

FACULDADE DE BIOCÊNCIAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR  
DOUTORADO EM BIOLOGIA CELULAR E MOLECULAR

**FERNANDA ROSA SAWITZKI**

**FENOTIPAGEM POR DNA - VARIANTES EM GENES HUMANOS QUE REGULAM A  
PIGMENTAÇÃO DE OLHOS E DE PELE: ANÁLISE FENOTÍPICA E GENOTÍPICA DE INDIVÍDUOS  
SULBRASILEIROS PARA FINS FORENSES**

Porto Alegre  
2017

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica  
do Rio Grande do Sul

FERNANDA ROSA SAWITZKI

**FENOTIPAGEM POR DNA - VARIANTES EM GENES HUMANOS QUE REGULAM A  
PIGMENTAÇÃO DE OLHOS E DE PELE: ANÁLISE FENOTÍPICA E GENOTÍPICA DE INDIVÍDUOS  
SULBRASILEIROS PARA FINS FORENSES**

Tese apresentada como requisito para a  
obtenção do grau de Doutor pelo  
Programa de Pós-Graduação em Biologia  
Celular e Molecular da Faculdade de  
Biotecnologia da Pontifícia Universidade  
Católica do Rio Grande do Sul

ORIENTADORA: PROFA. DRA. CLARICE SAMPAIO ALHO

Porto Alegre  
2017

FERNANDA ROSA SAWITZKI

Tese apresentada como requisito para a  
obtenção do grau de Doutor pelo  
Programa de Pós-Graduação em Biologia  
Celular e Molecular da Faculdade de  
Biotecnologia da Pontifícia Universidade  
Católica do Rio Grande do Sul

Aprovada em: 07 de novembro de 2017

BANCA EXAMINADORA:

---

Profa. Regina Maria Barreto Cicarelli (UNESP)

---

Dr. Rafael Scorsatto Ortiz (Perito DPF)

---

Prof. Maurício Reis Bogo (PPGBCM/PUCRS)

Porto Alegre  
2017

## Ficha Catalográfica

S271f Sawitzki, Fernanda Rosa

Fenotipagem por DNA - Variantes de genes humanos que regulam a pigmentação de olhos e pele : Análise fenotípica e genotípica de indivíduos sulbrasilieiros para fins forenses / Fernanda Rosa Sawitzki . – 2017.

79 f.

Tese (Doutorado) – Programa de Pós-Graduação em Biologia Celular e Molecular, PUCRS.

Orientadora: Profa. Dra. Clarice sampaio Alho.

1. Forense. 2. Pigmentação Humana. 3. Fenotipagem. 4. SNP. I. Alho, Clarice sampaio. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS  
com os dados fornecidos pelo(a) autor(a).

Bibliotecário responsável: Marcelo Votto Texeira CRB-10/1974

## **AGRADECIMENTOS**

Primeiramente, gostaria de agradecer à minha orientadora Prof<sup>a</sup>. Dr<sup>a</sup>. Clarice Sampaio Alho, pela confiança depositada em mim e por nossa parceria durante os últimos oito anos.

Aos queridos colegas e amigos do Laboratório de Genética Humana e Molecular por toda a compreensão, apoio e principalmente por toda a amizade e conversas durante os mais diversos momentos da minha trajetória acadêmica e pessoal.

Agradeço especialmente minha família, que sempre foi a base de tudo me acompanhando e incentivando o meu crescimento profissional, pelo carinho incondicional e por nunca me deixar desistir nos momentos de maior dificuldade.

## APRESENTAÇÃO

Um ramo da genética forense promissor para auxiliar a investigação policial identificando vítimas e/ou suspeitos é a predição de características externas visíveis (CEV) via o estudo do DNA. A análise de genes direta e indiretamente envolvidos no processo de síntese dos pigmentos eumelanina (castanho-preto) e feomelanina (vermelho-amarelo) tem poder preditivo, e serviria para inferir fenótipo de coloração de pele e olhos do doador de amostras biológicas de identificação desconhecida.

Na pele, em maior ou menor quantidade, a melanina é produzida nos melanócitos (camada basal da epiderme) e depositada nos queratinócitos (onde impede lesões no DNA causadas pela luz ultravioleta). Nos olhos, a melanina é produzida pelos melanócitos do epitélio pigmentar da íris, depositando-se no estroma da íris cujas variações de densidade resultam também nas diferenças de cor.

Avaliando amostras de indivíduos com pele e olhos claros ou pele e olhos escuros, em indivíduos de ancestralidade europeia e africana, respectivamente, se têm encontrado associações positivas entre a herança genética e a manifestação da pigmentação. Tais estudos associativos entre genótipo e o fenótipo de coloração, contudo, têm dificuldade em isolar o *background* étnico e, por não ter sido ainda completamente elucidada a implicação efetiva no fenótipo de cada variante alélica, alguns alelos ora indicados como geradores de fenótipo claro ou escuro podem estar sendo apenas marcadores de ancestralidade. Uma estratégia para contornar essa dificuldade é realizar tais estudos em populações com indivíduos claros e indivíduos escuros, tendo ambos um grau de mistura étnica estabelecida em decorrência do histórico de ocupações geográficas feitas por imigrantes de origens contrastantes. No Brasil esse perfil pode ser encontrado, sendo que indivíduos mesmo apresentando o fenótipo claro ou o fenótipo escuro têm, em geral, grau de miscigenação maior do que aqueles pertencentes a populações europeias e africanas de referência.

Com o propósito da predição de CEV, estabeleceu-se no laboratório de Genética Humana e Molecular da PUCRS um painel de marcadores moleculares presentes em sete genes humanos (*HERC2*, *OCA2*, *SLC45A2*, *SLC24A5*, *TYR*, *TYRP1* e *MC1R*), os quais estão diretamente envolvidos com a síntese de melanina. Construiu-se um sistema de genotipagem por SNaPshot que inclui oito SNPs (rs12913832, rs4778138, rs1426654, rs16891982, rs2733832, rs8045560; rs1042602, rs916977) presentes nestes genes, cujas variantes alélicas foram fortemente associadas à presença ou não de melanina na pele e no olho de indivíduos residentes em populações europeias e africanas já estudadas.

O presente trabalho teve como objetivo analisar quanti e qualitativamente o grau de pigmentação de olho (íris) e de pele de indivíduos sulbrasilieiros (habitantes dos estados do Rio Grande do Sul, Santa Catarina e Paraná), na intenção de identificar as possíveis relações dessas variáveis com os alelos dos genes *HERC2*, *OCA2*, *SLC45A2*, *SLC24A5*, *TYR*, *TYRP1* e *MC1R*, os quais estão envolvidos diretamente na síntese de melanina. O propósito final do estudo é gerar dados, metodologias de análise e protocolos laboratoriais capazes de auxiliar as unidades de perícia brasileiras a prever o fenótipo de pigmentação nas investigações forenses.

## RESUMO

Os genes relacionados a características externamente visíveis (EVC) apresentam grande importância para a previsão de cores de pele e olho em humanos. Neste estudo, testamos a capacidade de um conjunto de SNPs em genes relacionados à síntese de pigmento melanina para prever cores de pele e olho em uma população sul-brasileira. Inicialmente, utilizamos o modelo Cinderela - Tiana - Neve Branca para analisar a distribuição alélica de oito SNPs bialélicos associados à produção da melanina. Os padrões Cinderela e Tiana possuem respectivamente baixo e alto conteúdo de melanina na pele e nos olhos; Cinderela tem pele branca e olhos azuis (baixo conteúdo de melanina, LMC) e Tiana tem pele e olhos escuros (alto conteúdo de melanina, HMC). A investigação comparativa entre frequências de variantes genéticas em indivíduos com Cinderela-Like versus Tiana-Like pode indicar qual variante polimórfica está associada à síntese de melanina na pele e nos olhos. De forma coordenada, os estudos com indivíduos Branca de Neve-Like podem ser informativos para revelar a expressão tecido-específica, uma vez que esses indivíduos têm pele branca (LMC) e olhos escuros (HMC). Com base em frequências de alelos de diferentes populações humanas, o nome alelo "L" foi utilizado para os alelos associados a populações de baixo conteúdo de melanina (indivíduos Cinderela-Like) e o nome alelo "H" foi utilizado para os alelos associados a populações de alto conteúdo de melanina (Indivíduos Tiana-Like). A distribuição alélica de oito SNPs mostrou que 100% de indivíduos com Cinderela-Like (N= 73) tinham menos de oito alelos H e 82% de indivíduos Tiana-Like (N= 61) tinham oito ou mais alelos H. O valor AUC (*Area Under the Receiver Operating Characteristic Curve*) foi de 0,99 e o cálculo de PGL (*Pathway Genetic Load*) e GP (Probabilidade Genética) mostrou que o conjunto de SNP apresentou concordância de 93% e 91% entre genótipo e fenótipo, respectivamente. As análises discriminantes fatoriais (FDA) realizadas no grupo Branca de Neve-Like (pele clara e olhos escuros; N= 116) mostraram associação positiva entre SNPs rs16891982 (SLC45A2), rs8045560 (MC1R), rs1426654 (SLC24A5), rs2733832 (TYRP1) e rs1042602 (TYR) e o cluster de Cinderela para o fenótipo da pele, e associação positiva entre SNPs rs4778138 (OCA2), rs12913832 (HERC2) e rs916977 (HERC2) e o cluster Tiana para o fenótipo dos olhos. Por fim, estudamos 436 sujeitos sul-brasileiros com diferentes cores de pele e de olhos para construir um modelo de regressão logística multinomial capaz de prever o fenótipo. O modelo foi testado em 40 indivíduos aleatórios para medir a eficiência da predição. Os dados apresentaram concordância de 93% entre o fenótipo previsto e observado, mostrando a previsão fenotípica bem-sucedida do modelo em indivíduos sul-brasileiros. Isso é importante, uma vez que o Brasil tem uma população etnicamente mista, na qual a estrutura genômica pode favorecer os efeitos epistáticos e pleiotrópicos não observados em populações mais homogêneas. As análises aqui apresentadas são uma importante contribuição para a área da fenotipagem por DNA.

## PALAVRAS-CHAVE

Forense; Pigmentação Humana; Fenotipagem; SNP



## ABSTRACT

Genes related to externally visible characteristics (EVC) present a great importance for the prediction of skin and eye colors in humans. In this study, we tested the ability of a set of SNPs in pigment-related genes to predict skin and eye colors in an admixed south Brazilian population. First, we used the Cinderella - Tiana - Snow White model to analyze the allelic distribution of eight biallelic SNPs in pigment-related. Cinderella and Tiana patterns have respectively low and high melanin content in skin and eyes; Cinderella has white skin and blue eyes (low melanin content; LMC) and Tiana has dark skin and eyes (high melanin content; HMC). Comparative investigation between frequencies of genetic variants in Cinderella-Like *versus* Tiana-Like subjects may indicate which polymorphic variant is associated with melanin synthesis in skin and eyes. Coordinately, studies with Snow White-Like subjects may be informative to reveal any tissue-specific expression, since these individuals have both white skin (LMC) and dark eyes (HMC). Based on allele frequencies of different human populations, allele "L" was used for the alleles associated with low melanin content populations (Cinderella-like subjects), and allele "H" was used for the alleles associated with high melanin content populations (Tiana-like subjects). Allelic distribution of eight SNPs showed that 100% of Cinderella-like subjects (N= 73) had less than eight H alleles, and 82% of Tiana-Like subjects (N= 61) had eight or more H alleles. The AUC (Area Under the Receiver Operating Characteristic Curve) value was 0.99, and the calculation of PGL (Pathway Genetic Load) and GP (Genetic Probability) showed that the SNP set presented 93% and 91% concordance between DNA genotype and phenotypes, respectively. Factorial discriminant analyses (FDA) performed in the Snow White group (light skin and dark eyes; N= 116) showed an association between SNPs rs16891982 (SLC45A2), rs8045560 (MC1R), rs1426654 (SLC24A5), rs2733832 (TYRP1), and rs1042602 (TYR) and the Cinderella cluster for skin phenotype, and an association between SNPs rs4778138 (OCA2), rs12913832 (HERC2), and rs916977 (HERC2) and the Tiana cluster for eye phenotype. Lastly, we studied 436 South Brazilian subjects with different skin and eye colors to construct a multinomial logistic regression model able to predict phenotype. The model was tested in 40 random subjects to measure the efficiency of the prediction. The data presented 93% concordance between the predicted and the observed phenotype, showing the successful phenotype prediction of the model in South Brazilian individuals. This is important since Brazil has an ethnically mixed population in which genomic structure may favor epistatic and pleiotropic effects not observed in populations that are more homogeneous. The analyses presented here are an important contribution to forensic DNA phenotyping scenario.

## KEYWORDS

Forensic; Human Pigmentation; Phenotype; SNaPshot; SNP

## ABREVIATURAS

AIM: do inglês, *Ancestry Informative Marker*

ASIP: do inglês, *Agouti signaling protein*

AUC : do inglês, *Area Under the Receiver Operating Characteristic Curve*

CEV: Características Externas Visíveis

CREB : do inglês, *cyclicAMP response-element binding protein*

DQ : dopaquinona

EVT: Traços externamente visíveis; do inglês *externally visible traits*

FDP: fenotipagem de DNA forense; do inglês *Forensic DNA Phenotyping*

GP : do inglês, *Genetic Probability*

HERC2: do inglês, *HECT domain and RCC1-Like domain-containing protein 2*

HMC: Alto conteúdo de melanina; do inglês *High Melanin Content*

LMC: Baixo conteúdo de melanina; do inglês *Low Melanin Content*

MATP: do inglês, *Membrane associated transporter protein*

MC1R: do inglês, *melanocortin 1 receptor*

MITF : do ingles, *Microphthalmia-associated Transcription Factor*

OCA2: do inglês, *Oculocutaneous albinism type II*

PCR: do inglês, *Polymerase Chain Reaction*

PGL : do inglês *Pathway Genetic Load*

SLC24A5 : do inglês, *Solute Carrier Family 24, member 5*

SLC45A2: do inglês, *Solute Carrier Family 45, member 2*

SNP: do inglês, *Single Nucleotide Polymorphism*

STR: do inglês, *Short of Tandem Repeat*

TYR: do inglês, *Tyrosinase*

TYRP1: do inglês, *Tyrosinase-related protein 1*

UV: ultravioleta

VNTR: do inglês, *Variable Number of Tandem Repeat*

## SUMÁRIO

1- REFERENCIAL TEÓRICO	01
1.1- Predição de Fenótipo	03
1.2- Biologia da Pigmentação	04
1.2.1- Pigmentação da Pele	05
1.2.3- Pigmentação do Olho	06
1.2.4- Pigmentação do Cabelo	08
1.2.5- Melanogênese	08
Eumelanogênese	09
Feomelanogênese	09
1.3- Pigmentação e Ancestralidade	10
1.4- Genes e SNPs	12
1.4.1- Gene <i>HERC2</i> , SNPs rs916977 – rs12913832	14
1.4.2- Gene <i>OCA2</i> , SNP rs4778138	15
1.4.3- Gene <i>SLC24A5</i> , SNP rs1426654	16
1.4.4- Gene <i>SLC24A5</i> , SNP rs16891982	16
1.4.5- Gene <i>TYR</i> , SNP rs1042602	17
1.4.6- Gene: <i>TYRP1</i> , SNP rs2733832	18
1.4.7- Gene: <i>MC1R</i> , SNP: rs8045560	18
1.5- Frequências alélicas dos SNPs em populações Africanas e Europeias	19
2- OBJETIVOS	20
3- BIBLIOGRAFIA	21
ARTIGO CIENTÍFICO 1	25
ARTIGO CIENTÍFICO 2	43
CONSIDERAÇÕES FINAIS	66
ANEXOS	67

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18

**CAPÍTULO 1**  
**REFERENCIAL TEÓRICO E OBJETIVO**

## 1- REFERENCIAL TEÓRICO

Historicamente identificação humana foi baseada, na maioria dos casos, em métodos provenientes das áreas de antropologia forense, odontologia forense, radiologia e biologia forense (Alonso *et al.*, 2005). As principais características que auxiliam na identificação ou exclusão individual, são as morfologias do esqueleto, arcada dentária, impressões datiloscópicas, cicatrizes, implantes médicos e obviamente o perfil de DNA (Budowle *et al.*, 2005). Os constantes avanços na área de biológica molecular vêm aumentando a variedade de marcadores disponíveis para fins forenses. Apesar do genoma ser 99.9% compartilhado entre os seres humanos, existem variações em suas sequências, que podem ser usados tanto para individualizar quanto para correlacionar indivíduos. Ao longo dos anos foram implantadas diferentes técnicas a fim de caracterizar e individualizar pessoas, com estudos envolvendo marcadores genéticos e bioquímicos (STRs, VNTRs, polimorfismos dos sistemas HLA, ABO e de outras proteínas séricas) (Kobachuk, 2012).

A análise de marcadores polimórficos de DNA é atualmente reconhecida mundialmente como uma técnica forense padrão para a investigação de uma ampla variedade de crimes. O material biológico coletado no local de crime é processado para obtenção do perfil de DNA que, ao ser comparado com o perfil de DNA de suspeitos, pode auxiliar no estabelecimento de uma conexão entre o criminoso e o local de crime, ou ainda, na eliminação de suspeitos durante as investigações (Bond, 2007). Conjuntos autossômicos, altamente polimórficos conhecidos como STRs (*Short tandem repeat*) são marcadores moleculares utilizados na identificação humana em casos forenses há muitos anos (Jobling *et al.*, 2004) e, na atualidade, o uso dos marcadores STR é muito vasto no sistema forense, sendo aplicado por exemplo, nos casos de identificação criminal e em testes de paternidade ou parentesco (Yoshida *et al.*, 2011). A confiabilidade elevada dos resultados de análises de STRs permitem que tais análises sejam incorporadas como provas nos processos tanto da esfera penal quanto da esfera cível. Além disto, análises conjuntas de marcadores moleculares do tipo SNPs (*single nucleotide polymorphism*) podem também auxiliar na identificação humana, pois permitem a obtenção de perfis genômicos únicos para cada indivíduo (Kobachuk, 2012). Nos casos em que a genotipagem convencional (STR) não pode auxiliar na identificação, seria esperado que predições genéticas de características externamente visíveis (CEV) de um indivíduo pudessem ser de grande valia nas investigações e rastreamento de suspeitos. Avanços recentes na genética têm identificado uma gama de marcadores, do tipo SNP, úteis para a predição dessas características (Kayser e Schneider, 2009; Kayser, 2015).

