Lékařská fakulta Masarykovy university

Interní hematologická a onkologická klinika LF MU a FN Brno

Centrum molekulární biologie a genové terapie

Molekulární diagnostika vybraných dědičných nemocí

Habilitační práce

RNDr. Lenka Fajkusová, CSc.

Brno 2014

Poděkování

Velké díky patří všem současným i bývalým kolegům z Centra molekulární biologie a genové terapie, kteří se podíleli na diagnostice dědičných nemocí a přispěli tak k publikovaným výsledkům, úspěšnému řešení řady grantových projektů a hlavně k potvrzení diagnózy u tisícovek pacientů s nejrůznějšími typy dědičných monogenních nemocí. Dále děkuji vedení Centra molekulární biologie a genové terapie IHOK prof. Šárce Pospíšilové a prof. Jiřímu Mayerovi za možnost pracovat na problematice dědičných nemocí a volnost při výběru výzkumných i diagnostických témat.

Obsah

1 Abstrakt 4
2 Dědičné neuromuskulární nemoci 5
2.1 Facioskapulohumerální svalová dystrofie 6
2.2 Duchennova svalová dystrofie, pletencové svalové dystrofie 10
3 Dědičné kožní nemoci 71
3.1 Epidermolysis bullosa 72
3.2 Ichtyózy 75
4 Dědičné metabolické nemoci 94
5 Závěr 143
6 Použitá literatura 144

1 Abstrakt

Habilitační práce nazvaná “Molekulární diagnostika vybraných dědičných nemocí” je rozdělena do tří tématických oddílů. První část je zaměřena na molekulární podstatu a genetickou diagnostiku dědičných neuromuskulární nemocí, konkrétně na facioskapulohumerální svalovou dystrofii (FSHD), Duchennovu svalovou dystrofii a pletencovou svalovou dystrofii. Je zde popsán vývoj poznatků týkajících se molekulární podstaty FSHD, tj. asociace delece makrosatelitní repetice v subtelomerní oblasti a klinického projevu FSHD. V rámci této části je prezentováno 7 prací zaměřených na Duchennovu svalovou dystrofii, pletencovou svalovou dystrofii, kongenitální myotonii a spinální svalovou atrofii.

Druhá část habilitační práce je věnována dědičných kožním nemocem a to epidermolysis bullosa a ichtyózám. Jsou prezentovány dvě práce, první s tématikou epidermolysis bullosa dystrophica, druhá s tématikou epidermolysis bullosa simplex. Vzhledem k tomu, že problematice dědičných ichtyóz se pracoviště Centra molekulární biologie a genové terapie (CMBGT) věnuje teprve od roku 2013, publikace vztahující k tomuto onemocnění jsou v současné době připravovány. Kromě vlastní molekulárně genetické diagnostiky jsou zde stejně jako v předchozí části shrnuty i dosavadní poznatky v oblasti molekulární patologie jmenovaných chorob.

Třetí část habilitační práce je zaměřena na molekulární problematiku metabolických nemocí. V této části je po krátkém úvodu prezentováno 6 prací vztahujících se k problematice molekulárně genetické diagnostiky familiální hypercholesterolémie, hyperfenylalaninémie a kongenitální adrenální hyperplazie.

U všech prezentovaných prací je Lenka Fajkusová korespondující autor.

2 Dědičné neuromuskulární nemoci

Neuromuskulární nemoci jsou velmi rozsáhlou a různorodou skupinou chorob, která je spojena s narušením funkce svalů. V CMBGT se zbýváme molekulárně genetickou diagnostikou neuromuskulárních nemocí od roku 2002 a v současné době standardně nabízíme diagnostiku Duchennovy/Beckerovy svalové dystrofie (analýza genu *DMD*), facioskapulohumerální svalové dystrofie (stanovení delece makrosatelitní repetice D4Z4 v oblasti 4q35), pletencové svalové dystrofie (analýza genů *CAPN3*, *FKRP*, *SGCA* a *ANO5*), myotonické dystrofie 1 (stanovení expanze mikrosatelitní repetice v genu *DMPK*), myotonické dystrofie 2 (stanovení expanze mikrosatelitní repetice v genu *ZNF9*), non-dystrofické myotonie (analýza mutací v genech *CLCN1* a *SCN4A*) a spinální svalové atrofie (analýza genu *SMN1*).

Molekulární diagnostika zahrnuje v závislosti na typu hledané mutace různé metodické přístupy, jedním z nich je polymerázová řetězová reakce (PCR) ve spojení se sekvenční analýzou, tedy metoda, která umožňuje identifikaci mutací malého rozsahu. Tento metodický přístup je hojně využíván v diagnostice; dále se uplatňují různé modifikace PCR, např. *repeat-primed* PCR (pro detekci expanzí krátkých repetitivních sekvencí) nebo *multiplex ligation dependent probe amplification* (MLPA, pro detekci rozsáhlých delecí/duplikací); pro diagnostiku nemocí spojenou se změnou počtu repetitivních sekvencí se užívá i Southern blot a hybridizace s radioaktivně značenou sondou. U řady nemocí se genetické diagnostické přístupy mění v čase, příkladem může být Duchennova svalová dystrofie. Původně jsme vzhledem k velikosti genu *DMD* (79 exonů) analyzovali mRNA pomocí reverzní transkripce, PCR a PTT (*protein truncation test*), pak jsme přešli na analýzu DNA pomocí MLPA a klasické PCR-sekvenace, v současné době využíváme amplikonové sekvenování nové generace (DMD MASTR assay, Multiplicom).

Klinické manifestace neuromuskulárních nemocí se často mohou překrývat bez ohledu na genetickou příčinu nemoci a určit vhodný gen pro molekulárně genetickou diagnostiku na základě klinického obrazu pacienta (často i ve spojení s výsledkem patologické analýzy svalové tkáně) může být problematické. Z tohoto důvodu jsme zavedli nový diagnostický přístup, který umožní rychlou a relativně ekonomickou analýzu genů asociovaných s neuromuskulárními nemocemi: *sequence capture and targeted resequencing* (SeqCap-TR). SeqCap je proces, který zachytí a obohatí vybrané oblasti genomové DNA v jednom

kroku (místo nutnosti stovek až tisíců jednotlivých PCR) a ve spojení s technikami sekvenování nové generace umožní identifikaci kauzální mutace/mutací. Námi zavedený metodický přístup využívající SeqCap EZ Choice Library (Roche NimbleGene) a sekvenování na přístroji GS Junior System (Roche) nebo MiSeq (Illumina) umožňuje analýzu exonů a přilehlých intronových sekvencí 42 genů (1020 exonů, 280 kb) asociovaných se svalovými dystrofiemi, kongenitálními svalovými dystrofiemi, kongenitálními myopatiemi, distálními myopatiemi a dalšími myopatiemi (dle http://www.musclegenetable.org/). Seznam vybraných nemocí a genů zahrnuje Duchennovu svalovou dystrofii (*DMD*), Emery-Dreifussovu svalovou dystrofii (*EMD, FHL1, LMNA*), pletencové svalové dystrofie (*MYOT, LMNA, CAV3, CAPN3, DYSF, SGCG, SGCA, SGCB, SGCD, TCAP, TRIM32, FKRP, TTN, POMT1, ANO5, FKTN, POMT2, POMGNT1*); geny spojené s kongenitálními svalovými dystrofiemi (*LAMA2, LARGE, SEPN1, COL6A1, COL6A2, COL6A3, ITGA7, DNM2*); kongenitálními myopatiemi, distálními myopatiemi a dalšími myopatiemi (*NEB, TPM3, ACTA1, TPM2, TNNT1, CFL2, RYR, MTM1, BIN, CRYAB, DES, LAMP2, PABPN1*). Uvedený metodický přístup byl již aplikován u 70 pacientů a u poloviny z nich jsme identifikovali kauzální mutace spojené s onemocněním. U části vzorků jsme detekovali mutace, ale ještě ověřujeme jejich kauzalitu analýzou DNA rodinných příslušníků, vyhodnocením vlivu mutace na strukturu proteinu pomocí *in silico* metod a zastoupením zjištěné genetické varianty v populaci.

V oblasti diagnostiky neuromuskulárních nemocí spolupracujeme nejen s klinikami v rámci Fakultní nemocnice Brno (Klinika dětské neurologie, Neurologická klinika, Oddělení lékařské genetiky, Patologicko-anatomický ústav), ale intenzivní spolupráce je i s Fakultní nemocnicí Motol, Všeobecnou fakultní nemocnicí Praha, Fakultní nemocnicí Ostrava, Fakultní nemocnicí Olomouc a Thomayerovou nemocnicí.

2.1 Facioskapulohumerální svalová dystrofie

V řadě případů výzkumu dědičných nemocí můžeme sledovat zajímavý vývoj poznatků a jejich interpretací, které se dramaticky mění v čase. Tím nám poskytují nejen možnost nahlížet na problematiku daného onemocnění z řady různých hledisek, ale současně upozorňují na relativitu současného chápání molekulární podstaty dané choroby. Velmi

dobrým příkladem v tomto směru je vývoj poznání molekulárních příčin facioskapulohumerální svalové dystrofie (FSHD).

FSHD je autozomálně dominantní onemocnění a je třetí nejběžnější dědičnou svalovou dystrofií [1]. Prvním poznatkem směřujícím k objasnění molekulární podstaty FSHD byla vazba na kontrakci (snížení počtu) subtelomerických tandemových repetic D4Z4 na chromosomu 4q35 pod prahovou úroveň 11 kopií [2]. Na tomto zjištění bylo možno založit diagnostický postup, v němž lze pomocí restrikčního štěpení, pulsní gelové elektroforézy a následné Southernovy hybridizace změřit počet kopií repetic D4Z4 v 4q35 lokusu. Tento počet je 11-100 u standardní DNA a u pacientů 1-10, velikost jedné repetice D4Z4 je 3,3 kb. Nicméně, asi u 5% FSHD pacientů (tzv. FSHD2) se zkrácení repetic D4Z4 nevyskytuje [3].

Později bylo zjištěno, že součástí repetice D4Z4 je kandidátní gen, který by mohl kódovat protein se dvěma homeodoménami (*DUX4, double homeobox 4*) [4], ale jeho souvislost s patogenezí FSHD nebyla objasněna. Sekvence homologní k D4Z4 byly identifikovány také na krátkých ramenech několika dalších, především akrocentrických chromozómů [5]. FSHD je však spojena jedině s kontrakcí repetic D4Z4 na chromozómu 4, zatímco kontrakce D4Z4 na chromozómu 10 s FSHD spojena není. Monosomie subtelomerické oblasti 4q35, čili haploinsuficience v oblasti obsahující D4Z4, rovněž FSHD nevyvolává [6]. Stejně tak hybridní repetice vznikající subtelomerickou výměnou mezi chromozomy 4 a 10 (pokud celkový počet repetic na chromozomu 4 neklesne pod 11), nebo repetice translokované z chromozómu 4 na jiné chromozómy nevedou k onemocnění [7],[8]. Závislost delece D4Z4 na specifickém okolí chromozómu 4q35 ukazovala na epigenetickou podstatu FSHD. Současně se objevily hypotézy o tom, že repetice D4Z4 mají úlohu nekódujících regulačních sekvencí a jejich kontrakce vede k zvýšení exprese genů v jejich okolí až do vzdálenosti několika megabází [9],[10]. Hypotéza o příčinné úloze této tzv. poziční variability exprese (*position effect variegation*) v patogenezi FSHD nebyla jednoznačně prokázána, vznikla však řada studií, které na tuto atraktivní epigenetickou interpretaci navázaly. Mezi nimi je možno uvést např. práci [11], v níž autoři ukázali pomocí fluorescenční hybridizace *in situ*, že standardní subtelomerický lokus 4q35 je v buněčném jádře během interfáze lokalizován specificky do heterochromatinových kompartmentů na periférii jádra. Ukázalo se však také, že tato lokalizace je velmi stabilní a zůstává zachována i v myoblastech pacientů s FSHD. Autoři na základě svých výsledků spekulovali o možné izolátorové funkci repetic D4Z4, které

chrání proximální geny v sousedním euchromatinu před ještě silnější represí heterochromatinem distálních β-satelitních sekvencí [8] a tvoří hraniční pásmo mezi těmito epigeneticky rozdílnými kompartmenty. Podporou pro takové spekulace byla i studie, která identifikovala dvě alelické varianty uspořádání subtelomery 4q: 4qA nesoucí v sousedství D4Z4 β-satelitní sekvence, a 4qB bez těchto sekvencí [12]. Ačkoli obě varianty jsou v populaci zastoupeny podobně, FSHD je spojena výlučně se zkrácením D4Z4 v alele 4qA, zatímco stejné zkrácení u alely 4qB k FSHD nevede.

Další podporu epigenetických efektů jako příčinných faktorů onemocnění přinesly studie methylačního stavu sekvencí D4Z4. Ty jsou u všech FSHD pacientů (jak u těch, kteří vykazují kontrakci repetic D4Z4, tak i u tzv. FSHD2 bez kontrakce repetic D4Z4) hypomethylovány, tj. jeví významné snížení úrovně methylace cytosinů v CpG dinukleotidech, na které jsou D4Z4 repetice poměrně bohaté [13]. Zjištění hypomethylace, jako faktoru obvykle spojeného s aktivací exprese, inspirovalo analýzy dalších chromatinových značek, zejména methylací a acetylací histonů. Tyto studie ukázaly charakteristiky 4q35 chromatinu odpovídající spíše inaktivnímu euchromatinu (případně fakultativnímu heterochromatinu) než konstitutivnímu heterochromatinu [14],[15]. Pozornost výzkumníků se proto zaměřila na možnou roli částečné delece D4Z4 při aktivaci genů *DUX4* přítomných v elementech D4Z4, jejichž exprese se zdá být normálně omezena na stadium časné embryogeneze [16], zatímco mimo toto stadium byla zjištěna u FSHD myoblastů, kde má pro-apoptotické účinky [17],[18]. Vzhledem k tomu, že protein DUX4 se chová jako transkripční aktivátor [17], může hrát významnou roli při aktivaci exprese dalších genů. Byly detekovány dva typy transkriptů *DUX4*, z nichž první je přepisován z vnitřních repetic, zatímco druhý je přepisován z nejvíce distální D4Z4 jednotky. S touto jednotkou sousedí pLAM sekvence, která transkriptu poskytuje polyadenylační signál a podílí se tak na stabilitě vzniklé mRNA a vzniku proteinu DUX4. Vzhledem k tomu, že pLAM sekvence se vyskytuje pouze na chromozomech 4qA, tedy chromozomech asociovaných s FSHD, je její přítomnost zásadní v patofyziologii FSHD [17].

Deregulace exprese genů sousedících s D4Z4 v 4q35 lokusu (*FRG1, FRG2* a *ANT1*) se zdála snad nejpravděpodobnější příčinou onemocnění, zvýšenou expresi těchto genů ve svalových buňkách FSHD pacientů se ale nepodařilo reprodukovatelně prokázat a ani zvýšená exprese navozená u myší nevedla k projevům svalové dystrofie [19],[20],[21],[22],[23]. Spekulovalo se rovněž o možné deregulaci dalších genů v patogenezi FSHD. Podkladem k tomu je fakt, že každá z jednotek D4Z4 obsahuje 27 pb

dlouhý vazebný element DBE (D4Z4-*binding element*), k němuž se váže komplex proteinů YY1, HMGB2 a nukleolin [10]. Vzhledem k tomu, že se uvedené proteiny účastní řady dalších procesů a interakcí (např. remodelace struktury chromatinu, transkripční aktivace nebo represe), nerovnováha vzniklá úbytkem jejich vazebných míst může deregulovat i geny lokalizované zcela mimo 4q35 lokus. Jakkoli je tato hypotéza zajímavá, její testování představuje velmi komplexní problém a zatím neexistují experimentální důkazy, které by ji podporovaly [8].

Další zajímavý posun ve zkoumání příčin FSHD přineslo zjištění, že repetice D4Z4 jsou cílovým místem pro vazbu PcG (*Polycomb Group*) proteinů, které obecně umlčují expresi genů a udržují chromatin v transkripčně inaktivním stavu. Tím představují významnou složku epigenetické paměti u diferencovaných buněk. U FSHD pacientů je snížena vazba PcG na zbylé repetice D4Z4, stejně jako výskyt dalších výše zmíněných epigenetických značek typických pro umlčený chromatin (methylace cytosinu, modifikace histonů – např. methylace H3K9me3, deacetylace) [24],[25]. Ztráta represivní struktury chromatinu pak aktivuje transkripci dlouhé nekódující RNA (lncRNA), nazvané DBE-T. Počátek transkripce DBE-T je v proximální oblasti před D4Z4 úsekem. Tato RNA zůstává asociovaná s chromatinem, k němuž přitahuje protein Ash1L (protein skupiny *Trithorax*, TrxG), jehož funkce je opačná než u PcG. Vazba tohoto proteinu vede k aktivační modifikaci ve FSHD lokusu. Tím může dojít k aktivaci nejen samotných *DUX4* genů, ale i dalších genů v 4q35 lokusu. Změna v počtu kopií D4Z4 tak může fungovat jako epigenetický přepínač mezi Polycomb a Trithorax drahami.

Stále ovšem zbývalo vyřešit otázku, jak dochází k aktivaci exprese genů *DUX4* u FSHD2, kde ke kontrakci repetic nedochází, ačkoli zde podobně jako u FSHD1 dochází k hypomethylaci repetic D4Z4 (viz výše). Ukázalo se, že v těchto případech s hypomethylací D4Z4 kosegreguje mutace v genu pro jeden z proteinů SMC komplexů (*Structural maintenance of chromosomes*), SMCHD1. Snížení hladiny SCHMD1 v kosterním svalstvu v důsledku mutace v jedné z alel *SMCHD1* má za následek navození exprese *DUX4* i bez kontrakce D4Z4 repetic. SMCHD1 tak funguje jako epigenetický modifikátor metastabilní epialely D4Z4 a jako příčinný faktor FSHD2 [26].

Lze tedy uzavřít, že aktivace genu *DUX4* je základní příčinou patogeneze FSHD [27]. Obsah tohoto proteinu je zvýšen v myoblastech a svalových vláknech FSHD pacientů a narušuje buněčnou diferenciaci. Ke všem uvedeným genetickým a epigenetickým

faktorům, které se na této aktivaci podílejí nebo mohou podílet, lze přidat nejméně jeden další, a sice tzv. poziční efekt telomer (*telomere position effect*, TPE). Podobně jak oblast D4Z4 působí jako izolátor a brání aktivaci proximálních genů i vlastních *DUX4* genů, přičemž jejím zkrácením se tato izolátorová a represívní funkce ztrácí, samotná oblast D4Z4 je pod represívním pozičním vlivem svého distálního souseda – telomer. Jejich efekt záleží obecně na délce telomer a na vzdálenosti od nich, jak bylo zjištěno na řadě modelových organismů od kvasinek po savce. Nedávná studie ukázala, že exprese *DUX4* je skutečně nepřímo úměrná délce telomer [28]. Zkracováním telomer se jejich umlčovací efekt zmírňuje a exprese *DUX4* vzrůstá až desetinásobně. Je tedy možné, že FSHD je vůbec první známou lidskou nemocí u níž TPE přispívá k jejímu fenotypu v závislosti na věku.

V CMBGT se molekulárně genetickou diagnostikou FSHD zabýváme od roku 2005. Celkem jsme vyšetřili 280 pacientů a u 132 (47%) z nich jsme pomocí restrikčního štěpení, pulsní gelové elektroforézy a Southernovy hybridizace onemocnění potvrdili. V současné době zavádíme i analýzu genu *SMCHD1* a stanovení metylačního stavu repeticí D4Z4.

2.2 Duchennova svalová dystrofie, pletencové svalové dystrofie

Na rozdíl od FSHD bylo stanovení molekulární podstaty Duchennovy/Beckerovy svalové dystrofie (DMD/BMD) relativně přímočarou záležitostí. DMD a její mírnější alelická varianta BMD jsou výsledkem mutací genu, který kóduje protein dytrofin (*DMD*). Dystrofin je cytoplazmatický protein, který je součástí membránově vázaného proteinového komplexu, který propojuje cytoskelet svalového vlákna s extrabuněčnou matrix a umožňuje tak ochranu svalové buňky před mechanickým stresem během svalové kontrakce a dilatace [29]. Gen *DMD* je lokalizován na chromozomu Xp21. Asi 60% mutací genu *DMD* je tvořeno delecemi, 5% duplikacemi a 35% mutacemi malého rozsahu [30].

Obecně platí, že mutace způsobující předčasné ukončení translace jsou spojeny s fenotypem DMD a mutace mající za následek vznik proteinu s vnitřní delecí/duplikací (C-konec zůstává zachován) jsou spojeny s mírnějším fenotypem BMD [31]. V diagnostické praxi je však možné se setkat s případy, kdy fenotyp pacienta neodpovídá nalezenému genotypu a v takovém případě je pak nutné hledat vysvětlení zjištěné neshody. Z naší

laboratorní praxe, která zahrnuje 85 pacientů s identifikovanou mutací malého rozsahu a provedenými klinicko-geneticko-patologickými korelacemi, je možné uvést několik případů demonstrujících uvedenou neshodu. Za všechny případ pacienta s identifikovanou *nonsense* mutací, kdy očekávaným fenotypem byla DMD, ale pacient měl klinický projev odpovídající BMD. Analýza DNA byla doplněna analýzou mRNA a bylo zjištěno, že ve svalových buňkách pacienta jsou dva typy mRNA genu *DMD*: mRNA nesoucí identifikovanou *nonsens*e mutaci a mRNA obsahující deleci exonu, ve kterém se tato mutace nachází. *In silico* analýza ukázala, že v místě mutace je v případě standardní DNA lokalizována sekvence zvaná *exon splicing enhancer* (místo vazby sestřihových faktorů), mutace má pak za následek změnu chování této oblasti v průběhu sestřihu mRNA, tj. nezařazení exonu do výsledné mRNA. Vzhledem k tomu, že delece tohoto exonu neruší čtecí rámec translace, vzniká protein s vnitřní delecí. Svalová vlákna uvedeného pacienta obsahují tedy dva typy *DMD* transkriptů (transkript s *nonsense* mutací a transkript s delecí) a teoreticky i dva typy *DMD* proteinů (zkrácený protein bez C-konce, ten ale pravděpodobně nevzniká, protože mRNA je rozložena procesem *nonsense mediated mRNA decay*, a protein s vnitřní delecí, který je částečně funkční). Dalším příkladem nutnosti spojení analýzy DNA a mRNA jsou pacienti, u nichž pomocí MLPA a sekvenční analýzy exonů a přilehlých intronových oblastí nejsou zjištěny mutace. Pokud ale klinické projevy a výsledky patologické analýzy svalové tkáně svědčí pro DMD,

provádí se analýza mRNA, která může identifikovat inzerce intronových částí genu *DMD* do výsledné mRNA. Tyto inzerce většinou vznikají mutací, která hluboko uvnitř intronu vytvoří sestřihové místo a vnese část intronové sekvence tzv. pseudoexon do výsledné mRNA [publikace *Neuromuscul Disord. 2009 Nov;19(11); Neuromuscul Disord. 2001 Mar;11(2)*].

Pletencové svalové dystrofie (*limb girdle muscular dystrophies*, LGMD) jsou spojeny s mutacemi v nejméně 8 genech s autosomálně dominantní dědičností (LGMD1A-1H) a 16 genech s autosomálně recesivní dědičností (LGMD2A-Q). LGMD mohou mít podobný klinický obraz jako DMD nebo i FSHD. Základním diferenciálně diagnostickým testem je pohlaví, hodnota svalové kreatinkinázy, imunohistochemické vyšetření svalových proteinů, výskyt onemocnění v rodině; konečné potvrzení diagnózy umožňuje ale jenom molekulárně genetická diagnostika [32],[33],[34]. Stanovit gen, který se bude v rámci LGMD1 nebo LGMD2 analyzovat, je ale obtížné, protože klinické projevy většiny typů LGMD jsou podobné a imunohistochemická analýza se provádí jen u vybraných proteinů,

navíc jejich přítomnost nebo deficit nemusí souviset s výskytem kauzálních mutací v kódujícím genu. Často se volí postup, kdy se na základě literárních dat vyberou geny, u kterých je výskyt mutací v dané populaci nejpravděpodobnější. V CMBGT se zabýváme molekulárně genetickou diagnostikou LGMD2 od roku 2004, kdy jsme začali analyzovat gen *CAPN3* (LGMD2A), a v průběhu času jsme pak přidali geny *FKRP* (LGMD2I), *SGCA* (LGMD2D), *ANO*5 (LGMD2L). Analýza uvedených genů byla aplikována v souboru 218 nepříbuzných pacientů a četnost výskytu jednotlivých forem LGMD2 byla 32.6% (71 pacientů) pro LGMD2A, 4.1% (9 pacientů) pro LGMD2I, 2.8% (6 pacientů) pro LGMD2D a 1.4% (3 pacientů) pro LGMD2L. V současné době jsme ale již schopni díky technologiím SeqCap-TR analyzovat většinu genů asociovaných s LGMD2 i LGMD1 [publikace *Neuromuscul Disord. 2007 Feb;17(2); Neuromuscul Disord. 2004 Oct;14(10); Autosomal recessive limb-girdle muscular dystrophies in the Czech Republic (submitted)*].

V případě LGMD je zajímavé srovnání podobnosti klinických projevů na jedné straně a různosti funkce proteinů, jejichž geny jsou mutovány, na straně druhé. Jako příklad je možné uvést kalpain-3 (*CAPN3*), což je svalově specifická proteáza specificky štěpící své substráty [35]. Kalpain-3 se ve svalových buňkách vyskytuje volně nebo ukotven v cytoskeletu sarkomery, jeho úkolem je degradovat poškozené proteiny a přispívat tak k efektivnímu fungování svalu. Dalšími enzymy souvisejícími LGMD jsou glykosyltransferáza FKRP (*fukutin related protein*) [36] nebo ubikvitinligáza TRIM32 [37]. Naproti tomu s LGMD je spojena i řada strukturních proteinů účastnících se tvorby membránově vázaného komplexu s dystrofïnem (sarkoglykany alfa-delta, *SGCA-SGC*D) [38], ovlivňujících procesy související s reparací svalové membrány (dysferlin, *DYSF*) [39], podílejících se na kontrakci a dilataci svalového vlákna (titin, *TTN*) [40] nebo vytvářejících transmembránové iontové kanály (anoktamin 5, *ANO5*) [41].

Molekulárně genetická diagnostika a korelace výsledného genotypu s klinickými projevy pacienta a výsledky analýzy svalové tkáně je komplexní proces. Tento proces z pohledu molekulárního biologa musí zahrnovat analýzu typů identifikovaných mutací (*missense, nonsense, splicing*, delece/duplikace aj.); u *missense* mutací analýzu fyzikálních a chemických vlastností aminokyselin původních a mutantních a vyhodnocení mutací pomocí *in silico* přístupů; důležité je rovněž vyhodnocení mutací v souvislosti s lokalizací dané aminokyseliny ve struktuře proteinu resp. v jeho funkčních doménách. Vyhodnocení mutace znamená současně i analýzu publikovaných studií a databází souvisejících s daným genem. V některých případech je pro konečné potvrzení kauzality mutace nutné provést

její funkční analýzu, tj. zjistit jestli protein nesoucí identifikovanou mutaci plní svoji funkci. Tento přístup založený na expresi mutantních proteinů, analýze jejich buněčné lokalizace a funkce zavádíme v současné době v souvislosti s genem kódujícím chloridový kanál typu 1 (*CLCN1*), jehož mutantní formy jsou zodpovědné za onemocnění Thomsenova/Beckerova kongenitální myotonie, a v souvislosti s genem pro LDL receptor (*LDLR*) a tedy familiální hypercholesterolémií.

Publikace s tématikou neuromuskulárních nemocí

1) Autosomal recessive limb-girdle muscular dystrophies in the Czech Republic. Stehlíková K, Skálová D, Zídková J, Mrázová L, Vondráček P, Mazanec R, Voháňka S, Haberlová J, Hermanová M, Zámečník J, Souček O, Ošlejšková H, Dvořáčková N, Solařová P, Fajkusová L. Submitted *(L. Fajkusová jako korespondující autor)*

2) Point mutations in Czech DMD/BMD patients and their phenotypic outcome. Sedlácková J, Vondrácek P, Hermanová M, Zámecník J, Hrubá Z, Haberlová J, Kraus J, Maríková T, Hedvicáková P, Vohánka S, Fajkusová L. Neuromuscul Disord. 2009 Nov;19(11):749-53. *(L. Fajkusová jako korespondující autor)*

3) Novel dystrophin mutations revealed by analysis of dystrophin mRNA: alternative splicing suppresses the phenotypic effect of a nonsense mutation. Fajkusová L, Lukás Z, Tvrdíková M, Kuhrová V, Hájek J, Fajkus J. Neuromuscul Disord. 2001 Mar;11(2):133-8. *(L. Fajkusová jako 1. a korespondující autor)*

4) Quantitative analysis of CAPN3 transcripts in LGMD2A patients: involvement of nonsense-mediated mRNA decay. Stehlíková K, Zapletalová E, Sedlácková J, Hermanová M, Vondrácek P, Maríková T, Mazanec R, Zámecník J, Vohánka S, Fajkus J, Fajkusová L. Neuromuscul Disord. 2007 Feb;17(2):143-7. *(L. Fajkusová jako 1. a korespondující autor)*

5) Mutations in Czech LGMD2A patients revealed by analysis of calpain3 mRNA and their phenotypic outcome. Chrobáková T, Hermanová M, Kroupová I, Vondrácek P, Maríková T, Mazanec R, Zámecník J, Stanek J, Havlová M, Fajkusová L. Neuromuscul Disord. 2004 Oct;14(10):659-65. *(L. Fajkusová jako korespondující autor)*

6) CLCN1 Mutations in Czech Patients with Myotonia Congenita, In Silico Analysis of Novel and Known Mutations in the Human Dimeric Skeletal Muscle Chloride Channel. Skálová D, Zídková J, Voháňka S, Mazanec R, Mušová Z, Vondráček P, Mrázová L, Kraus J, Réblová K, Fajkusová L. PLoS One. 2013 Dec 11;8(12):e82549. *(L. Fajkusová jako korespondující autor)*

7) Analysis of point mutations in the SMN1 gene in SMA patients bearing a single SMN1 copy.Zapletalová E, Hedvicáková P, Kozák L, Vondrácek P, Gaillyová R, Maríková T,

Kalina Z, Jüttnerová V, Fajkus J, Fajkusová L. Neuromuscul Disord. 2007 Jun;17(6):476-81. *(L. Fajkusová jako korespondující autor)*

Autosomal recessive limb-girdle muscular dystrophies in the Czech Republic

Kristýna Stehlíková[a,b], Daniela Skálová[a,b], Jana Zídková[a,b], Lenka Mrázová[c], Petr Vondráček[c], Radim Mazanec[d], Stanislav Voháňka[e], Jana Haberlová[f], Markéta Hermanová[g], Josef Zámečník[h], Ondřej Souček[i], Hana Ošlejšková[c], Nina Dvořáčková[j], Pavla Solařová[k], Lenka Fajkusová[a,b,*]

[a] Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Černopolní 9, 613 00 Brno, Czech Republic
[b] Central European Institute of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic
[c] Department of Child Neurology, University Hospital Brno, Černopolní 9, 613 00 Brno, Czech Republic
[d] Department of Neurology, Second Faculty of Medicine, Charles University and University Hospital Motol, V Úvalu 84, 150 06 Prague, Czech Republic
[e] Department of Neurology, University Hospital Brno, Jihlavská 20, 625 00 Brno, Czech Republic
[f] Department of Child Neurology, Second Faculty of Medicine, Charles University and University Hospital Motol, V Úvalu 84, 150 06 Prague, Czech Republic
[g] First Department of Pathological Anatomy, Faculty of Medicine, Masaryk University and St. Anne´s University Hospital, Pekařská 53, 656 91 Brno, Czech Republic
[h] Department of Pathology and Molecular Medicine, Second Faculty of Medicine, Charles University and University Hospital Motol, V Úvala 84, 150 06 Prague, Czech Republic
[i] Institute of Pathology, University Hospital Brno, Jihlavská 20, 625 00 Brno, Czech Republic
[j] Departmemt of Medical Genetics, University Hospital Ostrava, 17. Listopadu 1790, 708 52 Ostrava, Czech Republic
[k] Department of Medical Genetics, University Hospital Hradec Králové, Sokolská 581, 500 05 Hradec Králové, Ostrava

* Corresponding author:
Lenka Fajkusová
Central European Institute of Technology

Masaryk University
Kamenice 5
CZ-62500 Brno, Czech Republic
e-mail: lenka.fajkus@gmail.com
Phone: +420-532234625
Fax: +420-532234623

Autosomal recessive limb-girdle muscular dystrophies in the Czech Republic

Abstract

Introduction: Limb-girdle muscular dystrophies (LGMD) include a number of disorders with heterogeneous etiology. In this study, we determined the frequency of LGMD subtypes within a cohort of Czech LGMD2 patients using mutational analysis of the *CAPN3*, *FKRP*, *SGCA*, and *ANO5* genes.
Results: Mutations of the *CAPN3* gene are the most common cause of LGMD2 and in the set of 218 Czech probands with suspicion of LGMD2, mutations in this gene were identified in 71 patients. Totally, we detected 37 different mutations of which 12 have been described only in Czech LGMD2A patients. The mutation c.550delA is the most frequent among our LGMD2A probands and was detected in 47.1% of *CAPN3* mutant alleles. The frequency of particular forms of LGMD2 was 32.6% for LGMD2A (71 probands), 4.1% for LGMD2I (9 probands), 2.8% for LGMD2D (6 probands), and 1.4% for LGMD2L (3 probands).
Further, we present first results of a new approach established in the Czech Republic for diagnostics of neuromuscular diseases: sequence capture and targeted resequencing. Using this approach, we identified patients with mutations in the *DYSF* and *SGCB* genes.
Conclusions: We characterised a cohort of Czech LGMD2 patients on the basis of mutation analysis of genes associated with the most common forms of LGMD2 in the European population and subsequently compared the occurrence of particular forms of LGMD2 among countries on the basis of our results and published studies.

Background

Limb-girdle muscular dystrophies (LGMD) are characterised by wide clinical and genetic heterogeneity. Considering this, achieving a precise diagnosis can be difficult and requires a comprehensive clinical and laboratory approach. LGMD are known to be associated with mutation/mutations in at least 8 genes with autosomal dominant inheritance (LGMD1A-1H) and 16 genes with autosomal recessive inheritance (LGMD2A-Q). It is helpful to take into account the geographical and ethnic origins of patients in differential diagnosis, since the relative local frequency of the different forms of LGMD varies considerably [1],[2],[3],[4]. The clinical course of LGMD ranges from severe to milder forms with different age of onset and progression even within the same family [5]. Diagnosis of LGMD relies on a combination of clinical findings, results of histopathological examination of muscle biopsy (including the analysis of muscle proteins using immunohistochemistry and/or immunoblotting), followed by DNA sequencing to identify the primary mutation, which is essential for the provision of genetic and prognostic counselling.

LGMD2A, caused by mutations in the calpain-3 gene (*CAPN3*), is probably the most frequent form of LGMD2 in Europe. The *CAPN3* gene encodes a muscle-specific member of the family of Ca^{2+}-activated neutral proteases that is important for muscle remodelling [6],[7],[8]. In calpainopathy, a marked clinical heterogeneity has been observed. While null type gene mutations are usually associated with absence of calpain-3 protein in muscle and a severe phenotype, missense-type mutations are associated with an unpredictability of their effect at both the protein and the phenotype levels [9],[10],[11].

Mutations in the fukutin-related protein gene (*FKRP*) result in two distinct allelic diseases: severe congenital muscular dystrophy and LGMD2I [12]. The missense mutation c.826C>A, p. Leu276Ile is particularly common in LGMD2I patients and has been reported to confer a relatively mild phenotype when present in the homozygous state [13],[14],[15]. The *FKRP* gene encodes a putative Golgi-resident glycosyltransferase that is involved in posttranslational glycosylation of the α-dystroglycan.

LGMD2C-2F are due to mutations in the genes encoding the components of the sarcoglycan (SG) complex. The SG complex, composed of 5 glycoproteins (α-, β-, γ-, δ-, ε-SG), is a member of the dystrophin-associated glycoprotein (DAG) complex localised to the sarcolemma of muscle fibres, which acts as a link between the extracellular matrix and

the cytoskeleton, confers structural stability, and protects the sarcolemma from mechanical stress developed during muscle contraction. The clinical phenotype of sarcoglycanopathies ranges from a severe Duchenne-like dystrophy to a relatively mild LGMD [16]. Although the relative frequency of mutations in the different sarcoglycan genes varies from population to population, α-sarcoglycanopathy (LGMD2D) appears to be the most frequent [17].

Recently, recessive mutations in the anoctamin 5 gene (*ANO5*) were identified as a cause of LGMD2L and non-dysferlin Miyoshi myopathy [18],[19]. *ANO5* encodes a member of the anoctamin family of proteins which contains eight transmembrane domains. Dominant mutations in the *ANO5* gene are associated with the skeletal disorder gnathodiaphyseal dysplasia [20]. While the role of ANO5 is unknown, ANO1 and ANO2, which share significant sequence homology with ANO5, are known to be calcium-activated chloride channels [21],[22],[23]. In the study by [24], analysis of *ANO5* was performed in a group of 59 British and German probands, and the mutation c.191dupA was found in 15 patients, homozygously in 11 and in compound heterozygosity with another *ANO5* variant in the rest. An intragenic SNP and an extragenic microsatellite marker were in linkage disequilibrium with the mutation, suggesting a founder effect in the populations analysed.

In this study, we determined the frequency of LGMD subtypes within a cohort of Czech LGMD2 patients using mutation analysis of the *CAPN3*, *FKRP*, *SGCA*, and *ANO5* genes. We present also two patients with mutations in the gene encoding dysferlin (*DYSF*) and a patient with mutations in the gene encoding β-sarcoglycan (*SGCB*). These mutations were identified using Sequence capture and targeted resequencing (SeqCap-TR). SeqCap is a process for the capture and enrichment of selected genomic regions from full genomic DNA in a single step which, in association with targeted resequencing, allows to focus on genomic regions of interest to discover causative mutations.

Patients and methods

Patients

For analysis of genes associated with LGMD2, patients´ blood or DNA were sent to the Centre of Molecular Biology and Gene Therapy, University Hospital Brno (as the only one laboratory performing LGMD genetic testing in the Czech Republic) from Departments of

Neurology and Medical Genetics of University Hospitals located in Brno, Prague, Ostrava, Hradec Králové. Patients were included into the study on the basis of fulfilment of the following criteria: 1) a clinical phenotype characterized by progressive muscle weakness affecting primarily or predominantly pelvic and/or shoulder girdle muscles and 2) dystrophic or myopathic features at muscle biopsy (if performed). Pathological analyses of muscle tissues were performed at the Institute of Pathology, University Hospital Brno, Brno and the Department of Pathology and Molecular Medicine, University Hospital Motol, Prague. Detailed clinical information was requested retrospectively on the basis of positive results of DNA analysis. Diagnoses as Duchenne muscular dystrophy, spinal muscular atrophy type III, facioscapulohumeral muscular dystrophy were excluded using DNA analysis, late-onset Pompe disease using the enzyme assay. All analyses were done in accord with the standards of the Committee on Human Experimentation of the University Hospital Brno.

Analysis of muscle tissue

Muscle tissues were routinely examined by conventional histopathological methods. Muscle protein analyses were performed using immunohistochemistry and immunoblotting as described in our previous studies [25],[26],[27].

DNA analysis

Genomic DNA was extracted from peripheral blood leukocytes by the standard salting-out method and amplified. Primers for amplification of all exons and adjacent intron sequences are available on request, as well as the conditions of particular PCRs. PCR products were sequenced directly using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and analysed on the ABI PRISM 310 or ABI 3130xl Genetic Analyzer (Applied Biosystems). The resulting sequences were compared with the reference sequences of *CAPN3* (NCBI NG_008660.1), *FKRP* (NCBI NG_008898.1), *SGCA* (NCBI NG_008889.1), *ANO5* (NCBI NG_015844.1), *DYSF* (NCBI NG_008694.1), and *SGCB* (NCBI NG_008891.1). To find out whether the detected sequence variations were described previously, we used the literature and the Leiden Muscular Dystrophy Pages (LMDP, http://www.dmd.nl/nmdb/home.php?action=switch_db). All novel missense mutations were screened in a control panel consisting of DNA from 150 healthy Czech

individuals. *In silico* approaches, PolyPhen (http://genetics.bwh.harvard.edu/pph2/index.shtml), SIFT (http://sift.jcvi.org/www/SIFT_BLink_submit.html), and PON-P (http://bioinf.uta.fi/PON-P/), were used for predicting effects of novel missense mutations on the function and the structure of proteins. For prediction of effects of mutations on splicing of pre-mRNA, the *in silico* tools NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) and SpliceView (http://zeus2.itb.cnr.it/~webgene/wwwspliceview.html) were used. The nomenclature of mutations was implemented according to the current HGVS recommendations (http://www.hgvs.org/mutnomen/). In patients with one mutation in the *CAPN3* gene, detection of deletions/duplications was performed using multiplex ligation-dependent probe amplification (MLPA). We used the SALSA MLPA P176 CAPN3 probemix kit (MRC Holland) according to the manufacturer's guidelines. Some *CAPN3* mutations were described also on the mRNA level in our previous studies [25],[26],[27].
We started with LGMD2 DNA analysis in 2002 when sequencing analysis of the *CAPN3* gene was introduced and subsequently, other genes were integrated. Patients were analysed step by step for mutations in the *CAPN3*, *FKRP*, *SGCA*, and *ANO5* genes. In *CAPN3* and *SGCA*, all exons and adjacent intron regions were sequenced. In *FKRP* and *ANO5*, we analysed only exons with the occurrence of the most common mutations p.Leu276Ile and c.191dupA, respectively. In case of patients heterozygous for mentioned mutations, analysis of all exons of a relevant gene was performed. The mentioned genes were selected on the basis of published results related to higher percentage representation of patients with these types of LGMD2 in European populations [28],[29],[30].

Sequence capture and targeted resequencing (SeqCap-RT)

For identification of mutations associated with neuromuscular disorders, we introduced a solution-based capture method SeqCap EZ Choice Library (Roche NimbleGene) and targeted resequencing on the GS Junior System (Roche) in 2013. A custom capture array was designed to capture exons and adjacent intron sequences of 42 genes (1020 exons, 280 kb) known to be associated with muscular dystrophies (Group 1), congenital muscular dystrophies (Group 2), congenital myopathies (Group 3), distal myopathies (Group 4), and other myopathies (Group 5), according to http://www.musclegenetable.org/ . The list of selected diseases and genes includes Duchenne muscular dystrophy (*DMD*), Emery-Dreifuss muscular dystrophy (*EMD*, *FHL1*, *LMNA*), limb-girdle muscular dystrophy

(*MYOT*, *LMNA*, *CAV3*, *CAPN3*, *DYSF*, *SGCG*, *SGCA*, *SGCB*, *SGCD*, *TCAP*, *TRIM32*, *FKRP*, *TTN*, *POMT1*, *ANO5*, *FKTN*, *POMT2*, *POMGNT1*); and additionally genes associated with congenital muscular dystrophies (*LAMA2*, *LARGE*, *SEPN1*, *COL6A1*, *COL6A2*, *COL6A3*, *ITGA7*, *DNM2*); congenital myopathies, distal myopathies, and other myopathies (*NEB*, *TPM3*, *ACTA1*, *TPM2*, *TNNT1*, *CFL2*, *RYR*, *MTM1*, *BIN*, *CRYAB*, *DES*, *LAMP2*, *PABPN1*). Probes for the target regions were designed by Roche NimbleGen according to the protocol NimbleGen Sequence Capture Custom Designs. Sequence capture was performed according to the manufacturer's instructions (NimbleGen SeqCap EZ Choice Library LR User's Guide) with modifications of DNA fragmentation and hybridisation. We replaced nebulization by shearing using the S220 Focused Ultrasonicator (Covaris) (3µg DNA, peak incident power 140 W, duty factor 10 %, 200 cycles per burst and treatment time 80 s) and replaced three days hybridization by hybridization according to the NimbleGen technical note Double Capture: High Efficiency Sequence Capture of Small Targets for use in SeqCap EZ Library Applications on 454 Sequencing Systems. Emulsion PCR and sequencing on the GS Junior System (Roche) were performed according to the emPCR Amplification Method Manual - Lib-L and the Sequencing Method Manual. Sequencing data were evaluated using the software Sequence Pilot (JSI Medical Systems)**.**

In this study, we present only results in relation to patients in who mutations in genes associated with LGMD2 were identified. Overall results and results related to other genes are a subject of another study.

Results and discussion

Considering the wide clinical and genetic variability of LGMD, determination of particular types of LGMD is a comprehensive process. The muscle biopsy still represents an important and economical step in diagnosis, however, protein changes documented by immunohistochemistry can be secondary or not pronounced and so a definite diagnosis requires genetic analysis. Mutations of the *CAPN3* gene are the most common cause of LGMD2 and in the set of 218 Czech probands with suspicion of LGMD2, two mutations in this gene were identified in 67 patients. In 4 patients, only one *CAPN3* mutation was determined, and sequence analysis was completed by MLPA but no gene deletions/duplications were found (Table 1). Totally, we identified 37 different mutations

of which 12 have been described only in Czech LGMD2A patients. The most frequent mutation among our LGMD2A probands is c.550delA which was detected in 65 mutant alleles from the total number of 138 (47.1 %). Patients´ clinical and pathological findings (when available) are presented in Table 1S (Additional file 1).

Probands negative for *CAPN3* mutations (151) were analysed for the most common mutation p. Leu276Ile in the *FKRP* gene. The homozygous occurrence of this mutation was identified in 7 patients. In two patients heterozygous for p.(Leu276Ile), we were able to identify a second mutation – patient 79 carries the mutation p.(Pro316Arg) described in the study of [12], and patient 80 has the new mutation p.(Trp359Ser) (Table 2).

In 142 probands without mutations in the *CAPN3* or *FKRP* genes, we performed analysis of the *SGCA* gene. Mutations were detected in 6 patients (Table 2). The most common mutation among our LGMD2D patients is p.(Arg77Cys), which was present in 4 mutant alleles. We identified also one new *SGCA* mutation, c.303dupA (patient 85).

Probands negative for *CAPN3*, *FKRP*, and *SGCA* mutations (136) were screened for the most common mutation c.191dupA in the *ANO5* gene. This mutation was found heterozygously in 3 patients, and subsequent sequencing analysis of all exons and adjacent intron regions detected the mutation p.(Arg758Lys) in two of them (Table 2). The mutation p.(Arg758Lys) was described in combination with c.191dupA in other studies [24],[31] and associated phenotypes corresponded to distal non-dysferlin Miyoshi myopathy, unlike our patient 88 whose phenotype matches rather LGMD2L (a detailed description of the phenotype of patient 89 was not available). Analysis of the mutation p.(Arg758Lys) was subsequently performed in all LGMD2 patients with unconfirmed diagnosis at the DNA level but this mutation was not detected. In patient 87 carrying c.191dupA on one allele, we did not identify a second mutation, only the polymorphism p.(Leu322Phe) described in LMDP. Patients´ clinical and pathological findings are presented in Table 2S (Additional file 2).

The aim of the study was to evaluate the relative proportion of the most frequent types of LGMD2 identified in European countries in Czech probands with the suspicion of LGMD2. The results show that the frequency of the forms of LGMD2 analysed is 32.6% for LGMD2A (71 probands), 4.1% for LGMD2I (9 probands), 2.8% for LGMD2D (6 probands), and 1.4% for LGMD2L (3 probands). These results indicate that there is good agreement between the frequency of particular forms of LGMD2 in the Czech Republic and in Italy (LGMD2A 31.1%, LGMD2I 7.4%, LGMD2D 8.4%, LGMD2L 2.1% [28]; and LGMD2A 28.4%, LGMD2I 6.4%, LGMD2D 8.3% [2]), in contrast to studies from

Denmark where the frequency of LGMD2A is significantly lower (12.1%) and that of LGMD2I significantly higher (38.4%) [29] (Table 3).

In 2013, we introduced the SeqCap-TR method for genetic diagnostics of neuromuscular disorders, and using this approach we identified two patients with LGMD2B (mutations in the *DYSF* gene) and a patient with LGMD2E (mutations in the *SGCB* gene). Patient 90 is the compound heterozygote for two *DYSF* mutations (genotype p.[Gln1278*];[Asp1837Tyr]). Compound heterozygous and homozygous occurrence of p.(Gln1278*) was described in the study of [32], in the case of a homozygous patient an initial phenotype corresponded to Miyoshi myopathy. The mutation c.5509G>T, p.(Asp1837Tyr) was not identified so far, but the mutation c.5509G>A, p.(Asp1837Asn) was described in several studies. In the study of [33], a patient carried the genotype p.[Asp1837Asn];[Trp1968*] and the phenotype of LGMD2B, and in the study of [34] and [35] patients with the genotype p.[Asp1837Asn];[Tyr522*] had Miyoshi myopathy. In cases of homozygous occurrence of the mutation, phenotypes of Miyoshi myopathy [36],[37],[38] and also of LGMD2B [38] were observed. The phenotype of our patient 90 corresponded rather with LGMD2B. Patient 91 carries 4 mutations in the *DYSF* gene – p.(Ala170Glu), p.(Arg204*), p.(Val374Leu), and a novel mutation p.(Trp1969Cys). The mutation p.(Ala170Glu) is described in the LMDP as unknown pathogenic, probably pathogenic, and pathogenic; and the mutation p.(Val374Leu) as probably not pathogenic and pathogenic. The evaluation of these missense mutations using *in silico* tools is described in Table S3 (Additional file 3) together with the evaluation of selected missense mutations identified in the genes analysed. DNA analysis was performed in patient´s parents; her father carries p.(Arg204*) and p.(Val374Leu), her mother carries p.(Ala170Glu) and p.(Trp1969Cys). On the basis of *in silico* tools and inheritance in the patient´s family, we suppose that p.(Arg204*) and p.(Trp1969Cys) are causal mutations. Patient 92 is homozygous for the *SGCB* mutation p.(Ser114Phe) described in the LMDP in association with LGMD2E. The patients´ clinical and pathological findings are shown in Table 2S (Additional file 2). All novel missense mutations were tested on 150 control DNAs and none of them were detected.

Conclusions

We characterised a cohort of Czech LGMD2 patients on the basis of mutation analysis of genes associated with the most common forms of LGMD2 in the European population and compared these results with published studies. This study expands our previous results obtained in the Czech LGMD2A population [25],[26],[27] as well as the mutation spectrum in other types of LGMD2, and further provides information concerning clinical and genetic correlations.

Additional file

Additional file 1: Table S1. Mutations and pathological-clinical findings identified in Czech LGMD2A probands
Additional file 2: Table S2. Mutations and pathological-clinical findings identified in Czech LGMD2I, 2D, 2L, 2B, and 2E probands
Additional file 3: Table S3. *In silico* prediction of effects of selected missense mutations

List of abbreviations

DAG complex, Dystrophin-associated glycoprotein complex;
LGMD, Limb-girdle muscular dystrophy;
LMDP, Leiden muscular dystrophy pages;
MLPA, Multiplex ligation-dependent probe amplification
SG complex, Sarcoglycan complex;
SeqCap-TR, Sequence capture and targeted resequencing;

Competing interests

The authors declare that they have no competing interests.

Authors´ contributions

KS, DS, JZ carried out the molecular genetic studies. LM, PV, RM, SV, JH, HO, ND, PS performed clinical diagnostics of LGMD patients. MH, JZ, OS were involved in pathological analysis of muscle tissue. LF participated in design study and the manuscript preparation.

Acknowledgements

This work was funded by the project of the Internal Grant Agency of the Czech Ministry of Health (NT/14574-3); the projects of the Czech Ministry of Education "CEITEC – Central European Institute of Technology" (CZ.1.05/1.1.00/02.0068) and SuPReMMe (CZ.1.07/2.3.00/20.0045); and the project of the Czech Ministry of Health for conceptual development of research organization 65269705 (University Hospital Brno, Brno, Czech Republic) and MH – DRO, University Hospital Motol, Prague, Czech Republic 00064203 . We would like to thank the physicians from Departments of Neurology and Medical Genetics in the Czech Republic (Drs. P. Doležalová, M. Forgáč, R. Gaillyová, D. Grečmalová, A. Gřegořová, M. Havlová, T Honzík, P. Ješina, Z. Kalina, K. Kalous, V. Křivková, R. Kutějová, M. Langová, T. Maříková, Š. Prášilová, D. Polendová, M. Soukupová, J. Staněk, M. Ševčíková, E. Šilhánová, S. Širůčková, D. Tenora, J. Zvolská) for providing us with their patients' blood samples.

References

1. Kang PB, Feener CA, Estrella E, Thorne M, White AJ, Darras BT, Amato AA, Kunkel LM: **LGMD2I in a North American population**. *BMC Musculoskelet Disord* 2007, **8**:115.
2. Guglieri M, Magri F, D'Angelo MG, Prelle A, Morandi L, Rodolico C, Cagliani R, Mora M, Fortunato F, Bordoni A *et al*: **Clinical, molecular, and protein correlations in a large sample of genetically diagnosed Italian limb girdle muscular dystrophy patients**. *Hum Mutat* 2008, **29**(2):258-266.
3. Lo HP, Cooper ST, Evesson FJ, Seto JT, Chiotis M, Tay V, Compton AG, Cairns AG, Corbett A, MacArthur DG *et al*: **Limb-girdle muscular dystrophy: diagnostic evaluation, frequency and clues to pathogenesis**. *Neuromuscul Disord* 2008, **18**(1):34-44.
4. van der Kooi AJ, Frankhuizen WS, Barth PG, Howeler CJ, Padberg GW, Spaans F, Wintzen AR, Wokke JH, van Ommen GJ, de Visser M *et al*: **Limb-girdle muscular dystrophy in the Netherlands: gene defect identified in half the families**. *Neurology* 2007, **68**(24):2125-2128.
5. Zatz M, Vainzof M, Passos-Bueno MR: **Limb-girdle muscular dystrophy: one gene with different phenotypes, one phenotype with different genes**. *Curr Opin Neurol* 2000, **13**(5):511-517.
6. Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, Bourg N, Brenguier L, Devaud C, Pasturaud P, Roudaut C *et al*: **Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A**. *Cell* 1995, **81**(1):27-40.
7. Sorimachi H, Kinbara K, Kimura S, Takahashi M, Ishiura S, Sasagawa N, Sorimachi N, Shimada H, Tagawa K, Maruyama K *et al*: **Muscle-specific calpain, p94, responsible for limb girdle muscular dystrophy type 2A, associates with connectin through IS2, a p94-specific sequence**. *J Biol Chem* 1995, **270**(52):31158-31162.
8. Keira Y, Noguchi S, Minami N, Hayashi YK, Nishino I: **Localization of calpain 3 in human skeletal muscle and its alteration in limb-girdle muscular dystrophy 2A muscle**. *J Biochem* 2003, **133**(5):659-664.
9. Fanin M, Fulizio L, Nascimbeni AC, Spinazzi M, Piluso G, Ventriglia VM, Ruzza G, Siciliano G, Trevisan CP, Politano L *et al*: **Molecular diagnosis in LGMD2A: mutation analysis or protein testing?** *Hum Mutat* 2004, **24**(1):52-62.
10. Piluso G, Politano L, Aurino S, Fanin M, Ricci E, Ventriglia VM, Belsito A, Totaro A, Saccone V, Topaloglu H *et al*: **Extensive scanning of the calpain-3 gene broadens the spectrum of LGMD2A phenotypes**. *J Med Genet* 2005, **42**(9):686-693.
11. Saenz A, Leturcq F, Cobo AM, Poza JJ, Ferrer X, Otaegui D, Camano P, Urtasun M, Vilchez J, Gutierrez-Rivas E *et al*: **LGMD2A: genotype-phenotype correlations based on a large mutational survey on the calpain 3 gene**. *Brain* 2005, **128**(Pt 4):732-742.
12. Brockington M, Blake DJ, Prandini P, Brown SC, Torelli S, Benson MA, Ponting CP, Estournet B, Romero NB, Mercuri E *et al*: **Mutations in the fukutin-related protein gene (FKRP) cause a form of congenital muscular dystrophy with secondary laminin alpha2 deficiency and abnormal glycosylation of alpha-dystroglycan**. *Am J Hum Genet* 2001, **69**(6):1198-1209.
13. Mercuri E, Brockington M, Straub V, Quijano-Roy S, Yuva Y, Herrmann R, Brown SC, Torelli S, Dubowitz V, Blake DJ *et al*: **Phenotypic spectrum associated with mutations in the fukutin-related protein gene**. *Ann Neurol* 2003, **53**(4):537-542.

14. Walter MC, Petersen JA, Stucka R, Fischer D, Schroder R, Vorgerd M, Schroers A, Schreiber H, Hanemann CO, Knirsch U *et al*: **FKRP (826C>A) frequently causes limb-girdle muscular dystrophy in German patients**. *J Med Genet* 2004, **41**(4):e50.
15. Frosk P, Greenberg CR, Tennese AA, Lamont R, Nylen E, Hirst C, Frappier D, Roslin NM, Zaik M, Bushby K *et al*: **The most common mutation in FKRP causing limb girdle muscular dystrophy type 2I (LGMD2I) may have occurred only once and is present in Hutterites and other populations**. *Hum Mutat* 2005, **25**(1):38-44.
16. Angelini C, Fanin M, Freda MP, Duggan DJ, Siciliano G, Hoffman EP: **The clinical spectrum of sarcoglycanopathies**. *Neurology* 1999, **52**(1):176-179.
17. Straub V, Bushby K: **The childhood limb-girdle muscular dystrophies**. *Semin Pediatr Neurol* 2006, **13**(2):104-114.
18. Jarry J, Rioux MF, Bolduc V, Robitaille Y, Khoury V, Thiffault I, Tetreault M, Loisel L, Bouchard JP, Brais B: **A novel autosomal recessive limb-girdle muscular dystrophy with quadriceps atrophy maps to 11p13-p12**. *Brain* 2007, **130**(Pt 2):368-380.
19. Bolduc V, Marlow G, Boycott KM, Saleki K, Inoue H, Kroon J, Itakura M, Robitaille Y, Parent L, Baas F *et al*: **Recessive mutations in the putative calcium-activated chloride channel Anoctamin 5 cause proximal LGMD2L and distal MMD3 muscular dystrophies**. *Am J Hum Genet* 2010, **86**(2):213-221.
20. Tsutsumi S, Kamata N, Vokes TJ, Maruoka Y, Nakakuki K, Enomoto S, Omura K, Amagasa T, Nagayama M, Saito-Ohara F *et al*: **The novel gene encoding a putative transmembrane protein is mutated in gnathodiaphyseal dysplasia (GDD)**. *Am J Hum Genet* 2004, **74**(6):1255-1261.
21. Schroeder BC, Cheng T, Jan YN, Jan LY: **Expression cloning of TMEM16A as a calcium-activated chloride channel subunit**. *Cell* 2008, **134**(6):1019-1029.
22. Yang YD, Cho H, Koo JY, Tak MH, Cho Y, Shim WS, Park SP, Lee J, Lee B, Kim BM *et al*: **TMEM16A confers receptor-activated calcium-dependent chloride conductance**. *Nature* 2008, **455**(7217):1210-1215.
23. Stephan AB, Shum EY, Hirsh S, Cygnar KD, Reisert J, Zhao H: **ANO2 is the cilial calcium-activated chloride channel that may mediate olfactory amplification**. *Proc Natl Acad Sci U S A* 2009, **106**(28):11776-11781.
24. Hicks D, Sarkozy A, Muelas N, Koehler K, Huebner A, Hudson G, Chinnery PF, Barresi R, Eagle M, Polvikoski T *et al*: **A founder mutation in Anoctamin 5 is a major cause of limb-girdle muscular dystrophy**. *Brain* 2011, **134**(Pt 1):171-182.
25. Chrobakova T, Hermanova M, Kroupova I, Vondracek P, Marikova T, Mazanec R, Zamecnik J, Stanek J, Havlova M, Fajkusova L: **Mutations in Czech LGMD2A patients revealed by analysis of calpain3 mRNA and their phenotypic outcome**. *Neuromuscul Disord* 2004, **14**(10):659-665.
26. Hermanova M, Zapletalova E, Sedlackova J, Chrobakova T, Letocha O, Kroupova I, Zamecnik J, Vondracek P, Mazanec R, Marikova T *et al*: **Analysis of histopathologic and molecular pathologic findings in Czech LGMD2A patients**. *Muscle Nerve* 2006, **33**(3):424-432.
27. Stehlikova K, Zapletalova E, Sedlackova J, Hermanova M, Vondracek P, Marikova T, Mazanec R, Zamecnik J, Vohanka S, Fajkus J *et al*: **Quantitative analysis of CAPN3 transcripts in LGMD2A patients: involvement of nonsense-mediated mRNA decay**. *Neuromuscul Disord* 2007, **17**(2):143-147.
28. Magri F, Del Bo R, D'Angelo MG, Sciacco M, Gandossini S, Govoni A, Napoli L, Ciscato P, Fortunato F, Brighina E *et al*: **Frequency and characterisation of**

anoctamin 5 mutations in a cohort of Italian limb-girdle muscular dystrophy patients. *Neuromuscul Disord* 2012, **22**(11):934-943.
29. Sveen ML, Schwartz M, Vissing J: **High prevalence and phenotype-genotype correlations of limb girdle muscular dystrophy type 2I in Denmark**. *Ann Neurol* 2006, **59**(5):808-815.
30. van der Kooi AJ, Ten Dam L, Frankhuizen WS, Straathof CS, van Doorn PA, de Visser M, Ginjaar IB: **ANO5 mutations in the Dutch limb girdle muscular dystrophy population**. *Neuromuscul Disord* 2013, **23**(6):456-460.
31. Sarkozy A, Hicks D, Hudson J, Laval SH, Barresi R, Hilton-Jones D, Deschauer M, Harris E, Rufibach L, Hwang E *et al*: **ANO5 gene analysis in a large cohort of patients with anoctaminopathy: confirmation of male prevalence and high occurrence of the common exon 5 gene mutation**. *Hum Mutat* 2013, **34**(8):1111-1118.
32. Krahn M, Beroud C, Labelle V, Nguyen K, Bernard R, Bassez G, Figarella-Branger D, Fernandez C, Bouvenot J, Richard I *et al*: **Analysis of the DYSF mutational spectrum in a large cohort of patients**. *Hum Mutat* 2009, **30**(2):E345-375.
33. De Luna N, Freixas A, Gallano P, Caselles L, Rojas-Garcia R, Paradas C, Nogales G, Dominguez-Perles R, Gonzalez-Quereda L, Vilchez JJ *et al*: **Dysferlin expression in monocytes: a source of mRNA for mutation analysis**. *Neuromuscul Disord* 2007, **17**(1):69-76.
34. Matsuda C, Hayashi YK, Ogawa M, Aoki M, Murayama K, Nishino I, Nonaka I, Arahata K, Brown RH, Jr.: **The sarcolemmal proteins dysferlin and caveolin-3 interact in skeletal muscle**. *Hum Mol Genet* 2001, **10**(17):1761-1766.
35. Tagawa K, Ogawa M, Kawabe K, Yamanaka G, Matsumura T, Goto K, Nonaka I, Nishino I, Hayashi YK: **Protein and gene analyses of dysferlinopathy in a large group of Japanese muscular dystrophy patients**. *J Neurol Sci* 2003, **211**(1-2):23-28.
36. Rosales XQ, Gastier-Foster JM, Lewis S, Vinod M, Thrush DL, Astbury C, Pyatt R, Reshmi S, Sahenk Z, Mendell JR: **Novel diagnostic features of dysferlinopathies**. *Muscle Nerve* 2010, **42**(1):14-21.
37. Cagliani R, Magri F, Toscano A, Merlini L, Fortunato F, Lamperti C, Rodolico C, Prelle A, Sironi M, Aguennouz M *et al*: **Mutation finding in patients with dysferlin deficiency and role of the dysferlin interacting proteins annexin A1 and A2 in muscular dystrophies**. *Hum Mutat* 2005, **26**(3):283.
38. Ueyama H, Kumamoto T, Nagao S, Masuda T, Horinouchi H, Fujimoto S, Tsuda T: **A new dysferlin gene mutation in two Japanese families with limb-girdle muscular dystrophy 2B and Miyoshi myopathy**. *Neuromuscul Disord* 2001, **11**(2):139-145.
39. Fanin M, Nascimbeni AC, Aurino S, Tasca E, Pegoraro E, Nigro V, Angelini C: **Frequency of LGMD gene mutations in Italian patients with distinct clinical phenotypes**. *Neurology* 2009, **72**(16):1432-1435.
40. Fanin M, Nascimbeni AC, Tasca E, Angelini C: **How to tackle the diagnosis of limb-girdle muscular dystrophy 2A**. *Eur J Hum Genet* 2009, **17**(5):598-603.
41. Boito CA, Melacini P, Vianello A, Prandini P, Gavassini BF, Bagattin A, Siciliano G, Angelini C, Pegoraro E: **Clinical and molecular characterization of patients with limb-girdle muscular dystrophy type 2I**. *Arch Neurol* 2005, **62**(12):1894-1899.
42. Hanisch F, Grimm D, Zierz S, Deschauer M: **Frequency of the FKRP mutation c.826C>A in isolated hyperCKemia and in limb girdle muscular dystrophy type 2 in German patients**. *J Neurol* 2010, **257**(2):300-301.

Table 1. Mutations identified in Czech LGMD2A probands

No.	Mutations (cDNA level)	Mutations (protein level)
1 - 23	c.550delA/c.550delA	p.(Thr184Argfs*36)/p.(Thr184Argfs*36)
24, 25	c.550delA/c.245C>T	p.(Thr184Argfs*36)/p.(Pro82Leu)
26	c.550delA/c.328C>T	p.(Thr184Argfs*36)/p.(Arg110*)
27	c.550delA/c.509A>G	p.(Thr184Argfs*36)/p.(Tyr170Cys)
28	c.550delA/c.598_612del	p.Thr184Argfs*36/p.Phe200_Leu204del
29	c.550delA/c.1043delG	p.(Thr184Argfs*36)/p.(Gly348Valfs*4)
30	c.550delA/c.1069C>T	p.(Thr184Argfs*36)/p.(Arg357Trp)
31	c.550delA/**c.1451T>C**	p.(Thr184Argfs*36)/**p.(Leu484Pro)**
32	c.550delA/c.1465C>T	p.(Thr184Argfs*36)/p.(Arg489Trp)
33, 34	c.550delA/c.1468C>T	p.Thr184Argfs*36/p.(Arg490Trp)
35	c.550delA/**c.1470delG**	p.(Thr184Argfs*36)/**p.(Arg490Argfs*6)**
36, 37	c.550delA/**c.1722delC**	p.Thr184Argfs*36/**p.Ser575Leufs*20**
38	c.550delA/c.1823G>A	p.(Thr184Argfs*36)/p.(Arg608Lys)
39	c.550delA/c.1981delA	p.Thr184Argfs*36/p.Ile661*
40	c.550delA/**c.2245A>C**	p.(Thr184Argfs*36)/**p.(Asn749His)**
41	**c.1A>G**/c.865C>T	**p.(Met1Val)**/p.(Arg289Trp)
42	c.133G>A/c.133G>A	p.Ala45Thr/p.Ala45Thr
43	c.146G>A/c.1069C>T	p.(Arg49His)/p.(Arg357Trp)
44	**c.224A>G/c.224A>G**	**p.Tyr75Cys/p.Tyr75Cys**
45	c.245C>T/c.245C>T	p.Pro82Leu/p.Pro82Leu
46	c.245C>T/**c.1800+1G>A**	p.(Pro82Leu)/**splicing**
47	c.245C>T/**c.1855C>T**	p.(Pro82Leu)/**p.(Gln619*)**
48	c.245C>T/c.2314_2317del	p.Pro82Leu/p.Asp772Asnfs*3
49	c.509A>G/c.509A>G	p.(Tyr170Cys)/p.(Tyr170Cys)
50	c.598_612del/c.598_612del	p.(Phe200_Leu204del)/p.(Phe200_Leu204del)
51	c.598_612del/c.640G>A	p.(Phe200-Leu204del)/p.(Gly214Ser)
52	c.598_612del/**c.2245A>C**	p.Phe200-Leu204del/**p.Asn749His**
53	c.1043delG/**c.1094G>A**	p.(Gly348Valfs*4)/**p.(Trp365*)**
54	c.1043delG/c.1343G>A	p.(Gly348Valfs*4)/p.(Arg448His)
55	c.1194-9A>G/**c.1800+1G>A**	splicing/**splicing**
56	c.1194-9A>G/c.2393C>A	splicing/p.(Ala798Glu)

57	c.1250C>T/c.1250C>T	p.(Thr417Met)/p.(Thr417Met)
58	c.1322G>A/c.1322G>A	p.Gly441Asn/p.Gly441Asn
59	c.1322delG/c.2114A>G	p.(Gly441Valfs*22)/p.(Asp705Gly)
60	c.1343G>A/**c.2093G>A**	p.(Arg448His)/**p.(Arg698His)**
61	c.1468C>T/c.2314_2317del	p.(Arg490Trp)/p.(Asp772Asnfs*3)
62	**c.1788_1793del**/c.2242C>T	**p.Lys597_Lys598del**/p.Arg748*
63 - 66	c.598_612del/c.1746-20C>G	p.(Phe200_Leu204del)/splicing
67	**c.614T>C**/c.1746-20C>G	**p.(Leu205Pro)**/splicing
68, 69	c.550delA/-	p.(Thr184Argfs*36)/-
70, 71	c.598_612del/-	p.(Phe200_Leu204del)/-

Mutations in bold letters were detected only in Czech LGMD2A patients.

Table 2. Mutations identified in Czech LGMD2I, 2D, 2L, 2B, and 2E probands

No.	Gene	Mutation (cDNA level)	Mutation (protein level)
72 - 78	*FKRP*	c.826C>A/c.826C>A	p.(Leu276Ile)/p.(Leu276Ile)
79	*FKRP*	c.826C>A/c.947C>G	p.(Leu276Ile)/p.(Pro316Arg)
80	*FKRP*	826C>A/**c.1076G>C**	p.(Leu276Ile)/**p.(Trp359Ser)**
81	*SGCA*	c.229C>T/c.229C>T	p.(Arg77Cys)/p.(Arg77Cys)
82, 83	*SGCA*	c.157+1G>A/c.850C>T	splicing/p.(Arg284Cys)
84	*SGCA*	c.229C>T/c.739G>A	p.(Arg77Cys)/p.(Val247Met)
85	*SGCA*	c.290A>G/**c.303dupA**	p.(Asp97Gly)/**p.(Gln101Glnfs*4)**
86	*SGCA*	c.229C>T/c.308T>C	p.(Arg77Cys)/p .(Ile103Thr)
87	*ANO5*	c.191dupA/*c.966A>T*	p.(Asn64Lysfs*15)/*p.(Leu322Phe)*
88, 89	*ANO5*	c.191dupA/c.2272C>T	p.(Asn64Lysfs*15)/p.(Arg758Cys)
90	*DYSF*	c.3832C>T/**c.5509G>T**	p.(Gln1278*)/**p.(Asp1837Tyr)**
91	*DYSF*	c.509C>A/**c.5907G>C** c.610C>T/c.1120G>C/	p.(Ala170Glu)/**p.(Trp1969Cys)** p.(Arg204*)/p.(Val374Leu)/
92	*SGCB*	c.341C>T/c.341C>T	p.(Ser114Phe)/p.(Ser114Phe)

Mutations in bold letters were detected only in Czech LGMD2 patients. The variant c.966A>T written in italics is probably a nucleotide polymorphism (LMDP, dbSNP-rs7481951).

Table 3. Frequency of LGMD2A, 2I, 2D, 2L in European countries

Litera-ture	Population	Number of patients analysed in the study	Number and percentage of patients with LGMD2A	Number and percentage of patients with LGMD2I	Number and percentage of patients with LGMD2D	Number and percentage of patients with LGMD2L
This study	Czech	218 LGMD2 probands	71 (67 had two mutations, 4 one); **32.6%**	9; **4.1%**	6; **2.8%**	3; **1.4%**
[28]	Italy	190 LGMD probands	59; **31.1%**	14; **7.4%**	16; **8.4%**	4; **2.1%**
[2]	Italy	155 LGMD probands	44 (30 had two mutations, 14 one); **28.4%**	10; **6.4%**	13; **8.4%**	NI
[39]	Italy	550 patients with LGMD, myopathy, or asymptomatic hyperCKemia	102; **18.5%**	16; **2.8%**	37; **6.7%**	NI
[40]	Italy	519 patients with LGMD, myopathy, or asymptomatic hyperCKemia	94 (76 had two mutations, 18 one); **18.1%**	NI	NI	NI
[10]	Italy	530 patients with muscular dytrophy	141 (104 had two mutations and 37 one); **26.6%**	NI	NI	NI
[41]	Italy	214 probands with muscular dystrophy or myopathy	NI	13 (9 had two mutations, 4 one); **6.1%**	NI	NI
[42]	Germany	98 probands with LGMD2 and 102 probands with asymptomatic or minimally symptomatic hyperCKemia	NI	7; **3.5%**	NI	NI
[29]	Denmark	99 LGMD2 patients	12; **12.1%**	38; **38.4%**	NI	NI
[4]	Dutch	67 LGMD probands	14; **20.9%**	5; **7.5%**	NI	NI
[30]	Dutch	68 LGMD probands	NI	NI	NI	11; **16.2%**

NI: no information

Neuromuscular Disorders 19 (2009) 749–753

Contents lists available at ScienceDirect

Neuromuscular Disorders

journal homepage: www.elsevier.com/locate/nmd

Point mutations in Czech DMD/BMD patients and their phenotypic outcome

Jana Sedláčková [a,b], Petr Vondráček [c], Markéta Hermanová [d], Josef Zámečník [e], Zuzana Hrubá [a], Jana Haberlová [f], Josef Kraus [f], Tat'ána Maříková [g], Petra Hedvičáková [g], Stanislav Voháňka [h], Lenka Fajkusová [a,b,*]

[a] Centre of Molecular Biology and Gene Therapy, University Hospital Brno and Masaryk University, Brno, Czech Republic
[b] Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czech Republic
[c] Department of Child Neurology, University Hospital Brno, Brno, Czech Republic
[d] Institute of Pathology, University Hospital Brno, Brno, Czech Republic
[e] Department of Pathology and Molecular Medicine, Charles University, 2nd Medical Faculty and University Hospital Motol, Prague, Czech Republic
[f] Department of Paediatric Neurology, Charles University, 2nd Medical Faculty and University Hospital Motol, Prague, Czech Republic
[g] Institute of Biology and Medical Genetics, Charles University, 2nd Medical Faculty and University Hospital Motol, Prague, Czech Republic
[h] Clinic of Neurology, University Hospital Brno, Brno, Czech Republic

ARTICLE INFO

Article history:
Received 31 March 2009
Received in revised form 31 July 2009
Accepted 24 August 2009

Keywords:
Duchenne muscular dystrophy
DMD
BMD
Point mutation
Nonsense mediated mRNA decay

ABSTRACT

Duchenne and Becker muscular dystrophies (DMD/BMD) are associated with mutations in the DMD gene. We determined the mutation status of 47 patients with dystrophinopathy without deletion or duplication in the DMD gene by screening performed by reverse transcription-PCR, protein truncation test, and DNA sequencing. We describe three patients with a mutation creating a premature termination codon (p.E55X, p.E1110X, and p.S3497PfsX2) but with a mild phenotype, which present three different ways of rescuing the DMD phenotype. In one patient we detected the insertion of a repetitive sequence *AluYa5* in intron 56, which led to skipping of exon 57. Further, using quantitative analysis of *DMD* mRNA carrying various mutated alleles, we examine levels of mRNA degradation due to nonsense mediated mRNA decay. The quantity of dystrophin mRNA is different depending on the presence of a mutation leading to a premature termination codon, and position of the analysed mRNA region with respect to its 5′ end or 3′ end. Average relative amounts of *DMD* mRNAs carrying a premature termination codon is 48% and 17%, when using primers amplifying the 5′ and 3′ cDNA regions, respectively.

1. Introduction

Dystrophin-associated muscular dystrophies range from the severe Duchenne muscular dystrophy (DMD) to the milder Becker muscular dystrophy (BMD). Both are the result of mutations in the gene that encodes dystrophin, the *DMD* gene is located on the X-chromosome where it occupies a region about 2.3 Mb long. The total coding sequence of 11,058 bp is distributed over 79 exons. Approximately 60% of DMD/BMD mutations are deletions and 5% duplications of one or more exons. These deletions and duplications can be detected using multiplex ligation dependent probe amplification (MLPA) or multiplex PCRs at the DNA level, or by means of reverse transcription-PCR (RT-PCR) at the mRNA level [1–3]. Most deletions, as well as pathogenic point mutations associated with the DMD phenotype, disrupt the translational reading frame. An effective mutation–detection method to monitor the integrity of the reading frame is the protein truncation test (PTT), which applies a combination of RT-PCR, *in vitro* transcription and *in vitro* translation [4].

Eukaryotic cells have evolved various mechanisms to recognize and eliminate functionally aberrant RNAs. Nonsense mediated mRNA decay (NMD) is one such type of quality-control mechanism that selectively degrades mRNAs with premature termination codons (PTCs), thus preventing the synthesis of potentially deleterious truncated proteins [5]. It is possible that NMD also induces degradation of dystrophin mRNAs containing mutations generating PTCs, but the level of this degradation has not been determined so far.

In this study, we apply RT-PCR, PTT, and DNA sequencing to scan dystrophin mRNA and/or DNA of 43 DMD/BMD patients and 4 female carriers without a deletion or duplication in the *DMD* gene. Further, using quantitative analysis of *DMD* mRNA carrying various mutated alleles, we examine levels of mRNA degradation due to NMD. The patients with mutations conserving the translational reading frame had relative amount of *DMD* mRNA approximately 118% and 78% (using primers complementary to 5′ end and 3′ end

* Corresponding author. Address: Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Černopolní 9, CZ-625 00 Brno, Czech Republic. Tel.: +420 532 234 625; fax: +420 532 234 623.
E-mail address: lfajkusova@fnbrno.cz (L. Fajkusová).

0960-8966/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.nmd.2009.08.011

of *DMD* cDNA, respectively) and patients with PTC mutations 48% and 17%, respectively.

2. Materials and methods

2.1. Patients

The patients presented in this study met the clinical criteria of DMD or BMD phenotype, respectively, including results of histopathological and immunohistochemical analysis of muscle tissue.

2.2. Immunohistochemical analysis, Western blots, and RNA analysis

Immunohistochemical analysis of muscle proteins, as well as RNA isolation from muscle tissue, reverse transcription, amplification of dystrophin mRNA, and PTT have been described in our previous studies [6,7].

2.3. DNA analysis

Isolation of DNA was performed according to the standard salting-out method. Primers for amplification of all exons of the *DMD* gene and adjacent intron regions are given in the Supplementary file including annealing temperature and Mg^{2+} concentration. PCRs contained 200 ng of DNA, 1× PCR GeneAmp Gold Buffer, 0.625 U AmpliTaq Gold polymerase (Applied Biosystems, Foster City, CA, USA), 0.2 mM dNTPs, 0.5 μM primers, and $MgCl_2$ at concentrations listed in Supplementary Table 1. Amplifications were performed under the following cycling conditions: 95 °C/7 min, followed by 35 cycles of 95 °C/30 s, annealing temperature as given in the Supplementary file/30 s, and 72 °C/30 s; and final step 72 °C/10 min. The PCR products were purified and sequenced on an ABI PRISM 310 sequencer (Applied Biosystems).

2.4. Real-time PCR

One hundred nanograms of RNA was used for RT using the SuperScript™ First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA). The cDNA synthesis reactions were primed by random hexamers. RT products were amplified using FastStart SYBR Green Master (Rox) (Roche Diagnostics, Mannheim, Germany) under the following cycling conditions: 95 °C/10 min, followed by 40 cycles of 95 °C/15 s, 60 °C/30 s, and 72 °C/60 s. Fluorescence was monitored at 72 °C and the specificity of amplification was evaluated by melting curve analysis. The analysis of each RNA sample was performed in two parallel PCRs. Amplification of the *DMD* mRNA was performed using three pairs of primers that amplified three different regions of dystrophin mRNA. These regions were chosen to be located near the 5′ end (exons 12–14), in the middle (exons 34–35), and near the 3′ end (exons 67–68) of the dystrophin cDNA so that the long dystrophin transcript was probed at several different segments. Primer sequences are shown in Supplementary Table 2. The data obtained were analysed by the $2^{-\Delta\Delta C_t}$ method that evaluates the target gene expression with respect to the reference gene expression and normalizes this value relative to expression in a calibrator sample [8]. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as the reference gene and the mRNA of a patient with the in-frame exon 48 deletion was used as the calibrator.

3. Results

3.1. Analysis of dystrophin mRNA and DNA

Analysis of dystrophin mRNA and/or DNA was performed in 47 patients with dystrophinopathy (34 DMD, 9 BMD, and 4 carrier females). The results of these analyses together with the phenotype and immunohistochemical results are shown in Table 1. Twenty-eight mutations are described in these patients for the first time. At present, all mutations are registered at the Leiden muscular dystrophy pages (www.dmd.nl).

From 43 male patients, nine are BMD (patients 1, 9, 14, 15, 16, 17, 22, 32, and 43). Patient 1 has the mutation c.163G>T (p.E55X). Besides a transcript with this mutation, a transcript with the frame-shift deletion of exons 3–7 (c.94_649del) was detected. Patient 9 carries the mutation c.1483-8T>G, which changes splicing and results in a transcript with the in-frame deletion of exon 13 (c.1483_1602del). Four other BMD patients have mutations associated with a defective splicing of exon 25. Patient 14 has the mutation c.3328G>T (p.E1110X) that changes an exon splicing enhancer sequence inside exon 25 (predicted by Rescue-ESE program, http://genes.mit.edu/burgelab/rescue-ese/) and causes alternative splicing; an apparent in-frame alternative splicing (c.3277_3432del) was also detected at the mRNA level, besides the transcript with the nonsense mutation [6]. Patient 15 carries the mutation c.3432G>A that changes the splicing of exon 25 and generates mRNA with the exon 25 deletion. A similar situation is seen in patients 16 and 17; mRNA containing the exon 25 deletion is generated as a result of the mutation c.3432+1G>A or c.3432+2T>C, respectively, and the standard full-length transcript is also detected in both cases. Patient 22 has the mutation c.4675-11A>G, which changes splicing and, as well as the standard mRNA, a transcript with insertion of a part of intron 33 is present (c.4674_4675ins10). In the case of patient 32, the mutation c.7660+4delA causes alternative splicing of exon 52 and both transcripts are generated (c.[=, 7543_7660del]). Patient 43 has the mutation c.10489delT, detected only at the DNA level (a muscle biopsy was not available).

