2. LÉKAŘSKÁ FAKULTA
UNIVERZITA KARLOVA V PRAZE

ÚSTAV BIOLOGIE A LÉKAŘSKÉ GENETIKY

CYSTICKÁ FIBRÓZA A NOVOROZENECKÝ SCREENING CYSTICKÉ FIBRÓZY V ČR

MUDr. Miroslava Balaščaková

Školitel:
Prof. MUDr. Milan Macek, DrSc.

DIZERTAČNÍ PRÁCE

Praha 2010

PODĚKOVÁNÍ

Ráda bych poděkovala svému školiteli Prof. MUDr. Milanu Mackovi Jr. DrSc. za odborné vedení a vstřícný přístup v průběhu mého postgraduálního studia.

Doc. MUDr. Felixu Votavovi za laskavé vedení, odbornou pomoc a ochotu se vším mi poradit.

Dále děkuji Doc. MUDr. Věře Vávrové a MUDr. Veronice Skalické a RNDr. Daně Zemkové z CF Centra FN Motol za skvělou spolupráci, cenné rady a laskavé poskytnutí klinických a antropometrických dat pacientů s cystickou fibrózou.

Dále děkuji svým kolegyním a kolegům z laboratoře Centra Cystické fibrózy nejen za spolupráci, ale především za milou atmosféru a přátelskou podporu, které si velmi vážím.

Také děkuji své rodině, zejména manželovi za podporu během celého studia.

OBSAH

1. ÚVOD

1.1. Cystická fibróza

1.1.1. Definice

Cystická fibróza (CF; OMIM #217900; patří mezi nejčastější závažné autozomálně recesivní onemocnění u evropských populací. Údaje o jejím výskytu se však značně liší. Obecně se udává incidence v evropské populaci 1:2500 -1: 4500 (1,2). V České republice byla incidence CF stanovena dřívějšími epidemiologickými (3,4,5) a následně molekulárně genetickými metodami (6,7) na 1:2736 novorozenců.

Jedná se o multiorgánové onemocnění primárně postihující plíce, exokrinní funkci pankreatu, játra, střeva a mužský reprodukční systém. Dosud není k dispozici kauzální léčba, nicméně délka a kvalita života pacientů s CF se v posledních desetiletích významně zlepšila. Zatímco v 50. letech minulého století umírala většina dětí v kojeneckém věku, v současnosti se ve vyspělých zemích nemocní mohou dožít i 50 let. Zlepšuje se i kvalita života nemocných. V České republice podle dat z národního registu (www.cfregistr.cz) je medián věku 25,3 roků a např. v USA činí medián 36,9 let. Intenzivní léčba na jedné straně prodlužuje střední délku života, ale na druhou stranu přispívá k rozvoji sekundárních komplikací CF (8).

1.1.2. Klinická charakteristika cystické fibrózy; klasická a atypická forma

Toto multiorgánové onemocnění je charakterizováno progresivním postižením sino-pulmonálního systému, postižením vývoje mužského reprodukčního traktu dané kongenitální bilaterální obstrukcí vas deferens (CBAVD; 9), insuficiencí zevní sekrece pankreatu a diagnosticky vysokou koncentrací chloridů v potu. Vedle těchto čtyřech základních příznaků, které by vždy měly upozornit na diagnózu, však CF postihuje téměř všechny systémy organismu. Jaterní onemocnění může být prvním projevem nemoci ve formě protrahované novorozenecké žloutenky nebo jaterní steatózy; může progredovat do jaterní cirhózy, podílející se přibližně na 3,5% mortalitě. Na CF vázaný diabetes (CFRD; 10) se vyskytuje poměrně často po 10. roce věku. Sekundárně je přítomna osteoporóza, vyskytují se vaskulitidy, artropatie, onemocněni ledvin (nefrolitiáza, onemocněni ledvin vyvolané antibiotiky), amyloidóza a řada dalších (1,2,8). Na prvním místě je třeba ale zmínit, že CF se nejčastěji projevuje chronickým, obstrukčním sino-pulmonálním

onemocněním, které je dané přítomností hustého hlenu v dýchacích cestách. Rozvíjí se chronická infekce způsobená hlavně bakteriálními kmeny: *Staphylococcus aureus, Pseudomonas aeruginosa, Burgholderia cepacia,* které vedou k tvorbě bronchiektázií a rozvoji respiračního selhání. Časté jsou i opakované sinusitidy a tvorba nosních polypů. V důsledku postižení gastrointestinálního systému se již prenatálně objevuje pankreatická exokrinní dysfunkce, která způsobuje po narození steatoreu a poruchu vstřebávání živin (tzv. pankreatická insuficience). Vlivem toho je narušen stav výživy, růst a vývoj nemocných. Někteří pacienti však mohou mít reziduální funkci pankreatu zachovanou a jsou tedy pankreaticky suficientní.

Neonatální mekoniový ileus se vyskytuje u cca 13% novorozenců s CF a je často prvním příznakem této choroby (11, 14). Z klinických příznaků mohou CF signalizovat metabolické poruchy (hypoelektrolytemický šok, metabolická alkalóza) nebo hypoproteinemie s edémy, a to zejména v kojeneckém věku. Průběh může být různě závažný a o délce života rozhoduje především progresivním postižení plic vyvolané chronickým zánětem, který se podílí na 85% z celkové mortality. Z tohoto důvodu rozlišujeme „klasickou“ a atypickou (2-10% případů) formu CF. Postižení jedinci s klasickou formou mají vyjádřeny dýchací a trávicí obtíže a hladiny chloridů v potu jsou zvýšeny nad 60 mmol/L (1,2,8). Pacienti s atypickou formou mají zpravidla chloridy v potu mezi 30-60 mmo/l, jsou pankreaticky suficientní a jejich obtíže jsou mírnější event. monosymptomatické např. ve formě opakovaných zánětů vedlejších nosních dutin s nosními polypy, postižením plic či chronické pankreatitidy. U dospělých mužů se může atypická forma CF projevit pouze jako porucha reprodukce ve formě obstrukční azoospermie vzniklé na podkladě CBAVD a hraničními hodnotami chloridů v potu.

Vzhledem k tomu, že se u atypické formy CF ve vyšším věku mohou další obtíže rozvinout, je třeba tyto pacienty dlouhodobě dispenzarizovat. Ze stručného výčtu klinických projevů je patrné, že klinická diagnostika je komplexní a mnohdy obtížná. I základní příznaky jsou často nenápadné a nemusí být vyjádřeny, nebo se mohou rozvinout v pozdějším věku (1,2,15). Na CF se zaměříme vždy i na základě pozitivní rodinné anamnézy výskytu tohoto onemocnění. CF je velmi variabilní onemocnění se značnou interfamiliární a intrafamiliární variabilitou. Variabilita CF se nejvíce projevuje v rozdílné závažnosti průběhu sino-pulmonálního onemocnění (1,12). Korelace genotypu *CFTR* genu

s fenotypem v poslední době prokázaly, že nejnižší variabilita klinického průběhu CF je v potní žláze a v mužském reprodukčním traktu, o něco vyšší v pankreatu a nejvyšší v plicích (1,15). CF tedy zahrnuje relativně široké spektrum příznaků, které mají navíc variabilní věk nástupu, rozdílnou míru progrese a klinického vývoje (8,15).

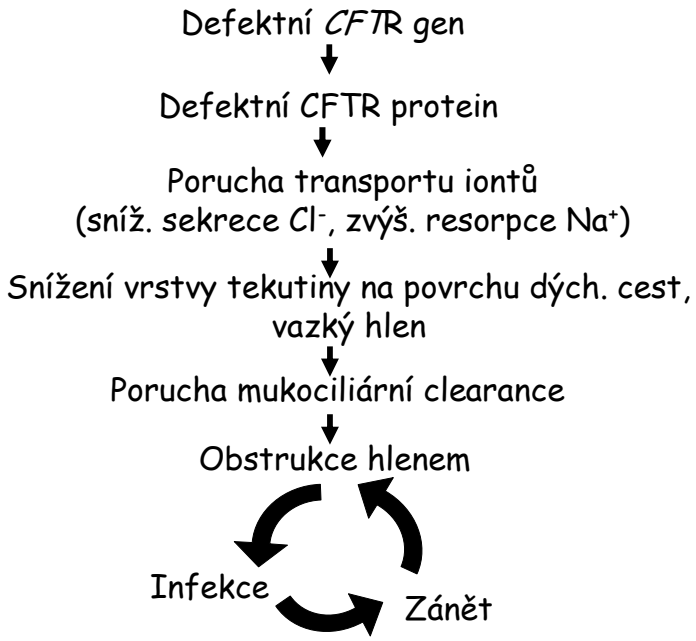
1.1.3. Patofyziologie

Patogeneza CF je orgánově specifická a projevuje se v různé intenzitě u jednotlivých cílových orgánů. Přes intenzivní výzkum v této oblasti nebylo jednotné vysvětlení distribuce a regulace CFTR exprese v postižených orgánech dosud zcela vysvětleno. *CFTR* je u člověka *in vivo* specificky exprimován na povrchu většiny epiteliálních tkání, které korelují s orgány postiženými cystickou fibrózou. V poslední době byla také popsána i non-epiteliální *CFTR* exprese např. v kardiomyocytech, lymfocytech, endoteliálních buněk cév. (16,17,18). Je jisté, že CFTR protein funguje zajména jako chloridový kanál na apikální membráně buněk, zodpovídá za transport elektrolytů a vody epitelialními membránami a současně reguluje funkci dalších proteinů. Porucha jeho funkce omezuje sekreci chloridů a podporuje resorpci natria, čímž se mění složeni i fyzikalně-chemické vlastnosti sekretů na epiteliálním povrchu buněk. Starší hypotéza rozvoje patologie předpokláda v dýchacich cestách vysoký obsah soli v periciliární tekutině (PCL) v důsledku poruchy resorpce, zprostředkované apikálně lokalizovaným chloridovým kanálem CFTR, v lumen dýchacích cest - tj. podobně jako je tomu ve vývodech potních žlaz. Vysoký obsah soli degraduje aktivitu antimikrobialních peptidů a vede k chronické infekci (19).

Naproti tomu druhá, častěji uznávaná hypotéza předpokladá nízký obsah soli v důsledku neprůchodnosti chloridoveho kanálu a omezené sekreci chloridových iontů z nitra buňky (1). Vzniká tak porucha sekrece chloridů a zvýšená absorpce natria a osmotickou silou i vody, což vede k dehydrataci periciliarní tekutiny a poruše mukociliarni clearance (viz. Schéma 1). Zahuštěný hlen působí obstrukci dýchacích cest, usídlí se v něm bakterie a dochází k chronické infekci – viz. *Schéma 1*. Byla popsána úloha proteinu CFTR jako internalizačního- fagocytárního receptoru pro *Pseudomonas aeruginosa* a *Burgholderia cepacia*. Reakcí na infekci je nepřiměřená zanětlivá odpověď, jejímž důsledkem je ireverzibilní poškození epitelu dýchacích cest proteázami, prozánětlivými cytokiny a oxidačním stresem. Stejný mechanismus poruchy sekrece chloridových iontů (Cl^-) a

bikarbonátových iontů (HCO_3^-) spolu s hyperabsorbcí sodných iontů (Na^+) je popisován v epiteliálních tkáních pankreatu a střeva.

Schéma 1. Patofyziologie respiračního onemocnění CF



1.1.4. Genetická podstata cystické fibrózy

1.1.4.1. Gen *CFTR*

Padesát let od doby prvního popisu nemoci (Andersenová, 1938; 20) zůstala biochemická podstata nemoci neznámá. V roce 1983 byl poprvé popsán základní defekt impermeability buněčných membrán pro chloridy (21,22), který nastartoval další výzkum a pátrání po genu zodpovědném za CF. Nejprve byl gen pro CF lokalizován na chromozom 7 pomocí segregačně ověřené vazby s genem pro *paraoxanázu (PON)*, což je polymorfní alela u mnoha lidských populací. Brzy poté byl identifikován anonymní maker *DOCRI-917*, který segregoval s lokusem CF ve vzdálenosti přibližně 15 cM (1 cM odpovídá ~$1x10^6$ pb). Tento marker byl lokalizován na dlouhé raménko chromozomu 7 do lokusu *7q31.2*. (112.1 Mpb). Upřesnění lokalizace genu pro CF umožnilo objev dalších vázaných SNP (single nucleotide polymorphisms) jako *metH* a *D7S8*, které jsou v daleko užší vazbě (< 1cM) s lokusem CF (22,23). Objev extragenových SNP markerů vázaných ke genu pro CF měl i

významný praktický dopad. Umožnil totiž nepřímou molekulárně genetickou prenatální diagnostiku a odhalení nosičů CF. A to zjištěním segregace markerů CF vyšetřením probanda a jeho rodičů, obligátních heterozygotů (25).

V roce 1989 byl s konečnou platností určen gen *CFTR (Cystic Fibrosis Transmembrane Conductance Regulátor,* překládáno jako regulátor transmembránové vodivosti iontů), s konečnou platností určen pozičním klonováním kombinací techniky „chromosome jumping" s fyzikálním mapováním, izolací potenciálních exonových sekvencí a genetickou analýzou segregace kandidátních regionů v rodinách s CF (26, 27, 28). *CFTR* gen má 27 exonů a rozprostírá se v oblasti přibližně 250 kpb a vytváří *CFTR*-mRNA 6.5 kpb dlouhou (26, 27). Dosud však není přesně známo jaké genetické a negenetické faktory regulují transkripci genu CFTR. Protisměrná oblast k *CFTR* je sice bohatá na guanin a cytosin, ale neobsahuje klasický box TATA. Pozitivní a negativní transkripční faktory byly identifikovány v pozici *cis* a v poslední době i v pozici *trans*. Tyto transkripční faktory jsou evolučně konzervované. Bylo také prokázáno, že transkripce genu *CFTR začíná na mnoha místech genu CFTR*, tj. například na methioninu v exonu 4. Funkční význam tohoto alternativního startu transkripce není zatím znám. Transkripce je také výrazně tkáňově specifická, protože většina neepiteliálních buňek má pouze minimální množství CFTR-mRNA. V současné době je proto snaha nalézt tkáňově specifické transkripční faktory, které by zodpovídaly za rozdílnou expresi CFTR proteinu v různých tkáních a tak objektivizovaly variabilní průběh onemocnění (1,18).

1.1.4.2. Protein CFTR

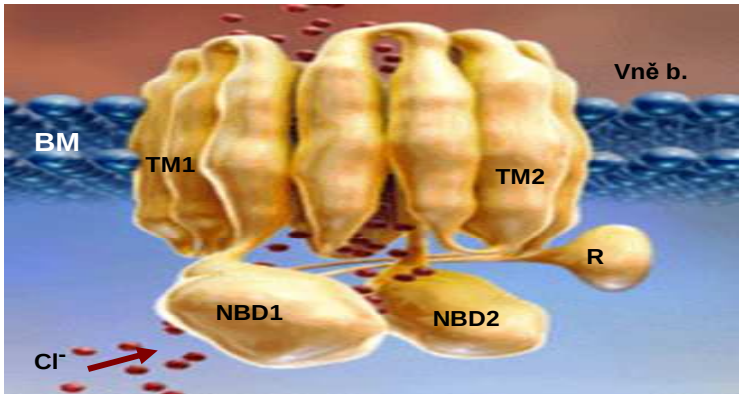
Protein CFTR se skládá z 1480 aminokyselin a patří do skupiny tzv. „ATP-binding cassette" transmembránových transportních proteinů. Plní především funkci kanálu, přes který jsou transportovány ionty (hlavně chloridy, ale i bikarbonát; 29,30,31). Současně s tím však reguluje funkci resorpčního na amilorid senzitivního epitelového sodíkového kanálu (ENaC-amiloride sensitive epithelial sodium channel; 32,33) a pomocí aktivní exkrece ATP, přes interakci s P-glykoproteinem, ovlivňuje i funkci dalšího důležitého chloridového kanálu – ORCC (outwardly rectified chloride channel; 34,35,36).

CFTR protein obsahuje pět základních domén. Dva opakující se transmembránové segmenty: TM1 a TM2 (Transmembrane Domains; synonymum MSD1 a MSD2 –

„membrane spanning domain“) tvoří vlastní kanál v apikální buněčné membráně. V cytoplasmě jsou lokalizované domény vázající aktivované nukleotidy, tj. hlavně ATP (adenosintrifosfát) – tzv. NBD1 a NBD2 (Nucleotide Binding Domains) a nepárová regulační doména "R" (Regulatory Domain), která funguje jako ventil a otevírá/zavírá vlastní kanál**.** Schematicky je CFTR protein znázorněn na Obr. č. 1.

CFTR protein funguje hlavně jako chloridový kanál, který je aktivovaný cAMP prostřednictvím fosforylace "R" doména proteinkinázou A a následnou hydrolýzou ATP na povrchu NBD1 a NBD2. Domény NBD1 a NBD2 po aktivaci změní konformaci a „vertikálním posunem“ fosforylované regulační domény „R“ se kanál otevře. Tak mohou Cl^- ionty proudit přes epiteliální membránu buněk (32,33).

Obr.č. 1. Schéma terciární struktury proteinu CFTR



Legenda: TM1, TM2 – transmembránová doména 1,2; NBD1, NBD2 – nucleotide binding doména; R – regulační doména; sekrece chloridových iontů přes membránu.

1.1.4.3. Úloha ENaC kanálu

Vedle defektní sekrece chloridů, ztráta funkce CFTR proteinu zvyšuje absorpci sodíku aktivací ENaC kanálu (37,38). ENaC je heteromerický transmembránový protein složený ze 3 podjednotek (α, β, γ). Tyto 3 jednotky jsou kódovány třemi různými geny. *SCNN1A* (sodium channel nonvoltage-gated 1, alpha); *SCNN1B* (sodium channel nonvoltage-gated 1,

beta) a *SCNN1G* (sodium channel nonvoltage-gated 2, gamma) (39,40,41). V současnosti zůstává molekulární mechanismus regulace mezi ENaC kanálem a CFTR kanálem stále neobjasněn. Nicméně je všeobecně přijato, že se podílí na patofyziologii CF plicního onemocnění (37). Funkční studie prokázaly, že společná regulace a aktivita ENaC a CFTR proteinů determinuje složení periciliální tekutiny (PCL) v plicích (42,43). Sekrece chloridových iontů pomocí CFTR kanálu zvyšuje objem PCL, zatímco absorpce sodíkových iontů zprostředkovaná ENaC kanálem ji redukuje. Z toho vyplývá, že vyvážená součinnost mezi kanály ENaC a CFTR je významná pro udržení optimálního objemu PCL (42,44).

U klasické formy CF s patogenními mutacemi na obou alelách CFTR genu dochází k depleci PCL, což je vysvětlováno nedostatečnou sekrecí chloridových iontů skrz kanál CFTR a hyperabsorbcí sodíkových iontů pomocí kanálu ENaC (44).

1.1.5. Laboratorní diagnostika cystické fibrózy

Diagnostika CF se opírá především o klinický obraz. Nejjednodušší a nejčasnější je diagnostika v případě pozitivní rodinné anamnézy výskytu CF a/anebo pozitivního výsledku prenatální diagnostiky. U cca. 13% se CF v novorozeneckém věku projeví mekoniovým ileem (45).

Klinické podezření je potřeba potvrdit laboratorním průkazem poruchy proteinu CFTR. Ten spočívá ve zjištění vysoké koncentrace chloridů v potu (tzv. potní test), v průkazu přítomnosti 2 patogenních mutací *CFTR* genu v pozici *trans*, případně ve sníženém transepiteliálním rozdílu nosních potenciálů. U cca. 15% novorozenců se CF diagnostikuje na základě pozitivní rodinné anamnézy a/anebo mekoniového ileu.

Důležitým mezníkem v historii CF je objev vysoké koncentrace solí především chloridů v potu u pacientů s CF. Pot dětí s CF obsahuje 5 x víc solí než pot zdravých dětí, proto se v minulosti doporučoval dotaz na to, jak chutná polibek na čelo novorozence „kiss your baby“. Pro praktické použití této skutečnosti se vyvinula jednoduchá metoda stanovení koncentrace chloridů v potu podle Gibsona a Cooka (46,47). Tzv. potní test se provádí výhradně stimulací pocení pilokarpinovou iontoforézou na předloktí pacienta. Pot nejméně o objemu 100 mg se sbírá do filtračního papírku případně do plastikové kapiláry (v minimálním množství 15 ml) a následně se provádí kvantitativní analýza koncentrace

chloridů. Hodnoty nad 60 mmol/L zjištěné v laboratoři s obratem pacientů nejméně 150/rok, svědčí s velkou pravděpodobností pro CF. Za hraniční se považují hodnoty v mezích 30-60 mmol/L. Koncentrace chloridů v potu je zvýšena již při narození, ale je obtížné ji správně stanovit u novorozenců a shromáždit potřebné množství potu > 75-100 mg. Z tohoto důvodu má test dobrou výpovědní hodnotu po 2. týdnu věku donošeného novorozence (48). Vyšetření transepiteliálního rozdílu nosních potenciálů se provádí vzhledem k náročnosti a obtížné interpretovatelnosti pouze na několika pracovištích v Evropě.

U všech pozitivních potních testů se následně provádí molekulárně genetické vyšetření.

1.1.6. Mutace genu *CFTR*

Celosvětově bylo v genu *CFTR* dosud popsáno více než 1700 sekvenčních variací viz aktuální stav v Cystic Fibrosis Genetic Analysis Consortium Dabatase – CFGAC; http://www.genet.sick-kids.on.ca/cftr/ (49). Genetickou analýzou v rodinách nemocných s CF byla nalezena hlavní a nejčastější mutace F508del, delece 3 pb, která vede ke ztrátě fenylalaninu (50,1). Nachází se v exonu 10 a porušuje tak funkci domény NBD1 proteinu CFTR. Tato mutace má společný původ a nachází se celosvětově v průměru v 70 % testovaných alel genu *CFTR*. V Evropě její frekvence vykazuje sestupný gradient od severovýchodu na jihozápad.

Převážná většina dalších mutací je privátních a pouze 30 mutací (1717-1G>A, 1898+1G>A, 2184delA, 2789+5G>A, 3120+1G>A, 3659delC, 3849+10kbC>T, 621+1G>T, 711+1G>T, A455E, F508del, I507del, G542X, G551D, G85E, N1303K, R1162X, 1161delC, S549N, R117H, R334W, R347P, R553X, R560T,W1282X, 1078delT, 394delTT, CFTRdele2,3(21kb), 3272-26A>G a D1152H) se vyskytuje s četností více než 0,1 %, mnohé z těchto mutací jsou populačně specifické (51). V Evropě v zásadě narůstá heterogenita mutací od severu na jihovýchod (52). Patogenetický potenciál mutací byl vyšetřen expresními studiemi u necelého 1 % mutací a tak „kauzální vztah“ většiny mutací k rozvoji CF byl stanoven pouze jejich asociací s genotypem choroby u pacienta, u něhož byly původně identifikovány (19, 53). Z hlediska typu mutací se nejčastěji vyskytují mutace vedoucí k záměně aminokyseliny (tzv. „missense“ mutace). Kromě nich jsou

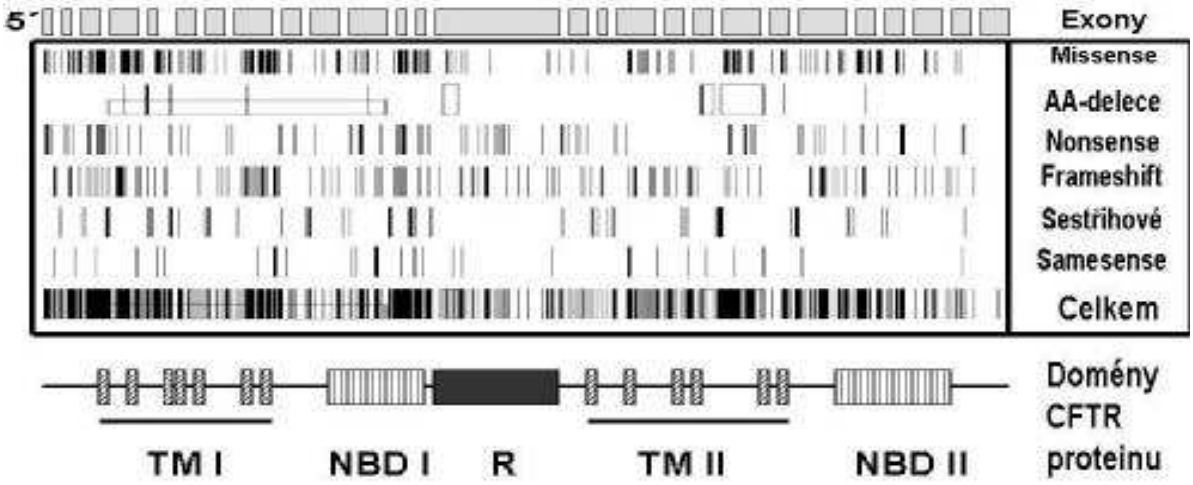
diagnostikovány nukleotidové substituce vedoucí k zavedení předčasného stop/terminačního kodónu („stop/nonsense“ mutace), posuny čtecího rámce („frameshift“ mutace) a sestřihové mutace („splicing“ mutace), včetně delecí/ inzercí jednotlivých kodonů, které nevedou k posunu čtecího rámce.

Větší delece nebo inzerce zahrnující celé exony a přilehlé intronové sekvence se vyskytují v 2-3 % všech případů. V naprosté většině případů se jedná o mutace germinální, zpravidla s ancestrální povahou (54), tj. vyskytují se v dané populaci dlouhodobě, a jsou důsledkem dávných historických selekčních procesů (55). Somatické mutace se v genu *CFTR* nebyly dosud nalezeny (1,49).

Z hlediska molekulárně genetické diagnostiky je podstatné, že efekt mutací na jednotlivé složky „fenotypu“ CF je tkáňově specifický, přičemž nejcitlivější je mužský reprodukční trakt (vas deferens/chámovod), pankreas, potní žláza a vedlejší nosní dutiny (56). Bronchiální postižení je na vlivu genu *CFTR* závislé nepřímo, díky výraznému vlivu prostředí (tj. infekcí) (57) a dalších genetických modifikátorů jako například *TGFbeta1, MBL, α_1 antitrypsin* atd., uplatňující se nepřímými a dosud ne zcela potvrzenými mechanismy (33, 53,58).

Na *Obr. č. 2*. je uvedena distribuce jednostlivých mutací v rámci celého genu *CFTR*. Většina mutací je rozložena rovnoměrně po celé kódující sekvenci a postihuje tak všechny jednotlivé funkční domény proteinu CFTR (mutace v exonech 3, 4, 6A a 7 postihují funkci první transmembránové domény; v exonech 9-12 porušují funkci NBD 1; v exonu 13 postihují funkcí regulační domény; v exonech 14A, 15, 16, 71A, 17B, 18 druhé transmembránové domény a v exonech 19-22 - NBD 2.

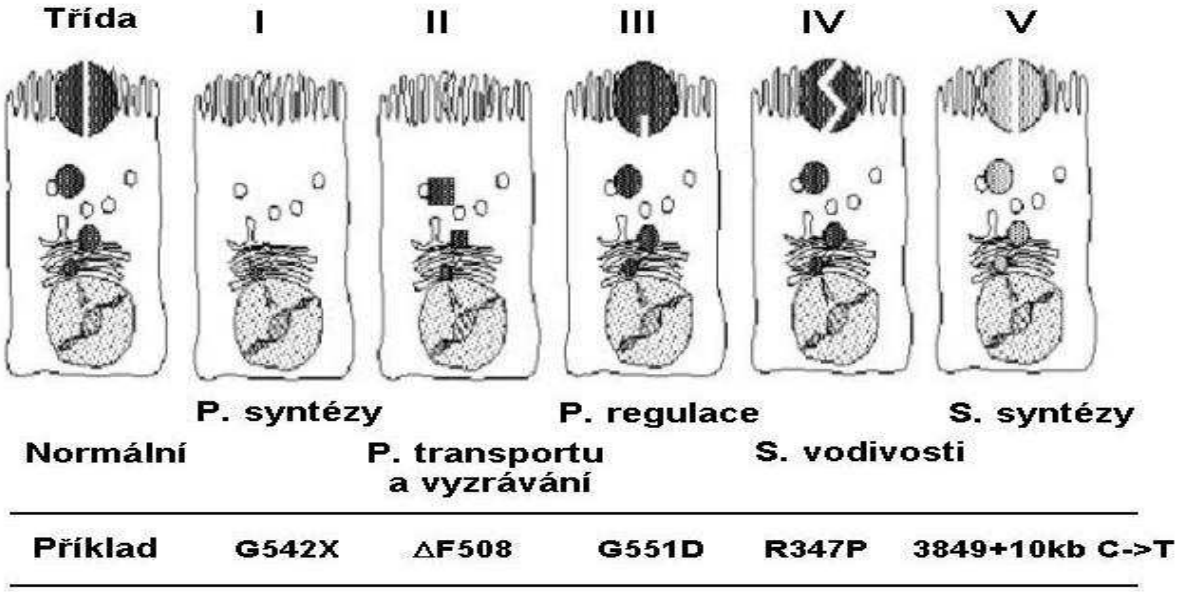
Obr.č. 2. Rozložení mutací v celé kódující sekvenci genu *CFTR* (upraveno podle Tsui, 1995, převzato z Vávrová a kol.2006)



Legenda: TM1 a TM2 - transmembranové domény, NBD1 a NBD2- „nucleotide binding“ domény, R – regulační doména, AA – delece, které nepoškozují čtecí rámec, jako např. F508del (Vávrová et al 2006)

Vzhledem k tomu, že je velmi obtížné u jednotlivých mutací určit jejich molekulární patogenezi a tak předpovědět jejich klinický dopad, byla snaha je rozdělit do tříd. Jak je znázorněno na *Obr. č. 3*. *CFTR* mutace se rozdělují do pěti základních tříd (59), podle toho na jakém stupni zasáhnou tvorbu, intracelulární transport, aktivaci a funkci proteinu CFTR. Třídy I.-III. jsou spojeny se závažnými formami, zatímco Třídy IV. a V. jsou asociovány s mírnější formu onemocnění. Třída I. syntetizuje zkrácený protein (především „nonsense“ mutace /G542X/ a „frameshift“ mutace /2143delT/ mutace), Třída II. způsobuje poruchu post-translačního zpracování (glykosylace) a intracelulárního transportu proteinu CFTR v Golgiho aparátu a hladkém endoplazmatické reticulu (např. mutace F508del), Třída III. způsobuje defektní regulaci nebo aktivaci chloridového kanálu (např. mutace G551D), Třída IV. snižuje průnik iontů chloridovým kanálem (např. mutace R347P) a Třída V. (např. mutace 3849+10kb C->T) je spojena se sníženou syntézou proteinu CFTR (1,2,18).

Obr. č. 3. Základní patogenetické třídy mutací genu *CFTR*



Legenda: P - porucha funkce, S – snížení funkce proteinu CFTR, převzato z Vávrová a kol. 2006.

Nicméně rozdělení alterací sekvence genu *CFTR* na polymorfismy, potenciálně patogenní varianty a patogenní mutace je spíše arbitrární, protože mezi nimi existuje z funkčního hlediska jakési kontinuum. Některé polymorfismy mohou snížit funkci proteinu CFTR řádově o několik procent, ale pokud je jich v genu *CFTR* více, může být jejich účinek aditivní. Vytvoří se tak „polyvariantní haplotyp“, který může fungovat jako patogenní mutace (60, 61,62).

Populačně genetická charakteristika mutací genu *CFTR* u české populace

Podle dostupných informací z českého registru (www.cfregistr.cz) je k roku 2007 sledováno celkem 471 pacientů (224 mužů a 247 žen). Z toho 38.2% je dospělých.
Z hlediska genotypů se jedná u 46.3% o homozygoty pro mutaci F508del, 35% pacientů je složených heterozygotů pro mutace F508del a jinou patogenní mutaci, 5.7% má mutaci F508del a druhou mutaci neidentifikovanou a 9.3% jsou kombinovaní heterozygoté pro non-F508del mutace, 1.5% jsou složení heterozygoté pro non-F508del mutaci a dosud neidentifikovanou mutaci a pouze 2.1% pacientů nemá nalezenou žádnou patogenní mutaci v genu *CFTR*.

V letech 1990 – 2009 bylo v CF Centru FN Motol vyšetřeno celkem 574 pacientů s CF. Celkem u nich bylo nalezeno 83 různých mutací. V *Tab. č. 1* je uveden přehled všech nalezených mutací u 1133 z celkem 1148 alel. K dosažení 98,7% záchytnosti přispělo také zavedení skenovací metody HRM, kvantifikační metody MLPA a sekvenace celého genu *CFTR*.

Tab. č. 1. Přehled mutací detekovaných v CF Centru FN Motol u českých pacientů

Mutace	**N**	**%**	**Mutace**	**N**	**%**
F508del (c.1520_1522delTCT)	768	66.90	1249-1G>A (c.1117-1G>A)	1	0.09
c.CFTRdele2.3/21kb/	67	5.84	2622+1G>A (c.2490+1G>A)	1	0.09
G551D (c.1652G>A)	34	2.96	1811+1G>C (c.1679+1G>C)	1	0.09
N1303K (c.3909C>G)	29	2.53	185+1G>A (c.53+1G>A)	1	0.09
G542X (c.1624G>T)	23	2.00	R1158X (c.3472C>T)	1	0.09
3849+10kbC>T (c.3717+12191C>T)	20	1.74	R75X (c.223C>T)	1	0.09
1898+1G>A (c.1766+1G>A)	15	1.31	2184delA (c.2052delA)	1	0.09
2143delT (c.2012delT)	11	0.96	3238delA (c.3106delA)	1	0.09
R347P (c.1040G>C)	9	0.78	296+1G>A (c.164+1G>A)	1	0.09
3272-26A>G (c.3140-26A>G)	8	0.70	2837delG (c.2705delG)	1	0.09
W1282X (c.3846G>A)	7	0.61	3840delT (c.3708delT)	1	0.09
R117H (c.350G>A)	6	0.52	2919delG (c.2787delG)	1	0.09
2789+5G>A(c.2657+5G>A)	6	0.52	2174delA (c.2042delA)	1	0.09
I336K (c.1007T>A)	6	0.52	2183delAA>G (c.2051_2052delAA)	1	0.09
R553X (c.1657C>T)	6	0.52	I1366N (c.4097T>A)	1	0.09
574delA (c.442delA)	6	0.52	L1324P (c.3971T>C)	1	0.09
S945L (c.2834C>T)	5	0.44	L1335F (c.4003C>T)	1	0.09
R1162X (c.3484C>T)	5	0.44	L1335P (c.4004T>C)	1	0.09

Mutace	**N**	**%**	**Mutace**	**N**	**%**
3141del9(c.3009_3017delAGCTATAGC)	4	0.35	1898+1G>C (c.1766+1G>C)	1	0.09
Y122X (c.366T>A)	4	0.35	L558S (c.1673T>C)	1	0.09
4374+1G>T (c.4242+1G>T)	4	0.35	M1101R (c.3302T>G)	1	0.09
621+1G>T (c.489+1G>T)	4	0.35	M952I (c.2856G>C)	1	0.09
2184insA (c.2052_2053insA)	4	0.35	R1066C (c.3196C>T)	1	0.09
1717-1G>A (c.1585-1G>A)	4	0.35	R170H (c.509G>A)	1	0.09
R334W (c.1000C>T)	3	0.26	R347H (c.1040G>A)	1	0.09
3600+2insT (c.3468+2_3468+3insT)	3	0.26	S1118F (c.3353C>T)	1	0.09
R709Q (c.2126G>A)	3	0.26	S42F (c.125C>T)	1	0.09
Delece exonu 13-14a	3	0.26	E585X (c.1753G>T)	1	0.09
G85E (c.254G>A)	2	0.17	Q1476X (c.4426C>T)	1	0.09
3659delC (c.3528delC)	2	0.17	R851X (c.2551C>T)	1	0.09
E92X (c.274G>T)	2	0.17	Q372X (c.1114C>T)	1	0.09
W57G (c.169T>G)	2	0.17	1341A+1G>A (c.1209+1G>A)	1	0.09
D1152H (c.3454G>C)	2	0.17	Y301C (c.902A>G)	1	0.09
Delece exonů 1-10	2	0.17	R1066G (c.3196C>G)	1	0.09
Delece exonu 2	2	0.17	R117C (c.349C>T)	1	0.09
Delece exonů 22-23	2	0.17	R764X (c.2290C>T)	1	0.09
Duplikace exonů 19	2	0.17	Y563H (c.1687T>C)	1	0.09
711+3A>G (c.579+3A>G)	2	0.17	E92K (c.274G>A)	1	0.09
L997F (c.2991G>C)	2	0.17	I807M (c.2421A>G)	1	0.09
G1047R (c.3139G>C)	1	0.09	Q1313X (c.3937C>T)	1	0.09
G178E (c.533G>A)	1	0.09	4259del5 (c.4127_4131delTGGAT)	1	0.09
G27R (c.79G>A)	1	0.09	neznámá mutace	15	1.31

N= počet alel, % = procentuální zastoupení jednotlivých mutací, celkem bylo vyšetřeno 1092 alel

1.2. Novorozenecký screening

1.2.1. Historie a vývoj novorozeneckého screeningu v ČR a zahraničí

Novorozeneckým screeningem se rozumí aktivní celoplošné vyhledávání chorob v jejich preklinickém stádiu pomocí stanovení koncentrace specifické látky (event. průkazem mutace) v suché kapce krve. Ta je odebírána všem novorozencům z patičky na filtrační papír. V současnosti se základní odběr provádí mezi 48. až 72. hodinou života.

Za zakladatele NS je považován profesor Robert Guthrie, který v roce 1963 vynalezl a následně i zavedl do celoplošného provádění jednoduchou, levnou a spolehlivou semikvantitativní metodu měření koncentrace fenylalaninu (založenou na principu inhibice růstu bakterií) pro NS fenylketonurie (63,64). V následujících letech se začal tento NS celosvětově rychle rozšiřovat. NS fenylketonurie se stal historicky prvním NS i v ČR (65,66,67). Pravidelně, celoplošně je prováděn již od roku 1975. Hlavním smyslem screeningu PKU je včasný zahájením léčby (dieta s monitorací hladin fenylalaninu) zabránit ireverzibilnímu poškození CNS.

Od roku 1985 se začal provádět NS kongenitální hypotyreózy (68,69), který je založen na měření koncentrace tyroxinu a/nebo tyreotropinu v suché kapce krve pomocí radio-, fluoro- či enzymo- imunoeseje. Jako třetí je v ČR prováděn od roku 2006 NS kongenitální adrenální hyperplazie - ve světě je datován od konce 70. let minulého století (18). Tento NS je založen na měření koncentrace 17-hydroxyprogesteronu v suché kapce krve (fluoro- či enzymo- imunoesejí).

Díky vědeckému a technologickému pokroku se počet chorob, které je možno diagnostikovat ze suché kapky krve na filtračním papírku stále zvyšuje. Významným momentem bylo zavedení metodiky tandemové hmotnostní spektrometrie (MS/MS), která provádí komplexní rozbor celého spektra analytů rozdělených dle hmotnostního čísla (71,72). Tato metodika se využívá k paralernímu vyšetření dědičných metabolických poruch (DMP). Hlavním smyslem je zachytit takové choroby, u kterých lze terapeutickými postupy předejít život ohrožujícímu energetickému či metabolickému rozvratu a/nebo dlouhodobému poškození zdraví (především CNS). Vyšetřuje se celkem 9 metabolických vad: leucinóza, glutarová acidurie typu I, izovalerová acidurie, deficit acyl-CoA dehydrogenázy mastných kyselin se středně dlouhým, dlouhým a velmi dlouhým řetězcem,

deficit karnitinpalmitoyltransferázy I, II a deficit karnitinacylkarnitintranslokázy.

Současný novorozenecký screening v České republice

Dne 1. října 2009 byl v ČR zahájen rozšířený NS. Je vyšetřováno celkem 13 vrozených onemocnění. Celkové riziko, onemocnění novorozencc některou ze 13 vyšetřovaných chorob činí 1:1200; některou z endokrinních chorob 1:2900; některou z dědičných metabolických poruch 1:4000.

Novorozenecký screening v zahraničí

Z pohledu Evropy je pro systémy NS v jednotlivých státech charakteristická značná roztříštěnost a nejednotnost ve všech částech logistiky systémů (72-75). Příčinou je, že se systémy NS vyvíjely na základě konkrétních historických, geografických, medicínských i socio-ekonomických podmínek daného státu / regionu, přičemž tyto vlivy i nadále rozhodují o fungování celého systému. Neexistuje jeden univerzální, nejlepší systém NS. Evropskou nejednotnost je možné dokumentovat odlišným spektrem screenovaných chorob v jednotlivých státech nebo různým počtem screeningových laboratoří. V posledních letech je pozorován trend rozšiřování spektra testovaných nemocí a zavádění nových forem novorozeneckého screeningu. V USA se počet pravidelně celoplošně vyšetřovaných onemocnění v jednotlivých státech pohybuje od 13 v Pensylvánii po 53 v Jižní Karolíně, ve většině států je mezi 30-45.

V současné době se začínají uplatňovat nové technologie typu multiplexních „microarrays“ pro rozbor analytu a/nebo DNA. Velikost a jednoduchost použití těchto nástrojů naznačuje další možný směr vývoje NS – diagnostiku přímo u lůžka novorozence (76-80).

1.2.2. Principy a etika novorozeneckého screeningu

Zahájení celoplošného screeningu si vyžádalo formulování obecných pravidel pro jeho provádění, což učinili na zadání Světové zdravotnické organizace Wilson a Junger již v roce 1968. Předpoklady pro zahájení celoplošného screeningu lze shrnout do následujících bodů (81):

1. *Průběh nemoci a její mechanismy musí být známé.*
2. *Choroba představuje významný zdravotně sociální problém.*

3. *Choroba je častá, má v dané populaci určitou incidenci.* (Frekvence nemoci je relativní a je ovlivněna především klinickou závažností choroby, možnostmi léčby. Hranici pro NS lze nyní spatřovat v incidenci mezi 1:50 000 až 1:200 000).
4. *Onemocnění má latentní fázi. Zachycení choroby v jejím časném, presymptomatickém stadiu umožňuje taková terapeutická opatření, která zásadním způsobem pozitivně ovlivní průběh choroby či dokonce sníží její mortalitu. Terapeutická opatření musí být dostupná a zajistitelná pro všechny zachycené jedince***.** (Péče o pacienty je většinou soustředěna do specializovaných center).
5. *Existuje spolehlivý screeningový test.* tj. choroba je v preklinickém stadiu detekovatelná obecně uznaným laboratorním testem v suché kapce krve s dostatečnou senzitivitou a specifitou. Hlavními měřítky věrohodnosti screenigového testu jsou:
    - obecně přijatá hranice negativity („cut off“ limit)
    - vybalancování hodnoty falešné negativity na jedné straně a hodnoty falešné pozitivity tedy zátěže zdravé části novorozenecké populace na straně druhé (tyto hodnoty jsou vyjádřeny jako negativní prediktivní hodnota (NPV) resp. pozitivní prediktivní hodnota (PPV).
6. *Společnost je schopna zajistit provádění laboratorního testu u všech svých novorozenců po stránce organizační a ekonomické.*
7. *Přínos NS má být pozitivní vůči jeho nákladům* (pozitivní poměr „benefit/costs“).
8. *Zavedený NS je kontinuální proces, jeho věrohodnost a účinnost* – tj. efektivita musí být trvale sledována a vyhodnocována.

Na tato klasická kritéria navazují z pohledu dnešní společnosti nové názory. Ve vyspělých zemích se věnuje velké úsilí prosazení práva pacienta postiženého vzácnou chorobou být diagnostikován a léčen způsobem odpovídajícím současným možnostem medicíny.
Z pohledu rodiny je tlak na právo znát diagnózu co nejdříve. Je však otázkou názoru, zda je indikován NS neléčitelné choroby pouze z důvodu znalosti diagnózy a informované

reprodukce. Tyto fakta vedou k etické diskuzi na fóru odborných společností o event. revizi základních předpokladů NS. Rozhodování o zavedení některých nových NS je komplikované a musí být učiněno *až po proběhlé pilotní studii v podmínkách konkrétního státu či regionu.*

S narůstající praktickou zkušeností s NS nabývá na stále větším významu hranice akceptovatelné zátěže zdravé části populace falešnou pozitivitou. Jedná se o řešení etického dilematu, tj. do jaké míry jsme oprávněni zatížit zdravé novorozence a jejich rodiny na úkor záchytu skutečně nemocného jedince. Cílem je pochopení situace a přístupu k problému, k tomu aby se minimalizovala stigmatizace rodiny. Velmi nápomocná je předcházející adekvátní informovanost rodičů v rámci získávání informovaného souhlasu s NS (82).

1.2.3. Novorozenecký screening cystické fibrózy

1.2.3.1. Hlavní argumenty pro provádění NSCF

Argumenty lze shrnout do následujících bodů:

- Průběh nemoci a její mechanismy jsou známé
- Choroba je významným celospolečenským socio-ekonomickým problémem (z důvodu finančně náročné léčby).
- Je relativně častá.
- Včasná diagnostika zkvalitňuje a lze předpokládat, že i prodlužuje život pacientů; tj. zlepšení stavu nutrice, zlepšení růstu, méně agresivní terapie; *navíc nedávné studie dokazují, že pacienti s CF diagnostikovaní do 2 měsíců mají dlouhodobě lepší klinický stav a příznivější průběh choroby s vyhlídkou delšího přežití (83, 84).* Klinicky je však v tomto věku obtížné CF diagnostikovat (2).
- Včasná diagnostika je prevencí úmrtí z nediagnostikované choroby – např. v metabolickém solném rozvratu kojenců, či rychlou progresí plicního postižení.
- Předpokládané náklady NS jsou nižší než náklady léčby komplikací choroby z pozdní diagnostiky (83).

- Včasná diagnostika redukuje stres rodiny s chronicky nemocným dítětem s nejasnou příčinou obtíží při pozdní diagnostice, a tím i obecně zvyšuje důvěru ve zdravotnický systém.
- Včasná diagnostika pomáhá optimalizovat genetické poradenství a informovanou reprodukci v rodinách s prvorozeným dítětem s CF. *Při absenci NS CF není ojedinělá situace, kdy je CF diagnostikována u staršího sourozence až na základě průkazu choroby u mladšího sourozence.*
- Při celoplošném provádění NS CF lze *efektivněji soustředit pacienty do klinických center, které mají s nemocí zkušenosti a kde jim je zajištěna komplexní péče*. Tím lze odstranit nerovnoměrnost v péči mezi pacienty.

1.2.3.2 Protokoly novorozeneckého screeningu CF

Snahy o zavedení NSCF jsou známy od 50. let minulého století. Byla vyzkoušena celá řada metod od otiskových testů prokazujících vysoký obsah solí v potu po průkaz albuminu v mekoniu různými metodami, které však nebyly spolehlivé. V roce 1979 (85) se prokázalo, že v krvi novorozenců s CF je zvýšená koncentrace imunoreaktivního trypsinogenu/trypsinu (IRT). Tento analyt se dal stanovit i ze zaschlé kapky krve na filtračním papíru, který se využíval pro jiné NS. Trypsin je v rovnovážném stavu se svým prekurzorem – trypsynogenem a většina metodik měří koncentrace obou látek. Proč je IRT v krvi novorozenců s CF zvýšený? Je známo, že pankreas pacientů s CF je již prenatálně poškozen blokádou acinárních duktů vazkým hlenem. Ty produkují trypsinogen a blokáda způsobí, že prostoupí do krve ve vyšší koncentraci (86). S postupujícím věkem a atrofizací pankreatu produkce enzymu postupně klesá (87-90), a tak se pro pozdější diagnostiku nehodí. IRT je dostatečně stabilním a specifickým markerem, ale má horší specifitu. Vysokou koncentraci IRT má i část zdravé populace – zejména nosiči mutace genu *CFTR*, vliv má etnický původ (91, 92). Dále se uvádí, že může být zvýšen následkem perinatální hypoxie, u chromozomálních aberací (Downův syndrom), některých vrozených vývojových vad ledvin, srdce a neurální trubice (87-90). V této souvislosti se uvádí, že cca. 25% novorozenců s CF a mekoniovým ileem má normální (nižší než „cut off“) hodnotu IRT.

Zatím jediným možným východiskem jak zvýšit specifitu NSCF je zavádění dvoustupňových event. vícestupňových protokolů. V prvním stupni jsou vyselektováni

rizikoví novorozenci se zvýšenou koncentrací IRT a teprve až ve druhém stupni jsou vyhledávni pacienti s CF pomocí dalšího testování.

Použité protokoly se dají rozdělit do několika skupin: IRT/IRT, IRT/DNA, IRT/DNA/IRT a IRT/PAP.

Protokol **IRT/IRT** spočívá v novém odběru opakování měření koncentrace IRT s odstupem 2-8 týdnů u vyselektovaných dětí. Takto se sníží množství falešně pozitivních (FP) výsledků na 0,03%. Nevýhodou je neúměrně vysoký počet „recallů“ (tj. opakovaných odběrů podle nastaveného „cut off“ asi 0,5-2%) a tak vystavení rodiny nadbytečnému stresu. Tento fakt znehodnocuje tento model NSCF, dá se však použít v případech, kdy legislativa neumožňuje DNA analýzu u novorozenců, nebo je v dané populaci nízká záchytnost mutací genu *CFTR* díky značné populační heterogenitě.(například díky přistěhovalectví). U dětí s hypertrypsinogenémií i ve druhém stupni se následně provádí potní test. PPV hodnota se po druhém odběru IRT zvyšuje až na 78% (93).

