UNIVERZA V LJUBLJANI
BIOTEHNIŠKA FAKULTETA

Vanja KASTELIC

VPLIV POLIMORFIZMA POSAMEZNIH NUKLEOTIDOV NA PIGMENTACIJSKE SPREMEMBE NA OČEH IN LASEH PRI SLOVENSKI POPULACIJI

DOKTORSKA DISERTACIJA

Ljubljana, 2013

UNIVERZA V LJUBLJANI
BIOTEHNIŠKA FAKULTETA

Vanja KASTELIC

VPLIV POLIMORFIZMA POSAMEZNIH NUKLEOTIDOV NA PIGMENTACIJSKE SPREMEMBE NA OČEH IN LASEH PRI SLOVENSKI POPULACIJI

DOKTORSKA DISERTACIJA

ASSOCIATION OF SINGLE - NUCLEOTIDE POLYMORPHISMS WITH THE PREDICTION OF EYE AND HAIR COLOUR IN SLOVENE POPULATION

DOCTORAL DISSERTATION

Ljubljana, 2013

Na podlagi Statuta Univerze v Ljubljani ter po sklepu Senata Biotehniške fakultete in sklepa Senata Univerze z dne 6. maja 2010 (po pooblastilu s 6. seje Senata Univerze v Ljubljani z dne 20. januarja 2009) je bilo potrjeno, da kandidatka izpolnjuje pogoje za neposreden prehod na doktorski Podiplomski študij bioloških in biotehniških znanosti ter opravljanje doktorata znanosti s področja genetike.

Celotna raziskava je bila opravljena v Nacionalnem forenzičnem laboratoriju Ministrstva za notranje zadeve Republike Slovenije v Ljubljani. Uvodna računalniška in statistična obdelava podatkov je bila izvedena v Inštitutu za forenzične preiskave v Krakovu, Poljska, (Institute of Forensic Research, Section of Forensic Genetics, Krakow, Poland). Za mentorico je bila imenovana prof. dr. Katja Drobnič.

Mentorica: prof. dr. Katja Drobnič

Komisija za oceno in zagovor:

Predsednik:	prof. dr. Radovan KOMEL Univerza v Ljubljani, Medicinska fakulteta
Član:	prof. dr. Damjan GLAVAČ Univerza v Ljubljani, Medicinska fakulteta
Član:	prof. dr. Katja DROBNIČ Nacionalni forenzični laboratorij, Ministrstvo za notranje zadeve Republike Slovenije

Datum zagovora: 19. december 2013

Delo je rezultat lastnega raziskovalnega dela. Izjavljam, da so vsa vključena znanstvena dela enaka kot v znanstvenih publikacijah objavljena verzija. Podpisana se strinjam z objavo svoje naloge v polnem tekstu na spletni strani Digitalne knjižnice Biotehniške fakultete. Izjavljam, da je naloga, ki sem jo oddala v elektronski obliki, identična tiskani verziji.

Vanja Kastelic

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

ŠD Dd
DK UDK: 575.21:572.52:340.6 (043.3)=163.6=111
KG SNP/amelogeninski gen/gen SRY/pigmentacijske spremembe/fenotipske lastnosti/forenzična genetika/biološke sledi/analize DNA/določanje spola/barva oči/barva las/populacija/Slovenija
AV KASTELIC, Vanja, univ. dipl. mikr.
SA DROBNIČ, Katja (mentorica)
KZ SI-1000, Ljubljana, Jamnikarjeva 101
ZA Univerza v Ljubljani, Biotehniška fakulteta, Podiplomski študij bioloških in biotehniških znanosti, področje Genetika
LI 2013
IN VPLIV POLIMORFIZMA POSAMEZNIH NUKLEOTIDOV NA PIGMENTACIJSKE SPREMEMBE NA OČEH IN LASEH PRI SLOVENSKI POPULACIJI
TD Doktorska disertacija
OP VIII, 57 str., 3 sl., 2 pril., 73 vir.
IJ sl
JI sl/en
AI Vidna oziroma fenotipska lastnost pri ljudeh, ki jo je moč nedvoumno napovedati na podlagi DNA označevalcev, je vsekakor spol. Za določanje spola se dandanes v komercialnih kompletih najpogosteje uporablja amelogeninski gen. Populacijske študije pa so odkrile relativno visoko stopnjo delecij na kopiji tega gena na kromosomu Y, kar privede do napačne določitve spola. Za pravilnejšo napoved spola smo skonstruirali nov par začetnih oligonukleotidov znotraj novega genetskega označevalca na genu *SRY* in ga uvedli v komercialni komplet AmpFlSTR® SGM™ Plus amplification (AB). Validacijska študija (115 moških) je potrdila visoko zanesljivost in občutljivost uvedbe novega para začetnih oligonukleotidov v komercialni komplet. Rezultat analize je bil zanesljiv že pri vnosu 25 pg DNA. Naslednji pomembni fenotipski lastnosti pri ljudeh za razreševanje kaznivih dejanj sta tudi pigmentacijska obarvanost oči in las. Za obarvanost oči so odgovorne številne spremembe v genih, ki so posledica različnih polimorfizmov posameznih nukleotidov (SNP-jev). Obarvanost las pa je posledica kombinacije polimorfizma posameznih nukleotidov in zunanjih dejavnikov. Namen raziskave je bil razviti nov in zanesljiv molekularni komplet za določanje obarvanosti oči in las neznane osebe iz različnih bioloških sledi za forenzične in druge aplikacije. Na začetku raziskave smo izbrali dvanajst SNP-jev znotraj šestih pigmentacijskih genov. Razvili smo zanesljiv in občutljiv komplet, ki omogoča sočasno pomnoževanje vseh dvanajstih SNP-jev že iz 65 pg DNA. Frekvence alelov oziroma poli-/monomorfnost SNP-jev kompleta so potrdile, da je slovenska populacija del širše evropske kavkazijske populacije. Z analizo binarne logistične regresije smo ugotovili, da so SNP-ji rs12913832, rs1393350, rs1800407, rs1805008, in rs7495174 statistično najbolj specifični za slovensko populacijo (104 prostovoljcev), ki smo jih uporabili v analizah za preverjanje uspešnosti napovedi za posamezno pigmentacijsko obarvanost. Za napoved obarvanosti oči in las smo izbrali in primerjali dva statistična napovedna modela, model logistične regresije in naivni Bayesov model. Izmed petih SNP-jev najbolj uspešno napoved za pigmentacijsko obarvanost oči in las podaja SNP rs12913832, medtem ko je napoved z uporabo preostali štirih SNP-jev manj uspešna. Ocena uspešnosti napovedi, izražena s površino pod krivuljo ROC (AUC), je 1.0 za modre in 0.832 za rjave oči ter 0.913 za blond in 0.832 za temno rjave/črne, kar kaže na zelo visoko uspešnost napovedi, ki je bila določena za 24 slovenskih prostovoljcev. Uspešnost modela logistične regresije za napoved obarvanosti oči pri slovenski populaciji smo še dodatno preverili na edinem do sedaj objavljenem validiranem napovednem modelu forenzičnega kompleta IrisPlex, ki vsebuje šest SNP-jev. Napovedni model temelji na podatkih 3804 nizozemskih prostovoljcev. Uspešnost napovedi, izražena z AUC, je v tem primeru 0.966 za modre in 0.913 za rjave oči, določena za 105 slovenskih prostovoljcev. To kaže na podobno uspešnost napovedi, kot je bila ocenjena za širšo evropsko populacijo.

KEY WORDS DOCUMENTATION

DN DD
DC UDC: 575.21:572.52:340.6 (043.3)=163.6=111
CX SNP/amelogenin gene/sry gene/human pigmentation/phenotype prediction/forensic genetic/biological traces/DNA analysis/ gender determination/eye color/hair color/ population/ Slovenia
AU KASTELIC, Vanja
AA DROBNIČ, Katja (supervisor)
PP SI-1000 Ljubljana, Jamnikarjeva 101
PB University of Ljubljana, Biotechnical Faculty, Postgraduate Study of Biological and Biotehnical Sciences, Field: Genetic
PY 2013
TI ASSOCIATION OF SINGLE - NUCLEOTIDE POLYMORPHISMS WITH THE PREDICTION OF EYE AND HAIR COLOUR IN SLOVENE POPULATION
DT Doctoral dissertation
NO VIII, 57 p., 3 fig., 2 ann., 73 ref.
LA sl
AL sl/en
AB The human externally visible characteristic that so far is most accurately predictable with DNA markers is human gender. A length differences between the X-chromosomal and the Y-chromosomal copy of the amelogenin gene is useful, but not always correct for DNA-based sex determination, nevertheless this marker is included in most of the commercially available kits. The irregularities could occur because of amelogenin deletions and that why we included new sex-specific SRY marker in the one of most usable commercial kit (AmpFlSTR® SGM™ Plus amplification (AB). Validation study (115 men) confirms that incorporation of the new primer set for SRY marker in the commercial kit in very reliable and sensitive. Reliable results were obtained even from as little input of DNA as 25 pg. Besides human gender we investigated two additional externally visible characteristic – human eye and hair colour. The coloration of eye and hair is primarily associated with the number of independently contributing SNPs, the genetic effect of each selected SNP and also of some non-genetic influences, regarding hair colour. We limited on twelve SNPs inside six pigmentation genes and developed a single multiplex genotyping assay. This assay is very robust, reliable and sensitive for multiplex genotyping of all 12 SNPs, even from 65 pg of DNA. The allele frequencies of all SNPs somehow confirmed that Slovene population is part of Caucasian population. With further binary logistic regression analysis we concluded that five (rs12913832, rs1393350, rs1800407, rs1805008, rs7495174) out of twelve SNPs are most statistical specific for our Slovene sample population (104 volunteers). For the evaluation of the accuracy of the prediction of the eye and hair colour, based on five mentioned SNPs, we used logistic regression model and naive Bayesian model and compared them. For the both models the eye and hair colour were categorized into three and into two states. From five SNPs the rs12913832 determine the highest predictive value for both eye and hair colour and the other four SNPs had also influence on better prediction, of course not as obvious as mentioned one. The overall prediction accuracies expressed by the area under the receiver characteristic operating curves (AUC) were 1.0 for blue and 0.832 for brown eyes and also 0.913 for blond and 0.832 for dark brown/black hair colour (using logistic regression model), which is highly optimistic, but indicated the prediction for only 24 Slovenian volunteers (the rest of volunteers were used for model-building set). Additionally for better presentation of prediction accuracy for eyes colour we used all Slovene volunteers (105 volunteers) and used prediction model of the only validated IrisPlex assay, which is based on genotype and phenotype data from 3804 Dutch volunteers. Prediction accuracies expressed by the AUC were found to be 0.966 for blue eyes and 0.913 for brown eyes for Slovene population. This indicates prediction accuracy comparable to previously established for other European populations.

KAZALO VSEBINE

KLJUČNA DOKUMENTACIJSKA INFORMACIJA III
KEY WORDS DOCUMENTATION IV
KAZALO VSEBINE V
KAZALO ZNANSTVENIH DEL VI
KAZALO SLIK VII
KAZALO PRILOG VII
OKRAJŠAVE IN SIMBOLI VIII
1 PREDSTAVITEV PROBLEMATIKE IN HIPOTEZE 1
1.1 GEN *SRY* V SPOLNEM KROMOSOMU Y 2
1.2 POLIMORFIZEM POSAMEZNIH NUKLEOTIDOV IN PIGMENTACIJSKE ZNAČILNOSTI 4
1.2.1 Biopolimer melanin 5
1.2.2 SNP-ji v forenzičnih preiskavah 6
1.2.3 Raziskanost SNP-jev v pigmentacijskih genih 8
1.3 STATISTIČNA NAPOVEDNA MODELA 9
1.3.1 Naivni Bayesov model 9
1.3.2 Model logistične regresije 11
1.3.3 Preverjanje uspešnosti statističnih napovednih modelov s pomočjo ROC analize 12
1.4 HIPOTEZE 13
2 ZNANSTVENA DELA 14
2.1 VALIDACIJA SRY OZNAČEVALCA ZA PRIMERE FORENZIČNIH PREISKAV 14
2.2 HKRATNO POMNOŽEVANJE MINISEKVENČNIH ZAČETNIH OLIGONUKLEOTIDOV: POVEZAVA MED PETIMI SNP-JI TER OBARVANOSTJO OČI IN LAS V SLOVENSKI POPULACIJI TER PRIMERJAVA BAYESOVEGA MODELA IN MODELA LOGISTIČNE REGRESIJE 20
2.3 DOLOČANJE OBARVANOSTI OČI V SLOVENSKI POPULACIJI Z UPORABO SNP JEV KOMPLETA IRISPLEX 29
2.4 DOLOČITEV ZUNANJEGA VIDEZA LJUDI S PREISKAVAMI DNK 36
3 RAZPRAVA IN SKLEPI 41
3.1 RAZPRAVA 41
3.1.1 Gen *SRY* v spolnem kromosomu Y 41
3.1.2 Polimorfizem posameznih nukleotidov in pigmentacijske lastnosti 42
3.2 SKLEPI 46
4 POVZETEK (SUMMARY) 48
4.1 POVZETEK 48
4.2 SUMMARY 49
5 VIRI 52
ZAHVALA
PRILOGA

KAZALO ZNANSTVENIH DEL

Kastelic V., Budowle B., Drobnič K. 2009. Validation of SRY marker for forensic casework analysis. Journal of Forensic Sciences, 54, 3: 551–555 14

Kastelic V., Drobnič K. 2012. A single-nucleotide polymorphism (SNP) multiplex system: the association of five SNPs with human eye and hair colour in the Slovenian population and comparison using a Bayesian network and logistic regresion model. Croatian Medical Journal, 53, 5: 401–408 20

Kastelic V., Branicki W., Pośpiech E., Draus-Barini J., Drobnič K. 2013. Prediction of eye colour in the Slovenian population using the IrisPlex SNPs. Croatian Medical Journal, 54, 4: 381-386 29

Kastelic V., Drobnič K. 2012. Določitev zunanjega videza ljudi s preiskavami DNK. Revija za kriminalistiko in kriminologijo, 63, 3: 225–228 36

KAZALO SLIK

Slika 1: Prikaz humanega kromosoma Y z lokacijo gena SRY (Y chromosome, 2008) 3

Slika 2: Sintezna pot dveh različnih vrst melanina iz osnovne aminokisline tirozin (Sturm in sod., 2001) 6

Slika 3: Podaljševanje posameznega nukleotida s pomočjo podaljševanja začetnega nukleotida (ABI…, 2010) 8

KAZALO PRILOG

Priloga A: Pomnoženega gena SRY (96 bp) z uporabo novega para začetnih oligonukleotidov v komercialnem kompletu AmpFlSTR® NGM™ (AB)

Priloga B: Dovoljenja založnikov za objavo člankov

OKRAJŠAVE IN SIMBOLI

AB	podjetje Applied Biosystems
AIM	manjkajoč pri melanomu (angl. absent in melanoma)
ACTH	adrenokortikotropični hormon (angl. adrenocorticotropic hormone)
AUC	površina pod krivuljo ROC (angl. area under ROC curve)
cAMP	ciklični adenozinmonofosfat (angl. cyclic adenosine monophosphate)
DAM	moški z delecijo v amelogen. genu (angl. deleted amelogenin males)
DHI	dihidroksiindol (angl. dihydroxyindole)
DHICA	dihidroksiindol karboksilna kislina (angl. dihydroxyindole carboxylic acid)
DNA	deoksiribonukleinska kislina (angl. deoxyribonucleic acid)
dNTP	deoksinukleotidtrifosfat
DOPA	dihidroksifenilalanin (angl. dihydroxyphenylalanine)
HW	Hardy-Weinbergovo (HW) ravnovesje
LD	vezano neravnovesje (angl. linkage disequilibrium)
LOH	izguba heterozigotnosti (angl. loss of heterozygosity)
LR	razmerje verjetij (angl. likelihood ratio)
MSH	melanocite-stimulirajoči hormon (angl. melanocyte-stimulating hormone)
NRY	nerekombinirajoč del kromosoma Y (angl. non-recombining Y)
PCR	verižna reakcija s polimerazo (angl. polymerase chain reaction)
ROC	krivulja občutljivosti in specifičnosti (angl. receiver operating characteristic curve)
RFU	relativne fluorescenčne enote (angl. relative fluorescence units)
SBE	podaljševanje posamezne baze oziroma nukleotida (angl. single base extension)
SNP	ponovitve posameznega nukleotida (angl. single nucletide polymorphism)
SRY	lokus, ki določa spol (angl. sex determining region)
STR	male tandemne ponovitve (angl. short tandem repeat)
TDF	faktor, ki določa spol (angl. testis determining factor)
VNTR	variabilno število tandemnih ponovitev (angl. variable num. tandem repeat)
TP	pravilna pozitivna napoved (angl. true positive)
TN	pravilna negativna napoved (angl. true negative)
FP	napačno pozitivna napoved (angl. false positive)
FN	napačno negativna napoved (angl. false negative)
TPR	resnično pozitivni rezultati (angl. true positive rate)
TNR	lažno pozitivni rezultati (angl. true negative rate)
TYR	tirozin (angl. tyrosine)
TYRP1, TYRP2	tirozin fosfataza 1 oziroma 2 (angl. tyrosine phosphatases)
ZFX, ZFY	cinkov prst na lokusu X oziroma Y (angl. zinc finger protein locus)

1 PREDSTAVITEV PROBLEMATIKE IN HIPOTEZE

Osnovni namen forenzičnih genetskih preiskav je individualizacija biološke sledi, kot so kri, semenska tekočina, slina, epitelne celice, kosti, zobje…, zavarovane med ogledi krajev kaznivih dejanj ali v zvezi z njimi, lahko pa so del preiskav množičnih grobišč, množičnih nesreč, pogrešanih oseb in drugo (Drobnič, 2004).

Preiskave DNA imajo danes v forenzičnih preiskavah nedvomno nenadomestljivo vlogo. Razvoj molekularne genetike je omogočil razlikovanje osebe od vseh drugih na podlagi razlik v nekodirajočih delih humane DNA, ki jo imenujemo odpadna DNA (Fowler in sod., 1988). Največji del med njimi predstavljajo ponavljajoča se zaporedja DNA, ki jih delimo na minisatelite (VNTR-označevalce) in mikrosatelite (STR-označevalce) (Tautz in sod., 1984). Slednji se danes najpogosteje pojavljajo v rutinskih forenzičnih genetskih preiskavah in skupaj z amelogeninskim označevalcem za spol predstavljajo forenzični genetski standard nacionalnih ali kriminalističnih evidenc DNA oseb in bioloških sledi. Zaradi milijonskega števila profilov DNA v različnih evidencah DNA in izrednega tehnološkega razvoja so avtosomalni STR-označevalci postali nenadomestljivi v rutinskih forenzičnih preiskavah (Asplen, 2009).

Vendar pa se za različne primere, kot so preiskave spolnih deliktov, ugotavljanje identitete umorjenega, arheološke raziskave in drugo, ki jih zaradi različnih težav ni mogoče razrešiti le z analizo avtosomalnih STR-označevalcev, uveljavljajo tudi druge forenzične genetske preiskave. To so genetske preiskave **STR-označevalcev oziroma genov na spolnih kromosomih** in analize mitohondrijske DNA (Lessing in sod., 2005), v novejšem času pa tudi **analize polimorfizmov posameznih nukleotidov** (SNP) za napovedovanje zunanjega videza posameznika ali njegovega biogeografskega izvora (Budowle, 2004; Kayser in Schneider, 2009).

V začetku smo našo raziskavo usmerili v izboljšanje zanesljivosti genetskega označevalca za določanje spola, ki predstavlja enega glavnih fenotipskih značilnostih ljudi. Pravilnejše določanje spola bo temeljilo na vpeljavi novega spolnega genetskega označevalca v do sedaj najbolj razširjene komercialne komplete. Pri vseh komercialnih kompletih določanje spola temelji na preiskavi polimorfizma amelogeninskega gena, ki je podvržen številnim delecijam (Cadenas in sod., 2007) oziroma se mutacije pojavljajo na mestih prileganja začetnih oligonukleotidov, ki so vključeni v komercialne komplete (Santos in sod., 1998; Roffey in sod., 2000; Thangaraj in sod., 2002). Skonstruirali smo nov par začetnih oligonukleotidov znotraj gena *SRY* z namenom, da bo mogoče nov genetski označevalec SRY uspešno vključiti v komercialne komplete za določanje identitete posameznikov, ki jih uporabljamo v rutinskih preiskavah. Validacijske študije smo izvedli za komercialni komplet AmpFlSTR® SGM™ Plus amplification (AB) (AmpFlSTR…, 1997) v skladu z zahtevami znanstvenega forenzičnega združenja SWGDAM (The Scientific Working Group on DNA Analysis Methods).

V nadaljevanju doktorske naloge smo se osredotočili na povsem novo in malo raziskano področje raziskav, to je določanje dveh fenotipskih značilnostih ljudi, in sicer pigmentacijske obarvanosti oči in las. Omenjeni in ostale (npr. pigmentacijska obarvanost kože, struktura obraza, plešavost in tudi višina posameznika) fenotipske značilnosti so poligenske in po dosedanjih raziskavah ima na njih največji vpliv polimorfizem posameznih nukleotidov (SNP) (Budowle, 2004; Kayser in Schneider, 2009). Namen raziskovanja SNP-označevalcev je bil razviti nov občutljiv in zanesljiv molekularni genetski komplet za določanje obarvanosti oči in las neznane osebe (oškodovanec, storilec kaznivega dejanja, neznano truplo) iz bioloških sledi za forenzično uporabo. Napovedovanje obarvanosti oči in las predstavlja velik potencial za prihodnost. Rezultati teh genetskih analiz bi lahko nadomestili ali dopolnili vlogo očividcev kaznivih dejanj, saj je poročanje posameznega očividca vedno zelo subjektivno in močno odvisno od več dejavnikov. Tako bodo te informacije o zunanjem videzu posameznika vodile predvsem k zmanjšanju kroga potencialnih osumljencev in s tem usmerjeno vodile kriminalistične preiskovalce do najverjetnejših storilcev, predvsem v primerih težjih kaznivih dejanj. V zadnjem delu raziskav smo se zaradi tega osredotočili na izbiro statističnih napovednih modelov za analizo podatkov. Najbolj razširjena statistična napovedna modela sta naivni Bayesov model in model logistične regresije. Njuno uspešnost napovedi za obarvanosti oči in las smo ocenili z ROC analizo oziroma s pomočjo kriterija AUC (Liu in sod., 2009). Na koncu smo ocenili, kateri od statističnih napovednih modelov je bolj uspešen ter zaradi neraziskanosti področja uspešnost napovedi primerjali z že objavljenimi rezultati nekaterih raziskav (Ruiz in sod., 2012; Walsh in sod., 2013).

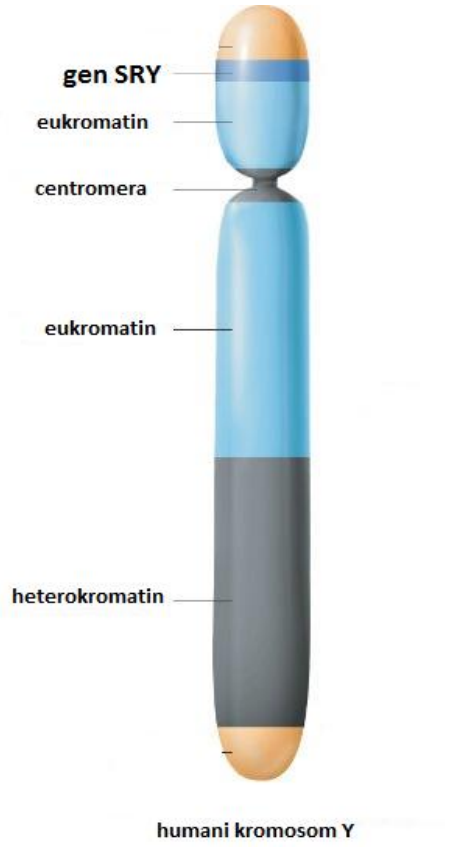
1.1 GEN *SRY* V SPOLNEM KROMOSOMU Y

V forenzičnih preiskavah določanje spola temelji predvsem na variabilnosti med kopijama amelogeninskega gena, odgovornega za razvoj zobne sklenine, na kromosomih X in Y. Pri pomnoževanju amelogeninskega gena s specifičnimi začetnimi oligonukleotidi, vključenimi v komercialne komplete, nastane produkt dolžine 106 bp na kromosomu X in dolžine 112 bp na kromosomu Y (Sullivan in sod., 1993). Kljub razširjeni uporabi amelogeninskega gena pa rezultati ne podajajo vedno pravilnega spola donorja posamezne biološke sledi (Santos in sod., 1998; Roffey in sod., 2000; Steinlechner in sod., 2002). Do napak pri določanju spola pride predvsem zaradi strukturnih nepravilnosti znotraj kromosoma Y. Tako se npr. delecije znotraj amelogeninskega gena na kromosomu Y (*AMELY*) kažejo v izgubi produkta pomnoževanja na omenjenem kromosomu in s tem v napačni napovedi spola osebe. Odstotek delecij na omenjenem genu se giblje od 0,018 % do 8,0 %, zaradi česar so rezultati bioloških preiskav premalo zanesljivi (Santos in sod., 1998; Roffey in sod., 2000; Steinlechner in sod., 2002; Thangaraj in sod., 2002; Brinkman, 2002; Mitchell in sod., 2006; Chang in sod., 2007; Turrina in sod., 2011). V slovenski populaciji se za enkrat ta delecija pojavlja pri približno 0,012 % moških (Drobnič, 2006). Torej so

rezultati, po katerih se določa spol pri genetskih forenzičnih preiskavah, premalo zanesljivi, kar kaže na potrebo po vpeljavi drugih, zanesljivejših genetskih označevalcih.

Mnogo raziskovalcev je začelo vključevati druge genetske označevalce v identifikacijske teste za ugotavljanje spola. Med njimi so lokusi STR, specifični le za kromosom Y – najbolj raziskani so mikrosateliti DYS19, DYS385, DYS388, DYS389, DYS390 in drugi (Kayser in sod., 1997). Uporabljajo se tudi homologni geni v spolnih kromosomih, kot sta npr. lokusa *ZFX* in *ZFY* (Reynolds in Varlaro, 1996) ter gen *SRY* na kromosomu Y (Sullivan in sod., 1993; Dennis Lo in sod., 1998; Santos in sod., 1998; Singh in sod., 1999; Kastelic in sod., 2009, Tozzo in sod., 2013).

Genetski označevalec v genu *SRY* (angl. sex determining region) je eden od najprimernejših predvsem zaradi svoje evolucijske ohranjenosti (McElreavy in sod., 1992).



Slika 1: Prikaz humanega kromosoma Y z lokacijo gena SRY (Y chromosome, 2008)
Figure 1: Location of SRY gene on human Y chromosome (Y chromosome, 2008)

Produkt gena *SRY* je protein SRY oziroma transkripcijski faktor TDF, ki z vezavo na DNA spremeni njeno obliko. S tem se sproži kaskada genetskih interakcij in izločanje specifičnih hormonov, kar povzroči, da se potencialne ženske gonade pretvorijo v moške testise (McElreavy in sod., 1992). Deluje s pozitivno in negativno regulacijo, in sicer v zgodnjem razvoju zarodka inducira razvoj moških specifičnih telesnih znakov in obenem inhibira razvoj ženskih (Haqq in sod., 1993).

1.2 POLIMORFIZEM POSAMEZNIH NUKLEOTIDOV IN PIGMENTACIJSKE ZNAČILNOSTI

Pigmentacijska obarvanost oči in las, ki tudi določata zunanji videz posameznika, sta zelo kompleksni lastnosti, saj nanju vpliva večje število različnih genov, med katerimi potekajo medsebojne interakcije, hkrati pa prihaja tudi do interakcije teh genov z okoljem (Mertens, 2009). Ravno kompleksnost omenjenih lastnosti najbolj otežuje raziskanost posameznih genov in njihovo posredno vlogo pri izražanju fenotipskih značilnostih. Do sedaj je poznanih že mnogo genov, ki vplivajo na posamezno značilnost, najbolj pa stopajo v ospredje geni, ki so vključeni v razvoj specifične pigmentacijske obarvanosti predvsem oči in las. Prve uspešne preiskave na področju pigmentacijskih genov so izvedli na miših in določili približno 120 genov, ki bi lahko vplivali na obarvanost dlake, medtem ko so jih le približno 30 povezali tudi z redkimi pigmentacijskimi obolenji pri ljudeh. Med njimi je do sedaj najbolj raziskanih enajst genov, ki bolj specifično vplivajo na pigmentacijske značilnosti in jih je moč uporabiti tudi za forenzične preiskave (Branicki in sod., 2008; Valenzuela in sod., 2010). Najbolj raziskan pigmentacijski gen je bil sprva gen *MC1R*, ki pri ljudeh vpliva na razvoj rdečkaste obarvanosti las in na svetlo polt. Sedaj pa sta po raziskanosti in pomembnosti bolj pomembna gena *OCA2* in *HERC2*, ki sta ključna pri izražanju obarvanosti očesne šarenice in obarvanosti las (Visser in sod., 2012). Znani pa so še mnogi drugi pigmentacijski geni – *SLC45A2*, *SLC24A5*, *SLC24A4*, *KITLG*, *TYRP1*, *TYR*, *DCT* in *IRF4* – ki so tudi predmet obsežnejših preiskav v povezavi s pigmentacijskimi značilnostmi (Kayser in Schneider, 2009).