Patient 18 has an unusual combination of two inherited neuromuscular disorders, DMD and Charcot-Marie-Tooth (CMT) disease type 1A. DNA analysis detected a duplication containing the *PMP22* (peripheral myelin protein 22) gene located within the region 17p11.2-p12. Besides this mutation, the mutation c.3609_3612delTAAAinsCTT (p.K1204LfsX11) was determined at the DNA and mRNA levels in the *DMD* gene [9]. In patient 34, we identified the deletion of exon 57 (c.8391_8547del) at the mRNA level. Analysis of DNA showed the insertion of the repetitive sequence *AluYa5* that was localised in intron 56 (described in Fig. 1). The presence of two pseudoexon insertions (c.264_265ins 265-595_265-464 and c.9563_9564ins9564-484_9564-432) was determined in patients 2 and 39, respectively. In both cases, DNA sequencing of the appropriate intron showed a point mutation creating a novel splice site. We detected two intronic point mutations resulting in a pseudoexon inclusion in mature mRNA: the mutation c.265-463A>G in intron 4 (patient 2) and c.9564-427T>G in intron 65 (patient 39).

3.2. Quantitative analysis of dystrophin transcripts

Quantitative analyses of *DMD* mRNA were performed in patients mentioned in Table 2 using real-time PCR as described in Section 2. Results of these analyses are shown in Fig. 2. Using primers 12Fx14R, the patients with mutations conserving the translational reading frame had an average relative amount of *DMD* mRNA 118% and patients with a PTC mutation 48%. Using primers 34Fx35R, these values were 83% and 40%, respectively, and using primers 67Fx68R the corresponding values were 78% and 17%, respectively.

4. Discussion

The BMD patient 1 has the mutation c.163G>T (p.E55X) and besides this mutation, a transcript with the frame-shift deletion of exons 3–7 was detected. Winnard et al. [10] also described a

Table 1
DMD/BMD patients, detected mutations and results of immunohistochemical analysis.

Patient	Phenotype	Localisation of mutation (exon)	Mutation detected at cDNA level	Mutation detected at DNA level	Mutation at protein level*	Immunohistochemical labelling using antibodies DYS1,2,3
1	**BMD**	**3**	**c.[163G>T, 94_649del]**	**Not performed**	**p.[E55X, F32_D217del]**	**Not performed**
2	**DMD**	**Intron 4**	**c.264_265ins265-595_265-464**	**c.265-463A>G**	**p.V89MfsX10**	**Negative**
3	DMD	7	c.583C>T	Not performed	p.R195X	Negative
4	DMD	7	c.583C>T	Not performed	p.R195X	Negative
5	DMD	Intron 7	c.531 649del	c.649+5G>A	p.P178CfsX2	DYS2 and DYS3 negative, DYS1 present at 20% of muscle fibres
6	DMD	Intron 7	Not performed	c.649+5G>A	p.P178CfsX2	Not performed
7	DMD	10	c.1062G>A	Not performed	p.W354X	Negative
8	**DMD**	**11**	**c.1324C>T**	**Not performed**	**p.Q442X**	**Negative**
9	**BMD**	**Intron 12**	**c.1483_1602del**	**c.1483-8T>G**	**p.V495_K534del**	**DYS1 negative, DYS2 and DYS3 decreased intensity**
10	**DMD**	**18**	**c.2276T>G**	**Not performed**	**p.L759X**	**Negative**
11	DMD	18	c.2281G>T	Not performed	p.E761X	Negative
12	**DMD**	**Intron 20**	**c.2622_2623insAG**	**c.2623-3C>G**	**p.D875RfsX17**	**Negative**
13	DMD	21	c.2797C>T	Not performed	p.Q933X	Negative
14	**BMD**	**25**	**c.[3328G>T, 3277_3432del]**	**Not performed**	**p.[E1110X, L1093_Q1144del]**	**Decreased intensity**
15	**BMD**	**25**	**c.3277_3432del**	**c.3432G>A**	**p.L1093_Q1144del**	**DYS1 and DYS2 normal, DYS3 decreased intensity**
16	BMD	Intron 25	c.[=, 3277 3432del]	c.3432+1G>A	p.[=, L1093 Q1144del]	Normal
17	**BMD**	**Intron 25**	**c.[=, 3277_3432del]**	**c.3432+2T>C**	**p.[=, L1093_Q1144del]**	**DYS3 negative, DYS1 and DYS2 decreased intensity**
18	**DMD+CMT1A**	**27**	**c.3609_3612del TAAAinsCTT**	**Not performed**	**p.K1204LfsX11**	**Not performed**
19	DMD	29	Not performed	c.3982C>T	p.Q1328X	Not performed
20	**DMD**	**30**	**c.4099C>T**	**Not performed**	**p.Q1367X**	**Negative**
21	**DMD**	**33**	**c.4638delA**	**Not performed**	**p.A1547LfsX2**	**Negative**
22	**BMD**	**Intron 33**	**c.[=, 4674_4675ins10]**	**c.4675-11A>G**	**p.[=, V1559FfsX20]**	**DYS1 negative, Dys2 and DYS3 decreased intensity**
23	DMD	34	c.4729C>T	Not performed	p.R1577X	Negative
24	**DMD**	**36**	**c.5031C>A**	**Not performed**	**p.Y1677X**	**Negative**
25	**DMD**	**40**	**c.5653C>T**	**Not performed**	**p.Q1885X**	**Negative**
26	**DMD**	**40**	**c.5662delG**	**Not performed**	**D1888IfsX3**	**Negative**
27	DMD	43	c.6283C>T	Not performed	p.R2095X	Not performed
28	DMD	44	c.6292C>T	Not performed	p.R2098X	Not performed
29	**DMD**	**47**	**c.6905G>A**	**Not performed**	**p.W2302X**	**Negative**
30	**DMD**	**48**	**Not performed**	**c.6931dupG**	**p.E2311GfsX7**	**Not performed**
31	DMD	52	c.7657C>T	Not performed	p.R2553X	Negative
32	**BMD**	**Intron 52**	**c.[=, 7543_7660del]**	**c.7660+4delA**	**p.[=, A2515LfsX22]**	**Decreased intensity**
33	**DMD**	**56**	**c.8374A>T**	**Not performed**	**p.K2792X**	**Negative**
34	**DMD**	**Intron 56**	**c.8391_8547del**	**c.8391-19_8391-18insAluYa5**	**p.S2798PfsX6**	**Negative**
35	DMD	58	c.8608C>T	Not performed	p.R2870X	Negative
36	**DMD**	**59**	**Not performed**	**c.8854C>T**	**p.Q2952X**	**Not performed**
37	**DMD**	**60**	**c.8947G>T**	**Not performed**	**p.G2983X**	**Not performed**
38	DMD	62	c.9182G>A	Not performed	p.W3061X	Negative
39	**DMD**	**Intron 65**	**c.9563_9564ins9564-484_9564-432**	**c.9564-427T>G**	**p.R3190NfsX6**	**Negative**
40	**DMD**	**70**	**c.10094C>A**	**Not performed**	**p.S3365X**	**Negative**
41	DMD	70	c.10108C>T	Not performed	p.R3370X	Negative
42	DMD	70	c.10141C>T	Not performed	p.R3381X	Negative
43	**BMD**	**74**	**Not performed**	**c.10489delT**	**p.S3497PfsX2**	**Not performed**
44F	Clinically affected	8	c.[721C>T]+[=]	Not performed	p.[Q241X]+[=]	Mosaic of normal and negative muscle fibres
45F	Not affected	12	c.[1465C>T]+[=]	Not performed	p.[Q489X]+[=]	Normal
46F	**Clinically affected**	**32**	**Not performed**	**c.[4384C>T]+[=]**	**p.[Q1462X]+[=]**	**Mosaic of normal and negative muscle fibres**
47F	**Not affected**	**55**	**c.[8198_8199del AG]+[=]**	**Not performed**	**p.[E2733AfsX11]+[=]**	**Normal**

The nucleotide numbering corresponds to the cDNA reference sequence on www.dmd.nl. The nomenclature is according to HGVS recommendation (http://www.hgvs.org). Mutations that were detected in this work for the first time are given in bold letters. F, female; =, normal transcript; *, protein changes are deduced theoretically.

BMD patient with the deletion of exons 3–7; antibodies directed against the amino-terminus and the 5′ end of exon 8 did not detect dystrophin in muscles of this patient, but dystrophin was detected with an antibody directed against the 3′ end of exon 8. The authors supposed that dystrophin production is initiated at a new start codon in exon 8, and a similar mechanism could have acted in the case of our patient.

Patient 43 carries the mutation c.10489delT leading to a frameshift in the C-terminal region, but showed only a mild BMD phenotype. The mRNA analysis was not performed because a muscle biopsy was not available. The rare protein-truncating mutations that occur near the 3′ end of the *DMD* gene can result in variable phenotypes, suggesting that these truncated proteins are capable of partially rescuing the DMD phenotype [11].

cttttcaatggaattgttagaatcatcaattacactt
ctagatattctgacatggtac *gctgctgttctttttt*
(T)$_N$ GAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAG
TGCAGTGGCGGGATCTCGGCTCACTGCAAGCTCCGCCT
CCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCAAG
TAGCTGGGACTACAGGCGCCCGCCACTACGCCCGGCTA
ATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTT
TTAGCCGGGATGGTCTCGATCTCCTGACCTCGTGATCC
GCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG
TGAGCCACCGCGCCCGGCC *gctgctgttctttttt* ca
gGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT
CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACA

Fig. 1. *AluYa5* insertion into intron 56 of the *DMD* gene. The insertion is localised in intron 56, 18 nucleotides before exon 57, and is in antisense orientation relative to the *DMD* gene. The *Alu* element is composed of a 281 bp of *Alu* repeat element with a polyA tract (the exact number of A residues was not determined) and 15 bp direct repeats flanking the inserted element. The direct repeats were duplicated during the *AluYa5* insertion. This sequence presents the sense strand. The lower case – the intron sequence; the upper case – the exon sequence; the italics – direct repeats; the upper case underlined – *AluYa5* insertion.

In patient 34, the deletion of exon 57 (c.8391-8547del) was detected at the mRNA level, and analysis of DNA showed the insertion of a repetitive sequence *AluYa5* in intron 56 (18 nucleotides before exon 57) in an antisense direction. To date, 5 other reports have described *de novo Alu* insertions into introns in an antisense orientation and in close proximity to the affected exon; these induce full exon skipping or alternative splicing of the downstream exon [12]. In the case of the *DMD* gene, one *Alu*-like retrotransposition has been described and was associated with X-linked dilated cardiomyopathy; this *Alu*-like sequence rearrangement occurred 2.4 kb downstream from the 5′ end of intron 11 of the dystrophin gene. This rearrangement activated one cryptic splice site in intron 11 and produced an alternative transcript containing the *Alu*-like sequence and a part of the adjacent intron 11, spliced between exons 11 and 12, which contained numerous stop codons in reading frame of the *Alu*-like sequence. Only the mutant mRNA was detected in the heart muscle, but in the skeletal muscle it coexisted with the normal one [13,14].

The NMD pathway recognizes and degrades mRNAs that contain a premature termination codon, and is one of many

Table 2
DMD/BMD patients selected for quantitative analysis of *DMD* transcripts.

Patient	Mutation detected at cDNA level	Mutation at protein level	Localisation of mutation (exon)	Impact of mutation on the ORF	Phenotype
del48.1	c.6913 7098del	p.V2305 Q2366del	48 deletion	Non-PTC	BMD
del48.2	c.6913 7098del	p.V2305 Q2366del	48 deletion	Non-PTC	BMD
del13-29	c.1483 4071del	p.V495 E1357del	13–29 deletion	Non-PTC	BMD
del48-51	c.6913 7542del	p.V2305 K2514del	48–51deletion	Non-PTC	BMD
pat17	c.[=, 3277 3432del] (DNA: 3432+2T>C)	p.[=,L1093 Q1144del]	Intron 25	Non-PTC	BMD
del49-50	c.7099 7309del	p.E2367LfsX9	49–50 deletion	PTC	DMD
del61	c.9085 9163del	p.V3029RfsX34	61deletion	PTC	DMD
pat11	c.2281G>T	p.E761X	18	PTC	DMD
pat13	c.2797C>T	p.Q933X	21	PTC	DMD
pat20	c.4099C>T	p.Q1367X	30	PTC	DMD
pat24	c.5031C>A	p.Y1677X	36	PTC	DMD
pat31	c.7657C>T	p.R2553X	52	PTC	DMD
pat33	c.8374A>T	p.K2792X	56	PTC	DMD
pat37	c.8947G>T	p.G2983X	60	PTC	DMD
pat39	c.9563 9564ins9564-484 9564-432 (DNA: c.9564-427T>G)	p.R3190NfsX6	Intron 65	PTC	DMD
pat41	c.10108C>T	p.R3370X	70	PTC	DMD

PTC, mutation creating a premature termination codon; non-PTC, mutation not creating a premature termination codon.

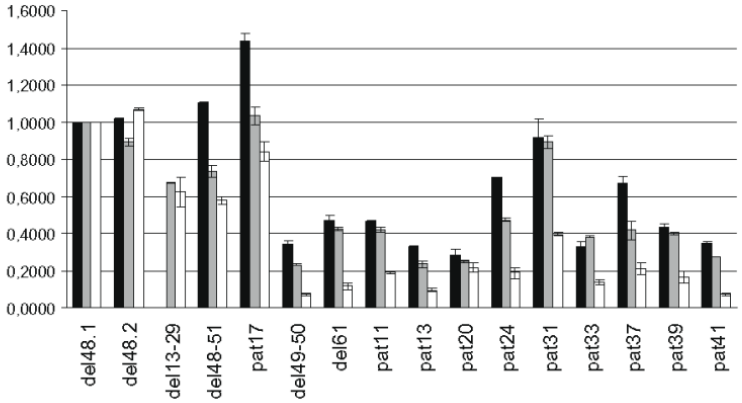


Fig. 2. Quantitative analysis of *DMD* transcripts in individual patients. Identification of samples is given on axis *x*. *y*-axis scale shows relative levels of mRNA transcripts with respect to the calibrator sample. Black columns show relative levels of the mRNA region amplified by primers 12F and 14R (5′ end region), grey columns correspond to products generated by 34F and 35R primers (central region) and white columns represent the levels of 5′ end region obtained using 67F and 68R primers. Error bars show standard deviations.

quality-control processes that have evolved to ensure the exclusive production of functional proteins. The efficiency of NMD appears to vary depending on the gene, the position of the PTCs, and the tissue type. We observed NMD in the case of the *CAPN3* gene whose mutated forms are associated with limb girdle muscular dystrophy 2A. We determined that patients with a single PTC mutation (a missense mutation or in-frame deletion present in the second allele) have the *CAPN3* mRNA level decreased to 28–58% of the reference value, and that patients with PTC mutations on both alleles have *CAPN3* transcript levels of only about 2% [15]. Similar studies dealing with NMD have been performed in other genes [16–18].

We performed quantitative analysis of *DMD* transcripts using real-time PCR and the $2^{-\Delta\Delta C_t}$ method for data evaluation. Using the pair of primers 12Fx14R and 67Fx68R, the patients with mutations conserving the translational reading frame had a relative amount of *DMD* mRNA of 118% and 78% and those with a PTC mutation 48% and 17%, respectively. The quantity of dystrophin mRNA is different depending on whether it is analysed using primers complementary to its 5′ end or 3′ end. Our results indicate, that levels of the 5′ end mRNA regions mostly exceed those of the 3′ end regions. These differences can hardly be explained merely by ongoing and time-consuming transcription of the dystrophin gene (which may generate certain gradient of abundance of the 5′ end region over the 3′ end region [19]) and/or co-transcriptional splicing [20], since these differences are subtracted by relative evaluation using calibrator sample. These differences are rather due to prevalent 3'-to-5' exonucleolytic digestion of mRNA by cellular mechanisms that control mRNA turnover (see [21,22], for a review). Although these pathways generally include both 5′ → 3′ and 3′ → 5′, as well as endonucleolytic modes of digestion, the 3′ → 5′ digestion, which is mediated by exosome and triggered by deadenylation in non-PTC transcripts, clearly prevails in some of our results. In PTC-containing mRNA molecules, even more marked decrease can be observed in abundance of 3′ end regions, presumably due to the ongoing NMD, which may use also deadenylation-independent pathways of 3′ → 5′ transcript degradation [23]. Our interpretation of the origin of the differences is indirectly supported also by the earlier observation of equal levels of both 5′ and 3′ RT-PCR products both in normal controls, and in patients suspected for DMD without any detected mutation [24]. In conclusion, our present study illustrates the importance of NMD as a functional checkpoint in the multi-level process of dystrophin gene expression.

Acknowledgements

This work was supported by the project MSMT LC06023, and institutional support MSM0021622415 and VZ FNM 00064203.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.nmd.2009.08.011.

References

[1] Schouten JP, McElgunn CJ, Waaijer R, et al. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 2002;30:e57.
[2] Beggs AH, Koenig M, Boyce FM, Kunkel LM. Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. Hum Genet 1990;86:45–8.
[3] Roberts RG, Barby TF, Manners E, Bobrow M, Bentley DR. Direct detection of dystrophin gene rearrangements by analysis of dystrophin mRNA in peripheral blood lymphocytes. Am J Hum Genet 1991;49:298–310.
[4] Roest PA, Roberts RG, Sugino S, van Ommen GJ, den Dunnen JT. Protein truncation test (PTT) for rapid detection of translation-terminating mutations. Hum Mol Genet 1993;2:1719–21.
[5] Frischmeyer PA, Dietz HC. Nonsense-mediated mRNA decay in health and disease. Hum Mol Genet 1999;8:1893–900.
[6] Fajkusova L, Lukas Z, Tvrdikova M, et al. Novel dystrophin mutations revealed by analysis of dystrophin mRNA: alternative splicing suppresses the phenotypic effect of a nonsense mutation. Neuromuscul Disord 2001;11:133–8.
[7] Chrobakova T, Hermanova M, Kroupova I, et al. Mutations in Czech LGMD2A patients revealed by analysis of calpain3 mRNA and their phenotypic outcome. Neuromuscul Disord 2004;14:659–65.
[8] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) method. Methods 2001;25:402–8.
[9] Vondracek P, Hermanova M, Sedlackova J, et al. Charcot-Marie-Tooth neuropathy type 1A combined with Duchenne muscular dystrophy. Eur J Neurol 2007;14:1182–5.
[10] Winnard AV, Mendell JR, Prior TW, Florence J, Burghes AH. Frameshift deletions of exons 3–7 and revertant fibers in Duchenne muscular dystrophy: mechanisms of dystrophin production. Am J Hum Genet 1995;56:158–66.
[11] Kerr TP, Sewry CA, Robb SA, Roberts RG. Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay? Hum Genet 2001;109:402–7.
[12] Lev-Maor G, Ram O, Kim E, et al. Intronic Alus influence alternative splicing. PLoS Genet 2008;4:e1000204.
[13] Ferlini A, Muntoni F. The 5′ region of intron 11 of the dystrophin gene contains target sequences for mobile elements and three overlapping ORFs. Biochem Biophys Res Commun 1998;242:401–6.
[14] Ferlini A, Galie N, Merlini L, et al. A novel Alu-like element rearranged in the dystrophin gene causes a splicing mutation in a family with X-linked dilated cardiomyopathy. Am J Hum Genet 1998;63:436–46.
[15] Stehlikova K, Zapletalova E, Sedlackova J, et al. Quantitative analysis of CAPN3 transcripts in LGMD2A patients: involvement of nonsense-mediated mRNA decay. Neuromuscul Disord 2007;17:143–7.
[16] Schimpf S, Fuhrmann N, Schaich S, Wissinger B. Comprehensive cDNA study and quantitative transcript analysis of mutant OPA1 transcripts containing premature termination codons. Hum Mutat 2008;29:106–12.
[17] Nogales-Gadea G, Rubio JC, Fernandez-Cadenas I, et al. Expression of the muscle glycogen phosphorylase gene in patients with McArdle disease: the role of nonsense-mediated mRNA decay. Hum Mutat 2008;29:277–83.
[18] Boyer J, Crosnier C, Driancourt C, et al. Expression of mutant JAGGED1 alleles in patients with Alagille syndrome. Hum Genet 2005;116:445–53.
[19] Tennyson CN, Shi Q, Worton RG. Stability of the human dystrophin transcript in muscle. Nucleic Acids Res 1996;24:3059–64.
[20] Wee KB, Pramono ZAD, Wang JL, MacDorman KF, Lai PS, Yee WC. Dynamics of co-transcriptional pre-mRNA folding influences the induction of dystrophin exon skipping by antisense oligonucleotides. PLOS one 2008;3:1–14.
[21] Parker R, Song H. The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol 2004;11:121–7.
[22] Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. Nat Rev Mol Cell Biol 2007;8:113–26.
[23] Mitchell P, Tollervey D. An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3′ → 5′ degradation. Mol Cell 2003;11:1405–13.
[24] Hamed SA, Hoffman EP. Automated sequence screening of the entire dystrophin c DNA in Duchenne dystrophy: point mutation detection. Am J Med Genet B Neuropsychiatr Genet 2006;141B:44–50.

PERGAMON

Neuromuscular Disorders 11 (2001) 133–138

www.elsevier.com/locate/nmd

Novel dystrophin mutations revealed by analysis of dystrophin mRNA: alternative splicing suppresses the phenotypic effect of a nonsense mutation

Lenka Fajkusová[a], Zdeněk Lukáš[b], Miroslava Tvrdíková[a], Viera Kuhrová[a], Jiří Hájek[b], Jiří Fajkus[c,*]

[a]*Research Institute of Child Health, Černopolní 9, CZ-66262 Brno, Czech Republic*
[b]*Faculty Hospital Brno, Černopolní 9, CZ-66262 Brno, Czech Republic*
[c]*Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, CZ-61265 Brno, Czech Republic*

Received 7 February 2000; received in revised form 8 May 2000; accepted 21 June 2000

Abstract

The complete dystrophin mRNA sequence has been analyzed in 20 Duchenne muscular dystrophy and Becker muscular dystrophy patients. In 13 cases, deletions in mRNA were detected using reverse transcription-polymerase chain reaction and in another seven cases, point mutations were found using the protein truncation test. Sixteen patients diagnosed with Duchenne muscular dystrophy showed the presence of deletions or of nonsense point mutations. From four patients with the Becker muscular dystrophy phenotype, three cases were associated with deletions conserving the translational frame and one was associated with a nonsense mutation E1110X. In the case of the E1110X mutation, an alternative splicing of dystrophin mRNA (3485–3640del) was detected in this patient which included the E1110X mutation site (nucleotide 3536) and did not change the translation reading frame. Individual nonsense point mutations were characterized by sequence analysis, which showed five novel mutations with respect to those reported in the Cardiff Human Gene Mutation Database http://uwcm.web.cf.ac.uk/uwcm/mg/hgmd0.html and the Leiden muscular dystrophy pages http://www.dmd.nl/.

Keywords: Duchenne muscular dystrophy; Becker muscular dystrophy; Dystrophin mRNA; Protein truncation test; Reverse transcription-polymerase chain reaction; Mutation

1. Introduction

Duchenne and Becker muscular dystrophy (DMD and BMD, respectively), are allelic X-linked diseases caused by expression of defective dystrophin, a protein of molecular weight 427 kDa which is bound to the cytoplasmic membrane of smooth, heart and skeletal muscle cells [1]. The dystrophin gene is located on the X-chromosome where it occupies a region about 2.3 Mb long. The total dystrophin coding sequence of 11 220 bp is distributed over 79 exons. About 60% of cases of DMD and BMD result from gross deletions, 5% from duplications and the remaining 35% from point mutations of the dystrophin gene [2]. Generally, a precise correlation exists between a patient's clinical state and the dystrophin gene defect. Mutations which cause a translational frameshift, and therefore give rise to a protein missing its C-terminal region, are tightly correlated with a severe DMD phenotype. On the other hand mutations conserving the frame of translation and, consequently, resulting in a protein with internal deletions or duplications, are usually correlated with a milder BMD phenotype [3]. Large deletions and duplications can be detected using multiplex polymerase chain reaction (PCR) at the DNA level or by means of reverse transcription-PCR (RT-PCR) at the mRNA level [4–7]. It has been shown that most pathological point mutations connected with the DMD phenotype disrupt the translational reading frame, giving rise to truncated proteins. An effective mutation-detection method to monitor the integrity of the open reading frame is the protein truncation test (PTT). This rapid and sensitive method applies a combination of RT-PCR, in vitro transcription and in vitro translation to a sample of mRNA or total RNA [8]. As a number of tissue-specific isoforms of dystrophin, as well as variant splicing of ectopic transcripts has been detected, results of

* Corresponding author. Tel.: +420-5-4151-7199; fax: +420-5-4121-1293.
E-mail address: fajkus@ibp.cz (J. Fajkus).

0960-8966/01/$ - see front matter
PII: S0960-8966(00)00169-3

Table 1
Overview of patients

Patient	Phenotype	Born	Year of diagnosis	Family history	Wheelchair bound
1	DMD	1990	1995	Negative	No
2	DMD	1996	1998	Negative	No
3	DMD	1991	1996	Negative	No
4	DMD	1991	1995	Negative	No
5	DMD	1992	1996	Negative	No
6	DMD	1986	1991	Negative	Yes
7	DMD	1984	1988	Negative	Yes
8	DMD	1990	1995	Negative	No
9	DMD	1996	1998	Negative	No
10	DMD	1993	1996	Negative	No
11	DMD	1990	1994	Negative	No
12	DMD	1991	1996	Positive	No
13	DMD	1995	1998	Negative	No
14	DMD	1993	1997	Negative	No
15	DMD	1989	1992	Positive	Yes
16	DMD	1988	1992	Positive	Yes
17	BMD	1993	1999	Negative	No
18	BMD	1983	1991	Positive	No
19	BMD	1988	1998	Positive	No
20	BMD	1979	1990	Negative	No

transcript analysis have to be assessed in relation to the source material used [8]. Here we apply RT-PCR and PTT to scan the whole coding sequence of dystrophin in muscle tissues of our group of DMD and BMD patients and characterize the mutations detected, and we discuss these in relation to the patients' phenotypes. A case of alternative splicing which partially suppresses the phenotypic effect of a nonsense mutation is described.

2. Material and methods

2.1. Patients

The patients were classified according to the standard clinical symptoms, the level of creatine kinase, histopathological examination and immunocytochemical analysis of muscle sections. A brief description of patients is given in Table 1. All patients are unrelated.

2.2. Immunohistochemical analysis

Analysis of dystrophin mRNA was performed in muscle tissues of patients diagnosed with DMD or BMD by previous immunodetection of dystrophin in sections of muscle tissue. Muscle biopsies were snap frozen in a propane-butane mixture cooled by liquid nitrogen. Cryosections were examined by conventional histological, histochemical and immunohistochemical methods [9]. Four monoclonal antibodies were used during this study: one to β-spectrin and three to dystrophin. The monoclonal antibodies to dystrophin are specific for epitopes in the N-terminal domain (amino acids 308–351, corresponding to exons 10–12, DYS3), for the rod domain (amino acids 1181–1388, corresponding to exons 26–30, DYS1) and for the C-terminal domain (amino acids 3669–3685, corresponding to exon 79, DYS2). The avidin-DH biotinylated horseradish peroxidase method was used for immunostaining. Spectrin is localized at the periphery of muscle fibres like dystrophin and labeling of this protein is thus an excellent control for muscle membrane integrity in fibres which appear negative for dystrophin labeling. All four antibodies were from Novocastra Laboratories, UK.

2.3. RNA isolation and RT-PCR

Muscle tissues stored in liquid nitrogen or on dry ice were homogenized in a mortar and pestle under liquid nitrogen and the resulting powder was used for total RNA preparation by the RNeasy Total RNA kit (Qiagen). RT-PCR was performed with the Titan™ One Tube RT-PCR System (Boehringer–Mannheim). The total dystrophin mRNA was divided into ten regions which were amplified in ten RT-PCR reactions (Table 2). The products of primary PCRs of each region were amplified in ten secondary PCRs using the Expand High Fidelity PCR System (Boehringer–Mannheim). Products of individual PCRs were electrophoretically separated in agarose gels and visualized in UV light after ethidium bromide staining. Primer sequences were taken from the literature [7,10]. In cases where a deletion overlapped a boundary between adjoining regions (which resulted in lack of product in secondary PCR), it was necessary to carry out RT-PCR and secondary PCR with a proper combination of primers so that the deletion was located between the primer sequences (a 'through-amplification').

In some cases, deletions were characterized by DNA sequencing of purified PCR products on an ABI PRISM 310 sequencer (Perkin–Elmer).

Table 2
Partitioning of dystrophin mRNA for RT-PCR and PTT examination

Amplified region of mRNA	Position of amplified region on dystrophin cDNA	Corresponding exons or exon parts
1	212–1194	1–10
2	1133–2396	9–18
3	2342–3594	17–25
4	3527–4810	25–33
5	4766–5897	33–40
6	5844–6583	40–44
7	6431–7657	43–51
8	7616–8869	51–58
9	8786–10046	58–68
10	9974–11296	67–79

2.4. Protein truncation test

For performing PTT, one of the primers for secondary PCR has to contain 5′-terminal 40 nt sequence with a bacteriophage T7 promoter and an eukaryotic translation initiation signal [10]. In vitro transcription and in vitro translation were performed by the TNT T7 Coupled Reticulocyte Lysate System (Promega). [^{35}S] Methionine (Amersham) was added to the reaction for autoradiographic detection of translation products after separation by 15% discontinuous sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Regions producing truncated or no proteins were sequenced. DNA sequencing of purified PCR products was performed on an ABI PRISM 310 sequencer (Perkin–Elmer).

3. Results

3.1. Immunohistochemical analysis

Muscle tissues of 20 DMD and BMD patients were examined. Biopsy findings were characterized in all cases as a typical myogenic pattern characterized by profound structural alterations of muscle fibres, presence of regenerating fibres and degrees of both endomysial and perimysial fibrosis.

Neither necrotic with negative reactivity for spectrin nor regenerating muscle fibres were classified. The immunohistochemical labeling patterns obtained using antibodies DYS1 and DYS2 were divided into six categories according to the most prominent features observed: (1): no labeling on any muscle fibres or labeling on occasional muscle fibres

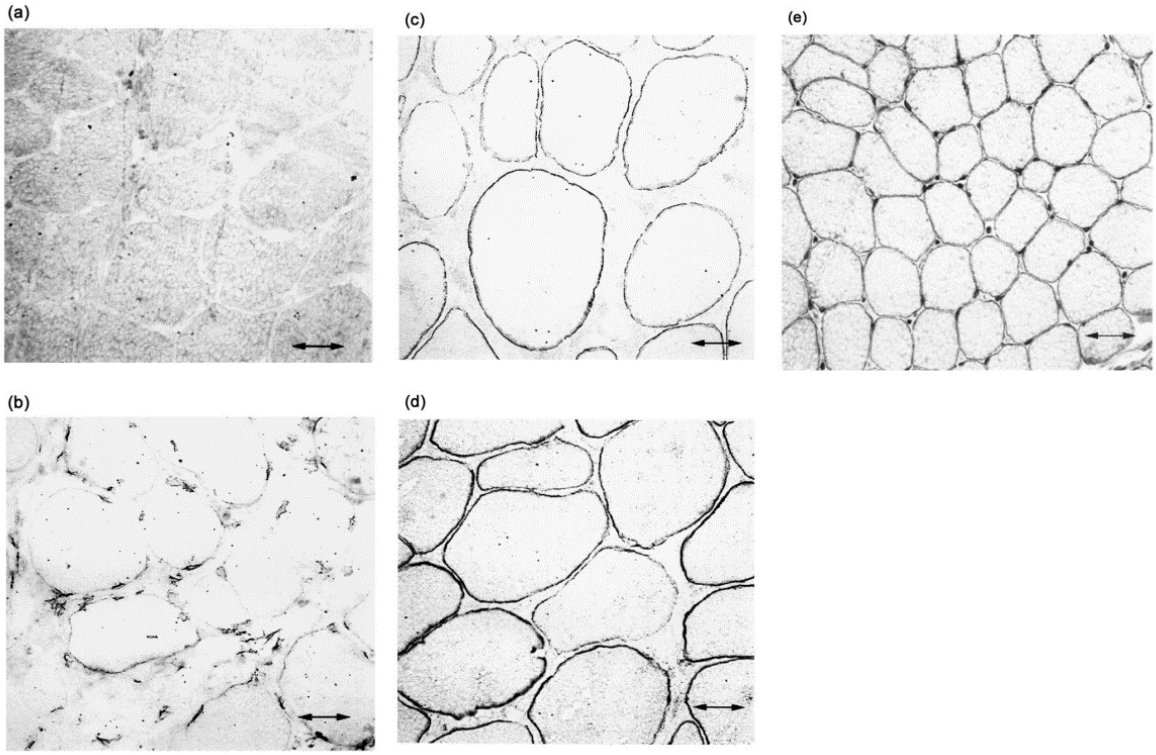


Fig. 1. Immunohistochemical localization of dystrophin in the muscles of DMD/BMD patients using C-terminal antibody (original magnification 200×, scale bars = 50 μm), DMD patient 2 (a), DMD patient 6 (b), BMD patient 20 (c), BMD patient 18 (d) and control muscle (e).

Table 3
Results of analyses of dystrophin mRNA

Patient	Phenotype	Mutation detected by analysis of dystrophin mRNA	Characterization of mutation by DNA sequencing	Exon(s) affected
1	DMD	Deletion in region 7	6647–6822del	45
2	DMD	Deletion in region 7 and 8	6647–8235del	45–54
3	DMD	Deletion in region 7 and 8	6647–7868del	45–52
4	DMD	Deletion in region 7 and 8	6823–7750del	46–51
5	DMD	Deletion in region 7 and 8	6647–7868del	45–52
6	DMD	Deletion in region 7 and 8	Not performed	
7	DMD	Deletion in region 7 and 8	Not performed	
8	DMD	Deletion in region 8 and 9	7869–9145del	53–59
9	DMD	Deletion in region 1	Not performed	
10	DMD	Deletion in region 10	Not performed	
11	DMD	Point mutation in region 1	791 C → T Arg (CGA) → STOP (TGA) R195X	7
12	DMD	Point mutation in region 1	791 C → T Arg (CGA) → STOP (TGA) R195X	7
13	DMD	Point mutation in region 2	1532 C → T Gln (CAA) → STOP (TAA) Q442X	10
14	DMD	Point mutation in region 3	2484 T → G Leu (TTA) → STOP (TGA) L759X	18
15	DMD	Point mutation in region 10	10302 C → A Ser (TCA) → STOP (TAA) S3365X	70
16	DMD	Point mutation in region 10	10349 C → T Arg (CGA) → STOP (TGA) R3381X	70
17	BMD	Deletion in region 7	7121–7306del	48
18	BMD	Deletion in region 7	6647–7120del	45–47
19	BMD	Deletion in regions 7 and 8	6647–8425del	45–55
20	BMD	Point mutation in region 3 associated with alternative splicing	3536 G → T Glu (GAA) → STOP (TAA) E1110X	25
			Alt. splic. 3485–3640del	25

(generally <5% of fibres) (patients 1–5, 7–10 and 12–16) (Fig. 1a); (2): sectional labeling of decreased intensity on <20% of muscle fibres (patients 6 and 11) (Fig. 1b); (3): labeling of decreased intensity on approximately 50% of muscle fibres (patient 17); (4): labeling of decreased intensity on approximately 80% of muscle fibres (patient 20) (Fig. 1c); (5): labeling is uniform, but a few muscle fibres are labeled at decreased intensity (patients 18 and 19) (Fig. 1d); (6): labeling at normal intensity (control muscle) (Fig. 1e).

The immunohistochemical labeling patterns obtained using DYS3 antibody were very heterogeneous in patients with a DMD diagnosis (from 0 to 90% of positive fibres). In BMD patients, the pattern obtained using DYS3 was similar to immunodetection patterns of DYS1 and DYS2.

3.2. Analysis of dystrophin mRNA

The results of the corresponding mRNA analyses are given in Table 3.

In 13 patients, deletions were found by means of RT-PCR. Ten of them were diagnosed as DMD (six deletions were characterized by DNA sequencing confirming a translational frameshift), and in the other three patients diagnosed as BMD, DNA sequencing showed deletions

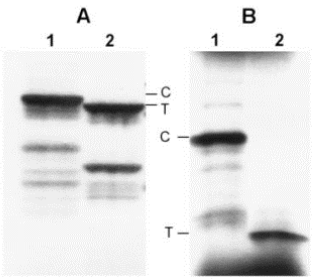


Fig. 2. Example PTTs of dystrophin mRNA: (A) region 3. Lane 1, control PTT; lane 2, PTT of patient 20; (B) region 1. Lane 1, control PTT; lane 2, PTT of patient 12. Control (C) and truncated (T) products are marked.

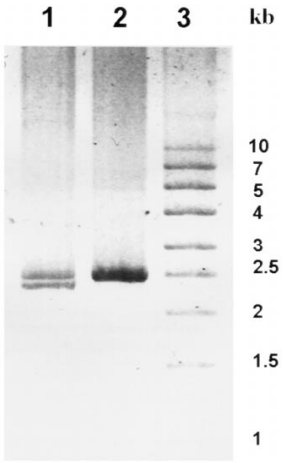


Fig. 3. Fragments of DNA detected after amplification of dystrophin mRNA region 3-4. Lane 1, PCR product of patient 20; lane 2, control PCR product; lane 3, size marker.

keeping the translation frame. Nonsense point mutations were detected in seven patients. Six of these were diagnosed as DMD and one as BMD.

The results of PTT in patients 11–16 and 20 showed the presence of point mutations which resulted in termination of translation (see Fig. 2 for example). Their exact location was determined by sequence analysis.

Notably, in patient 20 diagnosed as BMD, molecular genetic analysis revealed the presence of the nonsense mutation E1110X, corresponding rather to a DMD phenotype. Further analysis showed the presence of an alternative mRNA splicing: amplification of dystrophin mRNA region 3–4 (positions 2342–4810 on dystrophin cDNA, see Table 2) resulted in two RT-PCR products, while a single product was amplified from a control sample (Fig. 3). Both PCR products were excised from the agarose gel after electrophoretic separation and the purified DNAs were directly sequenced. The sequence analysis of the longer PCR product revealed the presence of point mutation E1110X compared to control mRNA, while the shorter PCR product showed a deletion (3485–3640del) including the point mutation E1110X (nucleotide position 3536) and keeping the translation reading frame.

4. Discussion

In this work we used a combination of RT-PCR, PTT and DNA sequencing to detect and characterize mutations of dystrophin mRNA in our DMD and BMD patients. Molecular changes corresponding to DMD or BMD diagnosis were found in all 20 mRNA samples isolated from muscle tissues – 13 cases of deletions (two unrelated patients had an identical deletion) and seven cases of point mutations (two unrelated patients had an identical point mutation) were detected.

We also describe four novel mutations in the dystrophin gene (R195X, Q442X, L759X and E1110X). The mutation S3365X was described by our group previously [11].

The results of these molecular genetic analyses were correlated with the clinical state of individual patients: 16 patients with a more severe form of disease, DMD, displayed deletions (patients no. 1–10; in patients 1–5 and 8, DNA sequencing was performed which confirmed the change in translation reading frame) or nonsense point mutations (patients no. 11–16) giving rise to a STOP codon and thus causing premature termination of translation.

Analysis of mRNA in the four patients diagnosed with BMD, the milder form of the disease, showed the following results. Patients 17–19 showed deletions of different sizes which do not affect the translation frame. Patient 20 showed a nonsense point mutation which would normally result in a DMD phenotype. The BMD phenotype of this patient was explained by the presence of an alternatively spliced mRNA, in which the deletion (3485–3640del) bridges the non-sense mutation and thus partially suppresses its effect. Immunohistochemical analysis was performed in the patient's age of 18 and it was his first immunohistochemical testing. A generalized low level of dystrophin, a mild decrease of dystrophin associated proteins (DAPs) and a strong staining for utrophin have been observed. Exon skipping has been observed in a number of human genetic disease mutations and attributed to the disruption of exonic splicing enhancers (see [12] for review). Mechanism of induction of partial skipping of the dystrophin exon 27 harbouring a nonsense mutation E1211X (G3839T) has been studied using in vitro splicing system [13]. The authors showed that the introduction of a T nucleotide to the purine-rich splicing enhancer suppresses its activity. We therefore presume a similar mechanism in case of nonsense mutation E1110X (G3536T transversion) found in patient 20, which occurs at the beginning of purine-rich tract of exon 25.

The overview of our results shows that 11 of 13 deletions detected by dystrophin mRNA analysis in BMD/DMD patients are located within regions 7–9 (patients 1–8 and 17–19). From these, two unrelated patients (3 and 5) had identical deletions and six started at the same nucleotide position (6647 bp) of dystrophin cDNA (patient 1–3, 5, 18, 19). The starting point of the deletion in patient 8 coincides with the endpoint of the deletion in patients 3 and 5 in dystrophin cDNA (7869 bp). Similarly, the deletion startpoint in patient 4 coincides with the deletion endpoint in patient 1 (6823 bp). Preferential breakpoints in the dystrophin gene have been observed also in the previous reports, (e.g. [7,14]). This concentration of mutation hotspots at such foci may reflect specific features of chromatin structure of the dystrophin gene, possibly its arrangement in loop domains, as described recently in the *hprt* gene [15].

Immunohistochemical analysis of dystrophin and several other proteins (α, β, γ, δ-sarkoglycans and α, β-dystroglycans), which are involved in multiprotein complex formation with dystrophin was performed. Their strong reduction

in the sarcolemma was observed in all cases where dystrophin lacking the C-terminal was present corresponding to published results [16]. The loss of the DAPs in the sarcolemma causing disruption of linkage between the subsarcolemmal cytoskeleton and extracellular matrix is presumed to be the cause of the severe phenotype of DMD patients. The reduction of dystrophin and DAPs was milder or negligible in BMD patients indicating that the dystrophin rod domain (where deletions associated with BMD mostly occur) is not crucial for interaction with the DAPs.

Acknowledgements

The authors thank Š. Vejválková and F. Cisárik for clinical collaboration, J. Zaorálková for performing immunodetection assays and R. Hancock for reviewing the manuscript. Sequencing service of J. Klánová from the Laboratory of Plant Molecular Physiology, Masaryk University Brno is appreciated. This work was supported by projects of IGA MZCR 3700-3, IGA MZCR M/19-3 and IGA MZCR NA/5227-3.

References

[1] Zubrzycka-Gaarn EE, Bulman DE, Karpati G, et al. The Duchenne muscular dystrophy gene product is localized in sarcolemma of human skeletal muscle. Nature 1988;333(6172):466–469.

[2] Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell 1987;50:509–517.

[3] Monaco AP, Bertelson CJ, Liechti-Gallati S, Moser H, Kunkel LM. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics 1988;2(1):90–95.

[4] Chamberlain JS, Gibbs RA, Ranier JF, Nguyen PN, Caskey CT. Deletion screening of the Duchenne muscular dystrophy locus via multiplex PCR amplification. Nucleic Acids Res 1988;16:11141–11156.

[5] Beggs A, Koenig H, Boyce FM, Kunkel LM. Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. Hum Genet 1990;86:45–48.

[6] Covone AE. Screening Duchenne and Becker muscular dystrophy patients for deletions in 30 exons of the dystrophin gene by three-multiplex PCR. Am J Hum Genet 1992;51:675–677.

[7] Roberts RG, Barby TFM, Manners E, Bobrow M, Bentley DR. Direct detection of dystrophin gene rearrangements by analysis of dystrophin mRNA in peripheral blood lymphocytes. Am J Hum Genet 1991;49:298–310.

[8] Cooper DN, Berg LP, Kakkar VV, Reis J. Ectopic (illegitimate) transcription: new possibilities for the analysis and diagnosis of human genetics disease. Ann Med 1994;26:9–14.

[9] Lukáš Z, Bednařík J, Fajkusová L, et al. Strategy for diagnosis of muscular dystrophies using methods to detection of muscle cytoskeletal components. Elec J Pathol Histol 1999;5.2:992–1007.

[10] Gardner RJ, Bobrow M, Roberts RG. The identification of point mutations in Duchenne muscular dystrophy patients by using reverse-transcription PCR and the protein truncation test. Am J Hum Genet 1995;57(2):311–320.

[11] Fajkusová L, Pekařík V, Hájek J, Kuhrová V, Blažková M, Fajkus J. Characterization of two non-sense mutations in the human dystrophin gene. J Neurogenet 1998;12(3):183–189.

[12] Blencowe BJ. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. Trends Biochem Sci 2000;25(3):106–110.

[13] Shiga N, Takeshima Y, Sakamoto H, et al. Disruption of the splicing enhancer sequence within exon 27 of the dystrophin gene by a non-sense mutation induces partial skipping of the exon and is responsible for Becker muscular dystrophy. J Clin Invest 1997;100(9):2204–2210.

[14] Danieli GA, Mioni F, Muller CR, Vitiello L, Mostacciuolo ML, Grimm T. Patterns of deletions of the dystrophin gene in different European populations. Hum Genet 1993;91(4):342–346.

[15] Fajkus J, Nicklas JA, Hancock R. DNA loop domains in a 1.4-Mb region around the human hprt gene mapped by cleavage mediated by nuclear matrix-associated topoisomerase II. Mol Gen Genet 1998;260:410–416.

[16] Matsumura K, Campbell KP. Dystrophin-glycoprotein complex: its role in the molecular pathogenesis of muscular dystrophies. Muscle Nerve 1994;17(1):2–15.

PERGAMON

Neuromuscular Disorders 17 (2007) 143–147

www.elsevier.com/locate/nmd

Quantitative analysis of CAPN3 transcripts in LGMD2A patients: Involvement of nonsense-mediated mRNA decay

Kristýna Stehlíková [a], Eva Zapletalová [a,b], Jana Sedláčková [a], Markéta Hermanová [c], Petr Vondráček [d], Tat'ána Maříková [e], Radim Mazanec [f], Josef Zámečník [g], Stanislav Voháňka [h], Jiří Fajkus [b], Lenka Fajkusová [a,b,*]

[a] *University Hospital Brno, Centre of Molecular Biology and Gene Therapy, Brno, Czech Republic*
[b] *Masaryk University, Faculty of Science, Department of Functional Genomics and Proteomics, Brno, Czech Republic*
[c] *University Hospital Brno, Institute of Pathology, Brno, Czech Republic*
[d] *University Hospital Brno, Department of Child Neurology, Brno, Czech Republic*
[e] *University Hospital Motol, Institute of Biology and Medical Genetics, 2nd Faculty of Medicine, Prague, Czech Republic*
[f] *University Hospital Motol, Clinic of Child Neurology, 2nd Faculty of Medicine, Prague, Czech Republic*
[g] *University Hospital Motol, Department of Pathology and Molecular Medicine, 2nd Faculty of Medicine, Prague, Czech Republic*
[h] *University Hospital Brno, Clinic of Neurology, Brno, Czech Republic*

Received 10 July 2006; received in revised form 6 September 2006; accepted 11 October 2006

Abstract

Limb girdle muscular dystrophy type 2A (LGMD2A) is caused by single or small nucleotide changes widespread along the CAPN3 gene, which encodes the muscle-specific proteolytic enzyme calpain-3. About 356 unique allelic variants of CAPN3 have been identified to date. We performed analysis of the CAPN3 gene in LGMD2A patients at both the mRNA level using reverse transcription-PCR, and at the DNA level using PCR and denaturing high performance liquid chromatography. In four patients, we detected homozygous occurence of a missense mutation or an in-frame deletion at the mRNA level although the DNA was heterozygous for this mutation in conjuction with a frame-shift mutation. The relationship observed in 12 patients between the quantity of CAPN3 mRNA, determined using real-time PCR, and the genotype leads us to propose that CAPN3 mRNAs which contain frame-shift mutations are degraded by nonsense-mediated mRNA decay. Our results illustrate the importance of DNA analysis for reliable establishment of mutation status, and provide a new insight into the process of mRNA decay in cells of LGMD2A patients.

Keywords: LGMD2A; CAPN3; Calpain-3; Real-time RT-PCR; Nonsense-mediated mRNA decay

1. Introduction

Limb girdle muscular dystrophies (LGMDs) are a clinically and genetically heterogeneous group of disorders characterised by progressive involvement of proximal limb girdle muscles. Autosomal recessive LGMDs of type 2 include at least 10 different genetic entities. LGMD2A, the most prevalent form, is caused by mutations in the CAPN3 gene which is localized at 15q15.1-q21.1 and encodes the proteolytic enzyme calpain-3 [1]. The CAPN3 gene comprises 24 exons and covers a genomic region of 50 kb. Its transcript is spliced to a 3.5 kb muscle-specific mRNA, which is translated to a 94 kDa protein. Calpain-3 is a member of the calpain

* Corresponding author. Tel.: +420 532234625. fax: +420 532234623.
E-mail address: lfajkusova@fnbrno.cz (L. Fajkusová).

0960-8966/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.nmd.2006.10.001

superfamily, which processes enzymes involved in signalling pathways, transcription factors, and cytoskeletal proteins thereby modulating their activities [2].

In our previous studies [3,4], we described results of CAPN3 gene analysis and discussed the relation between the mutations detected and the patients' phenotype. In this study, we completed our analyses on DNA (using PCR, denaturing high performance liquid chromatography, and sequencing) and mRNA (using reverse transcription, PCR, and sequencing) to have both analyses in each patient. In four patients carrying the genotypes P82L/T184RfsX36, P82L/D772delK773NfsX3, F200_L204del/ T184RfsX36 and R490W/T184RfsX36, mRNA analysis detected the homozygous occurrence of the missense mutation or the in-frame deletion, but DNA analysis detected heterozygous occurrence of these together with the frame-shift mutation (T184RfsX36 or D772delK773NfsX3) which was not detected by mRNA analysis. It is probable that the process termed nonsense mediated mRNA decay (NMD) induced degradation of mRNA carrying these frame-shift mutations [5–8]. Further, we performed quantitative analysis of CAPN3 mRNA in 12 LGMD2A patients: 2 patients had a genotype missense/missense mutation; 1 in-frame deletion/missense mutation; 3 missense/frame-shift mutation, 1 in-frame deletion/frame shift mutation, 1 in-frame deletion/nonsense mutation, 3 frame-shift/frame-shift mutation and 1 patient had the genotype in-frame deletion/intron mutation. We evaluated the relationship between the amount of CAPN3 mRNA and the genotype.

2. Patients and methods

2.1. Patients

Clinical findings for patients 2, 3, 4, 6–12 were presented in our previous works [3,4]. Patients 1 and 5 are newly described here. Age of disease onset was 8 and 5 years, respectively. At present, patient 1 is 33 years old and is wheel-chair bound since the age of 30. Patient 5 is 7 years old, and suffers from mild proximal muscle weakness in the upper and lower extremities. In both patients, immunohistochemical analysis of calpain-3 using Western blot showed 94 kDa band of a low intensity.

2.2. Immunohistochemical analysis, Western blot, RNA analysis, and DNA analysis

Immunohistochemical analysis of muscle proteins in tissue sections, Western blot of calpain-3 as well as RNA isolation from muscle tissue, reverse transcription, and amplification of calpain-3 mRNA, DNA isolation, amplification of all CAPN3 exons and DHPLC analysis have been described in our previous studies [3,4].

2.3. Real-time PCR

100 ng of RNA were used for reverse transcription (RT) using the SuperScript™ First-Strand Synthesis System (Invitrogen). The cDNA synthesis reactions were primed by random hexamers. RT products were amplified using SybrGreen PCR Core Reagents (Applied Biosystems) under the following cycling conditions: initial denaturation: 95°C/10 min; followed by 40 cycles: 95 °C/10 s, 55 °C/15 s, 72 °C/30 s. Fluorescence was monitored at 72 °C and specificity of amplification was evaluated by melting curve analysis. The analysis of each RNA sample was performed in two parallel PCRs in two or three independent experiments. Primer sequences are shown in Table 1. Amplification of the CAPN3 mRNA was performed using the set of three couples of primers that amplified three various regions of calpain-3 mRNA. Relative amount of CAPN3 mRNA was taken as the average of the results of relative quantifications of three different CAPN3 mRNA regions. The data obtained were analysed by the $2^{-\Delta\Delta Ct}$ method that evaluates the target gene expression with respect to the reference gene expression and normalizes this value relative to expression in a calibrator sample [9]. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), hypoxantine guanine phosphoribosyltransferase 1 (HPRT), and ABL1 protein tyrosine kinase (ABL) were used as the reference genes and the RNA of patient 1 was used as the calibrator.

3. Results

3.1. Analysis of calpain3 mRNA and DNA

In 12 LGMD2A patients, mRNA analysis of the CAPN3 gene was subsequently confirmed by DNA

Table 1
Primers used for real-time PCR

Primer	Sequence of forward primer (5′ → 3′direction)	Primer	Sequence of reverse primer (5′ → 3′direction)	Length of PCR product
GAPDH-F	cctgcaccaccaactgctta	GAPDH-R	gaggcagggatgatgttctg	178
HPRT-F	gcagactttgctttcctgg	HPRT-R	tcaagggcatatcctacaacaa	153
ABL-F	tggagataacactctaagcataactaaaggt	ABL-R	gatgtagttgcttgggaccca	123
CAPN-1F	tcctacgaagctctgaaagg	CAPN-1R	ttccataggtcatgttcgtg	183
CAPN-2F	gacgatgaccctgatgactc	CAPN-2R	aggcgttgtacaggaagaag	186
CAPN-3F	ttcatctgctgcttcgttag	CAPN-3R	gctttggaaatagagggtga	193

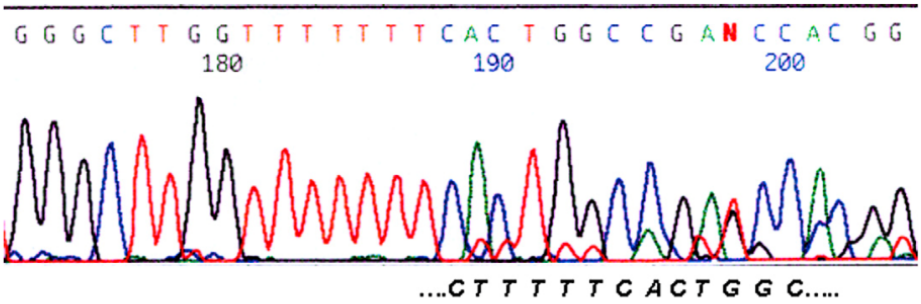


Fig. 1. Example of sequencing of the mRNA region with the in-frame mutation K595_K596. The sequence of the mRNA bearing standard sequence in this region, and mutation R748X further to the 3′ end, is represented by lower peaks and is given below in bold italics.

analysis. However, in patients 3, 4, 5, and 6 the results of mRNA and DNA analysis were not concordant. We detected homozygous occurence of the mutations P82L (patients 3 and 4), F200_L204del (patient 5) and R490W (patient 6) by mRNA analysis, but analysis at the DNA level did not confirm the homozygous status of these mutations. DHPLC analysis of all CAPN3 exons and subsequent sequencing detected the mutation T184RfsX36 in addition to the mutation P82L (patient 3), D772delK773NfsX3 in addition to P82L (patient 4), T184RfsX36 in addition to F200_L204del (patient 5) and T184RfsX36 in addition to R490W (patient 6). In the case of patient 12 (genotype K595_K596del/R748X), DNA as well as mRNA analysis gave concordant results, but it is evident (Fig. 1) that the amounts of mRNA carrying K595_K596del and R748X are not similar. In the case of patient 11, we detected heterozygous occurence of the in-frame deletion F200_L204del by mRNA analysis but no second mutation. DNA analysis detected the mutation in intron 13 (1746-20C > G) but this CAPN3 variant probably has no noticeable phenotypic effect (Leiden muscular dystrophy pages, www.dmd.nl). The mutation Y75C detected in case of patient 1 was not described so far. The tyrosine in the position 75 is the second amino acid of the calpain-3 catalytic domain. Sequence alignment of this region of several species (human, macaca, pig, mouse, rat, and chicken) confirmed the conservation of tyrosine in this position. Some species like sheep and bovine have another aromatic amino acid, phenylalanine, in this position. The phenotype of patient 1 is rather similar to phenotype of patients carrying mutations affecting translational reading frame. The results of mRNA and DNA analyses are shown in Table 2.

3.2. Real-time PCR

Quantitative analyses of CAPN3 mRNA using real-time PCR and the $2^{-\Delta\Delta Ct}$ method were evaluated in relation to the mRNAs of three housekeeping genes GAPDH, HPRT, and ABL. Results of these analyses are shown in Table 3. Relative to GAPDH, the patients with genotypes missense/missense or in-frame deletion/missense had an average amout of CAPN3 mRNA 1.07; the patients with genotype missense/frame-shift, in-frame deletion/frame-shift or in-frame deletion/nonsense had an average amount 0.39; and the patients with genotypes frame-shift/frame-shift had an average amount 0.03. Relative to HPRT, these values were 0.69; 0.41; and 0.01, respectively, and relative to ABL,

Table 2
Mutations detected in the set of LGMD2A patients

Patient	Mutation detected in allele 1	Position of mutation	Mutation detected in allele 2	Position of mutation	Age of disease onset
1	c.224A>G; p.Y75C	Exon 1	c.224A > G; p.Y75C	Exon 1	8
2	c.133G>A; p.A45T	Exon 1	c.133G > A; p.A45T	Exon 1	15
3	c.245C>T; p.P82L	Exon 1	c.550delA; p.T184RfsX36	Exon 4	12
4	c.245C>T; p.P82L	Exon 1	c.2314-2317del; p.D772delK773NfsX3	Exon 22	13
5	c.550delA; p.T184RfsX36	Exon 4	c.598-612del; p.F200_L204del	Exon 4	5
6	c.550delA; p.T184RfsX36	Exon 4	c.1468C > T; p.R490W	Exon 11	12
7	c.550delA; p.T184RfsX36	Exon 4	c.550delA; p.T184RfsX36	Exon 4	10
8	c.550delA; p.T184RfsX36	Exon 4	c.1981delA; p.I661X	Exon 17	7
9	c.550delA; p.T184RfsX36	Exon 4	c.1722delC; p.F574FfsX21	Exon 13	10
10	c.598-612del; p.F200_L204del	Exon 4	c.2245A > C; p.N749H	Exon 21	17
11	c.598-612del; p.F200_L204del	Exon 4	c.1746-20C > G	Intron 13	25
12	c.1783-1788del; p.K595_K596del	Exon 15	c.2242C > T; p.R748X	Exon 21	20

Table 3
Results of real-time PCR: Relative amount of CAPN3 transcripts relative to three housekeeping genes using CAPN3 mRNA of patient 1 as calibrator

Patient	Relative amount of CAPN3 mRNA with respect to			Average value
	GAPDH	HPRT	ABL	
1	1.00	1.00	1.00	1.00
2	1.2047 ± 0.0084	0.6462 ± 0.0045	0.5804 ± 0.0040	0.810 ± 0.010
3	0.285 ± 0.075	0.354 ± 0.082	0.257 ± 0.059	0.30 ± 0.13
4	0.313 ± 0.044	Not evaluated	0.238 ± 0.035	0.276 ± 0.056
5	0.435 ± 0.014	0.405 ± 0.012	0.374 ± 0.011	0.405 ± 0.022
6	0.604 ± 0.053	0.415 ± 0.037	0.732 ± 0.071	0.584 ± 0.087
7	0.0139 ± 0.0093	0.0066 ± 0.0044	0.0187 ± 0.0058	0.013 ± 0.012
8	0.0231 ± 0.0039	0.0097 ± 0.0038	0.0287 ± 0.0035	0.0205 ± 0.0067
9	0.0412 ± 0.0072	0.0118 ± 0.0021	0.0176 ± 0.0031	0.0235 ± 0.0081
10	0.995 ± 0.037	0.417 ± 0.015	0.569 ± 0.021	0.660 ± 0.045
11	0.795 ± 0.044	0.876 ± 0.049	1.092 ± 0.061	0.921 ± 0.089
12	0.295 ± 0.018	0.483 ± 0.035	0.638 ± 0.046	0.472 ± 0.061

0.72; 0.45; 0.02, respectively. When the results of all reference genes were taken into account, these values were 0.82; 0.41; and 0.02, respectively.

4. Discussion

The relative abundance and lifetime of any RNA molecule is governed by the balance between the rates of its synthesis and degradation. Increasing the destruction rate is perhaps the fastest means of modulating RNA level. Therefore, although the field of RNA decay is relatively young [10], its importance for regulation of RNA metabolism has become swiftly recognized as an apparent equivalent to proteasome-mediated degradation of proteins. NMD is a post-transcriptional process that rapidly degrades mRNA with premature translation termination codons. This limits the production of potentially toxic truncated proteins. The NMD pathway is initiated during translation when the responsible proteins discriminate premature from normal translation termination and mark the transcript for rapid decay. The distinction between a normal and a premature translation stop codon is made on the basis of its location with respect to the last exon–exon junction: If the termination codon is positioned >50–55 nucleotides upstream it is considered premature, and the mRNA is targeted for rapid decay [11]. NMD is an evolutionarily selected response because it reduces the severity of some genetic diseases caused by truncating dominant-negative mutations [12,13]. However NMD can also increase the severity of genetic diseases when truncating mutation does not destroy completely normal function. Thus, evidence suggests that NMD worsens the phenotype of many genetic diseases including e.g. Duchenne muscular dystrophy [14] and Ullrich's congenital muscular dystrophy [15,16]. Here we examined for the first time potential involvement of NMD in pathology of LGMD using a set of patients carrying various combinations of mutated alleles. To do this, we complemented our previous work [3,4] by performing DNA and mRNA analyses in all the patients presented and, especially, by quantitative analysis of their CAPN3 transcripts.

Positions of all frame-shift and nonsense mutations detected in our patients match the above mentioned exon–exon-junction rule [11] and thus constitute potential targets of NMD. In patients 3, 4, 5, and 6, mRNA analysis detected homozygous occurence of a mutation with no effect on translation reading frame, but DNA analysis detected a further frame-shift mutation. Degradation of mRNA carrying this frame-shift mutation was probably caused by NMD. It is interesting that in patient 12 mRNA analysis detected both the in-frame deletion and the nonsense mutation, but it is clear (Fig. 1) that the level of the mRNA carrying R748X is lower than that of the mRNA carrying K595_K596del. Further, we performed quantitative analysis of CAPN3 transcripts using real-time PCR using the $2^{-\Delta\Delta Ct}$ method for data evaluation; the amount of CAPN3 mRNA was determined with respect to GAPDH, HPRT, and ABL mRNAs, and then relative quantification was performed by reference to patient 1 chosen as the calibrator (quantity of CAPN3 mRNA = 1). Relative to GAPDH, a common housekeeping gene, the patients with genotypes missense/missense or in-frame/missense had an average amout of CAPN3 mRNA 1.07; the patients with a missense mutation or an in-frame deletion on one allele and a frame-shift mutation or a nonsense mutation on the second allele had an average amount 0.39; and the patients with genotypes frame-shift/frame-shift mutation had an average amount 0.03. Quantitative analysis of CAPN3 mRNA in 5 control samples (mRNA from patients without suspicion of inherited neuromuscular disease) showed relative transcript levels from 1.00 to 1.35.

While in the presence of a single frame-shift or a nonsense mutation, the level of CAPN3 transcripts

decreases to about 40%, its presence on both alleles results in just 2% of the reference value. Accordingly, in the case of a single frame-shift or nonsense mutated allele, the transcript of the other allele (maintaining the reading frame) is strongly overrepresented. These quantitative differences in CAPN3 transcript levels, which reflect precisely the type of mutations involved, provide the first evidence for a role of NMD in destruction of aberrant transcripts in LGMD patients. Our results thus illustrate the practical importance of DNA analysis for a reliable establishment of mutation status, and provide an interesting insight into the processes of mRNA decay in cells of LGMD2A patients.

Acknowledgements

This work was supported by IGA MHCR, Project No. NR/8087-3, and the institutional support MSM0021622415. The authors thank Professor Ronald Hancock, Hôtel-Dieu Hospital, Quebec, Canada, for the manuscript revision.

References

[1] Beckmann JS, Richard I, Hillaire D, Broux O, Antignac C, Bois E, et al. A gene for limb-girdle muscular dystrophy maps to chromosome 15 by linkage. CR Acad Sci III 1991;312:141 8.
[2] Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, Bourg N, et al. Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 1995;81:27 40.
[3] Chrobakova T, Hermanova M, Kroupova I, Vondracek P, Marikova T, Mazanec R, et al. Mutations in Czech LGMD2A patients revealed by analysis of calpain3 mRNA and their phenotypic outcome. Neuromuscul Disord 2004;14:659 65.
[4] Hermanova M, Zapletalova E, Sedlackova J, Chrobakova T, Letocha O, Kroupova I, et al. Analysis of histopathologic and molecular pathologic findings in Czech LGMD2A patients. Muscle Nerve 2006;33:424 32.
[5] Singh G, Lykke-Andersen J. New insights into the formation of active nonsense-mediated decay complexes. Trends Biochem Sci 2003;28:464 6.
[6] Lykke-Andersen J, Shu MD, Steitz JA. Communication of the position of exon exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science 2001;293:1836 9.
[7] Kim VN, Kataoka N, Dreyfuss G. Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon exon junction complex. Science 2001;293:1832 6.
[8] Le Hir H, Gatfield D, Izaurralde E, Moore MJ. The exon exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J 2001;20:4987 97.
[9] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402 8.
[10] Mitchell P, Petfalski E, Shevchenko A, Mann M, Tollervey D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3′→5′ exoribonucleases. Cell 1997;91:457 66.
[11] Maquat LE. Nonsense-mediated mRNA decay: splicing, translation, and mRNP dynamics. Nat Rev Mol Cell Biol 2004;5:89 99.
[12] Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE. Nonsense-mediated decay approaches the clinic. Nat Genet 2004;36:801 8.
[13] Inoue K, Khajavi M, Ohyama T, Hirabayashi S, Wilson J, Reggin JD, et al. Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. Nat Genet 2004;36:361 9.
[14] Kerr TP, Sewry CA, Robb SA, Roberts RG. Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay? Hum Genet 2001;109:402 7.
[15] Usuki F, Yamashita A, Higuchi I, Ohnishi T, Shiraishi T, Osame M, Ohno S. Inhibition of nonsense-mediated mRNA decay rescues the phenotype in Ullrich's disease. Ann Neurol 2004;55:740 4.
[16] Usuki F, Yamashita A, Kashima I, Higuchi I, Osame M, Ohno S. Specific inhibition of nonsense-mediated mRNA decay components, SMG-1 or Upf1, rescues the phenotype of Ullrich disease fibroblasts. Mol Ther 2006;14:351 60.

PERGAMON

Neuromuscular Disorders 14 (2004) 659–665

www.elsevier.com/locate/nmd

Mutations in Czech LGMD2A patients revealed by analysis of calpain3 mRNA and their phenotypic outcome

Táňa Chrobáková[a,b], Markéta Hermanová[c], Iva Kroupová[c], Petr Vondráček[d], Tat'ána Maříková[e], Radim Mazanec[f], Josef Zámečník[g], Jan Staněk[h], Miluše Havlová[i], Lenka Fajkusová[a,b,*]

[a]*Center of Molecular Biology and Gene Therapy, University Hospital Brno, Černopolni 9, CZ-62500 Brno, Czech Republic*
[b]*Laboratory of Functional Genomics and Proteomics, Faculty of Science, Masaryk University Brno, Brno, Czech Republic*
[c]*Institute of Pathology, University Hospital Brno, Brno, Czech Republic*
[d]*Clinic of Child Neurology, University Hospital Brno, Brno, Czech Republic*
[e]*Institute of Biology and Medical Genetics, 2nd Faculty of Medicine, University Hospital Motol, Prague, Czech Republic*
[f]*Clinic of Child Neurology, 2nd Faculty of Medicine, University Hospital Motol, Prague, Czech Republic*
[g]*Department of Pathology and Molecular Medicine, 2nd Faculty of Medicine, University Hospital Motol, Prague, Czech Republic*
[h]*Clinic of Child Neurology, University Hospital Ostrava, Ostrava, Czech Republic*
[i]*Clinic of Neurology, 1st Faculty of Medicine, General University Hospital, Prague, Czech Republic*

Received 3 February 2004; received in revised form 5 April 2004; accepted 13 May 2004

Abstract

Calpain3 (CAPN3, p94) is a muscle-specific nonlysosomal cysteine proteinase. Loss of proteolytic function or change of other properties of this enzyme (such as stability or ability to interact with other muscular proteins) is manifested as limb girdle muscular dystrophy type 2A (LGMD2A, calpainopathy). These pathological changes in properties of calpain3 are caused by mutations in the calpain3 gene. The fact that the human gene for calpain3 is quite long led us to analyse its coding sequence by reverse transcription-PCR followed by sequence analysis. This study reports nine mutations that we found by analysing mRNA of seven unrelated LGMD patients in the Czech Republic. Three of these mutations were novel, not described on the Leiden muscular dystrophy pages so far. Further, we observed a reduction of dysferlin in muscle membrane in five of our seven LGMD2A patients by immunohistochemical analysis of muscle sections.

Keywords: Limb girdle muscular dystrophy; LGMD2A; Calpain3; Dysferlin; mRNA

1. Introduction

Limb girdle muscular dystrophies (LGMDs) represent a heterogeneous group of genetic disorders characterised by progressive weakness of the pelvic and shoulder girdle muscles with a great variability in clinical course. At present 16 forms of LGMD are known (six autosomal dominant, 10 autosomal recessive). LGMD2A, the most prevalent recessive form, is caused by mutations in the CAPN3 gene which encodes the proteolytic enzyme calpain3. The human CAPN3 gene is localized at 15q15.1-q21.1, comprises 24 exons and covers a genomic region of 50 kb. Its transcript is spliced to a 3.5 kb muscle-specific mRNA, which is then translated to a 94 kD protein [1].

Calpain3 is a member of the calpain superfamily, which processes enzymes involved in signalling pathways, transcription factors and cytoskeletal proteins thereby modulating their activities. Calpain3 differs from its ubiquitous relatives (m- and μ-calpain) mainly in three additional sequences (NS, IS1 and IS2), a monomeric functional structure, an uncertain dependence on Ca^{2+} activation and a skeletal muscle-specific expression. It was shown that calpain3 undergoes autolysis and generates a small N-terminal fragment of 30 kDa and a large C-terminal fragment whose size ranges from 55 to 60 kDa during self-processing. Anderson et al. published that calpain3 is stable in intact human muscle, there being little decrease in

* Corresponding author. Address: Center of Molecular Biology and Gene Therapy, University Hospital Brno, Černopolni 9, CZ-62500 Brno, Czech Republic. Tel.: +420-53223-4625; fax: +420-53223-4623.
E-mail address: lfajkusova@fnbrno.cz (L. Fajkusová).

0960-8966/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.nmd.2004.05.005

the abundance of full-size protein detected in biopsy samples rested at room temperature for up to 8 h after the muscle had been removed [2]. This is in contrast to early report that showed rapid autolysis after translation of this enzyme [3]. The association of calpain3 with connectin (titin) is supposed to stabilise this enzyme in muscle tissue [4].

At present, more than 150 allelic variants are identified which spread along the whole sequence of the CAPN3 gene (Leiden muscular dystrophy pages, LMDP, http://www.dmd.nl/).

In the present study, we describe our approach to analysis of the CAPN3 gene of seven unrelated LGMD patients and discuss the relation between the mutations detected and the patients' phenotypes.

2. Patients and methods

2.1. Patients

Patients presented in this report (two females and five males) met the clinical criteria of a LGMD phenotype: autosomal recessive inheritance, progression of muscle weakness, dystrophic patterns in a muscle biopsy and an increased level of creatine kinase. Patients 1–5 and 7 had no confirmed family history of muscular dystrophy. Patient 6 had an affected younger brother. The age at onset varied from 6 to 25 years. All patients had elevated serum CK levels (437–10018 Iu/l). More detailed clinical information is given in Table 1. Informed consent was obtained from all patients.

2.2. Immunohistochemical analysis

Muscle tissues were snap frozen in a propane–butane mixture cooled in liquid nitrogen. Cryosections were examined by conventional histological, histochemical and immunohistochemical methods. Fifteen monoclonal antibodies were used for detection of proteins in muscle sections: three to dystrophin (NCL-DYS1, -DYS2, -DYS3), two to utrophin (NCL-DRP1, -DRP2) and dysferlin (NCL-Hamlet, -Hamlet-2), one to α-sarcoglycan (NCL-a-SARC), β-sarcoglycan (NCL-b-SARC), γ-sarcoglycan (NCL-c-SARC), δ-sarcoglycan (NCL-g-SARC), β-dystroglycan (NCL-b-DG), emerin (NCL-emerin), α2-merosin (NCL-merosin) and β-spectrin (NCL-SPEC1) (Novocastra).

Five micrometers thick tissue sections were incubated with primary antibodies diluted in antibody diluent (DAKO, S2022, ready to use). After repeated washes with phosphate-buffered saline (PBS), secondary biotinylated antibody from diagnostic kit LSAB+ (DAKO) was applied and then either streptavidin-biotin-peroxidase complex (LSAB+, DAKO) or streptavidin-biotin-Cy3 complex (Sigma) was used for immunohistochemical or immunofluorescence detection, respectively. Tissue sections were repeatedly washed in PBS and diaminobenzidine (DAKO) was applied as substrate in immunohistochemical samples. The slides were counterstained with hematoxylin. Immunohistochemical samples were evaluated by light microscopy and immunofluorescence samples by fluorescence microscopy (Nikon Eclipse E1000).

For Western blotting, cryopreserved muscle tissues were homogenized in lysis buffer with addition of protease inhibitors (Complete™ Mini, Roche). Samples were

Table 1
Clinical findings

Pat.	Sex	Age at onset	Age at present	CK, Iu/l	Clinical findings
1	M	15	48	438	At the age of 15, a pelvic girdle muscular weakness was noted. An axial musculature weakness followed by a lumbar hyperlordosis developed later. Pectoral and pelvic girdle muscular wasting and weakness, calf and peroneal atrophy, bilateral pes cavus deformities, Achilles tendon contractures and a waddling gait with braces are now present. The patient has been dependent for most of his activities for about 10 years
2	F	18	28	1500	At the age of 18, elevation of serum transaminases was found. Pelvic girdle muscular weakness, Achilles tendon contractures and steppage gait were noted 6 years later
3	F	6	7	10018	At the age of 6, elevation of serum transaminases was found. A mild calf pseudohypertrophy, pectoral girdle muscular wasting, generalized hypotonia, early fatiguing and gait disorder were observed
4	M	10	13	7463	At the age of 10, a mild calf pseudohypetrophy, pelvic girdle muscular wasting, lumbar hyperlordosis and Achilles tendon contractures developed
5	M	25	67	546	At the age of 25, a pelvic girdle muscular wasting was noted and a weakness of pectoral girdle muscles followed 5 years later. The patient has been wheelchair-dependent since 1998
6	M	8	20	2520	At the age of 8, elevation of serum creatine kinase and transaminases was found. A pelvic girdle muscular wasting and weakness with a slow progression developed, but the patient is still able to walk without braces
7	M	18	52	468	At the age of 18, a pectoral girdle muscular wasting was noted and a weakness of pelvic girdle muscles followed several years later. A gait disorder deteriorated markedly during the last 10 years, a waddling gait with a support of braces (10 m approximately)

F, female; M, male; CK, creatine kinase (year of analysis 2002/2003).

Table 2
Partitioning of calpain3 mRNA for RT-PCR

Amplified region of mRNA	Primers	Position of amplified region on calpain3 cDNA
1	5′-CCAAGTCAACATTGCTTCAG-3′ 5′-CAATGATAAATCGGGGATTC-3′	−256 → 340
2	5′-TTATGTGGACCCTGAGTTCC-3′ 5′-TTCCATAGGTCATGTTCGTG-3′	222 → 826
3	5′-TCCTACGAAGCTCTGAAAGG-3′ 5′-TGCAGATCTCCAACTTTGTG-3′	643 → 1243
4	5′-GTCAGAGGTCACGCCTACTC-3′ 5′-AGGCGTTGTACAGGAAGAAG-3′	991 → 1588
5	5′-GACGATGACCCTGATGACTC-3′ 5′-CTCCATGTCATCTCCTGCTA-3′	1402 → 2001
6	5′-AGGAAAGTGAGGAACAGCAA-3′ 5′-GCTTTGGAAATAGAGGGTGA-3′	1937 → 2536

electrophoretically separated on 10% SDS-polyacrylamide gels and Western-blotted. After blocking (Non-Fat Dry Milk, BIO-RAD) and washing with PBS, the membrane was incubated with monoclonal antibody to calpain3 (NCL-CALP-11B3 or NCL-CALP-12A2) and washed with PBS. Then the secondary antibody labelled by peroxidase was applied (RaMPx, DAKO) and detection by chemiluminescent reagent ECL (Amersham) was performed. The antibody NCL-CALP-11B3 reacts with calpain3 (94 kDa) plus pre- and post-autolysed forms of the ubiquitous calpains μ- and m- which produce a group of bands between 78 and 84 kDa on Western blots. The antibody NCL-CALP-12A2 reacts with calpain3 but not with μ- and m-calpain.

2.3. RNA isolation and RT-PCR

Muscle tissues stored in the RNA*later* RNA Stabilization Reagent (QIAGEN) were homogenized in a Pellet Pestle homogenizer (Sigma). RNA was prepared using the RNeasy Fibrous Tissue Mini Kit (QIAGEN). The total calpain3 mRNA was divided into six overlapping regions in order to analyse the whole coding region of the gene. Each region was reverse-transcribed (RT) and amplified using the Titan™ One Tube RT-PCR System (Roche) in individual reactions in the following conditions: reverse transcription of total RNA (50 °C for 30 min); cDNA amplification with calpain3 specific primers (Table 2) using initial denaturation (94 °C/2 min) followed by 10 PCR cycles (94 °C/30 s, 55 °C/30 s, 68 °C/1 min) and additional 25 cycles using the same parameters with a 5 s extension of the polymerisation step per cycle.

The products of individual PCRs were electrophoretically separated in 1.5% agarose gels and visualised in UV light after ethidium bromide staining. The product bands were excised, purified by the Gel Extraction Kit (QIAGEN) and analysed by DNA sequencing with the same primers on an ABI PRISM 310 sequencer (Perkin–Elmer).

2.4. Restriction analysis of the mutation R490W

The regions 4–5 (Table 2, position on calpain3 cDNA: 991 → 2001) were reverse-transcribed and amplified using the same conditions as in Section 2.3, except for the polymerisation step which was 2 min. Restriction enzyme digestion of the RT-PCR product was performed by the restriction endonuclease *Aci*I (NEB), analysed on 2% NuSieve agarose (FMC) and visualised in UV light after ethidium bromide staining.

3. Results

3.1. Immunohistochemical analysis

Muscle tissue samples were examined using immunohistochemistry and Western blot. The findings are summarized in Table 3. The level of calpain3 was assessed in six of

Table 3
Results of histology and immunohistochemical analyses of muscle tissues

Pat.	Histology of muscle tissue	Immunodetection of calpain3			Immunodetection of dysferlin in muscle sections
		94 kDa	60 kDa	30 kDa	
1	Myopathic; lobulation of type I fibres, size variability; endomysial and perimysial fibrosis	Absence	Absence	Absence	Normal
2	Myopathic; angular atrophic fibres, regression and regeneration	Absence	Moderate labelling	Absence	Not performed
3	Myopathic; size variability, regression and necrosis	Absence	Absence	Absence	Sporadic failure of expression on muscle fibres
4	Myopathic; size variability, regression and necrosis	Absence	Absence	Absence	80% Fibres with failure of expression
5	Myopathic; lobulation of type I fibres, size variability, angular atrophic fibres; endomysial fibrosis	Absence	Moderate labelling	Absence	30% Fibres with failure of expression
6	Myopathic; size variability	Not performed	Not performed	Not performed	30% Fibres with failure of expression
7	Myopathic; lobulation of type I fibres, size variability; endomysial and perimysial fibrosis	Absence	Strong labelling	Absence	80% Fibres with failure of expression

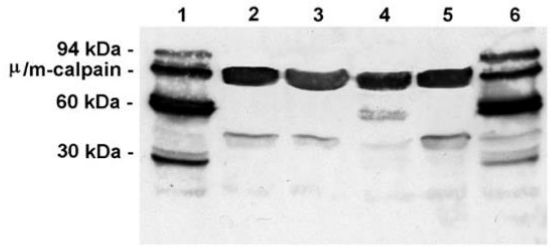


Fig. 1. The detection of calpain3 by Western blot (using monoclonal antibody NCL-CALP-11B3): lane 1, control; lane 2, patient 3; lane 3, patient 4; lane 4, patient 2; lane 5, patient 1; lane 6, control.

seven patients, showing the absence of 94 kDa protein in all cases (Figs. 1 and 2). Deficient expression of dysferlin in muscle sections was revealed in five patients by light microscopy and fluorescence microscopy (Fig. 3). Imunohistochemical analysis of other proteins (dystrophin, utrophin, α, β, γ, δ-sarcoglycans, β-dystroglycan, emerin, α2-merosin and β-spectrin) displayed normal results.

3.2. Analysis of calpain3 mRNA

We detected nine non-standard variants of the CAPN3 gene including three missense and four null mutations and two in-frame deletions. Three of these variants were novel with respect to the Leiden muscular dystrophy pages http://www.dmd.nl/; the first was missense (A45T), the second was an in-frame deletion (K595-K596del) and the third was a frameshift deletion (1722delC, F574fsX21). The localization of these mutations in individual exons of the gene and in domains of the calpain3 protein is specified in Table 4.

Some details about these three cases are worth noting:

In case no. 1, to demonstrate that the homozygous mutation A45T in the CAPN3 gene may be a cause of LGMD2A, the CAPN3 coding regions of several mammals (human, mouse, rat, bovine, sheep and porcine) were analysed by CLUSTALW Multiple sequence alignment and the conservation of alanine in this position among the species was confirmed. Checking for this mutation in

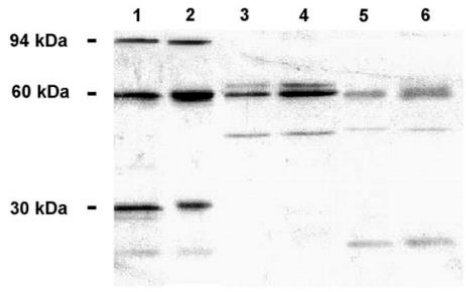


Fig. 2. The detection of calpain3 by Western blot (using monoclonal antibody NCL-CALP-12A2): lanes 1 and 2, control; lanes 3 and 4, patient 7; lanes 5 and 6, patient 5.

