

XXXI. OTDK Biológia Szekció

A humán írisz színének igazságügyi célú predikciója DNS vizsgálatok alapján

Készítette:

Erős Ágnes

ELTE-TTK, Biológiai Intézet, Genetikai Tanszék

Igazságügyi Szakértői és Kutató Intézetek, Budapesti Orvosszakértői Intézet;

DNS laboratórium

Témavezetők:

Dr. Pamzsav Horolma és Dr. Völgyi Antónia

2013.

Tartalomjegyzék

Bevezetés.....	3
Különböző szemszínek és kialakulásuk	3
Single Nucleotide Polymorphisms (SNPs)	4
Célkitűzések	5
Anyagok, módszerek és modellek.....	6
Mintavétel és feldolgozás.....	6
Kiválasztott SNP-k.....	6
A szemszínek kiértékelése.....	7
Az inhomogenitás küszöbértékének meghatározása	8
Makrók	9
Eredmények.....	10
Diszkusszió.....	13
Összefoglaló	15
Summary	16
Szerzői hozzájárulás.....	17
Felhasznált irodalom	18
Köszönetnyilvánítás	19

Bevezetés

Az emberek fenotípusos jellemzésére jól használhatók olyan tulajdonságok, mint a személyenként eltérő fokú pigmentálság. Ha megkérünk valakit, hogy jellemezzen egy másik személyt, általában az első mondatok között elhangzik az illető szem-, haj-, és esetleg bőrszíne (annak árnyalata) is. Egy-egy olyan mondat, hogy „Az illető világos bőrű, kék szemű és szőke hajú.”, vagy „Barna haja és kék szeme volt.” már sokat segít abban, hogy jobban is elképzelhessünk valakit.

Ezen tulajdonságok kialakítása mind a pigment termelésre, annak sejten belüli szabályozására vezethető vissza. A melanin a melanogenezis során alakul ki, melyben két formája különül el. Mindkettő előfordul a szemben, bőrben, és hajban is. De míg a bőrben és a hajban a melanin folyamatosan termelődik és kiválasztódik, addig az íriszben a melanociták citoplazmájában lévő melanoszómákban felhalmozódva található (Sturm és Frudakis, 2004). A fiziológiás változásoktól eltekintve bőrünk a környezeti viszonyoktól (napsütés) függően viszonylag gyorsan tud változni bizonyos korlátok között, míg hajunk színe főként az öregedés során változik meg, a melanin termelés folyamatos leállásával megőszülünk. Ellentétben e két tulajdonsággal, szemünk színe viszont viszonylag állandó marad, valamennyit fakul az évek során.

Különböző szemszínek és kialakulásuk

Az emberi szem íriszének színe nagyon polimorf fenotípust mutat. Genetikai tanulmányok alapján az alábbi kategóriákba sorolhatók: kék, szürke, zöld, mogyoróbarna, világos barna és sötét barna (Davenport és Davenport, 1907).

A szemszín kialakításáért felelős melanin pigment két alakja a sötét szemszín kialakításáért felelős eumelanin, amely barna-fekete, és a feomelanin, amely pedig vörös-sárga színű (Sturm és Frudakis, 2004). A szemszín fizikai alapjait az íriszben lévő melanociták eloszlása és tartalma határozza meg. Különböző szemszínű embereknél az íriszben lévő melanociták száma közel megegyezik, azonban a sötét szeműeknél más elhelyezkedésűek, és sokkal több melanint tartalmaznak. A kék szeműek íriszében a külső rétegekben kevés a melanin, így

mikor a fény áthalad rajtuk, több rövid hullámhosszú, kék fény szóródik szét róluk. Emiatt a kék szín főként a szerkezet következménye, és kevésbé fontosak a kémiai összetevők.

A világos (kék), és a sötét (barna) szemszín kialakítását régen egyszerű domináns-recesszív mendeli öröklődésűnek gondolták, de idővel egyre több gént írtak le, amelyek részt vesznek a szín kialakításában (Sturm és Frudakis, 2004; Mengel-From és mtsai, 2010). Valójában az írisz színe egy összetett poligénes folyamat eredménye, melyben olyan gének játszanak döntő szerepet, amelyek pl. a melanin termelésében vesznek részt (TYR, TYRP1), amelyek melanoszómális membrán asszociált transzporter fehérjék termeléséért felelősek (OCA2, MATP), vagy ezen gének szabályozói (pl. az OCA2 gént szabályozó HERC2).

A pigment termelést meghatározó génekhez hozzárendelhetők velük kapcsolatban lévő Single Nucleotide Polymorphism (SNP) lokuszok.

