhledem k tomu, že analýza genové exprese u pacientů s izolovaným defektem ATP syntázy nepřinesla jasné kandidátní geny pro přímé sekvenování, pokračovali jsme v dalším studiu. U osmi pacientů, jejich sourozenců a rodičů ze sedmi rodin jsme provedli vazebnou analýzu a 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 ve fibroblastech u všech pacientů sníženou expresi v porovnání s kontrolními buňkami. Sekvenováním tohoto genu byla u všech pacientů nalezena homozygotní substituce c.317-2A>G lokalizovaná v sestřihovém místě exonu 2. Dalšími studiemi bylo prokázáno, že tato mutace vede k abnormálnímu sestřihu a následné degradaci mRNA mechanismem „nonsense-mediated decay“. Pomocí PCR-RFLP analýzy jsme prokázali přítomnost homozygotní mutace u 23 pacientů z celkového počtu 25. Kauzalita této mutace byla prokázána komplementační

studií, obnovením funkce ATP syntázy po vnesení wild-type *TMEM70* do fibroblastů pacientů.

Identifikovali jsme TMEM70 jako protein účastnící se u vyšších eukaryot biogeneze ATP syntázy a prokázali jsme, že porucha *TMEM70* je relativně častá u pacientů s poruchou tvorby energie, především v romské populaci (příloha 3).

Další onemocnění, kterým jsme se v ÚDMP zabývali, byl již zmiňovaný Rotorův syndrom (RS, MIM237450). Toto onemocnění jsme začali studovat již v roce 2005 ve spolupráci s Institutem klinické a experimentální medicíny, kde byli zachyceni pacienti a kde byl následně díky zahraniční spolupráci shromážděn rozsáhlý soubor rodin a sporadických případů s touto diagnózou.

Rotorův syndrom je vzácné benigní autozomálně recesivní onemocnění, které se projevuje konjugovanou hyperbilirubinémií, koproporfyrinurií a sníženou absorpcí diagnostických sloučenin játry. U tohoto onemocnění se také vyskytuje prodloužená doba odbourávání nekonjugovaných anionických diagnostických barev (bromsulfathalein, indocyanová zeleň) a při cholescintigrafii nedochází k zobrazení jater a žlučovodů (Bar-Meir S. *et al.*, 1982; Fretzayas A. M. *et al.*, 1994).

Molekulární podstata RS nebyla dosud známa. Toto onemocnění svými projevy připomíná Dubin Johnsonův syndrom (DJS, MIM 237500), který je způsoben mutacemi v genu ABCC2. Proto vzhledem k ne zcela jasnému vymezení Rotorova syndromu a Dubin-Johnsonova syndromu, byla nejdříve testována hypotéza alelických variant stejného onemocnění reprezentovaných RS a DJS. Imunohistochemické nálezy potvrdily normální lokalizaci proteinu ABCC2 a sekvenování kódujících a promotorových oblastí genu *ABCC2* však nepotvrdilo přítomnost patogenní mutace v tomto genu.

Pro detekci změny počtu kopií exonů několika vybraných genů, včetně *ABCC2*, byl na našem pracovišti připraven čip pro komparativní genomovou hybridizaci. Ovšem ani u jednoho ze dvou pacientů s RS nebyly nalezeny změny v počtu kopií v žádném z 32 exonů *ABCC2* (příloha 3). Výsledky získané komparativní genomovou hybridizací, společně s dalšími mutačními a imunohistochemickými analýzami potvrdily, že v případě Rotorova syndromu a Dubin-Johnsonova syndromu se nejedná o alelické varianty jednoho onemocnění.

Proto jsme přistoupili ke genotypování pacientů z osmi dostupných rodin, ve kterých byl diagnostikován Rotorův syndrom. Homozygotní mapování odhalilo u pacientů ze všech

rodin jedinou homozygotní oblast na chromozomu 12. Po provedení analýzy změn počtu kopií byly u dvou haplotypů objeveny homozygotní delece- u prvního haplotypu delece v genu *SLCO1B3* a u druhého zasahovala delece 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. Pro zjištění závažnosti onemocnění byly provedeny funkční studie – imunochemická barvení jaterních biopsií pacientů prokázaly nepřítomnost proteinů OATP1B1 a OATP1B3 kódovanými geny *SLCO1B3* a *SLCO1B3*. Proteiny OATP1B3 a OATP1B1 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, některé léky, toxiny a jejich konjugáty (Hagenbuch B. and Meier P. J., 2004; Hagenbuch B. and Gui C., 2008).

Na základě těchto výsledků lze říci, že genetickou podstatou vzniku Rotorova syndromu jsou mutace genů *SLCO1B3* a *SLCO1B3*, které způsobují absenci proteinů OATP1B1 a OATP1B3. Tyto výsledky byly korelovány s výzkumem holandské skupiny z The Netherlands Cancer Institute v Amsterdamu, která se zabývala studiem myšího modelu s deficiencí proteinů Oatp1a/1b, myších homologů lidských proteinů OATP1B1 a OATP1B3. U Slco1a/1b-/- myši byly pozorovány zvýšené hodnoty plazmového bilirubinu, ty byly významně sníženy u Slco1a/1b; Abcc3-/- myši, čímž bylo prokázáno, že Abcc3 protein se podílí na zvýšených hodnotách bilirubinu v plazmě. Protein Abcc3 se podílí na vylučování konjugovaného bilirubinu z hepatocytů zpět do krve a proteiny Oatp1a a Oatp1b participují na transportu bilirubinu z krve zpět do hepatocytů. Vnesení lidského OATP1B3 nebo OATP1B1 proteinu do Slco1a/1b-/- myši vyvolalo normalizaci hladin bilirubinu v krvi i moči, což potvrzuje, že se oba lidské proteiny OATP1B3 i OATP1B1 podílejí na absorbci konjugovaného bilirubinu z plazmy v játrech.

Na základě výsledků těchto studií byla vyslovena hypotéza o exkreční dráze bilirubinu jakožto složitějším procesu než je jednosměrný transport bilirubinu z krve do žluče. Hypotéza vnáší do mechanismu transportu cyklus absorpce, sekrece do krve, reabsorpce a sekrece do žluče. Bilirubin vstupuje z krve do periportálních hepatocytů, kde je glukuronidován. Velká část konjugovaného bilirubinu v hepatocytech je vylučována prostřednictvím přenašeče ABCC3 zpět do krve, odkud je zpětně reabsorbována perivenózními hepatocyty za účasti proteinů OATP1B3 a OATP1B1, následný transport do žluče je zprostředkován transportérem

ABCC2. Jsou-li oba transportní proteiny OATP1B3 a OATP1BA defektní, dochází k poruše reabsorpce bilirubinu a k hromadění jeho konjugované formy v krvi, což je projevem Rotorova syndromu (příloha 4).

Tato studie je významná nejen objevením podstaty RS, ale má také širší klinický dopad. Přestože je výskyt kombinace defektů v obou genech velmi vzácný, výskyt mutací v jednom nebo druhém genu je v populaci 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 E. *et al.*, 2008).

Exomové sekvenování

Využití celoexomového sekvenování se velkou měrou podílí na objasnění genetické podstaty sporadických onemocnění a vzácných syndromů. U vzácných onemocnění 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 M. W. and Crowell S. L., 2000). Soubory pacientů s daným syndromem nebo onemocněním jsou často tvořeny pouze desítkami nepříbuzných jedinců. Z tohoto důvodu je dobré provádět celoexomové sekvenování rodičů a postiženého jedince, výsledný seznam kandidátních variant filtrovat s ohledem na možné typy dědičnosti a na přítomnost variant, které se u rodičů nevyskytují. Výsledky celoexomového sekvenování pacientů lze také porovnat s referenčními databázemi variant v populaci. Odhaduje se, že s výjimkou *de novo* vzniklých unikátních variant, jsou již v databázích přítomny téměř všechny varianty společně s jejich populačními frekvencemi (Abecasis G. R. *et al.*, 2010). Při hledání kauzální varianty pro vzácný sporadický případ je třeba nalézt buď unikátní variantu nepřítomnou v databázi, nebo variantu, jejíž frekvence přibližně odpovídá počtu popsaných případů.

Prvním projektem, ve kterém bylo v Ústavu dědičných metabolických poruch použito exomové sekvenovní, bylo studium molekulární podstaty adultní formy neuronální ceroidní lipofuscinózy (ANCL). V tomto projektu jsme využili kombinaci vazebné analýzy, expresní analýzy a exomového sekvenování. Vazebnou analýzou jsme nejprve identifikovali 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, nebyly nalezeny žádné potenciálně patogenní rozsáhlé delece nebo duplikace. Paralelně byla provedena analýza genové exprese DNA z leukocytů čtyř pacientů a čtyř kontrol. Tento experiment poskytl seznam rozdílně 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. I po těchto analýzách však zůstávala oblast zájmu příliš široká pro nalezení kauzálního genu. Bylo tedy přistoupeno k celoexomovému sekvenování jednoho pacienta na sekvenátoru SOLiD 4 v institutu CeGaT v Tubingen v Německu. Tato analýza objevila u tohoto jedince celkem 957 unikátních změn (SNPs a indel záměn).

Po propojení výsledků vazebné analýzy, expresní analýzy a exomového sekvenování se podařilo identifikovat jeden kandidátní gen *DNAJC5* ležící v kandidátní oblasti na chromozomu 20q13.33. Exprese tohoto genu byla v leukocytech pacienta ve srovnání s kontrolami statisticky významně zvýšena, navíc zde byla objevena unikátní heterozygotní mutace c.346_348delCTC (p. Leu116del). Segregace této záměny byla u ostatních členů rodiny ověřena Sangerovým sekvenováním.

DNAJC5 kóduje cysteine-string protein alpha (CSPα), membránový protein lokalizovaný v synaptických membránách neuronů. Jedná se o velmi konzervovaný protein, který se účastní inhibice neurodegenerativních procesů. Jeho deplece vede u myši a u *Drosophily* k neurodegeneraci a redukci délky života.

Funkčními studiemi na tkáňových kulturách a imunohistochemickým barvením ve tkáních jsme prokázali, že mutace v genu *DNAJC5* jsou příčinou adultní formy neuronální ceroidní lipofuscinózy. Zároveň byla prokázána neuroprotektivní funkce CSPα na modelu vzácného onemocnění a tímto způsobem 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 (příloha 5).

Identifikací mutací v genu *DNAJC5* byla objasněna zhruba čtvrtina případů NCL v našem souboru pacientů, rozhodli jsme se použít stejný přístup, popsaný v předchozí studii,

u rodiny s diagnózou autozomálně dominantní Kufsovy choroby (rodina UCL568, (Noskova L. *et al.*, 2011).

Nejprve byla provedena vazebná analýza, jejímž výsledkem byla identifikace 14 kandidátních oblastí s pozitivním LOD skóre na chromozómech 3, 4, 8, 9, 10, 13, 14, 16 a 19. Zároveň byla u jednoho pacienta provedena analýza změn počtu kopií pomocí Affymetrix GeneChip Mapping 6.0 Array přičemž nebyla nalezena žádná potenciálně patogenní rozsáhlá delece nebo duplikace. Bylo provedeno celoexomové sekvenování dvou pacientů a jednoho nepostiženého příbuzného na sekvenátoru SOLiD™4 v ÚDMP, za použití exomového obohacení pomocí Agilent SureSelect All Exome Kit. Sekvenováním bylo identifikováno 11723 jednonukleotidových záměn a 150 inzercí/delecí přítomných v heterozygotním stavu u obou pacientů a nepřítomných u zdravého příbuzného, z tohoto počtu pouze 65 jednonukleotidových záměn bylo 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 byla jako prioritní kandidátní záměna označena heterozygotní mutace c.509C>T (p.Ser170Phe) v genu presenilin 1 (*PSEN1*). Tato potenciálně patogenní mutace je přítomna v databázi dbSNP i v databázi Disease & Frontotemporal Dementia Mutation Database (http://www.molgen.ua.ac.be). *PSEN1* je jedním ze čtyř genů, u něhož jsou mutace spojovány s familiární Alzheimerovou chorobou s časným nástupem (Bertram L. and Tanzi R. E., 2012).

Pomocí funkční anotace variant, které byly nalezeny exomovým sekvenováním, byl u všech testovaných pacientů identifikován známý polymorfismus c.C173>T (rs17571) v genu kathepsin D (*CTSD*), který byl nedávno popsán jako rizikový faktor pro vznik Alzheimerovy choroby (Schuur M. *et al.*, 2011). U CLN10, kongenitální a pozdně infantilní varianty NCL, byly popsány změny v *CTSD* genu jako příčinné mutace vzniku tohoto onemocnění (Jalanko A. and Braulke T., 2009). Cílem dalšího výzkumu je zjistit, jaký vliv má tento polymorfismus na projevy a průběh onemocnění.

V této části studie nebyl objeven další kauzální gen pro adultní formu neuronální ceroidní lipofuscinózy, rodina zařazená do souboru byla nesprávně diagnostikována a byla u ní identifikována již známá mutace podmiňující Alzheimerovu chorobu. Metodický přístup ale potvrdil, že správnou kombinací různých metodických přístupů (vazebné analýzy a exomového sekvenování) je velmi vysoká pravděpodobnost rychlého a efektivního zjištění genetické podstaty daného onemocnění (Ehling R. *et al.*, 2013).

Dalším studovaným onemocněním byl GAPO syndrom (OMIM 230740) (Growth retardation, Alopecia, Pseudoanodontia, and progressive Optic atrophy), vzácné autozomálně recesivní onemocnění, které je popisováno jako komplex růstové retardace, alopecie, pseudoanodoncie s progresivní optickou atrofií. Doposud bylo popsáno kolem třiceti případů na celém světě (Demirgunes E. F. *et al.*, 2009; Goloni-Bertollo E. M. *et al.*, 2008; Kocabay G. and Mert M., 2009; Nanda A. *et al.*, 2010; Sayli B. S. and Gul D., 1996; Tipton R. E. and Gorlin R. J., 1984). Většina jedinců s tímto onemocněním pochází z příbuzenských svazků. Molekulární podstata tohoto syndromu nebyla dosud známa.

V ÚDMP jsme díky spolupráci s Klinikou dětského a dorostového lékařství měli možnost studovat pacienta s GAPO syndromem. Nejprve jsme provedli analýzu počtu kopií a homozygotní mapování pomocí Affymetrix GeneChip Mapping 6.0 Array. U pacienta ani u rodičů nebyly zjištěny žádné inzerce nebo delece, byly identifikovány dvě homozygotní oblasti na chromozomech 2 a 4. Následně bylo provedeno celoexomové sekvenování pacienta i rodičů na sekvenátoru SOLiD™4. Analýza exomových dat ukázala celkem 121 kandidátních variant splňujících kritérium autozomálně recesivní dědičnosti, pouze jedna varianta se vyskytovala v homozygotní oblasti vymezené homozygotním mapováním. Jednalo se o homozygotní záměnu v genu pro antrax toxin receptor 1 (*ANTXR1*) vedoucí k vytvoření předčasného stop kodonu v exonu 7 (c.C505>T; p.R169X). Mutace nebyla přítomna v žádných běžně dostupných databázích. Díky mezinárodní spolupráci se podařilo vytvořit rozsáhlejší soubor pacientů s tímto syndromem, u nichž byla nalezena homozygotní mutace v genu *ANTXR1* (c.C262>T; p.R88X a sestřihová mutace c.1435–12A>G).

ANTXR1 (TEM8, tumor endothelial marker 8) byl původně popsán jako tumorově specifický endoteliální marker, jehož exprese je zvýšená během procesu nádorové angiogeneze (Carson-Walter E. B. *et al.*, 2001; St Croix B. *et al.*, 2000). Zároveň byl popsán jako receptor pro toxin Bacillus anthracis (ATR, anthrax toxin receptor) (Bradley K. A. *et al.*, 2001). *ANTXR1* kóduje transmembránový glykoprotein typu I lokalizovaný na plazmatické membráně. Mezi jeho funkce patří zprostředkování interakce buňky s komponentami extracelulární matrix (Hotchkiss K. A. *et al.*, 2005), vazba ligandů k aktinovému cytoskeletu (Yang M. Y. *et al.*, 2011) a regulace buněčné adheze (Gu J. *et al.*, 2010; Verma K. *et al.*, 2011; Werner E. *et al.*, 2006).

Funkční studie provedené na fibroblastech a tkáních pacientů s předčasným stop kodonem ukázaly významné snížení množství transkriptu *ANTXR1* a absenci proteinu ve tkáních, ta byla potvrzena imunohistochemicky i pomocí analýzy Western blot. Imunofluorescenční analýza ukázala u GAPO pacientů výraznou změnu v organizaci aktinových cytoskeletárních mikrofilament. Lze předpokládat, že ANTXR1 je zásadní pro uspořádání aktinových vláken, narušení aktinové sítě by mohlo být příčinou patologických jevů popsaných u GAPO syndromu.

V literatuře byly popsány dva myší modely s mutacemi v myším homologu Antxr1. Vyřazení exonu 13 kódující transmembránovou doménu Antxr1 mělo za důsledek defekty ve vývoji řezáků a samičí sterilitu (Liu S. *et al.*, 2009). Celková delece Antxr1 vedla k abnormálnímu ukládání extracelulární matrix v mnoha tkáních a k dentální dysplázii (Cullen M. *et al.*, 2009).

Identifikace tří odlišných mutací v genu *ANTXR1*, funkční dopady jedné z mutací na tkáňové úrovni, údaje o myších modelech a jejich korelace s fenotypovými charakteristikami podporují kauzalitu mutací v *ANTXR1* u GAPO syndromu. Z klinického hlediska tato práce přinese možnost genetického testování v postižených rodinách, v rovině základního výzkumu poskytuje unikátní model pro studium funkce proteinu ANTXR1/TEM8 (příloha 6).

Zatím posledním studovaným onemocněním byla X-vázaná restriktivní hypertrofická kardiomyopatie, kterou jsme se zabývali ve spolupráci s Klinikou kardiologie Institutu experimentální a klinické medicíny v Praze. Výzkum byl zaměřen na rodinu se třemi muži postiženými hypertrofickou kardiomyopatií asociovanou s těžkou dysfunkcí levé komory srdeční.

Vzhledem ke známé genetické heterogenitě familiárních kardiomyopatií, jsme provedli celoexomové sekvenování na sekvenátoru SOLiD™4 u dvou postižených chlapců a jejich matky.

Ze sekvenačních dat jsme následným filtrováním získali jedinou potenciálně patogenní variantu - inzerci c.599_600insT v exonu 6 genu *FHL1*. (NM_001159702). *FHL1* kóduje four-and-a-half-LIM-domain protein 1 (FHL1), je transkribován do tří alternativně sestřižených izoforem, translatovaných do proteinů FHL1A, FHL1B a FHL1C. Nejčastěji se vyskytuje izoforma FHL1A, která je exprimována především v kosterním svalu, méně již ve svalu srdečním. Inzerce c.599_600insT v exonu 6 genu *FHL1* způsobuje posun čtecího

rámce translatovaných izoforem FHL1A a FHL1B, v obou případech vzniká předčasný stop kodón.

Mutace ve *FHL1* byly popsány u pěti různých X-vázaných myopatií- myopatie s redukujícími tělísky [MIM 300718], skapuloperonealní myopatie [MIM 300695], X-vázané myopatie s posturální svalovou atrofií [MIM 300696], "rigid spine" syndromu a Emery-Dreifussovy svalové dystrofie [MIM 310300] (Cowling B. S. *et al.*, 2011; Gueneau L. *et al.*, 2009; Knoblauch H. *et al.*, 2010; Shathasivam T. *et al.*, 2010; Windpassinger C. *et al.*, 2008).

K charakterizaci identifikované mutace jsme provedli expresní analýzu *FHL1* izoforem a potvrdili přítomnost FHL1 proteinů v tkáních pacientů a kontrol. Provedli jsme kvantitativní PCR a RT-PCR analýzu z celkové RNA a Western blot analýzu z homogenátů proteinů zmrazených srdečních tkání.

Tyto analýzy ukázaly, že identifikované mutace nemají vliv na transkripci, sestřih a stabilitu *FHL1* mRNA a potvrdily přítomnost jediného RT-PCR produktu odpovídajícího FHL1A izoformě. Sangerovým sekvenování jsme potvrdili sekvenci izoformy FHL1A a zároveň i přítomnost inzerce c.599_600insT vedoucí k posunu čtecího rámce. WB analýza srdečních homogenátů ukázala přítomnost imunoreaktivního proteinu odpovídajícího molekulární hmotnosti 27 kDa, předpokládaného zkráceného proteinu vzniklého posunem čtecího rámce. Imunoreaktivní protein o molekulové hmotnosti 32 kDa odpovídající FHL1A izoformě byl identifikován ve vysoké míře u kontrolních vzorků a u pacientů nebyl nalezen. Žádný imunoreaktivní protein o molekulové hmotnosti 22 kDa odpovídající izoformě FHL1C nebyl detekován ani u kontrol, ani u pacientů. Imunohistochemické barvení srdeční tkáně potvrdilo nepřítomnost všech forem imunoreaktivního proteinu FHL1 u pacientů.

Naše výsledky vypovídají o různých rolích FHL1 proteinů v kosterním a srdečním svalu. V kosterním svalstvu má FHL1 mnoho funkcí v procesu migrace myoblastů a jejich prodlužování, při aktivaci buněk a inhibici apoptózy myoblastů, regulaci hmoty kosterního svalstva, tvorby sarkomer a Notch signalizaci (Cowling B. S. *et al.*, 2011; Shathasivam T. *et al.*, 2010).

Klinické studie ukázaly, že pokud je mutace přítomna na C konci FHL1, myopatický fenotyp je méně závažný, než při přítomnosti na N konci. Bohužel jsme neměli k dispozici materiál pro studium exprese proteinu FHL1 v kosterním svalstvu našich pacientů.

Missense mutace mající za následek rozpad domén vázajících zinek, které jsou důležité pro formování terciální struktury FHL1, se projevují jako závažnější myopatie než myopatie

způsobené mutacemi vedoucí ke zkrácení FHL1 (Cowling B. S. *et al.*, 2011). Toto tvrzení bylo doloženo případem, kdy byl nahrazen vysoce konzervovaný cystein, který váže zinkové ionty v LIM3 doméně (c.625T>C; p.C209R). Přestože tato mutace postihuje stejný exon a LIM doménu, stejně jako mutace zjištěná u našeho případu, způsobuje nejen hypertrofickou kardiomyopatii, ale také Emery-Dreifussovu myopatii s cytoplasmatickými tělísky tvořenými chybně skládaným mutovaným proteinem FHL1 (Knoblauch H. *et al.*, 2010). Toto pozorování poskytuje alternativní hypotézu, že mutace, které vedou k redukci (nebo dokonce absenci) FHL1 proteinu způsobují méně závažná (nebo dokonce benigní, jako v našem případě) poškození kosterního svalstva, než missense mutace, jejichž patogenický efekt je způsoben chybným skládáním, self agregací a koagregací s proteiny, které se vážou do FHL1 (Cowling B. S. *et al.*, 2011; Gueneau L. *et al.*, 2009; Shathasivam T. *et al.*, 2010) (příloha 7).

Závěr

Moje práce významně přispěla k zavedení základních molekulárně biologických technik: PCR, klonování, sekvenování a exprese rekombinantních proteinů, celoexomového sekvenování a bioinformatické analýzy v laboratoři Ústavu dědičných metabolických poruch 1. lékařské fakulty UK a VFN v Praze. Zásadně jsem se podílela na zavedení metod genotypování, přípravy oligonukleotidových čipů, na optimalizaci přípravy vzorků a jejich fluorescenčního značení a hybridizačních experimentech, a na zavedení a rozvoji sekvenačních technik pro sekvenátory nové generace. Účastnila jsem se také bioinformatické analýzy výsledků.

Všechny tyto techniky dnes tvoří základní metodické zázemí pro kvalifikované studium molekulární podstaty širokého spektra vzácných onemocnění a umožňují rychlé přenášení metodických přístupů do dalších studií.

Souhrnnými výsledky této práce jsou:

1. Identifikace sekvence genu *ADSL* včetně jeho izoforem a identifikace mutací podmiňujících deficit ADSL. Byla provedena první podrobná studie pacientů s deficitem ADSL zahrnující podrobný klinický popis a biochemickou charakterizaci pacientů a vývoj nové diagnostické metody.
2. Zavedení čipu H-MitoArray pro studium mitochondriálních onemocnění, s jehož pomocí, v kombinaci s využitím vazebné analýzy a homozygotního mapování byl identifikován kauzální gen *TMEM70* pro vznik izolovaného deficitu ATP syntázy.
3. Identifikace kauzálního genu pro vznik mukopolysacharidózy typu IIIC - TMEM76, kombinací vazebné analýzy, genotypování, expresních DNA čipů a sekvenování.
4. Identifikace molekulární podstaty Rotorova syndromu s využitím metody komparativní genomové hybridizace, homozygotního mapování a sekvenování, nalezení delece a mutace v genech *SLCO1B1* a *SLCO1B3*.
5. Identifikace mutací v genu *DNAJC5* jako kauzální příčiny části případů s adultní formou neuronální ceroidní lipofuscinózy pomocí kombinace vazebné analýzy, analýzy genové exprese, analýzy změn v počtu kopií a celoexomového sekvenování.
6. 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.
7. Identifikace mutací v genu *ANTXR1* jako kauzální příčiny GAPO syndromu pomocí kombinace analýzy počtu kopií, homozygotního mapování a celoexomového sekvenování.

8. Identifikace a charakterizace mutace v *FHL1* genu způsobující X-vázanou restriktivní hypertrofickou kardiomyopatii pomocí celoexomového sekvenování a expresní analýzy.

Seznam příloh

Příloha 1:

Kmoch, S., **H. Hartmannova**, B. Stiburkova, J. Krijt, M. Zikanova and I. Sebesta. "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." Human Molecular Genetics 9, no. 10 (2000): 1501-1513.

Příloha 2:

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. A. Mayr, S. Seneca, J. Houstek and S. Kmoch. "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, (2008).

Příloha 3:

Cizkova, A., V. Stranecky, J. A. Mayr, M. Tesarova, V. Havlickova, J. Paul, R. Ivanek, A. W. 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 and S. Kmoch. "Tmem70 Mutations Cause Isolated Atp Synthase Deficiency and Neonatal Mitochondrial Encephalocardiomyopathy." Nature Genetics 40, no. 11 (2008): 1288-1290.

Příloha 4:

van de Steeg, E., V. Stranecky, **H. Hartmannova,** L. Noskova, M. Hrebicek, E. Wagenaar, A. van Esch, D. R. de Waart, Rpjo Elferink, K. E. Kenworthy, E. Sticova, M. al-Edreesi, A. S. Knisely, S. Kmoch, M. Jirsa and A. H. Schinkel. "Complete Oatp1b1 and Oatp1b3 Deficiency Causes Human Rotor Syndrome by Interrupting Conjugated Bilirubin Reuptake into the Liver." Journal of Clinical Investigation 122, no. 2 (2012): 519-528.

Příloha 5:

Noskova, L., V. Stranecky, **H. Hartmannova**, A. Pristoupilova, V. Baresova, R. Ivanek, H. Hulkova, H. Jahnova, J. van der Zee, J. F. Staropoli, K. B. Sims, J. Tyynela, C. Van Broeckhoven, P. C. G. Nijssen, S. E. Mole, M. Elleder and S. Kmoch. "Mutations in Dnajc5, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis." American Journal of Human Genetics 89, no. 2 (2011): 241-252.

Příloha 6:

Stranecky, V., A. Hoischen, **H. Hartmannova**, M. S. Zaki, A. Chaudhary, E. Zudaire, L. Noskova, V. Baresova, A. Pristoupilova, K. Hodanova, J. Sovova, H. Hulkova, L. Piherova, J. Y. Hehir-Kwa, D. de Silva, M. P. Senanayake, S. Farrag, J. Zeman, P. Martasek, A. Baxova, H. H. Afifi, B. St Croix, H. G. Brunner, S. Temtamy and S. Kmoch. "Mutations in Antxr1 Cause Gapo Syndrome." American Journal of Human Genetics 92, no. 5 (2013): 792-799.

Příloha 7:

Hartmannova, H., M. Kubanek, M. Sramko, L. Piherova, L. Noskova, K. Hodanova, V. Stranecky, A. Pristoupilova, J. Sovova, T. Marek, J. Maluskova, P. Ridzon, J. Kautzner, H. Hulkova and S. Kmoch. "Isolated X-Linked Hypertrophic Cardiomyopathy Caused by a Novel Mutation of the Four-and-a-Half Lim Domain 1 Gene." Circ Cardiovasc Genet, (2013).

Seznam zkratek

ADSL	adenylosukcinát lyáza
AICAR	aminoimidazol karboxamid ribotid
AMP	adenosinmonofosfát
ANCL	adultní forma neuronální ceroidní lipofuscinózy
ATP	adenosintrifosfát
bp	páry bazí
CCD	Charge-coupled device
cDNA	komplementární deoxyribonukleová kyselina
CGH	komparativní genomová hybridizace
cM	centimorgan
CNPs	copy number polymorphisms, polymorfismy v počtu kopií
CNVs	copy number variations, varianty v počtu kopií
DJS	Dubin Johnsonův syndrom
DNA	deoxyribonukleová kyselina
dNMP	deoxynukleosidmonofosfát
dNTP	deoxynukleosidtrifosfát
ddNTP	dideoxynukleosidtrifosfát
dsDNA	dvouvláknová deoxyribonukleová kyselina
EST	expressed sequence tags
GB	gigabite
gDNA	genomová deoxyribonukleová kyselina
Indel	inzerce/delece
LOD	logaritmus poměru pravděpodobnosti rekombinace
MB	megabite
MPSIIIC	mukopolysacharidóza typu IIIC
mtDNA	mitochondriální deoxyribonukleová kyselina
NCL	neuronální ceroidní lipofuscinózy
NGS	next-generation sequencing, sekvenování nové generace
NTP	nukleosidtrifosfát
PCR	polymerázová řetězová reakce
PP	pyrofosfát
RFLP	polymorfismy v délce restrikčních fragmentů
RNA	ribonukleotidová kyselina
RS	Rotorův syndrom
S-Ado	sukcinyladenosin
SAICAr	sukcinylaminoimidazolkarboxamidribosid
SAICAR	sukcinylaminoimidazolkarboxamidribotid
SAMP	sukcinyladenosinmonofosfát
SNPs	single nukleotide polymorfisms, jednonukleotidové záměny
STS	sequenced tag sites
STR	short tandem repeat, krátká tandemové repetice
ÚDMP	Ústav dědičných metabolických poruch
VFN	Všeobecná fakultní nemocnice
WB	western blot

Použitá literatura

Abecasis G. R., Cherny S. S., Cookson W. O*., et al.*, 2002, *Merlin--rapid analysis of dense genetic maps using sparse gene flow trees.* Nat Genet. **30**(1): p. 97-101.

Abecasis G. R., Altshuler D., Auton A*., et al.*, 2010, *A map of human genome variation from population-scale sequencing.* Nature. **467**(7319): p. 1061-1073.

Ansorge W. J., 2009, *Next-generation DNA sequencing techniques.* New Biotechnology. **25**(4): p. 195-203.

Ausseil J., Loredo-Osti J. C., Verner A*., et al.*, 2004, *Localisation of a gene for mucopolysaccharidosis IIIC to the pericentromeric region of chromosome 8.* J Med Genet. **41**(12): p. 941-945.

Bar-Meir S., Baron J., Seligson U*., et al.*, 1982, *99mTc-HIDA cholescintigraphy in Dubin-Johnson and Rotor syndromes.* Radiology. **142**(3): p. 743-746.

Bernstein B. E., Birney E., Dunham I*., et al.*, 2012, *An integrated encyclopedia of DNA elements in the human genome.* Nature. **489**(7414): p. 57-74.

Bertram L. and Tanzi R. E., 2012, *The genetics of Alzheimer's disease.* Prog Mol Biol Transl Sci. **107**: p. 79-100.

Botstein D., White R. L., Skolnick M*., et al.*, 1980, *Construction of a genetic linkage map in man using restriction fragment length polymorphisms.* Am J Hum Genet. **32**(3): p. 314-331.

Bradley K. A., Mogridge J., Mourez M*., et al.*, 2001, *Identification of the cellular receptor for anthrax toxin.* Nature. **414**(6860): p. 225-229.

Carson-Walter E. B., Watkins D. N., Nanda A*., et al.*, 2001, *Cell surface tumor endothelial markers are conserved in mice and humans.* Cancer Res. **61**(18): p. 6649-6655.

Collins F. S., 1995, *Positional cloning moves from perditional to traditional.* Nat Genet. **9**(4): p. 347-350.

Consortium T. I. H. G. M., 2001, *A physical map of the human genome.* Nature. **409**(6822): p. 934-941.

Cowling B. S., Cottle D. L., Wilding B. R*., et al.*, 2011, *Four and a half LIM protein 1 gene mutations cause four distinct human myopathies: a comprehensive review of the clinical, histological and pathological features.* Neuromuscul Disord. **21**(4): p. 237-251.

Croucher N. J., Fookes M. C., Perkins T. T*., et al.*, 2009, *A simple method for directional transcriptome sequencing using Illumina technology.* Nucleic Acids Res. **37**(22): p. e148.

Cullen M., Seaman S., Chaudhary A*., et al.*, 2009, *Host-derived tumor endothelial marker 8 promotes the growth of melanoma.* Cancer Res. **69**(15): p. 6021-6026.

Dear J. W., Lilitkarntakul P., and Webb D. J., 2006, *Are rare diseases still orphans or happily adopted? The challenges of developing and using orphan medicinal products.* British Journal of Clinical Pharmacology. **62**(3): p. 264-271.

Demirgunes E. F., Ersoy-Evans S., and Karaduman A., 2009, *GAPO syndrome with the novel features of pulmonary hypertension, ankyloglossia, and prognathism.* Am J Med Genet A. **149A**(4): p. 802-805.

Ehling R., Noskova L., Stranecky V.*, et al.*, 2013, *Cerebellar dysfunction in a family harboring the PSEN1 mutation co-segregating with a Cathepsin D variant p.A58V.* Journal of the Neurological Sciences. **326**(1-2): p. 75-82.

Fan J. B., Chen X., Halushka M. K.*, et al.*, 2000, *Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays.* Genome Res. **10**(6): p. 853-860.

Fodor S. P. A., Rava R. P., Huang X. C.*, et al.*, 1993, *Multiplexed biochemical assays with biological chips.* Nature. **364**(6437): p. 555-556.

Fretzayas A. M., Garoufi A. I., Moutsouris C. X.*, et al.*, 1994, *Cholescintigraphy in the diagnosis of Rotor syndrome.* J Nucl Med. **35**(6): p. 1048-1050.

Gemin A., Sweet S., Preston T. J.*, et al.*, 2005, *Regulation of the cell cycle in response to inhibition of mitochondrial generated energy.* Biochem Biophys Res Commun. **332**(4): p. 1122-1132.

Gilbert W. and Maxam A., 1977, *Chemical Methods for Sequencing DNA and Recognition of DNA-Protein Contacts.* Journal of Supramolecular Structure: p. 57-57.