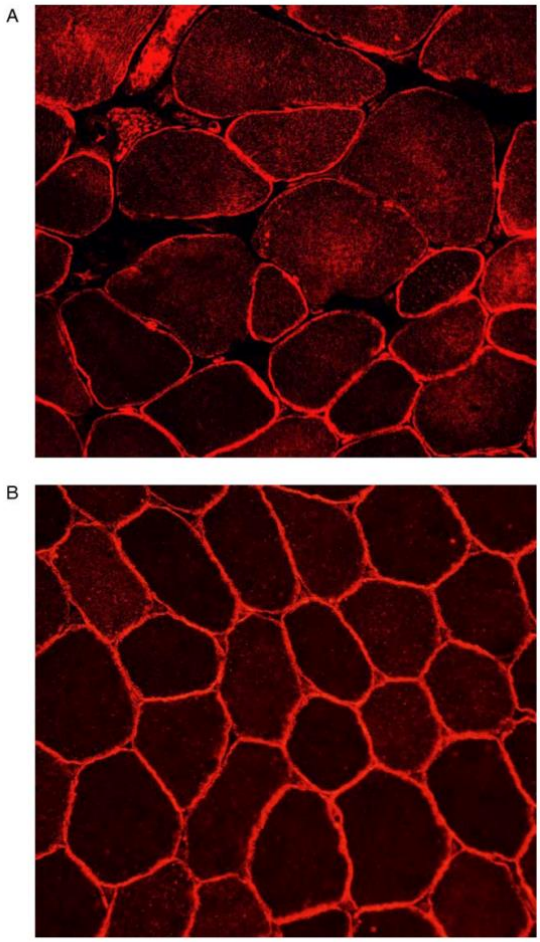


Fig. 3. Example of immunofluorescent detection of dysferlin in case of patient 3 (A) and a control sample (B).

a Czech control population of 100 control alleles was negative. The results of both analyses point to the connection between the mutation A45T and LGMD2A.

In case no. 5, only one mutation F200-L204del was detected. A second mutation was not found, although the whole coding region was examined.

In case no. 6, a homozygous occurrence of the mutation R490W (Fig. 4A) was detected by RNA analysis. However, sequencing of exon 11 in the patient's DNA confirmed the mutation at only one allele (Fig. 4B). The same results were obtained in the patient's brother. The homozygous presence of the mutation R490W at the mRNA level was further proved by restriction analysis of an RT-PCR product (Fig. 4C). We performed sequence analysis of the promoter and adjacent regions of the CAPN3 gene (nucleotide 34–1226, GenBank acc. no. AF209502) to detect the possible presence of a mutation affecting the expression of the allele without the mutation

Table 4
Results of analyses of calpain3 mRNA

Pat.	Mutation detected in allele 1	Position of mutation in the gene and the protein	Mutation detected in allele 2	Position of mutation in the gene and the protein
1	G(133) → A, A45T	Exon 1 Domain I	G(133) → A, A45T	Exon 1 Domain I
2	C(245) → T, P82L	Exon 1 Domain I	C(245) → T, P82L	Exon 1 Domain I
3	550delA, T184fsX36	Exon 4 Domain II	1981delA, I661X	Exon 17 Domain IV
4	550delA, T184fsX36	Exon 4 Domain II	1722delC, F574fsX21	Exon 13 Domain III
5	598–612del, F200-L204del	Exon 4 Domain II	Not detected	–
6	C(1468) → T, R490W	Exon 11 Domain III	Not detected	–
7	1783–1788del, K595–K596del	Exon 15 sequence IS2	C(2242) → T, R748X	Exon 21 Domain IV

R490W, but did not find any mutation. We also sequenced the coding region of the dysferlin gene of the same patient (due to the presence of about 30% fibres with failure of expression of dysferlin), but no mutation was found.

4. Discussion

Calpain3 consists of four domains conserved in conventional calpains with additional insertions of three unique regions, NS, IS1, and IS2. The mutations detected in our patients were localized in all domains including the IS2 sequence.

Domain I comprises a single alpha-helix anchored in a cavity of domain IV, thereby stabilizing the circular domain arrangement of the protein [5]. Patients 1 and 2 are homozygous for two different missense mutations that occur in this domain. The onset of the disease in these patients was at the age of 15 and 18 years, respectively.

Patients 3–5 are heterozygous for mutations in domain II that accounts for the protease activity of calpain3 [5]. Patients 3 and 4 are affected by two null mutations (a second mutation in domain IV and III). The onset of their disease was early, at the age of 6 and 10 years, respectively. The mutation 550delA occurring in both our patients is supposed to be the most frequent calpain3 defect in Europe [6,7].

Patient 5 is heterozygous for an in-frame deletion F200-L204del, and a second mutation was not found. This type of deletion is reported four times in the Leiden muscular dystrophy pages, once in connection with hyperCKemia (in this patient the mutation is combined with mutation in the β-sarcoglycan gene), once in an LGMD2A patient in which a second mutation has not been detected, and twice in LGMD2A patients in which a second mutation has been detected. Our patient 5 had the latest onset of disease and wheelchair confinement since the age of 62 years.

Domain III of calpain3 has a topology similar to the C2 domain found in various proteins, which is known to interact with Ca^{2+} [5]. Patients 4 and 6 are heterozygous for mutations in domain III. Patient 6 is homozygous for the missense mutation R490W at the RNA level; however, he is heterozygous for this mutation by the analysis of DNA.

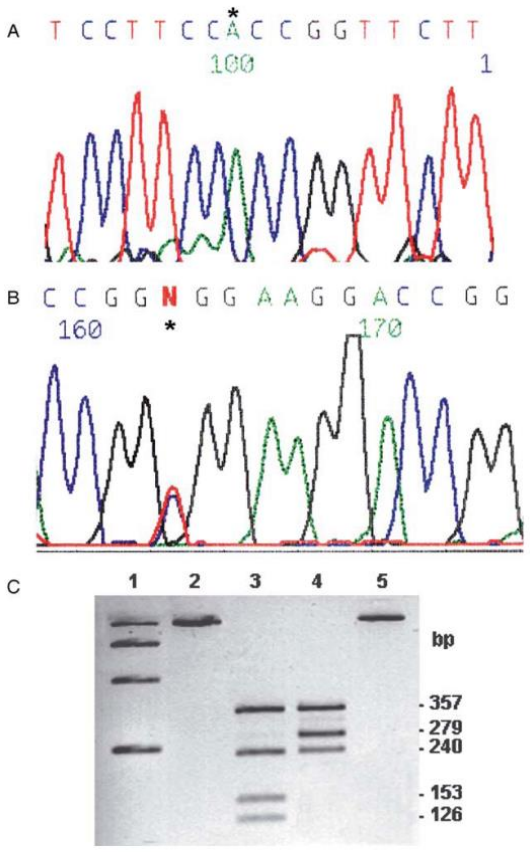


Fig. 4. The analysis of the mutation R490W detected in patient 6: RNA sequence analysis of the mutation R490W, the position of the mutation is marked by an asterisk (A). DNA sequence analysis of the mutation R490W, the position of the mutation is marked by an asterisk (B). Restriction analysis of RT-PCR product: homozygous occurrence of the mutation R490W yields a PCR product which is digested by restriction endonuclease *Aci*I into 357, 279 and 240 bp fragments, while the control PCR product is cleaved into 357, 240, 153 and 126 bp fragments. Lane 1; size markers, lane 2; non-restricted RT-PCR product (control); lane 3, RT-PCR product restricted by *Aci*I (control); lane 4, RT-PCR product restricted by *Aci*I (patient 6); lane 5, non-restricted RT-PCR product (patient 6) (C).

Table 5
Incidence of detected mutations

Pat.	Mutation detected in allele 1	Incidence	Mutation detected on allele 2	Incidence
1	G(133) → A, A45T	So far not described	G(133) → A, A45T	So far not described
2	C(245) → T, P82L	Brazil [11]	C(245) → T, P82L	Brazil [11]
3	550delA	France [1,16,17], the Netherlands [7,16], Russia [7], Slovenia (LMDP) Greece, Poland, UK [16], Italy [16], Turkey [17]	1981delA	the Netherlands (LMDP), France [7]
4	550delA		1722delC	So far not described
5	598–612del, F200-L204del	Germany [9], Bulgaria [7], Russia [6]	Not detected	
6	C(1468) → T, R490W	France [1], US [16], Portugal (LMDP)	Not detected	
7	1783–1788del, K595–K596del	So far not described	C(2242) → T, R748X	the Netherlands (LMDP), Brazil [11]

A similar case has previously been described by Chou [8]. Although we did not find a mutation in the promoter region, we are not able to eliminate the possibilities of 3′UTR modification or aberrant mRNA splicing which may cause RNA suppression [9]. The properties of mutant R490W were predicted by Ono; this enzyme does not undergo autolysis so rapidly and is more Ca^{2+}-dependent and although its association with connectin (titin) is normal, proteolytic function is rapidly abolished [10].

The IS1 and IS2 sequences are essential for autolysis of calpain3 [4]; IS2 is also responsible for interaction with connectin (titin) and further contains a nuclear translocation signal-like sequence [4,5]. Patient 7 has an in-frame deletion in IS2, and a second mutation in domain IV that is characterised like Ca^{2+}-binding domain containing five EF-hand motifs [5]. This second mutation introduces a premature stop codon in the coding region. Affection is not so severe as in the case of two null mutations; the onset of the disease was at the age of 18 years and at present this patient (at the age of 52 years) is confined to a wheelchair, but able to walk for about 10 m with support.

In an attempt to find correlations between the phenotype and genotype, the age of patients at the onset of disease and the type of their mutation was compared. The earliest age at onset was in patients 3 and 4 (heterozygous for two different out-of-frame deletions), and on the other hand the latest age at onset was in patient 6 (in-frame deletion on one allele, no second mutation detected). Although these results should be interpreted with caution due to the small number of patients and detection of only one mutation in two patients, and do not enable us to draw conclusions on relationships between the mutations found and clinical status, they are in agreement with published findings in which missense mutations and in-frame deletions are associated with a milder phenotype than nonsense mutations and out-of-frame deletions [11].

Western blot analysis of muscle calpain3 in LGMD2A patients can show a total, partial, or more rarely, no apparent deficiency, with no direct correlation between the amount of the protein and the severity of the phenotype [12]. In our patients Western blot analyses were performed in two different laboratories, using either NCL-CALP-11B3 antibody (Fig. 1), or the NCL-CALP-12A2 antibody (Fig. 2). We detected the absence of the 94 kDa fragment; absence (patients 1, 3 and 4), moderate labelling (patients 2 and 5) or strong labelling (patient 7) of the 60 kDa fragment; and the absence of the 30 kDa fragment. In all patients, we have observed faint labelling of a fragment between 60 kDa and 30 kDa that is apparently a product of self-processing of calpain3. The diagnostic significance of the observed bands is uncertain.

The analysis of calpain3 in other forms of LGMD revealed a secondary reduction in its expression in patients with LGMD2B (8 of 16 patients, [13]) and LGMD2J [14], suggesting an association between calpain3 and dysferlin and connectin (titin), respectively.

On the other hand, a secondary reduction of dysferlin has not been described in calpainopathy patients [15]. In the present study, we detected a reduction of dysferlin in muscle sections in five patients, no. 3 (sporadic failure of expression in muscle fibres), nos. 5 and 6 (30% fibres with failure of expression) and nos. 4 and 7 (80% fibres with failure of expression). Our results show that besides the secondary reduction of calpain3 in LGMD2B patients, the secondary reduction of dysferlin in LGMD2A patients occurs, observed here as the decreased immunohistochemical labelling in muscle membrane.

Six mutations that we detected were previously described in other countries, as summarized in Table 5. For example, the mutation P82L has been detected only in patients of Caucasian origin in Brazil so far [11], while the mutation 550delA is broadly represented in European populations (France, Netherlands, Russia) and also in Turkey.

Acknowledgements

We thank Ronald Hancock (Laval University Cancer Research Centre) for reviewing the manuscript and Jiri Fajkus (Institute of Biophysics, Czech Academy of Sciences) for useful comments on the manuscript.

This work was supported by the Czech Ministry of Health (CEZ MZ/98/0001/00209627 and IGA MH 8087-3

and 6506-3) and by the Czech Ministry of Education (MSM143100008).

References

[1] Richard I, Broux O, Allamand V, et al. Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 1995;81:27–40.
[2] Anderson LV, Davison K, Moss JA, et al. Characterization of monoclonal antibodies to calpain 3 and protein expression in muscle from patients with limb-girdle muscular dystrophy type 2A. Am J Pathol 1998;153:1169–79.
[3] Sorimachi H, Toyama-Sorimachi N, Saido TC, et al. Muscle-specific calpain, p94, is degraded by autolysis immediately after translation, resulting in disappearance from muscle. J Biol Chem 1993;268: 10593–605.
[4] Sorimachi H, Kinbara K, Kimura S, et al. Muscle-specific calpain, p94, responsible for limb girdle muscular dystrophy type 2A, associates with connectin through IS2, a p94-specific sequence. J Biol Chem 1995;270:31158–62.
[5] Khorchid A, Ikura M. How calpain is activated by calcium. Nat Struct Biol 2002;9:239–41.
[6] Pogoda TV, Krakhmaleva IN, Lipatova NA, Shakhovskaya NI, Shishkin SS, Limborska SA. High incidence of 550delA mutation of CAPN3 in LGMD2 patients from Russia. Hum Mutat 2000;15:295.
[7] Richard I, Roudaut C, Saenz A, et al. Calpainopathy—a survey of mutations and polymorphisms. Am J Hum Genet 1999;64:1524–40.
[8] Chou FL, Angelini C, Daentl D, et al. Calpain III mutation analysis of a heterogeneous limb-girdle muscular dystrophy population. Neurology 1999;52:1015–20.
[9] Haffner K, Speer A, Hubner C, Voit T, Oexle K. A small in-frame deletion within the protease domain of muscle-specific calpain, p94 causes early-onset limb-girdle muscular dystrophy 2A. Hum Mutat 1998;1:S298–S300.
[10] Ono Y, Shimada H, Sorimachi H, et al. Functional defects of a muscle-specific calpain, p94, caused by mutations associated with limb-girdle muscular dystrophy type 2A. J Biol Chem 1998;273: 17073–8.
[11] de Paula F, Vainzof M, Passos-Bueno MR, et al. Clinical variability in calpainopathy: what makes the difference? Eur J Hum Genet 2002;10: 825–32.
[12] Zatz M, de Paula F, Starling A, Vainzof M. The 10 autosomal recessive limb-girdle muscular dystrophies. Neuromuscul Disord 2003;13:532–44.
[13] Anderson LV, Harrison RM, Pogue R, et al. Secondary reduction in calpain 3 expression in patients with limb girdle muscular dystrophy type 2B and Miyoshi myopathy (primary dysferlinopathies). Neuromuscul Disord 2000;10:553–9.
[14] Haravuori H, Vihola A, Straub V, et al. Secondary calpain3 deficiency in 2q-linked muscular dystrophy: titin is the candidate gene. Neurology 2001;56:869–77.
[15] Pogue R, Anderson LV, Pyle A, et al. Strategy for mutation analysis in the autosomal recessive limb-girdle muscular dystrophies. Neuromuscul Disord 2001;11:80–7.
[16] Richard I, Brenguier L, Dincer P, et al. Multiple independent molecular etiology for limb-girdle muscular dystrophy type 2A patients from various geographical origins. Am J Hum Genet 1997;60: 1128–38.
[17] Dincer P, Leturcq F, Richard I, et al. A biochemical, genetic, and clinical survey of autosomal recessive limb girdle muscular dystrophies in Turkey. Ann Neurol 1997;42:222–9.

OPEN ACCESS Freely available online

PLOS ONE

CLCN1 Mutations in Czech Patients with Myotonia Congenita, *In Silico* Analysis of Novel and Known Mutations in the Human Dimeric Skeletal Muscle Chloride Channel

Daniela Skálová[1,2], Jana Zídková[1,2], Stanislav Voháňka[3], Radim Mazanec[4], Zuzana Mušová[5], Petr Vondráček[6], Lenka Mrázová[6], Josef Kraus[7], Kamila Réblová[2*], Lenka Fajkusová[1,2*]

1 Centre of Molecular Biology and Gene Therapy, University Hospital, Brno, Brno, Czech Republic, **2** Central European Institute of Technology, Masaryk University, Brno, Czech Republic, **3** Department of Neurology, University Hospital Brno, Brno, Czech Republic, **4** Department of Neurology, Charles University Second Faculty of Medicine and University Hospital Motol, Prague, Czech Republic, **5** Department of Biology and Medical Genetics, Charles University Second Faculty of Medicine and University Hospital Motol, Prague, Czech Republic, **6** Department of Child Neurology, University Hospital Brno, Brno, Czech Republic, **7** Department of Child Neurology, Second School of Medicine, Charles University and University Hospital Motol, Prague, Czech Republic

Abstract

Myotonia congenita (MC) is a genetic disease caused by mutations in the skeletal muscle chloride channel gene (*CLCN1*) encoding the skeletal muscle chloride channel (ClC-1). Mutations of *CLCN1* result in either autosomal dominant MC (Thomsen disease) or autosomal recessive MC (Becker disease). The ClC-1 protein is a homodimer with a separate ion pore within each monomer. Mutations causing recessive myotonia most likely affect properties of only the mutant monomer in the heterodimer, leaving the wild type monomer unaffected, while mutations causing dominant myotonia affect properties of both subunits in the heterodimer. Our study addresses two points: 1) molecular genetic diagnostics of MC by analysis of the *CLCN1* gene and 2) structural analysis of mutations in the homology model of the human dimeric ClC-1 protein. In the first part, 34 different types of *CLCN1* mutations were identified in 51 MC probands (14 mutations were new). In the second part, on the basis of the homology model we identified the amino acids which forming the dimer interface and those which form the Cl^- ion pathway. In the literature, we searched for mutations of these amino acids for which functional analyses were performed to assess the correlation between localisation of a mutation and occurrence of a dominant-negative effect (corresponding to dominant MC). This revealed that both types of mutations, with and without a dominant-negative effect, are localised at the dimer interface while solely mutations without a dominant-negative effect occur inside the chloride channel. This work is complemented by structural analysis of the homology model which provides elucidation of the effects of mutations, including a description of impacts of newly detected missense mutations.

Citation: Skálová D, Zídková J, Voháňka S, Mazanec R, Mušová Z, et al. (2013) *CLCN1* Mutations in Czech Patients with Myotonia Congenita, *In Silico* Analysis of Novel and Known Mutations in the Human Dimeric Skeletal Muscle Chloride Channel. PLoS ONE 8(12): e82549. doi:10.1371/journal.pone.0082549

Editor: Agustin Guerrero-Hernandez, Cinvestav-IPN, Mexico

Received August 22, 2013; **Accepted** October 26, 2013; **Published** December 11, 2013

Funding: This work was funded by the Czech Ministry of Education, the project "CEITEC – Central European Institute of Technology" (CZ.1.05/1.1.00/02.0068), the project SuPReMMe (CZ.1.07/2.3.00/20.0045), and the project of the Internal Grant Agency of the Czech Ministry of Health (NT/14574-3) and supported by the project (Ministry of Health, Czech Republic) for conceptual development of research organization 65269705 (University Hospital Brno, Brno, Czech Republic). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: kristina@physics.muni.cz (KR); lfajkusova@fnbrno.cz (LF)

Introduction

Myotonia congenita (MC) is a skeletal muscle disorder that causes myotonia, an abnormal delay in muscle relaxation after voluntary or evoked muscle contraction, as the prominent symptom. MC is due to mutations in the skeletal muscle chloride channel gene (*CLCN1*) that is located on chromosome 7q35, encompasses 35 kb of genomic DNA, contains 23 exons, and encodes the skeletal muscle chloride channel ClC-1 [1,2]. Mutations of *CLCN1* result in either autosomal dominant (Thomsen disease) or autosomal recessive (Becker disease) MC, and a subset have been found to cause both recessive and dominant MC (semidominant mutations) [3].

The ClC-1 protein is a homodimer with a separate ion pore within each subunit [4–7]. Voltage-dependent gating of ClC-1 has two components that are both activated by membrane depolarization: i) fast gating that occurs independently in each pore and ii) slow or common gating that operates in both pores of the dimer simultaneously [7]. Mutations causing recessive myotonia most likely affect properties of one subunit such as fast gating and/or conductance, leaving the other subunit of the heterodimer unaffected [7–10]. On the other hand mutations causing dominant myotonia affect properties of both subunits in the heterodimer and are associated with an alteration of the common gating process, which is preferentially confined to amino acids (AA) forming the dimer interface [7,11–13] but can also be affected by other regions of the structure [14–18]. In the dominant disease, a mutant monomer affects interaction with a wild type monomer subunit by exerting a dominant-negative effect, typically seen as a larger than expected shift in the voltage dependence of common gating towards more positive potentials when functional analyses of wild type-mutant heterodimers are performed [19]. The varied inheritance pattern of myotonia (recessive/dominant) appears to result from differential effects of a mutation on the channel dimer [7,8] however, detailed comprehensive structural analysis of effects of mutations in the ClC-1 protein has not been carried out.

First structural insight on ClC channels came from the discovery of bacterial homologues from *Escherichia coli* and *Salmonella typhimurium* by electron microscopy [15] and X-ray crystallography [20]. The first X-ray crystallographic structure of a eukaryotic ClC protein was determined by Feng et al. [21] for the Cl^-/H^+ transporter from *Cyanidioschyzon merolae* (CmClC). From the X-ray crystal structures, it became clear that individual subunits of the ClC dimer consist of a transmembrane (TM) and a cytosolic cystathionine beta-synthase (CBS) domain comprising 23 alpha-helices (A-V) and 5 beta-strands. The helices of the TM domain vary in length and are oriented somewhat obliquely relative to each other and to the plane of the membrane, with many of them not spanning the entire membrane.

Our present study addresses two points: 1) molecular genetic diagnostics of MC by analysis of the *CLCN1* gene, and 2) structural analysis of *CLCN1* missense mutations. In the first part, we analysed DNA of Czech patients with MC. In the second part, we built a homology model of the dimeric ClC-1 channel on the basis of the crystallographic structure of the CmClC transporter [21] and identified AA forming the dimer interface and those forming the Cl^- ion pathway. Further, in the literature we searched for mutations of those AA for which functional analyses were performed in order to assess the correlation between the localisation of a mutation and its functional impact. Subsequently, we carried out structural analysis of these mutations to elucidate their effect on the protein structure and function and simultaneously described the effect of novel missense mutations.

Materials and Methods

Ethics Statement

The study has been approved by the Ethical committee of University Hospital Brno and the Ethical committee of University Hospital Motol and has therefore been performed in accordance with ethical standards laid down in the 1964 Declaration of Helsinki. All participants provided their written informed consent which was approved by the committees.

Patients

For analysis of the *CLCN1* gene, patients were sent from Departments of Neurology and Medical Genetics within the Czech Republic. Detailed clinical information (age at onset, myotonia degree, myotonia distribution, specific clinical findings, muscle hypertrophy, muscle weakness, etc.) were requested retrospectively on the basis of positive results of DNA analysis. Out of the total number of 51 patients with mutation/mutations in *CLCN1*, valid clinical information was obtained in 34 patients.

DNA analysis

Genomic DNA was extracted from peripheral blood leukocytes by the standard salting-out method, and amplified by PCR. Primers for amplification of all exons and adjacent intron sequences are described in File S1 as well as the conditions of particular PCRs, see Table S1 in File S1. PCR products were sequenced directly using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and analysed on the ABI 3130Xl Genetic Analyzer (Applied Biosystems). The resulting sequences were compared with the *CLCN1* NCBI reference sequence (NG_009815.1). To find out whether the detected sequence variations were described previously, we used literature and databases such as the Leiden Open Variation Database (LOVD, http://chromium.liacs.nl/LOVD2/home.php?action=switch_db) and the Human Gene Mutation Database (HGMD, http://www.hgmd.cf.ac.uk/ac/index.php). All novel missense mutations were screened in a control panel consisting of DNA from 200 healthy Czech individuals. For prediction of effects of mutations on splicing of pre-mRNA, the *in silico* tools NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) and SpliceView (http://zeus2.itb.cnr.it/~webgene/wwwspliceview.html) were used. The nomenclature of mutations was implemented according to the current HGVS recommendations (http://www.hgvs.org/mutnomen/). In patients with one mutation in the *CLCN1* gene, detection of deletions/duplications was performed using Multiplex ligation-dependent probe amplification (MLPA). We used the SALSA MLPA kit P350-B1 CLCN1-KCNJ2 according to the manufacturer's guideline (MRC Holland). This kit contains probes and primers for all 23 *CLCN1* exons.

Homology modelling of the human ClC-1 dimer and structural analysis of the CLCN1 mutations

The crystallographic structure of the CmClC transporter from *Cyanidioschyzon merolae* (PDB code 3org, chain A, [21]) was used as a template for the human ClC-1 protein (we modeled

the TM domain, AA 120 to 593, using the I-TASSER server [22]). Regions 319-347, 408-420, and 430-460 which are disordered in the initial structure were not considered in detailed structural analyses. The human dimeric ClC-1 structure was obtained by superimposing the homology model of the human monomers and the dimeric CmClC crystal structure (PDB entry 3org, chains A and D) using the program VMD http://www.ks.uiuc.edu/Research/vmd/ [23]. The chloride pathway inside the ClC-1 model was identified using the software Mole*online* 2.0 http://mole.upol.cz/ [24] and the dimer interface by the program PyMOL (http://www.pymol.org, The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC). Structural figures were generated by PyMOL and VMD.

As in our previous study [25], we evaluated the impact of selected *CLCN1* missense mutations on the protein structure and function using *in silico* tools. In particular, we detected specific side chain contacts (H-bonds, aromatic interactions) of AA in the 3D structure of the wild type protein based on visual inspection using the VMD program. Loss of these contacts upon a mutation often results in destabilization of the protein structure. In addition, we measured the buriedness of wild type AA corresponding to a relative accessible surface area (RSA) of <15 % [26]. Replacement of buried AA is more likely to be associated with structural defects especially when volume, charge and polarity change upon a mutation, and thus for buried residues we measured these parameters. Volume change upon mutation was calculated according to [27], and a change with an absolute value of ≥30 $Å^3$ was considered destabilizing [28]. A charge change upon mutation was considered between charged and uncharged AA and a polarity change was considered between nonpolar (Leu, Ile, Phe, Trp, Cys, Met, Val), polar (Tyr, Pro, Ala, Thr, Gly, Ser), and very polar (His, Arg, Gln, Lys, Asn, Glu, Asp) AA [29].

Results

DNA analysis of the *CLCN1* gene

In this set of probands with a diagnosis of MC, we have 6 patients with one mutation detected in the *CLCN1* gene (Table 1). The mutation p.(Trp164Arg) identified in patient 1 is associated with dominant inheritance (the patient´s mother carrying the mutation also has MC). In another study [30], p.(Trp164Arg) was described in a patient with the genotype [p.(Trp164Arg)];[c.1471+1GA] but no information concerning MC in family members was included. The mutation p.(Met560Thr) was found in patient 2, whose parents died many years ago and probably were without any symptoms of myotonia. This mutation was also described in a patient with typical myotonia (genotype p.[Met560Thr];[Tyr261Cys]) and in his father with mild symptoms of myotonia (genotype p.[Met560Thr];[=]) [31]. The most frequent mutation among our MC probands is the semidominant mutation p.(Arg894*) which was detected in 38 disease alleles (39.6 %). This mutation can be expressed by the dominant or the recessive manner of inheritance. In a previous study [32], this mutation was dominant with late onset of symptoms (22-30 years), dominant with early onset of symptoms (7 years), or recessive. In patients 3-6, the mutation p.(Arg894*) was present as one mutation detected in the *CLCN1* gene; the patients´ parents were without clinical symptoms of MC, but detailed neurological examination was performed only in the parents of patient 6.

In 45 MC probands, two or more mutations were identified in the *CLCN1* gene (Table 1). Mutations detected in our patients and also described in the literature and/or the databases LOVD and HGMD in association with Becker disease include p.(Gln74*), p.(Arg105Cys), p.(Phe167Leu), p.(Gly190Arg), p.(Gly190Ser), p.(Trp303*), p.(Phe413Cys), p.(Met485Val), p.(Ala493Glu), p.(Arg894*), c.180+3AT, c.1437_1450del, c.1471+1GA, c.2284+5CT, and c.2364+2TA. The mutation c.1437_1450del, p.(Pro480Hisfs*24) was present in 18 disease alleles (18.8 %) and is thus the second most frequent *CLCN1* mutation detected in our probands. The mutations p.(His29Pro) and p.(Ala566Val) detected in patient 8 are also described in LOVD, but the MC type is not mentioned there and the pathogenicity is described as unknown (see below). The mutation p.(Ile290Met) was described as dominant [33–35] but in patient 18 it is present together with p.(Arg894*). Detailed neurological examination and DNA analysis were performed in the patient´s parents; the father carries p.(Arg894*) and is without MC symptoms, and the mother carries p.(Ile290Met) and has EMG-positive but clinically silent MC. In the set of our patients with Becker disease, we described 14 new mutations: 8 types are frame-shift, splicing, or nonsense [c.32delG, c.587delC, c.1044_1156del, c.1324_1325delAG, c.433+3AG, c.1401+3AT, c.2508+2TA, p.(Tyr257*)] and 6 types are missense [p.(Glu291Gln), p.(Tyr302Cys), p.(Thr432Arg), p.(Asn455Tyr), p.(Gly482Glu), p.(Ala493Val)]. Patient 39 carries two new mutations, p.(Gly482Glu) and the splicing mutation c.2508+2TA. A mutation occurring in the same codon, p.(Gly482Arg), was identified previously in a patient with Becker disease and the genotype p.[Gly482Arg];[Pro480Hisfs*24] [36]. All potential splicing mutations showed an identical effect, loss of a donor splice site, by using *in silico* tools. The presence of all novel missense mutations was tested in DNA from 200 controls and none of them were detected.

In five probands, we detected three sequence variants in *CLCN1*. The segregation of mutations in disease alleles was determined on the basis of DNA analysis of the parents of patient 8 (genotype p.[His29Pro; Tyr257*]; [Ala566Val]), patient 20 (genotype p.[Tyr302Cys; Thr432Arg]; [Pro480Hisfs*24]), and patient 21 (genotype p.[Tyr302Cys; Thr432Arg]; [Arg894*]). The phenotypes of these patients (if available) are described in File S1, see Table S2 in File S1. The most severe phenotypes were found in patients 8 and 14. Patient 8 (female, 39 years old, genotype p.[His29Pro; Tyr257*]; [Ala566Val]) suffers from moderate myotonia, permanent limb-girdle muscle weakness, and scoliosis. Patient 14 (male, 15 years old, genotype p.[Gly190Ser]; [Pro480Hisfs*24]) has severe myotonia, kyphoscoliosis, deformities of the feet, shortening of the Achilles tendons, and masseter and limb-girdle muscle weakness.

Structural analysis of the *CLCN1* mutations localised in the dimer interface and the Cl^- ion pathway

The homology model which we built comprises AA residues 120-593 and consists of helices B-R according to the template

Table 1. *CLCN1* mutations detected in Czech MC patients.

No. of patient	Phenotype	1st Mutation (cDNA and protein level)	2nd Mutation (cDNA and protein level)	3rd Mutation (cDNA and protein level)
1	TD	c.490T>C, p.(Trp164Arg)		
2	MC, isolated occurrence in family	c.1679T>C, p.(Met560Thr)		
3	MC, isolated occurrence in family	c.2680C>T, p.(Arg894*)		
4	MC, isolated occurrence in family	c.2680C>T, p.(Arg894*)		
5	MC, isolated occurrence in family	c.2680C>T, p.(Arg894*)		
6	MC, isolated occurrence in family	c.2680C>T, p.(Arg894*)		
7	BD	**c.32delG, p.(Gly11Valfs*66)**	c.2680C>T, p.(Arg894*)	
8	BD	c.86A>C, p.(His29Pro)	**c.771T>A, p.(Tyr257*)**	c.1697C>T, p.(Ala566Val)
9	BD	c.180+3A>T, splicing effect	c.220C>T, p.(Gln74*)	
10	BD	c.220C>T, p.(Gln74*)	**c.587delC, p.(Thr196Leufs*8)**	
11	BD	c.220C>T, p.(Gln74*)	c.1238T>G, p.(Phe413Cys)	
12	BD	c.313C>T, p.(Arg105Cys)	c.501C>G, p.(Phe167Leu)	c.1437_1450del, p.(Pro480Hisfs*24)
13	BD	**c.433+3A>G, splicing effect**	c.1437_1450del, p.(Pro480Hisfs*24)	
14	BD	c.568_569delinsTC, p.(Gly190Ser)	c.1437_1450del, p.(Pro480Hisfs*24)	
15	BD	c.568G>A, p.(Gly190Arg)	c.2680C>T, p.(Arg894*)	
16	BD	c.568G>A, p.(Gly190Arg)	c.2680C>T, p.(Arg894*)	
17	BD	c.803C>T, p.(Thr268Met)	c.1437_1450del, p.(Pro480Hisfs*24)	
18	BD	c.870C>G, p.(Ile290Met)	c.2680C>T, p.(Arg894*)	
19	BD	**c.871G>C, p.(Glu291Gln)**	**c.1478C>T, p.(Ala493Val)**	
20	BD	**c.905A>G, p.(Tyr302Cys)**	**c.1295C>G, p.(Thr432Arg)**	c.1437_1450del, p.(Pro480Hisfs*24)
21	BD	**c.905A>G, p.(Tyr302Cys)**	**c.1295C>G, p.(Thr432Arg)**	c.2680CT, p.Arg894*
22	BD	c.908G>A, p.(Trp303*)	c.1437_1450del, p.(Pro480Hisfs*24)	
23	BD	**c.1044-1156del, p.(Ala350Serfs*65)**	c.2284+5C>T, splicing effect	
24	BD	c.1238T>G, p.(Phe413Cys)	c.1437_1450del, p.(Pro480Hisfs*24)	
25	BD	c.1238T>G, p.(Phe413Cys)	c.2680C>T, p.(Arg894*)	
26	BD	c.1238T>G, p.(Phe413Cys)	c.2680C>T, p.(Arg894*)	
27	BD	c.1238T>G, p.(Phe413Cys)	c.2680C>T, p.(Arg894*)	
28	BD	c.1238T>G, p.(Phe413Cys)	c.2680C>T, p.(Arg894*)	
29	BD	**c.1324_1325delAG, p.(Ser442Profs*66)**	**c.1324_1325delAG p.(Ser442Profs*66)**	
30	BD	**c.1363A>T, p.(Asn455Tyr)**	**c.1401+3A>T, splicing effect**	c.2680CT, p.(Arg894*)
31	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.1437_1450del, p.(Pro480Hisfs*24)	
32	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.1437_1450del, p.(Pro480Hisfs*24)	
33	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.1437_1450del, p.(Pro480Hisfs*24)	
34	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.1453A>G, p.(Met485Val)	
35	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.2680C>T, p.(Arg894*)	
36	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.2680C>T, p.(Arg894*)	
37	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.2680C>T, p.(Arg894*)	
38	BD	c.1437_1450del, p.(Pro480Hisfs*24)	c.2680C>T, p.(Arg894*)	
39	BD	**c.1445G>A, p.(Gly482Glu)**	**c.2508+2T>A, splicing effect**	
40	BD	c.1471+1G>A, splicing effect	c.2680C>T, p.(Arg894*)	
41	BD	c.1471+1G>A, splicing effect	c.2680C>T, p.(Arg894*)	
42	BD	c.1478C>A, p.(Ala493Glu)	c.2364+2T>A, splicing effect	
43	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
44	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
45	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
46	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
47	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
48	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	

Table 1 (continued).

No. of patient	Phenotype	1st Mutation (cDNA and protein level)	2nd Mutation (cDNA and protein level)	3rd Mutation (cDNA and protein level)
49	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
50	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	
51	BD	c.2680C>T, p.(Arg894*)	c.2680C>T, p.(Arg894*)	

Mutations described by bold letters are not described previously. TD: Thomsen disease; BD: Becker disease.

doi: 10.1371/journal.pone.0082549.t001

structure (see Materials and Methods). Based on the model, we identified 40 residues forming the dimer interface and 43 residues forming the Cl^- ion pathway (Figure 1A and Table S3 in File S1). Residues Ser189, Glu232, and Tyr578 which are known to be essential for gating, coordination, and selectivity for Cl^- ions [21,37] were detected in the ion pathway (Figure 1A). The structural model shows that the central part of the Cl^- ion pathway comprises four glycines (residues 230, 233, 482, and 483) which probably shape this area. In addition, residues Cys278, Met485, Leu418, and Leu581 were found to be close to Cl^- ions superimposed from the template structure, and hence could be potentially involved in a coordination of these ions.

Using literature data, we performed searches for mutations of AA forming the dimer interface or the Cl^- ion pathway in our model for which functional analysis of the wild type-mutant ClC-1 heterodimer was performed. This was done in order to assess the correlation between the localisation of a mutation and the result of functional analysis (presence/absence of a dominant-negative effect) which indicate the type of MC.

In the case of AA localised in the dimer interface (see Table S3 in File S1), 15 positions are associated with 17 mutations (Table 2) for which functional analyses were performed. Fourteen mutations showed a pronounced dominant-negative effect of the mutation in the wild type-mutant heterodimer, 2 a weak dominant-negative effect, and one mutation no dominant-negative effect (Table 2). Structural analysis revealed that most of these mutations occur in positions which are buried in the interface. The positions 317 and 552 are exceptions, localised on the protein surface pointing into extracellular space. Some substitutions with a dominant-negative effect do not exhibit a significant structural defect, e.g. p.(Ile290Met) or p.(Phe306Leu), while others show changes of volume/charge/polarity, e.g. p.(Phe297Ser) or p.(Tyr302His) (Table 2). The mutation p.(Glu291Lys) shows the most significant structural defect and no dominant-negative effect was seen in this case. This mutation exhibits loss of contacts and simultaneously changes of volume and charge. In particular, structural analysis of our ClC-1 model reveals H-bonds between the side chain of Glu291 (positioned in the helix H) and the main chains of Val540 and Ser541 (positioned in the helix P) that stabilize the arrangement between the two helices at the interface (Figure S1 in File S1). Introduction of a positively charged lysine at position 291 prevents such interactions and probably results in protein misfolding. This is in agreement with the identified recessive inheritance of this mutation [36] and with a previous experimental study [34]. In that study, a complete abolishment

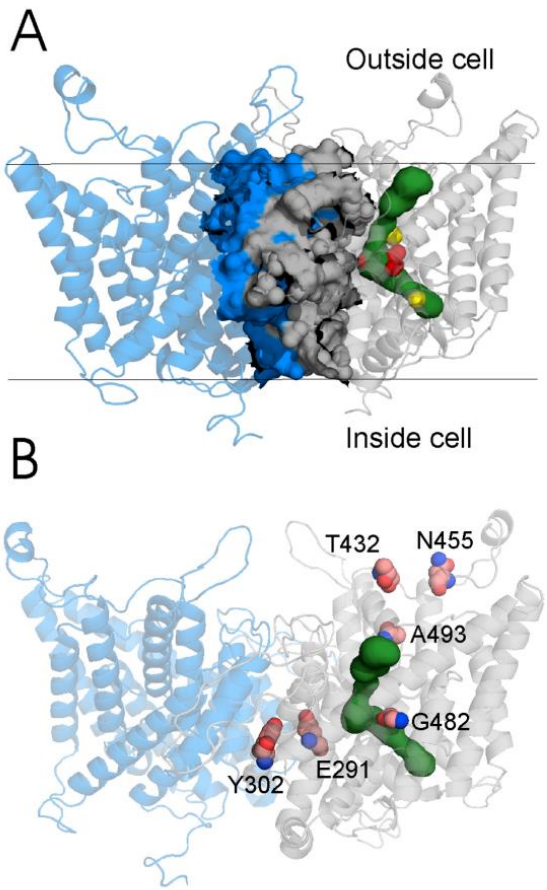


Figure 1. 3D structure of the homology model of the human ClC-1 dimer (transmembrane domains are in blue and gray) with the Cl^- channel (in green) visualized in one subunit. A shows highlighted interface and the channel with the key amino acid Glu 232 (red surface) and two Cl^- ions (yellow balls) which were superimposed from the original X-ray structure of the CmClC transporter. B shows 6 new missense mutations mapped in one subunit.

doi: 10.1371/journal.pone.0082549.g001

Table 2. Previously identified *CLCN1* mutations associated with AA localised directly on the dimer interface or along the Cl^- ion pathway in the dimeric ClC-1 model with functional analysis of the wild type-mutant heterodimer (see also AA in bold letters in Table S3 in File S1).

Mutation/Position in the secondary structure	Result of functional analysis, effect of mutation on wt-mut[$] heterodimer	Side chain interactions of AA in the starting structure	Buriedness of AA (RSA in %)[&]	Change of Volume (Å^3)[@]/ Charge/Polarity/ upon mutation
Mutations localized in the dimer interface				
p.(Leu283Phe)/α-helix H	Dominant-negative effect [16]	No	7	+23/No/No
p.(Ile290Met)/α-helix H	Dominant-negative effect [7,34,38]	No	7	-4/No/No
p.(Glu291Lys)/α-helix H	Protein misfolding and protein degradation [34]	H-bonds: Glu(OE1)-Ser541(N), Glu(OE2)-Val540(N)	9	+30/Yes/No
p.(Glu291Asp)/α-helix H	Dominant-negative effect [34]	H-bonds: Glu(OE1)-Ser541(N), Glu(OE2)-Val540(N)	9	-27/No/No
p.(Phe297Ser/β-strand 1	Dominant-negative effect [39]	No	9	-101/No/Yes
p.(Tyr302His)/α-helix I	Dominant-negative effect [40]	No	9	-35/Yes/Yes
c.(Trp303Arg)/α-helix I	Dominant-negative effect [39]	Clashes[%]	8	-54/Yes/Yes
p.(Phe306Leu)/α-helix I	Dominant-negative effect [39]	No	2	-23/No/No
p.(Phe307Ser)/α-helix I	Dominant-negative effect [8,39]	No	13	-101/No/Yes
p.(Thr310Met)/α-helix I	Dominant-negative effect [16]	No	7	+47/No/Yes
p.(Ala313Thr)/α-helix I	Dominant-negative effect [8,39]	No	4	+28/No/No
p.(Ala313Val)/α-helix I	Dominant-negative effect [39]	No	4	+51/No/Yes
p.(Arg317Gln)/α-helix I	Dominant-negative effect [34]	No	59	-
p.(Gln552Arg)/turn	Dominant-negative effect [34]	No	40	-
p.(Ile553Phe)/turn	Weak dominant-negative effect [41]	No	0	+23/No/No
p.(His555Asn)/turn	Dominant-negative effect [41]	No	16	-
p.(Ile556Asn)/α-helix Q	Weak dominant-negative effect [7,8]	No	0	-53/Yes/Yes
Mutations localized along the Cl- pathway				
p.(Gly230Glu)/turn	Weak dominant-negative effect, mutation does not shift the ClC-1 voltage dependence to positive voltages (as in fully dominant mutations) [36,42,43,44]	No	0	+78/Yes/Yes
p.(Gly233Ser)/α-helix F	No dominant-negative effect [45]	No	0	+29/No/No
p.(Arg421Cys)/α-helix L	Weak dominant-negative effect [40]	H-bond: Arg(NH2)-Phe279(O)	29	-
p.(Phe428Ser)/α-helix L	No dominant-negative effect [16]	Interaction of aromatic rings: Phe428-Phe351	20	-
p.(Met485Val)/α-helix N	No dominant-negative effect [8,9,36]	No	10	+23/No/No
p.(Thr550Met)/turn	Dominant negative effect [16]	No	47	-

[$] Wt-mut refers to wild type-mutant. [&] AA were considered buried if RSA < 15%. [@] Volume change with an absolute value of ≥ 30 Å^3 was considered destabilizing. [%] There are close contacts in this area so it is difficult to detect contacts for this residue

doi: 10.1371/journal.pone.0082549.t002

of channel activity was found in the homodimeric channel, while in the wild type-mutant heterodimeric channel observed currents were about 50% of wild type, consistent with the fact that p.(Glu291Lys) is inherited recessively. The position 291 is known to be associated with various effects according to the mutant AA. For example, p.(Glu291Asp) was found to have a dominant-negative effect in the wild type-mutant ClC-1 heterodimer (Table 2) probably due to conservation of negative charge and partial compensation of the contacts due to the similarity between glutamate and aspartate [34].

In the case of AA localised in the Cl^- ion pathway (Table S3 in File S1), 6 mutations out of 43 were found where functional analysis of the wild-type-mutant heterodimer was performed (Table 2). Five mutations have no pronounced dominant-negative effect in the wild type-mutant heterodimer; two of these are substitutions of glycine residues (positions 230 and 233 mentioned above) of which the first is part of a turn structure while the second one is positioned in an α-helix, typical for transmembrane helices where glycines take part in helix-helix interactions [46]. Substitutions of both glycines are more likely to change the local architecture or even cause misfolding [44,45]. Two other mutations are positioned in the external mouth of the channel (positions 421 and 428) and have single contacts in the starting model (Table 2) which are

Table 3. Structural analysis of newly identified *CLCN1* missense mutations.

Mutation/position in the secondary structure	Localization in the protein	Side chain interactions in the starting structure	Buriedness of AA (RSA in %)[&]	Change of Volume ($Å^3$)[@]/ charge/polarity/ upon mutation
p.(Glu291Gln)/α-helix H	Dimer interface	H-bonds: Glu(OE1)-Ser541(N), Glu(OE2)-Val540(N)	9	5/Yes/No
p.(Tyr302Cys)/α-helix I	Dimer interface	No	9	-85/No/Yes
p.(Thr432Arg)[*]/turn	On the surface	-	-	-
p.(Asn455Tyr)[*]/turn	On the surface	-	-	-
p.(Gly482Glu)/turn	Channel	No	7	+78/Yes/Yes
p.(Ala493Val)/α-helix N	Inside, but outside the channel and dimer interface	No	1	+51/No/Yes

[&] AA were considered buried if RSA < 15%. [@] Volume change with an absolute value of ≥ 30 $Å^3$ was considered destabilizing. [*] This residue is positioned on the surface in extracellular space in the region which is disordered in the X-ray structure. Thus, detailed structural description was not performed.

doi: 10.1371/journal.pone.0082549.t003

abolished upon mutation. The mutation p.(Met485Val) does not show significant structural defects (Table 2), but as mentioned above residue 485 probably coordinates Cl^- ions in the channel, hence p.(Met485Val) might directly impact the channel activity. The mutation p.(Thr550Met) with a dominant-negative effect is positioned in the external mouth of the channel very close to the dimer interface and does not show a significant structural defect (Table 2).

Structural analysis of the novel MC missense mutations

Six novel *CLCN1* missense mutations were identified in our patients (Tables 1, Table 3, and Figure 1B), all associated with recessive MC. The first, p.(Glu291Gln), occurs at AA position 291 (in the dimer interface, see above). This replacement of glutamate with uncharged glutamine most likely results in protein misfolding similarly to p.(Glu291Lys) which agrees with the observed recessive inheritance (Table 1). The mutations p.(Ala493Val) and p.(Gly482Glu) are positioned close to and in the Cl^- ion pathway, respectively, and the first will probably lead to a local structural defect (most likely affecting only the mutant subunit in the wild type-mutant heterodimer) while the second may induce larger structural perturbation due to replacement of the glycine residue in the turn structure with larger charged glutamate (which may lead to misfolding). Both these structural defects are in agreement with the observed recessive phenotype. Two other mutations [p.(Thr432Arg) and p.(Asn455Tyr)] are positioned on the surface of the protein (pointing into extracellular space) in the area which is disordered in the native structure. Therefore, we did not determine their structural impact but considering their positions on the exterior surface they will most likely cause only local structural changes, if any. The mutation p.(Thr432Arg) was identified in patients 20 and 21 on one allele together with another new mutation p.(Tyr302Cys) which occurs on the dimer interface and exhibits a change of volume and polarity (Table 3). Such defect was detected in mutations associated with dominant MC localized at the dimer interface (Table 2). In addition, AA position 302 is associated with mutation p.(Tyr302His) (Table 2) that shows a dominant-negative effect in the wild-type mutant heterodimer [40]. The dominant negative effect of p.(Tyr302His) would explain EMG myotonic "runs" shown by an individual carrying this mutation, but it is obviously not sufficient to cause MC clinical symptoms when present alone in heterozygous state [40]. In our patients 20 and 21, the inheritance pattern of MC corresponds with the recessive type however the detailed neurological examination of their parents was not performed.

Discussion

In this study, we present results of sequence analysis of the *CLCN1* gene performed in Czech patients with myotonia congenita. Mutations associated with the disease were identified in 51 probands; 14 mutations are new, 8 types are frame-shift, splicing, or nonsense and 6 are missense (Table 1). In the cohort of our patients, two mutations p.(Arg894*) and c.1437_1450del are significantly represented and account for 39.6% and 18.8% of mutant MC alleles, respectively. The third most frequent mutation is p.(Phe413Cys), detected in 6% of the disease alleles. This is in line with data for other European countries (Table S4 in File S1). In particular, all three mutations are the most frequently detected mutations in United Kingdom [39]. The mutation p.(Arg894*) is the most frequently observed mutation in Russia [47], Northern Scandinavia [48], and Denmark [32], but it is also significantly represented in Netherlands [49], Spain [40] and Italy [50]. The p.(Phe413Cys) belongs to one of the three most frequent mutations in Northern Scandinavia and Netherlands and the c.1437_1450del mutation to one of the three most frequent mutations in Russia and Denmark.

Further, we created a homology model of the human dimeric ClC-1 channel (Figure 1A) based on the crystallographic

structure of the CmClC transporter which has been shown to share many common features with the other members of the ClC family [21]. Using our *in silico* ClC-1 model, we mapped the new missense mutations onto the protein structure (Figure 1B) and analyzed their impacts. Further, based on our 3D model, we identified AA forming the dimer interface and those forming the Cl^- ion pathway (Figure 1A and Table S3 in File S1). In the next step, we performed searches for mutations of AA forming the dimer interface or the Cl^- ion pathway in our model for which functional analysis of the wild type-mutant ClC-1 heterodimer was performed. This was done in order to assess the correlation between the localisation of a mutation and the result of functional analysis (presence/absence of a dominant-negative effect) which indicate the type of MC. We did not consider clinical data for this correlation as they are often not complete and thus less reliable than the functional analysis.

Out of 40 residues forming the dimer interface, 15 positions were associated with 17 missense mutations for which functional analysis was performed (Table 2). As expected, these mutations were mainly associated with a dominant-negative effect. Structural analysis of the mutations in the homology model revealed that some mutations associated with a dominant-negative effect exhibit change of volume, charge and/or polarity while others do not show significant structural defects (Table 2). This indicates that even relatively conserved substitution at the interface impacts the dimeric structure and results in a dominant-negative effect, in accord with the fact that the dimer interface has a high shape complementarity index (similar to an antibody-antigen interface) [21] leading to low tolerance of substitutions in this area. The only exception considering a dominant-negative effect in the interface was the mutation p.(Glu291Lys) associated with the recessive type of MC. Structural analysis of our homology model showed that this mutation impacts protein structure significantly introducing combined effect (loss of contacts and change of volume and charge) which probably results in protein misfolding in agreement with an experimental study [34]. Mutations leading to a dominant-negative effect can however be found at various positions in the ClC-1 protein and are not associated only with the dimer interface [9,16]. These mutations probably induce rearrangements that propagate through the protein structure and affect the second subunit. Another example is a mutation with a dominant-negative effect p.(Thr550Met) (Table 2) positioned in the external mouth of the Cl^- ion pathway very close to the dimer interface in our model structure; even though this residue does not form part of the interface its proximity might be responsible for the observed dominant-negative effect. Mutations localized inside the Cl^- ion pathway were found to have no dominant-negative effect in the wild type-mutant heterodimer (Table 2). The *in silico* analysis suggested that these mutations can impose various defects, e.g. loss of ability to bind Cl^- ions or various structural disruptions.

Conclusions

In this set of Czech probands with MC, 51 were found with mutation(s) in the *CLCN1* gene, 14 of which were new. In the homology model of the human dimeric ClC-1 channel coded by the *CLCN1* gene, mutations of AA forming the dimer interface are prevalently associated with a dominant-negative effect and dominant inheritance, even though mutations leading to no dominant-negative effect and recessive inheritance can also be found there. We show that dominant/recessive mutations in this area differ in their impact on the protein structure. On the contrary, mutations of AA localized inside the Cl^- ion pathway were found to have no pronounced dominant negative effect. Structural analysis revealed that these mutations can directly affect the channel activity or the local structure of one subunit, or induce misfolding of the mutant subunit. Our results demonstrate structure-function relationships in the ClC-1 protein which are relevant to understanding the molecular pathogenesis of MC.

Supporting Information

File S1. Contains Table S1, S2, S3, S4 and Figure S1.
(DOC)

Acknowledgements

We would like to thank the physicians from Departments of Neurology and Medical Genetics in the Czech Republic (Drs. R. Beharka, V. Curtisová, N. Dvořáčková, E. Ehler, R. Gaillyová, M. Godava, D. Grečmalová,V. Gregor, A. Gřegořová, J. Haberlová, H. Jahnová, Z. Kalina, E. Kantorová, J. Latta, M. Machová, P. Munzar, L. Nováková, Š. Prášilová, M. Šenkeříková, M. Ševčíková, E. Šilhánová, D. Šišková, J. Šoukalová, P. Špalek, K.Veselá, J. Zvolská) for providing us with their patients' blood samples and clinical and laboratory data.

Author Contributions

Conceived and designed the experiments: LF KR DS SV. Performed the experiments: DS JZ ZM KR. Analyzed the data: LF KR DS SV JZ ZM RM PV LM JK. Contributed reagents/ materials/analysis tools: LF KR DS SV JZ ZM RM PV LM JK. Wrote the manuscript: LF KR. Collection of patients, neurological examination: SV RM PV LM JK.

References

1. Lorenz C, Meyer-Kleine C, Steinmeyer K, Koch MC, Jentsch TJ (1994) Genomic organization of the human muscle chloride channel ClC-1 and analysis of novel mutations leading to Becker-type myotonia. Hum Mol Genet 3: 941-946. doi:10.1093/hmg/3.6.941. PubMed: 7951242.
2. Lehmann-Horn F, Rüdel R (1996) Molecular pathophysiology of voltage-gated ion channels. Rev Physiol Biochem Pharmacol 128: 195-268. PubMed: 8791722.
3. Pusch M (2002) Myotonia caused by mutations in the muscle chloride channel gene CLCN1. Hum Mutat 19: 423-434. doi:10.1002/humu.10063. PubMed: 11933197.
4. Middleton RE, Pheasant DJ, Miller C (1994) Purification, reconstitution, and subunit composition of a voltage-gated chloride channel from Torpedo electroplax. Biochemistry 33: 13189-13198. doi:10.1021/bi00249a005. PubMed: 7947726.
5. Middleton RE, Pheasant DJ, Miller C (1996) Homodimeric architecture of a ClC-type chloride ion channel. Nature 383: 337-340. doi:10.1038/383337a0. PubMed: 8848046.
6. Ludewig U, Pusch M, Jentsch TJ (1996) Two physically distinct pores in the dimeric ClC-0 chloride channel. Nature 383: 340-343. doi:10.1038/383340a0. PubMed: 8848047.
7. Saviane C, Conti F, Pusch M (1999) The muscle chloride channel ClC-1 has a double-barreled appearance that is differentially affected in dominant and recessive myotonia. J Gen Physiol 113: 457-468. doi:10.1085/jgp.113.3.457. PubMed: 10051520.
8. Kubisch C, Schmidt-Rose T, Fontaine B, Bretag AH, Jentsch TJ (1998) ClC-1 chloride channel mutations in myotonia congenita: variable penetrance of mutations shifting the voltage dependence. Hum Mol Genet 7: 1753-1760. doi:10.1093/hmg/7.11.1753. PubMed: 9736777.
9. Wollnik B, Kubisch C, Steinmeyer K, Pusch M (1997) Identification of functionally important regions of the muscular chloride channel ClC-1 by analysis of recessive and dominant myotonic mutations. Hum Mol Genet 6: 805-811. doi:10.1093/hmg/6.5.805. PubMed: 9158157.
10. Weinreich F, Jentsch TJ (2001) Pores formed by single subunits in mixed dimers of different CLC chloride channels. J Biol Chem 276: 2347-2353. doi:10.1074/jbc.M005733200. PubMed: 11035003.
11. Aromataris EC, Rychkov GY, Bennetts B, Hughes BP, Bretag AH et al. (2001) Fast and slow gating of CLC-1: differential effects of 2-(4-chlorophenoxy) propionic acid and dominant negative mutations. Mol Pharmacol 60: 200-208. PubMed: 11408615.
12. Duffield M, Rychkov G, Bretag A, Roberts M (2003) Involvement of helices at the dimer interface in ClC-1 common gating. J Gen Physiol 121: 149-161. doi:10.1085/jgp.20028741. PubMed: 12566541.
13. Lossin C, George AL Jr (2008) Myotonia congenita. Adv Genet 63: 25-55. PubMed: 19185184.
14. Zhang J, Bendahhou S, Sanguinetti MC, Ptácek LJ (2000) Functional consequences of chloride channel gene (CLCN1) mutations causing myotonia congenita. Neurology 54: 937-942. doi:10.1212/WNL.54.4.937. PubMed: 10690989.
15. Mindell JA, Maduke M, Miller C, Grigorieff N (2001) Projection structure of a ClC-type chloride channel at 6.5 A resolution. Nature 409: 219-223. doi:10.1038/35051631. PubMed: 11196649.
16. Wu FF, Ryan A, Devaney J, Warnstedt M, Korade-Mirnics Z et al. (2002) Novel CLCN1 mutations with unique clinical and electrophysiological consequences. Brain 125: 2392-2407. doi:10.1093/brain/awf246. PubMed: 12390967.
17. Ma L, Rychkov GY, Hughes BP, Bretag AH (2008) Analysis of carboxyl tail function in the skeletal muscle Cl- channel hClC-1. Biochem J 413: 61-69. doi:10.1042/BJ20071489. PubMed: 18321245.
18. Ma L, Rychkov GY, Bretag AH (2009) Functional study of cytoplasmic loops of human skeletal muscle chloride channel, hClC-1. Int J Biochem Cell Biol 41: 1402-1409. doi:10.1016/j.biocel.2008.12.006. PubMed: 19135547.
19. Cederholm JM, Rychkov GY, Bagley CJ, Bretag AH (2010) Inter-subunit communication and fast gate integrity are important for common gating in hClC-1. Int J Biochem Cell Biol 42: 1182-1188. doi:10.1016/j.biocel.2010.04.004. PubMed: 20398785.
20. Dutzler R, Campbell EB, Cadene M, Chait BT, MacKinnon R (2002) X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. Nature 415: 287-294. doi:10.1038/415287a. PubMed: 11796999.
21. Feng L, Campbell EB, Hsiung Y, MacKinnon R (2010) Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. Science 330: 635-641. doi:10.1126/science.1195230. PubMed: 20929736.
22. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5: 725-738. doi:10.1038/nprot.2010.5. PubMed: 20360767.
23. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14: 33-38. doi:10.1016/0263-7855(96)00018-5. PubMed: 8744570.
24. Berka K, Hanak O, Sehnal D, Banas P, Navratilova V, et al. (2012) MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. Nucleic Acids Res 40: W222-W227.
25. Réblová K, Hrubá Z, Procházková D, Pazdírková R, Pouchlá S et al. (2013) Hyperphenylalaninemia in the Czech Republic: Genotype-phenotype correlations and in silico analysis of novel missense mutations. Clin Chim Acta 419: 1-10. doi:10.1016/j.cca.2013.01.006. PubMed: 23357515.
26. Chothia C (1976) The Nature of the Accessible and Buried Surfaces in Proteins. J Mol Biol 105: 1-12. doi:10.1016/0022-2836(76)90191-1. PubMed: 994183.
27. Zamyatnin AA (1972) Protein Volume in Solution. Prog Biophys Mol Biol 24: 107-123. doi:10.1016/0079-6107(72)90005-3. PubMed: 4566650.
28. Bordo D, Argos P (1991) Suggestions for "safe" residue substitutions in site-directed mutagenesis. J Mol Biol 217: 721-729. doi:10.1016/0022-2836(91)90528-E. PubMed: 2005621.
29. Grantham R (1974) Amino acid difference formula to help explain protein evolution. Science 185: 862-864. doi:10.1126/science.185.4154.862. PubMed: 4843792.
30. Ulzi G, Lecchi M, Sansone V, Redaelli E, Corti E et al. (2012) Myotonia congenita: novel mutations in CLCN1 gene and functional characterizations in Italian patients. J Neurol Sci 318: 65-71. doi:10.1016/j.jns.2012.03.024. PubMed: 22521272.
31. Gao F, Ma FC, Yuan ZF, Yang CW, Li HF et al. (2010) Novel chloride channel gene mutations in two unrelated Chinese families with myotonia congenita. Neurol India 58: 743-746. doi:10.4103/0028-3886.72163. PubMed: 21045501.
32. Colding-Jørgensen E, DunØ OM, Schwartz M, Vissing J (2003) Decrement of compound muscle action potential is related to mutation type in myotonia congenita. Muscle Nerve 27: 449-455. doi:10.1002/mus.10347. PubMed: 12661046.
33. Lehmann-Horn F, Mailänder V, Heine R, George AL (1995) Myotonia levior is a chloride channel disorder. Hum Mol Genet 4: 1397-1402. doi:10.1093/hmg/4.8.1397. PubMed: 7581380.
34. Pusch M, Steinmeyer K, Koch MC, Jentsch TJ (1995) Mutations in dominant human myotonia congenita drastically alter the voltage dependence of the CIC-1 chloride channel. Neuron 15: 1455-1463. doi:10.1016/0896-6273(95)90023-3. PubMed: 8845168.
35. Koty PP, Pegoraro E, Hobson G, Marks HG, Turel A et al. (1996) Myotonia and the muscle chloride channel: dominant mutations show variable penetrance and founder effect. Neurology 47: 963-968. doi:10.1212/WNL.47.4.963. PubMed: 8857727.
36. Meyer-Kleine C, Steinmeyer K, Ricker K, Jentsch TJ, Koch MC (1995) Spectrum of mutations in the major human skeletal muscle chloride channel gene (CLCN1) leading to myotonia. Am J Hum Genet 57: 1325-1334. PubMed: 8533761.
37. Dutzler R, Campbell EB, MacKinnon R (2003) Gating the selectivity filter in ClC chloride channels. Science 300: 108-112. doi:10.1126/science.1082708. PubMed: 12649487.
38. Accardi A, Pusch M (2000) Fast and slow gating relaxations in the muscle chloride channel CLC-1. J Gen Physiol 116: 433-444. doi:10.1085/jgp.116.3.433. PubMed: 10962018.
39. Fialho D, Schorge S, Pucovska U, Davies NP, Labrum R et al. (2007) Chloride channel myotonia: exon 8 hot-spot for dominant-negative interactions. Brain 130: 3265-3274. doi:10.1093/brain/awm248. PubMed: 17932099.
40. Mazón MJ, Barros F, De la Peña P, Quesada JF, Escudero A et al. (2012) Screening for mutations in Spanish families with myotonia. Functional analysis of novel mutations in CLCN1 gene. Neuromuscul Disord 22: 231-243. doi:10.1016/j.nmd.2011.10.013. PubMed: 22094069.
41. Burgunder JM, Huifang S, Beguin P, Baur R, Eng CS et al. (2008) Novel chloride channel mutations leading to mild myotonia among Chinese. Neuromuscul Disord 18: 633-640. doi:10.1016/j.nmd.2008.05.007. PubMed: 18579381.
42. George AL Jr, Sloan-Brown K, Fenichel GM, Mitchell GA, Spiegel R et al. (1994) Nonsense and missense mutations of the muscle chloride channel gene in patients with myotonia congenita. Hum Mol Genet 3: 2071-2072. PubMed: 7874130.
43. Steinmeyer K, Lorenz C, Pusch M, Koch MC, Jentsch TJ (1994) Multimeric structure of ClC-1 chloride channel revealed by mutations in dominant myotonia congenita (Thomsen). EMBO J 13: 737-743. PubMed: 8112288.

44. Fahlke C, Beck CL, George AL Jr (1997) A mutation in autosomal dominant myotonia congenita affects pore properties of the muscle chloride channel. Proc Natl Acad Sci U S A 94: 2729-2734. doi: 10.1073/pnas.94.6.2729. PubMed: 9122265.
45. Richman DP, Yu Y, Lee TT, Tseng PY, Yu WP et al. (2012) Dominantly inherited myotonia congenita resulting from a mutation that increases open probability of the muscle chloride channel CLC-1. Neuromolecular Med 14: 328-337. doi:10.1007/s12017-012-8190-1. PubMed: 22790975.
46. Javadpour MM, Eilers M, Groesbeek M, Smith SO (1999) Helix packing in polytopic membrane proteins: Role of glycine in transmembrane helix association. Biophys J 77: 1609-1618. doi:10.1016/S0006-3495(99)77009-8. PubMed: 10465772.
47. Ivanova EA, Dadali EL, Fedotov VP, Kurbatov SA, Rudenskaya GE, et al. (2012) The spectrum of CLCN1 gene mutations in patients with nondystrophic Thomsen's and Becker's myotonias. Russ J Genet+ 48: 952-961
48. Sun C, Tranebjaerg L, Torbergsen T, Holmgren G, Van Ghelue M (2001) Spectrum of CLCN1 mutations in patients with myotonia congenita in Northern Scandinavia. Eur J Hum Genet 9: 903-909. doi: 10.1038/sj.ejhg.5200736. PubMed: 11840191.
49. Trip J, Drost G, Verbove DJ, van der Kooi AJ, Kuks JBM et al. (2008) In tandem analysis of CLCN1 and SCN4A greatly enhances mutation detection in families with non-dystrophic myotonia. Eur J Hum Genet 16: 921-929. doi:10.1038/ejhg.2008.39. PubMed: 18337730.
50. Brugnoni R, Kapetis D, Imbrici P, Pessia M, Canioni E et al. (2013) A large cohort of myotonia congenita probands: novel mutations and a high-frequency mutation region in exons 4 and 5 of the CLCN1 gene. J Hum Genet 58: 581-587. doi:10.1038/jhg.2013.58. PubMed: 23739125.

PERGAMON

Neuromuscular Disorders 17 (2007) 476–481

www.elsevier.com/locate/nmd

Analysis of point mutations in the SMN1 gene in SMA patients bearing a single SMN1 copy

Eva Zapletalová [a,b], Petra Hedvičáková [c], Libor Kozák [a], Petr Vondráček [d], Renata Gaillyová [e], Tat'ána Maříková [c], Zdeněk Kalina [e], Věra Jüttnerová [f], Jiří Fajkus [b], Lenka Fajkusová [a,b,*]

[a] *University Hospital Brno, Centre of Molecular Biology and Gene Therapy, Brno, Czech Republic*
[b] *Masaryk University, Faculty of Science, Department of Functional Genomics and Proteomics, Brno, Czech Republic*
[c] *University Hospital Motol, Institute of Biology and Medical Genetics, 2nd Faculty of Medicine, Prague, Czech Republic*
[d] *University Hospital Brno, Department of Child Neurology, Brno, Czech Republic*
[e] *University Hospital Brno, Department of Medical Genetics, Brno, Czech Republic*
[f] *University Hospital Hradec Králové, Department of Medical Genetics, Hradec Králové, Czech Republic*

Received 21 November 2006; received in revised form 19 February 2007; accepted 4 March 2007

Abstract

Spinal muscular atrophy (SMA) is caused by homozygous deletion of the SMN1 gene in approximately 96% of cases. Four percent of SMA patients have a combination of the deletion or conversion on one allele and an intragenic mutation on the second one. We performed analysis of point mutations in a set of our patients with suspicion of SMA and without homozygous deletion of the SMN1 gene. A quantitative test determining SMN1 copy number (using real-time PCR and/or MLPA analysis) was performed in 301 patients and only 1 SMN1 copy was detected in 14 of them. When these 14 patients were screened for the presence of point mutations we identified 6 mutations, p.Y272C (in three patients) and p.T274I, p.I33IfsX6, and p.A188S (each in one case). The mutations p.I33IfsX6 and p.A188S were found in two SMAI patients and were not detected previously. Further, evaluation of the relationship between mutation type, copy number of the SMN2 gene and clinical findings was performed. Among our SMA patients with a SMN1 homozygous deletion, we found a family with two patients: the son with SMAII possesses 3 SMN2 copies and the nearly asymptomatic father has a homozygous deletion of SMN1 exon 7 and carries 4 SMN2 copies. Generally, our results illustrate that an increased SMN2 gene copy number is associated with a milder SMA phenotype.

Keywords: SMA; SMN1; SMN2; Real-time PCR; Copy number; Point mutation

1. Introduction

Spinal muscular atrophy (SMA), with an incidence of 1/6000–1/10,000 and a carrier frequency of 1/40, is the second most frequent lethal autosomal recessive disease in Europeans next to cystic fibrosis [1,2]. The SMA-determining gene, termed survival motor neuron (SMN), is present on 5q13 in two copies, a telomeric SMN1 gene and a centromeric SMN2 gene which are highly homologous and contain only five base-pair differences [3]. The SMN protein plays a crucial role in the generation of the pre-mRNA splicing machinery and thus in mRNA biogenesis [4]. Lefebvre concluded that the SMN gene has 8 exons [3]; however Burglen characterized the gene in more detail and showed that it consists of 9 exons [5] and in order not to confuse

* Corresponding author. Address: Centre of Molecular Biology and Gene Therapy, Cernopolní 9 CZ-62500, Brno, Czech Republic. Tel.: +420 532234625; fax: +420 532234623.
E-mail address: lfajkusova@fnbrno.cz (L. Fajkusová).

0960-8966/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.nmd.2007.03.003

previously published mutation data, exon 2 was referred to as exon 2a and 2b. The stop codon occurs in exon 7.

Exon 7 of the SMN1 gene is not detectable in approximately 96% of SMA patients, owing to either deletion of SMN1 or conversion of the SMN1 sequence to SMN2. Approximately 4% of patients have a combination of the deletion or conversion on one allele and an intragenic mutation on the second one. The centromeric SMN2 gene cannot compensate for the SMN1 defect because single nucleotide difference in exon 7 causes exon skipping in about 90% of SMN2 transcripts [6]. However, increased SMN2 gene copy number, which can occur as the result of gene conversion events, is associated with a milder SMA phenotype [7].

The C terminus of SMN, including the sequence encoded by exon 7 containing a highly conserved tyrosine/glycine-rich sequence (Y/G box), is required for SMN's nucleic acid and protein binding activity as well as for its oligomerization [8]. Several missense mutation clusters have been described in and around the Y/G box, including p.S262I, p.Y272C, p.T274I, p.G275S, and p.G279V. One of these mutations, p.Y272C, is associated with the most severe form of the disease (type SMAI), and among different subtle SMN1 mutations p.Y272C is the most frequent (20% of SMN1 mutations) [9].

We present results of molecular genetic analyses performed in the set of our patients who do not have homozygous deletion of the SMN1 exon 7, and in whom a quantitative test showed a single copy of this exon. By using long-range PCR, PCR and sequencing we identified six mutations (four different types). Two of these, p.I33IfsX6 and p.A188S, are novel ones not described so far. Further, we performed analysis of SMN2 copy number in these patients and evaluated relationships among the detected mutation, SMN2 copy number, and clinical manifestation of the disease. Moreover, we performed also analysis of SMN2 copy number in 70 patients with SMAI, II, and III. We found a family with two SMA patients where the son has been diagnosed as SMAII and possesses 3 SMN2 copies, and the nearly asymptomatic father has also homozygous deletion of SMN1 exon 7 and carries 4 SMN2 copies.

2. Patients and methods

2.1. Patients

Patients with supposed SMA diagnosis were tested for the presence of homozygous deletion of SMN1. This test is based on the single nucleotide differences in exons 7 and 8 that distinguish SMN1 and SMN2 [5]. Genomic DNA was isolated from peripheral blood by the salting-out method [10]. DNA concentration was initially determined from UV absorbance at 260 nm using a NanoDrop instrument (NanoDrop Technologies), and after dilution to ~10 ng/μl the exact concentration was established from UV absorbance.

2.2. Real-time PCR for determination of SMN1 gene copy number

Real-time detection of PCR products was performed with TaqMan hybridization probes labelled in the case of the SMN1 gene with FAM reporter dye, and in the case of a reference gene (the ALB gene, albumin) with JOE reporter dye. Sequences of primers and probes used are given in Table 1. Multiplex PCR was performed in a total volume of 25 μl, containing 50 ng of genomic DNA, 400 nM primers specific for the SMN1 gene, 400 nM primers specific for the ALB gene, 200 nM each of probes, and 1× TaqMan Universal PCR Master Mix (Roche). Real-time PCRs were carried out using a RotorGene 3000 instrument (Corbett Research). The following PCR conditions were used: 95 °C/10 min, followed by 40 cycles of 95 °C/15 s, and 60 °C/30 s. DNA samples were amplified in two parallel PCRs and each run contained two control genomic DNAs (the calibrators), the first with one SMN1 copy and the second with two SMN1 copies per genome. The number of SMN1 gene copies was determined using the comparative C_t method [11]. The result of this analysis determines the normalized SMN1 gene copy

Table 1
Primers and probes used for real-time PCR

Fragment	Name	Sequence (5′ → 3′ direction)
SMN1	SMN1-F	ACTTCCTTTTATTTTCCTTACAGGGTTTC
	SMN1-R	AATGCTGGCAGACTTACTCCTTAATTTAA
	SMN1-probe	FAM-ACAAAATCAAAAAGAAGGAAGGTGCTCACATTC-BHQ1
SMN2	SMN2-F	AATGCTTTTTAACATCCATATAAAGCT
	SMN2-R	CCTTAATTTAAGGAATGTGAGCACC
	SMN2-probe	FAM-TGATTTGTCTAAAACCC-MGB
Albumin	ALB-F	GCTGTCATCTCTTGTGGGCTGT
	ALB-R	ACTCATGGGAGCTGCTGGTTC
	ALB-probe	JOE-CCTGTCATGCCCACACAAATCTCTCC-BHQ1

number in the sample relative to the normalized copy number in the calibrator.

2.3. MLPA analysis for determination of SMN1 gene copy number

Multiplex ligation-dependent probe amplification (MLPA) analysis was performed according to the manufacturer's instructions (MRC-Holland). The SALSA MLPA KIT P021 SMA region was used which contains 37 probes, 16 of which are specific for the SMA region while the others are controls for other human genes. The SMN1-D01 and SMN1-D07 probes result in peaks at 270 and 295 nucleotides in the presence of SMN1 exons 7 and 8, respectively, and a reduced peak area indicates a reduction in the copy number of the SMN1 gene. After the MLPA reaction the samples were used for fragment analysis on an ABI PRISM 3100 Avant Genetic Analyser using GeneScan3.7 software (Applied Biosystems). Each reaction was analysed and the relative peak areas were calculated using software provided by NGRL Manchester. A ratio under 0.7 was taken as a sign of the presence of only 1 copy of the SMN1 gene.

2.4. Long-range PCR and PCRs for point mutation analysis

Long-range PCR of the region including exons 2a–6 of the SMN1 gene was carried out as described by Clermont et al. [12] and the product was visualized in a 0.8% agarose gel, excised, and extracted using a QIAquick Gel Extraction Kit (Qiagen). The eluted DNA was used for amplification of individual SMN1 exons. Sequences of primers and annealing temperatures are given in Table 2. PCRs were performed in a total volume of 25 μl, containing 1 μl of eluted long-range PCR product, 400 nM primers, 2.5 mM $MgCl_2$, 200 μM dNTPs, 1 U Taq polymerase (Fermentas), and 1× PCR buffer. The following temperature conditions were used: 95 °C/10 min, followed by 30 cycles of 95 °C/15 s, annealing temperature/30 s, and 72 °C/45 s. PCR products were extracted and sequenced on an ABI PRISM 310 sequencer (Applied Biosystems).

Exons 7 and 1 were amplified directly from genomic DNA. In the case of exon 7, primers specific for SMN1 could be used, while in case of exon 1 the primers amplified both SMN1 and SMN2 exon 1. PCR products were extracted and sequenced; moreover, the product of exon 1 was also cleaved with the restriction enzymes *Bst*NI and *Fok*I which can distinguish two detected mutations c.5C > G (p.A2G) [13] and c.43C > T (p.Q15X) [14] in exon 1.

2.5. Real-time PCR for determination of SMN2 gene copy number

The MGB® probe labelled with FAM reporter dye was used for determination of SMN2 gene copy number. Sequences of primers and the probe are described in the study of Anhuf et al. [15]. The ABL gene was used as a reference locus, and the primers and probe were the same as in the case of real-time PCR for determination of SMN1 gene copy number. PCR for SMN2 amplification was performed in a total volume of 25 μl, containing 50 ng of genomic DNA, 400 nM primers, 70 nM probe, and 1× TaqMan Universal PCR Master Mix (Roche); PCR for ABL amplification was performed in the same volume and contained 50 ng of genomic DNA, 400 nM primers, 200 nM probe, and 1× TaqMan Universal PCR Master Mix (Roche). DNA samples were amplified in two parallel PCRs. Further, two control DNAs with two and four SMN2 copies were amplified as calibrators for determina-

Table 2
Primers used for long-range PCR and PCR

Fragment	Name	Sequence (5′ → 3′ direction)	Length (bp)	Annealing (°C)
Exons 2a 6	SMN-LR-F	CCTTCCTTCTTTTTGATTTTGTCTGA	13219	63
	SMN-LR-R	TGTGTGGATTAAGATGACTC		
Exon 2a	SMN2A-F	TCCTTTCCAAATGAATAACGAGA	219	55
	SMN2A-R	TGTGTGGATTAAGATGACTCTTGG		
Exon 2b	SMN2B-F	AAGGACTAATGAGACATCCTTTGAA	259	60
	SMN2B-R	TGTGCACCACCCTGTAACAT		
Exon 3	SMN3-F	TTGCCCTCTTCAAAAGAAATG	299	55
	SMN3-R	TCTCTGCTTCCAGAAATTGAA		
Exon 4	SMN4-F	CAAAAGTTTCATGGGAGAGC	278	55
	SMN4-R	TTTTCTTAATCACACCCTTATAACAAA		
Exon 5	SMN5-F	CCCAAGGGATGTTCTACAATG	402	60
	SMN5-R	TTCCTATCATATTGAAATTGGTAAGTT		
Exon 6	SMN6-F	TGCAAGAGTAATTTAAGCCTCAGA	442	55
	SMN6-R	CCTCCCATATGTCCAGATTCTC		
Exon 7	SMN7-F	TCCTTTTATTTTCCTTACAGGGTTTC	301	58
	SMN7-R	GTTTCTTCCACATAACCAACCAG		
Exon 1	SMN1-F	GCGAGGCTCTGTCTCAAAAC	388	58
	SMN1-R	GGGTGCTGAGAGCGCTAATA		

tion of copy number using the comparative C_t method [11].

3. Results

3.1. Analysis of point mutations

The results concerning SMN1 point mutations were obtained from two hospitals, University Hospital Brno and University Hospital Motol. In University Hospital Brno, the quantitative test determining SMN1 copy number based on multiplex real-time PCR was used to screen 156 unrelated patients with suspicion of SMA and without homozygous deletion of the SMN1 gene. In 8 of these, 1 SMN1 copy was detected. In University Hospital Motol, the quantitative test was based on MLPA analysis; 145 unrelated patients without homozygous deletion of the SMN1 gene were tested and in 6 of these, 1 SMN1 copy was detected. By using long-range PCR, PCR and sequencing, we identified 6 mutations: p.Y272C (patients 1, 2, and 3), p.T274I (patient 4), p.I33IfsX6 (patient 5), and p.A188S (patient 6). The mutations p.I33IfsX6 (exon 2a) and p.A188S (exon 4) were not detected so far.

3.2. Analysis of SMN2 gene copy number

SMN2 copy number analysis was performed in patients with a genotype point mutation/exon 7 deletion, and also in 70 patients with homozygous deletion of exon 7. The results for patients with a point mutation are presented in Table 3 and the results relating to patients with the homozygous deletion are given in Table 4.

Table 4
Determination of SMN2 copy number in 70 SMA patients

Type of SMA	Number of patients	Copy number of the SMN2 gene			
		2 copies	3 copies	4 copies	5 copies
SMAI	10	7 (70%)	3 (30%)		
SMAII	40	1 (2.5%)	36 (90%)	3 (7.5%)	
SMAIII	20	1 (5%)	10 (50%)	8 (40%)	1 (5%)

We also performed SMN2 copy number analysis in the SMAII patient with homozygous deletion of exon 7 and in his nearly asymptomatic father who has also homozygous deletion of exon 7. The son was born in 1994 and has never been able to walk independently. His father was born in 1963 and had no problems until 1996, when he had an accident. In 1997, he was investigated for persistent difficulties and myopathy was found. Diagnosis of SMA was confirmed in 2005 by chance, when we analysed SMN1 copy number in parents of SMA patients. At present, he is still able to move without greater problems. Analysis of SMN2 copy number detected 3 copies in the affected son and 4 copies in his father.