Single Nucleotide Polymorphisms (SNPs)

Az SNP-k DNS szekvencia-variáció (pontmutációk), amelyek során egy bázispár a genomban egy másik bázispárra cserélődik. Akkor tekinthetjük a variációt polimorfizmusnak, ha az a populáció legalább 1%-ában megjelenik. Ezek a genomban előfordulhatnak nem kódoló régiókban (intronokban vagy gének között) és kódoló régiókban is.

A legtöbb SNP lokusznak két allélja van, így az adott populációt jellemezhetjük a minor allél frekvenciájával, mely a legalacsonyabb allél gyakorisága az adott lokuszon. Különböző földrajzi-, ill. etnikai csoportok között az SNP allélgyakoriságok eltérőek lehetnek. Az egyedek SNP eltéréseinek meghatározása adja az egyed genotípusát. Ezen mutációk teszik ki a humán genetikai variációk 90%-át, kb. 1000 bázispáronként van egy ismert SNP a humán genomban (The International HapMap Consortium, 2007; Human Genome Project). Az SNP-k generációról generációra való megváltozásának valószínűsége igen alacsony (a mutációs ráta: 10^{-8} /bázis/generáció), azaz a populációkban való SNP követés lineáris következtetéseket tesz lehetővé. Az SNP-k molekuláris markerként jól használhatók az orvostudományban betegség gének térképezésére, a farmakológiában a személyek közötti genetikai variabilitások miatt személyes gyógyszerfejlesztésre, továbbá igazságügyi genetika vizsgálatokban – különösen a nem kódoló régiókban lévő mutációk – jól használhatók ún. DNS-ujjlenyomatként.

A genomi DNS-ben lévő SNP-kről egyre több információ áll rendelkezésünkre, ezért az igazságügyi genetikában új módszerek vannak kidolgozás alatt. Ezen módszerek majd a rendőrség munkáját is segíthetik, pl. ismeretlen személyek megtalálásában, vagy olyan vizsgálatoknál, mikor a minta degradálódott, és nem megfelelő a hagyományos DNS profil vizsgálatok elvégzésére. Világszerte számos genetikai laboratórium is foglalkozik humán külső jellegek becslésével a tetthelyeken talált ismeretlen biológiai mintákból (Mengel-From és mtsai, 2009). E vizsgálatok közül a szemszínnel kapcsolatos kutatások ígéretesnek tűnnek (Liu és mtsai, 2009; Walsh és mtsai, 2010). Előzetes kutatások során több, főként szemszín gének intronjaiban lévő SNP-re is kimutatták, hogy információval rendelkeznek a szemszín meghatározására. Ezek közül kiválasztották azokat a mutációkat, melyek előfordulási gyakoriságából kiindulva több, mint 90%-os pontossággal tudták jósolni a szemszínt. Ezen információkat felhasználva az ISZKI BOI DNS laboratóriuma 2010-ben elindított egy kutatást egy saját, minél pontosabb információt szolgáltatató szemszín meghatározó predikciós modell kidolgozására.

A munkám során én több, különböző matematikai megközelítéssel készült makró (SNP allélok által szemszín becslésére használható Excelbe épített program) pontosságát ellenőriztem vissza. Felhasználtam a Walsh-féle IrisPlex nevű makrót, és 4 másik, 300 magyar minta alapján multinomiális logisztikus regressziós modell-építéssel készült makrót is.

Célkitűzések

A szemszín mérésére laborunkban már kidolgozott objektív számítógépes szoftver áll rendelkezésre, a saját bevallás mellett a mintavevő állítása és a szoftver kategorizálása határozza meg egyelőre azt a négy kategóriát, amelyet még becsülni szeretnénk.

A szemszín pigment projekt során célunk volt a magyar populációra lehető legpontosabban becslő szemszín predikciós makró létrehozása. Ennek részeként kutatásom során célul tűztem ki különböző íriszszínt becslő makrók tesztelését magyar mintákkal, az így kapott eredmények összevetését, és ez alapján a legpontosabban becslő módszer kiválasztását, esetleg továbbfejlesztésének megtervezését. Ezzel párhuzamosan a már meglévő, észak-európai minták alapján összeállított szemszín becslő makró (IrisPlex – Walsh és mtsai, 2010) tesztelését is elvégeztem magyar mintákkal.

Anyagok, módszerek és modellek

Mintavétel és feldolgozás

A kutatás mintavétellel kezdődött, mely során 366 magyar önkéntestől, döntően budapesti és pécsi egyetemistától vettünk szájnyalkahártya törletet. Emellett minden személy kitöltött egy kérdőívet, melyben saját véleményüket kérdeztük színkomplexiójukról, vagyis szem-, haj- és bőrszínükről, ill. származási adataikról. A színkomplexióra utaló adatokat a labor is kitöltötte összehasonlításként. Az adatok megőrzése, és szoftveres kiértékelése céljából standard beállításokkal fénykép is készült mind a három tulajdonságról. A további vizsgálatokban felhasznált szemszíneken belül négy kategória lett elkülönítve: zöld, kék-szürke, barna és inhomogén (később részletezve).