Protokol **IRT/DNA** – u novorozenců se jako druhý stupeň provádí molekulárně genetické vyšetření mutace/mutací genu *CFTR*. Pokud se prokáže mutace na obou alelách je pacient s rodiči pozván k edukačnímu a diagnosticko-terapeutickému pobytu v CF Centru. V případě nálezu jedné mutace, následuje potní test. Senzitivita tohoto postupu závisí na počtu testovaných mutací a populačním spektru (94).

Protokol **IRT/DNA/IRT** je alternativou k modelu IRT/DNA. U novorozenců s iniciální extrémně zvýšenou hodnotou IRT, i když nebyla prokázána v dalším stupni žádná mutace, se provádí „recall“. V případě, že IRT i v tomto měření přesáhne „cut off“ hodnotu, pozve se k potnímu testu. Problémem je vyšší počet recallů a nastavení hranice „cut off“ IRT (95).

Protokol **IRT/PAP je** zatím novou perspektivou a alternativním modelem k protokolu s DNA analýzou. U novorozenců s vysokým IRT v prvním kroku se stanovuje ze stejných vzorků koncentrace proteinu PAP (pancreatitis associated protein). PAP je sekreční protein, který je indukovaný stresem pankreatu, je zvýšený např. u akutní pankreatitidy. Moduluje průběh zánětu v pankreatu, stimuluje aktivitu makrofágů, uvádí se, že má protektivní účinek na tíži pankreatitidy. (96,97). Z dosud dostupných informací je dobře senzitivní a méně specifický, odobně jako IRT. Pokud ovšem je použit ve druhém stupni NS, specifita

se zvyšuje. V případě, že je koncentrace PAP zvýšená, je proband pozván k potnímu testu. Výhodou by mělo být udržení dobré senzitivity i specifity, nižší ekonomické náklady a odpadne problém s legislativou DNA testování (98,99). Nevýhodou je proti IRT/DNA vyšší falešná pozitivita. Co se týká falešné negativity, nejsou zatím dostatečné důkazy a závěry (98,99).(Pozn. od října roku 2009 probíhá v českých regionech pilotní studie, testující efektivitu tohoto protokolu v ČR v rámci projektu IGA MZ ČR).

Je třeba poznamenat, že žádný z novorozeneckých screeningů nemá 100% senzitivitu a specifitu a nulovou falešnou negativitu, takže diagnózu CF nelze na základě negativního screeningu jednoznačně zamítnout. To však platí obecně pro všechny typy screeningů.

2. PŘILOŽENÉ PUBLIKACE

Journal of Cystic Fibrosis 8 (2009) 224–227

Journal of Cystic Fibrosis

www.elsevier.com/locate/jcf

Case report

Pilot newborn screening project for cystic fibrosis in the Czech Republic: Defining role of the delay in its symptomatic diagnosis and influence of ultrasound-based prenatal diagnosis on the incidence of the disease

Miroslava Balaščaková [a], Andrea Holubová [a], Veronika Skalická [b], Dana Zemková [b], Petr Kračmar [d], Lucie Gonsorčíkova [b], Jana Čamajová [a], Tereza Piskáčková [a], Jan Lebl [b], Pavel Dřevínek [b], Vladimír Gregor [c], Věra Vávrová [b], Felix Votava [d,1], Milan Macek Jr. [a,*,1]

[a] *Department of Biology and Medical Genetics, Cystic Fibrosis Centre 2nd Medical School and Faculty Hospital Motol, Charles University, Prague, Czech Republic*
[b] *Department of Paediatrics, Cystic Fibrosis Centre 2nd Medical School and Faculty Hospital Motol, Charles University, Prague, Czech Republic*
[c] *Department of Medical Genetics, Faculty Thomayer Hospital, Charles University, Prague, Czech Republic*
[d] *Department of Paediatrics, 3rd Medical School and Faculty Hospital Královské Vinohrady, Charles University, Prague, Czech Republic*

Received 14 October 2008; received in revised form 12 January 2009; accepted 16 January 2009
Available online 8 February 2009

Abstract

The objective need for cystic fibrosis (CF) newborn screening (NBS) in the Czech Republic has recently been substantiated by a significant delay of its symptomatic diagnosis. This trend most likely resulted from the process of decentralisation of health care which led to the deterioration of care for patients who need specialised approaches. Applied newborn screening model (IRT/DNA/IRT) was efficacious enough to detect CF cases with median age at diagnosis of 37 days. The incidence of CF (1 in 6946 live births) ascertained in this project was lower than that established previously by epidemiological studies (1 in 2700–1 in 3300). However, adjustment for broadly applied ultrasound-based prenatal diagnosis (PND) in the 2nd trimester of pregnancy, that was performed within the period of the project (1/2/2005–2/11/2006), rendered an incidence estimate of 1 in 4023. This value is closer to that observed in other CF NBS programmes and reflects influence of PND on the incidence of CF.

Keywords: CF carriers; *CFTR* gene; Cystic fibrosis (OMIM 219700); Decentralisation of health care; DNA testing; Incidence; Immunoreactive trypsinogen; Newborn screening; Prenatal diagnosis

1. Delayed symptomatic diagnosis of CF

Early diagnosis of cystic fibrosis (CF; MIM 219700), i.e. during first 2 months of life, is considered a favourable prognostic factor for its overall course and significantly decreases parental anxiety due to delayed diagnosis of the disease [1,2]. Unfortunately, a significant delay in symptomatic diagnosis of CF in the period 1999–2004 has been observed in the Czech Republic (CZ). This evidence is based on registry data drawn from the Prague CF Centre. We compared data from the "1st period" (1993 to 1998; $n=90$), when paediatric care for children with specific needs (e.g. congenital heart disease, metabolic diseases or CF) still had been centrally organised

Abbreviations: ADG, age at diagnosis; CF, cystic fibrosis; *CFTR*, gene for the cystic fibrosis transmembrane conductance regulator protein; CZ, Czech Republic; IE, incidence estimate; IRT, immunoreactive trypsinogen; NBS, newborn screening; NBSR, newborn screening referral region; PND, prenatal diagnosis.

* Corresponding author. Department of Biology and Medical Genetics, 2nd Medical School and Faculty Hospital Motol, V Uvalu 84; Prague 5, CZ 15606, Czech Republic. Tel.: +420 224433501; fax: +420 224433520.
E-mail address: Milan.Macek.Jr@LFmotol.cuni.cz (M. Macek).

[1] Both senior authors contributed equally to the oversight of this project.

1569-1993/$ - see front matter © 2009 European Cystic Fibrosis Society. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.jcf.2009.01.002

[3,4], against the subsequent "2nd period" (1999 to 2004; $n=89$) when effects of health care decentralisation, a process that has been gradually implemented from mid-nineties [5], became apparent. Within the "1st period" the median age at diagnosis (ADG) was 0.45 years and 59/90 (65.6%) of patients were diagnosed by first year of age. Interestingly, ADG in our country during that period was similar to that reported in the developed world [2] and was not significantly different from the median ADG of 0.67 years in the period 1970–1992. However, during the "2nd period" the median of ADG increased significantly to 1.2 years in a cohort of 89 patients (chi-square; $p=0.0001$). Moreover, the percentage of patients diagnosed before 12 months of age decreased to 43/89 (48.3%). This negative trend has been accompanied by a simultaneous increase in the proportion of CF patients with their ADG being between 1 to 6 years. In the "1st period" this category accounted for 21/90 (23.3%), while in the "2nd period" it increased to 32/89 (36.0%) of all newly diagnosed cases. Observed higher ADG could not be attributed only to improved symptomatic diagnosis of "mildly presenting" pancreatic sufficient forms of CF between the ages of 1 and 6 years within the "2nd period". In the "1st period" 3/80 cases diagnosed before 6 years of age were pancreatic sufficient (PS), while in the "2nd period" this proportion was 6/75. Similarly, after 6 years of age 10 patients were diagnosed in "1st period", 5 of them were PS, while in the "2nd period" 14 patients were diagnosed and 6 of them were PS (all differences were not statistically significant).

Furthermore, 46 patients clinically diagnosed after their first year of life had at the time of diagnosis median body weight −0.8 (in SD-score; range −2.5; 1.2) and 11/46 of these patients had their weight below the 10th percentile using population specific standards [6]. Pulmonary effects of delayed diagnosis were assessed predominantly radiologically and clinically, since due to lower age of studied subjects lung function was examined only in 17 cases older than 5 years. The median FEV_1 in these patients was 74.0 (% predicted; range 32–112), while in 7/17 cases severe peripheral airways obstruction was detected (data not shown). One female with CF, who was diagnosed at the age of 5.9 years, died 1.6 years later despite intensive care.

We presume that during the time of centralised system of health care, presence of regional health authorities, annual provision of diagnostic/therapeutic guidelines to primary care paediatricians, including their regular training, had a positive effect in that they were knowledgeable of early CF clinical features [3,7]. However, regional devolution of health care together with strengthening of independent primary care [5], had an opposite effect leading to deterioration of care for patients who require specialised, expensive and thus centralised care [8], as we have observed in CF.

2. Newborn screening pilot project

This fact led us to initiate a pilot CF newborn screening (NBS) project (1/2/2005–1/11/2006) in order to assess the feasibility of achieving an unified and early diagnosis of CF in CZ [1,2,9]. The project was designed as an "IRT/DNA/IRT" scheme [9]. IRT was examined using Delphia Neonatal IRT kits (Wallac Oy^{Tm}, Turku, Finland) on the Delfia^{Tm}1232 fluorimeter. The biochemical laboratory successfully participates within the French external quality assessment (EQA) scheme for IRT (data not shown). An "opt-out concept" for individual participation [10] has been adopted by 92 collaborating district neonatal wards covering the Western part of CZ and accounting for approximately 62% of all live-births nationwide. However, due to irregular funding from the grant agency the course of the NBS pilot scheme was discontinuous and comprised in total 16 months. We screened 76,438 from the total of 119,028 (64.1%) newborns born within the CF NBS referral region (NBSR). For further analyses we only used data from within the period during which CF NBS was carried out. Concentrations of immunoreactive trypsinogen (IRT), with an arbitrarily fixed cut-off value of ≥ 75 ng/ml (IRT 99th percentile), were examined in dried blood spots (Guthrie cards) sampled within the 72nd–96th hour after birth. 799 (1.05% of all newborns) had IRT ≥ 75 ng/ml and were subjected to DNA testing. When the IRT was within the range of 75–150 ng/ml respective Guthrie cards were tested for the 5 most common CF-causing mutations occurring in Czech patients with the "classical form" of CF [p.F508del, c. CFTRdele2,3(21kb), p.N1303K, p.G551D, p.R553X] and comprising approximately 84% of all population specific *CFTR* alleles [11]. When IRT concentration was over 150 ng/ml or one of the initially tested mutations was detected in blood spots with IRT within the range 75–150 ng/ml, the "initial" panel was extended to 38 CF-causing mutations by the "CF38 Research Prototype Assay" (Roche Molecular Systems, Alameda, USA) accounting for 90.8% of population specific alleles (data not shown). The second sampling of a blood spot ("recall") was requested at the age 3–10 weeks when the initial IRT level was ≥ 200 ng/ml and no CF-causing mutations were detected. Newborns/infants were referred to a sweat test, performed by standard "Gibson-Cook" pilocarpine iontophoresis at our CF Centre facility, if the recall IRT value was not lower than the recall cut off level or in cases where we detected one or two CF-causing alleles in *trans* (i.e. in both homologous *CFTR* genes). A standard informed consent procedure was performed both prior to sweat testing (due to the pilot nature of this project) and as required prior to DNA testing [10]. Furthermore, all families with a newly diagnosed CF patient or with a detected CF carrier underwent genetic counselling and were offered "cascade" *CFTR* mutation screening to their direct relatives. We also offered sweat testing to the apparently healthy siblings of patients in order to detect potentially undiagnosed cases of CF.

Two *CFTR* mutations in *trans* were found in 11 cases, while only in 10 instances CF-causing alleles substantiated the diagnosis of CF. The diagnosis of CF could not be unambiguously established in one patient bearing the complex allele p.F508del/p.R117H-IVS8 T(7), who had mean sweat chloride concentration of 20.4 mM/L [12]. This patient was enlisted into the long-term monitoring at our CF centre and was not used for the calculation of observed incidence of the classical form of the disease. In 4/11 cases the diagnosis of CF was established clinically prior to completion of the NBS scheme: 1 case had meconium ileus (IRT was higher than the cut off value), 2 cases had bronchopneumonia, while one case

presented with failure to thrive, bronchopneumonia and hypochromic anaemia and all were homozygous for the p. F508del mutation.

A single CF allele was found in a total of 52 newborns, but we have not found couples where both parents were carriers by subsequent "cascade" DNA testing in their families. Sweat testing was finalised in all but 6 infants where their parents declined further examination. Sweat chloride concentration was positive (75 mM/L) in one newborn with the p.F508del mutation, where subsequent direct sequencing detected the p. I336K mutation [13], absent in the extended panel. Remaining 45 infants had a negative sweat test (<30 mM/L). In total, 11 CF infants (11/799; 1.38%) were diagnosed by NBS and the median ADG was 37 days (range 26–54 days). Interestingly, we also diagnosed previously unrecognized CF in 3 older siblings who had clinical signs of the disease (data not shown).

The frequency of unaffected CF carriers in the population of newborns with elevated IRT was 1 in 15.4 (6.5%), which is 1.66 times higher than the number of CF carriers in the random population—1 in 25.5 (3.9%) [14]. Diagnosed CF carriers have created an extra burden on genetic counselling and reflect ascertainment bias related to the testing of IRT "positives" [9,15]. Nevertheless, the majority of expected CF carriers (an estimate of 2990 cases), who were born during the time of the project, remained undetected.

In summary, we conclude that in CZ there is an objective need for CF NBS and that the applied algorithm is an efficacious tool for early, uniform and broad-based diagnosis of CF in our conditions.

3. Influence of ultrasound-based prenatal diagnosis on the incidence of CF

If NBS data were considered alone, the extrapolated incidence estimate (IE) of CF (1 in 6949 live births) would be substantially lower than the previously reported epidemiological IEs that range from 1 in 2700 to 1 in 3300 live births. These higher IEs resulted from long-term studies (performed during 1953–1963 and 1960–1967) of CF cases diagnosed either symptomatically (minority of patients) or by compulsory autopsy of all deceased newborns / infants from regions encompassed within the current NBSR [3,7]. However, we are aware of all limitations of these historic studies and present them only for general orientation.

Since the proportion of non-European immigrant births is not reported due to antidiscrimination laws, but its estimates are thus far marginal [16], the CF NBS IE of 1 in 6949 live births could be explained either by an ascertainment bias due to a smaller number of examined newborns or by false negativity of the applied algorithm. In order to rule out the latter bias we searched for "missed cases" of CF by contacting regional CF Centres, primary care physician associations or by examining the centralised "Birth Defects" registry. Thus, we are confident that no case of clinical diagnosed CF was reported in NBSR from within the period of our CF NBS pilot project.

Although population-based CF carrier screening [17] is non-existent in CZ, we were able to account for the influence of broadly applied ultrasound-based prenatal diagnosis (PND) screening for prenatal abnormalities, which is carried out in the 2nd trimester of pregnancy, on the IE of CF. In this "cascade" scheme mothers with foetuses with abdominal hyperechogenity are referred to DNA testing and in "positive" instances the father is examined as well in order to avoid unnecessary amniocentesis [18]. National "Birth Defects" registry provided evidence of eight additional new cases of "prenatal CF" from within NBSR and corresponding to the project timeline. All detected pregnancies were terminated according to the informed decision of parents [10] and in accordance with local legal provisions, but respective *CFTR* genotypes were not available. Thus, the adjusted CF NBS IE of 1 in 4023 live births is closer to IEs which were reported recently from other CF NBS programmes [15,17]. Our adjusted IE also reflects the effect of broad based provision of PND on the "final" IE of the disease [19]. Based on the above mentioned facts and accounting for the discontinuous character of our study, whereby seasonal birth distribution in CF is still a controversial subject [20], we believe that the only remaining bias could be due to the overall lower number of newborns screened. We hope to clarify this issue in future since the results of this project serve as a basis for the implementation of a nationwide CF NBS scheme.

Acknowledgement

Supported by IGA Czech Ministry of Health 9986-3, MSM0021620814 and VZFNM00064203.

References

[1] Sims EJ, Clark A, McCormick J, Mehta G, Connett G, Mehta A. United Kingdom Cystic Fibrosis Database Steering Committee. Cystic fibrosis diagnosed after 2 months of age leads to worse outcomes and requires more therapy. Pediatrics 2007;119:19–28.
[2] McKay K, Wilcken B. Newborn screening for cystic fibrosis offers an advantage over symptomatic diagnosis for the long term benefit of patients: the motion for. Paediatr Respir Rev 2008;9:290–4.
[3] Houštěk J, Vávrová V. Our experience with mucoviscidosis. Rev Méd Liege 1967;22:421–6.
[4] Šamánek M, Voříšková M. Infants with critical heart disease in a territory with centralized care. Int J Cardiol 1987;16:75–91.
[5] Earl-Slater A. Health-care reforms in the Czech Republic. J Manag Med 1996;10:13–22.
[6] Lhotská L, Bláha P, Vignerová J, Roth Z, Prokopec M. Vth nation-wide anthropological survey 1991 in the Czech Republic comparison with 1981 data. Cent Eur J Public Health 1994;2:95–9.
[7] Brunecký Z. The incidence and genetics of cystic fibrosis. J Med Genet 1972;9:33–7.
[8] Křížová E, Šimek J. Rationing of expensive medical care in a transition country: nihil novum? J Med Ethics 2002;28:308–12.
[9] Rock MJ, Hoffman G, Laessig RH, Kopish GJ, Litsheim TJ, Farrell PM. Newborn screening for cystic fibrosis in Wisconsin: nine-year experience with routine trypsinogen/DNA testing. J Pediatr 2005;147:S73–7.
[10] Křížová E, Šimek J. Theory and practice of informed consent in the Czech Republic. J Med Ethics 2007;33:273–7.
[11] Bobadilla JL, Macek Jr M, Fine JP, Farrell PM. Cystic fibrosis: a worldwide analysis of CFTR mutations: correlation with incidence data and application to screening. Hum Mutat 2002;19:575–606.
[12] Scotet V, Audrézet MP, Roussey M, Rault G, Dirou-Prigent A, Journel H, et al. Immunoreactive trypsin/DNA newborn screening for cystic fibrosis: should the R117H variant be included in CFTR mutation panels? Pediatrics 2006;118:e1523–9.

[13] Férec C, Verlingue C, Guillermit H, Quéré I, Raguénès O, Feigelson J, et al. Genotype analysis of adult cystic fibrosis patients. Hum Mol Genet 1993;2:1557–60.
[14] Macek Jr M, Macek Sr M, Krebsová A, Nash E, Hamosh A, Reis A, et al. Possible association of the allele status of the CS.7/HhaI polymorphism 5' of the *CFTR* gene with postnatal female survival. Hum Genet 1997;99: 565–72.
[15] Southern KW, Munck A, Pollitt R, Travert G, Zanolla L, Dankert-Roelse J, et al. On behalf of the ECFS CF Neonatal Screening Working Group. A survey of newborn screening for cystic fibrosis in Europe. J Cyst Fibros 2007;6:57–65.
[16] Fialová L, Pavlík Z, Vereš P. Fertility decline in Czechoslovakia during the last two centuries. Popul Stud 1990;44:89–106.
[17] Hale JE, Parad RB, Comeau AM. Newborn screening showing decreasing incidence of cystic fibrosis. N Engl J Med 2008;358:973–4.
[18] Muller F, Simon-Bouy B, Girodon E, Monnier N, Malinge MC, Serre JL. French Collaborative Group. Predicting the risk of cystic fibrosis with abnormal ultrasound signs of fetal bowel: results of a French molecular collaborative study based on 641 prospective cases. Am J Med Genet 2002;110:109–15.
[19] Scotet V, Audrézet MP, Roussey M, Rault G, Blayau M, De Braekeleer M, et al. Impact of public health strategies on the birth prevalence of cystic fibrosis in Brittany, France. Hum Genet 2003;113:280–5.
[20] Daigneault J, Aubin G, Simard F, De Braekeleer M. Birth distribution in cystic fibrosis in Saguenay-Lac-St-Jean, Quebec, Canada. J Med Genet 1991;9:613–4.

PŮVODNÍ PRÁCE

Čes.-slov. Pediat., 2007, roč. 62, č. 4, s. 187–195.

Novorozenecký screening cystické fibrózy v České republice: závěry pilotní studie

Holubová A.[1]*, Balaščáková M.[1]*, Skalická V.[2], Kračmar P.[3], Piskáčková T.[1], Vávrová V.[2], Zemková D.[2], Gonsorčíková L.[3], Lebl J.[2], Macek M.ml.[1], Votava F.[3]

Ústav biologie a lékařské genetiky UK 2. LF a FN Motol, Praha[1]
přednosta prof. MUDr. M. Macek ml., DrSc.
Pediatrická klinika UK 2. LF a FN Motol, Praha[2]
přednosta prof. MUDr. J. Lebl, CSc.
Klinika dětí a dorostu UK 3. LF a FN Královské Vinohrady, Praha[3]
přednosta doc. MUDr. F. Votava, PhD.

Věnováno k životnímu jubileu paní prof. MUDr. Olgy Hníkové, CSc., která novorozeneckému screeningu zasvětila velkou část svého profesního života.

Souhrn

Včasná diagnóza cystické fibrózy (CF) je jednoznačně považována za prognosticky příznivý faktor i za průkaz dobře fungující zdravotní péče. Zjištění o zhoršující se klinické diagnostice CF v posledním desetiletí nás vedla k tomu, že jsme zahájili pilotní studii novorozeneckého screeningu (NS) CF jako potenciálního nástroje zlepšení diagnostiky. V období mezi 1. 2. 2005 až 2. 11. 2006 byl změřen imunoreaktivní trypsin (IRT) v suchých kapkách krve (na tzv. „Guthrieho" screeningových kartičkách) u 76 438 novorozenců z oblasti pokrývající 62 % populace v ČR. U 799 novorozenců (1,05 %) byla zjištěna koncetrace IRT 75 ng/ml a vyšší. Tato skupina novorozenců byla považována za vysoce rizikovou z hlediska CF, a proto u nich byla provedena molekulárně-genetická analýza genu *CFTR*. Suché kapky s koncentrací IRT 75–150 ng/ml byly analyzovány na 5 nejčastějších mutací zahrnujících 83,8 % patologických alel v ČR. Suché kapky s koncentrací IRT >150 ng/ml a kapky s jednou nalezenou mutací ze skupiny s koncentrací IRT 75–150 ng/ml byly analyzovány na 38 mutací pokrývajících 90,8 % patologických alel genu *CFTR* v ČR. Dvě mutace v obou homologních genech *CFTR* byly nalezeny u 11 novorozenců. Jedna mutace byla nalezena u 53 novorozenců. U jednoho novorozence s 1 mutací byl potní test pozitivní (Cl^- >60 mmol/l); u 45 byl potní test negativní (Cl^- <30 mmol/l), u jednoho hraniční (Cl^- 30–40 mmol/l; toto dítě je dále sledováno) a u 6 nebyl potní test proveden. Pomocí NS bylo v uvedeném období diagnostikováno 12 dětí s cystickou fibrózou a zachyceno 52 heterozygotů – přenašečů. Medián věku při stanovení diagnózy pomocí NS byl 37 dnů. U 4 dětí byla CF diagnostikována klinicky před výsledkem NS. Zjištěná incidence CF 1:6369 je překvapivě nižší, než by odpovídalo dosud popisované incidenci v ČR (1:2736). Příčinou této diskrepance může být:

1. zatím stále relativně malý počet vyšetřených novorozenců vedoucí k chybě malých čísel;

2. falešná negativita použitého modelu screeningu, což je velmi málo pravděpodobné;

3. vliv prenatální diagnostiky a rozhodnutí rodiny provést selektivní přerušení těhotenství při průkazu CF u plodu.

Je zřejmé, že novorozenecký screening CF v podmínkách ČR je proveditelný a efektivní nástroj z hlediska jednotné a včasné diagnostiky CF. Otevřenou otázkou zůstává, zda byly skutečně zachycovány všechny postižené děti, v tomto ohledu jsou nezbytné další studie.

Klíčová slova: cystická fibróza (McKusick, OMIM 219700), novorozenecký screening, imunoreaktivní trypsin/trypsinogen, gen *CFTR*

*uvedené autorky přispěly stejným dílem k realizaci projektu

Summary

Newborn Screening for Cystic Fibrosis in the Czech Republic: Results of a Pilot Study

Early diagnosis of cystic fibrosis (CF) is considered as a favorable prognostic factor and as a hallmark of a properly functioning health care system. During the last decade evidence for deterioration of early clinical diagnosis of CF in the Czech Republic became evident. This alarming fact led us to commence a CF newborn screening (NBS) pilot project in order to assess its feasibility as an efficient tool for uniform and early diagnosis of CF in our country. Concentrations of immunoreactive trypsin (IRT) in dried blood spots (Guthrie cards) were examined in 76438 newborns born between 1. 2. 2005 and 2. 11. 2006 within a region comprising 62% of the entire Czech population. IRT concentration of 75 ng/ml and higher was found in 799 newborns, i.e. 1.05% of the total. Since these newborns were considered at high risk of CF and panels of population specific mutations of the gene *CFTR* were analysed from respective blood spots. The five most common mutations of gene *CFTR* (comprising 83.8% of alleles in CF patients in CR) were investigated if the concentration of IRT was within the range of 75–150 ng/ml. In addition, 38 *CFTR* gene mutations (comprising 90.8% of CF alleles in CR) were examined in blood spots where the concentration of IRT was more than 150 ng/ml and/or if one *CFTR* gene mutation was found in a dried blood spot with a lower IRT concentration.

Two *CFTR* gene mutations were found in 11 newborns, confirming the diagnosis of CF. A single *CFTR* mutation was found in 53 newborns. All such cases were subjected to sweat testing, which was finished in all but 6 newborns. The sweat test was positive (Cl^- level >60 mmol/l) in one newborn with one *CFTR* gene mutation, in one newborn borderline levels were detected (Cl^- level 30–40 mmol/l, which indicated his long-term follow-up) and in other 45 newborns the sweat test was negative (Cl^- level <30 mmol/l) – unaffected carries.

Over the course of our project 12 infants were diagnosed as having CF. The median of age at diagnosis was 37 days. Interestingly, 4 of these cases were diagnosed clinically prior the results of NBS. If data of NBS would have been used alone the extrapolated incidence of the disease (1:6369 newborns) was surprisingly lower than that established epidemiologically (1:2736 newborns). This discrepancy could be explained by:

1. a relatively small number of examined newborns, i.e. error due to a small number of studied subjects;

2. false negativity of utilised screening algorithm, which is however unlikely;

3. impact of prenatal diagnosis where parents have opted for selected termination of pregnancy.

Authors conclude that the NBS for CF in conditions of the Czech Republic is an effective tool for early and broad based diagnosis of CF. It also complements the current deterioration of early clinical diagnosis of the disease. Nevertheless, an unanswered question remains if really all cases of incident CF were ascertained, which necessitates further studies.

Key words: cystic fibrosis (OMIM 219700), newborn screening, immunoreactive trypsin/trypsinogen, *CFTR* gene, molecular genetic testing

H.

Úvod

Cystická fibróza (CF; registrační číslo OMIM 219700) je autozomálně recesivně dědičné onemocnění postihující primárně plíce, slinivku břišní, střeva, játra, potní žlázy a reprodukční systém. Řada dalších orgánů bývá postižena sekundárně, narušen je i stav výživy, růst a vývoj nemocných. Incidence CF v české populaci byla stanovena na 1:2736 novorozenců [1, 2, 3, 4]. Klinický průběh může být různě závažný a o délce života rozhoduje především postižení plic chronickým zánětem a stav výživy. Přestože dosud není k dispozici kauzální léčba, délka a kvalita života pacientů s CF se v posledních desetiletích významně zlepšila a je úměrná celkové úrov zdravotní péče v daném státě [5]. V současnos se nemocní běžně dožívají 4. i 5. dekády živo a např. v USA, podle aktuální zprávy Americl nadace pro CF (www.cff.org), činí střední dél života pacientů 36,8 let. Rozhodujícím faktore prognózy a adekvátní péče o pacienty s CF včasnost diagnózy [6, 7]. Již v roce 1970 by popsáno [8], že včasná léčba presymptomatic diagnostikovaných pacientů významně prodluž je jejich dobu přežití a kvalitu života. Nedávn studie dokazuje, že pacienti s CF diagnostikov ní do 2 měsíců věku mají dlouhodobě lepší klini ký stav a příznivější průběh choroby [9].

Pro včasnou diagnostiku choroby se nabi využít metody novorozeneckého screeningu (N

Hlavní argumenty pro provádění NS CF lze shrnout do následujících bodů [9, 10, 11, 12, 13]:

- Choroba je významným celospolečenským problémem.
- Je relativně častá.
- Včasná diagnostika zkvalitňuje a lze předpokládat, že i prodlužuje život pacientů (zlepšení stavu nutrice, zlepšení růstu, prevence plicního postižení).
- Včasná diagnostika je prevencí úmrtí z nediagnostikované choroby – např. v metabolickém solném rozvratu kojenců.
- Předpokládané náklady NS jsou menší než náklady léčby komplikací choroby z pozdní diagnostiky.
- Včasná diagnostika redukuje stres rodiny s chronicky nemocným dítětem s nejasnou příčinou obtíží při pozdní diagnostice a tím i obecně zvyšuje důvěru ve zdravotnický systém.
- Včasná diagnostika pomáhá optimalizovat genetické poradenství a informovanou reprodukci v rodinách s prvorozeným dítětem s CF. Při absenci NS CF není ojedinělá situace, kdy je CF diagnostikována u staršího sourozence až na základě průkazu choroby u mladšího sourozence.
- Při celoplošném provádění NS CF lze efektivněji soustřeďovat pacienty do klinických center a tím odstraňovat event. nerovnoměrnosti v péči.

Z hlediska včasnosti diagnostiky CF v České republice jsme bohužel v posledních letech svědky velmi nepříznivého vývoje [14]. Do roku 1998 byl medián věku při diagnóze 0,58 roků. V období let 1999–2005 došlo k signifikantnímu zvýšení mediánu věku při stanovení diagnózy na 1,2 roky. Navíc pacienti přicházeli ve velmi špatném stavu výživy a/nebo s nevratnými plicními změnami.

Uvedené skutečnosti nás vedly k zahájení pilotní studie NS CF jako možného nástroje zlepšení diagnostiky tohoto onemocnění v České republice.

Snahy o celoplošné vyšetřování novorozenců s cílem presymptomatické diagnostiky CF se datují od začátku druhé poloviny 20. století. Významnější pokrok nastal až ke konci 20. století se zavedením tzv. vícestupňového screeningu. V prvním stupni (kroku) je v suché kapce krve na filtračním papírku (tzv. „Guthrieho" screeningové kartičce) změřen imunoreaktivní trypsin (IRT). Pomocí tohoto biochemického parametru s vysokou senzitivitou, ale bohužel nízkou specificitou, je vyselektována skupina novorozenců s vyšším rizikem CF. Teprve až v této skupině novorozenců jsou ve druhém stupni vyhledáni pacienti s CF pomocí dalšího analytického kroku. V současné době je to především přímým vyšetřením populačně specifického spektra mutací v genu *CFTR*. Molekulárně-genetické vyšetření DNA se provádí bez potřeby dalšího odběru krve, přímo v té suché kapce krve na screeningové kartičce, kde byla zjištěna vysoká koncentrace IRT. Tento model NS CF se nazývá vícestupňový screening metodou IRT/DNA, resp. IRT/DNA/IRT, pokud do modelu screeningu je zahrnuto opakování odběru suché kapky krve (tzv. „recall") po několika týdnech života u dětí s extrémně vysokou koncentrací IRT v prvním odběru a s negativním molekulárně-genetickým vyšetřením. Podrobný přehled metod NS CF, jeho dokázaných i předpokládaných předností i možných nevýhod byl předmětem samostatného souborného referátu [15]. Jednoznačný, odbornou veřejností obecně přijatý model NS CF dosud nebyl ustanoven a v jednotlivých státech se screeningové modely dosud odlišují.

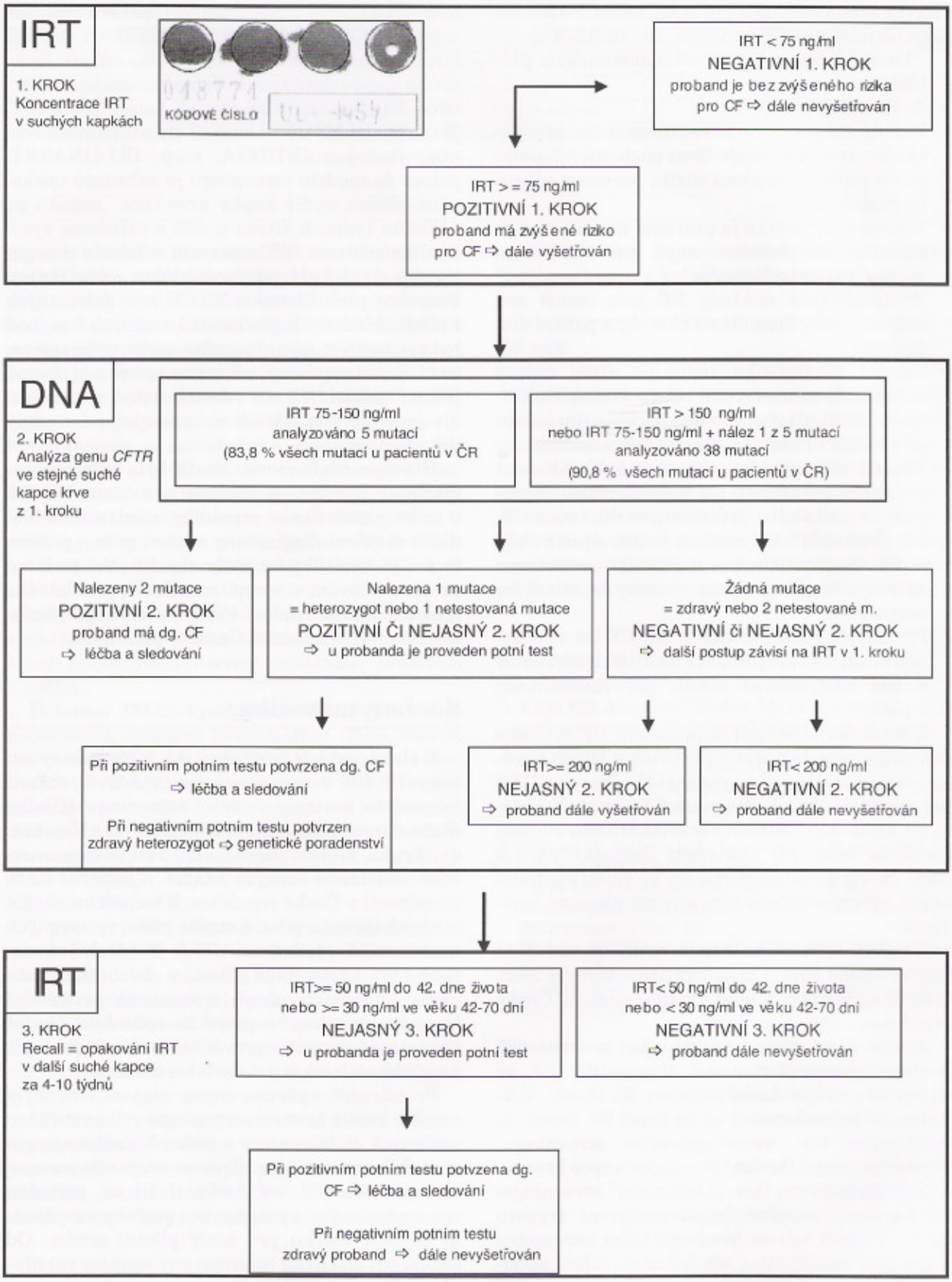
Hlavním cílem pilotní studie bylo ověřit proveditelnost a účinnost zvoleného modelu NS CF v podmínkách České republiky, především z hlediska zlepšení diagnostiky a tím i péče o pacienty s CF. Výsledky by měly sloužit jako podklad pro rozhodování o event. zavedení pravidelného celoplošného provádění tohoto screeningu Ministerstvem zdravotnictví České republiky.

Soubor, metodika

V období od 1. 2. 2005 do 2. 11. 2006 bylo vyšetřeno 76 438 novorozenců ze „spádové" oblasti laboratoře novorozeneckého screeningu Kliniky dětí a dorostu UK 3. LF a FN Královské Vinohrady, Praha. Oblast představuje 100 novorozeneckých oddělení v českých krajích a pokrývá 62 % porodnosti v České republice. Z technicko-ekonomických důvodů pilotní studie počet vyšetřených novorozenců představuje 71,3 % všech narozených (107 149) v dané oblasti a období (diskontinuita finančních zdrojů a omezená personální kapacita grantového projektu způsobila časově ohraničené výpadky provádění screeningu např. na přelomu roků či v době letní dovolené).

Primárním vyšetřovaným materiálem byly zaschlé kapky krve na screeningových kartičkách zaslaných do laboratoře v rámci běžného novorozeneckého screeningu. Krev se odebírala novorozencům mezi 72.–96. hodinou života, metodou tzv. suché kapky z patičky, bez potřeby navyšování objemu vzorku pro účely pilotní studie. Od rodičů byl získáván informovaný souhlas rozšířením informování o běžném novorozeneckém screeningu, způsobem obvyklým na každém jednotlivém novorozeneckém pracovišti, o informaci o této studii, s možností odmítnutí účasti.

Koncentrace IRT byla měřena pomocí fluoro-



Obr. 1. Schéma protokolu pilotní studie novorozeneckého screeningu cystické fibrózy v České republice, model IRT/DNA/IRT.

imunoeseje komerčním kitem Delfia® Neonatal IRT firmy Wallac Oy, Turku, Finsko. Výsledná fluorescence byla vyhodnocena pomocí fluorometru Delfia 1232 stejného výrobce.

Ve stejných suchých kapkách krve, ve kterých byla zjištěna zvýšená koncentrace IRT, byla následně provedena molekulárně-genetická analýza souboru nejčastějších, populačně specifických mutací v genu *CFTR* následujícími metodami: mutace F508del pomocí PCR-fragmentační analýzy za použití fluorescenčně značených primerů [16]; mutace *CFTR*dele2,3(21kb) byla vyšetřena pomocí PCR spojovacího fragmentu zlomové oblasti této delece [17] a mutace G551D, R553X a N1303K pomocí standardní fragmentové analýzy při štěpení amplikonů specifickými restrikčními endonukleázami (RFLP) [18, 19]. Zbývající analyzované mutace byly vyšetřeny soupravou CF 38 (Research Prototype Assay, Roche Molecular Systems, USA) [20, 21].

V suchých kapkách krve s koncentrací IRT 75–150 ng/ml bylo vyšetřeno 5 nejčastějších mutací (F508del, *CFTR*dele2.3(21kb), G551D, R553X, N1303K), které zahrnují 83,8 % patologických alel u pacientů s CF v ČR. Suché kapky s koncentrací IRT >150 ng/ml a kapky s jednou nalezenou mutací ze skupiny s koncentrací IRT 75–150 ng/ml byly analyzovány na 38 mutací (F508del, I507del, G551D, G542X, N1303K, 1717-1G>A, W1282X, R553X, R347P, R334W, 3849+10kbC>T, 621+1G>T, A455E, S549N, R560T, CFTRdele2.3(21kb), R117H, 2143delT, 175insT, E60X, G85E, 394delTT, Y122X, I148T, 711+1G>T, 1078delT, 1811+1.6kbA>G, 1898+1G>A, 2183AA>A, 2184delA, 2789+5G>A, 3120+1G>A, 3272-26A>G, Y1092X, R1162X, 3659delC, 3905insT, S1251N, včetně variant I506V, I507V, F508C a IVS8-T(5/7/9)). Toto spektrum pokrývá 90,8 % patologických alel u pacientů s CF v ČR.

V případě nálezu dvou nebo jediné mutace v genu *CFTR* byl proband (poté, co jeho rodina podepsala informovaný souhlas) pozván k potnímu testu pomocí pilokarpinové iontoforézy. Při potvrzení laboratorní diagnózy CF potním testem byl pacient zařazen do dlouhodobé klinické péče a antropologického sledování v rámci příslušného Centra CF. Při potvrzení heterozygotního stavu, tj. nosičství CF, bylo rodičům probanda nabídnuto specializované genetické poradenství. Tato služba byla rovněž automaticky poskytnuta rodinám, kde bylo diagnostikováno první dítě s CF. V případě negativního molekulárně-genetického vyšetření v kapkách krve s velmi vysokým IRT (200 ng/ml a více) bylo vyžádáno (opět po podpisu informovaného souhlasu rodiči pacienta) opakování náběru suché kapky, tzv. recall. V případě, že nedošlo k poklesu IRT, byl proband pozván k provedení potního testu. Schéma zvoleného modelu NS CF je uvedeno na obrázku 1.

Výsledky

Výsledky pilotní studie jsou sumarizovány v tabulce 1. Koncentrace IRT v suchých kapkách krve 75 ng/ml a vyšší byla zjištěna celkem u 799 novorozenců, což představuje 1,05 % všech vyšetřených. V těchto suchých kapkách krve byla provedena molekulárně-genetická analýza genu *CFTR*. U 683 probandů s koncentrací IRT 75–150 ng/ml bylo primárně vyšetřeno 5 mutací. U 116 probandů s koncentrací IRT >150 ng/ml bylo primárně vyšetřeno 38 mutací genu *CFTR* v té samé suché kapce krve, kde byla zjištěna vysoká koncentrace IRT. 37 mutací bylo testováno i u 43 probandů s koncentrací IRT 75–150 ng/ml a nálezem 1 mutace v genu *CFTR*.

Obě mutované alely genu *CFTR* byly nalezeny u 11 novorozenců. Jedna mutovaná alela genu *CFTR* byla nalezena u 53 novorozenců.

U jednoho novorozence s nálezem jedné mutace byl potní test pozitivní (Cl^- >60 mmol/l). Následným rozšířeným molekulárně-genetickým vyšetřením byla prokázána mutovaná i druhá alela (mutace I336K, která nebyla zahrnuta ve spektru 38 primárně vyšetřovaných mutací).

U 45 dětí s 1 nalezenou mutací byl potní test negativní (Cl^- <30 mmol/l), rodiny podstoupily genetické poradenství. U 1 probanda byl výsledek potního testu hraniční (Cl^- 30–40 mmol/l) a tento je dále sledován. U 6 nebyl potní test proveden (3x odmítnut v rámci informovaného souhlasu, 1 proband zemřel na vrozený nefrotický syndrom před provedením testu a 2 kojenci zatím potní test nepodstoupili). Všechny probandy s 1 mutací a nepotvrzenou CF považujeme při sumarizaci výsledků za zdravé heterozygoty.

U 38 novorozenců s velmi vysokou koncentrací IRT v prvním kroku (200 ng/ml a více), u kterých nebyla nalezena žádná mutace v genech *CFTR*, byl ve věku 3–10 týdnů vyžádán opakovaný odběr suché kapky krve (tzv. „recall"). Koncentrace IRT poklesla v opakovaném odběru do negativních hodnot (tj. pod 50 ng/ml ve věku do 42. dne, resp. pod 30 ng/ml ve věku 42–70 dní) u 18 z nich. Ve 4 případech nedošlo k poklesu IRT do negativních hodnot a u 2 nebyl odběr získán v požadovaném čase. U těchto 6 kojenců byl „recall" považován za pozitivní a byli pozváni k provedení potního testu. Zatím byl proveden u 4 z nich, s negativním výsledkem. „Recall" byl 5x v rámci informovaného souhlasu odmítnut a v 9 případech není vyšetření ukončeno. 38 opakovaných odběrů („recallů") v naší kohortě vyšetřených novorozenců představuje četnost potřeby opako-

Tab. 1. Výsledky pilotní studie novorozeneckého screeningu CF v České republice.

Počet vyšetřených novorozenců (n):	76 438
Počet probandů s IRT v suché kapce => 75 ng/ml (n):	799
Počet diagnostikovaných pacientů s CF (n):	12 (11x na základě nálezu 2 mutací, 1x na základě nálezu 1 mutace + potní test)
Incidence CF ve screenované kohortě novorozenců:	1:6369
Genotypy zachycených pacientů:	4x F508del/F508del 2x F508del/G551D 1x F508del/2143delT 1x F508del/I336K 1x F508del/N1303K 1x F508del/CFTRdele2,3 1x F508del/3849+10kbC>T 1x F508del/R117H
Věk při diagnóze (dny) medián: rozsah:	37 26–54
Počet probandů s nálezem 1 mutace (heterozygoti) genu *CFTR* (n):	52
Incidence heterozygotů (v populaci novorozenců s IRT => 75 ng/ml):	1 : 15,4
Výsledky potního testu u heterozygotů (n):	Negativní (Cl v potu <30 mmol/l): 45 Hraniční (Cl v potu 30–40 mmol/l): 1 Potní test zatím neproveden: 6
Počet „recallů" – žádná mutace + IRT >200 ng/ml (n):	38
Výsledky „recallů" (n):	Negativní: 18 Pozitivní: 6 Odmítnut: 5 Dosud neproveden: 9
Výsledky potního testu u pozitivních „recallů" (n):	Negativní (Cl v potu <30 mmol/l): 4 Potní test dosud neproveden: 2

Tab. 2. Koncentrace IRT v jednotlivých skupinách probandů. P-hodnota byla vypočítána pomocí Mannova-Whitneyova testu.

V náhodně vybrané části populace novorozenců, n = 6433 (medián, rozsah):	20,2 (2–389,7) ng/ml
Skupina A: probandi s IRT > = 75 ng/ml a žádnou mutací v *CFTR* genech, n = 735 (medián, rozsah):	95,7 (75–910) ng/ml
Skupina B: probandi s IRT > = 75 ng/ml a 1 mutací v *CFTR* genu (= heterozygoti), n = 52 (medián, rozsah):	90,2 (75–424) ng/ml
Skupina C: probandi s IRT > = 75 ng/ml a 2 mutacemi v *CFTR* genech (= pacienti s CF), n = 12 (medián, rozsah):	149,2 (79–352,7) ng/ml
P-hodnota – skupina C versus skupina A:	0,05
P-hodnota – skupina C versus skupina B:	0,07
P-hodnota – skupina B versus skupina A:	0,68

vání screeningu pro nejasný výsledek (tzv. „recall rate") 0,05 %.

Celkem jsme pomocí novorozeneckého screeningu ve vyšetřené kohortě novorozenců diagnostikovali 12 dětí s cystickou fibrózou, což představuje incidenci 1:6369. Medián věku při stanovení diagnózy pomocí NS činil 37 dnů (rozsah 26–54). V 8 případech byla diagnóza stanovena presymptomaticky, ve 4 případech v době výsledku NS již měly postižené děti diagnózu stanovenou na základě klinických příznaků (mekoniový ileus, bronchopneumonie).

Budeme-li předpokládat, že všech 52 probandů s nalezenou 1 mutací genu *CFTR* jsou zdraví heterozygoti, tak frekvence heterozygotů v populaci novorozenců s IRT 75 ng/ml a více byla 1:15,4 (6,5 %).

V tabulce 2 jsou uvedeny mediány koncentrací IRT v jednotlivých skupinách probandů. Statisticky významný vyšel rozdíl mezi koncentrací IRT pacientů s CF a ostatních novorozenců s IRT 75 ng/ml a vyšším. Testovat významnost rozdílů koncentrace IRT heterozygotů či pacientů s CF vůči běžné populaci novorozenců nemá logické

opodstatnění, protože primárně byli k DNA analýze vybíráni probandi s koncentrací IRT nad 98,95. percentilem.

Diskuse

Měření koncentrace IRT v prvním kroku NS CF slouží k vyselektování populace novorozenců s vyšším rizikem pro CF. Vysoká koncentrace tohoto pro pankreas specifického proteinu je způsobena jeho únikem do cirkulace z poškozených buněk acinů pankreatu při blokádě pankreatických vývodných cest hyperviskozním hlenem [22, 23]. S postupující atrofií pankreatu koncentrace IRT v krvi pacientů s CF klesá [24, 25]. Vysokou koncentraci IRT má však i část zdravých novorozenců, vliv má etnický původ, IRT může být zvýšeno následkem hypoxie, vyšší je i u heterozygotů pro CF, při virovém onemocnění, u chromosomálních aberací a při některých vrozených vývojových vadách ledvin, srdce a neurální trubice [26, 27, 28, 29]. Naopak pacienti s CF mohou mít nízkou koncentraci IRT ve screeningových suchých kapkách krve v případě mekoniového ileu [30] nebo při nezralosti (porod před 32. týdnem gestace) [24]. Zásadní otázkou v modelu NS CF je nastavení hodnoty selekční, „cut-off", koncentrace IRT v prvním kroku. Jinými slovy jde o rozhodnutí, jaká část běžné populace novorozenců bude považována za rizikovou pro CF a bude u nich následně provedena analýza mutací v genu *CFTR*. V tomto bodě se modely NS CF v jednotlivých státech značně liší. Hodnotu koncentrace IRT pro „cut-off" lze nalézt mezi 90.–99,5. percentilem, tzn. že DNA vyšetření se provádí u 0,5–10 % populace novorozenců [31, 32, 33, 34]. V posledních letech lze pozorovat tendenci k poklesu procenta novorozenecké populace, u kterých je DNA analýza prováděna. Jako selekční „cut-off" koncentrace IRT převládá 99.–99,5. percentil [35]. V našem modelu byla DNA vyšetřena u 1,05 % novorozenců, odpovídající percentil IRT pro „cut-off" byl tedy 98,95. Nabízí se logická otázka, zda takovéto nastavení „cut-off" nemůže být zdrojem falešné negativity NS CF. Jednoznačnou odpověď z naší studie nelze dát, lze se pouze opírat o citované zahraniční zkušenosti, kde převládá názor, že DNA testování probandů s IRT nad 99.–99,5. percentilem je dostatečně senzitivní z hlediska záchytu pacientů. Zvětšování počtu testovaných novorozenců nevede k vyšší senzitivitě, pouze se zvyšuje počet zachycených zdravých heterozygotů a významně stoupají ekonomické náklady na screening.