4. Discussion

Although the majority of SMA cases results from homozygous deletion of SMN1, the list of SMN1 point

Table 3
Results of DNA analyses and clinical findings

Patients	Year of birth	Mutation, localization	SMN2 copy number	Phenotype
1	93	c.815A > G, p.Y272C, exon 6	2	*SMAI*. At the age of 4 months severe peripheral hypotonia, floppy infant, delayed motor milestones. At the age of 6 months only supine position, no rolling, muscle weakness and wasting. Died at the age of 16 months due to bronchopneumonia
2	90	c.815A > G, p.Y272C, exon 6	3	*SMAII*. At the age of 5 months delayed motor milestones, peripheral hypotonia. At the age of 2 years only sitting, not standing. Respiratory insufficiency since 9 years of age. No disease information since 1999
3	82	c.815A > G, p.Y272C, exon 6	2	*SMAII*. At the age of 6 months delayed motor milestones. At the age of 2 years only sitting, not standing. At the age of 10 years tendon contractures and progression of skeletal deformities. Respiratory insufficiency and recurrent pneumonias since 9 years of age. No disease information since 2003
4	87	c.821C > T, p.T274I, exon 6	2	*SMAIII*. Normal motor milestones until 18 months of age. Walking since 11 months. At the age of 2 years unable to walk, only sitting. Progression of scoliosis, repeated elongations of Achille's tendons. At the age of 14 years severe skeletal deformities, multiple tendon contractures, progression of muscle weakness and wasting. No disease information since 2001
5	02	c.98delT, p.I33IfsX6, exon 2a	2	*SMAI*. Severe peripheral hypotonia since birth, global respiratory insufficiency, no further disease information
6	91	c.562G > T, p.A188S, exon 4	Not performed	*SMAI*. Severe peripheral hypotonia since birth. Died at the age of 5 months due to bronchopneumonia

mutations is increasing (Leiden muscular dystrophy pages, www.dmd.nl.). The SMN protein is a component of large macromolecular complexes that are found both in cytoplasmic and nuclear compartments. Critical insight into SMN function came from findings that the protein interacts with Sm proteins, core components of small nuclear ribonucleoproteins (snRNP). Cytoplasmic SMN plays an essential role in snRNP biogenesis and is required for the transport of the snRNP complex into nucleus and in the nucleus, SMN protein is required for regenerating an active splicing complex. Shpargel et al., analysed the abilities of eight SMN mutations to rescue Sm core assembly [16] and five SMAI alleles (including p.Y272C and ΔEx7) showed only low levels of Sm core assembly activity, whereas both of the SMAIII alleles (including p.T274I) functioned similarly to the wild-type construct. Moreover, the p.Y272C and ΔEx7 mutations have been shown to disrupt SMN oligomerization with a concomitant or downstream defect in Sm protein binding. In our set of SMA patients, we identified the mutation p.Y272C in association with 2 SMN2 copies in two patients (the first was SMAI and the second SMAII) and the same mutation in association with 3 SMN2 copies in one SMAII patient. The mutation p.T274I was detected together with 2 SMN2 copies in one SMAIII patient.

Further, we found two mutations, p.I33IfsX6 and p.A188S which are described here for the first time; the frame-shift mutation p.I33IfsX6 together with 2 SMN2 copies was detected in a SMAI patient, and the missense mutation p.A188S also in a SMAI patient where unfortunately the SMN2 copy number could not be determined.

The region of the SMN protein coded by the exon carrying p.A188S has not been believed to belong to the functionally most important SMN domains, but a PSI-BLAST search over all non-redundant GenBank CDS using BLOSUM62 matrix (gap costs existence = 11, ext. = 1) shows that the sequence in the 49 amino acid-long region surrounding this site (positions 164–212 in the human SMN protein sequence, QVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPP PPMPGPRLGPGKPGL) is highly conserved throughout all mammalian sequences, with the only exemption of *Felis catus* (cat) with a T in position 188. The A in position 188 is apparently a part of a conserved APWNSFLPPPPP motif and in a secondary-structure prediction (Protean, DNASTAR, HITACHI) is located at the boundary between hydrophilic and hydrophobic regions, probably forming a turn region between two coils. Searching for structural motifs in this region of the protein, between a TUDOR domain (positions 90–149) and a low complexity domain (positions 194–251) using SMART in normal mode (http://smart.embl-heidelberg.de) revealed no known structural domains or segments of known function.

We tested the possibility of a modification of pre-mRNA splicing owing to the p.A188S mutation using the programme ESEfinder (http://rulai.cshl.edu/tools/ESE), a web-based resource that facilitates rapid analysis of exon sequences to identify putative ESEs (exon splicing enhancer sequences) responsive to the human SR proteins SF2/ASF, SC35, SRp40, and SRp55, and to predict whether exonic mutations disrupt such elements. This test showed that the p.A188S mutation abolishes responsivity to SC35 and creates the possibility of SRp40 binding.

Our results concerning correlations between SMN2 copy number and clinical findings are in agreement with literature data [7] that increased copy number is connected with milder SMA phenotype; in the case of the family with two SMA patients, we detected 3 SMN2 copies in the SMAII patient and 4 copies in his nearly asymptomatic father. However, this rule is not general because we found 2 SMN2 copies in SMAIII patients and 4 copies in SMAI patients, so evidently factors additional to SMN genes are involved in the resulting clinical manifestation, pointing to the necessity of further research to gain a deeper understanding of the molecular causes of SMA.

Acknowledgements

This work was supported by IGA MHCR, project 1A/8608-4, and the institutional support MSM0021 622415, MSMT-LC06023 and MZO00064203. The authors would like to thank Professor Ronald Hancock, Hôtel-Dieu Hospital, Quebec, Canada, for the manuscript revision.

References

[1] Melki J, Lefebvre S, Burglen L, burlet P, Clermont O, Millasseau P, et al. De novo and inherited deletions of the 5q13 region in spinal muscular atrophies. Science 1994;264:1474–7.
[2] Ogino S, Wilson RB, Gold B. New insights on the evolution of the SMN1 and SMN2 region: simulation and meta-analysis for allele and haplotype frequency calculations. Eur J Hum Genet 2004;12:1015–23.
[3] Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L, et al. Identification and characterization of a spinal muscular atrophy-determining gene. Cell 1995;80:155–65.
[4] Pellizzoni L, Charroux B, Dreyfuss G. SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. Proc Natl Acad Sci USA 1999;96:11167–72.
[5] Burglen L, Lefebvre S, Clermont O, Burlet P, Viollet L, Crueud C, et al. Structure and organization of the human survival motor neurone (SMN) gene. Genomics 1996;32:479–82.
[6] Cartegni L, Krainer AR. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat Genet 2002;30:377–84.
[7] Parsons DW, McAndrew PE, Monani UR, et al. An 11 base pair duplication in exon 6 of the SMN gene produces a type I spinal

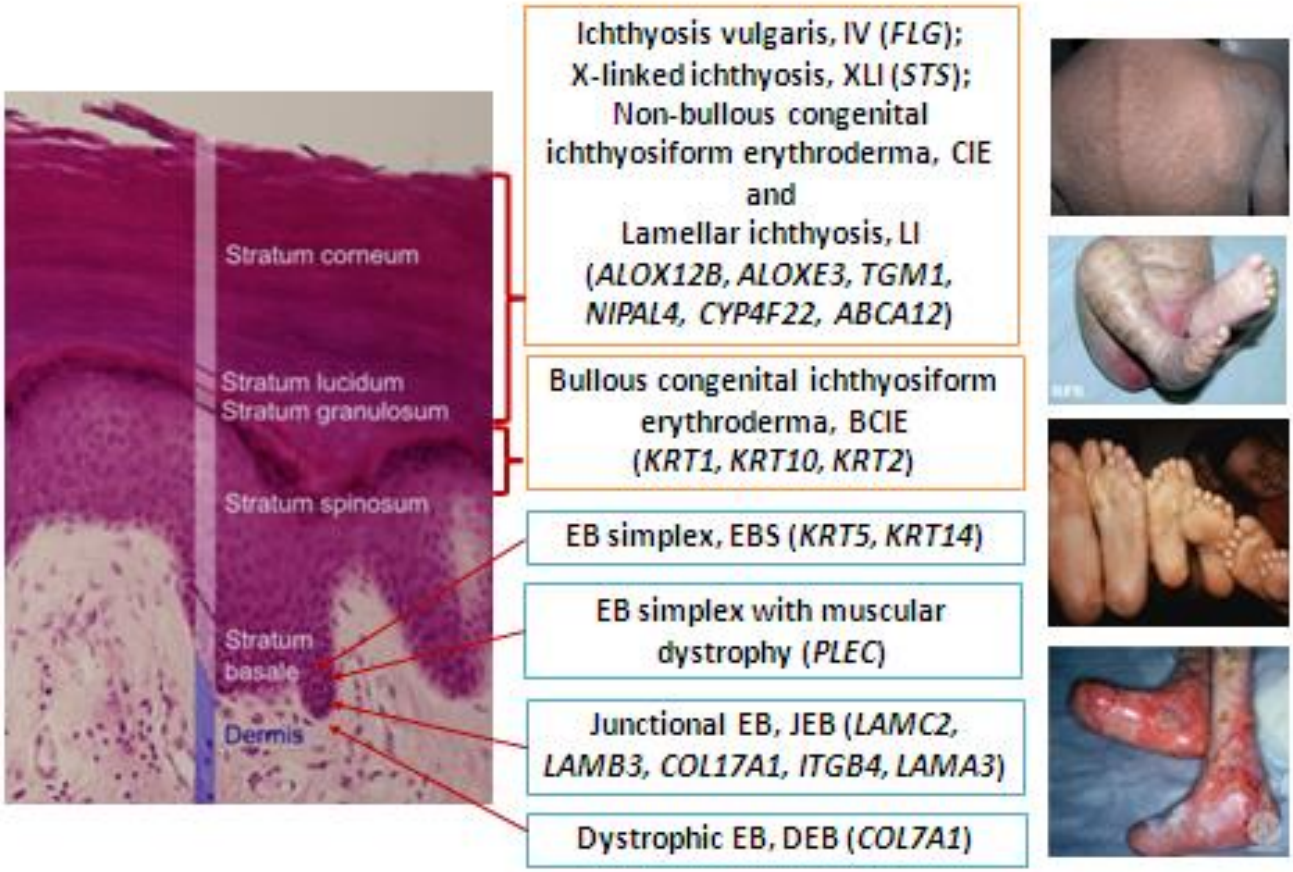
muscular atrophy (SMA) phenotype: further evidence for SMN as the primary SMA-determining gene. Hum Mol Genet 1996;5:1727 32.
[8] Wang J, Dreyfuss G. Characterization of functional domains of the SMN protein in vivo. J Biol Chem 2001;276:45387 93.
[9] Wirth B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). Hum Mutat 2000;15:228 37.
[10] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
[11] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.
[12] Clermont O, Burlet P, Benit P, et al. Molecular analysis of SMA patients without homozygous SMN1 deletions using a new strategy for identification of SMN1 subtle mutations. Hum Mutat 2004;24:417 27.
[13] Parsons DW, McAndrew PE, Iannaccone ST, Mendell JR, Burghes AH, Prior TW. Intragenic telSMN mutations: frequency, distribution, evidence of a founder effect, and modification of the spinal muscular atrophy phenotype by cenSMN copy number. Am J Hum Genet 1998;63:1712 23.
[14] Wirth B, Herz M, Wetter A, Moskau S, Hahnen E, Rudnik-Schoneborn S, et al. Quantitative analysis of survival motor neuron copies: identification of subtle SMN1 mutations in patients with spinal muscular atrophy, genotype phenotype correlation, and implications for genetic counseling. Am J Hum Genet 1999;64:1340 56.
[15] Anhuf D, Eggermann T, Rudnik-Schoneborn S, Zerres K. Determination of SMN1 and SMN2 copy number using TaqMan technology. Hum Mutat 2003;22:74 8.
[16] Shpargel KB, Matera AG. Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. Proc Natl Acad Sci USA 2005;102:17372 7.

3 Dědičné kožní nemoci

Genodermatózy jsou dědičná onemocnění kůže, která jsou charakterizována velkou klinickou a genetickou variabilitou. V CMBGT se molekulárně genetickou diagnostikou genodermatóz zabýváme od roku 2005, kdy jsme zavedli diagnostiku epidermolysis bullosa dystrophica (analýza mutací v genu *COL7A1*) a epidermolysis bullosa simplex (analýza mutací v genech *KRT5* a *KRT14*). V roce 2013 jsme pak začali s molekulárně genetickou diagnostikou ichtyózy vulgaris (analýza mutací v genu *FLG*), X-vázané ichtyózy (stanovení delece genu *STS*), autozomálně recesivních kongenitálních ichtyóz (analýza mutací v genech *ALOX12B, ALOXE3*, *NIPAL4, CYP4F22, TGM1*) a incontinentia pigmenti (analýza mutací v genu *NEMO*).

Nejčastějšími metodickými přístupy v genetické diagnostice genodermatóz jsou PCR a sekvenční analýza. V současné době se snažíme problematiku dědičných onemocnění kůže postihnout v co možná nejširším měřítku a zavádíme, podobně jako v případě neuromuskulárních nemocí, techniky sekvenování nové generace (SeqCap-TR) pro analýzu všech genů asociovaných s vybranými typy nemocí.

V oblasti genodermatóz spolupracujeme především s Kožním oddělením Pediatrické kliniky Fakultní nemocnice Brno, Oddělením lékařské genetiky Fakultní nemocnice Brno a 1. Patologicko-anatomickým ústavem Fakultní nemocnice u sv. Anny. V těsné spolupráci s uvedenými pracovišti provádíme genetickou analýzu pacientů s podezřením na některou z výše uvedených nemocí a následně korelace mezi genetickými, klinickými a patologickými nálezy.



Obrázek 1: Schématické znázornění lokalizace proteinů asociovaných s onemocněním epidermolysis bullosa a ichtyóza v kožním řezu; obrázky klinických projevů epidermolysis bullosa a ichtyóza

3.1 Epidermolysis bullosa

Epidermolysis bullosa (EB) je onemocnění charakterizované extrémně křehkou kůží a puchýři. EB je tradičně dělena do tří hlavních typů v závislosti na lokalizaci místa vzniku puchýře: EB simplex (EBS, puchýř vzniká v epidermis), junkční EB (JEB, puchýř vzniká na úrovni bazální membrány), dystrofická EB (DEB, puchýř vzniká v dermis těsně pod bazální membránou). Vzhledem k tomu, že se v současné době zabýváme molekulárně genetickou diagnostikou pouze DEB a EBS, další část bude podrobněji pojednávat jen o těchto typech EB.

Dystrofická EB může být děděna autosomálně dominantně nebo recesivně. Oba typy dědičnosti DEB jsou spojeny s mutacemi v genu, který kóduje kolagen typu VII (*COL7A1*) [42]. Fenotypové projevy onemocnění mají rozsah od mírné "*nail-only*" dominantní DEB až po nejzávažnější recesivní DEB s generalizovanými puchýři a jizvením, které vede k fúzi prstů na rukou a nohou, kromě kožních projevů mohou být postiženy i sliznice [43]. U dominantní DEB patogenní mutace zahrnují heterozygotní, dominantně negativní substituce glycinu. Recesivní DEB je způsobena kombinací mutací typu *nonsense, splicing, frame-shift, missense*. Některé případy recesivní DEB jsou také spojeny se substitucemi glycinu, které ale nemají klinický dopad, pokud druhá alela nenese mutaci [44],[45].

Kolagen typu VII vytváří kotvící fibrily, které spojují dermis s bazální membránou. Gen *COL7A1* kóduje alfa1-řetězec prokolagenu typu VII [proa1(VII)]. Tři řetězce proa1(VII) vytváří monomer uspořádaný do trojité kolagenní šroubovice s nekolagenními N- a C-koncovými doménami, následuje vznik antiparalelního dimeru ze dvou monomerů, odštěpení C-koncových domén a laterální agregace dimerů za vzniku kotvících fibril [46],[47]. Pro kolagenní trojšroubovici je charakteristická opakující se sekvence Glycin-X-Y, kde X a Y představují často lyzin nebo prolin. Přesné zachování glycinu v každé třetí pozici je požadováno z důvodu těsného sbalení tří řetězců proa1(VII). Glycin, jako nejmenší aminokyselina, je ukrytý ve středu trojšroubovice a jakákoli záměna této aminokyseliny vede k narušení struktury [48]. Kolagenní doména je několikrát přerušena krátkými nonhelikálními úseky, které jsou důležité pro intramolekulární flexibilitu proteinu. Klinicko-genetické korelace popsané v literatuře i provedené u našich 61 pacientů s mutacemi v genu *COL7A1* většinou odpovídají předpokládu, že kombinace dvou mutací vytvářejících předčasný terminační kodón (PTC) má závažnější klinický dopad než kombinace *missense* a PTC mutace. V případě srovnání klinických projevů pacientů se substitucí glycinu a PTC mutací a pacientů s *missense* mutací (jiná než substituce glycinu) a PTC mutací závažnější klinický projev mají pacienti nesoucí záměnu glycinu [publikace *J Dermatol Sci. 2010 Aug;59(2)*].

EB simplex je způsobena cytolýzou bazálních keratinocytů. Na základě závažnosti klinických příznaků jsou rozlišovány tři základní autosomálně dominantní subtypy EBS. U lokalizované EBS se puchýře často nachází jen na dlaních a ploskách nohou. U vzácné generalizované EBS se puchýře vyskytují po celém těle, ale průběh onemocnění je mírnější

než u nejzávažnější formy EBS Dowling-Meara. Všechny subtypy EBS jsou způsobeny mutacemi v genu kódujícím keratin 5 (*KRT5*) nebo keratin 14 (*KRT14*) [49],[50],[51].

Keratin 5 a keratin 14 jsou strukturní proteiny vytvářející keratinová intermediární filamenta [49],[52], která poskytují buňce mechanickou pevnost. Na základě fyzikálních a chemických vlastností jsou keratiny rozděleny do dvou skupin: keratiny typu I (KRT9-KRT24) a keratiny typu II (KRT1-KRT8) [53]. Základní strukturní jednotkou intermediárních filament je heterodimer složený z keratinu typu I a keratinu typu II [54] . V bazálních keratinocytech jsou přirozenými partnery pro vznik heterodimeru keratin 5 a keratin 14 stáčející se kolem sebe za vzniku paralelního dimeru, ten vytváří antiparalelní tetramer, následuje vznik protofilament, protofibril a intermediárních filament. Charakteristickým rysem všech keratinů je jejich centrální alfa-helikální tyčinková doména rozdělená do čtyř helikálních segmentů (1A, 1B, 2A, 2B) třemi nonhelikálními úseky a ohraničená N- a C-koncovými globulárními doménami. Na počátku segmentu 1A a na konci segmentu 2B se vyskytují dvě vysoce konzervované oblasti, *helix initiation peptide* a *helix termination peptide*, mutace lokalizované v těchto sekvencích jsou spojeny s nejzávažnější formou EBS-DM. Naproti tomu mutace v nonhelikálních oblastech jsou asociovány spíše s mírnějšími formami EBS [55],[56] . Klinicko-genetické korelace zjištěné u našich 28 pacientů s mutacemi v genu *KRT5* nebo *KRT14* odpovídají těmto předpokladům [publikace *Br J Dermatol. 2010 May;162(5)*].

Publikace

1) Analysis of the COL7A1 gene in Czech patients with dystrophic epidermolysis bullosa reveals novel and recurrent mutations. Jerábková B, Kopecková L, Bucková H, Veselý K, Valícková J, Fajkusová L. J Dermatol Sci. 2010 Aug;59(2):136-40.*(L. Fajkusová jako korespondující autor)*

2) Keratin mutations in patients with epidermolysis bullosa simplex: correlations between phenotype severity and disturbance of intermediate filament molecular structure. Jerábková B, Marek J, Bucková H, Kopecková L, Veselý K, Valícková J, Fajkus J, Fajkusová L. Br J Dermatol. 2010 May;162(5):1004-13.*(L. Fajkusová jako korespondující autor)*

3.2 Ichtyózy

Ichtyózy jsou heterogenní skupinou nemocí charakterizovanou abnormálním šupinatěním kůže. Šest hlavních klinických subtypů je známo u nesyndromatických ichtyóz; počínaje nejzávažnější formou zvanou harlekýnská ichtyóza (HI), přes lamelární ichtyózu (LI), nebulózní kongenitální erytrodermickou ichtyózu (CIE), bulózní kongenitální erytrodermickou ichtyózu (BCIE), X-vázanou ichtyózu (XLI), až po nejmírnější ichtyózu vulgaris (IV).

HI, CIE a LI se společně řadí do skupiny zvané autozomálně recesivní kongenitální ichtyózy (ARCI). Je známo několik genů odpověných za toto onemocnění: *TGM1* (*transglutaminase-1*) [57],[58], *ABCA12* (*ATP-binding binding cassette, subfamily A, member 12*) [59], *NIPAL4* (*NIPA-like domain-containing 4*) [60], *CYP4F22* (*cytochrome P450, family 4, subfamily F, polypeptide 22*) [61], *ALOX12B* (*12-lipoxygenase, R type*), *ALOXE3* (*lipoxygenase-3*) [62]; v poslední době byl tento seznam rozšířen o další nové geny: *PNPLA1* (*patatin-like phospholipase domain-containing protein 1*) [63], *LIPN* (*lipase family, member N*) [64], *CERS3* (*ceramide synthase 3*) [65]. BCEI je puchýřnatá ichtyóza s autosomálně dominantní dědičností a mutacemi v genech *KRT1* (keratin 1), *KRT10* (keratin 10), *KRT2* (keratin 2) [66], [67]. Publikovaná prevalence v evropských populacích a Severní Americe je u ARCI i BCIE 1⁄200000. XLI a IV se s prevalencí 1/2000-6000 resp. 1/250-1000 řadí mezi tzv. běžné ichtyózy. XLI postihuje mužské pohlaví, molekulární příčinou jsou mutace v genu *STS* (steroidní sulfatáza) [68]. IV je autozomálně semidominantní onemocnění, molekulární příčinou jsou mutace v genu *FLG* (filaggrin) [69].

Patologické mechanismy ichtyóz jsou spojeny s poruchou kožní bariéry, tato funkce kůže je lokalizovaná v nejvrchnější vrstvě, tzv. stratum corneum (SC). Keratinocyty jsou dominujícím typem buněk v epidermis, proliferují v bazální vrstvě epidermis (stratum basale) a v průběhu epidermální diferenciace postupují rozdílnými epidermálními vrstvami (stratum spinosum, stratum granulosum) směrem k povrchu (stratum corneum). Keratinocyty lokalizované ve SC nazýváme korneocyty. Prostor mezi jednotlivými korneocyty je vyplněn intercelulárními lipidy (ceramidy, cholesterol, volné mastné kyseliny), které jsou organizovány do lamelárních vrstev [70],[71],[72]. Intercelulární lipidové vrstvy brání nadměrnému odpařování a jejich narušení vede k patologickému

nárustu transepidermální ztráty vody [73]. Funkce kůže jako ochranné bariéry před vnějšími vlivy a pro zabezpečení vhodného vnitřního prostředí je umožněna třemi hlavními komponentami SC. Mezi tyto komponenty se řadí již zmíněné intercelulární lipidové vrstvy, dále pak zrohovatělá buněčná obálka korneocytu a degradační produkty keratinu a filaggrinu vyplňující cytoplazmu korneocytu.

Je velmi zajímavé sledovat vzájemné propojení jednotlivých genů resp. proteinových produktů těchto genů v jejich funkčních drahách, které se podílejí na epidermální diferenciaci. Významnou úlohu při tvorbě intercelulárních lipidových vrstev hrají proteiny ABCA12, ALOX12R, ALOXE3, CYP4F22, NIPAL4, STS, které přímo souvisejí s transportem a/nebo metabolismem lipidů v epidermis. Vytváření intercelulárních lipidových vrstev je vysoce komplikovaná série procesů. Lipidy jsou baleny do lamelárních tělísek ve stratum granulosum a předány exocytózou do mezibuněčného prostoru SC. Důležitou molekulou hrající roli při transportu lipidů v lamelárních tělískách je ATP-vázající kazetový transporter ABCA12 [74],[75],[76],[77]. Lipoxygenázy ALOX12B a ALOXE3 jsou dioxygenázy oxidující komplexní lipidy: ALOX12B oxiduje O-linoleoyl-ω-hydroxyacyl-sphingosin (ceramid 1) na hydroperoxidový derivát, který je následně konvertován prostřednictvím ALOXE3 na 9R,10R-epoxy-11E-13R-hydroxylinoleoyl-ω-hydroxyacyl-sphingosin, následuje hydrolýza na ω-hydroxyacyl-sphingosin (ω-hydroxyceramid). Některé ω-hydroxyceramidy jsou dále hydrolyzovány na ω-hydroxy-mastné kyseliny [78]. Produkty uvedné metabolické dráhy (ω-hydroxyceramid a ω-hydroxy-mastné kyseliny) jsou pak prostřednictvím transglutaminázy-1 vázány k cytoplazmatické vrstvě zrohovatělé buněčné obálky (viz. níže).

Dalšími proteiny hrajícími roli v metobolismu lipidů v epidermis jsou NIPAL4 a CYP4F22. NIPAL4 je pravděpodobně membránový receptor pro další produkty metabolických drah ALOX12B a ALOXE3, pro tzv. hepoxiliny, které hrají roli v mnoha buněčných procesech [78],[79],[80], rovněž se spekuluje o podílu NIPAL4 na fúzi lamelárních tělíšek s plazmatickou membránou [81],[82]. CYP4F22 je součástí široké rodiny cytochromů P450. Tato monooxygenáza pravděpodobně katalyzuje hydroxylaci metabolických produktů hepoxylinu a předpokládá se, že výsledné látky hrají roli v hydrataci kůže.

Transglutamináza-1 je odpovědná za vzájemné provázání (*croslink*) proteinů v cytoplazmatické vrstvě zrohovatělé buněčné obálky a rovněž za tvorbu lipidové obálky, tj.

vazbu ω-hydroxyceramidů a ω-hydroxy-mastných kyselin k proteinům zrohovatělé buněčné obálky [83],[84],[85]. TGM1 se tedy podílí na tvorbě zrohovatělé buněčné obálky korneocytu a lipidové obálky korneocytu, které jsou současně strukturami, na podkladě kterých se formují interbuněčné lipidové vrstvy [86].

Pacienti s mutací v genech *ABCA12, ALOX12R, ALOXE3, CYP4F22, NIPAL4, TGM1* trpí většinou lamelární ichtyózou nebo kongenitální erytrodermickou ichtyózou. Fenotypové projevy LI a CIE se ale mohou vzájemně překrývat a nebyly nalezeny ani silnější korelace mezi genotypem a fenotypem. Postižené děti se často rodí obaleny tzv. koloidní mebránou. Poté co se koloidní membrana sloupne (během prvních týdnů života), projeví se šupinatění kůže, které se u jednotlivých pacientů odlišuje rozsahem, barvou a stupněm přilnavosti. Spektrum klinického obrazu je v rozmezí od jemných, bílých šupin na erytrodermické kůži (CIE) po tmavé, velké šupiny s téměř neviditelnou kožní erytrodermií (LI). Dalšími klinickými projevy mohou být alopecie, hyperkeratóza, hyperlinearita dlaní a plosek nohou, postižení nehtů, hypohydróza, netolerance k teplu. Navíc se fenotypy mohou měnit v průběhu času a v odpovědi na léčbu [87]. Klinicky mohou být rozlišovány další minoritní subtypy ARCI.

X-vázaná ichtyóza je spojena s defekty genu *STS* [68] a s abnormální akumulací cholesterolsulfátu ve SC. Steroidní sulfatáza je koncentrována v lamelárních tělískách a sekretována do intercelulárního prostoru SC spolu s dalšími lipidovými hydrolázami [88]. V intercelulárním prostoru STS degraduje cholesterolsulfát a vytváří cholesterol. Progresivní úbytek cholesterolsulfátu umožňuje degradaci korneodesmosomu a normální odlupování kůže [88]. Deficit STS vede k malformaci interbuněčných lipidových vrstev a omezuje degradaci korneodesmozomu, což má za následek zadržování korneocytů a následnou hyperkeratózu [88].

Keratin1 a keratin 10 jsou strukturní proteiny vytvářející intermediární filamenta epiteliálních buněk. Mutace v genech *KRT1* a *KRT10* způsobují bulózní kongenitální erytrodermickou ichtyózu [89],[90],[66]. BCIE má autosomálně dominantní model dědičnosti. Většina kauzálních mutací je typu *missense* a jsou lokalizovány na začátku nebo na konci alfa-helikální tyčinkové domény, tzv. *helix initiation and helix termination peptide* (viz. epidermolysis bullosa simplex). BCIE je typ ichtyózy, který se vyznačuje vznikem puchýřů na erytrodermickém podkladě. Po perinatálním období, tvorba puchýřů

ustupuje a je patrná generalizovaná hyperkeratóza. Podobný, ale mírnější průběh má BCIE způsobená mutacemi v genu *KRT2* [67],[91], [92].

V granulární vrstvě epidermis se vytvářejí keratohyalinní granula obsahující velký polyprotein profilaggrin (>400 kDa). N-terminální doména profilaggrinu hraje roli během terminální epidermální diferenciace, kdy je odštěpena a transportována do jádra, kde se podílí na enukleaci keratinocytu ve SC. Přesná funkce C-terminální domény její známa, ale zřejmě bude souviset s metabolismem profilaggrinu. Profilaggrin obsahuje 10-12 téměř identických filaggrinových repetic (každá o velikosti 324 aminokyselin). V průběhu terminální diferenciace keratinocytu je profilaggrin v SC štěpen na 10-12 filaggrinových peptidů, které následně agregují s keratinovými filamenty. V důsledku těchto pochodů korneocyt získává zpolštěný tvar. Filaggrin dále degraduje na jednotlivé aminokyseliny. Histidin, aminokyselina vyskytující se ve filaggrinu s vysokou frekvencí, je metabolizován na trans-urokanovou kyselinu a ta společně s dalšími organickými kyselinami pomáhá udržovat kyselé pH vrchní vrstvy kůže, což je důležité pro funkční aktivitu řady enzymů [93],[94]. Gen *FLG* kódující profilaggrin má velikost 23 kb, cDNA tohoto genu je rozložena do 3 exonů. Exon 3 je velmi těžké analyzovat z důvodu jeho velikosti a vysoce repetitivní povahy (obsahuje 10-12 repeticí, velikost jedné repetice je 1.2 kb) [95] . Ichtyóza vulgaris je autozomálně semidominantní onemocnění s nekompletní penetrancí a variabilní expresivitou, klinické projevy jsou velmi různé i v rámci rodiny. Jedinci se dvěma mutacemi v genu *FLG* mají závažnější klinické projevy než s jednou mutací. IV se projevuje v časném dětství, progreduje do puberty, pak je většinou tendence ke zlepšování klinických projevů. Onemocnění je charakterizováno jemnými, mírně adherentními šupinami, bílé nebo šedé barvy, závažnost klinických příznaků je závislá i na prostředí. Často se vyskytuje hyperlinearita dlaní plosek nohou.

V CMBGT jsme začali s molekulárně genetickou diagnostikou ichtyóz v roce 2013 a postupně jsme zavedli analýzu genů *ALOX12B, ALOXE3, TGM1, NIPAL4, CYP4F22* (autosomálně recesivní kongenitální ichtyózy); *STS* (X-vázaná ichtyóza); *FLG* (ichtyóza vulgáris). V současné době máme 41 nepříbuzných pacientů s identifikovanými kauzálními mutacemi: 10 pacientů má mutace v *ALOX12B* (27,7%), 5 pacientů v *ALOXE3* (13,8%), 4 pacienti v *NIPAL4* (11,1%), 4 pacienti v *CYP4F22* (11,1%) a 2 pacienti v *TGM1* (5,6%) genech. Mutace v genu *STS* spojené s XLI byly zjištěny u 6 pacientů (11,1%) a mutace v genu *FLG* spojené s IV u 9 pacientů (19,4%). Vzhledem k tomu, že řada mutací identifikovaných u českých pacientů je nová, v literatuře nepopsaná, provádíme v současné

době podrobnou analýzu korelace genetických, klinických a patologických nálezů. Výsledky získané na základě interdisciplinární spolupráce mezi molekulárními biology, dermatology a patology jsou velmi hodnotné a zatím nebylo publikováno mnoho prací zabývajících se touto problematikou v oblasti ichtyóz. Doposud bylo popsáno 10 rozdílných typů mutací v genu *NIPAL4*. My jsme v souboru našich pacientů našli 5 typů mutací, z toho 4 nové, včetně delecí a duplikací, které ještě v genu *NIPAL4* nebyly u žádné populace identifikovány, podobná situace je i v případě genu *CYP4F22*. Věříme, že výsledky získané u našeho souboru pacientů významně přispějí k rozšíření dosavadních znalostí v této oblasti.

Journal of Dermatological Science 59 (2010) 136–150

Contents lists available at ScienceDirect

Journal of Dermatological Science

journal homepage: www.elsevier.com/jds

Letter to the Editor

Analysis of the *COL7A1* gene in Czech patients with dystrophic epidermolysis bullosa reveals novel and recurrent mutations

Epidermolysis bullosa (EB) is a clinically and genetically heterogeneous group of heritable skin disorders. Fine et al. separated EB into 4 major types – epidermolysis bullosa simplex, junctional epidermolysis bullosa, dystrophic epidermolysis bullosa (DEB), and Kindler syndrome – on the basis of distinguishing ultrastructural sites of blister formation [1]. Inheritance patterns of DEB may be autosomal dominant (DDEB) or autosomal recessive (RDEB). Both DDEB and RDEB result from mutations in the type VII collagen gene (*COL7A1*) [2]. DDEB is usually associated with glycine substitutions within the collagenous domain of type VII collagen. In RDEB, the allelic variants include "silent" glycine substitutions (mutations manifested in a recessive state when inherited with another mutation), non-glycine missense variants, nonsense mutations, splice site mutations, deletions, and insertions [3].

We performed DNA analysis of *COL7A1* in 6 DDEB and 27 RDEB probands at the EB Centre, University Hospital Brno, Czech Republic. From 27 RDEB patients; 17 patients suffered from RDEB-severe generalised (RDEB-sev gen), 3 patients had RDEB-generalised other (RDEB-O), 5 patients had RDEB-inversa (RDEB-I), 1 patient suffered from RDEB-acral (RDEB-ac), and 1 patient had RDEB-pretibial (RDEB-Pt). In case of RDEB patient 12, DNA was unavailable (patient died 10 years ago) and so DNA from his parents was examined. The promoter region (from -449 nucleotide) and 118 exons of the *COL7A1* gene, as well as adjacent intron regions, were amplified and sequenced.

Twenty-nine different sequence variants were found, nine of which have not been reported previously (Table 1). The most common mutation was the transition c.425A>G. This mutation was detected in 10 probands in a heterozygous state, and in 3 probands in a homozygous state (29.6% of mutant alleles). In the study of Csikos et al. [4], *COL7A1* mutations were analysed in 43 unrelated patients with DEB phenotypes from the registries of DEBRA Hungary and DEBRA Germany and the mutation c.425A>G was identified in 10 of them (11 of 86 alleles, 12.8%). The transition c.425A>G at the −2 donor splice site of exon 3 cause aberrant splicing and at least two abnormal transcripts with premature termination codons are generated downstream of this mutation [5]. Other common mutations detected in our set of DEB patients were p.Gly2049Glu (5 RDEB probands), p.Arg2069Cys (5 RDEB probands), and p.Arg1343X (4 RDEB probands).

The novel mutations comprise "silent" glycine substitutions (p.Gly1845Arg, p.Gly2296Glu, and p.Gly2557Arg), splice site mutations (c.3894+1G>A, c.5856+1G>A, and c.6751-2delAG), the deletion c.4556delG, the insertion c.5644insA, and the missense mutation p.Lys1981Arg. The presence of the last-named mutation was analysed in the control group using PCR-RFLP (Restriction Fragment Length Polymorphism). The mutation was not found in any of 200 control alleles (data not shown).

In patients 26 and 27, we detected only one *COL7A1* mutation (on the basis of literature data associated with RDEB-sev gen [5] and RDEB-inversa [6], respectively), the second mutation was not found. It is possible that these patients could have a *COL7A1* rearrangement or a sequence change within the primer binding sites to prevent PCR amplification or a pathogenic mutation within an intron not detected by our sequencing approach. Results of *COL7A1* gene analysis were correlated with clinical, electron-microscopical, and immunohistochemical findings and some of these correlations are shown in Table 1.

A missense mutation of Lys has not been described in DEB association so far. The patient's phenotype associated with p.Lys1981Arg is milder in comparison with patients' phenotypes associated with substitutions of Gly and Arg detected in our DEB patients and corresponds with the subtype RDEB-acral with distinct affliction of fingers (Fig. 1). Type VII collagen is the main constituent of anchoring fibrils where it is present as a homotrimer composed of three identical alpha chains [7]. The central collagenous domain of type VII collagen consists of characteristic Gly-X-Y repeat sequences. The Gly-X-Y repeat is a prerequisite for the formation of the collagen triple helix, which is stabilised by the presence of hydroxyproline and hydroxylysine [8]. The hydroxyl groups of hydroxylysine residues have two important functions: they serve as attachment sites for carbohydrate units and they are crucial for the stability of the intramolecular and intermolecular collagen crosslinks. It is possible that Lys1981, localised in the first Y-position of 38-triplet Gly-X-Y stretch flanked by non-collagenous sequences of 39 and 6 amino acids, has a part in assembly of collagen fibrils, and so mutations in this position will be associated with DEB phenotype.

In patients 16 and 17, we detected heterozygous and homozygous occurrence of the novel mutation c.6751-2delAG, respectively. Patient 17 has Czech citizenship but is a descendant of unrelated Russian parents, patient 16 is descendant of Czech parents. It seems that the mutation c.6751-2delAG could be specific for the Slavonic population. A similar mutation c.6751-2delA was found by Posteraro et al. in an Italian patient. This mutation affected the acceptor splice site of intron 85 and led to aberrant splicing of exon 86 [9].

In summary, this study represents a high mutation detection rate in the *COL7A1* gene in Czech DEB families. Besides mutations detected also in other countries, we described mutations specific for our patients. In the set of our new mutations, the mutation p.Lys1981Arg seems to be most interesting. The reason is that (i) a missense mutation of Lys was not described in DEB patients so far and (ii) the phenotype associated with p.Lys1981Arg is milder in comparison with patient phenotypes associated with Gly and Arg substitutions detected in our DEB patients.

Table 1
Mutation data and phenotypes of Czech RDEB and DDEB patients.

No	Subtype of EB	Age	Mutation on cDNA level	Mutation on protein level	Mutation on cDNA level	Mutation on protein level	Affliction of skin, hair, nails; pseudosyndactyly	Other afflictions
1	RDEB-sev gen	18	c.425A>G	Splice site	c.425A>G	Splice site	Aplasia cutis, generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, defluvium, loss of nails	Microstomia, ankyloglossia, oral cavity erosions, corneal erosions
2	RDEB-sev gen	41	c.425A>G	Splice site	c.425A>G	Splice site	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, defluvium, loss of nails, spinocellular carcinoma	Microstomia, ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosions
3	RDEB-sev gen	7	c.425A>G	Splice site	c.425A>G	Splice site	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, pruritus, defluvium	Microstomia, ankyloglossia, oral cavity erosions, oesophageal stenosis
4	RDEB-sev gen	4	c.425A>G	Splice site	c.682+1G>A	Splice site	Aplasia cutis, generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, pruritus, loss of nails	Ankyloglossia, oral cavity erosions, dysphagia
5	RDEB-sev gen	25	c.425A>G	Splice site	c.3551-3T>G	Splice site	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, pruritus, defluvium, loss of nails, spinocellular carcinoma	Microstomia, oral cavity erosions, oesophageal stenosis, corneal erosions
6	RDEB-sev gen	5	c.425A>G	Splice site	c.4027C>T	p.Arg1343X	Generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, defluvium, loss of nails	Ankyloglossia, oral cavity erosions
7	RDEB-sev gen	14	c.425A>G	Splice site	c.6146G>A	p.Gly2049Glu	Aplasia cutis, generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, defluvium, loss of nails	Ankyloglossia, oral cavity erosions, oesophageal stenosis
8	RDEB-sev gen	16	c.425A>G	Splice site	c.6146G>A	p.Gly2049Glu	Generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, pruritus, onychodystrophy	Microstomia, oral cavity erosions, oesophageal stenosis, corneal erosions
9	RDEB-sev gen	12	c.425A>G	Splice site	c.6187C>T	p.Arg2063Trp	Aplasia cutis, generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, pruritus, defluvium, loss of nails	Microstomia, ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosions
10	RDEB-sev gen	14	c.682+1G>A	Splice site	c.1826C>G	p.Ser609X	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, defluvium, loss of nails	Microstomia, ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosions
11	RDEB-sev gen	21	c.2638del25	PTC	c.4027C>T	p.Arg1343X	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, defluvium, loss of nails	Microstomia, ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosions
12	RDEB-sev gen	22	c.4027C>T	p.Arg1343X	**c.7669G>A**	**p.Gly2557Arg**	Aplasia cutis, generalised blistering, pseudosyndactyly, loss of nails	Extracutaneous involvement, corneal dystrophy
13	RDEB-sev gen	18	c.6081insC	PTC	**c.4556delG**	**PTC**	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, defluvium, loss of nails	Ankyloglossia, oral cavity erosions, corneal erosion, dysphagia
14	RDEB-sev gen	21	c.6146G>A	p.Gly2049Glu	**c.5644insA**	**PTC**	Aplasia cutis, generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, pruritus, defluvium, loss of nails	Microstomia, ankyloglossia, oral cavity erosions, oesophageal stenosis
15	RDEB-sev gen	34	c.6146G>A	p.Gly2049Glu	**c.5856+1G>A**	**Splice site**	Generalised blistering, atrophic scarring, skin contractures, pruritus, defluvium, loss of nails, spinocellular carcinoma	Ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosion
16	RDEB-sev gen	34	c.6146G>A	p.Gly2049Glu	**c.6751-2delAG**	**Splice site**	Generalised blistering, atrophic scarring, skin contractures, pseudosyndactyly, loss of nails	Ankyloglossia, oral cavity erosions, oesophageal stenosis
17	RDEB-sev gen	6	**c.6751-2delAG**	**Splice site**	**c.6751-2delAG**	**Splice site**	Aplasia cutis, generalised blistering, atrophic scarring, skin contractures, partial pseudosyndactyly, pruritus, loss of nails	Ankyloglossia, oral cavity erosions
18	RDEB-I	9	c.6205C>T	p.Arg2069Cys	c.425A>G	Splice site	Predominant blistering in intertriginous, lumbosacral, and axial distribution; atrophic scarring	Microstomia, oral cavity erosions, dysphagia
19	RDEB-I	2	c.6205C>T	p.Arg2069Cys	c.425A>G	Splice site	Aplasia cutis, mild blistering in intertriginous, lumbosacral, and axial distribution; atrophic scarring	Oral cavity erosions

Table 1 (*Continued*)

No	Subtype of EB	Age	Mutation on cDNA level	Mutation on protein level	Mutation on cDNA level	Mutation on protein level	Affliction of skin, hair, nails; pseudosyndactyly	Other afflictions
20	RDEB-I	57	c.6205C>T	p.Arg2069Cys	**c.3894+1G>A**	**Splice site**	Predominant blistering in intertriginous, lumbosacral, and axial distribution; atrophic scarring; defluvium, onychodystrophy; basocellular carcinoma	Ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosions
21	RDEB-I	29	c.6205C>T	p.Arg2069Cys	c.4018C>T	p.Arg1340X	Aplasia cutis, predominant blistering in axial distribution, atrophic scarring, partial pseudosyndactyly	Microstomia, ankyloglossia, oral cavity erosions, dysphagia
22	RDEB-I	27	c.6205C>T	p.Arg2069Cys	c.4027C>T	p.Arg1343X	Predominant blistering in intertriginous, lumbosacral, and axial distribution; atrophic scarring; defluvium, loss of nails	Oral cavity erosions, oesophageal stenosis
23	RDEB-O	25	c.425A>G	Splice site	**c.5533G>A**	**p.Gly1845Arg**	Mild generalised blistering, atrophic scarring, onychodystrophy	Ankyloglossia, oral cavity erosions, oesophageal stenosis, corneal erosions
24	RDEB-ac	36	c.497insA	PTC	**c.5942A>G**	**p.Lys1981Arg**	Mild generalised blistering predominant in acral and knee distribution; atrophic scarring; partial pseudosyndactyly; defluvium; loss of nails	Oral cavity erosions, dysphagia
25	RDEB-O	15	c.1573C>T	p.Arg525X	**c.6887G>A**	**p.Gly2296Glu**	Aplasia cutis, generalised blistering, atrophic scarring, pruritus, onychodystrophy	Microstomia, ankyloglossia, oral cavity erosions, dysphagia, corneal erosions
26	RDEB-O	6	c.425A>G	Splice site	Not found		Generalised blistering, atrophic scarring	Not detected
27	RDEB-Pt	17	c.7864C>T	p.Arg2622Trp	Not found		Mild localised blistering, atrophic scarring, onychdystrophy	Not detected
28	DDEB	7	c.6007G>A	p.Gly2003Arg			Aplasia cutis	Not detected
29	DDEB-gen	2	c.6119G>A	p.Gly2040Asp			Mild generalised blistering, atrophic scarring, onychodystrophy	Oral cavity erosions
30	DDEB-gen	30	c.6127G>C	p.Gly2043Arg			Mild generalised blistering, atrophic scarring	Not detected
31	DDEB	1	c.6190G>A	p.Gly2064Arg			Localised mild blistering	Oral cavity erosions
32	DDEB	1	c.6208G>A	p.Gly2070Arg			Aplasia cutis, localised mild blistering, onychodystrophy	Not detected
33	DDEB-gen	21	c.6227G>A	p.Gly2076Asp			Aplasia cutis, mild generalised blistering, onychodystrophy, loss of nails	Not detected

Note: Mutations depicted by bold letters are novel, not described so far; No: number of patient; Subtype of EB: the subtype of EB determined in the patient on the basis of clinical findings; Age: age at phenotype assessment.

Fig. 1. The clinical photographs of patient 24 (detected mutations: p.Lys1981Arg/c.497insA). (A and B) Atrophic skin in acral areas and knees. (C) Loss of toe-nails, hypopigmentation, small hemorrhagic crusts. (D and E) Loss of finger-nails, semiflectional position of fingers, partial pseudosyndactyly (the right hand is more afflicted), sporadically small erosions and crusts. The photographs document that the novel p.Lys1981Arg mutation is associated with a milder phenotype (in comparison with patient phenotypes associated with Gly and Arg substitutions).

Acknowledgements

This work was funded by the Internal Grant Agency of the Czech Ministry of Health (NR9346-3) and the Czech Ministry of Education (LC06023 and MSM0021622415).

References

[1] Fine JD, Eady RA, Bauer EA, Bauer JW, Bruckner-Tuderman L, Heagerty A, et al. The classification of inherited epidermolysis bullosa (EB): report of the Third International Consensus Meeting on Diagnosis and Classification of EB. J Am Acad Dermatol 2008;58(6):931–50.
[2] Christiano AM, Greenspan DS, Hoffman GG, Zhang X, Tamai Y, Lin AN, et al. A missense mutation in type VII collagen in two affected siblings with recessive dystrophic epidermolysis bullosa. Nat Genet 1993;4(1):62–6.
[3] Dang N, Murrell DF. Mutation analysis and characterization of COL7A1 mutations in dystrophic epidermolysis bullosa. Exp Dermatol 2008;17(7):553–68.
[4] Csikos M, Szocs HI, Laszik A, Mecklenbeck S, Horvath A, Karpati S, et al. High frequency of the 425A–>G splice-site mutation and novel mutations of the COL7A1 gene in central Europe: significance for future mutation detection strategies in dystrophic epidermolysis bullosa. Br J Dermatol 2005;152(5):879–86.
[5] Gardella R, Belletti L, Zoppi N, Marini D, Barlati S, Colombi M. Identification of two splicing mutations in the collagen type VII gene (COL7A1) of a patient affected by the localisata variant of recessive dystrophic epidermolysis bullosa. Am J Hum Genet 1996;59(2):292–300.
[6] Gardella R, Castiglia D, Posteraro P, Bernardini S, Zoppi N, Paradisi M, et al. Genotype–phenotype correlation in italian patients with dystrophic epidermolysis bullosa. J Invest Dermatol 2002;119(6):1456–62.
[7] Burgeson RE. Type VII collagen, anchoring fibrils, and epidermolysis bullosa. J Invest Dermatol 1993;101(3):252–5.
[8] Kivirikko KI. Collagens and their abnormalities in a wide spectrum of diseases. Ann Med 1993;25(2):113–26.
[9] Posteraro P, Pascucci M, Colombi M, Barlati S, Giannetti A, Paradisi M, et al. Denaturing HPLC-based approach for detection of COL7A1 gene mutations causing dystrophic epidermolysis bullosa. Biochem Biophys Res Commun 2005;338(3):1391–401.

Barbora Jeřábková[a,b,1]
[a]*Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Brno, Czech Republic*
[b]*Division of Functional Genomics and Proteomics, Dept. Exp. Biology, Faculty of Science, Masaryk University, Brno, Czech Republic*

Lenka Kopečková[1]
Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Brno, Czech Republic

Hana Bučková
Pediatric Clinic, University Hospital Brno, Masaryk University,Brno, Czech Republic

Karel Veselý
Institute of Pathologic Anatomy, St Anne's University Hospital, Brno, Czech Republic

BJD
British Journal of Dermatology

Keratin mutations in patients with epidermolysis bullosa simplex: correlations between phenotype severity and disturbance of intermediate filament molecular structure

B. Jeřábková,*† J. Marek,† H. Bučková,‡ L. Kopečková,* K. Veselý,§ J. Valíčková,‡ J. Fajkus†¶ and L. Fajkusová*†

*Centre of Molecular Biology and Gene Therapy and ‡Department of Pediatric Dermatology of 1st Pediatric Clinic, Faculty of Medicine, Masaryk University, University Hospital Brno, Černopolní 9, CZ-62500 Brno, Czech Republic
†Faculty of Science, Institute of Experimental Biology, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic
§Institute of Pathologic Anatomy, St Anne's University Hospital, Pekařská 53, CZ-65691 Brno, Czech Republic
¶Institute of Biophysics, Královopolská 135, CZ-61265 Brno, Czech Republic

Correspondence
Lenka Fajkusová.
E-mail: lenkafajkusova@volny.cz

Accepted for publication
11 December 2009

Key words
epidermolysis bullosa, intermediate filaments, keratin mutations, molecular dynamics, protein structure

Conflicts of interest
None declared.

DOI 10.1111/j.1365-2133.2009.09626.x

Summary

Background Epidermolysis bullosa simplex (EBS) is an inherited skin disorder caused by mutations in the keratin 5 (KRT5) and keratin 14 (KRT14) genes, with fragility of basal keratinocytes leading to epidermal cytolysis and blistering.
Objectives In this study, we characterized mutations in KRT5 and KRT14 genes in patients with EBS and investigated their possible structure–function correlations.
Materials and methods Mutations were characterized using polymerase chain reaction (PCR) and DNA sequencing. Further, to explore possible correlations with function, the structural effects of the mutations in segment 2B of KRT5 and KRT14 and associated with EBS in our patients, as well as those reported previously, were modelled by molecular dynamics with the aid of the known crystal structure of the analogous segment of human vimentin.
Results We have identified mutations in the KRT5 and KRT14 genes in 16 of 23 families affected by EBS in the Czech Republic. Eleven different sequence variants were found, of which four have not been reported previously. Novel mutations were found in two patients with the EBS-Dowling–Meara variant (EBS-DM) [KRT14-p.Ser128Pro and KRT14-p.Gln374_Leu387dup(14)] and in three patients with localized EBS (KRT14-p.Leu136Pro and KRT5-p.Val143Ala). Molecular dynamics studies show that the mutations p.Glu411del and p.Ile467Thr perturb the secondary alpha-helical structure of the mutated polypeptide chain, the deletion p.Glu411del in KRT14 has a strong but only local influence on the secondary structure of KRT14, and the structural impact of the mutation p.Ile467Thr in KRT5 is spread along the helix to the C-terminus. In all the other point mutations studied, the direct structural impact was significantly weaker and did not destroy the alpha-helical pattern of the secondary protein structure. The changes of 3-D structure of the KRT5/KRT14 dimer induced by the steric structural impact of the single point mutations, and the resulting altered inter- and intramolecular contacts, are spread along the protein helices to the protein C-terminus, but the overall alpha-helical character of the secondary structure is not destroyed and the atomic displacements induced by mutations cause only limited-scale changes of the quaternary structure of the dimer.
Conclusions The results of molecular modelling show relationships between patients' phenotypes and the structural effects of individual mutations.

Epidermolysis bullosa (EB) is a rare genetic disease characterized by extremely fragile skin and recurrent skin and mucous membrane blistering after mild or nonmechanical trauma. Traditionally, EB is classified in three major categories according to the level of dermoepidermal separation in the basement membrane zone: (i) epidermolysis bullosa simplex (EBS); (ii) junctional epidermolysis bullosa (JEB); and (iii) dystrophic epidermolysis bullosa (DEB). EBS is characterized by separation of the skin above the basement membrane due to cytolysis of basal keratinocytes, and has now been subdivided into basal and suprabasal types.[1] Three main autosomal dominant subtypes of basal EBS are distinguished based on the severity of blistering. In EBS localized (EBS-loc), previously called the EBS Weber–Cockayne variant (EBS-WC), blistering is usually limited to the hands and feet. In rare EBS, another generalized form [EBS-gen non-Dowling–Meara (DM)], previously classified as the EBS Koebner variant (EBS-K), more widespread blistering is observed but it is usually milder than in the most severe variant EBS-DM. EBS-DM is characterized by generalized herpetiform blistering especially in the neonatal and infant periods. These subtypes of EBS are caused by mutations in either the keratin 5 (*KRT5*) or the keratin 14 (*KRT14*) genes. Most keratin mutations are inherited in an autosomal dominant manner; however, rare recessive keratin 14 mutations have been reported, too.[2–7]

Keratins are a group of structural proteins that polymerize to form keratin intermediate filaments. They are divided into type I (KRT9–KRT24) or type II (KRT1–KRT8) proteins according to their physical and chemical properties,[8] and the basic structural unit of intermediate filaments is a heterodimer of one type I keratin and its corresponding type II partner.[9–11] A characteristic feature of all intermediate filament proteins is the central alpha-helical rod domain, which is divided into four helices (1A, 1B, 2A and 2B) by three short nonhelical linker domains and terminated by globular head domains. The sequence conservation across different intermediate filament proteins is particularly high within the head and tail helices. The first conserved region spans 26 residues corresponding to a part of the 1A segment, and the second is situated at the end of the 2B segment and spans 32 residues. Both of these regions are involved critically in the head to tail overlap that occurs between consecutive heterodimers.[12,13] Keratins 5 and 14 are natural partners that dimerize by coiled-coil interactions, and EBS-DM is mainly due to mutations located within the two hotspot regions that encode the helix initiation and termination peptide sequences in the segments 1A and 2B, respectively. The effects of mutations p.Glu477Lys (in the helix termination peptide of keratin 5), p.Arg125His and p.Arg125Cys (both in the helix initiation peptide of keratin 14) confirm the significance of the corresponding nucleotides as mutational hotspots in EBS-DM.[7,14] To model the consequences for heterodimer structure and function of KRT5 and KRT14 mutations associated with EBS, Smith *et al.*[15] used as a starting structure the previous model of Liovic *et al.*[16] together with crystallographic data for the analogous segment 1A of human vimentin,[17] replacing a number of residues, one at a time, with mutations observed in the corresponding segment of the keratin proteins.

In this study, we performed DNA analysis of the *KRT5* and *KRT14* genes in 23 EBS probands and their family members. A causative mutation was detected in 16 probands, while no mutation was found in seven probands showing *de novo* events with localized blistering. Further, we modelled the structural effects of mutations in segment 2B of KRT5 and KRT14, which are associated with EBS, using the crystal structure of the corresponding fragment of human vimentin and molecular dynamics simulations to reveal correlations between phenotypes and structural effects.

Materials and methods

Patients

We diagnosed 23 families as having EBS at the EB Centre, University Hospital Brno, Czech Republic; an unambiguous EBS-DM phenotype was observed in 11 individuals and 12 showed predominantly localized blisters on the hands and feet. All studies have been approved by the Ethical Board of the University Hospital Brno, and all patients and their family members gave their written, informed consent for DNA analysis.

Ultrastructural analysis

EBS cases were diagnosed clinically and confirmed by transmission electron microscopy examination and/or by indirect immunofluorescence mapping of protein components of the dermoepidermal junction zone from the leading edge of a fresh blister using techniques described previously.[18–20] Specifically, a panel of nine antibodies recognizing cytokeratin 5 (clone XM26; Novocastra, Newcastle Upon Tyne, U.K.), cytokeratin 14 (clone LL002; Novocastra), integrin α6 (clone NKI-GoH3; Chemicon International Inc., Billerica, MA, U.S.A.), integrin β4 (clone 3E1; Chemicon International Inc.), collagen type IV (clone PHM12; Novocastra), collagen type VII (clone LH7.2; Novocastra), laminin 5 (clone GB3; Abcam plc, Cambridge, U.K.), BPAG1 (polyclonal; Santa Cruz Biotechnology Inc., Santa Cruz, CA, U.S.A.), and collagen type XVII (polyconal; Santa Cruz Biotechnology Inc.) was used for the classification and subclassification of EB cases.

Analysis of mutations in the *KRT5* and *KRT14* genes

Genomic DNA was isolated from peripheral blood by the salting-out method. Nine exons of the *KRT5* gene and eight exons of the *KRT14* gene as well as adjacent intron regions were amplified using primers listed in Table 1. Long-range polymerase chain reaction (PCR) of the *KRT14* gene was performed prior to amplification of single exons to avoid amplification of the highly homologous *KRT14* pseudogene, using primers and cycling conditions as described by Wood *et al.*[21] This

Table 1 Primers and annealing temperatures for amplification of the KRT5 and KRT14 coding regions

Gene	Exon	Forward primer	Reverse primer	Annealing temperature (°C)
KRT5	1	tgggtaacagagccaccttc	gcctcctctagtgctgatgg	60
	1	cttgtggagtgggtggctat	ggcatttatttcagacccaca	60
	2	tgggaggcaccttagtgagt	ccaaaggaaggcatggtagt	58
	3	ttcccactgcaaaagtaggc	aaacagggatgattggcact	60
	4	agcctgcagctatgctctct	ccattcaaatctttcactcattg	60
	5	caatgagtgaaagatttgaatgg	gtcgtcatggcctagaatcc	60
	6	tcactgcctgtgaactttgg	tctagctgcgtgtgtttagca	60
	7	agcccctcagaaggagacac	accccacagtagagcagctt	60
	8	cgaatcatgaggatgggagt	gaggaaacactgcttgtga	60
	9	ctccaacaccagcagatcaa	gcttgcactgaagccagag	60
	9	ctaggtggtgggctcagtgt	tctgcaattggcttggtcta	65
KRT14	1	ccaaggggaatggaaagtg	aaggccaccaccatatcct	60
	1	ggcctgtctgtctcatcctc	ggcatgaattgttcccagaa	60
	2	aagttttgtgggggaaaagg	cagggagttttcatgcaccta	65
	3	gcactgtgttcaaccacgcc	tcctgtctcagcctcccaag	60
	4	gatggtgggagggcagtatt	atgccattcacaccagaagg	60
	5	aaggtgggtgtcatttgagg	agtgagtgtggccgttctct	60
	6	agagaacggccacactcact	aaggacaggatcattagatacatgg	60
	7	agtcttggccctccccttag	cactagagctcagcccctca	60
	8	tcaccttcttggcctcctta	gtgcatgacaagctccaaag	60

amplification was performed using the Expand Long Template PCR System (Roche Diagnostics GmbH, Mannheim, Germany) and contained 250 ng DNA, 1 × PCR buffer, 2·5 U Long Template enzyme mix, 2 mmol L^{-1} $MgCl_2$, 0·25 mmol L^{-1} dNTP and 0·3 μmol L^{-1} primers in 50-μL reactions. PCR reactions for individual exons contained 250 ng of DNA, 1× PCR GeneAmp Gold Buffer, 1·25 U AmpliTaq Gold polymerase (Applied Biosystems, Warrington, U.K.), 2 mmol L^{-1} $MgCl_2$, 0·2 mmol L^{-1} dNTP and 0·5 μmol L^{-1} primers. Amplifications were performed under the following cycling conditions: 95 °C for 7 min, followed by 35 cycles of 95 °C for 30 s, annealing temperature as given in Table 1 for 30 s, and 72 °C for 30 s. The PCR products were purified and sequenced on an ABI PRISM 310 DNA-sequencer (Applied Biosystems).

Molecular modelling of the structural effects of *KRT5* and *KRT14* mutations

The crystal structure of human vimentin fragments (PDB entries 1gk4, chains A and B)[17] was used as the input structure for each of our mutated heterodimers. The starting structures of keratin 5/14 dimers were prepared using the SWISS-MODEL server (http://swissmodel.expasy.org)[22] and with 1gk4 chains A and B as templates for keratin 5 (amino acid residues 399–476 and 14 (amino acid residues 344–421). Molecular dynamics simulations were carried out using the SANDER and PMEMD modules of the AMBER 9 software package[23] with the parm99 force-field.[24] All of the simulated structures were built up using the protocol of Mishra *et al.*,[25] slightly modified as follows: all hydrogens were added using the Xleap program from the Amber package; the system was solvated with TIP3P water molecules added as WAT3PBOX cubes to form a truncated octahedron water box; and to maintain overall electrostatic neutrality and isotonic conditions during simulation, Na^+ and Cl^- ions were added to the system. For details about the number of water molecules and ions, see Table 2.

In the next stage of input preparation, a stepwise equilibration of the systems was performed. First, the positions of water molecules and Cl^- and Na^+ ions were energy minimized. The equilibration of the solvent proceeded on a 30-ps-long simulated heating from 5 to 298 K, a 100-ps-long molecular dynamics simulation of ions and water molecules at 298 K, and a 30-ps-long simulated cooling of ions and waters back to 5 K. In the next steps, the equilibration continued with several sequential energy minimizations of keratin 5/14 atoms with gradually decreasing protein force constants. After this relaxation, the complete system (protein atoms, ions and water molecules) with no restraints was heated in three

Table 2 Solvation of polypeptides with individual mutations

Name of polypeptide	Number of ions Na^+/Cl^-	Number of water molecules
Vimentin (1gk4.A + 1gk4.B)	106/90	70 157
KRT5/KRT14	95/87	65 309
KRT5-p.Leu463Pro/KRT14	98/91	66 114
KRT5-p.Ile467Leu/KRT14	94/87	64 250
KRT5-p.Ile467Thr/KRT14	94/87	66 846
KRT5-p.Thr469Pro/KRT14	96/89	66 821
KRT5/KRT14-p.Leu408Met	95/88	67 278
KRT5/KRT14-p.Glu411del	94/88	66 417
KRT5/KRT14-p.Ala413Thr	96/89	68 379
KRT5/KRT14-p.Tyr415Cys	94/87	65 961
KRT5/KRT14-p.Tyr415His	95/88	66 682

50-ps-long molecular dynamics steps from 5 to 150 K, from 150 to 250 K and from 250 to 298 K, respectively. The final step of the preparation stage of molecular dynamics model was a 50-ps-long equilibration of the whole system at 298 K. All molecular dynamics simulations were run at a constant pressure of 1 atm.

The production phases of molecular dynamics calculations were run at a temperature of 298·15 K and a pressure of 1 atm. Two fs time integration steps were used, together with particle-mesh Ewald methods[26] for electrostatic interactions as implemented in the PMEMD module of Amber, with the cut-off of 9 Å for nonbonding interactions. Coordinates were stored every 2 ps, and all PMEMD calculations with one keratin 5/14 mutant were performed for at least 1000 ps. Analyses of the molecular dynamics simulations were carried out using the PTRAJ module of Amber. Molecular graphics were generated using MOLSCRIPT.[27]

Results

KRT5 and *KRT14* mutations in patients with epidermolysis bullosa simplex

Mutations in keratin genes were examined in 16 of 23 families with EBS. Eleven different sequence variants were found, four of which have not been reported previously. Twelve mutations were found in the KRT14 gene and four in the KRT5 gene (Table 3[14,28–32]), most of which occur in highly conserved domains 1A and 2B, except for KRT5-p.Val143Ala which is in the V1 head domain. In 11 patients, we detected mutations that have been reported previously. KRT14-p.Met119Thr was detected in EBS-DM patient 1 as a *de novo* event. KRT14-p.Arg125Cys was identified in EBS-DM patients 2–5, and in patient 2's family (affected brother, mother and mother's sister) and patient 5's family (affected father); patient 3's parents were unaffected and did not carry this mutation and patient 4's affected father was unavailable for analysis. KRT14-p.Arg125His was determined in EBS-DM patients 6 and 7; in patient 6 it was probably inherited (the affected mother was not available for examination) but in patient 7 it was a *de novo* event. KRT14-p.Glu411del was identified in patient 11 with localized EBS, and in his affected father and father's mother. KRT14-p.Tyr415His was detected in EBS-DM patient 12 as a *de novo* event, KRT5-pGlu170Lys in patient 15 (localized EBS) and her affected father, and KRT5-p.Glu477Lys in EBS-DM patient 16 but not in either healthy parent.

We found novel mutations in five patients, two with EBS-DM [KRT14-p.Ser128Pro and KRT14-p.Gln374_Leu387dup(14)], and three with localized EBS (KRT14-p.Leu136Pro and KRT5-p.Val143Ala). KRT14-p.Ser128Pro (in patient 8 and her affected father) and KRT14-p.Gln374_Leu387dup(14) (in patient 10 and his affected sister, the affected father was unavailable for

Table 3 Mutations and clinical phenotypes found in patients with epidermolysis bullosa simplex (EBS)

Patient	Mutation at cDNA	Mutation at protein	Domain of protein	Keratin	EBS subtype	Blister site	EM/AM	Family history	References
1	c.356T>C	p.Met119Thr	1A, HIP	14	DM	Generalized	−/−	De novo	28
2	c.373C>T	p.Arg125Cys	1A, HIP	14	DM	Generalized	+/−	Familial	29
3	c.373C>T	p.Arg125Cys	1A, HIP	14	DM	Generalized	+/+	De novo	29
4	c.373C>T	p.Arg125Cys	1A, HIP	14	DM	Generalized	+/−	Familial[a]	29
5	c.373C>T	p.Arg125Cys	1A, HIP	14	DM	Generalized	+/+	Familial	29
6	c.374G>A	p.Arg125His	1A, HIP	14	DM	Generalized	+/−	Familial[a]	29
7	c.374G>A	p.Arg125His	1A, HIP	14	DM	Generalized	+/+	De novo	29
8	**c.382T>C**	**p.Ser128Pro**	**1A, HIP**	**14**	**DM**	**Generalized**	**+/−**	**Familial**	**NM**
9	**c.407T>C**	**p.Leu136Pro**	**1A, HIP**	**14**	**Localized**	**Multiple scars**	**−/−**	**Familial**	**NM**
10	**c.1120_1161dup(42)**	**p.Gln374_Leu387dup(14)**	**1A, HIP**	**14**	**DM**	**Generalized**	**+/+**	**Familial**	**NM**
11	c.1231_1233delGAG	p.Glu411del	2B	14	Localized	Palmoplantar	+/+	Familial	30
12	c.1243T>C	p.Tyr415His	2B	14	DM	Generalized	−/−	De novo	31
13	**c.428T>C**	**p.Val143Ala**	**V1 head**	**5**	**Localized**	**Palmoplantar**	**+/−**	**Familial**	**NM**
14	**c.428T>C**	**p.Val143Ala**	**V1 head**	**5**	**Localized**	**Palmoplantar**	**−/−**	**Familial**	**NM**
15	c.508G>A	p.Glu170Lys	1A, HIP	5	Localized	Palmoplantar	−/−	Familial	32
16	c.1429G>A	p.Glu477Lys	2B, HTP	5	DM	Generalized	+/−	De novo	14
17	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	
18	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	
19	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	
20	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	
21	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	
22	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	
23	Not detected	Not detected	–	–	Localized	Palmoplantar	+/−	De novo	

[a]Affected relatives but unavailable for molecular genetic examination. AM, antigen immunohistochemical mapping (+ performed, − not performed); EM, electron microscopy (+ performed, − not performed); HIP, helix initiation peptide; HTP, helix termination peptide; NM, novel mutation; bold type, mutations detected for the first time.

examination) are both associated with a typical EBS DM phenotype of relatively severe blistering with lesions at other body sites besides the hands and feet. KRT14 p.Leu136Pro was found in 30 year old patient 9 (the affected mother was not available for examination) who, at examination, had multiple scars on the dorsum of the hands and atrophic scars in the lumbosacral area. Affected members of family 13 (localized EBS) were found to have the novel mutation KRT5 p.Val143Ala; patient 13 showed blistering localized to the palms and soles in infancy, which was gradually replaced by hyperkeratosis. The same mutation was identified in patient 14 with localized EBS and his affected father. Clinical photographs of patients with EBS carrying these novel mutations are shown in Figure 1 and immunofluorescence and electron microscopy findings in Figure 2. These mutations were not found in any of 200 alleles in an unrelated control group, as evaluated using restriction fragment length polymorphisms (PCR RFLP) with restriction by StuI for KRT14 p.Ser128Pro, *Msp*I for KRT14 p.Leu136Pro and *Hae*III for KRT5 p.Val143Ala, and PCR and agarose electrophoresis for KRT14 p.Gln374 Leu387dup(14) (data not shown).

Besides mutations associated with EBS, we detected missense mutations in our patients that were also found in their healthy parents and/or healthy control samples, and/or were not identified in affected relatives; these were p.Gly138Glu, p.Asp197Glu, p.Arg374Trp, p.Ser528Gly and p.Gly543Ser in the *KRT5* gene and p.Arg41Cys and p.Cys18Ser in the *KRT14* gene. The polymorphisms p.Arg374Trp, p.Arg41Cys and p.Cys18Ser have not been described thus far.

Effects of *KRT5* and *KRT14* mutations on keratin dimer structure

To examine the effects of the above mutations on the three dimensional (3 D) structure of keratin dimers, we constructed models of 77 residue long helices of KRT5/KRT14 based on the close similarity of their amino acid sequences to those of a fragment of human vimentin together with the known crystal structure of this fragment (PDB entries 1gk4, chains A and B). These models were merged with an octahedral box containing approximately 64 000–67 000 water molecules and approximately 100 Na^+ and Cl^- ions to maintain overall electrostatic neutrality and isotonic conditions in the system. To obtain the energy minimized structure, a multistep equilibration of the model systems with molecular dynamics runs of more than 200 ps long in total were performed before the production of the molecular dynamics simulations. Simulation of the relaxed systems was performed under physiological conditions (a constant pressure of 1 atm and a temperature of 298·15 K). For the purpose of comparison, we performed complete multi

Fig 1. (a) Partly haemorrhagic blisters, with erosions on the inflammatory skin on the sole and haemorrhagic crust under the fingers, in a 2-year-old patient with epidermolysis bullosa simplex (EBS) with the mutation KRT14-p.Ser128Pro. (b) Fresh and several older clear blisters on the palm and fingers in the same patient at age 4 years. (c) Partly haemorrhagic blisters in a herpetiform distribution with inflammatory surrounding in cubital localization in a 6-year-old patient with EBS with the mutation KRT14-p.Glu374_Leu387dup(14). (d) Extensive hyperkeratosis in the palm of the same patient. (e) Multiple tiny linear white scars on the dorsum of the hand in a 30-year-old patient with EBS with the KRT14-p.Leu136Pro mutation.

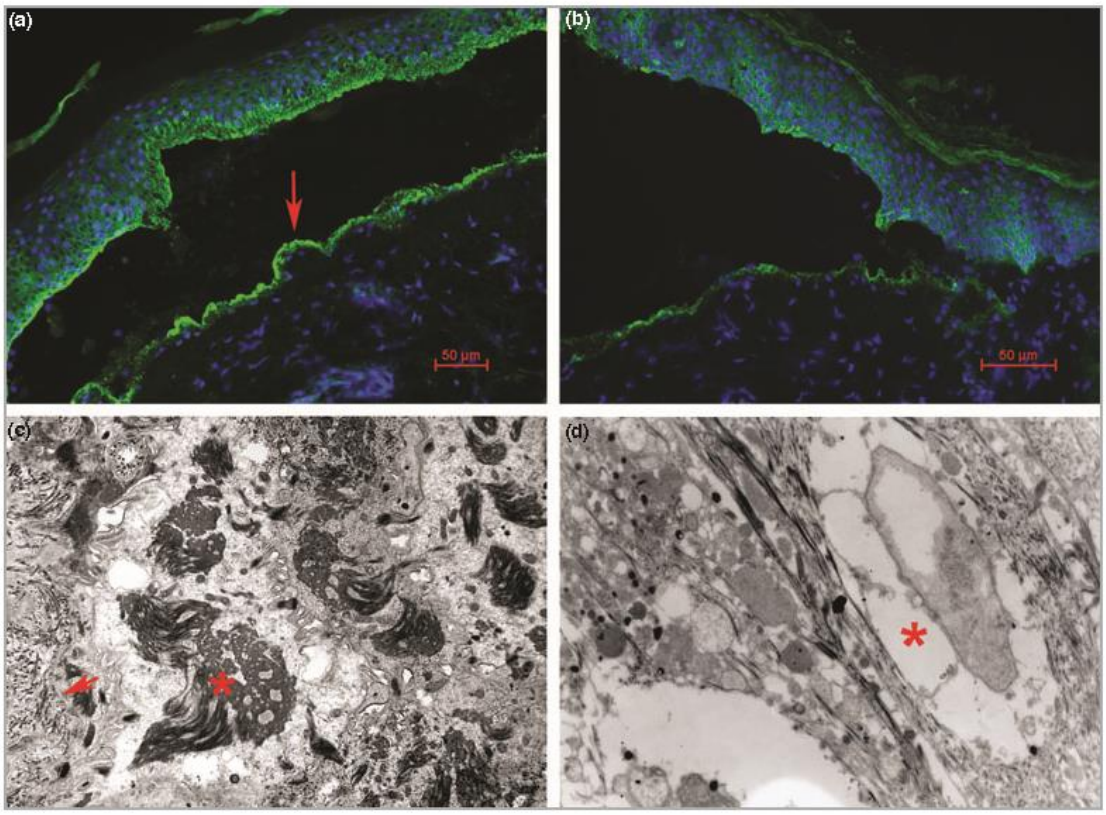


Fig 2. (a) Immunofluorescence microscopy findings in the skin of the patient with epidermolysis bullosa simplex (EBS) with the mutation KRT14-p.Gln374_Leu387dup(14) using monoclonal antibody to keratin 14, with positive staining of basal and suprabasal keratinocytes including the blister floor (arrow). (b) Immunofluorescence labelling of a skin biopsy from the same patient using monoclonal antibody to keratin 5 with a distribution pattern similar to (a). (c) Electron microscopy findings in the skin of the patient with EB with mutation KRT14-p.Ser128Pro. Ultrastructural changes are pronounced clumping of curled tonofilaments (*) in basal and suprabasal keratinocytes above the dermoepidermal junction (arrow) characteristic of EBS-Dowling Meara. Some subtle cytolysis is also present. Original magnification × 6000. (d) Electron microscopy findings in the skin of the patient with EB with mutation KRT14-p.Val143Ala. Ultrastructurally, advanced cytolysis of keratinocytes with vacuolization of cytoplasm (*) is seen. Original magnification × 6000.

step molecular dynamics simulations not only with the mutated keratin 5/14 dimers, but also with the original, non mutated KRT5/KRT14, as well as with the crystallized frag ment of vimentin.

The resulting models of the KRT5/KRT14 mutations obtained by this molecular dynamics approach resemble pri marily the 3 D structure of the corresponding fragment of vi mentin. The two 77 amino acid long peptide chains are folded into two more or less parallel alpha helices with 21 screw threads, and the mutations studied provide a source of disturbances to this reference structure. We performed molec ular modelling of mutations p.Leu463Pro, p.Ile467Leu, p.Ile467Thr and p.Thr469Pro located in KRT5, and mutations p.Leu408Met, p.Glu411del, p.Ala413Thr, p.Tyr415Cys and p.Tyr415His in the case of KRT14. The mutation p.Leu463Pro was described in a patient with the EBS K phenotype,[33] p.Ile467Leu in a patient with the EBS WC phenotype,[34] and p.Ile467Thr[35,36] and p.Thr469Pro[30] in patients with the EBS DM phenotype. The mutation p.Leu408Met,[37] as well as the in frame deletion p.Glu411del,[30] was associated with the EBS WC phenotype. p.Ala413Thr was detected in a patient with EBS K; analysis of 50 healthy control samples showed that this mutation is polymorphic (11% of control samples contained this mutation), despite the fact that it lies in the highly con served helix termination motif.[38] The mutation p.Tyr415Cys was found to result in the EBS WC phenotype[39] whereas p.Tyr415His was shown to cause the more severe EBS K phe notype.[7,31]

Using the classification of KRT5/KRT14 mutations of Smith et al.,[15] we can divide the effects of single mutations into two groups (1 and 5 of Smith). In the first group, we observed in two cases a disturbance of the secondary alpha helical struc ture of the protein chain (Fig. 3). In KRT14, p.Glu411del divides the wild type alpha helix structure into two parts con nected by an almost straight junction (amino acid 410–413), and the deletion p.Glu411del has a strong, but only local,

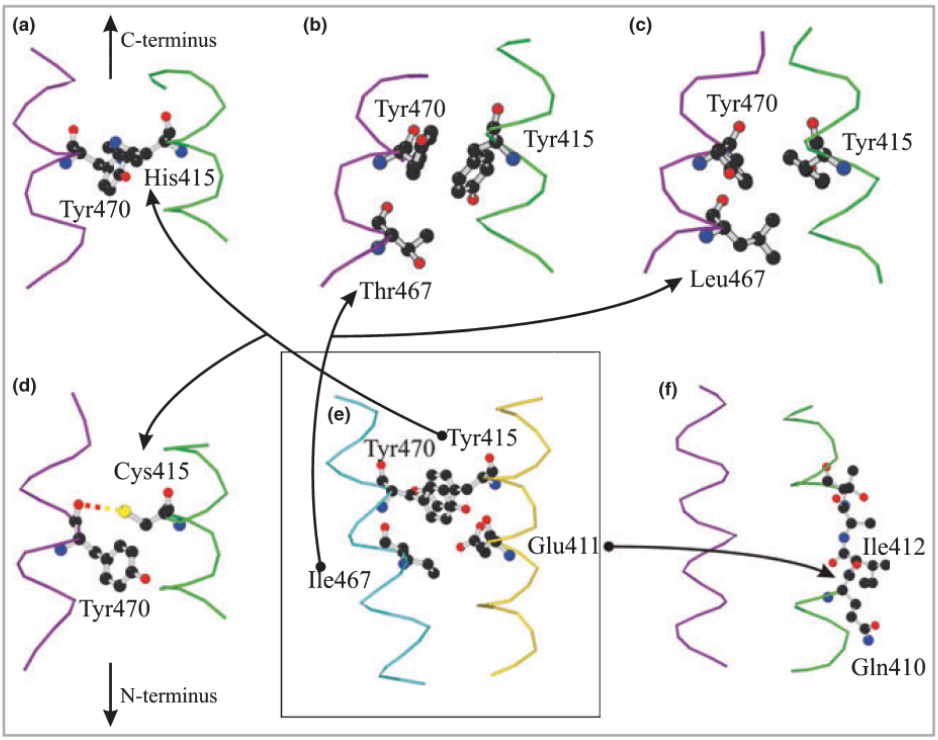


Fig 3. Model structures of KRT5/KRT14 dimers generated by molecular dynamics. The models of mutated KRT5 (violet)/KRT14 (green) dimers (panels a–d and f) are compared with that of wild-type KRT5 (aquamarine)/KRT14 (yellow) dimer (panel e) generated under the same conditions and annealing protocol. The wild-type structure (e) includes residues 455–475 from keratin 5 and residues 400–420 from keratin 14. Mutations in codon 467 of KRT5 and in codon 415 of KRT14 are shown in a more detailed view (panels a–d). Notable amino acid residues are shown by 'ball-and-stick' form. Arrows point to mutated sites. (a) KRT14-p.Tyr415His. Quaternary structure is destabilized by the lack of a half of the Tyr415–Tyr470 pair, which stabilizes the wild-type structure. (d) KRT14-p.Tyr415Cys. The stabilizing effect of the missing Tyr415–Tyr470 pair is partially compensated by the S-H...O hydrogen bond between Cys415 and carbonyl group of Tyr470. (b) KRT5-p.Ile467Thr. The impact of this mutation is spread along the KRT5 helix to the protein C-terminus. Note the shifts of the carbonyls (drawn in red) and the unstacked nonparallel pair of tyrosines Tyr415 and Tyr470. (c) KRT5-p.Ile467Leu. An example of a mutation with very small structural impact (contrary to the previous case). Note the almost parallel stacked residues Tyr415 and Tyr470 in both mutated and wild-type dimers. (f) KRT14-p.Glu411del. A mutation with strong local influence on the secondary structure. The missing residue Glu411 divides the original wild-type alpha-helix structure into two helical parts connected by an almost straight junction.

influence on the secondary structure. In KRT5, the structural impact of p.Ile467Thr is spread along the helix to the protein's C-terminus (see the shifts of carbonyls and the unstacked nonparallel pair of tyrosines Tyr415 and Tyr470 in Figure 3). In all the other point mutations studied, the direct structural impact was significantly weaker and did not destroy the alpha-helical pattern of the secondary protein structure. However, the replacement of one amino acid typically alters the involvement of side chains of other amino acids in inter- and intramolecular contacts, and consequently could cause distortion of the quaternary KRT5/KRT14 structure.

Discussion

Four novel mutations in keratin genes were identified in this study, three in the KRT14 gene and one in the KRT5 gene. The mutation p.Ser128Pro affects the 13th amino acid residue of the KRT14-1A domain, a region termed the helix initiation motif where serine is highly conserved within the family of type I keratins and where most EBS-DM mutations are located. No substitutions of serine residue at this position have been described in type I keratins, but its deletion was observed in KRT16 causing type 1 pachyonychia congenita[40] and in KRT17 causing type 2 pachyonychia congenita.[41] Another mutation localized in the 1A domain of KRT14 was p.Leu136Pro. The Leu136 is highly conserved within the family of type I and type II keratins. A different mutation at this codon (KRT14-p.Leu136Gln) was reported previously in relation to the EBS-WC phenotype.[42] The duplication p.Gln374_Leu387dup(14) is localized in the highly evolutionarily conserved 2B domain of keratin 14 and the mutation p.Val143Ala in the nonhelical V1 head domain of keratin 5. Comparison of the correspond-

ing amino acids at position 143 with other type II keratins shows that the valine residue is highly conserved and present in all 26 type II keratins; the substitution KRT5-p.Val143Asp was described in an EBS patient.[43]

We detected mutations in 16 of 23 unrelated patients suspected of having EBS (70%). Rugg *et al.*[34] identified *KRT5*/*KRT14* mutations in 20 of the 43 families registered as affected by EBS in Scotland, while Pfendner and Uitto[44] found *KRT5*/*KRT14* mutations in 18 of 57 EBS probands analysed in the global EB community in Philadelphia. Careful examination revealed that in 16 of the remaining 39 cases the referral diagnosis of EBS was not correct or was not consistent with electron microscopic or immunofluorescence findings, whereas 14 cases were found to harbour mutations in the plectin gene (*PLEC1*).[45] In the remaining nine cases, the electron microscopic and/or immunohistochemical data supported the diagnosis of EBS, but no mutations in the *KRT5*, *KRT14* or *PLEC1* genes could be identified. Analysis of *PLEC1* was not performed in our patients with EBS without *KRT5*/*KRT14* mutations but their phenotype did not correspond to those associated with *PLEC1* mutations (i.e. EB with muscular dystrophy, EBS of the Ogna type or EB with pyloric atresia). No mutations were found in our patients 17–23, who were clinically diagnosed as having localized EBS but were sporadic cases with no family history of EBS. It is possible that other genes may be associated with this phenotype.