Os SNPs, são polimorfismos que representam a classe mais abundante em humanos, eles podem servir tanto para a discriminação individual como para a predição CEV. Sua

1 abundância é a principal razão do grande interesse atual no campo forense; e o mapeamento  
2 do genoma humano tornou possível o desenvolvimento de mapas de genótipos e de haplótipos  
3 que podem caracterizar fenotipicamente um ser humano (Stacey *et al.*, 2002).

4 Entre as CEVs, a variação de pigmentação em humanos tem se tornado alvo de  
5 pesquisas. Mutações polimórficas do tipo SNPs, podem determinar substituições de  
6 aminoácidos na proteína, alterando as propriedades funcionais da proteína traduzida e sendo  
7 expressa em fenótipos distintos (O'rahilly, 2009). Os traços fenotípicos mais promissores para a  
8 identificação forense são aqueles relacionados à pigmentação da pele, olhos e cabelos, por  
9 serem características muito marcantes, de fácil visualização e constituírem um dos fenótipos  
10 mais variáveis na população humana (Jablonski e Chaplin, 2000; Sturm e Teasdale, 2001; Parra,  
11 2007; Kayser e Schneider, 2009; Kayser, 2015)). O desenvolvimento de técnicas quantitativas,  
12 tais quais a aferição dos valores de RGB (Red, Green, Blue) e/ou HSV (Hue saturation value), para  
13 mensurar a pigmentação proporcionou uma maior objetividade na classificação da ampla  
14 variação de pigmentação existente, permitindo identificar de forma mais precisa os genes que  
15 exercem diferentes influências sobre essa característica (Jablonski, 2004).

16 Apesar de ser promissora a utilização de marcadores genéticos de traços fenotípicos na  
17 identificação humana para fins forenses, a descrição de genes e do mecanismo pelo qual eles  
18 influenciam a definição das características fenotípicas é de difícil elucidação, uma vez que essas  
19 características apresentam um padrão de herança complexo, por serem determinadas por  
20 múltiplos genes (poligenia) e por sofrerem forte influência do meio ambiente (Sturm e Larsson,  
21 2009). Não obstante, estudos demonstram que a base genética da variação normal da  
22 pigmentação é passível de ser decifrada (Van Daal, 2008) se forem avaliadas as variantes  
23 genéticas adequadas (Liu *et al.*, 2009). Pesquisadores estão, neste momento, tentando  
24 identificar quais são, portanto, elas.

### 27 **1.1- Predição de Fenótipo**

28  
29 Estudos de associação genótipo-fenótipo sobre a cor do cabelo, a cor da íris e a  
30 pigmentação da pele têm pesquisado variações de SNPs em genes diretamente ou  
31 indiretamente envolvidos na síntese de pigmentos (Sturm, 2009; Walsh *et al.*, 2011a; Walsh *et al.*,  
32 2011b; Walsh *et al.*, 2013; Kayser, 2015). Todos eles sustentam que há uma necessidade  
33 grande por parte dos órgãos policiais de possuir uma base de dados com padrões genéticos que  
34 possam auxiliar na solução de crimes. Objetiva-se excluir ou não excluir indivíduos, em fase de  
35 investigação criminal, buscando identificar suspeitos ou vítimas (Pulker *et al.*, 2007). A predição  
36 de fenótipo torna-se de extrema relevância, especialmente no caso de não se ter informações

1 sobre o perfil genético de referência (para confronto de alelos de STRs), ou seja, em tais casos a  
2 amostra questionada pode ser estudada para fins da predição do fenótipo do indivíduo (Kayser  
3 *et al.*, 2011).

4 Através da predição de fenótipo seria possível definir as CEV e, portanto, a gama de  
5 suspeitos ou indivíduos investigados poderia ser reduzida e/ou direcionada, ou seja, a amostra  
6 se tornaria uma espécie de “testemunha biológica” do caso. Historicamente tem-se muitos  
7 casos em que tal predição ajudaria nesse sentido, exemplos são o caso do estupro que ocorreu  
8 em uma escola francesa onde realizou-se uma operação de coleta de amostras de DNA de 527  
9 homens, com o objetivo de encontrar o estuprador de uma aluna, já que a vítima não conseguia  
10 identificar o agressor (fonte: [www.bbc.com](http://www.bbc.com)) e o caso do assassinato da adolescente Naomi  
11 Smith, onde 800 homens se submeteram a coleta de DNA com a finalidade de comparação com  
12 o material genético encontrado no corpo da vítima (fonte: [www.independent.co.uk](http://www.independent.co.uk)). Se a  
13 tecnologia de predição de fenótipo já estivesse elucidada e a ponto na época das ocorrências,  
14 diversos indivíduos não precisariam ter se submetido a coleta de material genético o que,  
15 inclusive, daria celeridade ao deslinde do feito.

## 16 17 **1.2- Biologia da Pigmentação**

18  
19 Acredita-se que existam mais de 120 genes envolvidos nas vias de pigmentação, os quais  
20 agem em diferentes estágios do processo de produção da melanina (Sturm, 2006; Branicki *et al.*,  
21 2009). Tais genes agem em caráter quantitativo, influenciando cumulativamente nas diferentes  
22 tonalidades de coloração (Tully, 2007). Contudo, cerca de apenas uma dezena de genes têm  
23 alelos que definitivamente determinam os padrões extremos de pigmentação melânica, isto é,  
24 os padrões escuro e claro.

25 A compreensão da biologia da pigmentação depende do entendimento dos eventos  
26 intramelanossomais. Os melanossomas (compartimentos subcelulares, produzidos pelos  
27 melanócitos, que sintetizam e estocam polímeros de melanina) são os principais responsáveis  
28 pela coloração dos tecidos humanos (Sulem *et al.*, 2007) por causa do seu conteúdo de  
29 melanina. Melaninas são produzidas em dois tipos distintos quimicamente: eumelanina  
30 (coloração castanho-preto) e feomelanina (coloração amarelo-avermelhada) (Ito *et al.*, 2006).  
31 Eumelaninas são escuras, opacas e altamente polimerizadas (formato oval), enquanto que as  
32 feomelaninas, por possuírem sulfidril e cisteína em seus passos de conjugação, são mais leves  
33 e menos polimerizadas (formato esférico), o que ocasiona uma deposição desforme no lúmen.  
34 O acúmulo, a transferência, o conteúdo e a forma destes componentes estão diretamente  
35 ligados à intensidade da coloração de pele e de olho (Sturm, 2006; Kondo e Hearing, 2011).

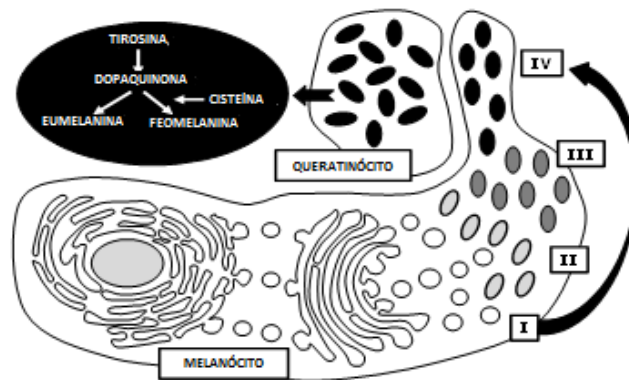
### 1 **1.2.1- Pigmentação da Pele**

2

3 A melanina é gradualmente produzida nos melanossomas que, quando maduros, são  
4 transportados ao longo das projeções dendríticas do melanócitos até os queratinócitos  
5 adjacentes da pele (Sturm *et al.*, 2008). O número de melanócitos entre os diferentes tons de  
6 pele é constante, mas peles mais escuras apresentam melanossomas mais densos e  
7 individualmente dispersos, enquanto peles mais claras apresentam melanossomas menos  
8 densos e menores (Sturm e Duffy, 2012).

9 Os melanossomos desenvolvem gradualmente o pigmento enquanto amadurecem  
10 (Slominski *et al.*, 2004; Lin e Fisher, 2007; Ho *et al.*, 2011). Após sintetizado, o pigmento é  
11 passado aos queratinócitos circundantes da pele e do folículo piloso (Sturm, 2006) e, por fim,  
12 para as células do epitélio pigmentar da retina (Kondo e Hearing, 2011). O acúmulo de melanina  
13 e a distribuição do melanossomos variam quantitativamente entre indivíduos de diferentes  
14 grupos étnicos (Alaluf *et al.*, 2002; Sturm, 2006). Os melanossomos são divididos em quatro  
15 estágios de maturação (I a IV), determinados pela sua estrutura, quantidade, qualidade e arranjo  
16 da melanina produzida (Kushimoto *et al.*, 2001; Costin e Hearing, 2007) (Figura 1). O  
17 melanossomo no estágio I é uma organela esférica, com ausência da enzima tirosinase e de  
18 componentes internos estruturais. No estágio II a organela apresenta uma estrutura ovoide,  
19 atividade de TYR e a expressão de uma proteína da matriz estrutural, conhecida como PMEL17,  
20 necessária para a produção da matriz fibrilar interna, e pela deposição mínima de melanina  
21 (Berson *et al.*, 2001; Sturm, 2006; Costin e Hearing, 2007; Kondo e Hearing, 2011). No estágio  
22 III, a síntese de melanina é iniciada e o pigmento é depositado sobre fibras internas. No estágio  
23 IV de desenvolvimento, o melanossomo apresenta formato elíptico ou elipsoide. Estes estágios  
24 de desenvolvimento referem à eumelanossomos (que sintetizam a eumelanina), e nos  
25 feomelanossomos (que sintetizam a feomelanina) o processo é similar com a diferença que  
26 durante o processo de maturação, há ausência de fibrilas (Costin e Hearing, 2007; Brenner e  
27 Berking, 2010).





1  
 2 **FIGURA 1** - Esquema de um melanócito: O mecanismo para a produção da melanina. Os quatro  
 3 estágios (I, II, III e IV) de desenvolvimento dos melanossomos são mostrados deslocando-se em  
 4 direção à periferia da célula. À esquerda (ampliado) está um melanossomo maduro (estágio IV)  
 5 onde se vê a via de síntese de eumelanina e feomelanina de forma simplificada.

6

### 7 **1.2.2- Pigmentação no Olho**

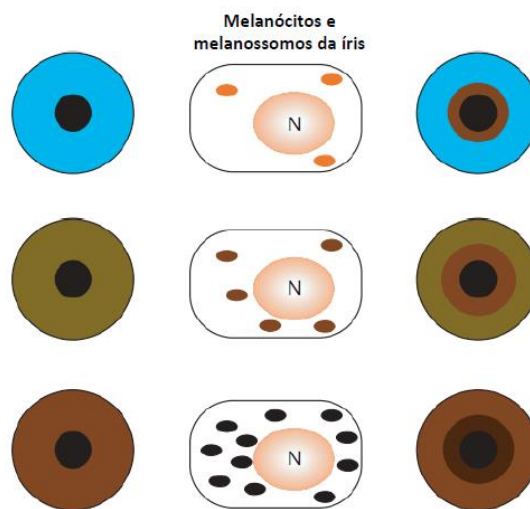
8

9 Os melanócitos dos olhos têm duas origens distintas, os que se encontram no epitélio  
 10 pigmentoso da íris (IPE) que têm origem neuroectodérmica, e os melanócitos estromais, que  
 11 têm a mesma origem embrionária dos melanócitos dérmicos (oriundos da crista neural), e  
 12 migram através do trato uveal durante o desenvolvimento. Na íris de coloração castanha existe  
 13 uma abundância de melanócitos e de melanina na camada basal anterior e no estroma,  
 14 enquanto que na íris de coloração azul essas camadas contêm muito pouco de melanina (Sturm  
 15 e Larsson, 2009).

16 A íris, região ocular composta por tecido conectivo e muscular, é responsável pelo  
 17 controle da entrada da luz incidente através de sua abertura central denominada pupila,  
 18 regulando a formação das imagens na retina. Essa região é subdividida em cinco camadas, sendo  
 19 a mais externa denominada borda anterior, seguida do estroma, esfíncter e músculos  
 20 dilatadores e, mais internamente, o epitélio posterior pigmentado (Sturm e Frudakis, 2004).  
 21 Dessas, a borda anterior e o estroma são as principais camadas contribuintes para a variação da  
 22 coloração da íris (Wilkerson *et al.*, 1996), uma vez que o epitélio posterior pigmentado apresenta  
 23 pigmentação em todos as cores observadas, não contribuindo de forma relevante para a  
 24 variação fenotípica final (Sturm e Frudakis, 2004). Íris de coloração marrom apresentam  
 25 melanócitos com alto conteúdo de melanina na borda anterior e no estroma, enquanto em íris  
 26 azuis, essas camadas possuem baixo teor de melanina, permitindo maior passagem da luz e  
 27 consequente refletância de ondas curtas azuis pelas fibras de colágeno presentes. No estroma,

1 os dendritos dos melanócitos são geralmente paralelos à superfície da íris, os quais tendem a se  
2 agrupar na borda anterior, e correspondem a aproximadamente 66% das células componentes  
3 dessas camadas, independentemente da coloração da íris (Wilkerson *et al.*, 1996).

4 Ao contrário do que ocorre na pele, onde a melanina é continuamente produzida e  
5 secretada, os melanossomos da íris não são secretados dos melanócitos, mas sim retidos e  
6 acumulados no citoplasma destes. As variações na coloração observadas são resultado das  
7 diferentes quantidades e tipos de melanossomos presentes nos melanócitos, sendo a proporção  
8 de eumelanina/feomelanina geralmente maior em colorações mais escuras, e colorações mais  
9 claras apresentando maior concentração do pigmento feomelanina (Sturm e Larsson, 2009).  
10 Sabe-se que, apesar de todas as colorações de olhos possuírem um número similar de  
11 melanócitos, os olhos azuis contêm um mínimo de pigmentos e poucos melanossomos; os olhos  
12 verdes são resultado de moderados níveis de pigmentos, intensidade de melanina e número de  
13 melanossomos, enquanto olhos castanhos são resultado de níveis altos de melanina e grande  
14 quantidade de partículas melanossomais (Sturm e Frudakis, 2004). Entretanto, estudos acerca  
15 das proporções eumelanina/feomelanina que possam vir a determinar a cor dos olhos são ainda  
16 inconclusivos (Figura 2).



30 **FIGURA 2** - Pigmentação dos olhos. Diferenças entre os tipos e quantidades de melanossomos  
31 nos melanócitos dos olhos mais claros aos mais escuros - Olhos azuis contêm um mínimo de  
32 pigmentos e poucos melanossomos; olhos verdes são resultado de moderados níveis de  
33 pigmentos, intensidade de melanina e partículas melanossomais e olhos castanhos/pretos  
34 resultam de alto nível de pigmentos e grande quantidade de partículas melanossomais (retirado  
35 de Sturm e Frudakis, 2004).

36

### 1.2.3- Pigmentação nos Cabelos

Quase todos os mamíferos têm cabelo com uma grande diversidade de padrões de cores. O cabelo é um “mini-órgão”, que os mamíferos obtiveram durante a evolução para cobrir a superfície da pele. Os mamíferos mais velhos muitas vezes podem ser distinguidos dos mamíferos mais novos por suas mudanças associadas ao envelhecimento na aparência, incluindo o envelhecimento dos cabelos (Nishimura, 2011).

O cabelo humano pertence a um grupo de proteínas denominadas  $\alpha$ - queratinas, que se distinguem de outras proteínas por seu alto teor de pontes de dissulfeto (S-S) provenientes do aminoácido cistina. Os melanócitos são originalmente derivados da crista neural e se localizam nos folículos pilosos para pigmentar o cabelo (Nishimura, 2011). Os melanócitos dos folículos capilares proliferam repetidamente e se diferenciam para originar a pigmentação do cabelo em todos os ciclos. Uma vez que os melanócitos colonizam a matriz do cabelo na parte inferior de cada folículo piloso, durante o desenvolvimento ou durante o ciclo do cabelo, eles se diferenciam em melanócitos totalmente maduros que produzem pigmentos de melanina em uma organela chamada melanosoma. Os melanosomas são então transferidos para queratinócitos precorticais adjacentes que produzem o eixo do cabelo, devido a isso, durante cada ciclo de cabelo, o cabelo crescente torna-se pigmentado. Os melanócitos morrem e desaparecem da matriz do cabelo quando os folículos pilosos regredem, mas reaparecem neste local durante a fase crescente do folículo piloso. A alteração no cabelo mais perceptível com a idade é a despigmentação do fio, resultando no aparecimento dos fios brancos. Isto ocorre devido à perda de atividade do melanócito na fase anágena de produção do fio no bulbo capilar. (Neste & Tobin, 2004).

### 1.2.4- Melanogênese

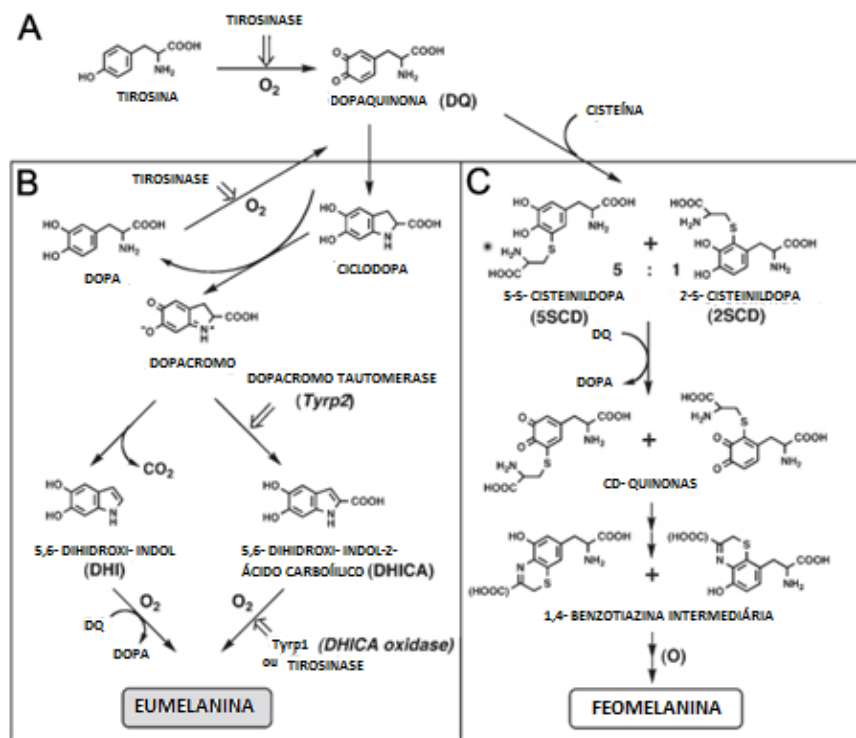
O regulador positivo da síntese de melanina mais importante é o receptor de melanocortina 1 (MC1R). Na sua via, o MC1R ativa a CREB (*cyclicAMP response-element binding protein*) que aumenta a expressão de MITF (*Microphthalmia-associated Transcription Factor*) e sua ativação, por fosforilação, estimula a transcrição de tirosinase (TYR), de proteína 1 relacionada a tirosinase (TYRP1) e de dopacromo tautomerase (DCT) (Liu *et al.*, 2013). O primeiro passo na melanogênese é a criação da dopaquinona (DQ) a partir da oxidação de tirosina pela tirosinase (TYR) com 3,4-dihidroxi-fenilalanina (DOPA) como agente intermediário (Kondo e Hearing, 2011) (Figura 3A).