V zgodnjih raziskavah pigmentacijskih značilnosti je bil eden pomembnejših genov gen *OCA2*, ki so mu v več raziskavah pripisovali posredno vlogo pri določanju rjavih in/ali modrih oči pri ljudeh. Gen *OCA2* kodira protein, ki ima funkcijo membranskega proteina (Sturm in Frudakis, 2004). Funkcija proteina še ni v celoti raziskana, naj pa bi bil vpleten pri dobavljanja substratov pri biosintezi melanina (Eiberg in sod., 2008).

Čeprav so omenjeni gen dolgo imeli za enega glavnih dejavnikov pri pigmentacijski obarvanosti oči, je nedavno kar nekaj raziskovalcev znotraj gena *HERC2* določilo vsaj dve pomembni spremembi, ki naj bi bili ključni za pigmentacijsko obarvanost oči. Funkcija proteina je še neznana, vendar pa naj bi bile spremembe znotraj gena *HERC2*

ključne pri regulaciji ekspresije gena *OCA2* in s tem ključne pri določanju obarvanosti oči pri ljudeh (Sulem in sod., 2007; Eiberg in sod., 2008; Kayser in sod., 2008).

Kljub obširni raziskanosti omenjenih dveh genov pa raziskave na področju pigmentacijskih genov segajo v leto 1995, ko je Velverde s sodelavci iskal povezave med pigmentacijsko obarvanostjo las in kože s spremembami v genu *MC1R* (Velverde in sod., 1995). Prav spremembe tega gena naj bi vplivale na pigmentacijske značilnosti, kot so rdeči lasje, svetlejši ten kože, prisotnost pegic ter močna občutljivost kože na UV sevanje. Spremembe gena *MC1R* obenem najpogosteje povezujejo tudi s povečanim tveganjem za tumorskim obolenjem kožnih celic (Harding in sod., 2000; Kanetsky in sod., 2006; Branicki in sod., 2007). Gen *MC1R* se izraža v različnih celicah, ki sodelujejo pri pigmentacijskem sistemu, vendar pa sama funkcija gena še ni v celoti znana (Rees, 2000).

Ravno tako sta bila že leta 2006 gena *SLC24A5* in *SLC45A2* vključena v skupino humanih pigmentacijskih genov. Gen *SLC45A2* kodira protein, ki sodeluje pri sintezi melanina v melanosomih in naj bi obenem sodeloval tudi kot membranski prenašalec. Mutacije tega gena v številnih raziskavah povezujejo tudi s človeškimi obolenji albinizma različnih oblik (Soejima in Koda, 2007).

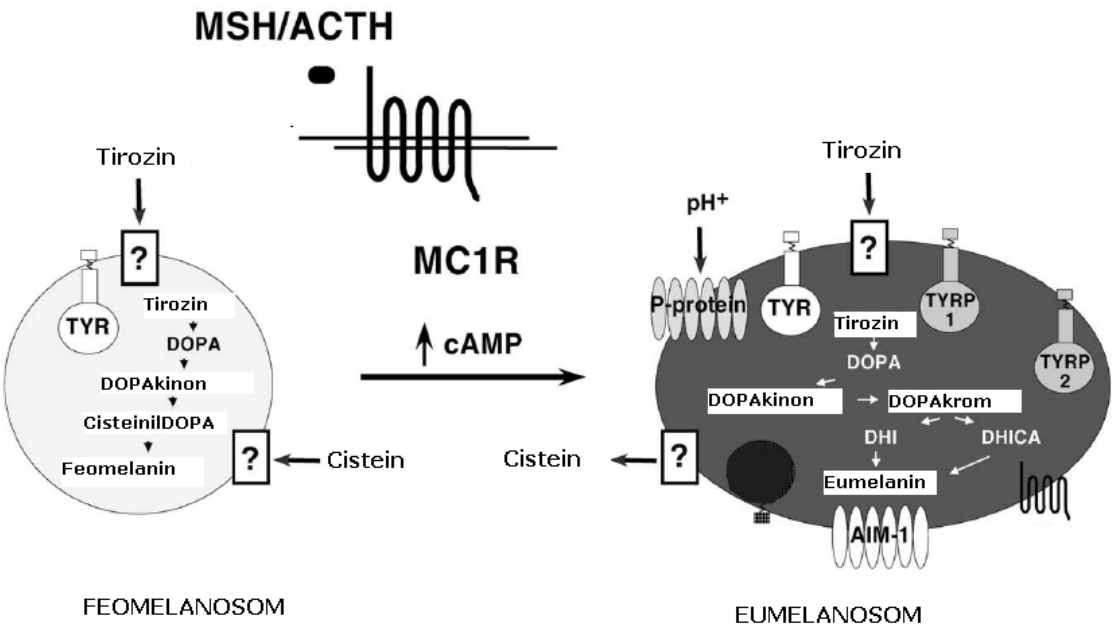
Pri melanogenezi je eden ključnih faktorjev tudi gen *TYR*, ki sodeluje pri prvih dveh reakcijah biosinteze melanina. Humani gen *TYR* se nahaja na kromosomu 11 in je dolg več kot 50 kb. Pri ljudeh, obolelih za okulokutanim albinizmom tipa 1 (OCA1), so do sedaj odkrili do 100 mutacij tega gena. Mutacije lahko zmotijo normalno sintezo melanina in s tem vplivajo na pigmentacijsko obarvanost oči, las in kože (Nakamura in sod., 2002).

1.2.1 Biopolimer melanin

Vsi zgoraj omenjeni pigmentacijski geni s svojimi spremembami vplivajo na pigmentacijske značilnosti ljudi, ki so odvisne predvsem od števila, tipa in celične porazdelitve melanosomov znotraj pigmentacijskih celic (melanocit ali keratinocit). Melanosomi so specializirani organeli, ki vsebujejo ključni encim za nastanek obeh oblik melanina in sodelujejo pri njegovem transportu (Passeron in sod., 2005). Biopolimer melanin nastane z oksidacijo aminokisline tirozin in se v posameznih pigmentacijskih celicah nahaja v dveh oblikah, in sicer v obliki rjavo-črnega eumelanina in rdeče-rumenega feomelanina, ki na različne načine absorbirata svetlobo (Sturm in sod., 2001). Torej tako oblika melanina kot tudi številčnost in porazdelitev melanosmov v pigmentacijskih celicah, ki se nahajajo predvsem v očesni stromi šarenice (melanocite) oziroma v korteksu lasnega stebla in v kožnem bazalnem nivoju med

dermisom in epidermisom (keranocite), ključno vplivajo na končno pigmentacijsko obarvanost oči, las in kože pri ljudeh (Sturm, 1998).

Makromolekula melanina je polimer indol 5,6-kinona ter 5,6 dihidroksiindol-2-karboksilne kisline. Oksidacija aminokisline tirozina poteka v prvih dveh korakih s pomočjo encima tirozinaza, in sicer s hidroksilacijo tirozina do 3,4-dihidroksifenilalanina (DOPA) in oksidacijo DOPA do DOPA-kinona. Eumelanin nato nastane preko metabolitov DOPA-kromov, feomelanin pa preko metabolitov 5-S-cisteinil-DOPA. Izomerizacija DOPA-kroma do 5,6 dihidroksiindol-2-karboksilne kisline (DHICA) je katalizirana s strani DOPA-krom tautomeraze, oksidacijo DHICA pa omogoča DHICA-oksidacijski encim (Slika 1) (Sturm in sod., 2001).



Slika 2: Sintezna pot dveh različnih vrst melanina iz osnovne aminokisline tirozin (Sturm in sod., 2001)
Figure 2: The synthetic pathway of two types of melanin from the tyrosine (Sturm in sod., 2001)

LEGENDA; MSH - melanocite-stimulirajoči hormon, ACTH - adrenokortikotropični hormon, MC1R – melanokortin 1 receptor, cAMP - ciklični adenozinmonofosfat, TYR – tirozin, TYRP1, TYRP2 - tirozin fosfataza 1 oziroma 2, AIM - manjkajoč pri melanomu , DOPA – dihidroksifenilalanin

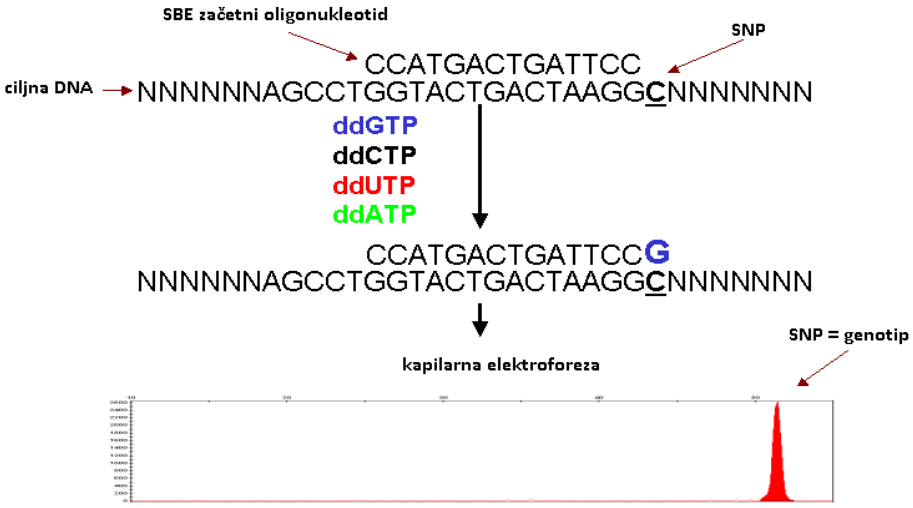
1.2.2 SNP-ji v forenzičnih preiskavah

Vse predhodno omenjene spremembe v posameznih pigmentacijskih genih in njihov vpliv na izražanje drugih sosednjih genov in obenem na sintezo najpomembnejše pigmentacijske molekule melanina se nanašajo na spremembe v genih, ki so v obliki substitucij (zamenjave med bazami), insercij (dodatek ene baze) ali delecij (izguba ene

baze). Torej so te spremembe tako imenovani polimorfizmi posameznih nukleotidov (SNP), ki so zelo številčni v humanem genomu in zato kljub njihovi preprostosti in dokaj nizki informativnosti posameznega SNP-ja zelo zanimivi za čedalje obširnejše preiskave. Ocenjeno je, da obstaja več kot 10 milijonov možnih mest v humanem genomu, kjer se lahko pojavi sprememba v eni bazi in jim pripisujejo kar 90-odstotno odgovornost za razlikovanje med posamezniki na nivoju DNA. Med drugim pa imajo SNP-ji zelo nizko mutacijsko stopnjo (velikostnega razreda 10^{-8}) (Kidd in sod., 2006), ki je približno sto tisočkrat manjša kot pri STR označevalcih (danes najpogosteje uporabljeni v forenzičnih preiskavah), in so obenem tako bolj primerni za določevanje evolucijskega izvora posameznika, dokazovanje starševstva in tudi za določevanje fenotipskih značilnostih ljudi (Budowle, 2004). SNP-je je možno analizirati s pomočjo več različnih metod in možna je hkratna analiza več SNP-jev, kar omogoča večjo fleksibilnost in avtomatiziranost analiz (Brookes, 1999).

Večje zanimanje za raziskovanje SNP-jev se je začelo leta 2000, ko sta dva večja konzorcija (Human Genome Project in International HapMap Consortium) začela z raziskavami in objavami SNP-jev. Sedaj objavljena zbirka obsega že 53.558.214 specifičnih sprememb humanega genoma (dbSNPSummary, 2012), od tega jih kar 85,0 % predstavljajo SNP-ji (Giampoli in sod., 2012). Ravno s podatki iz omenjene zbirke SNP-jev in njihovimi alelnimi frekvencami, specifičnimi za posamezno populacijo, smo lahko primerjali alelne frekvence oziroma njihovo poli-/monomorfnost s podatki, pridobljenimi za našo populacijo. Slovensko populacijo smo primerjali z alelnimi frekvencami CEU, ki predstavljajo evropsko populacijo. Podatki izvirajo od prebivalcev zvezne države Utah (ZDA), ki so potomci prebivalcev iz severne in zahodne Evrope (Giampoli in sod., 2012).

Za analize SNP-jev je razvitih in se še razvija mnogo različnih genotipizacijskih metod, ki temeljijo na različnih molekularnih mehanizmih in z njimi povezanimi detekcijskimi metodami. V naši raziskavi smo za analizo SNP-jev uporabili metodo SNaPshot™ ob uporabi komercialnega kompleta SNaPshot™ Multiplex Kit (AB) (ABI…, 2010). Metoda SNaPshot™ je že uveljavljena kot občutljiva, ponovljiva, robustna in fleksibilna metoda ter cenovno dokaj primerna za forenzične preiskave. Komercialna metoda SNaPshot™ (AB) je osnovana na začetnem minisekveniranju, ki mu sledi elektroforezno ločevanje produktov in hkratna fluorescenčna analiza le-teh. Pri SNaPshot™ metodi so neoznačeni začetni oligonukleotidi izbrani tako, da se vežejo tik pred mestom točkovne mutacije, medtem ko se na samo polimorfno mesto med pomnoževanjem veže posamezni fluorescenčno označeni ddNTP. Ustrezno kapilarno ločevanje produktov po barvi in dolžini je omogočeno z dodajanjem nespecifičnih nukleotidnih repkov na 5'- konec minisekvenčnih oligonukleotidnih začetnikov, ki se med seboj razlikujejo za nekaj baznih parov (Sanchez in Borsting, 2003; Sobrino in sod., 2005).



Slika 3: Podaljševanje posameznega nukleotida s pomočjo podaljševanja začetnega nukleotida (ABI…, 2010)
Figure 3: Single nucleotide primer extension (ABI…, 2010)

1.2.3 Raziskanost SNP-jev v pigmentacijskih genih

Prve objave na področju pigmentacijskih genov, ki segajo že v leto 1995, so takrat najbolj raziskani gen *MC1R* in njegove SNP-je – rs1805005, rs1805007, rs1805008 – povezovale z že prej omenjeno rdečkasto obarvanostjo las, svetlejšim tenom kože ter z močno občutljivostjo kože na UV sevanje (Velverde in sod., 1995). Vendar pa so te preiskave deloma obstale, saj so ti polimorfizmi vezani na manjši del populacije, in sicer s severnejšega predela Evrope in njihove potomce (Box in sod., 1997; Smith in sod., 1998).

Kljub temu da so SNP-jem (rs7170989, rs7495174, rs1800407 in rs1667394) znotraj gena *OCA2* do nedavnega pripisovali največji pomen pri pigmentacijski obarvanosti oči in las, pa so študije SNP-jev rs12913832 in rs1129038 znotraj gena *HERC2* (Visser in sod., 2012) močno spremenile mnenja. Nakazujejo namreč, da bi lahko ravno gen *HERC2*, če ne morebiti ravno omenjeni SNP rs12913832, imel poglavitno vlogo pri pigmentacijski obarvanosti oči, las in kože pri ljudeh. Raziskave so tako dognale, da je SNP rs12913832, ki leži v intronu 86 gena *HERC2* (to je 21 kb navzgor nad promotorjem gena *OCA2*), glavni ojačevalec ekspresije gena *OCA2*. T-alel SNP-ja rs12913832 tako okrepi transkripcijske faktorje, ki v kombinaciji s kromatinsko zanko spodbudijo transkripcijo gena *OCA2*. To omogoči povečano sintezo melanina in s tem povezano temnejšo pigmentacijsko obarvanost. Po drugi strani pa C-alel SNP-ja rs12913832 tako spremeni vezavno mesto transkripcijskih faktorjev, da se ne tvori ustrezna oblika kromatinske zanke in zato ne pride do intenzivnejše sinteze melanina. Do sedaj so bile ravno te raziskave najbolj intenzivne in usmerjene predvsem v

potrditev, kako lahko posamezni SNP vpliva na specifične pigmentacijske značilnosti, vendar se moramo zavedati, da na specifično pigmentacijsko značilnost vpliva še mnogo drugih genov in njihovi polimorfizmi posameznih nukleotidov (Visser in sod., 2012).

Za pigmentacijski gen *SLC45A2* sta najbolj raziskana polimorfizma rs26722 in rs16891982. SNP rs16891982 naj bi imel posredni vpliv, torej vpliv v interakciji z drugimi pigmentacijskimi geni oziroma SNP-ji, na temnejšo obarvanost las (Branicki in sod., 2008). V prvih raziskavah so velik vpliv na pigmentacijsko obarvanost, in sicer na svetlejšo polt, pripisovali tudi genu *SLC24A5* in SNP-ju rs1426654 (Soejima in Koda, 2007). Vendar pa so nadaljnje preiskave potrdile, da je omenjeni SNP fiksiran v celotni evropski populaciji in torej neinformativen za določanje večine pigmentacijskih značilnostih. V povezavi z določenimi oblikami albinizma so bile opravljene tudi mnoge analize gena *TYR* in analize SNP-jev rs1393350 in rs1126809, ki naj bi vplivala na svetlejšo polt in svetlejšo pigmentacijsko obarvanost las in oči (Nakamura in sod., 2002; Soejima in Koda, 2007).

1.3 STATISTIČNA NAPOVEDNA MODELA

Za oceno uspešnosti napovedovanja pigmentacije obarvanostioči in las 12 SNP-jev smo uporabili dva statistična napovedna modela, in sicer naivni Bayesov model in model logistične regresije. Pri tem smo v obeh primerih za učno skupino uporabili 80 naključno izbranih prostovoljcev (77,0 %) z znanimi genotipskimi in fenotipskimi lastnostmi. Oba statistična napovedna modela sta bila nato ocenjena s pomočjo preostanka 24 prostovoljcev (23,0 %) oziroma testne skupine z znanimi genotipskimi in fenotipskimi lastnostmi. Uspešnost napovedi je pri statističnih napovednih modelih v veliki meri odvisna od pravilne kategorizacije, ki pa je zaradi subjektivnega kvalitativnega ocenjevanja obarvanosti oči in las zelo otežena. Pri ocenjevanju smo pigmentacijsko obarvanost oči in las razdelili v tri oziroma dve kategoriji. Posameznike smo po obarvanosti oči razdelili najprej v tri kategorije: modra, rjava in vmesna obarvanost, v nadaljevanju pa le v dve kategoriji: svetla in temnejša obarvanost. Glede na obarvanost las smo jih sprva ravno tako porazdelili v tri kategorije: blond, temno blond/svetlo rjava in temno rjava/črna obarvanost, v nadaljevanju pa v le dve kategoriji: svetla in temnejša obarvanost (Pośpiech in sod., 2012).

1.3.1 Naivni Bayesov model

Naivni Bayesov model napoveduje verjetnost za fenotipsko značilnost posameznika (npr. obarvanost oči in las) z uporabo pogojne odvisnosti od vrednosti neodvisnih spremenljivk oziroma SNP-jev. Pri tem predpostavlja medsebojno pogojno neodvisnost med temi neodvisnimi spremenljivkami. Naivni naravi Bayesovega modela navkljub, se je v realnih primerih izkazal za dokaj uspešnega, njegova prednost pa je tudi ta, da

potrebuje majhno učno skupino za uspešnost napovedi (Halloran, 2009). Bayesov model torej omogoča izračun verjetnosti za obarvanost oči in las posameznika z znanim genotipom na podlagi predhodno določne posteriorne verjetnosti, pridobljene s pomočjo učne skupine. Na podlagi znanih genotipskih in fenotipskih lastnosti učne skupine smo določili pogojno verjetnost določenega genotipa za specifično obarvanost oči oziroma las (Pośpiech in sod., 2012).

Za posameznike iz testne skupine smo nato lahko na podlagi znanega genotipa petih SNP-jev določili posteriorno verjetnost oziroma napoved verjetnosti, da imajo specifično obarvanost oči oziroma las. Pri tem je bila uporabljena Bayesova enačba (1) kot kvocient dveh pogojnih verjetnosti; verjetnost opaženega genotipa, ki podaja posameznikovo specifično obarvanost oči oziroma las Pr (Mj/Hk=x), in seštevek verjetnosti opaženega genotipa, ki podaja posameznikovo drugačno obarvanost oči oziroma las $\sum_{k=1}^{n} \Pr(Mj / Hk)$.

Bayesova enačba (1) vsebuje tudi apriori komponento, torej razmerje verjetnosti, da ima posameznik v specifični populaciji značilno obarvanost oči oziroma las Pr ($H_{k=x}$) in seštevek verjetnosti, da ima posameznik v specifični populaciji drugačno obarvanost oči oziroma las $\sum_{k=1}^{n} \mathbf{Pr}(\boldsymbol{Hk})$. A ker nam izvor posameznika v forenzičnih preiskavah po navadi ni v celoti znan, apriori verjetnosti ne moremo upoštevati in je verjetnost Pr(Hk)=1/n za vse odtenke obarvanosti oči oziroma las enaka. Torej so bile apriori verjetnosti določene kot 0.33 pri kategoriziranju v tri skupine obarvanosti in kot 0.5 pri kategoriziranju v dve skupini obarvanosti, saj smo se s tem opredelili, da nam populacijski izvor posameznika ni znan (Pośpiech in sod., 2012).

Bayesova enačba:

$$\Pr(H_{k=x}/M_j = \frac{\Pr\left(\frac{M_j}{H_{k=x}}\right)\Pr(H_{k=x})}{\sum_{k=1}^{n}\Pr\left(\frac{M_j}{H_k}\right)\Pr(H_k)} \quad \ldots(1)$$

Mj – specifičen multilokusni genotip petih SNP-jev
Hk – specifična obarvanost oči oziroma las posameznika

Tako smo za 24 posameznikov v testni skupini določili posteriorno verjetnost oziroma napoved verjetnosti (angl. prediction probability) za specifično obarvanost oči oziroma las. Potem pa smo ocenili uspešnost napovedi omenjenega modela s pomočjo ROC analize, ki je opisana v nadaljevanju.

1.3.2 Model logistične regresije

Za razliko od naivnega Bayesovega modela sam diskriminativni model logistične regresije ne predpostavlja pogojne neodvisnosti med neodvisnimi spremenljivkami oziroma SNP-ji. Model logistične regresije je namreč osnovan tako, da napove izid na osnovi povezanosti neodvisnih spremenljivk in pripravi napovedno enačbo (Halloran, 2009). Torej smo uporabili logistično regresijo, kjer več neodvisnih spremenljivk (genotip posameznika, ki ga predstavlja pet specifičnih SNP-jev) hkrati vpliva na več odvisnih spremenljivk, in sicer multinominalni (obarvanost oči oziroma las, razdeljena v tri kategorije) oziroma binarni model (obarvanost oči oziroma las, razdeljena v dve kategoriji) (Liu in sod., 2009; Walsh in sod., 2011b).

V spodnjem primeru predstavljamo le primer, ko ima odvisna spremenljivka (Y) tri vrednosti (npr. modra, rjava in vmesna kategorija obarvanosti oči). Standardna multinominalna logistična regresija se za npr. modre oči izraža s funkcijo *logit* takole (2) (Liu in sod., 2009):

$$logit\left(P\left(y = modre\ oči\ \because\ x_1 \ldots x_5\right)\right) = \ln\frac{\pi_1}{\pi_3} = \alpha_1 + \sum \beta(\pi_1)_k x_k \quad \ldots(2)$$

x_k – je številka manj pogostega alela (angl. minor allele) določenega k-tega SNP-ja (k = 1-5)
π_1 – verjetnost, da ima posameznik npr. modre oči

V omenjenem primeru, ko ima odvisna spremenljivka Y tri vrednosti, izračunamo še dve *logit* funkciji (za rjavo in vmesno kategorijo obarvanosti oči), pri katerih je tretji *logit* odvisen od prvih dveh. Vse izračune *logit* funkcij smo opravili s pomočjo statističnega programa SPSS 19.0 (SPSS Inc., Chicago, IL, USA) in s tem ovrednotili modelna parametra, alfa (α) in beta (β), z uporabo enake učne skupine, kot smo jo uporabili pri prvem, naivnem Bayesovem modelu (Liu in sod., 2009).

Na podlagi določenih modelnih parametrov smo določili napoved verjetnosti (angl. prediction probability) za specifično obarvanost oči oziroma las pri vsakem posamezniku v testni skupini. Npr. pri obarvanosti oči, ki je razdeljena v tri kategorije, smo na osnovi logistične regresije določili verjetnosti, da je posameznik rjavook (π_1), modrook (π_2) oziroma da spada v vmesno kategorijo (π_3), na osnovi naslednje enačbe (3) (Liu in sod., 2009):

$$\pi_1 = \frac{\exp(\alpha_1 + \sum \beta(\pi_1)_k x_k}{1 + \exp(\alpha_1 + \sum \beta(\pi_1)_k x_k) + \exp(\alpha_2 + \sum \beta(\pi_2)_k x_k)} \quad \ldots(3)$$

x_k – je številka manj pogostega alela (angl. minor allele) določenega k-tega SNP-ja
α in β – modelna parametra

Potem smo ponovno ocenili uspešnost napovedi modela logistične regresije s pomočjo ROC analize.

1.3.3 Preverjanje uspešnosti statističnih napovednih modelov s pomočjo ROC analize

Verjetnosti napovedi za posamezno obarvanost oči oziroma las pri vsakem posamezniku testne skupine smo s pomočjo kontingenčne tabele (angl. confusion table) velikosti 2 x 2 razvrstili v štiri možne vrste izidov, ki so;

- pravilna pozitivna napoved – TP,
- pravilna negativna napoved – TN,
- napačna pozitivna napoved – FP in
- napačna negativna napoved – FN.

Za pravilno napovedano obarvanost smo na podlagi predhodnih objav (Liu in sod., 2009) v vseh primerih uporabili mejno vrednost verjetnost napovedi > 0.7 (angl. treshold). Torej smo pri celotni testni skupini za posamezno obarvanost oči oziroma las določili enega od možnih izidov. Npr. za modrooko skupino posameznikov smo najprej prešteli vse modrooke, ki so bili napovedani pravilno (TP), nato smo prešteli vse modrooke posameznike, ki niso bili napovedani kot modrooki (FN), nato smo prešteli vse posameznike, ki niso imeli modrih oči, vendar so bili napovedani kot modrooki (FP), na koncu smo prešteli še vse posameznike, ki niso imeli modrih oči in so bili napovedani kot nemodrooki (TN) (Liu in sod., 2009).

V povezavi z zgoraj opisanimi izidi je pri raziskavah SNP-jev najpogosteje uporabljani kriterij za preverjanje uspešnosti statističnih napovednih modelov predstavljen z naslednjimi vrednostmi:

- občutljivost (angl. sensitivity)
- specifičnost (angl. specificity),
- pozitivna napovedna vrednost – PPV in
- negativna napovedna vrednost – NPV

S pomočjo teh vrednosti smo nato s pomočjo programa SPSS 19.0 (SPSS Inc., Chicago, IL, USA) izvedli ROC analizo, s katero smo želeli grafično ponazoriti uspešnost dveh statističnih napovednih modelov. ROC krivulja ponazarja odnos specifičnosti in občutljivosti pri posameznem modelu. Na y osi grafa se nahaja občutljivost oziroma resnično pozitivni rezultati (angl. true positive rate – TPR), na x osi pa lažno pozitivni rezultati (angl. true negative rate – TNR) oziroma 1 – specifičnost. Tako model z visoko uspešnostjo napovedi dosega visoko občutljivost (veliko resnično pozitivnih rezultatov) in ima visoko specifičnost (majhno število lažno pozitivnih rezultatov). Pri modelih z bolj uspešno napovedjo se krivulja približuje levemu gornjemu kotu, kjer imamo veliko število resnično pozitivnih rezultatov in majhno število lažno pozitivnih rezultatov. Tako lahko vizualno ocenimo uspešnost napovedi posameznega modela in ju primerjamo med seboj (Zhu in sod., 2010).

