

# PCR-multiplex systém pro analýzu Y chromosomálních mikrosatelitních polymorfismů *DYS449*, *DYS456*, *DYS458* a *DYS464*

Ehler E.<sup>1, 2</sup>, Marvan R.<sup>1, 2</sup>, Mazura I.<sup>1, 2</sup>

<sup>1</sup>Katedra antropologie a genetiky člověka,  
Přírodovědecká fakulta Univerzity Karlovy, Praha

<sup>2</sup>Evropské centrum pro medicínskou informatiku,  
statistiku a epidemiologii-EuroMISE centrum

## Souhrn

DNA profilování – včetně určování pohlaví jedince – za pomoci mikrosatelitních markerů je v současnosti běžně užívaná genetická metoda studia člověka. Efektivní technikou produkující genetické údaje je amplifikace několika mikrosatelitních lokusů v jedné PCR reakci. Na tomto místě prezentujeme PCR-multiplex systém pro analýzu čtyř Y-STR polymorfních míst. Jedná se o *DYS449*, *DYS456*, *DYS458* a *DYS464*. Tyto lokusy jsme vybrali s ohledem na jejich dokumentovanou vysokou diverzitu v euroamerické populaci (10) a s ohledem na jejich absenci v komerčně dostupných analytických soupravách. Celý PCR-multiplex systém je designován pro fragmentační analýzu metodou kapilárové elektroforézy na semi-automatickém genetickém analyzátoru za použití značení pouze jednou fluorescentní barvou. Hlavními oblastmi využití těchto polymorfismů by měly být forenzní a lidská populační genetika.

**Klíčová slova:** Y-STR – NRY – chromosom Y – PCR-multiplex

## Summary

**PCR-Multiplex System for Analysis of Y-Chromosomal Microsatellite Polymorphisms *DYS449*, *DYS456*, *DYS458*, and *DYS464***

DNA profiling – inclusive sex determination – with microsatellite markers is currently a commonly used genetic method of studying humans. An efficient technique of producing the genetic data is amplification of multiple microsatellites in a single PCR reaction. Here we introduce a novel PCR-multiplex system for analysis of four polymorphic Y-STRs. Specifically, these are *DYS449*, *DYS456*, *DYS458*, and *DYS464*. These loci were chosen because of their reported high diversity in Euroamerican population (10), as well as their absence in the commercial analytical kits at the time of beginning of this study. Our objective was to design this PCR-multiplex for use of fragmentation analysis by electrophoresing samples on a capillary semi-automated genetic analyzer applying only one fluorescent dye. The PCR system we propose, may be notably used in fields such as forensic and human population genetics.

**Key words:** Y-STR – NRY – Y chromosome – PCR-multiplex

*Soud. Lék., 51, 2006, No. 2, p. 30–34*

## Úvod

Výzkum lidské genetické variability má dlouhou tradici. V minulosti bylo zkoumáno početné množství populací za použití velkého množství rozličných genetických markerů (2). V poslední době se populační genetické studie zaměřují

především na haploidní non-rekombinantní lokusy lidského genomu (mtDNA a NRY – „*non-recombinant part of Y chromosome*“) zejména z důvodů snazší interpretace získaných výsledků (11, 12, 13, 15, 16, 19). V oblasti NRY se setkáváme zvláště se dvěma druhy genetických polymorfismů. Jedná se o bialelické polymorfismy a Y-STR. Bialelické polymorfismy, někdy také nazývané *binární* nebo *indel polymorfismy*, jsou bodové mu-

tace nebo malé inserce a delece v nekódujících oblastech NRY. Jsou využívány především v evolučních a fylogenetických studiích pro jejich relativně pomalé mutační rychlosti (1, 4, 6, 17). Byl vytvořen systém haploskupin založený na přítomnosti a nepřítomnosti specifických bialelických mutací těchto haploskupin, který je využíván při zkoumání historie a genetické příbuznosti lidských populací (18).