## Eumelanogênese

Após a produção de DQ a sua espontânea ciclização acontece, dando origem a ciclodopa, que então faz uma reação de redução com uma molécula de DQ, resultando em uma molécula de DOPAcromo e uma de DOPA (Land *et al.*, 2003). O DOPAcromo, por sua vez, é decomposto espontaneamente por descarboxilação em pH neutro para formar 5,6-dihidroindole (DHI) e ácido 5,6-dihidroindole-2-carboxílico (DHICA) numa razão de 70:1 (Palumbo *et al.*, 1987). A enzima DOPAcromo tautomerase (DCT) tautomeriza o DOPAcromo produzindo somente DHICA (Palumbo *et al.*, 1991). DHI e DHICA são oxidadas e polimerizadas para dar origem a eumelanina (Figura 3B).

## Feomelanogênese

A feomelanogênese se dá pela adição de cisteína na oxidação da DQ para produzir dois isômeros principais de DOPAcisteinila (CD): a 5SCD e a 2SCD. Ocorre uma troca redox entre a CD e a DQ para produzir CD-quinonas e DOPA. A desidratação da CD-quinona faz com que ocorra a ciclização dando origem a orto-quinona (QI) que é rearranjada e forma o intermediário 1,4-benzotiazina, que se polimeriza dando origem a feomelaninas (Napolitano *et al.*, 1994; Napolitano *et al.*, 1999; Napolitano *et al.*, 2000; Greco *et al.*, 2009; Wakamatsu *et al.*, 2009) (Figura 3C).



1 **FIGURA 3** – Via biossintética da produção de eumelanina/feomelanina a partir de tirosina. As  
2 atividades da TYR, TYRP1 e DCT estão envolvidas na produção da eumelanina, mas apenas TYR  
3 e cisteína são necessários para a produção de feomelanina a partir de DQ (retirado de Ito e  
4 Wakamatsu, 2010).

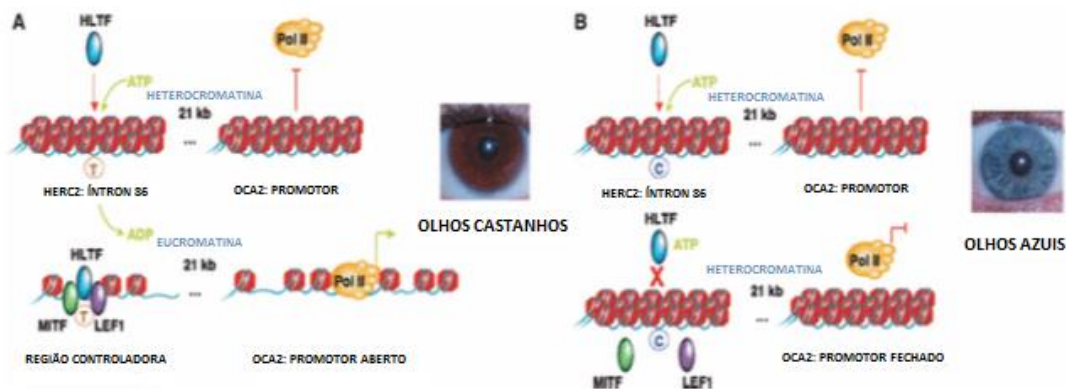
### 6 **1.3 - Pigmentação e Ancestralidade**

8 Sabe-se que a pigmentação de pele e de olhos está ligada à evolução. Os seres humanos,  
9 diferente de outros primatas e mamíferos, não possuem pelos em todo o corpo. Uma explicação  
10 bem aceita à ausência de pelos é de que seria difícil a transpiração em lugares quentes e, assim,  
11 a regulação da temperatura seria mais eficiente na carência de pelos. Porém, a sobrevivência do  
12 indivíduo diante da quantidade reduzida de pelos na superfície do corpo implicaria na  
13 necessidade de uma proteção contra a radiação ultravioleta (UV). A principal via de proteção  
14 contra a luz UV é a presença de melanina na superfície da pele, a qual é capaz de absorver a  
15 radiação eletromagnética no comprimento das ondas ultravioleta resguardando, assim, as  
16 moléculas de DNA, de proteínas e de outras macromoléculas (Rees, 2004; Liu *et al.*, 2013).

17 Observando a diferença da pigmentação e a localização geográfica das populações  
18 humanas, pode-se estabelecer uma relação entre radiação ultravioleta, perda dos pelos e  
19 manutenção da pigmentação da pele pela evolução: a pele negra oferece uma proteção maior  
20 em ambientes com alta incidência de radiação UV, como em locais de baixa latitude, enquanto  
21 a pele clara facilita a biossíntese de vitamina D em regiões de latitudes elevadas, onde a  
22 incidência de radiação UV é mediana ou baixa (Soejima *et al.*, 2007). Sendo assim, acredita-se  
23 que a história da transformação do fenótipo africano de pigmentação negra até o leque de  
24 variação melânica presente nos seres humanos de hoje se deu, pelo menos em parte, pela  
25 pressão seletiva da radiação solar, fixando-se o clareamento da pele mais em europeus e  
26 asiáticos, do que nos africanos (Liu *et al.*, 2013).

27 Diante da relação fortemente positiva entre pigmentação e origem étnica, poder-se-ia  
28 questionar se a presença de determinado alelo seria apenas um marcador de ancestralidade ou  
29 se, de fato, poderia ele estar diretamente envolvido na determinação fenotípica da  
30 pigmentação. Na tentativa de solucionar esta questão, muitos genes têm sido estudados, e os  
31 produtos proteicos e/ou a expressão de suas variantes alélicas têm sido associados à síntese ou  
32 não de melaninas. Porém, no presente momento é possível identificar apenas um número  
33 restrito de SNPs em genes os quais, de fato, estão diretamente envolvidos na síntese de  
34 pigmentos de pele e de olhos, independentemente da origem étnica da população em estudo  
35 (Kenny *et al.*, 2012). Por exemplo, como será citado a seguir, o SNP rs12913832 do gene *HERC2*,  
36 o qual se encontra a 21kb *upstream* do promotor de *OCA2*, regula a expressão de *OCA2* pelo

1 dobramento da cromatina que permite o ancoramento dos fatores de transcrição de *OCA2*. O  
 2 produto da expressão de *OCA2*, a proteína P, é essencial para a síntese de melanina. O alelo T  
 3 (ou alelo A, considerando a fita complementar) do SNP rs12913832 ao permitir a abertura de  
 4 cromatina, e o recrutamento dos fatores de transcrição de *OCA2*, induz à coloração mais escura,  
 5 enquanto o alelo C (ou alelo G, considerando a fita complementar) mantém a cromatina mais  
 6 fechada, tendo menos eficácia no recrutamento destes fatores, ocasionando uma coloração  
 7 mais clara (Figura 4).



8  
 9

10 **FIGURA 4-** Modelo de determinação da cor de iris castanho (A) e azul (B) influenciado pelo  
 11 SNP rs12913832 do gene *HERC2*, segundo Sturm e Larsson, 2009.

12

13 Mesmo se localizando no interior de segmentos gênicos chave para a produção de  
 14 melanina, poucos foram os alelos, contudo, já reconhecidos como sendo os próprios  
 15 responsáveis pelo efeito variante de síntese de pigmento. Por não ter sido ainda completamente  
 16 elucidada a implicação efetiva no fenótipo de cada variante alélica, elas podem estar associadas  
 17 à pele/olho escuro ou à pele/olho claro apenas por estarem em desequilíbrio de ligação com  
 18 uma variante, de fato, fenótipo-efetiva. De todas as formas, em geral, têm sido usados para  
 19 explicar a coloração escura aqueles alelos cuja frequência está mais elevada nas populações  
 20 africanas negras (alelos originais/ancestrais) e os alelos derivados, cuja frequência está mais  
 21 elevada nas populações europeias, para explicar a coloração clara. O desafio na análise da  
 22 maioria dos SNPs ainda é saber se o marcador é de fenótipo ou de ancestralidade. No intuito de  
 23 focar nos marcadores de fenótipo, têm sido retirados dos estudos associativos e/ou de predição  
 24 de fenótipo aqueles alelos derivados (isto é, mutações polimórficas mais recentes) cujas  
 25 frequências estão mais aumentadas nas populações africanas, como é o caso do alelo derivado  
 26 (mais recente) do SNP rs26722 do gene *SLC45A2* (regulador do pH que controla a via da síntese  
 27 de melanina, no interior do melanócito). Algumas populações africanas de fenótipo escuro  
 28 mostram a presença de um alelo derivado raro (alelo A; Lys272) deste SNP e, a despeito do seu

1 fenótipo claro, populações caucasianas têm em frequência o alelo original deste SNP (alelo G;  
2 Glu272) similares às de aborígenes australianas escuros (frequência de G ~0,98) (Nakayama *et*  
3 *al.*, 2006). Devido ao fato de o alelo original (G; Glu272) estar presente em grupos populacionais  
4 com diferentes fenótipos de pigmentação (caucasianos claros e aborígenes escuros), e ao fato  
5 de populações africanas de cor escura conterem ambas variantes (Lys272 e Glu272), Graf *et al.*  
6 (2007) concluíram que este polimorfismo não estaria associado com a determinação de  
7 pigmentação, mas sim que se trataria de uma variação associada à distribuição/isolamento  
8 geográfico. Assim, o alelo raro A surgiu como uma mutação recente apenas em africanos, e não  
9 interferiu na expressão da cor.

#### 12 **1.4 - Genes e SNPs envolvidos na determinação de pigmentação de pele e olhos**

14 Mais de uma centena de genes com efeitos conhecidos já foram descritos em modelos  
15 animais e seus homólogos identificados em humanos (IFPCS, 2009). Esses genes podem ser  
16 divididos em grupos funcionais de acordo com a influência que seus produtos protéicos exercem  
17 na produção e regulação da pigmentação como segue: 1- genes que codificam fatores de  
18 crescimento e transcrição para o controle do crescimento e diferenciação dos melanoblastos  
19 em melanócitos; 2- genes que codificam proteínas componentes dos melanossomos; 3- genes  
20 que controlam a biossíntese de organelas relacionadas; 4- genes que determinam a produção  
21 de eumelanina *versus* feomelanina e 5- genes envolvidos no transporte dos melanossomos.

22 Determinados SNPs nos genes *HERC2*, *OCA2*, *SLC24A5*, *SLC45A2*, *TYR*, *TYRP1* e *MC1R*,  
23 foram associados com a manifestação da pigmentação humana, sendo os três primeiros os  
24 responsáveis por grande parte da variação de coloração entre os continentes (Liu *et al.*, 2013).  
25 Com base na literatura disponível, um estudo de revisão do nosso grupo selecionou oito SNPs  
26 em genes candidatos ao estudo de previsão de pigmentação em diversas populações, por  
27 estarem significativamente relacionados com a manifestação da pigmentação em seres  
28 humanos. A base do nosso estudo foi: Tully, 2007; Sulem *et al.*, 2007; Sturm *et al.*, 2008; Giardina  
29 *et al.*, 2011; Walsh *et al.*, 2011a; Walsh *et al.*, 2011b, Donnelly *et al.*, 2012; Allwood *et al.*, 2013;  
30 Liu *et al.*, 2013; Walsh *et al.*, 2013. No Quadro 1, estão apresentadas informações acerca dos  
31 genes e SNPs por nós selecionados e, na Figura 5, se pode visualizar a localização de cada  
32 produto gênico ou estrutura em questão. Adicionalmente, no “ANEXO 1” há informações  
33 detalhadas sobre cada SNP. A seguir, serão descritos dados disponíveis sobre estes SNPs  
34 resultantes de estudos populacionais.

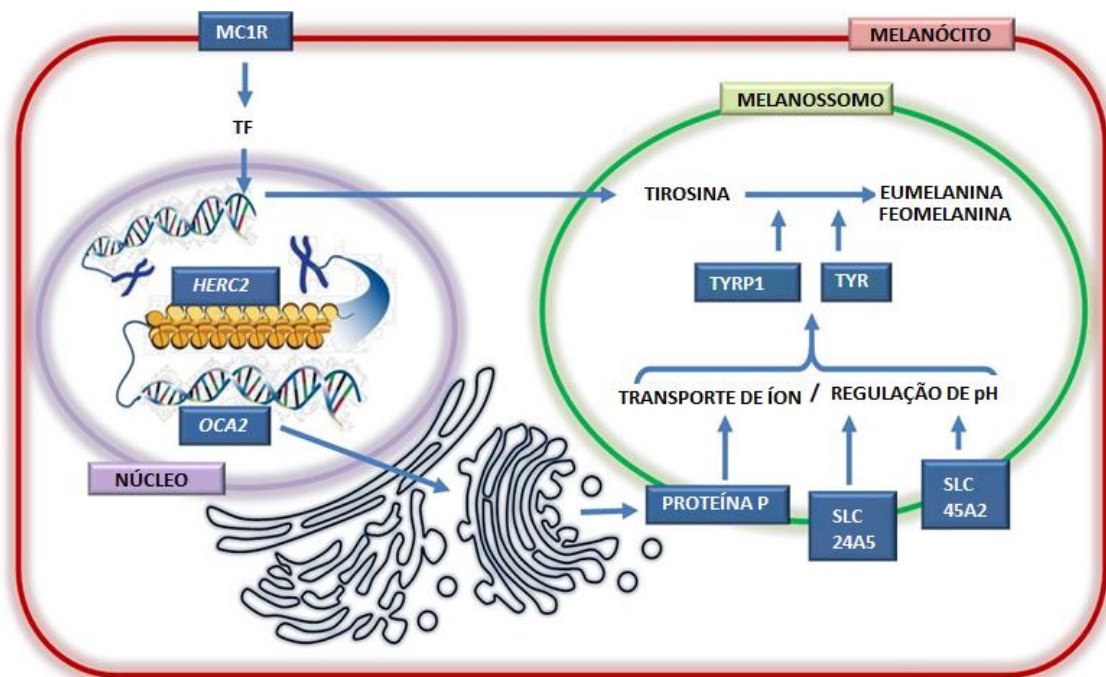
#### 36 **QUADRO 1- Genes e SNPs candidatos à predição de pigmentação em humanos**

1

Gene	Lócus	Proteína	rs	Varição	SNP
<b>HERC2</b>	15q13.1	HERC2 (350 aa)	rs916977	Íntron 12	A>G
			rs12913832	Íntron 86	A>G
<b>OCA2</b>	15q11.2	Proteína P (814 aa)	rs4778138	Intron 1	G>A
<b>SLC24A5</b>	15q21.1	NCKX5 (500 aa)	rs1426654	Ala111Thr	331G>A
<b>SLC45A2</b>	5p13.2	MATP (530 aa)	rs16891982	Phe374Leu	1122C>G
<b>TYR</b>	11q14	TYR	rs1042602	Ser192Tyr	C>A
<b>TYRP1</b>	9p23	TYRP1 (537 aa)	rs2733832	intragênica	C>T
<b>MC1R</b>	16q24.3	MC1R (317 aa)	rs8045560	intragênica	C>T

2

3



4

5

6 **FIGURA 5-** Indicação na estrutura celular dos genes e/ou seus produtos proteicos candidatos  
7 para predição de pigmentação em humanos.

8

9

#### 1.4.1- Gene *HERC2*, SNPs rs916977 – rs12913832

10

11 O gene *HERC2* (lócus 15q13.1) codifica a proteína *HECT domain - RCC1-Like domain-*  
12 *containing protein 2* (ou *E3 ubiquitin-protein ligase HERC2*) envolvida no tráfego de proteínas.  
13 Ainda que a função do gene *HERC2* seja desconhecida, a sequência de DNA no interior de *HERC2*  
14 desempenha um papel estrutural no genoma e na regulação de pigmentação (Donnelly *et al.*,  
15 2012). De acordo com Sturm *et al.* (2008), a proteína *HERC2* não está diretamente envolvida na  
16 síntese de pigmento, mas a sequência de DNA dentro da região do gene *HERC2* controla a  
17 expressão do gene *OCA2* o que, por sua vez, controla a produção de melanina (Sturm *et al.*,



1 2008). Tais sequências no interior de *HERC2* são *enhancers* do promotor do gene *OCA2* (Visser  
2 *et al.*, 2012).

3 Visser *et al.* (2012) mostraram que o SNP rs12913832 de *HERC2*, que se encontra a 21kb  
4 *upstream* do promotor de *OCA2*, regula a expressão deste gene pelo dobramento da cromatina  
5 que gera um *loop* de longo alcance que acaba por promover a expressão de *OCA2*. Esta atividade  
6 é mediada pelos fatores de transcrição HLTF, LEF1 e MITF. Porém, diferentes versões alélicas de  
7 *HERC2* impedem o desenrolamento da cromatina, o que reduz (mas não impede) a ligação dos  
8 fatores de transcrição de *OCA2*, reduzindo conseqüentemente sua taxa transcricional. O alelo T  
9 (ou alelo A, considerando a fita complementar) do SNP rs12913832 permite a abertura da  
10 cromatina e consegue recrutar os fatores de transcrição de *OCA2*, o que permite a produção de  
11 melanina e ocasiona uma coloração mais escura, enquanto o alelo C (ou alelo G, considerando  
12 a fita complementar) apresenta uma baixa no recrutamento destes fatores, impedindo a síntese  
13 plena de pigmento e ocasionando uma coloração mais clara. Assim, via controle da expressão  
14 de *OCA2*, o gene *HERC2* tem o papel predominante na determinação de cor. Dados  
15 populacionais indicam que os SNPs rs916977 e rs12913832, localizados no íntron 12 e íntron 86,  
16 respectivamente, são melhores preditores de cor de olho do que três SNPs já descritos no  
17 interior do íntron 1 do próprio gene *OCA2* (rs7495174, rs6497268 e rs11855019). Estudos em  
18 outras populações também indicam o SNP rs12913832 como sendo o SNP com maior efeito na  
19 predição da cor de olhos (Allwood *et al.*, 2013).