1.4 HIPOTEZE

Hipoteze, ki smo jih postavili, so bile naslednje:

Določanje spola donorjev bioloških sledi v forenzičnih analizah lahko temelji tudi na paru novih začetnih oligonukleotidov znotraj gena *SRY* z amplifikacijskim produktom dolžine 96 baznih parov, ki se pojavi le pri moških.

Hkratno pomnoževanje novega para začetnih oligonukleotidov lahko poteka v hkratni PCR reakciji ob prisotnosti drugih enajstih začetnih oligonukleotidov ter ob pogojih PCR, ki jih narekuje komercialni validirani komplet AmpFlSTR® SGM™ Plus amplification (AB) (AmpFlSTR…, 1997). Vključitev novega para začetnih oligonukleotidov pa naj bi bila uspešna tudi pri ostalih komercialnih kompletih npr. AmpFlSTR® NGM™ amplification (AB) (Priloga A).

Analize dvanajstih polimorfizmov posameznih nukleotidov (SNP-jev) v povezavi s pigmentacijskimi značilnostmi ljudi bodo v novem kompletu potekale z novim naborom začetnih oligonukleotidov, primernih za hkratno pomnoževanje v hkratni klasični PCR reakciji (24 parov začetnih oligonukleotidov) in v hkratni minisekvenčni SNaPshot™ reakciji (12 parov začetnih oligonukleotidov).

Frekvence alelov oziroma poli-/monomorfnost dvanajstih SNP-jev kompleta bodo potrdile, da je slovenska populacija del širše evropske kavkazijske populacije (primerjava s podatki CEU na portalu NCBI SNP).

Variacije izbranih dvanajstih SNP-jev določajo specifično pigmentacijsko obarvanost oči in las v slovenski populaciji.

Variacije šestih SNP-jev (štirih SNP-jev v sklopu našega kompleta in dodatnih dveh SNP-jev) določajo specifično pigmentacijsko obarvanost oči v slovenski populaciji.

Izbrana statistična napovedna modela, naivni Bayesov model in model logistične regresije, bosta uspešno napovedala posamezno obarvanost oči in las za slovensko populacijo.

Uspešnost napovedi modela logistične regresije za specifično obarvanost oči bo primerljiva z objavljenimi raziskavami, ki zajemajo širšo evropsko populacijo.

2 ZNANSTVENA DELA

2.1 VALIDACIJA SRY OZNAČEVALCA ZA PRIMERE FORENZIČNIH PREISKAV

Validation of SRY marker for forensic casework analysis

V. Kastelic, B. Budowle in K. Drobnič

Journal of Forensic Sciences, 2009, 54, 3: 551–555
doi: 10.1111/j.1556-4029.2009.01007.x

Večina analiz v forenzičnih preiskavah temelji na preiskavah lokusov STR na avtosomalnih kromosomih, vendar sta spolna kromosoma ključna pri določanju spola posameznika. V forenzičnih preiskavah določanje spola temelji predvsem na preiskavi amelogeninskega gena. Čeprav je ta označevalec dokaj zanesljiv, se zadnje čase pojavljajo vprašanja o nedvoumnosti dobljenih rezultatov, saj je v literaturi opisanih vedno več primerov mutacij, katerim je podvržen amelogeninski gen zaradi katerih je določitev spola napačna. Mnogo raziskovalcev je začelo vključevati tudi druge genetske označevalce v identifikacijske teste za ugotavljanje spola, med drugimi tudi označevalce znotraj gena *SRY*, ki leži na kromosomu Y. Zato smo tudi mi znotraj njega skonstruirali nov par začetnih oligonukleotidov ter del gena uspešno pomnožili tudi v hkratni reakciji pomnoževanja znotraj enega bolj uporabljenih forenzičnih komercialnih kompletov. Za možnost vključitve novega genetskega označevalca SRY v komercialni komplet AmpFlSTR® SGM™ Plus amplification (AB) smo izvedli ustrezne validacijske študije. Validacijske študije so tako vključevale ponovljivost, občutljivost, spolno specifičnost in študije mešanih vzorcev za potrebe pomnoževanja novega para začetnih nukleotidov v hkratni reakciji PCR, ki jih pogojuje komercialni komplet. Izbrani SRY-označevalec se je izkazal za zanesljivega, saj smo produkt dolžine 96 bp uspešno pomnožili pri vseh 115 testiranih moških in med začetnimi oligonukleotidi nismo zaznali navzkrižnih pomnoževanj ali dimerov. Obenem pa tudi za občutljivega, saj smo 96 bp dolg produkt pomnoževanja zaznali tudi pri minimalnem začetnem vnosu DNA (25.0 pg). Nespecifičnega pomnoževanja nismo opazili niti pri 10 ng DNA ženske osebe. Označevalec SRY smo zaznali tudi v mešanih sledeh, ko je ženska DNA prevladovala nad moško v razmerju 16:1. Produkt pomnoževanja gena *SRY* smo dokazali tako tudi pri moških, ki imajo delecijo v amelogeninskem genu, oziroma pri moških, ki so bili določeni kot ženska oseba, ko smo za dokazovanje spola uporabili le amelogeninski genetski označevalec. Nov par začetnih oligonukleotidov je primeren za vključitev v komercialni komplet AmpFlSTR® SGM™ Plus amplification (AB). S vključitvijo novega para začetnih oligonukleotidov bi se zmanjšala možnost napačne napovedi spola donorja posamezne biološke sledi, saj bi komercialni kompleti tako vsebovali dva označevalca za moški spol. SRY-označevalec bi obenem služil kot notranja kontrola.

J Forensic Sci, May 2009, Vol. 54, No. 3
doi: 10.1111/j.1556-4029.2009.01007.x
Available online at: www.blackwell-synergy.com

TECHNICAL NOTE

Vanja Kastelic,[1] *B.S.; Bruce Budowle,*[2] *Ph.D.; and Katja Drobnič,*[1] *Ph.D.*

Validation of *SRY* Marker for Forensic Casework Analysis

ABSTRACT: Determining the gender of the source of forensic DNA evidence is based on the amelogenin test. However, at times the assay may not be indicative of gender assignment, because of deletions at the amelogenin site. Previously, we described successful coamplification of a marker residing within the *SRY* gene with the short tandem repeat markers from two commercially available human identification kits. The study herein addresses the validation of primers for the target *SRY* gene regarding specificity, sensitivity, and robustness. Among 115 unrelated male Slovenians no null allele was observed. Repeatable and reliable results were obtained from as little as 25 pg of template DNA, indicating a high sensitivity of detection for the assay. No polymerase chain reaction product was observed even at a concentration of 10 ng/μL of template female DNA. Additionally, the male specific marker could be detected in mixed male and female samples down to a ratio of 1:16.

KEYWORDS: forensic science, DNA typing, validation, sex determination, amelogenin gene, *SRY* gene, mutation

Determining the gender of the source of a forensic DNA sample at times can be informative in various forensic investigations, especially in sexual assault cases. Sex determination is routinely performed by amplification by the polymerase chain reaction (PCR) of a region of the amelogenin. The assay typically generates a 106 bp long fragment from the X chromosome and a 112 bp long fragment from the Y chromosome (1). In the forensic field, the amelogenin (*AMEL*) gender test is carried out as part of a multiplex assay using commercially available identification kits, such as AmpFlSTR® SGM™ Plus kit (Applied Biosystems, Foster City, CA) and PowerPlex® 16 System (Promega Corp., Madison, WI).

However, several studies have shown that the amelogenin gender test may not always be concordant with true male gender in forensic casework or in prenatal diagnosis (2–5). This discrepancy is because of the structural variability within the Y chromosome (6–9). Deletions of *AMELY* can result in no amplification product and these null *AMELY* alleles can occur up to about 8.0% in some population groups (3,5,7,8,10–17). In the Slovenian male population *AMELY* null alleles are infrequent occurring in one out of 8300 male individuals (17). Using the YFiler™ kit (Applied Biosystems), this *AMELY* null male also was null at the DYS458 locus. Thus, the data support that the null allele is likely the result of a larger deletion on the short arm of the Y chromosome.

To reduce the potential interpretation difficulties in the few cases where gender misinterpretation may be problematic, some authors have investigated using genetic markers lying in the sex-determination region Y (*SRY*) on the Y chromosome (6,7,9). In these studies, the *SRY* assay was performed as an additional singleplex PCR or in combination with primers for *AMEL*. However, this approach requires an additional assay subsequent to *AMEL* and short tandem repeat (STR) typing, thus consuming more evidentiary material, as well as being laborious and time consuming.

Drobnič (17) reported successful amplification of a novel marker residing in the *SRY* gene which results in a 96 bp long PCR product and be incorporated into either the AmpFlSTR® SGM™ Plus (Applied Biosystems) or PowerPlex® 16 System (Promega Corp.) identification kits. Thus, it is feasible to incorporate the *SRY* gene assay into any routine *AMEL* and STR analysis. Moreover, the small size of the *SRY* amplicon provides two benefits for forensic DNA testing. First, male gender determination can be successful when typing degraded forensic samples, at least as successful as that for *AMEL*. Second, because of its short length, the *SRY* amplification product does not migrate with any of the *AMEL* or STR alleles in the multiplex STR kits.

The present study was undertaken to perform some validation studies on the *SRY* marker for use in forensic cases. Validation of the *SRY* marker was performed in accordance with the recommendations of the SWGDAM revised validation guidelines (18). The validation studies included repeatability, sensitivity, gender specificity, and mixture studies.

Materials and Methods

Quantification and PCR Amplification

Quantification of DNA was conducted using the Quantifiler™ Human DNA Quantification Kit (Applied Biosystems) with 2.0 μL of DNA extract on the ABI Prism 7000 Sequence Detection System (Applied Biosystems).

Amplification of DNA was performed using the *SRY* primers under AmpFlSTR® SGM™ Plus (Applied Biosystems) manufacturer's recommendations as reported previously (17). A singleplex DNA amplification was carried out in a total volume of 25 μL containing 5 U/μL AmpliTaq Gold® Polymerase, 0.2 μM forward *SRY* primer, 0.24 μM reversed *SRY* primer and 10.0 μL 10× Gold® STR buffer (Applied Biosystems) in a Perkin-Elmer 9600 thermal cycler (Applied Biosystems). A singleplex DNA amplification was

[1]Forensic Science Centre, Ministry of the Interior, Vodovodna 95, Ljubljana, Slovenia.
[2]FBI Laboratory, Quantico, VA 22135.

Received 13 Mar. 2008; and in revised form 15 May 2008; accepted 14 July 2008.

used in all studies except for the mixture study. Mixture studies were carried out as multiplex amplification. The same amount of *SRY* primers as used in a singleplex reaction was coamplified with the AmpFlSTR® SGM™ reaction mixture in a total volume of 25 μL following the procedures described in the technical manual. Ten microliters of appropriately diluted DNA were added to each tube so that the final template input range was 0.1–1.0 ng.

DNA Typing

Amplified product was combined with the Genscan-500 ROX internal line standard and loaded on an ABI Prism® 310 Genetic Analyzer as described by the manufacturer (Applied Biosystems). Samples were injected for 5 sec at 15 kV and electrophoresis was conducted at 15 kV and 60°C with Performance Optimized Polymer 4 (POP™ 4, Applied Biosystems). Data from samples amplified using AmpFlSTR® SGM™ Plus PCR Amplification kit were collected using ABI Prism Collection software version 3.7 with virtual filter set F. Results were analyzed using GeneScan® 3.7 analysis software. Using Genotyper® version 3.7 analysis software (Applied Biosystems), STR allele designations were made based on comparison with the allelic ladder. The *SRY* allele calls were made manually using the amplicon length determined using GeneScan® 3.7 analysis software.

Repeatability

DNA samples were obtained from buccal swabs taken from 115 unrelated male individuals from our casework. The samples were extracted using the chelex extraction method (19). The samples were prepared by serial dilution from samples of known concentration. DNA was added in each PCR at a concentration range of 0.5 to 1.0 ng/μL and amplified under AmpFlSTR® SGM™ Plus manufacturer's recommendations. The samples were analyzed three times by the same operator using the same ABI Prism® 310 Genetic Analyzer (Applied Biosystems).

Sensitivity Studies

Varying amounts of male control DNA 007 (Applied Biosystems) (ranging from 0.025 to 1.0 ng/μL) and of male DNA casework samples (ranging from 0.0625 to 1.0 ng/μL) were amplified to determine the minimum amount of input DNA that could be used to obtain a *SRY* profile. Each quantity of control and DNA casework sample were tested three times.

Gender-Specificity Studies

A male control cell line DNA (9948), female cell line control DNA (9947A) (Promega Corp.) and one male and one female DNA sample from our casework were amplified at a concentration of 1.0 ng/μL, respectively.

Singleplex reactions using our designed primer set for the *SRY* locus were performed with female control DNA 9947A (Promega Corp.) at concentrations of 1, 5, and 10 ng/μL.

Mixture Studies

Male–female mixture studies were performed on five different sets of male and female DNAs from our casework at ratios of 1:1, 1:2, 1:4, 1:8, and 1:16. Each mixture was tested twice. The amount of female casework DNA was held constant at 1.5 ng, while the amount of male casework DNA varied from 1500 down to 93.7 pg.

Results and Discussion

A validation study was carried out to define some limitations of *SRY* typing of forensic specimens using the *SRY* marker developed by Drobnič (17). The validation studies included repeatability, sensitivity, gender specificity, and mixture analyses.

Repeatability

The amplification and typing of the *SRY* marker was successful for all 115 male samples and there were no discrepancies with gender assignment. The amplification was carried out in a singleplex reaction. All 115 male samples were tested at three different capillary electrophoresis conditions but with the same ABI Prism® 310 Genetic Analyzer (Applied Biosystems) and yielded sizes of 94.44 ± 0.07 bp, 94.47 ± 0.18 bp, and 94.36 ± 0.06 bp. These lengths in bp are slightly lower than the known 96 bp size of the amplicon. Such differences are well known when using capillary electrophoresis (20). However, the results demonstrate the more important factor for genetic typing that the precision is exceedingly high. The new primer set for the *SRY* marker enables precise and repeatable results for male gender determination as a single system. Because no null allele was observed there is strong support that the primer binding sites are conservative and that the *SRY* marker is a good candidate for a reliable male gender test for forensic purposes. However, with a sample of 115 males a null allele can occur at a frequency as high as 3% (at 95% confidence level) and not be detected. Further testing should be done on a larger number of male samples to obtain a better assessment of potential primer binding site variants with the *SRY* marker region.

Sensitivity Studies

Amplification and correct typing was achieved at all concentrations ranging from 0.125 to 1.0 ng/μL of the male control DNA 007 (Applied Biosystems) and the male DNA casework sample. The amplification was carried out in a singleplex reaction. While the optimal quantity of template DNA was 1.0 ng, conclusive typing of the *SRY* marker was effective over a wide range of input template DNA, making the test applicable for challenged forensic DNA samples. The lower limit of template male control DNA 007 (Applied Biosystems) that yielded *SRY* PCR product with a peak height over 260 relative fluorescent units (RFU) was 50 pg, and as little as 25 pg of template DNA were sufficient to generate a *SRY* peak with a height of 80 RFUs (Fig. 1). The data were consistent over most replicates and support that the assay for the *SRY* marker is very sensitive for pure, nondegradated DNA. However, at 62.5 pg of casework male DNA some drop-out was observed (Fig. 2). Therefore, reliable male gender determination for routine forensic samples could be observed at a DNA quantity as low as 125 pg.

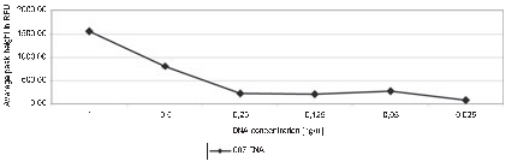


FIG. 1 *Sensitivity studies using consecutive dilutions of genomic of male control DNA 007 (Applied Biosystems) from 1000 pg down to 25 pg, analyzed on the ABI Prism® 310 Genetic Analyzer.*

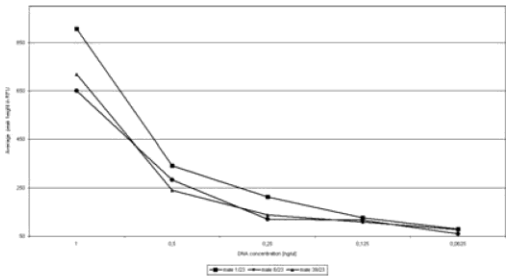


FIG. 2—*Sensitivity studies using consecutive dilutions of three unrelated DNA casework samples from 1000 pg down to 62.5 pg, analyzed on the ABI Prism® 310 Genetic Analyzer. The figure shows the amplification of three different male DNA casework samples.*

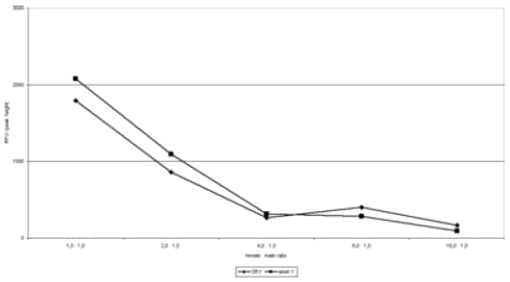


FIG. 4—*Mixture studies—comparison of peak heights of* SRY *gene profile and* AMELY *gene profile in admixed samples with constant female DNA concentration and decreasing concentration of male DNA.*

Gender-Specificity Studies

Two female DNA samples were tested for cross-reactivity with the *SRY* marker assay in a singleplex amplification. The assay failed to produce detectable *SRY* product from female control DNA even at concentrations of 5 and 10 ng/μL. The results of the failure of amplification of *SRY* product from female control DNA 9947A (Promega Corp.) and a female DNA casework sample at the concentration 1.0 ng/μL are shown in Fig. 3. The data support that the new set of *SRY* primers is highly specific for the *SRY* gene on the Y chromosome.

Mixture Studies

The mixture study was carried out in a multiplex reaction. *SRY* primers were coamplified with STR primers from the AmpFlSTR® SGM™ kit. The presence of a high background of female DNA in a sample had no impact on amplification of *SRY* marker down to the tested ratio of 1:16 (93.7 pg male DNA:1.5 ng female DNA). The decrease in peak height from approximately 1800 RFU to approximately 160 RFU of the *SRY* component is concomitant with a reduction of male DNA concentration in the mixed samples (Fig. 4). These results are consistent with those observed for the

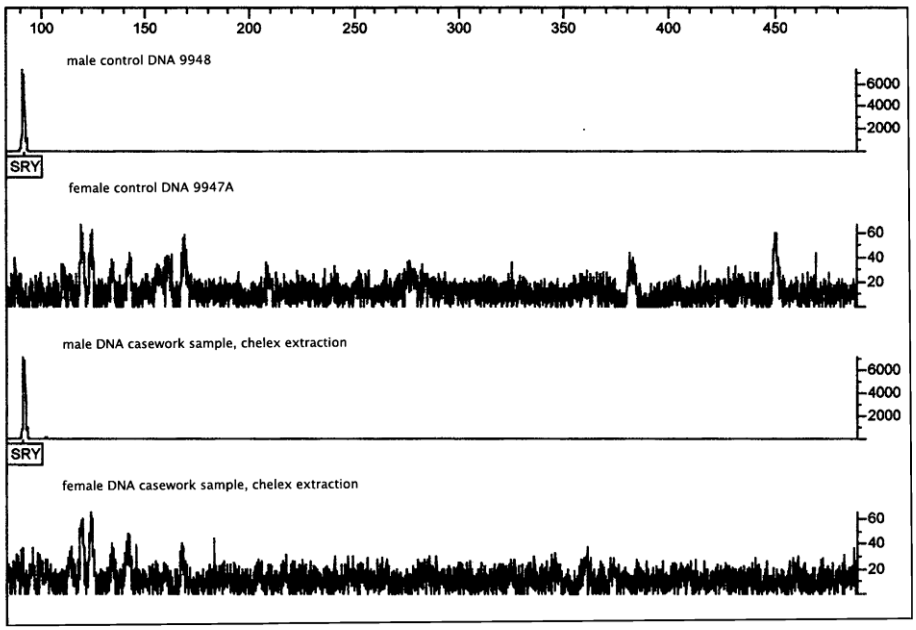


FIG. 3—*Electropherograms of male control DNA 9948—1 ng/μL, of female control DNA 9947A—1 ng/μL (Promega Corp.) and of male—1 ng/μL and female—1 ng/μL DNA sample from our casework.*

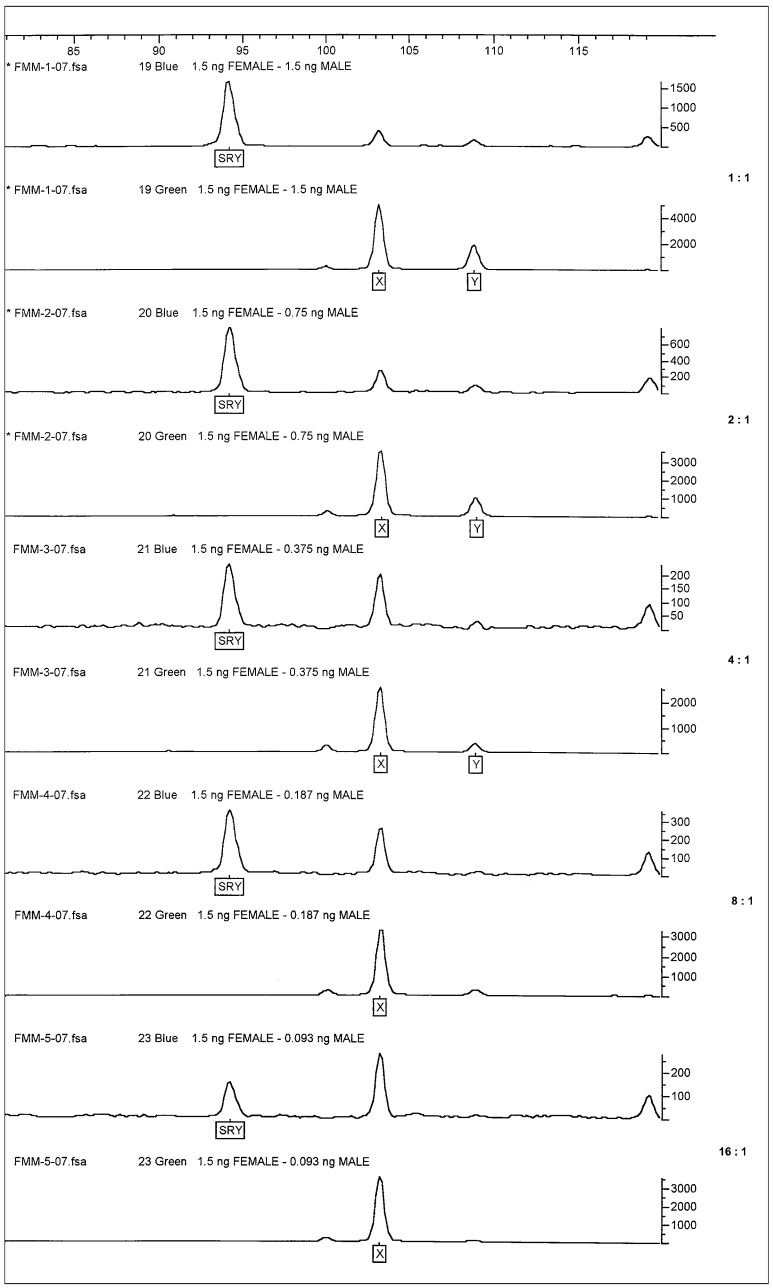


FIG. 5—*Mixture studies—amplification of female and male DNA casework sample. Partial electropherogram that capture amelogenin (AmpFlSTR® SGM™ Plus kit [Applied Biosystems]) and* SRY *genes—profiles are shown from top to bottom, with constant female concentration of DNA (1.5 ng) and decreasing concentration of male DNA: 1:1 (1.5 ng male DNA), 2:1 (750 pg male DNA), 4:1 (375 pg male DNA), 8:1 (187.5 pg male DNA), and 16:1 (93.7 pg male DNA).*

AMELY (data not shown). The robustness and sensitivity of the *SRY* assay in mixed samples is demonstrated at mixture ratios of 1:16 with a total of 93.7 pg male DNA in comparison with the *AMELY* allele assay (Fig. 5). *AMELY* was only detected successfully at a mixture ratio of 4:1 with 375 pg male DNA. Therefore, mixtures with low amounts of male DNA amidst high concentrations of female DNA can be typed with the *SRY* male gender marker assay.

Conclusions

The validation studies reported herein support that the *SRY* male gender marker developed by Drobnič (17) is sensitive, reliable, and can be used in concert with commercially available human STR identification kits to successfully type DNA derived from forensic samples. The *SRY* marker assay as a singleplex or included in multiplex kits can serve as an adjunct to standard gender typing. Future studies should include large population scale *SRY* marker analyses to determine whether drop-out is sufficiently low for forensic gender typing.

Acknowledgments

The authors would like to thank Karmen Čirič for help with the sample collection and her laboratory assistance.

References

1. Sullivan KM, Mannucci A, Kimpton CP, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis X–Y homologous gene amelogenin. BioTechniques 1993;5:636–41.
2. Santos FR, Pandya A, Tyler-Smith C. Reliability of DNA-based sex tests. Nat Genet 1998;18:103.
3. Roffey PE, Eckhoff CI, Kuhl JL. A rare mutation in the amelogenin gene and its potential investigation ramifications. J Forensic Sci 2000;45(5):1016–19.
4. McKeown B, Sickley J, Riordan A. Gender assignment by PCR of the *SRY* gene: an improvement on amelogenin. Progress in For Gen 8. Amsterdam: Elsevier, 2000:433–5.
5. Steinlechner M, Berger B, Niederstätter H, Parson W. Rare failures in the amelogenin sex test. Int J Legal Med 2002;116:117–20.
6. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 2003;423:825–37.
7. Thangaraj K, Reddy AG, Singh L. Is the amelogenin reliable for gender identification in forensic casework and prenatal diagnosis? Int J Legal Med 2002;116:121–3.
8. Lattanzi W, Giacomo MC, Lenato GM, Chimienti G, Voglino G, Resta N, et al. A large interstitial deletion encompassing the amelogenin gene on the short arm of the Y chromosome. Hum Genet 2005;116:395–401.
9. Jobling MA, Lo CC, Turner DJ, Bowden GR, Lee AC, Wue Y, et al. Structural variation on the short arm of the human Y chromosome: recurrent multigene deletions encompassing amelogenin Y. Hum Mol Gen 2007;16(3):307–16.
10. Cadenas AM, Regueiro M, Gayden T, Singh N, Zhivotovsky LA, Underhill PA, et al. Male amelogenin dropouts: phylogenetic context, origins and implications. Forensic Sci Int 2007;166:155–63.
11. Chang YM, Burgoyne LA, Both K. Higher failures of amelogenine sex test in an Indian population group. J Forensic Sci 2003;48(6):1309–13.
12. Chang YM, Perumal R, Keat PY, Yong RYY, Kuehn DLC, Burgoyne L. A distinct Y-STR haplotype for amelogenin negative males characterized by a large Yp11.2 (DYS458-Msy1-Amel-Y) deletion. Forensic Sci Int 2007;166:115–20.
13. Kashyap VK, Sahoo S, Sitalaximi T, Trivedi R. Deletions in the Y-derived amelogenin gene fragment in the Indian population. BMC Medical Genetics 2006;7:2350–57.
14. Michael A, Brauner P. Erroneous gender identification by amelogenin sex test. J Forensic Sci 2004;49(2):258–9.
15. Mitchell RJ, Kreskas M, Baxter E, Buffalino L, van Oorschot RAH. Amelogenin Y negative males: multiple origins. Progress in For Gen 11. Amsterdam: Elsevier, 2006;274–6.
16. Mitchell RJ, Kreskas M, Bexter E, Buffalino L, Van Oorschot RA. An investigation of sequence deletions of amelogenin (*AMELY*), a Y-chromosome locus commonly used for gender determination. Ann Hum Biol 2006;33(2):227–40.
17. Drobnič K. A new primer set in a SRY gene for sex identification. Progress in For Gen 11. Amsterdam: Elsevier, 2006;268–70.
18. Scientific Working Group on DNA Analysis Methods (SWGDAM). Revised validation guidelines. Forensic Sci Commun 2004;6(3), http://www.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm (accessed 5/03/08).
19. Walsh PS, Metzger DA, Higuchi R. Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 1991;10:506–13.
20. Lazaruk K, Walsh PS, Oaks F, Gilbert D, Rosenblum BB, Menchen S, et al. Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis 1998;19(1):86–93.