Gitschier J., Wood W. I., Goralka T. M.*, et al.*, 1984, *Characterization of the human factor VIII gene.* Nature. **312**(5992): p. 326-330.

Goloni-Bertollo E. M., Ruiz M. T., Goloni C. B.*, et al.*, 2008, *GAPO syndrome: three new Brazilian cases, additional osseous manifestations, and review of the literature.* Am J Med Genet A. **146A**(12): p. 1523-1529.

Gregory S. G., Barlow K. F., McLay K. E.*, et al.*, 2006, *The DNA sequence and biological annotation of human chromosome 1.* Nature. **441**(7091): p. 315-321.

Gu J., Faundez V., and Werner E., 2010, *Endosomal recycling regulates Anthrax Toxin Receptor 1/Tumor Endothelial Marker 8-dependent cell spreading.* Exp Cell Res. **316**(12): p. 1946-1957.

Gueneau L., Bertrand A. T., Jais J. P.*, et al.*, 2009, *Mutations of the FHL1 gene cause Emery-Dreifuss muscular dystrophy.* Am J Hum Genet. **85**(3): p. 338-353.

Gusella J. F., Wexler N. S., Conneally P. M.*, et al.*, 1983, *A polymorphic DNA marker genetically linked to Huntington's disease.* Nature. **306**(5940): p. 234-238.

Hagenbuch B. and Meier P. J., 2004, *Organic anion transporting polypeptides of the OATP/ SLC21 family: phylogenetic classification as OATP/ SLCO superfamily, new nomenclature and molecular/functional properties.* Pflugers Arch. **447**(5): p. 653-665.

Hagenbuch B. and Gui C., 2008, *Xenobiotic transporters of the human organic anion transporting polypeptides (OATP) family.* Xenobiotica. **38**(7-8): p. 778-801.

Heemstra H. E., van Weely S., Buller H. A.*, et al.*, 2009, *Translation of rare disease research into orphan drug development: disease matters.* Drug Discovery Today. **14**(23-24): p. 1166-1173.

Hotchkiss K. A., Basile C. M., Spring S. C.*, et al.*, 2005, *TEM8 expression stimulates endothelial cell adhesion and migration by regulating cell-matrix interactions on collagen.* Exp Cell Res. **305**(1): p. 133-144.

Hrebicek M., Mrazova L., Seyrantepe V.*, et al.*, 2006, *Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome).* American Journal of Human Genetics. **79**(5): p. 807-819.

Hrebicek M., Jirasek T., Hartmannova H.*, et al.*, 2007, *Rotor-type hyperbilirubinaemia has no defect in the canalicular bilirubin export pump.* Liver International. **27**(4): p. 485-491.

Hubbard T., Barker D., Birney E.*, et al.*, 2002, *The Ensembl genome database project.* Nucleic Acids Res. **30**(1): p. 38-41.

Jaeken J. and Vandenberghe G., 1984, *AN INFANTILE AUTISTIC SYNDROME CHARACTERIZED BY THE PRESENCE OF SUCCINYL PURINES IN BODY-FLUIDS.* Lancet. **2**(8411): p. 1058-1061.

Jaeken J., Wadman S. K., Duran M.*, et al.*, 1988, *Adenylosuccinase deficiency: an inborn error of purine nucleotide synthesis.* Eur J Pediatr. **148**(2): p. 126-131.

Jaeken J., Van den Bergh F., Vincent M. F.*, et al.*, 1992, *Adenylosuccinase deficiency: a newly recognized variant.* J Inherit Metab Dis. **15**(3): p. 416-418.

Jalanko A. and Braulke T., 2009, *Neuronal ceroid lipofuscinoses.* Biochim Biophys Acta. **1793**(4): p. 697-709.

Jarvie T., 2005, *Next generation sequencing technologies.* Drug Discovery Today: Technologies. **2**(3): p. 255-260.

Kent W. J., Sugnet C. W., Furey T. S.*, et al.*, 2002, *The human genome browser at UCSC.* Genome Res. **12**(6): p. 996-1006.

Knoblauch H., Geier C., Adams S.*, et al.*, 2010, *Contractures and hypertrophic cardiomyopathy in a novel FHL1 mutation.* Ann Neurol. **67**(1): p. 136-140.

Kocabay G. and Mert M., 2009, *GAPO syndrome associated with dilated cardiomyopathy: an unreported association.* Am J Med Genet A. **149A**(3): p. 415-416.

Kong X., Murphy K., Raj T.*, et al.*, 2004, *A combined linkage-physical map of the human genome.* Am J Hum Genet. **75**(6): p. 1143-1148.

Krijt J., Kmoch S., Hartmannova H.*, et al.*, 1999, *Identification and determination of succinyladenosine in human cerebrospinal fluid.* Journal of Chromatography B. **726**(1-2): p. 53-58.

Kwok S. C., Ledley F. D., DiLella A. G.*, et al.*, 1985, *Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalanine hydroxylase.* Biochemistry. **24**(3): p. 556-561.

Link E., Parish S., Armitage J.*, et al.*, 2008, *SLCO1B1 variants and statin-induced myopathy--a genomewide study.* N Engl J Med. **359**(8): p. 789-799.

Liu L., Li Y., Li S.*, et al.*, 2012, *Comparison of Next-Generation Sequencing Systems.* Journal of Biomedicine and Biotechnology. **2012**: p. 11.

Liu S., Crown D., Miller-Randolph S*., et al.*, 2009, *Capillary morphogenesis protein-2 is the major receptor mediating lethality of anthrax toxin in vivo.* Proc Natl Acad Sci U S A. **106**(30): p. 12424-12429.

Maaswinkel-Mooij P. D., Laan L. A., Onkenhout W*., et al.*, 1997, *Adenylosuccinase deficiency presenting with epilepsy in early infancy.* J Inherit Metab Dis. **20**(4): p. 606-607.

Mandal S., Guptan P., Owusu-Ansah E*., et al.*, 2005, *Mitochondrial regulation of cell cycle progression during development as revealed by the tenured mutation in Drosophila.* Dev Cell. **9**(6): p. 843-854.

Mardis E. R., 2008, *Next-generation DNA sequencing methods*, in *Annual Review of Genomics and Human Genetics*, Annual Reviews: Palo Alto. p. 387-402.

Mardis E. R., 2008, *The impact of next-generation sequencing technology on genetics.* Trends in Genetics. **24**(3): p. 133-141.

Margulies M., Egholm M., Altman W. E*., et al.*, 2005, *Genome sequencing in microfabricated high-density picolitre reactors.* Nature. **437**(7057): p. 376-380.

Matise T. C., Sachidanandam R., Clark A. G*., et al.*, 2003, *A 3.9-centimorgan-resolution human single-nucleotide polymorphism linkage map and screening set.* Am J Hum Genet. **73**(2): p. 271-284.

Maxam A. M. and Gilbert W., 1977, *A new method for sequencing DNA.* Proceedings of the National Academy of Sciences. **74**(2): p. 560-564.

Mouchegh K., Zikanova M., Hoffmann G. F*., et al.*, 2007, *Lethal fetal and early neonatal presentation of adenylosuccinate lyase deficiency: observation of 6 patients in 4 families.* J Pediatr. **150**(1): p. 57-61 e52.

Nachman M. W. and Crowell S. L., 2000, *Estimate of the mutation rate per nucleotide in humans.* Genetics. **156**(1): p. 297-304.

Nanda A., Al-Ateeqi W. A., Al-Khawari M. A*., et al.*, 2010, *GAPO syndrome: a report of two siblings and a review of literature.* Pediatr Dermatol. **27**(2): p. 156-161.

Noskova L., Stranecky V., Hartmannova H*., et al.*, 2011, *Mutations in DNAJC5, Encoding Cysteine-String Protein Alpha, Cause Autosomal-Dominant Adult-Onset Neuronal Ceroid Lipofuscinosis.* American Journal of Human Genetics. **89**(2): p. 241-252.

Nyren P. and Lundin A., 1985, *ENZYMATIC METHOD FOR CONTINUOUS MONITORING OF INORGANIC PYROPHOSPHATE SYNTHESIS.* Analytical Biochemistry. **151**(2): p. 504-509.

Pease A. C., Solas D., Sullivan E. J*., et al.*, 1994, *Light-generated oligonucleotide arrays for rapid DNA sequence analysis.* Proc Natl Acad Sci U S A. **91**(11): p. 5022-5026.

Petes T. D. and Botstein D., 1977, *Simple Mendelian inheritance of the reiterated ribosomal DNA of yeast.* Proc Natl Acad Sci U S A. **74**(11): p. 5091-5095.

Pospíšilová Š., Tichý B., and Mayer J., 2009, *Sekvenování lidského genomu – technologie nové generace aneb budeme rutinně sekvenovat lidské genomy?* ČASOPIS LÉKAŘŮ ČESKÝCH. **148**(7): p. 296-302.

Remuzzi G. and Garattini S., 2008, *Rare diseases: what's next?* Lancet. **371**(9629): p. 1978-1979.

Riordan J. R., Rommens J. M., Kerem B*., et al.*, 1989, *Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA.* Science. **245**(4922): p. 1066-1073.

Robson K. J., Beattie W., James R. J*., et al.*, 1984, *Sequence comparison of rat liver phenylalanine hydroxylase and its cDNA clones.* Biochemistry. **23**(24): p. 5671-5675.

Rohn J., 2013, *Billions spent on rare diseases.* Nat Biotech. **31**(5): p. 368-368.

Royer-Pokora B., Kunkel L. M., Monaco A. P*., et al.*, 1986, *Cloning the gene for an inherited human disorder--chronic granulomatous disease--on the basis of its chromosomal location.* Nature. **322**(6074): p. 32-38.

Salerno C., Crifo C., and Giardini O., 1995, *ADENYLOSUCCINASE DEFICIENCY - A PATIENT WITH IMPAIRED ERYTHROCYTE ACTIVITY AND ANOMALOUS RESPONSE TO INTRAVENOUS FRUCTOSE.* Journal of Inherited Metabolic Disease. **18**(5): p. 602-608.

Sanger F., Donelson J. E., Coulson A. R*., et al.*, 1973, *Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage fl DNA.* Proc Natl Acad Sci U S A. **70**(4): p. 1209-1213.

Sanger F. and Coulson A. R., 1975, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.* Journal of Molecular Biology. **94**(3): p. 441-448.

Sanger F., Air G. M., Barrell B. G*., et al.*, 1977, *Nucleotide sequence of bacteriophage phi X174 DNA.* Nature. **265**(5596): p. 687-695.

Sanger F., Nicklen S., and Coulson A. R., 1977, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A. **74**(12): p. 5463-5467.

Sayli B. S. and Gul D., 1996, *GAPO syndrome in Turkiye.* Am J Med Genet. **65**(3): p. 252-253.

Shathasivam T., Kislinger T., and Gramolini A. O., 2010, *Genes, proteins and complexes: the multifaceted nature of FHL family proteins in diverse tissues.* J Cell Mol Med. **14**(12): p. 2702-2720.

Shelton D. N., Chang E., Whittier P. S*., et al.*, 1999, *Microarray analysis of replicative senescence.* Curr Biol. **9**(17): p. 939-945.

Schena M., Shalon D., Davis R. W*., et al.*, 1995, *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science. **270**(5235): p. 467-470.

Schieppati A., Henter J. I., Daina E*., et al.*, 2008, *Why rare diseases are an important medical and social issue.* Lancet. **371**(9629): p. 2039-2041.

Schuur M., Ikram M. A., van Swieten J. C*., et al.*, 2011, *Cathepsin D gene and the risk of Alzheimer's disease: a population-based study and meta-analysis.* Neurobiol Aging. **32**(9): p. 1607-1614.

Sip M., Bystricka D., Kmoch S*., et al.*, 2010, *Detection of viral infections by an oligonucleotide microarray.* Journal of Virological Methods. **165**(1): p. 64-70.

Smigielski E. M., Sirotkin K., Ward M.*, et al.*, 2000, *dbSNP: a database of single nucleotide polymorphisms.* Nucleic Acids Research. **28**(1): p. 352-355.

St Croix B., Rago C., Velculescu V.*, et al.*, 2000, *Genes expressed in human tumor endothelium.* Science. **289**(5482): p. 1197-1202.

Stockl P., Hutter E., Zwerschke W.*, et al.*, 2006, *Sustained inhibition of oxidative phosphorylation impairs cell proliferation and induces premature senescence in human fibroblasts.* Exp Gerontol. **41**(7): p. 674-682.

Stone T. W., Roberts L. A., Morris B. J.*, et al.*, 1998, *Succinylpurines induce neuronal damage in the rat brain.* Adv Exp Med Biol. **431**: p. 185-189.

Strathdee C. A., Gavish H., Shannon W. R.*, et al.*, 1992, *Cloning of cDNAs for Fanconi's anaemia by functional complementation.* Nature. **356**(6372): p. 763-767.

Tipton R. E. and Gorlin R. J., 1984, *Growth retardation, alopecia, pseudo-anodontia, and optic atrophy--the GAPO syndrome: report of a patient and review of the literature.* Am J Med Genet. **19**(2): p. 209-216.

Tucker T., Marra M., and Friedman J. M., 2009, *Massively Parallel Sequencing: The Next Big Thing in Genetic Medicine.* American Journal of Human Genetics. **85**(2): p. 142-154.

Valik D., Miner P. T., and Jones J. D., 1997, *First U.S. case of adenylosuccinate lyase deficiency with severe hypotonia.* Pediatr Neurol. **16**(3): p. 252-255.

Van den Bergh F., Vincent M. F., Jaeken J.*, et al.*, 1993, *Residual adenylosuccinase activities in fibroblasts of adenylosuccinase-deficient children: parallel deficiency with adenylosuccinate and succinyl-AICAR in profoundly retarded patients and non-parallel deficiency in a mildly retarded girl.* J Inherit Metab Dis. **16**(2): p. 415-424.

Venter J. C., Adams M. D., Myers E. W.*, et al.*, 2001, *The sequence of the human genome.* Science. **291**(5507): p. 1304-1351.

Verma K., Gu J., and Werner E., 2011, *Tumor endothelial marker 8 amplifies canonical Wnt signaling in blood vessels.* PLoS One. **6**(8): p. e22334.

Voznyi Ya V., Karpova E. A., Dudukina T. V.*, et al.*, 1993, *A fluorimetric enzyme assay for the diagnosis of Sanfilippo disease C (MPS III C).* J Inherit Metab Dis. **16**(2): p. 465-472.

Werner E., Kowalczyk A. P., and Faundez V., 2006, *Anthrax toxin receptor 1/tumor endothelium marker 8 mediates cell spreading by coupling extracellular ligands to the actin cytoskeleton.* J Biol Chem. **281**(32): p. 23227-23236.

Wheeler D. L., Barrett T., Benson D. A.*, et al.*, 2006, *Database resources of the National Center for Biotechnology Information.* Nucleic Acids Res. **34**(Database issue): p. D173-180.

Windpassinger C., Schoser B., Straub V.*, et al.*, 2008, *An X-linked myopathy with postural muscle atrophy and generalized hypertrophy, termed XMPMA, is caused by mutations in FHL1.* Am J Hum Genet. **82**(1): p. 88-99.

Wu R. and Kaiser A. D., 1968, *Structure and base sequence in the cohesive ends of bacteriophage lambda DNA.* Journal of Molecular Biology. **35**(3): p. 523-537.

Yang M. Y., Chaudhary A., Seaman S*., et al.*, 2011, *The cell surface structure of tumor endothelial marker 8 (TEM8) is regulated by the actin cytoskeleton.* Biochim Biophys Acta. **1813**(1): p. 39-49.

Zeman J., Krijt J., Stratilova L*., et al.*, 2000, *Abnormalities in succinylpurines in fumarase deficiency: Possible role in pathogenesis of CNS impairment.* Journal of Inherited Metabolic Disease. **23**(4): p. 371-374.

Zikanova M., Krijt J., Hartmannova H*., et al.*, 2005, *Preparation of 5-amino-4-imidazole-N-succinocarboxamide ribotide, 5-amino-4-imidazole-N-succinocarboxamide riboside and succinyladenosine, compounds usable in diagnosis and research of adenylosuccinate lyase deficiency.* Journal of Inherited Metabolic Disease. **28**(4): p. 493-499.

Zikanova M., Skopova V., Hnizda A*., et al.*, 2010, *Biochemical and structural analysis of 14 mutant adsl enzyme complexes and correlation to phenotypic heterogeneity of adenylosuccinate lyase deficiency.* Hum Mutat. **31**(4): p. 445-455.

Human Molecular Genetics, 2000, Vol. 9, No. 10 **1501–1513**

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

Stanislav Kmoch[1,+], Hana Hartmannová[1], Blanka Stibůrková[1], Jakub Krijt[1], Marie Zikánová[1] and Ivan Šebesta[1,2]

[1]Institute for Inherited Metabolic Disorders, Ke Karlovu 2, 120 00 Prague 2, and [2]Department of Clinical Biochemistry, Charles University 1st School of Medicine and General Faculty Hospital, 120 00 Prague, Czech Republic

Received 11 February 2000; Revised and Accepted 4 April 2000 DDBJ/EMBL/GenBank accession nos. AF067853, AF067854, AF106656

Adenylosuccinate lyase (ADSL) is a bifunctional enzyme acting in *de novo* purine synthesis and purine nucleotide recycling. ADSL deficiency is a selectively neuronopathic disorder with psychomotor retardation and epilepsy as leading traits. Both dephosphorylated enzyme substrates, succinylaminoimidazole-carboxamide riboside (SAICAr) and succinyladenosine (S-Ado), accumulate in the cerebrospinal fluid (CSF) of affected individuals with S-Ado/SAICAr concentration ratios proportional to the phenotype severity. We studied the disorder at various levels in a group of six patients with ADSL deficiency. We identified the complete *ADSL* cDNA and its alternatively spliced isoform resulting from exon 12 skipping. Both mRNA isoforms were expressed in all the tissues studied with the non-spliced form 10-fold more abundant. Both cDNAs were expressed in *Escherichia coli* and functionally characterized at the protein level. The results showed only the unspliced ADSL to be active. The gene consists of 13 exons spanning 23 kb. The promotor region shows typical features of the housekeeping gene. Eight mutations were identified in a group of six patients. The expression studies of the mutant proteins carried out in an attempt to study genotype–phenotype correlation showed that the level of residual enzyme activity correlates with the severity of the clinical phenotype. All the mutant enzymes studied *in vitro* displayed a proportional decrease in activity against both of their substrates. However, this was not concordant with strikingly different concentration ratios in the CSF of individual patients. This suggests either different *in vivo* enzyme activities against each of the substrates and/or their different turnover across the CSF–blood barrier, which may be decisive in determining disease severity.

INTRODUCTION

Adenylosuccinate lyase [adenylosuccinase (ADSL); EC 4.3.2.2.] is an enzyme acting in two pathways of purine nucleotide metabolism. It catalyses the conversion of succinylaminoimidazole carboxamide ribotide (SAICAR) into aminoimidazole carboxamide ribotide (AICAR) in the purine *de novo* synthesis pathway, and the formation of adenosine monophosphate (AMP) from adenylosuccinate (S-AMP) in the purine nucleotide cycle. ADSL deficiency (McKusick 103050) was first recognized by Jaeken and van den Berghe (1) in 1984 in patients with hereditary psychomotor retardation and autism and remains so far the only inherited deficiency of *de novo* purine synthesis recognized in humans. To date, about 50 patients have been diagnosed world-wide and reports on about half of them have been published (1–8). The disease usually appears within the first months of life with neurological involvement. Affected individuals later show various degrees of psychomotor retardation often accompanied by various combinations of epilepsy, hypotonia, growth retardation and muscular wasting. A whole spectrum of behavioural changes such as autism, aggressiveness and self-mutilation have been also described (1,3). The non-specific nature of the neurological symptomatology results in the diagnosis relying on biochemical detection of dephosphorylated substrates of ADSL: SAICA riboside (SAICAr) and succinyladenosine (S-Ado) in body fluids. For their detection, several simple screening procedures have been developed (9–12). To better understand the ADSL function, enzymic assays and metabolic functional studies were performed in different tissues of ADSL-deficient patients with the following results. First, ADSL activity has been repeatedly proven to be decreased with both ADSL substrates, SAICAR and S-AMP, which confirms the bifunctional role of the enzyme (13,14). Secondly, residual enzyme activities varied between different tissues and the enzyme deficiency was not always generalized, suggesting the existence of either hitherto unrecognized tissue-specific ADSL isoforms or tissue-specific gene expression regulation (15). Thirdly, functional studies in fibroblasts showed neither a significant reduction in the rate of *de novo* purine synthesis nor any abnormality in purine nucleotide concentrations (16).

[+]To whom correspondence should be addressed. Tel: +420 2 24920294; Fax: +420 2 24919392; Email: skmoch@lf1.cuni.cz

Neither of these biochemical observations could be correlated with the severity of the clinical phenotype nor provide any explanation for the pathogenetic mechanism of selective neuronopathic features of the disease. The only biochemical finding that seems to correlate with the severity of the symptoms is the proportion of the accumulating S-Ado and SAICAr in body fluids, particularly in the cerebrospinal fluid (8,14).

Molecular studies of the *ADSL* gene were started by the work of Van Keuren (17) who assigned the *ADSL* gene to chromosome 22 using a somatic cell hybrid panel. The human liver *ADSL* cDNA (GenBank accession no. S60710) was cloned in 1992 (18) and the gene was mapped to chromosome 22q13.1––q13.2 (19). The nucleotide sequence showed 94% and 85% identity to the mouse (20) and avian *ADSL* cDNAs (21), respectively, and encoded a protein 459 amino acids long. The enzyme proved to have a homotetrameric structure with a relative molecular mass of ~52 kDa per subunit (18,22,23). However, several indications suggested discrepancy in the cDNA sequence. The C→A exchange in the third nucleotide of the cDNA sequence, accession no. X65867, deposited in the GenBank database, creates an alternative initiation codon located in-frame, 75 nucleotides upstream from the one reported by Stone (18). If translated from this first initiation codon the protein would have an additional 25 amino acids at the N-terminus. The question of the correct position of the initiation codon was also stressed by expression of a recombinant protein transcribed from the latter initiation codon. This resulted, even under standard experimental conditions, in an insoluble and inactive enzyme (22). Despite this discrepancy, cDNA mutation analysis was performed successfully in a dozen ADSL-deficient patients (4,18,24,25). These studies showed broad allelic heterogeneity with 14 mutations identified. The allelic heterogeneity suggests correlation with phenotypic variability; however, no convincing data providing a particular link between them have been presented so far.

Therefore, the present study had the following objectives of (i) identification, cloning, expression and characterization of the correct, full-length *ADSL* cDNA sequence and of its possible isoform(s), (ii) elucidation and description of the *ADSL* gene structure and its complete genomic sequence, (iii) identification of *ADSL* mutations in a series of six patients and (iv) expression and kinetic characterization of the mutant enzyme proteins, including the S-Ado/SAICAr ratio variance in an attempt to elucidate the genotype–phenotype correlations.

RESULTS

Identification of a complete *ADSL* cDNA sequence

To identify the complete *ADSL* cDNA sequence, muscle and lymphocyte full-length cDNAs were prepared, using the CapFinder technology (Clontech, Palo Alto, CA). The full-length cDNAs were ligated to GenomeWalker Adaptors (Clontech) and overlapping 5′ and 3′ *ADSL* cDNA ends were amplified using the *ADSL* cDNA and GenomeWalker Adaptor specific primers. Sequence analysis of the resulting fragments showed a novel 52 bp sequence at the 5′-end of the *ADSL* cDNA. (Fig. 1, GenBank accession no. AF067853). The sequence contained both initiation codons identical to those described previously (GenBank accession no. X65867).

Alternative splicing of *ADSL* mRNA

Two PCR products of different lengths were consistently generated by the RT–PCR amplification of the full-length *ADSL* cDNA. To reveal their identities the PCR products were cloned and sequenced. Sequence analysis showed the absence of 177 bp corresponding to exon 12, in the shorter fragment. This mRNA isoform (GenBank accession no. AF067854) can be translated in the same reading frame as the non-spliced *ADSL* mRNA but the synthesized protein would lack 59 amino acids (amino acids 397–456). The tissue-specific patterns of expression of both *ADSL* isoforms were studied using PCR analysis of the Clontech Human Multiple Tissue cDNA (MTC) Panel I. The results showed both isoforms to be expressed in all tissues studied in approximately the same relative amount with the unspliced form being ~10-fold more abundant (Figure 2).

Expression of *ADSL* isoforms in *Escherichia coli*

To characterize functionally the identified sequences, both cDNAs were cloned and expressed. The resulting fusion proteins were purified, cleaved with factor Xa, analysed by SDS–PAGE and the activities of both forms were measured using both ADSL substrates, S-AMP and SAICAR. The results showed that the full-length cDNA encodes an active protein. The product of the alternatively spliced isoform, although stable, was entirely inactive (Fig. 3).

Structure and complete genomic sequence of the human *ADSL* gene

To define the *ADSL* genomic structure, the known genomic organization of the murine *ADSL* gene (20) was used to predict conserved splicing sites within the human *ADSL* cDNA. Primers located within the predicted *ADSL* exons were designed and used in both a genome walking strategy (Universal GenomeWalker, Clontech) and for direct amplification of *ADSL* fragments from genomic DNA. The PCR fragments generated by genome walking were gel purified, cloned and sequenced. The fragments generated by genomic PCR were either cloned and sequenced as described above, or filter purified and sequenced using the *ADSL* cDNA-specific primers.

With this approach, genomic fragments covering the *ADSL* promoter and all of the exons including flanking intronic sequences were isolated and sequenced. The human *ADSL* gene was found to consist of 13 exons. Exon lengths, their positions within the cDNA and the sequences of all of the exon–intron boundaries are shown in Table 1. All the intron–exon splice junctions conform to the GT/AG consensus sequence. The genomic organization of the human *ADSL* gene showed identity to that of the murine gene.

To complete the whole gene sequence, the identified sequences were analysed using the BLASTN homology search in the Chromosome 22-specific Blast server database (www.sanger.ac.uk/HGP/Chrom_blast_server.shtml). Several sequences showing significant homologies were identified and allowed the reconstruction of the whole human *ADSL* genomic sequence which spans over 23 kb (GenBank accession no. AF106656).

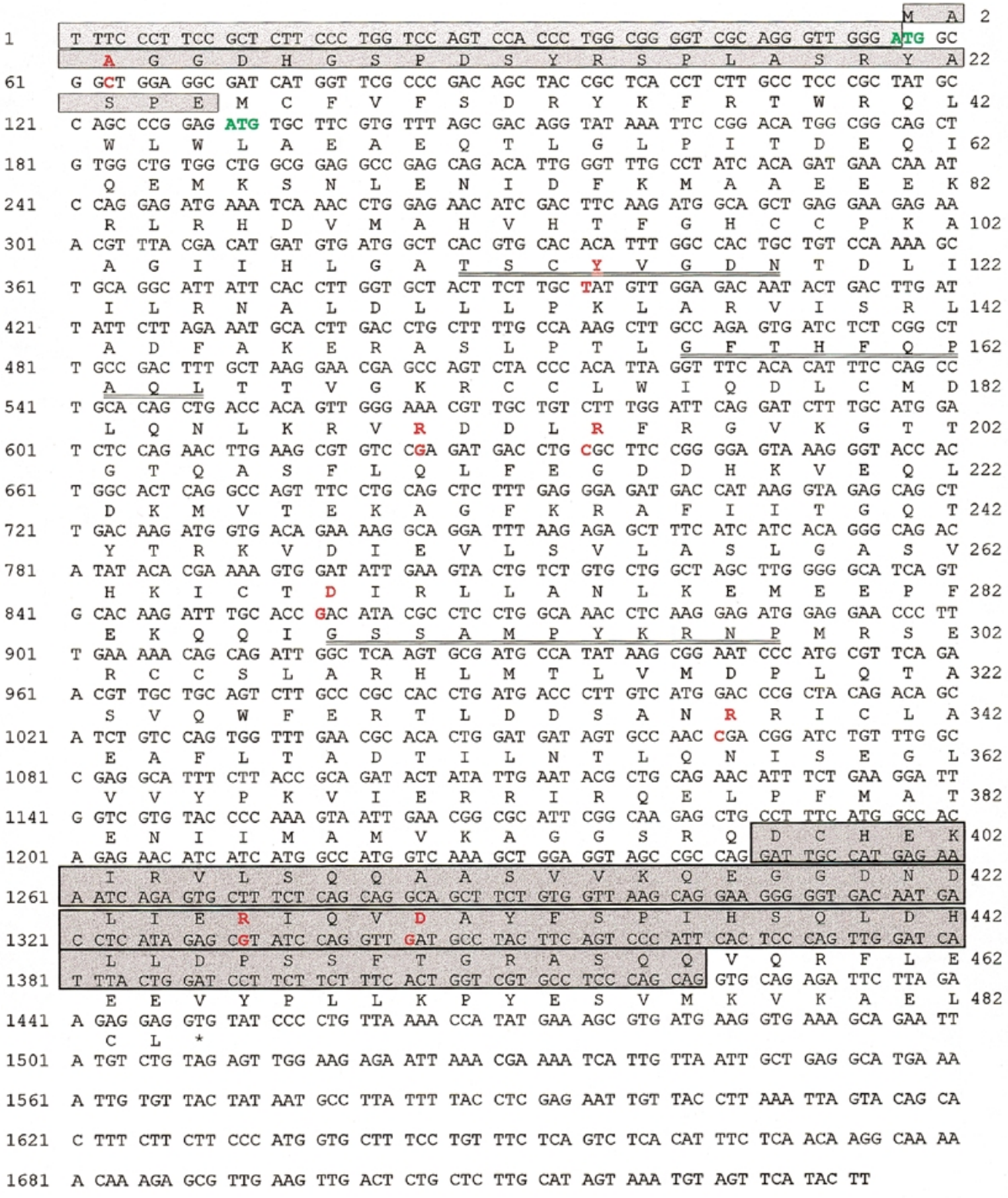


Figure 1. Nucleotide sequence and predicted protein product of the full-length human *ADSL* cDNA. New 5′ cDNA sequence, the resulting N-terminal protein sequence and the alternatively spliced exon12 are shown in shaded boxes. Alternative initiation codons M1, M2 are shown in green. Positions of individual mutations found in this study are shown in red. The three highly conserved regions within the fumarase superfamily are double underlined.

The lengths of the introns range from 146 to 4536 bp (Table 1, Fig. 4). Using the databases searches, we identified several typical promoter elements such as Oct 1, AP1, HSF2, Pbx-1a, E2F and Sp1 sites up to 500 bp upstream of the 5′-end

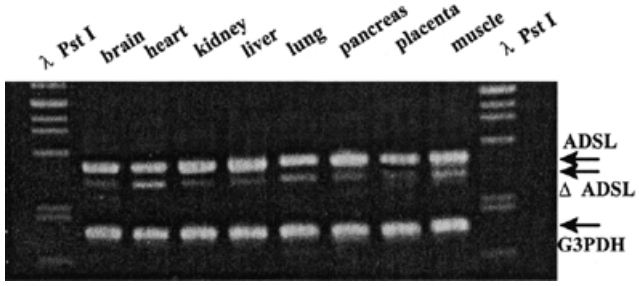


Figure 2. RT–PCR analysis of *ADSL* splicing variants in normalized human tissue-specific cDNA pools. *G3PDH* is a control mRNA.

of the cDNA sequence reported here. Canonical TATA and CAAT boxes were apparently absent (Fig. 4).

Mutation analysis of patients with ADSL deficiency

Full-length cDNA from both patient and control lymphocytes or fibroblasts were prepared and sequenced, with the following mutations identified: 63C→T (A3V), 395T→C (Y114H), 624G→A (R190Q), 635C→T (R194C), 857G→A (D268N), 1332G→A (R426H) and 1343G→A (D430N). The individual genotypes together with clinical and biochemical findings are listed in Table 2. All the mutations identified by cDNA sequencing, except for 63C→T, were confirmed by individually designed PCR–RFLP tests carried out on both cDNA and gDNA (Table 3). The 63C→T mutation, found in patient 2, appeared to be homozygous according to cDNA analyses (sequencing and PCR–RFLP). However, the PCR–RFLP analysis of the gDNA ruled out the homozygosity and showed the patient to be heterozygous for it. To isolate and analyse the second allele, cDNA from the patient was cloned and an allele-specific PCR assay was used to discriminate between 63C and 63T alleles. In total, 280 clones were analysed and only mutated 63T alleles were found. To identify the missing mutation in patient 2, a panel of intronic primer pairs was designed in order to amplify and sequence the *ADSL* promoter and all the coding regions (Table 4). Using this approach a heterozygous C→T transition in exon 9 (1064C→T on cDNA) was found. This mutation creates a premature stop codon in the corresponding mRNA (337R→X) which leads to its nonsense-mediated decay.

Characterization of identified mutations

To test the functional consequences of the individual mutations, the corresponding cDNAs were cloned and the mutated recombinant proteins expressed, isolated and characterized. Both the fusion proteins and proteins without the affinity tag were prepared and analysed on SDS–PAGE. Except for the 114H, all the proteins were stable (Fig. 3A and B). The catalytic activities of both protein forms were assayed using both ADSL substrates. Activities of the mutated enzyme proteins ranged from 0 to 140% of the wtADSL activity (Fig. 3C and D). The mutants had activities similar to or even higher than the wild type and compared with wild-type enzyme displayed modest changes in thermal stability and temperature dependence of ADSL activity. Significant effect was observed for the 430N mutant at 45°C and 55°C (*t*-test) (Fig. 5A and B). All of the mutants also showed changes in some of the kinetic parameters (Table 5). Mutant 190Q showed difference in K_M for S-AMP and V_{lim} for both substrates and mutants 430N and 3V showed lowered K_M for SAICAR (*z*-test).

Genotype–phenotype analysis

Clinical, biochemical and molecular data of individual patients are shown in Table 2. The patients were ordered according to the severity of their symptoms. The S-Ado/SAICAr concentration ratios in cerebrospinal fluid (CSF) or urine correlate with the severity of symptoms. To explain the variations in S-Ado/SAICAr concentration ratios, the ADSL activities were measured in patients' skin fibroblasts. The residual activities ranged from 20 to 60% of controls and were proportionally decreased with both substrates. No correlation between the residual activities and severity of symptoms was observed. The S-AMP/SAICAR activity ratios in individual patients were 0.8–1.2 against 1.44 ± 0.16 in controls and did not reflect the S-Ado/SAICAr concentration ratios found in CSF or urine. The same applied to the activities of the expressed mutated recombinant proteins. The residual activities were again decreased proportionately, with the S-AMP/SAICAR activity ratios ranging from 0.6 to 1.0 (wtADSL 1.3 ± 0.2) (Fig. 6B).

A trend suggesting a relationship between the phenotype severity and the predicted residual activities of individual genotypes was calculated as a mean of the homoallelic *in vitro* activities and is shown in Fig. 6A.

DISCUSSION

The main goal of the study was to characterize basic molecular aspects of ADSL pathobiology with special emphasis on the clinical and biochemical heterogeneity observed in this disease.