Jedenáct pacientů s CF bylo diagnostikováno přímo v druhém stupni screeningu na základě nálezu 2 mutací genu *CFTR*. Jeden pacient byl zachycen na základě potního testu při průkazu jedné patologické alely. Následnou rozšířenou analýzou byla na druhé alele nalezena vzácná mutace I336K pomocí přímého sekvenování genu *CFTR*. Medián věku při stanovení diagnózy jednoznačně svědčí pro včasnost záchytu pacientů. Otázkou zůstává význam záchytu pacienta s mírnou formou choroby (F508del/R117H), jejíž průběh může být variabilní nebo až prakticky asymptomatický. Někteří autoři [36] považují tuto mutaci spíše za benigní variantu a nedoporučují ji zařazovat do panelu NS CF. Důležité je i provedení takzvaného reflexního testu u této mutace vyšetřením konstituce potenciálně patogenní varianty IVS8 T(n), kde nález komplexní alely „5T-R117H" je spojen s relativně horším klinickým průběhem CF než u kombinace „7T-R117H" nebo „9T-R117H".

Není bez zajímavosti, i když se nejedná o přímý výsledek NS CF, že na základě diagnózy CF u 3 probandů z naší kohorty byla následně diagnostikována CF i u jejich 3 starších sourozenců. Ti byli zachyceni ve věku 3 až 13 let a měli příznaky CF!

Všichni diagnostikovaní pacienti byli zařazeni do klinických center pro CF k léčbě, klinickému a antropologickému sledování a rodiny podstoupily genetické poradenství.

Zjištěná incidence CF 1:6369 je překvapivě nižší, než by odpovídalo dosavadní, empirické incidenci v ČR (1:2736) [1, 2, 3, 4]. Možnou příčinu této diskrepance lze hledat:

1. v zatím stále relativně malém počtu vyšetřených novorozenců, tj. v chybě malých čísel;
2. ve falešné negativitě screeningu, což je však velmi málo pravděpodobné;
3. ve vlivu prenatální diagnostiky CF, kdy se v pozitivních případech rodiny rozhodly pro selektivní přerušení těhotenství.

Z dostupných informací bylo zjištěno, že v uvedeném období bylo ve spádové oblasti pilotní studie těhotenství přerušeno z důvodů CF v 8 případech. Takto korigovaná incidence by pak činila 1:3821, což odpovídá i nálezům z poslední doby v ostatních evropských zemích provádějících NS CF. Zatím nebyl zjištěn případ klinicky diagnostikované choroby, která by unikla záchytu screeningem.

Četnost heterozygotů ve skupině novorozenců s elevovaným IRT 1:15,4 (6,5 %) je téměř 2x vyšší než v běžné české populaci (1:26, 3,8 %) [3, 4]. Vedlejším „produktem" NS CF je tedy i rozšiřování genetického poradenství a informované reprodukce. Na druhé straně je zřejmé, že 52 zachycených heterozygotů představuje jen cca 2 % všech heterozygotů ve vyšetřované kohortě novorozenců (použijeme-li populačně genetické zákony) a 98 % heterozygotů (odhadem 2900 jedinců)

uniklo záchytu, protože jejich IRT byl pod „cut-off". Pochopitelně detekci heterozygotů nelze považovat za hlavní cíl NS CF.

Četnost opakování náběru suché kapky pro nejasný výsledek screeningu („recall rate") 0,05 % je ve srovnání s jinými screeningovými programy extrémně nízká (např. u kongenitální adrenální hyperplazie bývá „recall rate" 0,3–0,5 %, u kongenitální hypotyreózy a fenylketonurie 0,1–0,2 %). Tato hodnota neznehodnocuje NS CF nepřiměřenou zátěží zdravé části populace a ani nebude při event. celoplošném pravidelném provádění NS CF významněji zvyšovat kumulativní „recall rate" s ostatními screeningy.

Závěr

Výsledky naší studie považujeme za důkaz proveditelnosti novorozeneckého screeningu cystické fibrózy v podmínkách České republiky i jeho efektivity z hlediska včasnosti diagnostiky této choroby. Jeho celoplošné pravidelné provádění by bylo zásadním nástrojem zlepšení péče o pacienty s CF a zkvalitněním zdravotního systému v našem státě. Nezodpovězenou otázkou zůstává, zda použitý model NS CF skutečně zachytil všechny postižené děti.

Práce vznikla za podpory grantu IGA MZ ČR 1A/8236-3.

Literatura

1. **Houštěk J, Vávrová V.** K výskytu cystické fibrosy pankreatu v ČSSR. Čs. Pediat. 1962;17:445–451.
2. **Houštěk J, Vávrová V.** Výskyt a klinické formy mukoviscidosy. Prakt. Lék. 1969;49:219–222.
3. **Macek M Jr, Krebsová A, Macek M, et al.** Current possibilities of prenatal and postnatal molecular genetic diagnosis of cystic fibrosis in Czech and Slovak Republics. Čes.-slov. Pediat. 1997;52:557–564.
4. **Macek M Jr, Vávrová V.** Cystická fibróza. Model komplexní patogeneze u monogenních onemocnění. Postgraduální medicína 2002;4:530–534.
5. **Davis PB.** Cystic fibrosis since 1938. Am. J. Respir. Crit. Care Med., 2006;173:475–482.
6. **Farrell PM, Li Z, Kosorok MR, et al.** Bronchopulmonary disease in children with cystic fibrosis after early or delayed diagnosis. Am. J. Respir. Crit. Care Med. 2003;168:1100–1108.
7. **Wang SS, O'Leary LA, Fitzsimmons SC, et al.** The impact of early cystic fibrosis diagnosis on pulmonary function in children. J. Pediatr. 2002;141:804–810.
8. **Shwachman H, Redmond A, Khaw KT.** Studies in cystic fibrosis. Pediatrics 1970;46:335–343.
9. **Sims EJ, Clark A, McCormick J, et al.** Cystic fibrosis diagnosed after 2 months of age leads to worse outcomes and requires more therapy. Pediatrics 2007;119:19–28.
10. **Wagener JS, Farrell PM, Corey M.** A debate on why my state should or should not conduct newborn screening for cystic fibrosis. Pediatr. Pulmonol. 2001;32:385–396.
11. **Farrell MH, Farrell PM.** Newborn screening for cystic fibrosis: ensuring more good than harm. J. Pediatr. 2003;143:707–712.
12. **Castellani C.** Evidence for newborn screening for cystic fibrosis. Paediatr. Respir. Rev. 2003;4:278–284.
13. **Wagener JS, Sontag MK, Sagel SD, et al.** Update on newborn screening for cystic fibrosis. Curr. Opin. Pulm. Med. 2004;10:500–504.
14. **Vávrová V, Zemková D, Skalická V, Votava F.** Problémy v diagnostice cystické fibrózy – potřeba novorozeneckého screeningu. Čes.-slov. Pediat. 2006;12:703–709.
15. **Dittertová L, Vávrová V, Skalická V, a kol.** Novorozenecký screening cystické fibrózy. Čes.-slov. Pediat. 2005;6:373–378.
16. **Macek M Jr, Ladanyi L, Burger J, Reis A.** Missense variations in the cystic fibrosis gene: heteroduplex formation in the F508C mutation. Am. J. Hum. Genet. 1992;5:1173–1174.
17. **Dörk T, Macek M Jr, Mekus F, et al.** Characterization of a novel 21-kb deletion, CFTRdele2,3(21 kb), in the CFTR gene: a cystic fibrosis mutation of Slavic origin common in Central and East Europe. Hum. Genet. 2000;3:259–268.
18. **Kerem BS, Zielenski J, Markiewicz D, et al.** Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. Proc. Natl. Acad. Sci. USA. 1990;21:8447–8451.
19. **Friedman KJ, Highsmith WE Jr, Silverman LM.** Detecting multiple cystic fibrosis mutations by polymerase chain reaction-mediated site-directed mutagenesis. Clin. Chem. 1991;5:753–755.
20. **Kawasaki E, Saiki R, Erlich H.** Genetic analysis using polymerase chain reaction-amplified DNA and immobilized oligonucleotide probes: reverse dot-blot typing. Methods Enzymol. 1993;218:369–381.
21. **Kawasaki E, Chehab FF.** Analysis of gene sequences by hybridization of PCR-amplified DNA to covalently bound oligonucleotide probes. The reverse dot blot method. Methods Mol. Biol. 1994;28:225–236.
22. **Crossley JR, Elliot RB, Smith PA.** Dried blood spot screening for cystic fibrosis in newborn. Lancet 1979;8114:472–474.
23. **Lecoq I, Brouard J, Laroche D, et al.** Blood immunoreactive trypsinogen concentrations are genetically determined in healthy and cystic fibrosis newborns. Acta Paediatr. 1999;88:338–341.
24. **Bourguignon JP, Deby-Dupont G, Reuter A, et al.** Variations in dried blood spot immunoreactive trypsin in relation to gestational age and during the first week of life. Eur. J. Pediatr. 1986;144:547–549.
25. **Rock MJ, Mischler EH, Farrell PM, et al.** Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. Pediatrics 1990;85:1001–1007.
26. **Laroche D, Travert G.** Abnormal frequency of delta F508 mutation in neonatal transitory hypertrypsinaemia. Lancet 1991;8732:55.
27. **Pederzini F, Mastella G, Giglio L, et al.** Dried blood spot immunoreactive trypsin (IRT) values. Eur. J. Pediatr. 1988;147:663.
28. **Moya EF, Brocklebank JT, Littlewood JM, et al.** High serum immunoreactive trypsin not caused by cystic fibrosis. Arch. Dis. Child. Fetal. Neonatal. Ed. 1998;78:F78.
29. **Ravine D, Francis RI, Danks DM.** Non-specific elevation of immunoreactive trypsinogen in sick infants. Eur. J. Pediatr. 1993;152:348–349.

30. **Rusakow LS.** Immunoreactive trypsinogen levels in infants with CF complicated by meconium ileus. Screening 1993;2:13–18.
31. **Kaye CI.** Newborn screening fact sheets. Pediatrics 2006;118:934–963.
32. **Comeau AM, Parad RB, Dorkin HL, et al.** Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. Pediatrics 2004;113:1573–1581.
33. **Ranieri E, Lewis BD, Gerace RL, et al.** Neonatal screening for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis: four years' experience. BMJ 1994;6942:1469–1472.
34. **Telleria-Orriols JJ, Alonso-Ramos MJ, Garrote-Adrados JA, et al.** Neonatal screening for cystic fibrosis. An. Esp. Pediatr. 2002;57:60–65.
35. **Southern KW, Munck A, Pollit R, et al.** A survey of newborn screening for cystic fibrosis in Europe. J. Cyst. Fibros. 2006;Epub ahead of print.
36. **Scotet V, Audrézet MP, Rousseey M, et al.** Immunoreactive Trypsin/DNA Newborn Screening for Cystic Fibrosis: Schould the R117H Variant Be Included in CFTFR Mutation Panels. Pediatrics 2006;118:1523–1529.

Došlo: 3. 11. 2006
Přijato: 5. 1. 2007

Doc. MUDr. Felix Votava, PhD.
Klinika dětí a dorostu UK 3. LF
FN Královské Vinohrady
Šrobárova 50
100 34 Praha 10
e-mail: votava@fnkv.cz

ELSEVIER

Journal of Cystic Fibrosis 6 (2007) 57 – 65

Journal of Cystic Fibrosis

www.elsevier.com/locate/jcf

A survey of newborn screening for cystic fibrosis in Europe

Kevin W. Southern [a], Anne Munck [b], Rodney Pollitt [c], Georges Travert [d], Luisa Zanolla [e], Jeannette Dankert-Roelse [f], Carlo Castellani [g,*]
on behalf of the ECFS CF Neonatal Screening Working Group

[a] *University of Liverpool, Royal Liverpool Children's Hospital, Liverpool, United Kingdom*
[b] *AFDPHE, Paris, France*
[c] *Neonatal Screening Laboratory, Children's Hospital, Sheffield, UK*
[d] *Service de Biophysique, CHU de Caen, Caen, France*
[e] *Cardiology Department, Verona, Italy*
[f] *Atrium Medical Centre, Department of Pediatrics, Heerlen, The Netherlands*
[g] *Cystic Fibrosis Center, Azienda Ospedaliera Verona, piazzale Stefani 1, 37126 Verona, Italy*

Received 6 April 2006; received in revised form 9 May 2006; accepted 11 May 2006
Available online 25 July 2006

Members of European Cystic Fibrosis Society Neonatal Screening Working Group who contributed to the survey

Aitken David	Newborn Screening Laboratory, Institute of Medical Genetics, Glasgow, UK
Alonso Ramos Maria Jesus	Genetic Screening Service, Castilla-Leon, Spain
Antonozzi Italo	Genetic and metabolic disease section, University Hospital, Rome, Italy
Balascakova Miroslava	Institute of Biology and Medical Genetics, Charles University Prague-2, School of Medicine, Czech Republic
Bernardi Filippo	Clinica Pediatrica S. Orsola Malpighi, Bologna, Italy
Bignamini Elisabetta	Pediatric Cystic Fibrosis Center, Turin, Piedmont, Italy
Bodamer Olaf	University Children Hospital, Vienna, Austria
Bradley Don	Newborn Screening Service, Wales, UK
Bucci Micaela	Clinica Pediatrica S. Orsola Malpighi, Bologna, Italy
Burroni Massimo	Centro Screening Neonatale Regionale, Ospedale di Fano, Fano, Italy
Calvin Jacqui	Addenbrooke's Hospital, Cambridge, UK
Cerone Roberto	Dipartimento di Pediatria—Istituto G. Gaslini, Genova, Italy
Colon Mejeras Cristobal	Unidad de Deteccion Precoz Neonatal, Departemento de Pediatria, Santiago de Compostela, Spain
Corbetta Carlo	Laboratorio di Riferimento regionale per lo Screening Neonatale, Ospedale dei Bambini "V. Buzzi", Milano, Italy
Eichler Irmgard	Cystic Fibrosis Center AKH, Vienna, Austria
Gartner Silvia	Cystic Fibrosis Center, Catalonia, Spain
Holubova Andrea	Institute of Biology and Medical Genetics, Charles University Prague-2, School of Medicine, Czech Republic
Iapichino Luciana	Centro Regionale Fibrosi Cistica, Palermo, Italy
Kracmar Petr	Department of Pediatrics, Charles Univ. Prague-3, School of Medicine, Czech Republic
Leavy Anne	Royal Belfast Hospital for Sick Children, Belfast, UK
Lelli Alessandra	Italian Red Cross Central Laboratory, Rome, Italy
Leslie Hilary	Royal Belfast Hospital for Sick Children, Belfast, UK
Lilliu Franco	Malattie del Metabolismo e Screening Neonatale, Centro Microcitemie, Cagliari, Italy
Lucie Dittertova	Department of Pediatrics, Charles Univ. Prague-3, School of Medicine, Czech Republic
Macek Milan	Institute of Biology and Medical Genetics, Charles University Prague-2, School of Medicine, Czech Republic
Oltarzewski Mariusz	Neonatal Screening Laboratory, Warsaw, Poland
Pagliardini Severo	Centro Screening Neonatali, Ospedale Infantile Regina Margherita, Torino, Italy
Piskackova Tereza	Institute of Biology and Medical Genetics, Charles University Prague-2, School of Medicine, Czech Republic
Provenzano Ettore	Cystic Fibrosis Center, Soverato, Calabria, Italy
Reid Alastair	Royal Belfast Hospital for Sick Children, Belfast, UK
Restagno Gabriella	Dipartimento di Patologia Clinica, S.C. Genetica Molecolare, Turin, Italy

* Corresponding author. Tel.: +39 45 8072293; fax: +39 45 8072042.
E-mail address: carlo.castellani@azosp.vr.it (C. Castellani).

doi:10.1016/j.jcf.2006.05.008

Sands Dorata	Institute of Mother and Child, Warsaw, Poland
Scott Nigel	Northampton General Hospital NHS Trust, Northampton, UK
Shapiro Leonie	St James' University Hospital, Leeds, UK
Skalicka Veronika	Department of Pediatrics, Charles Univ. Prague-2, School of Medicine, Czech Republic
Taccetti Giovanni	Cystic Fibrosis F Centre, Meyer Hospital, Florence, Italy
Vavrova Vera	Department of Pediatrics, Charles Univ. Prague-2, School of Medicine, Czech Republic
Votava Felix	Department of Pediatrics, Charles Univ. Prague-3, School of Medicine, Czech Republic
Zemkova Dana	Department of Pediatrics, Charles Univ. Prague-2, School of Medicine, Czech Republic

Abstract

Background: Cystic fibrosis (CF) is a recessively inherited condition caused by mutation of the *CFTR* gene. Newborn infants with CF have raised levels of immuno-reactive trypsinogen (IRT) in their serum. Measurement of IRT in the first week of life has enabled CF to be incorporated into existing newborn screening (NBS) blood spot protocols. However, IRT is not a specific test for CF and NBS therefore requires a further tier of tests to avoid unnecessary referral for diagnostic testing. Following identification of the *CFTR* gene, DNA analysis for common CF-associated mutations has been increasingly used as a second tier test. The aim of this study was to survey the current practice of CF NBS programmes in Europe.

Method: A questionnaire was sent to 26 regional and national CF NBS programmes in Europe.

Results: All programmes responded. The programmes varied in number of infants screened and in the protocols employed, ranging from sweat testing all infants with a raised first IRT to protocols with up to four tiers of testing. Three different assays for IRT were used; in the majority (24) this was a commercially available kit (Delfia™). A number of programmes employed a second IRT measurement in the 4th week of life (as the IRT is more specific at this point). Nineteen programmes used DNA analysis for common *CFTR* mutations on samples with a raised first IRT. Three programmes used a second IRT measurement on infants with just one recognised mutation to reduce the number of infants referred for sweat testing. Referral to clinical services was prompt and diagnosis was confirmed by sweat testing, even in infants with two recognised mutations in most programmes. Subsequent clinical pathways were less uniform. Multivariate analysis demonstrated a relationship between the age of diagnosis and the timing of the first IRT. More sweat tests were undertaken if the first IRT was earlier and the diagnosis was later.

Conclusions: Annually these programmes screen approximately 1,600,000 newborns for CF and over 400 affected infants are recognised. The findings of this survey will guide the development of European evidence based guidelines and may help new regions or nations in the development and implementation of NBS for cystic fibrosis.

Keywords: Cystic fibrosis; Neonatal screening; Europe; IRT; Diagnosis

1. Introduction

Cystic fibrosis (CF) is caused by mutation of the *Cystic Fibrosis Transmembrane Conductance Regulator* (*CFTR*) gene [1]. In parts of Europe, the incidence can be as high as 1 in every 1000 births with Caucasians affected more frequently than other ethnic groups [2]. The heterogeneous nature of the condition to some extent reflects the large number of mutations that can affect the *CFTR* gene and the impact of other gene modifiers [3]. With the "classic" severe phenotype, outlook without treatment is poor with death frequently occurring in the first decade of life [4]. Prognosis has improved significantly with early and active treatment of chest infection and an holistic approach to management with particular attention to nutritional well being [5]. However, even with modern treatment regimens death during childhood is still occasionally reported [6,7].

Infants with a "severe" phenotype often have a short asymptomatic neonatal period and recognition in the 1970s that CF infants have high serum immunoreactive trypsinogen (IRT) levels, prompted the suggestion that CF may be a suitable condition for newborn screening (NBS) [8]. A raised IRT in the first week of life is a sensitive test for CF but not specific and a second "tier" of testing avoids an inappropriate number of families presenting to clinical services for definitive diagnostic testing [9,10]. In early protocols, the second tier involved repeating the IRT measurement at 3–4 weeks when a raised level is more specific for the condition [10]. A second tier by DNA analysis has become more widely used since recognition of the *CFTR* gene in 1989. Advantages of DNA analysis are that it can be undertaken on the original blood spot sample and provides a specific result with the recognition of two CF causing mutations [11–13]. A potential disadvantage of DNA analysis is carrier recognition [14,15]. Increased sweat chloride concentration remains an important diagnostic marker of CF, particularly when preliminary DNA analysis is negative or not fully informative, and sweat testing is an integral component of CF newborn screening [16].

The European CF Society has established a Working Group to examine and co-ordinate newborn screening for CF across Europe. Objectives include mapping the current situation, producing guidelines on critical issues and

facilitating the implementation of CF NBS. More than 70 experts from various European countries have been invited to take part in this project. In order to achieve a comprehensive picture of the situation in Europe, a questionnaire was circulated to European CF NBS programmes to survey current practice.

2. Materials and methods

The questionnaire was developed by a focus group and divided into sections (Fig. 1): screening protocol, sample collection (who collects and how, collection day), immunoreactive trypsinogen (levels, centiles, when tested, how tested), genetic analysis (what mutations, techniques, informed consent issues), sweat test (suggested age, positive and borderline values, techniques), diagnosis (diagnostic criteria, communication to the family), follow-up (clinical protocols, segregation issues), data storing (informatics tools, card storage), epidemiology (numbers of screened newborns, identified CF cases, carriers, false positives, false negatives). Aside from false negatives respondents were asked to reflect on their current practice.

The questionnaire was sent to all established CF NBS programmes in Europe known to the focus group. CF physicians and societies in all European countries were approached to achieve as complete an inclusion of programmes as possible. Recipients were contacted and encouraged to complete the questionnaire.

PROTOCOL DESCRIPTION

a.1 Please describe your CF neonatal screening protocol, and specify if it is an established programme or a pilot study. The sequential steps leading from IRT to diagnosis should be briefly outlined.

SAMPLE COLLECTION

b.1 Please briefly describe how the blood samples for IRT are collected and spotted, and by whom.
b.2 Which other (routine as well as experimental) neonatal screening analyses are performed on the collected bloood samples?
b.3 How many blood spots are collected per newborn?
b.4 What percentage of samples results inadequately smeared on the card, thus not allowing IRT determination?
b.5 How are the parents informed about the screening procedures and aims? Do they have to give their consent? If so, written or oral consent?
b.6 At what age (range) do you advice the maternity wards staff to collect the sample for IRT (heel prick for the Guthrie card)?
b.7 Please give details about dried blood spots mailing to the screening laboratory (wrapping, average delivery time, etc.)
b.8 If your screening strategy includes resampling (recalling IRT-positive neonates for a second blood spot), at which age (range) is the second sample collected, who contacts the family, and what is told to the parents?
b.9 Comments (if any)

IRT

c.1 What is the IRT technique you use?
c.2 What is the level (microgr/l) of your current cutoff for IRT at birth?
c.3 What is the centile of your current cutoff for IRT at birth, and how did you calculate it?
c.4 Do you periodically adjust your IRT cutoff level?
c.5 If your screening strategy includes resampling, what is the level (microgr/l) of your current IRT cutoff for the second sample?
c.6 If your screening strategy includes resampling, what is the centile of your current IRT cutoff for the second sample, and how did you calculate it?
c.7 Are there any cases where you proceed to your protocol second step even though IRT at birth was below the cutoff? If so, in which cases (positive family history, meconium ileus, others?)
c.8 Do you perform any internal IRT quality control?
c.9 Do you participate in any external (national or international) IRT quality control program?
c.10 Comments (if any)

GENETIC ANALYSIS

d.1 Do you test IRT-positive neonates for CF mutations?
d.2 How many and which mutations do you test for?
d.3 What is the frequency of these mutations in your population?
d.4 What technique do you use for genetic analysis?
d.5 Do you resample or sweat test neonates with extremely elevated IRT when no CF mutations are identified?
d.6 Comments (if any)

SWEAT TEST

e.1 From which age do you perform sweat test?
e.2 What technique do you use?
e.3 Do you perform sweat tests in only one laboratory? If not, in how many?
e.4 How many sweat tests per year are performed per lab?
e.5 What is the minimum amount of sweat you need in order to consider the test reliable (mgs)?
e.6 What is your chloride borderline range?
e.7 Do infants have a confirmatory sweat test even if two mutations are detected?
e.8 Comments (if any)

Fig. 1. Questionnaire.

DIAGNOSIS

f.1 What are your criteria for making a CF diagnosis in a screened newbron?
f.2 Do you perform further diagnostic procedures in case a CF diagnosis is uncertain, and if so, which diagnostic procedures do you use?
f.3 Who communicates the diagnosis to the parents?
f.4 How long does it take between birth and diagnosis communication to the family?
f.5 Do you communicate the diagnosis by letter, phone, personal contact, or how else?
f.6 What are the main concept you convey to the parents when the diagnosis is communicated?
f.7 Do you refer parents of identified CF newborns for genetic counselling? When do you refer them?
f.8 Are parents of incidentally identified carriers notified? Are they referred for genetic counselling? How many of them accept the offer of genetic counselling?
f.9 Comments (if any)

FOLLOW UP

g.1 To what sort of clinical institution is the family referred after the diagnosis?
g.2 What is the average time span between diagnosis communication and the first clinical evaluation?
g.3 Are identified CF newborns admitted to the hospital for clinical evaluation?
g.4 Are precautions taken to prevent early acquisition of *Pseudomonas aeruginosa* in the hospital? If so, please specify them.
g.5 Are identified CF newborns included in specific follow-up protocols and/or treated according to special guidelines?
g.6 Are borderline chloride cases included in any specific follow-up protocol? Please, briefly describe it.
g.7 Comments (if any)

DATA AND CARDS STORING

h.1 Please describe briefly your database for data processing and storing.
h.2 How and for how long do you store Guthrie cards?

EPIDEMIOLOGICAL DATA

i.1 For how long has your CF neonatal screening programme been running?
i.2 How many neonates do you screen per year?
i.3 How many neonates per year have birth IRT above your chosen cutoff?
i.4 How many sweat tests related to neonatal screening do you perform per year?
i.5 How many affected neonates do you detect per year?
i.6 How many heterozygous neonates do you detect per year?
i.7 How many false negatives are you aware of since the beginning of the programme?
i.8 Are CF patients missed by neonatal screening reported to you?
i.9 What have been the causes of missed cases (IRT below cutoff in first/second sample, other mutations than tested for, or other; please specify)
i.10 Comments (if any)

Fig. 1 (*continued*).

When not otherwise specified quantitative data are reported as median, with interquartile ranges in brackets. Univariate statistics results are not reported when less than 50% answers were obtained. Multivariate analysis takes into account outliers.

3. Results

Twenty-six questionnaires were sent in 2004/2005 and all were returned: 7 from the UK, 1 (nationwide) from France, 12 from Italy, 3 from Spain, 1 (nationwide) from Austria, 1 from Poland and 1 from the Czech republic (Table 1). Some questions were not answered, either because they did not apply to the programme or the data were not available to the person completing the questionnaire.

3.1. Descriptive results

3.1.1. Programme size and performance

IRT measured from a blood spot sample taken during the first week of life was the initial step of all programmes (IRT-1). Subsequently, a wide variety of protocols were reported ranging from moving directly to sweat testing infants with a raised IRT-1 to protocols involving four steps (Table 1). The number of infants screened ranged from 8000 to 800,000. From these data, we could calculate that incidence ranges from 1/2250 to 1/10,500 (Table 2).

These services have been working for an average of 10 years (range 9 months to 31 years), screen more than 1,600,000 newborns per year, and every year detect more than 400 affected neonates with an overall median calculated incidence of 1/3500 (Table 2). The median number of infants screened each year was 30,000 (18,000–54,000), the median number of raised IRT-1, 295 (148–825) and the median number of sweat tests, 70 (20–129). Subsequently, the median number of CF cases identified per year was 9 (5–14) and carriers 17 (13–25). The median number of false negatives reported was 2 (1–5), but only 15 programmes answered the question. This number was dependent on both duration of screening programme and the communication of late diagnosis by clinical services. False negatives were caused by IRT-1 below the cut-off in the majority of reports (Table 3).

Median age at diagnosis was 37 days (32–50), with significant difference between programmes (Table 2). There were clear associations between age at diagnosis and timing of IRT-1 and resampling IRT (IRT-2) cut-off (see multivariate analysis).

Table 1
Description of NBS programmes included in survey

	Area	2nd tier	3rd tier	4th tier	Details
Two-tier protocols					
I	Liguria	ST	–	–	
Three-tier protocols					
CZ	Western Czech republic	MUT	ST	–	PS; started February 2005
UK	Wales	MUT	ST	–	
UK	Northern Ireland	IRT-2	ST	–	
UK	Leeds Halifax Jersey	MUT IRT-2	ST	–	
I	Emilia Romagna	IRT-2	ST	–	
I	Calabria	IRT-2	ST MUT	–	PS; survey answers based on 10 months experience
I	Sardinia	ST	MUT	–	PS; survey answers based on 9 months experience
I	Lombardy	MUT IRT-2	ST	–	IRT-2 if IRT-1 > 97.5° centile MUT if IRT-1 > 99° centile
I	Marche	MUT IRT-2	ST	–	PS; IRT-2 if IRT-1 > 97.5° centile MUT if IRT-1 > 99.8° centile
I	Tuscany	MP IRT-2	ST	–	
I	Piedmont	MUT IRT-2	ST	–	PS; IRT-2 if IRT-1 > 98.6° centile MUT if IRT-1 > 99.6° centile
I	Lazio 1	IRT-2	ST MUT	–	
I	Lazio 2, Umbria	MUT IRT-2	ST	–	
I	Western Sicily	MUT IRT-2	ST	–	
A	Austria	IRT-2	ST	–	
SP	Catalunya	IRT-2	MUT ST	–	PS
SP	Castilla-Leon	IRT-2 MUT	ST		
SP	Galice	IRT-2	MUT ST	–	PS
Four-tier protocols					
F	France	MUT	IRT-2	ST	IRT-2 if MUT-tive
PL	Poland	MUT	IRT-2	ST	PS 1999–2003; IRT-2 if MUT-tive
UK	South Yorkshire East Midlands	MUT	IRT-2	ST	IRT-2 if 1 mutation or IRT-1 > 99.9th centile
UK	Scotland	MUT	IRT-2	ST	IRT-2 if 1 mutation or no mutations and non-Caucasian
UK	Northamptonshire	MUT	IRT-2	ST	IRT-2 if 1 or no mutations
UK	East Anglia	MUT	IRT-2	ST	IRT-2 if 1 mutation or IRT-1 > 99.9th centile
I	Veneto Trentino Alto-Adige	MUT MP	IRT-2	ST	IRT-2 if MUT and MP-tive and IRT-1 twice the cutoff

1st tier is always IRT (IRT-1).
Abbreviations: IRT-2 = IRT resampling; MUT = genetic analysis; MP = meconium proteins; ST = sweat test; NA = not available; PS = pilot study.
More than one test per tier is considered if tests are performed at the same time.

3.1.2. Practical issues

In all but one programme, the CF NBS protocol was integrated into the current blood spot screening programme (median number of conditions screened, 4 (3–5)). Consent for CF screening was reported in 11/26 programmes (written in three). Four blood spots (4–6) were collected in most programmes and approximately 1% of these samples were inadequate (0.5–1.53). Blood spot samples were collected from day 3 (3–5) to day 5 (4–7) and delivered promptly to the screening laboratory in most cases (1–3 days) but in one centre, this took longer than 6 days. In the majority of programmes, nurses or midwives were responsible for obtaining blood spot samples. In three programmes, doctors were responsible and in one, parents were given the opportunity to obtain the sample themselves following simple instructions.

In programmes in which a second IRT sample (IRT-2) was taken (programmes that did not use DNA analysis on the first sample or IRT-2 to reduce number of sweat tests for infants with one mutation recognised), the test was organised from day 27 (21–28) to day 28 (27–30). Reasons given to parents for a second heel prick blood sample were "more blood needed to complete test" in 8 and "result unclear" in 2 centres. One centre reported an inadequate sample. Clearer information was given from 4 centres ("CF cannot be excluded" or "IRT-1 elevated").

3.1.3. Laboratory issues

In the majority of programmes [24], IRT measurement was undertaken using the Delfia™ technique (heterogeneous time resolved fluorometric assay). In 12 of these, an automated version was used (Autodelfia™). Two centres used a radio-labelled immunoassay (RIA) and two, an enzyme-linked immunospecific assay (ELISA). The national programme in France uses both the Delfia technique (12 laboratories) and RIA (11 laboratories). The cut-off for IRT-1 was set at 70 ng ml^{-1} (60–70) or above the centile of 99% (99–99.5). In 18 programmes, the laboratory would regularly review the IRT-1 cut-off by examining the population mean and spread. The median cut-off for IRT-2 was 50 ng ml^{-1} (40–56). All laboratories employed an internal quality control process, although nine were not involved in external quality control. Eight programmes store the blood spot card for an unlimited time, the remaining store them for a median of 7.5 years (3.6–18.8).

Table 2
Number of newborns screened and case recognition

Areas		Newborns screened per year	CF incidence (calculated)	Age at diagnosis (weeks)
Two-tier protocols				
I	Liguria	11,000	1/4400	8–9
Three-tier protocols				
CZ	Western Czech republic *	45,500	1/9100	4–6
UK	Wales	32,500	1/2700	<4
UK	Northern Ireland	23,000	1/2850	4–6
UK	Leeds Halifax Jersey	11,000	1/2750	3–6
I	Emilia Romagna	33,000	1/4700	8–9
I	Calabria	16,000	–	6–9
I	Sardinia	14,000	–	17
I	Lombardy	92,000	1/4600	3–5
I	Marche	13,000	1/5200	8–9
I	Tuscany	30,000	1/3500	6
I	Piedmont	37,000	1/2650	6
I	Lazio 1	28,500	1/3150	NA
I	Lazio 2, Umbria	33,000	NA	NA
I	Western Sicily	20,000	1/2500	6
A	Austria	80,000	1/3500	5–6
SP	Catalunya	62,500	1/5700	7–10
SP	Castilla-Leon	18,000	1/4000	3–12
SP	Galice	21,000	1/10,500	4–5
Four-tier protocols				
F	France	800,000	1/4700	5
PL	Poland	90,000	1/5000	4–6
UK	South Yorkshire East Midlands	55,000	1/2450	NA
UK	Scotland	54,000	1/2700	From 3 upwards
UK	Northamptonshire	8000	1/2250	3–8
UK	East Anglia	25,000	1/2800	3–6
I	Veneto Trentino Alto-Adige	52,000	1/4150	3–6

* Figures collected after the questionnaire circulation, not included in data analysis.

Nineteen of the 26 programmes incorporated DNA analysis as the second tier of the CF screening protocol. Median number of mutations examined was 31 (30–31), in most cases by oligonucleotide ligation assay (OLA) [14], although a range of molecular biology techniques were used (commercially available kits, Innogenetics™ (4) and Elucigene™ (3); denaturing gradient gel electrophoresis (DGGE) (1); in-house kit (5) and sequencing in one pilot study). It was reported that these panels covered a median of 82% (76.8–86.3) of *CFTR* mutations in the screened population.

3.1.4. Processing a positive screening result

The majority of centres reported that they would not undertake a sweat test until day 28 (14–30) of life, but in some infants as young as 1 week were referred. Eighteen laboratories measured sweat chloride concentration, 3 measured conductivity, 3 measured both. Generally one laboratory undertook these measurements for each programme (1–3) and the median number of sweat tests at each laboratory was 300 (150–400) per year (including sweat tests not related to NBS). A collection of 55 mg (50–94) was the accepted minimum weight of sweat in centres measuring chloride concentration. A sweat chloride between 40 (30–40) and 60 (60–60) mmol l^{-1} was considered a borderline sweat chloride result (7 considered 30 to be the lower limit). In 13 programmes, a sweat test was undertaken even if two CF associated mutations were recognised (in four, never, and in six, sometimes). In cases where a diagnosis was uncertain following sweat test a variety of strategies were reported including no further action [2], repeat sweat test [6], extended DNA analysis [7], clinical investigations (stool analysis for malabsorption [6], chest radiograph [1], respiratory cultures [6]) and nasal potential difference measurement [1]. A positive diagnosis was communicated in person to the parents and in most cases by a specialist in CF (in some cases by a General Paediatrician [6] or a Geneticist [2]). All programmes that involved DNA analysis [19] informed families of identification of carrier status and the majority of programmes [21] referred parents of identified CF infants for genetic counselling.

Clinical referral following a positive screen was generally the following day (0–2) and in most cases to a specialist CF centre [22]. In 11 centres, this would involve an admission for clinical evaluation. Eighteen centres provided information regarding segregation, which was undertaken in 8/18 with separate days for clinic visits. Three centres had strict inpatient segregation with separate rooms, face masks for consultations and no mixing of patients.

3.2. Multivariate analysis

Age at diagnosis showed a significant inverse correlation with the timing of IRT-1 ($p<0.001$, Fig. 2) and with the cut-off value for IRT-2 ($p<0.01$, Fig. 3). A lower age at diagnosis was also associated with DNA analysis as a second tier test ($p<0.015$, Fig. 4) and use of external quality control for the IRT assay ($p<0.015$, Fig. 5). A higher incidence of CF

Table 3
Combined results (presented as median and interquartile ranges)

Years of screening	7 (3.6–18.8)
Screened per year	30,000 (18,000–54,000)
IRT +ive per year	295 (148–825)
Sweat tests per year (as part of programme)	70 (20–129)
Number of CF cases per year	9 (5–14)
Carriers per year	17 (13–25)
False negatives since start of programme	2 (1–5)
False negatives per year	0.3
False negatives reported?	Yes: 15 (57.7%) No: 5 (19.2%) Not always: 2 (7.7%)
Causes of false negatives	First IRT low: 17 (65.4%) No mutations recognised: 1 (3.9%) Second IRT low: 4 (15.4%) Negative sweat test: 1 (3.9%)

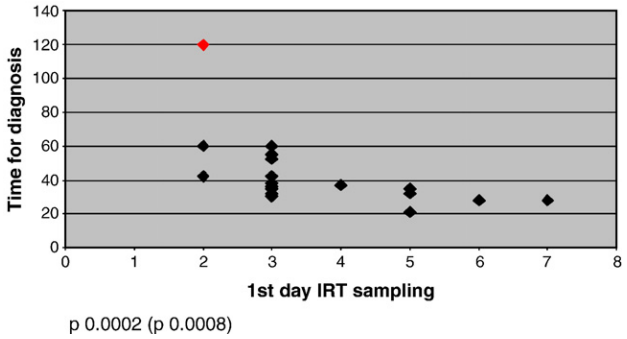


Fig. 2. Relationship between time to diagnosis and day on which IRT-1 is taken.

detected by the NBS programme was associated with an earlier IRT-1 ($p<0.03$) and a lower cut-off for IRT-2 ($p<0.02$).

4. Discussion

This is the first detailed survey of NBS for CF in Europe. The results demonstrate varied practice across the continent with protocols ranging from a sweat test following raised IRT-1 to those involving three or four tiers of testing. This variability reflects a number of issues (1) the complexity of NBS for CF, (2) geographical considerations, and (3) local circumstances. For example, in a geographically small region with limited molecular genetics and a good sweat test service, there might be an argument to moving straight to sweat test. However, in a geographically large region, strategies to reduce the number of sweat tests (for example, repeating the IRT on infants with one *CFTR* mutation recognised) may be appropriate [17].

The aim of newborn screening is to recognise index cases whilst causing minimal anxiety to the general public. This is particularly pertinent to CF screening where the weight of evidence that NBS improves long-term outcome is not as clear as for other conditions such as phenylketonuria [18]. Early nutritional benefits from screening do not appear to be maintained and a recent report suggested poorer chest radiograph appearance in screened children associated with an earlier acquisition of *Pseudomonas aeruginosa* lung infection [19–21]. However, our expectations for infants

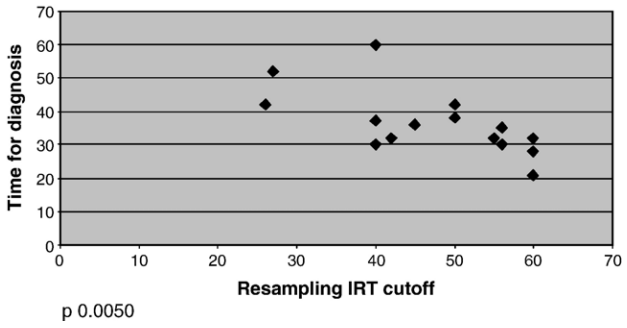


Fig. 3. Relationship between time to diagnosis and the cut-off value set for the resampling IRT (IRT-2).

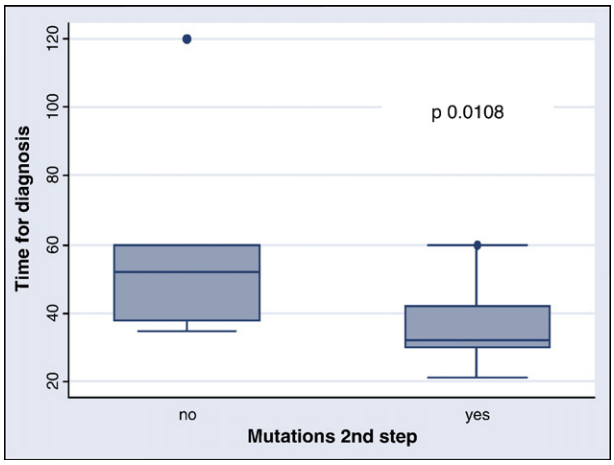


Fig. 4. Relationship between the time to diagnosis and the incorporation of DNA analysis into the second tier of testing.

born with CF have changed over the past two decades and the recent report of improved cognitive function in children with better vitamin E levels at diagnosis is of particular relevance [22,23]. In addition, NBS reduces the exposure of families to what is often a long and stressful diagnostic journey and gives couples the opportunity to make informed reproductive decisions [24–26]. The general consensus is that there is sufficient evidence to support NBS for CF and following systematic review, the US Center for Disease Control and Prevention concluded, "the health benefits to children with CF outweigh the risk of harm and justify screening for CF" [27]. Findings from this survey may aid regions or countries in their development and implementation of NBS for CF.

The heterogeneous nature of CF means that there will never be a perfect NBS protocol. Even incorporation of DNA analysis into a protocol does not remove the often-complicated clinical interface between the screening programme and the eventual diagnosis. The majority of programmes (over 70%) used DNA analysis as a second step

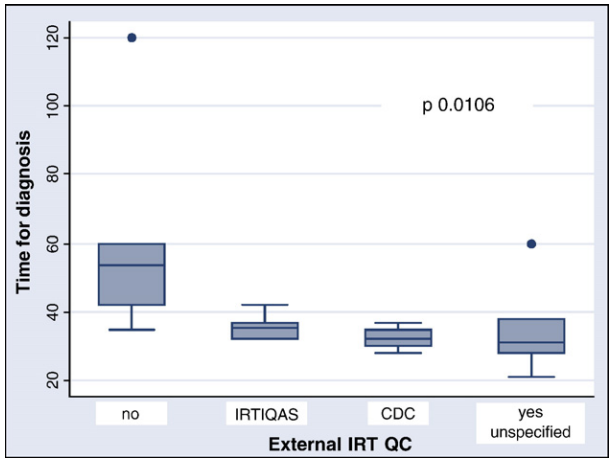


Fig. 5. Relationship between time to diagnosis and the use of an IRT quality assurance system.

following a raised IRT-1. Sometimes, this was just for the commonest CF causing mutation, phe508del, although in most cases a panel of 31 mutations was employed. Evidence from the screening programme in Massachusetts suggests that increasing the number of mutations on the screening panel does not necessarily improve case recognition but does result in increased carrier recognition [15]. However, in areas with a low frequency of phe508del (for example, South Italy), these data may not apply. Individual areas need to assess the implication of restricting the number of mutations that are included in the second tier panel on case recognition. In one programme *CFTR* gene sequencing was used, but only as a pilot study. It is difficult to envisage that this could represent a practical option for new programmes with available technology. Carrier recognition is a potential source of stress to families [14,28]. All centres employing DNA analysis reported carrier status to families.

It is clear that NBS for CF can be incorporated into ongoing NBS programmes although a significant frequency of inadequate samples was reported (median 1%). How these repeat samples are processed was not determined and it is not clear from this survey what increase in inadequate samples results from adding CF NBS to an existing programme. IRT analysis requires a better quality of blood spot sample than conventional NBS tests for phenylketonuria and congenital hypothyroidism and training in this area is imperative for regions implementing NBS for CF. In the majority of programmes, blood spot samples were collected by midwives or nurses. A second IRT sample was taken in the fourth week of life in programmes without DNA analysis or in some programmes with DNA analysis to reduce the number of sweat test requests on infants with only one *CFTR* mutation recognised (families of infants with one mutation and IRT-2 below the cut-off are given information regarding the carrier status and the low risk for CF). IRT was generally measured using the Delfia™ or Autodelfia™ techniques, with a median cut-off of 70 ng ml^{-1} for IRT-1 and 50 ng ml^{-1} for IRT-2. Most programmes reported using the top 1 or 0.5% of infants for the IRT-1 cut-off, adjusting this figure by continuously monitoring the population mean for the laboratory. All laboratories used internal quality control, but the present lack of a robust external quality control is highlighted by the fact that nine programmes were not involved in any form of external quality assurance scheme [29].

Sweat testing is a key component of the NBS protocols for CF. In the majority of programmes, a sweat test was undertaken even if two CF associated mutations were recognised. The identification of a physiological abnormality not only supports the molecular genetics but may also help the family come to terms with this diagnosis. A sweat chloride of 40–60 mmol l^{-1} was considered equivocal in most centres, although 30 was the lower limit in seven. There were a variety of responses as to how these infants should be subsequently investigated. The majority of programmes reported that sweat tests were being undertaken in one centre only; however, in geographically diverse regions this may not be possible and strategies to reduce sweat test numbers (such as a second IRT when one mutation is recognised) may be valid. The Association of Clinical Biochemists in the UK recently produced a consensus document recommending that 50 sweat tests was the minimum a centre should be undertaking in a year to maintain minimal experience to achieve adequate standards (http://www.acb.org.uk/site/guidelines.asp). Most centres continue to measure sweat chloride concentration (considered the "gold standard" [30]) however obsolete apparatus for the filter paper method of sweat collection may require centres to switch to capillary methods of sweat collection.

It is imperative that regions developing NBS for CF have clearly defined clinical referral pathways to maximise the positive impact of screening and minimise negative outcomes. Clear clinical referral pathways were evident in some programmes, most often through CF specialist services. It was not clear from the responses to these questionnaires that such pathways were always in place; however, this may reflect the fact that respondents were generally based in the screening labs and not in CF clinics. However, a key component of NBS for CF is the interface between the positive result and the subsequent specialist care that these infants need. It is evident from this survey that clear pathways are needed with respect to processing a positive result.

The programmes recorded in this survey cover geographically distinct regions with differing gene frequencies, making comparison of performance inappropriate. In addition, the programmes range in size from large national to small regional and whilst some are well established, others are in a pilot stage of development. However, some conclusions can be drawn regarding factors that might impact on performance, as determined by the average age of diagnosis. Some clear relationships are evident from multivariate analysis, most notably that an earlier collection of IRT-1 is associated with an older age of diagnosis. It is difficult to ascribe cause and effect to these relationships, which may reflect the underlying medical system and geographical variables. However, the higher incidence of CF in areas with earlier IRT-1 measurements merits further investigation. This relationship is supported by the finding that earlier IRT-1 is associated with an increased number of sweat tests (possibly leading to some delay). This relationship is maintained after removal of an outlier programme that reported an average age at diagnosis of 4 months (Fig. 2) and the French screening programme, which is the largest reported and performs a high number of sweat tests. DNA analysis as the second tier of testing was related to an earlier age of diagnosis. Previous reports have suggested improved screening performance with DNA analysis, but there are no randomised controlled trial data to support this [17,31].

This survey collected data from 26 programmes screening approximately 1,600,000 newborns each year in Europe for CF and identifying over 400 affected infants (a median

incidence of approximately 1 in 3500). Significant variation in NBS protocols exists, to some degree reflecting significant geographical differences. As the evidence for NBS for CF increases further regional and national programmes will become established. These data may guide the implementation of such programmes and will form the basis of consensus guidelines from the European CF Society Working Group on newborn screening.

Acknowledgments

Neonatal Screening in Czech Republic is supported by IGA MZCR 1A/8236-3, 00000064203 (6112) and CEPS.