Additional information and reprint requests:
Vanja Kastelic
Forensic Science Centre
Ministry of the Interior
Vodovodna 95
Ljubljana, Slovenia
E-mail: vanja.kastelic@policija.si

2.2 HKRATNO POMNOŽEVANJE MINISEKVENČNIH ZAČETNIH OLIGONUKLEOTIDOV: POVEZAVA MED PETIMI SNP-JI TER OBARVANOSTJO OČI IN LAS V SLOVENSKI POPULACIJI TER PRIMERJAVA BAYESOVEGA MODELA IN MODELA LOGISTIČNE REGRESIJE

A single-nucleotide polymorphism (SNP) multiplex system: the association of five SNPs with human eye and hair colour in the Slovenian population and comparison using a Bayesian network and logistic regresion model

V. Kastelic, K. Drobnič
Croatian Medical Journal, 2012, 53, 5: 401–408
doi.org/10.3325/cmj.2012.53.401

Raziskava vključuje gene, za katere so dosedanje študije potrdile njihov večji vpliv na pigmentacijske značilnosti ljudi. Znotraj šestih genov smo izbrali in analizirali dvanajst SNP-jev. Začetne oligonukleotide smo izbrali oziroma skonstruirali, tako da smo jih lahko nato združili v hkratni reakciji PCR z namenom določevanja pigmentacijske obarvanosti oči in las za slovensko populacijo. V raziskavi je sodelovalo 105 prostovoljcev. Variabilnost polimorfizmov posameznih nukleotidov smo analizirali s pomočjo uvedbe nove metode SNaPshot^{TM} in produkte pomnoževanja analizirali s kapilarno elektroforezo. Na podlagi pridobljenih genotipskih rezultatov smo s pomočjo dveh statističnih napovednih modelov (naivnega Bayesovega modela in modela logistične regresije) preverjali uspešnosti napovedi za posamezno pigmentacijsko obarvanost oči in las v slovenski populaciji. Pri primerjavi obeh statističnih napovednih modelov lahko zaključimo, da sta glede na rezultate oba modela primerna, saj smo za njiju določili primerljive rezultate glede uspešnosti napovedi, vendar smo višje vrednosti AUC, občutljivosti in specifičnosti za obarvanost oči in las dosegli z uporabo modela logistične regresije. To pomeni, da je model logistične regresije delno uspešnejši pri napovedovanju obarvanosti oči oziroma las za izbrani vzorec slovenske populacije. Pri razdelitvi barve oči v tri kategorije smo določili vrednosti AUC 1.0 za modrooke, vrednost AUC za rjavo barvo oči je bila 0.832 in za vse vmesne barve oči 0.747. Pri razdelitvi barve las v tri kategorije smo najbolj uspešno napovedali obarvanost pri posameznikih z blond lasmi (AUC=0.913) in s temno rjavimi/črnimi lasmi (AUC=0.832) ter malo manj uspešno pri posameznikih z blond/s svetlo rjavimi lasmi (AUC=0.723). V naslednjem koraku smo obarvanost oči in las prerazporedili v dve kategoriji (svetlejšo in temnejšo pigmentacijsko obarvanost), a s tem nismo bistveno vplivali na uspešnost napovedi modelov. Vrednost AUC za obarvanost oči je bila tako 0.99 in za obarvanost las 0.93 (enaka za temnejšo in svetlejšo obarvanost), ob podanih spremljajočih podatkih za občutljivost in specifičnost. Predstavljeni komplet za hkratno pomnoževanje se je v naši raziskavi izkazal za dovolj robustnega in občutljivega za pravilno napoved homo-/heterozigotnosti vseh dvanajstih SNP-jev v naboru za celoten vzorec slovenske populacije.

Croat Med J. 2012;53:401-8
doi: 10.3325/cmj.2012.53.401

A single-nucleotide polymorphism (SNP) multiplex system: the association of five SNPs with human eye and hair color in the Slovenian population and comparison using a Bayesian network and logistic regression model

Vanja Kastelic[1], Katja Drobnič[1,2]

[1]National Forensic Laboratory, General Police Directorate, Police, Ministry of the Interior, Ljubljana, Slovenia

[2]Faculty of Criminal Justice and Security, University of Maribor, Ljubljana, Slovenia

Aim To analyze two phenotype characteristics – eye and hair color – using single-nucleotide polymorphisms (SNPs) and evaluate their prediction accuracy in Slovenian population.

Methods Twelve SNPs (*OCA2* – rs1667394, rs7170989, rs1800407, rs7495174; *HERC2* – rs1129038, rs12913832; *MC1R* – rs1805005, rs1805008; *TYR* – rs1393350; *SLC45A2* – rs16891982, rs26722; *SLC24A5* – rs1426654) were used for the development of a single multiplex assay. The single multiplex assay was based on SNaPshot chemistry and capillary electrophoresis. In order to evaluate the accuracy of the prediction of eye and hair color, we used the logistic regression model and the Bayesian network model, and compared the parameters of both.

Results The new single multiplex assay displayed high levels of genotyping sensitivity with complete profiles generated from as little as 62 pg of DNA. Based on a prior evaluation of all SNPs in a single multiplex, we focused on the five most statistically significant in our population in order to investigate the predictive value. The two prediction models performed reliably without prior ancestry information, and revealed very good accuracy for both eye and hair color. Both models determined the highest predictive value for rs12913832 ($P < 0.0001$), while the other four SNPs (rs1393350, rs1800407, rs1805008, and rs7495174) showed additional association for color prediction.

Conclusion We developed a sensitive and reliable single multiplex genotyping assay. More samples from different populations should be analyzed before this assay could be used as one of the supplemental tools in tracing unknown individuals in more complicated crime investigations.

Received: May 18, 2012

Accepted: October 1, 2012

Correspondence to:
Vanja Kastelic
National Forensic Laboratory
General Police Directorate, Police
Ministry of the Interior
Vodovodna 95
Ljubljana 1000, Slovenia
vanja.kastelic@policija.si

Height, face structure, pigmentation of the eye, hair, and skin, the presence of freckles, and male baldness make up human externally visible characteristics (EVC). To be able to predict eye and hair color based solely on biological material left behind at a crime scene or obtained from dismembered missing persons, or even of disaster victims, is one of the major expectations from the routine forensic work in the near future (1). However, genetic understanding of human appearance is still in its infancy, mainly due to the fact that all EVCs are polygenic traits. This means that yields from a large number of different genes and the expression of these genes are further influenced by mutual interactions and environmental interactions (2). Above all, molecular mechanisms and functional protein assays must also be considered in order to really understand how allelic variation in pigmentation genes could result in such a diversity of phenotypes in different human populations (3). The human eye (iris) and hair color are one of the most highly polymorphic phenotypes in people of European origin. The non-brown iris colors and red hair are generally features of European origin resulting from positive selection in early European history. There are several hypotheses for positive selection that mainly occurred in the Baltic region and Northern Europe. These are most likely: UV exposure causing skin cancer, vitamin D deficiency, and even sexual selection (4,5). Most EVCs are complex traits with many genes and single nucleotide polymorphism (SNP) variations, so the right combination of SNPs is crucial for the correct prediction of eye and hair color. Several genome-wide association studies (GWAS) for pigmentation have revealed that SNPs within the *HERC2, OCA2, MC1R, SLC24A5, SLC45A2, TYR,* and *ASIP* (4,6-16) genes were most strongly associated with eye and hair color in European populations. The latest data have shown that the main iris color variation is associated with a highly evolutionarily preserved region in the *HERC2* gene or within the short sequence between the *HERC2* and *OCA2* genes. It is assumed that these regions represent a regulatory region controlling the constitutive expression of *OCA2* (4,11,12). As for iris color, it has also been explained that red hair color is mainly associated with polymorphisms in the *MC1R* gene (13,17). On the other hand, the variations of genes such as *SLC24A5, SLC45A, HERC2,* and *ASIP* seem to be responsible for influencing the shades of hair color from blond to black (18,19).

In order to correctly predict human eye and hair color from genetic data for the Slovenian population, we compared two alternative prediction models that are nowadays used most often in this field of forensics – the Bayesian network model and the logistic regression model. These models were developed and compared on the basis of the informative SNPs selected from our single multiplex assay.

MATERIAL AND METHODS

Sample collection

The research included 105 unrelated Slovenian volunteers, 70 male and 35 female donors. Buccal swabs of the adult volunteers were collected in 2008 using a SAFE® Box kit (ForensiX, Prionics AG, Zurich, Switzerland).

The eye and hair color of each volunteer were defined according to descriptions provided by the volunteers and through our observer grading. In the small number of elderly volunteers and volunteers with dyed hair, we used self-assessment for establishing natural hair color phenotypes. We then categorized each volunteer's eye and hair color into three groups. Eye color was defined as blue (44.7%), intermediate (25.7%), and brown (29.6%). For classification into blue and brown eye color, the eye must have been clearly composed of one color, regardless of its intensity. However, the classification into intermediate eye color included all other individuals with green or hazel eye colors, or those with two or more pigments within the iris (peripapillary rings or different colored spots). The hair color was defined as blond (5.7%), dark blond/light brown (41.0%), and dark brown/black (52.4%).

Red hair color was not included because our research group included only one red-haired person due to the small number of red-haired persons in the Slovenian population. A demonstration of prediction accuracy for red hair color based on only one person's data are not statistically acceptable, so we excluded this person and limited the study to the three hair color groups. The study was approved by the National Medical Ethics Committee in Slovenia. Before donating a sample, all the volunteers signed written consents for the use of their DNA solely for scientific research.

DNA extraction and quantification

DNA was extracted from all the samples using Chelex extraction (20). Isolated DNA was quantified using the real-time PCR method – ABI Prism 7500 Sequence Detection System (Applied Biosystems Inc., Foster City, CA, USA). The extraction yield ranged between 0.7 and 20 ng/μL per per-

son. This broad range can be explained by the unequal shedding of buccal cells among people.

SNP selection and multiplex design

Twelve of the SNPs (4,6-8,21) (Supplementary Table 1) that were associated with human eye and hair color in previous studies were analyzed in one multiplex assay. The PCR primers reported (Supplementary Table 2) were combined in one multiplex reaction and, as a result, the primers' melting temperatures were verified and the potential interactions were preliminarily checked using the AutoDimer software program (22). The length of PCR fragments was limited to between 104 and 238 bp in order to meet future forensic investigations due to degraded DNA. For the single multiplex PCR, a genomic DNA extract from each individual was amplified ranging from 0.03125 and 1.0 ng in 25-µL reactions containing 1× AmpFlSTR PCR reaction mix (Applied Biosystems), 8 mM $MgCl_2$, 0.1-1.7 µM of each primer, and 1-5U AmpliTaq Gold DNA polymerase (Applied Biosystems). The following cycle program was used: denaturation at 95°C for 10 minutes followed by 35 cycles of 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 30 seconds, followed by final incubation for 70 minutes at 72°C. PCRs were performed in a Perkin-Elmer 9600 thermal cycler (Applied Biosystems).

After PCR amplification, the excess PCR primers and ddNTPs were removed by addition of freshly prepared mix of shrimp alkaline phosphatase (1U/µL) (ABGene, Epsom, UK) and Exonuclease I (20U/µL) (ABGene) to 5 µL PCR products and incubation at 37°C for 1 hour followed by incubation at 75°C for 15 minutes.

Ten single base extension (SBE) primers were collected from previous reports and two were self-designed (Supplementary Table 1). To ensure varying lengths of SBE primer in order to adjust their mobility in capillary electrophoresis, GACT-tails were added to the 5′ ends of each SBE primer. A multiplex SBE reaction was performed, using an ABI Prism SNaPshot kit (Applied Biosystems) in a total reaction volume of 8 µL, containing 1 µL purified PCR products, 4 µL SnaPsho reaction mix (Applied Biosystems), and 0.01-1.7 µM of each of SBE primers and Milli-Q water. Thermal cycling for SBE was performed in a thermal cycler (Applied Biosystems), using a program for 30 cycles at 96°C for 10 seconds, at 50°C for 5 seconds, and at 60°C for 30 seconds. Excess fluorescently labeled ddNTPs were removed by adding shrimp alkaline phosphatase (1U/µL) (ABGene) and incubating them at 37°C for 45 minutes, followed by 15 minutes incubation at 75°C.

SBE products were then analyzed on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems) following standard protocol (23), with a 36-cm capillary array, POP-4 polymer (Applied Biosystems), and a 5-second injection at 1.5 kV. Allele calling was performed using GeneMapperID ver. 3.2 software (Applied Biosystems) and a bin set, according to our SBE product size, was designed for our multiplex to allow automation of genotyping. To test the overall sensitivity of the multiplex assay, all the samples were analyzed in duplicate; if the results were inconsistent, additional amplifications were made.

Analysis of approved allele calls

For each SNP locus, we determined the average peak height ratio for heterozygotes and homozygotes, and the standard deviation of the peak height was also provided. From some of the randomly selected electropherograms, the peak height of the highest background peak in each dye window was collected and an arbitrary maximum background level determined (blue and green dye – 110 RFUs, yellow dye – 80 RFUs, and red dye – 70 RFUs (relative fluorescent units). The background levels were used to calculate an arbitrary signal/noise ratio for homozygote allele calls. For heterozygote allele calls, the peak height ratio was calculated by dividing the peak height of the lower molecular weight allele by the peak height of the higher molecular weight allele for all SNPs. The right allele calls were determined for all 105 genotypes of twelve SNPs.

Reproducibility and sensitivity

Validation of the single multiplex was conducted on samples of two individuals. Multiplex PCR performance was assessed by analyses of dilution series of genomic DNA (0.03; 0.06; 0.12; 0.25; 0.50 1.00, and 2.00 ng), amplifying with 1-5 U AmpliTaq Gold DNA polymerase (Applied Biosystems) in the reaction mix. To ensure the consistency of genotyping, two of the samples with all dilution series were analyzed in duplicates and additionally the homozygote and heterozygote peak height ranges were noted for each locus.

Statistical analysis

In order to evaluate allele frequencies, locus-by-locus molecular variance (AMOVA), the Hardy-Weinberg equilibrium (HWE), and linkage disequilibrium (LD), we used Arlequin, version 3.1 (24). The *P* values were corrected using the Bonferroni correction for multiple testing

($P > 0.0045$) for the final determination of SNPs in linkage disequilibrium.

The statistical analysis for the prediction of phenotypes from genotypes was based on defining eye color as blue, intermediate, and brown, and hair color as blond, dark blond/light brown, and dark brown/black in the first step. Additionally we simplified the classification for both eye and hair color with only two stages: light and dark. For statistically relevant SNPs the probability values for phenotype prediction were first calculated based on multinomial/binary logistic regression (17,25) using SPSS 19.0 for Windows (SPSS Inc., Chicago, IL, USA) and second based on likelihood ratios (LRs) liable to the Bayesian approach using the excel macro of the Branicki group (26). In order to evaluate the predictive accuracy of both models, we randomly split our samples into a model-building set consisting of 80 individuals (77%) and a model verification set comprising the remaining 24 (33%) individuals. When working with the Bayesian network model, the *a priori* conditional probabilities of colors were entered as having the same value (0.33 – three color description, 0.5 – two color description), because we rarely have prior knowledge on the ancestry of the identified subject. For worldwide distribution, a threshold of 0.7 predicted eye or hair color probability was used for categorization. When the probability values were under this fixed threshold, the eye or hair color was predicted as undefined. This cut-off was based on the receiver operating characteristics (ROC) curve, or area under the receiver operating characteristic curve (AUC), derived from previous studies. AUC is the integral of ROC curves, which ranges from 0.5, representing a total lack of prediction, to 1.0, representing perfect prediction (25).

RESULTS

Multiplex design and protocol

The single multiplex assay was designed based on twelve PCR fragments of lengths less than 240 bp, and was therefore useful for casework DNA samples. Ten previously reported SBE primers and two newly designed primers were evenly separated by 5 bp in the region of 30-65 bp in length for precise marker differentiation. Allele peaks were only called if they were above 50 RFU in their respective size range within a custom-designed bin set for our single multiplex, using GeneMapperID, version 3.2 software.

This single multiplex assay worked optimally at 1 ng of template DNA using 1U of the AmpliTaq Gold DNA

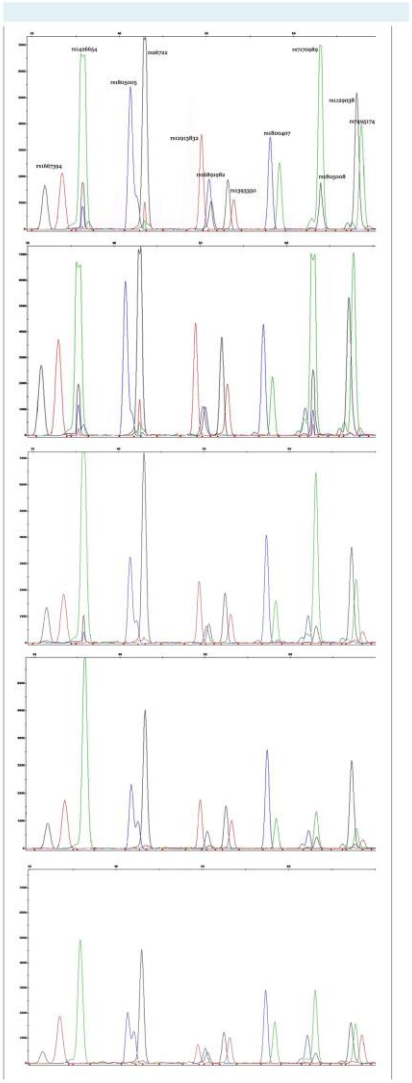


FIGURE 1. Sensitivity testing of our multiplex assay. Multiplex single base extension (SBE) products for 1.0; 0.5; 0.25; 0.125, and 0.0652 ng of DNA input for one random sample.

polymerase for PCR amplification. The sensitivity of the multiplex assay rised remarkably by increasing the concentration of the polymerase up to 5U. Using these amounts of polymerase for PCR amplification, the drop-out partially appeared at only 62 pg of DNA input or even less (Figure 1). For our assay, the homozygote and heterozygote peak heights for all SNPs were balanced inter-loci. Still, we could not achieve an inter-loci balance for all SNPs, which did not in any way affect the genotyping accuracy.

The homozygote and heterozygote average peak heights of each locus from all 105 samples of 1 ng of DNA input were calculated and the relations were from 1500 to 5500 RFU and from 500 to 3000 RFU, respectively. The heterozygote ratio for SNPs rs26722 and of rs16891982 could not be determined precisely and, as a result of their rarity in the European population, sufficient data could not be collected. In order to understand a more balanced profile for SNP rs16891982, we increased this final concentration from 0.25 μM up to 0.65 μM of the same SBE primer, and the results in the changed multiplex assay were the same as had been shown by the monoplex genotyping assay for the aforementioned SNP for all 105 samples.

Population analysis

In order to determine how well the Slovenian sample population reflected its ancestry, the allele frequencies for the twelve SNPs selected were compared with the data from the International HapMap consortium (27). The International HapMap project is a multi-country effort to identify and catalog genetic variants. For the evaluation of the Slovenian population we compared it with HapMap data of Utah residents with European ancestry (CEU). The allele frequency of our population was very similar to that of the HapMap data. A major difference was noticed for SNP rs16891982, assuming that this was for the same reason stated above – its rarity of heterozygote SNP profiles, and for this reason sufficient data could not be collected. The SNP rs1426654 was the only one that was monomorphic (homozygote for the allele A/A) and showing the same data as presented on the HapMap project (Supplementary Table 3). This indicated that rs1426654, derived threonine allele, was fixed in the European population.

We analyzed the fact that all 105 individuals were homozygotes for the derived allele A/A. Based on that, this marker was not included further in calculating the prediction accuracy for pigmentation color. From the 105 Slovenian samples we eliminated only one person, this being the only person with red hair in our population (0.9% of the Slovenian population). This low frequency of red hair color was expected for the general Slovenian population and additionally confirmed that red haired individuals were more prevalent in the Baltic region and Northern Europe populations (28).

The linkage disequilibrium (LD) between all the SNPs was calculated. As had been previously indicated (12), we established that SNPs rs1129038 and rs12913832 were in perfect linkage disequilibrium and could be considered a single haplotype owing to their close position on the chromosome 15. SNP allele haplotypes were significantly associated with eye and hair color, but due to linkage disequilibrium we decided to use only rs12913832 for further prediction information. We could also conclude that two SNPs, rs1667394 and rs7170989, were in linkage disequilibrium with SNP rs1800407 for the same reason as previously stated, and we also included only SNP rs1800407 in the statistical models. Finally, after using the Bonferroni correction for multiple testing, we did not detect any significant departures from the Hardy-Weinberg equilibrium (*P* values ranged from 0.1889 to 1.0000) for all the 11 analyzed SNPs (SNP rs1426654 was monomorphic).

Eye and hair color prediction models

In order to evaluate the predictive accuracy for eye and hair color, we established the SNP position that could be statistically significant in our study to determine the accurate prediction. In the prediction model, we did not include the markers that were monomorphic (rs1426654) or in linkage disequilibrium with other SNPs (rs1129038, rs1667394, and rs7170989). With the elimination of these four SNPs, we first used the eight remaining SNPs (rs1800407, rs7495174, rs12913832, rs1805008, rs1805005, rs1393350, rs26722, rs16891982) to simultaneously test for the effect on a dichotomous dependent variable with binary logistic regression (where eye color was classified as blue vs non-blue, intermediate vs non-intermediate, and brown vs non-brown; and hair color was classified as blond vs non-blond, dark blond/light brown vs non-dark blond/light brown, and dark brown/black vs non-dark brown/black). Based on the results we selected the five SNPs (rs1800407, rs7495174, rs12913832, rs1805008, rs1393350) that were statistically significant ($P < 0.05$) for the group for which the analysis was made. The multinomial/binary logistic regression model and the Bayesian network model (25,26) were developed and tested for eye and hair color accuracy prediction. The first variant for both models assumed

prediction of pigment color being divided into three groups: blue, intermediate, brown (eye color); and blond, dark blond/light brown, and dark brown/black (hair color). The second variants for both models were reduced to only two states – light and dark for eye and hair color.

For blue eye color using both statistical models for 24 samples, the prediction was completely accurate, when using three categorized eye colors. When using the logistic model for evaluation, high values were also obtained for brown eye color, with AUC of 0.832, but were least accurate for intermediate eye color, with AUC of 0.747. For blue eye color categorization, we got a 100% correct call rate (sensitivity), which means that all blue-eyed individuals were predicted correctly (24/24). The sensitivity was also high for intermediate (100%) eye color and relatively high for brown eye color (71%) (Table 1). The highest specificity was obtained for blue eye color (100%), which means that among all non-blue eye colored individuals, all of them were recognized correctly. High specificity was also obtained for brown (90%) eye color and much lower for intermediate (63%) eye color.

The AUC for eye color defined as light vs dark was calculated for both colors to equal 0.985. The sensitivity in this categorization of eye color was also high and even reached 100% for dark eye color and 88% for light eye color. All the described results were mostly summarized only for the multinomial logistic regression model. However, most values based on the Bayesian model variant were slightly lower, but still in the same range (Table 1).

The highest value for AUC, 0.913, was obtained for blond hair color and even for dark brown/black hair color, 0.832, when using the logistic model for hair color categorized into three groups. As indicated for eye color prediction, for hair color the AUC values were also slightly higher when using the logistic model in comparison to the Bayesian network model. Sensitivity (88%) and also specificity (95%) were highest for dark brown/black hair color and they could not be extended for the other two hair color groups, presumably due to their small size sample in the model verification set of 24 Slovenian volunteers.

The AUC for hair color, defined as light vs dark in the logistic model, was calculated for both colors to equal 0.929. The sensitivity was only highest for dark hair color (100%) when using the logistic model and was much lower for light hair color for the same reason (Table 2).

Predicting eye and hair color type separately using a multinomial logistic model yielded slightly higher accuracy compared to the Bayesian model. In any event, for both models the lowest AUC values were observed for intermediate eye color and for dark blond/light brown hair color.

TABLE 1. Parameters describing predictive accuracy of two developed eye color prediction models, divided into two variants for 24 Slovenian volunteers*

	Multinomial logistic regression/Bayesian network			Binary logistic regression/Bayesian network	
	blue	intermediate	brown	light	dark
AUC	1.000/1.000	0.747/0.632	0.832/0.685	0.985/0.924	0.985/0.924
Sensitivity (%)	100.0/100.0	100.0/x	71.0/14.0	88.0/86.0	100.0/94.0
Specificity (%)	100.0/100.0	63.0/100.0	90.0/100.0	88.0/100.0	x/100.0
PPV (%)	100.0/100.0	25.0/x	71.0/100.0	100.0/100.0	100.0/100.0
NPV (%)	100.0/100.0	100.0/50.0	90.0/54.0	88.0/86.0	x/94.0

*Abbreviations: AUC – area under the receiver operating characteristic (ROC) curves; PPV – positive predictive value; NPV – negative predictive value; x – zero denominator.

TABLE 2. Parameters describing predictive accuracy of two developed hair color prediction models, divided into two variants, for 24 Slovenian volunteers *

	Multinomial logistic regression/Bayesian network			Binary logistic regression/Bayesian network	
	blond	dark blond/light brown	dark brown/black	light	dark
AUC	0.913/0.714	0.723/0.445	0.832/0.543	0.929/0.878	0.929/0.878
Sensitivity (%)	x/x	50.0/x	88.0/79.0	14.0/71.0	100.0/88.0
Specificity (%)	x/100.0	80.0/88.0	95.0/100.0	100.0/100.0	100.0/100.0
PPV (%)	x/x	33.0/x	88.0/100.0	100.0/100.0	100.0/100.0
NPV (%)	x/100.0	89.0/50.0	95.0/89.0	14.0/67.0	100.0/88.0

*Abbreviations: AUC – area under the receiver operating characteristic (ROC) curves; PPV – positive predictive value; NPV – negative predictive value; x – zero denominator.

There is a possibility that the problem lies in imprecise color classification, which reflects uncertainties in distinguishing within intermediate eye color (green-eyed individuals and those with more pigments within the iris) and within different shades of hair color, which may be influenced by age-dependent hair color change during the entire lifetime of each individual (distinguishing between the dark blond and blond on one hand, and between light brown and brown on the other).

DISCUSSION

Eye and hair color represent human traits that have a potential to be predicted from genetic material with high reliability and used in expansive forensic investigations. In the present study, based on a new single multiplex of twelve SNPs, we confirmed the effect of five statistically most significant SNPs for eye and hair color in both statistical models. Specifically rs12913832 in *HERC2* showed the strongest association with both eye and hair color.

To illustrate the predictive performance of our single multiplex we focused on the five most statistically significant SNPs (rs1800407, rs7495174, rs12913832, rs1805008, and rs1393350). The other seven SNPs were not included, as being either monomorphic for our population or in linkage disequilibrium with other SNPs, or as just not being sufficiently statistically significant for prediction accuracy for Slovenian samples. For our study population, the only SNP being monomorphic was SNP rs1426654, as all volunteers were homozygote for the derived allele A/A. Our data were therefore inconclusive regarding the hypothesis of Velenzuela et al (19) that rs1426654 contributed to hair color variation. The single multiplex assay also included three SNPs (rs1129038 [4], rs7170989 [6], and rs1667394 [9]), which were previously recommended for eye and hair color prediction. For all these three SNPs, we concluded that they were in strong linkage disequilibrium with the other five most statistically informative SNPs, and it was therefore unlikely that their inclusion would have had any effect on the prediction power.

For genotyping accuracy of 105 Slovenian samples, this multiplex assay is for now designed to work optimally at 1 ng of template DNA, but its sensitivity has risen up to 62 pg by increasing the amount of AmpliTaq Gold DNA polymerase in reaction (5U). Our results on sensitivity and reliability of the multiplex could even be improved by increasing the injection and time voltage in capillary electrophoresis and also to use the polymer POP-7 instead of POP-4, due to its better mobility in the capillary of the ABI Prism 3130 Genetic Analyzer.

The single multiplex assay presented was a robust and sensitive DNA tool regarding amplification and regular determination of the homozygosity/heterozygosity for each SNP included. On the basis of these facts, the single multiplex assay could be suitable for the use in forensic casework due to its more efficient use of template DNA. Furthermore, even markers not significantly associated with a trait in a temporary study population can still independently contribute to the trait prediction in the enlarged Slovenian or even European population. After all, the growing number of known sequence variants that underlie the differences in human pigmentation may provide new SNPs with specific markers that we could include in our assay for a more accurate prediction of eye and hair color.

Acknowledgment We thank to all the volunteers who provided DNA samples and photographs of their eyes for this study. We would also like to thank Wojciech Branicki and Ewelina Pośpiech for their support in the statistical analysis of our results.

Funding None.

Ethical approval received from the National Medical Ethics Committee in Slovenia.