20 Adicionalmente, se observou que o alelo A do SNP rs1667394, localizado no íntron 4 do  
21 gene *HERC2*, tem um padrão de associação com cor de cabelo e de olhos em um gradiente de  
22 redução da pigmentação: menor frequência alélica em indivíduos com cabelos e olhos castanhos  
23 e com a maior frequência em indivíduos com cabelo loiro e olhos azuis. O alelo A é encontrado  
24 em uma frequência de 80-90% na população Norte da Europa (Sulem *et al.*, 2007).

#### 25 26 **1.4.2- Gene *OCA2*, SNP rs4778138**

27  
28 O gene que, ao conter mutações deletérias, ocasiona o albinismo oculocutâneo tipo 2  
29 (gene *OCA2*), codifica uma proteína de com 12 domínios transmembrana denominada proteína  
30 P (Sturm *et al.*, 2008). Acredita-se que a proteína P esteja envolvida no transporte de ânions, na  
31 regulação do pH melanossomal (Visser *et al.*, 2012), e que esteja envolvida no tráfego de  
32 proteínas internas como a tirosinase (TYR) e as proteínas associadas a tirosinase 1 (TYRP1)  
33 (Valenzuela *et al.*, 2011). Foi observado que a proteína P atuava na permuta de cátions  $\text{Na}^+/\text{H}^+$ ,  
34 e/ou como um transportador de glutamato. Ambas as funções indicam que este produto do  
35 gene *OCA2* está envolvido no fornecimento de substrato para a TYR. A proteína P pode ainda  
36 estar envolvida no tráfego intracelular da enzima TYR, durante a maturação dos melanossomos.

1 O pH do melanossomo regula a atividade da TYR afetando, portanto, uma série de fatores na  
2 síntese da melanina, como a taxa de produção de melanina, a proporção de  
3 eumelanina/feomelanina e a maturação de melanossomos nos melanócitos e células de  
4 melanoma. Resultados experimentais indicaram que a produção de melanina é ideal em  
5 ambiente celular com pH próximo a pH-neutro (cerca de pH= 6.8). Paralelamente, a síntese da  
6 melanina é quase ausente em pH menor que 5.5. Em um pH menos ácido, por outro lado, a  
7 melanogênese pode prosseguir sem qualquer alteração na abundância de proteínas  
8 melanogênicas. Outro fator envolvido ainda na regulação do pH é a enzima ATPase-vacuolar (V-  
9 ATPase), presente no interior do melanossomo, a qual atua na acidificação dos melanossomos  
10 (Kayser, 2015).

11 O *OCA2* por si só, independente de *HERC2*, também desempenha um papel na variação  
12 de coloração comum em olhos, cabelo e pele. Três SNPs que se encontram no intron 1 de *OCA2*  
13 mostraram ter uma forte associação com a coloração dos olhos (Liu *et al.*, 2013). Além disto, o  
14 SNP rs1800407 encontrado no exón 13 do gene *OCA2* ocasiona uma troca do tipo Arg419Gln na  
15 sequência polipeptídica da proteína P (Donnelly *et al.*, 2012), esta troca aumenta a penetrância  
16 do fenótipo de olhos azuis associados ao SNP rs12913832 (Sturm *et al.*, 2008; Walsh, *et al.*,  
17 2011a).

18

#### 19 **1.4.3- Gene *SLC24A5*, SNP rs1426654**

20

21 Baseado no que se conhece do gene *NCKX5* (Na<sup>+</sup>/Ca<sup>2+</sup>/K<sup>+</sup> exchanger 5), a regulação do  
22 pH melanossomal pode ser o papel da proteína codificada pelo gene *SLC24A5* (Tully, 2007). O  
23 gene *NCKX5* codifica uma proteína de 500 aminoácidos, que faz parte da família das proteínas  
24 envolvidas nas trocas de sódio/cálcio dependente de potássio, localiza-se próxima à membrana  
25 melanossomal e atua no transporte de membrana do melanossomo (Sturm, 2006; Norton *et al.*,  
26 2007; Liu *et al.*, 2013). Lamason *et al.* (2005) descobriram que *SLC24A5* está localizado em  
27 melanossomos ou seus precursores, e assumiram que ele poderia atuar acumulando Ca<sup>2+</sup> no  
28 melanossomo. O gene *SLC24A5* foi descoberto primeiramente em *zebra-fish*, no qual está  
29 relacionado ao fenótipo de uma listra clara significativa. O homólogo foi identificado e  
30 caracterizado em humanos. O gene humano mostrou que uma substituição de Ala111Thr,  
31 resultante do SNP rs1426654 (de Ala111-alelo G para Thr111-alelo A) tinha um efeito similar na  
32 pigmentação humana, isto é, conferia redução de melanina (fenótipo claro). De acordo com  
33 Ginger *et al.* (2008) o *SLC24A5* está associado com a rede trans-Golgi, eles também descobriram  
34 que o alelo associado com pele mais escura alelo G (Ala111) do SNP rs1426654 teve maior  
35 atividade de troca de íons em relação ao alelo A, o qual está associado com pele mais clara  
36 (Thr111). O alelo ancestral evolutivamente conservado (Ala111) do gene *SLC24A5* também foi

1 encontrado em alta frequência em populações africanas e asiáticas, e que o alelo derivado  
2 (Thr111) atingiu a fixação em europeus. Especificamente, a frequência para a variante Thr111  
3 varia entre 98,7 e 100% entre as amostras da população de origem europeia, enquanto Ala111  
4 tem uma frequência de 93-100% em amostras de africanos, nativos americanos e população do  
5 Leste Asiático da (Dimisianos *et al.*, 2008)

#### 6 7 **1.4.4- Gene *SLC45A2*, SNP rs16891982**

8  
9 Originalmente identificado como um antígeno do melanoma humano (AIM1) e  
10 recentemente renomeado carregador de soluto família 45, membro 2, o gene *SLC45A2* contém  
11 sete éxons que se estendem a uma região de aproximadamente 40kb e codifica a proteína MATP  
12 de 530 aminoácidos que se supõe ter 12 domínios transmembrana (Graf *et al.*, 2007). Mutações  
13 no gene *SLC45A2* causam uma forma de albinismo em humanos (OCA4) (Vierkötter *et al.*, 2012;  
14 Tully, 2007; Sturm, 2006). Graf *et al.* (2005) relataram dois SNPs em *SLC45A2* que não estavam  
15 associados a doenças, mas sim associados com a variação normal da pigmentação humana, e  
16 também a variações entre diferentes populações: Glu272Lys (rs26722) e Phe374Leu  
17 (rs16891982). A proteína MATP desempenha um papel crucial no processamento e tráfego  
18 intracelular da tirosinase (TYR), uma das enzimas cruciais e necessárias para síntese de melanina,  
19 assim como na via de outras proteínas melanossomais (Branicki *et al.*, 2009; Vierkötter *et al.*,  
20 2012).

21 O SNP rs16891982 gera o polimorfismo proteico Phe374Leu e regula a função de transporte  
22 da região transmembrana da proteína MATP. A variante Leu374 (alelo G) desempenha um papel  
23 importante no transporte de prótons, resultando em um pH ótimo intramelanossomal, o qual  
24 permite a atividade da tirosinase (TYR) e conseqüentemente a produção adequada da  
25 eumelanina (fenótipo pigmentado). A variante Phe374 (alelo C) pode alterar o transporte, o pH  
26 e a síntese do pigmento (Graf *et al.*, 2005; Tully, 2007; Vierkötter *et al.*, 2012). Nakayama *et al.*  
27 (2006) referiram o alelo C (Phe374) como alelo “tipo-Caucasiano”, já que ele resultou na redução  
28 da função da MATP, por alterar o tráfego intracelular de elementos melanossomais, criando um  
29 ambiente de decréscimo na produção de melanina. Estudando o mesmo gene, Graf *et al.* (2005)  
30 observaram que o alelo Lys272 do gene *SLC45A2* mostrou um significativo aumento nas  
31 populações asiáticas e afro-americanas quando comparadas a caucasianos. Contudo, as  
32 populações aborígenes australianas mostraram frequências alélicas similares às das populações  
33 de caucasianos. Devido ao fato do fenótipo de pigmentação dos dois últimos grupos  
34 populacionais serem bem diferentes, os autores sugeriram que o polimorfismo rs26722 não  
35 estaria associado com a determinação de pigmentação, mas apenas seria uma variação  
36 associada à distribuição ou isolamento geográfico.

1  
2 **1.4.5- Gene TYR, SNP rs1042602**  
3

4 O gene *TYR* transcreve a tirosinase, uma importante enzima para a fase inicial da  
5 melanogênese (Liu *et al.*, 2013). O gene *TYR* foi indicado como um loci para a susceptibilidade  
6 ao vitiligo generalizado (Jin *et al.*, 2012). Variações polimórficas em *TYR* estão diretamente  
7 associadas com alterações na coloração de olhos, cabelos e pele (Sulem *et al.*, 2007). Em  
8 especial, dois polimorfismos de *TYR*, rs1042602 (Ser192Tyr) e rs1126809 (Ala402Gly), têm  
9 mostrado frequências alélicas associadas a fenótipos de diferentes colorações: alelos que  
10 aparecem em alta frequência em populações europeias e estão ausentes em populações  
11 africanas (Stokowski *et al.*, 2007). As análises enzimáticas *in vitro* revelaram que existe uma  
12 redução de aproximadamente 40% na atividade catalítica de tirosinase devido à variação  
13 Ser192Tyr (Chaki *et al.*, 2011).  
14

15 **1.4.6- Gene TYRP1, SNP rs2733832**  
16

17 Localizado no cromossomo 9p23, o gene *TYRP1* codifica a proteína tirosinase 1, que faz  
18 parte do complexo da enzima tirosinase durante a produção de melanina (Parra, 2007; Liu *et al.*,  
19 2013). O impacto da proteína TYRP1, assim como a da DCT, na estabilidade da proteína TYR e na  
20 produção de outras enzimas envolvidas na formação catalítica de melanina mostra que  
21 alterações no gene *TYRP1* têm um papel importante na variação normal da pigmentação  
22 humana (Sturm, 2006). Há mais de dez anos, viu-se evidências que polimorfismos no *TYRP1*  
23 estão associados à coloração de olhos em europeus (Frudakis *et al.*, 2003). Estudos proteicos  
24 mostraram que a proteína codificada por *TYRP1* foi encontrada estando elevada 2,6 vezes na  
25 pele de africanos e de indianos (pigmentação escura) se comparada à pele de Mexicanos,  
26 Chineses e Europeus (pigmentação clara) (Alaluf *et al.*, 2003). Foi ainda observado que mutações  
27 raras (não polimórficas) em *TYRP1* são responsáveis pelo albinismo oculocutâneo tipo 3 (OCA3)  
28 (Sulem *et al.*, 2007).  
29

30 **1.4.7- Gene MC1R, SNP rs8045560**  
31

32 Até o momento, o gene *MC1R* é o mais bem caracterizado dentre os genes descritos que  
33 influenciam a variação normal de pigmentação em humanos (Gerstenblith *et al.*, 2007). Este  
34 gene apresenta um padrão peculiar de distribuição dos seus alelos entre as populações  
35 humanas: populações africanas e outras caracterizadas pela pele escura apresentam as menores  
36 variações alélicas desse gene, enquanto em populações asiáticas e europeias, esse gene é

1 altamente polimórfico (Makova e Norton, 2005). Estudos em diferentes populações têm  
2 demonstrado que a região codificadora do gene *MC1R* apresenta mais de 70 variantes já  
3 identificadas (Gerstenblith *et al.*, 2007), com impacto significativo no fenótipo de pigmentação  
4 neste grupo étnico (Harding *et al.*, 2000; Sturm, 2006).

5 O gene *MC1R* transcreve para o receptor de melanocortina 1 (*MC1R*), uma proteína-G  
6 transmembrana de sete passos localizada na membrana de melanócitos (Valenzuela *et al.*, 2011)  
7 e diretamente responsável pela regulação da síntese de eumelanina/feomelanina (Makova *et*  
8 *al.*, 2005). A ligação do hormônio paracrino alfa-estimulante dos melanócitos ( $\alpha$ -MSH) ao  
9 receptor *MC1R* causa uma cascata de sinalização de cAMP aumentando a produção de  
10 eumelanina (Makova *et al.*, 2005; Valenzuela *et al.*, 2011), enquanto que a ligação de seu  
11 antagonista, a proteína de sinalização agouti (ASIP), resulta em menos produção de eumelanina  
12 e no aumento da produção de feomelanina (Voisey *et al.*, 2006; Liu *et al.*, 2013).

### 14 **1.5 - Frequências alélicas dos SNPs envolvidos na determinação de pigmentação** 15 **de pele e olhos em populações Africanas e Europeias**

16  
17 O Quadro 2 apresenta um levantamento realizado com dados de indivíduos oriundos  
18 de populações Africanas e Europeias descritas nos bancos de dados ALFRED (Allele Frequency  
19 Database; National Science Foundation) e HAPMAP. ALFRED foi projetado para tornar os dados  
20 de frequência de alelos em amostras de populações humanas prontamente disponíveis para  
21 utilização pelas comunidades; e o projeto internacional HAPMAP é resultado de uma parceria  
22 de cientistas e agências de financiamento do Canadá, China, Japão, Nigéria, Reino Unido e  
23 Estados Unidos, que tem por finalidade desenvolver e manter um recurso público para ajudar  
24 pesquisadores a encontrar genes e estes SNPs resultantes de estudos populacionais.  
25 Consultando estas duas bases de dados, compilamos resultados de todas as populações já  
26 estudadas até maio de 2013. Todos os dados foram plotados em uma planilha prévia para,  
27 finalmente, se obter o número total de indivíduos já genotipados e a frequência de cada alelo.  
28 Estes dados estão apresentados no Quadro 2.

29  
30 **QUADRO 2.** Genes, SNPs (rs) e Alelos associados de pigmentação em humanos, segundo sua  
31 frequência nos conjuntos de populações Africanas e Europeias estudados nos projetos HAPMAP  
32 e ALFRED.

Gene (SNP) Alelo	PROJETO	FREQ AFRICANOS	FREQ EUROPEUS
<i>HERC2</i> (rs12913832): Alelo A	HapMap	1.000 (120/120)	0.208 (47/226)
	ALFRED	0.956 (4365/4568)	0.285 (7973/27945)
<i>HERC2</i> (rs916977): Alelo A	HapMap	0.950 (114/120)	0.133 (16/120)

	ALFRED	0.884 (1002/1134)	0.215 (756/3518)
<i>OCA2</i> (rs4778138): Alelo G	HapMap	0.733 (88/120)	0.092 (11/120)
	ALFRED	0.746 (1047/1404)	0.166 (945/5706)
<i>SLC24A5</i> (rs1426654): Alelo G	HapMap	0.987 (223/226)	0.000 (0/116)
	ALFRED	0.707 (2630/3720)	0.011 (49/4271)
<i>SLC452A</i> (rs16891982): Alelo C	HapMap	1.000 (114/114)	0.017 (2/116)
	ALFRED	0.919 (3478/3786)	0.071 (1840/25961)
<i>TYR</i> (rs1042602): Alelo C	HapMap	1.000 (120/120)	0.571 (129/226)
	ALFRED	0.940 (1757/1870)	0.705 (3544/5026)
<i>TYRP1</i> (rs2733832): Alelo C	HapMap	0.951 (215/226)	0.398 (90/226)
	ALFRED	0.879 (181/206)	0.419 (447/1066)
<i>MC1R</i> (rs8045560): Alelo C	HapMap	0.951 (215/226)	0.465 (105/226)
	ALFRED	0.909 (1225/1348)	0.438 (1624/3706)

1

2

FONTE: <http://hapmap.ncbi.nlm.nih.gov/>; <http://alfred.med.yale.edu/alfred/index.asp>.

3

1           **2- OBJETIVOS**

2

3           **2.1- Objetivo Geral**

4

5           Analisar quanti e qualitativamente o grau de pigmentação de olho (íris) e de pele em  
6 indivíduos sulbrasilieiros, na intenção de identificar as possíveis relações dessas variáveis com os  
7 alelos de genes envolvidos na síntese de melanina.

8

9           **2.1- Objetivos Específicos**

10

11           - Genotipar os SNPs em genes de pigmentação;

12           - Determinar as frequências alélicas e genotípicas para cada SNP nos dois grupos de  
13 estudo;

14           - Comparar os grupos de acordo com a relação genótipo-fenótipo.