Az adatok egy másik, párhuzamosan futó kutatásban (etnikai eredet becslése) is felhasználásra kerültek, így a mintavételt és a későbbi DNS kinyerést, pigment (SNP) meghatározást Bárány Gusztáv hallgatótársammal együtt végeztem.

A szájnyalkahártya törletből a genomi DNS kivonása és izolálása PrepFiler™ kittel történt a gyártói protokollt követve (Applied Biosystems Inc. Foster City, CA). A minták DNS koncentrációjának meghatározását ABI 7500 real-time PCR készüléken Quantifiler Human DNA Kittel a gyártói protokoll szerint végeztük (Applied Biosystems Inc. Foster City, CA).

Az SNP allélok felsokszorozását és meghatározását mintánként real-time PCR készüléken fluoreszcencia rezonancia energia átvitel (FRET) alapuló allél-specifikus oligonukleotid hibridizációs TaqMan próbával (Applied Biosystems, Foster City, CA) végeztük a gyártói protokoll szerint. A genotipizálást, vagyis az amplifikátum relatív fluoreszcencia értékeinek elemzését az ABI 7500 készülék segítségével saját SDS szoftverével készítettük.

Kiválasztott SNP-k

Irodalmi adatok alapján (Liu és mtsai, 2009, Walsh és mtsai, 2010) 20 SNP-re került a választás, melyekre a fentebb ismertetett módon meghatároztuk a 366 minta genotípusát. A kapott eredmények alapján asszociációs vizsgálatok segítségével egy matematikus az SNP-k számát tovább szűkítette, végül a Walsh és mtsai által az IrisPlex nevű makróban is használt 6

SNP-re, ezek az rs12913832, rs16891982, rs1800407, rs12896399, rs1393350 és rs12203592 lokuszok. E polimorfizmusok részletesebb tulajdonságai a 1. táblázatban láthatók. Egyszerűség kedvéért átneveztük őket és Pig. számokat kaptak.

1. táblázat:

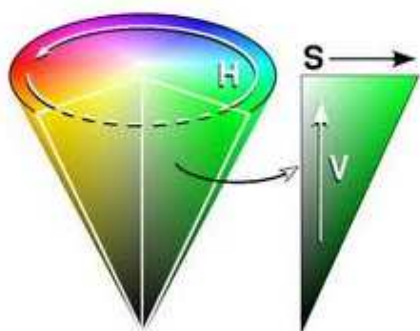
Pigment nevek	rs szám, SNP azonosító	allélok	ősi allél	minor allél	kromoszóma	gén
Pig2	rs12913832	A/G	A	A	15.	HERC2
Pig4	rs16891982	C/G	C	C	5.	SLC45A2(MATP)
Pig8	rs1800407	A/G	G	T	15.	OCA2
Pig12	rs12896399	G/T	G	T/G	14.	SLC24A4
Pig13	rs1393350	A/G	G	A	11.	TYR
Pig14	rs12203592	C/T	C	T	6.	IRF4 gén

Megjegyzés: az rs12896399-nál dán minták alapján a G, míg magyar minták alapján a T allél a minor allél.

A szemszínek kiértékelése

A szemszíneket négy kategóriába soroltam: kék/szürke, zöld, barna és inhomogén. Inhomogénnek mondjuk a kevert szemszínűeket.

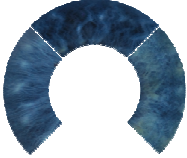
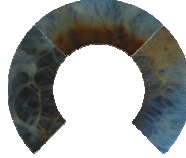
A szín meghatározására három-féle módszer alapján történt: az önkéntesek saját bemondása alapján (több színt is bejelölhettek), a mintavevő általi bemondás alapján és egy erre a célra fejlesztett számítógépes kiértékelő szoftver segítségével. Az első két módszert szubjektivitásuk és pontatlanságuk miatt elvetettem.



A szoftver (Pigment v1.0) a fénykép írisz részéből több mint 50000 pixelt felhasználva kiértékeli az íriszt: RGB (Red-Green-Blue) és HSV (Hue-Saturation-Value) értékeket számol (1.ábra), és megadja az értékek átlagát és szórását. Ezen adatok alapján adja meg a minta szemszínét. Inhomogénnek az a minta számít, aminél a Hue érték (színárnyalat) szórása nagyobb, mint 50%.

1. ábra: A HSV értékek meghatározása.

A szoftver értékelési módját az 2. ábra mutatja be.