References

[1] Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989;245 (4922):1066–73.
[2] Bobadilla JL, Macek Jr M, Fine JP, Farrell PM. Cystic fibrosis: a worldwide analysis of CFTR mutations—correlation with incidence data and application to screening. Hum Mutat 2002;19(6):575–606.
[3] Pilewski JM, Frizzell RA. Role of CFTR in airway disease. Physiol Rev 1999;79(1 Suppl):S215–55.
[4] Andersen D. Cystic fibrosis of the pancreas and its relation to celiac disease: a clinical and pathological study. Am J Dis Child 1938;56:135–42.
[5] Davis PB. Cystic fibrosis since 1938. Am J Respir Crit Care Med 2005;173(5):475–82.
[6] Doull IJ, Ryley HC, Weller P, Goodchild MC. Cystic fibrosis-related deaths in infancy and the effect of newborn screening. Pediatr Pulmonol 2001;31(5):363–6.
[7] Mastella G, Zanolla L, Castellani C, Altieri S, Furnari M, Giglio L, et al. Neonatal screening for cystic fibrosis: long-term clinical balance. Pancreatology 2001;1(5):531–7.
[8] Crossley JR, Elliott RB, Smith PA. Dried-blood spot screening for cystic fibrosis in the newborn. Lancet 1979;1(8114):472–4.
[9] Wesley AW, Smith PA, Elliott RB. Experience with neonatal screening for cystic fibrosis in New Zealand using measurement of immunoreactive trypsinogen. Aust Paediatr J 1989;25(3):151–5.
[10] Rock MJ, Mischler EH, Farrell PM, Wei LJ, Bruns WT, Hassemer DJ, et al. Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. Pediatrics 1990;85(6):1001–7.
[11] Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW. Neonatal screening for cystic fibrosis: addition of molecular diagnostics to increase specificity. Biochem Med Metab Biol 1993;49(2):200–11.
[12] Gregg RG, Wilfond BS, Farrell PM, Laxova A, Hassemer D, Mischler EH. Application of DNA analysis in a population-screening program for neonatal diagnosis of cystic fibrosis (CF): comparison of screening protocols. Am J Hum Genet 1993;52(3):616–26.
[13] Larsen J, Campbell S, Faragher EB, Gotz M, Eichler I, Waldherr S, et al. Cystic fibrosis screening in neonates-measurement of immunoreactive trypsin and direct genotype analysis for delta F508 mutation. Eur J Pediatr 1994;153(8):569–73.
[14] Parsons EP, Clarke AJ, Bradley DM. Implications of carrier identification in newborn screening for cystic fibrosis. Arch Dis Child Fetal Neonatal Ed 2003;88(6):F467–71.
[15] Comeau AM, Parad RB, Dorkin HL, Dovey M, Gerstle R, Haver K, et al. Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. Pediatrics 2004;113(6):1573–81.
[16] Parad RB, Comeau AM, Dorkin HL, Dovey M, Gerstle R, Martin T, et al. Sweat testing infants detected by cystic fibrosis newborn screening. J Pediatr 2005;147(3 Suppl):S69–72.
[17] Pollitt RJ, Dalton A, Evans S, Hughes HN, Curtis D. Neonatal screening for cystic fibrosis in the Trent region (UK): two-stage immunoreactive trypsin screening compared with a three-stage protocol with DNA analysis as an intermediate step. J Med Screen 1997;4(1):23–8.
[18] Wilfond BS, Parad RB, Fost N. Balancing benefits and risks for cystic fibrosis newborn screening: implications for policy decisions. J Pediatr 2005;147(3 Suppl):S109–13.
[19] Chatfield S, Owen G, Ryley HC, Williams J, Alfaham M, Goodchild MC, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. Arch Dis Child 1991;66(1 Spec No):29–33.
[20] Farrell PM, Kosorok MR, Laxova A, Shen G, Koscik RE, Bruns WT, et al. Nutritional benefits of neonatal screening for cystic fibrosis. Wisconsin Cystic Fibrosis Neonatal Screening Study Group. N Engl J Med 1997;337(14):963–9.
[21] Farrell PM, Li Z, Kosorok MR, Laxova A, Green CG, Collins J, et al. Bronchopulmonary disease in children with cystic fibrosis after early or delayed diagnosis. Am J Respir Crit Care Med 2003;168(9):1100–8.
[22] Koscik RL, Farrell PM, Kosorok MR, Zaremba KM, Laxova A, Lai HC, et al. Cognitive function of children with cystic fibrosis: deleterious effect of early malnutrition. Pediatrics 2004;113 (6):1549–58.
[23] Koscik RL, Lai HJ, Laxova A, Zaremba KM, Kosorok MR, Douglas JA, et al. Preventing early, prolonged vitamin E deficiency: an opportunity for better cognitive outcomes via early diagnosis through neonatal screening. J Pediatr 2005;147(3 Suppl):S51–6.
[24] Merelle ME, Huisman J, Alderden-van der Vecht A, Taat F, Bezemer D, Griffioen RW, et al. Early versus late diagnosis: psychological impact on parents of children with cystic fibrosis. Pediatrics 2003;111 (2):346–50.
[25] Campbell III PW, White TB. Newborn screening for cystic fibrosis: an opportunity to improve care and outcomes. J Pediatr 2005;147(3 Suppl):S2–5.
[26] Kharrazi M, Kharrazi LD. Delayed diagnosis of cystic fibrosis and the family perspective. J Pediatr 2005;147(3 Suppl):S21–5.
[27] Farrell PM, Lai HJ, Li Z, Kosorok MR, Laxova A, Green CG, et al. Evidence on improved outcomes with early diagnosis of cystic fibrosis through neonatal screening: enough is enough! J Pediatr 2005;147(3 Suppl):S30–6.
[28] Baroni MA, Anderson YE, Mischler E. Cystic fibrosis newborn screening: impact of early screening results on parenting stress. Pediatr Nurs 1997;23(2):143–51.
[29] Heeley ME, Travert G, Ferre C, Lemonnier F. The international quality assurance program (IRTIQAS) for the assay of immunoreactive trypsin in dried blood spots. Pediatr Pulmonol Suppl 1991;7:72–5.
[30] Baumer JH. Evidence based guidelines for the performance of the sweat test for the investigation of cystic fibrosis in the UK. Arch Dis Child 2003;88(12):1126–7.
[31] Gregg RG, Simantel A, Farrell PM, Koscik R, Kosorok MR, Laxova A, et al. Newborn screening for cystic fibrosis in Wisconsin: comparison of biochemical and molecular methods. Pediatrics 1997;99(6):819–24.

Journal of Cystic Fibrosis 8 (2009) 71 – 78

Journal of Cystic Fibrosis

www.elsevier.com/locate/jcf

A European consensus for the evaluation and management of infants with an equivocal diagnosis following newborn screening for cystic fibrosis

S.J. Mayell [a], A. Munck [b], J.V. Craig [a], I. Sermet [c], K.G. Brownlee [d], M.J. Schwarz [e], C. Castellani [f], K.W. Southern [a,*]

On behalf of the European Cystic Fibrosis Society Neonatal Screening Working Group [1]

* Corresponding author. Institute of Child Health, University of Liverpool, Royal Liverpool Children's Hospital, Eaton Road, Alder Hey, Liverpool, L12 2AP, United Kingdom. Tel.: +44 151 2525054; fax: +44 151 2525456.

E-mail address: kwsouth@liv.ac.uk (K.W. Southern).

[1] Cystic Fibrosis Specialists who contributed to the consensus:

Balascakova Miroslava	Institute of Biology and Medical Genetics, Charles University Prague-2, Czech Republic
Barben Juerg	Children's Hospital, St Gallen, Switzerland
Gabriel Bellon	Hopital Debrousse, Lyon, France
Brownlee Keith	Regional Paediatric CF Unit, St James University Hospital, Leeds, United Kingdom
Burrows Elinor	Royal Liverpool Children's Hospital, Liverpool, United Kingdom
Bush Andrew	Royal Brompton Hospital, London, United Kingdom
Castellani Carlo	Cystic Fibrosis Center, Azienda Ospedaliera Verona, Piazzale Stefani 1, 37126 Verona, Italy
Corbetta Carlo	Laboratorio di Riferimento regionale, Ospedale dei Bambini "V. Buzzi", Milano, Italy
Dankert-Roelse Jeannette	Atrium Medical Centre, Department of Pediatrics, Heerlen, The Netherlands
De Boeck Kristiane	Department of Pediatric Pulmonology, University Hospital, Leuven, Belgium
Desai Maya	Birmingham Children's Hospital, Birmingham, United Kingdom
Dodge John	University of Wales, Swansea, United Kingdom
Doull Iolo	Children's Hospital for Wales, Cardiff, United Kingdom
Eichler Irmgard	Cystic Fibrosis Center, University Children's Hospital, Vienna, Austria
Green Anne	Birmingham Children's Hospital, Birmingham, United Kingdom
Huet Frederic	Hôpital d'enfants, Dijon and AFDPHE, Paris, France
Holubova Andrea	Institute of Biology and Medical Genetics, Charles University Prague-2, Czech Republic
Iapichino Luciana	Centro Regionale Fibrosi Cistica, Palermo, Italy
Lebecque Patrick	Cliniques St Luc, Université Catholique de Louvain, Brussels, Belgium
Macek Milan	Institute of Biology and Medical Genetics, Charles University Prague-2, Czech Republic
Melotti Paola	Cystic Fibrosis Center, Azienda Ospedaliera di Verona, Verona, Italy
Munck Anne	Hôpital Robert Debré & AFDPHE, Paris, France
Padoan Rita	Cystic Fibrosis Service, Spedali Civili Brescia, Italy
Quattrucci Serena	Department of Paediatrics, University of Rome, Rome, Italy
Reid Alastair	Royal Belfast Hospital for Sick Children, Belfast, UK
Renner Sabine	Cystic Fibrosis Center, University Children's Hospital, Vienna, Austria
Roussey Michel	Hopital Sud, Rennes and AFDPHE, Paris, France
Satish Rao	Birmingham Children's Hospital, Birmingham, United Kingdom
Sands Dorota	Institute of Mother and Child, Warsaw, Poland
Seia Manuela	Laboratorio Genetica Molecolare, Fondazione Policlinico, Mangiagalli, Regina Elena, Milano, Italy
Sermet Isabelle	Hopital Necker Enfants Malades, Paris, France
Skalicka Veronika	Cystic Fibrosis Centre Prague, University Hospital Motol, Prague, Czech Republic
Southern Kevin	University of Liverpool, Royal Liverpool Children's Hospital, Liverpool, United Kingdom
Schwarz Martin	Regional Molecular Genetics Laboratory, St Mary's Hospital, Manchester, United Kingdom
Taylor Christopher	Sheffield Children's Hospital, United Kingdom
Taccetti Giovanni	Cystic Fibrosis Centre, Meyer Hospital, Florence, Italy
Tiddens Harm	Dept. Pediatric Pulmonology , Erasmus MC/Sophia children's Hospital, Rotterdam, Netherlands
Tümmler Burkhard	Klinische Forschergruppe, Medizinische Hochschule, Hannover, Germany
Vavrova Vera	Department of Pediatrics, Charles Univ. Prague-2, School of Medicine, Czech Republic
Votava Felix	Department of Pediatrics, Charles Univ. Prague-3, School of Medicine, Czech Republic
Weller Peter	Birmingham Children's Hospital, Birmingham, United Kingdom
Wilschanski Michael	Hadassah University Hospital, Jerusalem, Israel

1569-1993/$ - see front matter.
doi:10.1016/j.jcf.2008.09.005

[a] University of Liverpool, United Kingdom
[b] Hopital Robert Debre, AP-HP and AFDPHE, Paris, France
[c] Hopital Necker Enfants Malades, Paris, France
[d] Regional CF Unit, St James's University Hospital, Leeds, United Kingdom
[e] Regional Molecular Genetics Service, St Mary's Hospital, Manchester, United Kingdom
[f] Cystic Fibrosis Center, Azienda Ospedaliera, Verona, Italy

Received 2 June 2008; received in revised form 16 September 2008; accepted 17 September 2008
Available online 28 October 2008

Abstract

Screening newborns for cystic fibrosis (CF) is considered to be an ethical undertaking in regions with a significant incidence of the condition. Current screening protocols result in recognition of infants with an equivocal diagnosis. A survey of European practice suggested inconsistencies in the evaluation and management of these infants.

We have undertaken a consensus process using a modified Delphi method. This has enabled input of CF specialists from a wide geographical area to a rigorous process that has provided a clear pathway to a consensus statement. A core group produced 21 statements, which were modified over a series of three rounds (including a meeting arranged at the European CF Conference). A final document of 19 statements was produced, all of which achieved a satisfactory level of consensus. The statements cover four themes; sweat testing, further assessments and investigations, review arrangements and database.

This consensus document will provide guidance to CF specialists with established screening programmes and those who are in the process of implementing newborn screening in their region.

Keywords: Cystic fibrosis; Neonatal screening; Management; Sweat test; Gene analysis

1. Introduction

There is a good agreement that screening newborns for cystic fibrosis (CF) is a valid and ethical undertaking in regions such as Europe with a significant incidence of the condition [1]. Protocols for screening rely on the recognition that infants with CF have a high level of immuno-reactive trypsinogen (IRT) in their blood in the first week of life [2]. This test is sensitive but has poor specificity and therefore a second tier of investigations is necessary to identify those infants most at risk of CF [3]. In most newborn screening (NBS) protocols, this involves examining for common mutations of the *Cystic Fibrosis Transmembrane Conductance Regulator* (*CFTR*) gene which are associated with CF [4]. Such protocols will result in the recognition of carrier infants and data from well-established NBS programmes suggest that more carriers are recognised by NBS than might be expected for the incidence of CF in that population [5]. It has been concluded that this reflects the fact that carriers have a higher IRT than the non-carrier population and this is supported by data from population studies [6]. In most NBS protocols infants recognised as putative carriers have further assessment including a sweat test to exclude CF [4].

Another significant challenge of NBS for CF is the recognition of infants with an equivocal diagnosis [4,7]. This reflects the heterogeneous nature of the condition and poses a challenge to CF teams. Infants with one recognised *CFTR* mutation or persistent hypertrypsinaemia may have an intermediate sweat test result or an infant may be recognised with two *CFTR* mutations, one or both of which have unclear phenotypic consequences [7]. There is significant variability in the evaluation and management of these infants with an equivocal diagnosis [3,8]. We have used a modified Delphi method to form a consensus on the evaluation and management of these infants [9]. The Delphi method is a recognised technique that provides a formal strategy to gather expert opinion and form a consensus when there is a lack of high quality evidence on which to base practice. The method facilitates the inclusion of experts from a geographically disperse region and establishes a framework which makes it possible to clearly trace back how the group came to a decision.

2. Methods

Twenty-one preliminary statements were composed by a core group of experts in the field (CC, AM and KWS), taking into account the results of a survey of European practice [4]. The statements covered four thematic areas; sweat testing, further assessments and investigations, review arrangements and database. Two European Cystic Fibrosis Society (ECFS) working groups (the Diagnostic Network (ECFDN) and the Neonatal Screening Working Group) were approached and their members, which include clinicians, biochemists and geneticists, contacted by email. Additional invitations were made to increase multidisciplinary input. Consensus was determined *a priori* to be 80% of ratings providing agreement with the statement (considered sufficient for this type of study) [9].

For round one, specialists were asked to rate their opinion of each statement by choosing one of three options; 1) agree, 2) could agree if reworded or 3) disagree. Specialists choosing options 2) or 3) were asked for comments and also suggestions for alternative or modified statements.

After round one, statements not achieving consensus (or achieving consensus with provisos) were modified by the core

group, taking into account the comments and suggestions made by respondents. These modified statements were then circulated in round two together with the original statements, the group ratings from round one and a summary of comments. Individuals who replied to round one were included in round two.

Following round two, statements not achieving consensus were presented and discussed in the European Cystic Fibrosis Neonatal Screening Working Group meeting at the 30th European Cystic Fibrosis Society Conference, Belek, Turkey in 2007. This meeting involved members of the consensus group (although not all) and other CF Specialists. The entire consensus document was presented, but the focus of the meeting was on statements that had not achieved consensus. This facilitated an open discussion that enabled deeper reflection on the issues around these statements. Statements were subsequently revised by the core group taking into account the discussions at this meeting and comments already received from round two. The revised statements were again circulated in round three to all respondents together with the original statements, the level of agreement from round two and respondents' comments (Fig. 1). Four appendices were constructed to provide further information and background to the statements (Table 1) and cover a) Sweat Testing, b) Gene Testing, c) Clinical Features and d) Further Physiological Testing.

3. Results

3.1. Round one

Forty-one responses from specialists in 11 European countries were received for round one. A consensus of over 80% was achieved on twelve of 21 statements. A further five statements were approaching consensus with greater than 60% agreement. Four statements had poor level of agreement (<60%).

3.2. Round two

Nine statements not achieving consensus in round one and three that did with provisos were revised by the core group, following analysis of respondents' comments. Respondents were asked to rate their agreement with both the revised and modified statements. Thirty-eight responses were obtained following round two. A consensus of greater than 80% was achieved on a further ten statements. Consensus was not achieved in two.

3.3. Round three

The two statements not achieving consensus were discussed at the ECFS Screening Working Group Meeting. Taking into account this meeting and respondents comments, it was clear that modification and combination of four statements were required to obtain consensus, reducing the number of statements from twenty one to nineteen. Thirty-four responses were obtained to round three and consensus was achieved on all nineteen statements (Table 1). An algorithm was developed from the consensus statements (Fig. 2). Seven specialists who responded to round one did not respond to round three (17% attrition rate).

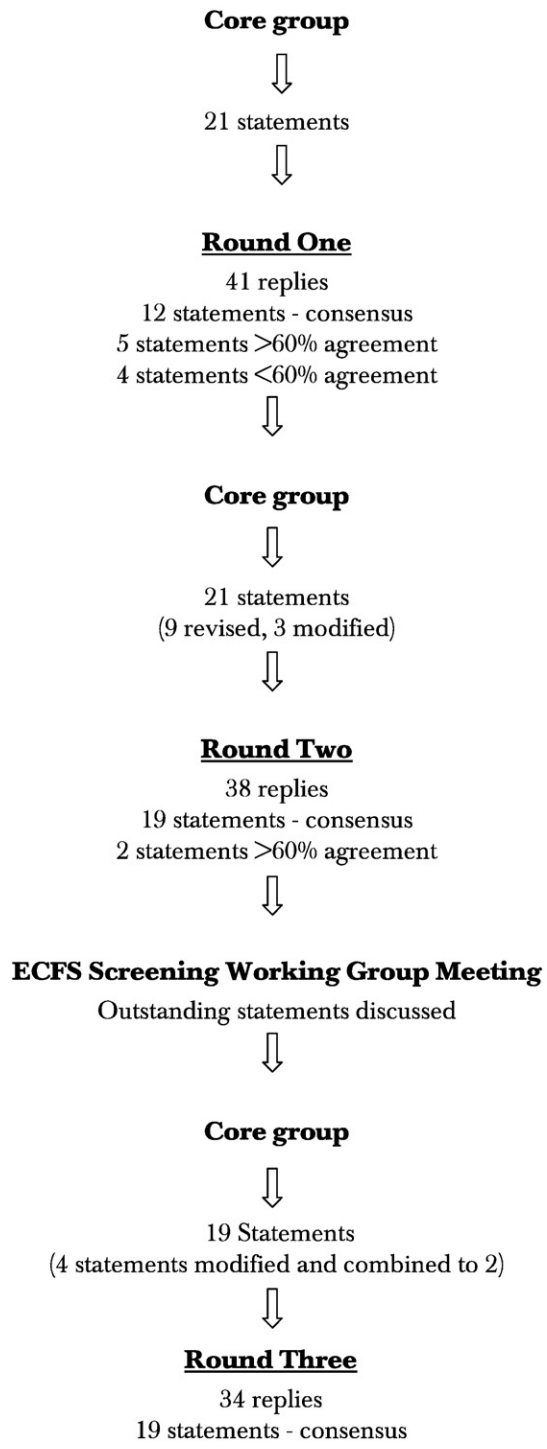


Fig. 1. The stages and outcomes of the modified Delphi method.

3.4. Specific issues and comments

3.4.1. Clinic size

Original statements for round one defined specialist CF clinic size greater than 100 patients and achieved consensus. Despite this, several respondents discussed the size of clinics

Table 1
Nineteen statements at the end of the consensus process

1	An infant with one or more raised IRT measurements and an equivocal sweat test (sweat $Cl^- \geq 30$ and <60 mmol L^{-1}) requires assessment and review in a specialist CF clinic (with >50 patients).
2	An infant with two *CFTR* gene changes (one of which has unclear clinical significance) and a normal sweat test requires assessment and review in a specialist CF clinic (>50 patients).
3	In these cases (infants from statements 1 and 2) a repeat sweat test should be undertaken in a centre with suitable experience (>150 sweat tests pa) of a validated technique for measuring sweat chloride (Appendix A).
4	Infants from statement 1, who have a normal repeat sweat test in an accredited centre (sweat $Cl^- <30$ mmol l^{-1}), do not require further clinical review (negative CF screening test).
5	Extended gene analysis must be undertaken in infants with two equivocal sweat tests and one or no *CFTR* mutations recognised (Appendix B).
6	Infants with one or more raised IRT measurements, one *CFTR* mutation and a normal sweat test ($Cl^- <30$) do not require extended gene analysis.
7	Infants with one or more raised IRT measurements, one *CFTR* mutation and a normal sweat test ($Cl^- <30$) do not require review in a CF clinic (negative CF screening test). Appropriate advice regarding carrier status should be given.
8	Infants with two equivocal sweat tests require detailed baseline assessment for respiratory disease (airways culture and chest radiograph). Further investigations may be indicated as determined by the clinical situation (for example, chest CT scan, and bronchoscopy).
9	Infants with two equivocal sweat tests require detailed baseline assessment for non-respiratory disease (fecal elastase). Other investigations as clinically indicated.
10	Infants with two equivocal sweat tests and any clinical evidence supportive of a CF diagnosis should have regular follow up in a CF specialist clinic (Appendix C).
11	Infants with two equivocal sweat tests, one or no CF causing mutations and no clinical evidence of CF should be *considered* for further investigation of a physiological defect in a centre with appropriate experience (Appendix D).
12	Infants with two equivocal sweat tests, one or no CF causing mutation, no clinical evidence of CF and evidence of ion transport defect on further testing should be followed up in a specialist CF clinic.
13	All infants with two equivocal sweat tests, one or no CF causing mutation, no clinical evidence of CF should be reviewed in a specialist CF clinic with a repeat sweat test at 6–12 months of age.
14	Infants with two *CFTR* gene changes (as statement 2) but a normal sweat test (at least one performed in a centre with adequate experience, as per statement 3) should have detailed clinical assessment (as 8 and 9).
15	Infants with two *CFTR* gene changes (as statement 2), a normal sweat test and any clinical evidence of CF should have regular follow up in a CF specialist clinic.
16	Infants with two *CFTR* gene changes (as statement 2), a normal sweat test and no clinical evidence of CF should be *considered* for further investigation of a physiological defect in a centre with appropriate experience (Appendix D).
17	Infants with two *CFTR* gene changes (as statement 2), a normal sweat test, no clinical evidence of CF but evidence of abnormal ion transport should have regular follow up in a specialist CF clinic.
18	Infants with two *CFTR* gene changes (as statement 2), a normal sweat test and no clinical evidence of CF should be reviewed in a specialist CF clinic with a repeat sweat test at 6–12 months.
19	Clinical and demographic information on all infants with an equivocal diagnosis should be entered onto a database or registry (pending consent from legal guardian).

(>50 or >100 patients) citing concerns that in some areas of Europe specialist CF clinics are traditionally smaller. Overall however it was felt beneficial to include a proposed number and the statements were modified to "specialist CF clinic (with >50 patients)", which again achieved consensus.

3.4.2. Sweat test

The sweat testing experience of the centre was suggested at greater than 150 sweat tests performed annually and this achieved consensus in round one (Statement 3). See also Appendix A.

3.4.3. Equivocal sweat test result

All infants considered by the guideline will have one or more raised IRT. When these infants subsequently have an equivocal sweat test result they require assessment and review in a specialist CF clinic with a repeat sweat test (Statement 1). If the sweat test in an accredited centre remains equivocal (chloride, 30–60 mmol L^{-1}) further investigation is required. This should include extended gene analysis if one or no *CFTR* mutations have been identified (Statement 5) and baseline assessment for respiratory and non-respiratory disease (Statements 8 and 9). If these infants show any clinical evidence supportive of the diagnosis of CF they should have regular follow up in a specialist CF clinic (Statement 10). If there is no clinical evidence of CF they should be reviewed in a specialist CF clinic with sweat test repeated between 6 and 12 months of age (Statement 13).

3.4.4. What constitutes a negative screening result?

Infants who have equivocal sweat test (chloride 30–60 mmol/l) that on repeat test is found to be normal (<30mmol/l) do not require further clinical review (Statement 4). This was agreed to be a negative screening result i.e. CF not suspected (85%, round one).

Infants who have one or more raised IRT measurements, one *CFTR* mutation and a normal sweat test have a negative screening test (Statement 7). This statement was approaching consensus in round one (71%). Following respondents comments, the statement was modified to advise discussion of carrier status and this achieved 92% consensus in round two.

3.4.5. Extended gene analysis

Extended gene analysis is indicated when an infant has had two equivocal sweat test results and only one or no *CFTR* mutation has been identified on the local common mutation panel (Statement 5, 85% consensus round one) (Appendix B). Extended gene analysis should not be performed in infants with one or more raised IRT, one *CFTR* mutation and a normal sweat test. These infants have a negative screening result (Statement 6, 80% consensus round one).

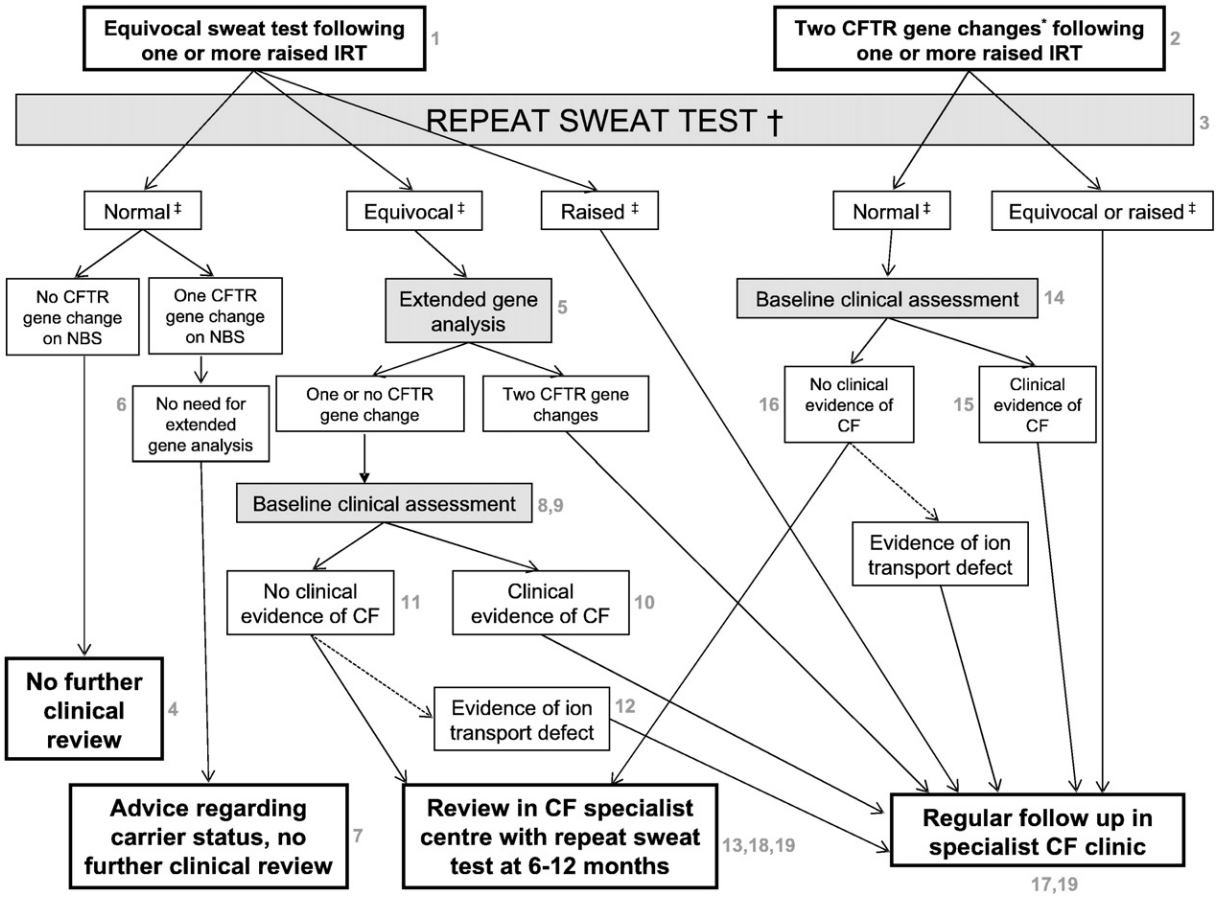


Fig. 2. The pathway of interventions that infants with an equivocal diagnosis may follow according to the results of this consensus process. The two distinct presentations of these infants (equivocal sweat test or two *CFTR* gene mutations of unclear clinical significance) represent the starting points at the top of the figure. The progress of the infant is then tracked following the repeat sweat test in an experienced centre. Subsequent interventions depend to some degree on the result of the repeat sweat test. The numbers indicate the consensus statement (Table 1) that corresponds to that part of the pathway. An important point to note is that infants who enter the pathway with an equivocal sweat test and then have a normal repeat sweat test do not require extended gene analysis or further clinical review (although some may require advice regarding carrier status). Clinical assessment for evidence of CF is considered important with respect to determining subsequent review arrangements (infants with any clinical evidence supporting a CF diagnosis should be seen in a specialist CF clinic). * One of which has unclear clinical significance. † In a centre with suitable experience (Appendix A). ‡ Normal = sweat Cl < 30 mmol l^{-1}, equivocal = sweat Cl ≥ 30 and < 60 mmol l^{-1}, raised = sweat Cl ≥ 60 mmol l^{-1}.

3.4.6. Baseline assessment for respiratory and non-respiratory disease

It was agreed that infants who had two equivocal sweat test results or who have two *CFTR* gene changes and a normal sweat test should have baseline clinical assessment (Statements 8, 9 and 14). In round one, baseline assessment for respiratory disease included CT scan and bronchoscopy and this statement achieved a poor level of agreement (41%). The statement was modified to include these investigations only when indicated by the clinical situation and 89% agreement was achieved in the second round (Statement 8). Similarly a significant number felt that some of the proposed non-respiratory investigations were unnecessary and this statement was changed accordingly (to faecal elastase with other investigations as clinically indicated). The statement achieved 92% agreement in round two (Statement 9).

3.4.7. Further investigation of physiological defect

A number of measurements of transepithelial salt transport exist that may help in investigating an equivocal diagnosis of CF. However, none have the face validity of sweat electrolyte measurement and are essentially extensions of research methodology. This was reflected in the variability of responses to statements concerning further electrophysiological investigation ranging from enthusiastic advocates to confirmed sceptics. It was agreed with 83% consensus in round two that infants who had two equivocal sweat tests and no clinical evidence of CF should be considered for further investigation of a physiological defect (Statement 11). If there is evidence of ion transport defect even in the absence of clinical evidence of CF with only one *CFTR* mutation these infants should be followed up in a specialist CF clinic (Statement 12).

It proved difficult to obtain a consensus regarding the subsequent management of infants who had no evidence of ion transport defect. In the absence of consensus, these infants are covered by Statement 13. This reflects a general anxiety that it would be inappropriate to exclude a diagnosis of CF on the basis of these measures alone, even when there is no clinical suspicion following baseline assessment.

3.4.8. Two CFTR gene mutations one or both of which have unclear clinical significance

The term *CFTR* gene change was used in the statements to highlight that these mutations have unclear phenotypic consequences. Subsequently, an ECFS consensus has been achieved on the use and interpretation of CF mutation analysis in clinical practice and has concluded that the term "mutation" should be used to define any molecular alteration in the DNA sequence of the *CFTR* gene [10]. Therefore for the purposes of this process the terms *CFTR* gene change and mutation are interchangeable. The dilemma remains that a number of frequently recognised *CFTR* gene mutations have unclear phenotypic consequence. The situation in CF NBS is further exacerbated by the fact that some *CFTR* mutations that are clearly "CF causing" can have little if any phenotypic consequence in the first years of life, related to the confounding impact of environment and other non-*CFTR* genes.

There was clear agreement that if these infants have any clinical evidence of CF they should have regular follow up in a specialist CF clinic, even with a normal sweat test result (Statement 15). If there is no clinical evidence of CF they should be considered for further investigation of ion transport defect (Statement 16) but regardless should be reviewed in a specialist CF clinic with sweat test repeated between 6 and 12 months of age (Statement 18). If there is evidence of abnormal ion transport these infants should have regular review in a specialist CF clinic (Statement 17).

4. Discussion

There is a good agreement that screening infants for CF is an ethical undertaking in regions with a significant incidence of the condition [1]. Unfortunately, current NBS protocols result in recognition of infants with an equivocal diagnosis of CF. To some degree this reflects the heterogeneity of the condition and the sensitivity of the IRT measurement. Infants with an equivocal diagnosis fall into two groups; those with intermediate sweat electrolytes (above the level expected for this age group), but no or one recognised *CFTR* gene mutation and those with two *CFTR* gene mutations, one or both of which have unclear long term phenotypic consequences.

Using a modified Delphi method, we have produced 19 statements that will act as a guide for CF teams in the evaluation and management of infants with an equivocal diagnosis following newborn screening. Strengths of this process have been the inclusive nature across a wide geographical area and a robust framework that enables clear identification of decision making pathways. From the consensus guideline, we have produced an algorithm to aid clinicians involved in CF NBS programmes. It was not the aim of this group to provide a diagnostic framework; rather to provide a pragmatic guideline for the management and evaluation of these infants. The end-points in the algorithm are therefore functional rather than categorical (Fig. 2).

Although the majority of statements attained consensus in the early phases of the process, some areas were more challenging. The group meeting was an essential part of the Delphi process to identify issues around these statements and achieve a final consensus. These guidelines can be used in established screening regions and in those with emerging programmes to guide the evaluation and management of this challenging clinical dilemma.

Acknowledgement

We thank Dr Phil Farrell for helpful comments on this manuscript.

Appendix A

Notes on sweat testing

1) Sweat collection in infants is challenging [11].
2) Sweat collection and analysis should be undertaken in a centre with adequate experience. CF physicians should be guided by national standards. If these are not available, consensus documents are available from the United Kingdom (http://acb.org.uk/docs/sweat.pdf) and North America [11]. These suggest that a laboratory should be undertaking at least 50 tests per year; however for infants with an initial equivocal result, the repeat sweat test should be done in a centre with more experience (>150 sweat tests per year, Statement 3).
3) Two equivocal sweat test results may be available on the same day that were undertaken in a suitably experienced centre, in which case the infant should proceed along the algorithm as described (Fig. 2).
4) Sweat chloride concentration remains the gold standard analytical measure to confirm a diagnosis of CF [11].
5) Sweat sodium should not be used [11].
6) Sweat conductivity may have a role in excluding a diagnosis of CF but does not have sufficient face validity in cases with an initial equivocal result [11].
7) Sweat electrolyte values fall over the first 4 weeks of life [12,13]. A sweat chloride value over 30 mmol L^{-1} should prompt clinical review and a repeat sweat test [11].

Appendix B

Notes on extended DNA analysis

1. Further investigation of these infants should be undertaken with close liaison with the local molecular genetics service. The extent of DNA analysis should reflect the clinical suspicion. Care should be taken in avoiding the situation where gene changes (mutations) are recognised with unclear phenotypic characterisation, although in most circumstances this will be unavoidable, particularly as laboratories move more quickly to comprehensive *CFTR* gene sequencing.
2. *CFTR* gene change is equivalent to *CFTR* mutation.
3. Further DNA analysis should be guided by the screening protocol (i.e., protocols that initially only examine for a limited panel of *CFTR* mutations would prompt further DNA analysis).
4. Infants recognised to be compound heterozygotes for R117H should have further characterisation of the poly T variant region (and TG repeats if found to be on a 5T background) [14,15]. Infants with R117H on a 7T background may have a normal or equivocal sweat test. Long term clinical outcome is

variable and the management of these infants requires specialist input (as per statements 2, 14, 15, 16, 17 and 18) [7,16].
5. Some *CFTR* mutations that are clearly "CF causing" (in particular, 3849+10 kb C>T) are associated with normal or equivocal sweat electrolyte values. Close liaison with the local molecular genetics service is needed to determine these infants.
6. Infants with persistent intermediate sweat electrolytes and clinical features (Appendix C) should have extensive DNA analysis after discussion with the local molecular genetics service.

Appendix C

Clinical features consistent with a diagnosis of CF following newborn screening

C.1. Respiratory

1. Symptoms
   a. Cough
   b. Wheeze
2. Clinical findings
   a. Over-expanded chest
   b. Crackles
   c. Wheeze
   d. Tachypnoeic
   e. Abnormal chest shape
3. Chest radiograph changes
   a. Overinflation
   b. Increased markings
   c. Areas of collapse or consolidation
4. Chest high resolution computerised tomogram (HRCT) changes
   a. Air trapping
   b. Early evidence of airway inflammation/bronchiectasis
5. Positive respiratory culture for characteristic CF pathogens
   a. Cough swab/plate
   b. Broncho-alveolar lavage

C.2. Non-respiratory

1. Clinical evidence of malabsorption
   a. Meconium ileus
   b. Poor weight gain
   c. Distended abdomen
   d. Loose offensive stool
   e. Poor head growth
   f. Rectal prolapse
2. Laboratory evidence of malabsorption
   a. Low fecal elastase (or chymotrypsin)
   b. Positive fat microscopy
   c. Low fat soluble vitamin levels
3. Radiological evidence of pancreatic disease
   a. Pancreatic calcification on Abdominal radiograph
   b. Pancreatic fibrosis on abdominal ultrasound scan
4. Liver disease
   a. Prolonged cholestatic jaundice
   b. Elevated liver enzymes (ALT/AST)
   c. Abnormal liver appearance on ultrasound scan
5. Salt loss
6. Absence of the vas deferens

Appendix D

Further physiological testing.

A number of electrophysiological techniques are available to demonstrate the salt transport defect that characterises CF. These are undertaken in specialist centres. None have the face validity of sweat testing or genotype analysis, but may provide useful additional information in equivocal cases. Some of these tests are particularly challenging in infants.

Test	Technical details	What it involves for the infant	Availability[a]
Nasal Potential Difference (PD)	Ion transport across airway epithelium can be assessed by measuring the baseline PD. The impact on the PD of perfusing different solutions and drugs provides further information to differentiate CF from non-CF recordings.	The exploring electrode is placed in the nose. A reference electrode is placed either subcutaneously or over abraded skin on the forearm. Solutions are perfused through the exploring electrode into the nose and can be swallowed.	Very few centres are able to undertake this measurement in infants although it is more widely available in older children and adults.
Intestinal Current Measurements (ICM)	A biopsy is mounted in the laboratory in a device (Ussing chamber) that enables measurement of transepithelial ion transport. Various aspects of ion transport can be examined.	Biopsy of rectal mucosa. This procedure is painless and well tolerated by young infants. Does not require general anaesthesia or sedation.	This technique requires a dedicated laboratory service with highly skilled technicians. Available in limited number of centres in Europe.
Small bowel biopsy	Similar measures of transepithelial transport processes can be undertaken in the laboratory on upper gastro-intestinal (GI) mucosal biopsies.	Upper GI biopsy; requires general anaesthesia in most cases.	Limited (only currently available in Sheffield, UK; contact Prof Chris Taylor).

[a]For details of centres in Europe that undertake appropriate electrophysiological investigations on infants, contact Dr Michael Wilschanski, chair of the European CF Society Diagnostic Network (michaelwil@hadassah.org.il).

References

[1] Farrell MP. Is newborn screening for cystic fibrosis a basic human right? J Cyst Fibros 2008;7:262–5.

[2] Crossley JR, Elliott RB, Smith PA. Dried-blood spot screening for cystic fibrosis in the newborn. Lancet Mar 3 1979;1(8114):472–4.

[3] Southern KW. Newborn screening for cystic fibrosis: the practical implications. J R Soc Med 2004;97(Suppl 44):57–9.

[4] Southern KW, Munck A, Pollitt R, Travert G, Zanolla L, Dankert-Roelse J, et al. A survey of newborn screening for cystic fibrosis in Europe. J Cyst Fibros Jan 2007;6(1):57–65.

[5] Massie RJ, Wilcken B, Van Asperen P, Dorney S, Gruca M, Wiley V, et al. Pancreatic function and extended mutation analysis in DeltaF508 heterozygous infants with an elevated immunoreactive trypsinogen but normal sweat electrolyte levels. J Pediatr Aug 2000;137(2):214–20.

[6] Castellani C, Picci L, Scarpa M, Dechecchi MC, Zanolla L, Assael BM, et al. Cystic fibrosis carriers have higher neonatal immunoreactive trypsinogen values than non-carriers. Am J Med Genet Jun 1 2005;135(2):142–4.

[7] Roussey M, Le Bihannic A, Scotet V, Audrezet MP, Blayau M, Dagorne M, et al. Neonatal screening of cystic fibrosis: diagnostic problems with CFTR mild mutations. J Inherit Metab Dis Aug 2007;30(4):613.

[8] Merelle ME, Nagelkerke AF, Lees CM, Dezateux C. Newborn screening for cystic fibrosis. Cochrane Database Syst Rev (Online) 2001(3):CD001402.

[9] Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess (Winchester, England). 1998;2(3):i–iv, 1–88.

[10] Castellani C, Cuppens H, Macek Jr M, Cassiman J, Kerem E, Durie P, et al. Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. J Cyst Fibros 2008;7:179–96.

[11] LeGrys VA, Yankaskas JR, Quittell LM, Marshall BC, Mogayzel Jr PJ. Diagnostic sweat testing: the Cystic Fibrosis Foundation guidelines. J Pediatr Jul 2007;151(1):85–9.

[12] Eng W, LeGrys VA, Schechter MS, Laughon MM, Barker PM. Sweat-testing in preterm and full-term infants less than 6 weeks of age. Pediatr Pulmonol 2005;40(1):64–7.

[13] Parad RB, Comeau AM, Dorkin HL, Dovey M, Gerstle R, Martin T, et al. Sweat testing infants detected by cystic fibrosis newborn screening. J Pediatr 2005;147(3):S69–72.

[14] Groman JD, Hefferon TW, Casals T, Bassas L, Estivill X, Des Georges M, et al. Variation in a repeat sequence determines whether a common variant of the cystic fibrosis transmembrane conductance regulator gene is pathogenic or benign. Am J Hum Genet Jan 2004;74(1):176–9.

[15] Cuppens H, Lin W, Jaspers M, Costes B, Teng H, Vankeerberghen A, et al. Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (Tg)m locus explains the partial penetrance of the T5 polymorphism as a disease mutation. J Clin Invest Jan 15 1998;101(2):487–96.

[16] Peckham D, Conway SP, Morton A, Jones A, Webb K. Delayed diagnosis of cystic fibrosis associated with R117H on a background of 7T polythymidine tract at intron 8. J Cyst Fibros Jan 2006;5(1):63–5.

RESEARCH ARTICLE

Human Mutation

OFFICIAL JOURNAL HGVS HUMAN GENOME VARIATION SOCIETY www.hgvs.org

Mutations in the Amiloride-Sensitive Epithelial Sodium Channel in Patients With Cystic Fibrosis-Like Disease

Abul Kalam Azad,[1] Robert Rauh,[2] François Vermeulen,[3] Martine Jaspers,[1] Judit Korbmacher,[2] Brigitte Boissier,[4] Laurence Bassinet,[4] Yann Fichou,[5] Marie des Georges,[6] Frauke Stanke,[7] Kris De Boeck,[3] Lieven Dupont,[8] Miroslava Balaščáková,[9] Lena Hjelte,[10] Patrick Lebecque,[11] Dragica Radojkovic,[1] Carlo Castellani,[12] Marianne Schwartz,[13] Manfred Stuhrmann,[14] Martin Schwarz,[15] Veronika Skalicka,[9] Isabelle de Monestrol,[10] Emmanuelle Girodon,[4] Claude Férec,[5] Mireille Claustres,[6] Burkhard Tümmler,[7] Jean-Jacques Cassiman,[1] Christoph Korbmacher,[2] and Harry Cuppens[1]*

[1]Center for Human Genetics, Katholieke Universiteit Leuven, Leuven, Belgium; [2]Institut für Zelluläre und Molekulare Physiologie, Universität Erlangen-Nürnberg, Erlangen, Germany; [3]Department of Pediatrics, UZ Gasthuisberg, Leuven, Belgium; [4]Service de Biochimie et Génétique, Hôpital Henri Mondor, Créteil, France; [5]Laboratoire de Génétique Moléculaire et d'Histocompatibilité, Centre Hospitalier Universitaire (CHU) de Brest, Brest, France; [6]Laboratoire de Génétique Moléculaire, Centre Hospitalier Universitaire (CHU) de Montpellier, Institut Universitaire de Recherche Clinique (IURC), Montpellier, France; [7]Klinische Forschergruppe Medizinische dHochschule, Hannover, Germany; [8]Department of Pulmonology, UZ Gasthuisberg, Leuven, Belgium; [9]Institute of Biology and Medical Genetics, Pneumology Clinic and Pediatrics Clinic, Charles University Prague, School of Medicine and University Hospital Motol, Prague, Czech Republic; [10]Stockholm Cystic Fibrosis (CF) Centre, Karolinska University Hospital, Huddinge, Sweden; [11]Pediatric Pulmonology and Cystic Fibrosis Unit, Cliniques St Luc, Université Catholique de Louvain, Brussels, Belgium; [12]Cystic Fibrosis Centre, Azienda Ospedaliera di Verona, Verona, Italy; [13]Department of Clinical Genetics, Rigshospitalet, Copenhagen, Denmark; [14]Institut für Humangenetik Medizinische Hochschule, Hannover, Germany; [15]Molecular Genetics Laboratory, Royal Manchester Children's Hospital, Manchester, United Kingdom

Communicated by Mark H. Paalman

Received 8 December 2008; accepted revised manuscript 18 February 2009.
Published online 3 March 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/humu.21011

ABSTRACT: We investigated whether mutations in the genes that code for the different subunits of the amiloride-sensitive epithelial sodium channel (ENaC) might result in cystic fibrosis (CF)-like disease. In a small fraction of the patients, the disease could be potentially explained by an *ENaC* mutation by a Mendelian mechanism, such as p.V114I and p.F61L in *SCNN1A*. More importantly, a more than three-fold significant increase in incidence of several rare *ENaC* polymorphisms was found in the patient group (30% vs. 9% in controls), indicating an involvement of *ENaC* in some patients by a polygenetic mechanism. Specifically, a significantly higher number of patients carried c.–55+5G>C or p.W493R in *SCNN1A* in the heterozygous state, with odds ratios (ORs) of 13.5 and 2.7, respectively.The p.W493R-*SCNN1A* polymorphism was even found to result in a four-fold more active ENaC channel when heterologously expressed in *Xenopus laevis* oocytes. About 1 in 975 individuals in the general population will be heterozygous for the hyperactive p.W493R-*SCNN1A* mutation and a cystic fibrosis transmembrane conductance regulator (*CFTR*) gene that results in very low amounts (0–10%) functional CFTR. These ENaC/CFTR genotypes may play a hitherto unrecognized role in lung diseases.
Hum Mutat 30:1093–1103, 2009. © 2009 Wiley-Liss, Inc.

KEY WORDS: CF; CFTR; ENaC; *SCNN1A*

Introduction

Loss-of-function mutations in both copies of the cystic fibrosis transmembrane conductance regulator (ATP-binding cassette subfamily C, member 7) (*CFTR*; MIM# 602421) gene of an individual cause cystic fibrosis (CF; MIM# 219700) [Kerem et al., 1989; Riordan et al., 1989]. *CFTR* encodes a cyclic adenosine monophosphate (cAMP)-regulated chloride channel [Rich et al., 1990; Riordan et al., 1989]. Apart from their involvement in classical CF, *CFTR* mutations also cause, or contribute to, the basic defect in congenital absence of the vas deferens (CAVD; MIM# 277180) [Chillón et al., 1995], disseminated bronchiectasis [Pignatti et al., 1995], chronic pancreatitis [Sharer et al., 1998], and so-called atypical CF. The latter refers to CF-like disease in which only one or a few organs are (mildly) affected. In the vast majority of classical CF patients both copies of the *CFTR* gene are mutated. However, in a minority of CF patients a mutation cannot be identified on both *CFTR* genes, i.e., in 1 to 2% of the CF patients of Northern Europe and in up to 8 to 10% in CF patients of Southern Europe (www.who.int/genomics/publications/en). In an even higher proportion of patients with CAVD, atypical CF, chronic pancreatitis, or disseminated bronchiectasis, a mutation cannot be identified on both *CFTR* genes [Claustres et al., 2000; Groman et al., 2005].

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Harry Cuppens, Center for Human Genetics, Katholieke Universiteit Leuven, Gasthuisberg O&N1 (602), Herestraat 49, B-3000, Leuven, Belgium. E-mail: harry.cuppens@med.kuleuven.be

There is substantial evidence that genes other than *CFTR* may cause CF or CF-like disease. Indeed, a German family was described with a CF patient without *CFTR* mutations, and with an unaffected healthy sister who inherited the same *CFTR* genes from her parents [Mekus et al., 1998]. In two American families, each with two affected sibs, no mutations could be found on both *CFTR* genes, the affected sibs inherited different parental *CFTR* genes [Groman et al., 2002].

Apart from the defective chloride secretion, loss of functional CFTR results in increased sodium absorption through the amiloride sensitive epithelial sodium channel (ENaC) in CF patients [Stutts et al., 1995]. The ENaC channel is a heteromeric transmembrane protein, composed of three subunits (α, β, and γ). These three subunits are encoded by three different genes: *SCNN1A* (sodium channel nonvoltage-gated 1, alpha); *SCNN1B* (sodium channel nonvoltage-gated 1, beta); and *SCNN1G* (sodium channel nonvoltage-gated 1, gamma) (MIM♯s 600228, 600760, and 600761, respectively) [Jasti et al., 2007]. At present, the molecular mechanisms of the regulatory relationship of the three subunits within ENaC, and between ENaC and CFTR, are incompletely understood and subject of considerable controversy. However, it is generally accepted that the coordinated regulation and activity of CFTR and ENaC determine the composition of the airway surface liquid (ASL) in the lungs. CFTR-mediated Cl^- secretion enhances ASL volume while ENaC mediated Na^+ absorption reduces it. Thus, an appropriate balance between CFTR and ENaC activity is essential for maintaining an optimal ASL volume; if not, disease will occur [Boucher, 2007; Knowles and Boucher, 2002].