Declaration of authorship VKI participated in designing the methods; analyzing, interpreting, and statistically evaluating the results and preparing the manuscript. KDI produced the idea, designed the study and contributed to the interpretation of the results, and revision of the manuscript for important intellectual content.

Competing interests All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

References

1 Jobling MA, Gill P. Encoded evidence: DNA in forensic analysis. Nat Rev Genet. 2004;5:739-51. Medline:15510165 doi:10.1038/nrg1455

2 Mertens G. Forensic DNA typing: quo vadis? The Open Forensic Science Journal. 2009;2:21-8. doi:10.2174/1874402800902010021

3 Sturm RA. Molecular genetics of human pigmentation diversity. Hum Mol Genet. 2009;18:R9-17. Medline:19297406 doi:10.1093/hmg/ddp003

4 Eiberg H, Troelsen J, Nielsen M, Mukkelsen A, Mengel-From J, Kjaer KW, et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Hum Genet. 2008;123:177-87. Medline:18172690 doi:10.1007/s00439-007-0460-x

5 Frost P. European hair and eye color: a case of frequency-dependent sexual selection? Evol Hum Behav. 2006;27:85-103.

doi:10.1016/j.evolhumbehav.2005.07.002

6 Duffy DL, Montgomery GW, Chen W, Zhao ZZ, Le L, James MR, et al. A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. Am J Hum Genet. 2007;80:241-52. Medline:17236130 doi:10.1086/510885

7 Soejima M, Koda Y. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. Int J Legal Med. 2007;121:36-9. Medline:16847698 doi:10.1007/s00414-006-0112-z

8 Mengel-From J, Børsting C, Sanchez JJ, Eiberg H, Morling N. Determination of cis/trans phase of variations in the MC1R gene with allele-specific PCR and single base extension. Electrophoresis. 2008;29:4780-7. Medline:19016241 doi:10.1002/elps.200800107

9 Sulem P, Gudbjartsson DF, Stacey SN. Genetic determination of hair, eye and skin pigmentation in Europeans. Nat Genet. 2007;39:1443-52. Medline:17952075 doi:10.1038/ng.2007.13

10 Oetting WS, Garrett SS, Brott M, King RA. P gene mutations associated with oculocutaneous albinism type II (OCA2). Hum Mutat. 2005;25:323-9. Medline:15712365 doi:10.1002/humu.9318

11 Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K, et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. Am J Hum Genet. 2008;82:411-23. Medline:18252221 doi:10.1016/j.ajhg.2007.10.003

12 Sturm RA, Duffy DL, Zhao ZZ, Leite FPN, Stark MS, Hayward NK, et al. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. Am J Hum Genet. 2008;82:424-31. Medline:18252222 doi:10.1016/j.ajhg.2007.11.005

13 Valverde P, Healy E, Jackson I, Rees JL, Thody AJ. Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. Nat Genet. 1995;11:328-30. Medline:7581459 doi:10.1038/ng1195-328

14 Sturm RA, Teasdale RD, Box NF. Human pigmentation genes: identification, structure and consequences of polymorphic variation. Gene. 2001;277:49-62. Medline:11602344 doi:10.1016/S0378-1119(01)00694-1

15 Box NF, Duffy DL, Chen W, Stark M, Martin NG, Sturm RA, et al. MC1R genotype modifies risk of melanoma in families segregating CDKN2A mutations. Am J Hum Genet. 2001;69:765-73. Medline:11500805 doi:10.1086/323412

16 Harding RM, Healy E, Ray AJ, Ellis NS, Flanagan N, Todd C, et al. Evidence for variable selective pressures at MC1R. Am J Hum Genet. 2000;66:1351-61. Medline:10733465 doi:10.1086/302863

17 Branicki W, Liu F, van Duijn K, Draus-Barini J, Pospiech E, Walsh S, et al. Model-based prediction of human hair color using DNA variants. Hum Genet. 2011;129:443-54. Medline:21197618 doi:10.1007/s00439-010-0939-8

18 Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, et al. Two newly identified genetic determinants of pigmentation in Europeans. Nat Genet. 2008;40:835-7. Medline:18488028 doi:10.1038/ng.160

19 Valenzuela RK, Henderson MS, Walsh MH, Garrison NA, Kelch JT, Cohen-Barak O, et al. Predicting phenotype from genotype: normal pigmentation. J Forensic Sci. 2010;55:315-22. Medline:20158590 doi:10.1111/j.1556-4029.2009.01317.x

20 Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 1991;10:506-13. Medline:1867860

21 Walsh S, Lindenbergh A, Zuniga SB, Sijen T, de Knijff P, Kayser M, et al. Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. Forensic Sci Int Genet. 2011;5:464-71. Medline:20947461 doi:10.1016/j.fsigen.2010.09.008

22 Vallone PM, Butler JM. AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques. 2004;37:226-31. Medline:15335214

23 Applied Biosystems. ABI PRISM®SNaPshot™ Multiplex Kit. Protocol. Foster City (CA): Applied Biosystems; 2010.

24 Excoffier L, Laval G, Schneider S. Arlequin (version 3.0) an integrated software package for population genetics data analysis. Evol Bioinform Online. 2005;1:47-50. Medline:PMC2658868

25 Liu F, van Duijn K, Vingetling JR, Hofman A, Uitterlinden AG, Janssens ACJW. Eye color and the prediction of complex phenotype from genotypes. Curr Biol. 2009;19:R192-3. Medline:19278628 doi:10.1016/j.cub.2009.01.027

26 Pospiech E, Draus-Barini J, Kupiec T, Wojas-Pelc A, Branicki W. Prediction of eye color from genetic data using Bayesian approach. J Forensic Sci. 2012;57:880-6. Medline:22372960 doi:10.1111/j.1556-4029.2012.02077.x

27 International HapMap Consortium. The international HapMap project. Nature. 2003;426:789-96. Medline:14685227 doi:10.1038/nature02168

28 Branicki W, Brudnik U, Kupiec T, Wolanska-Nowak P, Wojas-Pelc A. Determination of phenotype associated SNPs in the MC1R gene. J Forensic Sci. 2007;52:349-54. Medline:17316231 doi:10.1111/j.1556-4029.2006.00361.x

2.3 DOLOČANJE OBARVANOSTI OČI V SLOVENSKI POPULACIJI Z UPORABO SNP JEV KOMPLETA IRISPLEX

Prediction of eye colour in the Slovenian population using the IrisPlex SNPs

V. Kastelic, W. Branicki, E. Pośpiech, J. Draus-Barini, K. Drobnič
Croatian Medical Journal, 2013, 54, 4: 381-386
doi.org/10.3325/cmj.2013.54.381

Zaradi številčno manjšega vzorca slovenske populacije, ki smo ga predstavili v prejšnji raziskavi, smo za statistično bolj oprijemljivo in zanesljivejšo napoved obarvanosti oči naknadno uporabili statistični napovedni model edinega validiranega kompleta IrisPlex (Walsh in sod., 2011a). Komplet temelji na šestih najbolj informativnih SNP-jih in je z znanimi genotipskimi in fenotipskimi podatki 3804 nizozemskih prostovoljcev osnovan z metodo logistične regresije. Vseh 105 slovenskih prostovoljcev smo v tem primeru uporabili za testno skupino. Uspešnost napovedi specifične obarvanosti oči (razdeljena v tri kategorije) slovenske populacije smo nato primerjali s podatki, pridobljenimi v dveh večjih študijah. Poleg že analiziranih štirih SNP-jev znotraj našega kompleta (rs12913832, rs1800407, rs16891982 in rs1393350) smo v sodelovanju z Inštitutom forenzičnih preiskav v Krakovu analizirali še dodatna dva SNP-ja (rs12896399 in rs12203592) in tako zadostili modelu, ki je osnovan na skupnem genotipu šestih SNP-jev. IrisPlex komplet šestih SNP-jev se še ne uporablja v rutinskih forenzičnih genetskih preiskavah, saj je njegova validacija opravljena za večjo evropsko populacijo, pri kateri so alelne frekvence SNP-jev specifične in povsem drugačne od frekvenc ostalih populacij. Uspešnost napovedi oziroma vrednost AUC je bila v tem primeru za obarvanost oči pri slovenski populaciji 0.966 za modre oči, 0.913 za rjave oči in 0.796 za vmesno obarvanost oči. Največjo občutljivost smo določili za modro obarvanost oči, ki je bila kar 93,6 %, kar pomeni, da je bil takšen odstotek modrookih posameznikov pravilno napovedan. Uspešnost napovedi je bila primerljiva z uspešnostjo dveh raziskav, ki zajemata dva različna sklopa evropske populacije. Ruiz je s sodelavci (2012) v študiji zajel šest evropskih populacij in ocenili vrednosti AUC za modro (0.986) oziroma rjavo (0.978) obarvanost oči. Walsheva s sodelavci (2013) pa je zajela sedem evropskih populacij. AUC za modro je ocenila na 0.964 in za rjavo 0.956 obarvanost oči. Iz tega lahko zaključimo, da je komplet IrisPlex visoko informativen in primeren za zelo natančno napovedovanje modre in rjave obarvanosti oči posameznikov evropske populacije. Vendar se tudi tu kažejo pomanjkljivosti predvsem pri natančnejšemu napovedovanju obarvanosti za vmesno kategorijo oči, v kateri so zajeti različni odtenki od zelene do sive. Kot smo omenili, bi bilo tu smiselno posameznike porazdeliti v še manjše sklope, ki bi zajemali čim bolj med seboj podobne barvne odtenke oči, in obenem uvesti nove, bolj informativne SNP-je.

Croat Med J. 2013;54:381-6
doi: 10.3325/cmj.2013.54.381

Prediction of eye color in the Slovenian population using the IrisPlex SNPs

Vanja Kastelic[1], Ewelina Pośpiech[3], Jolanta Draus-Barini[2], Wojciech Branicki[2,3], Katja Drobnič[1]

[1]National Forensic Laboratory, General Police Directorate, Police, Ministry of the Interior, Ljubljana, Slovenia

[2]Institute of Forensic Research, Section of Forensic Genetics, Kraków, Poland

[3]Department of Genetics and Evolution, Institute of Zoology, Faculty of Biology and Earth, Kraków, Poland

Aim To evaluate the accuracy of eye color prediction based on six IrisPlex single nucleotide polymorphisms (SNP) in a Slovenian population sample.

Methods Six IrisPlex predictor SNPs (*HERC2* – rs12913832, *OCA2* – rs1800407, *SLC45A2* – rs16891982 and *TYR* – rs1393350, *SLC24A4* – rs12896399, and *IRF4* – rs12203592) of 105 individuals were analyzed using single base extension approach and SNaPshot chemistry. The IrisPlex multinomial regression prediction model was used to infer eye color probabilities. The accuracy of the IrisPlex was assessed through the calculation of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the area under the receiver characteristic operating curves (AUC).

Results Blue eye color was observed in 44.7%, brown in 29.6%, and intermediate in 25.7% participants. Prediction accuracy expressed by the AUC was 0.966 for blue, 0.913 for brown, and 0.796 for intermediate eye color. Sensitivity was 93.6% for blue, 58.1% for brown, and 0% for intermediate eye color. Specificity was 93.1% for blue, 89.2% for brown, and 100% for intermediate eye color. PPV was 91.7% for blue and 69.2% for brown color. NPV was 94.7% for blue and 83.5% for brown eye color. These values indicate prediction accuracy comparable to that established in other studies.

Conclusion Blue and brown eye color can be reliably predicted from DNA samples using only six polymorphisms, while intermediate eye color defies prediction, indicating that more research is needed to genetically predict the whole variation of eye color in humans.

Received: April 24, 2013

Accepted: July 29, 2013

Correspondence to:
Vanja Kastelic
National Forensic Laboratory,
General Police Directorate, Police,
Ministry of the Interior
Vodovodna 95a
1000 Ljubljana, Slovenia,
vanja.kkastelic@gmail.com

Prediction of human visible characteristics by genotyping informative polymorphisms in DNA opens up a new perspective in the forensic field. Multiple genes including *HERC2, OCA2, MC1R, SLC24A5, SLC45A2, TYR, TYRP1, ASIP, SLC24A4, TPCN2, KITLG,* and *IRF4* have been associated with eye, hair, and skin color in European populations and they have been used in studies dealing with eye color prediction (1-14). Variation of iris color depends on the content of eumelanine, a brown light-absorbing biopolymer, which is present in higher concentrations in brown-eyed individuals (15,16). Although eye color is evidently a continuous variable, it has been often classified into three categories – blue, brown, and intermediate (4,14). Eye color variability is particularly striking in European populations, constituting a highly differentiating trait of potential use in forensic investigations (7,14,17). Recent studies have shown that a significant fraction of human iris color variation can be explained by polymorphisms within a single region in the human genome, comprising the evolutionary conserved *HERC2* gene and the neighboring *OCA2* gene located on the chromosome 15. It is assumed that the level of expression of the known pigmentation gene – *OCA2* – is controlled by polymorphism rs12913832 on *HERC2* locus (18,19). The remaining genes that have been shown to contribute to eye color variation are *SLC24A4, SLC45A2, TYR,* and *IRF4* (4,20,21). However, their impact on eye color prediction is lower and it seems to vary between populations (8,14,22,23). Since such differences may potentially affect accuracy of prediction in various populations, we further addressed this issue and analyzed a population sample of individuals with defined eye color from Slovenia.

Several prediction models have already been proposed to be useful in eye color prediction (4,8,9,17,23,24). Here we used six IrisPlex predictors, which were selected by Liu et al (4) from a larger set of polymorphisms potentially influencing pigmentation in humans and included into the IrisPlex prediction system (4,13,17). The IrisPlex prediction model is based on a multinomial logistic regression method and uses phenotype and genotype data from 3804 Dutch individuals. Based on these data the model gives three probabilities for blue, brown, and intermediate eye color (13). From the obtained probabilities, the most probable iris color is predicted based on recommendations given in Walsh et al (13).

MATERIAL AND METHODS

Sample collection, DNA extraction, and quantification

The study population comprised 105 unrelated Slovenian volunteers, 70 male and 35 female, who signed a written consent for their DNA to be used in the project. The study was approved by the National Medical Ethics Committee of the Republic of Slovenia. The eye color was defined according to descriptions provided by the volunteers and our own grading. For confirmation and in order to prevent bias, photographs of each donor's eyes were taken. Participants were divided into three categories according to eye color: blue, intermediate, and brown. The intermediate group included individuals with green eyes (lighter phenotype), hazel eyes (darker phenotype), and with combination of two or more pigments within the iris, such as blue or green eye color with brown peripupillary rings. The blue and the brown group included the individuals with the eye color that was clearly composed of only one color including all the shades of this particular color. Buccal swabs were collected from all volunteers using a SAFE® Box kit (ForensiX, Prionics AG, Zurich, Switzerland). DNA was extracted from the samples using Chelex extraction (25). DNA extracts were quantified using the Quantifiler Human DNA Quantification Kit (Applied Biosystems Inc., Foster City, CA, USA) in accordance with the manufacturer's guidelines.

Single nucleotide polymorphisms (SNP) genotyping

Four SNPs (*HERC2* – rs12913832, *OCA2* – rs1800407, *SLC45A2* – rs16891982, and *TYR* – rs1393350) were genotyped previously as described in Kastelic et al (26). The remaining two IrisPlex SNPs (*SLC24A4* – rs12896399 and *IRF4* – rs12203592) were genotyped for the purpose of this study using the protocol described by Walsh et al (17). Marker details and primer sequences are listed in Supplementary Tables 1 and 2. All cleaned products were analyzed on the ABI Prism 3130 Genetic Analyzer (Applied Biosystems) using run parameters as described previously (17,26).

Model-based prediction of eye color and evaluation of its accuracy

On the basis of the formula provided by Liu et al (4) and implemented in the eye color prediction model of the IrisPlex system, three prediction probability values were generated for each of the three phenotype categories (blue, intermediate, and brown) (Supplementary Table 3). The overall prediction accuracy was assessed as previously explained by calculating area under the curve (AUC) values using SPSS 19.0 (SPSS Inc., Chicago, IL, USA) (26). The AUC is the integral of receiver operating characteristic (ROC) curve, and ranges from 0.5, which represents a total absence of prediction, to 1.0, which represents a perfect prediction. Additionally the values of sensitivity, specificity, positive predic-

tive value (PPV), and negative predictive value (NPV) were calculated according to Liu et al using prediction threshold at the ≥0.7 level, which has been determined to be the most appropriate (1). Inconclusive results (below the threshold 0.7) were considered as negative results. From these, false negatives were used to calculate sensitivity and true negatives were used to calculate specificity.

RESULTS

Characteristics of the study population

The frequency of blue eye color in the studied sample was 44.7% (47 samples) and the frequency of brown eye color was much lower and reached 29.6% (31 samples). The individuals were categorized in these two eye color groups only when the color was homogenous, regardless of the intensity. The frequency of individuals in the intermediate eye color group was relatively high, 25.7% (27 samples).

Prediction accuracy of the IrisPlex model

Prediction accuracy expressed by the AUC (Figure 1) was 0.966 for blue, 0.913 for brown, and 0.796 for intermediate eye color (Table 1). The highest sensitivity was obtained for blue eye color and reached 93.6%, which means that 93.6% (44/47 individuals) of the analyzed blue-eyed persons were predicted correctly. The sensitivity for brown eye color was lower and amounted to 58.1% (18/31 individuals). The specificity values for blue (93.1%) and brown (89.2%) eye colors were very high. This means that 93.1% of non-blue individuals and 89.2% of non-brown individuals were correctly recognized as non-blue and non-brown, respectively. The highest PPV was obtained for blue eye color at the 91.7% level. This value means that in all cases of assignments to blue eye color category, 91.7% individuals in fact had blue eye color. The PPV value for brown eye color was lower and reached 69.2% and the PPV value for intermediate eye color could not be determined due to the fact that in no cases the intermediate eye color was predicted as intermediate. The NPV was very high for these two eye color categories. For blue eye color, NPV was 94.7% and this means that out of all cases where individuals were classified to the non-blue-eyed category, 94.7% cases were correctly classified as non-blue-eyed individuals. NPV for brown eye color equaled 83.5% and for intermediate eye color 74.3% (Supplementary Figure). A complete lack of sensitivity (0%) was observed for intermediate eye colors. This means that out of the 27 samples tested, none was classified as intermediate. Four out of 27 intermediate samples were categorized as blue with very high probabilities (>0.9). Eight (probability values above 0.7) and as many as 14 (probability values above 0.5) out of 27 intermediate samples were categorized as brown (Supplementary Figure).

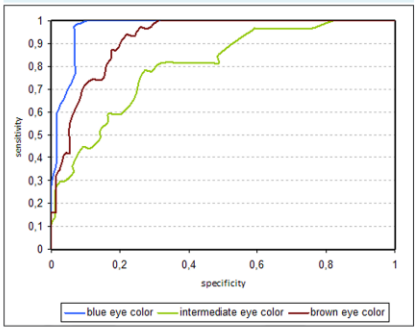


FIGURE 1. Receiver operating characteristic curve analysis of 105 Slovenian samples based on the IrisPlex prediction model.

DISCUSSION

In the studied population sample, blue eye color was present with the frequency 44.7%, while according to the Eupedia (http://www.eupedia.com/europe/maps_of_europe.shtml#eye_colour), the expected frequency of light eyed individuals in Slovenia should be between 50% and 79%. However, most of the individuals included in the intermediate category had green irises and blue irises with brown spots or peripupillary rings and therefore could be included in the group of light eyed people. Taking this into account, it can be said that the percentage of light-eyed individuals in the study was 56%, which is in accordance with the Eupedia.

TABLE 1. Parameters describing the accuracy of prediction with the IrisPlex model*

Parameter	Color		
	blue	intermediate	brown
Area under the receiver operating characteristic curve	0.966	0.796	0.913
Sensitivity (%)	93.6	0	58.1
Specificity (%)	93.1	100.0	89.2
Positive predictive value (%)	91.7	x*	69.2
Negative predictive value (%)	94.7	74.3	83.5

*x – zero in denominator.

The IrisPlex system includes six SNPs located on six genes (*HERC2* rs12913832, *OCA2* rs1800407, *SLC24A4* rs12896399, *SLC45A2* rs16891982, *TYR* rs1393350, and *IRF4* rs12203592), which are considered to be major genetic predictors of eye color (4,12-14). Numerous studies have confirmed that rs12913832 located on *HERC2* gene alone carries most of the eye color predictive information and is therefore the best known eye color predictor (4,5,8,14,22,27). We confirmed these results and also provided evidence that the CC genotype on rs12913832 was strongly associated with blue eye color (8,14,22,27). Among 49 Slovenian individuals carrying the CC genotype, 45 (91.8%) had blue eye color, which is comparable to other reports (93.4%) (14). In the remaining four cases, green eye color was observed and accordingly our study suggests that the CC genotype reliably predicts light eye colors. On the other hand, 12 (80%) out of 15 individuals carrying the TT genotype had brown eye color. It is worth noting that this proportion was higher (98%) in the larger EUREYE study (14). This difference may be a consequence of a relatively small sample set in this study, especially in brown eye color category, but also of differences in phenotype description among various studies. The latter is supported by phenotype distribution among CT genotype carriers – only 19 (46.3%) out of 41 individuals with the CT genotype had brown eyes in our population compared to 75.5% in Walsh et al (14). However, among the remaining 22 individuals, 2 had blue eye color but with intense brown peripupillary rings, 2 had green eye color, and 18 had hazel eye color. Overall, of the 56 individuals carrying the T allele either as homozygote or heterozygote state, 31 (55.4%) had brown eye color and even 49 (87.5%) had dark irises (brown or hazel). Notably, 8 (53.3%) individuals out of the 15 with a TT genotype had dark brown eye color and 4 (26.7%) had lighter brown eye color, which implies that other polymorphisms are important modifiers of eye color intensity. This study confirmed the prevailing role of the rs12913832 in the determination of blue and brown eye color. The significance of this position for human pigmentation seems to be undisputable and has been confirmed also by functional genome analysis (19). This huge effect on eye color detected in all the so far studied populations makes this SNP a key element of all eye color prediction methods. Moreover, it has been shown that this position also has influence on variation in hair and skin color (1,22). Indeed, rs12913832 is also an important element of hair color prediction models (1,3,17).

OCA2 SNP rs1800407, which ranked second among the best eye color predictors, has very low frequency of allele A (11.9%) and therefore may have had weak overall influence on variation in eye color in our population sample (4,12,13). The remaining IrisPlex predictors have been shown to have smaller effect on iris color variation but all six are implemented in the IrisPlex macro (13). It has been pointed out that IrisPlex can accurately predict blue and brown eye color while it is inefficient in the prediction of intermediate eye color and thus one should expect considerably lower prediction accuracy for this eye color category (Figure 1) (14). Indeed, the AUC values for blue and brown eye color categories in the Slovenian population were found to be very high and equaled 0.966 and 0.913, respectively. This result is similar to the AUC values obtained using multinomial logistic regression for a much larger group of seven European populations, which were 0.964 for blue and 0.956 for brown (14). Mengel-From et al (5) investigated *HERC2, OCA2,* and *SLC45A2* variation in 395 Danes using logistic regression and concluded that variation in *HERC2-OCA2* complex can be useful for reliable prediction of light and dark eye colors (5). Pośpiech et al (7) used Bayesian network built on 638 Poles and by testing 80 samples obtained AUC values of 0.783 and 0.583 for blue and brown color, respectively (7). These values were calculated from the scoring results rather than from probabilistic values, which is a much more conservative approach. They confirmed high sensitivity of prediction of blue eye color (80%) and lower for brown eye color (35%). They also concluded that eye color can be reliably predicted from the available DNA markers at the level light-dark, obtaining a high AUC value of 0.925 (7). A similar AUC value for light and dark eye color categories (AUC = 0.985) was obtained in our previous study involving the same Slovenian population and a different set of predictors (26). The accuracy values were also similar in the study of six European populations (with AUC values 0.986 and 0.978 for blue and brown, respectively) (8). In this latter work, a different model building data set was used, which could certainly influence the final AUC values. Recently, Allwood and Harbison (23) proposed a novel eye color prediction method utilizing classification tree approach and predicted blue and brown eye color with very high accuracy (23). Notably, all the mentioned studies reported serious difficulties with prediction of intermediate eye colors, indicating that regardless of the prediction method used (multinomial logistic regression, likelihood ratio, classification trees) accuracy parameters for intermediate eye color were very weak. This strongly suggests that the currently available eye color predictors are insufficient to reliably predict intermediate eye colors. In our study, the intermediate category was relatively large, comprising 27 of samples (25.7%), and a complete lack of sensitivity observed in this category confirms the previous

results indicating that more information is needed in order to predict intermediate phenotypes of iris color.

It is worth to mention that prediction of hair color from DNA is even more difficult. Branicki et al (1) and Walsh et al (17) both confirmed reliable prediction of red hair color, but prediction of other hair color categories is still affected by a relatively high error rate. Many problems with association studies aiming to find reliable predictors of various characteristics are caused by the continuous nature of the studied externally visible traits and the fact that loci with minor or weak effect on the phenotype are particularly difficult to discover in association studies. Therefore, there are different suggestions on phenotype description and proper eye color classification. For example Liu et al proposed a new digital method based on measuring continuous eye color variations using high-resolution digital full-eye photographs (28). Andersen et al (29) developed a Digital Iris Analysis Tool, which can be used to automatically identify and extract irises from high resolution digital images as well as calculate the so called Pixel Index of the Eye describing the eye color quantitatively. Such a detailed approach should reduce potential problems with correct categorization of eye colors. After all, further studies regarding additional genes and polymorphisms, as well as their interactions, contributing to variation in iris colors, especially non-blue and non-brown colors, should have a bearing on the improvement of their prediction accuracy (8,14,22,27). Multiple genes and their interactions are involved in the development of eye color variation, so further investigations of new pigmentation genes and SNP markers should be essential for more precise prediction of eye color, especially in the intermediate domain. We confirmed here that the model based prediction of eye color from DNA data was a reliable tool that can be useful in forensic investigations. Prediction of eye color from DNA extracted from biological traces may be used in criminal cases with unknown suspects. This tool may also supplement anthropological investigations by providing information about the eye color of a suspect. However, it is worth noting that forensic DNA phenotyping should be performed with care since it provides evidence that is associated with a relatively high error rate, particularly in case of some phenotypic categories like intermediate eye color.

In conclusion, our study contributed to the body of evidence on eye color phenotype variation across Europe, as well as on genotype distribution in the six eye color informative IrisPlex SNPs, providing data for Slovenian population sample. The obtained results confirmed the utility of the IrisPlex prediction model for accurate prediction of blue and brown eye colors. Further studies are needed to explain the remaining variation in human eye color and open up possibility for prediction of a complete spectrum of eye colors in humans.

Acknowledgments We thank the volunteers who provided DNA samples and eye photographs for this study.

Funding Ewelina Pośpiech and Wojciech Branicki received funding support from the European Union within the Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 285487.

Ethical approval The study was approved by the National Medical Ethics Committee of the Republic of Slovenia.

Declaration of authorship VK participated in designing the methods, analyzing and interpreting the results, and preparing the manuscript. EP participated in study conception and design, considerably contributed to statistical analysis and data interpretation, revised the manuscript, and gave the final approval for publication. JDB revised the data and calculations, and contributed to the manuscript preparation and revision. WB wrote substantial sections of the manuscript. KD produced the idea, designed the manuscript, and contributed to results interpretation and revision of the manuscript for important intellectual content.