Szín	Kék-szürke	Zöld	Barna	Inhomogén
Hue	180-240	60-180	0-40	SD>50
Saturation	<10	>10	>10	-
Kép				

2. ábra: A Pigment v1.0 szoftver értékelési alapjai

Az inhomogenitás küszöbértékének meghatározása

45 minta alapján kiválasztottam az inhomogénnek látszó szemeket (3. ábra), amely azt jelenti, hogy a szem színe „kevertnek” látszott a kép alapján, majd a makrók százalékos értékeit figyelembe véve megállapítottam az inhomogenitás küszöbértékét, vagyis a legkisebb százalékos értéket, amelynél még empirikusan megfigyelhető volt az inhomogenitás.



3. ábra: Példa inhomogén szemtípusra.

Makrók

Öt szemszín-predikciós makrót teszteltem vissza ugyanazzal a 66 mintával. Az első az IrisPlex nevű észak-európai minták alapján létrehozott 6 SNP-n alapuló makró volt (Walsh és mtsai, 2010), mely három kategóriába sorolta a szemszíneket: kék, barna és intermedier. Intermediereknek vette a kék (világos) és a barna (sötét) szemszín közötti árnyalatokat, átmeneteket. A másik négy makrót (Makró6, Makró6_fakt, Makró4, Makró4_fakt) 300 magyar minta alapján állította össze egy matematikus, aki a Matlab és R nevű programokban kiszámolta a makrók létrehozásához szükséges alfa és béta paramétereket. Az eredeti IrisPlex makróban használt számítási módszer továbbfejlesztését jelenti az ebben a dolgozatban tesztelt ún. faktoros változat, amely mind a 4 mind a 6 lokuszt tartalmazó verzióban is elkészült. Mind az öt makró ugyanazokat az SNP lokuszokat alkalmazza, ezek az rs12913832, rs1800407, rs12896399 és rs16891982 markereket a négyesek, és a fentiekén kívül a hatosok tartalmazzák még a rs1393350, rs12203592 markereket.

A makrók működésének jellemzéséhez, összehasonlításukhoz ROC-analízist használtam (Liu és mtsai, 2009). A 2. táblázat szerint az öt makróra külön-külön mintánként ellenőriztem minden kategóriára (kék/szürke, zöld, barna, inhomogén), hogy azt eltalálta-e.

3.táblázat: A becslés típusok osztályozása			
		Makró	
Szoftver		barna	nem barna
	barna	TP	FN
	nem barna	FP	TN

2. táblázat: A makrók által kapott eredmények osztályozása.

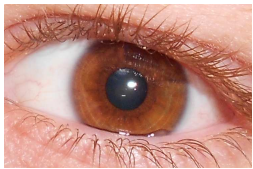
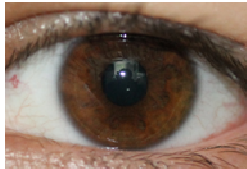
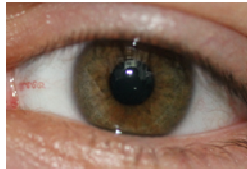
A kapott adatokból a következő értékeket számoltam ki (Liu és mtsai, 2009):

- 1) Érzékenység = $TP/(TP+FN) \times 100$. Megadja adott színre a helyes becslések százalékát a vizsgált kategórián belül.
- 2) Specifikusság = $TN/(TN+FP) \times 100$. A „nem igaz” kimenetek helyes értékelése százalékban megadva.
- 3) Pozitív jóslások értéke (PPV) = $TP/(TP+FP) \times 100$. A pozitív becslések közül a helyes jóslások százaléka.
- 4) Negatív jóslások értéke (NPV) = $TN/(TN+FN) \times 100$. A negatív becsléseken belül a helyes negatív becslések százaléka.
- 5) Pontosság (ACC) = $(TP+TN)/(TP+TN+FP+FN) \times 100$. A helyes találatok százaléka.

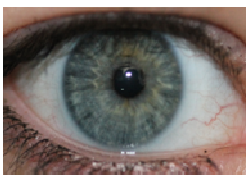
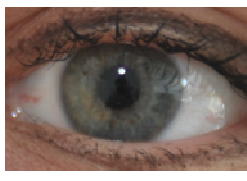
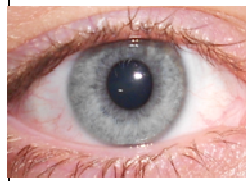
Betűkódok jelentése: TP: True Positive
 TN: True Negative
 FP: False Positive
 FN: False Negative

Eredmények

A szemszínre végzett predikciók eredményét, összevetésüket a szoftver által kapott szemszínnel az 4-5. ábrán illusztrálom, külön a kék/szürke és a barna színű szemek bemutatására.