15

16

17

### 3- BIBLIOGRAFIA CITADA NO REFERENCIAL TEÓRICO

- 1  
2  
3 A. Alonso, P. Martin, C. Albarran, P. Garcia, L. Fernandez de Simon, M. Jesus Iturralde, A.  
4 Fernandez-Rodrigues, I. Atienza, et al. Challenges of DNA profiling in mass disaster  
5 investigation. *Croat Med J.* 46 (2005) 540-548.
- 6 A. Napolitano, C. Costantini, O. Crescenzi, G. Prota. Characterization of 1,4- benzothiazine  
7 intermediates in the oxidative conversion of 5-S-cysteinyl-dopa to pheomelanins,  
8 *Tetrahedron Lett* 35 (1994) 6365–6368.
- 9 A. Napolitano, P. Di Donato, G. Prota, E.J. Land. Transient quinonimines and 1,4-benzothiazines  
10 of pheomelanogenesis: new pulse radiolytic and spectrophotometric evidence, *Free*  
11 *Radic Biol Med.* 27 (1999) 521–528.
- 12 A. Napolitano, P. Di Donato, G. Prota. New regulatory mechanisms in the biosynthesis of  
13 pheomelanins: rearrangement vs. redox exchange reaction routes of a transient 2H-1,4-  
14 benzothiazine-o-quinonimine intermediate, *Biochim Biophys Acta.* 1475 (2000) 47–54.
- 15 A. Palumbo, M. d'Ischia, G. Prota. Tyrosinase-promoted oxidation of 5,6-dihydroxyindole-2-  
16 carboxylic acid to melanin. Isolation and characterization of oligomer intermediates,  
17 *Tetrahedron*, 43 (1987) 4203–4206.
- 18 A. Palumbo, F. Solano, G. Misuraca, P. Aroca, J.C. Garcia-Borron, J.A. Lozano, G. Prota.  
19 Comparative action of dopachrome tautomerase and metal ions on the rearrangement  
20 of dopachrome, *Biochim Biophys Acta* 1115 (1991) 1–5.
- 21 A. Slominski, D. J. Tobin, S. Shibahara, J. Wortsman. Melanin pigmentation in mammalian skin  
22 and its hormonal regulation. *American Physiological Reviews* 84(2004) 1155–1228.
- 23 A. Van Daal, A. The genetic basis of human pigmentation, *Forensic Science International* 1(2008).  
24 541–543.
- 25 A. Vierkötter, U. Krämer, D. Sugiri, A. Morita, A. Yamamoto, N. Kaneko, M. Matsui, J. Krutmann.  
26 Development of lentigines in german and japanese women correlates with variants in  
27 the *SLC24A5*, *Gene Journal of Investigative Dermatology* 132 (2012) 733–736.
- 28 B. Budowle, F.R. Bieber, A.J. Eisenberg. Forensic aspects of mass disasters: strategic  
29 considerations for DNA-based human identification. *Leg Med (Tokyo)* 7(2005)230-243.
- 30 C.L. Wilkerson, N.A. Syed, M.R. Fisher, N.L. Robinson, H.L. Wallow, D.M. Albert. Melanocytes and  
31 iris color. Light microscopic findings, *Archives of Ophthalmology* 114(1996)437–442.
- 32 D.V. Neste; D.J. Tobin. Hair cycle and hair pigmentation: dynamic interactions and changes  
33 associated with aging, *Micron.* 35 (2004) 193-200.
- 34 E. Giardina, A. Spinella, G. Novelli. Past, present and future of forensic DNA typing,  
35 *Nanomedicine* 6 (2011) 257–270.
- 36 E.E. Kenny, N.J. Timpson, M.Sikora, M.C. Yee, A. Moreno-Estrada, C. Eng, S. Huntsman, E.G.  
37 Burchard, M. Stoneking, C.D. Bustamante, S. Myles. Melanesian blond hair is caused  
38 by an amino acid change in TYRP1, *Science* 336 (2012) 554. doi:  
39 10.1126/science.1217849
- 40 E.J. Land, S. Ito, K. Wakamatsu, P.A. Riley. Rate constants for the first two chemical steps of  
41 Eumelanogenesis, *Pigment Cell Res.* 16 (2003) 487–493.
- 42 E.J. Parra. Human pigmentation variation: Evolution, genetic basis, and implications for public  
43 health, *Year book of physical anthropology.* 50 (2007) 85-105.
- 44 E.K. Nishimura. Melanocyte stem cells: a melanocyte reservoir in hair follicles for hair and  
45 skin pigmentation, *Pigment Cell & Melanoma Research* 24 (2011) 401–410.
- 46 F. Liu, K. van Duijn, J.R. Vingerling, A. Hofman, A.G. Uitterlinden, A.C.J.W. Janssens, M. Kayser.  
47 Eye color and the prediction of complex phenotypes from genotypes, *Curr. Biol.* 19  
48 (2009) R192–R193.
- 49 F. Liu, B. Wenb, M. Kayser. Colorful DNA polymorphisms in humans, *Semin Cell Dev Biol* (2013).
- 50 G.B. Stacey, S.F. Schaffner, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. Defelice,  
51 A. Lochner, M. Faggart, S.N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E.S.  
52 Lander, M.J. Daly, D. Altshuler. The structure of haplotype blocks in the human genome,  
53 *Science* 296 (2002) 2225–2229.



- 1 G. Dimisianos, I. Stefanaki, V. Nicolaou, V. Sypsa, C. Antoniou, M. Poulou, O. Papadopoulos, H.  
2 Gogas, E. Kanavakis, E. Nicolaidou, A. D. Katsambas, A. J. Stratigos. A study of a single  
3 variant allele (rs1426654) of the pigmentation-related gene SLC24A5 in Greek subjects,  
4 *Experimental Dermatology* 18 (2009)175–7.
- 5 G. E Costin, V. J Hearing. Human skin pigmentation: melanocytes modulate skin color in response  
6 to stress, *The Journal of the Federation of American Societies for Experimental Biology*  
7 21 (2007) 976-994.
- 8 G. Greco, K. Wakamatsu, L. Panzella, S. Ito, A. Napolitano, M. D'Ischia. Isomeric cysteinyl dopas  
9 provide a (photo)degradable bulk component and a robust structural element in red  
10 human hair pheomelanin, *Pigment Cell Melanoma Res.* 22 (2009) 319–327.
- 11 G. Tully. Genotype versus phenotype: Human pigmentation, *Forensic Sci Int. Genet.* 1 (2007)  
12 100-104.
- 13 H. Ho, R. Kapadia, S. Al-Tahan, S.; et al. WIP1 Coordinates Melanogenic Gene Transcription and  
14 Melanosome Formation via TORC1 Inhibition, *The Journal of Biological Chemistry* 286, n.  
15 14. (2011) 12509-12523.
- 16 H.L. Norton, R.A Kittles, E. Parra, P. Mckeigue, X. Mao, K. Cheng, V. A. Canfield, D.G Bradley, B.  
17 Mcevoy, M. Shriver. Genetic evidence for the convergent evolution of light skin in  
18 Europeans and East Asian, *Mol Biol Evol* 24(2007) 710-22.
- 19 H. Pulker, M.V. Lareu, C. Phillips, A. Carracedo. Finding genes that underlie physical traits of  
20 forensic interest using genetic tools, *Forensic Sci Int Genet.* 1 (2007) 100-104.
- 21 INTERNATIONAL FEDERATION OF PIGMENT CELL SOCIETIES – IFPCS 2009: Disponível em:  
22 <http://www.espcr.org/micemut/> - Acessado em abril de 2013.
- 23 J. Berson, D. C. Harper, D. Tenza, G. Raposo, M.S. Marks. Pmel17 initiates premelanosome  
24 morphogenesis within multivesicular bodies, *Molecular Biology of the cell* 12(2001)  
25 3451-3464.
- 26 J.W. Bond. Value of DNA evidence in detecting crime. *J Forensic Sci* (2007) 128-136.
- 27 J. Graf, J. Voisey, I. Hughes, A.Van Daal. Promoter polymorphisms in the MATP (*SLC24A5*) gene  
28 are associated with normal skin color variation, *Hum. Mutat.* 28 (2007) 710-717.
- 29 J.L. Rees. The genetics of Sun sensitivity in humans, *Am. J. Hum. Genet.* 75 (2004) 739-751.
- 30 J.R.Landis, G.G Koch. The measurement of observer agreement  
31 for categorical data, *Biometrics* 33(1977)159-174.
- 32 J.S. Allwood, S.A. Harbison. SNP model development for the prediction of eye colour in New  
33 Zealand, *Forensic Science International: Genetics* 7 (2013) 444–452.
- 34 J. Voisey, M.C. Gomez-Cabrera, D.J. Smit, J.H. Leonard, R.A. Sturm, A. van Daal. A polymorphism  
35 in the agouti signalling protein (ASIP) is associated with decreased levels of mRNA  
36 *Pigment Cell Res.* 19 (2006) 226–231.
- 37 J.Y Lin, D.E Fisher. Melanocyte biology and skin pigmentation, *Nature* 445 (2007) 843-850.
- 38 K. Makova, H. Norton. Worldwide polymorphism at the *MC1R* locus and normal pigmentation  
39 variation in humans, *Peptides* 26 (2005) 1901-1908.
- 40 K. Nakayama, A. Soemantri, F.J. Bumbein Dashnyam, R. Ohtsuka, P. Duanchang, M. N. Isa, W.  
41 Settheetham-Ishida, S. Harihara, T. Ishida. Identification of novel functional variants of  
42 the melanocortin 1 receptor gene originated from asians, *Human Genet* 119(2006) 322-  
43 30.
- 44 K. Wakamatsu, K. Ohtara, S. Ito. Chemical analysis of late stages of pheomelanogenesis:  
45 conversion of dihydrobenzothiazine to a benzothiazole structure, *Pigment Cell*  
46 *Melanoma Res.* 22 (2009) 474–486.
- 47 K. Yoshida, K. Yayama, A. Hatanaka, K. Tamaki. Efficacy of extended kinship analyses utilizing  
48 commercial STR kit in establishing personal identification, *Leg. Med.* 13 (2011) 12–15.
- 49 L.D.G. Kobachuk. Estudo de frequências alélicas de dez locos STRs do cromossomo X na  
50 população do estado do Paraná e sua contribuição na identificação humana, (2012).
- 51 M.A. Jobling, P. Gill. Encoded evidence: DNA in forensic analysis, *Nature Rev. Genet.* 5 (2004)  
52 739–751.

- 1 M. Brenner, M. Berking. Grundlagen der Hautpigmentierung: Biochemie und regulation der  
2 melaninsynthese, *Der Hautarzt* 61(2010) 554-560.
- 3 M. Chaki, M. Sengupta, M. Mondal, A. Bhattacharya, S. Mallick, R. Bhadra, K. Ray. Molecular and  
4 functional studies of tyrosinase variants among Indian oculocutaneous albinism type 1,  
5 *Patients Journal of Investigative Dermatology* 131 (2011) 260–262.
- 6 M. Kayser, P. Schneider. DNA-based prediction of human externally visible characteristics in  
7 forensics: Motivations, scientific challenges, and ethical considerations, *Forensic Science*  
8 *International* 3(2009)154–161.
- 9 M. Kayser, P. Knijff. Improving human forensics through advances in genetics, genomics and  
10 molecular biology, *Nature Reviews* 12 (2011) 179-192.
- 11 M. Kayser. Forensic DNA Phenotyping: Predicting human appearance from crime scene material  
12 for investigative purposes. *Forensic Sci. Int. Genet.* (2015) 18, 33-48.
- 13 M.P. Donnelly, P. Paschou, E. Grigorenko, D. Gurwitz, C. Barta, R.B. Lu, O.V. Zhukova, J.J. Kim, M.  
14 Siniscalco, M. New, H. Li, S.L.B. Kajuna, V.G. Manolopoulos, W.C. Speed, A.J. Pakstis, J.R.  
15 Kidd, K.K. Kidd. A global view of the *OCA2-HERC2* region and pigmentation, *Hum Genet*  
16 131 (2012) 683-696.
- 17 M.R, Gerstenblith, A.M, Goldstein, M.C, Fargnoli, K. Peris, M.T Landi. Comprehensive evaluation  
18 of allele frequency differences of MC1R variants across populations, *Human Mutat*  
19 28(2007) 495-505.
- 20 M. Soejima, Y. Koda. Population differences of two coding SNPs in pigmentation-related genes  
21 *SLC24A5* and *SLC24A5*, *Int J Legal Med* 121 (2007) 36-39.
- 22 M. Visser, M. Kayser, R.J. Palstra. *HERC2* rs12913832 modulates human pigmentation by  
23 attenuating chromatin-loop formation between a long-range enhancer and the *OCA2*  
24 promoter *Genome Res.*22 (2012) 446-455.
- 25 N. Jablonski, G. Chaplin. The evolution of human skin coloration, *Journal of Human Evolution*, v.  
26 39(2000) 57-106.
- 27 N. Jablonski. The Evolution of Human Skin and Skin Color, *Annual Review of Anthropology* v. 33  
28 (2004) 585–623.
- 29 P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, K.P. Magnusson, A. Manolescu,  
30 A. Karason, A. Palsson, G. Thorleifsson, M. Jakobsdottir, S. Steinberg, S. Palsson, F.  
31 Jonasson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K.R. Benediksdottir, K.K.  
32 Aben, L.A. Kiemenev, J.H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir, K.  
33 Stefansson. Genetic determinants of hair, eye and skin pigmentation in Europeans, *Nat.*  
34 *Genet.* 39 (2007) 1443–1452.
- 35 R.A Sturm, R. Teasdale, N. Box. Human pigmentation genes: Identification, Structure and  
36 Consequences of Polymorphic Variation, *Gene* 277(2001) 49–62.
- 37 R.A Sturm, T.N Frudakis. Eye colour: portals into pigmentation genes andancestry, *Trends in*  
38 *Genetics* 20(2004) 327–332.
- 39 R.A. Sturm. A golden age of human pigmentation genetics, *Trends Genet.* 22 (2006) 464-468.
- 40 R.A. Sturm, D.L. Duffy, Z.Z. Zhao, F.P.N. Leite, M.S. Stark, N.K. Hayward, N.G. Martin, G.W.  
41 Montgomery, A single SNP in an evolutionary conserved region within intron 86 of the  
42 *HERC2* gene determines human blue-brown eye color. *The American Journal of Human*  
43 *Genetics* 82 (2008) 424-431.
- 44 R.A. Sturm. Molecular genetics of human pigmentation diversity, *Hum. Mol. Genet. Review Issue*  
45 18 (2009) R9-R17.
- 46 R.A. Sturm, M. Larsson. Genetics of human iris colour and patterns, *Pigment Cell Melanoma*  
47 *Res.* 22 (2009) 544-562.
- 48 R.A. Sturm, D.L. Duffy. Human pigmentation genes under environmental selection, *Genome*  
49 *Biololgy* (2012) 13:248. doi:10.1186/gb-2012-13-9-248
- 50 R.K. Valenzuela, S. Ito, K. Wakamatsu, M.H. Brilliant. Prediction model validation: normal human  
51 pigmentation variation, *J. Forenseci Res.* 2 (2011). doi:10.4172/2157-7145.1000139
- 52 R.L. Lamason, M.A. Mohideen, J.R. Mest, A.C. Wrong, H.L. Norton, M.C. Aros, M.J. Juryneec, X.  
53 Mao, V.R. Humphreville, J.E. Humbert, S. Sinha, J.L. Moore, P. Jagadeeswaran, W. Zhao,

1 G. Ning, I. Makalowska, P.M. McKeigue, D. O'Donnell, R. Kittles, E.J. Parra, N.J. Mangini,  
2 D.J. Grunwald, M.D. Sriver, V.A. Canfield, K.C. Cheng. SLC24A5, a putative cation  
3 exchanger, affects pigmentation in zebrafish and humans, *Science* 310(2005) 1782–  
4 1786.

5 R. M. Harding, E. Healy, A. J. Ray, N. S. Ellis, N. Flanagan, C. Todd, C. Dixon, A. Sajantila, I. J.  
6 Jackson, M. A. Birch-Machin, J. L. Rees. Evidence for Variable Selective Pressures at  
7 MC1R, *Am J Human Genet* 66(2000)1351-61.

8 R.P. Stokowski, P.V.K. Pant, T. Dadd, A. Fereday, D.A. Hinds, C. Jarman, W. Filsell, R.S. Ginger,  
9 M.R. Green, F.J. van der Ouderaa, D.R. Cox. A genomewide association study of skin  
10 pigmentation in a south Asian population, *The American Journal of Human Genetics*. 81  
11 (2007) 1119–1132.

12 R.S. Ginger, S. E. Askew, R.M. Ogborne, R.M Wilson, D. Ferdinando, T. Dadd , A.M Smith, S. Kazi,  
13 R.T Szerencsei, R.J Winkfein, P.P Schnetkamp, M.R Green. SLC24A5 encodes a trans-  
14 Golgi network protein with potassium-dependent sodium–calcium exchange activity  
15 that regulates human epidermal melanogenesis, *J Biol Chem* 283(2008) 5486–5495.

16 S. Alaluf, D. Atkins, K. Barrett, M. Blout, N. Carter, A. Heath. Ethnic variation in melanin content  
17 and composition in photoexposed and photoprotected human skin, *Pigment Cell Res*  
18 15(2002) 112-118.

19 S. Alaluf, K. Barrett, M. Blount, N. Carter. Ethnic variation in tyrosinase and *TYRP1* expression in  
20 photoexposed and photoprotected human skin, *Pigment. Cell. Res.* 16 (2003) 35-42.

21 S. Ito, K. Wakamatsu, Chemistry of Melanins. In: J.J. Nordlund, R.E. Boissy, V.J. Hearing, R.A. King,  
22 W.S. Oetting, J.P. Ortonne (Eds.). *The Pigmentary System. Physiology and*  
23 *Pathophysiology*, Blackwell Publishing; Oxford, 2006, pp. 282-310.

24 S. O'Rahilly. Human genetics illuminates the path to metabolic disease, *Nature* 462(2009) 307-  
25 14.

26 S. Walsh, A. Lidenbrgh, S.B. Zuniga, T. Sijen, P. Knijff, M. Kayser, K.N. Ballantyne. Developmental  
27 validation of the IrisPlex system: Determination of blue and brown iris colour for forensic  
28 intelligence, *Forensic Sci Int. Genet.* 5 (2011a) 464-471.

29 S. Walsh, F. Liu, K.N. Ballantyne, M. van Oven, O. Lao, M. Kayser. IrisPlex: a sensitive DNA tool  
30 for accurate prediction of blue and brown eye colour in the absence of ancestry  
31 information, *Forensic Sci Int. Genet.* (2011b). doi:10.1016/j.fsigen.2010.02.004

32 S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser. The  
33 HirisPlex system for simultaneous prediction of hair and eye colour from DNA, *Forensic*  
34 *Sci Int. Genet.* 7 (2013) 98-115.

35 T. Frudakis, M. Thomas, Z. Gaskin, K. Venkateswarlu, K.S. Chandra, S. Ginjupalli, S. Gunturi, S.  
36 Natrajan, V.K. Ponnuswamy, K.N. Ponnuswamy. Sequences associated with human iris  
37 pigmentation, *Genetics* 165 (2003) 2071–2083.

38 T. Kondo, V.J. Hearing. Update on the regulation of mammalian melanocyte function and skin  
39 pigmentation, *Expert Rev Dermatol.* 6 (2011) 97–108.

40 T. Kushimoto, V. Basrur, J. Valencia, J. Matsunaga, W.D. Vieira, V.J. Ferrans, J. Muller, E. Appella,  
41 V.J. Hearing. A model for melanosome biogenesis based on the purification and analysis  
42 of early melanosomes. *Proceedings of the National Academy of Sciences of the United*  
43 *States of America* 98 (2001) 10698-10703.

44 W. Branicki, U. Brudnik, A. Wojas-Pel. Interactions between *HERC2*, *OCA2* and *MC1R* May  
45 Influence Human Pigmentation Phenotype, *Annals of Hum Genet.* 73 (2009) 160-170.

46 Y. Jin, S.A. Birlea, P.R Fain, T.M. Ferrara, S. Ben, S.L. Riccardi, J.B. Cole, K. Gowan, P.J. Holland,  
47 D.C. Bennett, R.M. Luiten, A. Wolkerstorfer, J.P.W. van der Veen, A. Hartmann, S.  
48 Eichner, G. Schuler, N. van Geel, J. Lambert, E.H. Kemp, D.J. Gawkrödger, A.P. Weetman,  
49 A. Taieb, T. Jouary, K. Ezzedine, M.R. Wallace, W.T. McCormack, M. Picardo, G. Leone,  
50 A. Overbeck, N.B. Silverberg, R.A. Spritz. Genome-wide association analyses identify 13  
51 new susceptibility loci for generalized vitiligo, *Nature Genetics* 44 (2012) 676–680.