Identification of the full-length cDNA

To isolate and characterize the full-length *ADSL* cDNA we set up an original approach combining the advantages of the previously described methods for a full-length cDNA preparation (CapFinder technology, Clontech, Palo Alto, CA) with amplification of unknown flanking DNA sequence (vectorette PCR). Using this approach we identified a novel 5′-end cDNA sequence absent among more than 70 *ADSL* expressed sequence tags (ESTs) deposited currently in the human dbEST database, where the longest deposited EST sequence starts at nucleotide 35 of the novel sequence. This result illustrates the utility of our method for full-length cDNA preparation and its superiority over protocols currently used for full-length cDNA synthesis.

The enzyme protein

The 5′-end cDNA sequence identified contained the alternative initiation codon (M1) previously identified by Fon *et al.* (GenBank accession no. X65867). To characterize the full-length ADSL translated from the M1, we cloned and expressed the corresponding cDNA sequence in *E.coli*. The expressed enzyme protein was soluble, active and stable in contrast to the truncated protein translated from the M2, which was insoluble and inactive under standard experimental conditions (22). The

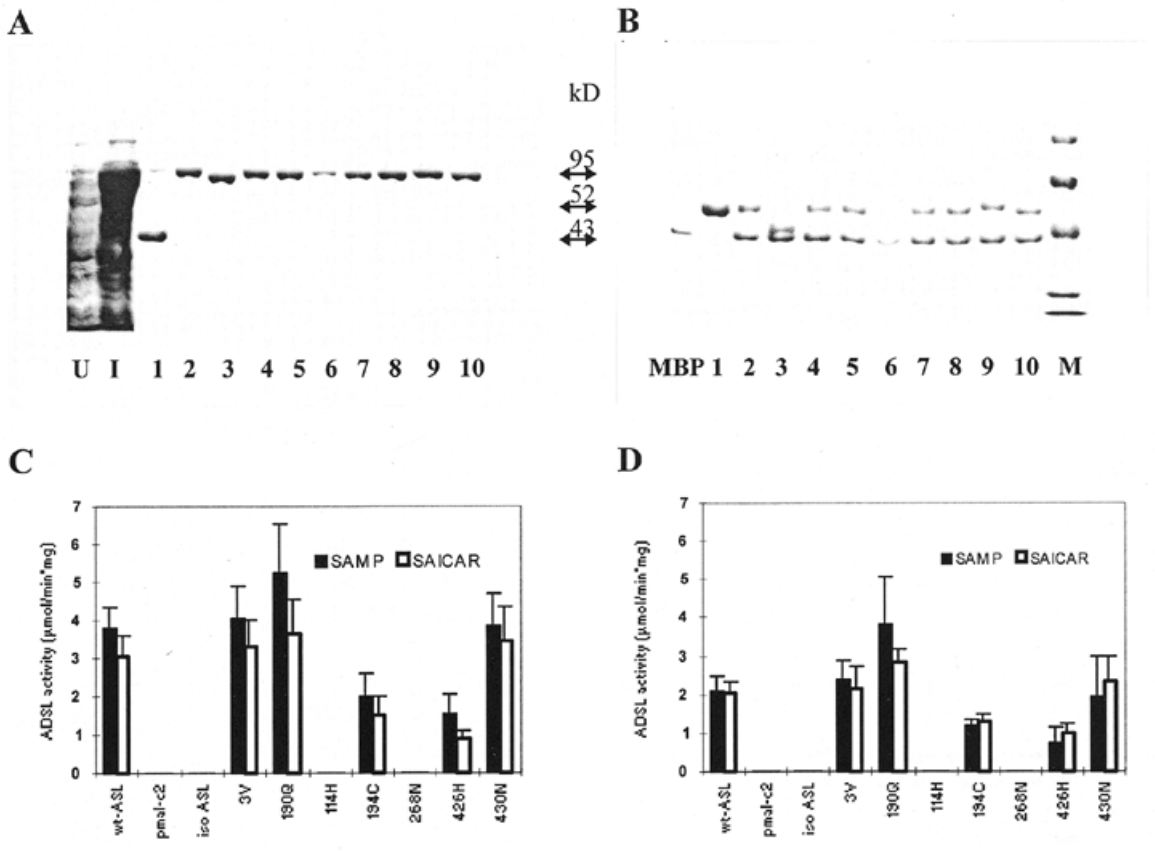


Figure 3. Expression, purification and activities of individual ADSL forms. (**A**) SDS–PAGE analysis of maltose affinity purified MBP–ADSL fusion proteins. U, non-induced cells; I, cells induced with 0.3 mM IPTG for 3 h; lane 1, MBP expressed from vector pmal-c2; lanes 2–10, individual MBP–ADSL fusion proteins (2-wtADSL, 3-ΔADSL, 4–3V, 5–190Q, 6–114H, 7–194C, 8–268N, 9–426H, 10–430N). (**B**) SDS–PAGE analysis of factor Xa-cleaved MBP–ADSL fusion proteins. MBP, maltose binding protein standard; lane 1, MBP expressed from pmal-c2 vector; lanes 2–10, individual ADSL proteins (2, wtADSL; 3, ΔADSL; 4, 3V; 5, 190Q; 6, 114H; 7, 194C; 8, 268N; 9, 426H; 10, 430N); M, SDS molecular weight marker (Pharmacia). (**C**) Activities of affinity-purified MBP–ADSL fusion proteins assayed using both ADSL substrates. (**D**) Activities of factor Xa-cleaved MBP–ADSL fusion proteins assayed using both ADSL substrates.

Table 1. Positions and nucleotide sequences of the *ADSL* exon–intron boundaries; individual exon, intron lengths

Exon	Position in cDNA[a]	Size (bp)	Acceptor splice site	Donor splice site	Intron	Size (bp)
1	1–208	208		GAGCAG/**gt**taaca	1	3115
2	209–412	204	ctgc**ag**/ACATTG	AATACT/**gt**aggc	2	3038
3	413–457	45	tcgt**ag**/GACTTG	CCAAAG/**gt**aagg	3	1130
4	458–537	80	ttgt**ag**/CTTGCC	TTTCCA/**gt**aagt	4	4536
5	538–709	172	tttc**ag**/GCCTGC	CATAAG/**gt**attc	5	224
6	710–756	47	tcca**ag**/GTAGAG	TAAGAG/**gt**aggt	6	1095
7	757–847	91	tcct**ag**/AGCTTT	CACAAG/**gt**gagt	7	780
8	848–917	70	cccc**ag**/ATTTGC	AGATTG/**gt**gagt	8	146
9	918–1065	148	gggc**ag**/GCTCAA	CAACCG/**gt**cagt	9	1345
10	1066–1156	91	ttct**ag**/ACGGAT	CCCAAA/**gt**aaga	10	1204
11	1157–1246	90	ctat**ag**/GTAATT	CGCCAG/**gt**ttgt	11	514
12	1247–1423	177	tccc**ag**/GATTGC	CAGCAG/**gt**aagc	12	1379
13	1424–1734	311	tggc**ag**/GTGCAG			

[a]Positions on cDNA are according to the AF067853 *ADSL* cDNA sequence.

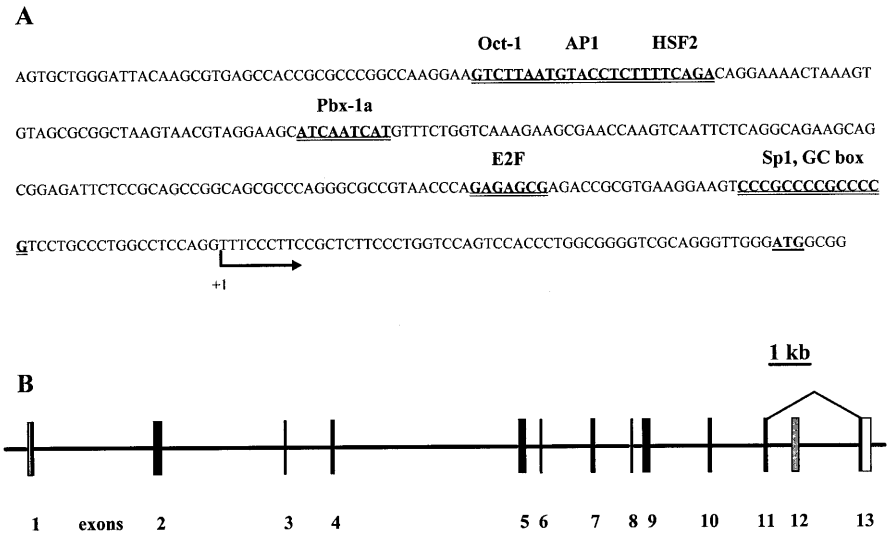


Figure 4. (**A**) Sequence of the 5′-flanking region of the human *ADSL* gene. Promotor element sequences are double underlined. The bent arrow indicates the position of the 5′-end of the cDNA sequence reported here. First initiation codon (M1) is in bold type. (**B**) Organization of the human *ADSL* gene. Vertical bars represent individual exons; 5′- and 3′-UTRs are in white (drawn to scale). Bent line indicates an alternative splicing of exon 12.

Table 2. Summary of clinical, biochemical and molecular data of individual patients

	sibs					
Patient	1	2	3	4	5	6
Gender	f	f	m	m	f	f
Onset	4th month	1st month	14th day	4th month	6th month	4th month
First sign	seizures	seizures	seizures	PMR	PMR	hypotonia
Other symptons	severe PMR	severe PMR, apnoea	PMR	hypotonia, hyperactivity	hypotonia, hyperactivity	no
Degree of involvement	+++++	+++++	++++	++	++	+
Present status and age	No psychomotor development, death at 15 months	9 years, coma vigil	5 years, hypotonia, deafness	11 years, erethic oligofrenia, aggressiveness	10 years, erethic oligofrenia, aggressiveness	hypotonia
Metabolites in CSF (μmol/l)						
S-Ado	269	126		260	283	
SAICAR	281	147	(urine)	125	127	(urine)
ratio S-Ado/SAICAr	1	0.9	1.2	2.2	2.1	2.6
Residual activity in fibroblasts (%)a						
S-AMP	nd	28 ± 4	37 ± 2	39 ± 2	37 ± 6	24 ± 2
SAICAR	nd	41 ± 1	53 ± 6	46 ± 2	46 ± 2	40 ± 1
S-AMP/SAICARb	nd	0.9 ± 0.1	0.9 ± 0.1	1.2 ± 0.2	1.1 ± 0.1	0.8 ± 0.0
ADSL mutations	R194C	A3V	R426H	Y114H	Y114H	R426H
	D268N	R337X	R426H	R190Q	R190Q	D430N

[a]ADSL activities in controls (n = 5), 2.23 ± 0.38 and 1.56 ± 0.28 nmol/min/mg protein with S-AMP and SAICAR, respectively.
[b]S-AMP/SAICAR activity ratios; controls (n = 5), 1.44 ± 0.16.
nd, not determined; PMR, psychomotor retardation.

kinetic parameters of the full-length enzyme (Table 5) were identical to those of the native enzyme isolated from human erythrocytes (26), rat muscle (27) and even to those of the truncated enzyme kept soluble under specific conditions (22). This implies that the identified 25 amino acid N-terminal ADSL sequence is most probably not essential for the enzyme cata-

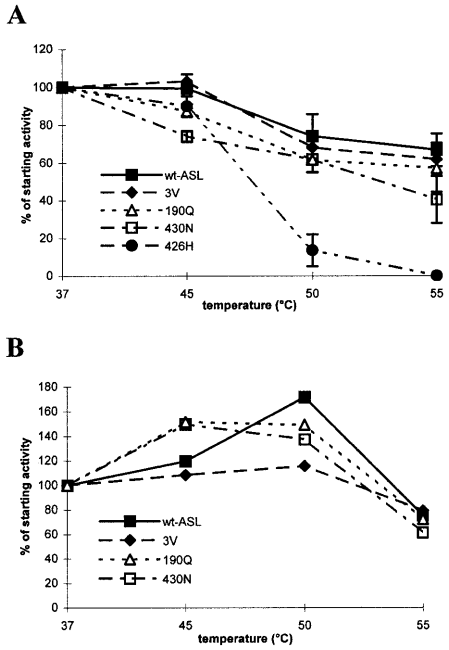


Figure 5. Thermal stability and temperature dependence of ADSL activity of mutated proteins displaying 'normal' ADSL activities. The starting activity represents the protein activity under standard reaction conditions (see Materials and Methods). (**A**) Thermal stability is defined as an enzyme activity measured at 37°C following incubation at a given temperature for 35 min (means of three experiments). (**B**) Temperature dependence of ADSL activity is defined as an enzyme activity measured at a given temperature after a 35 min incubation (results of single determination).

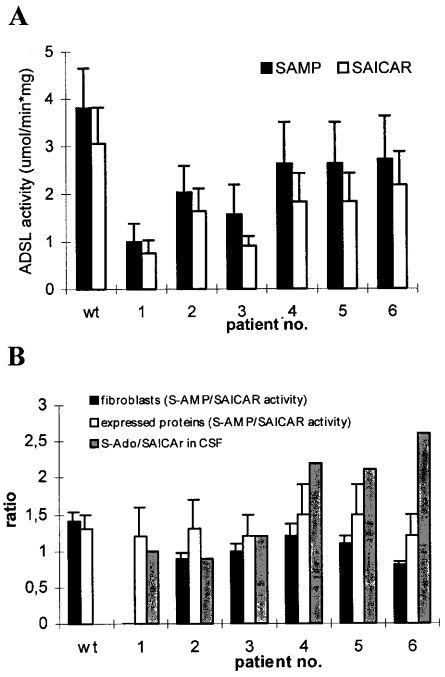


Figure 6. (**A**) Predicted residual ADSL activities of individual genotypes calculated as a mean of the homoallelic *in vitro* activities. (**B**) S-AMP/SAICAR activity ratios in fibroblasts and of expressed proteins compared with the S-Ado/SAICAr ratio found in the CSF of individual patients.

lytic competency but may be important for its structural stability. This is further corroborated by the nearly identical kinetic properties but altered temperature activation profiles (Table 5, Fig. 5B) of the A3V mutated protein present in the severely affected patient 2.

All the above corroborated by the identification of two mutations within the said N-terminal sequence (ref. 25 and this study) and a high degree of sequence identity within this part of the sequence with the murine *ADSL* cDNA (20) provides a strong argument that the native human protein starts with the M1 and is composed of 484 amino acids, the same as the murine ADSL.

ADSL isoforms

Two *ADSL* mRNA isoforms produced by an alternative splicing of exon 12 have been identified. The presence and the relative abundance of each transcript were studied in a normalized brain, heart, kidney, liver, lung, pancreas, placenta and skeletal muscle cDNA, using RT–PCR. The results showed that both transcripts are expressed in all tissues studied, in the same relative amounts with the unspliced form being ~10-fold more abundant. The exon 12 deletion preserves an open reading frame of the unspliced wtADSL mRNA. If translated, the enzyme would miss 59 amino acids (position 397–456 of the ADSL). To test the kinetic properties of the alternatively spliced isoform (ΔADSL) we cloned the Δex12 *ADSL* cDNA and expressed it as a recombinant protein. The ΔADSL was produced in an amount comparable to the wild-type ADSL. However, the protein was completely inactive and its kinetic parameters therefore could not be determined. The functional relevance of the ΔADSL isoform remains to be elucidated. Its mRNA is stable, is expressed in all major tissues and is most probably translated. Apparent is the question of whether the ΔADSL isoform could form mixed tetramers with full-sized subunits. If so, an array of enzymes with different activities would be possible, dependent on the composition of the tetramer. This may also impact relative expression in those individuals harbouring a mutant subunit, in which case an even larger number of subunit combinations would be possible, which may contribute to quantitative differences in ADSL activities across individuals. We intend to study the presence of the predicted protein isoform using a set of polyclonal anti-ADSL antibodies.

ADSL gene structure and genomic sequence

Complete genomic sequence and gene structure of *ADSL* were described here for the first time. To initiate analysis of the potential gene expression regulating mechanisms, the 500 bp upstream sequence from the 5′-end of the cDNA sequence reported here was computer-analysed using database searches. The promoter region, the promotor activity of which was established recently (28), has features typical of a housekeeping gene: a high G + C content, a high frequency of CpG nucleotides, and no TATA and CAAT boxes. No motifs suggesting tissue-specific expression regulation were found. Furthermore, knowledge of the genomic *ADSL* sequence allowed design of a panel of intronic primer pairs for PCR amplification and subsequent sequence analysis of the *ADSL*

Table 3. Primers and restriction enzymes used for mutation analysis and cloning of the *ADSL* cDNA

Purpose	5′-end position	5′–3′ sequence
Full-length cDNA amplification	6S	CTTCCGCTCTTCCCTGGTCC
	1542AS	AACAATGATTTTCGTTTAATTCTCTTCC
cDNA sequencing	101S	CCTCTTGCCTCCCGCTAT
(Cy5, fluorescein labeled)	474S	CTCGGCTTGCCGACTTTG
	823S	TAGCTTGGGGGCATCAGT
	1195S	GGCCACAGAGAACATCAT
	495AS	TTAGCAAAGTCGGCAAGC
	871AS	CAGGAGGCGTATGTCGGT
	1253AS	GGCAATCCTGGCGGCTAC
	1533AS	ATGATTTTCGTTTAA
pMAL-c2 cloning	56S	ATGGCGGCTGGAGGCGAT (used together with 1542AS)
Mutation-specific assays		
63C→T cDNA	35S	TGGCGGGGTCGCAGGGTTGGGATACCGG
	282AS	ATCTTGAAGTCGATGTTCTACCGGTTTG
		restriction enzyme: *Age*I
63C→T gDNA	35S	TGGCGGGGTCGCAGGGTTGGGATACCGG
	117AS	TAGCGGGAGGCAAGAGGT
		restriction enzyme: *Age*I
63C cDNA	41S	GTCGCAGGGTTGGGATGGC<u>C</u>GC
63T cDNA	41S	GTCGCAGGGTTGGGATGGC<u>C</u>GT
		allele-specific oligonucleotides used together with 282AS
395T→C cDNA	147S	GCGACAGGTATAAATTCCGGCCATGGC
		TAATCAAGTCAGTATTGTCTCCA<u>C</u>CAT
		restriction enzyme: *Nco*I
624G→A	341S	GGCCACTGCTGTCCAAAAGCTGCAGGC
	651AS	TTTACTCCCCGGAAGCGCAGGTCA<u>C</u>CT
		restriction enzyme: *Sty*I
635C→T gDNA	607S	GAACTTGAAGCGTGTCCGAGATGAC<u>A</u>TG
	intron 5	GTCCAGTGTGAGAATGCCTCAAC<u>A</u>TGT
		restriction enzyme: *Bsp*LU11I
857G→A gDNA	823S	TAGCTTGGGGGCATCAGT
	886AS	CTCCTTGAGGTTTGCCAGGAGCCATATGT
		restriction enzyme: *Esp*1396I
1332G→A gDNA	1250S	CCAACCGAAACTGGAAACCATATGCCATGAGAAAATCAGA
	1358S	GACTGAAGTAGGCATCAACCTG<u>C</u>ATA
		restriction enzyme: *Nde*I
1343G→A gDNA	1247S	GATTGCCATGAGAAAATCAG
	1423AS	CTGCTGGGAGGCACGACCAG
		restriction enzyme: *Mse*I

Mismatches used for a specific restriction site creation are underlined, the restriction enzymes used are indicated.

Table 4. Primers used for amplification of *ADSL* promotor, exons and flanking regions from gDNA.

Sequence	Product size (bp)	Strand	5′–3′ sequence
Promoter, exon 1	476	S	TAAAGTGTAGCGCGGCTAAGTA
		AS	GTCCTCACCGCCCAGCCCCTCA
Exon 2	403	S	TTTTTCCTTGGTGTCACTTCAT
		AS	GTCACTTCCTCAACTGGACAAC
Exons 3,4	3313	S	TTTCTGGCCAAAAAGTGTTATG
		AS	CTTCACCTCAAGGTTACTGTGG
Exons 5,6	660	S	TGGTTATTAAAAGAAGCAATGA
		AS	AGAAAAAGGCAAAAATTGTAAG
Exon 7	400	S	TTAACCTCCTGAGTTAAAGCAG
Exons 8,9	593	S	CTTTCATCAGCCTAGTCACAGC
		AS	GGAGACTGTAGAGATGCTCAAA
Exons 10,11	1602	S	ACAGTCAGTAGGGCAACTTGTT
		AS	TAAACACCTTCATCTGCTCTCC
Exon 12	399	S	AGACTGCATGGATGGGAAGTAT
		AS	CCATTCCCACCATACCTCCTAT
Exon 13	529	S	GTATCCCCTGACATTGGAAAAG
		AS	GACAGGACCAGAAATTTCTTGC
Overhangs for bi-directional sequencing of PCR products synthesized on the 5′-ends of above primers			
For T7 sequencing primer		S	GAAACAGCTATGACCATG
For M13 reverse sequencing primer		AS	ATACGACTCACTATAGGGC

Table 5. Kinetic properties of wild-type and three mutated ADSL proteins

	K_M (µmol/dm3)		V_{lim} (µmol/dm3/min)		k (mmol/g/min)	
	S-AMP	SAICAR	S-AMP	SAICAR	S-AMP	SAICAR
wtADSL	4.9 ± 0.5	3.6 ± 0.3	0.60 ± 0.01	0.32 ± 0.01	4.29	3.56
3V	5.1 ± 0.8	2.5 ± 0.3	0.66 ± 0.02	0.30 ± 0.01	4.71	3.33
190Q	8.0 ± 1.0	4.3 ± 0.5	1.06 ± 0.04	0.45 ± 0.02	7.57	5.00
430N	4.2 ± 0.8	2.3 ± 0.2	0.66 ± 0.02	0.29 ± 0.01	4.71	3.22

promotor and coding sequences. These data provide a basis for reliable postnatal and prenatal DNA-based diagnosis of ADSL deficiency on gDNA.

Mutation analysis

We identified eight mutations in a series of six patients, increasing the total amount of mutations identified within the *ADSL* gene to 21 [(4,18,24,25,28) and ADSL database (http://www.icp.ucl.ac.be/adsldb/mutations.html)]. All the mutations represent single base pair substitutions. These include 18 missense mutations, one nonsense mutation, one splice site mutation and one mutation in the 5′-UTR (28) of the mRNA. The most frequent mutation, R426H, has been detected in 10 homozygotes and six compound heterozygotes (ADSL database, refs 4,25,28 and this work) and thus accounts thus for more than 50% of the mutated alleles analysed so far. Only two other mutations, R190Q, found in the Czech and Belgian patients, and K246E, found in two unrelated Belgian patients (25), were present on more than one allele. The majority of the mutations, 14, occurred within exons 1, 2, 5 and 12. It is worth mentioning that except for the R426H mutation the other four mutations, located within the alternatively spliced exon 12, were found in the less severely affected patients. This emphasizes further the importance of studying the biological relevance of the ΔADSL isoform.

The allelic heterogeneity and different ethnic origin of the ADSL-deficient patients (Slavic, Romany, Moroccan, Turkish, Spanish, Italian, Dutch, Belgian, German, Australian, US and Brazilian) suggests that ADSL deficiency is panethnic and is probably more common than previously thought. Thanks to the simple and inexpensive screening procedures available (9–12),

ADSL deficiency should be tested in all children with congenital psychomotor retardation and neurological involvement. This approach succeeded in diagnosing five patients out of 2000 children screened for the disorder by our laboratory (3).

Expression studies of the mutated proteins

To see the functional effect of individual mutations we expressed seven relevant amino acid sequence-altering alleles in *E.coli* and measured enzyme activities of the affinity-purified enzyme proteins with each ADSL substrate. According to the residual activities, the mutant enzymes were divided into three groups: null mutations free of detectable enzymic activity (Y114H and D268N), severe mutations with enzyme activities substantially compromised (R194C and R426H) and mild mutations with activities comparable to that of the wild-type enzyme (A3V, R190Q and D430N). We propose the following mechanisms to explain the causative effect in the category of null mutations. In the case of the Y114H mutation it may be the pronounced protein instability (see Fig. 3). The wild-type tyrosine-114 resides in a highly conserved region (TSC**Y**VGDN), common to all members of the fumarase enzyme superfamily. Based on molecular modelling, this region was predicted to be essential for the *Bacillus subtilis* ADSL subunit assembly (29). Furthermore, the active site of the *B.subtilis* ADSL is formed by histidine–histidine residues pairing (29). One of these histidines, His68, is spatially close to this conserved region. Based on the homology with the *B.subtilis* enzyme it could be that the mutated histidine-114 residue in this region interferes with a proper active site formation. Both mechanisms, the hampered monomer association and improper active site formation, or even a combination of both can thus lead to the observed enzyme instability and inactivity. The diminished activity of the D268N mutant is easily explainable, as the wild-type aspartic acid residue resides in the active enzyme site and is directly involved in the catalytic mechanism where it serves as an electron donor of the acidobasic cleavage (22). The exchange of this acidic residue for the uncharged asparagine leads to active centre disruption but does not seem to interfere with the enzyme protein stability. The effect of the severe and mild mutations is very hard to explain using the experimental data. There is no doubt that the mutations compromising severely the enzymic activity are disease causing. To be able to make such a statement for the group of the mild mutations, temperature activation profiles and kinetic parameters of the expressed proteins were studied. All three mutants, compared with wild-type enzyme, showed modest changes in thermal stability and temperature dependence of ADSL activity suggesting the protein susceptibility to accelerated intracellular degradation.

Genotype–phenotype correlation

As in other metabolic disorders affecting the central nervous system, the relation between the enzyme defect and clinical symptoms is not well understood. Two main hypotheses of the pathophysiological mechanism in ADSL deficiency have been proposed. Impaired synthesis of purine nucleotides leading to purine and particularly adenine nucleotide deficiency was ruled out by the work of Van den Berghe *et al.* (16). On the other hand, there is evidence for interference of the accumulating succinylpurines with brain glucose metabolism in human patients (30), and of toxic effect of SAICAr on neurons in specific regions of rat hippocampus (31). An inverse relationship between the degree of clinical involvement and the excess of S-Ado over SAICAr has been proposed (8,14), providing further argument in favour of the toxic effect of succinylpurines.

Our attempts at correlating the severity of the clinical course with biochemistry can be summarized as follows. No strict correlation of genotype with phenotype could be drawn. With no exception all the patients carried at least one allele showing some residual activity, supporting the assumption that complete loss of ADSL activity is probably not compatible with life. The tendency in patients carrying a combination of milder mutations to be less affected suggests the pivotal role of the residual enzyme activity in determining severity of symptoms. However, we could not prove any difference in activities using either one of the two substrates with all the expressed enzymes as well as with the patients' fibroblast cell lines. Such a finding would represent an important link between the modulated enzyme activity and the varied body fluid S-Ado/SAICAr ratio, suggesting an inverse correlation with disease severity (8,14).

MATERIALS AND METHODS

Patients

Six patients from five unrelated families (four of Czech and one of American origin) were studied. A brief summary of the clinical and biochemical data is given in Table 2. More detailed case histories together with clinical and biochemical data of the Czech patients have already been published (3). The US case was published recently (5).

5′ and 3′ random amplification of cDNA ends (RACE)

To prepare the 'full-length' *ADSL* cDNA we used the CapFinder PCR cDNA Library Construction Kit (Clontech). Briefly, muscle and lymphocyte RNAs were reverse transcribed by MMLV reverse transcriptase in the presence of Clontech CDS/3′oligo(dT)$_{30}$ N$_{-1}$N; (N$_{-1}$ = A, C or G, N = A, C, G or T) and CapSwitch (template for reverse transcriptase after the switch at the 7-methylguanosine cap structure) oligonucleotides. Prepared cDNAs were amplified by LD-PCR using primers specific for CDS/3′ and CapSwitch oligonucleotides, so that only cDNA containing CapSwitch sequence on the 5′-end could be amplified. The cDNA was treated with T4 DNA polymerase and ligated into GenomeWalker Adaptor (Universal GenomeWalker kit, Clontech). 5′- and 3′-ends of *ADSL* cDNA were prepared by nested PCR using the *ADSL* and GenomeWalker Adaptor specific primers. The PCR products were purified and sequenced.

RT–PCR

Total RNA was isolated from peripheral blood lymphocytes, fibroblasts or skeletal muscle by the standard procedure (32). Poly(A)$^+$ RNA was isolated using the Oligotex Direct mRNA kit (Qiagen, Hilden, Germany). cDNA was reverse transcribed from total or poly(A)$^+$ RNA using the 1st-strand cDNA Synthesis kit (Clontech, Palo Alto, CA) and oligo(dT) primer. *ADSL* cDNA was amplified in PTC-200 DNA Engine (MJ

Research, Waltham, MA) in a reaction volume of 25 µl containing 2.5U Klen*Taq*1 polymerase (Ab Peptides, St Louis, MO), 0.1 U DeepVent polymerase (NEB, Beverly, MA), 200 µM dNTPs, 0.15 µM primers and 1.5 mM $MgCl_2$ under the following conditions: the initial denaturation at 94°C for 2 min followed by 30 cycles of denaturation at 94°C for 10 s, primer annealing at 59°C for 10 s and extension at 72°C for 1 min. Final chain elongation was performed at 72°C for 10 min.

The cDNAs used for cloning were amplified in a 50 µl reaction containing 1.2 U of cloned *Pfu* DNA polymerase (Stratagene, La Jolla, CA), 200 µM dNTPs, 0.10 µM primers and 2.5 mM $MgCl_2$ under the following conditions : the initial denaturation at 94°C for 1 min was followed by 30 cycles of denaturation (94°C for 10 s), primer annealing (55°C for 20 s) and extension (72°C for 4 min). Final chain elongation was carried out at 72°C for 10 min.

DNA sequencing

cDNAs prepared by RT–PCR were purified and concentrated on Microcon 100 microconcentrators (Amicon, Beverly, MA), while the plasmid DNA samples were prepared using the miniprep kits SNAP (Invitrogen, NV Leek, The Netherlands). Dideoxy cycle sequencing was performed in 7 µl containing 150 fmol of template, 1 pmol of the fluorescently labelled primer, 100 µM dNTP (dNTP:ddNTP ratio = 120), 5 mM $MgCl_2$ and Ampli*Taq* FS polymerase (PE Biosystems, Foster City, CA). Sequencing parameters were: initial denaturation (94°C for 2 min), 40 cycles of denaturation (94°C for 15 s), annealing of primers at specific temperature for 10 s and extension (72°C for 30 s), followed by 10 cycles of denaturation (94°C for 15 s) and extension at 72°C for 30 s. The samples were subsequently denatured in loading buffer (94°C for 3 min). Sequences were read either on the ALF sequencer—EMBL prototype (EMBL, Heidelberg, Germany) (fluorescein primers) or on the ALFExpress sequencer (Pharmacia, Uppsala, Sweden) (Cy5 primers). Sequences were analysed using the GeneSkipper software (EMBL). Both strands of cDNA were always sequenced. All the primers (Generi Biotech, Hradec Králové, Czech Republic) used are listed in Table 3.

Mutation analysis

To verify the nature of mutations and to be able to screen for individual mutations in affected families, we designed PCR–RFLP-based assays on genomic DNA. These assays were all based on the PCR-based introduction of either one or both, diagnostic and control, restriction enzyme recognition sequence sites. The primers and restriction enzymes used are listed in Table 3. The genomic DNA for PCR analysis was isolated either from blood samples (Qiagen, Hilden, Germany) or extracted from dried blood spots (33).

Genomic DNA sequencing

The PCRs were carried out according to standard procedure. Genomic DNA (100 ng) was amplified in 25 µl containing 2.5 U Klen*Taq*1 polymerase, 0.1 U DeepVent polymerase, 200 µM dNTPs and 0.15 µM primers. Amplification products were gel purified using the phenol-based method (34). Purified fragments were sequenced using the T7 and M13 primers. The dideoxy cycle sequencing was performed in a 7 µl reaction containing 150 fmol of template, 1 pmol of the Cy5 labelled primers, 100 µM dNTP (dNTP:ddNTP ratio = 120), 5 mM $MgCl_2$ and Ampli*Taq* FS polymerase. The sequences were read on ALFExpress sequencer. Sequences were analysed using the GeneSkipper software.

Alternative splicing

To reveal the identity of the shorter fragment we prepared the RT–PCR product as described above. After 30 cycles, 1 U *Taq* polymerase was added and the mixture was incubated for 15 min at 72°C. DNA was phenol–chloroform purified and ligated into pCRII-TOPO vector (Invitrogen). The OneShot cells (Invitrogen) were transformed according to the manufacturer's protocol and grown on Luria–Bertani (LB) plates containing 50 µg/ml ampicillin, X-gal and isopropyl 1-thio-β-D-galactoside (IPTG). The lithium minipreps were prepared from positive clones. The inserts were cleaved out of the vector by *Eco*RI and clones containing the fragment of interest were sequenced as described above. Sequence alignment was performed with the CLUSTAL W V1.5 program (35).

Tissue-specific expression patterns of ADSL isoforms

The tissue-specific expression pattern of *ADSL* mRNA isoforms was investigated by PCR, using the Human Multiple Tissue cDNA (MTC) Panel I (Clontech) as a template. *ADSL* cDNA was amplified as described above. A control fragment of glyceraldehyde-3-phosphate dehydrogenase (G3PDH) was amplified according to the manufacturer's protocol. The PCR products were analysed on 1% SeaKem agarose (FMC, Rockland, ME).

Expression of *ADSL* isoforms in *E.coli*

The *ADSL* cDNA was prepared and cloned into pCRII-TOPO vector as described above. The constructs of the correct sequence were used as templates for subsequent PCR amplification of *ADSL* cDNA using the 56S primer starting at the first ATG codon. PCR products were blunt end ligated into the pMAL-c2 vector (NEB). The constructs with the correct sequences were introduced into the *E.coli* strain DH5αF′IQ (Gibco, Paisley, UK) for fusion protein production. Briefly, a 600 ml culture of transformed bacteria was grown at 37°C in rich medium, 100 µg/ml ampicillin and 0.2% glucose to a density A_{600} of 0.5. IPTG was added to a final concentration of 0.3 mM and incubation continued for an additional 3 h. Bacteria were harvested by centrifugation at 3000 *g* for 10 min. The pellets were resuspended in 20 mM Tris–HCl pH 7.4, 200 mM NaCl and 1 mM EDTA buffer and sonicated four times for 15 s at 40 W. Crude lysates were obtained by centrifugation at 8000 *g* for 30 min. Fusion proteins were isolated from crude lysate on amylose affinity columns (NEB). The fusion proteins were cleaved overnight at 4°C using factor Xa (NEB). The ADSL activities were measured in crude lysate, isolated fusion protein and in cleaved fusion using the HPLC analysis of AMP and AICAR formed from both ADSL substrates, S-AMP (Sigma) and SAICAR (own synthesis, unpublished data). The reactions were run for 20 min at 37°C in 150 µl containing 10 mM Tris pH 7.5, 2 mM EDTA, 10 mM KCl, 1 mM DTT, 4% glycerol and 15 µg protein. Substrate concentrations were

0.14 and 0.09 mM for S-AMP and SAICAR, respectively. SDS–PAGE analyses were performed according to standard procedure: protein concentrations were determined using the Bradford assay (Sigma).