Druhou skupinou polymorfismů v oblasti NRY jsou Y-STR. Jedná se o klasickou mikrosatelitní DNA, tak jak ji známe z autosomálních oblastí, ovšem s tou výjimkou, že vzhledem k haploidnímu charakteru NRY se vyskytují u jedince pouze v jedné kopii (3, 5, 6). Do současné doby bylo popsáno 219 Y-STR (7), z nichž většina je polymorfních. Y-STR se dají využít ve forenzní genetice k odlišení paternálních linií, spíše než k identifikaci jednotlivce. Rovněž v kombinaci s bialelickými polymorfismy se používají v populačně genetických studiích.

V naší laboratoři se zabýváme výzkumem variability lidského chromosomu Y především pro účely forenzní a lidské populační genetiky. Na tomto místě představujeme nově vyvinutý systém pro získávání genetických informací ze čtyřech Y-STR polymorfních míst.

## Materiál a metody

V rámci našeho projektu jsme navrhli a prověřili PCR-tetraplex systém, kterým je možné testovat čtyři Y-STR polymorfismy, jež se amplifikují simultánně. Jsou to *DYS449*, *DYS456*, *DYS458* a *DYS464*. Tyto lokusy byly vybrány podle několika hledisek. Nejdůležitější skutečností bylo, že žádný z nich nebyl přítomen v komerčně dostupných analytických soupravách v době začátku naší práce. Dalšími hledisky byla vysoká diverzita lokusů v euroamerické populaci, tedy velký počet alel a jejich zhruba rovnoměrné zastoupení v populaci (10). Tyto údaje pro všechny 4 Y-STR lokusy jsou uvedeny v tabulce 1.

Pomocí programu *Primer3* (Whitehead Institute for Biomedical Research) (14) jsme navrhli unikátní dvojice oligonukleotidů (primerů) pro amplifikaci vybraných čtyřech polymorfismů pomocí metody PCR. Předpokládané velikosti PCR produktů jsme zvolili tak, aby se tyto nepřekrývaly a pro jejich následnou fragmentační ana-

Tab. 1. Popis použitých Y-STR polymorfismů.

Údaje podle (10).

Diverzita v EU/US populaci počítána podle (9)

| Y-STR         | Typ repeticce | Motiv                                                   | Počet opakování (n) | Diverzita v EU/US populaci |
|---------------|---------------|---------------------------------------------------------|---------------------|----------------------------|
| <i>DYS449</i> | tetra         | (TTTC) <sub>n</sub> N <sub>50</sub> (TTTC) <sub>n</sub> | 21-36               | 0.823                      |
| <i>DYS456</i> | tetra         | (AGAT) <sub>n</sub>                                     | 13-18               | 0.771                      |
| <i>DYS458</i> | tetra         | (GAAA) <sub>n</sub>                                     | 14-20               | 0.812                      |
| <i>DYS464</i> | tetra         | (CCTT) <sub>n</sub>                                     | 11-19               | 0.939                      |

Tab. 2. Informace o použitých primerech a očekávaných PCR produktech: F – forward primer, R – reverse primer, T<sub>m</sub> – teplota tání primerů.

Číslo v závorce u velikosti produktů udává, pro jaký počet repetice platí uvedená velikost. Očekávané velikosti produktů byly navrhovány s vědomím, že s velkou pravděpodobností budou objeveny nové alely.

| Y-STR         | Sekvence primerů             | Velikost primerů | T <sub>m</sub> | Velikost produktů |
|---------------|------------------------------|------------------|----------------|-------------------|
| <i>DYS449</i> | F-TGG AGT CTC TCA AGC CTG TT | 20               | 57.58          | 414 (29)          |
|               | R-ATG GCA GTC ACC TGT AAT CC | 20               | 57.45          |                   |
| <i>DYS456</i> | F-TCA GCC TGC AGA TGG TCT    | 18               | 57.90          | 326 (15)          |
|               | R-TTT TGA ACT CTT GGC CTC AA | 20               | 58.47          |                   |
| <i>DYS458</i> | F-GGT GGT GGA GGT TAC TGT GA | 20               | 58.41          | 233 (16)          |
|               | R-TTC CTG ACC TTG TGA TCC AG | 20               | 58.64          |                   |
| <i>DYS464</i> | F-TTT ACG AGC TTT GGG CTA TG | 20               | 58.07          | 532 (13)          |
|               | R-CCT CAA CGC CTC TTA ACT CA | 20               | 58.10          |                   |