52

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32

## **CAPÍTULO 2**

**MANUSCRITO DO ARTIGO CIENTÍFICO 1<sup>a</sup>**  
**MANUSCRITO DO ARTIGO CIENTÍFICO 2<sup>b</sup>**

<sup>a</sup>Os resultados do presente trabalho foram publicados no Journal of Forensic Research and Criminology.

<sup>b</sup>Os resultados do presente trabalho foram submetidos à Forensic Science International e estão formatados de acordo com as normas do periódico.

## Analysis of Eight SNPs in South Brazilian Subjects with Different Skin and Eye Melanin Content

Fernanda Rosa Sawitzki<sup>1</sup>, Rodrigo Rodenbusch<sup>1,2</sup>, Diego Wordell Gubert<sup>1</sup>, Deborah Soares Bispo Santos Silva<sup>1,3</sup>, Eduardo Filipe Avila Silva<sup>1,4</sup> and Clarice Sampaio Alho<sup>1\*</sup><sup>1</sup>Faculty of Biosciences, Laboratory of Human and Molecular Genetics, PUCRS, Brazil<sup>2</sup>Scientific and Technological Development Center (CDCT), State Foundation of Production and Research in Health (FEPPS), Porto Alegre-RS, Brazil<sup>3</sup>Forensic Sciences Institute, North Carolina State University, USA<sup>4</sup>Department of Federal Police, Porto Alegre-RS, Brazil

## Article Information

Received date: Jun 07, 2017

Accepted date: Aug 14, 2017

Published date: Aug 18, 2017

## \*Corresponding author

Clarice Sampaio Alho, Laboratório de Genética Forense, PUCRS. Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil, Tel: (+55) (51) 3320-3568; Email: csalho@pucrs.br

Distributed under Creative Commons CC-BY 4.0

**Keywords** Forensic; Human pigmentation; Phenotype; SNaPshot; SNP**Abbreviations** AUC: Area under the Receiver Operating Characteristic Curve; DNA: Deoxyribonucleic Acid; EVC: Externally Visible Characteristics; FDA: Factorial Discriminant Analyses; GP: Genetic Probability; HMC: High Melanin Content; LMC: Low Melanin Content; PCR: Polymerase Chain Reaction; PGL: Calculation of Pathway Genetic Load; RGB: Red, Green, Blue; SNP: Single Nucleotide Polymorphism; STR: Short Tandem Repeat

## Abstract

LMC (Low Melanin Content) and HMC (High Melanin Content) subjects have respectively low and high melanin content in both skin and eyes; LMC has white skin and blue eyes and HMC has dark skin and eyes. Comparative investigation between frequencies of genetic variants in LMC subjects versus HMC subjects may indicate which polymorphic variant is associated with melanin synthesis in skin and eyes. Coordinately, studies with Snow White-Like (SW) subjects may be informative to reveal any tissue-specific expression, since these individuals have white skin and dark eyes.

The LMC - HMC - SW model was used to analyze the allelic distribution of eight biallelic SNPs in pigment-related genes in admixed South Brazilian individuals. Based on allele frequencies of different human populations, allele "L" was used for the alleles associated with low melanin content populations (LMC subjects), and allele "H" was used for the alleles associated with high melanin content populations (HMC subjects). Allelic distribution of eight SNPs showed that 100% of LMC subjects (N=73) had less than eight H alleles, and 82% of HMC subjects (N=61) had eight or more H alleles. The AUC (Area under the Receiver Operating Characteristic Curve) value was 0.99, and the calculation of PGL (Pathway Genetic Load) and GP (Genetic Probability) showed that the SNP set presented 93% and 91% concordance between DNA genotype and phenotypes, respectively. Factorial Discriminant Analyses (FDA) performed in the SW group (light skin and dark eyes; N=116) showed a positive association between SNPs rs16891982 (SLC45A2), rs8045560 (MC1R), rs1426654 (SLC24A5), rs2733832 (TYRP1) and rs1042602 (TYR), and the LMC cluster for skin phenotype, and a positive association between SNPs rs4778138 (OCA2), rs12913832 (HERC2) and rs916977 (HERC2), and the HMC cluster for eye phenotype. The understanding of gene function in externally visible characteristics is important for the prediction of skin and eye colors in humans; the analyses presented here are an important contribution to the forensic DNA phenotyping scenario.

## Communication

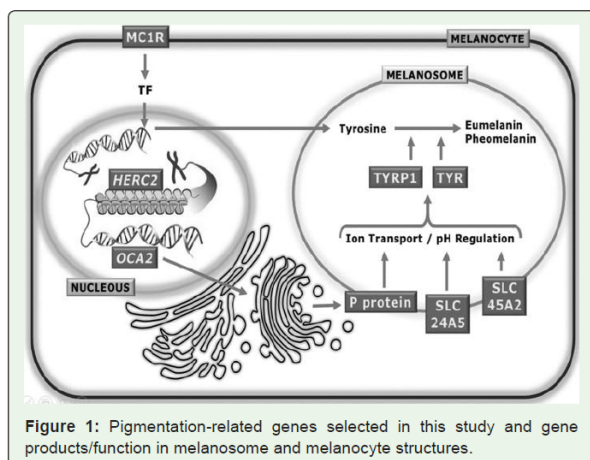
Forensic DNA Phenotyping (FDP) aims to predict appearance traits of a sample donor from DNA evidence left at a crime scene; a good FDP strategy may be especially useful in cases where the police have no other investigative leads, include reference DNA profiling. Skin and eye colors are important Externally Visible Characteristics (EVCs), since they are highly heritable genetic traits and are the most obvious and distinguishable externally visible characteristics to be used in human identification [1]. Hair melanin content is more prone to change by intrinsic or extrinsic factors; it is age-dependent, changing during late childhood, adolescence, adulthood, and old age, and becoming darker, gray or white at any given stage [1]. These changes, added to the occurrence of baldness, make hair a trait less useful to forensic DNA phenotyping applications.

There are hundreds of genes involved in the pigmentation process; however, for forensic purposes, a small number of SNPs have already been proved to be sufficient in the prediction of pigmentation patterns with high confidence [1]. In order to find a set of candidate SNPs able to predict EVCs, we first selected seven human genes linked to melanin synthesis (Figure 1): 1- *OCA2*: encodes the P protein involved in anion transport and in the melanosomal pH regulation, operating in the Tyrosinase (TYR) and the Tyrosinase Related Protein 1 (TYRP1) functions [2-5]. 2- *HERC2*: HERC domain and RCC1-Like domain 2 gene is located 10 Kb upstream of *OCA2* and it acts as a regulatory enhancer of *OCA2* region. The *HERC2* function is still unknown [1,2]. 3- *SLC45A2*: Solute carrier family 45 members 2 gene encodes MATP protein and regulates MATP function, with a crucial role in the trafficking processing and intracellular Tyrosinase (TYR) and protons transporting, controlling the intramelanosomal pH and the activity of Tyrosinase (TYR) [6]. 4- *MC1R*: transcribes the melanocortin 1 receptor (alpha melanocyte stimulating hormone receptor),

OPEN ACCESS

ISSN: 2574-2426

**How to cite this article** Sawitzki FR, Rodenbusch R, Gubert DW, Prado MJ, Silva DSBS and Alho CS. Analysis of Eight SNPs in Pigment-Related Genes in South Brazilian Subjects with High or Low Skin and Eye Melanin Content. SM J Forensic Res Criminol. 2017; 1(2): 1008.



a G-protein located in the melanocyte membrane which is directly responsible for eumelanin/phaeomelanin synthesis regulation [7]. 5- *SLC24A5*: solute carrier family 24, member 5, is responsible for accumulating  $\text{Ca}^{2+}$  in the melanosome [8]. 6- *TYRP1*: encodes the tyrosinase-related protein 1, which is necessary to melanin production [6-9]. 7- *TYR*: transcribes the tyrosinase, a crucial enzyme for the initial melanogenesis process [10]. After gene selection, SNPs were analyzed in each one of the genes using NCBI (National Center for Biotechnology; <https://www.ncbi.nlm.nih.gov/snp>). A large number of SNPs were identified per gene (*OCA2*: 44456; *HERC2*: 24579; *SLC45A2*: 5349; *MC1R*: 1238; *SLC24A5*: 2751; *TYRP1*: 3020; *TYR*: 13736) and SNP selection was done based on the following criteria: biallelic polymorphisms (i.e. no single insertion/deletion variants),

validation by 1000G (One Thousand Genomes Project; <http://www.internationalgenome.org>) and by HapMap (International Haplotype Map Project; <http://www.hapmap.ncbi.nlm.nih.gov>), and with global **Minor Allele Frequency (MAF)** distinguishable as a common polymorphism (i.e. with no rare variant).

Next, we examined the allelic frequencies data of each selected SNP in different human populations using ALFRED (Allele Frequency Database; <https://alfred.med.yale.edu>). In this step, we selected from ALFRED two sets of samples from extremely distinguishable subjects according to the melanin content in skin and eyes (subjects with low melanin content *versus* subjects with high melanin content), and discerned the allelic frequencies of each SNP in these two sets of samples. Thus, we selected less than 20 pre-candidate SNPs that independently could be able to distinguish those two sets of subjects (LMC *versus* HMC) by allelic frequencies. Simultaneously, we analyzed published studies with significant data on genotype-phenotype association. Finally, eight SNPs in seven genes were chosen. At each locus, allele L was used for each allele strongly associated with LMC subjects, and allele H for each allele strongly associated with HMC subjects. Table 1 presents the selected SNPs and their allele frequencies based on the populations frequencies available on ALFRED, and also data on sensibility, specificity, predictive values, relative ratio, and odds ratio (the database used to construct Table 1 is available in the supplemental file). These loci were tested in previously published genotype-phenotype association studies (Box 1). Since the eight SNPs in pigmentation-related genes presented remarkable results in all of the parameters, they were considered robust markers for color prediction.

1- HMC (High Melanin Content) subjects from populations of Africans, as available on ALFRED websites. 2- LMC (Low Melanin Content) subjects from populations of Europeans, as available on

**BOX 1:** Eight locus in seven pigmentation-related genes that were previously tested in genotype-phenotype association studies.

<ul style="list-style-type: none"> <li>• <b>OCA2:</b> The rs4778138 is a G&gt;A SNP in <i>OCA2</i> intron region strongly associated with eye color [6].</li> </ul>
<ul style="list-style-type: none"> <li>• <b>HERC2:</b> The rs12913832 is an A&gt;G SNP in <i>HERC2</i> intronic area, within <i>OCA2</i> enhancer region. The derived allele G of this SNP is associated with blue iris phenotype, being common in Europeans, particularly those of northwestern and eastern European descent [4]. The allele A (or allele T, if in the complementary DNA strand) allows chromatin opening and the <i>OCA2</i> transcription factors recruitment, which leads to darker iris, while the allele G (or allele C, if in complementary DNA strand) maintains the chromatin closed, being less effective in recruiting the aforementioned factors, which results in lighter iris [4]. The derived allele G of rs916977, other A&gt;G SNP, in about 10 Kb far from rs12913832, has been associated with blue eyes [1].</li> </ul>
<ul style="list-style-type: none"> <li>• <b>SLC45A2:</b> The rs16891982 is the 1122C&gt;G SNP in exon 5 of the <i>SLC45A2</i> gene. This SNP results in the non-synonymous substitution of Leucine (allele C; codon TTG on the coding/reverse DNA strand) by Phenylalanine (allele G; codon TTC on the coding/reverse DNA strand) at amino acid 374 (Leu374Phe). Rs16891982 is associated with human skin pigmentation normal variation, since it regulates the MATP function; the Leu374 variant (allele C – complementary to G on codon TTT on the coding/reverse DNA strand) plays an important role in the transport of protons, resulting in an optimum intramelanosomal pH, which allows the activity of tyrosinase (TYR) and, consequently, the adequate production of eumelanin (brown-black spectrum), while the derived Phe374 variant (allele G – complementary to C on codon TTC on the coding/reverse DNA strand) may change the transport, the pH and the synthesis of the pigment [7]. A strong association between rs16891982 and eye color has been reported [1].</li> </ul>
<ul style="list-style-type: none"> <li>• <b>MC1R:</b> The ancestral allele C of rs8045560 (C&gt;T SNP) is associated with darker skin populations [7].</li> </ul>
<ul style="list-style-type: none"> <li>• <b>SLC24A5:</b> The rs1426654 (G&gt;A SNP in the coding region of <i>SLC24A5</i> gene) results in non-synonymous substitution of Alanine (GCA) by Threonine (ACA) at amino acid 111 Ala111Thr. This SNP has shown evidence of natural selection; the derived allele A (Thr111) is associated with light skin pigmentation and is common in Europe, southwest Asia, and central Asia [1]. The ancestral allele (G allele; Ala111) is associated with darker skin [8].</li> </ul>
<ul style="list-style-type: none"> <li>• <b>TYRP1:</b> The ancestral allele C of rs2733832 (C&gt;T SNP) is associated with darker skin populations [6].</li> </ul>
<ul style="list-style-type: none"> <li>• <b>TYR:</b> The rs1042602 (C&gt;A SNP) results in the non-synonymous substitution of Serine (TCT) by Tyrosine (TAT) at amino acid 192 (Ser192Tyr) and in a reduction of about 40% in the catalytic activity of tyrosinase [10].</li> </ul>

**Citation:** Sawitzki FR, Rodenbusch R, Gubert DW, Prado MJ, Silva DSBS and Alho CS. Analysis of Eight SNPs in Pigment-Related Genes in South Brazilian Subjects with High or Low Skin and Eye Melanin Content. *SM J Forensic Res Criminol.* 2017; 1(2): 1008.

**Table 1:** Allele frequencies and other parameters of the selected SNPs based on population data from ALFRED.

RS (Gene; SNP)		HMC Populations <sup>1</sup>		LMC Populations <sup>2</sup>		P Value	Sensit	Specif	PV-1	PV-2	RR	OR
rs4778138 ( <i>OCA2</i> ; SNP G>A)	Allele H (G)	0.745	(1135/1524)	0.164	(957/5826)	<0.0001	0.744	0.836	0.542	0.926	7.31	14.81
	Allele L (A)	0.256	(390/1524)	0.836	(4870/5826)							
rs12913832 ( <i>HERC2</i> ; SNP A>G)	Allele H (A)	0.957	(4484/4688)	0.285	(8155/28286)	<0.0001	0.957	0.712	0.355	0.990	35.36	54.26
	Allele L (G)	0.043	(204/4688)	0.712	(20132/28286)							
rs16891982 ( <i>SLC452A</i> ; SNP C>G)	Allele H (C)	0.881	(3592/4076)	0.067	(1850/27594)	<0.0001	0.881	0.933	0.660	0.982	35.77	103.28
	Allele L (G)	0.119	(484/4076)	0.933	(25744/27594)							
rs8045560 ( <i>MC1R</i> ; SNP C>T)	Allele H (C)	0.915	(1440/1574)	0.440	(1730/3932)	<0.0001	0.915	0.560	0.454	0.943	7.92	13.68
	Allele L (T)	0.085	(134/1574)	0.560	(2202/3932)							
rs1426654 ( <i>SLC24A5</i> ; SNP G>A)	Allele H (G)	0.725	(2890/3984)	0.011	(52/4502)	<0.0001	0.725	0.988	0.982	0.803	4.98	226.12
	Allele L (A)	0.275	(1094/3984)	0.989	(4451/4502)							
rs2733832 ( <i>TYRP1</i> ; SNP C>T)	Allele H (C)	0.917	(396/432)	0.415	(537/1292)	<0.0001	0.917	0.584	0.424	0.954	9.33	15.47
	Allele L (T)	0.083	(36/432)	0.585	(755/1292)							
rs1042602 ( <i>TYR</i> ; SNP C>A)	Allele H (C)	0.946	(1982/2096)	0.702	(3614/5146)	<0.0001	0.946	0.298	0.354	0.930	5.11	7.37
	Allele L (A)	0.054	(114/2096)	0.298	(1532/5146)							
rs916977 ( <i>HERC2</i> ; SNP A>G)	Allele H (A)	0.890	(1116/1254)	0.214	(778/3638)	<0.0001	0.890	0.786	0.589	0.954	12.80	29.73
	Allele L (G)	0.110	(138/1254)	0.786	(2860/3638)							

ALFRED websites. ‘Allele H’: allele strongly associated with HMC people. ‘Allele L’: allele strongly associated with LMC people. Sensit: sensitivity. Specif: specificity. PV-1: predictive value for the presence of allele H in HMC populations. PV-2: Predictive Value for the presence of allele L in LMC populations. RR: Relative Ratio. OR: Odds Ratio.

To amplify the flanking SNP regions by multiplex PCR in a SNaPshot® System ABI Prism (Applied Biosystems), primer sequences were designed for each locus for conventional PCR and for the SNaPshot multiplex system (Table 2). Labeled products

were separated in an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) and analyzed in GeneMapper ID software version 3.2.1 (Applied Biosystem®).