Makró tesztelése barna szemszínre				
Mintakód	Px37	P306	P307	
Képek				
IrisPlex	kék	29,92%	0,16%	14,96%
	intermedier	25,26%	2,62%	13,98%
	barna	44,82%	97,23%	71,05%
	Predikció	TRUE	TRUE	TRUE
Makró6	zöld	2,67%	0,00%	0,76%
	szurkes-kek	5,29%	0,00%	2,76%
	inhomogén	14,44%	0,76%	31,88%
	barna	77,60%	99,23%	64,60%
	Predikció	TRUE	TRUE	TRUE
Makró6_or_fakt	zöld	5,76%	0,00%	0,70%
	szurkes-kek	4,72%	0,00%	2,39%
	inhomogén	21,74%	3,27%	27,56%
	barna	67,79%	96,73%	69,35%
	Predikció	TRUE	TRUE	TRUE
Makró4	zöld	1,97%	0,00%	0,53%
	szurkes-kek	5,21%	0,00%	3,03%
	inhomogén	12,87%	1,03%	37,49%
	barna	79,95%	98,96%	58,94%
	Predikció	TRUE	TRUE	TRUE
Makró4_or_fakt	zöld	4,31%	0,00%	0,44%
	szurkes-kek	4,71%	0,00%	2,59%
	inhomogén	21,01%	3,59%	31,26%
	barna	69,97%	96,41%	65,71%
	Predikció	TRUE	TRUE	TRUE
Szoftveres értékelés:	Barna	Barna	Barna	

4. ábra: A különböző makrók eredményei barna szemszínű mintákra.

Makró tesztelése szürkés-kék. kék szemszínre				
Mintakód	P304	Px10	Px15	
Képek				
IrisPlex	kék	93,73%	96,18%	95,00%
	intermedier	4,08%	2,49%	2,96%
	barna	2,19%	1,33%	2,04%
	Predikció	TRUE	TRUE	TRUE
Makró6	zöld	3,07%	2,82%	4,85%
	szurkes-kek	62,10%	77,33%	73,45%
	inhomogén	34,02%	19,37%	21,02%
	barna	0,81%	0,48%	0,68%
	Predikció	TRUE	TRUE	TRUE
Makró6_fakto	zöld	5,43%	0,00%	0,00%
	szurkes-kek	57,40%	77,01%	77,10%
	inhomogén	37,17%	22,99%	22,90%
	barna	0,00%	0,00%	0,00%
	Predikció	TRUE	TRUE	TRUE
Makró4	zöld	3,65%	3,37%	3,37%
	szurkes-kek	57,75%	73,79%	73,79%
	inhomogén	37,64%	22,26%	22,26%
	barna	0,97%	0,57%	0,57%
	Predikció	TRUE	TRUE	TRUE
Makró4_fakto	zöld	5,12%	0,00%	0,00%
	szurkes-kek	56,21%	75,31%	75,31%
	inhomogén	38,68%	24,69%	24,69%
	barna	0,00%	0,00%	0,00%
	Predikció	TRUE	TRUE	TRUE
Szoftveres értékelés:	Kék	Kék	Kék	

5. ábra: A különböző makrók eredményei kék szemszínű mintákra.

A makrók mind a négy (ill. három) kategóriára/színre adnak meg valószínűséget, melyek közül én azt az értéket választottam, amelyik a 4 (saját) vagy 3 (IrisPlex) közül a legmagasabb volt, ez alapján soroltam be az íriszek színét. Inhomogénnek akkor tekintettem a szemet, ha az adott írisz arra a kategóriára adta a legnagyobb valószínűségi értéket.

A visszatesztelt makrók összehasonlításának eredményeit a 3. táblázat mutatja be, a makrók közel hasonló értékkel becsülték a különböző szemszíneket.

Makró		Eredmények	Szemszín			
			barna	szürkés-kék	zöld	inhomogén
IrisPlex	B	érzékenység	89,47	100,00	0,00	0,00
		specifitás	85,71	69,09	100,00	100,00
		PPV	89,47	39,29	-	-
		NPV	85,71	100,00	92,42	81,82
		ACC	87,88	74,24	92,42	81,82
Makró 6	C	érzékenység	76,32	100,00	0,00	50,00
		specifitás	85,71	89,09	100,00	81,48
		PPV	87,88	64,71	-	37,50
		NPV	72,73	100,00	92,42	88,00
		ACC	80,30	90,91	92,42	75,76
Makró 6 faktor	D	érzékenység	78,95	90,91	0,00	50,00
		specifitás	85,71	90,91	100,00	79,63
		PPV	88,24	66,67	-	35,29
		NPV	75,00	98,04	92,42	87,76
		ACC	81,82	90,91	92,42	74,24
Makró 4	E	érzékenység	78,95	100,00	0,00	50,00
		specifitás	85,71	89,09	100,00	83,33
		PPV	88,24	64,71	-	40,00
		NPV	75,00	100,00	92,42	88,24
		ACC	81,82	90,91	92,42	77,27
Makró 4 faktor	F	érzékenység	73,68	100,00	0,00	50,00
		specifitás	85,71	89,09	100,00	79,63
		PPV	87,50	64,71	-	35,29
		NPV	70,59	100,00	92,42	87,76
		ACC	78,79	90,91	92,42	74,24

3. táblázat: a makrók eredményeinek összefoglalása (a szürke rész nem értékelhető).