Several authors have constructed 3-D computer models of the KRT5/KRT14 dimer to visualize the structural impact of point mutations. Liovic *et al.*[16] used the crystal structure of a leucine zipper in the transcriptional activator GCN4 of *Saccharomyces cerevisiae*[46] to construct a homology model of the 1A region of the KRT5/KRT14 dimer, and subsequently used this as input for molecular dynamics calculations for the wild-type dimer and two dimers containing pathogenic point mutations. The effects of a significantly larger number of point mutations (22) in the *KRT5* and *KRT14* genes were examined by Smith *et al.*[15] using these structural results[16] as the starting model, and this study showed that correlation between the disruption of the coiled-coil structure and the severity of the EBS is not simple.

Here we performed similar molecular modelling of the structural changes caused by mutations p.Leu463Pro, p.Ile467-Leu, p.Ile467Thr and p.Thr469Pro located in *KRT5* and mutations p.Leu408Met, p.Glu411del, p.Ala413Thr, p.Tyr415Cys and p.Tyr415His in *KRT14*. We discuss particularly the results for five mutations – KRT5-p.Ile467Leu, KRT5-p.Ile467Thr, KRT14-p.Glu411del, KRT14-p.Tyr415Cys and KRT14-p.Tyr415His – for the following reasons: (i) a couple of mutations in codon 467 (p.Ile467Leu and p.Ile467Thr) and in codon 415 (p.Tyr415Cys and p.Tyr415His) are associated with the different clinical symptoms (EBS-WC vs. EBS-DM or EBS-K); and (ii) the mutation p.Glu411del has a supposed strong impact on the protein structure but is connected with only a moderate EBS-WC phenotype, both in the literature and also in our patient 10. The effects of these mutations on protein structure (Fig. 3) demonstrate correlations between their structural and phenotypic consequences. Our molecular dynamics results cannot be compared directly with those already published,[15,16] because the mutations we studied are localized in the segment 2B of the KRT5/KRT14 dimer while the previous modelling was performed for the 1A segment. As a benchmark for our molecular dynamics simulations, we repeated the earlier calculations for point mutations localized at the same position of the 1A domain but associated with different phenotypes, using the same algorithm which was

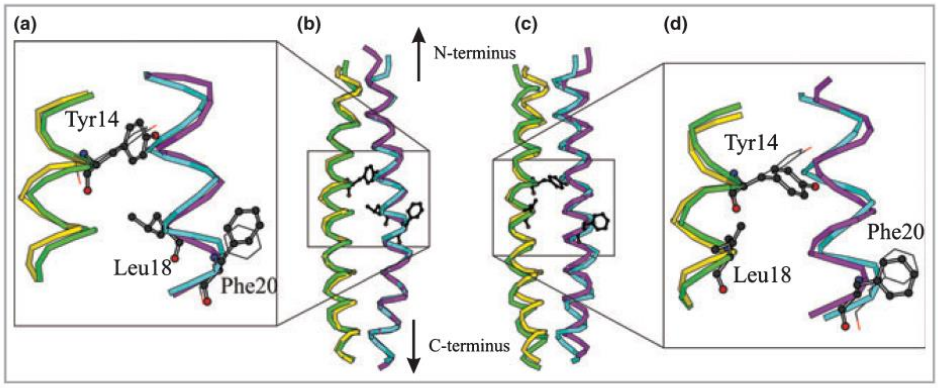


Fig 4. The model structures of KRT5/KRT14 dimers with mutation Val18Leu in the 1A region generated by molecular dynamics. The model structure of mutated KRT5 (violet)/KRT14 (green) dimer is superimposed onto the model of wild-type KRT5 (aquamarine)/KRT14 (yellow) structure generated under the same conditions and annealing protocol. The central panels (b) and (c) contain the whole model (residues 1–35), details on the left and right panels depict only residues 10–21 from keratin 5 (a) and residues 10–20 from keratin 14 (d). Altered residues in the mutated dimers are shown in ball-and-stick form, and the corresponding residues in the nonmutated dimer by thin bonds. (a, b) KRT5-p.Val18Leu. The mutation with a very small impact on the secondary helical structure, but with a significant influence, e.g. on angles of the aromatic side chains Phe20 of KRT5 or Tyr14 in KRT14. (c, d) KRT14-p.Val18Leu. An impact to the helical structure is considerably stronger here, but conformation changes in positions of side chains Phe20 of KRT5 are less notable.

used for multistep equilibration and final molecular dynamics modelling of the segment 2B with only a few obvious changes in input parameters: (i) the different primary sequence of the 1A region of KRT5/KRT14 was employed; (ii) the recently published structure of vimentin coil 1A (PDB entry 3G1E[47]) was used as an input to the SWISS-MODEL homology modelling of the 1A region; and finally (iii) the different range of applicability of 3G1E input model leads to a different number of protein atoms, ions and water molecules in our final model.

In contrast to previously published results,[16] our molecular dynamics calculations show a greater impact on the secondary structure for the mutation KRT14-p.Val133Leu associated with the EBS-WC phenotype (p.Val18Leu in the KRT14 1A segment) than for the mutation KRT5-Val186Leu associated with EBS-K phenotype (p.Val18Leu in the KRT5 1A segment) (Fig. 4). The mean distances in three dimensions between the corresponding C_α atoms in wild type and p.Val18Leu in KRT5 and KRT14 calculated by SUPERPOSE[48] are 0·1 and 0·5 Å, respectively. In contrast, in the tertiary structure of mutated 1A regions, significant changes of the positions of several bulky amino acids can be seen (e.g. Arg19 or Phe20) close to the point of mutation, with the p.Val18Leu mutation causing a more dramatic side-chain displacement in KRT5 than in KRT14 (Fig. 4). The results of our molecular dynamics modelling further suggest that the stability or disruption of the intermolecular stacking interaction between Tyr415 and Tyr470 due to a point mutation in the KRT5/KRT14 dimer could be related to the functional impact. Disturbance of this stacking could affect not only the stability of interaction between keratin chains, but also the accessibility of these tyrosines to kinases and/or phosphatases. It has been proposed[49] that mutation-associated conformational changes may decrease or increase the ability of intermediate filaments to serve as substrates for kinases, which regulate intermediate filament dynamics by modulating their intrinsic properties of solubility, conformation and filament organization, and thereby contribute to disease pathogenesis.

What's already known about this topic?

- KRT5 and KRT14 mutations are responsible for epidermal bullosa simplex

What does this study add?

- Novel mutations in KRT5 and KRT14 genes are associated with epidermal bullosa simplex
- To explore possible correlations with function, the structural effects of the mutations in segment 2B of KRT5 and KRT14 and associated with epidermal bullosa simplex in our patients, as well as those reported previously, were modelled by molecular dynamics

Acknowledgments

We thank Ronald Hancock, Laval University Cancer Research Centre, for critical reading of the manuscript. We thank the Meta Center (http://meta.cesnet.cz) for computer time. This work was funded by the Internal Grant Agency of the Czech Ministry of Health (NR9346), the Czech Ministry of Education (projects LC06023 and MSM0021622415), and the Czech Academy of Sciences (AV0Z50040507, AV0Z50040702).

References

1 Fine JD, Hintner H. *Life with Epidermolysis Bullosa: Etiology, Diagnosis, Multidisciplinary Care and Therapy*. Wien: Springer Verlag, 2009.
2 Hovnanian A, Pollack E, Hilal L *et al.* A missense mutation in the rod domain of keratin 14 associated with recessive epidermolysis bullosa simplex. *Nat Genet* 1993; **3**:327–32.
3 Rugg EL, Morley SM, Smith FJ *et al.* Missing links: Weber–Cockayne keratin mutations implicate the L12 linker domain in effective cytoskeleton function. *Nat Genet* 1993; **5**:294–300.
4 Chan Y, Anton-Lamprecht I, Yu QC *et al.* A human keratin 14 'knockout': the absence of K14 leads to severe epidermolysis bullosa simplex and a function for an intermediate filament protein. *Genes Dev* 1994; **8**:2574–87.
5 Chan YM, Yu QC, LeBlanc-Straceski J *et al.* Mutations in the nonhelical linker segment L1-2 of keratin 5 in patients with Weber–Cockayne epidermolysis bullosa simplex. *J Cell Sci* 1994; **107** (Pt 4):765–74.
6 Jonkman MF, Heeres K, Pas HH *et al.* Effects of keratin 14 ablation on the clinical and cellular phenotype in a kindred with recessive epidermolysis bullosa simplex. *J Invest Dermatol* 1996; **107**:764–9.
7 Hut PH, v d Vlies P, Jonkman MF *et al.* Exempting homologous pseudogene sequences from polymerase chain reaction amplification allows genomic keratin 14 hotspot mutation analysis. *J Invest Dermatol* 2000; **114**:616–19.
8 Schweizer J, Bowden PE, Coulombe PA *et al.* New consensus nomenclature for mammalian keratins. *J Cell Biol* 2006; **174**:169–74.
9 Fuchs E. The cytoskeleton and disease: genetic disorders of intermediate filaments. *Annu Rev Genet* 1996; **30**:197–231.
10 Hatzfeld M, Weber K. The coiled coil of *in vitro* assembled keratin filaments is a heterodimer of type I and II keratins: use of site-specific mutagenesis and recombinant protein expression. *J Cell Biol* 1990; **110**:1199–210.
11 Steinert PM, Marekov LN, Fraser RD *et al.* Keratin intermediate filament structure. Crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *J Mol Biol* 1993; **230**:436–52.
12 Herrmann H, Aebi U. Intermediate filaments and their associates: multi-talented structural elements specifying cytoarchitecture and cytodynamics. *Curr Opin Cell Biol* 2000; **12**:79–90.
13 Wu KC, Bryan JT, Morasso MI *et al.* Coiled-coil trigger motifs in the 1B and 2B rod domain segments are required for the stability of keratin intermediate filaments. *Mol Biol Cell* 2000; **11**:3539–58.
14 Stephens K, Ehrlich P, Weaver M *et al.* Primers for exon-specific amplification of the KRT5 gene: identification of novel and recurrent mutations in epidermolysis bullosa simplex patients. *J Invest Dermatol* 1997; **108**:349–53.
15 Smith TA, Steinert PM, Parry DA. Modeling effects of mutations in coiled-coil structures: case study using epidermolysis bullosa simplex mutations in segment 1a of K5/K14 intermediate filaments. *Proteins* 2004; **55**:1043–52.

16 Liovic M, Stojan J, Bowden PE *et al.* A novel keratin 5 mutation (K5V186L) in a family with EBS-K: a conservative substitution can lead to development of different disease phenotypes. *J Invest Dermatol* 2001; **116**:964–9.
17 Strelkov SV, Herrmann H, Geisler N *et al.* Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. *EMBO J* 2002; **21**:1255–66.
18 Jonkman MF, de Jong MC, Heeres K *et al.* 180-kD bullous pemphigoid antigen (BP180) is deficient in generalized atrophic benign epidermolysis bullosa. *J Clin Invest* 1995; **95**:1345–52.
19 Bruckner-Tuderman L, Winberg JO, Anton-Lamprecht I *et al.* Anchoring fibrils, collagen VII, and neutral metalloproteases in recessive dystrophic epidermolysis bullosa inversa. *J Invest Dermatol* 1992; **99**:550–8.
20 Bergman R. Immunohistopathologic diagnosis of epidermolysis bullosa. *Am J Dermatopathol* 1999; **21**:185–92.
21 Wood P, Baty DU, Lane EB *et al.* Long-range polymerase chain reaction for specific full-length amplification of the human keratin 14 gene and novel keratin 14 mutations in epidermolysis bullosa simplex patients. *J Invest Dermatol* 2003; **120**:495–7.
22 Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 2003; **31**:3381–5.
23 Case DA, Cheatham TE 3rd, Darden T *et al.* The Amber biomolecular simulation programs. *J Comput Chem* 2005; **26**:1668–88.
24 Wang JM, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 2000; **21**:1049–74.
25 Mishra NK, Kulhánek P, Snajdrová L *et al.* Molecular dynamics study of *Pseudomonas aeruginosa* lectin-II complexed with monosaccharides. *Proteins* 2008; **72**:382–92.
26 York DM, Darden TA, Pedersen LG. The effect of long-range electrostatic interactions in simulations of macromolecular crystals – a comparison of the Ewald and truncated list methods. *J Chem Phys* 1993; **99**:8345–8.
27 Kraulis PJ. Molscript – a program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 1991; **24**:946–50.
28 Shemanko CS, Mellerio JE, Tidman MJ *et al.* Severe palmo-plantar hyperkeratosis in Dowling–Meara epidermolysis bullosa simplex caused by a mutation in the keratin 14 gene (KRT14). *J Invest Dermatol* 1998; **111**:893–5.
29 Coulombe PA, Hutton ME, Letai A *et al.* Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: genetic and functional analyses. *Cell* 1991; **66**:1301–11.
30 Muller FB, Kuster W, Wodecki K *et al.* Novel and recurrent mutations in keratin KRT5 and KRT14 genes in epidermolysis bullosa simplex: implications for disease phenotype and keratin filament assembly. *Hum Mutat* 2006; **27**:719–20.
31 Rugg EL, Baty D, Shemanko CS *et al.* DNA based prenatal testing for the skin blistering disorder epidermolysis bullosa simplex. *Prenat Diagn* 2000; **20**:371–7.
32 Yasukawa K, Sawamura D, McMillan JR *et al.* Dominant and recessive compound heterozygous mutations in epidermolysis bullosa simplex demonstrate the role of the stutter region in keratin intermediate filament assembly. *J Biol Chem* 2002; **277**:23670–4.
33 Dong W, Ryynanen M, Uitto J. Identification of a leucine-to-proline mutation in the keratin 5 gene in a family with the generalized Kobner type of epidermolysis bullosa simplex. *Hum Mutat* 1993; **2**:94–102.
34 Rugg EL, Horn HM, Smith FJ *et al.* Epidermolysis bullosa simplex in Scotland caused by a spectrum of keratin mutations. *J Invest Dermatol* 2007; **127**:574–80.
35 Irvine AD, McKenna KE, Bingham A *et al.* A novel mutation in the helix termination peptide of keratin 5 causing epidermolysis bullosa simplex Dowling–Meara. *J Invest Dermatol* 1997; **109**:815–16.
36 Irvine AD, McLean WH. Human keratin diseases: the increasing spectrum of disease and subtlety of the phenotype–genotype correlation. *Br J Dermatol* 1999; **140**:815–28.
37 Schuilenga-Hut PH, Vlies P, Jonkman MF *et al.* Mutation analysis of the entire keratin 5 and 14 genes in patients with epidermolysis bullosa simplex and identification of novel mutations. *Hum Mutat* 2003; **21**:447.
38 Hattori N, Komine M, Kaneko T *et al.* A case of epidermolysis bullosa simplex with a newly found missense mutation and polymorphism in the highly conserved helix termination motif among type I keratins, which was previously reported as a pathogenic missense mutation. *Br J Dermatol* 2006; **155**:1062–3.
39 Ciubotaru D, Bergman R, Baty D *et al.* Epidermolysis bullosa simplex in Israel: clinical and genetic features. *Arch Dermatol* 2003; **139**:498–505.
40 Smith FJ, McKusick VA, Nielsen K *et al.* Cloning of multiple keratin 16 genes facilitates prenatal diagnosis of pachyonychia congenita type 1. *Prenat Diagn* 1999; **19**:941–6.
41 Terrinoni A, Smith FJ, Didona B *et al.* Novel and recurrent mutations in the genes encoding keratins K6a, K16 and K17 in 13 cases of pachyonychia congenita. *J Invest Dermatol* 2001; **117**:1391–6.
42 Glasz-Bona A, Medvecz M, Sajo R *et al.* Easy method for keratin 14 gene amplification to exclude pseudogene sequences: new keratin 5 and 14 mutations in epidermolysis bullosa simplex. *J Invest Dermatol* 2009; **129**:229–31.
43 Yasukawa K, Sawamura D, Goto M *et al.* Epidermolysis bullosa simplex in Japanese and Korean patients: genetic studies in 19 cases. *Br J Dermatol* 2006; **155**:313–17.
44 Pfendner E, Uitto J. Plectin gene mutations can cause epidermolysis bullosa with pyloric atresia. *J Invest Dermatol* 2005; **124**:111–15.
45 Pfendner E, Rouan F, Uitto J. Progress in epidermolysis bullosa: the phenotypic spectrum of plectin mutations. *Exp Dermatol* 2005; **14**:241–9.
46 O'Shea EK, Klemm JD, Kim PS, Alber T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* 1991; **254**:539–44.
47 Meier M, Padilla GP, Herrmann H *et al.* Vimentin coil 1A-A molecular switch involved in the initiation of filament elongation. *J Mol Biol* 2009; **390**:245–61.
48 Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 2004; **60**:2256–68.
49 Omary MB, Ku NO, Tao GZ *et al.* 'Heads and tails' of intermediate filament phosphorylation: multiple sites and functional insights. *Trends Biochem Sci* 2006; **31**:383–94.

4 Dědičné metabolické nemoci

V této části habilitační práce jsou prezentovány publikace týkající se familiální hypercholesterolémie, hyperfenylalaninémie a kongenitální adrenální hyperplazie. Ve všech případech se jedná o závažná onemocnění, u nichž včasná intervence umožňuje účinné blokování rozvinutí asociovaných klinických příznaků. Pro všechny uvedené nemoci je CMBGT jediným pracovištěm v ČR, které provádí jejich molekulárně genetickou diagnostiku. Problematice metabolických nemocí se L. Fajkusová věnuje od roku 2007, po tragické smrti kolegy Libora Kozáka.

Výsledky molekulárně genetické diagnostiky familiální hypercholesterolémie (FH) jsou prezentovány ve dvou publikacích [*Atherosclerosis. 2012 Aug;223(2)* a *Atherosclerosis. 2011 May;216(1)*]. V publikaci „*Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia*“ [*BMC Med Genet. 2010 Jul 27;11*] jsou charakterizovány body zlomu u 8 rozsáhlých genových delecí/duplikací identifikovaných v genu pro LDL receptor (*LDLR*). Výsledky ukázaly, že 6 přeuspořádání je výsledkem nealelické homologní rekombinace (*nonallelic homologous recombination,* NAHR) mezi *Alu* sekvencemi a 2 přeuspořádání jsou výsledkem nehomologního spojování konců (*nonhomologous end joining,* NHEJ). Naše práce jako první popisuje NHEJ jako mechanismus vzniku delecí/duplikací v genu *LDLR*.

Výsledky molekulárně genetické diagnostiky hyperfenylalaninémie (HPA) jsou shrnuty v práci „*Hyperphenylalaninemia in the Czech Republic: genotype-phenotype correlations and in silico analysis of novel missense mutations*“ [*Clin Chim Acta. 2013 Apr 18;419*]. Kromě spektra mutací identifikovaných u pacientů s HPA a jejich klinických projevů, jsou provedeny i podrobné *in silico* analýzy zjištěných *missense* mutací a korelace výsledků těchto analýz s klinickou závažností dané mutace.

V případě kongenitální adrenální hyperplazie (CAH) jsou prezentovány dvě práce [*Eur J Med Genet. 2011 Mar-Apr;54(2)* a *Int J Mol Med. 2010 Oct;26(4)*], které jednak shrnují výsledky molekulárně genetických analýz a korelace zjištěných genetických a klinických nálezů a dále identifikují nové typy chimérních genů vzniklých rekombinací mezi genem *CYP21A2* a pseudogenem *CYP21A1P*.

Publikace

1) The molecular basis of familial hypercholesterolemia in the Czech Republic: spectrum of LDLR mutations and genotype-phenotype correlations. Tichý L, Freiberger T, Zapletalová P, Soška V, Ravčuková B, Fajkusová L. Atherosclerosis. 2012 Aug;223(2):401-8. *(L. Fajkusová jako korespodující autor)*

2) An APEX-based genotyping microarray for the screening of 168 mutations associated with familial hypercholesterolemia. Dušková L, Kopečková L, Jansová E, Tichý L, Freiberger T, Zapletalová P, Soška V, Ravčuková B, Fajkusová L. Atherosclerosis. 2011 May;216(1):139-45. *(L. Fajkusová jako korespondující autor)*

3) Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia. Goldmann R, Tichý L, Freiberger T, Zapletalová P, Letocha O, Soska V, Fajkus J, Fajkusová L. BMC Med Genet. 2010 Jul 27;11:115. *(L. Fajkusová jako korespondující autor)*

4) Hyperphenylalaninemia in the Czech Republic: genotype-phenotype correlations and in silico analysis of novel missense mutations. Réblová K, Hrubá Z, Procházková D, Pazdírková R, Pouchlá S, Zeman J, Fajkusová L. Clin Chim Acta. 2013 Apr 18;419:1-10.*(L. Fajkusová jako korespondující autor)*

5) Chimeric CYP21A1P/CYP21A2 genes identified in Czech patients with congenital adrenal hyperplasia. Vrzalová Z, Hrubá Z, Hrabincová ES, Vrábelová S, Votava F, Kolouśková S, Fajkusová L. Eur J Med Genet. 2011 Mar-Apr;54(2):112-7. *(L. Fajkusová jako korespondující autor)*

6) Identification of CYP21A2 mutant alleles in Czech patients with 21-hydroxylase deficiency. Vrzalová Z, Hrubá Z, St'ahlová Hrabincová E, Pouchlá S, Votava F, Kolousková S, Fajkusová L. Int J Mol Med. 2010 Oct;26(4):595-603. *(L. Fajkusová jako korespondující autor)*

Atherosclerosis 223 (2012) 401–408

Contents lists available at SciVerse ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

The molecular basis of familial hypercholesterolemia in the Czech Republic: Spectrum of *LDLR* mutations and genotype–phenotype correlations

Lukáš Tichý [a,b], Tomáš Freiberger [c,b], Petra Zapletalová [a], Vladimír Soška [d,e], Barbora Ravčuková [c], Lenka Fajkusová [b,*]

[a] Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Brno, Czech Republic
[b] Central European Institute of Technology, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic
[c] Centre for Cardiovascular Surgery and Transplantation, Brno, Czech Republic
[d] 2nd Clinic of Internal Medicine, Masaryk University, Brno, Czech Republic
[e] Department of Biochemistry, Masaryk University, Brno, Czech Republic

ARTICLE INFO

Article history:
Received 28 December 2011
Received in revised form
10 May 2012
Accepted 11 May 2012
Available online 23 May 2012

Keywords:
Familial hypercholesterolemia
LDLR
APOB
LDL cholesterol
Lipid profile

ABSTRACT

Background: Familial hypercholesterolemia (FH), a major risk for coronary heart disease, is predominantly associated with mutations in the genes encoding the low-density lipoprotein receptor (*LDLR*) and its ligand apolipoprotein B (*APOB*).
Results: In this study, we characterize the spectrum of mutations causing FH in 2239 Czech probands suspected to have FH. In this set, we found 265 patients (11.8%) with the *APOB* mutation p.(Arg3527Gln) and 535 patients (23.9%) with a *LDLR* mutation. In 535 probands carrying the *LDLR* mutation, 127 unique allelic variants were detected: 70.1% of these variants were DNA substitutions, 16.5% small DNA rearrangements, and 13.4% large DNA rearrangements. Fifty five variants were novel, not described in other FH populations. For lipid profile analyses, FH probands were divided into groups [patients with the *LDLR* mutation (*LDLR*+), with the *APOB* mutation (*APOB*+), and without a detected mutation (*LDLR* /*APOB*)], and each group into subgroups according to gender. The statistical analysis of lipid profiles was performed in 1722 probands adjusted for age in which biochemical data were obtained without FH treatment (480 *LDLR*+ patients, 222 *APOB*+ patients, and 1020 *LDLR* /*APOB* patients). Significant gradients in i) total cholesterol (*LDLR*+ patients > *APOB*+ patients *LDLR* /*APOB* patients) ii) LDL cholesterol (*LDLR*+ patients > *APOB*+ patients *LDLR* /*APOB* patients in men and *LDLR*+patients > *APOB*+ patients >*LDLR* /*APOB* patients in women), iii) triglycerides (*LDLR* /*APOB* patients > *LDLR*+ patients > *APOB*+ patients), and iv) HDL cholesterol (*APOB*+ patients > *LDLR* /*APOB* patients *LDLR*+ patients) were shown.
Conclusion: Our study presents a large set of Czech patients with FH diagnosis in which DNA diagnostics was performed and which allowed statistical analysis of clinical and biochemical data.

1. Introduction

Familial Hypercholesterolemia (FH), a major risk factor for coronary heart disease (CHD), is an autosomal dominant disorder associated with mutations in the genes encoding the low-density lipoprotein receptor (*LDLR*) [1], apolipoprotein B (*APOB*) [2], and the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) [3].

The clinical phenotype of FH is caused predominantly by mutations in the *LDLR* or *APOB* genes. The *LDLR* gene is located on chromosome 19p13.2, spans 45 kb, contains 18 exons, and encodes a protein that is responsible for clearing of LDL cholesterol from plasma [4]. *LDLR* mutations have been reported along the whole length of the gene in FH patients from around the world. At present, the number of identified unique *LDLR* allelic variants is over 1100: 65% of the variants are DNA substitutions, 24% small DNA rearrangements (<100 bp) and 11% large DNA rearrangements (>100 bp) http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR [5]. In marked contrast to the *LDLR* gene, only one mutation of the *APOB* gene [p.(Arg3527Gln)] is common in Europe [2,6].

FH is characterised by raised serum LDL cholesterol levels, which lead to accelerated atherosclerosis, premature coronary heart disease and occasionally gives rise to tendon xanthomas, xanthelasmas, and arcus lipoides corneae. The frequency of FH in most populations is about 1/500, and so it is possible to predict that

* Corresponding author. Tel.: +420 549494003; fax: +420 549492654.
E-mail address: lfajkusova@fnbrno.cz (L. Fajkusová).

0021-9150/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.atherosclerosis.2012.05.014

approximately 20,000 people could be affected by FH in the Czech Republic. The early identification and treatment of FH probands and their affected relatives with effective lipid-lowering agents is important as this has been shown to significantly reduce both coronary morbidity and mortality [7].

The aim of this study was to analyse (1) the frequency of the mutation p.(Arg3527Gln) in the *APOB* gene; (2) the frequency and spectrum of mutations in the *LDLR* gene; (3) lipid profiles of patients with the *APOB* mutation, the *LDLR* mutation, and patients without detected mutation; and (4) lipid profiles of patients with the *LDLR* mutation generating premature termination codon (PTC) and patients with the *LDLR* mutation not generating PTC in Czech FH probands.

2. Material and methods

2.1. Subjects

Two thousands two hundred and thirty nine unrelated patients with the diagnosis of FH, submitted to the database of the MedPed project in the Czech Republic, were analysed for the presence of mutations in the *APOB* and *LDLR* genes. The experimental research reported in this study has been performed with the approval of the Ethical Committee of the General University Hospital in Prague, Czech Republic, and all patients gave their informed consent for participation in the study. The cohort of probands in our study included (1) patients with untreated total and/or LDL cholesterol serum levels above the 95th percentile of age, sex and population-specific values (1723 patients); (2) patients with elevated total and/or LDL cholesterol in serum but untreated levels unavailable or not exceeding the 95th percentile of age, sex and population-specific values, and, in addition, with high clinical suspicion of FH based on personal history and/or family history of premature coronary heart disease and/or elevated total and LDL cholesterol serum levels in the first degree relatives (516 patients).

2.2. Strategy for identification of FH mutations

Genomic DNA was extracted from peripheral blood leukocytes by a standard procedure [8]. DNA analysis of FH patients was divided into several consecutive steps: (1) PCR-RFLP (Restriction Fragment Length Polymorphism) detection of the most common mutation in the *APOB* gene [p.(Arg3527Gln)]; (2) PCR-RFLP detection of the most common mutations in the *LDLR* gene determined on the basis of a pilot study [p.(Gly592Glu), p.(Asp266Glu), p.(Cys209Tyr), and p.(Arg416Trp)]; (3) PCR-sequencing of *LDLR* exon 4 (the exon with the markedly most frequent occurrence of mutations, http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR); (4) MLPA (Multiple Ligation dependent Probe Amplification) of all *LDLR* exons; (5) PCR-sequencing of the promoter and *LDLR* exons 1, 5, 6, 9, 10, 12, 14; and (6) PCR-DHPLC (Denaturing High Performance Liquid Chromatography) of *LDLR* exons 2, 3, 7, 8, 11, 13, 15, 16, 17, and 18, followed by sequencing of regions which tested positively. In case of detection of a mutation, further DNA analysis continued when (1) a phenotypic manifestation was serious and possibly associated with the presence of two FH mutations or (2) a detected mutation was new with an effect on the protein structure and function which was difficult to predict. This diagnostic process is common in FH molecular genetic testing worldwide [9,10].

Primers for amplification of all exons of the *LDLR* gene and adjacent intron regions are available by request, as well as annealing temperatures, DHPLC temperatures, composition and temperature conditions of individual PCR reactions. The PCR products were analysed on the ProStar Helix DHPLC System (Varian) and/or purified and sequenced on the ABI PRISM 310 DNA-sequencer (Applied Biosystems). MLPA was performed using SALSA MLPA KIT P062-C1 LDLR (MRC-Holland) according to the manufacturer's instruction and analysed on the CEQ 8000 Genetic Analysis System (Beckman Coulter). Data were evaluated using the Coffalyser software (MRC-Holland).

2.3. Laboratory analysis

Lipid analyses were performed in local certified laboratories with appropriate quality controls and using identical analytic methods. Serum total cholesterol and triglyceride (TG) levels were measured using fully automated enzymatic methods, and HDL cholesterol was determined with the same method after the precipitation of apolipoprotein B containing lipoproteins with polyanions. LDL cholesterol was calculated using the Friedewald formula but only if TG levels were <4.5 mmol/l; when TG levels exceeded 4.5 mmol/l, LDL cholesterol values were considered missing [11]. All blood sampling for lipid profile analyses was performed after at least a 12-h fast and without any lipid modification treatment.

2.4. Statistical analysis

Continuous variables were described using mean and standard deviation or geometric mean and 95% confidence interval. Statistical significance of differences between and among groups of patients was analysed by the independent *t*-test, ANOVA, or maximal likelihood chi-square test. Due to significant differences in age between/among groups of patients, age adjustment of lipid characteristics was computed using linear regression prior to comparison of patients' groups. Adjustment for age was performed to the mean age in the studied cohort, it means 40 years.

3. Results

3.1. Molecular genetic analysis of the APOB and LDLR genes

In the set of 2239 probands with a clinical diagnosis of FH, we detected 265 patients with the *APOB* mutation p.(Arg3527Gln). The rest of the patients were analysed for the presence of mutations in the *LDLR* gene. In the 1st step of the *LDLR* analysis [PCR-RFLP detection of the mutations p.(Gly592Glu), p.(Asp266Glu), p.(Cys209Tyr), p.(Arg416Trp) and p.(Pro424_Asn425ins32)], mutations were determined in 240 probands; in the 2nd step (PCR-sequencing of exon 4) in 60 probands; in the 3rd step (MLPA) in 57 probands; in the 4th step (PCR-sequencing of the promoter and exons 1, 5, 6, 9, 10, 12, 14) in 126 probands; and in the 5th step (PCR-DHPLC of exons 2, 3, 7, 8, 11, 13, 15, 16, 17, 18) in 52 probands. In total, we detected mutations in 35.7% of patients suspected to be FH, 11.8% had the *APOB* mutation and 23.9% the *LDLR* mutation. The DNA substitutions and small DNA rearrangements found in Czech FH patients are shown in Table 1 (clear polymorphisms were not included), and large DNA rearrangements in Table 2. In 535 probands carrying a *LDLR* mutation, 127 unique allelic variants were detected: 70.1% of these variants were DNA substitutions, 16.5% small DNA rearrangements, and 13.4% large DNA rearrangements. Fifty five variants were new, and not described in other FH populations.

In silico approaches, called PolyPhen 2 (http://genetics.bwh.harvard.edu/pph2/index.shtml) and Refined SIFT (http://sift.jcvi.org/), were used for predicting effects of missense mutations on the LDLR protein. Although helpful, it should be remembered that these computer prediction programs provide only an indication that an amino acid substitution may affect the biological activity of the mature protein. Using these approaches, 93.8% of missense

Table 1
DNA substitutions and small DNA rearrangements in the *LDLR* gene found in Czech FH patients.

No.	Mutation at cDNA level	Mutation at protein level	Localization	No. of probands	Frequency	PolyPhen 2 result (Hum Div)	PolyPhen 2 result (Hum Var)	SIFT result
1	**c.-153C>T**[a]		Promoter	1	0.19	–	–	–
2	**c.-149C > A**		Promoter	7	1.31	–	–	–
3	**c.-120C>T**[a]		Promoter	1	0.19	–	–	–
4	c.58G > A	p.(Gly20Arg)	1	11	2.06	Benign	Benign	Not tolerated
5	**c.68-2A > T**	**Splicing defect**	Intron 1	1	0.19	–	–	–
6	**c.100T > G**	**p.(Cys34Gly)**	2	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
7	c.131G > A	p.(Trp44*)	2	7	1.31	–	–	–
8	**c.246C>A**[b]	**p.(Cys82*)**	3	1	0.19	–	–	–
9	**c.318_319insC**[b]	**p.(Lys107Glnfs*23)**	4	4	0.75	–	–	–
10	c.325T > C	p.(Cys109Arg)	4	2	0.37	Probably damaging	Probably damaging	Not tolera-ted
11	**c.347_367del21**	**p.(Cys116_Ile121del)**	4	1	0.19	–	–	–
12	c.369_370delTC	p.(Ser123fs*6)	4	1	0.19	–	–	–
13	**c.388T>C**[b]	**p.(Ser130Pro)**	4	2	0.37	Benign	Benign	Not tolera-ted
14	**c.420G>C**[b]	**p.(Glu140Asp)**	4	5	0.93	Probably damaging	Probably damaging	Not tolera-ted
15	c.427T > C	p.(Cys143Arg)	4	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
16	**c.445G > T**	**p.(Gly149Cys)**	4	1	0.19	Probably damaging	Possibly damaging	Not tolera-ted
17	**c.472delT**[b]	**p.(Ser158Profs*48)**	4	1	0.19	–	–	–
18	c.501C > A	p.(Cys167*)	4	1	0.19	–	–	–
19	**c.515A > G**	**p.(Asp172Gly)**	4	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
20	c.527G>T[b]	p.(Gly176Val)	4	5	0.93	Probably damaging	Probably damaging	Tolera-ted
21	c.530C > T[b]	p.(Ser177Leu)	4	7	1.31	Probably damaging	Probably damaging	Not tolera-ted
22	c.542C > G[b]	p.(Pro181Arg)	4	2	0.37	Probably damaging	Probably damaging	Tolera-ted
23	**c.578delA**	**p.(Asp193Alafs*13)**	4	2	0.37	–	–	–
24	**c.625_626dupTG**[b]	**p.(Cys209fs*57)**	4	1	0.19	–	–	–
25	c.626G>A[b]	p.(Cys209Tyr)	4	18	3.36	Probably damaging	Probably damaging	Not tolera-ted
26	c.654_656del[b]	p.(Gly219del)	4	6	1.12	–	–	–
27	c.662A > G	p.(Asp221Gly)	4	12	2.24	Probably damaging	Probably damaging	Not tolera-ted
28	c.676T > C	p.(Ser226Pro)	4	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
29	c.681C > G	p.(Asp227Glu)	4	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
30	c.682G > A	p.(Glu228Lys)	4	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
31	c.691T > G	p.(Cys231Gly)	4	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
32	c.693C > A	p.(Cys231*)	4	1	0.19	–	–	–
33	c.761A > C[b]	p.(Gln254Pro)	5	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
34	c.796G > A	p.(Asp266Asn)	5	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
35	c.798T > A[b]	p.(Asp266Glu)	5	88	16.45	Probably damaging	Probably damaging	Not tolera-ted
36	**c.808T>C**[b]	**p.(Cys270Arg)**	5	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
37	**c.815A > C**	**p.(Asn272Thr)**	5	2	0.37	Benign	Benign	Tolera-ted
38	**c.817+1G > A**	**Splicing defect**	Intron 5	1	0.19	–	–	–
39	c.828C > A	p.(Cys276*)	6	6	1.12	–	–	–
40	**c.846C > A**	**p.(Phe282Leu)**	6	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
41	c.862G > T	p.(Glu288*)	6	1	0.19	–	–	–
42	**c.880_881delAA**[b]	**p.(Lys294Serfs*6)**	6	1	0.19	–	–	–
43	**c.896delC**[b]	**p.(Ala299Valfs*71)**	6	1	0.19	–	–	–
44	c.910G > A	p.(Asp304Asn)	6	2	0.37	Probably damaging	Possibly damaging	Not tolera-ted
45	**c.917C>G**[b]	**p.(Ser306*)**	6	6	1.12	–	–	–
46	**c.919G > A**	**p.(Asp307Asn)**	6	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
47	c.939C > A	p.(Cys313*)	6	2	0.37	–	–	–
48	**c.940+1G > C**	**Splicing defect**	Intron 6	1	0.19	–	–	–
49	**c.947A > C**	**p.(Asn316Thr)**	7	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
50	c.949G > T	p.(Glu317*)	7	2	0.37	–	–	–
51	c.970delG[b]	p.(Gly324Alafs*46)	7	2	0.37	–	–	–
52	**c.977C > G**	**p.(Ser326Cys)**	7	1	0.19	Probably damaging	Possibly damaging	Not tolera-ted
53	c.986G > A	p.(Cys329Tyr)	7	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
54	c. 1013G > A	p.(Cys338Tyr)	7	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
55	**c.1024G > T**	**p.(Asp342Tyr)**	7	1	0.19	Possibly damaging	Benign	Not tolera-ted
56	c.1048C > T	p.(Arg350*)	7	2	0.37	–	–	–
57	c.1055G > A	p.(Cys352Tyr)	7	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
58	**c.1053_1060dup**[b]	**p.(Asp354fs*19)**	7	1	0.19	–	–	–
59	**c.1061A > C**	**p.(Asp354Ala)**	8	9	1.68	Probably damaging	Probably damaging	Not tolera-ted
60	**c.1091G > C**	**p.(Cys364Ser)**	8	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
61	c.1133A > C	p.(Gln378Pro)	8	1	0.19	Benign	Benign	Not tolera-ted
62	**c.1151_1159delAGCTGGACC**	**p.(Gln384_Asp386del)**	8	1	0.19	–	–	–
63	c.1205_1207delTCT	p.(Phe403del)	9	4	0.75	–	–	–
64	c.1217G > C	p.(Arg406Pro)	9	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
65	c.1222G > A	p.(Glu408Lys)	9	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
66	**c.1223A > T**	**p.(Glu408Val)**	9	3	0.56	Probably damaging	Probably damaging	Not tolera-ted
67	c.1246C > T[b]	p.(Arg416Trp)	9	22	4.11	Probably damaging	Possibly damaging	Not tolera-ted
68	c.1247G > C	p.(Arg416Pro)	9	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
69	**c.1272_1273ins96**[b]	**p.(Pro424_Asn425ins32)**	9	9	1.68	–	–	–
70	c.1291G > A	p.(Ala431Thr)	9	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
71	c.1358+2T > A	Splicing defect	Intron 9	6	1.12	–	–	–
72	**c.1367_1376delTTGACAGAGC**	**p.(Leu456Profs*48)**	10	1	0.19	–	–	–

(*continued on next page*)

Table 1 (*continued*)

No.	Mutation at cDNA level	Mutation at protein level	Localization	No. of probands	Frequency	PolyPhen 2 result (Hum Div)	PolyPhen 2 result (Hum Var)	SIFT result
73	**c.1377_1380delCCAC**	**p.(Ala459fs*47)**	10	1	0.19	–	–	–
74	**c.1394A > G**	**p.(Tyr465Cys)**	10	1	0.19	Probably damaging	Possibly damaging	Not tolera-ted
75	c.1414G > T	p.(Asp472Tyr)	10	4	0.75	Probably damaging	Possibly damaging	Not tolera-ted
76	c.1444G > A	p.(Asp482Asn)	10	3	0.56	Probably damaging	Probably damaging	Not tolera-ted
77	c.1474G > A	p.(Asp492Asn)	10	5	0.93	Probably damaging	Probably damaging	Not tolera-ted
78	**c.1510A > T**	**p.(Lys504*)**	10	1	0.19	–	–	–
79	c.1529C > T	p.(Thr510Met)	10	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
80	**c.1538G > A**	**p.(Arg513Lys)**	10	1	0.19	Benign	Benign	Tolera-ted
81	**c.1552A>G**[b]	**p.(Lys518Glu)**	10	1	0.19	Benign	Benign	Tolera-ted
82	**c.1533dupA**	**p.(Phe512Ilefs*24)**	10	2	0.37	–	–	–
83	c.1567G > A	p.(Val523Met)	10	7	1.31	Probably damaging	Probably damaging	Not tolera-ted
84	**c.1662_1669dup**[b]	**p.(Thr557Cysfs*3)**	11	4	0.75	–	–	–
85	**c.1672G > T**	**p.(Glu558*)**	11	1	0.19	–	–	–
86	**c.1721G > T**	**p.(Arg574Leu)**	12	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
87	**c.1748A > G**	**p.(His583Arg)**	12	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
88	**c.1751delC**[b]	**p.(Ser584fs*81)**	12	1	0.19	–	–	–
89	c.1775G>A[b]	p.(Gly592Glu)	12	103	19.25	Probably damaging	Probably damaging	Not tolera-ted
90	c.1822C > T	p.(Pro608Ser)	12	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
91	**c.1829C > G**	**p.(Ser610Cys)**	12	3	0.56	Probably damaging	Probably damaging	Not tolera-ted
92	c.1845+1G > T	Splicing defect	Intron 12	1	0.19	–	–	–
93	c.1864G > A	p.(Asp622Asn)	13	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
94	**c.1865A > G**	**p.(Asp622Gly)**	13	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
95	c.1897C > T	p.(Arg633Cys)	13	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
96	c.1998G>A[b]	p.(Trp666*)	14	1	0.19	–	–	–
97	c.1999T > C	p.(Cys667Arg)	14	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
98	**c.2023G>A**[b]	**p.(Gly675Ser)**	14	3	0.56	Possibly damaging	Possibly damaging	Not tolera-ted
99	c.2043C > A	p.(Cys681*)	14	1	0.19	–	–	–
100	c.2054C > T[b]	p.(Pro685Leu)	14	2	0.37	Probably damaging	Probably damaging	Not tolera-ted
101	**c.2068delC**[b]	**p.(His690Thrfs*19)**	14	12	2.24	–	–	–
102	c.2093G > T	p.(Cys698Phe)	14	2	0.37	Probably damaging	Probably damaging	Not tolera-ted
103	c.2096C > T	p.(Pro699Leu)	14	1	0.19	Probably damaging	Probably damaging	Not tolera-ted
104	c.2101G > A	p.(Gly701Ser)	14	1	0.19	Benign	Benign	Not tolera-ted
105	**c.2140+2T > C**	**Splicing defect**	Intron 14	1	0.19	–	–	–
106	c.2389G > A	p.(Val797Met)	16	2	0.37	Benign	Benign	Tolera-ted
107	**c.2390-1G > A**	**Splicing defect**	Intron 16	1	0.19	–	–	–
108	**c.2396T > G**	**p.(Leu799Arg)**	17	2	0.37	Possible damaging	Possible damaging	Not tolera-ted
109	c.2416_2417insG[b]	p.(Val806Glyfs*11)	17	2	0.37	–	–	–
110	c.2479G > A	p.(Val827Ile)	17	2	0.37	Probably damaging	Probably damaging	Not tolera-ted

[a] Mutations published in our previous study [12].
[b] Mutations published in our previous study [13], mutations marked by bold letters were not described in other countries so far.

mutations detected in our FH patients (60 from 64) had an adverse effect on the LDLR protein according at least one of the programs used, and the remaining 6.2% missense variants were predicted to be "benign" and "tolerated". Nevertheless, the carriers of these "non-deleterious" variants were included in the group of patients with the *LDLR* mutation because (1) another suspected sequence variant was not found in the *LDLR* gene and (2) family segregation analyses were performed, and although the number of examined family members was too small for reliable evaluation, lipid profiles typical for FH segregated with the mutation.

Table 2
Large DNA rearrangements in the *LDLR* gene detected in Czech FH patients.

No.	Deletion/duplication	No. of probands	Frequency
1	Promoter_ex1del	1	0.19
2	Promoter_ex2del[a]	1	0.19
3	Exon3_12del[a]	1	0.19
4	Exon5_10del[a]	3	0.56
5	Exon7_8del	1	0.19
6	Exon7_10del	2	0.37
7	**Exon7_12del**	1	0.19
8	Exon9_14del[a]	12	2.24
9	Exon9_15del[a]	9	1.68
10	**Exon11_17del**	1	0.19
11	Exon13_18del	1	0.19
12	Exon16_17del	1	0.19
13	Exon17_3′UTRdel	1	0.19
14	Exon2_6dup[a]	13	2.43
15	**Exon2_15dup**	1	0.19
16	Exon4_8dup[a]	2	0.37
17	Exon16_18dup[a]	6	1.12

[a] Mutations published in our previous study [14]; rearrangements marked by bold letters were not described in other countries so far.

3.2. Statistical analysis of lipid values

In Tables 3 and 4, we present statistical comparisons of lipid profiles of our FH probands (females and males separately) adjusted for age with the *APOB* mutation (*APOB*+), with the *LDLR* mutation (*LDLR*+), without a detected mutation (*LDLR*−/*APOB*−) (Table 3), with the *LDLR* PTC-mutation, and with the *LDLR* non PTC-mutation (Table 4). These analyses were performed in 1722 probands in whom biochemical data were obtained without any FH treatment (480 *LDLR*+ patients, 222 *APOB*+ patients, and 1020 *LDLR*−/*APOB*− patients). These large cohorts of FH probands allow us to perform statistical analyses, the results of which show significant gradients in total cholesterol (*LDLR*+ > *APOB*+ = *LDLR*−/*APOB*−), triglycerides (*LDLR*−/*APOB*− > *LDLR*+ > *APOB*+), and HDL cholesterol (*APOB*+ > *LDLR*−/*APOB*− = *LDLR*+) without differences between genders. In the case of LDL cholesterol, the situation was mildly different. We detected the significant gradient *LDLR*+ > *APOB*+ > *LDLR*−/*APOB*− in women and *LDLR*+ > *APOB*+ = *LDLR*−/*APOB*− in men. Lipid profile analyses also revealed that patients with mutations associated with PTC and/or large gene rearrangements have higher total cholesterol and LDL cholesterol than patients with missense mutations but this effect was significant only in the case of analysis of all patients adjusted

Table 3
Lipid characteristics of patients adjusted for age according to diagnosis and sex stratification.

Characteristics[c]	*APOB*+	*LDLR*+	*LDLR*−/*APOB*−	*p*[d]
Men	$N = 50$	$N = 193$	$N = 320$	
Age	29.0 (19.8)[a]	32.6 (18.0)[a]	39.2 (17.3)[b]	<0.001
LDL cholesterol	5.98 (0.98)[a]	6.61 (1.39)[b]	6.00 (1.45)[a]	<0.001
Total cholesterol	8.07 (1.00)[a]	8.78 (1.64)[b]	8.32 (1.46)[a]	0.001
Triglycerides[e]	1.01	1.42	1.85	<0.001
	(0.88; 1.15)[a]	(1.32; 1.52)[b]	(1.76; 1.94)[c]	
HDL cholesterol	1.49 (0.33)[a]	1.33 (0.32)[b]	1.37 (0.31)[b]	0.008
Women	$N = 172$	$N = 287$	$N = 700$	
Age	35.2 (20.8)[a]	36.2 (18.1)[a]	47.4 (17.3)[b]	<0.001
LDL cholesterol	6.21 (1.26)[b]	7.31 (1.77)[c]	5.90 (1.29)[a]	<0.001
Total cholesterol	8.44 (1.32)[a]	9.60 (1.91)[b]	8.27 (1.32)[a]	<0.001
Triglycerides[c]	1.14	1.35	1.71	<0.001
	(1.07; 1.22)[a]	(1.27; 1.42)[b]	(1.65; 1.77)[c]	
HDL cholesterol	1.62 (0.42)[a]	1.53 (0.39)[b]	1.56 (0.39)[b]	0.050

a,b: the same letters denotes groups of patients without statistically significant differences.
[c] Described by mean (standard deviation).
[d] Statistical significance evaluated by ANOVA.
[e] Described by geometric mean (95% confidence interval), statistical significance evaluated by ANOVA based on log transformed data.

for age and sex (data not shown) and men adjusted for age. No statistically significant differences were observed in the female group but the same trends as in males were present.

Further, we performed evaluation of cardiovascular events and tendon xanthomas occurrence in our FH patients (Table 5). The reliable clinical data were available from 998 probands older than 30 years of age. Cardiovascular disease, defined as presence of coronary heart disease, peripheral artery disease, ischemic stroke, or transient ischemic attack, were present in 164 of them (16.4%). Cardiovascular diseases were more frequent in *LDLR*+ group (20.8%) compared with *LDLR*−/*APOB*− group (16.0%) and *APOB*+ group (7.8%), but only the difference between *LDLR*+ and *APOB*+ group reached statistical significance. Tendon xanthomas were present in 38 patients (3.8%). The highest occurrence was found in *LDLR*+ group (13.1%), lower in *APOB*+ group (5.2%), and the lowest in *LDLR*−/*APOB*− group (0.7%). The frequencies differed significantly between particular groups of patients (for details see Table 5).

In the set of our FH patients, we detected 3 patients with both *LDLR* alleles mutated of which 2 were compound heterozygotes, and 1 was a true homozygote. The lipid profiles and clinical finding of these patients were not included in the analyses in Tables 3–5.

4. Discussion

In FH, DNA testing has clinical utility in (1) determination of the unequivocal diagnosis and (2) identification of affected relatives. For diagnostic purposes, it is critically important to determine whether a recognized sequence variant has a pathogenic effect on the LDLR protein structure and/or function with caution in order to avoid false positive diagnosis. Definite criteria are used for ascertainment of pathogenicity of sequence changes in cases when functional studies for particular mutations are not performed. The cosegregation of the *LDLR* variant with the FH phenotype inside of the proband's family is one of these criteria. Another criterion that can be applied for determination of pathogenicity of mutations is by using the computer programs PolyPhen and SIFT. We performed analysis of missense mutations detected in our FH patients using these program tools and the pathogenicity of four sequence variants [p.(Asn272Thr), p.(Arg513Lys), p.(Lys518Glu), and p.(Val797Met)] was indefinite (Table 1). By using PolyPhen 2 and Refined SIFT, these mutations were evaluated as "benign" and simultaneously "tolerated", respectively. For confirmation of these findings, we used the computer program SNPs3D (http://www.snps3d.org) and these were predicted as "benign" once again. LDL cholesterol, total cholesterol, and other clinical findings in probands carrying these mutations did not differ from probands with other *LDLR* mutations. Further, family screening was used to assess the pathogenicity of these gene variants, but unfortunately the number of examined family members was too small for reliable analyses.

Table 4
Lipid characteristics of patients adjusted for age according to *LDLR* mutation and sex stratification.

Characteristics[a]	Non PTC-mutation	PTC-mutation	*p*[b]
Men	$N = 139$	$N = 51$	
LDL cholesterol	6.38 (1.19)	7.27 (1.71)	<0.001
Total cholesterol	8.60 (1.54)	9.31 (1.82)	0.008
Triglycerides[c]	1.43 (1.31; 1.55)	1.37 (1.22; 1.54)	0.615
HDL cholesterol	1.35 (0.32)	1.30 (0.32)	0.412
Women	$N = 220$	$N = 64$	
LDL cholesterol	7.22 (1.72)	7.69 (1.93)	0.065
Total cholesterol	9.49 (1.81)	10.01 (2.23)	0.059
Triglycerides[c]	1.41 (1.32; 1.50)	1.18 (1.06; 1.33)	0.012
HDL cholesterol	1.54 (0.40)	1.52 (0.39)	0.666

[a] Described by mean (standard deviation).
[b] Statistical significance evaluated by *t*-test.
[c] Described by geometric mean (95% confidence interval), statistical significance evaluated by *t*-test based on log transformed data.

Twenty-five variants have been reported in the promoter region of the *LDLR* gene (http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR). In the set of Czech FH probands, we detected three promoter mutations. In two of these (c.-153C > T and c.-120C > T), we performed functional characterization through the use of transfection of HepG2 cells and the luciferase reporter assay [12]. In the mutation c.-149C > A, the functional characterization was not done but this mutation (as well as c.-153C > T) affects the sterol-regulatory-element binding site and thus presumably has the same or a similar effect as c.-153C > T. Further, we detected 7 types of *LDLR* mutations that disrupt the exon–intron consensus splice sequences. The causal effect on mRNA splicing was confirmed using the *in silico* prediction tools NetGene2 (http://www.cbs.dtu.dk/services/NetGene2) and NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html).

Analysis of DNA by MLPA was found to be a simple, rapid and robust mean of detecting major rearrangements in *LDLR*. Totally, we detected 17 types of large gene rearrangements (13 deletions and 4 duplications), 3 of which are novel not described in the literature so far. In eight types of gene rearrangements, we performed sequence analysis of deletion/duplication breakpoints and determined that non-allelic homologous recombination (NAHR) between *Alu* elements is involved in 6 events, while non-homologous end joining (NHEJ) is implicated in 2 rearrangements. Our study thus described for the first time NHEJ as a mechanism involved in genomic rearrangements in the *LDLR* gene [14].

Across nations, the spectrum of *LDLR* gene mutations is mostly complex. In populations such as Ashkenazi [15] and Sephardic [16] Jews, Icelanders [17], or Finns [18] a few mutations predominate due to founder effects. In countries such as France [19], Poland [20], Germany [21], Netherlands [22], or Austria [23] many more mutations produce a highly heterogeneous picture. In the Czech Republic, we found 127 different mutations, three of which account for 39.8% of FH cases. In Table 6, we present the most frequent *LDLR* mutations and their frequencies in European countries adjacent to the Czech Republic for which data concerning the DNA analysis of *LDLR* are known. It is notable that the most frequent mutation in Czech FH patients p.(Gly592Glu) is also the most common in Poland and that the second most common mutation p.(Asp266Glu) is the most common in Germany and Austria.

The frequency of detected mutations in patients with FH varies from 30 to 90% in recent studies and depends mainly on the clinical

Table 5
Occurrence of cardiovascular events and tendon xanthomas.

Clinical occurrence	All ($N = 998$)	*APOB*+ ($N = 77$)	*LDLR*+ ($N = 221$)	*LDLR*−/*APOB*− ($N = 700$)	p (*APOB*+ vs. *LDLR*+)	p (*APOB*+ vs. *LDLR*−/*APOB*−)	p (*LDLR*+ vs. *LDLR*−/*APOB*−)
CHD	130 (13.0%)	5 (6.5%)	40 (18.1%)	85 (12.1%)	0.014	0.142	0.024
PAD	19 (1.9%)	0 (0.0%)	7 (3.2%)	12 (1.7%)	0.114	0.247	0.185
Stroke	41 (4.1%)	1 (1.3%)	12 (5.4%)	28 (4.0%)	0.126	0.235	0.363
Any cardiovascular event	164 (16.4%)	6 (7.8%)	46 (20.8%)	112 (16.0%)	0.010	0.057	0.098
Tendon xantomas							
All	38 (3.8%)	4 (5.2%)	29 (13.1%)	5 (0.7%)	0.056	0.005	<0.001
<45 years	$N = 148$	$N = 20$	$N = 66$	$N = 62$			
	6 (4.1%)	1 (5.0%)	5 (7.6%)	0 (0.0%)	0.692	0.077	0.027
>45 years	$N = 850$	$N = 57$	$N = 155$	$N = 638$			
	32 (3.8%)	3 (5.3%)	24 (15.5%)	5 (0.8%)	0.048	0.002	<0.001

Only patients older than 30 years of age with reliable clinical data were included. Data for tendon xanthomas are presented for group of all patients and for patients younger/older than 45 years separately. CHD = coronary heart disease (includes myocardial infarction, angina pectoris, coronary artery bypass grafting or percutaneous coronary intervention); PAD = peripheral artery disease (peripheral arterial bypass graft or peripheral percutaneous transluminal angioplasty); stroke includes ischemic stroke and transient ischemic attack. Statistical significance evaluated using maximal likelihood chi-square test.

criteria used for diagnosis [24,25]. In this study, the detection rate of *LDLR* and *APOB* mutations was 35.7% of the total number of probands estimated to be FH. In the group of probands selected on the basis of the criterion (1) and (2) (see Material and Methods 2.2.), we detected mutations in 40.6% patients (28.0% had the *LDLR* mutation and 12.6% the *APOB* mutation) and 19.8% patients (10.5% with the *LDLR* mutation and 9.3% with the *APOB* mutation), respectively. It was impossible to identify a mutation associated with FH for 64.3% of the patients. Possible explanations are (1) gene defects associated with raised LDL-cholesterol levels hide in other genes involved in lipid metabolism, or (2) patients are clinically misdiagnosed with FH. To address the point 1, we performed sequence analysis of the *PCSK9* gene in 40 selected probands suspected to be FH and without the *APOB* mutation or a *LDLR* mutation but a causal *PCSK9* mutation was not detected. Concerning the point 2, we decided to choose clinical and biochemical criteria for inclusion of patients in the FH set which enable us to detect mutations with a high sensitivity, although with a lower specificity.

Not surprisingly and in concordance with other authors, we demonstrated higher age-adjusted total and LDL cholesterol levels in *LDLR*+ patients compared with *APOB*+ and *LDLR*−/*APOB*− patients [26–30]. While total cholesterol in both males and females and LDL cholesterol in males did not show any significant difference between *APOB*+ and *LDLR*−/*APOB*− groups, LDL cholesterol in females was significantly higher in the *APOB*+ compared with *LDLR*−/*APOB*− group in our study. Considering TG and HDL cholesterol levels, data in the literature are more conflicting. No differences in TG and HDL concentrations were reported in large cohorts between *LDLR*+, *APOB*+ and *LDLR*−/*APOB*− groups in the studies [27,28], while higher TG and lower HDL cholesterol was demonstrated in *LDLR*− compared to *LDLR*+ patients in another study [29]. HDL cholesterol was higher in our *APOB*+ patients compared with other two groups, but no differences were shown in HDL cholesterol between the *LDLR*+ and *LDLR*−/*APOB*− groups. However, we showed a clear gradient in TG levels with the highest concentrations in the *LDLR*−/*APOB*−, lower in the *LDLR*+ and the lowest in the *APOB*+ group. The differences in total cholesterol, LDL cholesterol and TG levels between *LDLR*+ and *LDLR*−/*APOB*− patients may support the assumption that another cause of dyslipidemia rather than the *LDLR* mutation is present, although not identified, in our *LDLR*−/*APOB*− group of patients. This observation underlines the relevance of genetic testing in FH for clinical practice and for screening purposes where early and correct diagnosis is very important for early and effective treatment, and also for research where the proper definition of study subjects is essential.

Table 6
The most frequent mutations of the *LDLR* gene detected in selected European countries and their frequency.

Country, references	Total amount of detected *LDLR* variants	Three most frequent mutations (frequency)	Total % of three most frequent mutations from all mutation
Czech Republic	127	p.(Gly592Glu) (19.3) p.(Asp266Glu) (16.4) p.(Cys209Tyr) (4.1)	39.8
Poland, [20]	71	p.(Gly592Glu) (22.5) Exon4_8dup (9.5) p.(Asp221Gly) (5.3)	37.3
Germany, [21]	77	p.(Asp266Glu) (3.7) p.(Glu228*) (3.1) c.313 +1G > A/C (3.1)	9.9
Austria, [23]	100	p.(Asp266Glu) (11.1) p.(Asp200Gly) (5.9) p.(Asp157Asn) (4.6) p.(Val485del47) (4.6)	26.2

Concerning PTC- and non-PTC-mutation status, most studies showed higher total and LDL cholesterol levels [31–34] and higher frequency of cardiovascular events [31,33], in PTC-mutation carriers. The occurrence of tendon xanthomas [35] and the tendency to lower intima-media thickness of the common carotid artery [34] were reported to be more frequent in PTC-mutation carriers. However, several studies failed to show higher total cholesterol and LDL cholesterol levels in patients with *LDLR* PTC-mutations [35,36]. In our study, we were able to demonstrate statistically significant differences in total and LDL cholesterol levels between PTC and non PTC-mutation carriers only in males, although similar trends were present also in females. A similar heterogeneity in published studies can be noted in comparison of HDL cholesterol in patients with PTC and non PTC-mutations. We found no differences in HDL levels in concordance with [32,37], while [26,31,36] showed lower HDL cholesterol in PTC-mutation carriers. Concerning TG levels, Gudnason et al. [32] reported higher levels in non PTC-mutation carriers compared with PTC-mutation carriers, but most studies showed no differences [26,31,36,37]. In our study, no significant differences were detected between the groups in males, but higher TG levels were noticed in female non PTC-mutation carriers.

The most serious consequence of FH is a premature clinical manifestation of atherosclerosis. The occurrence of cardiovascular events and tendon xanthomas was 16.4% and 3.8%, respectively, in the set of our FH patients older than 30 years of age. The highest incidence of both was recorded in the *LDLR*+ group (20.8% and 13.1%), in which also the highest level of LDL cholesterol was demonstrated. Besides a key role for LDL cholesterol, other risk

factors were described to be involved in clinical manifestation of atherosclerosis and tendon xanthomas occurrence in FH subjects, like gender, smoking, hypertension, diabetes, lifestyle, other lipoprotein phenotypes, sequence variations of genes having a role in lipid metabolism, etc. [35,38–41].

In conclusion, our study presents a large set of Czech patients with FH diagnosis in which DNA diagnostics was performed. This allowed us to accomplish statistical analysis of clinical and biochemical data and to find significant relationships between gene defects and clinical status of the relevant patients. Based on our experience with the multistep FH diagnostic protocols, we believe that the most suitable FH diagnostic procedure for our population comprises (1) detection of the most common mutation in the *APOB* gene, (2) direct sequencing of the *LDLR* gene, and (3) MLPA of the *LDLR* gene. Alternatively, the genotyping DNA microarray covering mutations detected in a population can be used, followed by a direct sequencing and MLPA in samples negative after microarray analysis. In practice, we have tested such approach [42], but this protocol does not seem to be faster and cheaper compared with the protocol in that *LDLR* sequencing is performed in the second step of analysis [after detection of the mutation p.(Arg3527Gln) in *APOB*].

Conflict of interest

None.

Acknowledgements

We would like to thank the physicians from the national (M. Vrablík, prof. R. Češka) and regional centres of the Czech MedPed project (V. Bláha, M. Budíková, J. Buryška, R. Cífková, L. Dlouhý, L. Dostálová-Kopečná, H. Halámková, J. Hyánek, J. Hyjánek, Z. Krejsová, J. Macháček, Š. Malá, J. Malý, V. Miláček, J. Mraček, H. Podzimková, D. Povalačová, H. Rosolová, F. Stožický, E. Šipková, L. Ťoukálková, Z. Urbanová, H. Vaverková, S. Zemek, A. Žák) and other physicians participating in this project for providing us with their patients' blood samples and clinical and laboratory data, and O. Letocha and M. Plotěná for technical assistance. This work was funded by the Czech Ministry of Education (MSMT 2B08060) and the project "CEITEC – Central European Institute of Technology" (CZ.1.05/1.1.00/02.0068).

References

[1] Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. Hum Mutat 1992;1(6):445–66.
[2] Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc Natl Acad Sci U S A 1989 Jan;86(2):587–91.
[3] Abifadel M, Varret M, Rabes JP, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 2003 Jun;34(2):154–6.
[4] Tolleshaug H, Goldstein JL, Schneider WJ, Brown MS. Posttranslational processing of the LDL receptor and its genetic disruption in familial hypercholesterolemia. Cell 1982 Oct;30(3):715–24.
[5] Leigh SE, Foster AH, Whittall RA, Hubbart CS, Humphries SE. Update and analysis of the University College London low density lipoprotein receptor familial hypercholesterolemia database. Ann Hum Genet 2008 Jul;72(Pt 4):485–98.
[6] Gaffney D, Reid JM, Cameron IM, et al. Independent mutations at codon 3500 of the apolipoprotein B gene are associated with hyperlipidemia. Arterioscler Thromb Vasc Biol 1995 Aug;15(8):1025–9.
[7] Versmissen J, Oosterveer DM, Yazdanpanah M, et al. Efficacy of statins in familial hypercholesterolaemia: a long term cohort study. BMJ 2008;337:a2423.
[8] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988 Feb 11;16(3):1215.
[9] Civeira F, Ros E, Jarauta E, et al. Comparison of genetic versus clinical diagnosis in familial hypercholesterolemia. Am J Cardiol 2008 Nov 1;102(9):1187–93. 93 e1.
[10] Taylor A, Wang D, Patel K, et al. Mutation detection rate and spectrum in familial hypercholesterolaemia patients in the UK pilot cascade project. Clin Genet Jun; 77(6):572–580.
[11] Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972 Jun;18(6):499–502.
[12] Francova H, Trbusek M, Zapletalova P, Kuhrova V. New promoter mutations in the low-density lipoprotein receptor gene which induce familial hypercholesterolaemia phenotype: molecular and functional analysis. J Inherit Metab Dis 2004;27(4):523–8.
[13] Kuhrova V, Francova H, Zapletalova P, et al. Spectrum of low density lipoprotein receptor mutations in Czech hypercholesterolemic patients. Hum Mutat 2001 Sep;18(3):253.
[14] Goldmann R, Tichy L, Freiberger T, et al. Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia. BMC Med Genet. 11: 115.
[15] Meiner V, Landsberger D, Berkman N, et al. A common Lithuanian mutation causing familial hypercholesterolemia in Ashkenazi Jews. Am J Hum Genet 1991 Aug;49(2):443–9.
[16] Leitersdorf E, Reshef A, Meiner V, et al. A missense mutation in the low density lipoprotein receptor gene causes familial hypercholesterolemia in Sephardic Jews. Hum Genet 1993 Mar;91(2):141–7.
[17] Gudnason V, Sigurdsson G, Nissen H, Humphries SE. Common founder mutation in the LDL receptor gene causing familial hypercholesterolaemia in the Icelandic population. Hum Mutat 1997;10(1):36–44.
[18] Koivisto UM, Viikari JS, Kontula K. Molecular characterization of minor gene rearrangements in Finnish patients with heterozygous familial hypercholesterolemia: identification of two common missense mutations (Gly823–>Asp and Leu380–>His) and eight rare mutations of the LDL receptor gene. Am J Hum Genet 1995 Oct;57(4):789–97.
[19] Marduel M, Carrie A, Sassolas A, et al. Molecular spectrum of autosomal dominant hypercholesterolemia in France. Hum Mutat. Nov; 31(11): E1811–E1824.
[20] Chmara M, Wasag B, Zuk M, et al. Molecular characterization of Polish patients with familial hypercholesterolemia: novel and recurrent LDLR mutations. J Appl Genet. 51(1): 95–106.
[21] Nauck MS, Koster W, Dorfer K, et al. Identification of recurrent and novel mutations in the LDL receptor gene in German patients with familial hypercholesterolemia. Hum Mutat 2001 Aug;18(2):165–6.
[22] Fouchier SW, Defesche JC, Umans-Eckenhausen MW, Kastelein JP. The molecular basis of familial hypercholesterolemia in The Netherlands. Hum Genet 2001 Dec;109(6):602–15.
[23] Widhalm K, Dirisamer A, Lindemayr A, Kostner G. Diagnosis of families with familial hypercholesterolaemia and/or Apo B-100 defect by means of DNA analysis of LDL-receptor gene mutations. J Inherit Metab Dis 2007 Apr;30(2):239–47.
[24] Graham CA, McIlhatton BP, Kirk CW, et al. Genetic screening protocol for familial hypercholesterolemia which includes splicing defects gives an improved mutation detection rate. Atherosclerosis 2005 Oct;182(2):331–40.
[25] Humphries SE, Norbury G, Leigh S, Hadfield SG, Nair D. What is the clinical utility of DNA testing in patients with familial hypercholesterolaemia? Curr Opin Lipidol 2008 Aug;19(4):362–8.
[26] Real JT, Chaves FJ, Ejarque I, et al. Influence of LDL receptor gene mutations and the R3500Q mutation of the apoB gene on lipoprotein phenotype of familial hypercholesterolemic patients from a South European population. Eur J Hum Genet 2003 Dec;11(12):959–65.
[27] Fouchier SW, Rodenburg J, Defesche JC, Kastelein JJ. Management of hereditary dyslipidaemia; the paradigm of autosomal dominant hypercholesterolaemia. Eur J Hum Genet 2005 Dec;13(12):1247–53.
[28] Fouchier SW, Kastelein JJ, Defesche JC. Update of the molecular basis of familial hypercholesterolemia in The Netherlands. Hum Mutat 2005 Dec; 26(6):550–6.
[29] van Aalst-Cohen ES, Jansen AC, Tanck MW, et al. Diagnosing familial hypercholesterolaemia: the relevance of genetic testing. Eur Heart J 2006 Sep; 27(18):2240–6.
[30] Van Gaal LF, Peeters AV, De Block CE, de Leeuw IH, Thiart R, Kotze MJ. Low-density lipoprotein receptor gene mutation analysis and clinical correlation in Belgian hypercholesterolaemics. Mol Cell Probes 2001 Dec;15(6):329–36.
[31] Bertolini S, Cantafora A, Averna M, et al. Clinical expression of familial hypercholesterolemia in clusters of mutations of the LDL receptor gene that cause a receptor-defective or receptor-negative phenotype. Arterioscler Thromb Vasc Biol 2000 Sep;20(9):E41–52.
[32] Gudnason V, Day IN, Humphries SE. Effect on plasma lipid levels of different classes of mutations in the low-density lipoprotein receptor gene in patients with familial hypercholesterolemia. Arterioscler Thromb 1994 Nov;14(11):1717–22.
[33] Gaudet D, Vohl MC, Couture P, et al. Contribution of receptor negative versus receptor defective mutations in the LDL-receptor gene to angiographically assessed coronary artery disease among young (25–49 years) versus middle-aged (50–64 years) men. Atherosclerosis 1999 Mar;143(1):153–61.
[34] Tonstad S, Joakimsen O, Stensland-Bugge E, Ose L, Bonaa KH, Leren TP. Carotid intima-media thickness and plaque in patients with familial hypercholesterolaemia mutations and control subjects. Eur J Clin Invest 1998 Dec; 28(12):971–9.

[35] Dedoussis GV, Skoumas J, Pitsavos C, et al. FH clinical phenotype in Greek patients with LDL-R defective vs. negative mutations. Eur J Clin Invest 2004 Jun;34(6):402–9.
[36] Chaves FJ, Real JT, Garcia-Garcia AB, et al. Large rearrangements of the LDL receptor gene and lipid profile in a FH Spanish population. Eur J Clin Invest 2001 Apr;31(4):309–17.
[37] Koeijvoets KC, Wiegman A, Rodenburg J, Defesche JC, Kastelein JJ, Sijbrands EJ. Effect of low-density lipoprotein receptor mutation on lipoproteins and cardiovascular disease risk: a parent-offspring study. Atherosclerosis 2005 May;180(1):93–9.
[38] Graham CA, McClean E, Ward AJ, et al. Mutation screening and genotype: phenotype correlation in familial hypercholesterolaemia. Atherosclerosis 1999 Dec;147(2):309–16.
[39] Pimstone SN, Sun XM, Du Souich C, Frohlich JJ, Hayden MR, Soutar AK. Phenotypic variation in heterozygous familial hypercholesterolemia: a comparison of Chinese patients with the same or similar mutations in the LDL receptor gene in China or Canada. Arterioscler Thromb Vasc Biol 1998 Feb;18(2):309–15.
[40] Souverein OW, Defesche JC, Zwinderman AH, Kastelein JJ, Tanck MW. Influence of LDL-receptor mutation type on age at first cardiovascular event in patients with familial hypercholesterolaemia. Eur Heart J 2007 Feb;28(3):299–304.
[41] Oosterveer DM, Versmissen J, Yazdanpanah M, Defesche JC, Kastelein JJ, Sijbrands EJ. The risk of tendon xanthomas in familial hypercholesterolaemia is influenced by variation in genes of the reverse cholesterol transport pathway and the low-density lipoprotein oxidation pathway. Eur Heart J 2010 Apr;31(8):1007–12.
[42] Duskova L, Kopeckova L, Jansova E, et al. An APEX-based genotyping microarray for the screening of 168 mutations associated with familial hypercholesterolemia. Atherosclerosis 2011 May;216(1):139–45.

Atherosclerosis 216 (2011) 139–145

Contents lists available at ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

An APEX-based genotyping microarray for the screening of 168 mutations associated with familial hypercholesterolemia

Lucie Dušková[a], Lenka Kopečková[a], Eva Jansová[a], Lukáš Tichý[a], Tomáš Freiberger[b], Petra Zapletalová[a], Vladimír Soška[c], Barbora Ravčuková[b], Lenka Fajkusová[a,d,*]

[a] University Hospital Brno, Centre of Molecular Biology and Gene Therapy, Černopolní 9, CZ-62500 Brno, Czech Republic
[b] Centre for Cardiovascular Surgery and Transplantation, Pekařská 53, CZ-656 91 Brno, Czech Republic
[c] St. Anne's University Hospital Brno, Department of Clinical Biochemistry, Pekařská 53, CZ-656 91 Brno, Czech Republic
[d] Masaryk University, Faculty of Science, Institute of Experimental Biology, Department of Functional Genomics and Proteomics, Kamenice 5, CZ-62500 Brno, Czech Republic

ARTICLE INFO

Article history:
Received 4 November 2010
Received in revised form 3 January 2011
Accepted 10 January 2011
Available online 21 January 2011

Keywords:
Familial hypercholesterolemia
LDLR gene
Microarray
APEX reaction

ABSTRACT

Objective: Familial hypercholesterolemia (FH) is an inborn disorder of lipid metabolism characterised by elevated plasma concentrations of low-density lipoprotein cholesterol and total cholesterol. This imbalance results in accelerated atherosclerosis and premature coronary heart disease. The early identification and treatment of FH patients is extremely important because it leads to significant reduction of both coronary morbidity and mortality. FH is transmitted in an autosomal dominant manner and associated predominantly with mutations in the genes encoding the low-density lipoprotein receptor (*LDLR*) and its ligand apolipoprotein B (*APOB*). To date, more than 1000 sequence variants have been described in the *LDLR* gene. In marked contrast to *LDLR*, only one *APOB* mutation is prevalent in Europe.
Methods and results: The aim of this study was, on the basis of data obtained by the molecular genetic analysis of 1945 Czech FH probands, to propose, generate, and validate a new diagnostic tool, an APEX (Arrayed Primer EXtension)-based genotyping DNA microarray called the FH chip. The FH chip contains the *APOB* mutation p.Arg3527Gln, all 89 *LDLR* point mutations and small DNA rearrangements detected in Czech FH patients, and 78 mutations frequent in other European and Asian FH populations. The validation phase revealed the sensitivity and specificity of this platform, 100% and 99.1%, respectively.
Conclusions: This FH chip is a rapid, reproducible, specific, and cost-effective tool for genotyping, and in combination with MLPA (multiple ligation-dependent probe amplification) represents a reliable molecular genetic protocol for the large-scale screening of FH mutations in the Czech population.

1. Introduction

Familial hypercholesterolemia (FH) is an inborn disorder of lipid metabolism characterised by elevated plasma concentrations of low-density lipoprotein cholesterol and total cholesterol. This imbalance gives rise to tendon xanthomas, xanthelasmas, arcus lipoides corneae, accelerated atherosclerosis, and premature coronary heart disease. The early identification and subsequent treatment of FH patients with effective lipid-lowering agents is extremely important, as this has been shown to significantly reduce both coronary morbidity and mortality [1]. MedPed (Make early diagnoses to Prevent early deaths in medical pedigrees) is an international project joining together experts from more than 30 countries of the world working in the field of diagnosis and treatment of adults and children with high cholesterol disorders. In the Czech Republic, the project is coordinated by the Czech Atherosclerosis Society.

FH is transmitted in an autosomal dominant manner and associated with mutations in the genes encoding the low-density lipoprotein receptor (*LDLR*, MIM#143890) [2] and its ligand apolipoprotein B (*APOB*, MIM#144010) [3]. Certain gain-of-function mutations in the gene encoding the proprotein convertase subtilisin kexin 9 (*PCSK9*, MIM#603766) also lead to FH, while other *PCSK9* mutations can cause hypocholesterolemia [4]. The frequency of heterozygous FH in most populations is about 1/500, while homozygous FH is rare (≤1/1,000,000). The *LDLR* gene is located on chromosome 19p13.2, spans 45 kb, contains 18 exons, and encodes a 160 kDa mature protein which regulates the uptake and degradation of LDL particles by hepatic cells [5]. To date, more than 1000 unique sequence variants have been identified in *LDLR*

* Corresponding author at: Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Černopolní 9, CZ-62500 Brno, Czech Republic.
Tel.: +420 549494003; fax: +420 549492654.
E-mail address: lenkafajkusova@volny.cz (L. Fajkusová).

0021-9150/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.atherosclerosis.2011.01.023

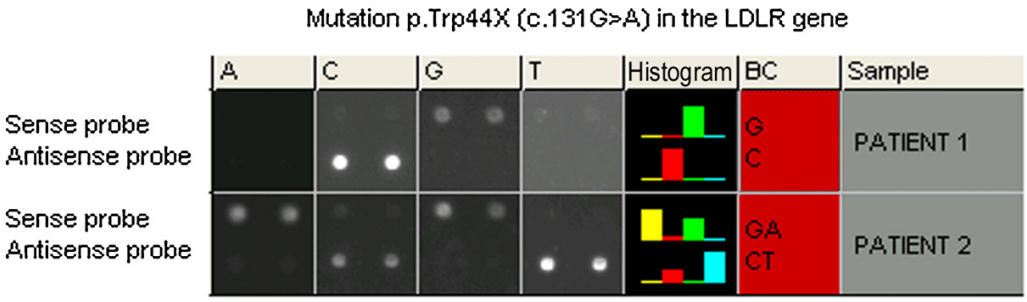


Fig. 1. Microarray APEX analysis of the mutation p.Trp44X (c.131G>A) in the *LDLR* gene using Genorama genotyping software. Signal intensities of the sense and antisense probe for the mutation p.Trp44X are detected in four channels which relate to fluorescently labeled dideoxynucleotides A, C, G, T. Signals of the sense and antisense probe are visualized in duplicate. The letters to the right of the histogram indicate the bases identified on each strand. Histogram for patient 1 represents the wild type reference for the mutation p.Trp44X, whereas patient 2 carries the mutant allele (additional signal in the A channel of the sense probe and in the T channel of the antisense probe).

(database website http://www.ucl.ac.uk/fh), including 65% of point mutations, 24% of small DNA rearrangements (<100 bp) and 11% of large DNA rearrangements (>100 bp). The detection rate of *LDLR* and *APOB* mutations varies from 30% to 60% of the total number of patients estimated to be FH, and depends mainly on the clinical criteria used for diagnosis [6].

The molecular genetic analysis of the *LDLR* gene utilizes mainly PCR-RFLP (restriction fragment length polymorphism) for the detection of the most common mutations, PCR-direct sequencing for the mutation scanning, and MLPA (multiplex ligation-dependent probe amplification) for the detection of large gene rearrangements. However, the molecular genetic analysis of the whole *LDLR* coding region is highly time-consuming considering the number of patients with a suspicion of FH. In the years 2000–2009, we performed molecular genetic analysis of 1945 unrelated Czech FH patients: 252 patients (13.0%) had the *APOB* mutation p.Arg3527Gln and 443 patients (22.8%) had an *LDLR* mutation. On the basis of these data, we proposed a new diagnostic tool, an APEX (Arrayed Primer EXtension)-based genotyping DNA microarray called the FH chip, to speed up mutation screening in Czech FH patients. We have selected a panel of 90 mutations detected in Czech FH patients and 78 mutations frequent in other populations for this microarray [7–30]. In APEX, mutation-specific probes immobilized on a chip are extended by one fluorescently labeled ddNTP (the site of the mutation) according to the sequence of a co-hybridized DNA fragment [31]. The FH chip was validated using DNA samples with known genotypes, and at the present time is used as the first step in molecular diagnostics of FH in our patients.

2. Materials and methods

2.1. Subjects

One thousand nine hundred and forty five unrelated patients with the diagnosis of FH, submitted to the database of the MedPed project in the Czech Republic, were analysed for the presence of mutations in the *APOB* and *LDLR* genes. The experimental research reported in this study has been performed with the approval of the Ethical Committee of the General University Hospital in Prague, Czech Republic, and all patients gave their informed consent with their participation in the study. The cohort of patients in our study included (1) 1552 patients with untreated total and/or LDL cholesterol serum levels above the 95th percentile of age, sex and population-specific values (group 1); (2) 393 patients with elevated total and/or LDL cholesterol in serum but untreated levels unavailable or not exceeding the 95th percentile of age, sex and population-specific values, and, in addition, with high clinical suspicion of FH based on personal history and/or family history of premature coronary heart disease and/or elevated total and LDL cholesterol serum levels in the first degree relatives (group 2).