In classical CF patients with loss-of-function mutations in both copies of the *CFTR* gene, the ASL depletion is explained by the lack of CFTR-mediated Cl^- secretion and accelerated Na^+ absorption. Interestingly, transgenic mice that overexpress SCNN1B in the lower airways have increased airway epithelial sodium absorption and present CF-like lung disease symptoms [Mall et al., 2004].

Mutations in ENaC are also known to be involved in two different human genetic diseases. Either too low or too high amounts of ENaC will thus cause disease. Activation of ENaC by mutations in either *SCNN1B* or *SCNN1G* cause Liddle's syndrome (MIM♯ 177200) [Shimkets et al., 1994], an inherited form of salt-sensitive arterial hypertension with enhanced renal sodium retention. These dominant gain-of-function mutations lead to an enhanced channel activity by increasing the number of ENaC channels expressed at the cell surface and by increasing the channel's open probability [Firsov et al., 1996; Snyder et al., 1995]. Loss-of-function mutations in *SCNN1A*, *SCNN1B*, or *SCNN1G* cause autosomal recessive pseudohypoaldosteronism type I (PHA-I; MIM♯ 264350) [Chang et al., 1996], characterized by severe renal salt-wasting and arterial hypotension. Moreover, in PHA-I patients, the reduced reabsorptive capacity of the lungs leads to an increased ASL volume, which often results in recurrent respiratory problems [Kerem et al., 1999]. Indeed, PHA-I patients were found with chronic lung disease that resembles that of CF in the absence of common *CFTR* mutations [Marthinsen et al., 1998; Schaedel et al., 1999].

Given the observation that sodium hyperabsorption through ENaC is part of the basic CF pathology, and that mice that overexpress SCNN1B present CF-like disease, mutations in the genes encoding ENaC may potentially explain disease in patients with CF-like disease in whom a mutation cannot be identified on both *CFTR* genes. We therefore investigated whether mutations in any of the subunits of ENaC could be involved in the latter patients.

Materials and Methods

Study Population

The patients originated from different CF clinics across Europe. These included both typical CF patients and patients having a phenotype of atypical CF. Originally these patients were studied in three independent studies that were finally combined in one study. In the first study, a total of 31 patients (Belgian, ten; Swedish, four; German, six; Czech, three; and French (Montpellier), eight) were included. Of these, 18 patients did not carry a CF-causing mutation on either of their two *CFTR* genes, while 13 patients carried a CF-causing mutation on one *CFTR* gene. A total of 45 French (Paris and Brest) patients were included in the two other studies. Among the latter, 29 patients did not carry a *CFTR* mutation at all, while 16 patients carried a *CFTR* mutation in one *CFTR* gene. A total of 234 true healthy control individuals (Belgian 85, Czech 149) without any reported disease were included. Also, because of p.F508del homozygosity, 683 CF patients (Danish 156, Czech 121, Belgian 104, British 101, Italian 101, and German 100) with explained CF disease were used as controls. The studies were approved by the ethical committees of the participating centers.

PCR and Sequencing

All coding exons and their exon/intron junctions of the *SCNN1A*, *SCNN1B*, and *SCNN1G* genes were sequenced in the 31 patients of the first study. Moreover, the noncoding exon 1 of *SCNN1A* and *SCNN1G* were also sequenced. PCR was performed in a 50-µl solution containing 1x GeneAmp® Gold buffer (Applied Biosystems, Foster City, CA), 0.2 mM dNTPs (GE Healthcare Europe GmbH, Munich, Germany), 2.5 mM $MgCl_2$, 2.5 U AmpliTaq® Gold (Applied Biosystems), 10 pmol of each primer (see Supp. Table S1 for the oligonucleotide sequences) and 100 ng genomic DNA. The PCR conditions were as follows: initial denaturation at 95°C for 5 min; 40 cycles of denaturation at 95°C for 30 sec, annealing at 60°C for 30 sec, extension at 72°C for 1 min, and a final extension step at 72°C for 5 min. Since the region containing exon 3 of *SCNN1A* could not be amplified in a single amplicon, this region was amplified in two amplicons (3a and 3b). The amplification mixture for amplicon 3b contained 1.5 mM $MgCl_2$ and 1 M betaine (Sigma, St. Louis, MO).

PCR products were purified on Montage™ PCR centrifugal filter devices (Millipore, Billerica, MA). Both DNA strands were sequenced with the BigDye® terminator cycle sequencing kit v3.1 (Applied Biosystems) according to the manufacturer's protocol. Most of the primers used for sequencing were identical to the primers used for PCR, only a minority of them was nested (see Supp. Tables S1 and S2 for the oligonucleotide sequences). The sequencing reactions were run on an ABI 3130xl genetic analyzer (Applied Biosystems) and were analyzed with SeqScape® software v2.5 (Applied Biosystems). All nucleotides of the exons and the flanking intronic 30 nucleotides were analyzed.

For the sequencing of the 45 patients of the two other studies, all the coding regions of the three *ENaC* genes were analyzed using primers and conditions described by Sheridan et al. [2005].

Multiplex PCR and Single-Nucleotide Primer Extension Assay

Four different multiplex PCR reactions (see Supp. Table S3 for the oligonucleotide sequences) were developed for two sets of single-nucleotide primer extension assays, which included all the

mutations found after sequencing of the three subunits of *ENaC* in the first study. The SNaPshot™ Multiplex kit (Applied Biosystems) was used according to the manufacturer's protocol. The length of the primers varied between 24 bp and 81 bp, allowing discrimination of the different SNPs in the SNaPshot reactions (see Supp. Table S4 for the oligonucleotide sequences). The reaction mixture was run on an ABI 3130xl genetic analyzer (Applied Biosystems) and the electropherograms were analyzed with the GeneScan® software v3.7 (Applied Biosystems). The mutation nomenclature follows the journal guidelines (www.hgvs.org/mutnomen) as described in the footnote to Table 1.

Mutagenesis

Human cDNA of the *SCNN1A*, *SCNN1B*, and *SCNN1G* subunits were synthesized from total RNA, extracted from primary cultured nasal epithelial cells, and were each cloned in pcDNA3 (Invitrogen, Carlsbad, CA). Specific mutations found during sequencing were introduced with the QuickChange® XL site-directed mutagenesis kit (Stratagene, La Jolla, CA) according to the recommendations of the manufacturer. The mutagenesis constructs were transformed into XL1 blue cells by electroporation. Plasmid DNA was isolated with the NucleoBond® PC 500 plasmid DNA purification kit (Macherey-Nagel, Düren, Germany). The inserts and insert vector junctions of the different expression vectors were completely verified by sequencing.

Pulse-Chase, Western Blot, and Endoglycosidase Analysis

Pulse-chase, Western blot, and endoglycosidase analysis were performed analogously as previously described [Cuppens et al., 1998; Jaspers et al., 1988]. For pulse-chase experiments, *ENaC* expression vectors were transfected in COS-1 cells. For Western blot analysis, human bronchial epithelial (HBE) cells (16HBE14o-) [Forbes et al., 2003] were cultured in collagen-coated tissue cultures. Anti-α-ENaC antibody (1 mg/ml) (Calbiochem, San Diego, CA), goat-anti-rabbit-HRP (1/2,000 dilution) (Dako, Glostrup, Denmark) and endoglycosidase H (Endo H) (Roche Diagnostics GmbH, Mannheim, Germany) were used.

Isolation of Oocytes and Injection of cRNA

Oocytes were obtained from adult female *Xenopus laevis* as described previously [Kraus et al., 2007; Rauh et al., 2006; Wielputz et al., 2007]. Defolliculated stage V–VI oocytes were injected with cRNA for *SCNN1A*, *SCNN1B*, and *SCNN1G* (0.5–1.0 ng/subunit) dissolved in RNase-free water with a total volume of 46 nl per oocyte. cRNA was synthesized from full-length cDNA using T7 RNA polymerases (mMessage mMachine; Ambion, Austin, TX). Since cRNA quality can differ from batch to batch, and therefore can influence the protein expression level, matched cRNAs were transcribed in parallel (with the exception of the cRNAs for p.S82C-*SCNN1B* and p.G589S-*SCNN1B*), and were stored in aliquots at –80°C until further use. Injected oocytes were stored at 19°C in modified Barth's solution (in mM: NaCl 85, KCl 1, $NaHCO_3$ 2.4, $Ca(NO_3)_2$ 0.3, $CaCl_2$ 0.4, $MgSO_4$ 0.8, HEPES 10, pH 7.4) for *SCNN1B* mutants, or in ND96 (NaCl 96, KCl 2, $CaCl_2$ 1.8, $MgCl_2$ 1, HEPES 5, pH 7.4) for *SCNN1A* or *SCNN1G* mutants, each supplemented with penicillin (100 U/ml) and streptomycin (100 µg/ml) (Sigma, Taufkirchen, Germany) to prevent bacterial overgrowth.

Two-Electrode Voltage Clamp

Whole-cell currents were routinely measured 2 days after injection using the two-electrode voltage clamp technique as described previously [Kraus et al., 2007; Rauh et al., 2006; Wielputz et al., 2007]. The oocytes were superfused with ND96 (see above) and clamped at a holding potential of –60 mV. Amiloride-sensitive whole-cell currents (ΔI_{ami}) were obtained by washing out amiloride (2 µM) with amiloride-free ND96, and subtracting the whole-cell currents measured in the presence of amiloride from the corresponding whole-cell currents recorded in the absence of amiloride. ΔI_{ami} values were normalized to the mean value of the matched wild-type ENaC-expressing control group.

Nasal Potential Difference Measurements

Nasal PD was determined by standard criteria as described [Knowles et al., 1995; Schuler et al., 2004]. The PD is measured between a fluid-filled exploring bridge on the nasal mucosa and a reference bridge on the skin of the forearm [Middleton et al., 1994]. Perfusion of nasal mucosa was performed with the following solutions: 1) Ringer's solution containing Na 147.5 mEq/l, K 4 mEq/l, Cl 156 mEq/l, Ca 4.5 mEq/l; 2) Ringer's solution containing 100 µM Amiloride; 3) chloride-free solution containing Na 147.5 mEq/l, K 4 mEq/l, Cl 0 mEq/l, Ca 4.5 mEq/l, and 100 µM Amiloride; 4) chloride-free solution as in 3) plus 10 µM isoproterenol.

Statistical Analysis

Database management and statistical analysis was performed with the SAS statistical package, release 9.1.3 (SAS Institute Inc., Cary, NC). Statistical tests were considered significant when their type I error was <0.05. For highly relevant associations, Bonferroni corrections were performed. Tests for Hardy-Weinberg equilibrium were performed for each SNP, separately for the patient and control group using a Fisher's exact test. Differences in genotype frequencies between patients and controls were tested for each SNP with the use of a chi-squared test with 1 degree of freedom, or by the Fisher exact probability test when more than 25% of the cells had a count of <5. A series of tests assuming an additive, dominant, or recessive genetic model were performed for each of the SNPs. Genotypic odds ratios (ORs) and 95% confidence intervals (CIs) were estimated.

Results

Patients With CF or CF-Like Disease Without Two *CFTR* Mutations

The clinical criteria for inclusion of patients in this study were the presence of CF or CF-like lung symptoms, as well as the presence of a positive or borderline sweat test and/or the presence of one CF-causing *CFTR* mutation. Originally these were three independent studies; in the end, all of the data from the three studies were combined. In the first study, 31 patients were recruited from different centers over Europe. In the two other studies, 45 patients were recruited from two centers in France (Paris and Brest). A CF-causing *CFTR* mutation in the coding region and exon/intron junctions of both *CFTR* genes was excluded in these patients; therefore, only patients truly carrying one or no CF-causing *CFTR* mutations were included—either after sequencing of the complete *CFTR* coding region and its exon/intron junctions, or after mutation scanning of the complete *CFTR*

Table 1. Summary of the Detected Mutations and Their Allelic Frequencies*

Subunit/GenBank Accession ID	SNP reference	Mutation	NT Variant	NT Position	Location	Allelic frequency (mutated allele)								
						Patients ($n = 152$)	True controls ($n = 352$)	CF controls ($n = 1294$)[a]	**Belgium ($n = 194$)**	**Czech Republic ($n = 188$)**	**Germany ($n = 200$)**	**Denmark ($n = 308$)**	**Italy ($n = 202$)**	**United Kingdom ($n = 202$)**
SCNN1A	rs10849447	c.–93A>G	A>G	–93	Exon 1	0.480	0.490	0.510	**0.488**	**0.478**	**0.495**	**0.543**	**0.415**	**0.624**
NT_009759.15	rs13306617	c.–55+5G>C	G>C	–55+5	Intron 1	0.026	0.003	0.002	/	**0.014**	**0.005**	/	/	/
NM_001038.4	rs61758858	c.–54–14C>T	C>T	–54–14	Intron 1	0.008	/	/	/	/	/	/	/	/
	rs13306619	p.P33P	C>T	99	Exon 2	0.007	/	0.004	/	/	**0.016**	/	/	/
	rs61758859	p.F61L	C>A	183	Exon 2	0.007	/	/	/	/	/	/	/	/
	rs61759861	p.V114I	G>A	340	Exon 2	0.007	/	/	/	/	/	/	/	/
	rs61759862	p.L180L	G>T	540	Exon 3	0.026	0.024	0.030	**0.033**	**0.014**	**0.016**	**0.043**	**0.016**	**0.037**
	rs61759925	p.R181W	C>T	541	Exon 3	0.026	0.023	0.026	**0.026**	**0.013**	**0.010**	**0.042**	**0.015**	**0.035**
	rs11542844	p.A334T	G>A	1000	Exon 6	0.046	0.039	0.039	**0.027**	/	**0.020**	**0.071**	**0.040**	**0.038**
	rs5742912	p.W493R	T>C	1477	Exon 10	0.040	0.018	0.015	**0.018**	/	**0.015**	**0.010**	**0.015**	**0.040**
	rs2228576	p.T663A	A>G	1987	Exon 13	0.605	0.648	0.708	**0.724**	**0.671**	**0.717**	**0.696**	**0.720**	**0.720**
SCNN1B	rs35731153	p.S82C	C>G	245	Exon 2	0.007	0.009	0.007	**0.011**	/	**0.015**	**0.003**	/	**0.011**
NT_010393.15	rs238547	p.P93P	T>C	279	Exon 2	0.467	0.517	0.537	**0.534**	**0.487**	**0.551**	**0.529**	**0.605**	**0.506**
NM_000336.2	rs61759915	c.777–5T>C	T>C	777-5	Intron 4	0.007	/	0.005	/	**0.014**	**0.010**	**0.003**	/	**0.012**
	rs250563	p.F293F	C>T	879	Exon 5	0.066	0.067	0.080	**0.090**	**0.097**	**0.075**	**0.091**	**0.069**	**0.060**
	rs61759916	p.I515I	C>T	1545	Exon 13	0.007	/	0.002	**0.006**	/	/	/	/	**0.007**
	rs61759926	p.G589S	G>A	1765	Exon 13	0.007	/	/	/	/	/	/	/	/
	rs61759917	p.D629D	C>T	1887	Exon 13	0.007	/	0.001	/	/	/	**0.003**	/	/
SCNN1G	rs5734	p.Y129Y	T>C	387	Exon 3	0.276	0.309	0.298	**0.281**	**0.342**	**0.321**	**0.284**	**0.255**	**0.321**
NT_010393.15	rs5735	p.I158I	T>C	474	Exon 3	0.282	0.317	0.309	**0.275**	**0.363**	**0.319**	**0.283**	**0.300**	**0.476**
NM_001039.3	rs5737	p.G183G	C>T	549	Exon 3	0.046	0.066	0.066	**0.056**	**0.106**	**0.077**	**0.056**	**0.060**	**0.055**
	rs5738	p.E197K	G>A	589	Exon 3	0.013	/	0.006	/	/	**0.016**	**0.003**	**0.005**	/
	rs5740	c.1176+14A>G	A>G	1176+14	Intron 7	0.711	0.801	0.811	**0.844**	**0.790**	**0.758**	**0.830**	**0.815**	**0.821**
	rs12708649	c.1373+29T>C	T>C	1373+29	Intron 9	0.200	0.225	0.226	**0.275**	**0.223**	**0.222**	**0.216**	**0.195**	**0.235**
	rs13306653	c.1432–7G>A	G>A	1432–7	Intron 10	0.197	0.211	0.222	**0.276**	**0.170**	**0.222**	**0.214**	**0.195**	**0.240**
	rs5723	p.L649L	C>G	1947	Exon 13	0.197	0.234	0.232	**0.275**	**0.224**	**0.222**	**0.219**	**0.195**	**0.288**

*NT_009759.15 and NT_010393.15 were used as reference sequence. Nucleotide numbering reflects cDNA numbering, with +1 as the A of the ATG translation initiation codon in the reference sequences NM_001038.4, NM_000336.2, and NM_001039.3, and the initiation codon is codon 1. p.L180L and p.R181W in *SCNN1A* involved two neighboring nucleotides, which were always observed in *cis*.

[a]F508del/F508del controls were pooled together from the F508del/F508del CF patients from individual countries shown in bold columns.

NT, nucleotide; rs, NCBI reference SNP; n, highest number of alleles for which a SNP genotyping was successful; /, allele not found in tested samples.

coding region and its exon/intron junctions using denaturing gradient gel electrophoresis (DGGE) or denaturing high-performance liquid chromatography (dHPLC). The sensitivity of these scanning assays may not reach 100%, so a *CFTR* mutation might have escaped detection. Therefore, the *CFTR* coding region, including the exon/intron junctions, was completely sequenced in a given patient of the latter group when a highly relevant *ENaC* mutation was found, to rule out a *CFTR* mutation being missed by the screening assays. The *CFTR* analysis also included the search for large deletions or duplications by multiplex ligation-dependent probe amplification (MLPA) or quantitative fluorescent multiplex PCR (QFM-PCR). The overall clinical characteristics of the patients are given in Supp. Table S5.

An Increased Proportion of *ENaC* Mutations in Patients With Unexplained CF or CF-Like Disease

The coding region and exon/intron junctions of the three *ENaC* subunits were sequenced in the first 31 patients. Based on the observed CF-like lung disease in mice that overexpress SCNN1B, we first analyzed the *SCNN1B* gene. This analysis identified seven mutations in the tested *SCNN1B* genes in a small number of patients. We then sequenced the other two *ENaC* subunits, *SCNN1A* and *SCNN1G*, in which we identified 11 and eight additional mutations, respectively. All 26 identified sequence variants, and their location in the *SCNN1A*, *SCNN1B*, and *SCNN1G* genes, are shown in Figure 1; their allelic frequencies in the three combined studies are shown in Table 1. Multiplex single-nucleotide primer extension assays were developed to screen all of these 26 *ENaC* sequence variants. The prevalence of these 26 *ENaC* sequence variants was then determined in controls; i.e. 234 controls with no lung disease, as well as in 683 p.F508del *CFTR* homozygotes, from different corners of Europe (Table 1).

Of these 26 *ENaC* mutations, seven were not described before in the literature or in public databases. Two of the seven mutations, i.e. p.V114I and p.F61L in *SCNN1A*, were only found once in patients and not at all in controls.

Besides the 26 sequence variants from the first study, we identified four additional new sequence variants in the two other studies of 45 CF patients (Fig. 1): the missense mutations p.R81C (c.241C>T) (rs61759860) in *SCNN1A*, p.P368A (c.1102C>G) (rs61759918) in *SCNN1G*, and the silent mutations p.S145S (c.435C>T) (rs62639702) and p.S212S (c.636C>T) (rs16977041) in *SCNN1G*. These mutations were not analyzed in detail regarding their prevalence in controls and physiological consequences.

The prevalence of the different variations were comparable between the different control populations, so that they could be pooled and compared to the patient group. Except for the p.F293F locus in *SCNN1B* ($P<0.0001$), no deviation from the Hardy-Weinberg equilibrium was found for the different loci in controls. It is not clear if a technical artifact or a biological mechanism is responsible for the observed linkage disequilibrium for p.F293F in *SCNN1B*.

For the combined group of mutations that have a minor allele frequency (MAF) <2.5% in controls (in *SCNN1A*: c.–55+5G>C, c.–54–14C>T, p.P33P, p.F61L, p.L180L-R181W, p.V114I, and p.W493R; in *SCNN1B*: p.S82C, c.777–5T>C, p.I515I, p.G589S, and p.D629D; and p.E197K in *SCNN1G*), a cumulative frequency of 30% in the studied patient group was found, which is significantly increased compared to the cumulative frequency of 9% found in controls ($P<0.0001$). Specifically, for the c.–55+5G>C locus in *SCNN1A*, a higher number of patients (four patients, 5.3%) were heterozygous for the C allele compared to controls (0.4%) ($P<0.002$), with an OR of 13.5 (95% CI, 3.0–61.3) (Table 2). The cumulative frequency of hyperactive *ENaC* mutations p.W493R-*SCNN1A*, p.L180L-R181W-*SCNN1A*,

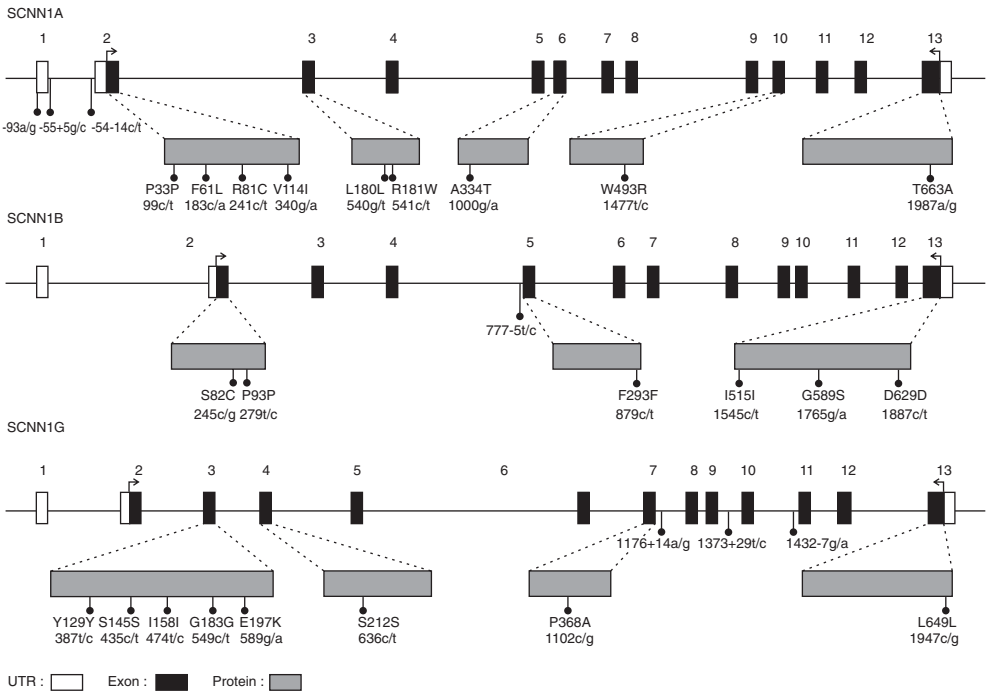


Figure 1. Schematic diagram of the different *SCNN1A*, *SCNN1B*, and *SCNN1G* mutations found in this study. The coding exons and their flanking 30 nucleotides in the introns were analyzed. Translation initiation and termination are shown by arrows. Open boxes refer to the untranslated region (UTR), black boxes refer to coding exon, and gray boxes refer to protein.

and p.V114I-*SCNN1A*, as shown below, was also significantly increased in the patient group (14%) vs. controls (7%) (P = 0.01). The most relevant functional variant, p.W493R-*SCNN1A*, was found in six patients (8%), at a more than two-fold increased incidence compared to controls (3.1%) ($P<0.03$), with an OR of 2.7 (95% CI, 1.1–6.8) (Table 2).

With regard to SNPs having a MAF > 2.5%, homozygosity for the A allele at the c.1176+14A>G locus in *SCNN1G* was found at a three-fold incidence in patients (11.8%) vs. controls (4.1%) ($P<0.002$), with an OR of 3.2 (95% CI, 1.4–6.9), while homozygosity for the G allele was found at a reduced incidence in patients (54% vs. 66%, $P<0.04$) (Table 2). This association was also found to a lesser extent for other loci that were in strong linkage disequilibrium with c.1176+14A>G. These associations were more pronounced in patients that did not carry a *CFTR* mutation at all ($P<0.0007$; $P<0.02$ after Bonferroni correction), compared to the CF carriers (Table 2).

From a genetic point of view, all these findings suggest an involvement or susceptibility of *ENaC* mutations in disease in some of these patients.

Hyperactive and Hypoactive ENaC Mutations

In a next step we investigated whether the different missense mutations affect amiloride-sensitive whole-cell currents mediated by ENaC heterologously expressed in *Xenopus laevis* oocytes. For this purpose, expression vectors for each of the *ENaC* subunits were generated, starting from cDNA that was converted from RNA isolated from a primary cell line of human nasal epithelial cells. The different missense mutations were inserted in the corresponding expression vector. Since p.T663A-*SCNN1A* is a frequent variant, each missense mutation was introduced on the p.T663A-*SCNN1A* background as found in the patient. When the *cis* or *trans* configuration of the p.T663A-*SCNN1A* background could not be determined, the missense mutation was generated on both the p.T663-*SCNN1A* and p.A663-*SCNN1A* background. Oocytes injected with wild-type cRNA of *SCNN1A*, *SCNN1B*, and *SCNN1G* were used as controls. For every tested mutant, the mutant subunit cRNA was injected, together with the wild-type cRNAs of the other two subunits. The p.T663A background of *SCNN1A*, as observed in the patient, was also taken into consideration for *SCNN1B* or *SCNN1G* mutations.

Compared to wild-type ENaC, a significantly higher amiloride-sensitive whole-cell current (ΔI_{ami}) was found for ENaC harboring p.W493R, p.R181W, or p.V114I in the SCNN1A subunit, with an increase to about 450% ($P<0.0001$), 175% ($P<0.0001$), and 150% ($P=0.0004$), respectively. ENaC carrying p.F61L or p.A334T in the SCNN1A subunit showed significantly lower ΔI_{ami} values, which were decreased by about 90% ($P<0.0001$) and 45% ($P<0.0001$), respectively. p.T663A in *SCNN1A* (in contrast to the findings of Samaha et al. [2004]); p.S82C and p.G589S in *SCNN1B*; and p.E197K in *SCNN1G* did not significantly affect sodium transport (Fig. 2; Table 3).

The SCNN1A protein steady-state levels were studied in pulse-chase experiments (Supp. Fig. S1). Except for p.F61L-A663-SCNN1A, we could not detect any significant change in the protein steady-state levels of the SCNN1A mutants compared to that of the controls p.T663-SCNN1A or p.A663-SCNN1A. The reduced protein expression of p.F61L-A663-SCNN1A in vitro is in agreement with the reduced functional expression of p.F61L-SCNN1A in the *Xenopus laevis* oocyte expression system.

In vivo functional studies could be performed in two patients who carried the most hyperactive mutation p.W493R-*SCNN1A* mutation, by means of nasal potential difference (NPD) measurements using standard methods [Knowles et al., 1995; Schuler et al., 2004]. The ENaC channel can be blocked by amiloride, so perfusion with amiloride provides information about ENaC activity. The NPD of a patient with classic CF is remarkably different from that of controls. In CF patients, the basal PD is much higher (more negative) (values are mean ± standard deviation [SD]. CF: –31.1 ± 16.6 mV; controls: –15.2 mV ± 6.9 mV), the amiloride response is exaggerated (CF: +19.4 ± 11.7 mV; controls: +6.5 ± 2.6 mV), and there is little or no response to chloride free and isoproterenol solutions (CF: +0.08 ± 3.7 mV; controls: –16.3 ± 9.2 mV).

One patient, who was heterozygous for p.W493R in *SCNN1A* and p.F508del in *CFTR* (Fig. 3), had a basal PD of –25 mV, an amiloride-response of +20 mV, and a response to chloride removal of –15 mV. Thus, this patient had a rather negative basal potential, however still within the normal range, a strong amiloride-response never seen in controls, and a chloride secretion reflecting the heterozygous state of p.F508del-*CFTR* with one unaffected CFTR gene sufficient to produce functional CFTR channels. The second individual, who was heterozygous for p.W493R-*SCNN1A* and the milder *CFTR* polymorphisms p.E528E and the TG12-T7 haplotype located in intron 8, had a basal PD of –19 mV, which is at the higher end of the normal range, a pronounced amiloride response of +18 mV, which is well above the control range, and a normal chloride secretion (–20 mV). These in vivo findings are consistent with our in vitro expression studies and support our conclusion that the p.W493R-*SCNN1A* mutation causes hyperactivity of ENaC.

The detailed clinical characteristics of the patients carrying the p.W493R-*SCNN1A* mutation are shown in Supp. Table S6.

Of special interest was the p.V114I-*SCNN1A* mutation, which was found in the German CF patient without a CF-causing *CFTR* mutation, and with a healthy sibling who inherited the same parental *CFTR* genes [Mekus et al., 1998]. The clinical features of this patient were described in detail by Mekus et al. [1998], and are summarized in Supp. Table S6. Only the p.V114I-*SCNN1A* mutation was found in this patient, while p.V114I-*SCNN1A* was not found in his healthy sibling. Moreover, p.V114I-*SCNN1A* was not found in the healthy parents, and nonpaternity was excluded. p.V114I-*SCNN1A* was thus a de novo mutation and was not found in 1,646 control alleles. Moreover, p.V114 was completely conserved in the orthologs and paralogs of *ENaC* subunit genes (Supp. Fig. S2).

Discussion

We investigated whether *ENaC* mutations may contribute to the pathophysiology of CF or CF-like disease in patients in whom only one *CFTR* gene is mutated, or that carry no mutations at all in the *CFTR* coding region and its exon/intron junctions. A total of 30 sequence variants in *ENaC* genes were found in 76 patients. Only the hyperactive p.V114I and hypoactive p.F61L mutations in *SCNN1A* were observed once in our patients, but not in controls. These mutations thus potentially cause disease by a Mendelian mechanism. Interestingly, the hyperactive p.V114I-*SCNN1A* mutation was identified in the German patient that provided the first evidence that mutations in other genes than *CFTR* may cause CF-like disease [Mekus et al., 1998], and now provides evidence that a mutant *SCNN1A* gene causes CF-like disease in this patient.

The cumulative frequency of mutations that have a minor allele frequency of <2.5% in controls had a more than three-fold significantly increased cumulative frequency (30%) in the studied

Table 2. Distribution of Genotypes in Patients vs. Controls*

	All patients (CF carriers + non-CF carriers)				Patients (CF carriers)				Patients (non-CF carriers)			
p.T663A		AA	GG+GA			AA	GG+GA			AA	GG+GA	
Patients	$n=76$	14 (18.4)	62 (81.6)	$P<0.03$	$n=29$	5 (17.2)	24 (82.8)	**$P<0.22$**	$n=47$	9 (19.2)	38 (80.9)	$P<0.05$
Controls	$n=757$	77 (10.2)	680 (89.8)		$n=757$	77 (10.2)	680 (89.8)		$n=757$	77 (10.2)	680 (89.8)	
c.–93A>G		**AA+GA**	**GG**			AA	GG+GA			AA+GA	GG	
Patients	$n=76$	64 (84.2)	12 (15.8)	**$P<0.06$**	$n=29$	2 (6.9)	27 (93.1)	$P<0.03$	$n=47$	42 (89.4)	5 (10.6)	$P<0.02$
Controls	$n=716$	532 (74.3)	184 (25.7)		$n=716$	176 (24.6)	540 (75.4)		$n=716$	532 (74.3)	184 (25.7)	
c.–55+5G>C		CC+GC	GG			CC+GC	GG			CC+GC	GG	
Patients	$n=76$	4 (5.3)	72 (94.7)	$P<0.002$	$n=29$	1 (3.5)	28 (96.5)	**$P<0.14$**	$n=47$	3 (6.4)	44 (93.6)	$P<0.004$
Controls	$n=730$	3 (0.4)	727 (99.6)		$n=730$	3 (0.4)	727 (99.6)		$n=730$	3 (0.4)	727 (99.6)	
OR (95% CI)		13.5 (3.0 – 61.3)				8.7 (0.9–85.8)				16.5 (3.2–84.2)		
p.W493R		CC+CT	TT			CC+CT	TT			CC+CT	TT	
Patients	$n=76$	6 (7.9)	70 (92.1)	$P<0.03$	$n=29$	2 (6.9)	27 (93.1)	**$P<0.25$**	$n=47$	4 (8.5)	43 (91.5)	**$P<0.07$**
Controls	$n=739$	23 (3.1)	716 (96.9)		$n=739$	23 (3.1)	716 (96.9)		$n=739$	23 (3.1)	716 (96.9)	
OR (95% CI)		2.7 (1.1–6.8)										
c.1176+14A>G		AA	GG+GA			AA	GG+GA			AA	GG+GA	
Patients	$n=76$	9 (11.8)	67 (88.2)	$P<0.002$	$n=29$	2 (6.9)	27 (93.1)	**$P<0.35$**	$n=47$	7 (14.9)	40 (85.1)	$P<0.0007$
Controls	$n=785$	32 (4.1)	753 (95.9)		$n=785$	32 (4.1)	753 (95.9)		$n=785$	32 (4.1)	753 (95.9)	
OR (95% CI)		3.2 (1.4–6.9)								4.1 (1.7–9.9)		
c.1176+14A>G		AA+GA	GG			AA+GA	GG			AA+GA	GG	
Patients	$n=76$	35 (46.1)	41 (53.9)	$P<0.04$	$n=29$	12 (41.4)	17 (58.6)	**$P<0.42$**	$n=47$	23 (48.9)	24 (51.1)	$P<0.04$
Controls	$n=785$	268 (34.1)	517 (65.9)		$n=785$	268 (34.1)	517 (65.9)		$n=785$	268 (34.1)	517 (65.9)	
c.1432–7G>A		**GG**	**AA+GA**			GG	AA+GA			**GG**	**AA+GA**	
Patients	$n=76$	52 (68.4)	24 (31.6)	**$P<0.24$**	$n=29$	15 (51.7)	14 (48.3)	**$P<0.28$**	$n=47$	37 (78.7)	10 (21.3)	$P<0.02$
Controls	$n=757$	467 (61.7)	290 (38.3)		$n=757$	467 (61.7)	290 (38.3)		$n=757$	467 (61.7)	290 (38.3)	
p.L649L		CC	GG+GC			CC	GG+GC			**CC**	**GG+GC**	
Patients	$n=76$	51 (67.1)	25 (32.9)	**$P<0.22$**	$n=29$	14 (48.3)	15 (51.7)	**$P<0.21$**	$n=47$	37 (78.7)	10 (21.3)	$P<0.01$
Controls	$n=756$	453 (59.9)	303 (40.1)		$n=756$	453 (59.9)	303 (40.1)		$n=756$	453 (59.9)	303 (40.1)	
c.1373+29T>C		**CC+CT**	**TT**			CC+CT	TT			**CC+CT**	**TT**	
Patients	$n=60$	19 (31.7)	41 (68.3)	**$P<0.25$**	$n=17$	9 (52.9)	8 (47.1)	**$P<0.26$**	$n=43$	10 (23.3)	33 (76.7)	$P<0.04$
Controls	$n=808$	317 (39.2)	491 (60.8)		$n=808$	317 (39.2)	491 (60.8)		$n=808$	317 (39.2)	491 (60.8)	
p.P93P		TT	CC+CT			TT	CC+CT			TT	CC+CT	
Patients	$n=75$	25 (33.3)	50 (66.7)	$P<0.04$	$n=29$	12 (41.4)	17 (58.6)	$P<0.02$	$n=46$	13 (28.3)	33 (71.7)	**$P<0.40$**
Controls	$n=783$	179 (22.9)	604 (77.1)		$n=783$	179 (22.9)	604 (77.1)		$n=783$	179 (22.9)	604 (77.1)	

*Data are shown for all patients grouped, for patients that carry no *CFTR* mutation at all, and for patients that carry one *CFTR* mutation. A series of tests assuming an additive, dominant, or recessive genetic model were performed for each of the SNPs. Only the most significant genetic model is shown. Only significant loci are shown for which the MAF allele was observed in at least two individuals of the total group of patients. Moreover, only loci are shown for which a significant different distribution of genotypes was found in at least one of the patient group models tested. For a given shown SNP, no significant association may therefore be found in certain patient group models, which are in bold. If no Hardy Weinberg equilibrium was found for a given SNP in a given group of individuals, the genotypes are in bold. For the most significant associations, ORs are given with their 95% CIs.

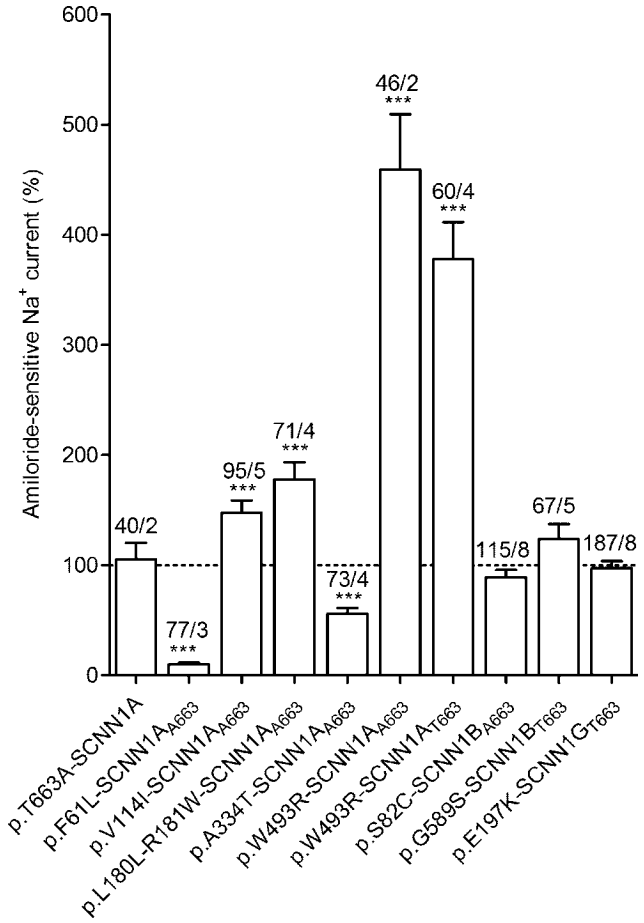


Figure 2. ENaC-mediated amiloride-sensitive Na^+ currents detected with two-electrode voltage clamp. *Xenopus laevis* oocytes were measured two days after cRNA injection. Amiloride-sensitive whole-cell Na^+ currents were normalized to the mean current value measured in matched wild-type ENaC-expressing control oocytes and are given in percentage values. The dotted line at 100% indicates the control level. Group sizes of control and mutant ENaC-expressing groups were similar in individual experiments. Numbers above columns indicate the individual number of oocytes measured and the number of different batches of oocytes used. Index on x-axis indicates p.A663-background or p.T663-background on SCNN1A used for the mutant and the corresponding controls. In the first column p.A663-SCNN1A was compared to p.T663-SCNN1A.

patient group. These mutations are thus potentially involved in disease in these patients by a polygenetic mechanism. Among these, the most relevant functional variant p.W493R-*SCNN1A* was found at a more than two-fold significantly increased incidence in patients (8%).

From a genetic point of view, all these observations thus suggest an involvement of *ENaC* in disease in some of these patients. However, based on the functional findings, it is very difficult to develop a mechanistic explanation of the effect of the different mutations, especially in the light that either too low or too high amounts of ENaC can already result in two different diseases, which will be even further complicated if the CFTR protein is implicated.

Our observations that *ENaC* mutations can cause CF-like disease are in agreement with another study, in which the *SCNN1B* gene was completely sequenced in 20 patients in whom no mutation could be identified on both *CFTR* genes [Sheridan et al., 2005]. Three of the four identified *SCNN1B* mutations were loss-of-function mutations (p.P267L, c.1670–2A>G, and p.E539K), while only one mutation (p.G294S) resulted in a gain-of-function.

We mainly found hyperactive mutations in the *SCNN1A* gene, while transgenic mice overexpressing SCNN1A do not present with CF-like lung disease [Mall et al., 2004]. A likely explanation is that the amount of different ENaC subunits that is present in the cell is likely not equal. Indeed, the amount of a given subunit depends on many factors at the transcriptional and translational level. One subunit will thus be the limiting factor for the generation of functional ENaC channels. In this regard, we found a four-fold higher amount of *SCNN1A* transcripts compared to *SCNN1B* transcripts (our unpublished results). SCNN1B might thus be the limiting factor in the formation of functional ENaC channels, such that overexpression results in more ENaC activity and disease in transgenic mice [Mall et al., 2004]. Conversely, overexpression of SCNN1A has no effect, since it is already present at higher levels than stoichiometrically needed, so that further overexpression results in no higher ENaC activity. On the other hand, gain-of-function *SCNN1A* mutations, as identified in this study, result in increased activity by a qualitative rather than quantitative mechanism.

Of special interest were patients in whom one copy of the *CFTR* gene was found to be mutated, with the other *CFTR* gene being unaffected. Indeed, CFTR is part of a CFTR network, which is (almost) completely nonfunctional when a disease-causing mutation is found on both *CFTR* genes. However, in patients in whom

Table 3. Electrophysiological Characterization of Mutant ENaC Expressed in *Xenopus laevis* Oocytes*

Variant	Background *SCNN1A*	Δ_{ami} (%)	SE	Number of oocytes	Number of frogs	Number of cRNA batches	*P* value
SCNN1A							
p.T663A		105.4	15.0	40	2	2	0.7736
p.F61L	p.A663	10.1	1.3	77	3	2	<0.0001
p.V114I	p.A663	147.7	11	95	5	2	0.0004
p.L180L-R181W	p.A663	177.9	15.5	71	4	2	<0.0001
p.A334T	p.A663	56	5.4	73	4	2	<0.0001
p.W493R	p.A663	459.1	50.4	46	2	2	<0.0001
p.W493R	p.T663	377.7	33.7	60	4	2	<0.0001
SCNN1B							
p.S82C	p.A663	88.9	6.8	115	8	1	0.2138
p.G589S	p.T663	123.6	13.7	67	5	2	0.1196
SCNN1G							
p.E197K	p.T663	97.3	6.4	187	8	4	0.6983

*The ENaC-mediated amiloride-sensitive whole-cell Na^+ current (Δ_{ami}) was normalized to the mean Δ_{ami} of wild-type ENaC-expressing (same background) control oocytes. Numbers of control and mutant expressing oocytes were similar in all experiments. Batches of oocytes were derived from different donors (frogs). With the exception of p.S82C-SCNN1B, at least two batches of cRNA were transcribed in parallel to avoid artifacts due to different cRNA quality.
SE, standard error.

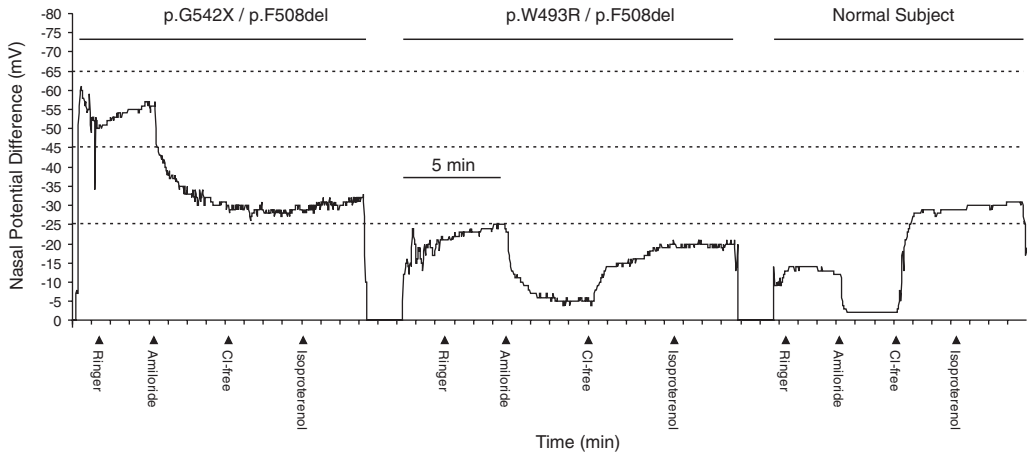


Figure 3. NPD measurements. The NPD was measured between a Ringer's solution–filled exploring bridge on the nasal mucosa of the floor of the nasal cavity, and a reference bridge on the abraded skin of the forearm in a CF patient compound heterozygous for p.F508del and p.G542X in *CFTR*, a patient heterozygous for p.W493R-*SCNN1A* and p.F508del-*CFTR*, and a normal subject. Arrows indicate the time points where perfused solutions were switched, in the following sequence: solution 1, Ringer's solution; solution 2, Ringer's solution + amiloride; solution 3, chloride-free solution + amiloride; solution 4, chloride-free solution + amiloride + isoproterenol.

a mutation is only found on one *CFTR* gene, about 50% of the CFTR network is still functional. While the majority of these individuals are healthy CF carriers, we hypothesize that in some of these individuals the remainder of the CFTR network may be rendered nonfunctional by a mutation in at least one other member of the CFTR network, such as ENaC.

In the *Xenopus laevis* oocyte expression system, p.W493R-SCNN1A was found to result in a more than four-fold higher ENaC activity. In individuals that are heterozygous for p.W493R-SCNN1A, this mutation is therefore predicted to translate in a two-fold higher ENaC activity. The magnitude of this gain-of-function effect is similar to that observed for Liddle's syndrome mutations in the same expression system [Firsov et al., 1996]. In our study, the gain-of-function effect of the p.W493R-*SCNN1A* mutation could be confirmed in vivo using nasal PD measurements in two patients carrying this mutation. Indeed, an increased amiloride response never seen in controls, but only in CF patients, was observed in these two patients. The basal nasal potential difference in these two patients, however, only partially reflected the hyperactive effect of the p.W493R-*SCNN1A* mutation. Possibly, the p.W493R-*SCNN1A* mutation might only have its deleterious consequences in nonbasal conditions. The effect was also only partially reflected in the sweat test values of the six patients that were found to carry the p.W493R-*SCNN1A* mutation (Supp. Table S6). It is known that sweat test values, especially borderline values in milder patients, may change to more pathological sweat values at older ages. For some of these patients, the limited numbers of sweat tests were obtained at older ages. The hyperactive effect of the p.W493R-*SCNN1A* mutation might thus not be representatively reflected in the obtained sweat tests. This would especially apply if the overall effect of the hyperactive *ENaC* mutations might have tissue-specific presentations.

Given the CF carrier incidence of about 3.3% (1/30), and the fact that about 3.1% of the individuals in the general population carry the hyperactive p.W493R-*SCNN1A* mutation, about 1 in 3,900 individuals is expected to be heterozygous for both a CF-causing *CFTR* mutation and p.W493R-*SCNN1A*. This is higher than the incidence of p.F508del homozygous CF newborns (1/7,350). Moreover, the T5 polymorphism in the *CFTR* gene has an allelic frequency of about 5% in the general population [Cuppens et al., 1994] and only results in about 10% functional CFTR [Chu et al., 1993]. Therefore, about 1 out of 975 individuals will be heterozygous for a CF-causing mutation or T5 in *CFTR* and p.W493R-*SCNN1A*. In these individuals, CFTR function will thus be reduced by about 50% and ENaC function will be doubled. No disease will occur when these mutations are found alone in a heterozygous state. It is tempting to speculate that the p.W493R-*SCNN1A* mutation in combination with a loss-of-function mutation in one copy of the *CFTR* gene may cause or predispose to CF or CF-like disease. These genotypes might possibly cause disease by a not fully-penetrant mechanism. Indeed, other modulating genes and environmental factors may determine the penetrance of a given genotype [Vanscoy et al., 2007]. In fact, *CFTR* and the three *ENaC* subunits form an oligogenetic system. The presentation of a functional mutation in one of these genes might therefore even depend on an epistatic effect of another mutation/polymorphism in the other *ENaC* subunit genes or the *CFTR* gene. Possibly, p.W493R-*SCNN1A* might thus be an important risk factor in development of other lung diseases, which should be investigated further.

Our findings will be also clinically relevant for nasal PD measurements in patients, since functional *ENaC* polymorphisms are expected to affect the nasal PD results. More and more atypical patients are referred to the clinic for nasal PD analysis, but often nasal PD traces are found that are neither normal nor typical for CF. Our functional studies provide a likely explanation for these findings, and knowledge of the presence of functional polymorphisms as found in this study may be helpful for the correct interpretation of these assays.

The actual mechanism by which ENaC functions is still unclear. The functional mutations found in this study might provide insight. Recently, the structure of the chicken acid-sensing ion channel 1 (ASIC1) has been determined [Jasti et al., 2007]. ASICs are voltage-independent, proton-activated receptors, which are in amino acid sequence related to the epithelial sodium channels. Based on the finding that ASIC1 is a homotrimer, ENaC is now believed to be a heterotrimeric protein, in which the SCNN1A, SCNN1B, and SCNN1G subunits are each found once [Jasti et al., 2007]. All found functional ENaC mutations very likely contact a twisted β-sheet, which is connected to the second transmembrane helix (TM2)

(Supp. Fig. S3). This suggests a crucial function for this twisted β-sheet for sodium transport through the ENaC channel.

It should be noted that in a considerable number of our patients no involvement of ENaC was found, implicating additional genes/proteins to be involved in disease. Potential proteins are any protein that is part of the complex CFTR interactome [Wang et al., 2006].