A barna szemszín meghatározásában az IrisPlex makró pontosabb, a kék szemszínre viszont az általunk létrehozott makrók becsülnek jobban. Zöld szemszínű egyén kevés volt, így egyelőre ezen személyek adatai alapján nem tudtam elvégezni a kiértékelést. Nagyobb mintaszámú visszatesztelés szükséges a makró „zöld szem” meghatározásának kiértékeléséhez.

Az inhomogenitás küszöbértékei a következők:

- Makró6: 28,3%
- Makró6_fakt: 24,6%
- Makró4: 36,0%
- Makró4_fakt: 31,3%

A küszöbérték meghatározás után csak egy minta tévesen lett homogénnek becsülve, a kép alapján nem volt.

Diszkusszió

Minden tesztelt makró szemszín becslése hasonlóképpen, 70-100%-os pontossággal ($ACC=70-100\%$) adta meg a felhasznált minták szemszínét a genotípusuk alapján. Nem emelhető ki egyetlen makró sem a megbízhatóság szempontjából.

A kutatási projekt egyik célja volt a holland csoport által kifejlesztett IrisPlex makró továbbfejlesztése, egy pontosabban becselő és több kategóriát (szemszínt) használó program kifejlesztése, amely igazságügyi célra a nyomozó hatóságok segítésére használható. A létrehozott négy makró szemszín becslésére jól alkalmazható, az eddigi eredmények szerint nem becsülnek pontosabban az IrisPlex-nél, viszont több kategóriát különítenek el. Szemszínekre lebontva a barna (sötét) szemszínt pontosabban becsülte a holland csoport által kifejlesztett IrisPlex, a magyar változatoknál, a kék szemszín megállapításánál viszont a laborunkban kifejlesztett változatok bizonyultak eredményesebbnek. A visszateszteléshez használt alacsony mintaszám miatt kevés példánk volt a zöld szemszínre, így ezeket az adatokat a kiértékelés során nem tudtam használni.

Felmerül a kérdés, hogy azokban az esetekben, amikor az inhomogenitás nem dominál, hanem helyette valamelyik szín a jellemző (legmagasabb százalékos érték a makróban) viszont a szemszín nem látszik homogén kéknek, zöldnek vagy barnának, akkor hogyan válasszuk meg azt a határt, amelyik egy nem homogén, de dominánsan például kék szemszínű egyén íriszét jellemzi. 45 minta alapján kiválasztottam az inhomogénnek látszó szemeket, amely azt jelenti, hogy a szem színe „kevertnek” látszott – az íriszben több szín is felfedezhető volt – és ez alapján a makrók százalékos értékeit figyelembe véve megállapítottam az inhomogenitás küszöbértékét. Ez a 6 SNP lokuszt tartalmazó makrók esetében megközelítőleg 20, a 4 SNP lokuszt tartalmazó makrók esetében megközelítőleg 30 százalék volt. Ez azt jelenti, hogy az írisz domináns szemszíne mellett a makró hitelesen becsli meg az inhomogenitást a fenti küszöbértékeket figyelembe véve. Fontos megjegyezni, hogy a küszöbérték megállapítására felhasznált 45 személy kevés, több minta visszatesztelésére lenne szükség a pontosabb küszöb megállapításához. Az egyik minta esetében az inhomogenitás százalékos értéke a makró szerint a küszöbérték felett volt mind a 4 makró esetében, vagyis a szemszín nem volt inhomogén a képek szerint, viszont a makrók

szerint igen. Ez azt jelenti, hogy az inhomogenitás becslésében a makró tévedhet, a makrók ACC értéke szerint ez átlagosan 75%, a küszöbérték megállapítása után kiszámolt érték pedig 98% volt. A fentiek miatt feltételezhetjük, hogy a küszöbértéket helyesen határoztam meg, viszont a szemszín „reprodukciója” csak a genotípus alapján némi tévedéssel működik az inhomogenitás tekintetében. Meg kell jegyezni, hogy barna szemszín esetében az inhomogenitás – bár jelen lehet – nem látszik, mivel a fényt elnyeli a szem. Elképzelhető, hogy számítógépes eljárással a fotókból az inhomogenitást elő lehet hívni, ennek kidolgozása is terveink szerint szerepel.