2.2. Strategy for the identification of mutation spectrum in Czech FH patients

DNA analysis of FH patients was divided into several consecutive steps: (1) PCR-RFLP detection of the most common mutation in the *APOB* gene (p.Arg3527Gln); (2) PCR-RFLP detection of the most common mutations in the *LDLR* gene (p.Gly592Glu, p.Asp266Glu, and p.Arg416Trp); (3) PCR-sequencing of *LDLR* exon 4 (the exon with the most frequent occurrence of mutations in Czech FH patients); (4) MLPA analysis of all *LDLR* exons; (5) PCR-sequencing of the promoter and *LDLR* exons 1, 5, 6, 9, 10, 12, 14; and (6) PCR-DHPLC (denaturing high performance liquid chromatography) of *LDLR* exons 2, 3, 7, 8, 11, 13, 15, 16, 17, and 18, followed by sequencing of regions which tested positively. Further DNA analysis in cases where a mutation was found depended on the personal and family history of hypercholesterolemia; it continued in cases when (1) a phenotypic manifestation could be associated with the presence of two FH mutations or (2) a detected missense mutation was new with an effect on the protein structure and function which was hard to predict. This diagnostic process is common in FH molecular genetic testing worldwide.

2.3. Oligonucleotide probes and the FH chip

Two specific oligonucleotide probes (sense and antisense) were designed for every mutation selected for detection by the FH microarray according to the *LDLR* sequence (http://www.ucl.ac.uk/fh). These oligonucleotides have 3′OH ends immediately adjacent to target nucleotides, which can be defined as those at which the mutations were found. The melting temperatures and self-annealing secondary structures were considered during design of oligonucleotide probes. One or two mismatches were incorporated into the 5′ ends or internal parts of some probes with a high potential for secondary structure formation. The melting temperature of the oligonucleotide probes ranges from 61.5 °C to 65.5 °C, and the length varies from 18 to 34 bases. In a few cases, the same probe determines two or more different *LDLR* mutations. In total, there are 286 different specific oligonucleotide probes (143 sense and 143 antisense) and 4 different self-elongating probes (serving as internal positive controls of the APEX reaction) on the microarray. These probes enable the detection of 168 mutations associated with FH. Oligonucleotide probes were synthesized with 5-prime 6-carbon amino linkers (Metabion International AG, Martinsried, Germany), diluted to

Table 1
Mutations in *LDLR* and *APOB* genes identified in Czech FH patients.

No. of mutation	Gene	Mutation (protein level)	Mutation (cDNA level)	Exon (intron)
1	LDLR		c.-120C>T	Promoter
2	LDLR		c.-153C>T	Promoter
3	LDLR	p.Gly20Arg	c.58G>A	1
4	LDLR	p.Cys34Gly	c.100T>G	2
5	LDLR	p.Trp44X	c.131G>A	2
6	LDLR	p.Cys82X	c.246C>T	3
7	LDLR	p.Lys107GlnfsX23	c.318_319insC	4
8	LDLR	p.Cys109Arg	c.326G>A	4
9	LDLR	p.Cys116_Ile122del	c.347_367del	4
10	LDLR	p.Ser123fsX6	c.369_370delTC	4
11	LDLR	p.Ser130Pro	c.388T>C	4
12	LDLR	p.Glu140Asp	c.420G>T	4
13	LDLR	p.Cys143Arg	c.427T>C	4
14	LDLR	p.Gly149Cys	c.445G>T	4
15	LDLR	p.Ser158ProfsX48	c.472delT	4
16	LDLR	p.Cys167X	c.501C>A	4
17	LDLR	p.Asp172Gly	c.515A>G	4
18	LDLR	p.Gly176Val	c.527G>T	4
19	LDLR	p.Ser177Leu	c.530C>T	4
20	LDLR	p.Pro181Arg	c.542C>G	4
21	LDLR	p.Asp193AlafsX13	c.578delA	4
22	LDLR	p.Cys209fsX57	c.625_626dupTG	4
23	LDLR	p.Cys209Tyr	c.626G>A	4
24	LDLR	p.Glu228Lys	c.628G>A	4
25	LDLR	p.Gly218del	c.652_654delTGG	4
26	LDLR	p.Asp221Gly	c.662A>G	4
27	LDLR	p.Ser226Pro	c.676T>C	4
28	LDLR	p.Asp227Glu	c.681C>G	4
29	LDLR	p.Cys231Gly	c.691T>G	4
30	LDLR	p.Cys231X	c.693C>A	4
31	LDLR	p.Gln254Pro	c.761A>C	5
32	LDLR	p.Asp266Asn	c.796G>A	5
33	LDLR	p.Asp266Glu	c.798T>A	5
34	LDLR	p.Cys270Arg	c.808T>C	5
35	LDLR	p.Asn272Thr	c.815A>C	5
36	LDLR	splicing defect	c.817+1G>A	Intron 5
37	LDLR	p.Cys276X	c.828C>A	6
38	LDLR	p.Phe282Leu	c.846C>T	6
39	LDLR	p.Glu288X	c.862G>T	6
40	LDLR	p.Lys294SerfsX6	c.880_881delAA	6
41	LDLR	p.Ala299ValfsX71	c.896delC	6
42	LDLR	p.Asp304Asn	c.910G>A	6
43	LDLR	p.Ser306X	c.917C>G	6
44	LDLR	p.Asp307Asn	c.919G>A	6
45	LDLR	p.Cys313X	c.939C>A	6
46	LDLR	p.Asn316Thr	c.947A>C	7
47	LDLR	p.Glu317X	c.949G>T	7
48	LDLR	p.Gly324AlafsX46	c.970delG	7
49	LDLR	p.Cys338Tyr	c.1013G>A	7
50	LDLR	p.Asp342Tyr	c.1024G>T	7
51	LDLR	p.Arg350X	c.1048C>T	7
52	LDLR	p.Asp354Ala	c.1061A>C	8
53	LDLR	p.Asp354fsX19	c.1053_1060dup	8
54	LDLR	p.Cys364Ser	c.1091G>C	8
55	LDLR	p.Gln378Pro	c.1133A>C	8
56	LDLR	p.Gln384Pro387delinsPro	c.1151_1159del9	8
57	LDLR	p.Phe403del	c.1205_1207delTCT	9
58	LDLR	p.Glu408Val	c.1223A>T	9
59	LDLR	p.Arg416Trp	c.1246C>T	9
60	LDLR	p.Arg416Pro	c.1247G>C	9
61	LDLR	p.Pro424_Asn425ins32	c.1272_1273ins96	9
62	LDLR	splicing defect	c.1358+2T>A	Intron 9
63	LDLR	p.Asp472Tyr	c.1414G>T	10
64	LDLR	p.Asp482Asn	c.1444G>A	10
65	LDLR	p.Asp492Asn	c.1474G>A	10
66	LDLR	p.Arg513Lys	c.1538G>A	10
67	LDLR	p.Lys518Glu	c.1552A>G	10
68	LDLR	p.Val523Met	c.1567G>A	10
69	LDLR	p.Thr557CysfsX3	c.1662_1669dup	11
70	LDLR	p.Glu558X	c.1672G>T	11
71	LDLR	p.Ser584fsX81	c.1751delC	12
72	LDLR	p.Gly592Glu	c.1775G>A	12
73	LDLR	p.Pro608Ser	c.1822C>T	12
74	LDLR	p.Ser610Cys	c.1829C>G	12
75	LDLR	splicing defect	c.1845+1G>A	Intron 12
76	LDLR	splicing defect	c.1845+1G>T	Intron 12

Table 1 (*Continued*)

No. of mutation	Gene	Mutation (protein level)	Mutation (cDNA level)	Exon (intron)
77	LDLR	p.Trp666X	c.1998G>A	14
78	LDLR	p.Cys667Arg	c.1999T>C	14
79	LDLR	p.Gly675Ser	c.2023G>A	14
80	LDLR	p.Pro685Leu	c.2054C>T	14
81	LDLR	p.His690ThrfsX19	c.2064delC	14
82	LDLR	p.Cys698Phe	c.2093G>T	14
83	LDLR	p.Pro699Leu	c.2096C>T	14
84	LDLR	p.Gly701Ser	c.2101G>A	14
85	LDLR	p.Val797Met	c.2389G>A	16
86	LDLR	splicing defect	c.2390-1G>A	Intron 16
87	LDLR	p.Leu799Arg	c.2396T>G	17
88	LDLR	p.Val806GlyfsX11	c.2416_2417insG	17
89	LDLR	p.Val827Ile	c.2479G>A	17
90	APOB	p.Arg3527Gln	c.10580G>A	26

50 μM in 100 mM carbonate buffer (pH 9.0), and spotted (each in duplicate) onto the microarray slide coated with 3-aminopropyl-trimethoxysilane plus 1,4-phenylenedi-isothiocyanate with a BioRad VersArray (BioRad Laboratories, Hercules, CA). The spotted FH chips with dimensions of 24 mm × 60 mm × 0.15 mm were blocked with 1% ammonia solution and stored at 4 °C. The FH microarrays were produced by Asper Biotech Ltd., Tartu, Estonia.

2.4. Template preparation for the FH chip

The ability of the FH chip to discriminate between mutant and wild-type alleles of the *LDLR* gene was tested during the validation phase. DNA samples of FH patients with known genotypes were used for validation of 89 *LDLR* mutations and one *APOB* mutation whose detection is possible on the FH chip. For validation of 78 mutations frequent in other populations, mutated primers were designed and PCR products carrying the analysed mutations were generated to test the correct function of microarray probes.

One exon of *APOB*, 15 exons of *LDLR*, and the *LDLR* promoter (all harboring the analysed mutations) were amplified in 4 independent PCRs using primers listed in Supplementary File. Amplification reactions comprised 250 ng of DNA, 0.4–1.6 μmol/l primers (Invitrogen, CA, USA), 0.2 mmol/l dNTPs with 20% of deoxyuridine triphosphate (dUTP) (Roche, Basel, Switzerland), 1.5 mmol/l Mg^{2+}, 2.5 units of *Taq* polymerase, and 1× *Taq* polymerase buffer (Fermentas International, Vilnius, Lithuania) in a total volume of 50 μl. The multiplex PCRs were performed with the initial denaturation step at 94 °C for 7 min, followed by 30 cycles at 94 °C for 30 s, 58 or 62 °C for 60 s, 72 °C for 30 s, and final extension at 72 °C for 7 min. Multiplex PCR products were pooled and purified by QIAquick PCR Purification kits (Qiagen GmBH, Hilden, Germany) into a volume of 40 μl. Fragmentation of PCR products was performed by 1.5 units of Thermolabile uracil N-glycosylase (Epicentre Technologies, Madison, WI) and degradation of unincorporated dNTPs by 1 unit of Shrimp alkaline phosphatase (GE Healthcare UK Ltd., Amersham, UK) at 37 °C for 1 h followed by heat denaturation at 95 °C for 30 min. The fragmentation of target DNA was checked on a 2% agarose gel.

2.5. APEX reaction

Before the APEX reaction, the FH chip was washed in ddH_2O at 95 °C and equilibrated at 58 °C for 15 min. The APEX mix contained 35 μl of fragmented and denatured PCR products, 100 pmol of Cy5-ddUTP, 100 pmol of FITC-ddGTP, 100 pmol of TexasRed-5-ddATP, 115 pmol of Cy3-ddCTP (Asper Biotech Ltd., Tartu, Estonia), 5 units of Thermo Sequenase DNA polymerase, and 1× Thermo Sequenase reaction buffer (GE Healthcare UK Ltd., Amersham, UK) in a total volume of 50 μl. The APEX mixture was applied on the pre-warmed FH chip, covered by a LifterSlip (Thermo Scientific, Waltham, USA), and incubated at 58 °C for 25 min. After the APEX reaction, the microarray was washed once for 3 min in 0.3% Alcanox solution (Sigma–Aldrich, St. Louis, USA) at room temperature, and 3 times in ddH_2O at 95 °C. The chip was scanned under 12 μl of SlowFade Antifade solution (Invitrogen, CA, USA), covered by a cover glass, with a Genorama™ QuattroImager 003 (Genorama Ltd., Tartu, Estonia) [32]. Sequence variants were determined using Genorama™ 4.5 genotyping software (Genorama Ltd., Tartu, Estonia). The APEX reaction is simple, based on Sanger's dideoxysequencing. One of 4 terminators (ddATP, ddCTP, ddGTP, and ddUTP) labeled with a different fluorescent dye is incorporated into the 3′OH end of each immobilized probe according to the complementary sequence of the hybridized DNA fragment. Sequence variants are determined according to which ddNTP is linked to each probe [32] (Fig. 1).

3. Experimental results

In the set of 1945 probands with the clinical diagnosis of FH, we detected 252 patients (13%) with the *APOB* mutation p.Arg3527Gln, of whom 206 (13.3%) belonged to the group 1 and 46 (11.7%) to the group 2. The rest of the patients (1693) were analysed for the presence of mutations in the *LDLR* gene: 406 patients (20.9%) carried a point mutation or a small DNA rearrangement, and 37 patients (1.9%) had a large DNA rearrangement. With regard to the clinical diagnosis of FH, *LDLR* gene mutations were detected in 406 patients of group 1 (26.2%) and 37 patients of group 2 (9.4%). On the basis of these results, we designed a genotyping DNA microarray for the simultaneous screening of known *LDLR* and *APOB* mutations to speed up the molecular genetic diagnostics of FH.

The present version of the FH chip contains the *APOB* mutation p.Arg3527Gln, all *LDLR* point mutations and small DNA rearrangements detected in Czech FH patients (89 types, Table 1) and 78 *LDLR* mutations frequent in other European and Asian FH populations (Table 2).

After the first validation phase, DNA analysis of 90 Czech patients carrying all *APOB* and *LDLR* mutations detected in the Czech FH population and 78 *LDLR*-mutated DNA samples carrying mutations frequent in other populations, 21 oligonucleotide probes (from the total number of 286) provided no or a weak fluorescence signal and so were redesigned. After the second validation phase, the analysis of DNA samples carrying mutations tested by the redesigned probes, 6 probes still remain with negative signals (c.1-206C>T_sense, p.G20R_sense, p.Ser130Pro_antisense, p.Cys143X_sense, p.C167X_sense, p.Asp386Glu_sense). Despite this, the appropriate mutations are reliably determined by the pairing probe on the microarray. The fact that some oligonucleotide probes do not provide a fluorescent signal is likely caused by formation of secondary structures and/or the presence of repetitive sequences [33].

Table 2
LDLR mutations frequent in different populations selected for the FH chip.

No. of mutation	Gene	Mutation (protein level)	Mutation (cDNA level)	Exon (intron)	Reference
1	LDLR		c.1-206C>T	Promoter	[7,8]
2	LDLR		c.1-156C>T	Promoter	[7]
3	LDLR		c.1-142C>T	Promoter	[7]
4	LDLR		c.1-138T>C	Promoter	[7]
5	LDLR		c.1-138delT	Promoter	[7]
6	LDLR		c.1-137C>T	Promoter	[7]
7	LDLR		c.1-136C>T	Promoter	[7]
8	LDLR	p.Met1Val	c.-21A>G	1	[9]
9	LDLR	p.Cys27Trp	c.81C>G	2	[7]
10	LDLR	p.Cys27X	c.81C>A	2	[10]
11	LDLR	p.Glu31X	c.91G>T	2	[7]
12	LDLR	p.Gln33X	c.97C>T	2	[11]
13	LDLR	p.Cys39MetfsX13	c.114_115insA	2	[7,12]
14	LDLR	slicing defect	c.191-2A>G	Intron 2	[13]
15	LDLR	p.Trp87Gly	c.259T>G	3	[7,14]
16	LDLR	p.Ser99X	c.296C>G	3	[7,8,15]
17	LDLR	p.Glu101Lys	c.301G>A	3	[7]
18	LDLR	p.Glu101X	c.301G>T	3	[7]
19	LDLR	p.Glu101SerfsX115	c.301delG	3	[7]
20	LDLR	splicing defect	c.313+1G>A	Intron 3	[7,9]
21	LDLR	splicing defect	c.313+1G>C	Intron 3	[7,9]
22	LDLR	splicing defect	c.313+1G>T	Intron 3	[16]
23	LDLR	splicing defect	c.313+2T>C	Intron 3	[7,13]
24	LDLR	p.Cys116Arg	c.346T>C	4	[7]
25	LDLR	p.Ser123Thr	c.367T>A	4	[17]
26	LDLR	p.Phe126Tyr	c.377T>A	4	[17]
27	LDLR	p.Glu140Lys	c.418G>A	4	[18]
28	LDLR	p.Glu140X	c.418G>T	4	[19]
29	LDLR	p.Cys143X	c.429C>A	4	[7]
30	LDLR	p.Cys148Arg	c.442T>C	4	[7]
31	LDLR	p.Phe200Leu	c.600C>A	4	[9]
32	LDLR	p.Phe200LeufsX199	c.600delC	4	[13]
33	LDLR	p.Glu228X	c.682G>T	4	[7,9,11]
34	LDLR	p.Glu228Gln	c.682G>C	4	[20]
35	LDLR	p.Ser286Arg	c.858C>A	6	[7,17]
36	LDLR	p.Ser306Leu	c.917C>T	6	[13]
37	LDLR	p.Pro309LysfsX49	c.925_931delCCCATCA	6	[7,11]
38	LDLR	p.Lys311Arg	c.932A>G	6	[7,13]
39	LDLR	p. Lys311fsX59	c.933delA	6	[7]
40	LDLR	p.Lys311ArgfsX20	c.932_933delAA	6	[21]
41	LDLR	p.Cys313Trp	c.939C>G	6	[7]
42	LDLR	p.Cys329Gly	c.985T>G	7	[22]
43	LDLR	p.Cys329Tyr	c.986G>A	7	[11,13,23]
44	LDLR	p.Cys338Ser	c.1012T>A	7	[18]
45	LDLR	p.Cys338Arg	c.1012T>C	7	[24]
46	LDLR	p.Cys338Gly	c.1012T>G	7	[25]
47	LDLR	p.Gln366Pro	c.1097A>C	8	[26]
48	LDLR	p.Gln366Arg	c.1097A>G	8	[7,8]
49	LDLR	p.Gly373Asp	c.1118G>A	8	[19]
50	LDLR	p.Gly373Val	c.1118G>T	8	[16]
51	LDLR	p.Cys379Tyr	c.1136G>A	8	[12]
52	LDLR	p.Gln384X	c.1150C>T	8	[7,27]
53	LDLR	p.Asp386Glu	c.1158C>G	8	[7,27]
54	LDLR	p.Ler401His	c.1202T>A	9	[7,11]
55	LDLR	p.Leu414Arg	c.1241T>G	9	[23]
56	LDLR	p.Ile423Thr	c.1268T>C	9	[7]
57	LDLR	p.Asp433His	c.1297G>C	9	[28,29]
58	LDLR	p.Asp433Tyr	c.1297G>T	9	[25]
59	LDLR	p.T434M	c.1301C>T	9	[17]
60	LDLR	p.T434K	c.1301C>A	9	[17]
61	LDLR	p.T434R	c.1301C>G	9	[12,13]
62	LDLR	splicing defect	c.1359-1G>A	Intron 9	[7,13]
63	LDLR	p.Ile473fsX64	c.1415_1418dup	10	[7]
64	LDLR	p.Asp482His	c.1444G>C	10	[7]
65	LDLR	p.Gly549Asp	c.1646G>A	11	[7]
66	LDLR	p.Gly549Val	c.1646G>T	11	[12,14]
67	LDLR	p.Asn564Asp	c.1690A>G	11	[7,13,20]
68	LDLR	p.Asn564His	c.1690A>C	11	[14]
69	LDLR	splicing defect	c.1845+2T>C	Intron 12	[18]
70	LDLR	p.Trp666Leu	c.1997G>T	14	[7]
71	LDLR	p.Cys681X	c.2043C>A	14	[7]
72	LDLR	p.Cys681Trp	c.2043C>G	14	[30]
73	LDLR	p.Cys698Tyr	c.2093G>A	14	[29]
74	LDLR	splicing defect	c.2312-3C>A	Intron 15	[28]

Table 2 (*Continued*)

No. of mutation	Gene	Mutation (protein level)	Mutation (cDNA level)	Exon (intron)	Reference
75	LDLR	p.Val797Leu	c.2389G>T	16	[25]
76	LDLR	p.Lys811X	c.2431A>T	17	[18,28]
77	LDLR	p.Asn825Lys	c.2475C>A (G)	17	[12]
78	LDLR	p.Gly844Asp	c.2531G>A	17	[7]

The sensitivity of the FH chip to detect selected mutations was verified by analyzing 168 DNA samples with a known genotype. The overall sensitivity for all tested mutations (true positive/true positive + false negative) was 100%. The specificity of the FH chip (true negative/true negative + false positive) was 99.1%. Two false positives were found in a total of 230 true negative samples. A blank sample (without DNA) was used to test each FH microarray batch for unspecific signals from probes. False positives can be explained by the use of the genotyping software – when the signal is intermediate between negative and positive, the software can scores this as positive.

Further, we performed a blind screening of 237 patients clinically diagnosed as FH but with unknown genotype. Using the FH chip and MLPA, a mutation was detected in 63 of them (26.6%): 16 patients (6.8%) carried the *APOB* mutation, 39 patients (16.5%) a *LDLR* mutation, and 8 patients (3.4%) a large gene deletion or duplication. Every nucleotide change determined on the FH chip was verified by direct sequencing to confirm the accuracy of the chip. In patients with no identified mutation on the FH chip or by MLPA, DNA analysis of the whole coding region of *LDLR* (sequencing and DHPLC) continues.

4. Discussion

The purpose of this study was (1) to develop a DNA genotyping microarray for the detection of mutations associated with FH in the Czech population, (2) to evaluate the specificity and sensitivity of the FH chip using previously genotyped DNA samples, and (3) to apply this platform for the identification of new FH patients. *LDLR* mutations spread throughout the whole gene and the molecular genetic testing of 18 exons and the promotor area is time-consuming, especially considering the amount of patients with possible FH diagnosis.

In 2005, Tejedor and co-workers developed a DNA microarray called the Lipochip which enables the detection of mutations causing FH in the Spanish population [34]. The Lipochip has been gradually extended and at the present time it enables the detection of about 250 mutations in the *LDLR*, *APOB*, and *PCSK9* genes (http://www.progenika.com). In the Lipochip, each mutation is analysed by two pairs of oligonucleotide probes specific for the wild-type and the mutant sequence. Target bases are always located in central positions of oligonucleotide probes. The ability of allele-specific probes to differentiate the sequence variants depend on hybridization conditions, nucleotide sequences which flank a site of mutation, and secondary structures of probes. The spectrum of selected *LDLR* mutations on the Lipochip and the microarray technology and methodology were different from those in our study. We compared 191 *LDLR* mutations present on the Lipochip and published by Alonso et al. [35] with *LDLR* mutations incident in the Czech FH population, and found that only 17 mutations were identical for both groups.

We therefore prepared and subsequently validated a genotyping DNA microarray which is suitable for the detection of mutations associated with FH in the Czech population. There are 90 mutations detected in Czech FH patients and 78 mutations frequent in European and Asian FH populations on the FH chip. The FH chip is able to detect mutations with high sensitivity and specificity in a short time period (8 h including the template preparation phase) and with a low cost (29 Euro for the chip). It is possible to extend the FH chip by addition of new oligonucleotide probes for the screening of other mutations at anytime. At the other side, the APEX-based microarray has some limitations – it is able to detect only known mutations (whose detection probes are spotted on the chip) and is not able to detect large gene rearrangements. For these reasons, patients without a detected mutation on the FH chip are subsequently analysed using MLPA (detection of large gene deletions and duplications). In patients with no identified mutation on the FH chip or by MLPA, DNA analysis of the whole coding region of *LDLR* (sequencing and DHPLC) continues.

The APEX-based microarray is a robust technique for screening of known mutations, and at the present time is used in many genetic tests (http://www.asperbio.com/genetic-tests); for example for the detection of mutations in the *ATP7B* gene [36], in the *CFTR* gene [37], and in the *TP53* gene [33]. Total costs and the time of analysis are important arguments for APEX-based large-scale population screening. FH diagnosis is usually made on the basis of clinical and laboratory findings which are often obtained later in life, and the determination of diagnosis in younger patients can be difficult. Genetic analysis provides unequivocal diagnosis and therefore the initiation of treatment as soon as possible. The FH chip represents a reproducible, sensitive, and specific tool for the screening of mutations and together with MLPA can become the basic step in molecular genetic diagnostics of FH in the Czech population.

Conflict of interest

None.

Acknowledgements

We would like to thank the physicians from the national (M. Vrablík, Prof. R. Češka) and regional centres of the Czech MedPed project (V. Bláha, M. Budíková, J. Buryška, R. Cífková, L. Dlouhý, L. Dostálová-Kopečná, H. Halámková, J. Hyánek, J. Hyjánek, Z. Krejsová, J. Macháček, Š. Malá, J. Malý, V. Miláček, J. Mraček, H. Podzimková, D. Povalačová, H. Rosolová, F. Stožický, E. Šipková, L. Ťoukálková, Z. Urbanová, H. Vaverková, S. Zemek, A. Žák) and other physicians participating in this project for providing us with their patients' blood samples and clinical and laboratory data, and J. Porupková and M. Plotěná for technical assistance. This work was funded by the Czech Ministry of Education (projects 2B08060 and MSM0021622415).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.atherosclerosis.2011.01.023.

References

[1] Neil A, Cooper J, Betteridge J, et al. Reductions in all-cause, cancer, and coronary mortality in statin-treated patients with heterozygous familial hypercholesterolaemia: a prospective registry study. Eur Heart J 2008;29:2625–33.
[2] Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. Hum Mutat 1992;1:445–66.

[3] Soria LF, Ludwig EH, Clarke HR, et al. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc Natl Acad Sci U S A 1989;86:587–91.
[4] Abifadel M, Varret M, Rabes JP, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 2003;34:154–6.
[5] Goldstein JL, Hobbs HH, Brown MS. Familial hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. The metabolic and molecular bases of inherited disease. New York: McGraw-Hill Book Co; 2001. p. 2863–913.
[6] Taylor A, Wang D, Patel K, et al. Mutation detection rate and spectrum in familial hypercholesterolaemia patients in the UK pilot cascade project. Clin Genet 2010;77:572–80.
[7] Dedoussis GV, Schmidt H, Genschel J. LDL-receptor mutations in Europe. Hum Mutat 2004;24:443–59.
[8] Lind S, Rystedt E, Eriksson M, et al. Genetic characterization of Swedish patients with familial hypercholesterolemia: a heterogeneous pattern of mutations in the LDL receptor gene. Atherosclerosis 2002;163:399–407.
[9] Nauck MS, Koster W, Dorfer K, et al. Identification of recurrent and novel mutations in the LDL receptor gene in German patients with familial hypercholesterolemia. Hum Mutat 2001;18:165–6.
[10] Laurie AD, Scott RS, George PM. Genetic screening of patients with familial hypercholesterolaemia (FH): a New Zealand perspective. Atheroscler Suppl 2004;5:13–5.
[11] Zakharova FM, Damgaard D, Mandelshtam MY, et al. Familial hypercholesterolemia in St-Petersburg: the known and novel mutations found in the low density lipoprotein receptor gene in Russia. BMC Med Genet 2005;6:6.
[12] Garcia-Garcia AB, Real JT, Puig O, et al. Molecular genetics of familial hypercholesterolemia in Spain: ten novel LDLR mutations and population analysis. Hum Mutat 2001;18:458–9.
[13] Fouchier SW, Defesche JC, Umans-Eckenhausen MW, Kastelein JP. The molecular basis of familial hypercholesterolemia in The Netherlands. Hum Genet 2001;109:602–15.
[14] Jensen HK. The molecular genetic basis and diagnosis of familial hypercholesterolemia in Denmark. Dan Med Bull 2002;49:318–45.
[15] Leren TP, Tonstad S, Gundersen KE, et al. Molecular genetics of familial hypercholesterolaemia in Norway. J Intern Med 1997;241:185–94.
[16] Graham CA, McIlhatton BP, Kirk CW, et al. Genetic screening protocol for familial hypercholesterolemia which includes splicing defects gives an improved mutation detection rate. Atherosclerosis 2005;182:331–40.
[17] Mihaylov VA, Horvath AD, Savov AS, et al. Screening for point mutations in the LDL receptor gene in Bulgarian patients with severe hypercholesterolemia. J Hum Genet 2004;49:173–6.
[18] Maruyama T, Miyake Y, Tajima S, et al. Common mutations in the low-density-lipoprotein-receptor gene causing familial hypercholesterolemia in the Japanese population. Arterioscler Thromb Vasc Biol 1995;15:1713–8.
[19] Bertolini S, Cantafora A, Averna M, et al. Clinical expression of familial hypercholesterolemia in clusters of mutations of the LDL receptor gene that cause a receptor-defective or receptor-negative phenotype. Arterioscler Thromb Vasc Biol 2000;20:E41–52.
[20] Fouchier SW, Kastelein JJ, Defesche JC. Update of the molecular basis of familial hypercholesterolemia in The Netherlands. Hum Mutat 2005;26:550–6.
[21] Nissen H, Lestavel S, Hansen TS, et al. Mutation screening of the LDLR gene and ApoB gene in patients with a phenotype of familial hypercholesterolemia and normal values in a functional LDL receptor/apolipoprotein B assay. Clin Genet 1998;54:79–82.
[22] Amsellem S, Briffaut D, Carrie A, et al. Intronic mutations outside of Alu-repeat-rich domains of the LDL receptor gene are a cause of familial hypercholesterolemia. Hum Genet 2002;111:501–10.
[23] Mak YT, Pang CP, Tomlinson B, et al. Mutations in the low-density lipoprotein receptor gene in Chinese familial hypercholesterolemia patients. Arterioscler Thromb Vasc Biol 1998;18:1600–5.
[24] Maruyama T, Yamashita S, Matsuzawa Y, et al. Mutations in Japanese subjects with primary hyperlipidemia—results from the Research Committee of the Ministry of Health and Welfare of Japan since 1996. J Atheroscler Thromb 2004;11:131–45.
[25] Lombardi MP, Redeker EJ, Defesche JC, et al. Molecular genetic testing for familial hypercholesterolemia: spectrum of LDL receptor gene mutations in The Netherlands. Clin Genet 2000;57:116–24.
[26] Damgaard D, Larsen ML, Nissen PH, et al. The relationship of molecular genetic to clinical diagnosis of familial hypercholesterolemia in a Danish population. Atherosclerosis 2005;180:155–60.
[27] Humphries SE, Cranston T, Allen M, et al. Mutational analysis in UK patients with a clinical diagnosis of familial hypercholesterolaemia: relationship with plasma lipid traits, heart disease risk and utility in relative tracing. J Mol Med 2006;84:203–14.
[28] Yu W, Nohara A, Higashikata T, et al. Molecular genetic analysis of familial hypercholesterolemia: spectrum and regional difference of LDL receptor gene mutations in Japanese population. Atherosclerosis 2002;165:335–42.
[29] Mozas P, Castillo S, Tejedor D, et al. Molecular characterization of familial hypercholesterolemia in Spain: identification of 39 novel and 77 recurrent mutations in LDLR. Hum Mutat 2004;24:187.
[30] Chiu CY, Wu YC, Jenq SF, Jap TS. Mutations in low-density lipoprotein receptor gene as a cause of hypercholesterolemia in Taiwan. Metabolism 2005;54:1082–6.
[31] Tonisson N, Zernant J, Kurg A, et al. Evaluating the arrayed primer extension resequencing assay of TP53 tumor suppressor gene. Proc Natl Acad Sci U S A 2002;99:5503–8.
[32] Kurg A, Tonisson N, Georgiou I, et al. Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. Genet Test 2000;4:1–7.
[33] Kringen P, Bergamaschi A, Due EU, et al. Evaluation of arrayed primer extension for TP53 mutation detection in breast and ovarian carcinomas. Biotechniques 2005;39:755–61.
[34] Tejedor D, Castillo S, Mozas P, et al. Reliable low-density DNA array based on allele-specific probes for detection of 118 mutations causing familial hypercholesterolemia. Clin Chem 2005;51:1137–44.
[35] Alonso R, Defesche JC, Tejedor D, et al. Genetic diagnosis of familial hypercholesterolemia using a DNA-array based platform. Clin Biochem 2009;42:899–903.
[36] Gojova L, Jansova E, Kulm M, Pouchla S, Kozak L. Genotyping microarray as a novel approach for the detection of ATP7B gene mutations in patients with Wilson disease. Clin Genet 2008;73:441–52.
[37] Schrijver I, Oitmaa E, Metspalu A, Gardner P. Genotyping microarray for the detection of more than 200 CFTR mutations in ethnically diverse populations. J Mol Diagn 2005;7:375–87.

Goldmann et al. BMC Medical Genetics 2010, 11:115
http://www.biomedcentral.com/1471-2350/11/115

BMC Medical Genetics

RESEARCH ARTICLE Open Access

Genomic characterization of large rearrangements of the *LDLR* gene in Czech patients with familial hypercholesterolemia

Radan Goldmann[1], Lukáš Tichý[1], Tomáš Freiberger[2], Petra Zapletalová[1], Ondřej Letocha[1], Vladimír Soška[3], Jiří Fajkus[4], Lenka Fajkusová[1,4*]

Abstract

Background: Mutations in the *LDLR* gene are the most frequent cause of Familial hypercholesterolemia, an autosomal dominant disease characterised by elevated concentrations of LDL in blood plasma. In many populations, large genomic rearrangements account for approximately 10% of mutations in the *LDLR* gene.

Methods: DNA diagnostics of large genomic rearrangements was based on Multiple Ligation dependent Probe Amplification (MLPA). Subsequent analyses of deletion and duplication breakpoints were performed using long-range PCR, PCR, and DNA sequencing.

Results: In set of 1441 unrelated FH patients, large genomic rearrangements were found in 37 probands. Eight different types of rearrangements were detected, from them 6 types were novel, not described so far. In all rearrangements, we characterized their exact extent and breakpoint sequences.

Conclusions: Sequence analysis of deletion and duplication breakpoints indicates that intrachromatid non-allelic homologous recombination (NAHR) between *Alu* elements is involved in 6 events, while a non-homologous end joining (NHEJ) is implicated in 2 rearrangements. Our study thus describes for the first time NHEJ as a mechanism involved in genomic rearrangements in the *LDLR* gene.

Background

Familial hypercholesterolemia (FH) is an autosomal dominant disease, caused predominantly by variants in the low density lipoprotein receptor (*LDLR*) gene. Pathogenic alternations in the LDLR protein cause a lack of functional receptors for LDL particles on the liver cell surface and give rise to increased serum LDL-cholesterol levels. The high LDL-cholesterol level frequently gives rise to tendon xanthomas, xanthelasmas, arcus lipoides corneae, and accelerated atherosclerosis resulting from cholesterol deposition in the arterial wall, thereby increasing the risk of premature coronary heart disease. The frequency of heterozygous FH in most populations is about 1/500, homozygous FH is rare (≤ 1/1000,000) [1]. The identification and treatment of FH patients and their affected relatives with effective lipid-lowering agents is important and as this has been shown to significantly reduce both coronary morbidity and mortality [2,3]. Genetic testing is the preferred diagnostic method in FH families because it provides an unequivocal diagnosis [1,4,5]. The *LDLR* gene is localized at 19p13.2, is composed of 18 exons spanning 45 kb, the transcript is 5.3 kb long and encodes a peptide containing 860 amino acids [6]. *LDLR* mutations have been reported along the whole length of the gene in FH patients from around the world. At present, the number of identified unique *LDLR* allelic variants is over 1000: 65% of the variants are DNA substitutions, 24% small DNA rearrangements (< 100 bp) and 11% large DNA rearrangements (> 100 bp) http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR and [7].

Genesis of large DNA rearrangements in the *LDLR* gene is frequently associated with *Alu* elements, which are highly abundant in this particular locus [6,8,9].

* Correspondence: lenkafajkusova@volny.cz
[1]University Hospital Brno, Centre of Molecular Biology and Gene Therapy, Černopolní 9, CZ-62500 Brno, Czech Republic
Full list of author information is available at the end of the article

Publication of the human genome DNA sequence has revealed that there are 98 *Alu* repeats within the *LDLR* gene (95 in intronic sequences and 3 in the 3'untranslated region) and *Alu* repeats accounted for 65% of *LDLR* intronic sequences [10].

Alu is the most abundant short interspersed nuclear element (SINE) of the human genome, occupying 10% of the genome content with a copy number estimated to be at least 1.3 million [11]. Consensus *Alu* sequence is approximately 300 bp in length, and consists of two similar, but distinct monomers. The longer right *Alu* monomer contains a 31 bp insert absent from the left *Alu* monomer. A functional RNA polymerase III promoter is present in the left monomer, but is absent from the right monomer [12,13]. *Alu* sequences are regarded as retrotransposons that have inserted into the human genome via a single-stranded RNA intermediate generated by RNA pol III transcription [14]. The *Alu* dimer is usually followed by a 3'A-rich region, a typical feature of SINEs, and the two monomers are separated by a middle A-rich region, an obvious remnant of an ancestral monomeric *Alu*'s 3'A-rich tail [15].

Throughout *Alu* evolution, the source gene(s) accumulated mutations that were incorporated into the new copies made, creating new *Alu* subfamilies. Therefore, the *Alu* family is composed of a number of distinct subfamilies characterized by a hierarchical series of mutations that result in a series of subfamilies of different ages [16-20].

Alu repeat dispersion throughout the genome offers many opportunities for homologous recombinations. Nonallelic homologous recombination (NAHR) is the most common mechanism underlying disease associated genome rearrangements. NAHR can use either region-specific low-copy repeats or repetitive sequences (e.g., *Alu*) as homologous recombination substrates [21,22]. Another recombination mechanism causing genomic disorders is nonhomologous end joining (NHEJ). This process involves the double strand breakage of DNA followed by end joining in the absence of extensive sequence homology [23-25]. NHEJ is associated with very short stretches of sequence identity (a few bp) between the two ends of the breakpoint junctions [22,26,27].

In this study, we present results of analyses of large genomic rearrangements in Czech FH patients detected using Multiple Ligation dependent Probe Amplification (MLPA). In set of 1441 unrelated FH patients, large genomic rearrangements were detected in 37 probands. We found 8 different types of rearrangements, from them 6 types were novel, not described so far. In all rearrangements, we characterized their exact extent and breakpoint sequences. The results showed that 6 events were products of NAHR between *Alu* repeat sequences. The remaining 2 events apparently originated from NHEJ.

Methods

Patients

One thousand nine hundred and forty five probands with probable or definite diagnosis of FH, submitted to the database of the MedPed (Make Early Diagnoses to Prevent Early Deaths) project in the Czech Republic, were included into the study. MedPed is an international project joining together experts from more than 30 countries of the world. In the Czech Republic, the project is coordinated by the Czech Society for Atherosclerosis. Experimental research reported in this study has been performed with the approval of the Ethical Committee of the General University Hospital in Prague, the Czech Republic, and is in compliance with the Helsinki Declaration. All patients gave their informed consent with their participation in the study, which is a part of each patient's personal documentation. The text of the informed consent is available at: http://www.athero.cz/user_data/zpravodajstvi/obrazky/File/medped/informovany_souhlas.pdf The patient file in our study include a) patients with untreated total and/or LDL cholesterol serum levels above the 95th percentile of age, sex and population specific values; b) patients with elevated total and LDL cholesterol in serum but untreated levels unavailable or not exceeding the 95th percentile of age, sex and population specific values, and, in addition, with high clinical suspicion of FH based on personal history and/or family history of premature coronary heart disease and/or elevated total and LDL cholesterol serum levels in the first degree relatives.

DNA analysis of FH patients is divided into several consecutive steps: 1) PCR-RFLP detection of the most common mutation in the *APOB* gene (p.Arg3527Gln) [28,29]; 2) PCR-RFLP detection of the most common mutations in the *LDLR* gene (p.Gly592Glu, p.Asp266Glu, and p.Arg416Trp); 3) PCR-sequencing of *LDLR* exon 4 (the exon with the greatest occurrence of mutations in Czech FH patients); 4) MLPA analysis of all *LDLR* exons; 5) PCR-sequencing of the promoter and *LDLR* exons 1, 5, 6, 9, 10, 12, 14; and 6) PCR-denaturing high performance liquid chromatography of *LDLR* exons 2, 3, 7, 8, 11, 13, 15, 16, 17, and 18, followed by sequencing of positively tested regions.

The break of DNA analysis in case of a mutation finding depends on personal and family history of hypercholesterolemia, the presence of tendon xanthomas, xanthelasmas, early coronary artery disease and premature coronary heart disease. The DNA analysis continues in cases when i) a phenotype manifestation could be associated with the presence of two *LDLR* mutations or ii) a detected missense mutation is new with hardly

predicted effect on the protein structure and function. This diagnostic process is common in FH diagnostics [30,31]. Only data obtained by MLPA analysis are present in this study.

Analysis of deletion and duplication breakpoints in the *LDLR* gene

DNA was isolated according to the standard salting-out method. MLPA was performed using SALSA MLPA KIT P062-C1 LDLR (MRC-Holland), according to the manufacturer's instruction, and analysed on CEQ 8000 Genetic Analysis System (Beckman Coulter). To characterize the precise locations of genomic breakpoints, a number of amplifications and PCR product analyses were performed. Primers for initial long-range amplifications are given in Table 1 together with approximate sizes of PCR fragments of mutated alleles and nested primers for precise determination breakpoints using DNA sequencing. Long-range PCR were performed using Expand Long Template PCR System Kit (Roche) and PCR amplifying fragments around breakpoints using AmpliTaq Gold polymerase (Applied Biosystems). PCR products were purified and sequenced on ABI PRISM 310 DNA-sequencer (Applied Biosystems). Repetitive sequences were identified using RepeatMasker version-3.1.5 available at http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker.

Results

For DNA analysis, 1945 FH probands were selected: 252 probands (13,0%) had the *APOB* mutation; 186 probands (9,6%) had the mutation p.Gly592Glu or p. Asp266Glu or p.Arg416Trp; 66 probands (3,4%) had a mutation in exon 4. 1441 patients were analyzed by MLPA and in 37 probands (1,9%) a deletion/duplication was detected. At present time, the DNA analysis continues in 1404 patients.

Using MLPA, we found 8 types of large genomic rearrangements - 5 deletions and 3 duplications (Table 2). Six types of rearrangements were novel, so far not described: exon2_6dup, exon3_12del, exon4_8dup, exon5_10del, exon9_15del, and exon16_18dup (The terminology used should be read e.g., in exon2_6dup as duplication of exon 2 to exon 6). Using long-range PCR, PCR, and DNA sequencing, we analysed breakpoints of deletions and duplications identified in our FH patients. In Table 2, we show correct sizes of deletions and duplications together with terms of repetitive elements surrounding breakpoints. Schematic illustration of recombination events are given in Figure 1 and 2. As new rearrangements, we denote deletions/duplications which have not been described in literature so far in terms of exons involved. In this denotation, we do not take into account the exact sequence position of breakpoints determined in this work.

NAHR was detected in six DNA rearrangements (promoter_exon2del, exon2_6dup, exon3_12del, exon9_14del, and exon9_15del, exon16_18dup). In four NAHRs (promoter_exon2del, exon2_6dup, exon9_14del, and exon16_18dup), extensive sequence identity was detected between the breakpoints. In all four cases, the rearrangements were caused by recombination between consensus *Alu* repeats and novel complete recombinant *Alu* sequence was formed in the mutation breakpoint. In contrast, sequence identity around breakpoints of rearrangements exon3_12del and exon9_15del was not

Table 1 Primers for *LDLR* breakpoint analysis

Mutation at cDNA level	*Primers for long-range PCR (5'♦ 3'direction)*	*Size* (kb)*	*Primers for precise breakpoint determination (5'♦ 3'direction)*
promoter_ex2del	F: TGTCGCAAATGGCATAAGGAA R: CGGATTTGCAGGTGACAGACA	2.0	F: AAGGCTGCAGTGAAGTATGATGG R: GAGACGGAGTCTCACTCTGTCG
exon2_6dup	F: AGTTCAAGTGTCACAGCGGC R: GTCTTGGCACTGGAACTCGT	8.0	F: AGTTCAAGTGTCACAGCGGC R: CAAGGTTGGCGTTTTTCATATT
exon3_12del	F: CCAGAAGATTCCAGAAATTTCCAG R: CCTTTCTCCTTTTCCTCTCTCTCA	3.5	F: TGGCTCACTGCAAGCTCCG R: AGGCTGGAGTCCAGTGGTACC
exon4_8dup	F: CAAGTGCCAGTGTGAGGAAGG R: CCCTTGGAACACGTAAAGACCC	2.5	F: CACGTGACTTCAAGGGGTTAAAG R: TTCTCTAAAATGCTTGGGACCA
exon5_10del	F: CACCTGCATCCCCCAGCTGTGGGC R: TGGCTGGGACGGCTGTCCTGCGAAC	3.0	F: TTTGTACAGACACAGGCTGGTC R: CAGATGTCACCTGACAGGTACAG
exon9_14del	F: GGAGTGACTTCAAGGGGTTAAAG R: AGGTGGCTCAGGCTGGGC	0.5	F: GGAGTGACTTCAAGGGGTTAAAG R: AGGTGGCTCAGGCTGGGC
exon9_15del	F: CACGTGATCGTCCCGCCTA R: AAATTCTTGTCAACCTACTTGTGC	0.8	F: AAATTCTTGTCAACCTACTTGTGC R: CACGTGATCGTCCCGCCTA
exon16_18dup	F: CGTGAACATCTGCCTGGAGTC R: TCTTCTCATTTCCTCTGCCAGC	3.0	F: TCGTGTGTGTTGGGATGGGA R: ACCCCAGCCCCCAAACTAAA

F - forward primer; R - reverse primer; * size of PCR fragment of mutated allele. The genomic sequence of the *LDLR* gene was obtained from http://www.ucl.ac.uk/ldlr/LOVDv.1.1.0/refseq/LDLR_codingDNA.html.

Table 2 Genomic characteristics of deletion and duplication breakpoints in the *LDLR* gene in Czech FH patients

Mutation at cDNA level	Mutation at DNA level	Deletion/ duplication size	Recombination mechanism	Repetitive element 5'/ class/family	Repetitive element 3'/ class/family	No. of probands
promoter_2exondel	c. -1823_190 +566del	13186bp	NAHR	AluY/SINE/Alu	AluY/SINE/Alu	1
exon2_6dup	**c. 67+3968_940 +296dup**	14228bp	NAHR	AluSx1/SINE/Alu	AluSx3/SINE/Alu	9
exon3_12del	**c.190+984_1846-1160del**	17604bp	NAHR	FLAM_A/SINE/Alu	AluY/SINE/Alu	1
exon4_8dup	**c.314-446_1187-386dup**	8119bp	NHEJ	AluSx1/SINE/Alu	MER83/LTR/ERV1	1
exon5_10del	**c. 695-67_1586 +371del**	7636bp	NHEJ	AluJo/SINE/Alu	AluSx1/SINE/Alu	4
exon9_14del	c.1186+700_2141-545del	10291bp	NAHR	AluYa5/SINE/Alu	AluY/SINE/Alu	10
exon9_15del)	**c.1187-169_2312-790del**	14110bp	NAHR	AluJb/SINE/Alu	AluSx1/SINE/Alu	8
exon16_18dup)	**c.2311 +1941_*1216dup**	7248bp	NAHR	AluYb8/SINE/Alu	AluSq2/SINE/Alu	3

Newly described rearrangements are in bold letters; NAHR: nonallelic homologous recombination; NHEJ: nonhomologous end joining.

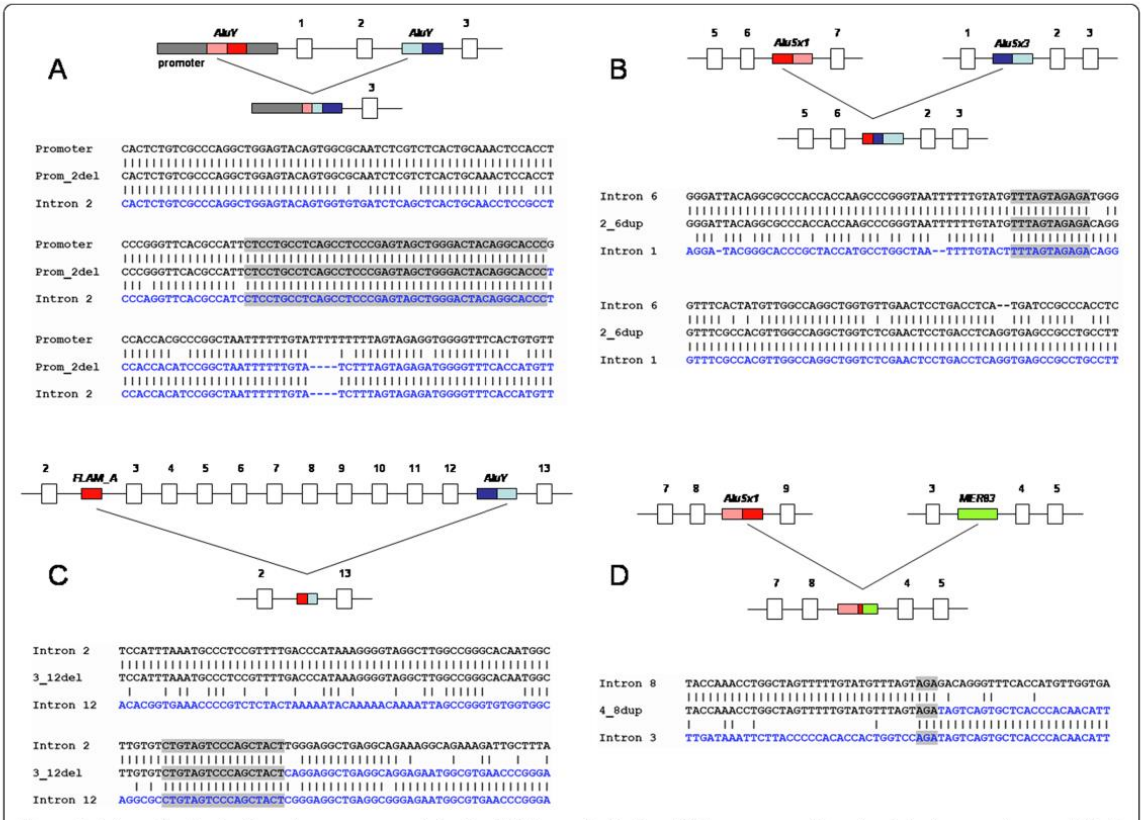


Figure 1 Schematic illustration of rearrangements in the *LDLR* gene including DNA sequence of breakpoints. **A**: promoter_exon2del, **B**: exon2_6dup, **C**: exon3_12del and **D**: exon4_8dup. Consensus *Alu* sequences are depicted as red and blue boxes, their monomer subunits are given in dark and light tones. Sense orientation is marked by a darker tone of the first monomer of the *Alu* consensus sequence, the opposite order marks antisense orientation. MER83 repeat is depicted as a green box. Grey boxes represent sequence overlaps between 5'end and 3' end of the reference sequence.

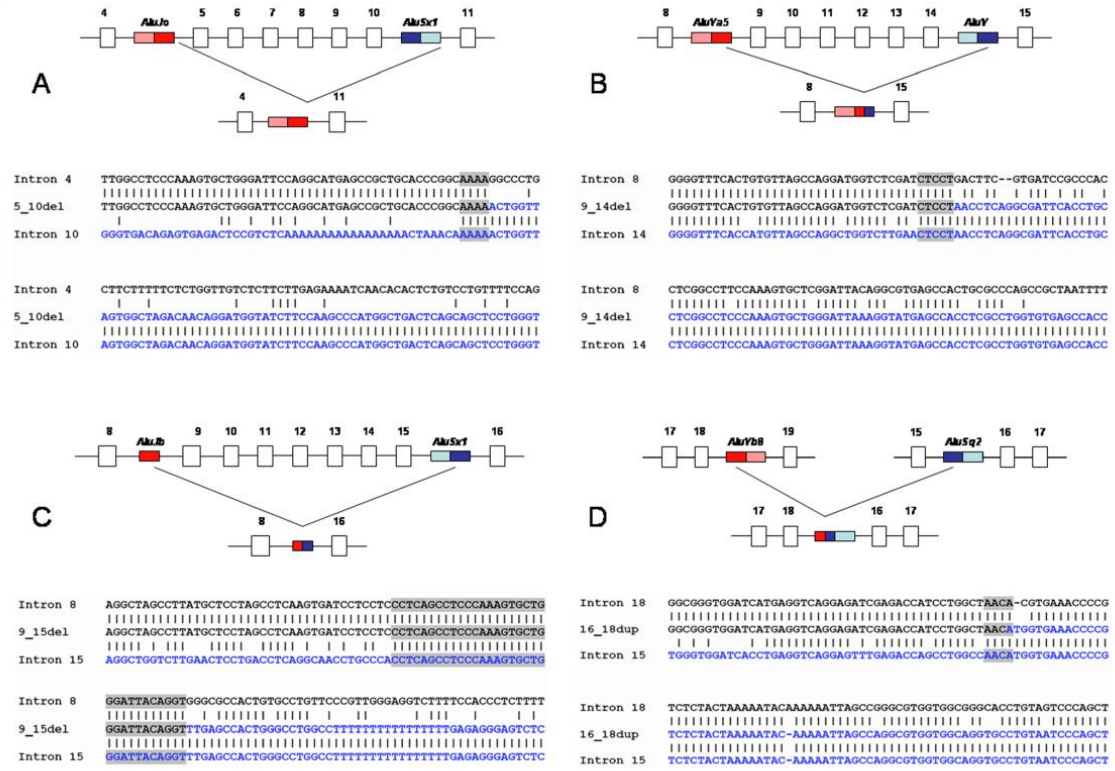


Figure 2 Schematic illustration of rearrangements in the *LDLR* gene including DNA sequence of breakpoints. **A**: exon5_10del, **B**: exon9_14del, **C**: exon9_15del and **D**: exon16_18dup. Consensus *Alu* sequences are depicted as red and blue boxes, their monomer subunits are given in dark and light tones. Sense orientation is marked by a darker tone of the first monomer of the *Alu* consensus sequence, the opposite order marks antisense orientation. Grey boxes represent sequence overlaps between 5' end and 3' end of the reference sequence.

so extensive like in previous cases. These mutations were caused by recombination between an *Alu* repeat in monomer status and a consensus *Alu* repeat (dimer status). The recombination between *FLAM_A* (free left *Alu* monomer, size: 133 bp) and *AluY* (size: 315 bp) was detected in exon3_12del. The recombination between *AluJb* (size: 137 bp) and *AluSx1* (size: 311 bp) was identified in exon9_15del. The deletion breakpoints of both consensus *Alu* repeats were localised in right monomer and so novel complete monomer recombinant *Alu* sequence was formed in the mutation breakpoint. Promoter_exon2del, exon2_6dup, exon9_14del and exon9_15del were formed between *Alu* repeats in the antisense orientation, exon3_12del and exon16_18dup between *Alu* repeats in the sense orientation.

NHEJ was detected in two DNA rearrangements (exon5_10del and exon4_8dup). In exon5_10del, the breakpoint localized in intron 4 was present at the end of the *AluJo* repeat in antisense orientation, and the breakpoint localized in intron 10 was at the end of the *AluSx1* repeat in the sense orientation. In exon4_8dup, the breakpoint in intron 3 was localized in the *MER83* repeat (*ERV1* family repeat) and the breakpoint in intron 8 in the *AluSx1* sequence. There is no sequence homology between these repeats.

Discussion

The 117 large DNA rearrangements are listed on http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR [7]: 100 deletions and 17 duplications. In the view of 98 *Alu* repeats within the *LDLR* gene [10], it is probable that DNA rearrangement breakpoints are located inside of *Alu* repetitive sequences. In the set of our FH patients, we detected 37 large DNA rearrangements in the *LDLR* gene and performed the precise characterization of breakpoints in all types of deletions and duplications. Results define most of breakpoints inside of *Alu* repeats (except one localised in *ERV1* repeat) and NAHR and NHEJ as responsible for these rearrangements. Our results thus demonstrate that *Alu* mediated recombination leads to massive disturbances in the structural and functional integrity of the *LDLR* gene region.

Promoter_exon2del is 13186 bp long and was detected in one Czech FH proband. Approximately 20 kb and 18 kb deletions of promoter, exon 1 and 2 were described previously [32,33]. Exon3_12del detected in one Czech FH proband was not described previously but deletions involving exon 3 were identified (exon3del [34], exon3_5del [35], exon3_6del [34], exon3_8del [36], and exon3_10del [37]). Exon5_10del was found in 4 Czech FH probands and was not described previously. Deletions encompassing exon 5 (exon5del and exon5_6del) were detected in studies [34,38], respectively. Exon9_14del was detected in 10 Czech FH probands. Niessen et al. found exon9_14del in Danish FH patients and performed also analysis of breakpoints. The correct size of the deletion described by Niessen was 9713 bp and both deletion breakpoints were localised in repetitive elements *AluSq* [39]. Exon9_14del determined in our FH probands sized 10291 bp and breakpoints were localised in repetitive elements *AluYa5* and *AluY*. Exon9_10del and exon9_12del were also detected in literature [40]. Exon9_15del was present in 8 Czech FH probands and was not described previously.

All duplications detected in our FH patients were new, not described so far. Exon2_6dup was detected in 9 Czech FH probands, exon4_8dup was found in one Czech FH proband, and exon16_18dup was determined in 3 Czech FH probands. The duplications exon2_8dup, exon4_5dup, and exon16_17dup were described [41-43].

It is interesting that exon2_6dup and exon9_15del were not described in literature and on http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR, but in Czech FH patients these are relatively frequent (9 and 8 probands, respectively).

In the above mentioned work, Nissen et al. [39] described 5 genomic deletions in the *LDLR* gene and defined the breakpoints of each deletion. The five deletions were flanked by *Alu* elements, supporting a mutation mechanism involving unequal homologous recombination between highly similar *Alu* elements. The deletion exon13_15del described by Nissen et al. was flanked by two *AluSg* elements and 15 bp had been inserted at the site of the deleted DNA. This short insertion did not show similarity to any interspersed repeats or any other DNA sequence in the LDLR gene. However, the sequence shows partial homology to several sites in human genome. It is possible to speculate, that in this particular case, the final sequence arrangement has been generated by a more complex mechanism of double strand break repair, involving several recombination steps (e.g., resection and invasion of one DNA strand to a site of a partial homology and its elongation, which was not followed by the single-strand annealing step of homologous recombination, but instead by synthesis-dependent NHEJ). Alternatively, this kind of deletion could have been produced by NHEJ alone, without previous steps of homologous recombination). However, this cannot be clearly distinguished from the final sequence.

In this respect it should be mentioned that deletion and duplication spectra as the outcomes of recombination events in a given genomic locus are influenced not only by the DNA sequence context in the region itself (e.g., abundance and orientation of repeats and their variability) [44,45], but also by epigenetic factors. It corresponds to the fact that it is chromatin template, not a naked DNA, which is a subject of recombination. In the particular case of *Alu* repeats, the role of heterochromatic marks such as DNA methylation, or histone H3K9 methylation in suppression of recombination by these elements has been suggested in recent studies [46-48].

Conclusions

Eight different types of large genomic rearrangements were detected in the *LDLR* gene, from them 6 types were novel, not described so far. Sequence analysis of deletion and duplication breakpoints indicates that both intrachromatid non-allelic homologous recombination (NAHR), and non-homologous end joining (NHEJ) are involved in *LDLR* genomic rearrangements. While NAHR has been described in relation to the *LDLR* gene, this study as the first describes NHEJ in *LDLR* genomic rearrangements.

Acknowledgements

We would like to thank to physicians from the national (M. Vrablík, prof. R. Češka) and regional centres of the Czech MedPed project (V. Bláha, M. Budíková, J. Buryška, R. Cífková, L. Dlouhý, L. Dostálová-Kopečná, H. Halámková, J. Hyánek, J. Hyjánek, Z. Krejsová, J. Macháček, Š. Malá, J. Malý, V. Miláček, J. Mraček, H. Podzimková, D. Povalačová, H. Rosolová, F. Stožický, E. Šipková, L. Ťoukálková, Z. Urbanová, H. Vaverková, S. Zemek, A. Žák) and other physicians participating in this project for providing us with their patients' blood samples and clinical and laboratory data, and to J. Porupková and M. Plotená for technical assistance. This work was funded by the Czech Ministry of Education (projects 2B08060, LC06023, and MSM0021622415).

Author details

[1]University Hospital Brno, Centre of Molecular Biology and Gene Therapy, Černopolní 9, CZ-62500 Brno, Czech Republic. [2]Centre for Cardiovascular Surgery and Transplantation, Pekařská 53, CZ-656 91 Brno, Czech Republic. [3]St. Anne's University Hospital Brno, Department of Clinical Biochemistry, Pekařská 53, CZ-656 91 Brno, Czech Republic. [4]Masaryk University, Faculty of Science, Institute of Experimental Biology, Department of Functional Genomics and Proteomics, Kamenice 5, CZ-62500 Brno, Czech Republic.

Authors' contributions

RG performed detection of deletions/duplications in the *LDLR* gene using MLPA and detailed characterization of breakpoints, LT performed detection of deletions/duplications in the *LDLR* gene using MLPA and administrated database of patients, PZ and OL performed molecular analysis of the *LDLR* gene using PCR-RFLP, DNA sequencing and denaturing high performance liquid chromatography. TF and VS performed clinical examination, selection of patients with suspicion for familial hypercholesterolemia, and blood collections for DNA isolation. LF designed and coordinated the study and has been involved in evaluation of results and manuscript preparation. JF

has been involved in evaluation of results and manuscript preparation. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 12 January 2010 Accepted: 27 July 2010
Published: 27 July 2010

References

1. Goldstein JL, Hobbs HH, Brown MS: **Familial hypercholesterolemia.** *The Metabolic and Molecular Bases of Inherited Disease* New York: McGraw-Hill Medical Publishing DivisionScriver CR, Baudet AL, Sly WS, Valle D 2001, 2863-2910.
2. Neil A, Cooper J, Betteridge J, Capps N, McDowell I, Durrington P, Seed M, Humphries SE: **Reductions in all-cause, cancer, and coronary mortality in statin-treated patients with heterozygous familial hypercholesterolaemia: a prospective registry study.** *Eur Heart J* 2008, **29(21)**:2625-2633.
3. Versmissen J, Oosterveer DM, Yazdanpanah M, Defesche JC, Basart DC, Liem AH, Heeringa J, Witteman JC, Lansberg PJ, Kastelein JJ, *et al*: **Efficacy of statins in familial hypercholesterolaemia: a long term cohort study.** *BMJ* 2008, **337**:a2423.
4. Civeira F: **Guidelines for the diagnosis and management of heterozygous familial hypercholesterolemia.** *Atherosclerosis* 2004, **173(1)**:55-68.
5. Marks D, Thorogood M, Neil HA, Humphries SE: **A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia.** *Atherosclerosis* 2003, **168(1)**:1-14.
6. Hobbs HH, Russell DW, Brown MS, Goldstein JL: **The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein.** *Annu Rev Genet* 1990, **24**:133-170.
7. Leigh SE, Foster AH, Whittall RA, Hubbart CS, Humphries SE: **Update and analysis of the University College London low density lipoprotein receptor familial hypercholesterolemia database.** *Ann Hum Genet* 2008, **72(Pt 4)**:485-498.
8. Hobbs HH, Lehrman MA, Yamamoto T, Russell DW: **Polymorphism and evolution of Alu sequences in the human low density lipoprotein receptor gene.** *Proc Natl Acad Sci USA* 1985, **82(22)**:7651-7655.
9. Hobbs HH, Brown MS, Goldstein JL: **Molecular genetics of the LDL receptor gene in familial hypercholesterolemia.** *Hum Mutat* 1992, **1(6)**:445-466.
10. Amsellem S, Briffaut D, Carrie A, Rabes JP, Girardet JP, Fredenrich A, Moulin P, Krempf M, Reznik Y, Vialettes B, *et al*: **Intronic mutations outside of Alu-repeat-rich domains of the LDL receptor gene are a cause of familial hypercholesterolemia.** *Hum Genet* 2002, **111(6)**:501-510.
11. Deininger PL, Batzer MA: **Alu repeats and human disease.** *Mol Genet Metab* 1999, **67(3)**:183-193.
12. Fuhrman SA, Deininger PL, LaPorte P, Friedmann T, Geiduschek EP: **Analysis of transcription of the human Alu family ubiquitous repeating element by eukaryotic RNA polymerase III.** *Nucleic Acids Res* 1981, **9(23)**:6439-6456.
13. Willis IM: **RNA polymerase III. Genes, factors and transcriptional specificity.** *Eur J Biochem* 1993, **212(1)**:1-11.
14. Weiner AM, Deininger PL, Efstratiadis A: **Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information.** *Annu Rev Biochem* 1986, **55**:631-661.
15. Mighell AJ, Markham AF, Robinson PA: **Alu sequences.** *FEBS Lett* 1997, **417(1)**:1-5.
16. Shen MR, Batzer MA, Deininger PL: **Evolution of the master Alu gene(s).** *J Mol Evol* 1991, **33(4)**:311-320.
17. Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiewicz E, Zuckerkandl E: **Standardized nomenclature for Alu repeats.** *J Mol Evol* 1996, **42(1)**:3-6.
18. Roy AM, Carroll ML, Kass DH, Nguyen SV, Salem AH, Batzer MA, Deininger PL: **Recently integrated human Alu repeats: finding needles in the haystack.** *Genetica* 1999, **107(1-3)**:149-161.
19. Roy AM, Carroll ML, Nguyen SV, Salem AH, Oldridge M, Wilkie AO, Batzer MA, Deininger PL: **Potential gene conversion and source genes for recently integrated Alu elements.** *Genome Res* 2000, **10(10)**:1485-1495.
20. Slagel V, Flemington E, Traina-Dorge V, Bradshaw H, Deininger P: **Clustering and subfamily relationships of the Alu family in the human genome.** *Mol Biol Evol* 1987, **4(1)**:19-29.
21. Shaw CJ, Lupski JR: **Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease.** *Hum Mol Genet* 2004, **13(Spec No 1)**:R57-64.
22. de Smith AJ, Walters RG, Coin LJ, Steinfeld I, Yakhini Z, Sladek R, Froguel P, Blakemore AI: **Small deletion variants have stable breakpoints commonly associated with alu elements.** *PLoS One* 2008, **3(8)**:e3104.
23. Linardopoulou EV, Williams EM, Fan Y, Friedman C, Young JM, Trask BJ: **Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication.** *Nature* 2005, **437(7055)**:94-100.
24. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev D, Carriero NJ, Du L, *et al*: **Paired-end mapping reveals extensive structural variation in the human genome.** *Science* 2007, **318(5849)**:420-426.
25. Perry GH, Ben-Dor A, Tsalenko A, Sampas N, Rodriguez-Revenga L, Tran CW, Scheffer A, Steinfeld I, Tsang P, Yamada NA, *et al*: **The fine-scale and complex architecture of human copy-number variation.** *Am J Hum Genet* 2008, **82(3)**:685-695.
26. Chan CY, Kiechle M, Manivasakam P, Schiestl RH: **Ionizing radiation and restriction enzymes induce microhomology-mediated illegitimate recombination in Saccharomyces cerevisiae.** *Nucleic Acids Res* 2007, **35(15)**:5051-5059.
27. Lee K, Lee SE: **Saccharomyces cerevisiae Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes microhomology-mediated end joining.** *Genetics* 2007, **176(4)**:2003-2014.
28. Innerarity TL, Weisgraber KH, Arnold KS, Mahley RW, Krauss RM, Vega GL, Grundy SM: **Familial defective apolipoprotein B-100: low density lipoproteins with abnormal receptor binding.** *Proc Natl Acad Sci USA* 1987, **84(19)**:6919-6923.
29. Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ: **Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100.** *Proc Natl Acad Sci USA* 1989, **86(2)**:587-591.
30. Civeira F, Ros E, Jarauta E, Plana N, Zambon D, Puzo J, Martinez de Esteban JP, Ferrando J, Zabala S, Almagro F, *et al*: **Comparison of genetic versus clinical diagnosis in familial hypercholesterolemia.** *Am J Cardiol* 2008, **102(9)**:1187-1193, 1193 e1181.
31. Taylor A, Wang D, Patel K, Whittall R, Wood G, Farrer M, Neely RD, Fairgrieve S, Nair D, Barbir M, *et al*: **Mutation detection rate and spectrum in familial hypercholesterolaemia patients in the UK pilot cascade project.** *Clin Genet* 2009, **77(6)**:572-80.
32. Garuti R, Lelli N, Barozzini M, Tiozzo R, Ghisellini M, Simone ML, Li Volti S, Garozzo R, Mollica F, Vergoni W, *et al*: **Two novel partial deletions of LDL-receptor gene in Italian patients with familial hypercholesterolemia (FH Siracusa and FH Reggio Emilia).** *Atherosclerosis* 1996, **121(1)**:105-117.
33. Chaves FJ, Real JT, Garcia-Garcia AB, Puig O, Ordovas JM, Ascaso JF, Carmena R, Armengod ME: **Large rearrangements of the LDL receptor gene and lipid profile in a FH Spanish population.** *Eur J Clin Invest* 2001, **31(4)**:309-317.
34. Fouchier SW, Kastelein JJ, Defesche JC: **Update of the molecular basis of familial hypercholesterolemia in The Netherlands.** *Hum Mutat* 2005, **26(6)**:550-556.
35. Chang JH, Pan JP, Tai DY, Huang AC, Li PH, Ho HL, Hsieh HL, Chou SC, Lin WL, Lo E, *et al*: **Identification and characterization of LDL receptor gene mutations in hyperlipidemic Chinese.** *J Lipid Res* 2003, **44(10)**:1850-1858.
36. Fouchier SW, Defesche JC, Umans-Eckenhausen MW, Kastelein JP: **The molecular basis of familial hypercholesterolemia in The Netherlands.** *Hum Genet* 2001, **109(6)**:602-615.
37. Bertolini S, Garuti R, Lelli W, Rolleri M, Tiozzo RM, Ghisellini M, Simone ML, Masturzo P, Elicio NC, Stefanutti C, *et al*: **Four novel partial deletions of LDL-receptor gene in Italian patients with familial hypercholesterolemia.** *Arterioscler Thromb Vasc Biol* 1995, **15(1)**:81-88.
38. Hobbs HH, Brown MS, Goldstein JL, Russell DW: **Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia.** *J Biol Chem* 1986, **261(28)**:13114-13120.
39. Nissen PH, Damgaard D, Stenderup A, Nielsen GG, Larsen ML, Faergeman O: **Genomic characterization of five deletions in the LDL receptor gene in Danish Familial Hypercholesterolemic subjects.** *BMC Med Genet* 2006, **7**:55.

Clinica Chimica Acta 419 (2013) 1–10

Contents lists available at SciVerse ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

Hyperphenylalaninemia in the Czech Republic: Genotype–phenotype correlations and *in silico* analysis of novel missense mutations

Kamila Réblová [a,*,1], Zuzana Hrubá [b,1], Dagmar Procházková [c], Renata Pazdírková [d], Slávka Pouchlá [b], Lenka Fajkusová [a,b,*]

[a] Central European Institute of Technology, Masaryk University, Brno, Czech Republic
[b] Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Brno, Czech Republic
[c] Department of Pediatrics, Medical Faculty of Masaryk University and University Hospital Brno, Brno, Czech Republic
[d] Department of Children and Adolescents, Third Faculty of Medicine, Charles University and Faculty Hospital Kralovske Vinohrady, Prague, Czech Republic

ARTICLE INFO

Article history:
Received 23 November 2012
Received in revised form 10 January 2013
Accepted 16 January 2013
Available online 26 January 2013

Keywords:
Hyperphenylalaninemia
Phenylalanine hydroxylase
Genotype–phenotype relationship
Molecular modelling

ABSTRACT

Background: Hyperphenylalaninemia (HPA) is one of the most common inherited metabolic disorders caused by deficiency of the enzyme phenylalanine hydroxylase (PAH). HPA is associated with mutations in the *PAH* gene, which leads to reduced protein stability and/or impaired catalytic function. Currently, almost 700 different disease-causing mutations have been described. The impact of mutations on enzyme activity varies ranging from classical PKU, mild PKU, to non-PKU HPA phenotype.
Methods: We provide results of molecular genetic diagnostics of 665 Czech unrelated HPA patients, structural analysis of missense mutations associated with classical PKU and non-PKU HPA phenotype, and prediction of effects of 6 newly discovered HPA missense mutations using bioinformatic approaches and Molecular Dynamics simulations.
Results: Ninety-eight different types of mutations were indentified. Thirteen of these were novel (6 missense, 2 nonsense, 1 splicing, and 4 small gene rearrangements). Structural analysis revealed that classical PKU mutations are more non-conservative compared to non-PKU HPA mutations and that specific sequence and structural characteristics of a mutation might be critical when distinguishing between non-PKU HPA and classical PKU mutations. The greatest impact was predicted for the p.(Phe263Ser) mutation while other novel mutations p.(Asn167Tyr), p.(Thr200Asn), p.(Asp229Gly), p.(Leu358Phe), and p.(Ile406Met) were found to be less deleterious.

1. Introduction

Hyperphenylalaninemia (HPA, MIM# 261600) is the most common inborn error of amino acid (AA) metabolism in Europeans. HPA is transmitted by the autosomal recessive mode of inheritance caused by mutations of the phenylalanine hydroxylase (PAH) gene. In the liver, PAH metabolizes L-phenylalanine (Phe) to L-tyrosine (Tyr) using (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH4) as cofactor. Failure in this conversion leads to an increase of Phe in the body fluids and severe mental retardation unless Phe intake is restricted. Almost 700 different disease-causing mutations in the *PAH* gene have been identified and can be found in the mutation databases PAH and HGMD (http://www.pahdb.mcgill.ca/, http://www.hgmd.cf.ac.uk/ac/index.php), most of which are point mutations scattered throughout the whole *PAH* gene. Mutations vary in their impact on enzyme activity, causing a range of clinical phenotypes. HPA patients can be classified based on their off-diet diagnostic plasma Phe levels as classical PKU (Phe above 1200 μmol/L), mild PKU (Phe between 600 and 1200 μmol/L) or non-PKU HPA (Phe between 120 and 600 μmol/L) [1].

In a subset of HPA patients, Phe concentration is manageable with pharmacological doses of BH4 with either limited or no dietary restriction [2–5]. For patients whose disease responds to BH4, cofactor therapy is an attractive means to increase patient compliance. A number of studies have documented that HPA patients with milder phenotypes are more likely to benefit from BH4 therapy [4–8]. Mechanism of BH4 responsiveness is multifactorial [3]. Current data suggest the most common mechanism by which BH4 rescues PAH function is by acting as a pharmacological chaperone promoting proper enzyme folding, which consequently reduces enzyme degradation [9–11]. About 75% of PAH mutations characterized by high residual activity have been found to be associated with BH4 responsiveness both *in vitro* [12–14] and *in vivo* [4,15]. For this reason, *PAH* genotyping has utility not only in disease categorization (classical PKU, mild PKU, non-PKU HPA) but also

* Corresponding authors. Tel.: +420 532234625; fax: +420 532234623.
E-mail addresses: kristina@physics.muni.cz (K. Réblová), lenkafajkusova@volny.cz (L. Fajkusová).
[1] K. Réblová and Z. Hrubá equally contributed to this manuscript.

0009-8981/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cca.2013.01.006

in predicting the potential for response to cofactor therapy [16,17]. However, there are also reports showing inconsistencies in correlation between BH4 responsiveness and genotype [16,17].

The *PAH* gene, mapped on 12q23.2, consists of 13 exons encompassing 171 kb. The full-length *PAH* cDNA encodes a protein of about 52 kDa (452 AAs) that is assembled as a homotetramer in the mature form. Each monomer consists of three functional domains: an N-terminal regulatory domain (residues 1–142); a catalytic domain (residues 143–410) that includes binding sites for Fe^{3+} ion, which is reduced to the active Fe^{2+} form upon binding of cofactor; and a C-terminal oligomerization domain (residues 411–452) with dimerization (residues 411–426) and tetramerization motifs (residues 427–452) [18]. It is now established that in most HPA cases, the loss of PAH function is due to structural distortion leading to decreased stability [10,19,20], folding deficiency [21], and/or increased susceptibility toward aggregation and degradation [10,22] of PAH mutant proteins. The availability of the crystal structures of PAH [18,23] makes it possible to model missense mutations and their effects on the protein structure [24]. Missense mutations affecting protein structure and stability are referred to as "structural" and are frequently associated with change of AA size, charge, polarity, and more importantly, with a loss of structure-maintaining contacts such as stacking interaction and H-bonding.

In this study, we present results of molecular genetic diagnostics of 665 unrelated Czech HPA patients, genotype-phenotype correlations, and results of the BH4-loading test in selected patients. Further, using bioinformatic approaches and molecular modelling, we carried out structural analysis of selected HPA missense mutations associated predominantly with the PKU or non-PKU HPA phenotype and explored the pathogenicity of six novel HPA missense mutations discovered in our patients. In addition, we utilized widely used mutation prediction tools, such as Polyphen-2 [25], SIFT [26], SNPs3D [27] and FOLDX [28] in order to predict the impact of novel missense mutations.

2. Materials and methods

2.1. DNA analysis

Genomic DNA was extracted from peripheral blood leukocytes by the standard salting-out method and amplified. Primers for amplification of all exons and adjacent intron sequences and the conditions of particular PCRs are available by request. PCR products were sequenced directly using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and analyzed on an ABI 3130Xl Genetic Analyzer (Applied Biosystems). The resulting sequences were compared with the *PAH* NCBI reference sequence (NG_008690.1). All novel missense mutations were analyzed on a control panel consisting of DNA from 200 healthy Czech individuals. Multiple Ligation dependent Probe Amplification (MLPA) was performed using the SALSA P055 PAH MLPA kit (MCR Holland). This kit contained 25 sets of probes; 13 were PAH specific, and the others were control standards from other human genes. In detail, the assay is described in our previous study [29] as well as long-range PCRs for detailed characterization of deletions detected by MLPA. The nomenclature of mutations was implemented according to the current HGVS recommendations (http://www.hgvs.org/mutnomen/).

2.2. BH4-loading test

The BH4 test was performed at two workplaces: 1) the Department of Pediatrics of the Faculty of Medicine, Masaryk University and the University Hospital Brno and 2) the Department of Children and Adolescents of the Third Faculty of Medicine, Charles University and the Faculty Hospital Kralovske Vinohrady, after obtaining informed consent from all participants or their parents. Four days before BH4 loading and during the entire testing period patients were recommended to relax the diet and consume Phe-rich food. A single dose of 20 mg/kg BH4 (Kuvan™, sapropterindihydrochloride) was administered orally after night fasting [7]. In the workplace 1, blood samples for Phe analysis were collected 0 (Phe 0), 4 (Phe 4), 8 (Phe 8), and 24 (Phe 24) hours after administration of BH4. In the workplace 2, blood samples for Phe analysis were collected 0 (Phe 0), 8 (Phe 8), 16 (Phe 16), and 24 (Phe 24) hours after administration of BH4, in some patients the second dose of BH4 was applied and blood samples was collected 32 (Phe 32) and 48 (Phe 48) hours after administration of the first BH4 dose. BH4 responsiveness was defined as a decrease in blood Phe by 30% or more during 24 h. Patients were evaluated as responders (R, reduction of blood Phe by 30% or more during 24 h), late-responders (LR, reduction of blood Phe below 30% during 24 h but this was changed after the second BH4 dose), and non-responders (NR, reduction of blood Phe below 30%).

2.3. Building of 3D protein structures of wild type and mutant PAH

The 3D model of PAH containing the complete catalytic (C) and tetramerization (T) domains was built based on the X-ray structures of truncated forms of human PAH (pdb codes: 1PAH [30] and 2PAH [30]). Based on superposition of these structures in the VMD program [31], we obtained a model comprising residues 117–452. The missing regulatory (R) domain was built using homology modelling with the Modeler program [32] taking the X-ray structure of the truncated form of rat PAH (pdb code: 1PHZ [23]) with R and C domains as a template. A model comprising residues 19–452 was then obtained by superimposing the homology model of human R and C domains and the structure with complete C and T domains over the C domain in VMD. Subsequently, we cut from this model a large portion of the helix Tα1 (α1-helix in the T domain), which participates in oligomerization. We expected this helix to be dynamic in Molecular Dynamic (MD) simulations without the other monomer subunits which could result in formation of incorrect contacts and structural perturbations. Thus, the final model subjected to the mutational analysis and MD simulations consisted only of residues 19–432. The model does not contain Fe^{3+} ion because modelling of divalent/trivalent cations suffers from insufficient parameterization [33]. Since the model was built based on X-ray structures lacking hydrogen atoms, we added these using the Xleap module of Amber 10 [34] which is a package of programs for MD simulations of nucleic acids and proteins. The protonation states of all histidines in the protein were set to allow formation of proper H-bonds. Hence, histidines 146 and 290 were δ protonated. The structure of mutants N167Y, T200N, D229G, F263S, L358F, and I406M carrying mutations p.(Asn167Tyr), p.(Thr200Asn), p.(Asp229Gly), p.(Phe263Ser), p.(Leu358Phe), and p.(Ile406Met), respectively, were generated by truncating the side chains in the wild type structure and building new side chains using the Xleap module.

2.4. Structural analysis of missense mutations

We evaluated 9 sequence and structure features for selected HPA missense mutations associated predominantly with PKU or non-PKU HPA phenotype, and for 6 novel HPA missense mutations discovered in our patients. These parameters characterize the mutations, i.e. they reveal possible structural or functional defects, which indicates the potential causality of a mutation [35,36].

Evaluated sequence and structure features: *1) Formation of specific side chain contacts* (H-bonds, salt bridges, stacking interactions) of AAs in the wild type (wt) structure (based on visual inspection of the 3D model of PAH using the VMD program). Loss of these contacts upon a mutation often results in destabilization of protein folding. *2) Occurrence of AAs on the inner surface of the active site* located in the cavity of the catalytic domain (based on visual inspection of the 3D model of PAH using the VMD program). Replacement of AAs at this site often results in impairment of the catalytic function of a protein. *3) Buriedness of*

AAs corresponding to the relative accessible surface area (RSA) ≤ 10% [37]. Replacement of buried AAs is more likely to be associated with structural defects. *4) Volume change upon mutation* calculated according to [38]. A volume change with an absolute value of ≥ 30 $Å^3$ was considered destabilizing [39]. The volume changes were taken into account only if AAs occurred simultaneously at buried positions, where they may lead to side chain overpacking or formation of cavities. *5) Charge change upon mutation* detected between charged and uncharged AAs. Charge change was taken into account only if AAs occurred simultaneously at a buried position, where they may more likely result in disruption of protein structure. *6) Polarity change upon mutation* detected between nonpolar (L, I, F, W, C, M, V, Y), polar (P, A, T, G, S), and very polar (H, R, Q, K, N, E, D) AAs [40]. Polarity change was considered only if AAs occurred simultaneously at a buried position. *7) Conservation of wt AAs* calculated using the MSV3D web application [41] which generates multiple sequence alignment (MSA) for PAH and homologous sequences. Based on the MSA, we detected conservation corresponding to >50% identity of wt AAs. High conservation indicates key positions which when mutated frequently cause severe structural and functional defects. *8) Presence of helix/turn breakers.* Mutations of wt Pro and Gly in turn structures and incorporation of Pro and Gly into helices were considered, both of which disrupt secondary structure conformation. *9) Mutational likelihood* detected using the Blosum62 matrix [42]. Mutations with negative scores were detected, which are less likely to occur in the context of evolution.