Expression of mutated *ADSL* in *E.coli*

All mutated cDNAs were first cloned into pCRII-TOPO vector as described above. Individual mutated alleles were sequenced and subsequently subcloned into the pMAL-c2wt*ADSL* construct, using the appropriate set of restriction enzymes: for alleles 395C, 624A and 635T, *Eco*72I–*Nhe*I; for alleles 857A, 1332A and 1343A, *Nhe*I–*Bse*RI. Allele 63T was prepared using the mismatched 56S primer (ATGGCGGTTGGAG-GCGAT). PCR product was blunt end cloned into the pmal-c2 vector as described above. Resulting clones were sequenced and introduced into *E.coli* strain DH5αF′IQ. The protein expression experiments and enzyme activity measurements were performed in 20 ml cultures essentially as described above. The expression experiments were performed five to eight times with the enzyme activity measurements run in duplicate.

Thermal stability experiments and kinetic characterization of the expressed proteins

Thermal stability is defined as an enzyme activity measured at 37°C following the incubation at given temperatures for 35 min. Temperature dependence of ADSL activity is defined as an enzyme activity measured at given temperature after a 35 min incubation. The starting activity represents the protein activity under standard reaction condition (20 min reaction at 37°C). The thermal stability experiments were performed in three expression experiments in duplicate for each temperature indicated. The temperature dependence of ADSL activity was studied in single experiment. The Michaelis–Menten kinetics were established for both substrates at 10 different concentrations ranging from 0.5 to 130 µM. The reaction conditions for enzyme activity measurements were the same as described above.

ADSL activities in cultured fibroblasts

ADSL activities were measured in five controls and five patients' cell lines essentially as described above. Measurements for each substrate were run in duplicate using a 100 µl reaction volume and 15 µg of protein extract.

Accession numbers

The following sequences have been deposited with the GenBank database: complete *ADSL* cDNA sequence under accession no. AF067853; the alternatively spliced (Δex12) *ADSL* cDNA isoform under accession no. AF067854; the *ADSL* genomic sequence under accession no. AF106656.

ACKNOWLEDGEMENTS

We thank Dr Zumrová, Dr Švehláková, Dr Valík and Dr Jones who referred their patients or cell lines for investigation; to M. Elleder for help in manuscript preparation, and V. Kožich and M. Hřebíček for helpful discussions. This work was supported by grant 3608-3 from the Grant Agency of Ministry of Health of the Czech Republic and partly by grants VS96-127 from the Ministry of Education and 301/00/0689 from the Grant Agency of the Czech Republic.

REFERENCES

1. Jaeken, J. and Van den Berghe, G. (1984) An infantile autistic syndrome characterised by the presence of succinylpurines in body fluids. *Lancet*, **ii**, 1058–1061.
2. Van den Bergh, F.A., Bosschaart, A.N., Hageman, G., Duran, M. and Tien Poll-The, B. (1998) Adenylosuccinase deficiency with neonatal onset severe epileptic seizures and sudden death. *Neuropediatrics*, **29**, 51–53.
3. Šebesta, I., Krijt, J., Kmoch, S., Hartmannová, H., Wojda, M. and Zeman, J. (1997) Adenylosuccinase deficiency: clinical and biochemical findings in 5 Czech patients. *J. Inherit. Metab. Dis.*, **20**, 343–344.
4. Maaswinkel-Mooij, P.D., Laan, L.A., Onkenhout, W., Brouwer, O.F., Jaeken, J. and Poorthuis, B.J. (1997) Adenylosuccinase deficiency presenting with epilepsy in early infancy. *J. Inherit. Metab. Dis.*, **20**, 606–607.
5. Valik, D., Miner, P.T. and Jones, J.D. (1997) First U.S. case of adenylosuccinate lyase deficiency with severe hypotonia. *Pediatr. Neurol.*, **16**, 252–255.
6. Salerno, C., Crifo, C. and Giardini, O. (1995) Adenylosuccinase deficiency: a patient with impaired erythrocyte activity and anomalous response to intravenous fructose. *J. Inherit. Metab. Dis.*, **18**, 602–608.
7. Jaeken. J., Wadman, S.K., Duran, M., van Sprang, F.J., Beemer, F.A., Holl, R.A., Theunissen, P.M., de Cock, P., van denBergh, F., Vincent, M.F. and Van denBerghe, G. (1988) Adenylosuccinase deficiency: an inborn error of purine nucleotide synthesis. *Eur. J. Pediatr.*, **148**, 126–131.
8. Van den Berghe, G., Vincent, M.F. and Jaeken, J. (1997) Inborn errors of the purine nucleotide cycle: adenylosuccinase deficiency. *J. Inherit. Metab. Dis.*, **20**, 193–202.
9. Šebesta, I., Shobowale, M., Krijt, J. and Simmonds, H.A. (1995) Screening tests for adenylosuccinase deficiency. *Screening*, **4**, 117–124.
10. Maddocks, J. and Reed, T. (1989) Urine test for adenylosuccinase deficiency in autistic children. *Lancet*, **i**, 158–159.
11. Laikind, P.K., Seegmiller, J.E. and Gruber, H.E. (1986) Detection of 5′-phosphoribosyl-4-(*N*-succinylcarboxamide)-5-aminoimidazole in urine by use of the Bratton-Marshall reaction**:** identification of patients deficient in adenylosuccinate lyase activity. *Anal. Biochem.*, **156**, 81–90.
12. Wadman, S.K., Duran, M. and Fabery de Jonge, H. (1986) Diagnosis of inherited adenylosuccinase deficiency by thin-layer chromatography of urinary imidazoles and by automated cation exchange column chromatography of purines. *Clin. Chim. Acta*, **156**, 279–287.
13. Van den Bergh, F., Vincent, M.F., Jaeken, J. and Van den Berghe, G. (1991) Radiochemical assay of adenylosuccinase: demonstration of parallel loss of activity toward both adenylosuccinate and succinylaminoimidazole corboxamide ribotide in liver of patients with the enzyme defect. *Anal. Biochem.*, **193**, 287–291.
14. Van den Bergh, F., Vincent, M.F., Jaeken, J. and Van den Berghe, G. (1993) Residual adenylosuccinase activities in fibroblasts of adenylosuccinase-deficient children: parallel deficiency with adenylosuccinate and succinyl-AICAR in profoundly retarded patients and non-parallel deficiency in a mildly retarded girl. *J. Inherit. Metab. Dis.*, **16**, 415–424.
15. Van den Berghe, G., Van den Bergh, F., Vincent, M.F. and Jaeken, J. (1993) The biochemical aspects of Asase deficiency. In Gresser, U. (ed.), *Molecular Genetics, Biochemistry and Clinical Aspects of Inherited Disorders of Purine and Pyrimidine Metabolism*. Springer Verlag, Berlin, pp. 140–143.
16. Van den Bergh, F., Vincent, M.F., Jaeken, J. and Van den Berghe, G. (1993) Functional studies in fibroblasts of adenylosuccinase-deficient children. *J. Inherit. Metab. Dis.*, **16**, 425–434.
17. Van Keuren, M.L., Hart, I.M., Kao, F.T., Neve, R.L., Bruns. G, A., Kurnit, D.M. and Patterson, D. (1987) A somatic cell hybrid with a single human chromosome 22 corrects the defect in the CHO mutant (Ade-I) lacking adenylosuccinase activity. *Cytogenet. Cell Genet.*, **44**, 142–147.
18. Stone, R.L., Aimi, J., Barshop, B.A., Jaeken, J., Van den Berghe, G., Zalkin, H. and Dixon, J.E. (1992) A mutation in adenylosuccinate lyase associated with mental retardation and autistic features. *Nature Genet.*, **1**, 59–63.

19. Fon, E.A., Demczuk, S., Delattre, O., Thomas, G. and Rouleau, G.A. (1993) Mapping of the human adenylosuccinate lyase (ADSL) gene to chromosome 22q13.1→q13.2. *Cytogenet. Cell Genet.*, **64**, 201–203
20. Wong, L.J., and O'Brien, W.E. (1995) Characterization of the cDNA and the gene encoding murine adenylosuccinate lyase. *Genomics*, **20**, 341–343.
21. Aimi, J., Badylak, J., Williams, J., Chen, Z.D., Zalkin, H. and Dixon, J.E. (1990) Cloning of a cDNA encoding adenylosuccinate lyase by functional complementation in *Escherichia coli*. *J. Biol. Chem.*, **5**, 9011–9014.
22. Stone, R.L., Zalkin, H. and Dixon, J.E. (1993) Expression, purification, and kinetic characterization of recombinant human adenylosuccinate lyase. *J. Biol. Chem.*, **268**, 19710–19716.
23. Casey, P.J. and Lowenstein, J.M. (1987) Purification of adenylosuccinate lyase from rat skeletal muscle by a novel affinity column. Stabilization of the enzyme, and effects of anions and fluoro analogues of the substrate. *Biochem. J.*, **246**, 263–269.
24. Verginelli, D., Luckow, B., Crifo, C., Salerno, C. and Gross, M. (1998) Identification of new mutations in the adenylosuccinate lyase gene associated with impaired enzyme activity in lymphocytes and red blood cells. *Biochim. Biophys. Acta*, **1406**, 81–84.
25. Marie, S., Cuppens, H., Heuterspreute, M., Jaspers, M., Tola, E.Z., Gu, X.X., Legius, E., Vincent, M.F., Jaeken, J., Cassiman, J.J. and Van den Berghe, G. (1999) Mutation analysis in adenylosuccinate lyase deficiency, eight novel mutations in the re-evaluated full ADSL coding sequence. *Hum. Mutat.*, **13**, 197–202.
26. Barnes, L.B. and Bishop, S.H. (1975) Adenylosuccinate lyase from human erythrocytes. *Int. J. Biochem.*, **6**, 497–503.
27. Casey, P.J. and Lowenstein J.M. (1987) Purification of adenylosuccinate lyase from rat skeletal muscle by a novel affinity collumn. Stabilization of the enzyme, and effects of anions and fluoro analogues of the substrate. *Biochem. J.*, **246**, 263–269.
28. Marie, S., Race, V., Nassogne, M.C., Vincent, M.F. and Van den Berghe G. (1999) A mutation in the 5′UTR of the ADSL gene in a patient with adenylosuccinate lyase deficiency. *Cell. Mol. Biol. Lett.*, **3**, 9.
29. Lee, T.T., Worby, C., Bao, Z.Q., Dixon, J.E. and Colman, R.F. (1999) His68 and His141 are critical contributors to the intersubunit catalytic site of adenylosuccinate lyase of *Bacillus subtilis*. *Biochemistry*, **38**, 22–32.
30. De Volder, A.G., Jaeken, J., Van den Berghe, G., Bol, A., Michel, C., Cogneau, M. and Goffinet, A.M. (1988) Regional brain glucose utilization in adenylosuccinase-deficient patients measured by positron emission tomography. *Pediatr. Res.*, **24**, 238–242.
31. Stone, T.W., Roberts, L.A., Morris, B.J., Duley, J.A. and Ogilvy, H.V. (1997) Are succinylpurines neurotoxic? *Clin. Biochem.*, **3**, P41, 169.
32. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, **162**, 156–159.
33. Gregersen, N., Blakemore, A.I., Winter, V., Andresen, B., Kolvraa, S., Bolund, L., Curtis, D. and Engel, P.C. (1991) Specific diagnosis of medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in dried blood spots by a polymerase chain reaction (PCR) assay detecting a point-mutation (G985) in the MCAD gene. *Clin. Chim. Acta*, **203**, 23–34.
34. Favre, D. (1992) Improved phenol-based method for the isolation of DNA fragments from low melting temperature agarose gels. *Biotechniques*, **13**, 25–26.
35. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.

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.

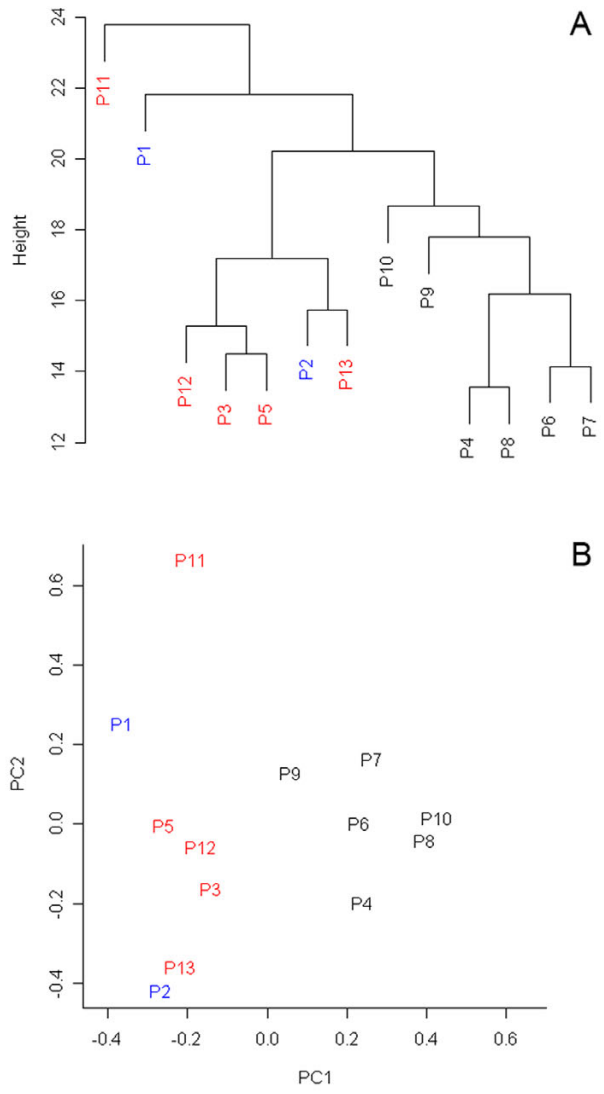


Figure 1

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

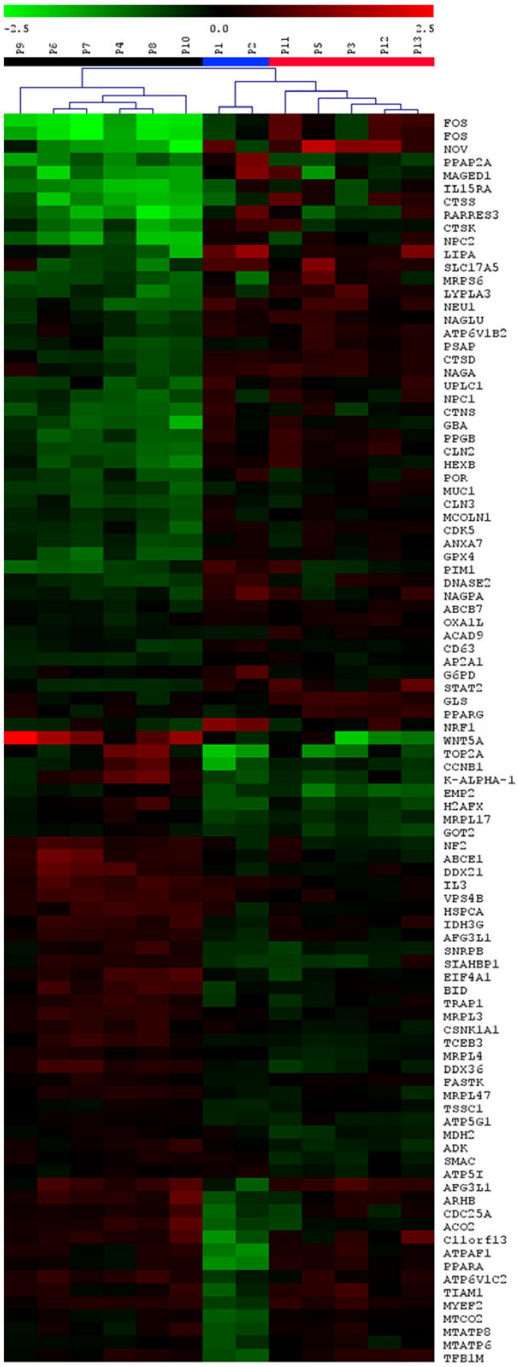


Figure 2
Differentially expressed genes defined by ANOVA analysis. Heatmap of genes detected as differentially expressed between defined patient groups using ANOVA analysis and unadjusted P < 0.01 significance level. The results are shown as Log2 ratio of relative gene expression signal in each patient 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-

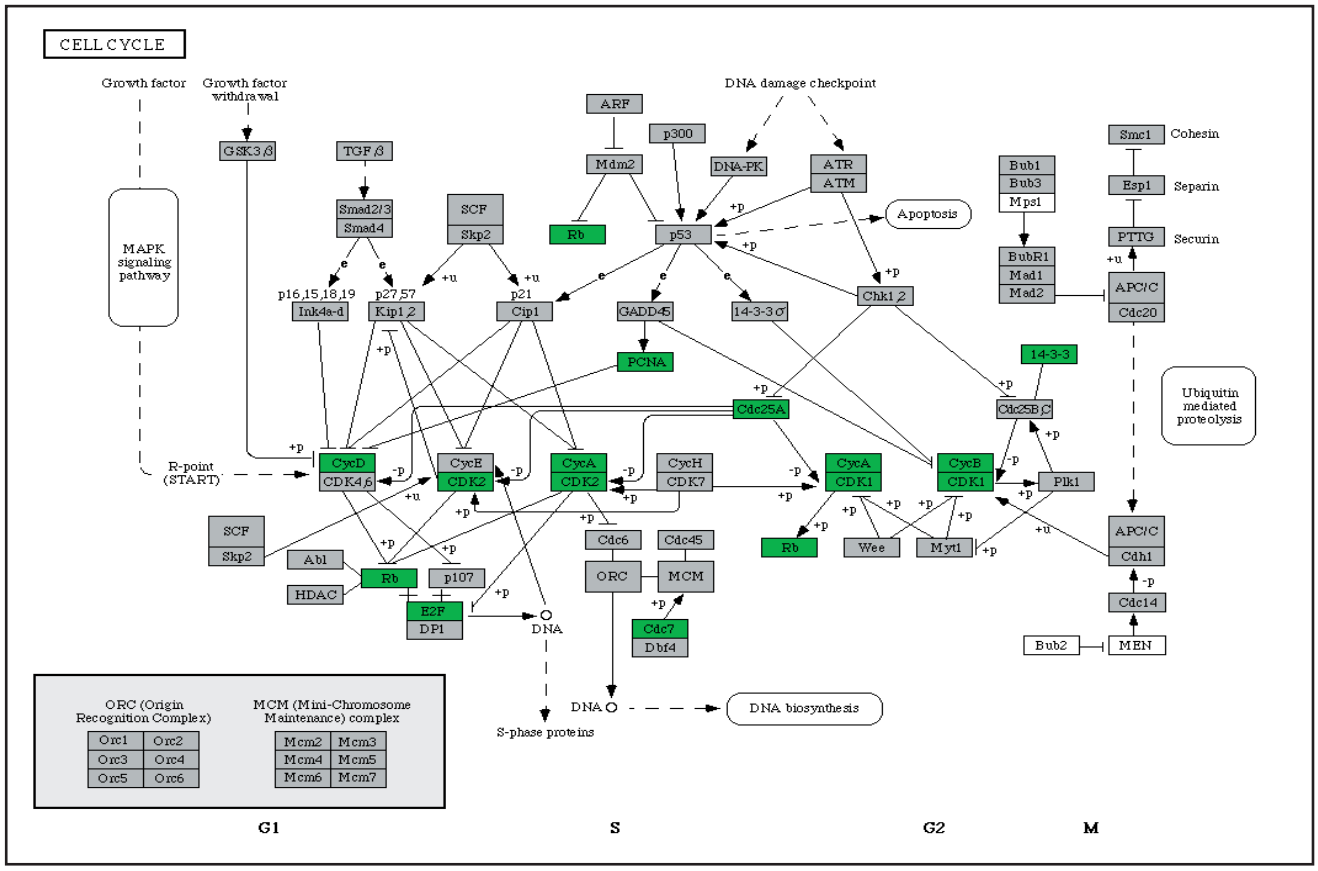


Figure 3
Gene expression changes detected in selected pathways in M group. General changes in cell cycle pathway detected in patients with mtDNA mutation (M group) using KEGGArray software.

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*.

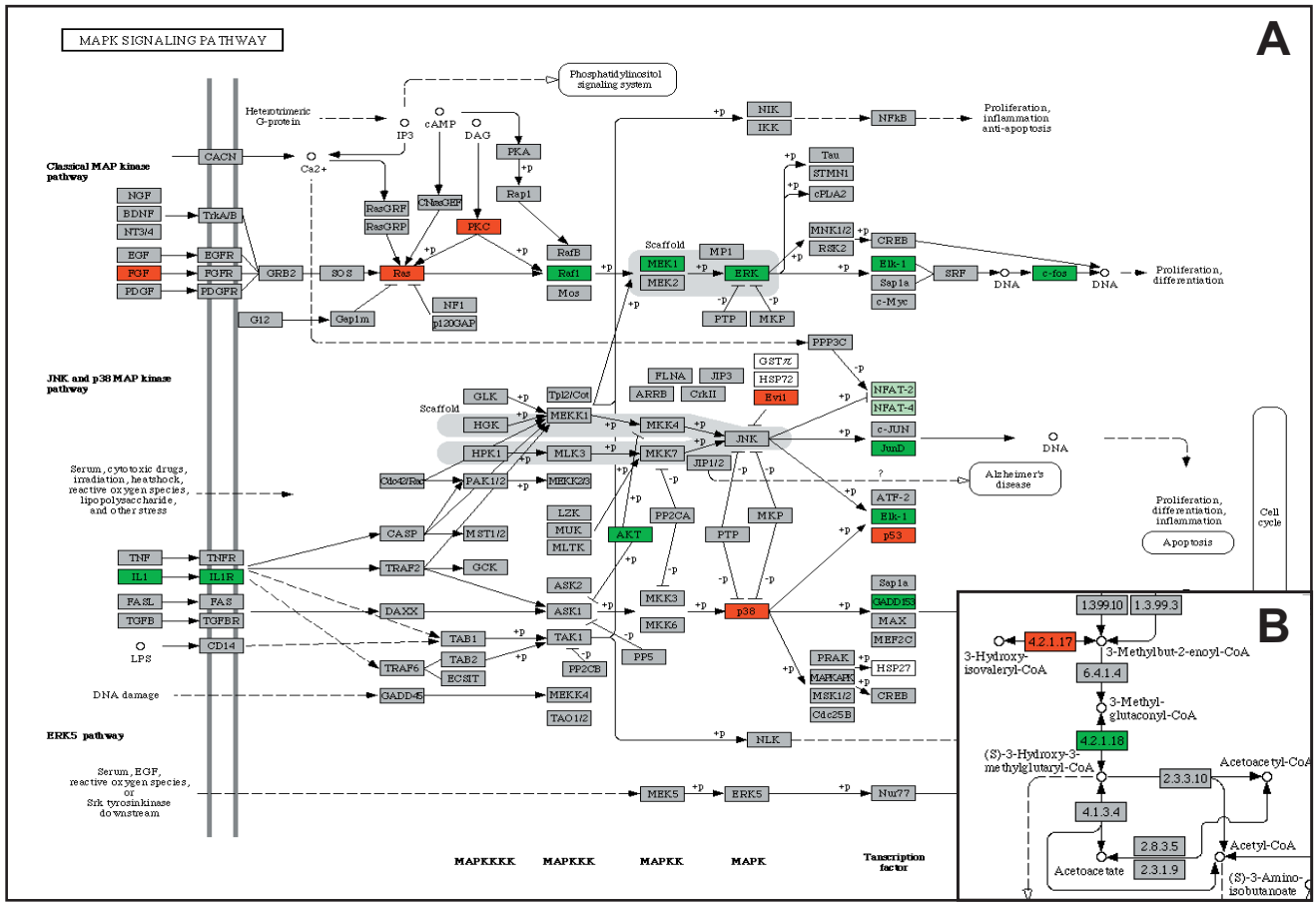


Figure 4

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.

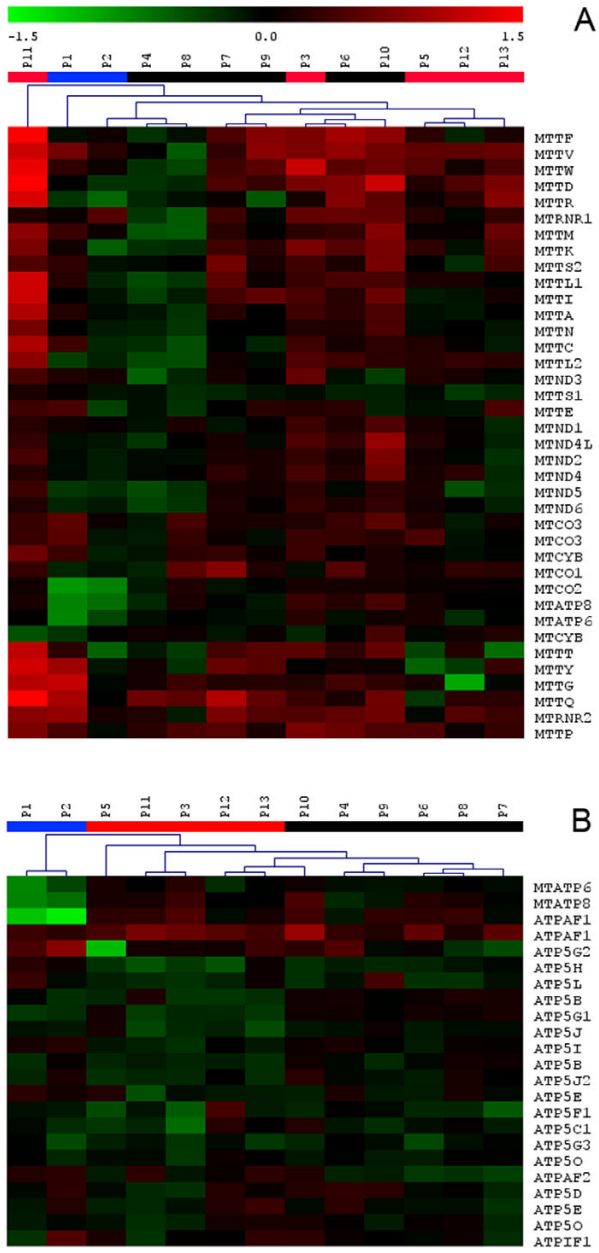


Figure 5

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

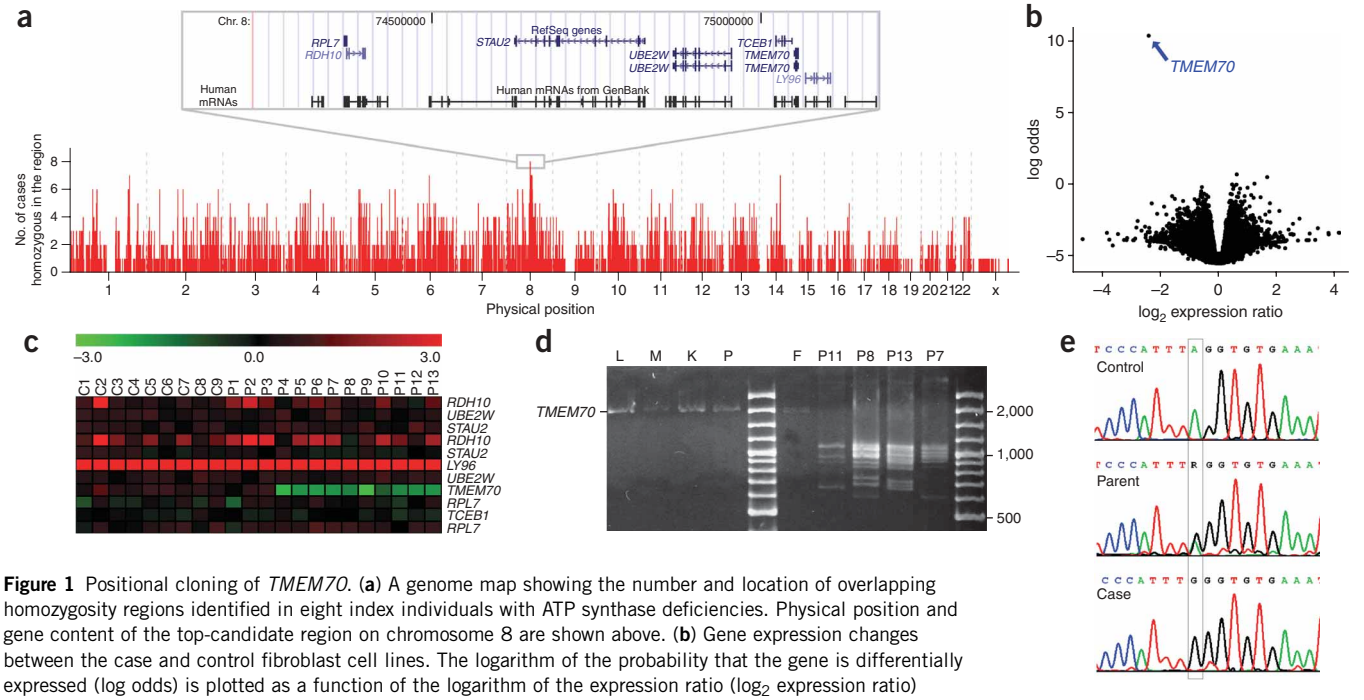


Figure 1 Positional cloning of *TMEM70*. (**a**) A genome map showing the number and location of overlapping homozygosity regions identified in eight index individuals with ATP synthase deficiencies. Physical position and gene content of the top-candidate region on chromosome 8 are shown above. (**b**) 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,

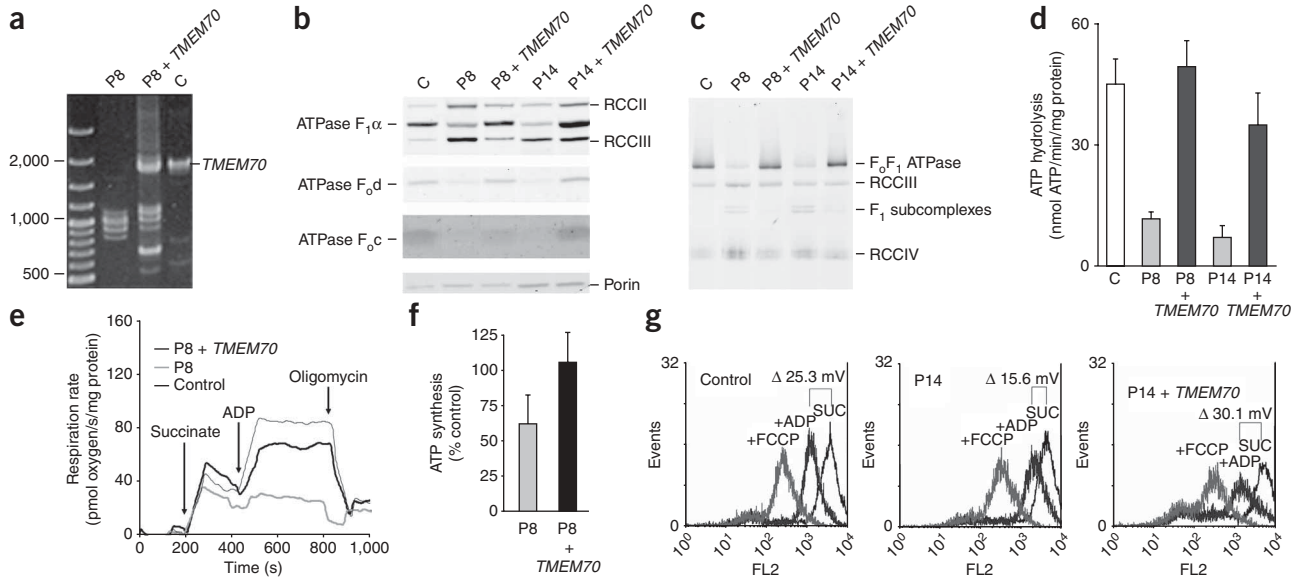


Figure 2 *TMEM70* complementation of ATP synthase deficiency. (**a**) *TMEM70* cDNA is present after transfection. (**b**) SDS-PAGE protein blot of fibroblasts shows a specific increase of the content of ATP synthase subunits relative to the respiratory chain complexes and porin. (**c**) BN-PAGE protein 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).

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.