In conclusion, in only a small fraction of the 76 patients studied, in whom a mutation cannot be identified on both *CFTR* genes, disease can be explained by a Mendelian model. However, given the significantly increased incidence of several *ENaC* polymorphisms in the patient group, *ENaC* seems to be involved in disease in some patients, possibly in combination with *CFTR*, by a polygenetic mechanism. Of special interest is the strongly hyperactive p.W493R-*SCNN1A* polymorphism, which is not that rare in the general population, that may predispose to disease, especially in combination with the frequent CF carrier status. Such genotypes may play a hitherto unrecognized role in lung diseases, which will be the basis for further studies.

Acknowledgments

These investigations were realized by grants from the Alphonse and Jean Forton Fund–Koning Boudewijn Stichting (2005 04 R7 115 BO); Het Fonds voor Wetenschappelijk Onderzoek Vlaanderen (G.0521.06 and 1.5.111.07), the Interuniversity Attraction Poles (IAP P6/05); the EU grant (QLG1-CT-2001-01005-CFPRONET), the Vaincre La Mucoviscidose (RC0532), and the Deutsche Forschungsgemeinschaft (SFB 423, TP A12). A.K.A. is a recipient of a doctoral fellowship award from the Interfaculty Council for Development Cooperation (IRO) of the Katholieke Universiteit Leuven, Belgium. J.J.C. holds the Arthur Bax and Anna Vanluffelen Chair of Human Genetics, KU Leuven, Belgium. We thank Steffen Fieuws for his advice in statistical analysis and Dieter Gruenert for providing the 16HBE14o- cell line. We also acknowledge Milan Macek, Jr., Věra Vávrová, Libor Fila, and Daniela Zemková for their assistance in clinical data collection and support. They were supported by grants VZFNM 00064203 and NR/9488-3.

References

Boucher RC. 2007. Cystic fibrosis: a disease of vulnerability to airway surface dehydration. Trends Mol Med 13:231–240.

Chang SS, Grunder S, Hanukoglu A, Rösler A, Mathew PM, Hanukoglu I, Schild L, Lu Y, Shimkets RA, Nelson-Williams C, Rossier BC, Lifton RP. 1996. Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. Nat Genet 12:248–253.

Chillón M, Casals T, Mercier B, Bassas L, Lissens W, Silber S, Romey M-C, Ruiz-Romero J, Verlingue C, Claustres M Nunes V, Férec C, Estivill X. 1995. Mutations in the cystic fibrosis gene in patients with congenital absence of the vas deferens. N Engl J Med 332:1475–1480.

Chu CS, Trapnell BC, Curristin S, Cutting GR, Crystal RG. 1993. Genetic basis of variable exon 9 skipping in cystic fibrosis transmembrane conductance regulator mRNA. Nat Genet 3:151–156.

Claustres M, Guittard C, Bozon D, Chevalier F, Verlingue C, Ferec C, Girodon E, Cazeneuve C, Bienvenu T, Lalau G, Dumur V, Feldmann D, Bieth E, Blayau M, Clavel C, Creveaux I, Malinge MC, Monnier N, Malzac P, Mittre H, Chomel JC, Bonnefont JP, Iron A, Chery M, Georges MD. 2000. Spectrum of CFTR mutations in cystic fibrosis and in congenital absence of the vas deferens in France. Hum Mutat 16:143–156.

Cuppens H, Teng H, Raeymaekers P, De Boeck C, Cassiman JJ. 1994. CFTR haplotype backgrounds on normal and mutant CFTR genes. Hum Mol Genet 3: 607–614.

Cuppens H, Lin W, Jaspers M, Costes B, Teng H, Vankeerberghen A, Jorissen M, Droogmans G, Reynaert I, Goossens M, Nilius B, Cassiman JJ. 1998. Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes: the polymorphic (TG)m locus explains the partial penetrance of the T5 polymorphism as a disease mutation. J Clin Invest 101:487–496.

Firsov D, Schild L, Gautschi I, Merillat AM, Schneeberger E, Rossier BC. 1996. Cell surface expression of the epithelial Na channel and a mutant causing Liddle syndrome: a quantitative approach. Proc Natl Acad Sci USA 93: 15370–15375.

Forbes B, Shah A, Martin GP, Lansley AB. 2003. The human bronchial epithelial cell line 16HBE14o-as a model system of the airways for studying drug transport. Int J Pharm 257:161–167.

Groman JD, Meyer ME, Wilmott RW, Zeitlin PL, Cutting GR. 2002. Variant cystic fibrosis phenotypes in the absence of CFTR mutations. N Engl J Med 347: 401–407.

Groman JD, Karczeski B, Sheridan M, Robinson TE, Fallin MD, Cutting GR. 2005. Phenotypic and genetic characterization of patients with features of "nonclassic" forms of cystic fibrosis. J Pediatr 146:675–680.

Jaspers M, de Strooper B, Spaepen M, van Leuven F, David G, van den Berghe H, Cassiman JJ. 1988. Post-translational modification of the beta-subunit of the human fibronectin receptor. FEBS Lett 231:402–406.

Jasti J, Furukawa H, Gonzales EB, Gouaux E. 2007. Structure of acid-sensing ion channel 1 at 1.9A resolution and low pH. Nature 449:316–323.

Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC. 1989. Identification of the cystic fibrosis gene: genetic analysis. Science 245:1073–1080.

Kerem E, Bistritzer T, Hanukoglu A, Hofmann T, Zhou Z, Bennett W, MacLaughlin E, Barker P, Nash M, Quittell L, Boucher R, Knowles MR. 1999. Pulmonary epithelial sodium-channel dysfunction and excess airway liquid in pseudohypoaldosteronism. N Engl J Med 341:156–162.

Knowles MR, Boucher RC. 2002. Mucus clearance as a primary innate defense mechanism for mammalian airways. J Clin Invest 109:571–577.

Knowles MR, Paradiso AM, Boucher RC. 1995. In vivo nasal potential difference: techniques and protocols for assessing efficacy of gene transfer in cystic fibrosis. Hum Gene Ther 6:445–455.

Kraus C, Reis A, Naehrlich L, Dotsch J, Korbmacher C, Rauh R. 2007. Functional characterization of a novel CFTR mutation P67S identified in a patient with atypical cystic fibrosis. Cell Physiol Biochem 19:239–248.

Mall M, Grubb BR, Harkema JR, O'Neal WK, Boucher RC. 2004. Increased airway epithelial Na+ absorption produces cystic fibrosis-like lung disease in mice. Nat Med 10:487–493.

Marthinsen L, Kornfalt R, Aili M, Andersson D, Westgren U, Schaedel C. 1998. Recurrent Pseudomonas bronchopneumonia and other symptoms as in cystic fibrosis in a child with type I pseudohypoaldosteronism. Acta Paediatr 87:472–474.

Mekus F, Ballmann M, Bronsveld I, Dork T, Bijman J, Tummler B, Veeze HJ. 1998. Cystic-fibrosis-like disease unrelated to the cystic fibrosis transmembrane conductance regulator. Hum Genet 102:582–586.

Middleton PG, Geddes DM, Alton EW. 1994. Protocols for in vivo measurement of the ion transport defects in cystic fibrosis nasal epithelium. Eur Respir J 7:2050–2056.

Pignatti PF, Bombieri C, Marigo C, Benetazzo M, Luisetti M. 1995. Increased incidence of cystic fibrosis gene mutations in adults with disseminated bronchiectasis. Hum Mol Genet 4:635–639.

Rauh R, Dinudom A, Fotia AB, Paulides M, Kumar S, Korbmacher C, Cook DI. 2006. Stimulation of the epithelial sodium channel (ENaC) by the serum- and glucocorticoid-inducible kinase (Sgk) involves the PY motifs of the channel but is independent of sodium feedback inhibition. Pflugers Arch 452:290–299.

Rich DP, Anderson MP, Gregory RJ, Cheng SH, Paul S, Jefferson DM, McCann JD, Klinger KW, Smith AE, Welsh MJ. 1990. Expression of cystic fibrosis transmembrane conductance regulator corrects defective chloride channel regulation in cystic fibrosis airway epithelial cells. Nature 347:358–363.

Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, and others. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245:1066–1073.

Samaha FF, Rubenstein RC, Yan W, Ramkumar M, Levy DI, Ahn YJ, Sheng S, Kleyman TR. 2004. Functional polymorphism in the carboxyl terminus of the alpha-subunit of the human epithelial sodium channel. J Biol Chem 279:23900–23907.

Schaedel C, Marthinsen L, Kristoffersson AC, Kornfalt R, Nilsson KO, Orlenius B, Holmberg L. 1999. Lung symptoms in pseudohypoaldosteronism type 1 are associated with deficiency of the alpha-subunit of the epithelial sodium channel. J Pediatr 135:739–745.

Schuler D, Sermet-Gaudelus I, Wilschanski M, Ballmann M, Dechaux M, Edelman A, Hug M, Leal T, Lebacq J, Lebecque P, Lenoir G, Stanke F, Wallemacq P, Tümmler B, Knowles MR. 2004. Basic protocol for transepithelial nasal potential difference measurements. J Cyst Fibros (Suppl 2):151–155.

Sharer N, Schwarz M, Malone G, Howarth A, Painter J, Super M, Braganza J. 1998. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. N Engl J Med 339:645–652.

Sheridan MB, Fong P, Groman JD, Conrad C, Flume P, Diaz R, Harris C, Knowles M, Cutting GR. 2005. Mutations in the beta-subunit of the epithelial Na+ channel in patients with a cystic fibrosis-like syndrome. Hum Mol Genet 14:3493–3498.

Shimkets RA, Warnock DG, Bositis CM, Nelson-Williams C, Hansson JH, Schambelan M, Gill Jr. JR, Ulick S, Milora RV, Findling JW, Canessal CM, Rossierl BC, Liftona RP. 1994. Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. Cell 79:407–414.

Snyder PM, Price MP, Mcdonald FJ, Adams CM, Volk KA, Zeiher BG, Stokes JB, Welsh MJ. 1995. Mechanism by which Liddles syndrome mutations increase activity of a human epithelial Na+ channel. Cell 83:969–978.

Stutts MJ, Canessa CM, Olsen JC, Hamrick M, Cohn JA, Rossier BC, Boucher RC. 1995. CFTR as a cAMP-dependent regulator of sodium channels. Science 269:847–850.

Vanscoy LL, Blackman SM, Collaco JM, Bowers A, Lai T, Naughton K, Algire M, McWilliams R, Beck S, Hoover-Fong J, Hamosh A, Cutler D, Cutting GR. 2007. Heritability of lung disease severity in cystic fibrosis. Am J Respir Crit Care Med 175:1036–1043.

Wang X, Venable J, LaPointe P, Hutt DM, Koulov AV, Coppinger J, Gurkan C, Kellner W, Matteson J, Plutner H, Riordan JR, Kelly JW, Yates 3rd. JR, Balch WE. 2006. Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis. Cell 127:803–815.

Wielputz MO, Lee IH, Dinudom A, Boulkroun S, Farman N, Cook DI, Korbmacher C, Rauh R. 2007. (NDRG2) stimulates amiloride-sensitive Na+ currents in *Xenopus laevis* oocytes and fisher rat thyroid cells. J Biol Chem 282:28264–28273.

Current Biology *18*, 1241–1248, August 26, 2008 ©2008 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2008.07.049

Report

Correlation between Genetic and Geographic Structure in Europe

Oscar Lao,[1,22] Timothy T. Lu,[2,22] Michael Nothnagel,[2] Olaf Junge,[2] Sandra Freitag-Wolf,[2] Amke Caliebe,[2] Miroslava Balascakova,[3] Jaume Bertranpetit,[4] Laurence A. Bindoff,[5] David Comas,[4] Gunilla Holmlund,[6] Anastasia Kouvatsi,[7] Milan Macek,[3] Isabelle Mollet,[8] Walther Parson,[9] Jukka Palo,[10] Rafal Ploski,[11] Antti Sajantila,[10] Adriano Tagliabracci,[12] Ulrik Gether,[13] Thomas Werge,[14] Fernando Rivadeneira,[15,16] Albert Hofman,[16] André G. Uitterlinden,[15,16] Christian Gieger,[17,18] Heinz-Erich Wichmann,[17,18] Andreas Rüther,[19] Stefan Schreiber,[19] Christian Becker,[20] Peter Nürnberg,[20] Matthew R. Nelson,[21] Michael Krawczak,[2,23] and Manfred Kayser[1,23,*]

[1]Department of Forensic Molecular Biology
Erasmus University Medical Center Rotterdam
3000 CA Rotterdam
The Netherlands

[2]Institut für Medizinische Informatik und Statistik
Christian-Albrechts University Kiel
D-24105 Kiel
Germany

[3]Institute of Biology and Medical Genetics
University Hospital Motol and 2nd School of Medicine
Charles University Prague
CZ 150 06, Prague 5
Czech Republic

[4]Unitat de Biologia Evolutiva
Pompeu Fabra University
08003 Barcelona, Catalonia
Spain

[5]Department of Neurology
Haukeland University Hospital and Institute of Clinical Medicine
University of Bergen
5021 Bergen
Norway

[6]Department of Forensic Genetics and Forensic Toxicology
National Board of Forensic Medicine
SE 581 33 Linkoping
Sweden

[7]Department of Genetics, Development, and Molecular Biology
Aristotle University of Thessaloniki
GR-540 06 Thessaloniki
Greece

[8]Laboratoire d'Empreintes Génétiques
EFS-RA site de Lyon
69007 Lyon
France

[9]Institute of Legal Medicine
Medical University Innsbruck
A-6020 Innsbruck
Austria

[10]Department of Forensic Medicine
University of Helsinki
Helsinki FIN-00014
Finland

[11]Department of Medical Genetics
Medical University Warsaw
02-007 Warsaw
Poland

[12]Istituto di Medicina Legale
University of Ancona
I-60020 Ancona
Italy

[13]Molecular Neuropharmacology Group and Center for Pharmacogenomics Department of Neuroscience and Pharmacology
University of Copenhagen
2200 Copenhagen
Denmark

[14]Research Institute of Biological Psychiatry and Center for Pharmacogenomics
Mental Health Center Sct. Hans
Copenhagen University Hospital
DK-4000 Roskilde
Denmark

[15]Department of Internal Medicine, Genetics Laboratory
Erasmus University Medical Center Rotterdam
3000 CA Rotterdam
The Netherlands

[16]Department of Epidemiology
Erasmus University Medical Center Rotterdam
3000 CA Rotterdam
The Netherlands

[17]Institute of Epidemiology
Helmholtz Zentrum München - German Research Center for Environmental Health
D-85764 Neuherberg
Germany

[18]Institute of Medical Informatics, Biometry and Epidemiology
Ludwig-Maximilians University
D-81377 Munich
Germany

[19]Institut für Medizinische Molekularbiologie
Christian-Albrechts University Kiel
D-24105 Kiel
Germany

[20]Cologne Center for Genomics and Institut für Genetik
University of Cologne
D-50674 Cologne
Germany

[21]Genetics
GlaxoSmithKline
Research Triangle Park, North Carolina 27709

[22]These authors contributed equally to this work
[23]These authors contributed equally to this work

*Correspondence: m.kayser@erasmusmc.nl

Summary

Understanding the genetic structure of the European population is important, not only from a historical perspective, but also for the appropriate design and interpretation of genetic epidemiological studies. Previous population genetic analyses with autosomal markers in Europe either had a wide geographic but narrow genomic coverage [1, 2], or vice versa [3–6]. We therefore investigated Affymetrix GeneChip 500K genotype data from 2,514 individuals belonging to 23 different subpopulations, widely spread over Europe. Although we found only a low level of genetic differentiation between subpopulations, the existing differences were characterized by a strong continent-wide correlation between geographic and genetic distance. Furthermore, mean heterozygosity was larger, and mean linkage disequilibrium smaller, in southern as compared to northern Europe. Both parameters clearly showed a clinal distribution that provided evidence for a spatial continuity of genetic diversity in Europe. Our comprehensive genetic data are thus compatible with expectations based upon European population history, including the hypotheses of a south-north expansion and/or a larger effective population size in southern than in northern Europe. By including the widely used CEPH from Utah (CEU) samples into our analysis, we could show that these individuals represent northern and western Europeans reasonably well, thereby confirming their assumed regional ancestry.

Table 1. European Subpopulation Summary Statistics

Subpopulation	Code	Total No. Samples	Final No. Samples*	Sex Ratio (M:F)
Norway (Førde)	NO	52	52	1.74
Sweden (Uppsala)	SE	50	46	all male
Finland (Helsinki)	FI	47	47	0.74
Ireland	IE	37	35	4.29
UK (London)	UK	197	194	8.85
Denmark (Copenhagen)	DK	60	59	1.22
Netherlands (Rotterdam)	NL	292	280	all female
Germany I (Kiel)	DE1	500	494	1.08
Germany II (Augsburg)	DE2	500	489	1.02
Austria (Tyrol)	AT	50	50	all male
Switzerland (Lausanne)	CH	134	133	0.81
France (Lyon)	FR	50	50	2.13
Portugal	PT	16	16	0.78
Spain I	ES1	83	81	1.02
Spain II (Barcelona)	ES2	48	47	0.71
Italy I	IT1	107	106	1.38
Italy II (Marches)	IT2	50	49	all male
Former Yugoslavia	YU	58	55	1.90
Northern Greece	EL	51	51	1.43
Hungary	HU	17	17	0.54
Romania	RO	12	12	1.00
Poland (Warsaw)	PO	50	49	all male
Czech Republic (Prague)	CZ	53	45	0.96
Total		2,514	2,457	

Total number of samples, final number of samples after data cleaning, and the sex ratio (male:female) of the final sample data set for each subpopulation. * is after stringent quality control.

Results and Discussion

According to current theory, the autosomal gene pool of extant human populations in Europe lacks sharp discontinuities [1, 2], with the exception of known isolates such as the Finns [6, 7]. For classical genetic markers including, for example, erythrocyte antigens, changes in population genetic structure have been observed to follow a predominantly southeast-northwest gradient [1, 2], thereby apparently matching the Pleistocene settlement of Europe, the Neolithic expansion from the Fertile Crescent, and (at least in part) the postglacial resettlement of Europe during the Mesolithic. Such gradient was also observed with particular haplogroups derived from the nonrecombining part of the Y chromosome (NRY), but other NRY data revealed additional population structure in Europe that has been associated with various demographic events in prehistoric, historic, and modern times [8–10]. In contrast, the European mitochondrial DNA pool has been found to be rather homogeneous [11]. Here, we investigated the genetic structure of the European population by using 309,790 single-nucleotide polymorphisms (SNPs) in 2,457 individuals, ascertained at 23 sampling sites (henceforth referred to as "subpopulations") in 20 different European countries. The data emerged from the genotyping of 2,514 European samples with the GeneChip Human Mapping 500K Array, followed by stringent quality control (see Table 1 and Experimental Procedures for details) and represent the largest Europe-wide genetic study to date.

First, we quantified the amount of information that each SNP could potentially provide about an individual's subpopulation affiliation by using the ancestry informativeness index *In* (Figure S1 available online) [12]. The maximum *In* value (0.09) was observed for rs6730157 in the *RAB3GAP1* gene located about 68 kb away from the Lactase (*LCT*) gene. Furthermore, nine of the 20 (45%) most ancestry-informative SNPs, and 17 of the top 100 (Table S1), were from the *LCT* region and previously showed signatures of a selective sweep in CEU (Centre d'Etude du Polymorphism Humain from Utah) samples [13]. The average *In* across markers was 0.0064 (standard deviation: 0.0032), which represents only 0.93% of the maximum possible *In* of 0.69 in our study. (Note that this maximum would be attained if a SNP was fixed for one allele in 12 subpopulations and for the other allele in the remaining 11 subpopulations).

Second, we performed a principal-component analysis (PCA) in which the first two PCs were found to account for 31.6% and 17.3%, respectively, of the total variation, an amount similar to that reported in previous studies [1, 5]. In our study, the first two PCs revealed a SNP-based grouping of European subpopulations that was strongly reminiscent of the geographic map of Europe (Figure 1; Figure S2). The first PC aligned subpopulations according to latitude, with the two Italian subpopulations at one end and the Finnish subpopulation at the other. The second PC tended to separate subpopulations more according to longitude, with the Finnish subpopulation showing the largest values and the Irish and UK subpopulations showing the lowest values. The apparent geographic footing of the two PCs received additional support from an observed statistically significant positive correlation (Pearson $r^2 = 0.632$, two-tailed $p < 10^{-15}$) between the genetic distance (Euclidian distance between the median first two eigenvectors of the PCA) and the geographic (great-circle) distance between the analyzed subpopulations.

Third, we searched for genetic barriers [14] in our dataset by using the same genetic and geographic distance matrices. This analysis identified two statistically significant barriers for the 23 subpopulations. One barrier was observed between the Finnish and all other subpopulations (first PC considering FI against the rest: $r^2 = 0.074$, two-tailed $p < 10^{-15}$; second PC considering FI against the rest: $r^2 = 0.33$, two-tailed $p < 10^{-15}$) and the other one between the two Italian and all other subpopulations (first PC considering IT1 and IT2 against the

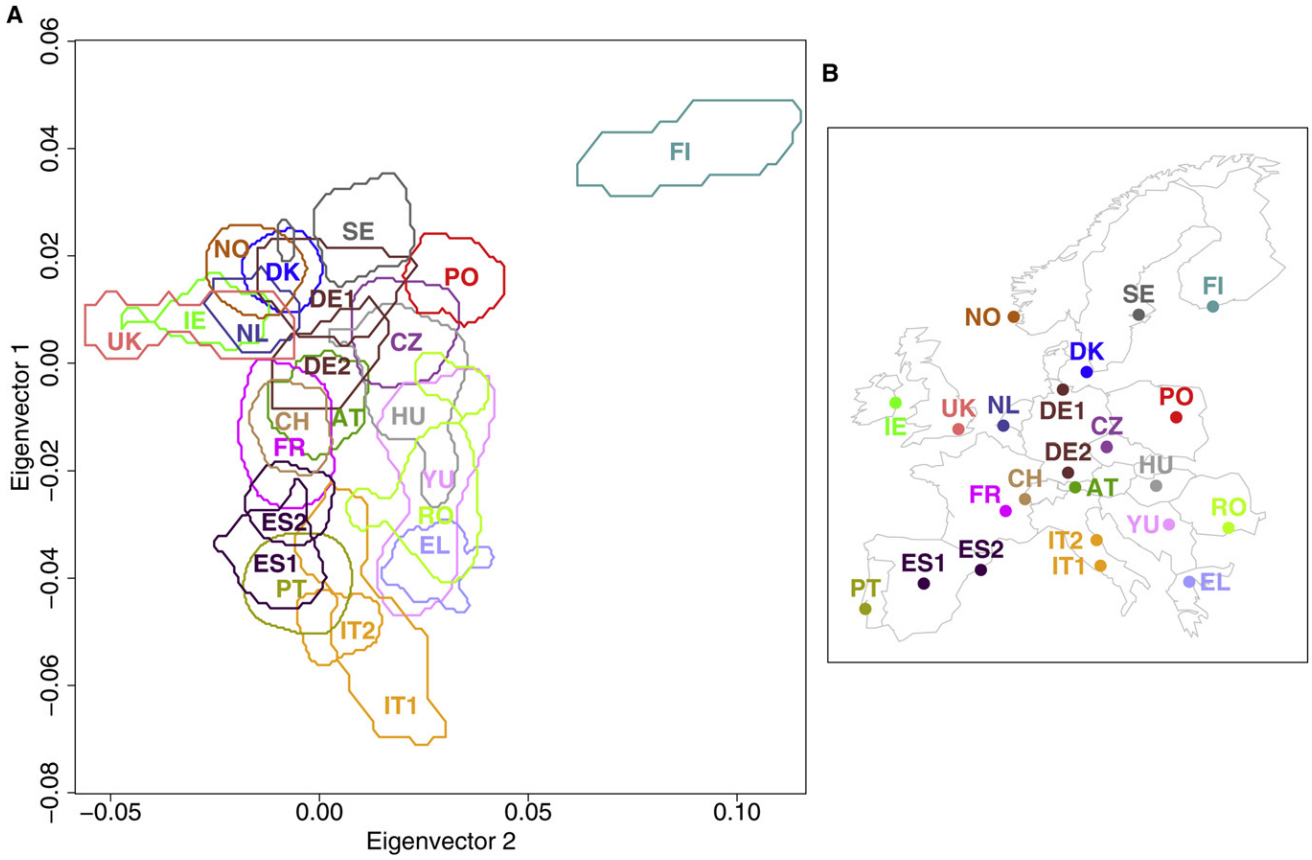


Figure 1. SNP-Based PCA of 2,457 European Individuals from 23 Subpopulations

(A) Kernel density plot of the first two dimensions of a SNP-based PCA using those 309,790 SNPs from the GeneChip Human Mapping 500K Array Set (Affymetrix) that passed quality control.

(B) Geographic distribution of the 23 subpopulations; capitals were used as the respective landmark if location information was either unspecific or lacking (see Table 1 for further sample details).

rest: $r^2 = 0.37$, two-tailed $p < 10^{-15}$; second PC considering IT1 and IT2 against the rest: $r^2 = 0.014$, two-tailed $p = 2.31 \times 10^{-9}$).

Fourth, we studied the geographic distribution of genetic diversity by computing mean heterozygosity and mean linkage disequilibrium (LD) based upon HR^2 [15] between markers at a distance < 10 kb for each subpopulation. Results from both analyses showed that the genetic diversity tended to be larger, and the LD smaller, in southern Europe as compared to northern Europe (Figure 2). Moreover, both analyses supported a genetic gradient of south-north orientation (r^2 adjusted for the number of data points between the mean observed heterozygosity and latitude: 0.76, $p = 3.80 \times 10^{-8}$; adjusted r^2 between HR^2 and latitude: 0.71, two-tailed $p = 4.33 \times 10^{-7}$) but not of west-east orientation (adjusted r^2 between heterozygosity and longitude: 0.03, two-tailed p = 0.416; adjusted r^2 between HR^2 and longitude: 0.099, two-tailed p = 0.078). Spatial autocorrelation analysis of both variables revealed statistically significant ($p < 0.05$) patterns compatible with a clinal distribution as indicated by the presence of positive and statistically significant autocorrelation values for small pair-wise distances and negative and statistically significant Moran's I values for large distances (see Figure 2). Bearing analysis [16] revealed for the heterozygosity measure the maximal angular correlations (r = 0.69) at 87° and the minimal (r = −0.153) at 165°, as well as for HR^2 the maximal at 55° (r = 0.67) and the minimal (r = −0.167) at 160°, thus also suggesting a south-to-north spatial distribution of both variable. These results are compatible with larger effective population sizes in the south than in the north of Europe and/or a population expansion from southern toward northern Europe. Hierarchical analysis of molecular variance (AMOVA) [17] revealed that clustering the individuals according to four geographic groups—north (NO, SE, FI), north-west/central (IE, UK, DK, NL, DE1, DE2, AT, CH, FR), east (HU, RO, PO, CZ), and south (PT, ES1, ES2, IT1, IT2, YU, EL)—explained an average of 0.17% (95% coefficient interval: 0.0% to 0.91%) of the total genetic variance, whereas individual subpopulation affiliation explained 0.25% (95% coefficient interval: 0.0% to 1.25%).

Overall, our study showed that the autosomal gene pool in Europe is comparatively homogeneous but at the same time revealed that the small genetic differentiation that is present between subpopulations is characterized by a significant correlation between genetic and geographic distance. Furthermore, the qualitative nature of these results is in close agreement with expectations based on human migration history in Europe. The major prehistoric waves of human migration in Europe followed south and southeastern to north and northwestern directions [1], including the first Paleolithic settlement of the continent by anatomically modern humans [18], most of the postglacial resettlement during the Mesolithic [19], and the farming-related population expansion during the Neolithic [18, 20]. Thus, both the level and the change in neutral autosomal

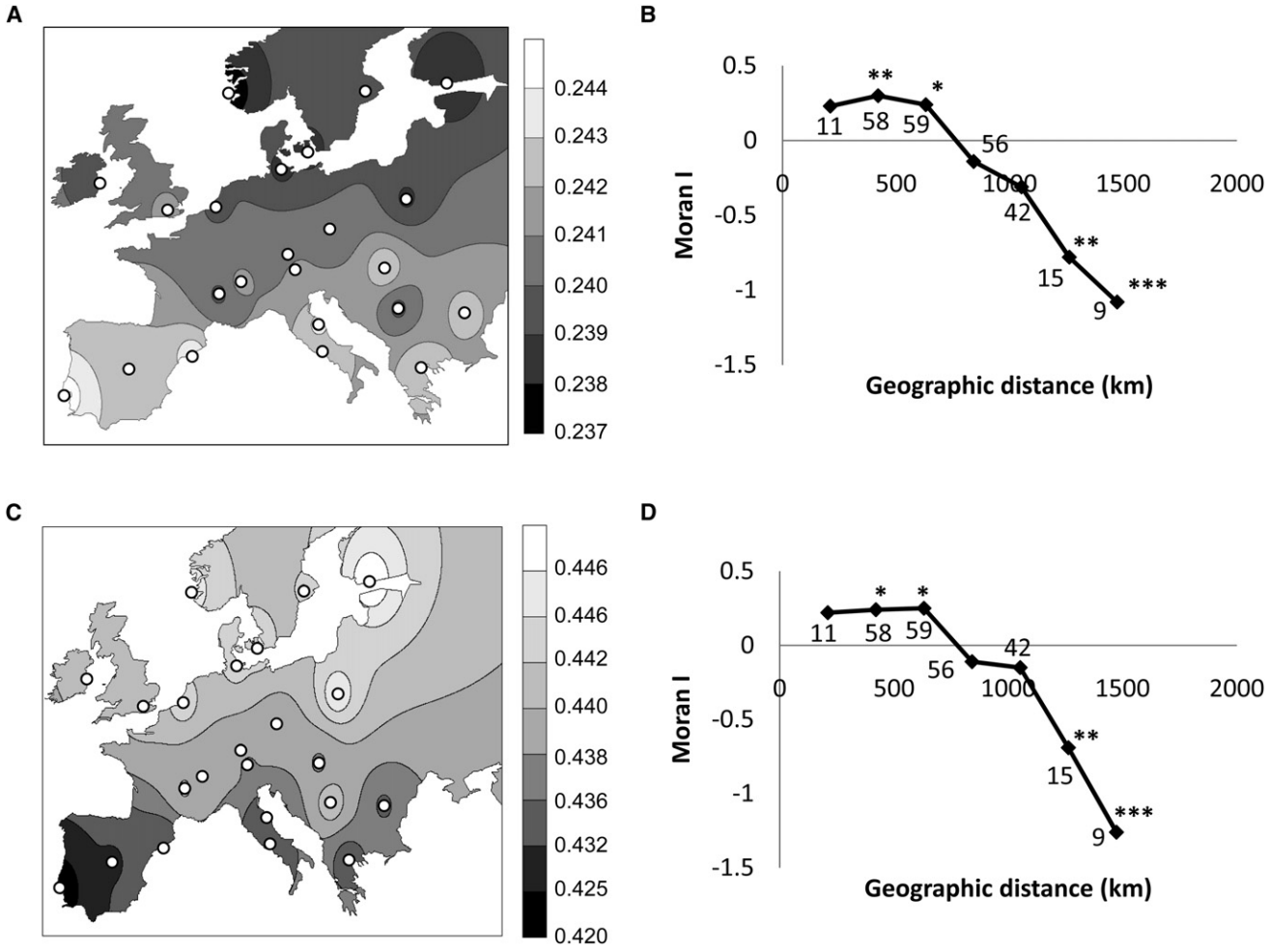


Figure 2. Geographic Distribution of Two Measures of Genetic Diversity across the European Population

(A and B) Isoline map (A) of Europe based on the mean observed heterozygosity in each of 23 European subpopulations with (B) corresponding spatial autocorrelogram.

(C and D) Isoline map (C) of Europe based on the mean observed linkage disequilibrium based on HR^2 in each of 23 European subpopulations with (D) corresponding spatial autocorrelogram. Both spatial autocorrelograms showed statistically significant departures from randomness ($p < 0.05$). For each distance class, the number of subpopulation pairs included and the statistical significance (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$) are provided.

variation in Europe can be expected to roughly follow southern-to-northern gradients as we observed, with the possible exception of population isolates as observed for the Finns. On the other hand, migration events in more recent (i.e., historic) times are presumed to have had a more homogenizing effect upon the previously established genetic landscape, as a result of their sporadic nature and haphazard geographic orientation [2]. This implies that genetic differences between extant European subpopulations can be expected to be small indeed. The genetic landscape described by the ~300,000 autosomal SNPs analyzed here closely resembles that previously obtained with 128 alleles from 49 classical markers (see Table 1.3.1 in [1]). This similarity is highlighted by a significant correlation ($r = 0.516$; two-tailed Mantel test $p = 0.0042$, performed with 10,000 Monte Carlo permutations) between the pair-wise F_{ST} values [21] computed for the 19 European subpopulations that overlapped between the two datasets (Danish, Dutch, Yugoslavian, Hungarian, Irish, Italian, Portuguese, Spanish, Swiss, English, German, Austrian, Finnish, French, Greek, Norwegian, Polish, Swedish, and Czechoslovakian). This notwithstanding, a stronger correlation between F_{ST} and great-circle geographic distances was observed for the subpopulations when the SNPs from our study were used ($r = 0.661$; two-tailed Mantel test $p = 0.00010$, performed with 10,000 Monte Carlo permutations) as compared to the classical markers ($r = 0.503$, two-tailed Mantel test $p = 0.00020$, performed with 10,000 Monte Carlo permutations).

Previous studies based on genome-wide SNP diversity reported differences between individuals of southern and northern/central European ancestry [3, 5, 6] and, to a lesser extent, between those of eastern and western European ancestry [3], which were not confirmed in our study. They mostly relied on the analysis of European Americans whose geographic assignment was determined from self-reported family records. Although genetic studies using European Americans can reveal important information about the genetic structure of the European ancestry of European Americans, caution must be exercised when drawing conclusions about the current genetic structure of Europe from European Americans because (1) European migrants may not have been representative of their country of origin, (2) the temporal difference introduced by sampling second- or third-generation descendants means

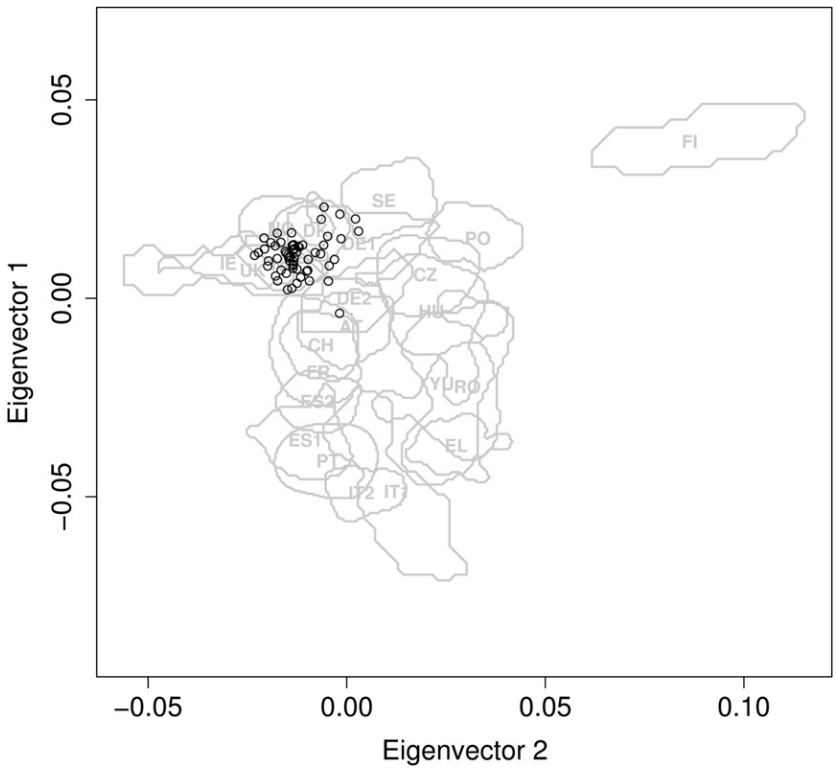


Figure 3. Position of CEPH-CEU Samples in a SNP-Based PCA Kernel-Density Plot of 23 European Subpopulations

CEU individuals (U.S. Americans of European descent from Utah) are plotted as open circles. For details, see Figure 1 and Table 1.

that allele-frequency estimates inevitably ignored recent population movements (i.e., WWII-related migrations), and (3) self-reported geographic origin is error prone [22]. Our study avoided these potential pitfalls by using large samples of individuals of genuinely European origin, as evidenced by the documentation of their respective place of birth or residence being in one of the named subpopulations, and with comprehensive continent-wide coverage.

It is of general interest to place the CEU samples, widely used in genetic epidemiological and population genetic studies as representing the European population, into the context of our findings. The CEPH-CEU panel comprises U.S. Americans who were collected in Utah in 1980 and who are assumed to have descended from migrants originating from northern and western parts of Europe [23]. The samples were also included in the International HapMap Project and formed the basis of selecting tagging SNPs used in current genome-wide association studies with Illumina SNP arrays. Whereas a previous study [3] confirmed the grouping of the CEPH-CEU samples with other northern and western European subpopulations, our study was capable of providing their most precise positioning on the European genetic map (Figure 3). It turned out that, while the CEPH-CEU panel was indeed largely representative of northwestern and central Europeans, parts of Scandinavia as well as southern and eastern Europe were not well represented by these samples (Figure 3). Estimated inflated false-positive rates for all subpopulations were largest in the Finns, followed by the two Italian subpopulations (see Table S2). This implies that researchers conducting genetic-association studies in at least these regions, using the CEPH-CEU samples as controls, may be at increased risk of false-positive associations. Our confirmation of the regional European origin of the CEPH-CEU samples also indicates that inferring the geographic origin of an unknown person from autosomal DNA markers, which is highly relevant in the forensic context, might now be feasible down to the level of European subregions, at least when a large number of genetic markers and a reference database, such as are applied here, are used.

Conclusions

Our comprehensive SNP genotype data from 23 European subpopulations, providing a dense coverage at both the geographic and genomic level and representing the largest Europe-wide genetic study to date, allowed us to describe the genetic structure of the European population with the highest resolution. Although the amount of differentiation within the European autosomal gene pool was found to be small, the existing genetic differences nevertheless correlated well with geographic distances. Furthermore, mean heterozygosity was larger, and mean linkage disequilibrium smaller, in southern than in northern European subpopulations, and both parameters exhibited a continuous clinal distribution across Europe. Overall, our results were compatible with expectations based on European population history, mainly the prehistoric population expansion from southern to northern Europe and/or a larger effective population size in the south as compared to the north of Europe. Our dataset also allowed placement of the widely used CEPH-CEU samples onto the European genetic landscape, essentially confirming their genetic ancestry in northern and western Europe.

Experimental Procedures

Samples and Genotyping

The GeneChip Human Mapping 500K Array Set (Affymetrix) was used to genotype 500,568 SNPs in 2,514 individuals from 23 different sampling sites (henceforth termed "subpopulations") located in one of 20 different European countries. Genotyping according to the instructions provided by the manufacturer was carried out at one of seven specialized centers: the Cologne Center for Genomics at the University of Cologne (Germany) for DE1, NO, SE, FI, AT, FR, ES2, IT2, EL, PO, and CZ; the Helmholtz Zentrum München - German Research Center for Environmental Health for DE2; the genetics laboratory of the Department of Internal Medicine, Erasmus MC (Netherlands) for NL; and the RH Microarray Centre Rigshospitalet, Copenhagen University Hospital (Denmark) for DK (see Table 1 for abbreviation explanations). Samples from the GlaxoSmithKline-sponsored POPRES project (IE, UK, CH, PT, ES1, IT1, YU, HU, and RO) were genotyped at Expression Analysis (Durham, NC, USA) and at Gene Logic (Gaithersburg, MD, USA) (see Table 1 for abbreviation explanations). Some samples belonged to existing control population studies, with detailed descriptions available elsewhere: KORA [24] for DE2, PopGen [25] for DE1, the Rotterdam Study [26–28] for NL, and POPRES (drawn from the LOLIPOP and CoLaus studies) for IE, UK, CH, PT, ES1, IT1, YU, HU, and RO [29–31]. Samples were drawn randomly from these pools or, in the case of POPRES, were ascertained on the basis of sample-size requirements. European migrants from non-European regions were not included in the initial analysis. For 11 of the subpopulations (NO, SE, FI, AT, FR, ES2, IT2, EL, PO, CZ, and DK), samples were

obtained from healthy unrelated volunteers: Norwegian samples (NO) from blood donors of the Førde region, Swedish samples (SE) from the Uppsala region [32], Finnish samples (FI) from the Helsinki area with parents and grandparents originating from various regions in Finland, Austrian samples (AT) from the Tyrol region with parents originating from Tyrol, French samples (FR) from blood donors of Lyon with parents originating from the Rhône Alpes area, Spanish samples (ES2) from Catalonia of blood donors from rural areas who speak Catalan as their mother tongue and who had regional Catalan ancestry for at least two generations [33], Italian samples (IT2) from blood donors of the upland of the Marches region [34], Greek samples (EL) from the north of the country [35], Polish samples (PO) from the Warsaw region of central Poland [36], Czech samples (CZ) from the central Bohemian region in and around Prague, and Danish samples (DK) from the Danish Blood Donor Corps in the Copenhagen area. In addition, GeneChip Human Mapping 500K Array data from CEPH-CEU samples were retrieved from the Affymetrix website (http://www.affymetrix.com).

Quality Assessment and Control Procedure

Array-based SNP genotypes were subjected to stringent quality control: First, each individual was required to have a genotype call rate $\geq$ 93%, with the dynamic model (DM) algorithm with a confidence score of 0.26, and a per-individual call rate $\geq$ 95% for all individuals genotyped by the same facility, with the Bayesian robust linear model with Mahalanobis distance classifier (BRLMM) algorithm with a confidence score of 0.5. The call rate was defined here as the proportion of unambiguous genotypes among either all SNPs (per-individual call rate) or all individuals (per-marker call rate), respectively. Markers that were monomorphic (1.4% of the total), that were located on the X chromosome (2.1%), or that had a per-marker call rate $\leq$ 90% in at least one genotyping facility (5.7%) were excluded, as were those showing a significant ($p \leq 0.05$) deviation from Hardy-Weinberg equilibrium (HWE) in at least one subpopulation (31.3%). HWE was tested by means of a χ^2 test, or by Fisher's exact test when the observed or expected number of a given genotype was less than 5. This method was preferred over others that have been shown to be more powerful [37] because the computational requirements of these methods increase exponentially with sample size and were thus too resource intensive for our study. The average proportion of heterozygous genotypes at X chromosomal markers was estimated per individual in order to detect false gender assignments. Male subjects can be expected to show X chromosomal heterozygosity proportions $\leq$ 1%, reflecting the overall genotyping error rate, and female subjects should show proportions near the average heterozygosity (26%) of the analyzed X chromosomal SNPs. Average identity-by-state (IBS) distances were calculated for a given set of markers as the average genetic dissimilarity between pairs of individuals. Analysis of IBS values within subpopulations allowed us to detect two types of outliers: (1) cognate relatives, i.e., individuals that were genetically more similar than expected to another member of the same subpopulation, and (2) "aliens," i.e., individuals that were far less genetically similar than expected to the rest of the subpopulation. Formally, cognate relatives were defined as pairs of individuals having a pair-wise IBS value larger than the so-called "Tukey outlier criterion" when compared with the rest of pairs of individuals of the same subpopulation, i.e., the median IBS plus three times the interquartile range (IQR) in that subpopulation. In this case, the partner with the lower call rate was excluded. Aliens were defined as individuals with at least 60% of their pair-wise IBS values below the median minus three times the IQR. These two criteria led to the exclusion of 56 individuals from further analysis (Table 1). One individual identified as female had an average proportion of heterozygous X chromosomal markers of only 0.6% and was thus excluded from further analysis. In total, quality control left 2,457 individuals (97.6%) and 309,790 markers (62.4%) for inclusion in subsequent analysis. AMOVA [17] was performed to ascertain the magnitude of variation attributable to the respective genotyping center or subpopulation. The mean amount of genetic variance explained among genotyping centers was 0.095% (95% confidence interval: 0% to 0.71%), whereas subpopulation affiliation explained 0.63% of the variance (95% confidence interval: 0% to 2.86%). As expected, the largest amount of genetic variation was explained by differences between individuals (99.72%; 95% confidence interval: 98.61% to 100.00%). Data are available on request from the authors according to the regulations of the participating studies and sample cohorts.

Statistical Data Analyses

The ancestry-informativeness index *In* was estimated for each marker as described eslewhere [12]. Principal-component analysis was performed with the *Eigensoft* program with the default settings [38]. Population-wise kernel densities were computed from the first two PCs with the *adehabitat* R package [39] and subjected to least-squares crossvalidation [40] that used 80% of individuals per subpopulation for training. Pearson correlation coefficients were computed for the genetic distance between the subpopulations (represented by the respective median over all individuals in that subpopulation of the first two eigenvectors) and the great-circle geographic distance. The statistical significance of these correlation coefficients was assessed by means of a Mantel test [41]. Barrier analysis was performed on the basis of the Monmonier's algorithm [14]. Locus-wise AMOVA [17] was conducted after clustering the European subpopulations by genotyping center as well as by the use of four geographic groups. Negative percentages of explained variation were settled to 0. Both mean heterozygosity and mean linkage disequilibrium computed by means of HR^2 [15] were computed with a subsample of ten individuals per population in order to adjust for possible influence of sample size [42]. Spatial autocorrelation and Bearing analyses were performed with the software PASSAGE 1.1 [43]. Isoline maps were performed with the Golden Surfer 8 software [44], with the inverse-distance method used for interpolation points. Isoline levels were defined to include the value of at least one of the 23 populations with intervals of 0.001 in the case of heterozygosity and 0.002 in the case of HR^2. For evaluation of the extent to which the CEPH-CEU samples are representative of the subpopulations used in the present study, marker-wise tests of association (Fisher's exact test) were performed each time with the CEPH-CEU samples as "controls" and a given subpopulation as "cases." The false-positive rate was defined as the percentage of markers yielding a p value < 0.05. If the CEPH-CEU samples were representative of a subpopulation, the false-positive rate would be around 0.05, whereas higher false-positive rates indicate that the CEPH-CEU samples may not be representative of the respective subpopulation.

Supplemental Data

Supplemental Data include two tables and two figures and can be found with this article online at http://www.current-biology.com/cgi/content/full/18/16/1241/DC1/.

Acknowledgments

All volunteers are gratefully acknowledged for sample donation. We thank the following colleagues for their help and support: J. Kooner and J. Chambers of the LOLIPOP study and D. Waterworth, V. Mooser, G. Waeber, and P. Vollenweider of the CoLaus study for providing access to their collections via the GlaxoSmithKline-sponsored Population Reference Sample (POPRES) project; K. King for preparing the POPRES data; M. Simoons, E. Sijbrands, A. van Belkum, J. Laven, J. Lindemans, E. Knipers, and B. Stricker for their financial contribution to the generation of the Rotterdam Study dataset; P. Arp, M. Jhamai, W. van IJken, and R. van Schaik for generating the Rotterdam Study dataset; T. Meitinger, P. Lichtner, G. Eckstein, and all other members of the Helmholtz Zentrum München genotyping staff for generating the KORA Study dataset; H. von Eller-Eberstein for management of the PopGen project; F.C. Nielsen, R. Borup, C. Schjerling, H. Ullum, E. Haastrup, and numerous colleagues at the Copenhagen University Hospital Blood Bank for assistance in making the Danish data available; and S. Brauer for DNA sample management. We would additionally like to thank Affymetrix for making the GeneChip Human Mapping 500K Array genotypes of the CEPH-CEU trios publicly available and the Centre d'Etude du Polymorphisme Humain (CEPH) for the original sample collection. We are grateful to three anonymous reviewers for their comments, which stimulated us to improve the manuscript. This work was supported by the Netherlands Forensic Institute to M.Ka.; Affymetrix to M.Ka. and M.Kr.; the German National Genome Research Network and the German Federal Ministry of Education and Research to H.-E.W., S.S., M.Kr., and P.N. (01GR0416 to P.N.); the Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, and the Munich Center of Health Sciences as part of LMUinnovativ to H.-E.W.; the Netherlands Organization for Scientific Research (NWO 175.010.2005.011) to A.G.U.; the European Commission to A.G.U. (GEFOS; 201865) and A.S. (LD Europe; QLG2-CT-2001-00916); the Czech Ministry of Health (VZFNM 00064203 and IGA NS/9488-3) to M.M.; Helse-Vest, Regional Health Authority Norway to L.A.B., the Swedish National Board of Forensic Medicine (RMVFoU 99:22, 02:20) to G.H.; and the Academy of Finland to A.S. (80578, OMLL) and J.P. (109265 and 111713). None of the funding organizations had any influence on the design, conduct, or conclusions of the study.