Mind a makrók pontosítására, mind a visszatesztelés szempontjából cél a mintaszám növelése. Az inhomogenitás pontosabb küszöbértékének meghatározásához is fontos lenne a tesztelendő minták számának növelése. Folyamatban van a módszerre vonatkozó AUC érték számolása ROC-analízissel, amellyel a becslés bizonytalanságára vonatkozó információt kapjuk meg.

Kevert szemszín esetén az IrisPlex makró nem tudja meghatározni, hogy a szín inhomogén-e, ezért a makró tévedésének az elméleti valószínűsége egy inhomogén szem esetében nagyobb. Gyakorlatban, a szemszín inhomogenitás nélküli kategóriákba sorolásnál azonban kiderült, hogy az IrisPlex nem téved többet, mint az általam tesztelt többi makró, viszont több információval szolgál. Terveink között szerepel a szemszín számítógépes rekonstrukciója genotípus alapján, eszerint egy általunk létrehozott és tesztelt szoftver virtuálisan létrehozná a szemet. Ebben fontos szerepe van az inhomogenitásnak, hiszen másképp képzelünk el egy inhomogén, majdnem „foltos” kék és barna íriszt, mint egy homogén kéket vagy barnát. Jóval „természetesebb” is a variációk virtuális generálásának eredménye is. Elképzelésünk szerint a nyomozó hatóságok felé a biológiai nyom valószínűsített szemszín százaléka valamint a meghatározás bizonytalansága mellett egy virtuális szemet is lehetne küldeni, ez alapján a keresett személy beazonosítása egyszerűbb lenne.

A faktorozás a matematikában a regresszió finomítására szolgál, a különbség a normál logisztikus regressziótól az, hogy az egyes SNP lokuszokat nem számként értelmezi, hanem minőségi változóként. Az elképzelés az volt, hogy a fentiek miatt a faktoros változatok jobban becslik majd a kategóriákat. A tesztelés eredménye szerint a faktorizálás nem változtatott a makró eredményességén, jelenleg ugyanolyan hatékonyan működik a 4 és a 6 SNP lokuszt tartalmazó faktoros változat, mint a nem faktoros párja. Mivel a faktorizálás nem segített jobb makrókat alkotni, valószínűleg a gyakorlatban a nem faktoros változatokat fogjuk alkalmazni, mivel a holland csoport által megalkotott makróval jobban összevethető, elméleti és gyakorlati szinten jobban magyarázható.

Összefoglaló

Az emberek közti biológiai variabilitás egyik jellemzője a szem-, haj- és bőr pigmentáltsága. A pigmenttermelést meghatározó génekhez hozzárendelhetők velük kapcsoltásban lévő Single Nucleotide Polymorphism (SNP) lokuszok. Az ISZKI Budapesti Orvosszakértői Intézet DNS laboratóriuma elindított egy projektet, amelynek célja a szemszín becsléséhez szükséges modellépítés, valamint predikciós makró kidolgozása volt. A projektben 300 magyar személy szemszínét határoztuk meg egy erre a célra kidolgozott értékelő szoftver segítségével. A pigmentáltsághoz való kapcsoltáguk alapján publikált adatokból 20 SNP lokusz került kiválasztásra, asszociációs tulajdonságuk alapján pedig a predikciós makró felállítására az alkalmazott SNP lokuszok számát maximum 6-ra szűkítettük. A Walsh-féle IrisPlex nevű makrót, és 4 másik, 300 magyar minta alapján multinomiális logisztikus regressziós modell-építéssel készült különböző összetételű és matematikai összefüggéseken alapuló makrókat 66 személyen teszteltem vissza. A saját makrók több kategóriát különítenek el, mint az IrisPlex, újdonság volt bennük az inhomogenitás figyelembevétele. Megfigyeléseink alapján a szem már 20%-os (6 SNP), 30%-os (4 SNP) inhomogenitás érték felett inhomogénnek látható.

A visszatesztelés eredménye szerint a makrók különböző mértékű pontossággal becsülték a szemszín.

Az IrisPlex a barna, a saját makrók a kék színt becsülték pontosabban. A zöld szemszínű minták alacsony reprezentáltsága miatt, a visszatesztelés erre a színre nem valósult meg. A faktoros és nem faktoros változatok között nem tapasztaltam jelentős különbséget.

A vizsgálat a hazai igazságügyi genetikai vizsgálati palettát bővíti, segítve ezzel a nyomozást ismeretlen eredetű humán biológiai minták tulajdonosainak szemszíneire vonatkozólag.