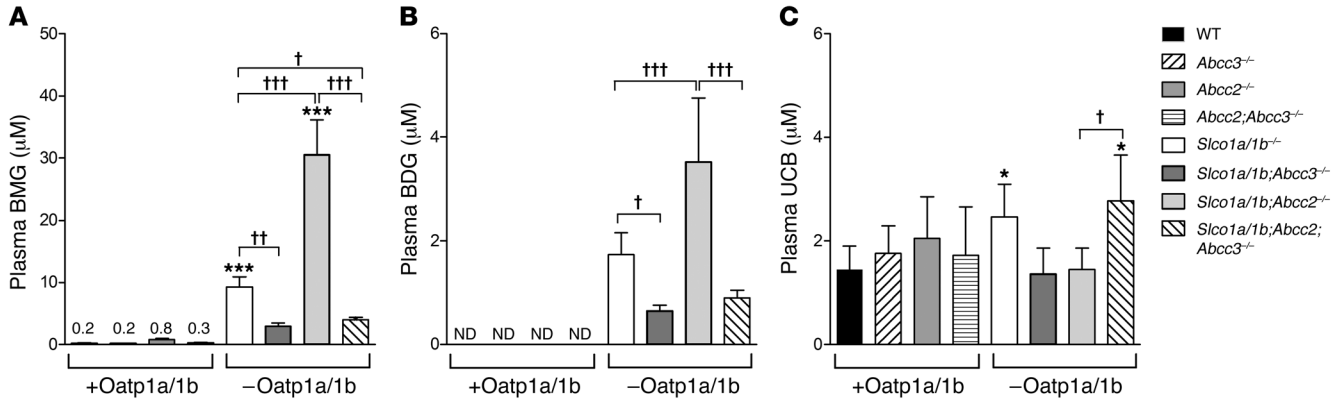


Figure 1

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: †P < 0.05, ††P < 0.01, †††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

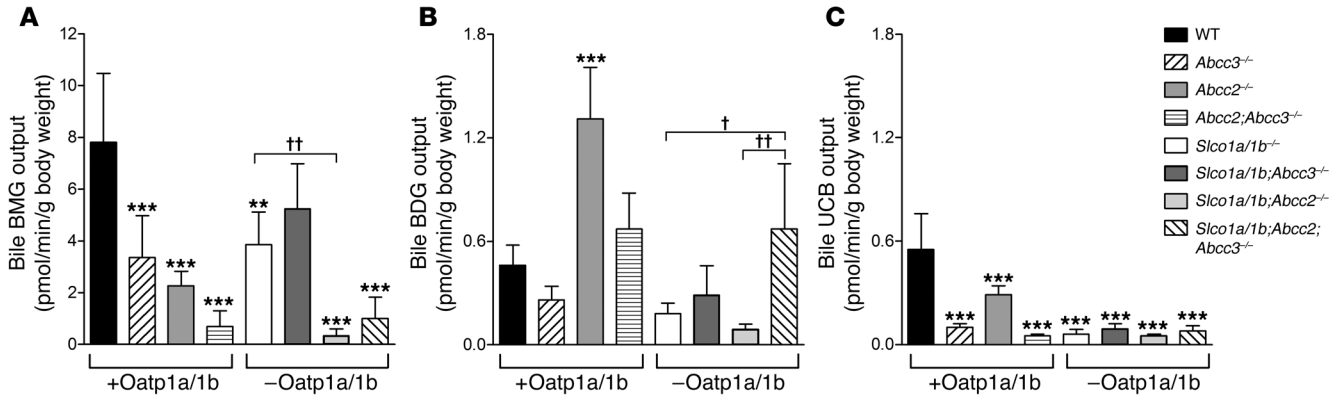


Figure 2

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

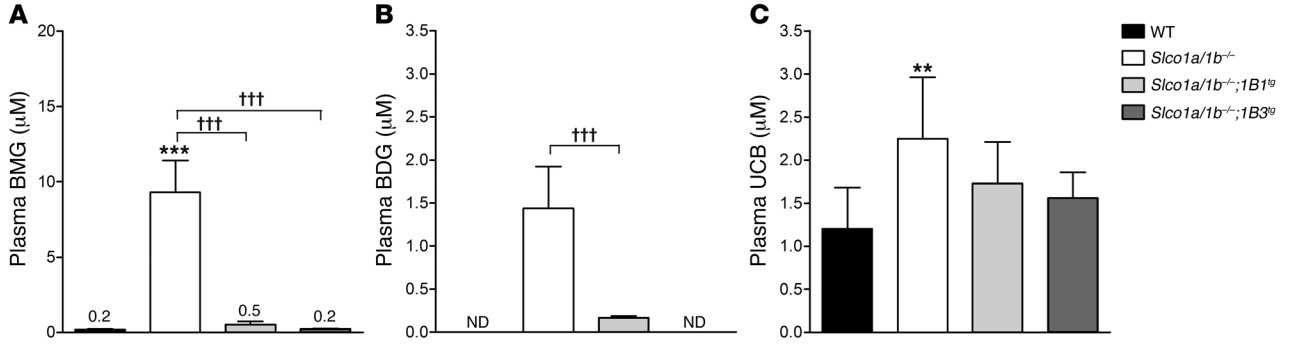


Figure 3
Increased plasma bilirubin glucuronide in *Slco1a/1b*$^{-/-}$ mice is reversed by human OATP1B1 and OATP1B3. (**A**) BMG, (**B**) BDG, and (**C**) UCB levels in plasma of male wild-type and *Slco1a/1b*$^{-/-}$ mice, and of the derived OATP1B1- and OATP1B3-transgenic strains (*Slco1a/1b*$^{-/-}$*;1B1*^{tg} 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

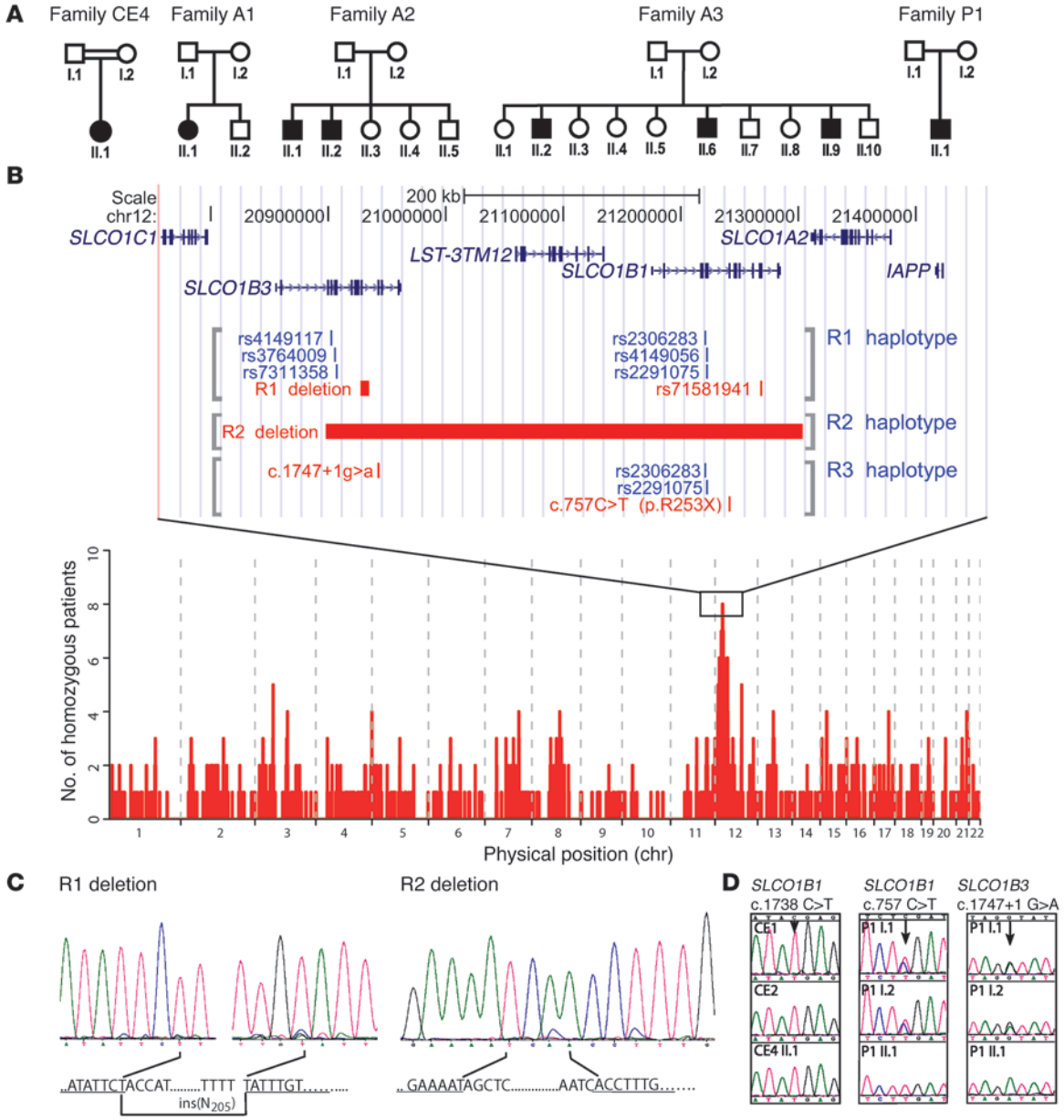


Figure 4

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

		Haplotype R1-linked mutations		Haplotype R2-linked mutations		Haplotype R3-linked mutations	
Subject	**Family status**	***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

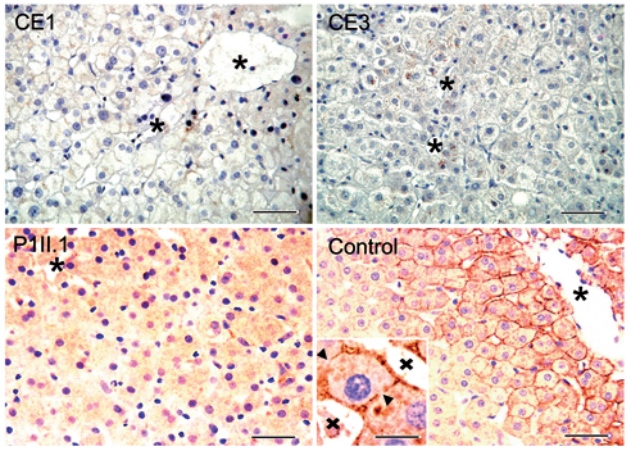


Figure 5
Liver expression of OATP1B proteins in RS subjects and control. With an anti-OATP1B1/3 antibody, basolateral membrane immunostaining of hepatocytes in centrilobular areas was intense in control. Asterisks indicate central veins, arrowheads bile canaliculi, and crosses sinusoids. OATP1B proteins 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.

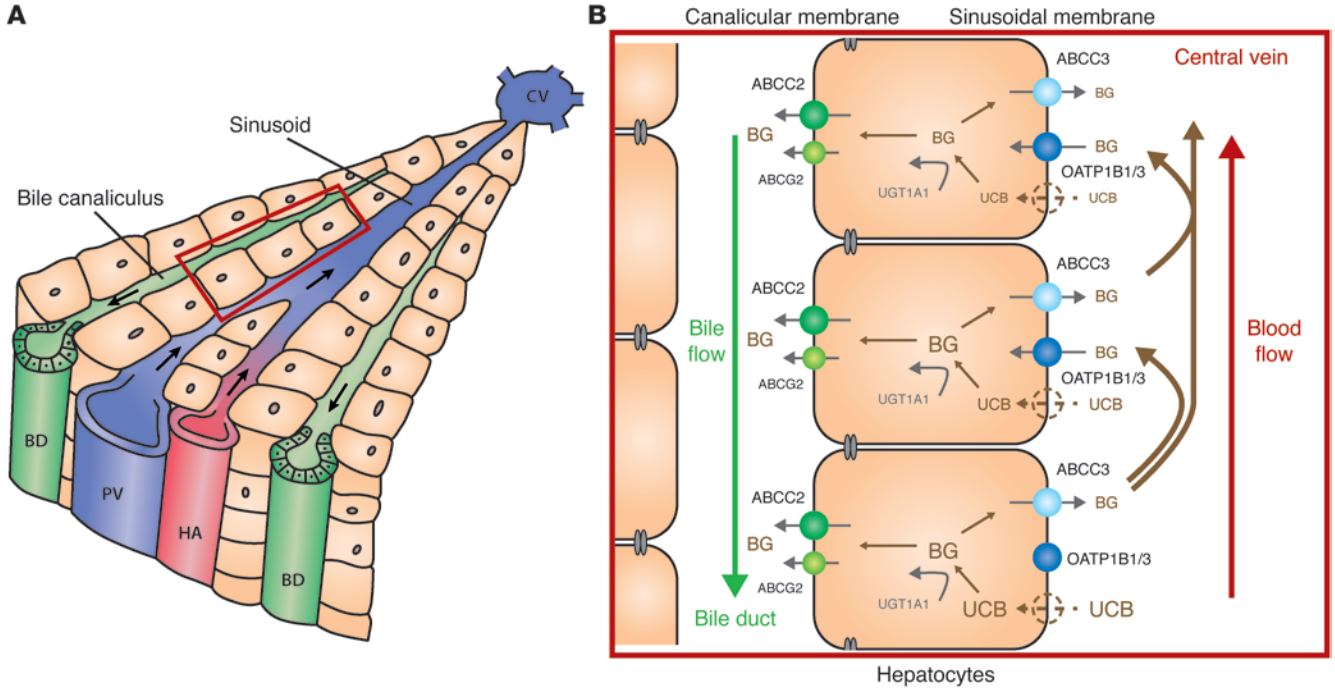


Figure 6
Hepatocyte hopping distributes the biliary excretion load of bilirubin glucuronides across the liver lobule. (**A**) Schematic of liver lobule. Hepatocytes are organized around portal tracts, with branches of the portal vein (PV), hepatic artery (HA), and bile ducts (BD). The PV and HA deliver 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.

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

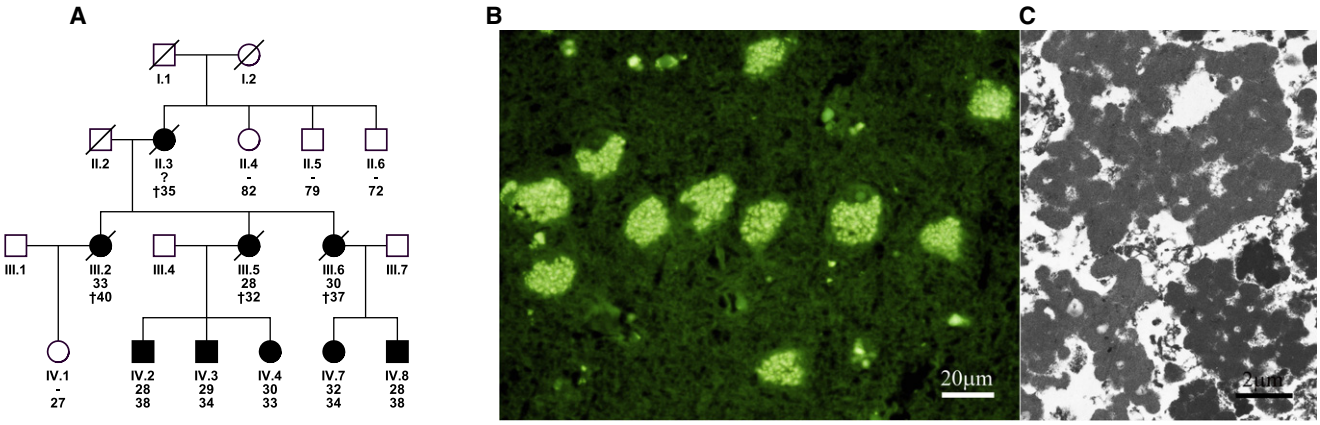


Figure 1. Pedigree and Neuropathology Findings in Family P1
(A) Pedigree of the Czech family. Black symbols denote affected individuals; open symbols denote unaffected individuals. Age of onset is shown above current age or age of death (indicated by †).
(B) Epifluorescence. Hippocampal pyramidal neurons 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

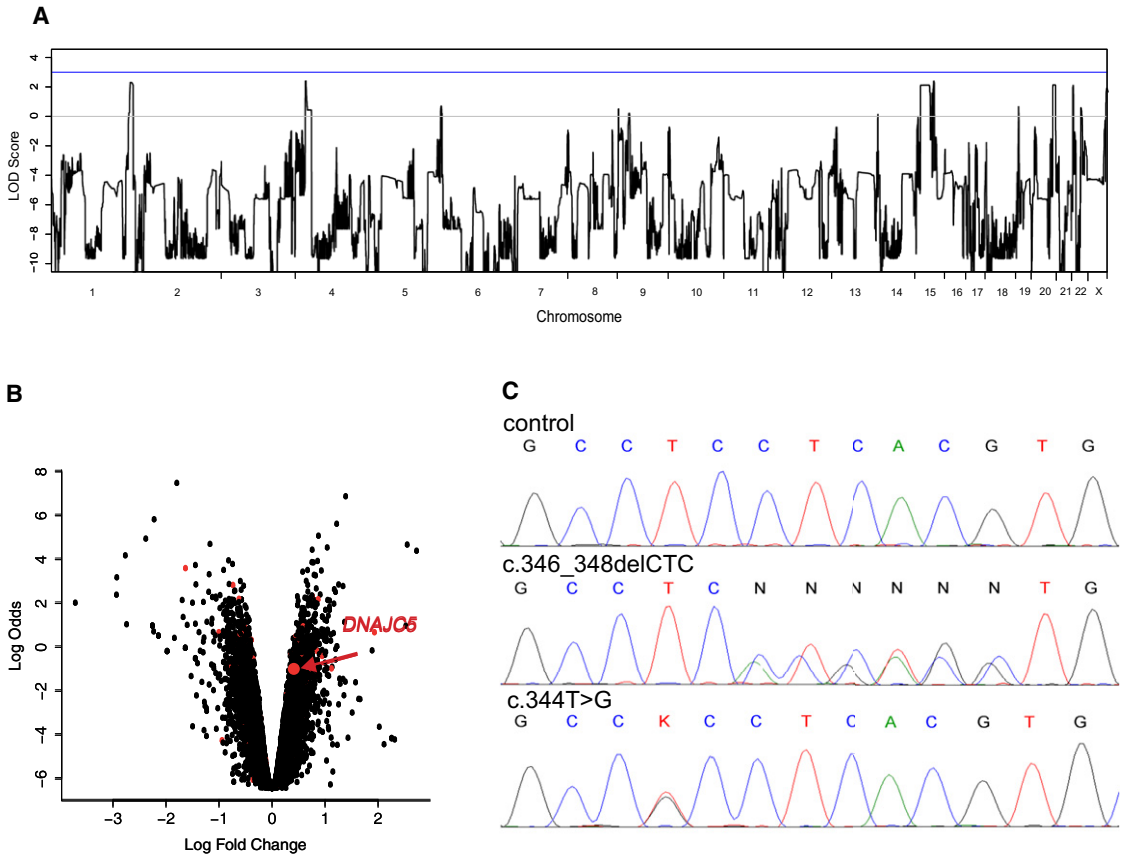


Figure 2. Identification of *DNAJC5* Mutations

(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

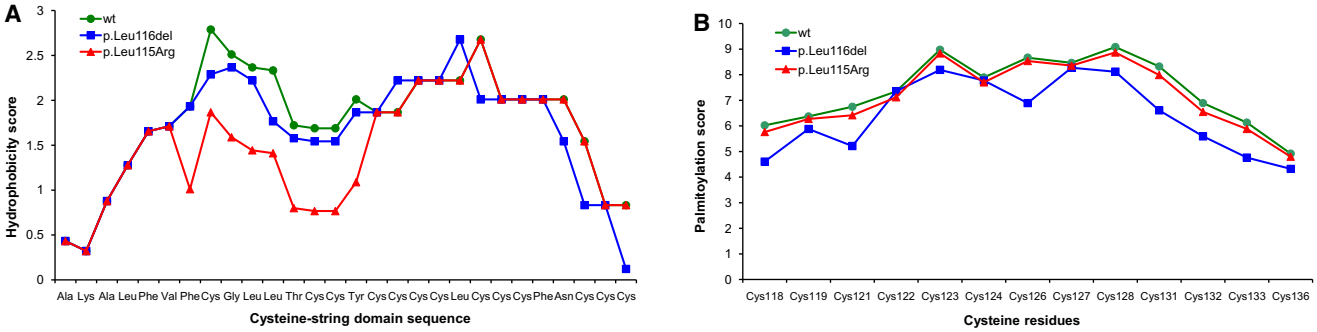


Figure 3. In Silico Analysis of Properties of the Cysteine-String Domain
(A) p.Leu115Arg mutation decreases the hydrophobicity of this domain, which is needed for initial binding to the ER.
(B) The p.Leu116del mutation decreases the palmitoylation score, that is, the confidence that cysteine residues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

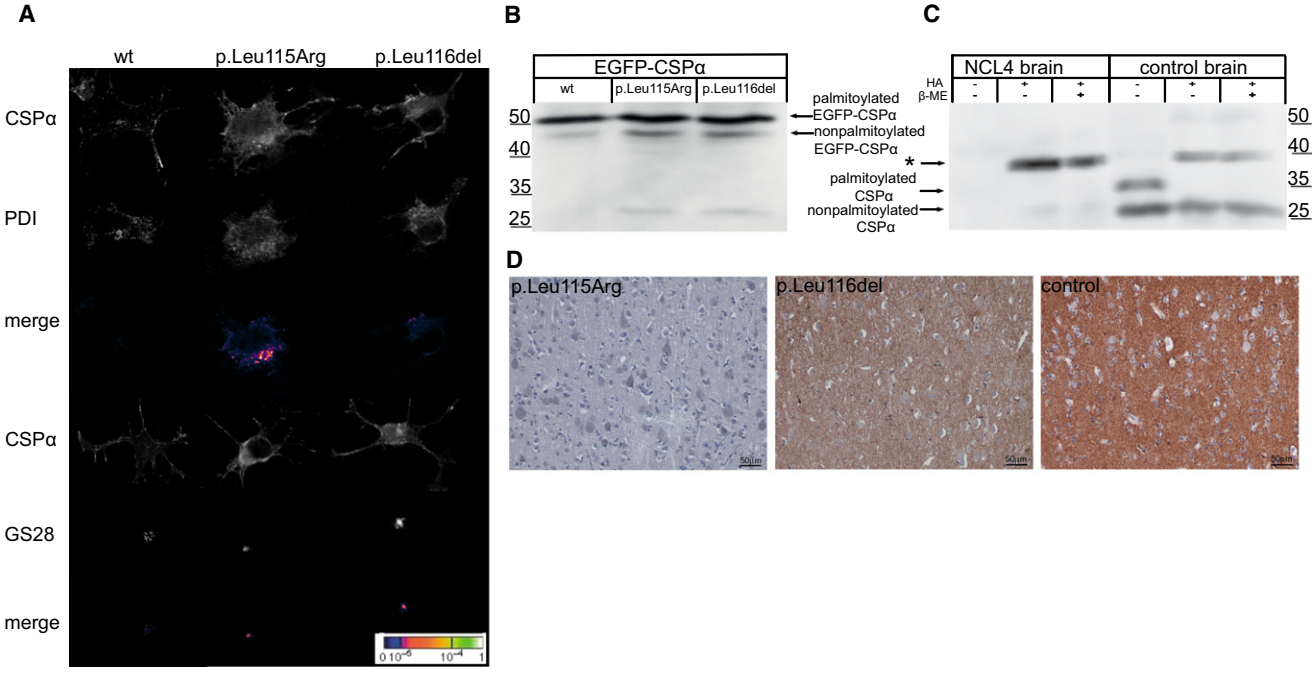


Figure 4. Characterization of Mutated CSPα
(A) Immunofluorescence analysis of transiently expressed EGFP-CSPα proteins in CAD5 cells showing prominent membrane localization of wild-type CSPα compared to the diffuse cytoplasmic staining and marked colocalization of mutated CSPα with endoplasmic reticulum 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

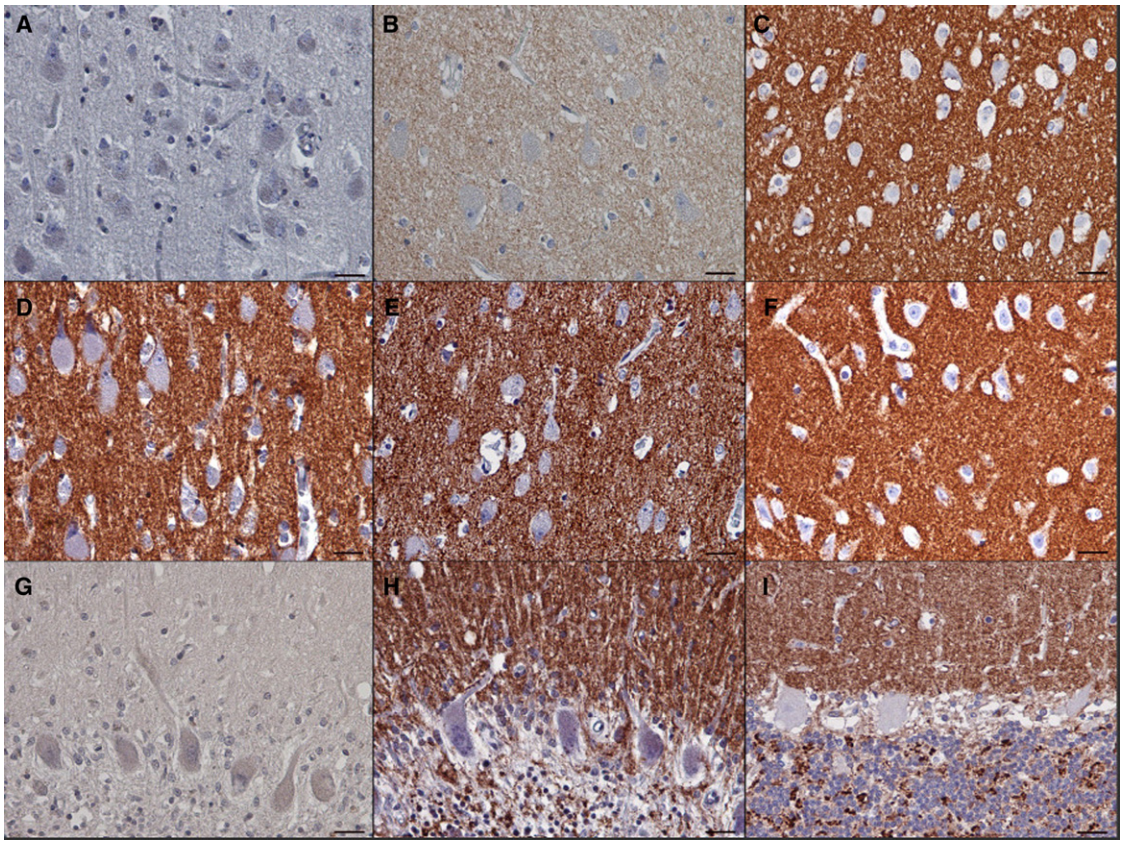


Figure 5. Brain Immunohistochemistry
(A–C) Detail of the CSPα staining in neuropil in the cerebral cortex that is absent in the individual with mutation p.Leu115Arg (A), decreased in the individual with mutation p.Leu116del (B), and strong in the age-matched control (C). Note the prominent neuronal storage, 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.

Please cite this article in press as: Stránecký et al., Mutations in *ANTXR1* Cause GAPO Syndrome, The American Journal of Human Genetics (2013), http://dx.doi.org/10.1016/j.ajhg.2013.03.023

REPORT

Mutations in *ANTXR1* Cause GAPO Syndrome

Viktor Stránecký,[1,9] Alexander Hoischen,[2,9] Hana Hartmannová,[1,9] Maha S. Zaki,[3] Amit Chaudhary,[4] Enrique Zudaire,[4] Lenka Nosková,[1] Veronika Barešová,[1] Anna Přistoupilová,[1] Kateřina Hodaňová,[1] Jana Sovová,[1] Helena Hůlková,[1] Lenka Piherová,[1] Jayne Y. Hehir-Kwa,[2] Deepthi de Silva,[5] Manouri P. Senanayake,[6] Sameh Farrag,[7] Jiří Zeman,[7] Pavel Martásek,[7] Alice Baxová,[8] Hanan H. Afifi,[3] Brad St. Croix,[4] Han G. Brunner,[2] Samia Temtamy,[3] and Stanislav Kmoch[1,*]

The genetic cause of GAPO syndrome, a condition characterized by growth retardation, alopecia, pseudoanodontia, and progressive visual impairment, has not previously been identified. We studied four ethnically unrelated affected individuals and identified homozygous nonsense mutations (c.262C>T [p.Arg88*] and c.505C>T [p.Arg169*]) or splicing mutations (c.1435–12A>G [p.Gly479Phefs*119]) in *ANTXR1*, which encodes anthrax toxin receptor 1. The nonsense mutations predictably trigger nonsense-mediated mRNA decay, resulting in the loss of ANTXR1. The transcript with the splicing mutation theoretically encodes a truncated ANTXR1 containing a neopeptide composed of 118 unique amino acids in its C terminus. GAPO syndrome's major phenotypic features, which include dental abnormalities and the accumulation of extracellular matrix, recapitulate those found in *Antxr1*-mutant mice and point toward an underlying defect in extracellular-matrix regulation. Thus, we propose that mutations affecting ANTXR1 function are responsible for this disease's characteristic generalized defect in extracellular-matrix homeostasis.

GAPO syndrome (MIM 230740) is the acronym for a complex disorder characterized by growth retardation, alopecia, pseudoanodontia, and, in many but not all cases, progressive optic atrophy.[1] Although variations of these phenotypes have been associated with other syndromes, their combination is unique to individuals with GAPO syndrome, and more than 30 cases of various ethnic origins have been described.[2–8] Most of the cases are from consanguineous parents, and inheritance patterns within these families have suggested that the disease is inherited as an autosomal-recessive trait. Altough affected individuals have no readily identifiable biochemical or endocrine abnormalities, histopathologic studies have revealed an abnormal accumulation of extracellular material,[9,10] and clinical presentation has shown predominant involvement of connective tissue (fibroblasts, chondrocytes, and osteoblasts), venous malformations, and heart, lung, and ocular abnormalities. These clinicopathologic changes point to a generalized defect in extracellular-matrix homeostasis. However, prior research has been unable to identify the genetic roots or reveal the basic molecular mechanisms responsible for GAPO syndrome.

To identify the genetic defect in GAPO syndrome, we performed genomic analysis in four unrelated and ethnically diverse families (Figure 1). The study was approved by institutional review boards, and the investigations were performed according to the Declaration of Helsinki principles. Adults provided informed consent, and the affected child provided assent with parental consent. Consents to publish clinical photographs in scientific journals were also obtained.

We analyzed a previously reported Czech family trio[11] (called CZE1) with one affected child (II-1 [Figures 1A and 1B]) who died from a heart attack at the age of 19 years, a previously reported Egyptian family (EGY1) with one affected child (V-3 [Figures 1C and 1D])[10] who died from renal failure at the age of 12 years, and two recently identified cases in families from Egypt (EGY2) (VI-4 [Figures 1E and 1F]) and Sri Lanka (SRI1) (III-1 [Figures 1G and 1H]). All four cases demonstrated the major clinical hallmarks of GAPO syndrome as summarized in Table 1.

Participants provided venous blood samples, and genomic DNA was isolated with standard technology. We first genotyped genomic DNA from all three Czech family members (i.e., both unaffected parents and the affected child) by using Affymetrix GeneChip Mapping 6.0 Arrays. We used data from both SNP and copy-number probes and identified in all three individuals copy-number alterations relative to a built-in reference as previously described.[12] In our analysis of the Czech proband, II-1, no rare or potentially disease-causing deletion or amplification larger than 10 Kb was revealed to be compatible with an expected autosomal-recessive inheritance model. Because ~0.85% of the proband genome was found to be autozygous, we estimated that the parents might be fifth-degree relatives. Accordingly, when we used the Affymetrix Genotyping

[1]Institute for Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague, 120 00 Prague 2, Czech Republic; [2]Department of Human Genetics, Nijmegen Center for Molecular Life Sciences, Institute for Genetic and Metabolic Disease, Radboud University Nijmegen Medical Center, 6500 HC Nijmegen, the Netherlands; [3]Clinical Genetics Department, National Research Centre, Cairo 12311, Egypt; [4]Tumor Angiogenesis Section, Frederick National Laboratory for Cancer Research, Frederick, MD 21702-1201, USA; [5]Faculty of Medicine, University of Kelaniya, Ragama 11010, Sri Lanka; [6]Department of Paediatrics, Faculty of Medicine, University of Colombo, Colombo 08, Sri Lanka; [7]Department of Pediatrics, First Faculty of Medicine, Charles University in Prague, 120 00 Prague 2, Czech Republic; [8]Department of Medical Genetics, First Faculty of Medicine, Charles University in Prague, 120 00 Prague 2, Czech Republic
[9]These authors contributed equally to this work
*Correspondence: skmoch@lf1.cuni.cz
http://dx.doi.org/10.1016/j.ajhg.2013.03.023.

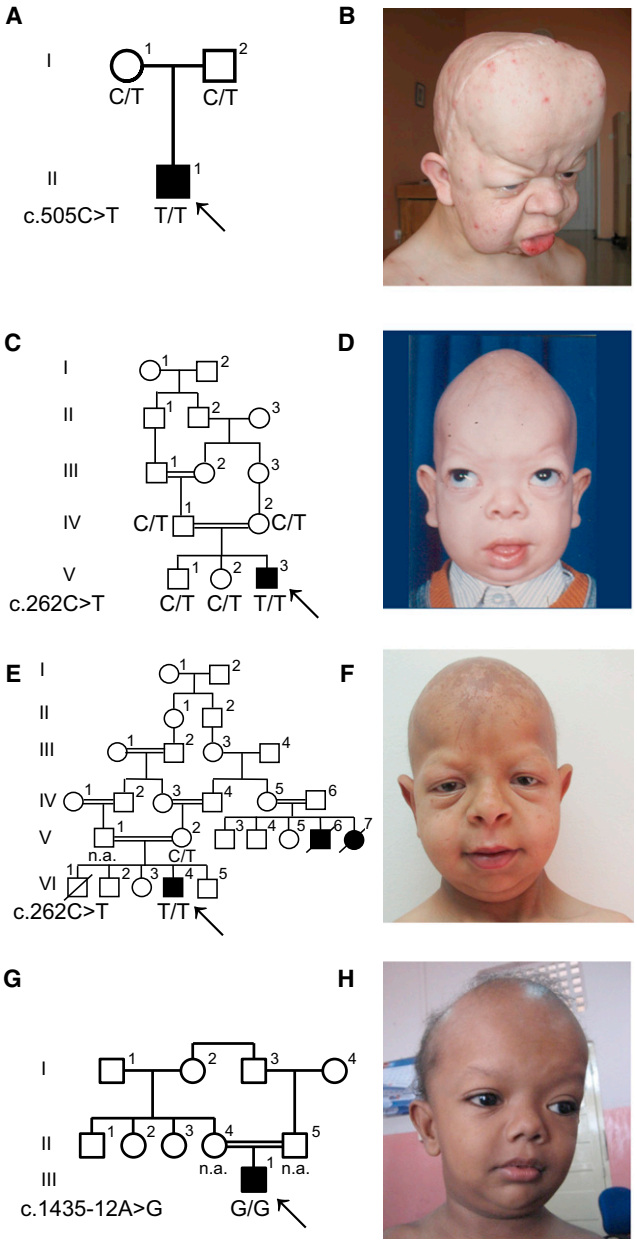


Figure 1. Family Pedigrees, Segregation of the *ANTXR1* Mutations, and Facial Appearance of the Probands with GAPO Syndrome

(A) Pedigree of the Czech family, CZE1.
(B) Facial appearance of the Czech proband, II-1, after neurosurgery due to a posthemorrhagic malatic lesion in the frontal cortex.
(C) Pedigree of Egyptian family EGY1.
(D) Facial appearance of Egyptian proband V-3.
(E) Pedigree of Egyptian family EGY2. Note extensive consanguinity and similarly affected relatives.
(F) Facial appearance of Egyptian proband VI-3.
(G) Pedigree of the Sri Lankan family, SRI1.
(H) Facial appearance of the Sri Lankan proband, III-1.
Black symbols denote affected individuals, and open symbols denote unaffected parents and siblings. "NA" indicates that DNA was not available for the investigation. The diagonal slash denotes deceased individuals. The arrows indicate the probands displayed in the corresponding pictures.

Console Software version 4.1 algorithm to compare values from the user's sample set and SNP-specific distributions derived from a reference set of 200 ethnically diverse individuals,[12] we identified in the proband sample two extended autozygous regions on chromosome 2 (chr2: 60,738,227–74,103,186) and chromosome 4 (chr4: 20,458,688–32,646,855), and they contained 114 and 29 genes, respectively (Figure S1A, available online).

To directly identify potential disease-causing mutations, we sequenced and analyzed the exomes of all three individuals from the Czech family as previously described.[13] In the resulting data set, we searched for variants that were either private or present in the internal exome database or in the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project Exome Variant Server with allele frequencies lower than 0.1% and whose genotypes were compatible with an expected autosomal-recessive model of the disease. This analysis revealed 121 candidate variants in proband II-1. However, the only relevant variant compatible with a recessive disorder was a homozygous nonsense mutation (c.505C>T [p.Arg169*]) in *ANTXR1*, encoding anthrax toxin receptor 1, also known as tumor endothelial marker 8 (TEM8) (RefSeq accession number NM_032208.2) (Table S1 and S2). This mutation is localized in one of the extended homozygous regions identified in the proband's genome and was inherited from both parents, who are heterozygous carriers. We confirmed the presence of the c.505C>T mutation in the parents and in the proband's genomic DNA by Sanger sequencing (Figure S2A). The identified mutation was not reported in dbSNP, 1000 Genomes, the Exome Variant Server, or an internal exome database (>120 exomes). It was absent in an additional 200 control samples that we analyzed with an XhoI-based restriction assay of the corresponding PCR-amplified genomic DNA fragments. To confirm the recurrence of *ANTXR1* mutations in another affected family, we sequenced *ANTXR1* genomic DNA of the proband (VI-4) from family EGY2 and identified a homozygous nonsense mutation (c.262C>T [p.Arg88*]) (Figure S2B), which was also localized in an apparently autozygous region (according to the homozygous genotypes for common SNPs present across the analyzed *ANTXR1* genomic sequence and quantitative-PCR results verifying the presence of both mutated alleles; Figure S3) and was not reported in dbSNP, 1000 Genomes, the Exome Variant Server, an internal exome database, or 200 control samples analyzed with a BsaJI-based restriction assay performed on PCR-amplified genomic DNA fragments.