Competing interests All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

References

1 Branicki W, Liu F, van Duijn K, Draus-Barini J, Pospiech E, Walsh S, et al. Model-based prediction of human hair color using DNA variants. Hum Genet. 2011;129:443-54. Medline:21197618 doi:10.1007/s00439-010-0939-8

2 Frudakis T, Thomas M, Gaskin Z, Venkateswarlu K, Chandra KS, Ginjupalli S, et al. Sequences associated with human iris pigmentation. Genetics. 2003;165:2071-83. Medline:14704187

3 Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS Genet. 2008;4:e1000074. Medline:18483556 doi:10.1371/journal.pgen.1000074

4 Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens ACJW. Eye color and the prediction of complex phenotype from genotypes. Curr Biol. 2009;19:R192-3. Medline:19278628 doi:10.1016/j.cub.2009.01.027

5 Mengel-From J, Børsting C, Sanchez JJ, Eiberg H, Morling N. Human eye colour and HERC2, OCA2 and MATP. Forensic Sci Int Genet. 2010;4:323-8. Medline:20457063 doi:10.1016/j.fsigen.2009.12.004

6 Pneuman A, Budimlija ZM, Caragine T, Prinz M, Wurmbach E. Verification of eye and skin color predictors in various populations. Leg Med (Tokyo). 2012;14:78-83. Medline:22284939 doi:10.1016/j.legalmed.2011.12.005

7 Pośpiech E, Draus-Barini J, Kupiec T, Wojas-Pelc A, Branicki W. Prediction of eye color from genetic data using Bayesian approach. J Forensic Sci. 2012;57:880-6. Medline:22372960

doi:10.1111/j.1556-4029.2012.02077.x

8 Ruiz Y, Phillips C, Gomez-Tato A, Alvarez-Dios J, Casares de Cal M, Cruz R, et al. Further development of forensic eye color predictive tests. Forensic Sci Int Genet. 2013;7:28-40. Medline:22709892 doi:10.1016/j.fsigen.2012.05.009

9 Spichenok O, Budimlija ZM, Mitchell AA, Jenny A, Kovacevic L, Marjanovic D, et al. Prediction of eye color and skin color in diverse populations using seven SNPs. Forensic Sci Int Genet. 2011;5:472-8. Medline:21050833 doi:10.1016/j.fsigen.2010.10.005

10 Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, et al. Genetic determination of hair, eye and skin pigmentation in Europeans. Nat Genet. 2007;39:1443-52. Medline:17952075 doi:10.1038/ng.2007.13

11 Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, et al. Two newly idnetified genetic determinants of pigmentation in Europeans. Nat Genet. 2008;40:835-7. Medline:18488028 doi:10.1038/ng.160

12 Walsh S, Lindenbergh A, Zuniga SB, Sijen T, de Knijff P, Kayser M, et al. Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. Forensic Sci Int Genet. 2011;5:464-71. Medline:20947461 doi:10.1016/j.fsigen.2010.09.008

13 Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet. 2011;5:170-80. Medline:20457092 doi:10.1016/j.fsigen.2010.02.004

14 Walsh S, Wollstein A, Liu F, Chakravarthy U, Rahu M, Seland JH, et al. DNA-based eye colour prediction across Europe with the IrisPlex system. Forensic Sci Int Genet. 2012;6:330-40. Medline:21813346 doi:10.1016/j.fsigen.2011.07.009

15 Sturm RA, Teasdale RD, Box NF. Human pigmentation genes: identification, structure and consequences of polymorphic variation. Gene. 2001;277:49-62. Medline:11602344 doi:10.1016/S0378-1119(01)00694-1

16 Passeron T, Mantoux F, Ortonne J. Genetic disorders of pigmentation. Clin Dermatol. 2005;23:56-67. Medline:15708290 doi:10.1016/j.clindermatol.2004.09.013

17 Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, et al. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet. 2013;7:98-115. Medline:22917817 doi:10.1016/j.fsigen.2012.07.005

18 Eiberg H, Troelsen J, Nielsen M, Mukkelsen A, Mengel-From J, Kjaer KW, et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Hum Genet. 2008;123:177-87. Medline:18172690 doi:10.1007/s00439-007-0460-x

19 Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. Genome Res. 2012;22:446-55. Medline:22234890 doi:10.1101/gr.128652.111

20 Soejima M, Koda Y. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. Int J Legal Med. 2007;121:36-9. Medline:16847698 doi:10.1007/s00414-006-0112-z

21 Duffy DL, Montgomery GW, Chen W, Zhao ZZ, Le L, James MR, et al. A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. Am J Hum Genet. 2007;80:241-52. Medline:17236130 doi:10.1086/510885

22 Branicki W, Brudnik U, Wojas-Pelc A. Interactions between HERC2, OCA2 and MC1R may influence human pigmentation phenotype. Ann Hum Genet. 2009;73:160-70. Medline:19208107 doi:10.1111/j.1469-1809.2009.00504.x

23 Allwood JS, Harbison S. SNP model development for the prediction of eye colour in New Zealand. Forensic Sci Int Genet. 2013;7:444-52. Medline:23597786 doi:10.1016/j.fsigen.2013.03.005

24 Hart KL, Kimura SL, Mushailov V, Budimlija ZM, Prinz M, Wurmbach E. Improved eye- and skin-color prediction based on 8 SNPs. Croat Med J. 2013;54:248-56. Medline:23771755 doi:10.3325/cmj.2013.54.248

25 Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR – based typing from forensic material. Biotechniques. 1991;10:506-13. Medline:1867860

26 Kastelic V, Drobnic K. A single-nucleotide polymorphism (SNP) multiplex system: the association of five SNPs with human eye and hair color in the Slovenian population and comparison using a Bayesian network and logistic regression model. Croat Med J. 2012;53:401-8. Medline:23100201 doi:10.3325/cmj.2012.53.401

27 Pospiech E, Draus-Barini J, Kupiec T, Wojas-Pelc A, Branicki W. Gene-gene interactions contribute to eye color variation in humans. J Hum Genet. 2011;56:447-55. Medline:21471978 doi:10.1038/jhg.2011.38

28 Liu F, Wollstein A, Hysi PG, Ankra-Badu GA, Spector TD, Park D, et al. Digital quantification of human eye colour highlights genetic association of three new loci. PLoS Genet. 2010;6:e1000934. Medline:20463881 doi:10.1371/journal.pgen.1000934

29 Andersen JD, Johansen P, Harder S, Christoffersen SR, Delgado MC, Henriksen ST, et al. Genetic analyses of the human eye colours using a novel objective method for eye colour classification. Forensic Sci Int Genet. 2013;7:508-15. Medline:23948321 doi:10.1016/j.fsigen.2013.05.003

2.4 DOLOČITEV ZUNANJEGA VIDEZA LJUDI S PREISKAVAMI DNK

Določitev zunanjega videza ljudi s preiskavami DNK

V. Kastelic, K. Drobnič

Revija za kriminalistiko in kriminologijo, 2012, 63, 3: 225–228

Forenzične preiskave dandanes omogočajo identifikacije bioloških sledi le na podlagi predhodnega poznavanja profilov DNA osumljencev oziroma oseb, ki se nahajajo v evidencah preiskav DNA. V Nacionalnem forenzičnem laboratoriju smo na področju identifikacije bioloških sledi zelo uspešni. Kljub temu so in bodo obstajala kazniva dejanja, pri katerih bioloških sledi s kraja zločina ne moremo povezati z nobeno od nam znanih oseb. Za reševanje nekaterih od teh primerov se v forenzične preiskave vpeljuje nova metoda. Na podlagi nove metode bomo lahko iz bioloških sledi določali zunanji izgled osebe oziroma njegove pigmentacije lastnosti, kot so barva oči in las. Rezultate teh genetskih analiz bi lahko imenovali »genetski očividec«, saj ima isto vlogo kot očividec kaznivega dejanja. V tujini so s tovrstnimi raziskavami razrešili že nekaj kriminalističnih primerov. Rutinsko pa jo izvaja le Nizozemski forenzični inštitut, in še ta le pri težjih kaznivih dejanjih. Raziskave na tem področju bodo torej služile predvsem za usmeritve kriminalističnih preiskav v zmanjšan krog morebitnih osumljencev, saj bo delno znan njihov zunanji izgled. Namen vpeljave tovrstnih raziskav v Nacionalni forenzični laboratorij policije je, da tudi v Sloveniji pristopimo k razreševanju težjih kaznivih dejanj z uporabo novih molekularnih genetskih metod, katerih rezultate bodo preiskovalci lahko uporabili kot novo analitično orodje za iskanje storilcev kaznivih dejanj.

Revija za kriminalistiko in kriminologijo / Ljubljana 63 / 2012 / 3, s. 225-228

Določitev zunanjega videza ljudi s preiskavami DNK

Vanja Kastelic*, Katja Drobnič**

Forenzične preiskave danes omogočajo ugotavljanje identitete osebe, ki je pustila biološko sled, le na podlagi predhodnega poznavanja profilov DNK osumljencev oziroma oseb, ki so v evidencah preiskav DNK. V Nacionalnem forenzičnem laboratoriju smo skozi več kot petnajstletno prakso postali na tem področju zelo uspešni. Kljub temu so in bodo obstajala kazniva dejanja, pri katerih bioloških sledi s kraja zločina ne moremo povezati z nobeno od nam znanih oseb. Za reševanje nekaterih od teh primerov se v forenzične preiskave vpeljuje nova metoda. Na njeni podlagi bomo lahko iz bioloških sledi določali zunanji videz osebe oziroma njene pigmentacijske lastnosti, kot so barva oči in las. Rezultate teh genetskih analiz bi lahko imenovali »genetski očividec«, saj imajo isto vlogo kot očividec kaznivega dejanja. V tujini so s tovrstnimi raziskavami razrešili že nekaj kriminalističnih primerov. Rutinsko pa jih izvaja le nizozemski forenzični inštitut, in še ta le pri težjih kaznivih dejanjih. Raziskave na tem področju bodo torej služile predvsem za usmeritve kriminalističnih preiskav v zmanjšan krog morebitnih osumljencev, saj bo delno znan njihov zunanji videz. Namen vpeljave tovrstnih raziskav v Nacionalni forenzični laboratorij policije je, da tudi v Sloveniji pristopimo k razreševanju težjih kaznivih dejanj z uporabo novih molekularnih genetskih metod, katerih rezultate bodo preiskovalci lahko uporabili kot novo analitično orodje za iskanje storilcev kaznivih dejanj.

Ključne besede: SNP-označevalci, vidne karakteristike človeka, biološke sledi

UDK: 343.983.2

1 Uvod

Ugotavljanje identitete storilcev kaznivih dejanj, žrtev masovnih nesreč, pogrešanih oseb ter sorodstvenih razmerij in naših prednikov danes skorajda ne mine brez uporabe forenzičnih genetskih preiskav, če so seveda na voljo ustrezne biološke sledi in primerjalni vzorci. Biološke sledi za uspešne genetske preiskave predstavljajo praktično vse biološke sledi, kot so kri, semenska tekočina, slina, epitelne celice, kosti, zobje in drugo. V rutinskih genetskih preiskavah se zdaj najpogosteje uporabljajo avtosomalni[1] STR-označevalci (angl. short tandem repeat)[2], ki skupaj z amelogeninskim označevalcem za spol[3] predstavljajo forenzični genetski standard nacionalnih ali kriminalističnih evidenc DNK oseb[4] in bioloških sledi (Asplen, 2009). Zaradi milijonskega števila profilov DNK[5] v različnih evidencah DNK in izrednega tehnološkega razvoja so avtosomalni STR-označevalci postali nenadomestljivi v rutinskih preiskavah. Vendar pa se za različne primere, kot so preiskave spolnih deliktov, ugotavljanje identitete umorjenega, arheološke raziskave in drugo, ki jih zaradi različnih težav ni mogoče razrešiti le z analizo avtosomalnih STR-označevalcev, uveljavljajo tudi druge forenzične genetske preiskave. Te so analize STR-označevalcev na spolnih kromosomih in analize mitohondrijske DNK[6] (Quintans, Alvarez-Iglesias, Phillips, Lareu in Carracedo, 2004; Lessing in sod., 2005). V novejšem času se uveljavljajo analize polimorfizmov posameznih nukleotidov (angl. single nucleotide polymorphism, SNP) za napovedova-

* Vanja Kastelic, univ. dipl. mikrobiol., kriminalistično-tehnična izvedenka v Nacionalnem forenzičnem laboratoriju, GPU Policija. MNZ.

** Katja Drobnič, redna profesorica za forenzično znanost, FVV, UM; nadzornica kakovosti v Nacionalnem forenzičnem laboratoriju, GPU Policija. MNZ.

1 Avtosom ali avtosomalni kromosom je vsak kromosom (nitasta struktura, sestavljena iz DNK in beljakovin) razen spolni. V človeški celici jih je 22, oštevilčeni so od 1 do 22 ter so pri moških in ženskah enaki. Spolna kromosoma sta dva, označena z Y in X (kromosom Y določa moški spol).

2 Genetski označevalec predstavlja področje na DNK, katerega polimorfizem uporabljamo za zaznavo različnih polimorfizmov: fenotipskih, biokemijskih ali DNK.

3 Amelogeninski označevalec je najpogosteje uporabljen fenotipski marker v forenzičnih genetskih preiskavah za določanje spola iz biološke sledi. Leži na spolnih kromosomih, njegova struktura je odvisna od spolnega kromosoma (X ali Y), na katerem leži.

4 Kriteriji za vnos oseb v evidence DNK so zakonsko določeni in različni med državami. Danes ima večina držav dovoljen vnos podatkov DNK osumljenih v evidenco DNK, ne glede na težo kaznivega dejanja.

5 Profil DNK je zbirka analiziranih klasifikacijskih značilnosti, ki ležijo na specifičnih področjih DNK (področja STR-označevalcev). Sestavljen je iz numeričnih in črkovnih oznak. Vsaka oznaka predstavlja tip oblike oziroma alel na analiziranem področju DNK.

6 Mitohondrijska DNK je DNK, ki je v mitohondrijih, to so organeli v vsaki celici, ki so odgovorni za "celično dihanje". Mitohondrijska DNK se navadno deduje samo po materini strani ne glede na spol potomca.

Revija za kriminalistiko in kriminologijo / Ljubljana 63 / 2012 / 3, s. 225-228

nje zunanjega videza posameznika ali njegovega biogeografskega izvora (Budowle, 2004; Kayser in Schneidet, 2009), ki je povsem nov molekularni genetski pristop; rezultati teh genetskih analiz pa vodijo kriminalistične preiskovalce do ožjega kroga ljudi, ki so jim bili pred tem povsem neznani.

2 Molekularno-genetsko ozadje novih preiskav

2.1 Polimorfizem posameznega nukleotida

Humani genom je med ljudmi enak v 99,9 odstotkih. Sestavljen je iz treh milijard baznih parov[7]. Ocenjeno je, da obstaja več kot 10 milijonov možnih mest v DNK-verigi, kjer se lahko pojavi sprememba v eni bazi oziroma polimorfizem posameznega nukleotida (SNP)[8]. Te spremembe so lahko v obliki substitucije (zamenjave med bazami), insercije (dodatek ene baze) ali delecije (izguba ene baze). SNP-ji so torej tista mesta v humanem genomu, kjer se posamezniki zelo razlikujemo med seboj. SNP-jem pripisujejo kar devetdesetodstotno odgovornost za razlikovanje med posamezniki na nivoju DNK. Te manjše spremembe našega genoma pa imajo lahko velik vpliv na zunanji videz posameznika (Budowle in van Daal, 2008).

Visoka zastopanost SNP-jev v humanem genomu je bil eden od razlogov za veliko zanimanje za njih že v preteklosti, kljub njihovi preprostosti in dokaj nizki informativnosti. Raznolikost večine SNP-jev namreč temelji le na dveh tipih oziroma bialelnem polimorfizmu[9]. Vendar pa je njihova dodatna prednost, da imajo zelo nizko mutacijsko stopnjo, kar pomeni, da isti tip SNP-ja ostaja enak preko več generacij. Zato so zelo primerni tudi za določanje sorodstvenih povezav in ugotavljanje evolucijskega izvora posameznika (Budowle, 2004). Poleg tega pa je analiza SNP-označevalcev zelo primerna za preiskave bioloških sledi z močno razgrajeno DNK, kar se je izkazalo za zelo uspešno tudi pri ugotavljanju identitete žrtev z ruševin terorističnega napada na stolpnici World Trade Centra v New Yorku (Mertens, 2009). Zdaj pa se je zanimanje za preiskave SNP-jev še povečalo zaradi možnosti njihove uporabe pri napovedovanju vidnih karakteristik ljudi.

[7] Bazni par je par nukleotidnih baz na sosednjih komplementarnih verigah DNK, ki sta med seboj povezani.

[8] Nukleotidi so osnovni gradniki DNK, ki vsebujejo fosfatni ostanek, sladkor in organske baze. Ločimo štiri tipe nukleotidov, ki se razlikujejo le v vrsti organske baze, ki ga sestavljajo.

[9] Bialelni polimorizem pomeni, da obstajata le dve obliki DNK na preiskovanem področju. Za posamezen STR-označevalec praviloma obstaja več kot deset oblik. Alel je alternativna oblika gena (zaporedja DNK) na določenem genetskem lokusu.

2.2 Zunanji videz posameznika

Zunanji videz posameznika oziroma lastnosti posameznika, ki bi jih bilo v bližnji prihodnosti mogoče rutinsko preiskovati v forenzične namene, so predvsem obarvanost oziroma pigmentacija oči, las in kože ter struktura obraza. Napovedovanje telesne višine pa se je izkazalo za manj zanesljivo (Kayser in de Knijff, 2011). Zunanji videz posameznika je seveda zapisan v njegovem genomu, kar v vsakdanjem okolju nakazujejo identični dvojčki s svojo podobo. Zunanji videz je zelo kompleksna lastnost, saj nanj vpliva večje število različnih genov, med katerimi potekajo medsebojne interakcije in njihove interakcije z okoljem (Mertens, 2009).

2.3 Pigmentacijske lastnosti ljudi

Raznolikost v obarvanosti oči, las in kože je rezultat različne oblike in količine melanina, ki je naš glavni pigment in se sintetizira v specializiranih veziklih (melanosomih) znotraj pigmentacijskih celic (melanocitah). Znani sta dve glavni obliki melanina, in sicer eumelanin, ki je obarvan rjavo-črno, ter feomelanin, ki je obarvan rumeno-oranžno. Večina raziskanih genov na tem področju deluje ravno na sintezo melenina oziroma sintezo melanosomov, kar je ključno pri izražanju pigmentacijskih lastnosti (Passeron, Mantoux in Ortonne, 2005). Ravno pravilen izbor določenih genov pa je ključen za natančnejše in pravilnejše napovedovanje zunanjega videza posameznika na podlagi bioloških sledi.

3 Analize novih označevalcev

Različne metode za analiziranje novih označevalcev oziroma SNP-jev se na področju določanja zunanjega videza posameznika v grobem delijo glede na indirekten in direkten pristop. Indirektni pristop vključuje preiskave specifičnih SNP-označevalcev, ki posameznika dokaj natančno uvršča v njegovo etnično skupino. Torej deloma nakazuje posameznikovo lastnost glede na poznavanje zunanjega videza njegovih prednikov. Znano je na primer, da je svetlejša polt bolj domena ljudi s severnega dela Evrope kot tistih z drugih predelov sveta (Budowle in van Daal, 2008).

Direktni pristop pa vključuje preiskovanje čim večjega števila genov in znotraj njih specifičnih SNP-označevalcev, ki skupno vplivajo na posamezno vidno lastnost posameznika. Do danes je večina raziskav SNP-označevalcev na področju zunanjega videza narejenih na pigmentacijskih genih, ki v medsebojnih interakcijah določajo specifično barvo oči, las in kože. Prve uspešne preiskave na področju t. i. pigmentacijskih genov so izvedli na miših, kjer so raziskovalci preučevali obarvanost njihove dlake. Pri miših so do sedaj določili približno 120 genov, medtem ko so jih le približno trideset povezali tudi z redkimi pigmentacijskimi obolenji pri ljudeh. Med njimi je do sedaj raziskanih okoli dvanajst genov, ki bolj specifično

vplivajo na pigmentacijske lastnosti in jih je mogoče uporabiti za forenzične preiskave (Branicki, Brudnik, Draus-Barini, Kupiec in Wojas-Pelc, 2008; Kayser in Schneider, 2009).

Dosedanje svetovne raziskave, ki so bile izvedene na omejenem številu populacijskih vzorcev, kažejo visoko stopnjo pravilne napovedi za pigmentacijske lastnosti, kot so barva oči in las. Vendar jih zaradi pomanjkanja populacijskih podatkov trenutno večina forenzičnih laboratorijev še ne uporablja rutinsko. Za dokončno potrditev posameznikove povezanosti s pigmentacijskimi lastnostmi je tako treba opraviti raziskave v čim več različnih populacijah, da bi se ugotovilo, če izbrani SNP-ji zadostijo kriterijema populacijske genetike. S tem bi bila omogočena neposredna primerjava rezultatov med različnimi populacijami po svetu.

4 Vpeljava nove metode v Nacionalni forenzični laboratorij

V Nacionalnem forenzičnem laboratoriju smo za določanje zunanjega videza posameznika s preiskavo SNP-označevalcev uporabili direktni pristop. V ta namen smo uvedli novo metodo, t. i. SNaPshot™ (Applied Biosystems), ki je po svetu precej razširjena za analize SNP-označevalcev (Sanchez in Borsting, 2003; Sobrino, Brion in Carracedo, 2005). Metoda je namreč zelo občutljiva, ponovljiva, robustna in fleksibilna ter cenovno primerna za analize v forenzičnih laboratorijih, saj temelji na enaki tehnologiji, kot jo že uporabljamo v našem laboratoriju. Za analizo z omenjeno metodo smo v našo raziskavo vključili dvanajst SNP-označevalcev za določanje barve oči in las, za katere je znano, da so zelo polimorfni znotraj evropske populacije. V nasprotju s predhodnimi raziskavami smo uspeli dvanajst SNP-označevalcev združiti v eno reakcijo. S tem smo zmanjšali porabo biološkega materiala pri analizi in hkrati skrajšali čas za analize za napovedovanje pigmentacijskih lastnosti.

5 Etični vidik novih preiskav

Test za napoved zunanjega videza mora temeljiti na lastnostih, ki so običajno opažene in jih je težko zakriti. Z etičnega vidika kljub temu, da so vidne s prostim očesom, nekaterih genetskih podatkov ne moremo imeti za javne, kajti mnogi geni za pigmentizacijo so tesno povezani z napovedovanjem rizičnosti za kožnega raka (Sturm, Teasdale in Box, 2001). Forenzična skupnost teh SNP-označevalcev ne sme uporabljati in jih tudi ne uporablja za napoved vidnih značilnosti, prav tako tudi ne SNP-označevalcev, ki so povezani s kakršnimi koli drugimi bolezenskimi stanji.

Že v začetkih omenjenih raziskav v povezavi s pigmentacijskimi lastnostmi ljudi so raziskovalci težili k čim manjšim povezovanjem z genskimi označevalci, ki lahko podajajo npr. specifična bolezenska stanja posameznika (Kayser in Schneider, 2009). Hkrati pa je treba poudariti, da do sedaj še niso odkrili, da bi bila pojavnost SNP-jev, vključenih v našo raziskavo, povezana s pojavnostjo neke bolezni, ampak le s pigmetacijskimi lastnostmi. Če bi se v prihodnje izkazalo, da je kateri izmed njih povezan z bolezenskim stanjem, ga bomo morali izločiti iz nadaljnje uporabe za določanja pigmentacijskih lastnosti.

6 Razprava

Z današnjo odprtostjo državnih mej in večjo pretočnostjo ljudi se lahko veča tudi število kaznivih dejanja, tudi najhujših, v katerih so nam storilci lahko popolnoma neznani. Najboljša načina za uspešno razreševanje kaznivih dejanj sta dva, in sicer nadaljnja izmenjava podatkov profilov DNK med državami, ki je v veljavi že nekaj let, ter uvajanje najnovejših forenzičnih metod v forenzične laboratorije. Ena od teh forenzičnih metod je zagotovo preiskovanje pigmentacijskih lastnosti posameznika iz bioloških sledi, najdenih na kraju kaznivega dejanja. Te preiskave se ponekod že uporabljajo v kriminalističnih preiskavah in so omogočile razrešiti že več različnih primerov kaznivih dejanj po vsem svetu (Forensic DNA ethics, 2012). Kot primer lahko navedemo uspešno raziskano kaznivo dejanje posilstva in umora 16-letne deklice na Nizozemskem že leta 1999; forenzično preiskavo je vodil Peter de Knijffe. Na prošnjo tamkajšnje policije in seveda na podlagi zavarovane semenske tekočine so lahko določili pigmentacijske lastnosti donorja omenjene biološke sledi, ki so pokazale delen zunanji videz storilca krutega kaznivega dejanja. S tem je predstavil možnost, kako bi lahko forenzične genetske preiskave služile pri bolj specifičnem usmerjanju preiskovanja kaznivih dejanj na podlagi zmanjšanja kroga osumljencev, saj so z njimi podane določene vidne karakteristike storilca. Obenem so preiskave pomirile etnična trenja na območju, kjer je bila deklica umorjena, saj so nakazale, naj bi bil storilec domačin in ne priseljenec iz vzhodnih držav. To pa je bil tudi povod za pobudo nizozemske vlade sodnim institucijam za zakonsko ureditev teh preiskav, kar se je zgodilo že maja 2003 (Forensic DNA ethics, 2012).

7 Sklepne misli

Vedno novejša znanja in tehnologije na področju genetike in molekularne biologije omogočajo učinkovitejše in hitrejše analize forenzičnih preiskav. Komaj dvajset let od prvih analiz DNK je forenzična tipizacija DNK ena ključnih pri obsodbi ali oprostitvi osumljenca, seveda ko nam je njegov DNK-profil znan. Napredek na področju preiskav DNK je neverjeten in mu sledimo tudi v Nacionalnem forenzičnem laboratoriju. Ravno zato smo, zaenkrat v raziskovalne namene, uvedli novo občutljivo metodo za ugotavljanje najverjetnejšega zunanjega

Revija za kriminalistiko in kriminologijo / Ljubljana 63 / 2012 / 3, s. 225-228

videza osebe oziroma njenih pigmentacijskih lastnosti oči in las. Naše dosedanje raziskave za slovensko populacijo so zelo vzpodbudne, saj smo pravilno obarvanost oči in las napovedali z verjetnostjo med 80 in 96 %. Informacije o zunanjem videzu posameznika bodo v prihodnje, ko bodo opravljene še dodatne validacijske študije, služile predvsem za zmanjšanje kroga mogočih osumljencev in s tem vodile kriminalistične preiskovalce do najverjetnejših storilcev, predvsem pri težjih kaznivih dejanjih.

8 Literatura

1. Asplen, C. (2009). ENFSI Survey on the DNA Profile Inclusion, *Removal and Retention of Member States Forensic DNA Database*: GTH-Governmental Affairs, Washington, str. 1–11.
2. Branicki, W., Brudnik, U., Draus-Barini, J., Kupiec, T., Wojas-Pelc, A. (2008). Association of the SLC45A2 gene with physiological human hair colour variation. *Journal of Human Genetics, 53*, str. 966–971.
3. Budowle, B. (2004). SNP typing strategies. *Forensic Science International*, 146 Suppl.: str. 139–142.
4. Budowle, B., van Daal, A. (2008). Forensically relevant SNP classes. *Biotehniques, 44*, str. 603–610.
5. Forensic DNA ethics (2012). Pridobljeno 25. januarja 2012, http://forensicdnaethics.com/resources/cases?start=8
6. Kayser, M., Schneider, P. M. (2009). DNA-based prediction of human externally visible characteristic in forensics: motivations, scientific challenges and ethical considerations. *Forensic Science International Genetics, 3*, str. 154–161.
7. Kayser, M., de Knijff, P. (2011). Improving human forensic through advances in genetics, genomics and molecular biology. *Nature Reviews – Genetics, 11*(12), str. 179–192.
8. Lessing, R., Zoledziewska, M., Fahr, K., Edelmann, J., Kostrzewa, M., Dobosz, T., Kleemann, W. J. (2005). Y-SNP genotyping – a new approach in forensic analysis. *Forensic Science International, 154*(2-3), str. 128–136.
9. Mertens, G. (2009). Forensic DNA Typing: Quo Vadis? *The open forensic science journal, 2*, str. 21–28.
10. Passeron, T., Mantoux, F., Ortonne, J. (2005). Genetic disorders of pigmentation. *Clinics in Dermatology, 23*(1), str. 56–67.
11. Quintans, B., Alvarez-Iglesias, V., Phillips, C., Lareu, M. V., Carracedo, A. (2004). Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Forensic Science International, 140*(2-3), str. 251–257.
12. Shancez, J., Borsting, C. (2003). Multiplex PCR and minisequencing of SNPs – a model with 35 Y chromosome SNPs. *Forensic Science International, 137*(1): str. 74–84.
13. Sobrino, B., Brion, M., Carracedo, A. (2005). SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Science International, 154*(2-3), str. 181–194.
14. Sturm, A. R., Teasdale, D. R., Box, F. N. (2001). Human pigmentation genes: identification, structure and consequences of polymorphic variation. *Gene, 277*(1-2), str. 49–62.