lyzu bylo možné použít značení pouze jednou fluorescenční barvou. Sekvence všech primerů byly testovány vůči sobě na potenciální vznik dimerů pomocí online volně dostupného *Oligo Analysis & Plotting Tool* (Operon), a ty, které se vzájemně doplňovaly ve více než čtyřech bázích, nebyly pro další práci použity. Sekvence primerů, jejich velikosti, teploty tání a očekávané velikosti produktů jsou uvedeny v tabulce 2. Pro fragmentační analýzu byl vždy levý primer označen 6-FAM (fluorescein) na jeho 5' konci. Primery vyrobila a fragmentační analýzu provedla firma GENERI-BIO-TECH, s.r.o. (ČR).

V první fázi výzkumu jsme testovali primery optimalizací PCR reakcí pro každý polymorfismus zvlášť. Následovala optimalizace celého PCR-tetraplex systému. Zaměřili jsme se především na úpravy koncentrací Mg<sup>2+</sup> iontů, reakčního pufru (K<sup>+</sup> iontů) a anelační teploty. Testovali jsme také vliv PCR enhanceru BSA (boviní sérový albumin, Promega, 0,1-0,8 µg/µl). Výnos PCR reakcí s BSA byl však výrazně horší, a proto jsme od přidávání tohoto reagentu upustili. Plně optimalizovaná PCR reakce pro všechny čtyři Y-STR polymorfismy fungovala za následujících podmínek: v reakčním objemu 10 µl za použití 1 µl lidské dvouvláknové DNA jako templátu

(200 ng/μl), izolované z periferní krve vysolovací metodou (8) – 3,0 mM MgCl<sub>2</sub>; 1,4x PCR pufr (70 mM KCl, 14 mM Tris-HCl); 250 μM dNTPs; 0,15 U Taq (vše TaKaRa) a 4 páry primerů (každý 200 μM).

Kontroly amplifikací byly prováděny na 2% agarózových gelech (Serva) v 0,5x TBE (tris-borátový/EDTA elektroforetický pufr, pH = 8,0) značených etidium bromidem (Sigma) při 80 V po dobu 55–75 min. Výsledky neuvedeny.

Tímto postupem byl zpracován soubor 50 náhodně vybraných jedinců mužského pohlaví z genové banky výzkumného centra EuroMISE centrum-Kardio.

## Výsledky a diskuse

Hlavním cílem našeho výzkumu bylo rozšířit možnosti použití Y-STR pro forenzní a populačně genetické výzkumné účely. K tomu jsme si vybrali 4 Y-STR polymorfismy (*DYS449*, *DYS456*, *DYS458*, *DYS464*). Jejich umístění na chromosomu Y ilustruje obr. 1. Z obrázku je rovněž patrné, že *DYS449*, *DYS456* a *DYS458* se vyskytují v oblasti NRY pouze v jedné kopii, ovšem *DYS464* ve čtyřech kopiích (*DYS464a*, *b*, *c*, *d*), a každá z těchto kopií má unikátní alelu *DYS464*, která se může, ale nemusí lišit od zbývajících třech alel lokusu. Tato skutečnost může působit menší potíže při zpracovávání údajů z fragmentační analýzy. Mohou totiž nastat tyto čtyři situace i) 4, ii) 3-1, iii) 2-2, iv) 1-1-1-1. Proto je nutné manuálně, podle výšky jednotlivých peaků, zhodnotit, jaké alely polymorfismu *DYS464* jsou přítomny v daném vzorku. Pro bližší informace viz (9).

Frekvence výskytu jednotlivých alel v našem zkoumaném souboru ilustrují grafy 1-4.