In parallel, DNA samples from two other cases (V-3 from family EGY1 and III-1 from family SRI1) were independently analyzed by exome sequencing performed essentially as above and as described previously.[14–17] As in a previous study,[14] autozygous regions were identified directly from the exome data of both samples. Strikingly, this resulted in two large overlapping regions of homozygosity on chromosome 2; the total overlap was a ~27 Mb region (chromosome 2: 43–70 Mb) containing 144 genes

Please cite this article in press as: Stránecký et al., Mutations in *ANTXR1* Cause GAPO Syndrome, The American Journal of Human Genetics (2013), http://dx.doi.org/10.1016/j.ajhg.2013.03.023

Table 1. Main Clinical Findings in the Four Studied Individuals with GAPO Syndrome

	Cases			
Features	**V-3 from EGY1**	**VI-4 from EGY2**	**II-1 from CZE1**	**III-1 from SRI1**
General				
Age at evaluation (years)	3	10	18	4
Gender	male	male	male	male
Parental consanguinity (first or second cousins)	+	+	−	+
Family history of similarly affected case	−	+	−	−
Height	–2 SDs	–3.7 SDs	–4 SDs	–4 SDs
Head circumference	–2 SDs	+2 SDs	+1 SD	–2 SDs
Bone age	delayed	mild delay	normal for age	delayed
Craniofacial				
Plagiocephaly	+	+	+	−
Frontal bossing	+	+	+	+
Broad forehead	+	+	+	+
Enlarged persistent anterior fontanel	+	+	+	+
Widely spaced eyes	+	+	+	+
Epicantus	+	+	−	+
Depressed nasal bridge	+	+	+	+
Short nose	+	+	+	+
Long philtrum	+	+	+	+
Thick and anteverted nares	+	+	+	+
Thick lower lip	+	+	+	+
Micrognathia	+	+	+	+
Pseudoanodontia	+	+	+	+
Skin and Hair				
Sparse scalp hair (alopecia)	+	+	+	+
Scalp pigmented with scars and papules	−	+	−	−
Sparse eyebrows and eyelashes	+	+	+	+
Ophthalmologic				
Megalocornea	+	+	+	NR
Nystagmus	+	−	−	+
Esotropia	+	−	+	−
Shallow anterior chamber	−	+	+	NR
Bilateral engorged tortuous retinal vessels	−	+	+	+
Bilateral optic atrophy	+	−	+	+
VEP (abnormal pattern)	+	−	+	NR
Other				
Umbilical hernia	+	+	+	−
Hyperextensible joints	+	−	−	+
Mild webbing between fingers	−	+	−	−

(*Continued on next page*)

Table 1. ***Continued***

Features	Cases			
	V-3 from EGY1	VI-4 from EGY2	II-1 from CZE1	III-1 from SRI1
Facial nerve palsy	+	–	–	–
MRI brain changes	bilateral high signal of deep white matter at deep parietal and occipital region and around the optic nerve	ND	ND	ND

Abbreviations are as follows: VEP, visual-evoked potential; NR, not recorded; and ND, not done.

(Figures S1B and S1C). The only gene harboring private or rare homozygous coding or splice-site variants within this overlapping region was *ANTXR1*. For the EGY1 case (V-3), we identified the same nonsense mutation (c.262C>T [p.Arg88*]) as for the EGY2 case (VI-4) (Figure S2C), whereas in the SRI1 case (III-1), we identified a substitution, c.1435–12A>G (Figure S2D). This latter variant is predicted by ESE finder[18,19] to generate an alternative strong splice acceptor site 11 nucleotides upstream of the last exon (Figure S4), and this would theoretically result in a frameshift of the complete reading frame of the last exon and proteosynthesis of a truncated ANTXR1 containing a neopeptide composed of 118 unique amino acids in its C terminus (p.Gly479Phefs*119) (Figure S5).

Two of the affected probands, V-3 and VI-4, from Egyptian families EGY1 and EGY2, respectively, harbor an identical c.262C>T [p.Arg88*] mutation. To determine whether these two probands might be distantly related and share a mutated chromosomal segment from a common ancestor, we examined *ANTXR1* intragenic SNP haplotypes obtained by exome sequencing (for V-3 from EGY1) and Sanger sequencing (for VI-4 from EGY2). This revealed that the c.262C>T mutations are present on two distinct haplotypes, indicating that these mutations most likely developed independently or that these families share a very old ancestral allele (Figure S6). The c.262C nucleotide belongs to a CpG doublet, making deamination of the cytosine a possible explanation for the recurrence of the mutation.

ANTXR1, also called TEM8, was initially identified as one of the tumor endothelial markers (TEMs) that displays elevated protein levels during tumor angiogenesis.[20,21] Soon after its discovery, it was independently identified as the anthrax toxin receptor (ATR).[22] Several variants of human ANTXR1 have been proposed to exist on the basis of the identification of rare alternative mRNA splice variants (Figure 2A). The biosynthesis of all known variants is driven by a common signal peptide (amino acids 1–27) and proceeds by cotranslational translocation in the endoplasmic reticulum. The full-length *ANTXR1* variant v1, (RefSeq NM_032208.2) is by far the most prevalent transcript found in databases of cDNA and expressed sequence tags. It encodes a single-pass type 1 transmembrane glycoprotein that has a molecular weight of approximately 85 kDa and that is composed of a predicted N-terminal extracellular sequence (amino acids 28–322) containing a von Willebrand type A domain (amino acids 44–215), a transmembrane domain (amino acids 322–342), and large cytoplasmic domain (amino acids 343–564) (isoform 1). Variant v2 (RefSeq NM_053034.2) encodes protein isoform 2, which is structurally similar to variant 1 but contains a much shorter cytoplasmic domain (amino acids 343–368).[22] Variant v3 (RefSeq NM_018153.3) encodes protein isoform 3, which does not contain the transmembrane or cytoplasmic domains and is predicted to be secreted.[23] Two other transcript variants have recently been identified: v4 (GenBank accession number JX424838.1), potentially encoding membrane-bound protein isoform 4, which, compared to isoform 1, lacks 36 aa residues in its cytoplasmic domain, and v5 (GenBank JX424839.1), potentially encoding secreted protein isoform 5, which, compared to isoform 3, has an alternative C-terminal sequence.[24] Because detecting cDNA for alternative splice variants v2–v5 is difficult in that it requires as many as 60 cycles of nested PCR[24] and because similar conserved variants in other species have not yet been described, it is currently unclear whether they represent transcriptional noise caused by inappropriate splicing events and whether the encoded protein isoforms are produced at sufficient endogenous levels needed to impact biological function. However, the full-length ANTXR1 isoform 1 has been shown to promote interaction between cells and various components of the extracellular matrix,[25,26] link extracellular ligands to the actin cytoskeleton,[25,27] and regulate cell spreading.[28–30]

In three of the four GAPO cases, the identified mutations introduce premature stop codons in *ANTXR1* mRNA, and in the fourth case (III-1 in SRI1), the mutation most likely results in a loss-of-function allele. From cases II-1 (CZE1) and VI-4 (EGY2), we studied skin fibroblast cell lines harboring the *ANTXR1* mutations encoding p.Arg169* and p.Arg88*. To characterize the molecular consequences of the identified mutations on *ANTXR1* mRNA expression, splicing, and stability, we isolated total RNA from two cases and control skin fibroblasts and performed RT-PCR and quantitative-PCR analyses. In fibroblasts from affected individuals, we found a single PCR product comparable in size to a control specimen (Figure 2B). Sanger sequencing demonstrated that the obtained PCR products corresponded to cDNA of the major transcript variant v1 of *ANTXR1* and independently confirmed the presence of the premature-stop-codon-encoding mutations previously identified in corresponding genomic DNA in affected

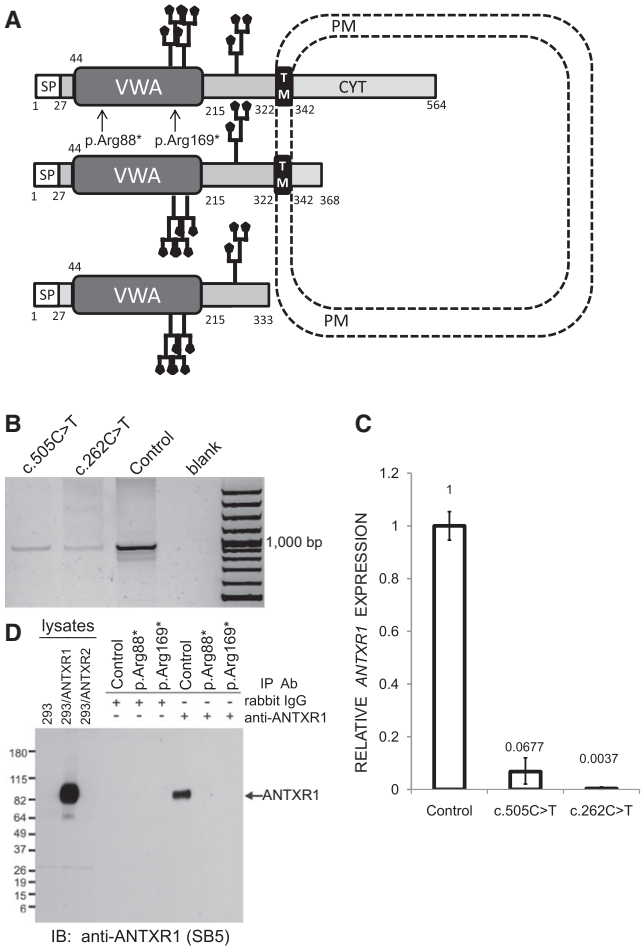


Figure 2. Effects of the Identified *ANTXR1* Mutations

(A) A schematic representation of ANTXR1 shows the protein structure, cellular topology, and location of the p.Arg88* and p.Arg169* substitutions. Two potential N-glycosylation sites are depicted. The numbers denote amino acid residues defining the boundaries of predicted ANTXR1 domains. Only ANTXR1 isoforms 1, 2, and 3 are depicted. Abbreviations are as follows: SP, signal peptide; VWA, Von Willebrand factor type A domain; TM, transmembrane domain; CYT, cytoplasmic domain; and PM, plasma membrane.

(B) *ANTXR1* cDNA analysis. Total RNA was isolated from pellets from a cultured skin fibroblast cell line with the use of TRIZOL solution (Invitrogen). RNA concentrations were determined spectrophotometrically at A260 nm by NanoDrop (NanoDrop Technologies), and RNA quality was verified with an Agilent 2100 bioanalyser, RNA Lab-on-a-Chip (Agilent Technologies). The first-strand cDNA synthesis was carried out with an oligo-dT primer and SuperScript III Reverse Transcriptase (Life Technologies). *ANTXR1* cDNA was PCR amplified from the synthesized first-strand cDNA with oligonucleotide primers designed to span and amplify all three *ANTXR1* variants in parallel (Table S2). Lanes 1 and 2 show reduced amounts of RT-PCR products from cases with p.Arg169* and p.Arg88* substitutions, and lane 3 shows the cDNA amount obtained under identical conditions from a control cell line, C. Lane 4 is a negative control. Lane 5 is a 100 bp DNA ladder.

(C) Relative expression levels of *ANTXR1* mRNA amounts normalized to glyceraldehydes-3-phosphate dehydrogenase (*GAPDH*) mRNA amounts in skin fibroblasts. Quantitative PCR was carried out on a StepOne Plus Real Time System (Applied Biosystems). The reactions were carried out in a 96-well plate in a 20 μl reaction volume containing 10 μl 2 × Maxima SYBR Green qPCR Master Mix (Thermo Scientific), 0.2 μM forward and reverse primer, and 5 ng cDNA. Data were analyzed by StepOne Software v.2.0. The comparative Ct (ΔΔ Ct) method was used for normalizing target-gene mRNA to *GAPDH* mRNA. The relative amounts of the *ANTXR1*cDNA were significantly reduced in cases compared to control samples. The means ± SD of three experiments performed in triplicate are shown.

(D) Immunoblot (IB) analysis of immunoprecipitated (IP) total-protein extracts showing absence of ANTXR1 in cultured skin fibroblasts from cases with p.Arg88* and p.Arg169* substitutions. Cultured cells were lysed in TNT lysis buffer (50 mM Tris [pH 7.5], 75 mM NaCl, and 1% Triton X-100 plus complete protease inhibitor cocktail [Roche]) and clarified by centrifugation. Protein extracts were quantified with a BCA assay (Pierce), normalized, and immunoprecipitated with a rabbit monoclonal ANTXR1 antibody (clone 37). This rabbit monoclonal antibody was produced as part of a collaboration between Epitomics and the National Cancer Institute and will be described in more detail elsewhere. After immunoprecipitation using protein A agarose, protein extracts were separated by SDS-PAGE, transferred to a PDVF membrane (Millipore), and detected by immunoblotting with SB5 mouse monoclonal ANTXR1 antibodies followed by HRP-conjugated anti-mouse or anti-rabbit F(ab')2 secondary antibodies (Jackson). Chemiluminescence was visualized with the ECL plus system (Amersham) according to the supplier's instructions. Lysates of 293 cells stably transfected with an empty vector (293), 293 cells stably expressing human ANTXR1 (293/ANTXR1), and 293 cells stably expressing ANTXR2 (293/ANTXR2) were used as negative, positive, and specificity controls, respectively. Equal protein amounts in the original lysates were immunoprecipitated in parallel with either control rabbit nonspecific IgG antibodies or rabbit ANTXR1 antibodies.

samples (data not shown). The amounts of *ANTXR1* cDNA were significantly reduced in GAPO compared to control samples, most likely through partial degradation of mutated transcripts via nonsense-mediated mRNA decay (NMD) (Figure 2C). Next, we examined whether any mRNAs might have escaped NMD and translated into any form of ANTXR1. We immunoprecipitated ANTXR1 from skin fibroblast lysates and culture media by using a rabbit monoclonal ANTXR1 antibody recognizing an epitope between amino acid residues 28 and 81 of ANTXR1 (A.C. and B.S.C., unpublished data) and then performed immunoblotting for ANTXR1. For immunodetection of ANTXR1, we used either SB5 mouse monoclonal antibodies recognizing the extracellular region of ANTXR1 (amino acid residues 82–145)[25,27,31] or the rabbit monoclonal antibody. This analysis revealed the presence of the full-length ANTXR isoform 1 in control fibroblasts but a complete loss of ANTXR1 in both affected fibroblasts (Figure 2D and data not shown). Specific chemiluminiscence signals indicating the presence of the secreted ANTXR1 isoform 3 were not detected in immunoprecipitates obtained from culture media of the control or GAPO cases (not shown). In parallel, we studied and were unable to detect any ANTXR1 in paraformaldehyde-fixed cultured control and GAPO fibroblasts by immunofluorescence analysis. This is in accordance with the results of our immunoblot analysis and the available RNaseq data, which both suggest that ANTXR1 levels and corresponding mRNA amounts are very low in these cells. Next, we used phalloidin staining to evaluate the actin cytoskeleton

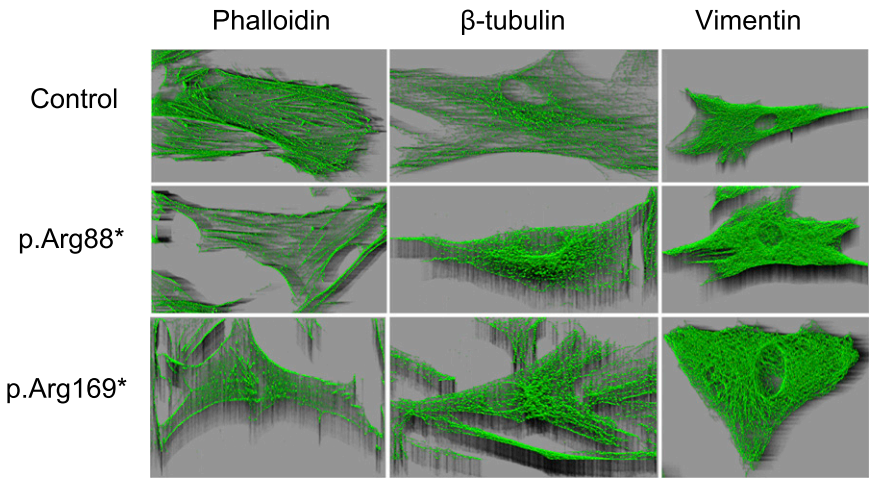


Figure 3. Immunofluorescence Analysis of Cultured Skin Fibroblasts

The cells were grown on 70 mm^2 glass chamber slides (Lab-Tek, Nalge Nunc International) for 48 hr. Then the cells were fixed with 4% paraformaldehyde in PBS, permeabilized in 0.1% TRITON, washed, blocked with 5% BSA in PBS, and incubated in a humidified chamber at 4°C overnight with mouse monoclonal β-tubulin antibody (Sigma) or Vimentin (V9) antibody (BioGenex). For fluorescence detection, species-specific secondary antibodies Alexa Fluor 488 or 555 (Molecular Probes, Invitrogen) were used. For actin staining, Alexa Fluor 488 Phalloidin (Molecular Probes, Invitrogen) was used. Slides were mounted in fluorescence mounting medium Immu-Mount (Shandon Lipshaw) and analyzed by confocal microscopy. *XYZ* images were sampled according to Nyquist criterion with the Nikon TE2000E C1si laser-scanning confocal microscope with a Nikon PlanApo objective (60×, numerical aperture 1.40) and 488 and 543 laser lines. Images were restored with a classic maximum-likelihood restoration algorithm in the Huygens Professional Software (SVI, Hilversum, the Netherlands). Phalloidin staining demonstrated remarkable alterations in the actin cytoskeletal network in cell lines from GAPO cases with p.Arg169* and p.Arg88* substitutions. No abnormalities in microtubules or intermediate filaments were detected with β-tubulin or vimentin staining, respectively.

because ANTXR1 has been previously shown to interact with actin. These studies revealed a striking reorganization of the actin cytoskeletal microfilaments specifically in the GAPO fibroblasts, but β-tubulin and vimentin staining of microtubules and intermediate filaments, respectively, were unaltered (Figure 3). This suggests that ANTXR1, a molecule mediating the coupling of extracellular ligands to the actin cytoskeleton, is crucial for actin assembly and that disruption of the actin network might be the major pathogenetic event leading to altered cell-adhesion properties and progressive extracellular-matrix buildup observed in individuals with GAPO syndrome.

Unfortunately, we did not have the opportunity to study mRNA processing, protein production, or the actin network in the fourth GAPO case with *ANTXR1* mutation c.1435–12A>G. In this case, the mutation theoretically affects splicing of *ANTXR1* mRNA and potentially encodes ANTXR1isoforms 1 and 2 with altered C-terminal cytoplasmic tails; if produced, these isoforms could potentially retain some biological functions. In addition to this, the mutation should not theoretically affect proteosynthesis of the secreted ANTXR1 isoform 3. This could help to explain the evidently milder clinical presentation of this case compared to the other three cases with nonsense mutations.

ANTXR1 is most highly produced in tumor endothelial cells and other tumor stromal cells, which might include both pericytes and fibroblasts.[25,32,33] *Antxr1*-mutant mice with targeted deletion of exon 13—encoding the transmembrane domain—are viable and progressively develop misaligned incisor teeth, and female mice are infertile.[34] In another mutant mouse model, complete *Antxr1* disruption due to removal of exon 1—encoding the start codon and signal peptide—leads to a moderate excess of extracellular matrix in many tissues, including the ovaries, uterus, skin, hair follicles, cranial sutures of the skull, and the periodontal ligament of the incisors, resulting in dental dysplasia.[31] These features are consistent with the clinical presentation of individuals afflicted with GAPO syndrome. It is also notable that some of the individuals with GAPO syndrome have infantile hemangiomas,[35] which have been associated with germline heterozygosity for missense mutations in *ANTXR1*[36] and dysfunction of the complex formed by VEGFR2, β1 integrin, and ANTXR1.[37] Generalized extracellular-matrix-homeostasis defects observed in GAPO-syndrome-affected individuals with *ANTXR1* mutations are similar to those found in individuals with juvenile hyaline fibromatosis (MIM 228600) and infantile systemic hyalinosis (MIM 236490), which are allelic disorders caused by mutations in anthrax toxin receptor 2 (*ANTXR2*), also known as capillary morphogenesis gene 2 (*CMG2*), an *ANTXR1* homolog.[38]

We conclude that our data, together with recapitulation of many of the phenotypic features characteristic of GAPO syndrome in *Antxr1*-mutant mice and individuals with *ANTXR2* mutations, strongly suggest involvement of *ANTXR1* mutations in the generalized extracellular-matrix-homeostasis defect characteristic of this disease. From a clinical perspective, our finding provides essential information for DNA testing in other families. In addition, autopsy tissues and cultured skin fibroblasts from these affected individuals represent an interesting cellular model and potential resource for detailed studies on the pathogenesis of individual clinical symptoms present in GAPO syndrome and studies focused on ANTXR1 functions in general.

Supplemental Data

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

Acknowledgments

This work was supported by Charles University institutional programs PRVOUK-P24/LF1/3, UNCE 204011, and SVV2012/2645; by the Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University (CZ.1.05/1.1.00/02.0109); by the European Regional Development Fund; by grant NT13116-4/2012 from the Ministry of Education and Ministry of Health of the Czech Republic; and by the intramural research program of the National Cancer Institute, National Institutes of Health, US Department of Health and Social Services. A.H. was supported by the Netherlands Organization for Health Research and Development (ZonMW 916-12-095). We thank clinical colleagues and families who contributed samples used in this study, as well as members of the Genomic Disorders Group Nijmegen for their technical support. A.C. and B.S.C. are coinventors of filed patents related to the development of ANTXR1 antibodies for cancer therapy.

Received: November 9, 2012
Revised: January 30, 2013
Accepted: March 28, 2013
Published: April 18, 2013

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://www.1000genomes.org/
ESE finder, http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
GenBank, http://www.ncbi.nlm.nih.gov/genbank/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq

References

1. Tipton, R.E., and Gorlin, R.J. (1984). Growth retardation, alopecia, pseudo-anodontia, and optic atrophy—the GAPO syndrome: report of a patient and review of the literature. Am. J. Med. Genet. *19*, 209–216.
2. Goloni-Bertollo, E.M., Ruiz, M.T., Goloni, C.B., Muniz, M.P., Valério, N.I., and Pavarino-Bertelli, E.C. (2008). GAPO syndrome: three new Brazilian cases, additional osseous manifestations, and review of the literature. Am. J. Med. Genet. A. *146A*, 1523–1529.
3. Sinha, R., Trikha, A., Laha, A., Raviraj, R., and Kumar, R. (2011). Anesthetic management of a patient with GAPO syndrome for glaucoma surgery. Paediatr. Anaesth. *21*, 910–912.
4. Nanda, A., Al-Ateeqi, W.A., Al-Khawari, M.A., Alsaleh, Q.A., and Anim, J.T. (2010). GAPO syndrome: a report of two siblings and a review of literature. Pediatr. Dermatol. *27*, 156–161.
5. Lei, S., Iyengar, S., Shan, L., Cherwek, D.H., Murthy, S., and Wong, A.M. (2010). GAPO syndrome: a case associated with bilateral interstitial keratitis and hypothyroidism. Clin. Dysmorphol. *19*, 79–81.
6. Castrillon-Oberndorfer, G., Seeberger, R., Bacon, C., Engel, M., Ebinger, F., and Thiele, O.C. (2010). GAPO syndrome associated with craniofacial vascular malformation. Am. J. Med. Genet. A. *152A*, 225–227.
7. Kocabay, G., and Mert, M. (2009). GAPO syndrome associated with dilated cardiomyopathy: an unreported association. Am. J. Med. Genet. A. *149A*, 415–416.
8. Demirgüneş, E.F., Ersoy-Evans, S., and Karaduman, A. (2009). GAPO syndrome with the novel features of pulmonary hypertension, ankyloglossia, and prognathism. Am. J. Med. Genet. A. *149A*, 802–805.
9. Wajntal, A., Koiffmann, C.P., Mendonça, B.B., Epps-Quaglia, D., Sotto, M.N., Rati, P.B., and Opitz, J.M. (1990). GAPO syndrome (McKusick 23074)—a connective tissue disorder: report on two affected sibs and on the pathologic findings in the older. Am. J. Med. Genet. *37*, 213–223.
10. Meguid, N.A., Afifi, H.H., Ramzy, M.I., Hindawy, A., and Temtamy, S.A. (1997). GAPO syndrome: first Egyptian case with ultrastructural changes in the gingiva. Clin. Genet. *52*, 110–115.
11. Baxova, A., Kozlowski, K., Obersztyn, E., and Zeman, J. (1997). GAPO syndrome (Radiographic clues to early diagnosis). Radiol. Med. (Torino) *93*, 289–291.
12. van de Steeg, E., Stránecký, V., Hartmannová, H., Nosková, L., Hřebíček, M., Wagenaar, E., van Esch, A., de Waart, D.R., Oude Elferink, R.P., Kenworthy, K.E., et al. (2012). Complete OATP1B1 and OATP1B3 deficiency causes human Rotor syndrome by interrupting conjugated bilirubin reuptake into the liver. J. Clin. Invest. *122*, 519–528.
13. Nosková, L., Stránecký, V., Hartmannová, H., Přistoupilová, A., Barešová, V., Ivánek, R., Hůlková, H., Jahnová, H., van der Zee, J., Staropoli, J.F., et al. (2011). Mutations in DNAJC5, encoding cysteine-string protein alpha, cause autosomal-dominant adult-onset neuronal ceroid lipofuscinosis. Am. J. Hum. Genet. *89*, 241–252.
14. Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, H.J., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., et al. (2011). Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. Am. J. Hum. Genet. *88*, 362–371.
15. Gilissen, C., Arts, H.H., Hoischen, A., Spruijt, L., Mans, D.A., Arts, P., van Lier, B., Steehouwer, M., van Reeuwijk, J., Kant, S.G., et al. (2010). Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. Am. J. Hum. Genet. *87*, 418–423.
16. Hoischen, A., van Bon, B.W., Gilissen, C., Arts, P., van Lier, B., Steehouwer, M., de Vries, P., de Reuver, R., Wieskamp, N., Mortier, G., et al. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat. Genet. *42*, 483–485.
17. Hoischen, A., van Bon, B.W., Rodríguez-Santiago, B., Gilissen, C., Vissers, L.E., de Vries, P., Janssen, I., van Lier, B., Hastings, R., Smithson, S.F., et al. (2011). De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. Nat. Genet. *43*, 729–731.
18. Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., and Krainer, A.R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res. *31*, 3568–3571.
19. Smith, P.J., Zhang, C., Wang, J., Chew, S.L., Zhang, M.Q., and Krainer, A.R. (2006). An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum. Mol. Genet. *15*, 2490–2508.
20. St Croix, B., Rago, C., Velculescu, V., Traverso, G., Romans, K.E., Montgomery, E., Lal, A., Riggins, G.J., Lengauer, C., Vogelstein, B., and Kinzler, K.W. (2000). Genes expressed in human tumor endothelium. Science *289*, 1197–1202.

21. Carson-Walter, E.B., Watkins, D.N., Nanda, A., Vogelstein, B., Kinzler, K.W., and St Croix, B. (2001). Cell surface tumor endothelial markers are conserved in mice and humans. Cancer Res. *61*, 6649–6655.
22. Bradley, K.A., Mogridge, J., Mourez, M., Collier, R.J., and Young, J.A. (2001). Identification of the cellular receptor for anthrax toxin. Nature *414*, 225–229.
23. Liu, S., and Leppla, S.H. (2003). Cell surface tumor endothelium marker 8 cytoplasmic tail-independent anthrax toxin binding, proteolytic processing, oligomer formation, and internalization. J. Biol. Chem. *278*, 5227–5234.
24. Vargas, M., Karamsetty, R., Leppla, S.H., and Chaudry, G.J. (2012). Broad expression analysis of human ANTXR1/TEM8 transcripts reveals differential expression and novel splice variants. PLoS ONE *7*, e43174.
25. Nanda, A., Carson-Walter, E.B., Seaman, S., Barber, T.D., Stampfl, J., Singh, S., Vogelstein, B., Kinzler, K.W., and St Croix, B. (2004). TEM8 interacts with the cleaved C5 domain of collagen alpha 3(VI). Cancer Res. *64*, 817–820.
26. Hotchkiss, K.A., Basile, C.M., Spring, S.C., Bonuccelli, G., Lisanti, M.P., and Terman, B.I. (2005). TEM8 expression stimulates endothelial cell adhesion and migration by regulating cell-matrix interactions on collagen. Exp. Cell Res. *305*, 133–144.
27. Yang, M.Y., Chaudhary, A., Seaman, S., Dunty, J., Stevens, J., Elzarrad, M.K., Frankel, A.E., and St Croix, B. (2011). The cell surface structure of tumor endothelial marker 8 (TEM8) is regulated by the actin cytoskeleton. Biochim. Biophys. Acta *1813*, 39–49.
28. Werner, E., Kowalczyk, A.P., and Faundez, V. (2006). Anthrax toxin receptor 1/tumor endothelium marker 8 mediates cell spreading by coupling extracellular ligands to the actin cytoskeleton. J. Biol. Chem. *281*, 23227–23236.
29. Gu, J., Faundez, V., and Werner, E. (2010). Endosomal recycling regulates Anthrax Toxin Receptor 1/Tumor Endothelial Marker 8-dependent cell spreading. Exp. Cell Res. *316*, 1946–1957.
30. Verma, K., Gu, J., and Werner, E. (2011). Tumor endothelial marker 8 amplifies canonical Wnt signaling in blood vessels. PLoS ONE *6*, e22334.
31. Cullen, M., Seaman, S., Chaudhary, A., Yang, M.Y., Hilton, M.B., Logsdon, D., Haines, D.C., Tessarollo, L., and St Croix, B. (2009). Host-derived tumor endothelial marker 8 promotes the growth of melanoma. Cancer Res. *69*, 6021–6026.
32. Davies, G., Rmali, K.A., Watkins, G., Mansel, R.E., Mason, M.D., and Jiang, W.G. (2006). Elevated levels of tumour endothelial marker-8 in human breast cancer and its clinical significance. Int. J. Oncol. *29*, 1311–1317.
33. Chaudhary, A., and St Croix, B. (2012). Selective blockade of tumor angiogenesis. Cell Cycle *11*, 2253–2259.
34. Liu, S., Crown, D., Miller-Randolph, S., Moayeri, M., Wang, H., Hu, H., Morley, T., and Leppla, S.H. (2009). Capillary morphogenesis protein-2 is the major receptor mediating lethality of anthrax toxin in vivo. Proc. Natl. Acad. Sci. USA *106*, 12424–12429.
35. Goucha, S., Fazaa, B., Ezzine, N., Jaber, K., Elandaloussi, H., Abid, R., and Kamoun, M.R. (2000). [GAPO syndrome]. Ann. Dermatol. Venereol. *127*, 501–504.
36. Jinnin, M., Medici, D., Park, L., Limaye, N., Liu, Y., Boscolo, E., Bischoff, J., Vikkula, M., Boye, E., and Olsen, B.R. (2008). Suppressed NFAT-dependent VEGFR1 expression and constitutive VEGFR2 signaling in infantile hemangioma. Nat. Med. *14*, 1236–1246.
37. Nicolae, C., and Olsen, B.R. (2010). Unexpected matrix diseases and novel therapeutic strategies. Cell Tissue Res. *339*, 155–165.
38. Hanks, S., Adams, S., Douglas, J., Arbour, L., Atherton, D.J., Balci, S., Bode, H., Campbell, M.E., Feingold, M., Keser, G., et al. (2003). Mutations in the gene encoding capillary morphogenesis protein 2 cause juvenile hyaline fibromatosis and infantile systemic hyalinosis. Am. J. Hum. Genet. *73*, 791–800.

Isolated X-Linked Hypertrophic Cardiomyopathy Caused by a Novel Mutation of the Four-and-a-Half LIM Domain 1 Gene

Hana Hartmannova, Milos Kubanek, Marek Sramko, Lenka Piherova, Lenka Noskova, Katerina Hodanova, Viktor Stranecky, Anna Pristoupilova, Jana Sovova, Tomas Marek, Jana Maluskova, Petr Ridzon, Josef Kautzner, Helena Hulkova and Stanislav Kmoch

Circ Cardiovasc Genet. published online October 10, 2013;
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1942-325X. Online ISSN: 1942-3268

DOI: 10.1161/CIRCGENETICS.113.000245

Isolated X-Linked Hypertrophic Cardiomyopathy Caused by a Novel Mutation of the Four-and-a-Half LIM Domain 1 Gene

Running title: *Hartmannova et al.; Isolated HCM due to FHL1 mutation*

Hana Hartmannova, MSc[1*]; Milos Kubanek, MD, PhD[2*]; Marek Sramko, MD[2]; Lenka Piherova, MSc[1]; Lenka Noskova MSc[1]; Katerina Hodanova, MSc, PhD[1]; Viktor Stranecky, MSc[1]; Anna Pristoupilova, MSc[1]; Jana Sovova, BS[1]; Tomas Marek, MD, PhD[2]; Jana Maluskova, MD[3]; Petr Ridzon, MD[4]; Josef Kautzner, MD, PhD[2]; Helena Hulkova, MD, PhD[1]; Stanislav Kmoch, MSc, PhD[1]

[1]Institute of Inherited Metabolic Disorders, Charles University in Prague – 1st Faculty of Medicine; [2]Department of Cardiology, Institute for Clinical & Experimental Medicine; [3]Department of Pathology, Institute for Clinical & Experimental Medicine; [4]Department of Neurology, Thomayer´s Hospital, Prague, Czech Republic

*contributed equally

Correspondence:

Milos Kubanek, MD, PhD

Department of Cardiology

Institute for Clinical and Experimental Medicine

Videnska 1958/9, Prague

Czech Republic

Tel: +420261365047

Fax: +420261362989.

E-mail: milos.kubanek@ikem.cz

Journal Subject Codes: [110] Congestive, Myocardial biology:[105] Contractile function

12

DOI: 10.1161/CIRCGENETICS.113.000245

Abstract:

Background - Hypertrophic cardiomyopathy (HCM) with severe left ventricular diastolic dysfunction (LVDD) has been associated with marked exercise intolerance and poor prognosis. However, molecular etiology of this phenotype remains unexplained in a large proportion of cases.