2.5. MD simulations of proteins with novel PAH missense mutations

Using the Amber package 10 [34] and the force field parm99SB [43], we ran 50 ns-long MD simulations for the 3D model of wt PAH and six mutant proteins (N167Y, T200N, D229G, F263S, L358F, and I406M). The structures were built in the Xleap module as described above. Prior to simulations, the protein structures were neutralized by several K^+ ions [44] and solvated by an octahedral SPC/E water box extending 10 Å away from the solute. To increase ionic strength, we added additional K^+ [44] and Cl^- [45] ions corresponding to 0.2 M concentration. Equilibration and production phases were carried using standard protocols [46]. Trajectories were analyzed using the Ptraj module of Amber and visualized using VMD.

2.6. Predictions of effects of PAH missense mutations using standard programs

We used Polyphen-2 [25], SIFT [26], SNPs3D [27], and FOLDX3.0 [28] in order to predict the impact of 6 novel missense mutations. The first three programs rely either solely upon sequence alignment (SIFT), or on sequence alignment with protein structural attributes (PolyPhen-2), or use sequence alignment with or without protein structural data in combination with a SupportVector Machine (SNPs3D). The forth program utilizes the 3D protein structure and empirical potential with different energy terms and predicts the difference in the Gibbs free energy of unfolding (ΔΔG) between wt and mutant proteins ($\Delta\Delta G = \Delta G(\text{mutant protein}) - \Delta G(\text{wt})$. Mutations leading to a change in ΔΔG of >1 kcal/mol are considered destabilizing and those leading to a change in ΔΔG of <−1 kcal/mol are considered stabilizing. The range 1 kcal/mol ≥ ΔΔG ≥ −1 kcal/mol is considered neutral. The assessment of mutations by each program was carried out with default settings.

3. Results

3.1. DNA analysis and genotype–phenotype correlations

We sequenced the exons and adjacent intron regions of the *PAH* gene in 665 unrelated HPA patients from the Czech Republic. When no mutation or one mutation was detected, MLPA analysis for detection of large gene deletions/duplications was performed. The complete genotype was detected in 658 patients, and in 7 patients only one mutation was identified. In total, 98 different types of mutations were identified. From these, 61 types were missense (62.2 %), 9 nonsense, 12 splicing, 13 small gene rearrangements, and 3 large gene rearrangements (Tables 1A and 1B). When phenotypic classes are considered, p.(Arg408Trp) is the most common mutation in the group of classical PKU patients and p.(Ala403Val) in non-PKU HPA patients (Table 1A). Twenty one types of mutations were observed only in Czech HPA patients, of which 8 were mentioned in our previous studies [29,47,48] and/or can be found in mutation databases PAH and HGMD (http://www.pahdb.mcgill.ca/, http://www.hgmd.cf.ac.uk/ac/index.php). Thirteen mutations described in this study are new and not described so far (6 missense, 2 nonsense, 1 splicing, and 4 small gene rearrangements). For prediction of the causality of missense mutations, especially of novel mutations, we used bioinformatic tools and carried out structural analysis employing MD simulations (see below). For prediction of effect of a novel splicing mutation on pre-mRNA splicing, the *in silico* tools NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) and SpliceView (http://zeus2.itb.cnr.it/~webgene/wwwspliceview.html) were used and both showed a deleterious effect for this mutation. Large gene deletions were identified in 41 disease alleles. In our previous study [29], we performed sequence analysis of deletion breakpoints and determined that non-allelic homologous recombination between *Alu* elements is involved in the deletion Ex3del4765, while non-homologous end joining is implicated in the deletions Ex5del955 and Ex5del4232ins268.

3.2. BH4 loading test

The outcome of the BH4 loading test performed in 46 HPA patients is summarized in Supplementary Tables 2A (workplace 1) and 2B (workplace 2). BH4 responsiveness was detected in 22 patients (9 non-PKU HPA, 11 mild PKU, and 2 classical PKU), non-responsiveness in 22 patients (1 mild PKU and 21 classical PKU), and 2 patients were described as late-responders (both classical PKU). In accordance with expectation, the responders in the BH4 test are patients with a mutation associated with non-PKU HPA or mild PKU phenotype but some exceptions were identified. Patient 1 with mutations associated with classical PKU (Supplementary Table 2B, the genotype p.[Phe39Leu];[Arg408Trp]) was evaluated in the BH4 test as responder. The mutation p.(Ile65Thr) associated in literature and in our patients with classical PKU was in patient 1 (Supplementary Table 2A, the genotype p.[Ile65Thr];[Arg261Gln]) BH4 responsive and in patients 13 and 15 (Supplementary Table 2B, the genotype p.[Ile65Thr];[Arg408Trp]) late-BH4 responsive and non-BH4 responsive, respectively. Patient 13 (Supplementary Table 2A, the genotype p.[Lys396Arg]; [Arg408Trp]) with mild PKU was scored as non-responsive, but we cannot rule out that this result was influenced by a low initial concentration of blood Phe.

3.3. Structural analysis of selected PAH missense mutations

For analysis of correlations between structural/functional properties of PAH mutations and their phenotypic manifestations, we divided the mutations into two groups i) mutations associated predominantly with the non-PKU HPA phenotype, and ii) mutations associated predominantly with the classical PKU phenotype (Table 2). We took into account only mutations with clear clinical manifestations and occurring in at least three unrelated patients. This selection was carried out on the basis of our data and of data from the PAH database (http://www.pahdb.mcgill.ca). Our analysis showed that AAs whose mutations lead to the classical PKU phenotype are more frequently i) involved in formation of specific structural contacts; ii) positioned in the active site of the protein; iii) buried in the interior of the protein; iv) conserved, v) associated with helix/turn breakers; and vi) replaced with evolutionary less-likely substitutions compared to AAs whose mutations lead to

Table 1A
Missense mutations detected in Czech HPA probands, their frequencies, and the associated phenotypes.

Mutation (cDNA level)	Mutation (protein level)	Number of alleles	Frequency	Phenotype according to database data (http://www.pahdb.mcgill.ca)	Phenotype according to our results
c.117C>G	p.(Phe39Leu)	4	0.301	Classical PKU	Classical PKU
c.143 T>C	p.(Leu48Ser)	25	1.879	Variant PKU/classical PKU	Mild PKU/classical PKU
c.194 T>C	p.(Ile65Thr)	17	1.278	Classical PKU	Classical PKU
c.193A>G	p.(Ile65Val)	4	0.301	Not mentioned	Non-PKU HPA
c.204A>T	p.(Arg68Ser)	2	0.15	Variant PKU	Mild PKU
c.250 G>T	p.(Asp84Tyr)	1	0.075	Not mentioned	Non-PKU HPA
c.311C>A	p.(Ala104Asp)	2	0.15	Variant PKU	Classical PKU
c.386A>G	p.(Asp129Gly)	2	0.15	Mild PKU (p.[Asp129Gly];[Asp129Gly])	Mild PKU
c.473 G>A	p.(Arg158Gln)	55	4.135	Classical PKU	Classical PKU
c.499A>T	**p.(Asn167Tyr)**	1	0.075	–	Non-PKU HPA
c.506 G>A	p.(Arg169His)	1	0.075	Non-PKU HPA	Non-PKU HPA
c.520A>G	p.(Ile174Val)	1	0.075	Not mentioned	Non-PKU HPA
c.527 G>T	p.(Arg176Leu)	1	0.075	Non-PKU HPA	Non-PKU HPA
c.529 G>C	p.(Val177Leu)	2	0.15	Not mentioned	Non-PKU HPA/mild PKU
c.533A>G	p.(Glu178Gly)	3	0.225	Non-PKU HPA	Non-PKU HPA
c.569 T>C	p.(Val190Ala)	2	0.15	Non-PKU HPA	Classical PKU
c.599C>A	**p.(Thr200Asn)**	1	0.075	–	Non-PKU HPA
c.601C>T	p.(His201Tyr)	4	0.301	Non-PKU HPA	Non-PKU HPA
c.631C>A	p.(Pro211Thr)	1	0.075	Classical PKU	Non/PKU HPA (p.[Pro211Thr];[Gly239Val; Ile269Leu])
c.673C>A	p.(Pro225Thr)	21	1.579	Not mentioned	Classical PKU
c.686A>G	**p.(Asp229Gly)**	1	0.075	-	Non-PKU HPA
c.716 G>T	p.(Gly239Val)	1	0.075	-	Classical PKU
c.805A>C[a]	p.(Ile269Leu)[a]			non-PKU HPA	
c.716 G>C	***p.(Gly239Ala)***	1	0.075	Classical PKU	Classical PKU
c.721C>T	p.(Arg241Cys)	2	0.15	Mild PKU	Mild PKU
c.728 G>A	p.(Arg243Gln)	11	0.827	Classical PKU	Classical PKU
c.733 G>A	***p.(Val245Ile)***	1	0.075	Not mentioned	Non-PKU HPA
c.734 T>C	p.(Val245Ala)	7	0.526	Non-PKU HPA	Non-PKU HPA
c.737C>T	p.(Ala246Val)	1	0.075	Not mentioned	Non-PKU HPA
c.754C>T	p.(Arg252Trp)	51	3.834	Classical PKU	Classical PKU
c.776C>T	p.(Ala259Val)	1	0.075	Classical PKU	Classical PKU
c.782 G>A	p.(Arg261Gln)	21	1.579	Classical PKU	Classical PKU
c.789C>G	p.(Phe263Leu)	1	0.075	Not mentioned	Classical PKU
c.788 T>C	**p.(Phe263Ser)**	1	0.075	-	Mild PKU (p.[Phe263Ser];[Tyr414Cys])
c.809 G>A	p.(Arg270Lys)	6	0.451	Variant PKU (p.[Phe410Cys];[Leu249Phe; Arg270Lys])	Mild PKU/classical PKU
c.842C>T	p.(Pro281Leu)	27	2.03	Classical PKU	Classical PKU
c.847A>T	p.(Ile283Phe)	16	1.203	Classical PKU	Classical PKU
c.889C>T	p.(Arg297Cys)	1	0.075	Non-PKU HPA	Non-PKU HPA
c.890 G>A	p.(Arg297His)	1	0.075	–	Non-PKU HPA
c.896 T>G	p.(Phe299Cys)	1	0.075	Classical PKU	Classical PKU
c.898 G>T	p.(Ala300Ser)	13	0.977	Non-PKU HPA	Non-PKU HPA
c.916A>G	p.(Ile306Val)	23	1.729	Non-PKU HPA	Non-PKU HPA/mild PKU
c.922C>T	p.(Leu308Phe)	2	0.15	Not mentioned	Mild PKU
c.974A>G	p.(Tyr325Cys)	6	0.451	Not mentioned	Classical PKU
c.1024 G>C	***p.(Ala342Pro)***	1	0.075	Not mentioned	Classical PKU
c.1042C>G	p.(Leu348Val)	2	0.15	Classical PKU	Classical PKU
c.1045 T>C	p.(Ser349Pro)	1	0.075	Classical PKU	Classical PKU
c.1074A>T	**p.(Leu358Phe)**	2	0.15	–	Non-PKU HPA
c.1139C>T	p.(Thr380Met)	5	0.376	Non-PKU HPA	Non-PKU HPA
c.1157A>G	p.(Tyr386Cys)	1	0.075	Classical PKU	Mild PKU (p.[Tyr386Cys];[Glu178Gly])
c.1159 T>C	p.(Tyr387His)	11	0.827	Not mentioned	Classical PKU
c.1162 G>A	p.(Val388Met)	2	0.15	Classical PKU	Classical PKU
c.1169A>G	p.(Glu390Gly)	3	0.225	Non-PKU HPA	Non-PKU HPA
c.1183 G>C	p.(Ala395Pro)	5	0.376	Classical PKU	Classical PKU
c.1187A>G	p.(Lys396Arg)	1	0.075	–	Mild PKU
c.1208C>T	p.(Ala403Val)	68	5.113	Non-PKU HPA, variant PKU	Non-PKU HPA
c.1218A>G	**p.(Ile406Met)**	2	0.15	–	Non-PKU HPA
c.1220C>T	p.(Pro407Leu)	1	0.075	Not mentioned	Mild PKU (p.[Pro407Leu]; [Tyr414Cys])
c.1222C>T	p.(Arg408Trp)	560	42.105	Classical PKU	Classical PKU
c.1241A>G	p.(Tyr414Cys)	21	1.579	Variant PKU/ classical PKU	Non-PKU HPA/Variant PKU/classical PKU
c.1243 G>A	p.(Asp415Asn)	1	0.075	Non-PKU HPA	Non-PKU HPA

Mutations mentioned in bold letters are novel and not described so far. Mutations mentioned in italic bold letters were also identified only in Czech HPA patients but were described in our previous studies and/or submitted to the PAH mutation database (http://www.pahdb.mcgill.cz). Phenotypes for particular mutations were derived from patients with genotypes where the second mutation was p.(Arg408Trp) or another mutation associated with complete loss of protein function. If a patient with such a genotype was not available, the described genotype is mentioned.

[a] This patient carries the mutations c.716 G>T, p.(Gly239Val) and c.805A>C, p.(Ile269Leu) on one allele.

non-PKU HPA. When considering buried AAs, we noticed that mutations associated with the classical PKU phenotype more frequently exhibit changes of volume, charge, and polarity than mutations associated with non-PKU HPA. In summary, each of 9 determined features was more abundant in the group of mutations associated with the classical PKU phenotype (Table 2). We recorded the number of determined

Table 1B
Nonsense and splicing mutations and small and large gene rearrangements detected in Czech HPA probands.

Mutation (cDNA level)	Mutation (protein level)	Number of alleles	Frequency
c.58C>T	p.(Gln20*)	2	0.15
c.561 G>A	p.(Trp187*)	3	0.225
c.612 T>G	p.(Tyr204*)	1	0.075
c.727C>T	p.(Arg243*)	36	2.707
c.781C>T	p.(Arg261*)	1	0.075
c.814 G>T	p.(Gly272*)	28	2.105
c.910C>T	**p.(Gln304*)**	1	0.075
c.781C>T	p.(Arg261*)	1	0.075
c.1161C>A	**p.(Tyr387*)**	1	0.075
c.168+5 G>C	**Splicing mutations**	1	0.075
c.441+1 G>A	Splicing mutations	1	0.075
c.441+5 G>T	Splicing mutations	7	0.526
c.441+6 T>A	Splicing mutations	3	0.225
c.509+1 G>A	Splicing mutations	4	0.301
c.842+1 G>A	Splicing mutations	5	0.376
c.913-7A>G	Splicing mutations	1	0.075
c.1066-11 G>A	Splicing mutations	48	3.609
c.1066-3C>T	Splicing mutations	12	0.902
c.1199+1 G>A	Splicing mutations	1	0.075
c.1200-8 G>A	***Splicing mutations***	4	0.301
c.1315+1 G>A	Splicing mutations	33	2.481
c.47_48delCT	p.(Ser16*)	1	0.075
c.116_118delTCT	p.(Phe39del)	25	1.879
c.165delT	p.(Phe55Leufs*6)	15	1.128
c.324_328delGCTTT	**p.(Glu108Aspfs*4)**	3	0.225
c.556delA	p.(Thr186Hisfs*9)	2	0.15
c.607dupT	**p.(Cys203Leufs*3)**	1	0.075
c.664_665delGA	p.(Asp222*)	1	0.075
c.667_674dup8	**p.(Gln226Thrfs*118)**	1	0.075
c.740delG	**p.(Gly247Alafs*94)**	1	0.075
c.810_814delACATG	***p.(His271Ilefs*10)***	1	0.075
c.822_832del	p.(Lys274Asnfs*5)	1	0.075
c.1282delC	p.(Gln428Serfs*24)	1	0.075
c.1355dupA	p.(*453Valext*36)	2	0.15
Ex3del4765		13	0.977
Ex5del955		3	0.225
Ex5del4232ins268		25	1.879

Mutations mentioned in bold letters are novel and not described so far. Mutations mentioned in italic bold letters were also identified only in Czech HPA patients but were described in our previous studies and/or submitted to the PAH mutation database (http://www.pahdb.mcgill.cz).

features for each mutation and averaged these values over all mutations in the corresponding group, i.e. either in the non-PKU HPA or the classical PKU group. This showed that mutations associated with non-PKU HPA change on the average 1.7 of 9 analyzed features, while this value for mutations associated with classical PKU is 4.3 (Table 2). The comparison of the 9 analyzed features is also depicted graphically in Fig. 1. Further, we recorded the combination of determined features for each mutation (see last column in Table 2). In particular, we detected 7 unique combinations in the group of non-PKU HPA mutations and 13 in the group of PKU mutations. None of the 7 combinations found for non-PKU HPA mutations was seen for mutations with the classical PKU phenotype.

3.4. Characterization of newly identified PAH missense mutations and prediction of their likely effects

We investigated intramolecular contacts (H-bonds, salt bridges, and stacking interactions) in the 3D model of PAH and complemented this information by analysis of contacts observed in MD simulations of wt and mutant proteins (Table 3, Fig. 2). These contacts stabilize the spatial arrangement of the protein and their loss may lead to structural defects. We do not provide detailed analysis of the performed simulations because the simulated proteins were basically stable in the course of 50 ns (Supplementary Fig. 1). In addition, we evaluated the 9 sequence and structure features (see Materials and methods) for these missense mutations.

p.(Asn167Tyr)

Asn167 is substituted by the larger Tyr, which should not impact the local arrangement because AA167 is located on the protein surface. Side chain of Asn167 does not form key interaction within the protein structure (Table 3). Despite the fact that the substitution Asn→Tyr is evolutionally less likely, the measured physico-chemical properties associated with this mutation are not significantly changed (Table 2). Based on our analysis, we predict that Tyr can substitute Asn167 without large structural defects and that patients carrying this mutation have the non-PKU HPA phenotype (like our patient with the genotype p.[Asn167Tyr];[Arg408Trp]).

p.(Thr200Asn)

Thr200 is substituted by Asn of similar size. MD simulation of the protein T200N showed that Asn can partially compensate for the contact formed by Thr (Table 3). The physico-chemical properties analyzed are not significantly changed upon mutation (Table 2). We predict that Asn can substitute Thr200 without large structural defects and that patients with this mutation have non-PKU HPA phenotype (in concordance with our patient with the genotype p.[Thr200Asn];[Arg243*]).

p.(Asp229Gly)

Asp229 is located on the protein surface. The conservation of this AA is 56%. Gly is a smaller residue with a lower polarity (uncharged compared to Asp) and is known as a helix breaker (Table 2). In addition, the substitution Asp229→Gly leads to abolishment of the salt bridge with Arg176. Surprisingly, this interaction was not stable in the MD simulation of the wt protein, which may indicate that this contact is not important for the spatial arrangement of the protein. This idea is supported by information from the PAH database, where a patient with the genotype p.[Arg176Leu];[Arg408Trp] and the non-PKU HPA phenotype is described [49]. Based on this information, we can hypothesize that due to the difference in side chains between Arg and Leu at position 176, Leu does not maintain the salt bridge with Asp229. This indirectly indicates that the presence of the salt bridge Asp229–Arg176 is not critical for the protein structure. Taking into account the available information, we predict that p.(Asp229Gly) does not impact significantly the protein structure and that patients with this mutation have the non-PKU HPA phenotype (like our patient with the genotype p.[Asp229Gly];[Arg408Trp]).

p.(Phe263Ser)

Phe263 is highly conserved (85%) and positioned in the active site in one of the cofactor binding regions [30]. The substitution Phe→Ser is unfavourable because Ser is a smaller residue and more polar (Table 2). In addition, the substitution Phe263→Ser leads to loss of the stacking interaction (Table 3). Based on our analysis, we predict that p.(Phe263Ser) causes significant impairment of the protein structure and function, and that patients carrying this mutation have the classical PKU phenotype (like our patient with the genotype p.[Phe263Ser];[c.1066-11 G>A]). The mutation p.(Phe263Leu) has been reported previously [50] but its phenotype-genotype relationship was not determined due to insufficient data. We assume that the impact of the mutation p.(Phe263Leu) on the protein structure is similar to the impact of our mutation, i.e. it destabilizes the protein structure due to the lost stacking interaction with Phe294 and/or reduces the binding affinity of the protein to BH4 and Fe^{3+} ion.

p.(Leu358Phe)

The conservation of Leu358 is 53%, which is slightly above our detection threshold. Leu and Phe have similar properties. Even though, Phe is larger, AA358 is located on the protein surface, hence the substitution Leu→Phe should not affect the local arrangement. In the

Table 2
Sequence and structure features determined for selected missense mutations detected in our HPA patients (see also Fig. 1).

Mutation–number of determined features	Specific side chain contacts of wt AA	Occurrence of AA in the active site	Buriedness of wt AA	Volume change upon mutation at buried position	Charge change upon mutation at buried position	Polarity change upon mutation at buried position	Conservation (identity of wt AA in MSA > 50%)	Presence of Helix/Turn breaker	Mutation likelihood based on Blosum62 (Negative score)	
Mutations predominantly associated with non-PKU HPA										
p.(Ile65Val) -1			9	27			43			001000000
p.(Arg169His) -1			43	20			47			100000000
p.(Glu178Gly) -1			96	78			45			000000001
p.(His201Tyr) -1			17	40			47			100000000
p.(Val245Ala) -5			5	51			63			011101100
p.(Ala300Ser) -2			0	0			86			001000100
p.(Ile306Val) -2			1	27			68			001000100
p.(Thr380Met) -2			73	47			53			000000101
p.(Ala403Val) -1			16	51			51			000001100
p.(Asp415Asn) -1			19	3			41			100000000
Average value is 1.7										
Mutations predominantly associated with classical PKU										
p.(Phe39Leu) -2			2	23			18			101000000
p.(Ile65Thr) -4			9	51			43			001101001
p.(Arg158Gln) -5			5	30			78			101110100
p.(Pro225Thr) -4			0	3			91			001000111
p.(Arg243Gln) -5			2	30			63			101110100
p.(Arg252Trp) -7			5	55			76			101111101
p.(Arg270Lys) -3			5	5			94			101000100
p.(Pro281Leu) -7			32[a]	54			92			011101111
p.(Ile283Phe) -2			0	23			25			011000000
p.(Phe299Cys) -6			0	84			67			111100101
p.(Tyr325Cys) -5			5	85			73			011100101
p.(Ser349Pro) -3			11	24			92			110000100
p.(Tyr387His) -5			5	40			22			101111000
p.(Ala395Pro) -4			2	24			62			001000111
p.(Arg408Trp) -3			15	54			60			100000101
Average value is 4.3										
Newly determined mutations associated with non-PKU HPA										
p.(Asn167Tyr) -2			68	80			34			100000001
p.(Thr200Asn) -1			55	2			44			100000000
p.(Asp229Gly) -4			40	51			56			100000111
p.(Leu358Phe) -1			32	23			53			000000100
p.(Ile406Met) -1			6	4			46			001000000
Average value is 1.8										
Newly determined mutation associated with classical PKU										
p.(Phe263Ser) -7			0	101			85			111101101

Grey fields show detected features (see Materials and methods for definition of features). The last column shows the combination of determined features for each mutation in binary code, i.e. 1 indicates presence of the feature while 0 indicates its absence. Each combination has a unique color. [a]Pro281 was considered buried because it is localized on the inner surface of the active site.

MD simulation of the mutant protein L358F, Phe358 formed a new stacking interaction with His201, which lies in Cα3 and is H-bonded to Thr200 (see above and Fig. 2). This original contact was not disrupted upon formation of the new stacking interaction between Phe358 and His201. We assume that the new stacking interaction may reduce the local flexibility of the protein, but should not significantly impact the protein function. We predict that patients carrying this mutation have the non-PKU HPA phenotype. Our non-PKU HPA patient carries this mutation together with the other non-PKU HPA mutation p.(Ile306Val).

p.(Ile406Met)

Ile406 is located near the protein surface. Both Ile and Met have a similar size and properties (Table 2). Ile406, as well as Met406, does not form specific interactions with surrounding residues and the substitution Ile→Met does not disrupt local protein arrangement. We predict that the impact of p.(Ile406Met) mutation on the protein architecture is small and that patients with this mutation have the non-PKU HPA phenotype (similarly to our patients with the genotype p.[Ile406Met];[Arg408Trp]).

The 9 sequence and structure features determined for 6 novel missense mutations are summarized in Table 2. The mutations p.(Asn167Tyr), p.(Thr200Asn), p.(Asp229Gly), p.(Leu358Phe), and p.(Ile406Met), predicted to cause non-PKU HPA and associated with a non-PKU HPA or mild PKU phenotype in our HPA patients, exhibit on the average 1.8 of the 9 analyzed features, which agrees well with the value of 1.7 obtained for mutations associated with the non-PKU HPA phenotype (Table 2). The mutation p.(Phe263Ser), predicted to cause classical PKU, exhibits 7 of 9 determined features, which can be seen for other mutations that are associated with classical PKU phenotype (Table 2). Further, for the five mutations predicted to cause the non-PKU HPA phenotype, we observed five unique combinations of determined features (see last column in Table 2). More importantly, three of these combinations have been detected for mutations predominantly associated with non-PKU HPA.

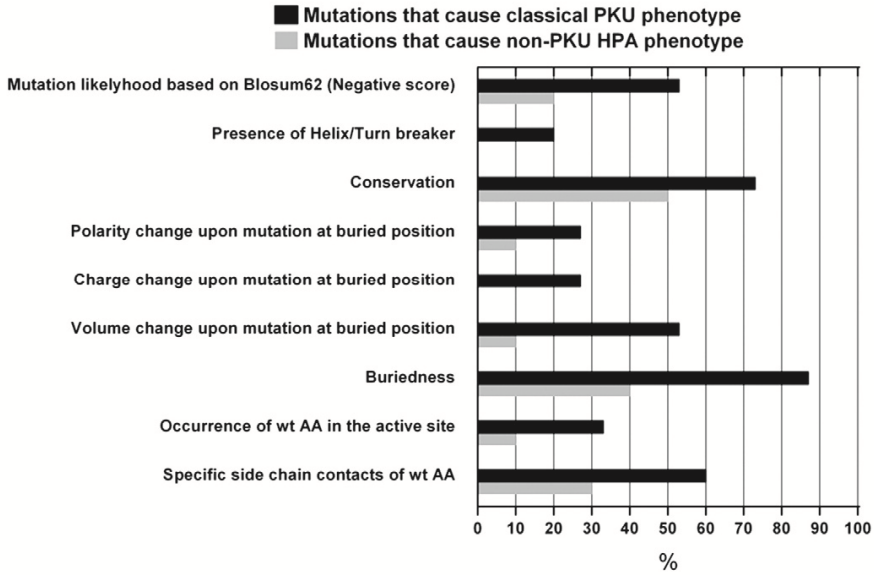


Fig. 1. Comparison of sequence and structure features determined for selected mutations associated with non-PKU HPA and the classical PKU phenotype.

3.5. Evaluation of newly identified PAH missense mutations using Polyphen-2, SIFT, SNPs3D, and FOLDX programs

The functional impact of novel missense mutations was evaluated using Polyphen2, SIFT, SNPs3D, and FOLDX programs (Supplementary Table 1). SIFT evaluated all novel mutations as "non-tolerated" all with score ≤0.05 which is indicative of deleterious substitution. SNPs3D indicated the mutation p.(Phe263Ser) as the most deleterious with a negative score of 3.6. This correlates with our structural analysis where we proposed that the mutation p.(Phe263Ser) has the largest impact on the PAH structure. Variance in predictions was seen for Polyphen-2, as the sequence database used in this software is dependent on the actual version of UniRef100 database which is continuously updated. This is a well known problem affecting all prediction methods relying on MSA for their conservation estimates (for more information see http://genetics.bwh.harvard.edu/pph2/dokuwiki/faq#inconsistent_predictions). Our first prediction using Polyphen-2 determined the novel mutations as either "benign" or "possibly damaging", with the exception of p.(Phe263Ser) that was indicated as "probably damaging". In the second prediction using Polyphen-2 with an updated database, mutations p.(Leu358Phe) and p.(Ile406Met) were predicted as "probably damaging" in contrast to "possibly damaging" in the first prediction. With the use of FOLDX all mutations were found to be either destabilizing or neutral (Supplementary Table 1). The significantly largest change of ΔΔG (4.14 kcal/mol) was detected for the mutation p.(Phe263Ser) in accord with SNPs3D analysis and our structural predictions. In summary, the first prediction with Polyphen-2 and the predictions with SNPs3D and FOLDX programs found the p.(Phe263Ser) mutation as the most deleterious which is strongly indicative. However, results for other mutations differ depending on the use of particular program, i.e. each program determines different order of deleteriousness (see Supplementary Table 1). This inconsistency is caused by limitations and different strategies of each method which are briefly explained in the 2.6. of Materials and methods section.

4. Discussion

We performed molecular genetic analysis of 665 unrelated Czech HPA patients. In 1323 identified mutant alleles, 98 types of mutations were detected (see Table 1A and 1B) and of these 21 types were described only in Czech HPA patients. In the Czech Republic, three mutations account for 51.3% of mutant HPA alleles [p.(Arg408Trp), 42.1%; p.(Ala403Val), 5.1%, and p.(Arg158Gln), 4.1%] in line with data for other Central European countries. Three most prevalent mutations account for i) 79.2% of the mutant alleles in Western Poland [p.(Arg408Trp), 68%; c.1066-11 G>A, 6%; c.1315+1 G>A, 5.2%] [51]; ii) 41.2% of the mutant alleles in Austria [p.(Arg408Trp), 22.8%; c.1315+1 G>A, 11.6%; p(Arg261Gln), 6,8%] [52]; iii) 41.3% of the mutant alleles in Southern Germany [p.(Arg408Trp), 23.2%; c.1315+1 G>A, 10.4%; p.(Ala403Val), 7.7%] and 42.2% of the mutant alleles in Western Germany [p.(Arg408Trp), 26.7%; c.1315+1 G>A, 10%; p.(Tyr414Cys), 5.5%] [53]. In 46 HPA patients, BH4 loading test was performed. In accordance with published data most of BH4 responders are patients with the mutation associated with non-PKU HPA or mild PKU phenotypes but there is no absolute correlation between genotype and BH4 responsiveness [3–8]. From the total amount of detected mutations missense mutations were responsible for 78% (1032, 60 types). Most of them have been characterized in structural analysis by Erlandesen and Stevensen [24] based on a composite model derived from X-ray structures of truncated forms of PAH. This strategy, molecular modelling of mutation effects *in silico*, was also used in other studies of HPA, e.g. when describing new mutations or complementing experimental data [10,54,55].

In the present study, we carried out molecular modelling of the consequences of HPA missense mutations in order to better understand their character and phenotypic manifestations. In particular, we classified our missense mutations into a group associated predominantly with the non-PKU HPA phenotype or a group associated predominantly with the classical PKU phenotype. Both groups were subjected to analysis of 9 sequence and structural features which indicate potential causality of a mutation (see Materials and methods for more details). Our analysis showed that the occurrence of these features is ~2.5 times more frequent for classical PKU mutations than for non-PKU HPA mutations (Table 2), confirming the hypothesis that the classical PKU mutations are more non-conservative than the non-PKU HPA mutations, which subsequently leads to more severe structural and functional defects in the PAH protein and a severe HPA phenotype. These conclusions are not absolutely unequivocal. In the group of non-PKU HPA mutations we found several with one

Table 3
Structural analysis of six novel missense mutations (see also Fig. 2).

Mutation	Localisation of mutated AA	Side chain contacts in the wt structure/development of these contacts in MD simulation of wt protein	Side chain contacts of mutated AA in the MD simulations of the mutant proteins
p.(Asn167Tyr)	The protein surface, at the end of Cα2b[a].	The H-bond Asn167(ND2)–Asp163(O) stabilizes Cα2b/ this H-bond fluctuates in the simulation.	New side chain H-bond between Tyr167 and Asp163(O) did not develop.
p.(Thr200Asn)	The protein surface, the end of Cα3.	The H-bond Thr200(OG1)–His201(ND1) stabilizes the Cα3 terminal part/this H-bond is stable in the simulation.	The new side chain H-bond Asn200(ND2)–His201(ND1) is formed replacing the original H-bond Thr200(OG1)–His201(ND1). The new H-bond fluctuated.
p.(Asp229Gly)	The protein surface, the Cα5.	Asp229 forms a salt bridge with Arg176 localized in the long turn structure between Cα2 and Cα3/ this salt bridge is not stable in the MD simulation.	The salt bridge with Arg176 is abolished.
p.(Phe263Ser)	Active site of the protein, cofactor binding region, Cβ2;	Phe263 stacks with Phe294 from Cα8/ this stacking interaction is stable.	The mutation disrupts the stacking interaction with Phe294. Large structural rearrangement of the active site was not seen in MD simulation (probably due to the short time scale).
p.(Leu358Phe)	The protein surface, the end of Cα11.	Leu358 is surrounded by Leu333, Ile340, and His201 from Cα3.	The new stacking interaction between rings of Phe358 and His201 is formed.
p.(Ile406Met)	Located near the protein surface, in the loop connecting the T and C domains.	Ile406 is surrounded by helices Cα8, Cα9, and Cα12.	No specific interaction between Met406 and surrounding residues was observed.

[a] Cα2b refers to α-helix 2b in the C domain, we used the secondary structure assignment of the human PAH sequence according to [24].

detected feature and simultaneously one with 5 detected features [p.(Val245Ala)] (Table 2). Similarly, in the group of classical PKU mutations there are some with 2 detected features [p.(Ile283Phe) and p.(Phe39Leu)] and some with 7 detected features [p.(Pro281Leu) and p.(Arg252Trp)] (Table 2). This shows that the range of detected features for individual mutations is wide in both groups, and hence the number of determined features is not necessarily a predictor of the phenotype. Further, this analysis revealed that non-PKU HPA and classical PKU mutations exhibit different combinations of determined features (see last column in Table 2). This may be partly due

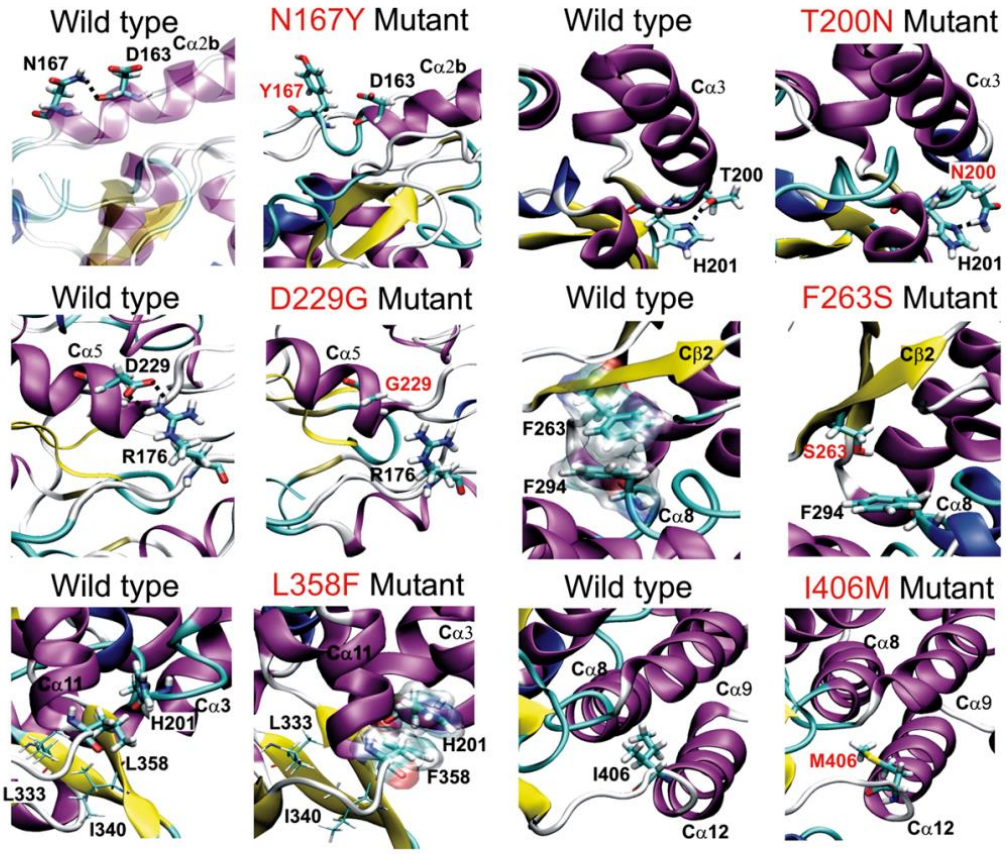


Fig. 2. Six detailed views of the tertiary structure of mutant PAH proteins carrying newly discovered missense mutations from MD simulations (right) and the corresponding region in the wt protein (left). Wild type and substituting AAs are highlighted in licorice representation and marked. H-bonds between residues are indicated by dashed black lines and stacked residues are in transparent surface representation. Key secondary structure elements are also indicated.

to the insufficient amount of data we used for this analysis (there are 10 mutations with the non-PKU HPA phenotype and 15 mutations with the classical PKU phenotype), but it may also indicate that certain combinations of parameters are more critical and lead to a severe HPA phenotype. In particular, the most frequent combination of detected features associated with the classical PKU phenotype is buriedness and the formation of specific side chain contact. In the group of non-PKU HPA mutations there is none with this combination while in the group of classical PKU mutations it is exhibited in half of them. This reflects the fact that AAs forming intramolecular contacts inside the protein can be hardly substituted without perturbation. Additional combination found exclusively for classical PKU mutations is the formation of specific side chain contact and the presence of AA in the active site. Severe phenotype is also frequently associated with buriedness and the change of i) volume, ii) charge or iii) polarity. This agrees with a previous study [56] where combinations of specific structural (stability) features were used to discriminate between disease and non-disease variants, and is also in accord with our observation that some combinations of detected features of new HPA mutations predicted to cause a non-PKU HPA phenotype (see below) are only seen for non-PKU HPA mutations (Table 2).

In this study, we also investigated the impact of the newly detected HPA missense mutations based on the analysis of 9 sequence and structure features and 50 ns-long MD simulations of PAH wt and mutant proteins. MD simulations are widely used in the study of biomolecules and provide atomistic insights into molecular flexibility and structural dynamics on ns→μs time scales [57–59]. Even though the time scale of our simulations was not sufficient to reveal protein unfolding or large structural changes due to mutations [60], it showed local rearrangements within the protein which indicated the impact of an AA substitution. Such an approach was recently applied in studies of effects of mutations in the E-box motif from group B *Streptococcus* [61,62] and of the human prion protein [62]. In spite of the fact that MD simulations are time-consuming compared to bioinformatic analysis available through various web applications, they provide a deeper view of the effects of AA substitution. However, they also have limitations given mainly by force field and time scale, and therefore the results of simulations should be taken with care. For instance utilization of MD simulations to recover PAH mutant structure in the presence of BH4 is barely feasible at this moment. Such calculations are limited by extensively long time scale (~μs) and by the quality of the force field. Without extensive testing they seem to be rather risky.

Based on these structural analyses, our clinical data, and phenotype-genotype information from the PAH mutation database the likely effects of novel mutations were predicted. In particular, we propose that the mutation p.(Phe263Ser) significantly impacts the structure of PAH so that patients with this mutation consequently have the classical PKU phenotype. Other novel mutations p.(Asn167Tyr), p.(Thr200Asn), p.(Asp229Gly), p.(Leu358Phe), and p.(Ile406Met) were found to be less deleterious, and we propose that they do not significantly affect the protein structure and function so that patients carrying these mutations consequently have a non-PKU HPA or mild PKU phenotype. Our results concerning the mutation p.(Phe263Ser) are in agreement with published data from a mouse model of HPA and with *in vitro* analysis of murine recombinant PAH proteins; the $Pah^{enu2/2}$ mouse carries the mutation p.(Phe263Ser) on both *Pah* alleles and is characterized by the classical PKU phenotype [63–65], and COS-7 cells expressing a recombinant murine protein carrying p.(Phe263Ser) show almost no PAH activity [66]. Structural consequences of novel mutations were also evaluated using standard prediction programs (Polyphen-2, SIFT, SNPs3D and FOLDX) where SNPs3D and FOLDX programs identified p.(Phe263Ser) as the most deleterious mutation in agreement with our structural predictions.

Acknowledgments

This study was supported by the project CEITEC (CZ.1.05/1.1.00/02.0068) and the project SuPReMMe (CZ.1.07/2.3.00/20.0045). The access to the MetaCentrum computing facilities provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" LM2010005 funded by the Ministry of Education, Youth, and Sports of the Czech Republic is acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cca.2013.01.006.

References

[1] Trefz FK, Scheible D, Frauendienst-Egger G, Korall H, Blau N. Long-term treatment of patients with mild and classical phenylketonuria by tetrahydrobiopterin. Mol Genet Metab 2005;86(Suppl. 1):S75–80.
[2] Kure S, Hou DC, Ohura T, et al. Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. J Pediatr 1999;135:375–8.
[3] Blau N, Erlandsen H. The metabolic and molecular bases of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. Mol Genet Metab 2004;82:101–11.
[4] Levy HL, Milanowski A, Chakrapani A, et al. Efficacy of sapropterin dihydrochloride (tetrahydrobiopterin, 6R-BH4) for reduction of phenylalanine concentration in patients with phenylketonuria: a phase III randomised placebo-controlled study. Lancet 2007;370:504–10.
[5] Muntau AC, Roschinger W, Habich M, et al. Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria. N Engl J Med 2002;347:2122–32.
[6] Bernegger C, Blau N. High frequency of tetrahydrobiopterin-responsiveness among hyperphenylalaninemias: a study of 1,919 patients observed from 1988 to 2002. Mol Genet Metab 2002;77:304–13.
[7] Fiege B, Blau N. Assessment of tetrahydrobiopterin (BH4) responsiveness in phenylketonuria. J Pediatr 2007;150:627–30.
[8] Burton BK, Grange DK, Milanowski A, et al. The response of patients with phenylketonuria and elevated serum phenylalanine to treatment with oral sapropterin dihydrochloride (6R-tetrahydrobiopterin): a phase II, multicentre, open-label, screening study. J Inherit Metab Dis 2007;30:700–7.
[9] Pey AL, Martinez A. The activity of wild-type and mutant phenylalanine hydroxylase and its regulation by phenylalanine and tetrahydrobiopterin at physiological and pathological concentrations: an isothermal titration calorimetry study. Mol Genet Metab 2005;86(Suppl. 1):S43–53.
[10] Gersting SW, Kemter KF, Staudigl M, et al. Loss of function in phenylketonuria is caused by impaired molecular motions and conformational instability. Am J Hum Genet 2008;83:5–17.
[11] Muntau AC, Gersting SW. Phenylketonuria as a model for protein misfolding diseases and for the development of next generation orphan drugs for patients with inborn errors of metabolism. J Inherit Metab Dis 2010;33:649–58.
[12] Aguado C, Perez B, Ugarte M, Desviat LR. Analysis of the effect of tetrahydrobiopterin on PAH gene expression in hepatoma cells. FEBS Lett 2006;580:1697–701.
[13] Staudigl M, Gersting SW, Danecka MK, et al. The interplay between genotype, metabolic state and cofactor treatment governs phenylalanine hydroxylase function and drug response. Hum Mol Genet 2011;20:2628–41.
[14] Erlandsen H, Pey AL, Gamez A, et al. Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. Proc Natl Acad Sci USA 2004;101:16903–8.
[15] Giovannini M, Verduci E, Salvatici E, Fiori L, Riva E. Phenylketonuria: dietary and therapeutic challenges. J Inherit Metab Dis 2007;30:145–52.
[16] Trefz FK, Scheible D, Gotz H, Frauendienst-Egger G. Significance of genotype in tetrahydrobiopterin-responsive phenylketonuria. J Inherit Metab Dis 2009;32:22–6.
[17] Zurfluh MR, Zschocke J, Lindner M, et al. Molecular genetics of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. Hum Mutat 2008;29:167–75.
[18] Fusetti F, Erlandsen H, Flatmark T, Stevens RC. Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. J Biol Chem 1998;273:16962–7.
[19] Knappskog PM, Flatmark T, Aarden JM, Haavik J, Martinez A. Structure/function relationships in human phenylalanine hydroxylase. Effect of terminal deletions on the oligomerization, activation and cooperativity of substrate binding to the enzyme. Eur J Biochem 1996;242:813–21.
[20] Waters PJ, Parniak MA, Akerman BR, Scriver CR. Characterization of phenylketonuria missense substitutions, distant from the phenylalanine hydroxylase active site, illustrates a paradigm for mechanism and potential modulation of phenotype. Mol Genet Metab 2000;69:101–10.
[21] Pey AL, Stricher F, Serrano L, Martinez A. Predicted effects of missense mutations on native-state stability account for phenotypic outcome in phenylketonuria, a paradigm of misfolding diseases. Am J Hum Genet 2007;81:1006–24.
[22] Gjetting T, Petersen M, Guldberg P, Guttler F. In vitro expression of 34 naturally occurring mutant variants of phenylalanine hydroxylase: correlation with metabolic phenotypes and susceptibility toward protein aggregation. Mol Genet Metab 2001;72:132–43.

[23] Kobe B, Jennings IG, House CM, et al. Structural basis of autoregulation of phenylalanine hydroxylase. Nat Struct Biol 1999;6:442–8.
[24] Erlandsen H, Stevens RC. The structural basis of phenylketonuria. Mol Genet Metab 1999;68:103–25.
[25] Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9.
[26] Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res 2001;11:863–74.
[27] Yue P, Melamud E, Moult J. SNPs3D: candidate gene and SNP selection for association studies. BMC Bioinformatics 2006;7:166.
[28] Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 2002;320:369–87.
[29] Kozak L, Hrabincova E, Kintr J, et al. Identification and characterization of large deletions in the phenylalanine hydroxylase (PAH) gene by MLPA: evidence for both homologous and non-homologous mechanisms of rearrangement. Mol Genet Metab 2006;89:300–9.
[30] Erlandsen H, Stevens RC. A structural hypothesis for BH4 responsiveness in patients with mild forms of hyperphenylalaninaemia and phenylketonuria. J Inherit Metab Dis 2001;24:213–30.
[31] Humphrey W, Dalke A, Schulten K. VMD—visual molecular dynamics. J Mol Graph Model 1996;14:33–8.
[32] Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
[33] Gresh N, Sponer JE, Spackova N, Leszczynski J, Sponer J. Theoretical study of binding of hydrated Zn(II) and Mg(II) cations to 5′-guanosine monophosphate. Toward polarizable molecular mechanics for DNA and RNA. J Phys Chem B 2003;107:8669–81.
[34] Case DA. AMBER 10. San Francisco: University of California; 2008.
[35] Mirkovic N, Marti-Renom MA, Weber BL, Sali A, Monteiro AN. Structure-based assessment of missense mutations in human BRCA1: implications for breast and ovarian cancer predisposition. Cancer Res 2004;64:3790–7.
[36] Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat 2001;17:263–70.
[37] Chothia C. The nature of the accessible and buried surfaces in proteins. J Mol Biol 1976;105:1–12.
[38] Zamyatin AA. Protein volume in solution. Prog Biophys Mol Biol 1972;24:107–23.
[39] Bordo D, Argos P. Suggestions for "safe" residue substitutions in site-directed mutagenesis. J Mol Biol 1991;217:721–9.
[40] Grantham R. Amino acid difference formula to help explain protein evolution. Science 1974;185:862–4.
[41] Luu TD, Rusu AM, Walter V, et al. MSV3d: database of human MisSense variants mapped to 3D protein structure. Database (Oxford) 2012, http://dx.doi.org/10.1093/database/bas1018.
[42] Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. PNAS 1992;89:10915–9.
[43] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–25.
[44] Dang LX, Kollman PA. Free energy of association of the K+: 18-crown-6 complex in water: a new molecular dynamics study. J Phys Chem 1995;99:55–8.
[45] Smith DE, Dang LX. Computer simulations of NaCl association in polarizable water. J Chem Phys 1994;100:3757–66.
[46] Spackova N, Reblova K, Sponer J. Structural dynamics of the Box C/D RNA kink-turn and its complex with proteins: the role of the A-minor 0 interaction, long-residency water bridges, and structural ion-binding sites revealed by molecular simulations. J Phys Chem B 2010;114:10581–93.
[47] Kozak L, Kuhrova V, Blazkova M, et al. Phenylketonuria mutations and their relation to RFLP haplotypes at the PAH locus in Czech PKU families. Hum Genet 1995;96:472–6.
[48] Kozak L, Blazkova M, Kuhrova V, Pijackova A, Ruzickova S, St'astna S. Mutation and haplotype analysis of phenylalanine hydroxylase alleles in classical PKU patients from the Czech Republic: identification of four novel mutations. J Med Genet 1997;34:893–8.
[49] Guldberg P, Henriksen FF, Thony B, Blau N, Güttlet F. Molecular heterogeneity of nonphenylketonuria hyperphenylalaninemia in 25 Danish patients. Genomics 1994;21:453–5.
[50] Takarada Y, Kalanin J, Yamashita K, Ohtsuka N, Kagawa S, Matsuoka A. Phenylketonuria mutant alleles in different populations: missense mutation in exon 7 of phenylalanine hydroxylase gene. Clin Chem 1993;39:2354–5.
[51] Dobrowolski SF, Borski K, Ellingson CC, Koch R, Levy HL, Naylor EW. A limited spectrum of phenylalanine hydroxylase mutations is observed in phenylketonuria patients in western Poland and implications for treatment with 6R tetrahydrobiopterin. J Hum Genet 2009;54:335–9.
[52] Sterl E, Paul K, Paschke E, et al. Prevalence of tetrahydrobiopterine (BH4)-responsive alleles among Austrian patients with PAH deficiency: comprehensive results from molecular analysis in 147 patients. J Inherit Metab Dis 2013;36:7–13.
[53] Aulehla-Scholz C, Heilbronner H. Mutational spectrum in German patients with phenylalanine hydroxylase deficiency. Hum Mutat 2003;21:399–400.
[54] Daniele A, Scala I, Cardillo G, et al. Functional and structural characterization of novel mutations and genotype–phenotype correlation in 51 phenylalanine hydroxylase deficient families from Southern Italy. FEBS J 2009;276:2048–59.
[55] Pey AL, Desviat LR, Gamez A, Ugarte M, Perez B. Phenylketonuria: genotype–phenotype correlations based on expression analysis of structural and functional mutations in PAH. Hum Mutat 2003;21:370–8.
[56] Yue P, Li Z, Moult J. Loss of protein structure stability as a major causative factor in monogenic disease. J Mol Biol 2005;353:459–73.
[57] Cheatham III TE. Simulation and modeling of nucleic acid structure, dynamics and interactions. Curr Opin Struct Biol 2004;14:360–7.
[58] Ditzler MA, Otyepka M, Sponer J, Walter NG. Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in. Acc Chem Res 2010;43:40–7.
[59] Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE. Long-timescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol 2009;19:120–7.
[60] Shaw DE, Maragakis P, Lindorff-Larsen K, et al. Atomic-level characterization of the structural dynamics of proteins. Science 2010;330:341–6.
[61] Cozzi R, Nuccitelli A, D'Onofrio M, et al. New insights into the role of the glutamic acid of the E-box motif in group B Streptococcus pilus 2a assembly. FASEB J 2012;26:2008–18.
[62] Guo J, Ning L, Ren H, Liu H, Yao X. Influence of the pathogenic mutations T188K/R/A on the structural stability and misfolding of human prion protein: insight from molecular dynamics simulations. Biochim Biophys Acta 2012;1820:116–23.
[63] McDonald JD, Charlton CK. Characterization of mutations at the mouse phenylalanine hydroxylase locus. Genomics 1997;39:402–5.
[64] Sarkissian CN, Boulais DM, McDonald JD, Scriver CR. A heteroallelic mutant mouse model: a new orthologue for human hyperphenylalaninemia. Mol Genet Metab 2000;69:188–94.
[65] Shedlovsky A, McDonald JD, Symula D, Dove WF. Mouse models of human phenylketonuria. Genetics 1993;134:1205–10.
[66] Gersting SW, Lagler FB, Eichinger A, et al. Pahenu1 is a mouse model for tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency and promotes analysis of the pharmacological chaperone mechanism in vivo. Hum Mol Genet 2010;19:2039–49.

European Journal of Medical Genetics 54 (2011) 112–117

Contents lists available at ScienceDirect

European Journal of Medical Genetics

journal homepage: http://www.elsevier.com/locate/ejmg

Original article

Chimeric *CYP21A1P/CYP21A2* genes identified in Czech patients with congenital adrenal hyperplasia

Zuzana Vrzalová [a], Zuzana Hrubá [a], Eva Šťahlová Hrabincová [a], Slávka Vrábelová [a], Felix Votava [b], Stanislava Koloušková [c], Lenka Fajkusová [a,d,*]

[a] University Hospital Brno, Centre of Molecular Biology and Gene Therapy, Černopolní 9, CZ-62500 Brno, Czech Republic
[b] Department of Pediatrics, 3rd Faculty of Medicine, Charles University and University Hospital Kralovské Vinohrady, Šrobárova 50, CZ-10034 Prague, Czech Republic
[c] Department of Pediatrics, 2nd Faculty of Medicine, Charles University and University Hospital Motol, V Úvalu 84, CZ-15006 Prague, Czech Republic
[d] Masaryk University, Faculty of Science, Institute of Experimental Biology, Dept. Functional Genomics and Proteomics, Kamenice 5, CZ-62500 Brno, Czech Republic

ARTICLE INFO

Article history:
Received 30 May 2010
Accepted 13 October 2010
Available online 21 October 2010

Keywords:
Congenital adrenal hyperplasia
Steroid 21-hydroxylase gene
21-Hydroxylase deficiency
Chimeric gene

ABSTRACT

Congenital adrenal hyperplasia (CAH) comprises a group of autosomal recessive disorders caused by an enzymatic deficiency which impairs the biosynthesis of cortisol and, in the majority of severe cases, also the biosynthesis of aldosterone. Approximately 95% of all CAH cases are caused by mutations in the steroid 21-hydroxylase gene (*CYP21A2*). The *CYP21A2* gene and its inactive pseudogene (*CYP21A1P*) are located within the HLA class III region of the major histocompatibility complex (MHC) locus on chromosome 6p21.3. In this study, we describe chimeric *CYP21A1P/CYP21A2* genes detected in our patients with 21-hydroxylase deficiency (21OHD). Chimeric *CYP21A1P/CYP21A2* genes were present in 171 out of 508 mutated *CYP21A2* alleles (33.8%). We detected four types of chimeric *CYP21A1P/CYP21A2* genes: three of them have been described previously as CH-1, CH-3, CH-4, and one type is novel. The novel chimeric gene, termed CH-7, was detected in 21.4% of the mutant alleles. Possible causes of *CYP21A1P/CYP21A2* formation are associated with 1) high recombination rate in the MHC locus, 2) high recombination rate between highly homologous genes and pseudogenes in the *CYP21* gene area, and 3) the existence of *chi*-like sequences and repetitive minisatellite consensus sequences in *CYP21A2* and *CYP21A1P* which play a role in promoting genetic recombination.

1. Introduction

Congenital adrenal hyperplasia (CAH) is a group of autosomal recessive disorders caused by an enzymatic deficiency which impairs the biosynthesis of cortisol and, in the majority of severe cases, also the biosynthesis of aldosterone. Approximately 95% of all CAH cases are connected with 21-hydroxylase deficiency (21OHD) [36]. In the adrenal cortex, the steroid 21-hydroxylase (21OH) converts 17-hydroxyprogesterone into 11-deoxycortisol and progesterone into 11-deoxycorticosterone. These steroids are subsequently converted into cortisol and aldosterone, respectively.

According to the severity of the disease, three clinical forms of CAH are distinguished: classical salt wasting (SW-CAH), classical simple virilising (SV-CAH), and nonclassical (NC-CAH) [28]. A severe deficiency of 21OH leads to SW-CAH. In addition to having a severe cortisol deficiency, patients with SW-CAH do not synthesise enough aldosteron, and therefore are not able to maintain sodium homeostasis. In the female foetus, an androgen excess causes variable degrees of external genitalia virilisation, and consequently results in genital ambiguity in newborn females [13,20,36]. A residual activity of 21OH (aldosterone is synthetised) results in SV-CAH with external genitalia virilisation in females and signs of precocious pseudopuberty developed before the 8th year in female and male patients [4,13,36]. NC-CAH is associated with a moderate deficiency of 21OH and is manifested in later childhood or adolescence by hirsutism and decreased fertility or precocious pseudopuberty [13]. Patients with NC-CAH secrete aldosterone normally and most male patients diagnosed after puberty are entirely asymptomatic [27,36].

The steroid 21-hydroxylase gene (*CYP21A2*) and its inactive pseudogene (*CYP21A1P*) are located within the HLA class III region of the major histocompatibility complex (MHC) locus on chromosome 6p21.3. In this region, four tandemly arranged genes – serine/threonine kinase *RP*, complement *C4*, steroid 21-hydroxylase *CYP21*, and tenascin *TNX* – are organised as a genetic unit designated as the

* Corresponding author. Centre of Molecular Biology and Gene Therapy, University Hospital Brno Černopolní 9, CZ-62500 Brno, Czech Republic. Tel.: +420 532234625; fax: +420 532234623.
E-mail address: lenkafajkusova@volny.cz (L. Fajkusová).

1769-7212/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmg.2010.10.005

RCCX module [3,38]. In the RCCX bimodular haplotype (69% of 6p21.3 chromosomes in Caucasians), duplication of the RCCX module occurs and the orientation of genes from telomere to centromere is *RP1–C4A–CYP21A1P–TNXA–RP2–C4B–CYP21A2–TNXB*. The sequence length of the *RP1–C4A–CYP21A1P–TNXA–RP2–C4B–CYP21A2–TNXB* gene cluster is about 120 kb. Three pseudogenes (*CYP21A1P*, *TNXA* and *RP2*) located between the two *C4* loci do not encode functional proteins [3,38].

The functional *CYP21A2* gene and the *CYP21A1P* pseudogene each contains 10 exons, spaced over 3.4 kb [14,34]. Intergenic recombinations are responsible for about 95% of mutations associated with 21OHD: 75% of the intergenic recombinations are represented by microconversion events which result in the transfer of mutations normally present in the pseudogene to the functional gene. The remaining 25% of the intergenic recombinations result in *CYP21A2* gene deletions and deletions involving the 3′end of *CYP21A1P* and the 5′end of *CYP21A2* [24,33]. In *RP1–C4A–CYP21A1P–TNXA–RP2–C4B–CYP21A2–TNXB*, 26- or 32-kb deletions encompassing *TNXA*, *RP2*, *C4B*, and partial sequences of *CYP21A1P* and *CYP21A2* produce chimeric *CYP21A1P/CYP21A2* genes. The size of a deletion (26- resp. 32-kb) depends on the presence or absence of an endogenous retrovirus sequence HERV-K (6.7 kb) in intron 9 [40] of *C4B*.

The incidence of classical CAH in central Europe is 1/10,000 and the carrier rate is 1/50 [19,20]. In this study, we describe chimeric *CYP21A1P/CYP21A2* genes detected in Czech 21OHD patients.

2. Patients and methods

2.1. Subjects

Patients were sent to our laboratory for molecular analysis of the *CYP21A2* gene after endocrine and clinical evaluation. DNA samples were obtained from peripheral blood leukocytes by the standard salting-out method. All studies have been approved by the Ethical Board of the University Hospital Brno, and all patients and their family members gave their informed consent for DNA analysis.

2.2. Amplification of chimeric CYP21A1P/CYP21A2 genes

The amplification of *CYP21A1P/CYP21A2* was performed by Expand long-template system kit (Roche Diagnostics GmbH, Germany) using the 5′ primer AF1 specific for *CYP21A1P* and the 3′ primer 21BR specific for *CYP21A2* (Table 1) [24]. The PCR contained 250 ng of genomic DNA, 1× PCR buffer, 2.6 U Enzyme mix polymerase, 1.75 mM $MgCl_2$, 350 μM dNTP, and 0.3 μM primers in 50-μl reaction. The amplification was performed under the following cycling conditions: 94 °C for 2 min; followed by 9 cycles of 94 °C for 30 s, 59 °C for 30 s, and 68 °C for 2.5 min; followed by 19 cycles of 94 °C for 30 s, 59 °C for 30 s, and 68 °C for 2.5 min with 20 s extension in each cycle; and a final extension of 68 °C for 7 min. The 3.5-kb PCR product was generated. The heterozygous or homozygous status of *CYP21A1P/CYP21A2* was distinguished using MLPA (see 2.5.) and specific amplification of *CYP21A2* [24]. In the routine analysis, we also used PCRs utilizing specific primers and *Eco*RI digestion to distinguished PCR products from *CYP21* genes. *Eco*RI digestion of the PCR product from *CYP21A2* produces two fragments (1,3 and 2,2 kb) whereas digestion of the PCR products from *CYP21A1P* and *CYP21A1P/CYP21A2* (CH-1, CH-3 and CH-7) produces three fragments (0,5; 0,8; 2,2 kb) [25].

2.3. Determination of the types of chimeric CYP21A1P/CYP21A2 genes

Nested PCR, PCR-restriction fragment length polymorphism (RFLP), and DNA sequencing were applied to determine point mutations (p.Pro30Leu, c.290-13A/C > G, p.Gly110ValfsX21, p.Ile172Asn, p.Val237Glu, p.Val281Leu, p.Leu307PhefsX6, p.Gln318X, and p.Arg356Trp) which are normally present in the pseudogene but can occur in the chimera [24,25]. The presence of such mutations inside *CYP21A1P/CYP21A2* determines the type of chimeric gene.

Table 1
The sequences of primers.

Primers	Sequence (5′–3′ direction)
AF1	CCCAGGTCGGGGCGGACACCC
21BR	AATTAAGCCTCAATCCTCTGCAGCG
30F	CAGTCTACACAGCAGGAGGGATGGC
30R	AGCAAGTGCAAGAAGCCCGGGGCAAGCTG
110F	ATCAGTTCCCACCCTCCAGCCCCGA
110R	AGGGCTGAGCGGGTGAGCTTC
172F	GAGGAATTCTCTCTCCTCACCTGCAGCATTA
172R	AGTTGTCGTCCTGCCAGAAAAGGA
237F	AGCAGGCCATAGAGAAGAGGGATCACATCG
237R	ATGCAAAAGAACCCGCCTCATAGC
281F	TGCAGGAGAGCCTCGTGGCAGG
281R	GACGCACCTCAGGGTGGTGAAG
356F	GCTGGGGCAGGACTCCACCCGA
356R	GTGGGGCAAGGCTAAGGGCACAACTGGC
Promotor/1F*	ATGTGGAACCAGAAAGCTG
Promotor/1R*	CAGCCCAAGGTGGAGCCTGTAGATG
2F*	CTCCAAGAGGACCATTGAGGAAGCC
2R*	GAGCGGGTGAGCTTCTTGTG
3F*	AAGCTCTTGGGGGGCATATC
3R*	GGCTACTGTGAGAGGCGAGG
4F*	GTCAGCCTCGCCTCTCACAG
4R*	CAGTTCAGGACAAGGAGAGGCT
5F*	AGCCCCTCCCTGAGCCTCTC
5R*	AGCCTCTCCCTCCACCCCAG
6F*	TGGGTTGTAGGGGAGAGGCT
6R*	TAGCAATGCTGAGGCCGGTA
7F*	TGCCACTCTGTACTCCTCTC
7R*	ACAGTGCTCAGAGCTGAGTG
8F*	CTCACCGGCACTCAGGCTCA
8R*	AAGGGGGCTGGAGTTAGAGGCT

F: forward primer; R: reverse primer; primers with asterisk* were used for PCR-sequencing; primers without asterisk were used for PCR-RFLP. The reference sequence is published on http://www.ncbi.nlm.nih.gov/nuccore/187895.

The *CYP21A1P/CYP21A2* product of long-template PCR (see 2.2.) was used as the DNA template for the following nested PCRs. These reactions contained 150 ng of the long-template PCR product, 1× PCR buffer, 1.0 U Taq DNA polymerase (Fermentas GmbH, Germany), 1.5 mM MgCl, 200 μM dNTP, and 0.5 μM primers in 25-μl reaction. The amplifications were performed under the following cycling conditions: 94 °C for 5 min; followed by 20 cycles of 94 °C for 30 s, 59 °C for 30 s, 72 °C for 30 s; and a final extension of 72 °C for 7 min. Primers 30F/30R and PstI restriction endonuclease digestion were used for the detection of p.Pro30Leu; primers 110F/110R and SacI digestion for the detection of c.290-13A/C > G and the same primers also for the detection of the 8-nucleotide deletion p.Gly110ValfsX21; primers 172F/172R and MseI digestion were applied for the detection of p.Ile172Asn; primers 237F/237R and TaqI digestion for the detection of p.Val237Glu; primers 281F/281R and ApaLI digestion for the detection of Val281Leu; primers 356F/356R and PstI digestion for the detection of p.Gln318X and the same primers and MscI digestion for detection of p.Arg356Trp. The p.Leu307PhefsX6 mutation was analysed by PCR with primers 7F/7R, and the PCR product was sequenced on the ABI PRISM 310 sequencer (Applied Biosystems, USA) [24,25]. Sequences of primers are showed in Table 1. The determination of the type of *CYP21A1P/CYP21A2* was also confirmed by MLPA (see 2.5.).

2.4. Analysis of chimeric CYP21A1P/CYP21A2 genes by DNA sequencing

All chimeric *CYP21A1P/CYP21A2* genes have the *CYP21A1P* promoter and p.Pro30Leu but differ in the presence of other

mutations and polymorphisms (Table 2). The sequence analysis of chimeric genes was performed to determine the breakpoints in *CYP21A1P–CYP21A2* conversion areas. However, the precise determination of breakpoints cannot be achieved because of the long sequence identity between *CYP21A1P* and *CYP21A2*.

The *CYP21A1P/CYP21A2* product of long-template PCR (see 2.2.) was used as the DNA template for the detection of mutations and polymorphisms in the chimera (including the promoter sequence). The following nested PCR reactions contained 500 ng of the long-template PCR product, 1× PCR buffer, 1.0 U Taq DNA polymerase (Fermentas GmbH, Germany), 1.5 mM $MgCl_2$, 150 μM dNTP, and 0.35 μM sequencing primers in 30-μl reaction. The sequences of primers are showed in Table 1. The amplifications were performed under the following cycling conditions: 94 °C for 5 min; followed by 20 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s; and a final extension of 72 °C for 7 min. PCR products were sequenced by Big Dye Terminator kit and analysed on the ABI PRISM 310 sequencer (Applied Biosystems, USA).

2.5. MLPA

Multiplex ligation-dependent probe amplification (MLPA) is a method for the detection of large gene deletions and duplications. For CAH diagnostics, we used SALSA MLPA KIT P050B CAH (MRS

Table 2
The sequence analysis of chimeric *CYP21A1P/CYP21A2* genes (CH-1, CH-3, CH-4, and CH-7).

Sequence position at cDNA and protein level	*CYP21A2*	*CYP21A1P*	CH-1	CH-3	CH-4	CH-7
c.1-209T > C	T	**C**	**C**	**C**	**C**	**C**
c.1-198C > T	C	**T**	**T**	**T**	**T**	**T**
c.1-(188_189)insT	–	**+T**	**+T**	**+T**	**+T**	**+T**
c.1-126C > T	C	**T**	**T**	**T**	**T**	**T**
c.1-113G > A	G	**A**	**A**	**A**	**A**	**A**
c.1-110T > C	T	**C**	**C**	**C**	**C**	**C**
c.1-103A > G	A	**G**	**G**	**G**	**G**	**G**
c.1-4C > T	C	**T**	**T**	**T**	**T**	**T**
c.89C > T (Pro30Leu)	Pro	**Leu**	**Leu**	**Leu**	**Leu**	**Leu**
c.289 + 9T > C	T	**C**	**C**	**C**	T	**C**
c.289 + 57T > G	T	**G**	**G**	**G**	T	**G**
c.289 + 84A > G	A	**G**	**G**	**G**	A	**G**
c.289 + 92A > G	A	**G**	**G**	**G**	A	**G**
c.289 + 100A > G	A	**G**	**G**	**G**	A	**G**
c.289 + 116A > G	A	**G**	**G**	**G**	A	**G**
c.289 + 127T > G	T	**G**	**G**	**G**	T	**G**
c.289 + 138T > C	T	**C**	**C**	**C**	T	**C**
c.289 + 144A > T	A	**T**	**T**	**T**	A	**T**
c.289+(155_156)insTCC	–	**+TCC**	**+TCC**	**+TCC**	–	**+TCC**
c.289 + 160C > A	C	**A**	**A**	**A**	C	**A**
c.290–94T > A	T	**A**	**A**	**A**	T	**A**
c.290-95 G > C	G	**C**	**C**	**C**	G	**C**
c.290-96G > T	G	**T**	**T**	**T**	G	**T**
c.290-91G > A	G	**A**	**A**	**A**	G	**A**
c.290-87A > G	A	**G**	**G**	**G**	A	**G**
c.290-88G > A	G	**A**	**A**	**A**	G	**A**
c.290-79G > T	G	**T**	**T**	**T**	G	**T**
c.290-74G > A	G	**A**	**A**	**A**	G	**A**
c.290-67C > A	C	**A**	**A**	**A**	C	**A**
c.290-48A > G	A	**G**	**G**	**G**	A	**G**
c.290-44G > T	G	**T**	**T**	**T**	G	**T**
c.290-38A > G	A	**G**	**G**	**G**	A	**G**
c.290-39C > G	C	**G**	**G**	**G**	C	**G**
c.290-13A/C > G	A/C	**G**	**G**	**G**	A/C	**G**
c.324C > G (Ser108Ser)	C	**G**	**G**	**G**	C	**G**
c.329_336del8 (Gly110ValfsX21)	–	**Gly110Val fsX21**	**Gly110Val fsX21**	**Gly110Val fsX21**	–	**Gly110Val fsX21**
c.515T > C (Ile172Asn)	Ile	**Asn**	Ile	**Asn**	Ile	**Asn**
c.547-15C > A	C	**A**	C	**A**	C	**A**
c.547-8T > C	T	**C**	T	**C**	T	**C**
c.549C > G (Asp183Glu)	Asp	**Glu**	Asp	**Glu**	Asp	**Glu**
c.702T > C (Asp234Asp)	T	**C**	T	**C**	T	**C**
c.707T > A (Ile236Asn)	Ile	**Asn**	Ile	**Asn**	Ile	**Asn**
c.710T > A (Val237Glu)	Val	**Glu**	Val	**Glu**	Val	**Glu**
c.716T > A (Met239Lys)	Met	**Lys**	Met	**Lys**	Met	**Lys**
c.735 + 12A > G	A	**G**	A	**G**	A	**G**
c.735 + 13C > T	C	**T**	C	**T**	C	**T**
c.744C > G (Leu248Leu)	C	**G**	C	**G**	C	C
c.841G > T (Val281Leu)	Val	**Leu**	Val	Val	Val	Val
c.920_921insT	–	**+T**	–	**+T**	–	–
c.936 + 11G > C	G	**C**	G	**C**	G	G
c.952C > T (Gln318X)	Gln	**X**	Gln	**X**	Gln	Gln
c.1066C > T (Arg356Trp)	Arg	**Trp**	Arg	Arg	Arg	Arg

Nucleotides derived from the *CYP21A1P* pseudogene are showed by bold letters.
Nomenclature for mutations is in agreement with Nomenclature for the description of sequence variations published on http://www.hgvs.org/mutnomen/. Nucleotides associated with changes in coding regions are numbered relative to cDNA sequence Genbank RefSeq NM_000500.5 (http://www.ncbi.nlm.nih.gov/nuccore/67906817), nucleotide 1 is the first A of the ATG initiation codon. The polymorphic insertion c.27_28insCTG has been omitted from this sequence. Nucleotides associated with changes in noncoding regions are numbered relative to cDNA sequence (Genbank RefSeq NM_000500.5) and the arrangement of particular intron sequences was taken from the gDNA sequence (Genbank M12792.1, http://www.ncbi.nlm.nih.gov/nuccore/187895).

Holland, the Netherlands). This kit contains probes and primers for analysis of five *CYP21A2* gene fragments located in the promotor region and exons 3, 4, 6, and 8. The established MLPA method detects deletions and duplications of the *CYP21A2* gene and chimeric *CYP21A1P*/*CYP21A2* genes, but also particular point mutations provided that these mutations are located in places of the MLPA probe binding (c.1-126CT, c.1-113GA, c.1-110TC, c.1-103AG, p.Gly110ValfsX21, p.Ile172Asn, p.Ile236Asn, p.Val237Glu, p.Met239Lys, and p.Gln318X). Additionally, MLPA kit contains 28 *CYP21A2* nonspecific probes that are control standards and sets for amplification of adjacent genes (*CREBL*, *C4B*, *TNXB*) and pseudogenes (*CYP21A1P*, *C4A*, *TNXA*). The assay was performed according to the manufacturers' recommendations. Amplification products were run on the CEQ8000 Genetic Analyzer (Beckman Coulter, USA). Detected changes at the MLPA level were correlated with results obtained using Expand long-template PCR, nested PCRs, RFLP, and DNA sequencing. Identified mutations were also confirmed by segregation analyses of parents' alleles.

3. Results

Chimeric *CYP21A1P*/*CYP21A2* genes are the most frequent mutation in Czech 21OHD patients – 171 mutant alleles (33.8%) out of the total number of 508 carry this type of gene rearrangement. Chimeric *CYP21A1P*/*CYP21A2* genes can be distinguished according to the different extent of a 5′ part of *CYP21A1P* attached to a 3′part of *CYP21A2*. In the set of Czech 21OHD patients, four types of chimeric *CYP21A1P*/*CYP21A2* genes were detected: CH-1, CH-3, CH-4, and CH-7 (Fig. 1). CH-1 (9.3% of mutant alleles) includes the promoter and exons 1-3 from *CYP21A1P* (the mutations p.Pro30Leu, c.290-13A/C > G, and p.Gly110ValfsX21 are present in the chimera). CH-3 (0.4% of mutant alleles) comprises the promoter and exons 1–8 from *CYP21A1P* (the mutations p.Pro30Leu, c.290-13A/C > G, p.Gly110ValfsX21, p.Ile172Asn, p.Ile236Asn, p.Val237Glu, p.Met239-Lys, p.Leu307PhefsX6, and p.Gln318X are present in the chimera; the p.Val281Leu mutation is not present). CH-4 (2.6% of mutant alleles) possesses the *CYP21A1P* promoter and p.Pro30Leu. The

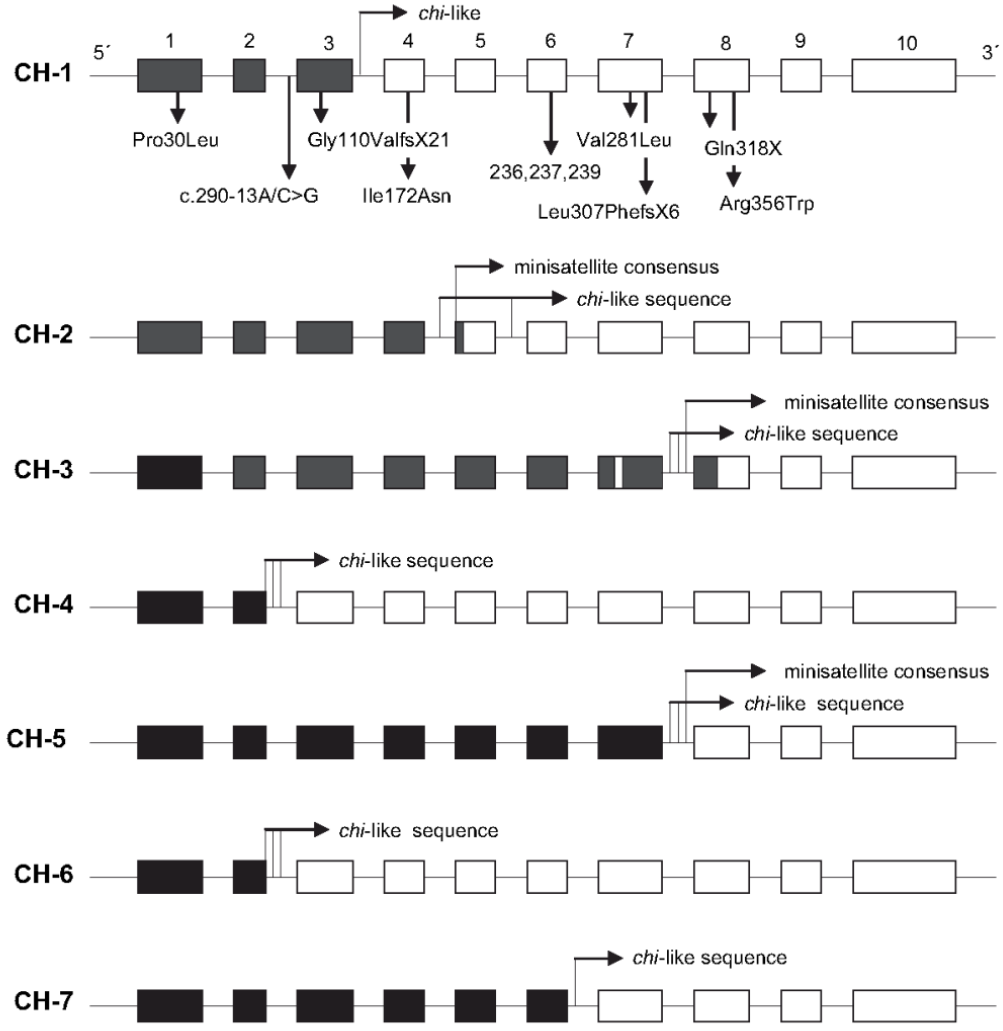


Fig. 1. Types of chimeric *CYP21A1P*/*CYP21A2* genes and localizations of *chi*-like sequences and minisatellite consensus sequences inside *CYP21A2* and *CYP21A1P*. The structure of the functional *CYP21A2* gene is depicted by white boxes; black boxes represent the nonfunctional *CYP21A1P* pseudogene. The arrows indicate mutations which exist in *CYP21A1P*. 236/237/239 depicts mutations p.Ile236Asn, p.Val237Glu, and p.Met239Lys; these mutations always occur together. The *CYP21A1P-CYP21A2* junction site is localised upstream of c.290-13A/C > G and between c.290-13A/C > G and p.Gly110ValfsX2 in case of CH-4 and CH-6, respectively.

most common type of a chimeric gene determined in our patients is a newly characterized chimeric gene CH-7 (21.4% of mutant alleles). CH-7 comprises the promoter and exon 1–6 from *CYP21A1P* (the mutations p.Pro30Leu, c.290-13A/C > G, p.Gly110ValfsX21, p.Ile172Asn, p.Ile236Asn, p.Val237Glu, p.Met239Lys are present in the chimera). The sequence analyses of chimeric genes determined in 5 patients with CH-1, CH-4, CH-7 and 2 patients with CH-3 are presented in Table 2.

CH-1, CH-3, and CH-7 are associated (in the homozygous state) with the SW-CAH phenotype. In these chimeric genes, c.290-13A/C > G (the splice-site mutation in intron 2) and p.Gly110ValfsX21 (the frame-shift mutation in exon 3) disrupt the production of any functional protein. In contrast, the *CYP21A1P–CYP21A2* junction site in CH-4 is located upstream of c.290-13A/C > G and the resulting chimeric gene differs from the functional gene by the presence of two mutations (the weak *CYP21A1P* promoter and p.Pro30Leu). The final protein, resulting from CH-4, retains minor 21-hydroxylase activity and is associated with NC-CAH or SV-CAH.

In the set of our 21OHD probands, we detected 7 patients with the combination CH-4 on one *CYP21* allele and the same or other chimeric gene on the second one: 1 proband with the genotype CH-4/CH-4 was clinically diagnosed as NC-CAH; 1 proband with CH-4/CH-1 and 4 probands with CH-4/CH-7 suffered from SV-CAH; and 1 proband with CH-4/CH-7 suffered from SW-CAH. The female SW-CAH patient with the genotype CH-4/CH-7 was diagnosed in the first week of life on the basis of metabolic disruption of the organism and external genitalia virilisation. Other patients with the genotype CH-4/CH-7 (1 girl and 3 boys) had the SV-CAH phenotype, with external genitalia virilisation in the female and signs of precocious pseudopuberty developed in all patients.

The MLPA technique has been used to estimate the copy number of *CYP21A2* and *CYP21A1P*. The determination of the *CYP21A1P* copy number has mostly no diagnostic importance for the determination of 21OHD because the number of *CYP21A1P* can vary, depending on the presence of one, two, three or even four RCCX modules on the chromosome [3]. In *CYP21A2*, mutations c.1-126CT, c.1-113GA, c.1-110TC, c.1-103AG (the promoter), p.Gly110ValfsX21 (exon 3), p.Ile172Asn (exon 4), p.Ile236Asn, p.Val237Glu, p.Met239Lys (exon 6), and p.Gln318X (exon 8) are located in places of the MLPA probe binding and so the presence of these mutations avoids the MLPA reaction. MLPA results obtained for an example in the genotype CH-4/CH-7 were: 0 copy of *CYP21A2* using the MLPA probe specific to the promoter; 1 copy of *CYP21A2* using probes specific to exons 3, 4, 6; and 2 copies of *CYP21A2* using the probe specific to exon 8.

4. Discussion

To date, 6 different chimeric *CYP21A1P/CYP21A2* genes (Fig. 1) have been found and characterised (CH-1, CH-2, CH-3 [22,23,26], CH-4 [21]; [18], CH-5 [35], CH-6 [8]). CH-1, CH-2, and CH-3 were described for the first time in Chinese patients in Taiwan, CH-4 and CH-5 in patients of Caucasian origin, and CH-6 in a patient of Italian origin. In Czech 21OHD patients, we identified four types of chimeric *CYP21A1P/CYP21A2* genes: three of them have been described previously (CH-1, CH-3, and CH-4), and one type is novel. The novel chimeric gene was designated as CH-7.