Summary

Biological variability among people can be described – among others - by the pigmentation of the eye, the hair and the skin. Genes determining pigmentation can be assigned to Single Nucleotide Polymorphism (SNP) loci. The DNA laboratory at the Institutes for Forensic Sciences, Institute of Forensic Medicine, Budapest has started a project aiming at building models to estimate eye colour and developing prediction models (so called macros). The project has determined the eye colour of 300 Hungarian people with the help of an evaluating software designed for this reason. 20 SNP loci have been chosen based on published data about their linkage to pigmentation. In order to set up the prediction model, we narrowed down the number of SNP loci to the maximum of 6, according to their associative features. I tested the following models on 66 people: an international prediction model entitled Walsh IrisPlex, and 4 other models with different compositions and mathematical correlations, which were based on multinomial logistic regression models from 300 Hungarian samples.

Our own models distinguish more categories than the IrisPlex. Furthermore, the novelty of them is that they take inhomogeneity into consideration. Based on our observations, the eye can be seen as inhomogeneous above 20% (with 6 SNP) and 30% (with 4 SNP) inhomogeneity value. The results of the testing suggest that the different models predicted eye colour with different accuracy. The IrisPlex estimated the brown colour, whereas our macros estimated the blue colour the most precisely. Due to the low representation of samples originating from green-eyed people, testing on this colour was not possible. I have not observed significant difference between varieties with different mathematical correlations. Our research provides greater opportunities for future genetic analysis and it helps determine the eye colour of people with samples of unknown human biological origin during investigation.

Szerzői hozzájárulás

1. Részt vettem 366 önkéntes mintaadó DNS mintájának izolálásában és feldolgozásában, genotipizálásában.
2. Kiértékeltem a szemszíneket a fényképek alapján egy számítógépes szoftverrel.
3. 66 mintával elvégeztem az öt makró visszatesztelését.
4. Kiértékeltem a makrók eredményeit, összevettem a szoftver által megadott szemszínekkel, majd összehasonlítottam a makrókat egymással.

Felhasznált irodalom

1. **Sturm** AR. and Frudakis TN. (2004) Eye colour: portals into pigmentation genes and ancestry. *Trends in Genetics* Vol. 20
2. **Davenport**, G.C. and Davenport, C.B. (1907). Heredity of eye-color in man. *Science* 26, 590–592
3. **Eiberg** H, Troelsen J., Nielsen M., Mikkelsen A., Mengel-From J., Kjaer K. W., Hansen L. (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. *Hum Genet* Vol. 123, 177-187
4. **Mengel-From** J., Břrstring C., Sanchez J. J., Eiberg H., Morling N. (2010) Human eye colour and HERC2, OCA2 and MATP. *Forensic Science International: Genetics* Vol. 4 323-328
5. **Liu** F., Duijn K. van, ívingerling J. R., Hofman A., Uitterlinden A. G., Janssens A. C. J. W., Kayser M. (2009) *Current Biology* Vol 19. No 5
6. **Walsh** S., Liu F., Ballantyne KN., Oven M., Lao O., Kayser M. (2010) IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information *Forensic Science International: Genetics*; 572
7. **Mengel-From†** J., Wong T: H., Morling N., Rees J. L., Jackson I. J. (2009) Genetic determinants of hair and eye colours in the Scottish and Danish populations
8. The International HapMap Consortium, 2007; *Nature* Vol 449 doi:10.1038/nature06258

Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőimnek, Dr. Pamzsav Horolmának, a Budapesti Orvosszakértői Intézet DNS laboratórium vezetőjének, és Dr. Völgyi Antónia igazságügyi-genetikus szakértőnek, hogy lehetővé tették számomra a laboratóriumi munka elvégzését, a háttérből segítettek munkámat, és bátran fordulhattam hozzájuk szakmai kérdésekkel az adatok feldolgozásakor és a dolgozat megírása során is.

Szeretnék még köszönetet mondani Dr. Mádi-Nagy Gergely matematikusnak a makrók elkészítéséért, és Sándor Gábor informatikusnak a Pigment v1.0 szoftver kifejlesztéséért.

Köszönöm hallgatótársam, Bárány Gusztáv együttműködését a rengeteg minta feldolgozásakor.

Köszönettel tartozom a labor többi dolgozójának is, akikhez mindig fordulhattam apróbb-cseprőbb kérdéseimmel a minták feldolgozásakor.

Ezen kívül köszönettel tartozom azoknak az önkéntes embereknek, akiknek a mintáit felhasználhattam a kutatásomban, az ő hozzájárulásuk nélkül ez a dolgozat nem jöhetett volna létre.

Végül pedig köszönettel tartozom családomnak, és barátaimnak is, akik a háttérből segítettek munkámat.