Determination of the Physical Appearance of Persons by Means of DNA Investigations

Vanja Kastelic, BSc. in Microbiol., criminal and technical expert in the National Forensic Laboratory, General Police Directorate. Ministry of the Interior

Katja Drobnič, Full Professor of Criminal Technics, Faculty of Criminal Justice and Security, University of Maribor; quality supervisor in the National Forensic Laboratory, General Police Directorate. Ministry of the Interior

Today, forensic investigation allows the identification of biological traces based only on knowledge of the DNA profiles of suspects or persons found in DNA investigation records. The National Forensic Laboratory has been very successful in the identification of biological traces. Nevertheless, there are, and will be, criminal offences in which biological traces from a crime scene cannot be linked to any known person. In order to solve some of these cases, a new method is being introduced to forensic investigations. Based on the new method, the physical appearance of the person, or their pigmentation characteristics, such as eye and hair colour, could be determined from biological traces. The results of these genetic analyses could be termed 'genetic eyewitnesses', as they will have the same role as eyewitnesses to criminal offences. Such investigations abroad have helped to solve a number of criminal cases. However, only the Netherlands Forensic Institute performs these routinely, and only for serious criminal offences. Research in this area will be mainly used to focus criminal investigations on a narrower range of suspects, as their physical appearance will be partly known. The introduction of such investigations to the National Forensic Laboratory is intended to enable Slovenia to approach solving serious criminal offences by using new molecular techniques, the results of which will be used by investigators as a new analytical tool to identify criminals.

Key words: SNP markers, visual characteristics of people, biological traces

UDC: 343.983.2

3 RAZPRAVA IN SKLEPI

3.1 RAZPRAVA

3.1.1 Gen *SRY* v spolnem kromosomu Y

Določanje spola donorja bioloških sledi v forenzičnih preiskavah je v nekaterih primerih ključnega pomena, predvsem pri raziskovanju kaznivih dejanj zoper spolno nedotakljivost (moškega/ženske). Vendar je bilo do sedaj opravljenih mnogo raziskav, pri katerih so določili mutacije v najbolj pogosto uporabljenemu amelogeninskemu genu *AMEL*, kar je posledično razlog za napačno določanje spola (Santos in sod., 1998; Roffey in sod., 2000; Steinlechner in sod., 2002; Drobnič, 2006).

Eden od možnih pristopov pri raziskavah, v katerih so raziskovalci skušali pravilno določiti spol donorja biološke sledi, je bil ta, da so, zato da so se izognili delecijam znotraj amelogeninskega gena, skonstruirali krajša začetna oligonukletoida in s tem pridobili dva produkta pomnoževanja, in sicer 80 bp na kromosomu X in 83 bp na kromosomu Y. Vendar se ta metoda kasneje ni izkazala za najprimernejšo (Haas-Rochholz in Weiler, 1997).

Primernejše so bile raziskave, pri katerih so se raziskovalci usmerili v konstruiranje različnih začetnih oligonukleotidov znotraj gena *SRY*, za katerega je značilno, da je močno evolucijsko ohranjen (Santos in sod., 1998; Dennis Lo in sod., 1998; Singh in sod., 1999; Tozzo in sod., 2013). Leta 1998 je Santos s sodelavci prvi skonstruiral par začetnih oligonukleotidov (SRY 93 I, SRY 93 II), ki da po pomnoževanju DNA fragment dolžine 93 baznih parov in se pojavi le pri moških. Zaporedji sta (Santos in sod., 1998):
SRY 93 I: **5'** ATA AGT ATC GAC CTC GTC GGA AG **3',**
SRY 93 II: **5'** GCA CTT CGC TGC AGA GTA CCG AAG **3'**.
Poleg omenjenih parov začetnih oligonukleotidov so bili nato skonstruirani in uporabljeni tudi drugi pari, ki so bistvenega pomena pri reševanju kaznivega dejanja zoper spolno nedotakljivost ter so pomembni tudi pri prenatalni diagnostiki bolezni, ki so specifične le za moške (Naito in sod., 1994; Dennis Lo in sod., 1998; Singh in sod., 1999; Drobnič, 2006; Tozzo in sod., 2013).

Za potrebe forenzičnih genetskih preiskav, ki danes temeljijo na različnih komercialnih kompletih, je bilo potrebno skonstruirati primerne začetne oligonukleotide, ki bi lahko bili vključeni v že uveljavljene komercialne komplete v hkratni reakciji PCR, pri standardnih pogojih PCR, ki jih narekuje posamezni uporabljeni komercialni komplet. Potreba po vključitvi dodatnega genetskega označevalca v obstoječe komercialne komplete sloni tudi na tem, da v kompletih ni prisotne notranje pozitivne kontrole, zato pri sami individualizaciji biološke sledi ni mogoče zagotovo trditi, ali gre za moško ali

žensko osebo. Posledično se kaže potreba po vključitvi dveh genetskih označevalcev v posamezni komplet, kar pa za enkrat še ni del prakse pri proizvajalcih komercialnih kompletov (Hedges in sod., 2003; Kastelic in sod., 2009).

Da ne bi prišlo do morebitne napačne interpretacije spola donorja specifične biološke sledi, smo tako kot drugi raziskovalci tudi mi vpeljali nov genetski označevalec SRY na kromosomu Y (Thangaraj in sod.; 2002; Jobling in sod., 1997) za potrebe pomnoževanja znotraj komercialnega kompleta AmpFlSTR® SGM™ Plus Amplification (AB) in opravili vse potrebne validacijske študije na podlagi priporočil SWGDAM (Scientific Working Group on DNA Analysis Methods). Zaporedji novih začetnih oligonukleotidov znotraj gena *SRY* sta:
96 I: FAM - AGC AGT CAG GGA GGC AGA TCA,
96 II: CCC CCT AGT ACC CTG ACA ATG TAT T (Drobnič, 2006).

Validacijske študije (Kastelic in sod., 2009) so potrdile, da je pomnoževanje gena *SRY* z novim parom začetnih oligonukleotidov zanesljivo in ponovljivo pri vseh 115 sorodno nepovezanih moških, vključenih v preiskavo, in uspešno že pri vnosu minimalne količine DNA (25.0 pg). Pomnoževanje gena *SRY* je potekalo v hkratni PCR reakciji skupaj z ostalimi enajstimi pari začetnih oligonukleotidov znotraj komercialnega kompleta AmpFlSTR® SGM™ Plus Amplification (AB), ki pa niso motili samega pomnoževanja gena *SRY*. Obenem vnos novega para začetnih oligonukleotidov ni vplival na pomnoževanje lokusov, ki so zajeti v omenjenem komercialnem kompletu (Kastelic in sod., 2009). Ravno tako smo potrdili, da vnos novega para začetnih oligonukleotidov ne vpliva na pomnoževanje ostalih 16 lokusov znotraj komercialnega kompleta AmpFlSTR® NGM™ amplification (AB), za katerega pa je treba opraviti še validacijske študije (Priloga A).

Produkt pomnoževanja gena *SRY* smo dokazali tako tudi pri moških, ki imajo delecijo v amelogeninskem genu, oziroma pri moških, ki bi bili določeni kot ženska oseba, če bi za dokazovanje spola uporabili le amelogeninski genetski označevalec. Določanje spola s hkratno uporabo obeh genetskih označevalcev, tako amelogeninskega gena kot tudi gena *SRY*, bi zagotovilo bolj verodostojne rezultate, medtem ko jih le sedaj uporabljeni amelogeninski gen ne (Kastelic in sod., 2009).

3.1.2 Polimorfizem posameznih nukleotidov in pigmentacijske lastnosti

V zadnjih letih so DNA preiskave bioloških sledi, pridobljenih pri različnih kaznivih dejanjih, postale ključne za napovedovanje zunanjega videza ljudi, niso pa še postale del rutinskih preiskav v forenzični genetiki. Možnost takega napovedovanja bi bila v prvi vrsti v veliko pomoč kriminalističnim preiskavam, predvsem v primerih, ko za specifično biološko sled ne bi mogli podati primernega DNA profila (v primeru nizke

koncentracije DNA) oziroma ko za omenjeni profil ne bi mogli potrditi ujemanja z osumljeno ali s katero drugo osebo iz evidenc preiskav DNA. Tako bi za potrebe kriminalističnih preiskav z napovedjo zunanjega videza posameznika, tako njegovega spola kot tudi specifičnih pigmentacijskih značilnosti, močno zmanjšal krog morebitnih osumljenih, na katere bi se lahko osredotočali kriminalisti (Kayser in Schneider, 2009).

Dosedanje študije na področju pigmentacijskih značilnosti so določile že mnogo genov in njihovih SNP-jev, predvsem s pomočjo raziskanosti pigmentacijskih obolenj pri človeku (npr. ekstremne oblike albinizma). Pri pigmentacijskih obolenjih so se raziskave usmerile predvsem v mutacije gena *OCA2*, ki je bil kmalu označen kot glavni gen, ki prispeva k specifični obarvanosti oči (Branicki in sod., 2008; Sturm in sod, 2008). Nato je sledil nov val raziskav, ki je gen *HERC2* in predvsem SNP rs12913832 označil za enega pomembnejših, če ne celo najpomembnejšega, ki prispeva k specifični obarvanosti oči, las in kože (Visser in sod., 2012). Kljub vsemu pa je znano, da so pigmentacijske lastnosti zelo kompleksne in da nanje vpliva večje število različnih genov, med katerimi so bolj raziskani *MC1R, TYR, TYRP1, SLC24A4, SLC45A2, IRF4* in *ASIP* (Nakamura in sod., 2002; Branicki in sod., 2007; Soejima in Koda, 2007;).

Od vseh omenjenih smo se pri raziskavi osredotočili na šest genov in znotraj njih na dvanajst SNP-jev (*OCA2* – rs7495174, rs7170989, rs1800407, rs1667394, *TYR* – rs1393350, *SLC45A2* – rs16891982, rs26722, *MC1R* – rs1805005, rs1805008 in *SLC24A5* – rs1426654, *HERC2* – rs12913832, rs112938), ki naj bi imeli eno ključnih vlog pri pigmentaciji oči in las (Branicki in sod., 2008; Kayser in Schneider, 2009). Za pridobljene genetske podatke vseh 105 prostovoljcev smo najprej določili, da so aleli enajstih SNP-jev (razen rs1426654) v Hardy-Weinbergovem (HW) ravnovesju. Ta podatek ni presenetljiv, saj je bil izbor sorodstveno nepovezanih prostovoljcev naključen. S primerjavo alelnih frekvenc SNP-jev s podatki CEU smo ugotovili, da je SNP rs1426654 monomorfen tako za slovensko kot za evropsko populacijo. Preostalih deset SNP-jev (za SNP rs1129038 ni podanih podatkov, je pa znano, da je v popolnem vezavnem neravnovesju s SNP-jem rs12913832) kaže zelo podobno frekvenco alelov, kot je predstavljena za evropsko populacijo. Za vseh dvanajst SNP-jev smo preverili tudi njihovo vezano dedovanje (LD) in ugotovili, da sta SNP-ja rs1129038 in rs12913832 v popolnem vezavnem neravnovesju; ravno tako sta SNP-ja rs1667394 in rs7170989 v vezavnem ravnovesju s SNP-jem rs1800407. Zaradi tega smo iz nadaljnjih statističnih analiz izločili vse tri omenjene SNP-je, saj ti predstavljajo enak haplotip posameznika. Obenem smo z metodo binarne logistične regresije ugotovili, da je za naš vzorec statistično značilnih pet SNP-jev ($P < 0,05$): rs12913832, rs1393350, rs1800407, rs1805008 in rs7495174, ter le te uporabili v nadaljnih analizah za preverjanje uspešnosti napovedi za posamezno pigmentacijsko obarvanost oči in las v slovenski populaciji (omejili smo se na 104 prostovoljce, saj je bil en posameznik izvzet, kajti je edini v vzorcu slovenske populacije, ki ima rdečkasto obarvane lase). S pomočjo dveh statističnih napovednih modelov, in sicer naivnega Bayesovega modela

in modela logistične regresije smo nato ovrednotili njuno uspešnost pri napovedi obarvanosti oči in las.

Pri primerjavi obeh statističnih napovednih modelov lahko zaključimo, da sta glede na rezultate oba modela primerna, saj smo za njiju določili primerljive rezultate glede uspešnosti napovedi, kljub številčno manjši učni skupini. Znano je namreč, da je Bayesov model uspešnejši pri napovedovanju, ko imamo opravka z manjšo učno skupino, njegova slabost v primeru študije, kot je naša, pa je, da ne upošteva korelacij med posameznimi neodvisnimi spremenljivkami (Halloran, 2009). Torej v primeru obarvanosti oči ne upošteva, da posamezni SNP-ji vzajemno vplivajo na fenotipske značilnosti ljudi in da med njimi obstaja močna korelacija. Povezanost med posameznimi neodvisnimi spremenljivkami oziroma SNP-ji pa upošteva model logistične regresije. Rezultati so bili skladni s predvidevanji. Pri modelu logistične regresije smo dobili višje vrednosti AUC, občutljivosti in specifičnosti za obarvanost oči in las. To pomeni, da je model logistične regresije uspešnejši pri napovedovanju obarvanosti oči oziroma las za izbran vzorec slovenske populacije, kljub številčno majhni učni skupini. V nadaljevanju so opisane le vrednosti, pridobljene s tem modelom (Kastelic in sod., 2012).

Za obarvanost oči smo s pomočjo modela logistične regresije za 24 prostovoljcev uspešno napovedali posamezno obarvanost oči, in sicer za modre (AUC=1.0) in rjave (AUC=0.832). Za npr. modrooke smo obenem določili stoodstotno občutljivost (kar pomeni, da smo za vse modrooke pravilno napovedali, da so res modrooki) in stoodstotno specifičnost (kar pomeni, da smo za vse posameznike, ki niso imeli modrih oči, pravilno napovedali, da res niso imeli modrih oči). Za vse vmesne barve oči smo določili najnižjo vrednost AUC in obenem tudi nižjo specifičnost. Vendar te vrednosti, ki govorijo o uspešnosti napovedi, niti niso presenetljive, saj vse dosedanje raziskave z znanimi SNP-ji niso bistveno natančneje napovedale obarvanosti (Walsh in sod., 2011b; Walsh in sod., 2013) za to vmesno skupino oči. Razlog je predvsem ta, da so v to skupino zajeti vsi posamezniki, ki imajo že samo obarvanost oči bolj kompleksno in raznoliko, npr. sivo, zeleno oziroma kombinacijo več barv (rjave z zelenim kolobarjem, modre s sivim kolobarjem …). Za to skupino obarvanosti oči bi bilo v prihodnje pomembno določiti še bolj specifičen nabor SNP-jev, obenem pa bi bilo znotraj te skupine potrebno posameznike še dodatno porazdeliti v manjše sklope, kjer so si obarvanosti oči med seboj bolj podobne, kot smo to lahko naredili pri skupinah posameznikov z modrimi oziroma rjavimi očmi. Ravno tako smo za obarvanost las uspešno napovedali obarvanost, predvsem za blond lase (AUC=0.913) in temno rjave/črne lase (AUC=0.832), malo manj uspešni smo bili pri napovedi temno blond/svetlo rjavih las (AUC=0.723). Pri tako opredeljeni obarvanosti las smo največjo občutljivost (88,0 %) in specifičnost (95,0 %) določili za posameznike s temno rjavimi/črnimi lasmi. To npr. pomeni, da smo za 88,0 % posameznikov v testni skupini s temno rjavimi/črnimi lasmi, pravilno napovedali tâko obarvanost las in da smo med

vsemi posamezniki, ki niso imeli temno rjavih/črnih las, za 95,0 % od njih pravilno določili, da nimajo temno rjavih/črnih las (Kastelic in sod., 2012).

Obenem smo za uspešnost napovedi preverili še pristop, pri katerem smo posamezno obarvanost oči in las opredelili le kot svetlejšo oziroma temnejšo. Vrednosti AUC pri tej kategorizaciji sta zelo visoki, in sicer za obarvanost oči je AUC = 0.99 in pri enaki kategorizaciji las je AUC = 0.93. Vendar pa je tak način opisovanja za posamezno pigmentacijsko obarvanost manj informativen, saj imamo podatke le o svetlejšem in temnejšem odtenku obarvanosti oči, ne pa o specifični barvi, kljub temu da sta vrednosti AUC in občutljivost zelo visoki (Kastelic in sod., 2012).

Za potrebe genotipiziranja šestih SNP-jev znotraj slovenske populacije smo skonstruirali nov hkratni komplet za pomnoževanje dvanajstih SNP-jev. Komplet se je izkazal za zelo zanesljivega, saj smo za vseh 105 slovenskih prostovoljcev pravilno določili genotip vseh dvanajstih SNP-jev. Obenem pa se je izkazal za zelo občutljivega, saj je bilo pomnoževanje uspešno tudi pri zelo majhni koncentraciji začetne DNA (62.0 pg), kar je zelo primerno pri zahtevnejših forenzičnih preiskavah, kjer imamo velikokrat opravka z biološkimi sledmi z zelo nizkimi koncentracijami DNA oziroma z DNA, ki je lahko že močno razgrajena (Kastelic in sod., 2012). Vse zgoraj omenjene vrednosti, ki opredeljujejo uspešnost napovedi za pigmentacijsko obarvanost oči in las, so zelo obetavne in bodo v prihodnje, z večjo raziskanostjo še večjega števila pigmentacijskih genov in njihovih SNP-jev, imele velik pomen v forenzičnih genetskih preiskavah. Obenem pa se je potrebno zavedati, da so dosedanje študije omejene predvsem na evropsko populacijo. Za potrditev vseh dosedanjih ugotovitev glede specifičnih SNP-jev bi bile potrebne dodatne validacijske študije, ki bi potrdile oziroma zavrnile, da sklop omenjenih SNP-jev vpliva na točno opredeljeno pigmentacijsko obarvanost oči in las za vse svetovne populacije podobno. Hkrati bi bile za potrditev uspešnosti napovedi pigmentacijske obarvanosti oči in las slovenske populacije potrebne nadaljnje validacijske študije s še večjim številom prostovoljcev, predvsem na področju obarvanosti oči pri kategorizaciji z vmesno obarvanostjo oči, kjer je zajetih več barvnih odtenkov – sive, zelene oziroma kombinacija več barv (rjave z zelenim kolobarjem, modre s sivim kolobarjem …). Torej bi bilo potrebno v nadaljevanju to kategorijo obarvanosti oči razdeliti v manjše sklope, kjer bi si bili barvni odtenki oči med seboj bolj podobni, morebiti tudi s pomočjo vnaprej pripravljene barve lestvice (Kastelic in sod., 2012; Walsh in sod., 2013).

V nadaljevanju raziskav smo se zaradi manjšega vzorca slovenske populacije odločili, da uspešnost napovedi za obarvanost oči določimo tudi s pomočjo modela logistične regresije, ki je bil izdelan za 3804 prostovoljcev nizozemske populacije, in zajema šest najbolj specifičnih SNP-jev za obarvanost oči (komplet IrisPlex) (Walsh in sod., 2011a). Uspešnost napovedi, izražena z vrednostjo AUC, je bila za slovensko populacijo v tem primeru 0.966 za modre oči, 0.913 za rjave oči in 0.796 za vmesno

skupino obarvanosti oči (Kastelic in sod., 2013). Predstavljeni rezultati o uspešnosti napovedi so primerljivi z napovedmi v sklopu preostalih dveh raziskav, ki zajemata širšo evropsko populacijo (Ruiz in sod., 2012; Walsh in sod., 2013). Iz tega lahko sklepamo, da je validiran komplet IrisPlex primeren za potrebe napovedi obarvanosti oči posameznikov v forenzičnih preiskavah predvsem v primerih, kjer imamo opravka z nižjo koncentracijo oziroma degradirano DNA (Walsh in sod., 2013).

3.2 SKLEPI

V sklopu preiskav za zanesljivejše napovedovanje spola donorjev specifičnih bioloških sledi smo s pomočjo novega para začetnih oligonukleotidov gena *SRY* le pri moških osebah lahko določili končni fragment dolžine 96 baznih parov.

Validacijske študije so potrdile, da je pomnoževanje gena *SRY* z novim parom začetnih oligonukleotidov občutljivo, zanesljivo in ponovljivo, tako v smislu samostojnega pomnoževanja kot tudi v smislu pomnoževanja v hkratni PCR reakciji v sklopu komercialnih kompletov. Pomnoževanje gena *SRY* smo potrdili in validirali v sklopu komercialnega kompleta AmpFlSTR® SGM™ Plus Amplification (AB). Obenem pa smo pomnoževanje gena *SRY* potrdili tudi v komercialnem kompletu AmpFlSTR® NGM™ Amplification (AB), ki se sedaj najbolj razširjeno uporablja v rutinskih forenzičnih preiskavah in vsebuje kar šestnajst parov začetnih oligonukleotidov (priloga A). Z vnosom novega para začetnih oligonukleotidov znotraj gena *SRY* v komercialne komplete bi skrajšali čas forenzičnih preiskav in obenem zanesljivejše potrdili spol donorja specifične biološke sledi.

Z novim naborom začetnih oligonukleotidov, primernih za hkratno pomnoževanje v hkratni klasični PCR reakciji (24 parov začetnih oligonukleotidov) in v hkratni minisekvenčni SNaPshot™ reakciji (12 parov začetnih oligonukleotidov), smo pri vseh 105 slovenskih prostovoljcih uspešno določili njihov haplotip za vseh dvanajst SNP-jev hkrati. Avtomatska analiza rezultatov, pridobljenih s kapilarno elektroforezo, je potekala s pomočjo računalniškega programa GeneMapper ID version 3.2.1, pri katerem smo morali predhodno definirati velikostni standard za vsakega od dvanajstih produktov pomnoževanja. Komplet z novim naborom začetnih oligonukleotidov je obenem zanesljiv, ponovljiv (vse analize so potekale v duplikatih in podale identičen rezultat) in zelo občutljiv, saj so validacijske študije kompleta pokazale, da smo celoten haplotip SNP-jev določili že pri vnosu minimalne količine DNA (62.0 pg).

Frekvence alelov dvanajstih SNP-jev smo primerjali s frekvencami, ki so objavljene na portalu NCBI SNP, in sicer za potomce severne in zahodne Evrope – CEU (prebivalce zvezne države Utah, ZDA). Alelne frekvence SNP-jev pri slovenski populaciji so bile podobne tistim, predstavljenim na portalu NCBI SNP – CEU. Edino za SNP rs1129038

na portalu NCBI SNP ni podanih podatkov, je pa zanj znano, da je v popolnem vezavnem neravnovesju s SNP-jem rs12913832 in lahko sklepamo, da so si tudi alelne frekvence obeh SNP-jev med seboj močno podobne. Torej lahko povzamemo, da na podlagi teh frekvenc slovenska populacija pripada širši evropski kavkazijski populaciji.

V naboru dvanajstih SNP-jev smo za slovensko populacijo lahko določili pet SNP-jev (rs1800407, rs7495174, rs12913832, rs1805008 in rs1393350), ki so statistično značilni ($P<0,05$) za obarvanost oči in las. Z vključitvijo teh SNP-jev v dva statistična napovedna modela (Bayesianov model, model logistične regresije) smo lahko večjo uspešnost napovedi ugotovili pri uporabi modela logistične regresije, saj ta upošteva povezavo med neodvisnimi spremenljivkami oziroma SNP-ji, česar pa nam naivni Bayesov model ne omogoča. Obenem smo ugotovili, da se uspešnost napovedi med seboj bistveno ne razlikuje, ko obarvanost oči oziroma las kategoriziramo v tri oziroma dve skupini, vendar pa je mnogo bolj informativna kategorizacija obarvanosti v tri skupine. Pri kategorizaciji v tri skupine je bila tako uspešnost napovedi oziroma vrednost AUC 1.0 za modre oči in 0.832 za rjave oči, manjša pa za vmesno (0.747) kategorijo obarvanosti oči, ki je zajemala različne barve od zelenih do sivih. Dokaj visoke vrednosti AUC smo določili tudi za obarvanost las, vendar z manjšo občutljivostjo in specifičnostjo. Tako je bila vrednost AUC 0.913 za blond barvo las, 0.832 za temno rjavo/črno barvo las in 0.723 za temno blond/svetlo rjavo barvo las.

Zaradi manjšega vzorca slovenske populacije smo naknadno ocenili uspešnost napovedi obarvanosti oči s pomočjo šestih najbolj specifičnih SNP-jev in na osnovi statističnega modela kompleta IrisPlex izdelanega za nizozemsko populacijo (model logistične regresije). S pomočjo tega modela smo določili zelo visoko občutljivost tako za modre kot tudi za rjave oči, in sicer za celoten vzorec slovenske populacije. Obenem je bila zelo visoka tudi uspešnost napovedi, izražena z AUC (0.966 za modre, 0.913 za rjave oči) (Kastelic in sod., 2013). Zaradi kompleksnosti vmesne kategorije, ki vsebuje več različnih obarvanosti oči (sivo, zeleno oziroma kombinacijo več barv (rjave z zelenim kolobarjem, modre s sivim kolobarjem …)), pa točnost napovedi ni bila tako zanesljiva. Sta pa omenjeni vrednosti AUC za modre in rjave oči zelo primerljivi z AUC vrednostmi, ki so bile določene z dvema preiskavama za širšo evropsko populacijo, ki sta vključevali sedem (Walsh in sod., 2013) oziroma šest evropskih populacij (Ruiz in sod., 2012). Vrednosti AUC so bile v tem primeru 0.964/0.986 za modre oči in 0.956/0.978 za rjave oči. Iz tega sledi, da je validiran komplet IrisPlex, ki še ni del rutinskih forenzičnih preiskav, zelo primeren za natančno napovedovanje modrih in rjavih oči predvsem posameznikov evropske populacije.

4 POVZETEK (SUMMARY)

4.1 POVZETEK

V forenzičnih preiskavah je poleg identifikacije posameznika pomembno določiti tudi spol te osebe. S tem takoj izločimo približno polovico populacije in deloma olajšamo kriminalistom nadaljnjo preiskavo. Danes je najbolj razširjen genetski označevalec za določanje spola amelogeninski gen. Začetni oligonukleotidi za pomnoževanje gena so tako prisotni v večini splošno uporabljenih komercialnih kompletov, kljub temu da raziskovalci že od leta 1998 opozarjajo na dokaj visoko stopnjo mutacij v amelogeninskem genu (Santos in sod., 1998; Roffey in sod., 2000; Steinlechner in sod., 2002; Thangaraj in sod., 2002; Brinkman, 2002; Drobnič, 2006; Mitchell in sod., 2006; Chang in sod., 2007; Turrina in sod., 2011). Raziskovalci so začeli uporabljati nove genetske označevalce za določanje spola. Eden od njih je gen *SRY*, za katerega smo v našem raziskovalnem delu skonstruirali nov par začetnih oligonukleotidov, katerega produkt pomnoževanja je dolg le 96 baznih parov in je določen le pri moških. Gen *SRY* smo uspešno pomnožili z novim parom začetnih oligonukleotidov tudi pri standardnih pogojih, ki jih narekuje proizvajalec uporabljenih komercialnih kompletov AmpFlSTR® SGM™ Plus Amplification in AmpFlSTR® NGM™ Amplification (AB) (priloga A). Obenem smo opravili obsežnejše validacijske študije za pomnoževanje gena *SRY* z uporabo omenjenega para začetnih oligonukleotidov znotraj komercialnega kompleta AmpFlSTR® SGM™ Plus Amplification (AB) in s tem predstavili možnost, kako v krajšem času zanesljiveje napovedati spol donorja specifične biološke sledi (Kastelic in sod., 2009).

Pri zunanjem videzu človeka se seveda najprej osredotočimo na njegov spol, izrazitejše pa so tudi druge značilnosti, kot so pigmentacijska obarvanost oči, las, kože, višina posameznika in njegova struktura obraza. Na vse omenjene lastnosti po sedanjih raziskavah najbolj specifično vplivajo polimorfizmi posameznih nukleotidov (SNP-ji). Za potrebe raziskav smo odvzeli brise ustnih sluznic 105 sorodstveno nepovezanim prostovoljcem iz slovenske populacije in obenem za raziskavo pridobili potrditev slovenske Komisije za medicinsko etiko. V tej raziskavi smo se osredotočili na dve lastnosti, in sicer obarvanost oči in las, ki sta za enkrat tudi najbolj raziskani, za razliko od obarvanosti kože. Za analizo omenjenih lastnosti smo se osredotočili na šest pigmentacijskih genov in na dvanajst SNP-jev znotraj njih in jih združili v nov komplet za hkratno pomnoževanje s polimerazo. Novi komplet se je izkazal za zelo občutljivega (primeren je npr. tudi za biološke vzorce z minimalno količino DNA (62.0 pg)) in zanesljivega, saj smo za vseh 105 prostovoljcev lahko določili celoten haplotip dvanajstih SNP-jev. S primerjavo frekvenc alelov s frekvencami za širšo evropsko populacijo (podatki CEU iz portala NCBI SNP) za posamezni SNP smo potrdili, da so frekvence za slovensko populacijo primerljive in s tem potrdili njen enak izvor oziroma pripadnost širši evropski kavkazijski populaciji (Kastelic in sod. 2012).