Genová diverzita (podle 9) pro náš soubor (N = 50) je následující:  $h_{449} = 0,8438$ ,  $h_{456} = 0,7682$ ,  $h_{458} = 0,7048$ ,  $h_{464} = 0,8194$ . Porovnáním našich zjištění s údaji uvedenými v tab. 1 můžeme vyvodit několik závěrů. Námi pozorovaná diverzita je u polymorfismů *DYS456*, *DYS458*, *DYS464* nižší pravděpodobně díky nízkému počtu zkoumaných jedinců. U polymorfismu *DYS449*

jsme zjistili vyšší diverzitu než autoři citované studie. To může být způsobeno i zachytem alely 31.2. V případě této alely se však může jednat o chybu metody (fragmentační analýza) a bude nezbytné zpracovat větší soubor s pokusem opět zachytit tuto hypotetickou alelu a tím ověřit její existenci. Dalším zajímavým polymorfismem je *DYS464*. V předchozím textu jsme již zmiňovali, že v NRY se vyskytují 4 kopie lokusu *DYS464*. Ovšem tyto kopie se dědí v paternální linii jako tzv. „single-gene locus“ díky absenci rekombinace. *DYS464* se tedy chová jako jeden velký STR „superlokus“ s charakteristikou (CCTT)<sub>n</sub>-N<sub>x</sub>-(CCTT)<sub>n</sub>-N<sub>x</sub>-(CCTT)<sub>n</sub>, kdy *n* nabývá hodnot 11-19. Nedostatkem současné metodologie je neschopnost odlišit, z kterého *DYS464* lokusu pochází která pozorovaná alela. Jinými slovy, například u alel 12-13-14-15 nejsme schopni odlišit, která z těchto alel se nachází na lokusu *DYS464a*, která na lokusu *DYS464b*, atd. Tento nedostatek by mohl být podstatný pouze v případě potřeby genetické identifikace dvou jedinců (nebo v případě NRY paternálních linií), kteří by se lišili nikoliv délkou alel na *DYS464*, ale pouze jejich pozicí. Například 12a-13b-14c-15d versus 12a-14b-13c-15d (písmena a, b, c, d uvádějí část *DYS464*, na které se konkrétní alela vyskytuje). V tomto případě by bylo řešením použít větší množství STR markerů a nalézt lokusy, v kterých se tyto dva jedinci odlišují. Naproti tomu velkou výhodou lokusu *DYS464* je jeho vysoká diverzita, pramenící právě ze čtyř kopií *DYS464* na každém chromosomu Y.

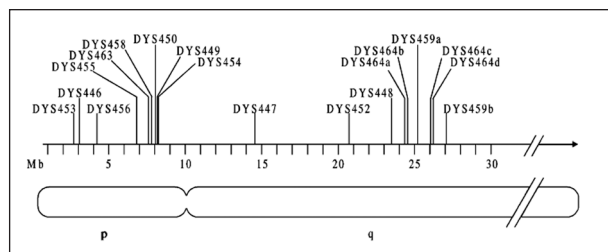
Jak již bylo uvedeno výše, předpokládáme také existenci dalších, vzácně se vyskytujících alel na všech čtyřech polymorfních místech. V současné době probíhá v naší laboratoři výzkum zaměřený na minoritní subpopulace na území střední Evropy, který by mohl některé tyto nové alely pospat.

Závěrem je třeba připomenout, že metodologie použitá v této práci významně šetří čas a laboratorní náklady. Čtyři lokusy jsou amplifikovány pomocí PCR za použití stejného množství template DNA jako při klasických amplifikacích každého lokusu zvlášť, a tak tento přístup umožňuje neplýtvat cenným genetickým materiálem vzorku, jehož obvykle bývá například ve forenzní genetické praxi velmi limitované množství.