Methods and Results - We performed whole exome sequencing as an initial genetic test in a large Czech family with three males affected by non-obstructive HCM with severe LVDD in end-stage disease. A novel frameshift mutation of four-and-a-half LIM domain 1 gene (*FHL1*) (c.599_600insT; p.F200fs32X) was detected in these individuals. The mutation does not affect transcription, splicing and stability of *FHL1* mRNA and results into production of truncated FHL1 protein, which is contrary to heart tissue homogenate not detectable in frozen tissue sections of myocardial biopsy of affected males. The identified mutation co-segregated also with abnormal electrocardiogram and with one case of apical HCM in heterozygous females. Although skeletal muscle involvement is a common finding in FHL1-related diseases, we could exclude myopathy in all mutation carriers.

Conclusions - We identified a novel *FHL1* mutation causing isolated HCM with X-chromosomal inheritance.

Key words: hypertrophic cardiomyopathy, diastolic function cardiomyopathy, diastolic heart failure, exome, gene mutation, four-and-a-half LIM domain 1 gene

DOI: 10.1161/CIRCGENETICS.113.000245

Introduction:

Severe left ventricular diastolic dysfunction (LVDD) in individuals with hypertrophic cardiomyopathy (HCM) has been associated with a poor exercise tolerance and prognosis both in adults and children with this disease [1-7]. Severe LVDD may occur either as an initial manifestation of HCM with limited left ventricular hypertrophy or it may represent advanced disease, an alternative to end-stage dilated HCM [6]. Recently, mutations of beta-myosin heavy chain, cardiac troponin I [7,8] and cardiac troponin T [9] have been identified as disease genes for HCM with severe LVDD.

We report on a large family with isolated X-linked HCM with severe LVDD co-segregating with a novel mutation of *FHL1* gene encoding four-and-a-half LIM domain protein 1 (FHL1). Over 30 mutations of *FHL1* have been associated with five different X-linked myopathies – reducing body myopathy [MIM 300718], scapuloperoneal myopathy [MIM 300695], X-linked myopathy with postural muscle atrophy [MIM 300696], rigid spine syndrome and Emery-Dreifuss muscular dystrophy [MIM 310300] with a variable cardiac involvement presenting either as dilated or hypertrophic cardiomyopathy [10-14]. In addition, a recent report described two novel variants of *FHL1* causing isolated HCM [15]. This report extends the spectrum of FHL1-related diseases by description of *FHL1* as a causal gene for isolated HCM with severe LVDD in advanced disease.

Methods:

1. Study population

The index patient was identified in a large Czech family with three males affected by HCM who developed severe LVDD and advanced heart failure in two cases. A detailed family history was obtained, and an extended pedigree was constructed (Figure 1) according to the guidelines [16].

Further investigations were approved by participating center's Institutional Review Boards, and were conducted according to the Declaration of Helsinki principles. Written, informed consent was obtained from all subjects.

2. Cardiologic screening

We examined 20 members of the above family. Medical records were acquired in the proband´s uncle who was treated in our institution and died after heart transplantation. The cardiologic screening included physical examination, electrocardiography and echocardiography. Echocardiography was performed by a single operator (T.M.) in accordance with guidelines of the American Society of Echocardiography [17,18]. M-mode, 2D images, conventional and tissue Doppler recordings were obtained using a Vivid 7 (GE Healthcare, Chalfont St Giles, UK). Left ventricular ejection fraction was assessed using Simpson's biplane method. Mitral inflow pattern was classified as restrictive in the presence of E-wave deceleration time < 120 ms or a ratio of early transmitral flow velocity to atrial flow velocity ≥ 2 associated with an E-wave deceleration time ≤150 ms [19].

The proband and his brother underwent cardiovascular magnetic resonance (Siemens Trio scanner, Siemens Medical Solutions, Erlangen, Germany). Ventricular volumes, ventricular mass and ejection fractions were obtained from standard cine images in short-axis (repetition time/echo time 65/1.2 ms; slice thickness of 8 mm without inter-slice gap) using a dedicated software Segment 1.8 (http://segment.heiberg.se) [20]. Images for visualization of delayed contrast enhancement were acquired 10-15 min after intravenous bolus injection of 0.2 mmol/kg of Gadobutrol (Gadovist, Bayer Schering, Germany) in short-axis and orthogonal long-axis planes using phase-sensitive inversion-recovery sequences (repetition time/echo time 690-850/3.2 ms; adjusted inversion time; slice thickness of 8 mm; inter-slice gap of 0.8 mm; in-plane

resolution of1.7 x 1.7 mm). A similar magnetic resonance study was available also in the proband´s uncle (II/3).

3. Molecular genetic analysis

Genomic DNA of all available individuals was extracted from whole blood samples using a standard technology. Genomic DNA of the proband´s uncle was available from a previous research project [21]. With respect to known genetic heterogeneity of familiar cardiomyopathies, the whole exome sequencing was performed as an initial genetic test. Exome sequencing was performed using 2 μg of DNA from three affected individuals (II/ 3, III/4, III/5). For DNA enrichment, individually bar-coded DNA libraries [22] and SureSelect All Exome Kit V4 (Agilent, Santa Clara, USA) were 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). A 50 base pair reads were aligned in color space to the reference genome (hg19) using NovoalignCS version 1.08 (Novocraft, Malaysia) allowing for up to six mismatches. Sequence variants in analyzed samples were identified using SAMtools package (version 0.1.8) [23]. The high confidence variants, (quality >50 and coverage >10X), were annotated using ANNOVAR Annotation tool (hg19). Only the sequence variants present in all three affected individuals 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. Identified genetic variants were filtered according to autosomal dominant or X-linked genetic model of the disease which was inferred from the pedigree structure and the resulting 26 candidate coding variants were evaluated in GeneDistiller (24), according to biological relevance of corresponding genes to cardiomyopathy. Candidate variants were visualized in Integrative

Genomics Viewer (IGV) - version 1.5.65. Mutation-bearing fragment of *FHL1* (NM_001159702) was 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 *FHL1* mutation with the phenotype was assessed.

4. Analysis of myocardial samples

Material collection

Myocardial specimens from hearts explanted during transplantation were available in the proband (III/4) and his uncle (II/3). Samples from all major anatomic locations of the proband´s heart were prospectively saved for histopathology, immunohistochemistry, electron microscopy and were also snap-frozen for molecular analyzes. Explanted heart from the male patient II/3 was available as formaldehyde fixed paraffin embedded tissue (FFPE) from all major anatomic heart locations and small snap-frozen samples for molecular analyzes from an another research project [21].

Control samples

Snap-frozen myocardial specimens were obtained from four hearts explanted during heart transplantation. Underlying cardiac disease was dilated cardiomyopathy in control 1 a 18-year-old female and control 2 a 52-year-old male. Ischemic heart disease had control 3 a 53-year-old male and control 4 a 52-year-old female.

Histology and immunohistochemistry

FFPE samples were sectioned (4 μm) and processed for histology and immunohistochemistry.

The paraffin sections were stained by standard histological techniques (hematoxylin-

eosin, trichrom stain). Immunohistochemistry of sarcomeric actin, desmin, lysosomal-associated membrane proteins (LAMP) 1 and 2, and beta subunit of mitochondrial ATP synthase (ATP B) was applied on selected paraffin sections. Immunohistochemical findings were evaluated in comparison with those seen in a set of previously examined myocardial specimens diagnosed for cardiomyopathies based on mitochondrial disorders, lysosomal storage disorders and amyloidosis, and for dilated or hypertrophic cardiomyopathy not otherwise specified for a genetic defect.

Immunohistochemical analysis of FHL1 protein

The monoclonal antibody detecting FHL1 (Ab58067, Abcam, Cambridge, UK) was tested in both paraffin and cryostat sections of myocardium and skeletal muscle including demasking of antigen epitopes in FFPE material. Specific and sensitive results were obtained only in unfixed frozen tissues. Immunohistochemical detection of FHL1 was therefore performed in cryostat sections of myocardium from the proband III/4. Frozen samples from male patient II/3 were no longer available for immunohistochemical analysis. Cryostat sections (5 µm) fixed in cold anhydrous acetone for 10 min at -20°C were dried for 20 min at room temperature, then underwent standard blocking procedures. Primary mouse monoclonal antibody (Ab58067, Abcam, Cambridge, UK) was applied diluted 1: 100 in 5% BSA, in PBS. Sections were incubated for 1 hour at 37°C with the primary antibody, then washed in PBS. Detection of bound primary antibody was achieved using Dako EnVision+ System-HRP Mouse (Dako, Glostrup, Denmark) with 3,3´- diaminobenzidine as substrate.

Electron microscopy

Samples fixed in 10% paraformaldehyde were subsequently embedded into an Epon-Araldite mixture, double stained, and examined using a Jeol 1200 electron microscope.

DOI: 10.1161/CIRCGENETICS.113.000245

***Qualitative (RT-PCR) analysis of FHL1* mRNA**

Total RNA was isolated from snap-frozen myocardial specimens using the TRIZOL solution (Invitrogen, Carlsbad, Ca). RNA concentration was determined spectrophotometrically at 260 nm by NanoDrop (NanoDrop Technologies, Wilmington, USA) and quality was checked on Agilent 2100 bioanalyser - RNA Lab-On-a-Chip (Agilent Technologies, Santa Clara, USA). Aliquots of isolated RNA were stored at –80°C until analysis. The first-strand cDNA synthesis was carried out using oligo-dT primer and SuperScript® III Reverse Transcriptase (Life Technologies, Carlsbad, USA). *FHL1* cDNA was PCR-amplified from the synthesised first-strand cDNA using oligonucleotide primers cFHL1_135U 5´-CCCATCGGTGCGGACTCCAAG-3´ and cFHL1_1013L 5´-TTTGGCACAGTCGGGACAATA-3´ designed to span and amplify all three *FHL1* mRNA isoforms in parallel. Obtained PCR products were sequenced using version 3.1 Dye Terminator cycle sequencing kit (Applied Biosystems, FosterCity, CA) with electrophoresis on an ABI 3500XL Avant Genetic Analyzer (Applied Biosystems).

***Quantitative (qPCR) analysis of FHL1* mRNA.**

The reactions were carried out in a 96-well plate in a 20-µl reaction volume containing 10 µl of 2×Maxima SYBR Green qPCR Master Mix (Thermo Scientific); 0.2 µM of cFHL1_135U and cFHL1_1013L primers and 5 ng of cDNA on a StepOne Plus Real Time System (Applied Biosystems). Data were analyzed by StepOne Software v 2.0. The comparative Ct (ΔΔ Ct) method was used to normalize *FHL1* mRNA to *GAPDH* mRNA amounts.

Western blot analysis of FHL1 protein

Frozen myocardial specimens were homogenized under liquid nitrogen, dissolved in 10mM Tris, 10mM KCl, 2mM EDTA, 4% glycerol, 1mM DTT, Complete Protease Inhibitor Cocktail (Roche

Diagnostics, Basel, Switzerland), centrifuged at 15 000 g for 30 min at 4°C, and assessed for protein content in the supernatant using the Bradford assay. Homogenate aliquots corresponding to 30 μg of total protein were resolved on 10% SDS-PAGE and transferred to PVDF membrane. Membranes were blocked by 5% BSA, 0.05% Tween 20 in PBS. FHL1 was visualised by incubation with mouse monoclonal antibodies potentially detecting all three FHL1 isoforms (Ab58067, Abcam, Cambridge, UK) diluted 1: 2000 in 5% BSA, 0.05% Tween 20 in PBS for 90 min, followed by incubation with goat anti-mouse HRP (Pierce), 1:20000 in 0.05% Tween 20 in PBS for 60 min and detection by SuperSignal West Pico Maximum Sensitivity Substrate (Pierce).

5. Neurological assessment

All living individuals with *FHL1* mutation underwent neurological assessment focused on detection of possible muscle disability (muscle atrophy, muscle strength testing), including standard routine examination of the central and peripheral nervous systems. In adult patients with *FHL1* mutation, nerve conduction studies of the motor and sensory nerves of lower extremities were performed to exclude neuropathy. Two muscles (left vastus medialis and left deltoid muscle) were investigated by needle electromyography. Electromyographic signals were evaluated using commercially available analyzer Medelec Synergy (Viasis Neurocare, Madison, Wisconsin, US) to detect possible muscle disease. In addition, all living individuals with FHL1 mutation had a biochemical analysis to assess creatine phosphokinase activity and myoglobin concentration in serum.

Results

Clinical features of the proband

The proband was a 31-year-old male (III/4) with a history of non-obstructive HCM since 18

years of age who presented with dyspnea in New York Heart Association (NYHA) functional class IV, right heart failure and new-onset atrial fibrillation with a rapid ventricular response. Symptoms of end-stage heart failure persisted even after successful cardioversion to sinus rhythm. Proband´s electrocardiogram showed signs of biatrial enlargement and left ventricular (LV) hypertrophy (Figure 2). In comparison with previous data (Table 1), the most recent echocardiography study demonstrated non-obstructive HCM with less prominent LV hypertrophy, preserved LV systolic function and severe LVDD. Cardiac catheterization revealed normal coronary arteries, low cardiac output, elevated filling pressures of both ventricles and borderline pulmonary artery pressures (Table 1). These findings were compatible with advanced heart failure due to non-obstructive HCM with severe LVDD. The patient underwent an uneventful heart transplantation six months later. His family history revealed diagnosis of HCM in his brother and uncle.

Pedigree analysis

Male patient III/5

Brother of the proband was diagnosed with non-obstructive HCM aged 17. At 30 years of age, he reported mild breathlessness. ECG showed 1st degree of atrioventricular block, signs of biatrial enlargement and LV hypertrophy (Figure 2). Echocardiographic findings were compatible with non-obstructive HCM with severe LVDD (Table 1).

Male patient II/3

Uncle of the proband experienced bilateral decompensation of heart failure aged 55 years. Subsequent cardiologic examination revealed non-obstructive HCM with severe LVDD disproportional to mild to moderate LV systolic dysfunction (ejection fraction 40-45%) (Table 1, Figure 3). Aged 59 years, he was admitted to our institution due to end-stage heart failure and

DOI: 10.1161/CIRCGENETICS.113.000245

underwent an urgent heart transplantation. Unfortunately, the postoperative course was complicated by severe right ventricular failure, hepatorenal failure and septicemia. The patient died 4 months after transplantation.

Female patient I/2

Grandmother of the proband was diagnosed with apical form of HCM during cardiologic screening in the family. This 83-year-old lady reported a history of myocardial infarction, diabetes and breathlessness NYHA class III; however, she did not experience any episode of heart failure. Echocardiography revealed apical hypertrophy of the left ventricle beginning in the mid-ventricular segment (septum 14 mm, lateral wall 12 mm) reaching 15-16 mm in the apical third (Figure 3). The patient refused magnetic resonance imaging and any other examinations.

Results of the cardiologic screening

Cardiologic examination in the remaining family members revealed electrocardiographic signs of LV hypertrophy (Figure 2) in two asymptomatic females heterozygous for *FHL1* mutation (II/1- aunt of the proband, age 63y; II/2- mother of the proband, age 53y). Echocardiography demonstrated normal findings in both females except of modest increase in LV mass (98g/m^2 and 90/m^2, respectively). Mild electrocardiographic abnormalities, such as mild depression of ST segments and either negative or biphasic T waves in leads III and aVF, were detected in two other females heterozygous for *FHL1* mutation (III/1, age 41y; III/8, age 34y) and one male hemizygous for *FHL1* mutation (IV/2, age 17y). All these subjects were asymptomatic and had normal echocardiographic findings. The results of the cardiologic screening were normal in the remaining members of the family including a female heterozygous for *FHL1* mutation (IV/1, age 20y) and a male heterozygous for *FHL1* mutation (IV/9, age 2y).

DOI: 10.1161/CIRCGENETICS.113.000245

Genetic and molecular biology analyses

To directly identify possible disease-causing mutation we performed exome sequencing in three affected individuals - (II/3, III/4, III/5). From the sequencing run we obtained from these samples 128, 151 and 109M of sequencing reads, respectively, of which on average 57 % we were able to map uniquely on the human genome reference sequence (hg19). Compared to the reference sequence we identified in these data 16486, 17953 and 16591 variants, respectively. Among them we identified 8515 single nucleotide variants (SNVs), (SNP/indel qual > 50) that were present in all three affected probands. From these variants, 69 SNVs (2 stop-gain, 46 non-synonymous SNVs and 21 synonymous SNVs) and 3 indels were either novel or present at frequencies lower than 0.01 in the 1000 Genomes, Exome Variant Server (http://evs.gs.washington.edu/EVS/) and internal exome database (> 200 exomes). Resulting 51 rare coding variants, 48 SNVs and 3 indels (Supplemental Table 1), were further evaluated in GeneDistiller (24) according to potential biological and clinical relevance of corresponding genes to cardiomyopathy. From this evaluation emerged as the only potentially causative variant the frameshift insertion c.599_600insT in exon 6 of *FHL1* (NCBI reference sequence: NM_001159702). *FHL1* encodes the four-and-a-half-LIM-domain protein 1, (FHL1). The gene is transcribed into three alternatively spliced mRNA isoforms *FHL1A*, *FHL1B* and *FHL1C*, encoding FHL1A, FHL1B and FHL1C proteins, respectively. *FHL1A* is by far the most abundant form. It is highly expressed in skeletal muscle, moderately in heart and to a much lesser extent in a wide array of other tissues [10,11]. Compared to *FHL1A* the other two isoforms are expressed at much lower abundance and demonstrate tissue specific expression patterns. *FHL1B* is specifically expressed in brain, whereas *FHL1C* is expressed in skeletal muscle and at lower levels in aorta, left atrium, left, and right ventricles of human heart [10,11,25]. The insertion encodes for a frameshifts in

DOI: 10.1161/CIRCGENETICS.113.000245

translation of *FHL1A* and *FHL1B* isoforms, that are in both cases followed soon thereafter by a novel stop codon predicting proteosynthesis of one sequentially identical truncated form of the FHL1A containing a neo-peptide composed of 32 aminoacids (200 HRCGGPVLLRGLLQELCGQEVC WMQEPHHWVW-232) on the C-terminal end (p.F200fs32X). Located in alternatively spliced exon 6, the identified insertion is not predicted to affect proteosynthesis of the FHL1C variant (Figure 4A). We confirmed the presence of the c.599_600insT variant of *FHL1* in genomic DNA of probands by Sanger sequencing (Figure 4B) and using the same technique we subsequently demonstrated segregation of the mutation with the phenotype of cardiac disease in the family (Figure 1).

To characterize molecular consequences of the identified mutation, we assessed expression of *FHL1* mRNA isoforms and presence of FHL1 proteins in patients and control tissues obtained from two females (C1F and C4F) and two males (C2M and C3M). We performed quantitative PCR and RT-PCR analysis of total RNA extracts and Western blot analysis of protein homogenates prepared from the snap-frozen myocardial specimens. Quantitative PCR analysis suggested that FHL1 transcription may be generally higher in females than in males and that the mutation and related dysfunction of FHL1 protein may stimulate FHL1 transcription in affected males. (Figure 4C). RT-PCR analysis revealed in all analyzed myocardial specimens presence of a single RT-PCR product of the size expected for *FHL1A* isoform (Figure 4D). The Sanger sequencing and sequence analysis demonstrated that the obtained RT-PCR products correspond to *FHL1A* mRNA isoform and independently confirmed in patients´ samples the presence of the frameshift insertion c.599_600insT identified previously in corresponding genomic DNA (not shown). In accordance with this finding, Western blot analysis of myocardial homogenates revealed in patient´s samples presence of the immune-

reactive protein of a molecular weight ~ 27 kDa corresponding to predicted molecular weight of the p.F200fs32X FHL1 protein. Immunoreactive protein of a molecular weight ~ 32 kDa, corresponding to predicted molecular weight of the FHL1A identified in control samples, was in patients´ samples absent. Immunoreactive protein of a molecular weight ~ 22 kDa, corresponding to predicted molecular weight of the FHL1C was not detected. Instead, the analysis revealed in all control´s and patient´s samples presence of the immunoreactive protein of a molecular weight ~ 25 kDa, which identity is unknown (Figure 4E). In contrast to Western blot, the immunohistochemical staining of cryostat sections of myocardium revealed in the proband III/4 absence of any immunoreactive FHL1 protein (Figure 4F) in comparison to distinct cross-striation pattern observed in control myocardium (Figure 4G).

Morphological findings in two explanted hearts

Male patient III/4 (proband):

The explanted heart weighted 346 grams (less than the upper limit of normal 400 grams). On gross pathology, both ventricles were non-dilated with relatively thin walls (left ventricle 9-12 mm, right ventricle 5 mm) and discrete focuses of fibrosis on incision. Both atria were dilated with limited amounts of tissue missing due to surgical reasons. Coronary vessels showed mild intimal thickening. Histopathology revealed hypertrophy of cardiomyocytes with anisonucleosis and hyperchromasia, interstitial and replacing fibrosis, subendocardial scars and vacuolization of cardiocytes, and mild sclerosing arteriosclerosis. Disarray of cardiomyocytes was present, involving 5-15% of the tissue section in the left ventricular wall (Figure 5A, C, E). Sarcomeric actin and desmin detected immunohistochemically revealed an usual cross-striation pattern without detectable inclusions, being slightly disordered in focuses of myocyte disarray (Figure 5F). There were no abnormalities of lysosomal and mitochondrial compartments, based on

immunohistochemistry. Electron microscopy showed prevalent regular organization of myofibrils and increased number of mitochondria, in subsarcolemmal localization or focally replacing myofibrils. Mitochondria did not show any substantial abnormalities in size, or shape and organization of their crists. Disorganized myofibrils in cardiocytes were found exceptionally (Figure 5G, H). Morphological findings were comparable in both ventricles.

Male patient II/3 (proband´s uncle)

The explanted heart weighted 578 grams (the upper limit of normal 400 grams). Gross pathology showed mild dilatation of both ventricles with thickened LV wall (10-18 mm), dilated atria and mild atherosclerotic plaques of coronary arteries. Histopathology demonstrated hypertrophy of cardiomyocytes, focal myocyte disarray less frequent than in the proband and marked interstitial and replacing fibrosis (Figure 5B, D). Immunohistochemical staining for sarcomeric actin and desmin showed patterns comparable with that in the proband. No abnormalities of lysosomal and mitochondrial systems were detected using immunohistochemistry. The histopathological changes involved both ventricles.

In summary, both patients, the proband (III/4) and his uncle (II/3) showed to a different degree of hypertrophy and disarray of cardiocytes together with interstitial fibrosis, which are features consistent with HCM at the histological level.

Neurological assessment

Mutations in FHL1 have been identified as the cause of several skeletal muscle diseases. In this respect, carriers of the c.599_600insT mutation underwent neurological examination. Neurological status of individuals with *FHL1* mutation was unremarkable, except of mild muscle hypotrophy in the proband. However, needle electromyographic recordings as well as nerve conduction studies showed normal findings. Similarly, biochemical markers of muscle injury in

peripheral blood had normal values in all individuals with *FHL1* mutation. Taken together, there were no signs of muscle disease in these individuals.

Discussion:

In this report, we describe a Czech family in which frameshift mutation c.599_600insT in exon 6 of *FHL1,* a gene located on chromosome X, co-segregated with non-obstructive HCM with severe LVDD in advanced disease in three hemizygous males. Penetrance of this mutation was age-dependent and included manifestation of HCM in hemizygous males in the second to sixth life decade and progression into advanced heart failure in two of these individuals in the fourth to sixth decennium. On the other hand, mild cardiac involvement in heterozygous females, consistent with random X-inactivation of mutated FHL1, included asymptomatic abnormalities of electrocardiogram with a modest increase in LV mass since the sixth decennium and one case of apical HCM in the ninth decennium. Microscopic features compatible with HCM were detected in both cases available for histopathological analyses. None of the individuals showed signs of skeletal muscle disease.

Molecular consequences of c.599_600insT *FHL1* mutation

The above mutation is either very rare or even private, as it has not been observed in any of the publically searchable human genome, exome or polymorphism and mutation databases as well as in our internal population specific human exome sequence database (> 200 exomes). The mutation does not affect synthesis and stability of *FHL1A* mRNA in heart tissue and the transcript seems to be translated into truncated 27 kDa form of the FHL1 protein which is probably missing LIM3 and LIM4 domains, and containing the neo-peptide composed of 32 aminoacids on C-terminus. In parallel, the mutation which is predicted to abolish proteosynthesis of FHL1A and FHL1B proteins, should not have impact on proteosynthesis of FHL1C. In our

study, we have not found any evidence for expression of *FHL1C* mRNA isoform and production of FHL1C protein (of predicted molecular weight) in analyzed control and patient derived myocardial specimens. Instead, we detected in all these samples presence of an immunoreactive protein of a molecular weight ~ 25 kDa. Origin of this protein is unknown and it remains to be investigated whether it represents post-translationally modified FHL1C isoform, proteolytically processed FHL1 protein, or simply nonspecificity of the antibodies used in this study. The amounts of the p.F200fs32X FHL1 protein in homogenates of patients myocardial specimens were comparable to amounts of FHL1A in control samples. Contrary to this observation, we found reduction of FHL1 signal in frozen tissue sections from patients specimens. Decrease in FHL1 staining in this material thus suggests that p.F200fs32X mutation results in proteosynthesis of truncated neo-protein, which has, either due to loss of LIM 3 and LIM4 domains and/or presence of the neo-peptide on its C-terminal end, altered structural properties limiting function and immunodetection of the mutant protein in the heart tissue.

Reasons for preserved skeletal muscle function in p.F200fs32X FHL1 mutation

Our finding of isolated HCM resulting from complete structural and functional deficiency of FHL1A demonstrates different biological and functional roles of FHL1 proteins in skeletal and heart muscle. In skeletal muscle, it has been shown that FHL1 may have multiple roles in myoblast migration and elongation, satellite cell activation, inhibition of myoblast apoptosis, regulation of skeletal muscle mass, sarcomere formation and Notch signaling [10,11]. Clinical studies revealed that C-terminal FHL1 mutations preserving synthesis of FHL1C are responsible for less severe myopathic phenotype than the N-terminal ones [10,26]. Importantly, FHL1C isoform seems to be expressed much stronger in human skeletal muscle than in myocardium [25]. Preserved expression of the FHL1C isoform in skeletal muscle may explain absence of

myopathy in our case. Unfortunately, we had no opportunity to analyze skeletal muscle specimens in our patients.

In addition, missense mutations disrupting zinc-binding residues critical for tertiary structure of FHL1 result in more severe myopathy than FHL1 truncations [10]. This can be illustrated by replacement of highly conserved cysteine binding a zinc-ion within LIM3 domain (c. 625T>C; p.C209R). Although this mutation affects the same exon and LIM domain as in our case, it causes not only hypertrophic cardiomyopathy but also Emery-Dreifuss myopathy with cytoplasmatic bodies suggesting misfolding of mutated FHL1 protein [13]. This observation provides an alternative hypothesis, that mutations leading to reduced amounts (or even absence) of FHL1 proteins are less deleterious (or even benign, as in our case) for skeletal muscle than missense mutations exerting their pathogenic effects through misfolding, self-aggregation and co-aggregation of FHL1-binding partners [10,11,14].

Practical implications

HCM with severe LVDD has been associated with mutations of myofilament proteins (beta-myosin heavy chain, cardiac troponin I, cardiac troponin T, alpha-cardiac actin) [7,8] and Z-disk proteins (myopalladin) [27]. However, myofilament mutations explained in the above mentioned study [7] approximately 50% cases of HCM with severe LVDD. In our study, phenotype of HCM with severe LVDD seemed to be associated rather with advanced disease than to be a primary phenotype of the above mentioned FHL1 mutation. As shown by Friedrich et al., *FHL1* mutations may explain etiology of isolated HCM in 2.5% of cases unexplained by traditional sarcomeric mutations [15]. In addition, a recent report of Binder et al. extends the spectrum of FHL1-related diseases by description of spongious HCM in individuals with X-linked myopathy with postural muscle atrophy [28]. In summary, mutation of *FHL1* should be suspected in

individuals with X-linked HCM. FHL1 mutations thus extend spectrum of X-linked HCM where belong Barth syndrome, Danon´s and Fabry´s disease. Exome sequencing seems to be the most practical approach in genetic dissection of familiar cardiomyopathies.

Study limitations:

Unfortunately, we did not have an opportunity to study the pathophysiology of p.F200fs32X FHL1 mutation in experimental models. However, perfect segregation of the mutation in 12 individuals, available evidence in other FHL1-related diseases, and absence of wild-type FHL1A protein in myocardial samples support involvement of the above mutation in the pathophysiology of HCM in our case. Finally, the disease severity may be determined not only by the gene mutation itself, but also by gene dosage and age. Further studies are needed to evaluate relationship between the above gene mutation and LVDD observed in our patients.

Conclusions:

We identified a novel mutation of *FHL1* (c.647_648 ins.T) causing isolated HCM with X-chromosomal inheritance and severe LVDD in advanced disease in three hemizygous males. Mild cardiac involvement in heterozygous females included asymptomatic abnormalities of electrocardiogram with a modest increase in LV mass since the sixth decennium and one case of apical HCM in the ninth decennium. None of the individuals showed signs of skeletal muscle disease. Mutations of *FHL1* should be suspected in individuals with X-linked HCM.

Funding Sources: The study was supported by Ministry of Health, Czech Republic - Conceptual Development of Research Organization („Institute for Clinical and Experimental Medicine – IKEM, IN 00023001“) and by the EU Operational Program Prague—Competitiveness: project CEVKOON (CZ.2.16/3.1.00/22126). S.K.et al. were funded by the by Charles University institutional programs PRVOUK-P24/LF1/3, UNCE 204011 and SVV2013/266504; and by BIOCEV – Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University“ (CZ.1.05/1.1.00/02.0109), from the European Regional Development Fund. Specific

support was provided by grant NT13116-4/2012 from the Ministry of Education and Ministry of Health of the Czech Republic.

Conflict of Interest Disclosures: None.

References:

1. Matsumura Y, Elliott PM, Virdee MS, Sorajja P, Doi Y, Mc Kenna WJ. Left ventricular diastolic function assessed using Doppler tissue imaging in patients with hypertrophic cardiomyopathy: relation to symptoms and exercise capacity. *Heart.* 2002;87:247-251.

2. Efthimiadis GK, Giannakoulas G, Parcharidou DG, Karvounis HI, Mochlas ST, Styliadis IH, et al. Clinical significance of tissue Doppler imaging in patients with hypertrophic cardiomyopathy. *Circ J.* 2007;71:897-903.

3. Biagini E, Spirito P, Rocchi G, Ferlito M, Rosmini S, Lai F, et al. Prognostic implications of the Doppler restrictive filling pattern in hypertrophic cardiomyopathy. *Am J Cardiol*. 2009;104:1727-1731.

4. McMahon CJ, Nagueh SF, Pignatelli RH, Denfield SW, Dreyer WJ, Price JF, et al. Characterization of left ventricular of left ventricular diastolic function by tissue Doppler imaging and clinical status in children with hypertrophic cardiomyopathy. *Circulation*. 2004;109:1756-1762.

5. Maskatia SA, Decker JA, Spinner JA, Kim JJ, Price JF, Jefferies JL, et al. Restrictive physiology is associated with poor outcome in children with hypertrophic cardiomyopathy. *Pediatr Cardiol.* 2012;33:141-149.

6. Harris KM, Spirito P, Maron MS, Zenovich AG, Formisano F, Lesser JR, et al. Prevalence, clinical profile, and significance of left ventricular remodeling in the end-stage phase of hypertrophic cardiomyopathy. *Circulation.* 2006;114:216-225.

7. Kubo T, Gimeno JR, Bahl A, Steffensen U, Steffensen M, Osman E, et al. Prevalence, clinical significance and genetic basis of hypertrophic cardiomyopathy with restrictive phenotype. *J Am Coll Cardiol.* 2007;49:2419-2426.

8. Mogensen J, Kubo T, Duque M, Uribe W, Shaw A, Murphy R, et al. Idiopathic restrictive cardiomyopathy is part of the clinical expression of cardiac troponin I mutations. *J Clin Invest.* 2003;111:209-216.

9. Menon SC, Michels VV, Pellikka PA, Ballew JD, Karst ML, Herron KJ, et al. Cardiac troponin T mutation in familial cardiomyopathy with variable remodeling and restrictive physiology. *Clin Genet.* 2008;74:445-454.

10. Cowling BS, Cottle DL, Wilding BR, D´Arcy CE, Mitchell CA, McGrath MJ. Four and a

half LIM protein 1 gene mutations cause four distinct human myopathies: A comprehensive review of the clinical, histological and pathological features. *Neuromuscul Disord*. 2011;21:237-251.

11. Shatasivam T, Kislinger T, Gramolini AO. Genes, proteins and complexes: the multifaceted nature of FHL family proteins in diverse tissues. *J Cell Mol Med*. 2010;14:2702-2720.

12. Windpassinger C, Schoser B, Straub V, Hochmeister S, Noor A, Lohberger B, et al. An X-linked myopathy with postural muscle atrophy and generalized hypertrophy, termed XMPMA, is caused by mutations in FHL1. *Am J Hum Genet.* 2008;82:88-99.

13. Knoblauch H, Geier C, Adams S, Budde B, Rudolph A, Zacharias U, et al. Contactures and hypertrophic cardiomyopathy in a novel FHL1 mutation. *Ann Neurol.* 2010;67:136-140.

14. Gueneau L, Bertrand AT, Jais JP, Salih MA, Stojkovic T, Wehnert M, et al. Mutations of the FHL1 gene cause Emery-Dreifuss muscular dystrophy. *Am J Hum Genet.* 2009;85:338-353.

15. Friedrich FW, Wilding BR, Reischmann S, Crocini C, Lang P, Charron P, et al. Evidence for FHL1 as a novel disease gene for isolated hypertrophic cardiomyopathy. *Hum Mol Genet.* 2012;21:3237-3254.

16. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MRG, Towbin JA. Genetic evaluation of cardiomyopathy- a Heart Failure Society of America Practice Guidelines. *J Card Fail.* 2009;15:83-97.

17. Lang RM, Bieri M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, et al. Recommendations for chamber quantification. *J Am Soc Echocardiogr.* 2005;18:1440-1463.

18. Quinones MA, Otto CM, Stoddard W, Waggoner A, Zoghbi WA. Recommendations for quantification of Doppler echocardiography. *J Am Soc Echocardiogr.* 2002;15:167-184.

19. Nishimura RA, Tajik AJ. Evaluation of diastolic filling of left ventricle in health and disease: Doppler echocardiography is the clinician´s Rosetta stone. *J Am Coll Cardiol*. 1997;30:8-18.

20. Heiberg E, Wigstrom L, Carlsson M, Bolger AF, Karlsson M. Time resolved three-dimensional automated segmentation of the left ventricle. *Proceedings of IEEE Computers in Cardiology, Lyon, France.* 2005;32:599-602.

21. Hubacek JA, Vymetalova Y, Bohuslavova R, Kocik M, Malek I. Detection of donor DNA after heart transplantation: how far could it be affected by blood transfusion and donor chimerism? *Transplant Proc.* 2007; 39:1593-1595.