In this study, first we analyzed 134 southern Brazilian subjects belonging to categories: LMC subjects (white skin and blue eyes; N = 73) and HMC subjects (dark skin and eyes; N = 61). Each skin color participant of was identified using Fitzpatrick score and the amount of red (R), green (G), and blue (B) pigments in an inner and hairless portion (below elbow) of the right arm using the *color analyzer* ACR-1023 (Instrutherm, São Paulo, Brazil). Each eye

**Table 2:** Primer sequences per locus for conventional PCR and for the SNaPshot multiplex system.

RS (Gene; SNP)	Primer Sequence
rs916977 ( <i>HERC2</i> ; SNP A>G)	F-5'- ttgtttggcaaacctccaca-3' / R-5'- tttccaattgccctgacat-3' (PCR Fragment: 200pb)
SNaPshot (35pb):	ggcaaacctccacagtgaggatgagcattgagtagaT/C
rs4778138 ( <i>OCA2</i> ; SNP G>A)	F-5'- ggattcaaaaagaagtctcaagg-3' / R-5'- gctctcctttgataccagca-3' (PCR Fragment: 146pb)
SNaPshot (42pb):	tattgaactgaatgaaagtgaataataacatatcaaaattgG/A
rs12913832 ( <i>HERC2</i> ; SNP A>G)	F-5'- cagctccatcaatgtgtca-3' / R-5'- ctgatgatagcgtgcaga-3' (PCR Fragment: 253pb)
SNaPshot (56pb):	ttcatggctctgtgtgatccaagaggcaggccagttcattgagcattaaA/G
rs16891982 ( <i>SLC452A</i> ; SNP C>G)	F-5'- ccaagtgtgtcagaccagaaa-3' / R-5'- cctcaacagcctccaatc-3' (PCR Fragment: 211pb)
SNaPshot (64pb):	aagacatcctaggagagagaagaacttaagaataaagtgaggaaacacggagttgatgcA/G
rs8045560 ( <i>MC1R</i> ; SNP C>T)	F-5'- aacgatgtttggtcagca-3' / R-5'- actcaaggcatctggaatgg-3' (PCR Fragment: 289pb)
SNaPshot (73pb):	caccaccccttttccatggggatctgcactcatctccaggaagatggtgggagataaccccgctgcC/T
rs1426654 ( <i>SLC24A5</i> ; SNP G>A)	F-5'- gcctccctcacccttcta-3' / R-5'- aggatggtgctaagccaat-3' (PCR Fragment: 425pb)
SNaPshot (90pb):	tagtgtaaacacatactcttaccattattaggcataacaatcattcattatgttcagccctggatgtctcaggatgttcaggcG/A
rs2733832 ( <i>TYRP1</i> ; SNP C>T)	F-5'- atgacctgctgttgaagt-3' / R-5'- ctcttgctgcatttcaa-3' (PCR Fragment: 334pb)
SNaPshot (98pb):	ccaaatgatcctatgttctctctgcacacctcacagatgagctcttggatgaatggctgaggagatacaatggtgaagacatttcattatgcC/T
rs1042602 ( <i>TYR</i> ; SNP C>A)	F-5'- actgcaagttggctttgg-3' / R-5'- gcttcatgggcaaatcaat-3' (PCR Fragment: 307pb)
SNaPshot (106pb):	ggccaaatgaaaaatgatcaacaccatgtttaacagacatcaatattatgacctcttctgtctgatcattattatgttcaatgatgactgtctgggggC/A

**Citation:** Sawitzki FR, Rodenbusch R, Gubert DW, Prado MJ, Silva DSBS and Alho CS. Analysis of Eight SNPs in Pigment-Related Genes in South Brazilian Subjects with High or Low Skin and Eye Melanin Content. *SM J Forensic Res Criminol.* 2017; 1(2): 1008.

Table 3: Phenotype data from 134 LMC and HMC Southern Brazilian subjects.

		LMC	HMC
<b>Total</b>	N (%)	73 (54.82)	61 (45.18)
<b>Sex</b>	Male [N (%)]	40 (54.80)	32 (52.46)
	Female [N (%)]	33 (45.20)	29 (47.54)
<b>Age</b>	Years [Mean (SD)]	36.6 (13.6)	30.4 (8.7)
<b>Skin Fitzpatrick Score</b>	Type 1 [N (%)]	30 (41.09)	0
	Type 2 [N (%)]	43 (58.91)	0
	Type 5 [N (%)]	0	26 (42.62)
	Type 6 [N (%)]	0	35 (57.38)
<b>Skin R-G-B Color</b>	Red [Mean (SD)]	88.4 (13.1)	49.1 (12.9)
	Green [Mean (SD)]	70.3 (13.2)	35.3 (10.2)
	Blue [Mean (SD)]	61.8 (13.8)	27.9 (8.2)
<b>Eye Type</b>	Light blue [N (%)]	50 (68.49)	0
	Dark blue [N (%)]	23 (31.51)	0
	Dark brown [N (%)]	0	39 (63.94)
	Black [N (%)]	0	22 (36.06)
<b>Eye R-G-B Color</b>	Red [Mean (SD)]	99.4 (12.5)	59.2 (19.1)
	Green [Mean (SD)]	95.9 (12.5)	32.6 (11.7)
	Blue [Mean (SD)]	137.9 (25.2)	42.8 (14.8)
<b>Auto Declared Origin</b>	European	66 (90.41)	13 (21.31)
	African	0	26 (37.70)
	Amerindian	0	2 (0.03)
	Mixed	4 (5.48)	10 (16.39)
	Unknown	3 (4.11)	10 (16.39)

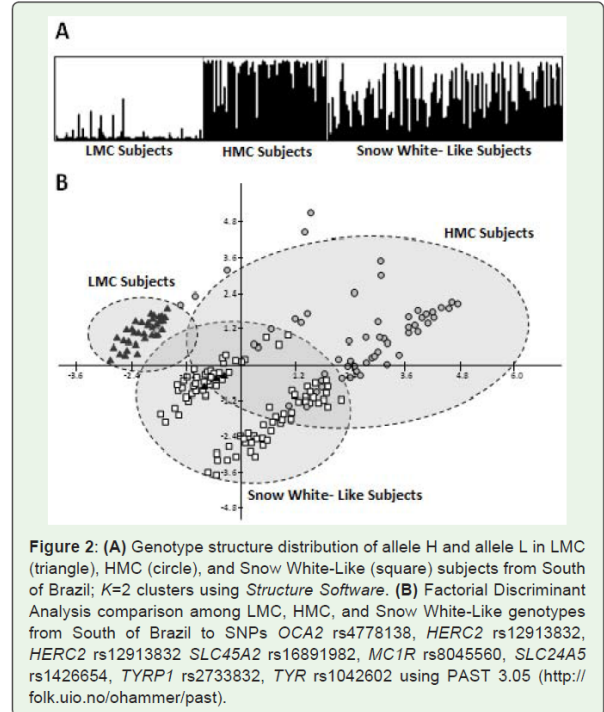


Figure 2: (A) Genotype structure distribution of allele H and allele L in LMC (triangle), HMC (circle), and Snow White-Like (square) subjects from South of Brazil;  $K=2$  clusters using *Structure Software*. (B) Factorial Discriminant Analysis comparison among LMC, HMC, and Snow White-Like genotypes from South of Brazil to SNPs *OCA2* rs4778138, *HERC2* rs12913832, *HERC2* rs12913832, *SLC45A2* rs16891982, *MC1R* rs8045560, *SLC24A5* rs1426654, *TYRP1* rs2733832, *TYR* rs1042602 using PAST 3.05 (<http://folk.uio.no/ohammer/past>).

color participant was classified and photographed and RGB values were measured in both eyes by COLOURS software [<https://www.colours-software-pvt-ltd>]. Phenotype data of these subjects are presented in table 3.

Table 4: Allele frequencies and other parameters of the selected SNPs in 134 South Brazilian subjects with HMC (N=61), or LMC (n=73) patterns.

RS (Gene; SNP)		HMC SB People <sup>1</sup>	LMC SB People <sup>2</sup>	P Value	Sensit	Specif	PV-1	PV-2	RR	OR
rs4778138 ( <i>OCA2</i> ; SNP G>A)	Allele H (G)	0.607 (74/122)	0.075 (11/146)	<0.0001	0.607	0.892	0.871	0.655	2.52	12.75
	Allele L (A)	0.393 (48/122)	0.925 (135/146)							
rs12913832 ( <i>HERC2</i> ; SNP A>G)	Allele H (A)	0.828 (101/122)	0.021 (3/146)	<0.0001	0.828	0.979	0.971	0.872	7.58	229.25
	Allele L (G)	0.172 (21/122)	0.979 (143/146)							
rs16891982 ( <i>SLC45A2</i> ; SNP C>G)	Allele H (C)	0.279 (34/122)	0.021 (3/146)	<0.0001	0.279	0.980	0.919	0.622	2.43	18.67
	Allele L (G)	0.721 (88/122)	0.993 (145/146)							
rs8045560 ( <i>MC1R</i> ; SNP C>T)	Allele H (C)	0.746 (91/122)	0.390 (57/146)	<0.0001	0.746	0.612	0.615	0.744	2.40	4.63
	Allele L (T)	0.254 (31/122)	0.616 (90/146)							
rs1426654 ( <i>SLC24A5</i> ; SNP G>A)	Allele H (G)	0.672 (82/122)	0.000 (0/146)	<0.0001	0.672	1.000	1.000	0.785	4.65	<1000.0
	Allele L (A)	0.328 (40/122)	1.000 (146/146)							
rs2733832 ( <i>TYRP1</i> ; SNP C>T)	Allele H (C)	0.770 (94/122)	0.479 (70/146)	<0.0001	0.770	0.521	0.573	0.731	2.13	3.64
	Allele L (T)	0.230 (28/122)	0.521 (76/146)							
rs1042602 ( <i>TYR</i> ; SNP C>A)	Allele H (C)	0.869 (106/122)	0.623 (91/146)	<0.0001	0.869	0.377	0.538	0.775	2.39	4.00
	Allele L (A)	0.131 (16/122)	0.377 (55/146)							
rs916977 ( <i>HERC2</i> ; SNP A>G)	Allele H (A)	0.697 (85/122)	0.027 (4/146)	<0.0001	0.697	0.973	0.955	0.794	4.65	82.13
	Allele L (G)	0.303 (37/122)	0.979 (143/146)							

Citation: Sawitzki FR, Rodenbusch R, Gubert DW, Prado MJ, Silva DSBS and Alho CS. Analysis of Eight SNPs in Pigment-Related Genes in South Brazilian Subjects with High or Low Skin and Eye Melanin Content. *SM J Forensic Res Criminol.* 2017; 1(2): 1008.



**Table 5:** Genotype frequencies and Chi-Square Test result of High Melanin Content (HMC), Low Melanin Content (LMC), HMC Subjects, LMC Subjects, and Snow White-Like (SW-L) subjects according HH (homozygote to allele H), LL (homozygote to allele L), and HL (heterozygote) per locus.

RS (Gene; SNP)		HMC	LMC	HMC Subj	LMC Subj	SW Subj
rs4778138 (OCA2; SNP G>A)	HH	0.55	0.03	0.36	0.00	0.15
	HL	0.38	0.27	0.49	0.15	0.54
	LL	0.06	0.70	0.15	0.85	0.31
	N	762	2913	61	73	116
	P			>0.0001	>0.0001	0.53
rs12913832 (HERC2; SNP A>G)	HH	0.92	0.08	0.75	0.00	0.47
	HL	0.08	0.40	0.15	0.04	0.00
	LL	0.00	0.52	0.10	0.96	116
	N	2344	14143	61	73	
	P			<b>0.262</b>	<0.0001	
rs16891982(SLC452A; SNP C>G)	HH	0.84	0.19	0.56	0.15	0.19
	HL	0.15	0.49	0.38	0.48	0.59
	LL	0.01	0.32	0.06	0.37	0.22
	N	787	1966	61	73	116
	P			>0.0001	<b>0.120</b>	
rs8045560 (MC1R; SNP C>T)	HH	0.84	0.00	0.05	0.15	0.19
	HL	0.15	0.13	0.46	0.48	0.59
	LL	0.01	0.87	0.49	0.37	0.22
	N	2038	13797	61	73	116
	P			>0.0001	<b>0.102</b>	
rs1426654 (SLC24A5; SNP G>A)	HH	0.52	0.00	0.46	0.00	0.00
	HL	0.40	0.02	0.43	0.00	0.20
	LL	0.08	0.98	0.11	1.00	0.80
	N			61	73	116
	P			>0.0001	<b>0.178*</b>	
rs2733832 (TYRP1; SNP C>T)	HH	0.84	0.17	0.58	0.22	0.25
	HL	0.15	0.49	0.39	0.52	0.55
	LL	0.01	0.34	0.03	0.26	0.20
	N	216	646	61	73	116
	P			>0.0001	<b>0.650</b>	
rs1042602 (TYR; SNP C>A)	HH	0.89	0.49	0.77	0.37	0.35
	HL	0.10	0.42	0.20	0.51	0.49
	LL	0.01	0.09	0.03	0.12	0.16
	N	1048	2573	61	73	116
	P			>0.0001	<b>0.845</b>	
rs916977 (HERC2; SNP A>G)	HH	0.79	0.04	0.49	0.00	0.24
	HL	0.20	0.34	0.41	0.05	0.47
	LL	0.01	0.62	0.10	0.95	0.29
	N	627	1819	61	73	116
	P			>0.0001	>0.0001	

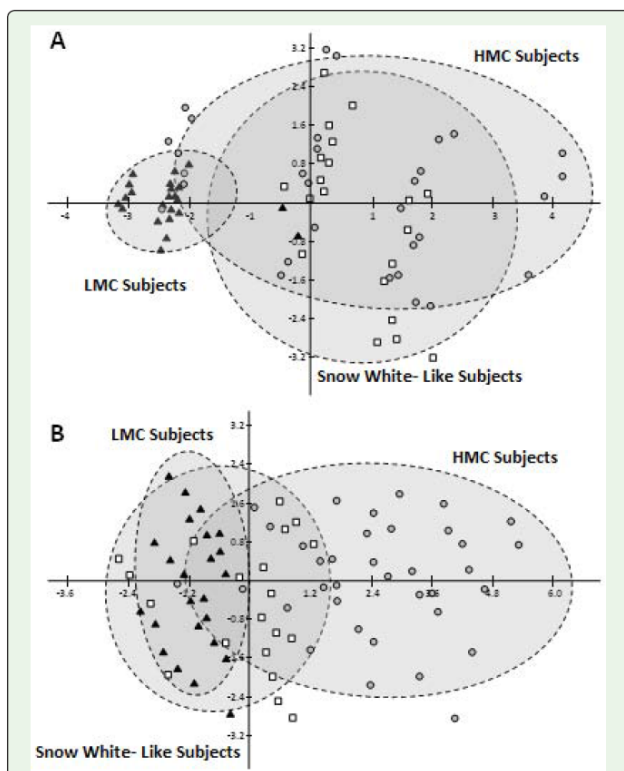
\*The association chi-square test was calculated with the allele data.

There was a highly significant association between alleles H and HMC Brazilian subjects, and alleles L and LMC subjects. All other parameters (sensitivity, specificity, predictive values, relative ratio, odds ratio) were equally remarkable (Table 4). To observe the cumulative distribution of alleles H and alleles L in the two phenotypically different color groups, we plotted the frequency data and noticed LMC subjects presented between zero and seven alleles H, while HMC subjects presented between four and 14 alleles H. No subject had 15 or 16 alleles H; the average of alleles H among the LMC subjects was 3.3 (SD=1.49), 96% (70/73) of them had less than five alleles H, and 100% had seven or less alleles H. The average of alleles H among the HMC subjects was 12.0 (SD=3.78), 82% (56/61) of them had eight or more alleles H. We also measured the ability of the SNP panel correctly classify those with and without the phenotypic trait (pigmentation) and obtained very high AUC (area under the receiver operating characteristic curve) value of 0.9908 in the prediction of both HMC and LMC phenotypes (SE= 0.0088; IC 95% = 0.99736-1.0). The weighted Pathway Genetic Load (PGL) scores, which tested multiple loci cumulative effect on the phenotype [11], showed HMC subjects ranged from 65.29 to 870.92 (median= 567.37; mean= 584.63; SD=262.46) and LMC from zero to 148.40 (median= 38.31; mean= 44.80; SD=30.93). There was a significant difference between the two groups (P<0.0001; Mann-Whitney Test). Genotype Probability (GP) in all 134 South Brazilian subjects showed that 80% (49/61) of HMC genotypes had the highest chance to belong to HMC group (rates from 1.95 to 8.27x10<sup>7</sup>; mean= 7.92 x 10<sup>6</sup>), and 100% (73/73) of LMC genotypes had the highest chance to belong to the LMC group (rates from 1.46x10<sup>4</sup> to 7.77x10<sup>12</sup>; mean= 2.02 x 10<sup>11</sup>). This GP analysis using the current eight SNPs genotyping system showed 91% (122/134) concordance between predicted and observed phenotype (methodology details are available on the supplemental file).

1- HMC SB People: South Brazilian subjects with dark skin and eyes. 2- LMC SB People: South Brazilian subjects with Light skin and blue eyes. We called as allele H for the allele strongly associated with people with high melanin content (HMC; from African populations, see table 1), and as allele L for the allele strongly associated with people with low melanin content (LMC; from European populations, see Table 1). Sensit: sensitivity. Specif: specificity. PV-1: predictive value to hazard of HMC people to have allele H. PV-2: predictive value to hazard of LMC people to have allele L. RR: Relative Ratio. OR: Odds Ratio.

It is still unclear when a pigment-related gene variant is expressed in both skin and iris, or when it has tissue-specific expression. In order to identify differences on genotype-phenotype associations, we studied 116 Snow White-Like subjects (SW – light skin and dark eyes) from South Brazil. Genotype data from HMC, LMC and SW subjects were verified by structuring [Suppl 4] and Factorial Discriminant Analysis (FDA). Snow White-Like subjects were grouped in an intermediate cluster, as expected since they have light skin as LMC subjects and dark eyes as HMC subjects (Figure 2).

Association Chi-square test was performed comparing SW genotypes with genotype distribution in ALFRED-HMC, ALFRED-LMC, HMC, and LMC Southern Brazilian samples. There were no significant differences between SW genotype frequencies and HMC data in *HERC2* rs12913832 (both have dark eyes), and there were no significant differences between SW genotype frequencies and LMC



**Figure 3:** Factorial Discriminant Analysis comparison among LMC (triangle), HMC (circle), and Snow White-Like (square) genotypes using PAST 3.05 (<http://folk.uio.no/ohammer/past>) to SNPs (A) *OCA2* rs4778138, *HERC2* rs12913832, *HERC2* rs12913832 and (B) *SLC45A2* rs16891982, *MC1R* rs8045560, *SLC24A5* rs1426654, *TYRP1* rs2733832, *TYR* rs1042602.

data in *SLC45A2* rs16891982, *MC1R* rs8045560, *SLC24A5* rs1426654, *TYRP1* rs2733832, *TYR* rs1042602 (both have light skin) (Table 5).

Based on these results, two definitive FDA were performed using two separate clusters of loci to group the ones able to express melanin in iris (*HERC2* rs12913832, *OCA2* rs4778138, *HERC2* rs916977, we added *OCA2* and *HERC2* SNPs since they are strongly linked to rs12913832), and the ones able to express melanin in skin (*SLC45A2* rs16891982, *MC1R* rs8045560, *SLC24A5* rs1426654, *TYRP1* rs2733832, and *TYR* rs1042602) (Figure 3). SNPs *HERC2* rs12913832, *OCA2* rs4778138, *HERC2* rs916977 clustered 90% of SW subjects with HMC group and *SLC45A2* rs16891982, *MC1R* rs8045560, *SLC24A5* rs1426654, *TYRP1* rs2733832, *TYR* rs1042602 clustered 82% of SW subjects with LMC group.