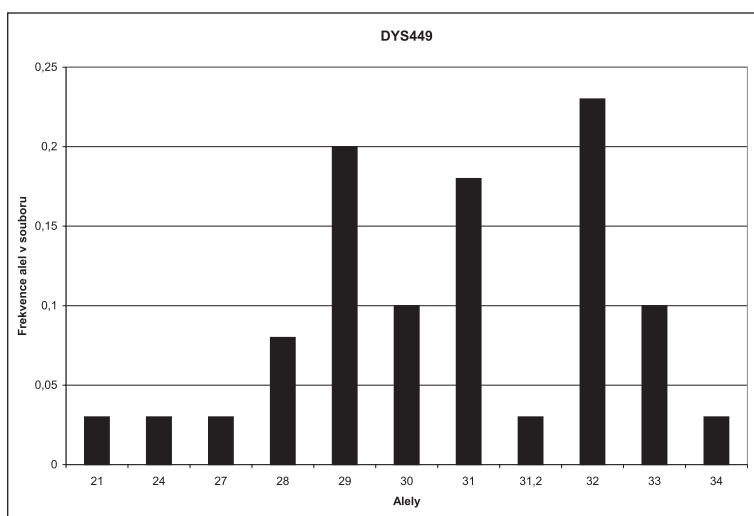
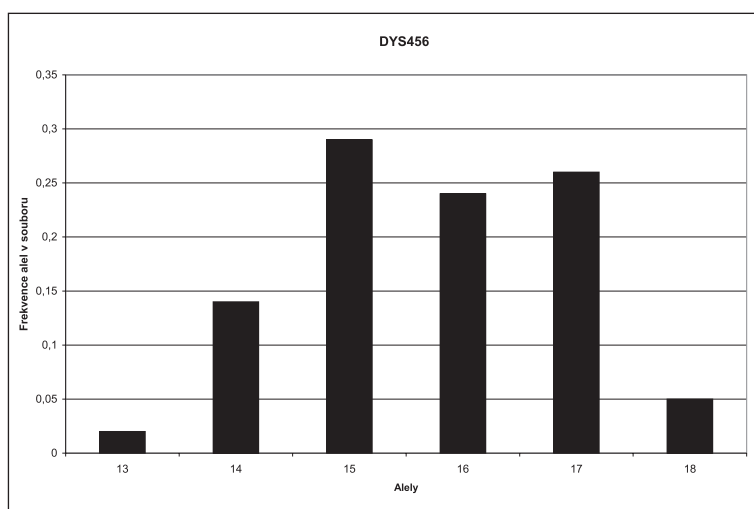
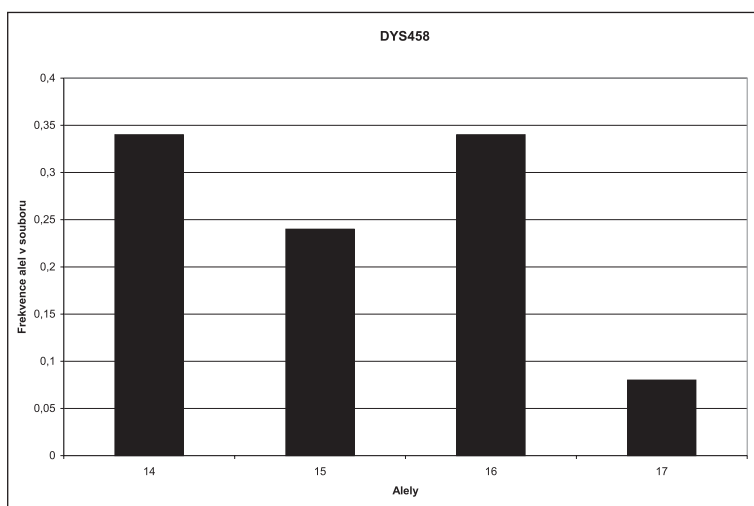
## Poděkování

Rádi bychom poděkovali všem kolegům, kteří se podíleli na sběru genetického materiálu v podobě krevních vzorků a na izolacích NK. náš dík patří také firmě GENERI BIOTECH, s.r.o. za provedenou fragmentační analýzu.

Tento projekt byl plně financován z prostředků výzkumného centra EuroMISE (č.grantu: LN00B107).