22. Harakalova M, Mokry M, Hrdlickova B, Renkens I, Duran K, van Roekel H, et al. Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective targeted next-generation sequencing. *Nat Protoc.* 2011;6:1870-1886.

23. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAM tools. *Bioinformatics.* 2009;25:2078-2079.

24. Seelow D, Schwarz JM, Schuelke M. GeneDistiller--distilling candidate genes from linkage intervals. *PLoS ONE.* 2008;3:e3874.

25. Ng EK, Lee SM, Li HY, Ngai SM, Tsui SK, Waye MM, et al. Characterization of tissue-specific LIM domain protein (FHL1C) which is an alternatively spliced isoform of a human LIM-only protein (FHL1). *J Cell Biochem*. 2001;82:1-10.

26.Schoser B, Goebel HH, Janisch I, Quasthoff S, Rother J, Bergmann M, et al. Consequences of mutations within the C-terminus of the FHL1 gene. *Neurology*. 2009;73:543-551.

27. Purevjav E, Arimura T, Augustin S, Huby AC, Takagi K, Nunoda S, et al. Molecular basis for clinical heterogeneity in inherited cardiomyopathies due to myopalladin mutations. *Hum Mol Genet*. 2012;21:2039-2053.

28. Binder JS, Wiedemann F, Schoser B, Nieman M, Machann W, Beer M, et al. Spongious hypertrophic cardiomyopathy in patients with mutations in the Four-and-a-Half LIM Domain Gene. *Circ Cardiovasc Genet.* 2012;5:490-502.

DOI: 10.1161/CIRCGENETICS.113.000245

Table 1: Clinical, electrocardiographic, echocardiographic, hemodynamic and laboratory characteristics of individuals with FHL1 mutation associated hypertrophic cardiomyopathy. To demonstrate progression of the disease, available previous observations in each individual were added into this table and labeled with age.

	Proband			Proband´s brother		Proband´s uncle		Proband´s grandmother
Age (years)	18	24	31	17	29	55	59	83
Clinical events	-	-	End-stage HF, HTx		-	HF hosp., NSVT, ICD	End-stage HF, HTx	History of myocardial infarction, diabetes
NYHA functional class	II	II	IV	I	II	III	IV	III
Rhythm	SR	SR	AF, SR	SR	SR	SR	AF	SR
ECG	LVH, LVS	LVH, LVS	BAE, LVH, LVS	LVE	PR 224 ms, BAE, LVH, LVS	LBBB	LBBB	PR 244 ms, QS V1-V3, LVS
LVEDD (mm)	42	45	47	46	48	47	48	52
LVEDD (mm/m)	22	23	24	25	26	26	27	33
Interventricular septum thickness (mm)	12-16	24	13	12-15	13	20	14	10 (apically 15-16 mm)
Posterior wall thickness (mm)	8	14	10	9	11	13	12	9
LV mass (g/m^2)	-	-	80	-	70	-	118	91
LVOT gradient	Absent	Absent	Absent	Absent	Absent	Absent	Absent	Absent
LVEF (%)	Normal	70	55	Normal	60	45	41	55
Restrictive mitral inflow pattern	Absent	Absent	Present	Absent	Present	Present	Present	Absent
E-wave deceleration time (ms)	-	-	60	-	65		80	170
E/A ratio			3.3		3.3		-	
Mitral E and A wave velocities (cm/s)			50; 15		60; 19		72;-	
Mean E/Em ratio	-	-	12	-	9.5		16	10.3

DOI: 10.1161/CIRCGENETICS.113.000245

Mitral regurgitation (scale)	Trivial	Trivial	Mild	None	Trivial	Mild	Moderate	Trivial
Left atrium diameter (mm)	37	42	49	37	43	52	60	37
LAVI (cm^3/m^2)	-	-	59	-	41	-	70	55
Mean right atrium pressure (mm Hg)	-	-	10	-	-	-	10	-
Right ventricle pressure (mm Hg)	-	-	29/13	-	-	-	45/4	-
Mean pulmonary artery pressure (mm Hg)	-	-	24	-	-	-	33	-
Mean pulmonary capillary wedge pressure (mm Hg)	-	-	18	-	-	-	20	-
Cardiac output (l/min)	-	-	3.3	-	-	-	3.9	-
Cardiac index (l/min.m2)	-	-	1.6	-	-	-	2.0	-
Pulmonary arterial resistance (Wood´s units)	-	-	1.8	-	-	-	3.3	-
B-type natriuretic peptide (ng/l)	-	-	1002	-	991	-	2324	221
Creatine phosphokinase (μkat/l)	-	-	1.22	-	2,21	-	3.0	2.66
Myoglobin (μg/l)	-	-	140	-	139	-	-	115

Abbreviations: AF- atrial fibrillation, BAE- biatrial enlargement, E/Em ratio- early mitral inflow velocity to peak mitral annulus velocity ratio, ICD- implantable cardioverter-defibrillator, HF- heart failure, HTx- heart transplantation, LAVI- left atrium volume index, LBBB- left bundle branch block, LVS- left ventricular strain, LVEF- left ventricular ejection fraction, LVH- left ventricular hypertrophy, LVOT- left ventricular outflow tract, NSVT- non-sustained ventricular tachycardia, SR- sinus rhythm.

The upper limit of normal: creatine phosphokinase 3.33 μkat/l, myoglobin 140 μg/l.

DOI: 10.1161/CIRCGENETICS.113.000245

Figure Legends:

Figure 1: Pedigree. Black symbols indicate affected males with hypertrophic cardiomyopathy with restrictive physiology and one affected female with apical hypertrophic cardiomyopathy. Half-filled symbols depicts individuals with an abnormal electrocardiogram and a normal echocardiogram. Plus signs indicate presence of the FHL1 c.599_600 insT mutation. Minus signs indicate absence of this mutation. Circles, females; squares, males; diagonal line, deceased; arrow, the proband.

Figure 2: Electrocardiograms in individuals with FHL1 mutation (labeled by patient numbers).

Figure 3: Images from magnetic resonance in three males with restrictive HCM and transoesophageal echocardiogram in female with apical HCM (labeled by patient numbers).

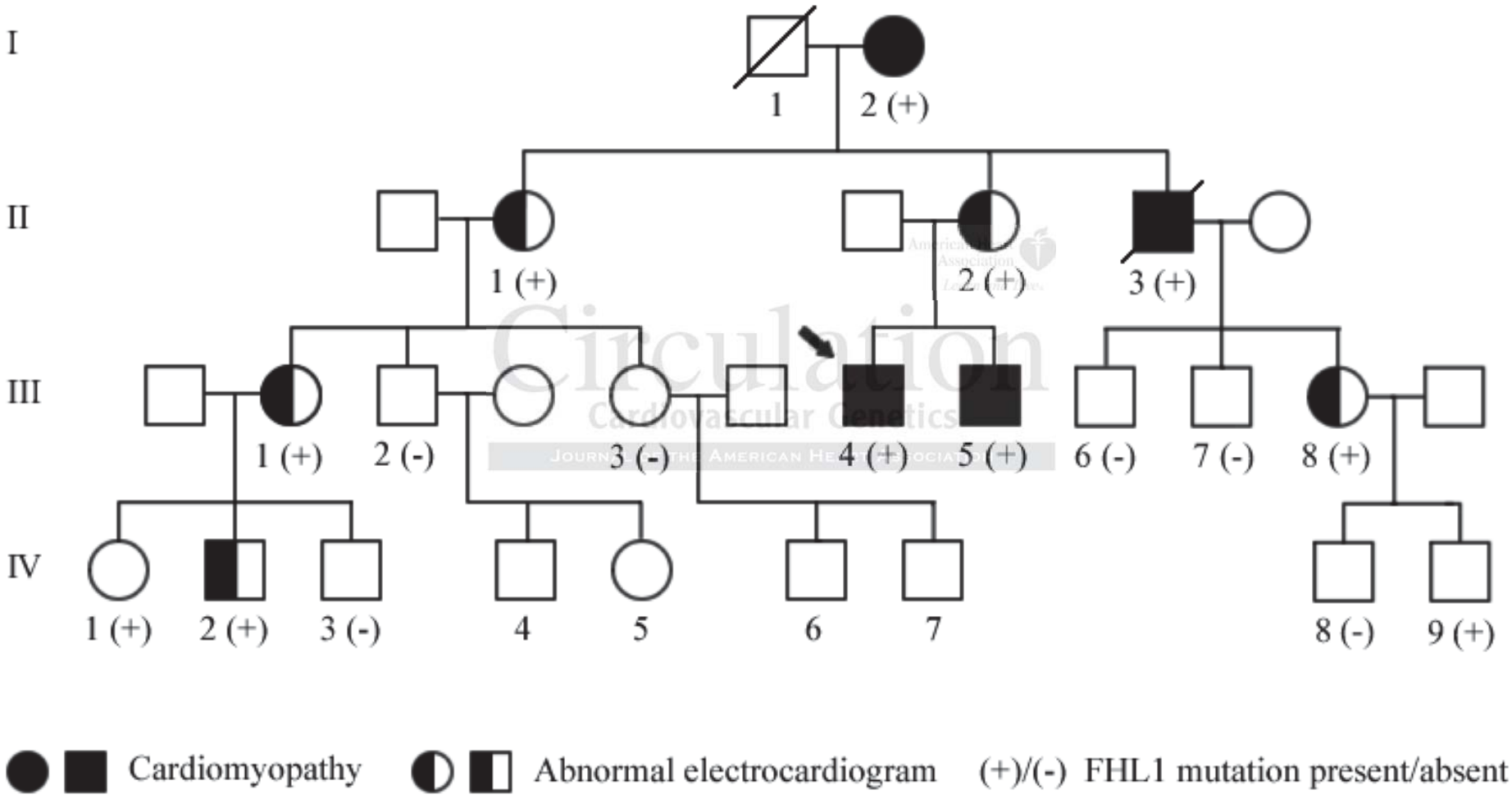
Figure 4: Effects of the identified *FHL1* mutation. (**A**) Schematic representations of *FHL1* genomic structure (*FHL1* gene), *FHL1* mRNA isoforms (*FHL1A*, *FHL1B* and *FHL1C*) and corresponding FHL1 protein variants. Light gray and dark gray boxes demonstrate alternatively and constitutively spliced exons, respectively. ATG – position of the initiation codon; TAA, TGA – position of the stop codons. Positions of the c.599_600 insT mutation and of the primers (U primer and L primer) used for RT-PCR and qPCR amplification are indicated. Blue boxes demonstrate FHL1 domains. LIM1/2, LIM1, LIM2 and LIM4 denotes individual LIM domains; NLS - nuclear localization signal; NES – nuclear export sequence; RBP-J_K - recombining binding protein immunoglobulin J kappa binding site. Protein domain recognized by employed

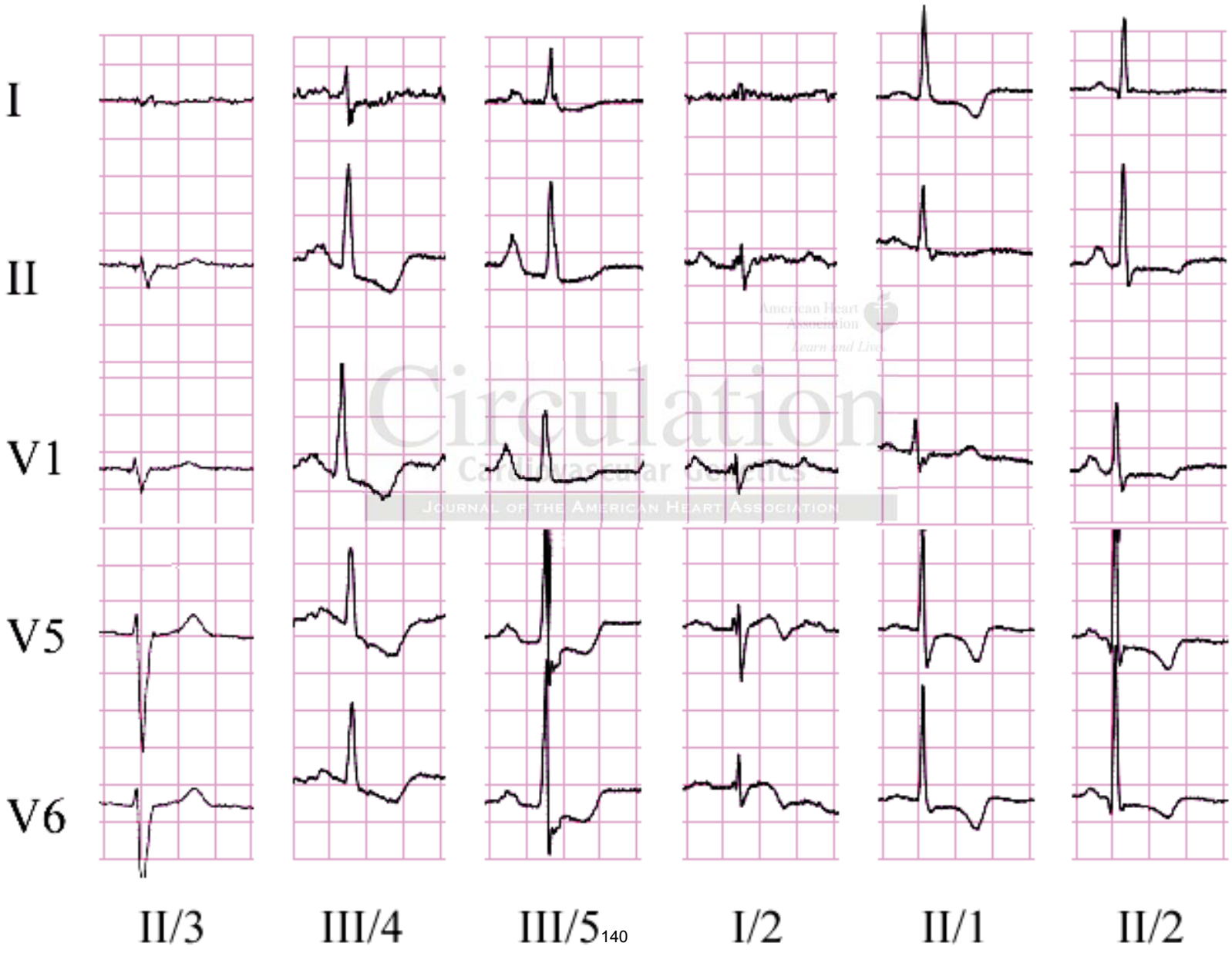
FHL1 antibody and predicted effects of the p.F200fs32X mutation on expression of FHL1 variants in FHL1 are indicated. (**B**) Chromatograms of *FHL1* genomic DNA sequences showing identified mutations in the Czech family. (Upper panel) Sequence of the proband, (middle panels) sequence showing heterozygous mutation in the mother, and (lower panel) sequence of an unaffected individual. (**C**) Relative amounts of *FHL1* mRNA normalized to amounts of Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH)* mRNA in myocardial specimens; C1-C4 denotes specimens from controls, P1 and P2 denotes specimens from probands III/4 and II/3, respectively; adjoined F or M denotes gender (female of male, respectively) of control individuals. The means ± SD of three experiments performed in triplicates are shown. (**D**) *FHL1A* cDNA analysis showing profiles of RT-PCR products amplified from total RNA isolated from snap-frozen myocardial specimens of control (C1-C4) and probands III/4 (P1) and II/3 (P2). A single RT-PCR product of the size of 699 base pairs expected for *FHL1A* isoform is detected. RT-PCR product of the size 512 bp expected for *FHL1C* isoform is absent; M - is a 100 base pair DNA ladder; B-blank. (**E**) Western blot analysis of homogenates prepared from snap-frozen myocardial specimens showing presence of the immune-reactive protein of a molecular weight ~ 27 kDa corresponding to predicted molecular weight of the p.F200fs32X FHL1 protein in samples from probands P1 and P2. Immunoreactive protein of a molecular weight ~ 32 kDa, corresponding to predicted molecular weight of the FHL1A identified in high abundance in control samples (C1-C4), is in patients´samples (P1, P2) absent. Specific immunoreactive protein of a molecular weight ~ 22 kDa corresponding to predicted molecular weight of the FHL1C is not detected in patients´ and control samples (C1-C4). Comparable staining intensity of Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) demonstrate that approximately equal protein amounts have been analysed. (**F,G**) Immunofluorescence analysis of cryostat sections of

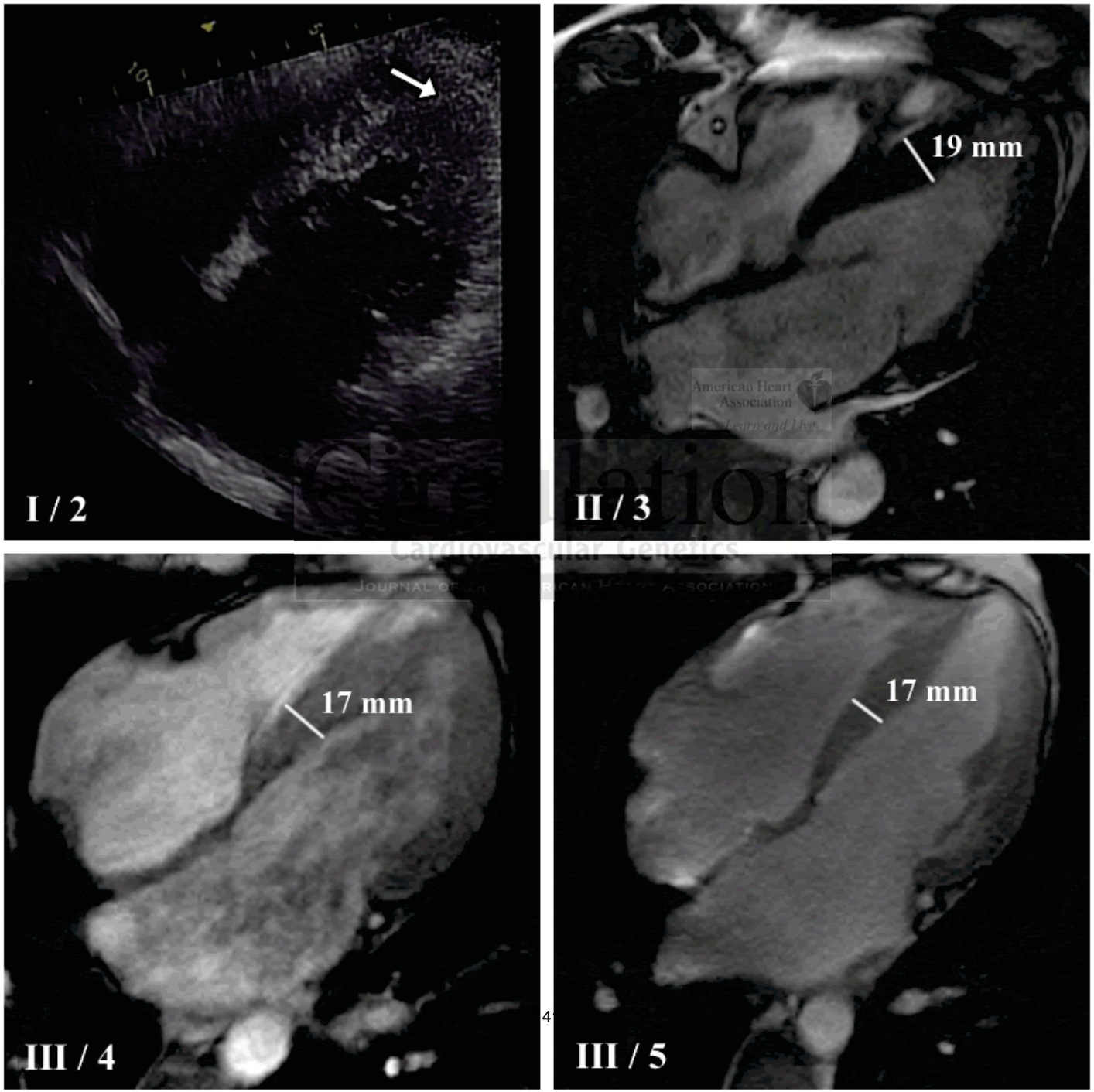
DOI: 10.1161/CIRCGENETICS.113.000245

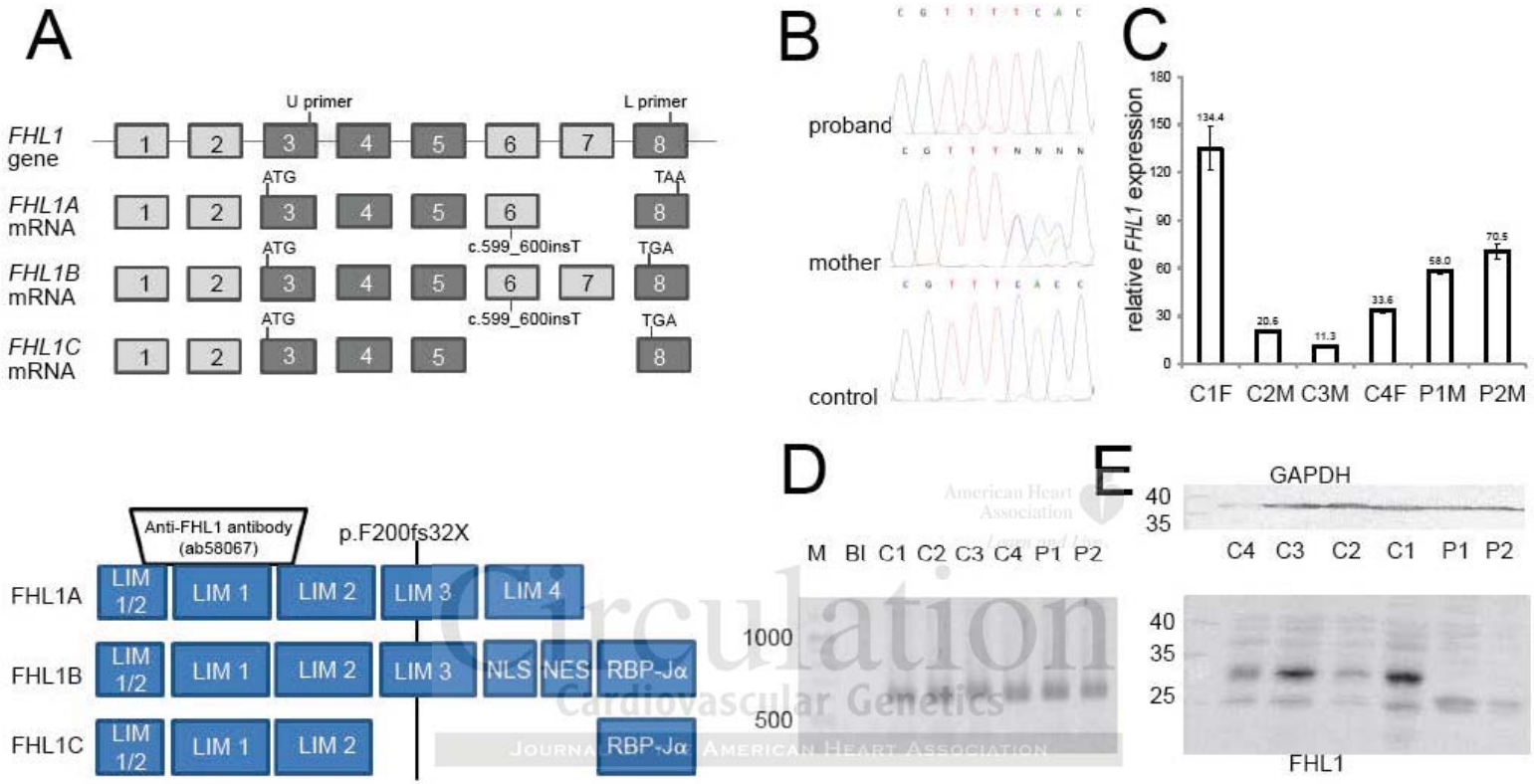
myocardium showing absence of immuno-reactive FHL1 in the proband III/4, (**F**) in comparison to distinct cross-striation pattern in control cardiocytes (**G**).

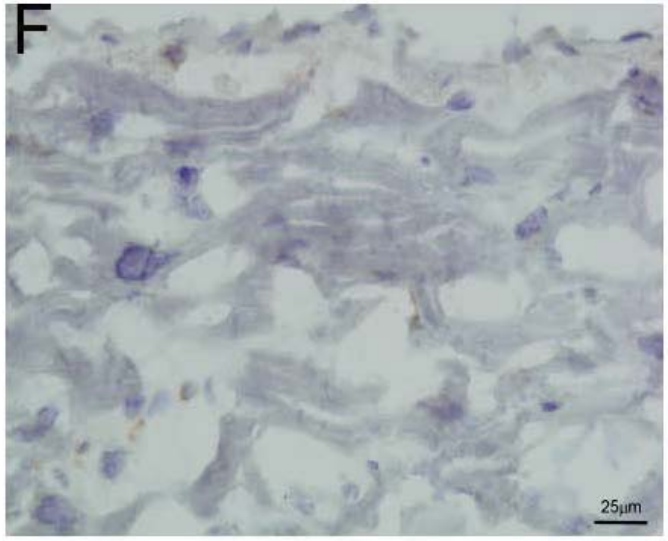
Figure 5: Histopathology and electron microscopy**.** (**A**) Hypertrophy and disarray of cardiocytes in patient III/4, H&E stain. (**B**) Combination of hypertrophic and regressive changes in myocardium in patient II/3, H&E stain. (**C**) Interstitial fibrosis in myocardium in patient III/4, trichrome stain. (**D**) Interstitial and replacing fibrosis in myocardium in patient II/3, trichrome stain. (**E)** Detail of enlarged hyperchromatic bizarre-shaped nuclei of hypertrophied cardiocytes. Patient III/4, H&E stain. (**F**) Immunohistochemical detection of desmin demonstrated in patient III/4. Normal cross-striation pattern was a prevalent feature with slightly disordered appearance in branched cardiocytes. (**G, H**) Electronograms showing findings in myocardium from patient III/4. Disoriented myofibrils in cardiocytes with extensive mechanical junctions are demonstrated in (G). Mitochondria are increased in number, replacing focally myofibrils (H).

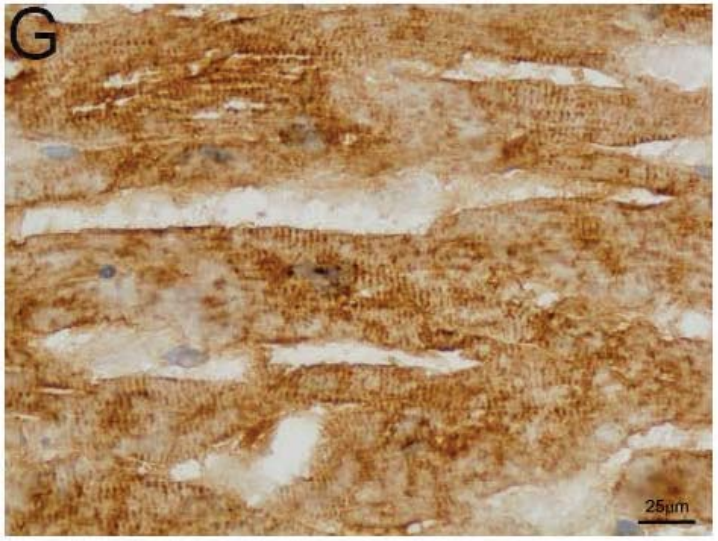


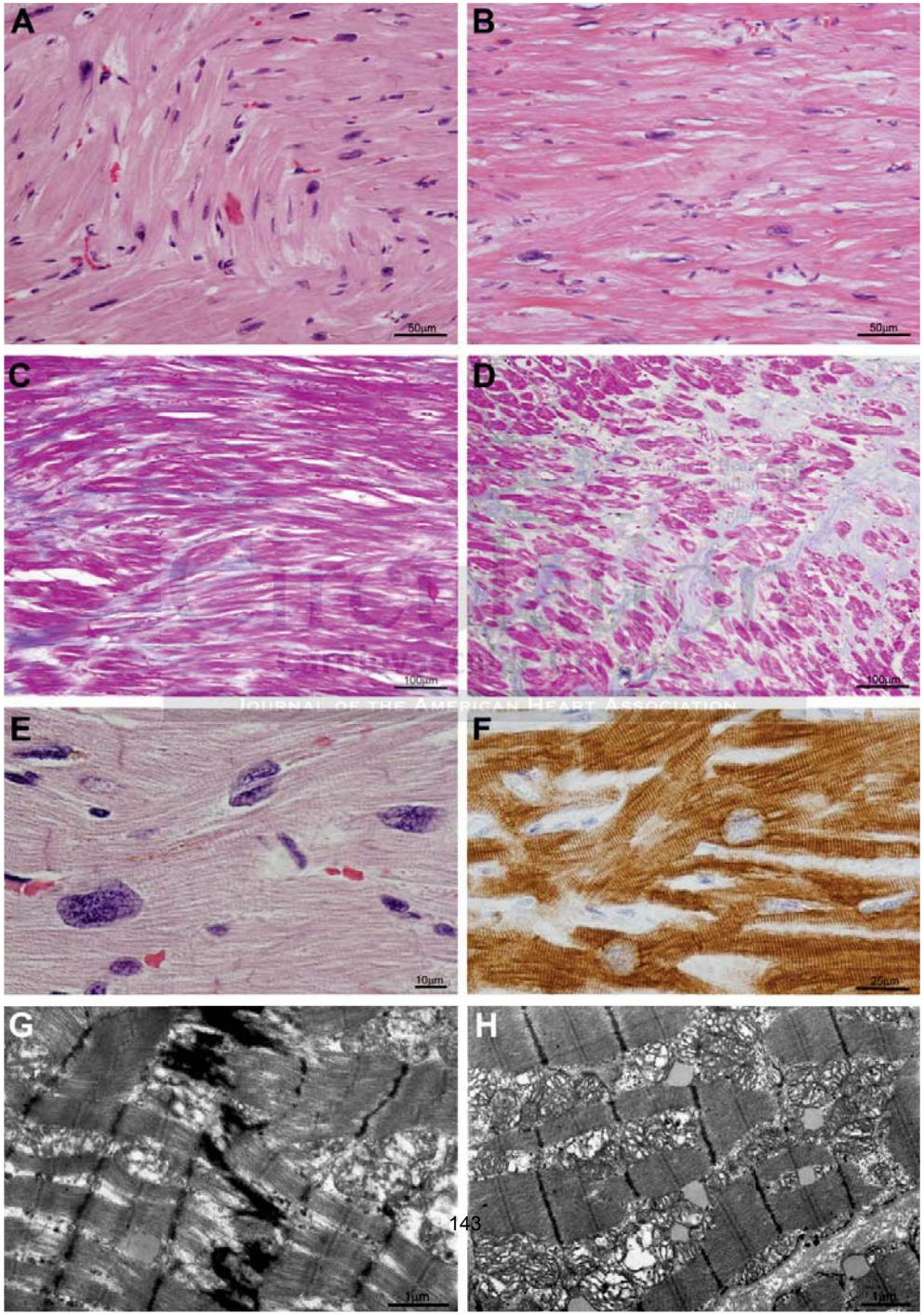



I / 2
II / 3
19 mm
III / 4
17 mm
III / 5
17 mm








A
50μm
B
50μm
C
100μm
D
100μm
E
10μm
F
25μm
G
1μm
H
1μm

Supplemental Material

Supplemental Table 1: Novel and rare genetic variants identified by exome sequencing as present in all three analyzed probands.

Chr	Position	Ref	Ob	Gene	Accession	cDNA change	AAA change	EESP 5400	SIFT	Poly Phen2
X	135290711	-	T	FHL1	NM_001159702	c.599_600 insT	p.F200fs			
1	161019039	A	G	ARHGAP30	NM_001025598	c.T1772C	p.L591P	0.0004	D	B
1	208072436	G	A	CD34	NM_001025109	c.C398T	p.T133I	0.0008	D	D
1	222801962	A	G	MIA3	NM_198551	c.A1400G	p.K467R		D	P
2	1497804	G	A	TPO	NM_175722	c.G1480A	p.G494S		D	D
2	103120120	T	A	SLC9A4	NM_001011552	c.T934A	p.L312M		T	B
2	145156524	T	C	ZEB2	NM_001171653	c.A2158G	p.I720V	0.0004	T	B
3	75786337	G	T	ZNF717	NM_001128223	c.C2437A	p.P813T			
6	26184044	C	-	HIST1H2BE	NM_003523	c.21delC	p.S7fs			
6	74019338	G	A	KHDC1	NM_001251874	c.C100T	p.Q34X			
6	116875504	A	G	FAM26D	UNKNOWN					
7	2284216	G	A	NUDT1	NM_002452	c.G7A	p.A3T	0.0001	T	B
7	6064315	C	T	EIF2AK1	NM_001134335	c.G1879A	p.G627S		D	NA
7	28844125	G	A	CREB5	NM_001011666	c.G595A	p.G199S	0.0001	T	NA
7	99786108	T	C	STAG3	NM_012447	c.T466C	p.S156P	0.0006	D	B
8	80915304	C	A	MRPS28	NM_014018	c.G325T	p.E109X		NA	NA
9	33794824	A	C	PRSS3	NM_001197097	c.A35C	p.K12T			
10	97919123	A	C	ZNF518A	UNKNOWN			0.0001		
10	102770082	T	G	PDZD7	NM_001195263	c.A2564C	p.N855T			
10	121652353	C	G	SEC23IP	NM_007190	c.C59G	p.T20S		T	P
11	4594559	-	G	C11orf40	NM_144663	c.285_286insC	p.D95fs			
11	49075621	G	A	TRIM64C	NM_001206631	c.C998T	p.A333V			
11	55110845	A	G	OR4A16	NM_001005274	c.A169G	p.M57V		D	P
12	121600323	G	A	P2RX7	NM_002562	c.G533A	p.R178Q		T	P
12	123780585	G	A	SBNO1	NM_001167856	c.C4052T	p.P1351L	0.0003	T	D
13	25672135	C	A	PABPC3	NM_030979	c.C1799A	p.S600Y		D	B
15	50215583	T	C	ATP8B4	NM_024837	c.A1751G	p.H584R		T	D
16	15690575	G	A	KIAA0430	NM_001184998	c.C5204T	p.S1735F		D	NA
17	48213398	C	T	PPP1R9B	UNKNOWN					
19	1051006	G	A	ABCA7	NM_019112	c.G2639A	p.R880Q	0.0007	D	D
19	53086185	T	G	ZNF701	NM_018260	c.T873G	p.I291M	0.0004	D	B
19	57027673	A	G	ZNF471	NM_020813	c.A63G	p.I21M		D	P

Individual variants are defined by chromosomal localization (Chr), their chromosomal position (Position), reference (Ref) and observed (Ob) alleles, gene symbol (Gene), accession number of the corresponding cDNA (Accession), resulting cDNA and amino acid changes, population frequency in exome sequence project (ESP5400) and predicted effect on protein function according to SIFT and Polyphen2 algorithms. In SIFT prediction D denotes

damaging effect and T denotes tolerating effect. In PolyPhen2 prediction B denotes benign effect, P denotes possibly damaging effect and D denotes probably damaging effect. All coordinates refer to hg19.