Received: May 2, 2008
Revised: July 9, 2008
Accepted: July 10, 2008
Published online: August 7, 2008

References

1. Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1994). The History and Geography of Human Genes (Princeton, NJ: Princeton University Press).
2. Sokal, R.R., Harding, R.M., and Oden, N.L. (1989). Spatial patterns of human gene frequencies in Europe. Am. J. Phys. Anthropol. *80*, 267–294.
3. Bauchet, M., McEvoy, B., Pearson, L.N., Quillen, E.E., Sarkisian, T., Hovhannesyan, K., Deka, R., Bradley, D.G., and Shriver, M.D. (2007). Measuring European population stratification with microarray genotype data. Am. J. Hum. Genet. *80*, 948–956.
4. Price, A.L., Butler, J., Patterson, N., Capelli, C., Pascali, V.L., Scarnicci, F., Ruiz-Linares, A., Groop, L., Saetta, A.A., Korkolopoulou, P., et al. (2008). Discerning the ancestry of European Americans in genetic association studies. PLoS Genet *4*, e236.
5. Tian, C., Plenge, R.M., Ransom, M., Lee, A., Villoslada, P., Selmi, C., Klareskog, L., Pulver, A.E., Qi, L., Gregersen, P.K., et al. (2008). Analysis and application of European genetic substructure using 300 K SNP information. PLoS Genet *4*, e4.
6. Seldin, M.F., Shigeta, R., Villoslada, P., Selmi, C., Tuomilehto, J., Silva, G., Belmont, J.W., Klareskog, L., and Gregersen, P.K. (2006). European population substructure: Clustering of northern and southern populations. PLoS Genet *2*, e143.
7. Sajantila, A., Salem, A.H., Savolainen, P., Bauer, K., Gierig, C., and Paabo, S. (1996). Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc. Natl. Acad. Sci. USA *93*, 12035–12039.
8. Roewer, L., Croucher, P.J., Willuweit, S., Lu, T.T., Kayser, M., Lessig, R., de Knijff, P., Jobling, M.A., Tyler-Smith, C., and Krawczak, M. (2005). Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum. Genet. *116*, 279–291.
9. Rosser, Z.H., Zerjal, T., Hurles, M.E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., et al. (2000). Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am. J. Hum. Genet. *67*, 1526–1543.
10. Kayser, M., Lao, O., Anslinger, K., Augustin, C., Bargel, G., Edelmann, J., Elias, S., Heinrich, M., Henke, J., Henke, L., et al. (2005). Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. Hum. Genet. *117*, 428–443.
11. Simoni, L., Calafell, F., Pettener, D., Bertranpetit, J., and Barbujani, G.V. (2000). Geographic patterns of mtDNA diversity in Europe. Am. J. Hum. Genet. *66*, 262–278.
12. Rosenberg, N.A., Li, L.M., Ward, R., and Pritchard, J.K. (2003). Informativeness of genetic markers for inference of ancestry. Am. J. Hum. Genet. *73*, 1402–1422.
13. Voight, B.F., Kudaravalli, S., Wen, X., and Pritchard, J.K. (2006). A map of recent positive selection in the human genome. PLoS Biol. *4*, e72.
14. Manni, F.C., Guérard, E., and Heyer, G.E. (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by "Monmonier's algorithm.". Hum. Biol. *76*, 173–190.
15. Sabatti, C., and Risch, N. (2002). Homozygosity and linkage disequilibrium. Genetics *160*, 1707–1719.
16. Falsetti, A.B., and Sokal, R.R. (1993). Genetic structure of human populations in the British Isles. Ann. Hum. Biol. *20*, 215–229.
17. Excoffier, L., Smouse, P.E., and Quattro, J.M.V. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics *131*, 479–491.
18. Belle, E.M., Landry, P.A., and Barbujani, G. (2006). Origins and evolution of the Europeans' genome: Evidence from multiple microsatellite loci. Proc Biol Sci *273*, 1595–1602.
19. Torroni, A., Bandelt, H.J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., et al. (2001). A signal, from human mtDNA, of postglacial recolonization in Europe. Am. J. Hum. Genet. *69*, 844–852.
20. Chikhi, L., Nichols, R.A., Barbujani, G., and Beaumont, M.A.V. (2002). Y genetic data support the Neolithic demic diffusion model. Proc. Natl. Acad. Sci. USA *99*, 11008–11013.
21. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. Evolution Int. J. Org. Evolution *38*, 1358–1370.
22. Burnett, M.S., Strain, K.J., Lesnick, T.G., de Andrade, M., Rocca, W.A., and Maraganore, D.M. (2006). Reliability of self-reported ancestry among siblings: Implications for genetic association studies. Am. J. Epidemiol. *163*, 486–492.
23. Dausset, J., Cann, H., Cohen, D., Lathrop, M., Lalouel, J.M., and White, R. (1990). Centre d'etude du polymorphisme humain (CEPH): Collaborative genetic mapping of the human genome. Genomics *6*, 575–577.
24. Lowel, H., Doring, A., Schneider, A., Heier, M., Thorand, B., Meisinger, C., and Group, M.K.S. (2005). The MONICA Augsburg surveys–basis for prospective cohort studies. Gesundheitswesen *67* (*Suppl 1*), S13–S18.
25. Krawczak, M., Nikolaus, S., von Eberstein, H., Croucher, P.J., El Mokhtari, N.E., and Schreiber, S. (2006). PopGen: Population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community Genet. *9*, 55–61.
26. Hofman, A., Breteler, M.M., van Duijn, C.M., Krestin, G.P., Pols, H.A., Stricker, B.H., Tiemeier, H., Uitterlinden, A.G., Vingerling, J.R., and Witteman, J.C. (2007). The Rotterdam Study: Objectives and design update. Eur. J. Epidemiol. *22*, 819–829.
27. Hofman, A., Grobbee, D.E., de Jong, P.T., and van den Ouweland, F.A. (1991). Determinants of disease and disability in the elderly: The Rotterdam Elderly Study. Eur. J. Epidemiol. *7*, 403–422.
28. Kayser, M., Liu, F., Janssens, A.C., Rivadeneira, F., Lao, O., van Duijn, K., Vermeulen, M., Arp, P., Jhamai, M.M., van Ijcken, W.F., et al. (2008). Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. Am. J. Hum. Genet. *82*, 411–423.
29. Kooner, J.S., Chambers, J.C., Aguilar-Salinas, C.A., Hinds, D.A., Hyde, C.L., Warnes, G.R., Gomez Perez, F.J., Frazer, K.A., Elliott, P., Scott, J., et al. (2008). Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. Nat. Genet. *40*, 149–151.
30. Nelson, M.R., Bacanu, S.A., Mosteller, M., Li, L., Bowman, C.E., Roses, A.D., Lai, E.H., and Ehm, M.G. (2008). Genome-wide approaches to identify pharmacogenetic contributions to adverse drug reactions. Pharmacogenomics J., in press. Published online February 26, 2008. 10.1038/tpj.2008.4.
31. Sandhu, M.S., Waterworth, D.M., Debenham, S.L., Wheeler, E., Papadakis, K., Zhao, J.H., Song, K., Yuan, X., Johnson, T., Ashford, S., et al. (2008). LDL-cholesterol concentrations: A genome-wide association study. Lancet *371*, 483–491.
32. Karlsson, A.O., Wallerstrom, T., Gotherstrom, A., and Holmlund, G. (2006). Y-chromosome diversity in Sweden - a long-time perspective. Eur. J. Hum. Genet. *14*, 963–970.
33. Plaza, S., Calafell, F., Helal, A., Bouzerna, N., Lefranc, G., Bertranpetit, J., and Comas, D. (2003). Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. Ann. Hum. Genet. *67*, 312–328.
34. Onofri, V., Alessandrini, F., Turchi, C., Fraternale, B., Buscemi, L., Pesaresi, M., and Tagliabracci, A. (2007). Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis. Int. J. Legal Med. *121*, 234–237.
35. Kondopoulou, H., Loftus, R., Kouvatsi, A., and Triantaphyllidis, C. (1999). Genetic studies in 5 Greek population samples using 12 highly polymorphic DNA loci. Hum. Biol. *71*, 27–42.
36. Ploski, R., Wozniak, M., Pawlowski, R., Monies, D.M., Branicki, W., Kupiec, T., Kloosterman, A., Dobosz, T., Bosch, E., Nowak, M., et al. (2002). Homogeneity and distinctiveness of Polish paternal lineages revealed by Y chromosome microsatellite haplotype analysis. Hum. Genet. *110*, 592–600.
37. Schaid, D.J., Batzler, A.J., Jenkins, G.D., and Hildebrandt, M.A. (2006). Exact tests of Hardy-Weinberg equilibrium and homogeneity of disequilibrium across strata. Am. J. Hum. Genet. *79*, 1071–1080.
38. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet *2*, e190.
39. Calenge, C. (2006). The package "adehabitat" for the R software: A tool for the analysis of space and habitat use by animals. Ecol. Modell. *197*, 516–519.

40. Silverman, B.W. (1986). Density Estimation for Statistics and Data Analysis (Boca Raton, Florida: Chapman & Hall / CRC Press).
41. Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. Cancer Res. *27*, 209–220.
42. Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R., et al. (2008). Genotype, haplotype and copy-number variation in worldwide human populations. Nature *451*, 998–1003.
43. Rosenberg, M.S. (2001). PASSAGE: Pattern Analysis, Spatial Statistics, and Geographic Exegesis. 1.1 Edition, A.S.U. Department of Biology, ed. (Tempe, AZ.).
44. Golden Software. (2007). Surfer Version 8.08.3267. Colorado, USA.

European Journal of Human Genetics (2009) 17, 967 – 975

www.nature.com/ejhg

ARTICLE

An evaluation of the genetic-matched pair study design using genome-wide SNP data from the European population

Timothy Tehua Lu[1,23], Oscar Lao[2,23], Michael Nothnagel[1], Olaf Junge[1], Sandra Freitag-Wolf[1], Amke Caliebe[1], Miroslava Balascakova[3], Jaume Bertranpetit[4], Laurence Albert Bindoff[5], David Comas[4], Gunilla Holmlund[6], Anastasia Kouvatsi[7], Milan Macek[3], Isabelle Mollet[8], Finn Nielsen[9], Walther Parson[10], Jukka Palo[11], Rafal Ploski[12], Antti Sajantila[11], Adriano Tagliabracci[13], Ulrik Gether[14], Thomas Werge[15], Fernando Rivadeneira[16,17], Albert Hofman[17], André Gerardus Uitterlinden[16,17], Christian Gieger[18], Heinz-Erich Wichmann[18,19], Andreas Ruether[20], Stefan Schreiber[20], Christian Becker[21], Peter Nürnberg[21], Matthew Roberts Nelson[22], Manfred Kayser[2,23] and Michael Krawczak*[,1,23]

[1]*Institut für Medizinische Informatik und Statistik, Christian-Albrechts-Universität Kiel, Kiel, Germany;* [2]*Department of Forensic Molecular Biology, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands;* [3]*Department of Biology and Medical Genetics, University Hospital Motol and 2nd School of Medicine, Charles University Prague, Prague, Czech Republic;* [4]*Institute of Evolutionary Biology (UPF-CSIC), CEXS-UPF-PRBB, Universitat Pompeu Fabra, Barcelona, Spain;* [5]*Department of Neurology, Haukeland University Hospital and Department of Clinical Medicine, University of Bergen, Bergen, Norway;* [6]*Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden;* [7]*Department of Genetics, Development and Molecular Biology, Aristotle University of Thessaloniki, Thessaloniki, Greece;* [8]*Laboratoire d'Empreintes Génétiques, EFS-RA site de Lyon, Lyon, France;* [9]*Department of Clinical Biochemistry and Center for Pharmacogenomics, University of Copenhagen, Copenhagen, Denmark;* [10]*Institute of Legal Medicine, Medical University Innsbruck, Innsbruck, Austria;* [11]*Department of Forensic Medicine, University of Helsinki, Helsinki, Finland;* [12]*Department of Medical Genetics, Warsaw Medical University, Warsaw, Poland;* [13]*Istituto di Medicina Legale, Università di Ancona, Ancona, Italy;* [14]*Molecular Neuropharmacology Group and Center for Pharmacogenomics, Department of Neuroscience and Pharmacology, University of Copenhagen, Copenhagen, Denmark;* [15]*Research Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Copenhagen University Hospital, and Center for Pharmacogenomics, University of Copenhagen, Copenhagen, Denmark;* [16]*Department of Internal Medicine, Genetics Laboratory, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands;* [17]*Department of Epidemiology and Biostatistics, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands;* [18]*Institute of Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health (GmbH), Neuherberg, Germany;* [19]*Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany;* [20]*Institut für Medizinische Molekularbiologie, Christian-Albrechts-Universität Kiel, Kiel, Germany;* [21]*Cologne Center for Genomics and Institut für Genetik, Universität zu Köln, Köln, Germany;* [22]*Genetics, GlaxoSmithKline, Research Triangle Park, NC, USA*

Genetic matching potentially provides a means to alleviate the effects of incomplete Mendelian randomization in population-based gene–disease association studies. We therefore evaluated the genetic-matched pair study design on the basis of genome-wide SNP data (309 790 markers; Affymetrix GeneChip Human Mapping 500K Array) from 2457 individuals, sampled at 23 different recruitment sites across

***Correspondence: Professor Dr M Krawczak, Institut für Medizinische Informatik und Statistik, Christian-Albrechts-Universität Haus 31, Arnold-Heller-Straße 3, Kiel 24105, Germany. Tel: +49 431 597 3200; Fax: +49 431 597 3193; E-mail: krawczak@medinfo.uni-kiel.de**
[23]These authors contributed equally to this work.
Received 26 September 2008; revised 3 December 2008; accepted 10 December 2008; published online 21 January 2009

Europe. Using pair-wise identity-by-state (IBS) as a matching criterion, we tried to derive a subset of markers that would allow identification of the best overall matching (BOM) partner for a given individual, based on the IBS status for the subset alone. However, our results suggest that, by following this approach, the prediction accuracy is only notably improved by the first 20 markers selected, and increases proportionally to the marker number thereafter. Furthermore, in a considerable proportion of cases (76.0%), the BOM of a given individual, based on the complete marker set, came from a different recruitment site than the individual itself. A second marker set, specifically selected for ancestry sensitivity using singular value decomposition, performed even more poorly and was no more capable of predicting the BOM than randomly chosen subsets. This leads us to conclude that, at least in Europe, the utility of the genetic-matched pair study design depends critically on the availability of comprehensive genotype information for both cases and controls.
European Journal of Human Genetics (2009) **17,** 967–975; doi:10.1038/ejhg.2008.266; published online 21 January 2009

Keywords: population structure; matching; association; ancestry; microarray

Introduction

In both classical epidemiology and clinical research, potential confounders are usually controlled for by one of two different means, matching or randomization. In genetic studies, however, including the large number of genome-wide association (GWA) studies that have recently been published,[1–3] only so-called 'Mendelian' randomization has been employed to control for genetic confounders, whereas matching by genotype has not played an important role.[4] Nevertheless, there has always been some awareness among genetic epidemiologists that Mendelian randomization may fail, thereby leading to false positive reports of disease genes or to biased effect size estimates.[5] One possible cause of such failure may be systematic differences in terms of the rate at which individuals with a particular phenotype or genotype are sampled from genetically distinct populations. Therefore, two statistical methods to retrospectively rectify genetic imbalances in case-control studies were developed in the late 1990s, both of which rely upon genotyping loci that are unrelated to the genetic variants under study (ie unlinked and not in linkage disequilibrium). The 'genomic control' approach[6] uses marker genotypes to correct the employed test statistic, whereas 'structured association'[7] infers the number of populations represented in a sample, and then assigns each individual to one of these populations with a certain probability.

With the possibility to effectively genotype large numbers of single nucleotide polymorphisms (SNPs) in large numbers of individuals, using microarray technology,[8] the effects of imperfect Mendelian randomization can, in principle, also be alleviated by genetic matching. If individuals from different samples such as cases and controls were as closely matched as possible in terms of their identity-by-state (IBS) status at a large number of SNPs, it may be surmised that most systematic population genetic differences would be eliminated between the ensuing sub samples. However, genetic matching would have to be based on markers from outside the genomic region under study to avoid over-matching. This implies that, in practise, repeated matching may be necessary if multiple or even GWA assessments are due. In any case, genetic matching could of course be accomplished efficiently with the use of genome-wide microarray data, but such a costly strategy may not be necessary if a set of 'best genetic match' (BGM) markers could be established in advance that are capable of capturing the major population genetic characteristics of relevant extant populations. Once a set of BGM markers has been found, it can be used in two ways: either to retrospectively confirm whether two samples of interest were genetically well-matched or to select members of matched samples prospectively, before any additional genotyping.

Recruitment of phenotypically well-characterized control samples is one of the major bottlenecks of genetic epidemiological and pharmacogenetic research. The use of common controls across different association studies has proven to be an efficient solution to this problem, pioneered at a local level by the Wellcome Trust Case Control Consortium (WTCCC),[3] and since adopted, for example, by the US-American Genetic Association Information Network (GAIN)[1] and the German National Genome Research Network ('Nationales Genomforschungsnetz', NGFN).[9] However, the number and geographical distribution of control samples required for the common controls approach to be feasible at a broader geographical level are currently unknown.

In the present study, we investigated three issues related to the genetic-matched pair study design, using genome-wide SNP data from across Europe: (1) the prospects of identifying a small subset of SNPs that accurately predict the 'best' genome-wide matching partner of a given

individual, (2) the distribution of 'best' genetic-matching partners between the European subpopulations and (3) the inter-individual variability in terms of the uniqueness of the 'best' genetic-matching partner. To this end, we analyzed the genotypes of 309 790 markers obtained from the GeneChip Human Mapping 500K Array Set in 2457 individuals, ascertained at one of 23 recruitment sites. The European population is important in this context, not only because of the historical interest in these people and their descendants in the Americas, Australia and elsewhere, but also because they are a major focus of both genetic epidemiological and pharmacogenetic research.[1,3]

Material and methods

Samples, genotyping and quality control

The GeneChip Human Mapping 500K Array (Affymetrix) was used to genotype 500 568 SNPs in 2514 individuals from 23 different sampling sites (henceforth, termed 'subpopulations'), distributed over 20 different European countries. Subpopulation sizes ranged from 12 to 500 individuals (Table 1). Sex ratios differed markedly between subpopulations, with some comprising only females or males, respectively. Genotyping was carried out at six different facilities. For further details, see Lao *et al*.[10]

Array-based SNP genotypes were subjected to stringent quality control as described earlier.[10] Briefly, markers, which had a genotype call rate $\geq 93\%$, were monomorphic, located on the X chromosome or had a per marker call rate $\leq 90\%$ in at least one genotyping facility were excluded, as were those showing a significant ($P<0.05$) deviation from Hardy–Weinberg equilibrium (HWE) in at least one subpopulation. Individuals deemed genetic outliers to their subpopulation of origin, based on low average IBS to the remaining individuals, were omitted from the respective subpopulation. In total, quality control left 2457 individuals (97.6%) and 309 790 markers (62.4%) for inclusion in subsequent analyses. The set of quality controlled markers will henceforth be referred to as marker set C. Ascertainment of a marker set for genetic matching was carried out with internal validation, using 2/3 of the members of each subpopulation (ie, 1638 randomly chosen individuals) as the training set, and using the remainder (819 individuals) as the validation set (Table 1).

All data were stored as either flat files or in a customized database with an interface to the R statistical software. All data analysis, except for the IBS estimation, was done in R version 2.4.1[11] using customized scripts. IBS calculations and selection of marker sets were carried out using custom C++ programs. All software is available from the authors on request.

Best genetic match marker set

For the ascertainment of a marker subset M of C that would allow us to identify 'best' genetic-matching partners, we will use a set-specific criterion, $\Delta(M)$ that is related to the IBS between given individuals and their matching partners, as selected on the basis of M (see below). In this context, we will use the term 'best overall match' (BOM) to denote that individual or group of individuals who maximize the average pair-wise IBS with the individual of interest for the complete marker set C. Ideally, we would want to ascertain a subset of markers that consistently lead to the selection of matching partners with an IBS with the reference individual that is close to the IBS between the reference individual and its BOM.

More formally, if the genotype (g), of a given SNP is encoded by the dose of one of its two alleles (ie, as 0, 1 or 2), then the IBS between any two individuals x and y equals $1-|g(x)-g(y)|/2$ for that SNP. Here, $g(x)$ and $g(y)$ denote the genotypes of x and y, respectively. For a marker set M, let $i_M(x,y)$ be the average IBS, taken over all markers in M, and let $i_M(x)$ denote the maximum $i_M(x,y)$, taken over all individuals y other than x. Finally, if $M \subseteq N$ are two nested marker sets, let $i_{M,N}(x)$ be the average $i_N(x,y)$ taken over all y for which $i_M(x,y) = i_M(x)$. For a marker set $M \subseteq C$, $\Delta(M)$ is defined as the average difference $|i_C(x)-i_{M,C}(x)|$, taken over all individuals x and weighted by the inverse of the size of the subpopulation to which x belongs.

We used forward selection from marker set C to ascertain marker sets that successively minimized the Δ criterion. The ensuing marker sets will be referred to as the best genetic match (BGM) marker sets. Upper and lower baselines for Δ were computed as follows. The upper baseline was obtained from randomly chosen marker sets of varying size (10–100 in steps of 10), with 1000 sets sampled for each set size value. The lower baseline was obtained from marker sets that theoretically should have captured most of the genetic variation present in the individuals under study, ie sets for which any additional marker would have been in strong linkage disequilibrium with the markers already included. Each chromosome was thus divided into bins of 20 kb, based on the mean swept radius of 500 kb estimated for the European population.[12,13] The swept radius is the distance at which the average association between two markers, measured by r^2, is reduced to approximately one-third (more precisely, e^{-1}) of its initial value. A bin size of 20 kb therefore ensures an average r^2 of $e^{-10/500} = 0.98$ between markers in the bin. Markers were then randomly selected from bins, one at a time, and Δ calculated for the resulting marker set. The described selection process was repeated 1000 times and the mean Δ value taken as the lower baseline, ie the expectation of Δ at r^2-based saturation.

Ancestry-sensitive marker set

To compare the BGM set, which focuses on inter-individual genetic variation with a marker set that was ascertained with the aim to highlight inter-population variation, we generated an ancestry-sensitive marker (ASM) set using the

Table 1 European subpopulation summary statistics

Subpopulation	*Code*	*No. samples*	*Final no. samples*	*No. training*
Norway (Førde)	NO	52	52 (0.63)	35
Sweden (Uppsala)	SE	50	46 (1.00)	31
Finland (Helsinki)	FI	47	47 (0.43)	31
Ireland	IE	37	35 (0.80)	23
UK (London)	UK	197	194 (0.90)	129
Denmark (Copenhagen)	DK	60	59 (0.56)	39
Netherlands (Rotterdam)	NL	292	280 (0.00)	187
Germany I (Kiel)	DE1	500	494 (0.52)	329
Germany II (Augsburg)	DE2	500	489 (0.51)	326
Austria (Tyrol)	AT	50	50 (1.00)	33
Switzerland (Lausanne)	CH	134	133 (0.44)	89
France (Lyon)	FR	50	50 (0.68)	33
Portugal	PT	16	16 (0.44)	11
Spain I	ES1	83	81 (0.51)	54
Spain II (Barcelona)	ES2	48	47 (0.43)	31
Italy I	IT1	107	106 (0.58)	71
Italy II (Marche)	IT2	50	49 (1.00)	33
Former Yugoslavia	YU	58	55 (0.65)	37
Northern Greece	EL	51	51 (0.59)	34
Hungary	HU	17	17 (0.35)	11
Romania	RO	12	12 (0.50)	8
Poland (Warsaw)	PO	50	49 (1.00)	33
Czech Republic (Prague)	CZ	53	45 (0.51)	30
Total		2514	2457	1638

Subpopulation, site of sample origin, with more specific location details given in parentheses; No. samples, total number of samples genotyped; Final no. samples, number of samples that passed stringent quality control, with proportion of males in parenthesis (for details, see text); No. training, size of the training set used for marker selection.

singular value decomposition (SVD) method with redundant marker reduction described by Paschou *et al*.[14,15] Global allele frequencies were used to interpolate missing data as suggested by the authors. Some 228 individuals were eliminated from the training set during PCA analysis with Eigensoft2[16] using the standard criterion of having an ancestry coefficient >6 standard deviations in at least one of the eigenvector axes. SVD was carried out with SVDLIBC (version 1.34, http://tedlab.mit.edu/~dr/SVDLIBC), a C library based on the SVDPACK library.[17] Rank-revealing QR matrix decomposition was carried out in Octave version 2.0.17[18] to reduce the redundancy of the first 5000 markers, ordered by the first SVD eigenvector. This resulted in a set of the same size (ie 100 markers) as the BGM set.

Distribution of best genetic match pairs

A count matrix was generated that contains, for each pair of subpopulations, the number of times an individual in the first subpopulation had their BOM in the second population. Cell counts were tested for a deviation from the null hypothesis that BOMs were drawn randomly from subpopulations using a two-tailed exact test as implemented in the R routine *binom.test*. A plot of directed graphs representing the relationships between individuals and their BOMs was generated using Graphviz.[19]

False positive rates

Thresholds for the false positive rates of population-based gene–disease associations in Europe were determined from contrived case-control experiments, using PLINK version 1.03[20] on all markers in set C (Fisher's exact test on allele frequencies). These mock studies were carried out for all pair-wise combinations of subpopulations, each time labeling one subpopulation as 'cases' and the other as 'controls'. The percentage of markers with *P*-values <0.05 was reported. As the variance of the *P*-value is inversely related to sample size, false positive rates were not estimated for subpopulations with sample sizes <20 (PT, HU and RO; see Table 1 for subpopulation abbreviations).

Results

Best genetic match and ancestry sensitive marker sets

Two subsets of markers (BGM and ASM) were ascertained from the complete marker set using either IBS-based forward selection or SVD with redundant marker reduction, respectively. As the decrease in Δ as a function of marker set size levelled off very rapidly (see Figure 1), BGM marker selection was terminated at 100 SNPs (Supplementary Table 1). For the sake of comparability, the ASM set was chosen so as to contain the same number of markers as the BGM set (Supplementary Table 2). Interestingly, the top 5000 markers of the provisional ASM set included various SNPs annotated to genes known to stratify the European

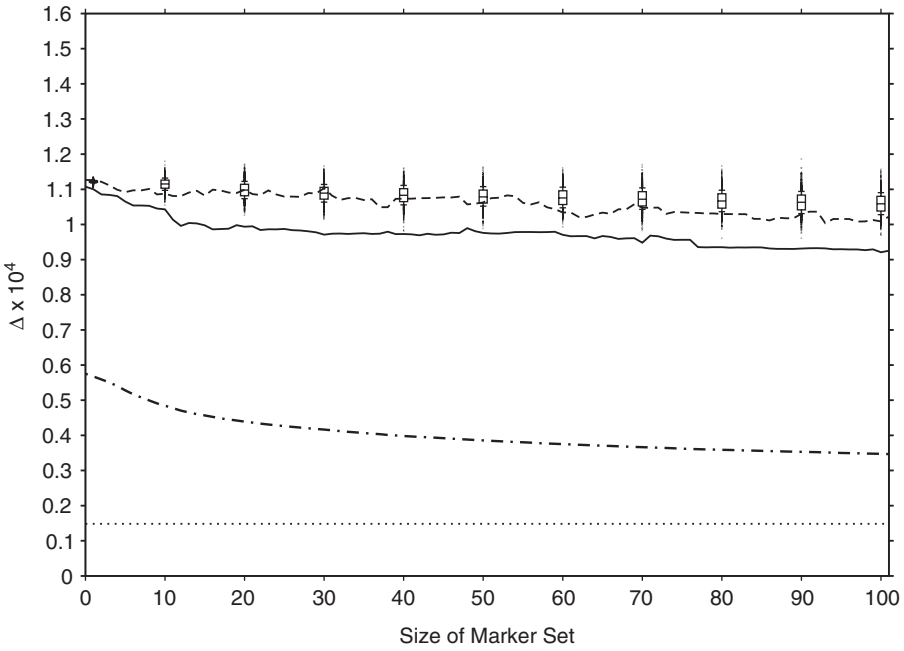


Figure 1 IBS-based forward selection of best genetic match (BGM) marker sets. The upper baseline for Δ is illustrated by box-whisker plots, each generated from 1000 random selections of a marker set of given size. The lower baseline for Δ (dotted line) is provided by a marker set for which any additional markers could be expected to be in strong linkage disequilibrium ($r^2 > 0.98$) with at least one marker already included in that set (for details, see text). Selection of the BGM marker sets is depicted by a solid line; the performance of ASM sets of various sizes is illustrated by a dashed line. All Δ values were calculated from the validation set of individuals. The training set Δ values obtained for the BGM marker sets are included for reference (dash-dotted line).

gene pool as a result of recent positive selection acting differently in different geographic regions, including *HERC2*[21] (ranked 7), *OCA2*[22] (ranked 33), *LCT*[23] (ranked 262) and *TYRP1*[24] (ranked 1138).

A graphical representation of the forward selection process leading to the BGM set is provided in Figure 1. In the validation set, the Δ criterion decreased by ~10% until it levelled off at ~20 markers, and decreased only marginally thereafter. Although forward selection on the training set showed a promising reduction in Δ value, the validation Δ for the 100 top markers comprising the BMG set was still at 9.3×10^{-5}, which is 14.3% lower than the upper (random) baseline but exceeds the lower baseline of 1.5×10^{-5} by a factor of six. This implies that the genome-wide similarity of two European individuals is hard to predict with sufficient accuracy on the basis of a small, specifically selected marker set, and that the little benefit that can be gained in this respect already arises from 100 markers or even fewer. By comparison, the capacity of the ASM set for BOM prediction was found to be indistinguishable from the upper (random) baseline, ie, it performed no better than randomly drawn marker sets.

Distribution of best overall matches (BOMs)

A significant amount of genetic similarity between the European subpopulations is revealed by an assessment of the subpopulation of origin of BOMs (Table 2). In a considerable proportion of cases (1868/2457 or 76.0%), the BOM of a given individual belonged to a different subpopulation than the individual itself. That this was particularly so when individuals or BOMs came from subpopulations with large sample sizes (DE1, DE2 and NL) was presumably due to the wider range of genetic diversity captured by these samples, but may also reflect their concurrent geographic location in central Europe. On the other hand, for some relatively isolated subpopulations (FI and IT2) the source of the BOM was mostly the subpopulation itself, reflecting their separation also seen in genetic barrier analysis and, in the case of the Finns, principle component analysis.[10] Closer inspection at the individual level revealed that some individuals were disproportionately more often selected as BOMs than others (Figure 2). Thus, of the 2457 individuals examined, 1860 (75.7%) were never deemed a BOM at all. This is significantly higher than the expected number (1553.3, 63.2%) if BOMs were drawn at random ($\chi^2 = 165.1$, 1 df, $P < 0.001$). At the same time, 120 individuals were chosen as BOMs at least five times, which is a significant excess over expectation (9.0, 0.36%, $\chi^2 = 1401.9$, 1 df, $P < 0.001$). The subpopulation of origin of the 10 most frequently ascertained BOMs was generally among those central Europeans who also had the largest sample size (DE1 five,

Table 2 Count matrix of BOM (best overall match) affiliation

	NO	*SE*	*FI*	*IE*	*UK*	*DK*	*NL*	*DE1*	*DE2*	*AT*	*CH*	*FR*	*PT*	*ES1*	*ES2*	*IT1*	*IT2*	*YU*	*EL*	*HU*	*RO*	*PO*	*CZ*	*Total*
NO	**8**	2	0	0	0	0	3	**25**	7	0	0	0	0	0	0	0	3	0	0	0	0	4	0	52
SE	6	1	0	0	2	2	4	22	5	0	2	0	0	0	0	0	2	0	0	0	0	0	0	46
FI	1	0	**39**	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	47
IE	1	0	0	4	**12**	3	3	3	3	0	0	2	0	0	0	0	4	0	0	0	0	0	0	35
UK	2	0	0	8	27	**23**	40	62	15	0	1	3	0	0	0	0	10	1	0	0	0	2	0	194
DK	1	0	0	0	0	**10**	13	23	7	0	0	0	0	0	0	0	2	0	0	0	0	3	0	59
NL	4	1	0	1	14	**45**	**94**	79	16	0	3	1	0	0	0	0	8	0	1	0	0	11	2	280
DE1	19	1	0	4	21	**74**	60	**230**	54	0	4	1	0	0	0	0	10	0	0	0	0	16	0	494
DE2	9	0	0	5	24	**68**	83	179	90	3	10	1	0	0	0	0	8	0	0	0	0	9	0	489
AT	0	0	0	0	3	**8**	7	15	13	0	0	0	0	0	0	0	3	0	0	0	0	1	0	50
CH	2	1	0	1	18	**15**	26	36	20	2	8	1	0	0	0	0	1	0	0	0	0	2	0	133
FR	0	0	0	1	3	4	11	16	8	0	1	2	0	0	0	0	1	1	0	0	0	2	0	50
PT	1	0	0	0	0	2	3	5	2	0	1	1	0	1	0	0	0	0	0	0	0	0	0	16
ES1	1	1	0	1	9	6	16	21	8	1	5	8	0	0	0	1	1	0	0	0	0	2	0	81
ES2	1	1	0	0	4	3	8	6	6	0	3	3	0	0	**9**	2	0	0	1	0	0	0	0	47
IT1	1	2	0	1	6	3	19	28	15	2	10	5	0	0	2	5	0	2	2	0	0	1	2	106
IT2	2	0	0	0	0	1	3	9	1	0	0	0	0	0	0	0	**33**	0	0	0	0	0	0	49
YU	3	1	0	0	4	0	8	20	7	0	1	0	0	0	0	0	1	**10**	0	0	0	0	0	55
EL	2	2	0	0	2	1	7	17	8	3	1	0	0	0	0	1	0	0	**7**	0	0	0	0	51
HU	1	0	0	0	1	2	4	8	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	17
RO	0	0	0	0	1	2	1	2	2	0	2	1	0	0	0	1	0	0	0	0	0	0	0	12
PO	3	0	0	0	1	**13**	2	16	5	0	0	0	0	0	0	0	0	0	0	0	0	**8**	1	49
CZ	0	0	0	2	2	3	11	15	3	1	1	1	0	1	0	0	1	0	0	0	0	0	4	45
Total	68	13	39	28	154	288	426	841	295	12	53	31	0	2	11	10	88	14	11	0	0	64	9	

Row, subpopulation of origin of reference individual; Column, subpopulation of origin of BOM of reference individual.
Underlined values are significantly higher than random expectation (*P*-value ≤ 0.05), bold values are statistically significant after Bonferroni correction (FWER ≤ 0.05).

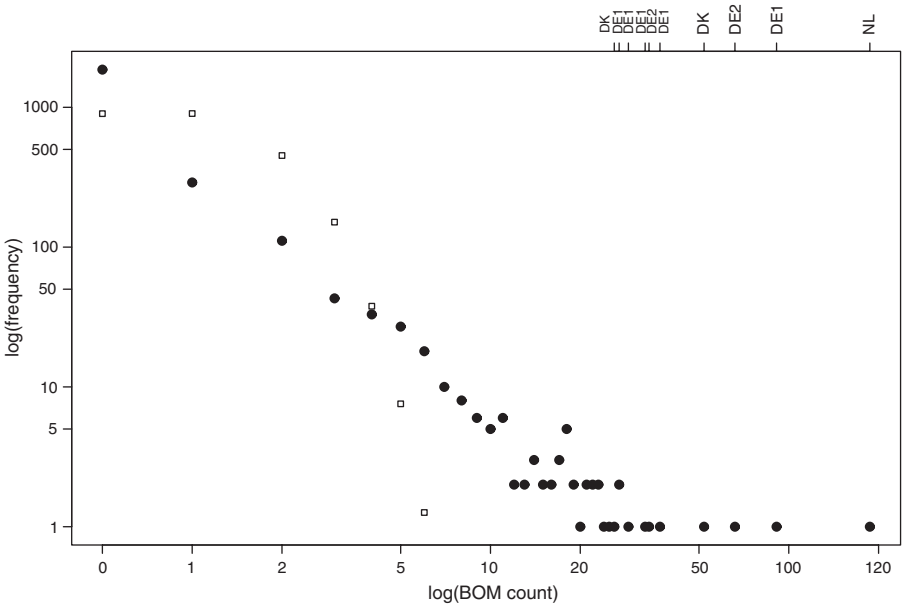


Figure 2 Distribution of the number of times an individual was deemed a BOM. The observed distribution is marked by circles. Also included is a Poisson distribution with the same mean as the sample mean (marked by squares), which approximately corresponds to the theoretical expectation if best overall matching (BOM) were selected at random. The codes of the subpopulation of origin of the 10 most frequently selected BOMs are given at the upper right edge of the plot.

DE2 two and NL one), with the notable exception of DK (59 individuals, yet holding two of the top 10 positions; Figure 2). Interestingly, barring of the 10 most frequently chosen BOMs left the number of times the BOM was found outside the subpopulation of origin of the individual of interest virtually unchanged (1862/2457 or 75.8%,

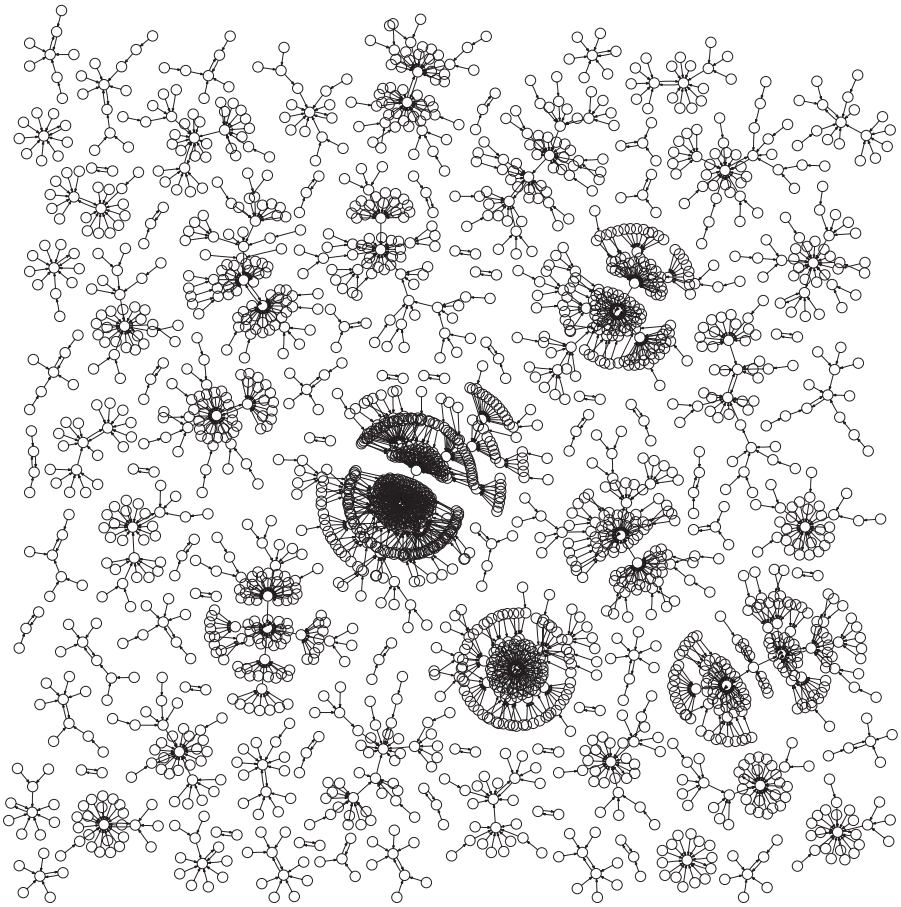


Figure 3 Directed graph illustrating the best overall matching (BOM) relationships between individuals. Circles represent individuals (2457 total) and arrows point towards the respective BOM. The most frequently selected BOM (centre of the plot) was selected for 187 individuals.

Supplementary Table 4). A graphical representation of the BOM relationships between individuals is provided in a directed graph illustrating the complexity of networks of matches (Figure 3).

False positive rates

Although it is admittedly unlikely that a researcher would actually carry out a population-based gene–disease association study in which cases and controls were sampled from different countries, without adjusting for population origin in one way or another, measurement of the false positive rates expected from such undertaking is of general interest as a gauge of the magnitude of stratification pertaining in the European population. Mock false positive rates for pairs of subpopulations (Supplementary Table 3) ranged from 0.039 (CZ and PO) to 0.208 (DE1 and IT1), with a median of 0.070. Subpopulations sampled from the same political country often had false positive rates indicative of little or no population stratification, although this was not always the case (DE1–DE2: 0.089). Many neighboring countries also had false positive rates close to those expected under the null hypothesis, indicating the absence of major population differences as well (eg UK-IE: 0.042, NL-DK: 0.051, EL-YU: 0.047, CH-AT: 0.039, FR-DE2: 0.051).

Discussion

This is the first study to evaluate the genetic (ie, IBS-) matched pair study design with genome-wide SNP data of a large number of European individuals from across the continent. The high number of best genetic-matching partners found in different subpopulations corroborates earlier reports of a considerable amount of genetic similarity between the European subpopulations,[4,10,14,25–27] particularly those in close geographic proximity. The surprising inter-individual variability observed in terms of the number of times a person was chosen as the best genetic-matching partner of others does not necessarily imply that the relationship between genetic and geographic distance in a

given sample hinges on a small number of people. Thus, when the most frequently chosen matching partners were barred in our analysis, the proportion of best matches found outside the subpopulation of origin of the respective index person remained virtually unchanged.

We observed that the best genetic-matching partner for a genome-wide marker set such as the Affymetrix GeneChip Human Mapping 500K Array cannot be predicted from a small, specifically selected subset of markers alone, but that the information required to make such predictions is distributed evenly across all markers. This leads us to conclude that, at least in Europe, the utility of the genetic-matched pair study design depends critically on the availability of comprehensive genotype information for both cases and controls. In practise, this would mean that shared controls should ideally be genotyped for all relevant genome-wide marker sets, thereby allowing the chromosome-specific choice of best matching partners for given case individuals on the basis of the remainder of the genome.

A distinction must obviously be made between ASM, collections of which have been described in recent papers,[14,25–28] and the BGM marker set that we attempted to generate. As the genetic within-subpopulation variation in Europe is much greater than the between-subpopulation variation, it is not unlikely for any two individuals from different subpopulations to be genetically more similar to each other than any two individuals from the same subpopulation. In this sense, an ASM marker set consists of markers that differentiate subpopulations, whereas a BGM marker set should contain variants that highlight genetic similarity at the individual level. Although the two concepts are complimentary, the marker sets fit to each task need not be the same, and the existence of one set does not necessitate the existence of the other. Obviously, markers that arose on early branches of the corresponding, region-specific coalescence tree of the extant Europeans would provide good ASM, but they cannot at the same time identify nearest neighbors at the tips of the tree. Such identification requires a much higher resolution of the tree topology, and therefore many more markers. Consequently, no adequately sized BGM set could be constructed in our study and the ASM set selected with established methodology was no more capable of identifying the best genetic-matching partner of an individual than a randomly chosen marker set.

Recently, two independent applications of genetic matching have been reported in the context of GWA studies,[4,29] both of which relied on information derived from PCA of genotypes to match individuals. In the first study, using US-American type 1 diabetes patients and German controls, Luca *et al*[4] carried out 'full' matching wherein matches consist of clusters of individuals that contain at least one case and one control. Matching was based upon a distance measure with the top eigenvectors as coordinates, weighted by the eigenvalues to exaggerate differences in dimensions of greater importance. In the second study, Heath *et al*[29] undertook a PCA on a large pan-European group of individuals and proposed a method to predict the population affiliation of a sample of unknown origin from the eigenvector matrix of its genotypes. As both methods are likely to reduce spurious genetic differences between cases and controls in disease association studies, basing their matching criteria on eigenvectors from PCA is strongly reminiscent of selecting ASM. However, as we have shown above, matching with ASM is less efficient than best overall genetic matching particularly in Europe, where the within-subpopulation genetic variation is known to be much greater than the between-subpopulation variation. Indeed, the conclusion by Luca *et al*[4] that some individuals remain 'unmatchable' by their approach is not surprising bearing in mind that ASM can only capture a miniscule proportion of the actual inter-individual genetic differences in a given population.

The false positive rates derived in our study from mock genetic case-control experiments represent an upper limit to the likely consequences of sharing samples in continent-wide scientific collaborations. In this respect, the rate estimates also rationalize collaborative genetic epidemiological and pharmacogenetic research in Europe; from the data we have compiled, it seems as if research projects combining cases from neighboring subpopulations and matching them against common control samples, such as those provided by the WTCCC,[3] GAIN[1] and NGFN,[9] may indeed be valid.

In conclusion, we found that the pattern of pair-wise genetic matching in the European population was more complex than anticipated. Best genetic matches occurred frequently across the continent in our study, and disproportionately often involved a small group of individuals. Ascertainment of a subset of markers that accurately predicts best overall genetic matches turned out to be infeasible.

Acknowledgements

All sample donors are gratefully acknowledged for their participation. We thank the following colleagues for their help and support: J Kooner and J Chambers of the LOLIPOP study and D Waterworth, V Mooser, G Waeber and P Vollenweider of the CoLaus study for providing access to their collections through the GlaxoSmithKline-sponsored Population Reference Sample (POPRES) project; K King for preparing the POPRES data; M Simoons, E Sijbrands, A van Belkum, J Laven, J Lindemans, E Knipers and B Stricker for their financial contribution to the Rotterdam study; P Arp, M Jhamai, W van IJken and R van Schaik for generating the Rotterdam study dataset; T Meitinger, P Lichtner, G Eckstein and all genotyping staff at the Helmholtz Zentrum München for generating the KORA study dataset; H von Eller-Eberstein for providing access to the PopGen data; R Borup, C Schjerling, H Ullum, E Haastrup and numerous colleagues at the Copenhagen University Hospital Blood Bank for making the Danish data available; and S Brauer for DNA sample management. We also wish to thank Affymetrix for making

the GeneChip Human Mapping 500K Array genotypes of the CEPH-CEU trios publicly available, and the Centre d'Etude du Polymorphisme Humain (CEPH) for the original sample collection. This work was supported by the Netherlands Forensic Institute (M Ka), Affymetrix (M Ka and M Kr), the German National Genome Research Network and the German Federal Ministry of Education and Research (H-EW, SS, M Kr and PN); the Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg and the Munich Center of Health Sciences as part of LMUinnovativ (H-EW), the Netherlands Organization for Scientific Research (AGU: NWO 175.010.2005.011), the European Commission (AGU: GEFOS; 201865, AS: LD Europe; QLG2-CT-2001-00916); the Czech Ministry of Health (MM: VZFNM 00064203 and IGA NS/9488-3), Helse-Vest, Regional Health Authority Norway (LAB), the Swedish National Board of Forensic Medicine (GH: RMVFoU 99:22, 02:20) and the Academy of Finland (AS: 80578, OMLL, JP: 109265 and 111713). None of the funding organization had any influence on the design, conduct or conclusions of the study.

References

1 GAIN Collaborative Research GroupManolio TA, Rodriguez LL, Brooks L *et al*: New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 32007; **9**: 1045–1051.
2 Hirschhorn JN: Genetic approaches to studying common diseases and complex traits. *Pediatr Res* 2005; **57**: 74R–77R.
3 The Wellcome Trust Case Control Consortium: Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661–678.
4 Luca D, Ringquist S, Klei L *et al*: On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. *Am J Hum Genet* 2008; **82**: 453–463.
5 Davey Smith G, Ebrahim S: What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* 2005; **330**: 1076–1079.
6 Devlin B, Roeder K: Genomic control for association studies. *Biometrics* 1999; **55**: 997–1004.
7 Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000; **155**: 945–959.
8 Wang WY, Barratt BJ, Clayton DG, Todd JA: Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 2005; **6**: 109–118.
9 Wichmann HE, Gieger C, Illig T, MONICA/KORA_Study_Group: KORA-gen – resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* 2005; **67**: 26–30.
10 Lao O, Lu TT, Nothnagel M *et al*: Correlation between genetic and geographic structure in Europe. *Curr Biol* 2008; **18**: 1241–1248.
11 R Development Core Team: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, 2008.
12 Morton NE, Zhang W, Taillon-Miller P, Ennis S, Kwok PY, Collins A: The optimal measure of allelic association. *Proc Natl Acad Sci USA* 2001; **98**: 5217–5221.
13 Wollstein A, Herrmann A, Wittig M *et al*: Efficacy assessment of SNP sets for genome-wide disease association studies. *Nucleic Acids Res* 2007; **35**: e113.
14 Paschou P, Drineas P, Lewis J *et al*: Tracing sub-structure in the European American population with PCA-informative markers. *PLoS Genet* 2008; **4**: e1000114+.
15 Paschou P, Ziv E, Burchard EG *et al*: PCA-correlated SNPs for structure identification in worldwide human populations. *PLoS Genet* 2007; **3**: 1672–1686.
16 Patterson N, Price AL, Reich D: Population structure and eigenanalysis. *PLoS Genet* 2006; **2**: e190.
17 Berry MW: Large scale singular value computations. *Int J Supercomput Appl* 1992; **6**: 13–49.
18 Eaton JW: *GNU Octave Manual*. Network Theory Unlimited: Bristol, 2002.
19 Gansner ER, North SC: An open graph visualization system and its applications to software engineering. *Softw Pract Exp* 2000; **30**: 1203–1233.
20 Purcell S, Neale B, Todd-Brown K *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
21 Kayser M, Liu F, Janssens AC *et al*: Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411–423.
22 Duffy DL, Montgomery GW, Chen W *et al*: A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. *Am J Hum Genet* 2007; **80**: 241–252.
23 Bersaglieri T, Sabeti PC, Patterson N *et al*: Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet* 2004; **74**: 1111–1120.
24 Voight BF, Kudaravalli S, Wen X, Pritchard JK: A map of recent positive selection in the human genome. *PLoS Biol* 2006; **4**: e72.
25 Bauchet M, McEvoy B, Pearson LN *et al*: Measuring European population stratification with microarray genotype data. *Am J Hum Genet* 2007; **80**: 948–956.
26 Price AL, Butler J, Patterson N *et al*: Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet* 2008; **4**: e236.
27 Seldin MF, Shigeta R, Villoslada P *et al*: European population substructure: clustering of northern and southern populations. *PLoS Genet* 2006; **2**: e143.
28 Tian C, Hinds DA, Shigeta R *et al*: A genomewide single-nucleotide-polymorphism panel for Mexican American admixture mapping. *Am J Hum Genet* 2007; **80**: 1014–1023.
29 Heath SC, Gut IG, Brennan P *et al*: Investigation of the fine structure of European populations with applications to disease association studies. *Eur J Hum Genet* 2008; **16**: 1413–1429.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)

3. DISKUSE K PŘILOŽENÝM PUBLIKACÍM

Ve své práci jsem se zaměřila na studium cystické fibrózy z hlediska:

1/ zavedení celoplošného novorozeneckého screeningu a zhodnocení jeho dopadu

2/ stanovení evropských doporučení pro sledování pacientů s atypickými formami onemocnění

3/ studia dalších genů podílejících se na atypických formách onemocnění

4/ populační genomiky české populace z hlediska novorozeneckého screeningu

1/ V České republice došlo v letech 1999-2004 k výraznému zhoršení klinické diagnostiky (100), které se vymykalo úrovni zdravotní péče ve vyspělých zemích. Medián věku při diagnóze se zdvojnásobil, podíl pacientů zachycených do 1 roku života poklesl. Pacienti diagnostikovaní mezi 1.-10. rokem života často přicházeli ve velmi špatném nutričním stavu a/anebo s ireverzibilními změnami plic.