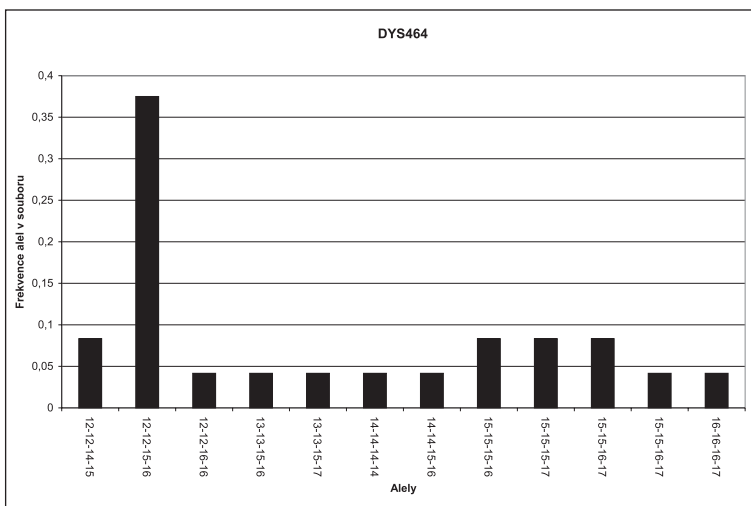


Obr. 1. Rozložení Y-STR na chromosomu Y (převzato z 10)

Graf 1. Zastoupení alel pro polymorfismus *DYS449*Graf 2. Zastoupení alel pro polymorfismus *DYS456*Graf 3. Zastoupení alel pro polymorfismus *DYS458*

## Literatura

**Brion, M., Salas, A., González-Neira, A., Lareu, M.V., Carracedo, A.:** Insight into Iberian population origins through the construction of highly informative Y-chromosome haplotypes using biallelic markers, STRs, and the MSY1 minisatellite. *Am. J. Phys. Anthropol.*, 122, 2003, s. 147–161. – 2. **Cavalli-Sforza, L.L., Menozzi, P., Piazza, A.:** The history and geography of human genes. Princeton: Princeton University Press, 1994. – 3. **Forster, P., Röhl, A., Lünemann, P. et al.:** A short tandem repeat-based phylogeny for the human Y chromosome. *Am. J. Hum. Genet.*, 67(1), 2000, s. 182–196. – 4. **Francalacci, P., Morelli, L., Underhill, P.A. et al.:** Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. *Am. J. Phys. Anthropol.*, 121(3), 2003, s. 270–279. – 5. **Heyer, E., Puzmirat, J., Dietjes, P., Bakker, E., de Knijff, P.:** Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum. Mol. Genet.*, 6(5), 1997, s. 799–803. – 6. **Jobling, M.A., Hurles, M., Tyler-Smith, Ch.:** Human evolutionary genetics: Origins, peoples and disease. New York: Garland Science, 2004. – 7. **Kayser, M., Kittler, R., Erler, A. et al.:** A comprehensive survey of human Y-chromosomal microsatellites. *Am. J. Hum. Genet.*, 74(6), 2004, s. 1183–1197. – 8. **Miller, S.A., Dykes, D.D., Polesky, H.F.:** A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, 16(3), 1988, s. 1215. – 9. **Nei, M.:** Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA*, 70, 1973, s. 3321–3323. – 10. **Redd, J.A., Agellon, A.B. Kearney V.A. et al.:** Forensic value of 14 novel STRs on the human Y chromosome. *Forensic Sci. Int.*, 130(2-3), 2002, s. 97–111. – 11. **Richards, M.B., Macaulay, V.A., Bandelt, H.J., Sykes, B.C.:** Phylogeography of mitochondrial DNA in Western Europe. *Ann. Hum. Genet.*, 62(3), 1998, s. 241–260. – 12. **Rootsi, S., Magri, Ch., Kivisild, T. et al.:** Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.*, 75(1), 2004, s. 128–137. – 13. **Rosser, Z.H., Zerjal, T., Hurles, M.E. et al.:** Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am. J. Hum. Genet.*, 67, 2000, s. 1526–1543. – 14. **Rozen, S. and Skaletsky H.J.:** Primer3 on the WWW for general users and for biologist programmers. In: Krawetz, S., Misener, S., eds. *Bioinformatics, methods and protocols: methods in molecular biology*. Totowa, NJ, Humana Press: 2000, 365–386. – 15. **Semino, O., Magri, Ch., Benuzzi, G. et al.:** Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the mediterranean area. *Am. J. Hum. Genet.*, 74(5), 2004,