Po predhodnih statistični analizah smo med dvanajstimi SNP-ji določili pet najbolj statistično informativnih (rs1800407, rs7495174, rs12913832, rs1805008 in rs1393350), s katerimi smo lahko za 24 prostovoljcev slovenske populacije – testno skupino (podatke preostalih 80 prostovoljcev smo uporabili za vzpostavitev statističnega modela – študijska skupina) določili vrednosti AUC 1.0 za modre oči, 0.832 za rjave oči in 0.747 za vmesno skupino obarvanosti oči, ter 0.913 za blond lase, 0.832 za temno rjavo/črno barvo las in 0.723 za temno blond/svetlo rjavo barvo las, in sicer z uporabo uspešnejšega modela logistične regresije (Kastelic in sod., 2012).

Zaradi premajhnega vzorca slovenske populacije, ki smo jo v prvem delu raziskave uporabili tako za študijsko kot tudi za testno skupino, smo v naslednjem koraku za uspešnost napovedi obarvanosti oči uporabili statistični model kompleta IrisPlex, izdelanega za nizozemsko populacijo. S tem modelom smo za vseh 105 slovenskih prostovoljcev določili vrednosti AUC 0.966 za modre oči in 0.913 za rjave oči, ki so zelo podobne vrednostim, določenimi za širšo evropsko populacijo (Ruiz in sod., 2012; Walsh in sod., 2013; Kastelic in sod., 2013).

Raziskave na področju napovedovanja pigmentacijskih značilnosti, predvsem pigmentacijski obarvanosti oči in las, so zelo obetavne in znanih je že nekaj tujih kriminalističnih primerov, ki so bili uspešno razrešeni na podlagi omenjenih analiz (Cases overview, 2012). Raziskave bodo torej služile predvsem za usmeritve kriminalističnih preiskav v zmanjšan krog morebitnih osumljencev z znanimi omenjenimi pigmentacijskimi značilnostmi, predvsem pri težjih kriminalnih dejanjih (Kayser in Schneider, 2009).

4.2 SUMMARY

In everyday forensic genetic investigations we firstly identify individual DNA profile and also his or her gender. This immediately eliminated about half of the researched population and partly facilitates further criminal investigations. Currently is the amelogenine gene the most widely used genetic marker for sex determination. Primers for amplification of this gene are present in the most commonly used commercial kits. Despite the fact that since 1998 there have been many researchers that alert on high degree of mutations (0.018% to 8.0%) in amelogenine gene on chromosome (Santos in sod., 1998; Roffey in sod., 2000; Steinlechner in sod., 2002; Thangaraj in sod., 2002; Brinkman, 2002; Drobnič, 2006; Mitchell in sod., 2006; Chang in sod., 2007; Turrina in sod., 2011). Because of that many researcher began to use new genetic markers for gender determination. One of them is *SRY* gene, that we also used in our investigation. For the region in the *SRY* gene we constructed a new pair of primers with the end amplification product of 96 bp. These new pair of primers for *SRY* gene was successfully amplified in multiplex PCR with other primers and under PCR conditions,

which are required by the manufacturer of two commercial kits AmpFlSTR® SGM^{TM} Plus Amplification in AmpFlSTR® NGM^{TM} Amplification (AB). We also conducted extensive validation studies (115 volunteers) for the amplification of SRY gene with the new pair of primers in the commercial kit AmpFlSTR® SGMTM Plus Amplification (AB) to present the opportunity of more reliable prediction of the gender of the donor for specific biological sample in shorter time (Kastelic in sod., 2009).

Beside human gender there are also other externally visible characteristics (EVC) to which we pay attention when perceiving other people. These EVCs could be pigmentation color of eyes, hair and skin, height of the individual and his facial structure. All characteristics are strong associated with specific single nucleotide polymorphisms (SNPs). For the purpose of our research the buccal swabs were taken from 105 Slovene volunteers, which all signed written consents for the use of their DNA solely for scientific research and the study was approved by National Medical Ethics Committee. In this study, we focused on two attributes, such as color of eyes and hair, which are for once the most investigated, as opposed to the color of the skin. For the analysis of these properties, we focused on six pigment genes and twelve SNPs within them and merged them into a new kit for the simultaneous amplification by polymerase. The new kit has proven to be highly sensitive (suitable, for for example even for biological samples with the minimum amount of DNA (62.0 pg)) and reliable, because we identify full haplotype for all twelve SNPs for each of 105 Slovene volunteers (all the investigations were made in duplicates). Comparing the frequencies of alleles frequencies with frequencies of wider European population (CEU data from the portal NCBI SNP) for each SNP, we confirmed that the Slovenian population has the origin like wider European Caucasian population (Kastelic in sod. 2012).

According to preliminary statistical analyzes the five out of twelve SNPs were statistically most informative (rs1800407, rs7495174, rs12913832, rs1805008 and rs1393350). With these five SNPs and for 24 Slovene volunteers (remaining 80 volunteers were used to establish a statistical model) we could determine AUC 1.0 for blue eyes, 0.832 for brown eyes and 0.747 for the intermediate group color eyes, and 0.913 for the blonde hair, 0.832 for dark brown/black hair and 0.723 for dark blonde/light brown hair color, when using statistical method of logistic regression (Kastelic in sod. 2012).

Due to the small sample size, which was in first research used to establish the statistical model, as well as for the evaluation of it, we additionally used the statistical model (IrisPlex kit) designed for the Dutch population for the more accurate prediction of the eye color for Slovene volunteers. With this model we determine AUC 0.966 for blue eyes and 0.913 for brown eyes of Slovene volunteers, which are very similar to definite AUC values for the general European population (Ruiz in sod., 2012; Walsh et al., 2013; Kastelic in sod., 2013).

Research on pigmentation characteristics, particularly pigmentation of eyes and hair color are very promising and There have already been some criminal cases, which have been successfully resolved on the basis of these analyzes (Cases overview, 2012). Research of pigmentation characteristics and many others externally visible characteristics will in near future serve mainly for guidance in criminal investigations, in order to decreased range of potential suspects with knowledge of these pigmentation characteristics or other EVCs, especially in serious crime insults (Kayser in Schneider, 2009).

5 VIRI

AmpFlSTR® SGM™ Plus Amplification. User's manual. 1997. Perkin-Elmer Applied Biosystems. Foster City (CA), The Perkin-Elmer Corporation: 20 str.

ABI PRISM®SNaPshot™ Multiplex Kit. Protocol. 2010. Foster City (CA), Applied Biosystems: 50 str.

Asplen C. 2009. ENFSI Survey on the DNA Profile Inclusion, Removal and Retention of Member Sates Forensic DNA Database. Warsaw, ENFSI: 11 str. http://www.enfsi.eu/sites/default/files/documents/enfsi_report_on_dna_legislation_in_europe_0.pdf (27. jan. 2012)

Box N.F., Wyeth J.R., O'Gorman L.E., Martin N.G., Sturm R.A. 1997. Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. Human Molecular Genetics, 6: 1891-1897

Branicki W., Brudnik U., Kupiec T., Wolanska-Nowak P., Wojas-Pelc A. 2007. Determination of phenotype associated SNPs in the MC1R gene. Journal of Forensic Science, 52, 2: 349-354

Branicki W., Brudnik U., Draus-Barini J., Kupiec T., Wojas-Pelc A. 2008. Association of the SLC45A2 gene with physiological human hair colour variation. Journal of Human Genetics, 53: 966-971

Branicki W., Liu F., van Duijn K., Draus-Barini J., Pośpiech E., Walsh S. 2011. Model–based prediction of human hair color using DNA variants. Human Genetics, 129: 443-454

Brinkmann B. 2002. Is the amelogenin sex test valid? International Journal of Legal Medicine, 116, 2: 63-63

Brookes A. J. 1999. The essence of SNPs. Gene, 234: 177-186

Budowle B. 2004. SNP typing strategies. Forensic Science International, 146: 139-142

Cadenas A.M., Regueiro M., Gayden T., Singh N., Zhivotovsky L.A., Underhill P.A., Herrera R.J. 2007. Male amelogenin dropouts: phylogenetic context, origins and implications. Forensic Science International, 166, 2-3: 155-163

Cases overview. 2012. Forensic DNAethics. Philadelphia, UPennDepartment of Medical Ethics: 1 str. http://forensicdnaethics.org/resources/cases/?start=4 (25. jan. 2012)

Chang Y.M., Perumal R., Keat P.Y., Yong R.Y., Kuehn D.L., Burgoyne L. 2007. A distinct YSTR haplotype for Amelogenin negative males characterized by a large Y(p) 11.2 (DYS458-MSY1-AMEL-Y) deletion. Forensic Science International,166: 115-120

dbSNPSummary. 2012. Bethesda,National Center forBiotechnology Information: 2 str. http://www.ncbi.nlm.nih.gov/projects/SNP (26. jun. 2012)

Dennis Lo Y.M., Tein M.S.C., Lau T.K., Haines C.J., Leung T.N., Poon P.M.K., Wainscoat J.S., Johnson P.J., Chang A.M.Z., Hjelm N.M. 1998. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for nonivasive prenetal diagnosis. Human Genetics, 62: 768-775

Drobnič K. 2004. Biološke sledi. V: Kriminalistika: uvod, taktika, tehnika. Maver D. (ur.). Ljubljana, Uradni list Republike Slovenije: 413–460

Drobnič K., 2006. A new primer set in a SRY gene forsex identification. V: Progress in forensic genetics 11: proceedings of the 21st International ISFG Congress, Ponta Delgada, Portugal, 13-16 September 2005. Amorim A., Corte-Real F., Morling N. (eds.). Amsterdam, New York, Elsevier: 269-270

Eiberg H., Troelsen J., Nielsen M., Mukkelsen A., Mengel-From J., Kjaer K.W. 2008. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Human Genetics, 123: 177-187

Excoffier L., Laval G., Schneider S. 2005. Arlequin ver. 3.0 an integrated software package for population genetics data analysis. Evolution Bioinformatics Online 1: 47-50

Fowler J.C.S., Burgoyne L.A., Scott A.C., Harding H.W. J. 1988. Repetative deoxyribonucleic acid and human genome variation – a concise review relavant to forensic biology. Journal of Forensic Science, 33, 5: 1111-1126

Giampaoli S., Chillemi G., Valerian F., Lazzaro D., Borro M., Gentile G., Simmaco M., Zanni G., Berti A., Romano Spica V. 2012. The SNPs in the human genetic blueprint era. New Biotehnology, 30, 5: 475-484

Halloran J. 2009. Classification: Naive Bayes vs Logistic regression. Honolulu, University of Hawaii at Manoa: 24 str. http://melodi.ee.washington.edu/~halloj3/classification.pdf (17. maj. 2013)

Haqq C.M., King C.Y., Donahoe P.K. 1993. SRY recognizes conserved DNA sites in sex-specific primer. Biochemistry, 90: 1097-1101

Harding R.M., Healy E., Ray A.J., Ellis N.S., Flanagan N., Todd C., Dixon C., Sajantila A., Jackson I.J., Birch-Machin M.A., Rees J.L. 2000. Evidence for variable selective pressures at MC1R. American Journal of Human Genetics, 66: 1351-1361

Haas-Rochholz H., Weiler G. 1997. Additional primer sets for an amelogenin gene PCR-based DNA-sex test. International Journal of Legal Medicine, 110: 312-315

Hedges D.J., Walker A.J., Callinan P.A., Shewale J.G., Sinha S.K., Batzer M.A. 2003. Mobile element-based assay for human gender determination. Analytical Biochemistry, 312: 77-79

Jobling M.A., Pandya A., Tyler-Smith C. 1997. The Y chromosome in forensic analysis and paternity testing. International Journal of Legal Medicine, 110: 118-124

Kanetsky P.A., Rebbeck T.R., Hummer A.J., Panossian S., Armstrong B.K., Kricker A., Marrett L.D., Millikan R.C., Gruber S.B., Culver H.A., Zanetti R., Gallagher R.P., Dwyer T., Busam K., From L., Mujumdar U., Wilcox H., Begg C.B., Berwick M. 2006. Population-based study of natural variation in the melanocortin-1 receptor gene and melanoma. Cancer Research, 66, 18: 9330-9337

Kastelic V., Budowle B., Drobnič K. 2009. Validation of SRY marker for forensic casework analysis. Journal of Forensic Science, 54, 3: 551-555

Kastelic V., Drobnič K. 2012. Single multiplex system of twelve SNPs: association of SNPs with human eye and hair colour in Slovene population. Croatian Medical Journal, 53, 5: 401-408

Kastelic V., Branicki W., Pośpiech E., Draus-Barini J., Drobnič K. 2013. Prediction of eye colour in the Slovenian population using the IrisPlex SNPs. Croatian Medical Journal, 54, 4: 381-386

Kayser M., Caglia A., Corach D., Fretwell N., Gehrig C., Graziosi G., Heidorn F., Herrmann S., Herzog B., Hidding M., Honda K., Jobling M., Krawczak M., Leim K., Meuser S., Meyer E., Oesterreich W., Pandya A., Parson W., Penacino G., Perez-Lezaun A., Piccinini A., Prinz M., Schmitt C., Schneider P.M., Szibor R., Teifel-Greding J., Weichhold G., de Knijff P., Roewer L. 1997. Evolution of Y-chromosomal STRs: a multicenter study. International Journal of Legal Medicine, 110: 134-140

Kayser M., Liu F., Janssens A. C., Rivadeneira F., Lao O., van Duijn K. 2008 Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. American Journal of Human Genetics, 82, 2: 411–23.

Kayser M., Schneider P. M. 2009. DNA-based prediction of human externally visible characteristic in forensics: motivations, scientific challenges and ethical considerations. Forensic Science International: Genetics, 3: 154-161

Kidd K.K., Pakstis A. J., Speed W.C., Grigorenko E.L., Kajuna S.L., Karoma N.J., Kungulilo S., Kim J.J., Lu R.B., Odunsi A., Okonofua F., Parnas J., Schulz L.O., Zhukova O.V., Kidd J.R. 2006. Developing a SNP panel for forensic identification of individuals. Forensic Science International, 164, 1: 20-32

Lessing R., Zoledziewska M., Fahr K., Edelmann J., Kostrzewa M., Dobosz T., Kleemann W.J. 2005. Y-SNP genotyping – a new approach in forensic analysis. Forensic Science International, 154: 128-136

Lewin B. 2000. Genes VII. 7th ed. New York, Oxford University Press: 990 str.

Liu F., van Duijn K., Vingetling J.R., Hofman A., Uitterlinden A. G., Janssens A.C.J.W. 2009. Eye color and the prediction of complex phenotype from genotypes. Current Biology, 19: 192-193

McElreavy K., Vilain E., Abbas N., Costa J.M., Souley N., Kucheria K., Boucekkine C., Thibaud E., Brauner R., Flamant F., Fellous M. 1992. XY sex reversal associated with a deletion 5' to the SRY "HMG box" in the testis-determining region. Medical Sciences, 89: 11016-11020

Mertens G. 2009. Forensic DNA Typing: Quo Vadis? The Open Forensic Science Journal, 2: 21-28

Mitchell R.J., Kreskas M., Baxter E., Buffalino L., Van Oorschot R.A. 2006. An investigation of sequence deletions of amelogenin (AMELY), a Y-chromosome locus commonly used for gender determination. Annals of Human Biology, 33: 227-240

Nakamura E., Miyamura Y., Matsunaga J., Kano Y., Dakeishi-Hara M., Tanita M., Kono M., Tomita Y. 2002. A novel mutation of the tyrosinase gene causing oculocutaneous albinism type I (OCA1). Journal of Dermatological Science, 28, 2: 102-125

Passeron T., Mantoux F., Ortonne J. 2005. Genetic disorders of pigmentation. Clinics in Dermatology, 23: 56-67

Pośpiech E., Draus-Barini J., Kupiec T., Wojas-Pelc A., Branicki W. 2012. Prediction of eye color from genetic data using Bayesian approach. Journal of Forensic Science, 57, 4: 880-886

Rees J.L. 2000. The melanocortin 1 receptor (MC1R): more than just red hair. Pigment Cell Research, 13: 135-140

Reynolds R., Varlaro J. 1996. Gender determination of forensic samples using PCR amplification of ZFX/ZFY gene sequences. Journal of Forensic Science, 41, 2: 279-286

Roffey P.E., Eckhoff C.I., Kuhl J.L. 2000. A rare mutation in the amelogenin gene and its potential investigative ramifications. Forensic Science, 45, 5: 1016-1019

Santos F.R., Pandya A., Tyler-Smith C. 1998. Reliability of DNA – based sex tests. Nature Genetics, 18, 2: 103-103

Sanchez J., Borsting C. 2003. Multiplex PCR and minisequencing of SNPs – a model with 35 Y chromosome SNPs. Forensic Science International, 137: 74-84

Singh L., Pathak N.H., Rachel A.J., Thangaraj K. 1999. Snake's eye view of Adam and Eve. Reproductive Imunology: 132-148

Sobrino B., Brion M., Carracedo A. 2005. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Science International, 154: 181-194

Smith R., Healy E., Siddiqui S., Flanagan N., Steijlen P.M., Rosdahl I., Jacques J.P., Rogers S., Turner R., Jackson I.J., Birch-Machin M.A., Rees J.L. 1998. Melanocortin 1 receptor variants in an Irish population. Journal of Investigative Dermatology, 111, 1: 119-122

Soejima M., Koda Y. 2007. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. Internal Journal of Legal Medicine, 121, 1: 36-39

Steinlechner M., Berger B., Niedestätter H. 2002. Rare failures in the amelogenin sex test. International Journal of Legal Medicine, 116: 117-120

Sturm R.A. 1998. Human pigmentation genes and theirnresponse to solar UV radiation. Mutation Research, 422, 1: 69-76

Sturm R.A., Teasdale R.D., Box N.F. 2001. Human pigmentation genes: identification, structure and consequences of polymorphic variation. Gene, 277: 49-62

Sturm R.A., Frudakis T.N. 2004. Eye colour: portals into pigmentation genes and ancestry. Trends in Genetics, 20, 8: 327-332

Sturm R.A., Duffy D.L., Zhao Z.Z., Leite F.P.N., Stark M.S., Hayward N.K. 2008. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. American Journalof Human Genetics, 82, 2: 424-431

Sulem P., Gudbjartsson D.F., Stacey S.N. 2007. Genetic determination of hair, eye and skin pigmentation in Europeans. Nature Genetics, 39, 12: 1443-1452

Sullivan K.M., Mannucci A., Kimpton C.P., Gill P. 1993. A rapid and quantitative DNA sex test: fluorescence based PCR analysis of X – Y homologous gene amelogenin. BioTechniques, 15, 4: 636-640

Tautz D., Renz M. 1984. Simple sequences are ubiquitous repetitive componentss of eukaryotic genomes. Nucleic Acids Research, 12, 10: 4127-4138

Thangaraj K., Reddy A.G., Singh L. 2002. Is the amelogenin gene reliable gender identification in forensic casework and prenatal diagnosis? International Journal of Legal Medicine, 116: 121-123

The International HapMap Consortium. 2003. International HapMap poject. Nature, 426: 789-796

Tozzo P., Giuliodori A., Corato S., Ponzano E., Rodriguez D., Caenazzo L. 2013. Deletion of amelogenin Y-locus in forensics: Literature revision and description of a novel method for sex confirmation. Journal of Forensic and Legal Medicine, 20, 5: 387-391

Turrina S., Filippini G., Voglino G., De Leo D. 2011. Two additional reports of deletion on the short arm of the Y chromosome. Forensic Science International: Genetics, 5: 242-246

Valenzuela R.K., Henderson M.S., Walsh M.H., Garrison N.A., Kelch J.T., Cohen-Barak O. 2010. Predicting phenotype from genotype: normal pigmentation. Journal of Forensic Science, 55, 2: 315-322

Velverde P., Healy E., Jackson I., Rees J.L., Thody A.J. 1995. Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. Nature Genetics, 11: 328-330

Visser M., Kayser M., Palstra R.J. 2012. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. Genome Research, 22, 3: 446-455

Walsh S., Lindenbergh A., Zuniga S.B., Sijen T., de Knijff P., Kayser M., Ballantyne K.N. 2011a. Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. Forensic Science International: Genetics, 5, 5: 464-471

Walsh S., Liu F., Ballantyne K.N., van Oven M., Lao O., Kayser M. 2011b. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Science International: Genetics, 5, 3: 170-180

Walsh S., Wollstein A., Liu F., Chakravarthy U., Rahu M., Seland J.H., Soubrane G., Tomazzoli L., Topouzis F., Vingerling J.R., Vioque J., Fletcher A.E., Ballantyne K.N., Kayser M. 2013. DNA-based eye colour prediction across Europe with the IrisPlex system. Forensic Science International: Genetics, 6: 330-340

Y chromosome. 2008. Conceptsof Genetics. Biology 3400 Genetics. Aurora,Aurora University: 1 str.
http://bio3400.nicerweb.com/Locked/media/ch07/Y_chromosome.html (26. maj 2011)

Zhu W., Zeng N., Wang N. 2010. Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS® Implementations. V: NESUG proceedings, November, 14-17, 2010 - Baltimore, Maryland.Eckler L. (ed.). Baltimore, NESUG: 1-9
http://www.nesug.org/Proceedings/nesug10/hl/hl07.pdf (3. mar. 2013)

ZAHVALA

Zahvaljujem se mentorici, prof. dr. Katji Drobnič, za strokovno vodenje.

Zahvaljujem se dr. Wojciechu Branickemu, Ewelini Pośpiech in Jolanti Draus-Barini z Inštituta forenzičnih preiskav v Krakovu za vso strokovno pomoč pri statistični obdelavi podatkov.

Še posebej hvala moji mami, ki verjame vame, in hvala za ves njen čas. Hvala tudi očiju, ki je na žalost že odšel.

Hvala sestri in vsem njenim.

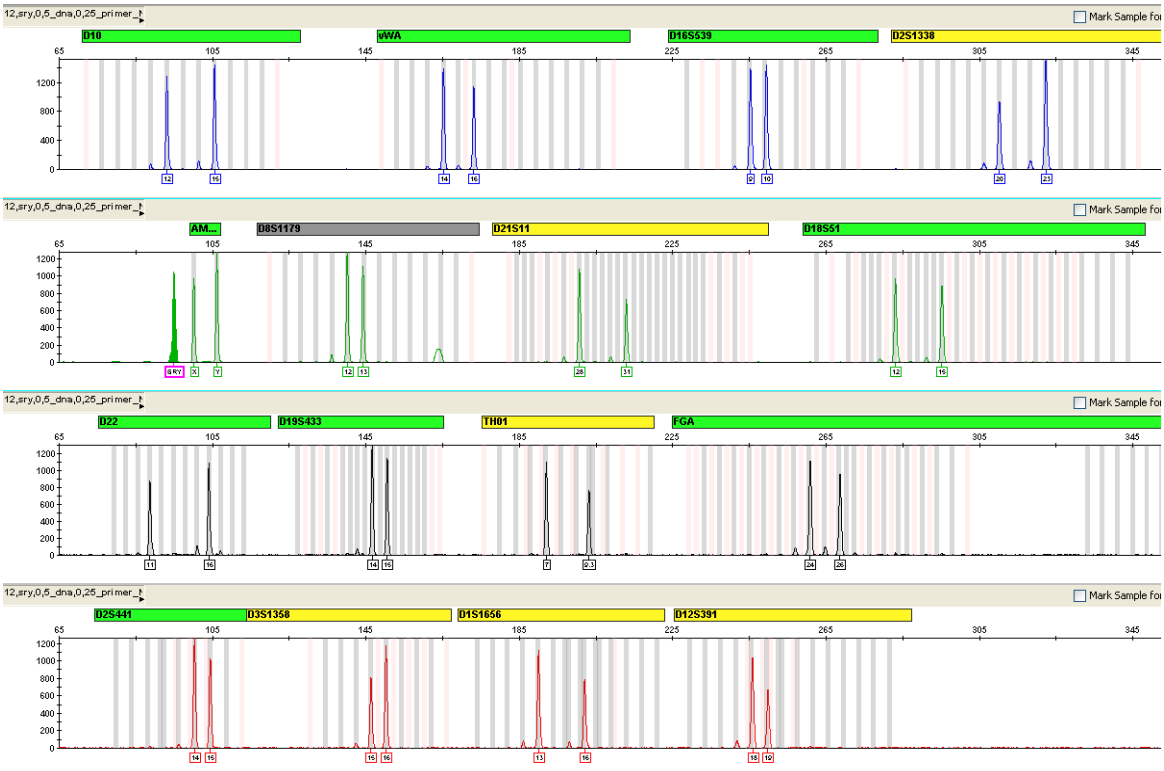
Hvala prijateljem, ki mi zapolnjujejo neprecenljivi prosti čas, in Jerci, ki si je še dodatno vzela čas zame.

Hvala mojemu Maticu in mojim trem angelčkom, Klasu, Lejli in Lorelai, ki mi izpolnjujejo življenje, in brez katerih nikakor ne bi šlo.

PRILOGA

Priloga A: Pomnoženega gena SRY (96 bp) z uporabo novega para začetnih oligonukleotidov v komercialnem kompletu AmpFlSTR® NGM™ (AB)

Leta 2010 smo v Nacionalnem forenzičnem laboratoriju začeli uporabljati nov komercialni komplet AmpFlSTR® NGM™ (AB), ki združuje 15 STR lokusov in amelogeninski lokus. Tudi za ta komercialni komplet smo preverili možnost hkratnega pomnoževanja z novimi začetnimi oligonukleotidi znotraj gena SRY. Tudi v tem primeru je bilo pomnoževanje uspešno.



Priloga A: Elektroferogram prikazuje amplifikacijski produkt pomnoženega gena SRY (96 bp) z uporabo novega para začetnih oligonukleotidov s sočasno uporabo komercialnega kompleta AmpFlSTR® NGM™ (AB) v vzorcu moške DNA.

Priloga B: Dovoljenja založnikov za objavo člankov

Validation of SRY marker for forensic casework analysis
V. Kastelic, B. Budowle in K. Drobnič
Journal of Forensic Sciences, 2009, 54, 3: 551–555

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Oct 22, 2013

This is a License Agreement between Vanja Kastelic ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3254070484283
License date	Oct 22, 2013
Licensed content publisher	John Wiley and Sons
Licensed content publication	Journal of Forensic Sciences
Licensed content title	Validation of SRY Marker for Forensic Casework Analysis
Licensed copyright line	© 2009 American Academy of Forensic Sciences
Licensed content author	Vanja Kastelic,Bruce Budowle,Katja Drobni
Licensed content date	Mar 16, 2009
Start page	551
End page	555
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Total	0.00 USD

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from

receipt

13. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

Wiley Open Access Terms and Conditions

Wiley publishes Open Access articles in both its Wiley Open Access Journals program [http://www.wileyopenaccess.com/view/index.html] and as Online Open articles in its subscription journals. The majority of Wiley Open Access Journals have adopted the Creative Commons Attribution License (CC BY) which permits the unrestricted use, distribution, reproduction, adaptation and commercial exploitation of the article in any medium. No permission is required to use the article in this way provided that the article is properly cited and other license terms are observed. A small number of Wiley Open Access journals have retained the Creative Commons Attribution Non Commercial License (CC BY-NC), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Online Open articles - Authors selecting Online Open are, unless particular exceptions apply, offered a choice of Creative Commons licenses. They may therefore select from the CC BY, the CC BY-NC and the Attribution-NoDerivatives (CC BY-NC-ND). The CC BY-NC-ND is more restrictive than the CC BY-NC as it does not permit adaptations or modifications without rights holder consent.

Wiley Open Access articles are protected by copyright and are posted to repositories and

Ostali prispevki zajeti v doktorski disertaciji so objavljeni v revijah, ki so prosto dostopne.

Kastelic V., Drobnič K. 2012. A single-nucleotide polymorphism (SNP) multiplex system: the association of five SNPs with human eye and hair colour in the Slovenian population and comparison using a Bayesian network and logistic regresion model. Croatian Medical Journal, 53, 5: 401–408
http://www.cmj.hr/2012/53/5/23100201.htm
http://www.cmj.hr/default.aspx?id=26

Kastelic V., Branicki W., Pośpiech E., Draus-Barini J., Drobnič K. 2013. Prediction of eye colour in the Slovenian population using the IrisPlex SNPs. Croatian Medical Journal, 54, 4: 381-386
http://www.cmj.hr/2013/54/4/23986280.htm
http://www.cmj.hr/default.aspx?id=26

Kastelic V., Drobnič K. 2012. Določitev zunanjega videza ljudi s preiskavami DNK. Revija za kriminalistiko in kriminologijo, 63, 3: 225–228
http://www.policija.si/images/stories/Publikacije/RKK/PDF/2012/03/RKK2012-03_Kastelic_Drobnic_DolocitevZunanjegaVidezaLjudi.pdf