The 5′ untranslated region of *CYP21A2*, responsible for transcriptional activity, is located in the first 167 nucleotides upstream of the ATG codon and contains binding sites for the specificity protein Sp-1 and adrenal-specific protein transcription factors [16,17]. In this fragment, the pseudogene promoter differs from the *CYP21A2* promoter in four nucleotides, located at −126, −113, −110, and −103 positions. These differences account for a lower affinity of the pseudogene promoter to transcription factors, and consequently reduce its transcriptional activity to 20% when compared with *CYP21A2* [5,6]. The p.Pro30Leu mutation is associated with NC-CAH and reduces the 21OH activity to 30–40% of the normal enzyme. In CH-4, p.Pro30Leu in synergy with the *CYP21A1P* promoter decreases the enzyme activity to 4–10% [2]. Generally, 21OHD is an autosomal recessive disease and most 21OHD patients are compound heterozygotes, in which phenotypes reflect a mutation that is predicted to cause a less severe impairment of the enzymatic activity. This approach to phenotype prediction has been shown to be correct but deviations to this correlation exist [12,31,36]. Phenotypes of Czech 21OHD patients correspond generally to this rule. At least, patients carrying homozygous CH-4 have NC-CAH and patients carrying other types of chimeric genes (with the splicing mutation c.290-13A/C > G) on both *CYP21* alleles suffer from SW-CAH.

Both *CYP21A1P* and *CYP21A2* are located within the RCCX gene module of chromosome 6p21.3, adjacent to and alternating with *C4A* and *C4B* encoding the fourth components of the serum complement located within the HLA complex. The existence of tandemly repeated genes, in conjunction with a high recombination rate within the MHC locus, predispose this region to frequent non-allelic homologous recombination events [22]. Indeed, intergenic recombinations are responsible for 95% of the mutations associated with 21OHD. The frequency of genetic recombination in this locus may be further promoted by the presence of specific sequences, namely the short octanucleotide *chi*-like sequences and the decanucleotide minisatellite consensus sequences [15,22]. *Chi* (crossover hotspot instigator) sites had been originally described as cis-acting motifs 5′-GCTGGTGG-3′around which the rate of Rec-promoted recombinations is elevated [30]. *Chi* sites were extensively studied in *E. coli*, where this sequence is recognised by RecBCD enzyme, which possesses combined nuclease and helicase activity. It unwinds dsDNA and performs exonuclease degradation of both its strands. However, the degradation of the strand containing the 5′-GCTGGTGG-3′ sequence is strongly reduced upon interaction with *chi* during translocation, while the helicase activity of RecBCD enzyme is unaffected [9,10]. *Chi* and *chi*-like sequences have been suggested to act as recombination hot spots also in higher eukaryotic organisms [15,32]. In humans, they were found to be overrepresented at many translocation breakpoints associated with inherited diseases and cancer [1,37,39]. They have also been implicated in the generation of hybrid alleles coding for some rarely occurring blood subgroups [29]. Interestingly, 9 *chi*-like sequences are present in *CYP21A1P* or *CYP21A2* genes and these are frequently found in close proximity to the translocation breakpoints [22]. Together with the other determinants at the DNA and chromatin structure level (for example, preferential topoisomerase II cleavage sites [11], or sequences that tend to form non-B-DNA structures [7]), these sequence motifs apparently play an important role in site-specific recombination in human cells.

Acknowledgements

This work was supported by the Internal Grant Agency of the Czech Ministry of Health (NR9308-3 and NS9981-3) and by the Czech Ministry of Education (projects LC06023, MSM0021622415).

References

[1] S.S. Abeysinghe, N. Chuzhanova, M. Krawczak, E.V. Ball, D.N. Cooper, Translocation and gross deletion breakpoints in human inherited disease and cancer I: nucleotide composition and recombination-associated motifs, Hum. Mutat. 22 (2003) 229–244.

[2] R.S. Araujo, A.E. Billerbeck, G. Madureira, B.B. Mendonca, T.A. Bachega, Substitutions in the CYP21A2 promoter explain the simple-virilizing form of 21-hydroxylase deficiency in patients harbouring a P30L mutation, Clin. Endocrinol. (Oxf). 62 (2005) 132–136.

[3] C.A. Blanchong, B. Zhou, K.L. Rupert, E.K. Chung, K.N. Jones, J.F. Sotos, W.B. Zipf, R.M. Rennebohm, C. Yung Yu, Deficiencies of human complement component C4A and C4B and heterozygosity in length variants of RP-C4-CYP21-TNX (RCCX) modules in caucasians. The load of RCCX genetic diversity on major histocompatibility complex-associated disease, J. Exp. Med. 191 (2000) 2183–2196.
[4] A.O. Brinkmann, Molecular basis of androgen insensitivity, Mol. Cell. Endocrinol. 179 (2001) 105–109.
[5] J. Bristow, S.E. Gitelman, M.K. Tee, B. Staels, W.L. Miller, Abundant adrenal-specific transcription of the human P450c21A "pseudogene", J. Biol. Chem. 268 (1993) 12919–12924.
[6] S.F. Chang, B.C. Chung, Difference in transcriptional activity of two homologous CYP21A genes, Mol. Endocrinol. 9 (1995) 1330–1336.
[7] N. Chuzhanova, J.M. Chen, A. Bacolla, G.P. Patrinos, C. Ferec, R.D. Wells, D.N. Cooper, Gene conversion causing human inherited disease: evidence for involvement of non-B-DNA-forming sequences and recombination-promoting motifs in DNA breakage and repair, Hum. Mutat. 30 (2009) 1189–1198.
[8] P. Concolino, E. Mello, A. Minucci, E. Giardina, C. Zuppi, V. Toscano, E. Capoluongo, A new CYP21A1P/CYP21A2 chimeric gene identified in an Italian woman suffering from classical congenital adrenal hyperplasia form, BMC. Med. Genet. 10 (2009) 72.
[9] D.A. Dixon, S.C. Kowalczykowski, Homologous pairing in vitro stimulated by the recombination hotspot, Chi, Cell 66 (1991) 361–371.
[10] D.A. Dixon, S.C. Kowalczykowski, The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme, Cell 73 (1993) 87–96.
[11] J. Fajkus, J.A. Nicklas, R. Hancock, DNA loop domains in a 1.4-Mb region around the human hprt gene mapped by cleavage mediated by nuclear matrix-associated topoisomerase II, Mol. Gen. Genet. 260 (1998) 410–416.
[12] A. Friaes, A.T. Rego, J.M. Aragues, L.F. Moura, A. Mirante, M.R. Mascarenhas, T.T. Kay, L.A. Lopes, J.C. Rodrigues, S. Guerra, T. Dias, A.G. Teles, J. Goncalves, CYP21A2 mutations in Portuguese patients with congenital adrenal hyperplasia: identification of two novel mutations and characterization of four different partial gene conversions, Mol. Genet. Metab. 88 (2006) 58–65.
[13] J. Goncalves, A. Friaes, L. Moura, Congenital adrenal hyperplasia: focus on the molecular basis of 21-hydroxylase deficiency, Expert. Rev. Mol. Med. 9 (2007) 1–23.
[14] Y. Higashi, H. Yoshioka, M. Yamane, O. Gotoh, Y. Fujii-Kuriyama, Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: a pseudogene and a genuine gene, Proc. Natl. Acad. Sci. U S A 83 (1986) 2841–2845.
[15] A.J. Jeffreys, V. Wilson, S.L. Thein, Hypervariable 'minisatellite' regions in human DNA, Nature 314 (1985) 67–73.
[16] N. Kagawa, M.R. Waterman, cAMP-dependent transcription of the human CYP21B (P-450C21) gene requires a cis-regulatory element distinct from the consensus cAMP-regulatory element, J. Biol. Chem. 265 (1990) 11299–11305.
[17] N. Kagawa, M.R. Waterman, Evidence that an adrenal-specific nuclear protein regulates the cAMP responsiveness of the human CYP21B (P450C21) gene, J. Biol. Chem. 266 (1991) 11199–11204.
[18] A.A. Killeen, K.S. Sane, H.T. Orr, Molecular and endocrine characterization of a mutation involving a recombination between the steroid 21-hydroxylase functional gene and pseudogene, J. Steroid. Biochem. Mol. Biol. 38 (1991) 677–686.
[19] P.F. Koppens, T. Hoogenboezem, H.J. Degenhart, Carriership of a defective tenascin-X gene in steroid 21-hydroxylase deficiency patients: TNXB -TNXA hybrids in apparent large-scale gene conversions, Hum. Mol. Genet. 11 (2002) 2581–2590.
[20] J. Kovacs, F. Votava, G. Heinze, J. Solyom, J. Lebl, Z. Pribilincova, H. Frisch, T. Battelino, F. Waldhauser, Lessons from 30 years of clinical diagnosis and treatment of congenital adrenal hyperplasia in five middle European countries, J. Clin. Endocrinol. Metab. 86 (2001) 2958–2964.
[21] D. L'Allemand, V. Tardy, A. Gruters, D. Schnabel, H. Krude, Y. Morel, How a patient homozygous for a 30-kb deletion of the C4-CYP 21 genomic region can have a nonclassic form of 21-hydroxylase deficiency, J. Clin. Endocrinol. Metab. 85 (2000) 4562–4567.
[22] H.H. Lee, The chimeric CYP21P/CYP21 gene and 21-hydroxylase deficiency, J. Hum. Genet. 49 (2004) 65–72.
[23] H.H. Lee, S.F. Chang, Y.J. Lee, S. Raskin, S.J. Lin, M.C. Chao, F.S. Lo, C.Y. Lin, Deletion of the C4-CYP21 repeat module leading to the formation of a chimeric CYP21P/CYP21 gene in a 9.3-kb fragment as a cause of steroid 21-hydroxylase deficiency, Clin. Chem. 49 (2003) 319–322.
[24] H.H. Lee, J.G. Chang, C.H. Tsai, F.J. Tsai, H.T. Chao, B. Chung, Analysis of the chimeric CYP21P/CYP21 gene in steroid 21-hydroxylase deficiency, Clin. Chem. 46 (2000) 606–611.
[25] H.H. Lee, H.T. Chao, H.T. Ng, K.B. Choo, Direct molecular diagnosis of CYP21 mutations in congenital adrenal hyperplasia, J. Med. Genet. 33 (1996) 371–375.
[26] H.H. Lee, Y.J. Lee, P. Chan, C.Y. Lin, Use of PCR-based amplification analysis as a substitute for the southern blot method for CYP21 deletion detection in congenital adrenal hyperplasia, Clin. Chem. 50 (2004) 1074–1076.
[27] M.I. New, Extensive clinical experience: nonclassical 21-hydroxylase deficiency, J. Clin. Endocrinol. Metab. 91 (2006) 4205–4214.
[28] M.I. New, R.C. Wilson, Steroid disorders in children: congenital adrenal hyperplasia and apparent mineralocorticoid excess, Proc. Natl. Acad. Sci. U S A 96 (1999) 12790–12797.
[29] M.L. Olsson, M.A. Chester, Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies, Transfus. Med. 11 (2001) 295–313.
[30] G.R. Smith, S.M. Kunes, D.W. Schultz, A. Taylor, K.L. Triman, Structure of chi hotspots of generalized recombination, Cell 24 (1981) 429–436.
[31] N.M. Stikkelbroeck, L.H. Hoefsloot, I.J. de Wijs, B.J. Otten, A.R. Hermus, E.A. Sistermans, CYP21 gene mutation analysis in 198 patients with 21-hydroxylase deficiency in the Netherlands: six novel mutations and a specific cluster of four mutations, J. Clin. Endocrinol. Metab. 88 (2003) 3852–3859.
[32] S. Tourmente, J.M. Deragon, J. Lafleuriel, S. Tutois, T. Pelissier, C. Cuvillier, M.C. Espagnol, G. Picard, Characterization of minisatellites in Arabidopsis thaliana with sequence similarity to the human minisatellite core sequence, Nucleic Acids Res. 22 (1994) 3317–3321.
[33] M.T. Tusie-Luna, P.C. White, Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase gene) and CYP21P involve different mechanisms, Proc. Natl. Acad. Sci. U S A 92 (1995) 10796–10800.
[34] P.C. White, D. Grossberger, B.J. Onufer, D.D. Chaplin, M.I. New, B. Dupont, J.L. Strominger, Two genes encoding steroid 21-hydroxylase are located near the genes encoding the fourth component of complement in man, Proc. Natl. Acad. Sci. U S A 82 (1985) 1089–1093.
[35] P.C. White, M.I. New, B. Dupont, HLA-linked congenital adrenal hyperplasia results from a defective gene encoding a cytochrome P-450 specific for steroid 21-hydroxylation, Proc. Natl. Acad. Sci. U S A 81 (1984) 7505–7509.
[36] P.C. White, P.W. Speiser, Congenital adrenal hyperplasia due to 21-hydroxylase deficiency, Endocr. Rev. 21 (2000) 245–291.
[37] R.T. Wyatt, R.A. Rudders, A. Zelenetz, R.A. Delellis, T.G. Krontiris, BCL2 oncogene translocation is mediated by a chi-like consensus, J. Exp. Med. 175 (1992) 1575–1588.
[38] Z. Yang, A.R. Mendoza, T.R. Welch, W.B. Zipf, C.Y. Yu, Modular variations of the human major histocompatibility complex class III genes for serine/threonine kinase RP, complement component C4, steroid 21-hydroxylase CYP21, and tenascin TNX (the RCCX module). A mechanism for gene deletions and disease associations, J. Biol. Chem. 274 (1999) 12147–12156.
[39] S. Yoshida, Y. Kaneita, Y. Aoki, M. Seto, S. Mori, M. Moriyama, Identification of heterologous translocation partner genes fused to the BCL6 gene in diffuse large B-cell lymphomas: 5′-RACE and LA – PCR analyses of biopsy samples, Oncogene 18 (1999) 7994–7999.
[40] C.Y. Yu, The complete exon–intron structure of a human complement component C4A gene. DNA sequences, polymorphism, and linkage to the 21-hydroxylase gene, J. Immunol. 146 (1991) 1057–1066.

Identification of *CYP21A2* mutant alleles in Czech patients with 21-hydroxylase deficiency

ZUZANA VRZALOVÁ[1], ZUZANA HRUBÁ[1], EVA ST'AHLOVÁ HRABINCOVÁ[1], SLAVKA POUCHLÁ[1], FELIX VOTAVA[2], STANISLAVA KOLOUSKOVÁ[3] and LENKA FAJKUSOVÁ[1,4]

[1]Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Cernopolní 9, CZ-62500 Brno; [2]Department of Pediatrics, Third Faculty of Medicine, Charles University and University Hospital Kralovské Vinohrady, Srobárova 50, CZ-10034 Prague; [3]Department of Pediatrics, Second Faculty of Medicine, Charles University and University Hospital Motol, V Úvalu 84, CZ-15006 Prague; [4]Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic

Received April 28, 2010; Accepted June 18, 2010

DOI: 10.3892/ijmm_00000504

Abstract. Congenital adrenal hyperplasia (CAH) is comprised of a group of autosomal recessive disorders caused by an enzymatic deficiency which impairs the biosynthesis of cortisol and, in most of the severe cases, also the biosynthesis of aldosterone. Approximately 90-95% of all the CAH cases are due to mutations in the steroid 21-hydroxylase gene (*CYP21A2*). In this study, the molecular genetic analysis of *CYP21A2* was performed in 267 Czech probands suspected of 21-hydroxylase deficiency (21OHD). 21OHD was confirmed in 241 probands (2 mutations were detected). In 26 probands, a mutation was found only in 1 *CYP21A2* allele. A set of 30 different mutant alleles was determined. We describe i) mutated *CYP21A2* alleles carrying novel point mutations (p.Thr168Asn, p.Ser169X and p.Pro386Arg), ii) mutated *CYP21A2* alleles carrying the novel chimeric gene designated as CH-7, which was detected in 21.4% of the mutant alleles, iii) an unusual genotype with a combination of the *CYP21A2* duplication, 2 point mutations and the *CYP21A2* large-scale gene conversion on the second allele, and (iv) a detailed analysis of the chimeric *CYP21A1P/CYP21A2* genes. In conclusion, our genotyping approach allowed for the accurate identification of the *CYP21A2* gene mutations in 21OHD patients and their families and provided some useful information on diagnosis and genetic counselling.

Correspondence to: Dr Lenka Fajkusová, Centre of Molecular Biology and Gene Therapy, University Hospital Brno, Cernopolní 9, CZ-62500 Brno, Czech Republic
E-mail: lenkafajkusova@volny.cz

Key words: congenital adrenal hyperplasia, steroid 21-hydroxylase, 21-hydroxylase deficiency, pseudogene, chimeric gene

Introduction

Congenital adrenal hyperplasia (CAH) is a group of autosomal recessive disorders caused by an enzymatic deficiency which impairs the biosynthesis of cortisol and, in most of the severe cases, also the biosynthesis of aldosterone. Approximately 90-95% of all the CAH cases are due to 21-hydroxylase deficiency (21OHD), and ~5-8% cases are due to 11-ß-hydroxylase deficiency (1). In the adrenal cortex, the steroid 21-hydroxylase (21OH) converts 17-hydroxyprogesterone into 11-deoxycortisol and progesterone into 11-deoxycorticosterone. These steroids are subsequently converted into cortisol and aldosterone, respectively (1).

The steroid 21OH gene (*CYP21A2*) and its inactive pseudogene (*CYP21A1P*) are located within the HLA class III region of the major histocompatibility complex locus on chromosome 6p21.3. The *CYP21A2* gene is 98% homologous to the *CYP21A1P* pseudogene in exons and 96% in introns (2,3). Together with the neighbouring genes (serine/threonine kinase RP, complement C4 and tenascin TNX), *CYP21* forms a genetic unit termed as the RCCX module (4,5). In the RCCX bimodular haplotype (69% of chromosome 6p21.3), the orientation of genes from telomere to centromere is *RP1-C4A-CYP21A1P-TNXA-RP2-C4B-CYP21A2-TNXB*. The 3 pseudogenes, *CYP21A1P*, *TNXA* and *RP2*, located between the 2 C4 loci, do not encode functional proteins (5).

According to the severity of the disease, 3 clinical forms of CAH have been distinguished: the classical salt-wasting (SW-CAH), the classical simple virilising (SV-CAH), and the non-classical CAH (NC-CAH) (6). If the patients are not timely diagnosed (by neonatal sreening) and treated, the severe deficiency in 21OH leads to SW-CAH. In addition to having a severe cortisol deficiency, patients with SW-CAH do not synthesise enough aldosteron and therefore, are not able to maintain sodium homeostasis. Thus, affected patients usually manifest hyponatraemia, hyperkalaemia and vomiting during the first 4 weeks of life. In the female foetus, the excess of androgen causes variable degrees of external genital

virilisation, and consequently, newborn females have genital ambiguity (1,7,8). The residual activity of 21OH (aldosterone is synthetized) results in SV-CAH with external genital virilisation in females and also results in signs of precocious pseudopuberty which develop by 8 years of age in female and male patients (1,8,9). NC-CAH is associated with a moderate deficiency in 21OH and manifests in later childhood or adolescence with hirsutism and decreased fertility or precocious pseudopuberty (8). Patients with NC-CAH secrete aldosterone normally and most of the male patients diagnosed after puberty are entirely asymptomatic (1,10). In order to reduce CAH-associated morbidity and mortality through early diagnosis and treatment, the neonatal screening of CAH based on the detection of 17-hydroxyprogesterone, has been introduced since 2006 in the Czech Republic (11).

The functional *CYP21A2* gene and the *CYP21A1P* pseudogene, each containing 10 exons, are spaced over 3.4 kb (2,3). The most common source of mutations, involving ~95% of mutant alleles, is the intergenic recombination between *CYP21A2* and *CYP21A1P*, whereas the remaining 5% of mutant alleles are represented by new mutations. The main *CYP21A2* defects are comprised of *CYP21A2* deletions, large-scale gene conversions of *CYP21A2* into a structure similar to *CYP21A1P*, and *CYP21A1P/CYP21A2* chimeric genes. These mutations are generated by the unequal crossing-over during meiosis. However, the mechanism of small-scale gene conversions, that result in short-range mutations in *CYP21A2* derived from *CYP21A1P*, is not yet well understood (12,13). Unequal crossing-over can cause a 30 kb deletion, involving the 3' end of *CYP21A1P*, all of *TNXA*, *RP2*, *C4B*, and the 5' end of *CYP21A2*, which produces a non-functional chimeric gene with 5' and 3' ends corresponding to *CYP21A1P* and *CYP21A2*, respectively (14,15).

The incidence of classical CAH in Central Europe is 1/10,000, with a gender ratio of 40% boys and 60% girls, and the carrier rate is 1/50 (7,16). The milder, non-classical form of CAH is much more common, with a prevalence of 1/100 in the general population (8). As *CYP21A2* analysis has been made available in many countries, studies on *CYP21A2* mutations in large national patient series have provided an insight into specific mutation distributions. In this study, we analyzed *CYP21A2* mutations and their frequencies in 21OHD patients in the Czech population, and compared the results with previous studies from other countries.

Patients and methods

Subjects. The patients were sent to our laboratory for the molecular analysis of the *CYP21A2* gene after endocrine and clinical evaluation. DNA samples were obtained from peripheral blood leukocytes and/or amniotic fluid by the standard salting-out method. All the studies were approved by the Ethics Board of the University Hospital Brno, and all the patients and their family members gave their written, informed consent for DNA analysis.

Detection of CYP21A2 point mutations derived from CYP21A1P and chimeric CYP21A1P/CYP21A2 genes. The direct screening of *CYP21A2* mutations required the specific amplification of the *CYP21A2* gene. This step was performed by PCR with primers specific for *CYP21A2*, which did not allow for the concomitant amplification of *CYP21A1P* (17). The product of this step was used as the DNA template for the nested amplifications of the *CYP21A2* gene fragments carrying the most frequent point mutations. The specific amplification of *CYP21A2* was performed by the Expand Long Template System kit (Roche Diagnostics, GmbH), and the nested amplifications were performed by Taq DNA Polymerase (Fermentas, GmbH). The PCR primers for the nested PCRs are shown in Table I. The detection of the most common point mutations was performed by long-template PCR and nested PCRs using restriction fragment length polymorphisms (RFLP) (p.Pro30Leu, p.Ser97fsX12, p.Gly110ValfsX21, p.Ile172Asn, p.Val237Glu, p.Val281Leu, p.Gln318X and p.Arg356Trp) and DNA sequencing (p.Leu307PhefsX6) (17). Long-template PCR was also used for the amplification of the chimeric *CYP21A1P/CYP21A2* gene. This step was performed using the 5' primer specific for *CYP21A1P*, and the 3' primer specific for *CYP21A2* (17). The nested PCRs and RFLPs were applied analogously to the detection of the most common point mutations in order to detect the type of chimeric *CYP21A1P/CYP21A2* gene (Fig. 1) (17).

Multiplex ligation-dependent probe amplification (MLPA). MLPA is a method used for the detection of large gene deletions and duplications. For the CAH diagnostics, we used the SALSA MLPA kit P050B CAH (MRS Holland, The Netherlands). This kit contains probes and primers for the analysis of 5 *CYP21A2* gene fragments located in the promoter region and exons 3, 4, 6 and 8. The established MLPA method detects deletions and duplications of the *CYP21A2* gene, as well as particular *CYP21A2* point mutations, provided that these mutations are located at probe binding sites (p.Gly110ValfsX21, p.Thr168Asn, p.Ser169X, p.Ile172Asn, p.Val237Glu, p.Arg316X and p.Gln318X). Additionally, the MLPA kit contains 28 *CYP21A2* non-specific probes that are the control standards and sets for the amplification of adjacent genes (*CREBL*, *C4B* and *TNXB*) and pseudogenes (*C4A* and *TNXA*). The assay was performed according to the manufacturer's recommendations. Genomic DNA (200 ng) was denatured at 98°C for 5 min, hybridised overnight at 60°C with the SALSA probemix, and treated with the Ligase-65 enzyme at 54°C for 15 min. The reactions were stopped by incubation at 98°C for 5 min. Subsequently, PCRs were performed with the specific SALSA PCR primers (cycling conditions were 35 cycles at 95°C for 30 sec, 60°C for 30 sec, and 72°C for 60 sec). The amplification products were run on the CEQ 8000 Genetic Analyzer (Beckman Coulter). Two healthy men were included in every analysis as the controls. The peak height of each analysed fragment was normalised to the peak height of the control sample. Relative peak heights were reduced to 40-60% with deletions and increased by 30-50% with duplications.

The results obtained by MLPA were correlated with the results obtained using long-template PCR, nested PCRs, RFLP and/or DNA sequencing. Identified mutations of the *CYP21A2* gene were also confirmed by the segregation of the parents' alleles. Detected *CYP21A2* deletions can also pose *CYP21A2* large-scale conversions into a structure similar to

Table I. Primers used for the amplification of the *CYP21A2* gene.

Primers	Localization	Sequence direction (5'- 3')
30F	Promoter	CAGTCTACACAGCAGGAGGGATGGC
30R	Exon 1	AGCAAGTGCAAGAAGCCCGGGGCAAGCTG
97/110F	Intron 2	ATCAGTTCCCACCCTCCAGCCCCGA
97/110R	Exon 3	AGGGCTGAGCGGGTGAGCTTC
172F	Exon 4	GAGGAATTCTCTCTCCTCACCTGCAGCATTA
172R	Intron 4	AGTTGTCGTCCTGCCAGAAAAGGA
236/237/239F	Exon 6	AGCAGGCCATAGAGAAGAGGGATCACATCG
236/237/239R	Intron 6	ATGCAAAAGAACCCGCCTCATAGC
281F	Exon 7	TGCAGGAGAGCCTCGTGGCAGG
281R	Exon 7	GACGCACCTCAGGGTGGTGAAG
318/356F	Intron 7	GCTGGGGCAGGACTCCACCCGA
318/356R	Exon 8	GTGGGGCAAGGCTAAGGGCACAACTGGC
1R*	Intron 1	AAGCAGCGTCAGCGGAGAGGG
2F*	Intron 1	TTGAGGCTGAGGTGGGAGGA
2R*	Intron 2	GCGGAGGTGACGGAGAGGGT
3F*	Intron 2	AAGCTCTTGGGGGGCATATC
3R*	Intron 3	GGCTACTGTGAGAGGCGAGG
4F*	Intron 3	GTCAGCCTCGCCTCTCACAG
4R*	Intron 4	CAGTTCAGGACAAGGAGAGGCT
5F*	Intron 4	AGCCCCTCCCTGAGCCTCTC
5R*	Intron 5	AGCCTCTCCCTCCACCCCAG
6F*	Intron 5	TGGGTTGTAGGGGAGAGGCT
6R*	Intron 6	TAGCAATGCTGAGGCCGGTA
7F*	Intron 6	TGCCACTCTGTACTCCTCTC
7R*	Intron 7	ACAGTGCTCAGAGCTGAGTG
8F*	Intron 7	CTCACCGGCACTCAGGCTCA
8R*	Intron 8	AAGGGGGCTGGAGTTAGAGGCT
9F*	Intron 8	AGTGAGGAAAGCCCGAGCCC
9R*	Intron 9	GTGGGTGGGGAGGCGTTCAG
10F*	Intron 9	AAAATGTGGTGGAGGCTGGT
10R*	3'UTR	ACGGGAGCAATAAAGGAGAAAC
RM1F	Promoter	TTCAGGCGATTCAGGAAGGC
RM1R	Exon 3	CTTTCCAGAGCAGGGAGT
RM2F	Exon 3	CGGACCTGTCCTTGGGAGACTAC
RM2R	3'UTR	TTTCAGCCCCACAGTGTAACAGG

F, forward primer; R, reverse primer. *Primers used for PCR-sequencing. No (*) primers used for PCR-RFLP (the names of the primers are consistent with the no. of codons carrying an analysed mutation, and primer 30F was also used as the forward primer for the sequencing of exon 1). The RM primers were used for the specific amplification of *CYP21A2* in order to perform the sequencing analysis of individual exons. For the cDNA reference sequence go to, http://www.ncbi.nlm.nih.gov/nuccore/187895.

CYP21A1P. This type of *CYP21A2* deletion was verified by PCR with primers specific to the *TNXB/TNXA* hybrid gene. (see below) (16).

Detection of CYP21A2 large-scale conversions into CYP21A1P associated with the TNXB/TNXA hybrid gene. The *TNXB/TNXA* hybrid gene was detected using PCR with the 5' primer specific to *TNXB* and the 3' primer specific to the both genes, *TNXB* and *TNXA* (16). In the presence of the *TNXB/TNXA* hybrid gene, the PCR product sized 2688 bp was generated in contrast to the PCR product sized 2808 pb resulting from the *TNXB* gene (for the schematic presentation see ref.16).

Detection of rare mutations in the CYP21A2 gene. DNA sequencing was applied for the detection of rare *CYP21A2* mutations in patients with a mutation identified only in 1 *CYP21A2* allele after basic DNA diagnosis, which included the detection of i) 9 point mutations derived from *CYP21A1P*, ii) chimeric genes, and iii) *CYP21A2* deletions and duplications. As 90-95% of the mutant alleles carried ≥1 discrete mutation, the samples carrying none of these mutations were presumed to be unaffected with >99% confidence (1). As a primary template for this analysis, we did not use long-template PCR product (see the detection of *CYP21A2* point mutations derived from *CYP21A1P* and chimeric *CYP21A1P/CYP21A2* genes) as various polymorphisms were identified

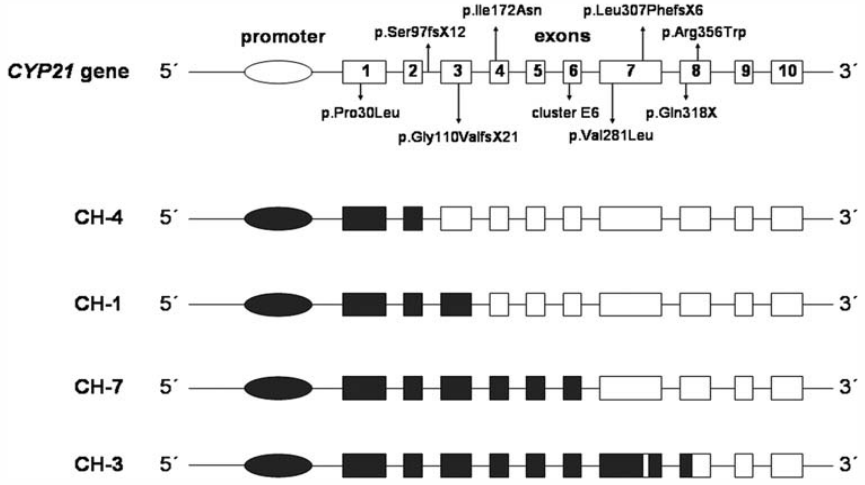


Figure 1. Types of chimeric *CYP21A1P/CYP21A2* genes identified in the Czech 21OHD patients. White boxes, structure of the functional *CYP21A2* gene; black boxes, non-functional *CYP21A1P* pseudogene; arrows, mutations existing in *CYP21A1P*. From *CYP21A1P*, the chimeric gene, CH-4, harbours the promoter, exon 1 (p.Pro30Leu) and exon 2. The chimeric gene, CH-1, bears exon 1, exon 2 (p.Ser97fsX12) and exon 3 (p.Gly110ValfsX21). The chimeric gene, CH-7, carries the promoter, exons 1-3, exon 4 (p.Ile172Asn), exon 5 and exon 6 (p.Ile236Asn, p.Val237Glu, Met239Lys). The chimeric gene, CH-3, harbours the promoter, exons 1-6, exon 7 (p.Leu307PhefsX6, but not Val281Leu) and exon 8 (p.Gln318X).

in the annealing sequences of the *CYP21A2* specific primers in some patients. For this reason, the detection of rare mutations was performed with 2 primary PCRs (primers are marked RM1 and RM2, Table I). The first amplified fragment was comprised of a region from the promoter to exon 3, and the second one of a region from exon 3 to 3'UTR. Both PCRs used *CYP21A2* specific primers which were complementary to an 8-bp segment in exon 3 deleted in *CYP21A1P* (18). PCR products were used as DNA templates for nested amplifications using sequencing primers (Table I). For the localization of primers for primary PCRs inside exon 3, DNA sequencing of this exon was performed using PCR product from long-template PCR (17). All the amplifications were performed by Taq DNA Polymerase (Fermentas). The purified PCR products were subsequently sequenced by the BigDye Terminator kit (Applied Biosystems) and analysed on an ABI PRISM 310 sequencer (Applied Biosystems).

Results

The molecular genetic analysis of the *CYP21A2* gene was performed in 267 Czech probands suspected of 21OHD. The diagnosis was confirmed in 241 of them (2 *CYP21A2* mutations were found). In 26 probands, a mutation was determined in only 1 *CYP21A2* allele (7 patients had SW-CAH, 8 had SV-CAH and 11 had NC-CAH). In the set of 21OHD probands, we determined 30 different mutant alleles (Table II).

CYP21A1P/CYP21A2 chimeric genes. The most frequent mutation, the chimeric *CYP21A1P/CYP21A2* gene, was found in 33.7% of mutant alleles. Four types of chimeric *CYP21A1P/CYP21A2* genes were detected in the Czech patients: CH-1, CH-3, CH-4 and CH-7 (Fig. 1). All the hybrid genes had a different extent of the *CYP21A1P* sequence attached to the 3' part of *CYP21A2*. The types CH-1, CH-3 and CH-4, have been previously described (12,19). The CH-4 type was found in 2.6% of the mutant alleles. This hybrid molecule differs from the functional *CYP21A2* gene by the *CYP21A1P* promoter sequence and the p.Pro30Leu mutation in exon 1. The CH-1 type is identical to the *CYP21A1P* gene in the promoter and exons 1-3, and the mutation p.Ile172Asn is not present in exon 4. This type of chimeric gene was detected in 9.3% of the mutant alleles. In addition, 1% of the patients had an atypical mutant allele with CH-1 and the p.Arg356Trp mutation in exon 8. Mutant alleles carrying CH-1 and p.Arg356Trp were confirmed by the segregation of the parents' alleles. The most common type of the chimeric gene was the newly characterized chimeric gene, CH-7 (21.4% of the mutant alleles). This chimeric gene involved the *CYP21A1P* sequence from the promoter up to exon 6, and the p.Val281Leu mutation was not present in exon 7. The least frequent chimeric gene, CH-3 (0.4% of mutant alleles) was identical to the *CYP21A1P* pseudogene from the promoter up to exon 8, and the p.Gln318X mutation was present in the chimeric gene in contrast to the p.Arg356Gln mutation. In all the cases, the mutant CH-3 allele did not contain the p.Val281Leu mutation.

CYP21A2 deletions and duplications. Total deletions of the *CYP21A2* gene were detected in 4.9% of the *CYP21A2* mutant alleles. The *CYP21A2* large-scale conversions into *CYP21A1P* generating the *TNXB/TNXA* hybrid gene represent the majority part of these (3.1% of the mutant alleles). Koppens *et al* described that *CYP21A2* large-scale conversions generating the *TNXA/TNXB* hybrid gene are associated with a presence of 2 *CYP21A1P* pseudogenes on the involved chromosome (13). The presence of 2 copies of the *CYP21A1P* pseudogenes on 1 chromosome was identified in all the

Table II. Types of *CYP21A2* mutant alleles and their frequency in Czech 21OHD probands.

Mutant allele	Mutation at cDNA level	Mutation at protein level	Localization of mutation	Phenotype	No. of alleles	Frequency (%)
1	Chimeric *CYP21A1P*/ *CYP21A2* gene (CH-4)		Promoter - exon 2	NC, SV	13	2,6
2	Chimeric *CYP21A1P*/ *CYP21A2* gene (CH-1)		Promoter - exon 3	SW	42	8,3
3	Chimeric *CYP21A1P*/ *CYP21A2* gene (CH-1) + p.Arg356Trp		Promoter - exon 3, exon 8	SW	5	1,0
4	Chimeric *CYP21A1P*/ *CYP21A2* gene (CH-7)		Promoter - exon 6	SW	109	21,4
5	Chimeric *CYP21A1P*/ *CYP21A2* gene (CH-3)		Promoter - exon 8	SW	2	0,4
6	*CYP21A2* deletion		*CYP21A2* gene	SW	9	1,8
7	*CYP21A2* large-scale gene conversion		*CYP21A2* gene	SW	16	3,1
8	*CYP21A2* duplication c.290-13A/C>G, c.952C>T	*CYP21A2* duplication p.Ser97fsX12, p.Gln318X	*CYP21A2* gene Intron 2, exon 8	SW	2	0,4
9	c.89C>T	p.Pro30Leu	Exon 1	NC	19	3,5
10	c.290-13A/C>G	p.Ser97fsX12	Intron 2	SW	122	24,00
11	c.329_336del8	p.Gly110ValfsX21	Exon 3	SW	6	1,2
12	**c.503C>A***	**p.Thr168Asn***	**Exon 4**	**NC**	**1**	**0,2**
13	**c.507C>A***	**p.Ser169X***	**Exon 4**	**SW**	**1**	**0,2**
14	c.515T>A	p.Ile172Asn	Exon 4	SV	57	11,2
15	c. 841G>T	p.Val281Leu	Exon 7	NC	47	9,3
16	c.918_919insT	p.Leu307PhefsX6	Exon 7	SW	1	0,2
17	**c.946C>T**	**p.Arg316X**	**Exon 8**	**SW**	**1**	**0,2**
18	c.952C>T	p.Gln318X	Exon 8	SW	18	3,5
19	c.1066C>T	p.Arg356Trp	Exon 8	SW	20	3,9
20	**c.1067G>A**	**p.Arg356Gln**	**Exon 8**	**SV**	**2**	**0,4**
21	**c.1157C>G***	**p.Pro386Arg***	**Exon 9**	**SW**	**1**	**0,2**
22	**c.1357C>T**	**p.Pro453Ser**	**Exon 10**	**NC**	**3**	**0,6**
23	**c.1375C>T**	**p.Pro459Ser**	**Exon 10**	**SV**	**1**	**0,2**
24	**c.1448G>C**	**p.Arg483Pro**	**Exon10**	**SV**	**2**	**0,4**
25	c.290_13A/C>G, c.329_336del8	p. Ser97fsX12, p.Gly110ValfsX21	Intron 2, exon 3	SW	1	0,2
26	c.290_13A/C>G, c.329_336del8, c.515T>A	p.Ser97fsX12, p.Gly110ValfsX21, p.Ile172Asn	Intron 2, exons 3 and 4	SW	1	0,2
27	c.707T>A, c.710 T>A, c.716T>A	p.Ile236Asn, p.Val237Glu, p.Met239Lys	Exon 6	SW	3	0,6
28	c.841G>T, c.952C>T	p.Val281Leu, p.Gln318X,	Exons 7 and 8	SW	1	0,2
29	c.841G>T, c.920_921insT, c.952C>T, c.1066C>T	p.Val281Leu, p.Leu307PhefsX6, p.Gln318X, p.Arg356Trp	Exons 7 and 8	SW	1	0,2
30	c.952C>T, c.1066C>T	p.Gln318X, p.Arg356Trp	Exon 8	SW	1	0,2

Rare *CYP21A2* mutations (not derived from *CYP21A1P*) are shown in bold letters. *These are novel mutations, and have not been described previously.

Table III. Correlation between genotype and phenotype in 142 Czech 21OHD patients.

Group	Mutant allele 1	Mutant allele 2	Phenotype	No. of probands	Frequency (%)
A	Chimeric gene	Chimeric gene	SW	18	12.7
	CH-4	Chimeric gene	SW	1	0.7
	Chimeric gene	*CYP21A2* deletion	SW	2	1.4
	CH-4	*CYP21A2* deletion	SW	3	2.1
	Chimeric gene	p.Ser97fsX12	SW	13	9.2
	Chimeric gene	p.Gly110ValfsX21	SW	2	1.4
	Chimeric gene	p.Val237Glu	SW	2	1.4
	Chimeric gene	p.Val281Leu*	SW	1	0.7
	Chimeric gene	p.Gln318X	SW	4	2.8
	CYP21A2 deletion	*CYP21A2* deletion	SW	3	2.1
	CYP21A2 deletion	dup[p.Ser97fsX12; p.Gln318X]	SW	1	0.7
	CYP21A2 deletion	p.Ser97fsX12	SW	6	4.2
	p.Pro30Leu*	p.Arg356Trp	SW	1	0.7
	p.Ser97fsX12	dup[p.Ser97fsX12; p.Gln318X]	SW	1	0.7
	p.Ser97fsX12	p.Ser97fsX12	SW	4	2.8
	p.Ser97fsX12	p.Ser169X	SW	1	0.7
	p.Ser97fsX12	p.Gln318X	SW	3	2.1
	p.Ser97fsX12	p.Arg356Trp	SW	1	0.7
	p.Ser97fsX12	p.Pro386Arg	SW	1	0.7
	p.Ser97fsX12, p.Gly110ValfsX21	p.Gly110ValfsX21	SW	1	0.7
	p.Val281Leu, p.Leu307PhefsX6, p.Gln318X, p.Arg356Trp	p.Gln318X	SW	1	0.7
	p.Arg356Trp	p.Arg356Trp	SW	2	1.4
B	Chimeric gene	p.Ser97fsX12	SW/SV	4	2.8
	CYP21A2 deletion	p.Ser97fsX12	SW/SV	1	0.7
	p.Ser97fsX12	p.Ser97fsX12	SW/SV	5	3.5
	p.Ser97fsX12	p.Arg356Trp	SW/SV	3	2.1
C	CH-4	Chimeric gene	SV	6	4.2
	CH-4	p.Ser97fsX12	SV	1	0.7
	Chimeric gene	p.Ile172Asn	SV	11	7.7
	Chimeric gene	p.Pro459Ser	SV	1	0.7
	CYP21A2 deletion	p.Ile172Asn	SV	2	1.4
	p.Ser97fsX12	p.Ile172Asn	SV	6	4.2
	p.Ile172Asn	p.Ile172Asn	SV	4	2.8
	p.Ile172Asn	p.Val237Glu	SV	1	0.7
	p.Ile172Asn	p.Val281Leu	SV	1	0.7
	p.Ile172Asn	p.Arg316X	SV	1	0.7
	p.Ile172Asn	p.Gln318X	SV	1	0.7
	p.Ile172Asn	p.Arg356Trp	SV	2	1.4
D	CH-4	CH-4	NC	1	0.7
	Chimeric gene	p.Pro30Leu	NC	2	1.4
	Chimeric gene	p.Thr168Asn	NC	1	0.7
	CH-4	p.Val281Leu	NC	1	0.7
	p.Pro30Leu	p.Pro30Leu	NC	1	0.7
	p.Pro30Leu	p.Ser97fsX12	NC	1	0.7
	p.Pro30Leu	p.Val281Leu	NC	1	0.7
	p.Ser97fsX12	p.Val281Leu	NC	3	2.1
	p.Ile172Asn	p.Val281Leu	NC	2	1.4
	p.Val281Leu	p.Val281Leu	NC	4	2.8
	p.Val281Leu	p.Gln318X	NC	1	0.7
	p.Val281Leu	p.Arg356Trp	NC	2	1.4

Group A, patients with SW-CAH; group B, patients with indefinite diagnosis, SV or SW-CAH; group C, patients with SV-CAH; group D, patients with NC-CAH. Chimeric gene, the *CYP21A1P/CYP21A2* chimeric gene (CH-1, CH-3 and CH-7 forms are not distinguished here). *Mutation which does not correspond to the patient's phenotype.

Table IV. *CYP21A2* allele frequencies in different European populations.

P	TA	Del/ con	p.Pro30 Leu	p.Ser97 fsX12	p.Gly 110Valf sX21	p.Ile 172 Asn	p.Val 281Leu	p.Gln 318X	p.Arg 356Trp	T (%)
Czech Republic	508	38.6	3.5	24.0	1.2	11.2	9.3	3.5	3.9	95.2
Central Europe (20)	696	34.1	3.7	31.2	1.0	14.5	3.4	2.6	2.4	92.9
Austria (21)	158	35.4	3.2	22.8	0.0	15.8	12.0	2.5	3.2	94.9
Hungary (20)	270	27.1	3.0	35.9	0.0	14.1	5.6	1.9	3.0	90.6
Southern Germany (22)	310	29.0	2.6	30.3	1.6	19.7	2.9	4.8	4.5	95.4
The Netherlands (23)	370	31.9	0.3	28.1	4.3	12.4	2.2	3.5	8.4	91.1
Spain (24)	266	5.6	1.5	6.0	1.1	2.3	63.2	2.3	0.8	82.7

P, population; TA, total no. of alleles; del, deletion; con, conversion; T, total frequency of given mutations.

Czech patients carrying a *CYP21A2* large-scale conversion into *CYP21A1P*.

Duplications of *CYP21A2* associated with a mutation on both gene copies, were found in 2 probands (0.4% of the mutated alleles). In both cases, the *CYP21A2* genes localized on 1 chromosome carried the p.Ser97fsX12 mutation, as well as the p.Gln318X mutation, respectively. The association between the *CYP21A2* duplication and the mentioned point mutations was confirmed by the segregation of the parents' alleles.

CYP21A2 point mutations. Small DNA rearrangements of the *CYP21A2* gene, derived and non-derived from *CYP21A1P*, were present in 56.8 and 2.4% of the mutant alleles, respectively. The most frequent point mutations were p.Ser97fsX12, p.Ile172Asn and p.Val281Leu. We also detected 7 types of alleles with ≥2 mutations in 1 *CYP21A2* gene.

DNA sequencing of the *CYP21A2* exons and adjacent intron regions was performed in 38 probands with 1 identified *CYP21A2* mutant allele. Point mutations non-derived from *CYP21A1P* were found in 12 patients (Table II). Three of the identified point mutations have not been described previously (p.Thr168Asn, p.Ser169X and p.Pro386Arg). Female patient 1 (NC-CAH) is a compound heterozygote for the p.Thr168Asn mutation (the paternal allele), and the chimeric gene CH-7 (the maternal allele). Female patients 2 and 3 carried a phenotype associated with SW-CAH and both sufferred from genital virilisation and a life-threatening salt-wasting crisis in the neonatal phase. Patient 2 carried the p.Ser169X mutation, and the p.Ser97fsX12 mutation (DNA of parents was unavailable for analysis). Patient 3 carried the p.Pro386Arg mutation (the paternal allele) and the p.Ser97fsX12 mutation (the maternal allele).

Phenotype-genotype correlations. The correlation between genotype and phenotype was presented for 142 patients (Table III) in which sufficient clinical and biochemical data were available. Seventy-two patients suffered from the classic SW-CAH form (Group A), 13 had an indefinite phenotype, in that it was impossible to distinguish the SW and SV-CAH form (Group B), 37 suffered from the classic SV-CAH form (Group C), and 20 had the NC-CAH form (Group D). We observed a good correlation between genotype and phenotype. Discrepancies between genotype and phenotype were found in 2 patients with the SW-CAH phenotype (both patients had a mutation associated with the more moderate phenotype). The first patient is a male compound heterozygote for the p.Val281Leu mutation and the chimeric gene, CH-7. The second one is a female compound heterozygote carrying the p.Pro30Leu and the p.Arg356Trp mutations.

Discussion

The gene encoding the steroid 21OH enzyme, *CYP21A2*, is considered to be one of the most polymorphic human genes. Point mutations and copy number variations, such as deletions and duplications, have been described in many populations (20-24). Using the analysis of the *CYP21A2* gene, we confirmed the referral diagnosis in 241 Czech unrelated patients suspected of 21OHD. *CYP21A2* mutations can be predicted to cause a certain phenotype (SW, SV and NC), on the basis of the reduction of 21OH enzymatic activity. As 21OHD is an autosomal recessive disease and most 21OHD patients are compound heterozygotes, the phenotype of a patient should reflect a mutation that is predicted to cause the least severe impairment of enzymatic activity. This approach to phenotype prediction has been shown to be correct, although slight deviations to this correlation exist (1,23,25). We checked for correlations between the genotypes and phenotypes in 142 probands with 21OHD. A discrepancy was observed in 2 patients with SW-CAH. The first one was a male proband carrying the p.Val281Leu mutation (associated with NC-

CAH) (1) and the CH-7 chimeric gene. The second one was a female proband carrying the p.Pro30Leu mutation (associated with NC-CAH) and the p.Arg356Trp mutation (associated with SV-CAH) (1). Both patients suffered from a salt wasting crisis in the neonatal period and were treated with fludrocortisone and hydrocortisone.

Duplications of *CYP21A2* associated with a mutation on both gene copies were found in 2 probands. In both cases, the *CYP21A2* duplication was associated with the mutations, p.Ser97fsX12 and p.Gln318X (the same mutant allele has been previously described) (24). Both probands with the *CYP21A2* duplication, p.Ser97fsX12 and p.Gln318X on 1 allele have SW-CAH phenotypes. The complete genotypes of these patients were dup[p.Ser97fsX12; p.Gln318X]/p.Ser97fsX12 and dup[p.Ser97fsX12; p.Gln318X]/*CYP21A2* deletion. Determination of the genotype, dup[p.Ser97fsX12; p.Gln318X]/*CYP21A2* deletion, was quite difficult. Using long-template PCR, nested PCRs, and PCR-RFLP, the mutations, p.Ser97fsX12 and p.Gln318X, were detected and this result was also confirmed by MLPA. MLPA showed 1 copy of exon 8 (the p.Gln318X mutation was located in the probe binding site), and other *CYP21A2* fragments detected by MLPA had normal peak sizes corresponding to 2 *CYP21A2* copies. Thus, we presumed that this patient's genotype was p.Ser97fsX12/p.Gln318X. Using long-template amplification with primers specific to the *TNX* genes, we discovered the presence of the *TNXB/TNXA* hybrid gene and thus the presence of the *CYP21A2* large-scale conversion into structure similar to *CYP21A1P*. In this patient, a rare genotype combination was characterized, which has not been described so far. For the verification and diagnostic accuracy of our DNA, we also performed DNA analyses on the patients' family members. The mutant allele carrying the *CYP21A2* duplication with p.Ser97fsX12 and p.Gln318X was inherited from the mother and the mutant allele with the *CYP21A2* deletion was inherited from the father. We then offered the prenatal diagnosis to this family. After DNA analysis of the amniotic fluid, we detected the *CYP21A2* duplication and the point mutations, p.Ser97fsX12 and p.Gln318X, (using PCR-RFLP and MLPA), and the *CYP21A2* large-scale gene conversion.

The most frequent mutations detected in the Czech 21OHD patients were the chimeric *CYP21A1P/CYP21A2* genes. To date, 6 different chimeric *CYP21A1P/CYP21A2* genes have been characterized (12,14,19,26) The CH-1, CH-3 and CH-4 chimeric genes were detected in the Czech population and in addition, 1 novel type (CH-7) was determined (Fig. 1). The 5' untranslated region of the *CYP21A2* gene responsible for the transcriptional activity, is located in the first 167 nucleotides upstream of the ATG codon and contains binding sites for the specificity protein, Sp-1, and adrenal-specific protein transcription factors (27,28). In this fragment, the pseudogene promoter differs from the *CYP21A2* promoter in 4 nucleotides, located at the -126, -113, -110 and -103 positions. These differences cause a lower affinity of the pseudogene promoter to the transcription factors and, consequently, reduce its transcriptional activity to 20% when compared to the *CYP21A2* gene (15,29). The mutation p.Pro30Leu is associated with NC-CAH, and reduces the 21OH activity to 30-40% of the normal enzyme. In the chimeric gene CH-4, the mutation p.Pro30Leu, in synergism with the promoter sequence substitutions, can decrease enzyme activity to 4-10% (30). In the set of our 21OHD probands, we detected 13 patients with at least 1 CH-4 chimeric gene. Three probands with the genotype CH-4/*CYP21A2* deletion, and 1 proband with the genotype CH-4/CH-7 suffered from SW-CAH, 7 probands suffered from SV-CAH and their genotypes were CH-4/CH-1, CH-4/CH-7 and CH-4/p.Ser97fsX12, and 2 probands were clinically diagnosed as NC-CAH and had the genotypes CH-4/p.Val281Leu and CH-4 in the homozygous state.

Furthermore, we detected 6 types of alleles with ≥2 mutations in 1 *CYP21A2* gene. In the case of mutant alleles 25, 26, 27, 29, and 30 (Table II), mutations present in *CYP21A1P* were probably transferred to *CYP21A2* as 1 recombination event. In the case of alleles 3 and 28, we presume that 2 recombination events took place (*CYP21A1P* mutations lying between transferred mutations are missing in these alleles).

In the set of our 21OHD patients, we detected 1 novel non-sense mutation (p.Ser169X) and 2 novel missense (p.Thr168Asn and p.Pro386Arg) mutations. The p.Thr168Asn mutation is associated with NC-CAH, and the p.Pro386Arg mutation with SW-CAH. A conservation of amino acid residues was determined by Robins *et al*, based on the comparison of multiple species-specific sequence variants (pig, dog, cow, sheep, mouse, rat, eel, and puffer fish) homologous to the human *CYP21A2* (31). Based on this result, both missense mutations (p.Thr168Asn and p.Pro386Arg) are located in a highly conserved region. The p.Thr168Asn mutation is situated in the central part of helix E and p.Pro386Arg in a loop between the ß-sheet 1-3 and helix K' (31). In addition, the p.Pro386Arg is located in the heme-binding site (31).

CYP21A2 mutations were detected in 508 alleles of unrelated patients suspected of 21OHD. A comparison with mutation frequencies in other European countries showed similar results in general (Table IV), especially with central European countries. In addition, we described i) mutated *CYP21A2* alleles carrying novel point mutations (p.Thr168Asn, p.Ser169X and p.Pro386Arg), ii) mutated *CYP21A2* alleles carrying the novel chimeric gene designated as CH-7, iii) an unusual genotype with a combination of the *CYP21A2* duplication and the *CYP21A2* large-scale gene conversion on the second allele, and iv) a detailed analysis of the chimeric *CYP21A1P/CYP21A2* genes. In conclusion, our genotyping approach allowed for the accurate identification of *CYP21A2* gene mutations in 21OHD patients and their families and provided some useful information on diagnosis and genetic counselling.

Acknowledgements

This study was supported by the Internal Grant Agency of the Czech Ministry of Health (grant nos. NR9308-3 and NS9981-3) and by the Czech Ministry of Education (project nos. LC06023, MSM0021622415 and MSM0021620814).

References

1. White PC and Speiser PW: Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Endocr Rev 21: 245-291, 2000.

2. White PC, Grossberger D, Onufer BJ, *et al*: Two genes encoding steroid 21-hydroxylase are located near the genes encoding the fourth component of complement in man. Proc Natl Acad Sci USA 82: 1089-1093, 1985.
3. Higashi Y, Yoshioka H, Yamane M, Gotoh O and Fujii-Kuriyama Y: Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: A pseudogene and a genuine gene. Proc Natl Acad Sci USA 83: 2841-2845, 1986.
4. Yang Z, Mendoza AR, Welch TR, Zipf WB and Yu CY: Modular variations of the human major histocompatibility complex class III genes for serine/threonine kinase RP, complement component C4, steroid 21-hydroxylase CYP21, and tenascin TNX (the RCCX module). A mechanism for gene deletions and disease associations. J Biol Chem 274: 12147-12156, 1999.
5. Blanchong CA, Zhou B, Rupert KL, *et al*: Deficiencies of human complement component C4B and C4B and heterozygosity in length variants of RP-C4-CYP21-TNX (RCCX) modules in caucasians. The load of RCCX genetic diversity on major histocompatibility complex-associated disease. J Exp Med 191: 2183-2196, 2000.
6. New MI and Wilson RC: Steroid disorders in children: Congenital adrenal hyperplasia and apparent mineralocorticoid excess. Proc Natl Acad Sci USA 96: 12790-12797, 1999.
7. Kovacs J, Votava F, Heinze G, *et al*: Lessons from 30 years of clinical diagnosis and treatment of congenital adrenal hyperplasia in five middle European countries. J Clin Endocrinol Metab 86: 2958-2964, 2001.
8. Goncalves J, Friaes A and Moura L: Congenital adrenal hyperplasia: Focus on the molecular basis of 21-hydroxylase deficiency. Expert Rev Mol Med 9: 1-23, 2007.
9. Brinkmann AO: Molecular basis of androgen insensitivity. Mol Cell Endocrinol 179: 105-109, 2001.
10. New MI: Extensive clinical experience: Nonclassical 21-hydroxylase deficiency. J Clin Endocrinol Metab 91: 4205-4214, 2006.
11. Strnadova KA, Votava F, Lebl J, *et al*: Prevalence of congenital adrenal hyperplasia among sudden infant death in the Czech Republic and Austria. Eur J Pediatr 166: 1-4, 2007.
12. Lee HH, Lee YJ and Lin CY: Pcr-based detection of the CYP21 deletion and TNXA/TNXB hybrid in the RCCX module. Genomics 83: 944-950, 2004.
13. Koppens PF, Hoogenboezem T and Degenhart HJ: Duplication of the CYP21A2 gene complicates mutation analysis of steroid 21-hydroxylase deficiency: Characteristics of three unusual haplotypes. Hum Genet 111: 405-410, 2002.
14. Lee HH: Chimeric CYP21P/CYP21 and TNXA/TNXB genes in the RCCX module. Mol Genet Metab 84: 4-8, 2005.
15. Chang SF and Chung BC: Difference in transcriptional activity of two homologous CYP21A genes. Mol Endocrinol 9: 1330-1336, 1995.
16. Koppens PF, Hoogenboezem T and Degenhart HJ: Carriership of a defective tenascin-X gene in steroid 21-hydroxylase deficiency patients: TNXB-TNXA hybrids in apparent large-scale gene conversions. Hum Mol Genet 11: 2581-2590, 2002.
17. Lee HH, Chang JG, Tsai CH, Tsai FJ, Chao HT and Chung B: Analysis of the chimeric CYP21P/CYP21 gene in steroid 21-hydroxylase deficiency. Clin Chem 46: 606-611, 2000.
18. Wedell A and Luthman H: Steroid 21-hydroxylase deficiency: Two additional mutations in salt-wasting disease and rapid screening of disease-causing mutations. Hum Mol Genet 2: 499-504, 1993.
19. L'Allemand D, Tardy V, Gruters A, Schnabel D, Krude H and Morel Y: How a patient homozygous for a 30-kb deletion of the C4-CYP 21 genomic region can have a nonclassic form of 21-hydroxylase deficiency. J Clin Endocrinol Metab 85: 4562-4567, 2000.
20. Dolzan V, Solyom J, Fekete G, *et al*: Mutational spectrum of steroid 21-hydroxylase and the genotype-phenotype association in Middle European patients with congenital adrenal hyperplasia. Eur J Endocrinol 153: 99-106, 2005.
21. Baumgartner-Parzer SM, Schulze E, Waldhausl W, *et al*: Mutational spectrum of the steroid 21-hydroxylase gene in Austria: Identification of a novel missense mutation. J Clin Endocrinol Metab 86: 4771-4775, 2001.
22. Krone N, Braun A, Roscher AA, Knorr D and Schwarz HP: Predicting phenotype in steroid 21-hydroxylase deficiency? Comprehensive genotyping in 155 unrelated, well defined patients from southern Germany. J Clin Endocrinol Metab 85: 1059-1065, 2000.
23. Stikkelbroeck NM, Hoefsloot LH, de Wijs IJ, Otten BJ, Hermus AR and Sistermans EA: CYP21 gene mutation analysis in 198 patients with 21-hydroxylase deficiency in The Netherlands: six novel mutations and a specific cluster of four mutations. J Clin Endocrinol Metab 88: 3852-3859, 2003.
24. Loidi L, Quinteiro C, Parajes S, *et al*: High variability in cyp21a2 mutated alleles in Spanish 21-hydroxylase deficiency patients, six novel mutations and a founder effect. Clin Endocrinol (Oxf) 64: 330-336, 2006.
25. Friaes A, Rego AT, Aragues JM, *et al*: CYP21A2 mutations in Portuguese patients with congenital adrenal hyperplasia: Identification of two novel mutations and characterization of four different partial gene conversions. Mol Genet Metab 88: 58-65, 2006.
26. Concolino P, Mello E, Minucci A, *et al*: A new CYP21A1P/CYP21A2 chimeric gene identified in an Italian woman suffering from classical congenital adrenal hyperplasia form. BMC Med Genet 10: 72, 2009.
27. Kagawa N and Waterman MR: cAMP-dependent transcription of the human CYP21B (P-450C21) gene requires a cis-regulatory element distinct from the consensus cAMP-regulatory element. J Biol Chem 265: 11299-11305, 1990.
28. Kagawa N and Waterman MR: Evidence that an adrenal-specific nuclear protein regulates the cAMP responsiveness of the human CYP21B (P450C21) gene. J Biol Chem 266: 11199-11204, 1991.
29. Bristow J, Gitelman SE, Tee MK, Staels B and Miller WL: Abundant adrenal-specific transcription of the human P450c21A 'pseudogene'. J Biol Chem 268: 12919-12924, 1993.
30. Araujo RS, Billerbeck AE, Madureira G, Mendonca BB and Bachega TA: Substitutions in the CYP21A2 promoter explain the simple-virilizing form of 21-hydroxylase deficiency in patients harbouring a P30L mutation. Clin Endocrinol (Oxf) 62: 132-136, 2005.
31. Robins T, Carlsson J, Sunnerhagen M, Wedell A and Persson B: Molecular model of human CYP21 based on mammalian CYP2C5: Structural features correlate with clinical severity of mutations causing congenital adrenal hyperplasia. Mol Endocrinol 20: 2946-2964, 2006.

5 Závěr

Habilitační práce předložená formou komentovaného souboru publikací se věnuje molekulárně genetické diagnostice dědičných nemocí. Uvedené výsledky byly získány v Centru molekulární biologie a genové terapie IHOK LF MU a Fakultní nemocnice Brno, výsledky před rokem 2002 ve Výzkumném ústavu zdraví dítěte.

V současné době se v Sekci vrozených genetických chorob CMBGT provádí standardně 44 genetických vyšetření, z toho 30 unikátně v rámci ČR. Kromě standardních metod DNA diagnostiky (PCR, klasická sekvenační analýza, fragmentační analýza, kvantitativní PCR) jsou do diagnostické praxe zavedeny i techniky jako pulzní gelová elektroforéza spojená se Southern blotem a hybridizací, amplikonové sekvenování nové generace nebo *sequence capture* a cílené re-sekvenování.

S rozvojem technik sekvenování nové generace stale aktuálněji vyvstává potřeba ověření kauzality identifikovaných mutací. Jednou z možností jak zjistit, jestli nalezená mutace souvisí s nemocí, je její analýza *in silico* přístupy jako jsou predikční programy (Polyphen-2, SIFT,SNPs3D and FOLDX) a molekulární modelování na základě známých krystalových struktur. Další možností je pak exprese mutantních proteinů a následná analýza jejich funkce. I tyto metodiky jsou na pracovišti rozvíjeny, protože spolu s výsledky klinickými, biochemickými, patologickými vytvářejí komplexní obraz o dané nemoci a její molekulární podstatě. Poznatky a zkušenosti z oblasti výzkumu a diagnostiky vrozených poruch jsou využívány rovněž v teoretické i praktické výuce na Masarykově universitě, především v rámci semestrálního kurzu Molekulární diagnostika dědičných nemocí, a jsou předmětem závěrečných prací studentů magisterských a doktorských studijních programů.

Publikace uvedené v habilitační práci prezentují výsledky získané analýzou DNA/mRNA genů asociovaných s daným typem nemoci a korelace těchto výsledků s klinickými, biochemickými a patologickými nálezy. Vzhledem k tomu, že molekulárně genetická diagnostika těchto nemocí se provádí v CMBGT jako v jediné laboratoři v rámci ČR, výsledky mapují mutační spektrum a frekvenci výskytu jednotlivých typů mutací v České republice. Zjištěné genotypy a asociované fenotypy jsou prezentovány i v mezinárodních databázích mutací a přispívají tak k utvoření celkového obrazu o složitosti dané nemoci.

6 Použitá literatura

1. Kissel JT (1999) Facioscapulohumeral dystrophy. Seminars in Neurology 19: 35-43.
2. van Deutekom JCT, Wijmenga C, Vantienhoven EAE, Gruter AM, Frants RR, et al. (1993) Fshd Associated DNA Rearrangements Are Due to Deletions of Integral Copies of a 3.2 Kb Tandemly Repeated Unit. Human Molecular Genetics 2: 2037-2042.
3. Gilbert JR, Stajich JM, Wall S, Carter SC, Qiu H, et al. (1993) Evidence for Heterogeneity in Facioscapulohumeral Muscular-Dystrophy (Fshd). American Journal of Human Genetics 53: 401-408.
4. Gabriels J, Beckers MC, Ding H, De Vriese A, Plaisance S, et al. (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene 236: 25-32.
5. Lyle R, Wright TJ, Clark LN, Hewitt JE (1995) Fshd-Associated Repeat, D4z4, Is a Member of a Dispersed Family of Homeobox-Containing Repeats, Subsets of Which Are Clustered on the Short Arms of the Acrocentric Chromosomes. Genomics 28: 389-397.
6. Tupler R, Berardinelli A, Barbierato L, Frants R, Hewitt JE, et al. (1996) Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. Journal of Medical Genetics 33: 366-370.
7. Lemmers RJLF, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJB, et al. (2002) Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. Nature Genetics 32: 235-236.
8. de Greef JC, Frants RR, van der Maarel SM (2008) Epigenetic mechanisms of facioscapulohumeral muscular dystrophy. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 647: 94-102.
9. Bickmore WA, van der Maarel SM (2003) Perturbations of chromatin structure in human genetic disease: recent advances. Human Molecular Genetics 12: R207-R213.
10. Gabellini D, Green MR, Tupler R (2002) Inappropriate gene activation in FSHD: A repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell 110: 339-348.
11. Tam R, Smith KP, Lawrence JB (2004) The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. Journal of Cell Biology 167: 269-279.
12. van Geel M, Dickson MC, Beck AF, Bolland DJ, Frants RR, et al. (2002) Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. Genomics 79: 210-217.
13. van Overveld PGM, Lemmers RJFL, Sandkuijl LA, Enthoven L, Winokur ST, et al. (2003) Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nature Genetics 35: 315-317.
14. Jiang GC, Yang F, van Overveld PGM, Vedanarayanan V, van der Maarel S, et al. (2003) Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Human Molecular Genetics 12: 2909-2921.
15. Yang F, Shao CB, Vedanarayanan V, Ehrlich M (2004) Cytogenetic and immuno-FISH analysis of the 4q subtelomeric region, which is associated with facioscapulohumeral muscular dystrophy. Chromosoma 112: 350-359.
16. Deschamps J, Meijlink F (1992) Mammalian homeobox genes in normal development and neoplasia. Crit Rev Oncog 3: 117-173.
17. Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, et al. (2007) DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of

PITX1. Proceedings of the National Academy of Sciences of the United States of America 104: 18157-18162.
18. Kowaljow V, Marcowycz A, Ansseau E, Conde CB, Sauvage S, et al. (2007) The DUX4 gene at the FSHDIA locus encodes a pro-apoptotic protein. Neuromuscular Disorders 17: 611-623.
19. Gabellini D, D'Antona G, Moggio M, Prelle A, Zecca C, et al. (2006) Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. Nature 439: 973-977.
20. Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R (2007) Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology 68: 569-577.
21. Laoudj-Chenivesse D, Carnac G, Bisbal C, Hugon G, Bouillot S, et al. (2005) Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. Journal of Molecular Medicine-Jmm 83: 216-224.
22. Deak KL, Lemmers RJLF, Stajich JM, Klooster R, Tawil R, et al. (2007) Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. Neurology 68: 578-582.
23. Lemmers RJLF, Osborn M, Haaf T, Rogers M, Frants RR, et al. (2003) D4F104S1 deletion in facioscapulohumeral muscular dystrophy - Phenotype, size, and detection. Neurology 61: 178-183.
24. Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E, et al. (2012) A Long ncRNA Links Copy Number Variation to a Polycomb/Trithorax Epigenetic Switch in FSHD Muscular Dystrophy. Cell 149.
25. Cabianca DS, Casa V, Gabellini D (2012) A novel molecular mechanism in human genetic disease A DNA repeat-derived lncRNA. Rna Biology 9: 1211-1217.
26. Lemmers RJ, Tawil R, Petek LM, Balog J, Block GJ, et al. (2012) Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. Nature Genetics 44: 1370-1374.
27. Jones TI, Chen JC, Rahimov F, Homma S, Arashiro P, et al. (2012) Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. Human Molecular Genetics 21: 4419-4430.
28. Stadler G, Rahimov F, King OD, Chen JC, Robin JD, et al. (2013) Telomere position effect regulates DUX4 in human facioscapulohumeral muscular dystrophy. Nat Struct Mol Biol 20: 671-678.
29. Zubrzycka-Gaarn EE, Bulman DE, Karpati G, Burghes AH, Belfall B, et al. (1988) The Duchenne muscular dystrophy gene product is localized in sarcolemma of human skeletal muscle. Nature 333: 466-469.
30. Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, et al. (1987) Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell 50: 509-517.
31. Monaco AP, Kunkel LM (1988) Cloning of the Duchenne/Becker muscular dystrophy locus. Adv Hum Genet 17: 61-98.
32. Guglieri M, Magri F, D'Angelo MG, Prelle A, Morandi L, et al. (2008) Clinical, molecular, and protein correlations in a large sample of genetically diagnosed Italian limb girdle muscular dystrophy patients. Hum Mutat 29: 258-266.
33. Fanin M, Nascimbeni AC, Aurino S, Tasca E, Pegoraro E, et al. (2009) Frequency of LGMD gene mutations in Italian patients with distinct clinical phenotypes. Neurology 72: 1432-1435.

34. van der Kooi AJ, Frankhuizen WS, Barth PG, Howeler CJ, Padberg GW, et al. (2007) Limb-girdle muscular dystrophy in the Netherlands: gene defect identified in half the families. Neurology 68: 2125-2128.
35. Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, et al. (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 81: 27-40.
36. Brockington M, Blake DJ, Prandini P, Brown SC, Torelli S, et al. (2001) Mutations in the fukutin-related protein gene (FKRP) cause a form of congenital muscular dystrophy with secondary laminin alpha2 deficiency and abnormal glycosylation of alpha-dystroglycan. Am J Hum Genet 69: 1198-1209.
37. Frosk P, Weiler T, Nylen E, Sudha T, Greenberg CR, et al. (2002) Limb-girdle muscular dystrophy type 2H associated with mutation in TRIM32, a putative E3-ubiquitin-ligase gene. Am J Hum Genet 70: 663-672.
38. Kirschner J, Lochmuller H (2011) Sarcoglycanopathies. Handb Clin Neurol 101: 41-46.
39. Liu J, Aoki M, Illa I, Wu C, Fardeau M, et al. (1998) Dysferlin, a novel skeletal muscle gene, is mutated in Miyoshi myopathy and limb girdle muscular dystrophy. Nat Genet 20: 31-36.
40. Haravuori H, Vihola A, Straub V, Auranen M, Richard I, et al. (2001) Secondary calpain3 deficiency in 2q-linked muscular dystrophy: titin is the candidate gene. Neurology 56: 869-877.
41. Sarkozy A, Hicks D, Hudson J, Laval SH, Barresi R, et al. (2013) ANO5 gene analysis in a large cohort of patients with anoctaminopathy: confirmation of male prevalence and high occurrence of the common exon 5 gene mutation. Hum Mutat 34: 1111-1118.
42. Christiano AM, Hoffman GG, Chung-Honet LC, Lee S, Cheng W, et al. (1994) Structural organization of the human type VII collagen gene (COL7A1), composed of more exons than any previously characterized gene. Genomics 21: 169-179.
43. Fine JD, Eady RA, Bauer EA, Bauer JW, Bruckner-Tuderman L, et al. (2008) The classification of inherited epidermolysis bullosa (EB): Report of the Third International Consensus Meeting on Diagnosis and Classification of EB. J Am Acad Dermatol 58: 931-950.
44. Dang N, Murrell DF (2008) Mutation analysis and characterization of COL7A1 mutations in dystrophic epidermolysis bullosa. Exp Dermatol 17: 553-568.
45. Almaani N, Liu L, Dopping-Hepenstal PJ, Lai-Cheong JE, Wong A, et al. (2011) Identical glycine substitution mutations in type VII collagen may underlie both dominant and recessive forms of dystrophic epidermolysis bullosa. Acta Derm Venereol 91: 262-266.
46. Burgeson RE (1993) Type VII collagen, anchoring fibrils, and epidermolysis bullosa. J Invest Dermatol 101: 252-255.
47. Kivirikko KI (1993) Collagens and their abnormalities in a wide spectrum of diseases. Ann Med 25: 113-126.
48. Persikov AV, Pillitteri RJ, Amin P, Schwarze U, Byers PH, et al. (2004) Stability related bias in residues replacing glycines within the collagen triple helix (Gly-Xaa-Yaa) in inherited connective tissue disorders. Hum Mutat 24: 330-337.
49. Hovnanian A, Pollack E, Hilal L, Rochat A, Prost C, et al. (1993) A missense mutation in the rod domain of keratin 14 associated with recessive epidermolysis bullosa simplex. Nat Genet 3: 327-332.

50. Rugg EL, Morley SM, Smith FJ, Boxer M, Tidman MJ, et al. (1993) Missing links: Weber-Cockayne keratin mutations implicate the L12 linker domain in effective cytoskeleton function. Nat Genet 5: 294-300.
51. Stephens K, Ehrlich P, Weaver M, Le R, Spencer A, et al. (1997) Primers for exon-specific amplification of the KRT5 gene: identification of novel and recurrent mutations in epidermolysis bullosa simplex patients. J Invest Dermatol 108: 349-353.
52. Bolling MC, Lemmink HH, Jansen GH, Jonkman MF (2011) Mutations in KRT5 and KRT14 cause epidermolysis bullosa simplex in 75% of the patients. Br J Dermatol 164: 637-644.
53. Schweizer J, Bowden PE, Coulombe PA, Langbein L, Lane EB, et al. (2006) New consensus nomenclature for mammalian keratins. J Cell Biol 174: 169-174.
54. Fuchs E (1996) The cytoskeleton and disease: genetic disorders of intermediate filaments. Annu Rev Genet 30: 197-231.
55. Herrmann H, Aebi U (2000) Intermediate filaments and their associates: multi-talented structural elements specifying cytoarchitecture and cytodynamics. Curr Opin Cell Biol 12: 79-90.
56. Wu KC, Bryan JT, Morasso MI, Jang SI, Lee JH, et al. (2000) Coiled-coil trigger motifs in the 1B and 2B rod domain segments are required for the stability of keratin intermediate filaments. Mol Biol Cell 11: 3539-3558.
57. Huber M, Rettler I, Bernasconi K, Frenk E, Lavrijsen SP, et al. (1995) Mutations of keratinocyte transglutaminase in lamellar ichthyosis. Science 267: 525-528.
58. Russell LJ, DiGiovanna JJ, Rogers GR, Steinert PM, Hashem N, et al. (1995) Mutations in the gene for transglutaminase 1 in autosomal recessive lamellar ichthyosis. Nat Genet 9: 279-283.
59. Lefevre C, Audebert S, Jobard F, Bouadjar B, Lakhdar H, et al. (2003) Mutations in the transporter ABCA12 are associated with lamellar ichthyosis type 2. Hum Mol Genet 12: 2369-2378.
60. Lefevre C, Bouadjar B, Karaduman A, Jobard F, Saker S, et al. (2004) Mutations in ichthyin a new gene on chromosome 5q33 in a new form of autosomal recessive congenital ichthyosis. Hum Mol Genet 13: 2473-2482.
61. Lefevre C, Bouadjar B, Ferrand V, Tadini G, Megarbane A, et al. (2006) Mutations in a new cytochrome P450 gene in lamellar ichthyosis type 3. Hum Mol Genet 15: 767-776.
62. Jobard F, Lefevre C, Karaduman A, Blanchet-Bardon C, Emre S, et al. (2002) Lipoxygenase-3 (ALOXE3) and 12(R)-lipoxygenase (ALOX12B) are mutated in non-bullous congenital ichthyosiform erythroderma (NCIE) linked to chromosome 17p13.1. Hum Mol Genet 11: 107-113.
63. Grall A, Guaguere E, Planchais S, Grond S, Bourrat E, et al. (2012) PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nat Genet 44: 140-147.
64. Israeli S, Khamaysi Z, Fuchs-Telem D, Nousbeck J, Bergman R, et al. (2011) A mutation in LIPN, encoding epidermal lipase N, causes a late-onset form of autosomal-recessive congenital ichthyosis. Am J Hum Genet 88: 482-487.
65. Eckl KM, Tidhar R, Thiele H, Oji V, Hausser I, et al. (2013) Impaired epidermal ceramide synthesis causes autosomal recessive congenital ichthyosis and reveals the importance of ceramide acyl chain length. J Invest Dermatol 133: 2202-2211.
66. Rothnagel JA, Dominey AM, Dempsey LD, Longley MA, Greenhalgh DA, et al. (1992) Mutations in the rod domains of keratins 1 and 10 in epidermolytic hyperkeratosis. Science 257: 1128-1130.

67. Rothnagel JA, Traupe H, Wojcik S, Huber M, Hohl D, et al. (1994) Mutations in the rod domain of keratin 2e in patients with ichthyosis bullosa of Siemens. Nat Genet 7: 485-490.
68. Webster D, France JT, Shapiro LJ, Weiss R (1978) X-linked ichthyosis due to steroid-sulphatase deficiency. Lancet 1: 70-72.
69. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, et al. (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 38: 337-342.
70. Bouwstra JA, Ponec M (2006) The skin barrier in healthy and diseased state. Biochim Biophys Acta 1758: 2080-2095.
71. Iwai I, Han H, den Hollander L, Svensson S, Ofverstedt LG, et al. (2012) The human skin barrier is organized as stacked bilayers of fully extended ceramides with cholesterol molecules associated with the ceramide sphingoid moiety. J Invest Dermatol 132: 2215-2225.
72. Masukawa Y, Narita H, Sato H, Naoe A, Kondo N, et al. (2009) Comprehensive quantification of ceramide species in human stratum corneum. J Lipid Res 50: 1708-1719.
73. Meguro S, Arai Y, Masukawa Y, Uie K, Tokimitsu I (2000) Relationship between covalently bound ceramides and transepidermal water loss (TEWL). Arch Dermatol Res 292: 463-468.
74. Akiyama M, Sugiyama-Nakagiri Y, Sakai K, McMillan JR, Goto M, et al. (2005) Mutations in lipid transporter ABCA12 in harlequin ichthyosis and functional recovery by corrective gene transfer. J Clin Invest 115: 1777-1784.
75. Kelsell DP, Norgett EE, Unsworth H, Teh MT, Cullup T, et al. (2005) Mutations in ABCA12 underlie the severe congenital skin disease harlequin ichthyosis. Am J Hum Genet 76: 794-803.
76. Akiyama M, Dale BA, Smith LT, Shimizu H, Holbrook KA (1998) Regional difference in expression of characteristic abnormality of harlequin ichthyosis in affected fetuses. Prenat Diagn 18: 425-436.
77. Akiyama M, Yoneda K, Kim SY, Koyama H, Shimizu H (1996) Cornified cell envelope proteins and keratins are normally distributed in harlequin ichthyosis. J Cutan Pathol 23: 571-575.
78. Zheng Y, Yin H, Boeglin WE, Elias PM, Crumrine D, et al. (2011) Lipoxygenases mediate the effect of essential fatty acid in skin barrier formation: a proposed role in releasing omega-hydroxyceramide for construction of the corneocyte lipid envelope. J Biol Chem 286: 24046-24056.
79. Krieg P, Furstenberger G (2014) The role of lipoxygenases in epidermis. Biochim Biophys Acta 1841: 390-400.
80. Munoz-Garcia A, Thomas CP, Keeney DS, Zheng Y, Brash AR (2014) The importance of the lipoxygenase-hepoxilin pathway in the mammalian epidermal barrier. Biochim Biophys Acta 1841: 401-408.
81. Dahlqvist J, Klar J, Hausser I, Anton-Lamprecht I, Pigg MH, et al. (2007) Congenital ichthyosis: mutations in ichthyin are associated with specific structural abnormalities in the granular layer of epidermis. J Med Genet 44: 615-620.
82. Dahlqvist J, Westermark GT, Vahlquist A, Dahl N (2012) Ichthyin/NIPAL4 localizes to keratins and desmosomes in epidermis and Ichthyin mutations affect epidermal lipid metabolism. Arch Dermatol Res 304: 377-386.
83. Hohl D (1990) Cornified cell envelope. Dermatologica 180: 201-211.
84. Greenberg CS, Birckbichler PJ, Rice RH (1991) Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues. FASEB J 5: 3071-3077.

85. Nemes Z, Marekov LN, Fesus L, Steinert PM (1999) A novel function for transglutaminase 1: attachment of long-chain omega-hydroxyceramides to involucrin by ester bond formation. Proc Natl Acad Sci U S A 96: 8402-8407.
86. Elias PM, Schmuth M, Uchida Y, Rice RH, Behne M, et al. (2002) Basis for the permeability barrier abnormality in lamellar ichthyosis. Exp Dermatol 11: 248-256.
87. Vahlquist A, Ganemo A, Virtanen M (2008) Congenital ichthyosis: an overview of current and emerging therapies. Acta Derm Venereol 88: 4-14.
88. Elias PM, Crumrine D, Rassner U, Hachem JP, Menon GK, et al. (2004) Basis for abnormal desquamation and permeability barrier dysfunction in RXLI. J Invest Dermatol 122: 314-319.
89. Chipev CC, Korge BP, Markova N, Bale SJ, DiGiovanna JJ, et al. (1992) A leucine----proline mutation in the H1 subdomain of keratin 1 causes epidermolytic hyperkeratosis. Cell 70: 821-828.
90. Cheng J, Syder AJ, Yu QC, Letai A, Paller AS, et al. (1992) The genetic basis of epidermolytic hyperkeratosis: a disorder of differentiation-specific epidermal keratin genes. Cell 70: 811-819.
91. McLean WH, Morley SM, Lane EB, Eady RA, Griffiths WA, et al. (1994) Ichthyosis bullosa of Siemens--a disease involving keratin 2e. J Invest Dermatol 103: 277-281.
92. Kremer H, Zeeuwen P, McLean WH, Mariman EC, Lane EB, et al. (1994) Ichthyosis bullosa of Siemens is caused by mutations in the keratin 2e gene. J Invest Dermatol 103: 286-289.
93. Sandilands A, Sutherland C, Irvine AD, McLean WH (2009) Filaggrin in the frontline: role in skin barrier function and disease. J Cell Sci 122: 1285-1294.
94. Rawlings AV, Harding CR (2004) Moisturization and skin barrier function. Dermatol Ther 17 Suppl 1: 43-48.
95. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, et al. (2007) Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. Nat Genet 39: 650-654.