Graf 4. Zastoupení alel pro polymorfismus *DYS464*

s.1023–1034. – 16. **Torroni, A., Huoponen, K., Francalacci, P. et al.**: Classification of European mtDNAs from an analysis of three European populations. *Genetics*, 144(4), 1996, s. 1835–1850. – 17. **Underhill, P.A., Shen, P., Lin, A.A. et al.**: Y chromosome sequence variation and the history of human populations. *Nat. Genet.*, 26(3), 2000, s. 358–361. – 18. **Y Chromosome Consortium (YCC)**: A nomenclature system for the tree of Y chromosomal binary haplogroups. *Genome Res.*, 12(2), 2002, s. 339–348. – 19. **Zhivotovsky, L.A., Underhill, P.A., Cinnioglu, C. et al.**: The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.*, 74(1), 2004, s. 50–61.

Mgr. Edvard Ehler

Katedra antropologie a genetiky člověka, PřF

UK v Praze

Viničná 7

Praha 2, 128 44

Tel.: (+420)221951621

E-mail: eda.ehler@seznam.cz

## ZPRÁVY

### Medicínské informační centrum projektů EU

Barbořáková H., Špunda M., Štětková P.

Medicínské informační centrum pro evropské projekty (MICEP) je nově vytvořeným specializovaným kontaktním bodem rámcových programů EU pro oblast zdravotnictví. Jedním z hlavních podnětů vedoucích k založení tohoto informačního centra byla nízká účast medicínských projektů výzkumu a vývoje v rámcových programech EU, které znamenají pro naši vědu nejen příležitost k získání finančních prostředků, ale zejména možnosti mezinárodní vědecko-výzkumné spolupráce mezi členskými zeměmi.

Centrum bylo založeno v lednu 2005 s podporou Univerzity Karlovy v Praze a její 1. lékařské fakulty. Finanční zdroje pro první roky činnosti poskytuje Ministerstvo školství, mládeže a tělovýchovy ČR společně s 1. LF UK. Hlavní náplní MICEP je poskytování informací o možnostech financování vědy, výzkumných a vývojových projektů ze zdrojů EU, zprostředkování kontaktů a asistence při přípravě projektů.

Služeb Medicínského informačního centra pro evropské projekty mohou využívat všechny subjekty s návazností na zdravotnictví (VŠ, nemocnice, výzkumné ústavy, malé a střední podniky (MSP) aj.). Činnost centra je koordinována s národním kontaktním bodem pro ČR, kterým je Národní informační centrum pro evropský výzkum (NICER), působící na půdě Akademie věd České republiky.

Očekávaným výsledkem činnosti Medicínského informačního centra pro evropské projekty bude lepší informovanost medicínských a ostatních zdravotnických výzkumných pracovišť o možnostech účasti v evropských projektech. V důsledku této lepší informovanosti a aktivního působení MICEP na konkrétní potenciální řešitele projektů rámcových programů v České republice a řešitele v partnerských zemích EU je možné očekávat i zlepšení bilance účasti České republiky v rámcových programech.

MICEP zahájilo činnost v květnu 2005 a otevřelo svou kancelář ve Faustově domě na Karlově náměstí v Praze. Aktuální informace o rámcových programech zveřejňuje na <http://micep.cuni.cz>.

Doc. ing. Miloslav Špunda, CSc.

Medicínské informační centrum projektů EU

Karlovo náměstí 40

120 00 Praha 2

fax: +420 224 965 843

e-mail: micep@cuni.cz