Přestože stále postrádáme kauzální léčbu, délka a kvalita života pacientů s CF se v posledních letech významně zlepšila a je úměrná celkové úrovni zdravotní péče v daném státě (101). Rozhodujícím faktorem prognózy a adekvátní péče je *včasnost diagnózy.* Tyto skutečnosti vedly k zahájení pilotní studie novorozeneckého screeningu (NSCF) s hlavním cílem - *zlepšit časnou diagnostiku a odstranit nerovnoměrnost péče*. Na rozdíl od klasicky prováděných NS, jako např. fenylketonurie nebo kongenitální hypothyreóza, je NSCF specifický tím, že vyžaduje vícestupňové testování a jako druhý krok často DNA analýzu mutací genu *CFTR. Ta vyžaduje dobrou znalost spektra populačně specifických mutací*.

V pilotní studii prováděné na území českých krajů v letech 2005-2006 jsme zvolili schéma IRT/DNA/IRT a *snažili se prokázat účinnost a proveditelnost tohoto protokolu v podmínkách České republiky.* Od rodičů byl získáván informovaný souhlas rozšířením informování o běžném novorozeneckém screeningu, způsobem obvyklým na každém jednotlivém novorozeneckém pracovišti, o informaci o této studii, s možností odmítnutí účasti. Primárním vzorkem byla zaschlá kapka krve na filtračním papírku, ze které se nejdříve (v prvním stupni) měřila koncentrace IRT pomocí fluoroimunoeseje. Klíčové bylo nastavení selekční „cut off“ hodnoty IRT, která se v různých státech či studiích poměrně

liší. V současnosti se nejvíce používá hodnota mezi 98.-99,5. percentilem, která by dle zahraničních zkušeností měla být dostatečně senzitivní a současně nezvyšovat ekonomické náklady dané následnou DNA analýzou. Schéma, využívající ve druhém stupni u selektovaných vzorků se zvýšenou koncentrací IRT, molekulárně genetické vyšetření vybraných mutací genu *CFTR,* jsme zvolili zejména proto, že spektrum a frekvence mutací jsou u české populace detailně zmapovány . Nejdříve bylo nutné zajistit a optimalizovat izolaci dostatečného množství DNA ze zaschlé kapky krve. Další zásadní otázkou byl výběr spektra mutací.

Dosud neexistují komerčně vyráběné kity určené pro novorozenecký screening, tzn. s častými mutacemi spojenými s klasickou formou CF a neexistuje žádný doporučený konsenzus. Zvolili jsme ekonomicky výhodnější a logický postup, kdy vzorky s IRT koncentrací 75-150 ng/ml byly analyzovány na 5 nejčastějších mutací se záchytností 83,8% patologických alel v ČR; a vzorky s IRT >150 ng/ml nebo již jednou nalezenou mutací byly vyšetřeny komerčním diagnostickým kitem na 38 mutací pokrývající 90,8% alel v ČR. Celkem bylo v uvedeném období zachyceno 12 pacientů s CF. Medián věku při stanovení diagnózy pomocí NS činil 37 dnů (rozsah 26-54), což jednoznačně svědčí pro včasnost záchytu pacientů. U jednoho pacienta s normálním potním testem a genotypem F508del/R117H –IVS-7T/7T), který je zpravidla spojen s velmi mírným a v dětství i asymptomatickým průběhem, je však sporné, zda je v tomto případě NSCF nějakým přínosem (102).

Nepřímým výsledkem a důkazem horší klinické diagnostiky je, že tři (z celkem pěti) starší sourozenci zachycených novorozenců měli příznaky CF, *ale na diagnózu se u nich pomyslelo až na základě výsledku z novorozeneckého screeningu* a jejich následného vyšetření v CF Centru. Všichni detekovaní pacienti byli zařazeni k léčbě a komplexní dispenzarizaci do CF center a jejich rodiny podstoupily genetické poradenství.

Námi zjištěná incidence pacientů v souboru 1: 6,369, pokud počítáme i pacienta s velmi mírným genotypem, *je v rozporu s předchozími epidemiologickými a molekulárně genetickými studiemi ze šedesátých let minulého století,* ve kterých byla stanovena incidence na 1: 2736. Domníváme se, že vysvětlením by mohlo být:

1. relativně malý počet vyšetřených jedinců narozených pouze na území českých krajů a také diskontinuita testování (z finančních důvodů daných grantovou agenturou IGA MZČR)
2. falešné negativitě výsledku – toto se jeví jako málo pravděpodobné, protože dosud nebyl zaznamenán ze sledované kohorty žádný pacient s klasickou formou CF
3. podíl prenatální diagnostiky a ukončení gravidity při postižení plodu CF

Poslední bod jistě významně přispěl ke snížení incidence a takto korigovaný výsledek činí 1: 3821, což odpovídá nálezům zahraničních studií. Dalším vysvětlením by mohlo být, že v evropských zemích je v současnosti pozorován trend klesající incidence výskytu CF (103) z ne zcela jasných důvodů zahrnujících mimo jiné mimoevropské přistěhovalectví a s tím související zvýšená porodnost oproti „domácím" populacím a/nebo široce prováděnou prenatální diagnostiku na základě ultrazvukového screeningu fetální hyperechogenity v dutině břišní (105). Je tedy možné, že vlivem NS a rozšířeného molekulárně genetického vyšetření u dalších pokrevných příbuzných, jakož i prenatální diagnostiky, nebude za pár desetiletí screening v novorozeneckém období potřeba.

Dále jsme se zabývali otázkou, zda pacienti časně zachycení pomocí NS klinicky profitují z časné diagnózy a zda vyžadují méně intenzivní a/nebo nákladnou terapii. Pro zodpovězení jsme porovnali kohortu pacientů zachycených NS s kontrolní skupinou CF pacientů stejné věkové kategorie klinicky diagnostikovaných (medián stanovení diagnózy 1 rok). Z antropometrických dat u (8/16) dětí v batolecím věku jsme zjistili, že děti ze screeningu jsou v lepším nutričním stavu (BMI, tělesná hmotnost) a průměrné i celkové náklady na léčbu jsou u kontrolní skupiny signifikantně vyšší. V úvahu je třeba brát i to, že se na tíži onemocnění i u klasické formy onemocnění mohou podílet modifikující faktory. Podrobné výsledky byly prezentovány formou posteru na mezinárodní konferenci ISNS (International Society of Newborn Screening) v roce 2009.

Vedlejším nálezem byl záchyt zdravých nosičů CF mutací ve skupině novorozenců se zvýšeným IRT. Z genetického hlediska se domníváme, že i tento nález a informace daná rodině jsou přínosné v souvislosti s relativně vysokou frekvencí nosičů v populaci (1:26 v ČR) a tedy možností informované reprodukce v případech, kdy např. oba rodiče nebo budoucí partner probanda budou nosiči. V souladu s evropskými studiemi jsme také

potvrdili, že pomocí NS je nacházeno více nosičů, než je předpokládaná četnost v dané populaci. To vypovídá o faktu, že nosiči CF mutací s normálním potním testem mohou mít zvýšenou koncentraci IRT aniž by byla prokázána pankreatická dysfunkce (104). Četnost opakování náběru suché kapky pro nejasný výsledek screeningu („recall rate“) 0,05% je v srovnání s jinými screeningovým programy extrémně nízká (např. u kongenitální adrenální hyperplazie bývá „recall rate“ 0,3-0,5%, u kongenitální hypotyreozy a fenylketonurie 0,1-0,2%). *Tato hodnota neznehodnocuje NS CF nepřiměřenou zátěží zdravé části populace a ani nebude při event. celoplošném pravidelném provádění NS CF významněji zvyšovat kumulativní „recall rate“ s ostatním screeningy.* Závěry naší studie jsme publikovali v původních článcích v *Česko-slovenské pediatrii* 2007; 62(4):187 a *Journal of Cystic Fibrosis* 2009; 8:224.

2/ Analýza NSCF v posledních letech v Evropě se zaměřuje nejen na vlastní technologii a analýzu výsledků, ale také na management a vytváření jednotných evropských doporučení. S narůstajícím množstvím zavádění různých protokolů NSCF na širokém území Evropy vyvstává řada nejasností zejména v oblasti hraničních výsledků a nejednoznačné diagnózy. Předmětem diskuse jsou záchyty probandů s jednou prokázanou mutací a hraničním potním testem nebo dvěma mutacemi, z nichž jedna nebo obě jsou spojeny s velmi mírným fenotypem, jako je například R117H. V tomto ohledu spolupracujeme na vytváření konsenzů s pracovní skupinou evropské společnosti pro cystickou fibrózu (Neonatal Screening Group of ECFS - www.ecfs.eu/ecfs_supported_initiatives/ecfs_neonatal_screening_working_group).

Je zřejmé, že vzhledem k odlišné legislativě a populačnímu spektru je obtížné vytvořit univerzální NSCF – jako např. zavedení jednotného „cut off“ pro IRT a/nebo jednotného spektra testovaných mutací v genu *CFTR*. Nicméně technologický pokrok v oblasti DNA testování – v podobě zavádění sekvenování celého genu *CFTR*, vede k obtížné interpretaci klinického dopadu pro pacienta. V souladu s poznatky této skupiny se domníváme, že v protokolech IRT/DNA je rozšířená DNA analýza *CFTR* genu indikována v případě probanda, který má nalezenou jednu jednoznačně patogenní mutaci, event. je bez nalezené mutace ze základního panelu a současně se u něho opakovaně prokáže hraniční potní test, který je proveden v referenční laboratoři. Dále je samozřejmě vhodná při klinickém

podezření na CF. Přehled závěrů je publikován v původních článcích, na kterých jsme se podíleli *Journal of Cystic Fibrosis 2007; 6: 57 a Journal of Cystic Fibrosis 2009; 8:*71 (spoluautorka v rámci mezinárodního konsorcia).

3/ V diagnostické praxi dochází k záchytu nejen CF s klasickými patologickými symptomy způsobené mutacemi v obou alelách, ale i k záchytu atypických forem onemocnění. Tyto klinické jednotky jsou z hlediska klinické prezentace podobné CF (v mezinárodní odborné literatuře nazývané také CF-like phenotype, *CFTR*-related disorders). Tito pacienti nesou mutaci pouze v jedné alele genu *CFTR* nebo dokonce ani po extenzivní molekulárně genetické analýze genu *CFTR* nejsou patogenní mutace nalezeny. Studium těchto pacientů je důležité z hlediska novorozeneckého screeningu, protože jsou zde převážně diagnostikování pacienti v preklinickém stadiu rozvoje onemocnění – a to na základě laboratorního nálezu, kde molekulární diagnostika mutací v genu *CFTR* hraje zásadní roli. Z tohoto důvodu je nutné studovat genetickou heterogenitu CF a jejích příbuzných forem, které mají jinou prognózu quoad vitam a vyžadují rozdílné terapeutické modality.

V motolském Centru CF jsou dlouhodobě sledováni pacienti s klasicky vyjádřenými respiračními příznaky CF a opakovaně zvýšeným nebo hraničním potním testem, u kterých byla nalezena pouze jedna patogenní mutace při úvodním testování s více než 90% populační záchytností mutací. Druhá mutace *CFTR* genu u některých pacientů dokonce ani pomocí sekvence celé kódující sekvence a následné MLPA analýzy intragenových přestaveb nebyla nalezena. Vyvstává tedy otázka *genetické heterogenity* u těchto pacientů.

K možnému objasnění genetické podstaty jejich onemocnění jsme se zúčastnili mezinárodní studie, ve které byly zkoumány mutace a polymorfismy v ENaC proteinu (na amilorid senzitivní epitelový sodíkový kanál). Hypotéza vycházela z poznatku, že sodíková hyperabsorbce zprostředkovaná ENaC kanálem v dýchacích cestách u nemocných je součástí patogeneze CF. Nabízí se tedy otázka, zda mohou být mutace v genech kódujících protein ENaC zodpovědné za rozvoj „CF-like“ onemocnění.

Výsledky této mezinárodní studie sice neprokázaly jasný patogenní potenciál mutací v kanálu ENaC na rozvoj respiračního onemocnění u těchto pacientů. Nicméně byl dokumentován jejich nepřímý efekt daný oligogenním účinkem variant v subjednotce

SCNN1A - proteinu ENaC, které byly nacházeny právě u pacientů s „CF-like“ onemocnění, a to v kombinaci s nosičstvím jedné mutace v genu *CFTR*. Přibližně 1 z 975 jedinců v obecné populaci je nosičem tzv. hyperaktivní alely - p.W493R-SCNN1A (z hlediska sodíkové resorpce) a současně i jedné patogenní mutace v genu *CFTR, což v důsledku transheterozygosity vytváří podklad dosud nevysvětlených „CF-like“ onemocnění*. Vliv této „hyperresporpční alely“ bylo možno prokázat i pomocí měření nosních potenciálů u studovaných pacientů a v heterologních expresních studiích ve vajíčcích *Xenopus laevis*.

Je zřejmé, že k dosažení statisticky signifikantních výsledků je potřeba studovat u takto vzácného fenoménu příslušné pacienty z celé Evropy a autorka této práce se podílela na jejich výběru, genetické analýze a dlouhodobé klinicko-genetické dispensarizaci. Podrobné výsledky této studie jsou publikovány v původním článku *Human Mutation* 2009; 30:1093.

4/ Dosud se předpokládalo, že CF populace jsou relativně homogenní z populačně genetického hlediska a převážně evropského původu. V souvislosti s vyšší incidencí CF zjištěnou pomocí pilotní studie NSCF, výběrem optimálního spektra mutací genu *CFTR* pro IRT/DNA protokoly novorozeneckého screeningu (viz. bod 1/) a souvislosti se současnými demografickými trendy v České republice vyvstala otázka, zda jsou české a moravské populace stále dostatečně homogenní z populačně genetického hlediska?

Z tohoto důvodu jsme se zúčastnili multicentrického evropského projektu, který analyzoval genetickou a geografickou homogenitu evropských populací. Populační analýza, za užití „Affymetrix GeneChip 500K“ genotypických dat (tvořených 309790 SNP polymorfismy) od 2,514 náhodně vybraných jedinců ze 23 evropských populací (z 20 různých zemí) na širokém území Evropy, potvrdila korelaci mezi geografickou a genetickou vzdáleností. Jedinou výjimkou byla finské populace, která se signifikantně odlišovala od ostatních evropských populací.

Výsledky této mezinárodní kolaborativní studie potvrdily původní hypotézu vycházející z poznatků populační historie o migraci a expanzi původních evropských (paleolitických a neolitických) populací. Jedná se rovněž o dosud největší evropskou genetickou studii, která potvrdila v rámci české a moravské populace populační

homogenitu a minimální intra-populační rozdíly. Tento výsledek výrazně zvyšuje šanci, že výběr mutací v genu *CFTR* s vysokou populační záchytností je optimalizován pro Českou republiku z hlediska IRT/DNA panelu novorozeneckého screenigu, a že nedochází k potenciální falešné negativitě z důvodu „nedostatečně širokého“ výběru mutací v genu *CFTR*. Autorka této práce přispěla výběrem anonymizovaných náhodných jedinců a jejich přesnou populačně genetickou charakteristikou. Výsledky této studie byly publikovány v původním článku v *Current Biology 2008; 18:1241*, společně s následnou komplexní statistickou analýzou v *Eur J Hum Genet 2009; 17:967*.

4. ZÁVĚR

Diagnostika a léčba CF od roku 1938, kdy byla CF vyčleněna jako samostatná klinická jednotka, velmi pokročila. Průměrné přežití pacientů s CF se každým rokem prodlužuje, na což má nesporný vliv také včasné stanovení diagnózy. Počet pacientů, kteří zemřou dříve, než se stanoví diagnóza CF se celosvětově odhaduje na cca. 5%. NSCF otevírá možnost zahájení včasné pankreatické a nutriční suplementace či presymptomatickou terapii dýchacích cest, *a to před rozvojem ireverzibilních změn*. Prokázali jsme, že pacienti zachycení v rámci NS ve věku 1-2 roky profitují, co se týká jejich nutričního stavu a vyžadují méně agresivní a nákladnou léčbu. Vzhledem k potřebné věkové kompliance ve vyšetřování plicních funkcí, jsme výsledky těchto parametrů zatím nehodnotili z důvodu krátkého období sledování.

Dosud neexistuje ideální test pro NSCF, který by byl hromadně použitelný a současně by nikdy nezachytil případy s nejasnou diagnózou CF. Přestože je NSCF problematický z hlediska populační heterogenity danou nastavením specifického „cut off“ koncentrace IRT a výběrem spektra mutací, dosud nebyl v naší studii prokázán falešně negativní případ při udržení racionálních ekonomických nákladů. Výsledky naší pilotní studie NSCF dokazují, že takto nastavený novorozenecký screening je proveditelný a efektivní jak z hlediska potřebnosti, tak včasnosti a uniformity diagnostiky. Výstupy naší pilotní byly podkladem pro odborné společnosti (Společnost lékařské genetiky, Pediatrická společnost ČLS JEP) a Ministerstvo Zdravotnictví (MZČR), které rozhodly o zavedení celoplošného pravidelného novorozeneckého screeningu CF (Věstník MZ ČR ze 12. srpna 2009, částka 6; http://ldmp.upol.cz/ns/vestnik.pdf), který jsme začali provádět od 1. října 2009. Od ledna 2010 se pro moravské kraje připojila i spolupracující pracoviště z FN Brno z Oddělení lékařské genetiky a Oddělení klinické biochemie a hematologie.

Naší prací jsme rovněž přispěli k zavedení účinného managementu a diagnostiky CF. Věříme, že pomocí novorozeneckého screeningu jsme dětem s CF výrazně pozitivně ovlivnili jejich start do života a ušetřili jejich rodiče, jejich slovy „diagnostického martyria“, v důsledku pozdě či chybně stanovené klinické diagnózy onemocnění. To otevírá pacientům s CF v České republice naději, že se dožijí zavedení kauzální léčby tohoto onemocnění v optimálním klinickém stavu.

5. PŘEHLED PUBLIKACÍ, ABSTRAKT A PŘEDNÁŠEK

Publikované původní články

1. Lao O, Lu TT, Nothnagel M, Junge O, Freitag-Wolf S, Caliebe A, Balascakova M, Bertranpetit J, Bindoff LA, Comas D, Holmlund G, Kouvatsi A, Macek M, Mollet I,Parson W, Palo J, Ploski R, Sajantila A, Tagliabracci A, Gether U, Werge T,Rivadeneira F, Hofman A, Uitterlinden AG, Gieger C, Wichmann HE, Rüther A,Schreiber S, Becker C, Nürnberg P, Nelson MR, Krawczak M, Kayser M. Correlation between genetic and geographic structure in Europe. Curr Biol. 2008 Aug 26;18(16):1241-8. (IF 10,777)
2. Mayell SJ, Munck A, Craig JV, Sermet I, Brownlee KG, Schwarz MJ, Castellani C, Southern KW; European Cystic Fibrosis Society Neonatal Screening Working Group. A European consensus for the evaluation and management of infants with an equivocal diagnosis following newborn screening for cystic fibrosis. J Cyst Fibros. 2009 Jan; 8(1):71-8. (IF 1,550)
3. Balascaková M, Holubová A, Skalická V, Zemková D, Kracmar P, Gonsorcíkova L,Camajová J, Piskácková T, Lebl J, Drevínek P, Gregor V, Vávrová V, Votava F, Macek M Jr. Pilot newborn screening project for cystic fibrosis in the CzechRepublic: defining role of the delay in its symptomatic diagnosis and influence of ultrasound-based prenatal diagnosis on the incidence of the disease. J CystFibros. 2009 May; 8(3):224-7. (IF 1,550)
4. Lu TT, Lao O, Nothnagel M, Junge O, Freitag-Wolf S, Caliebe A, Balascakova M, Bertranpetit J, Bindoff LA, Comas D, Holmlund G, Kouvatsi A, Macek M, Mollet I,Nielsen F, Parson W, Palo J, Ploski R, Sajantila, A, Tagliabracci A, Gether U, Werge T, Rivadeneira F, Hofman A, Uitterlinden AG, Gieger CWichmann HE, Ruether A, Schreiber S, Becker C, Nürnberg P, Nelson MR, Kayser M, Krawczak M. An evaluation of the genetic-matched pair study design using genome-wide SNP data from the European population. Eur J Hum Genet. 2009 Jul;17(7):967-75. (IF 3,925)
5. Azad AK, Rauh R, Vermeulen F, Jaspers M, Korbmacher J, Boissier B, Bassinet L, Fichou Y, des Georges M, Stanke F, De Boeck K, Dupont L, Balascaková M, Hjelte L, Lebecque P, Radojkovic D, Castellani C, Schwartz M, Stuhrmann M, Schwarz M, Skalicka V, de Monestrol I, Girodon E, Férec C, Claustres M, Tümmler B, Cassiman JJ, Korbmacher C, Cuppens H. Mutations in the amiloride-sensitive epithelial sodium channel in patients with cystic fibrosis-like disease. Hum Mutat. 2009 Jul; 30(7):1093-103. (IF 7,033)

Původní články bez impact faktoru

1. Pohunek P, Macek M, Balaščaková M, Pelikán L, Böhmová K, Skalická V, Brázová J, Vávrová V, Fischerová T, Zemková D. [Notes to the article Nekl R., Mockóva, A., Bierhanzlová, J., Kutilová J. "Meconium Ileus as a first symptom of cystic fibrosis in a newborn"]. Cas Lek Cesk. 2007;146(2):114. Czech. PubMed PMID: 17373104.

2. Holubová A*, Balaščaková M*, Skalická V, et al. Novorozenecký screening cystické fibrózy v České republice: závěry pilotní studie. Čes.-slov. Pediat. 2007;62(4): 187–195.

3. Balaščaková M*, Piskáčková T*, Holubová A, Raušová E, Kazárová V, Krebsová A, Koudová M, Štambergová A, Čamajová J, Norambuena P, Křenková P, Votava F, Vávrová V, Macek M st., Macek M ml. Současné metodické postupy a přehled preimplantační, prenatální a postnatální DNA diagnostiky cystické fibrózy v České republice. Čes.-slov. Pediat, 2008, roč. 63, č.2, s.62-75

Seznam publikovaných abstrakt:

1. Skalicka V, Fischerova T, Bohmova K, Balascakova M, Zemkova D, Votava F, Macek MJr, Vavrova V. Newborn screening (NBS) facilitates diagnosis of previously unrecognised CF in older siblings. 30th European Cystic Fibrosis Conference, Belek, Turkey, 13-16 June 2007. J Cyst Fibros 2007; 6, Supp. 1: S87. Poster

2. Balascakova M, Holubova A, Skalicka V, Kracmar P, Vavrova V, Zemkova D, Camajova J, Piskackova T, Macek MJr, Votava F. Cystic fibrosis newborn screening in the Czech Republic: overview of a pilot study. 30th European Cystic Fibrosis Conference, Belek, 2007. J Cyst Fibros 2007; 6, Supp. 1: S89. Poster

3. Zemkova D, Skalicka V, Balascakova M, Vavrova V, Kracmar P, Macek MJr, Votava F. Newborn screening in the Czech republic: an effective tool for early detection of CF. 30th European Cystic Fibrosis Conference, Belek, Turkey, 13-16 June 2007. J Cyst Fibros 2007; 6, Supp. 1: S89, Poster

4. Balascakova M, Holubova A, Skalicka V, Vavrova V, Zemkova D, Kracmar P, Piskackova T, Votava F, Macek MJr. Newborn screening for cystic fibrosis in the Czech Republic: observation of lower incidence of the disease. European Human Genetics Conference, June 16-19, 2007, Nice, France. European Journal of Human genetics. 2007, 15 (Suppl 1): S133.

5. Azad A, Rauh R, Vermeulen F, Jaspers M, Korbmacher J, Boissier B, Fichou Y, des Georges M, Stanka F, Nuytte H, De Boeck K, Dupont LJ, Balascakova M, Hjelte L, Radojkovic D, Hagens S, Fila L, de Monestrol I, Girodon E, Claude F, Claustres M, Tummler B, Cassiman J, Korbmacher C, Cuppens H. Mutations in the amiloride-sensitive epithelial sodium channel in patients with cystic fibrosis-like disease. Workshop 10. Diagnosis of CF & Identification of CF-related Diseases. The 22nd Annual North American Cystic Fibrosis Conference, Orlando, Florida, October 23-25, 2008. Pediatr Pulmonol 2008; 43(S31): 273.

6. Azad A, Jaspers M, Lelievre C, Radojkovic D, Yarden J, De Boeck K, Dupont LJ, Lebecque P, Pressler T, Schwartz M, Balascakova M, Fila L, Castellani C, Stuhrmann M, Cassiman J, Cuppens H. Analysis of the genes coding for the amiloride sensitive epithelial sodium channel as modifier genes for cystic fibrosis. The 22nd Annual North American Cystic Fibrosis Conference, Orlando, Florida, October 23-25, 2008. Pediatr Pulmonol 2008; 43(S31): 273.

7. Skalická V, Balaščaková M, Votava F, Zemková D, Vávrová V, Macek M Jr. Novorozenecký screening cystické fibrózy: české zkušenosti s novou možností

diagnostiky. Čes.-Slov. Pediatr 2008;63(7-8):411. Abstrakta z VIII. českého pediatrického kongresu s mezinárodní účastí, Jihlava. 2008

8. Skalicka V, Balascakova M, Zemkova D, Piskackova T, Holubova A, Kracmar P, Vavrova V, Macek Jr. M, Votava F , The decision making process leading towards implementation of a nationwide newborn screening for cystic fibrosis in the Czech Republic, sborník abstrakt z Mezinárodní konference novorozeneckého screeningu, ISNS, Praha 2009

9. M. Balascakova[1], V. Skalicka[1], P. Hodkova[2], T. Piskackova[1], L. Smolikova[1], P. Drevinek[1], K. Austova[1], A. Pickova[1], F. Votava[3], M. Macek Jr. Experience with harmonizing the Neonatal CF Screening (NSCF) management via model IRT/PAP/DNA in the Czech Republic, 33rd European Cystic Fibrosis Conference, Valencia, Spain, 16-19 June 2010. J Cyst Fibros 2010; Vol.9, Supp. 1: S24. Poster

10. P.Krenkova, T.Piskackova, M. Balascakova, A. Holubova, J. Camajova, M. Hancarova, L. Dvorakova, V. Skalicka,, A.. Stambergova, V. Vavrova, M.Macek Jr. Spectrum of CFTR Mutations in The Czech Republic, 33rd European Cystic Fibrosis Conference, Valencia, Spain, 16-19 June 2010. J Cyst Fibros 2010; Vol.9, Supp. 1: S7. Poster

***Seznam přednášek*:**

1. Balascakova M, Holubova A, Votava F, Skalicka V, Vavrova V, Zemkova D, Kracmar P, Libik M., Camajová J., Piskackova T, Macek M Jr Novorozenecký screening cystické fibrózy v české rebublice: průběžné výsledky pilotní studie**,** VII. český pediatrický kongres s mezinárodní účastí. Praha, 2006

2. Balaščaková M, Votava F., Holubová A, Skalická V, Vávrová V, Čamajová J, Piskáčková T, Kračmar P, Macek ml. M. Novorozenecký screening cystické fibrózy v ČR: souhrn pilotní studie. 6. Slovenský pediatrický kongres, Košice, 2007

3. Balaščaková M, Holubová A, Skalická V, Vávrová V, Zemková D, Kračmar P, Čamajová J, Libik M, Mullerova M, Piskáčková T, Votava F, Macek ml. M. Novorozenecký screening cystické fibrózy v ČR, vliv prenatální diagnostiky na snížení incidence. Celostátní sjezd Společnosti lékařské genetiky ČLS JWEP a 40. výroční cytogenetická konference. Praha, 2007

4. 4.Balaščaková M, Votava F, Holubová A, Skalická V, Vávrová V, Čamajová J, Piskáčková T, Zemková D, Kračmar P, Macek M. Jr., Novorozenecký screening cystické fibrózy v ČR, souhrn pilotní studie, XXVIII. Imunologické dny. České Budějovice, 2007

5. Balaščaková M, Skalická V, Holubová V, Zemková D, Kračmar P, Čamajová J, Libík M, Müllerová M, Piskáčková T, Macek M Jr, Vávrová V, Votava F. Novorozenecký screening cystické fibrózy v České republice. XXIV. Neonatologické dny. Olomouc, 2008

6. SEZNAM POUŽITÝCH ZKRATEK

ADG..... age at diagnosis, věk při stanovení diagnózy

CF..... cystická fibróza

CFTR ..... cystic fibrosis transmembrane conductance protein, transmembránový regulátor vodivosti

ENaC..... amiloride sensitive epithelial natrium channel, na amiloridu senzitivní epiteliální sodíkový kanál

IRT ..... imunoreaktivní trypsinogen;

PAP ..... pancreatitis associated protein, protein asociovaný s pankreatitidou

DNA..... DNA analýza, molekulárně genetické vyšetření genu *CFTR*

NSCF ..... novorozenecký screening cystické fibrózy

NS ..... novorozenecký screening

CFRD..... CF-related diabetaes, na CF vázaný diabetes mellitus

CBAVD ..... kongenitální bilaterální absence vas deferens

PCL ..... pericilial liquid, periciliální tekutina

HRM ..... hight resolution melting curve analysis, vysokorozlišovací analýza křivky tání

MLPA ..... multiplex ligation-dependent probe amplification

ADG..... věk diagnózy

PPV ..... pozitivní prediktivní hodnota

NPV ..... negativní prediktivní hodnota

MBL..... manose binding lecitin, lektin vážící manózu

NBD..... nukleotide binding domain, doména vázající nukleotid

GIT..... gastriontestinální trakt

7. CITOVANÁ LITERATURA

1. Welsh MJ, Ramsey BW. The Metabolic and Molecular Basis of Inherited Disease. 8th ed. New York: McGraw-Hill, 2001: 5121–5188
2. Vávrová V, a kol. Cystická fibróza. Praha, Grada ISBN 80-247-0531-1, s 520
3. Houštěk J, Vávrová V. K výskytu cystické fibrózy pankreatu v ČSSR. Čs. Pediatrie 1962; 445-451
4. Houšťek J, Vávrová V. Frequency and forms of mucoviscidosis in Czechoslovakia. Čs. Pediatr. 1965; 20“ 412-414
5. Houšťek J, Vávrová V. Nôtre éxperience propos de la mucoviscidose. Rev.Med Liège 1967; 22:421-426
6. Macek M Jr, Macek M, et al. Possible association of the allele status of the CS.7/Hal polymorfism 5 of the CFTR gene with postnatal female survival.Hum Genet 1997a; 52: 557-564
7. Macek M Jr , Vávrová V. Cystická fibróza. Model komplexní patogeneze u monogenních onemocnění. Postgrad Med 2002; 141: 28-34
8. Rosenstein BJ, Cutting GR. The diagnosis of cystic fibrosis: a consensus statement. Cystic Fibrosis Foundation Consensus Panel. *J Pediatr 1998; 132:* 589-595
9. Anguiano A, Oates RD, Amos JA. et al. Congenital bilateral absence of the vas deferens. *JAMA 1992; 267:* 1794-1797
10. Rodman HM, Doershuk CF. The interaction of two diseases: diabetes mellitus and cystic fibrosis. Medicine 1986; 65: 389
11. De Braekeleer M, et al. Is meconium ileus genetically determined or associated with a more severe evolution of cystic fibrosis? *J Med Genet 1998; 35:* 262-263
12. Wine JJ. The genesis of cystic fibrosis lung disease. *J Clin Ivest 1999; 103:* 309-312
13. Welsh MJ, et at: Cystic fibrosis. In: Scriver CR, Beaudet AL, Sly WS and Valle D. eds. The metabolic and molecular basis of inherited disease. VII. Ed. New York: *McGraw-Hill Inc.* Vol III 3799-3876, 1995
14. Kraemer R, Birrer P, Liechti-Gallati S. Genotype-phenotype association in infants with cystic fibrosis at the time of diagnosis. *Pediatr Res 1998; 44:* 920-926
15. Zielenski J. Genotype& phenotype in cystic fibrosis. *Respiration 2000; 67:* 117-133
16. Kulaksiz H. et al. Guanilin regulates chloride secretion in the human gallbladder via bile fluid. Gastroenterology 2004; 126:732-740
17. Tousson A. et.al. Characterisation of the CFTR expression and chloride channel aktivity in human endothelia. Am J Physiol 1998; 275:C1555-C1564
18. Bush et al. Cystic Fibrosis in the 21th Century, S. Kargel AG, 2006
19. Anderson MP et al: Demonstration that CFTR is a chloride channel by alteration of its anion selectivity. *Science 1990; 253:* 202-205

20. Andersen, DH. Cystic fibrosis of the pancreas and its relation to celiac disease. A clinical and pathological study. *Am J Dis Child. 1938; 56:* 344-399

21. Quinton PM, Bijman J: Higher bioelectric potentials due to decreased chloride absorption in the sweat glands of patients with cystic fibrosis. *New Engl J Med 1983; 308:*1185-1189

22. Quinton PM. et al .Chloride impermeability in cystic fibrosis. *Nature 1983; 301:* 421-422

23. Beaudet AL, Spence JE. et al. Experience with new DNA markers for the diagnosis of cystic fibrosis. *N Engl J Med 1988; 318:* 50

24. Lathrop GM, Farrall M, et al. Refined linkage map of chromosome 7 in the region of the cystic fibrosis gene. *Am J Hum Genet 1988; 42:* 38-45

25. Sedláček Z. Strategies in prenatal diagnosis of cystic fibrosis after the introduction of DNA analysis: initial experience.*Čas Lék čes 1990; 129: 683-688*

26. Riordan JR, Rommens JM, Kerem B et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science 1989; 245:* 1066-1073

27. Rommens JM, Iannuzzi MC, Kerem B et al. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science 1989; 245:* 1059-1065

28. Kerem B, Rommens JM, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science 1989; 245:* 1073-1080

29. Frizzell RA. Functions of the cystic fibrosis transmembrane conductance regulator protein. *Am J Respir Crit Care Med 1995; 151:* S54-S58

30. Schwiebert EM, Benos DJ et al. Cystic fibrosis: a multiple exocrinopathy caused by dysfunctions in a multifunctional transport protein. *Am J Med 1998; 104:* 576-590

31. Gray MA, O'Reilly C, Winpenny J, Argent B. Functional interactions of HCO3-with cystic fibrosis transmembrane conductance regulator. JOP. 2001 Jul;2(4Suppl):207-11.

32. Bronsveld I, Mekus F, Bijman J et al. Residual chloride secretion in intestinal tissue of deltaF508 homozygous twins and siblings with cystic fibrosis. The European CF Twin and Sibling Study Consortium. *Gastroenterology 2000;* 119:32-40

33. Zielenski J. Genotype& phenotype in cystic fibrosis. *Respiration 2000; 67:* 117- 133

34. Jovov B, Ismailov II, Benos DJ. Cystic fibrosis transmembrane conductance regulator is required for proteinkinase A activation of an outwardly rectified anion channel purified from bovine tracheal epithelia. *J Biol Chem 1995 270:* 1521-1528

35. Jovov B, Ismailov II, Berdiev BK, et al. Interaction between cystic fibrosis transmembrane conductance regulator and outwardly rectified chloride channeis. *J Biol Chem 1995; 270:* 29194-29200

36. Schwiebert EM, Flotte TR, Cutting GR, et al: Both CFTR and outwardly rectifying chloride channels contribute to whole cell chloride currents. *Am J Physiol 1994*

37. Ismailov II, Awayda MS, Jovov B, et al: Regulation of epithelial sodium channels by the cystic fibrosis transmembrane conductance regulator. *J Biol Chem 1996; 271:* 4725-4732

38. Stutts, MJ, Canessa, CM, Olsen JC et al: CFTR as a cAMP-dependent regulator of sodium channels. *Science 1995; 269:* 847-850

39. Canessa CM, Horisberger JD, Schild L, et al. Expression cloning of the epithelial sodium channel. *Kidney Intl 1995; 48:* 950-955

40. Lingueglia E, Voilley N, Waldman R, et al: Expression cloning of an epithelial chloride-sensitive Na+ channel. *FEBS 1993; 318:95-99*

41. Jasti J. et al. Structure of acid-sensing ion channel 1 at 1,9A resoluting and low pH. Nature 2007; 449:316-323

42. Boucher RC. Cystic fibrosis a disease of vulnerability to airwaysurface dehydratation. *Trends Mol Med 2007; 13:231-240*

43. Knowles RC, Boucher RC. Mucus clearence as a primary innate defense mechanism for mammalian airways. *J Clin Incest 2002; 109:571-577*

44. Mall M. Increased airway epithelial Na+ absorbtion produces cystic fibrosis-like lung disease in mice. *Nat Med 2004; 10:487-493*

45. Casaccia G, Trucchi A, Nahom A. et al. The impact of cystic fibrosis on neonatal intestinal obstruction: the need for prenatal/neonatal screening., *Pediatr Surg Int. 2003 Apr;19(1-2):75-8.*

46. Gibson LE, Cooke RE. et al. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocaropine iontophoresis. *Pediatrics 1959; 23: 545-549*

47. Rosenstein BJ, Cutting GR, et al. The diagnosis of cystic fibrosis: a consensus statement. Cystic Fibrosis Foundation Consensus Panel. *J Pediatr 1998; 132: 589-595*

48. Denning CR, Huang NN, Cuasay LR. et al. Cooperative study comparing three methods of performing sweat tests to diagnose cystic fibrosis. *Pediatrics 1980; 66: 752-757*

49. Bobadilla JL, Macek M Jr, et al. Cystic fibrosis: a worldwide analysis of *CFTR* mutations – correlations with incidence data and application to screening. *Hum. Mutat. 2002;19: 575–606.*

50. Kerem E, Corey M, Kerem B, et al. The relation between genotype and phenotype in cystic fibrosis-analysis of the most common mutation DF508. *N Engl J Med 1990; 323: 1517-1522*

51. The Molecular Genetic Epidemiology of Cystic Fibrosis. Report of a joint meeting of WHO/ECFTN/ICF(M)A/ECFS http://www.who.int/genomics/publications/ en/ and WHO Geneva, Human Genetics Programme WHO/HGN/CF/WG/04.02.

52. Serre JL, Simon-Bouy B,Mornet E, et al. Studies of RFLP closely linked to the cystic fibrosis locus throughout Europe lead to new considerations in population genetics. *Am. J. Hum. Genet. 1990;84: 449–454.*

53. Mickle JE, Cutting GR et al. Genotype phenotype relationships in cystic fibrosis. *Med. Clin. North Am. 2000;84: 597–607.*

54. Ferec C, Casals T, Chuzhanova N, et al. Gross genomic rearrangements involving deletions in the *CFTR* gene: characterization of six new events from a large cohort of hitherto unidentified cystic fibrosis chromosomes and meta-analysis of the underlying mechanisms. *Eur. J. Hum. Genet. 2006;14: 567–576.*

55. Morral N, Bertranpetit J, Estivill X, et al. The origin of the major cystic fibrosis mutation (delta F508) in European populations. *Nat. Genet. 1994;7: 169–175.*

56. Chiba-Falek O, Parad RB, et al.Variable levels of normal RNA in different fetal organs carrying a cystic fibrosis transmembrane conductance regulator splicing mutation. *Am. J. Respir. Crit. Care Med. 1999;159: 1998–2002*

57. Vanscoy LL, Blackman SM, Collaco JM, et al. Heritability of lung disease severity in cystic fibrosis. *Am. J. Respir. Crit. Care Med. 2007;175: 1036–1043.*

58. Cutting GR. Cystic fibrosis In Rimoin DL, Connor JM, Pyeritz RD (eds). Principles and Practice of Medical Genetics. New York: *Churchill Livingstone, 1997; 2685–2717.*

59. Tsui L-C. The spectrum of CF mutations. *Trends Genet 1992; 8:* 392-398

60. Niksic M, Romano M, Buratti E, et al. Functional analysis of cis-acting elements regulating the alternative splicing of human *CFTR* exon 9. *Hum. Mol. Genet.1999;8: 2339–2349.*

61. Cuppens H, Lin W, Jaspers M, et al. Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (TG)n locus explains the partial penetrance of the T5 polymorphism as a disease mutation. *J. Clin. Invest. 1998;101: 487–496.*

62. Disset A, Michot C, et al. A T3 allele in the *CFTR* gene exacerbates exon 9 skipping in vas deferens and epididymal cell lines and is associated with Congenital Bilateral Absence of Vas Deferens (CBAVD). *Hum. Mutat. 2005;25: 72–81.*

63. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics, 1963, 32, p. 338-343.*

64. Guthrie R. Screening for phenylketonuria. *Triangle, 1969, 3, p.104-109.*

65. Blehová B, Pažoutová M. et al. Zhodnocení screeningu fenylketonurie za šest let trvání laboratoře. *Československá Pediatrie, 1976, 7, s. 399-400.*

66. Čechák P, Hejcmanová L. et al. Výsledky screeningu hyperfenylalaninémií v českých zemích v letech 1970-2000. *Československá Pediatrie, 2001, 11, s. 667-670.*

67. Hyánek J., Hoza J., et al. Výsledky hromadného screeningu vrozených metabolických poruch u novorozenců. *Československá Gynekologie, 1972, 4, s. 208-209.*

68. Larsen PR, Merker A.,et al. Immunoassay of human TSH using dried blood samples. Journal *of Clinical Endocrinology & Metabolism, 1976, 5, p. 987-990.*

69. Irie, M., Enomoto, K., Enomoto, H., et al. Microdetermination of TSH in dried blood spot--its use in the mass-screening for congenital or juvenile primary hypothyroidism. *Nippon Naibunpi Gakkai Zasshi, 1976, 3, p. 243-247.*

70. Millington, D.S., Kodo, N., Norwood, D.L., et al. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. *Journal of Inherited Metabolic Disease, 1990, 3, p. 321-324.*

71. Wilcken, B., Wiley, V., Hammond, J., et al. Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *New England Journal of Medicine, 2003, 348, p. 2304-2312.*

72. Loeber, J.G. Neonatal screening, the European perspective. *Česko-slovenská Pediatrie, 2009, 4, s. 159-161.*

73. Loeber,, J.G. Neonatal screening in Europe; the situation in 2004. *Journal of Inherited Metabolic Disease, 2007, 4, p. 430-438.*

74. Bodamer, O.A., Hoffmann, G.F., Lindner, M. Expanded newborn screening in Europe 2007. *Journal of Inherited Metabolic Disease, 2007, 4, p. 439-444.*

75. Dhondt, J.L. Neonatal screening: from the 'Guthrie age' to the 'genetic age'. *Journal of Inherited Metabolic Disease, 2007, 4, p. 418-422.*

76. Waisbren, S.E. Expanded newborn screening: information and resources for the family physician. *American Family Physician, 2008, 7, p. 987-994.*

77. Alexander, D., van Dyck, P.C. A vision of the future of newborn screening. *Pediatrics, 5, p. 350-354.*

78. Green, N.S., Pass, K.A. Neonatal screening by DNA microarray: spots and chips. *Natural Review of Genetics, 2005, 2, p. 147-151.*

79. Tachibana, S., Suzuki, M., Asano, Y. Application of an enzyme chip to the microquantification of l-phenylalanine. *Anales of Biochemistry, 2006, 1, p. 72-78.*

80. Eising, S., Svensson, J., Skogstrand, K., et al. Type 1 diabetes risk analysis on dried blood spot samples from population-based newborns: design and feasibility of an unselected case-control study. *Paediatrics, Perinatology & Epidemiology, 2007, 6, p. 507-517.*

81. Wilson, J.M.C., Jungner, G. Principles and practise of screening for disease. Public Health Papers, 1968, 34, Geneva World Health Organizaion.

82. Lipstein, E.A., Perrin, J.M., Waisbren, S.E., et al. Impact of false-positive newborn metabolic screening results on early health care utilization. *Genetics in Medicine, 2009, 10, p. 716-721.*

83. Sims EJ, Clark A, McCormick J, Mehta G, Connett G, Mehta A. United Kingdom Cystic Fibrosis Database Steering Committee. Cystic fibrosis diagnosed after 2 months of age leads to worse outcomes and requires more therapy. *Pediatrics 2007;119:19–28.*

84. McKay K, Wilcken B. Newborn screening for cystic fibrosis offers an advantage over symptomatic diagnosis for the long term benefit of patients: the motion for. *Paediatr Respir Rev 2008;9:290–4*

85. Crossley JR, Elliot RB, Smith PA. Dried blood spot screening for cystic fibrosis in newborn. *Lancet. 1979;8114: 472-474.*

86. Lecoq I, Brouard J, Laroche D et al. Blood immunoreactive trypsinogen concentrations are genetically determined in healthy and cystic fibrosis newborns. *Acta Paediatr. 1999;88:338-341.*

87. Bourguignon JP, Deby-Dupont G, Reuter A et al. Variations in dried blood spot immunoreactive trypsin in relation to gestational age and during the first week of life. *Eur. J. Pediatr.1986;144:547-549.*

88. Rock MJ, Mischler EH, Farrell PM et al. Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. *Pediatrics. 1990;85:1001-1007.*

89. Laroche D, Travert G. Abnormal frequency of delta F508 mutation in neonatal transitory hypertrypsinaemia. *Lancet. 1991;8732:55.*

90. Pederzini F, Mastella G, Giglio L et al. Dried blood spot immunoreactive trypsin (IRT) values. *Eur. J. Pediatr. 1988;147:663.*

91. Moya EF, Brocklebank JT, Littlewood JM et al. High serum immunoreactive trypsin not caused by cystic fibrosis. *Arch. Dis. Child. Fetal. Neonatal. Ed. 1998;78: F78.*

92. Comeau AM, Parad RB, Dorkin HL et al. Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. *Pediatrics. 2004;113: s.1573-1581.*

93. Bonham JR, Downing M, Dalton A. Screening for cystic fibrosis: the practice and the debate. *Eur. J. Pediatr., 2003, 162, s. S42-S45.*

94. Comeau AM, Parad RB, Dorkin HL et al. Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. *Pediatrics. 2004;113: s.1573-1581.*

95. Corbetta, C., Seia, M., Bassotti, A., Ambrosioni, A., Giunta, A., Padoan, R. Screening for cystic fibrosis in newborn infants: results of a pilot programme based on a two tier protocol (IRT/DNA/IRT) in the Italian population. *J. Med. Screen., 2002, 9, s. 60-63*

96. Sarles J, Bertézene P, L Louarn, et. Al; Combining immunoreactive trypsinogen and pancreatitis-associated protein assays, a method of newborn screening for cystic fibrosis that avoids DNA analysis. *J Pediatr. 2005 Sep;147(3):302-5,*

97. Barthellermy S, Maurin N, Férec C et. al. Evaluation of 47,213 infants in neonatal screening for cystic fibrosis, using pancreatitis-associated protein and immunoreactive trypsinogen assays, *Arch Pediatr. 2001 Mar;8(3):275-81)*

98. Sarles J, Berthezene P, et.al. Combining immunoreactive trypsinogen and pancreatitis-associated protein assays, a method of newborn screening for cystic fibrosis that avoids DNA analysis. *J Pediatr. 2005 Sep;147(3):302-5.*

99. Sarles J, Barthellemy S, Ferec C, et al. Blood concentrations of pancreatitis associated protein in neonates: relevance to neonatal screening for cystic fibrosis. *Arch. Dis. Child Fetal Neonatal Ed., 1999, 80, s. F118-122. J. Med. Screen., 2002, 9, s. 60-63*

100. Vávrová. Problémy v diagnostice cystické fibrózy – potřeba novorozeneckého screeningu.*Čes-slov.Pediat.2006;12:703-709*

101. Davis PB. Cystic fibrosis since 1938. *Am. J. Respir. Crit. Care Med. 2006;173: 475-482*

102. Scotet V, Audrézet MP, Rousseey M et al. Immunoreactive Trypsin/DNA Newborn Screening for Cystic Fibrosis: Schould the R117H Variant Be Included in CFTR Mutation Panels. *Pediatrics. 2006;118:1523-1529.*

103. McCormick J, Mehta G, Olesen HV et al. Comparative demographics of the European cystic fibrosis population: a cross-sectional database analysis. *Lancet. 2010 Mar20;375(9719):1007-13*

104. Massie RJ, Wilcken B. Pancreatic function and extended mutation analysis in DeltaF508 heterozygous infants with an elevated immunoreactive trypsinogen but normal sweat electrolyte levels. *J Pediatr 2000 Aug;137(2):214-20*

105. Muller F, Simon-Bouy B, Girodon E, Monnier N, Malinge MC, Serre JL. French Collaborative Group. Predicting the risk of cystic fibrosis with abnormal ultrasound signs of fetal bowel: results of a Frenchmolecular collaborative study based on 641 prospective cases. *Am J Med Genet 2002;110:109–15.*