## Conclusion

The LMC - HMC - SW model was used to analyze eight biallelic SNPs in seven pigment-related genes in an admixture South Brazilian samples. Data revealed that SNPs rs16891982 (*SLC45A2*), rs8045560 (*MC1R*), rs1426654 (*SLC24A5*), rs2733832 (*TYRP1*), rs1042602 (*TYR*) were associated simultaneously with LMC and Snow White

clusters, indicating a link of these SNPs with skin phenotype. SNPs rs4778138 (*OCA2*), rs12913832 (*HERC2*), rs916977 (*HERC2*) were associated with HMC and Snow White clusters, demonstrating a link with eye phenotype. The understanding of gene tissue-specific expression in the externally visible characteristics tissue-specific is important for the prediction of skin and eye color in humans; we believe our analysis is an important contribution to the forensic DNA phenotyping scenario.

## Supplemental File

Methodology details are available on the Supplemental File.

## Acknowledgments

This work was supported by CAPES Brazil | Edital 25/2014 | Pró-Forenses, and CNPqBrasil | Edital INCT. Fernanda Rosa Sawitzki, Rodrigo Rodenbusch, Diego WordellGubert, Eduardo Felipe Ávila da Silva and Deborah Soares Bispo Santos Silva were supported by CAPES. The authors would like to thank to Mayara Jorgens Prado, Bruna Schroeder, PietraGraebin and Eduardo Wildemann Capelesso for collection and organization of phenotype data.

## References

- Kayser M. Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Sci Int Genet.* 2015; 18: 33-48.
- 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: 424-431.
- Kondo T, Hearing VJ. Update on the regulation of mammalian melanocyte function and skin pigmentation. *Expert Rev Dermatol.* 2011; 6: 97-108.
- 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: 446-455.
- Valenzuela RK, Ito S, Wakamatsu K, Brilliant MH. Prediction model validation: normal human pigmentation variation. *J Forensic Res.* 2011; 2: 139.
- Liu F, Wenb B, Kayser M. Colorful DNA polymorphisms in humans. *Semin Cell Dev Biol.* 2013; 24: 562-575.
- Gerstenblith MR, Goldstein AM, Fargnoli MC, Peris K, Landi MT. Comprehensive evaluation of allele frequency differences of *MC1R* variants across populations. *Human Mutat.* 2007; 28: 495-505.
- Vierkötter A, Krämer U, Sugiri D, Morita A, Yamamoto A, Kaneko N, et al. Development of lentigenes in German and Japanese women correlates with variants in the *SLC45A2* gene. *J Invest Dermatol.* 2012; 132: 733-736.
- Alaluf S, Atkins D, Barrett K, Carter N. Ethnic variation in tyrosinase and *TYRP1* expression in photoexposed and photoprotected human skin. *Pigment Cell Res.* 2003; 16: 35-42.
- Chaki M, Sengupta M, Mondal M, Bhattacharya A, Mallick S, Bhadra R, et al. Molecular and functional studies of tyrosinase variants among Indian oculocutaneous albinism type 1 patients. *J Invest Dermatol.* 2011; 131: 260-262.
- Huebinger RM, Garner HR, Barber RC. Pathway genetic load allows simultaneous evaluation of multiple genetic associations. *Burns.* 2010; 36: 787-792.

## 1- SUPPLEMENTAL FILE - MATERIAL AND METHODS

### ***Design, subjects, and approval.***

Buccal cells were collected using cotton swabs from random adult subjects (over 18 years old). Subjects were selected based on their skin and eye phenotype characteristics. This COLOR-genotyping project was approved by the Research Ethics Committee of the Pontifical Catholic University of Rio Grande do Sul (protocol #11-05722, Of.CEP-0295/12 and Of.CEP-1041/12), and the informed written consent to participate was obtained from all subjects.

### ***Skin Phenotyping.***

The skin color of each participant was identified using the Fitzpatrick score as Type 1 to Type 6 [1]. We also measured the skin amount of red (R), green (G), and blue (B) pigments in an inner (below elbow) and hairless portion of right arm using a battery-portable color analyzer equipment ACR-1023 (Instrutherm, São Paulo, Brazil) which employs a spectral analysis method to determine the color of the sample. In this system, each R, G, and B values range scale from zero to 1023; the total white color has R=1023, G=1023, and B=1023, and total black has R=0, G=0, and B=0. The RGB skin values were measured three times per person and the average was recorded. The equipment was calibrated before each use measuring a white plate. Subjects that reported having a recent or intense tan were excluded. During data collection, we used standardized forms and templates (see below).

### ***Eye Phenotyping.***

The eye color of each participant was identified using scores as Type 1 to Type 6: 1- light blue; 2- dark blue; 3- green; 4- hazel or light brown; 5-dark brown; 6-black. We photographed both eyes using a Nikon COOLPIX L110, with MacroZoom, ISO 80, with Flash at 5,500K (D55 illuminant) taking a picture with resolution of 4,000 X 3,000 pixels. To take a picture, the camera was fixed on a tripod with a distance of 6cm  $\pm$  0.5cm from the eye, and the focus was measured on the iris central area. The images were taken in a room with standardized artificial light conditions. Participants using contact lenses were required to remove them. We measured both eye amounts of R, G, and B pigments analyzing each photo with the software COLOURS [2]. A representative triangular area of the iris covering approximately 2,500 pixels was selected in each picture and the R, G, and B values of each pixel was measured in a range scale from zero

1 to 255, where the maximum value (255) was white and minimum value (zero) was black. To  
2 calculate the R value, the software uses the sum of red in each pixel of the selected area divided  
3 by 255 times the total of pixels in the selected area. For example (SUP-FIGURE 1): total of red in  
4 the selected area (26,4002) / 255 x total of pixels in the area (2542) = 26,4002 / (255 x 2,542) =  
5 0.41. The same mathematical formula is used to G and B calculations. In this system, the total  
6 white has R=255, G=255, and B=255 and black color has R=0, G=0, and B=0. The RGB eye values  
7 were measured in both eyes and the average was recorded. Using standard formulas, we  
8 converted the primary RGB values to obtain the values in the HSV (hue, saturation, value of  
9 brightness) system. HSV is a cylindrical-coordinate representation of points based in the RGB  
10 color model. This representation maps the values into a cylinder (color wheel); the angle around  
11 the central vertical axis corresponds to hue (H) and the distance from the axis corresponds to  
12 saturation (S). The height corresponds to a third value (V), the system's representation of the  
13 perceived luminance (brightness) in relation to the saturation. Potential subjects with  
14 *Heterochromia iridum*, when one iris has an excessive different color from the other, were  
15 excluded from the experiment.

16

### 17 **Genotyping.**

18

19 Genomic DNA was extracted with a NucleoSpin<sup>®</sup> Tissue kit (Macherey-Nagel Inc).  
20 Amplification of regions flanking the SNPs was done by multiplex PCR in a volume of 25 µL with  
21 2 to 10 ng genomic DNA, 0.2 µM of each primer (Table 1), and 1X of Qiagen Multiplex PCR Master  
22 Mix (Qiagen Inc, CA). Amplification consisted of 95°C for 5 min, followed by 30 cycles of 30 s at  
23 94°C, 90 s at 57°C, 90 s at 72°C, and a final extension for 10 min at 72°C. After the amplification,  
24 we performed enzymatic purification of the PCR products using USB<sup>®</sup> ExoSAP-IT<sup>®</sup> PCR Product  
25 Cleanup (Affymetrix<sup>®</sup>) according to the user's manual recommendation. For SNP analysis,  
26 SNaPshot<sup>®</sup> Multiplex System ABI Prism (Applied Biosystems<sup>®</sup>, São Paulo, Brazil) was used with  
27 eight SNaPshot primers. Reactions were performed in a final volume of 10 µL, containing 3.0 µL  
28 of purified multiplex PCR product, 5.0 µL of SNaPshot Multiplex Ready Reaction Mix, 1.0 µL of  
29 pooled SNaPshot primers (each primer at 1 µM final), and 1.0 µL of sterile water. Multiplex single  
30 base extensions were carried out for 28 cycles according to the following program: 10 seconds  
31 at 96°C, 5 seconds at 50°C, and 30 seconds at 60°C. SNaPshot products were then treated at  
32 37°C for 1 hour with 0.8 µL of shrimp alkaline phosphatase (1.0 U/µl) and 1.2 µL 10X SAP buffer  
33 reaction, added directly to 10 µL of the SNaPshot product. After heat inactivation of shrimp  
34 alkaline phosphatase for 15 minutes at 75°C, 1µL of the labeled products were mixed with 8.5  
35 µL of HiDi formamide and 0.5 µL of Genescan-120 LIZ size standard. They were then separated  
36 using an ABIPRISM<sup>®</sup> 3130xl Genetic Analyzer (Applied Biosystems) with POP-4 matrix and the

1 respective run parameters: injection voltage of 1.2 Kv, injection time of 23s, run voltage of 15Kv  
2 and run time 1200s, in a capillary of 36cm. The analysis was performed using GeneMapper ID  
3 software version 3.2.1 (Applied Biosystems). In order to confirm the genotyping system, we  
4 performed a sequencing analysis in ABIPRISM<sup>®</sup> 3130xl Genetic Analyzer (Applied Biosystems)  
5 using the same designed primers in 10% of our samples.

6

7 ***Candidate SNP parameters.***

8

9 At each locus, we dubbed 'allele H' each of the alleles strongly associated with subjects  
10 from African populations (individuals with High Melanin Content; HMC), and 'allele L' the alleles  
11 strongly associated with subjects from European populations (individuals with Low Melanin  
12 Content; LMC). Based on the SNP frequencies per locus available on the ALFRED database, we  
13 constructed a 2 x 2 table with number of alleles, where: the True Positive (TP; correctly  
14 identified) was allele H in HMC subjects, True Negative (TN; correctly rejected) was allele L in  
15 LMC subjects, False Positive (FP; incorrectly identified or Type I error) was allele H in LMC  
16 subjects; False Negative (FN; incorrectly rejected or Type II error) was allele L in HMC subjects.  
17 We then calculated the following parameters: Sensitivity (ability to identify positive results) = TP  
18 / (TP + FN); Specificity (ability to identify negative results) = TN / (FP + TN); Predictive value to  
19 allele H in HCM subjects = TP / (TP + FP); Predictive value for the presence of allele L in LMC  
20 subjects = TN / (FN + TN); Relative risk (ratio of the probability of allele H in HMC group to the  
21 probability of allele H in LMC group) = [TP / (TP + FP)] / [FN / (FN + TN)]; and Odds ratio (to  
22 identify the probability of exposure and non-exposure) = (TP x TN) / (FP x FN). We calculated the  
23 sensitivity, specificity and accuracy values joining the set of all 16 alleles. The Accuracy value, or  
24 the detection of how the test correctly identifies or excludes the predicted phenotype, was  
25 calculated as = (TP + TN) / (TP + FP + TN + FN).

26

### **Area under the Receiver Operating Characteristic curve.**

Using the threshold of 8 alleles H, the sensitivity, specificity, and accuracy values of our system were, respectively, 91.8% (56/61), 100% (73/73), and 96.3% (129/134). However, if the decision about the threshold were changed and a new threshold in 7 or 6 alleles H were used, for example, we would obtain different sensitivity, specificity, and accuracy values. Thus, we changed the decision threshold from zero to 16 alleles H and plotted the "sensitivity" values as "x" and "1 - specificity" values as "y" on a graph. The axes of this graph both range from zero to one because these are the limits of all possible combinations of "sensitivity" and "1 - specificity" values. The points of the curve inevitably passed through the graph origin (where "sensitivity" = 0 and "1 - specificity" = 0) and through the upper right corner of the graph (where "sensitivity" = 1 and "1 - specificity" = 1). The resulting graph was a curve which is called Receiver Operating Characteristic (ROC) curve since "it describes the inherent detection *characteristics* of the test" and since "the *receiver* of the test information can *operate* at any point on the curve by using an appropriate decision threshold" [3]. Based on the same assessment that Liu *et al.* (2009) [4] made in their work, we measured the ability of the set of SNPs to correctly classify those with and without the phenotypic trait evaluating the area under the ROC curve (AUC). A Receiver Operating Characteristic (ROC) curve was constructed by computing the sensitivity and specificity of the set of all 16 SNPs. The accuracy in this test was measured by the area under the ROC curve (AUC) with a parametric method using a maximum likelihood estimator. The AUC measures the ability of the set of SNPs to correctly classify those with and without the phenotype, as an overall measure for prediction accuracy. An area of 1 represents a perfect test; an area of 0.5 represents a worthless test. A recommended guide for classifying the accuracy would be 0.90-1 = excellent; 0.80-0.90 = good; 0.70-0.80 = fair; 0.60-0.70 = poor; 0.50-0.60 = fail.

### **Calculation of Pathway Genetic Load (PGL).**

Genotypes were scored as zero for the LL homozygotes, 1 for heterozygotes, and 2 for the HH homozygotes. Unweighted PGL was calculated by summing the numerical scores for genotypes across all SNPs as described in Huebinger *et al.* [5]. Individual PGL scores potentially ranged from zero (homozygous for alleles L at all loci) to 16 (homozygous for alleles H at all loci). PGL weighted scores for color were obtained by multiplying the allele H count (HH=2.0, HL=1.0, LL=0.0) per *locus* by its adjusted odds ratio (OR) – previously determined by ALFRED data. Weighted PGL scores per locus were then computed creating individual PGL scores. For example: to the following loci rs4778138 (*OCA*), rs12913832 (*HERC2*), rs16891982 (*SLC452A*), rs8045560

1 (*MC1R*), rs1426654 (*SLC24A5*), rs2733832 (*TYRP1*), rs1042602 (*TYR*), and rs916977 (*HERC2*), the  
2 complete genotype of subject HC123 was, respectively: LL – LL – LL – HL – LL – HL – HH – LL.  
3 Using respective OR values by locus, we obtained to HC123 subject this values: PGL-WS= (0.0 x  
4 14.81) + (0.0 x 54.26) + (0.0 x 103.28) + (1.0 x 13.38) + (0.0 x 226.12) + (1.0 x 15.47) + (2.0 x 7.37)  
5 + (0.0 x 29.73) = 43.59. PGL and other continuous data were compared by the Mann-Whitney U  
6 test. Categorical data was compared using chi-square. The BioEstat 5.0 Statistical Package was  
7 used for all calculations.

### 8 9 ***Calculation of Genotype Probability (GP).***

10  
11 We applied the follow model to predict color in the HMC or LMC phenotype categories  
12 based on the genotype probability (GP). From ALFRED allele frequencies, we calculated each  
13 expected genotype frequency to HH, HL, and LL genotypes, per locus, in HMC and LMC groups.  
14 We examined the complete genotype of each Cinderella-Like or Tiana-Like South Brazilian  
15 subject and calculated the probability of the complete genotype belonging to expected HMC  
16 and LMC groups. For example: to the following loci rs4778138 (*OCA*), rs12913832 (*HERC2*),  
17 rs16891982 (*SLC452A*), rs8045560 (*MC1R*), rs1426654 (*SLC24A5*), rs2733832 (*TYRP1*),  
18 rs1042602 (*TYR*), and rs916977 (*HERC2*) the complete genotype of subject HC123 was,  
19 respectively: LL – LL – LL – HL – LL – HL – HH – LL. Using each genotype frequency by locus  
20 obtained from HMC population, we had the following values for subject HC123: 0.06554 -  
21 0.00185 - 0.01416 - 0.15555 - 0.0756 - 0.15222 - 0.89492 - 0.01210. The product of these values  
22 was  $3.33 \times 10^{-11}$ , which means the probability of this complete genotype is present in the HMC  
23 population. Then, using each genotype frequency by locus obtained from LMC population, we  
24 had the following values for subject HC123: 0.69890 - 0.50694 - 0.87049 - 0.49280 - 0.97812 -  
25 0.48555 - 0.49280 - 0.61780. The product of these last values was: 0.022, which means the  
26 probability of this complete genotype is present in the LMC population. The genotype of the  
27 HC123 subject, by ratio of two probabilities, has  $6.6 \times 10^8$  times higher chance of belonging to  
28 the LMC population than to the HML population. This finding was in agreement with the actual  
29 HC123 phenotype data obtained: a male with light skin color (Type 2), dark blue eyes (Type 2),  
30 high RGB values to skin (317-239-200) and eyes (0.385-0.365-0.530), and of European origin. We  
31 applied this model to predict color in the HMC or LMC phenotype categories based on the  
32 genotype probability (GP) in all 134 South Brazilian subjects; we observed that 80% (49/61) of  
33 the genotypes of our HMC subjects had a higher chance of belonging to the HMC group than to  
34 the LMC group, with rates from 1.95 to  $8.27 \times 10^7$  (mean=  $7.92 \times 10^6$ ). Analyzing the genotypes  
35 of our LMC subjects, we noticed that 100% (73/73) had a higher chance of belonging to the LMC  
36 group, with rates from  $1.46 \times 10^4$  (in subject HC120, with seven alleles H) to  $7.77 \times 10^{12}$  (mean=

1 2.02 x 10<sup>11</sup>). This GP analysis using the current eight SNPs genotyping system showed 91%  
2 (122/134) concordance between predicted and observed HMC or LMC phenotype.

#### 3 4 ***Cluster analysis using STRUCTURE Software.***

5  
6 We used the eight SNPs and the STRUCTURE software to estimate clusters among  
7 subjects. The STRUCTURE Software v2.3.4 (<http://pritch.bsd.uchicago.edu/structure.html>)  
8 analyses the data plotting probabilistically the individuals in  $K$  populations by characterization  
9 of the allele frequencies set in each locus. The individuals may be assigned in one, two or more  
10 populations if their genotype shows an admixed pattern. Our data file was composed by: row of  
11 makers (SNP and gene data), column of labels of each individual, column of population data (1  
12 for HMC; 2 for LMC), and the genotype data. The ADMIXTURE MODEL and CORRELATED ALLELE  
13 FREQUENCIES MODEL were chosen because the admixture between populations is a common  
14 characteristic of real genetic data, since subjects may have recent ancestors in more than one  
15 population [6]. To estimate the  $K$  value, the data was analyzed with twenty replicates for  $K=1$  to  
16  $K=10$ , all runs were performed with 10,000 burn-in period and 100,000 Markov Chain Monte  
17 Carlo (MCMC) repeats after burn-in, ADMIXTURE MODEL, CORRELATED ALLELE. The  $K$  values  
18 were estimated using the  $\ln\Pr(X|K)$  values for each simulation. With these values, it is possible  
19 to calculate the probability for each  $K$  and estimate the best  $K$  [7]. The method used to estimate  
20 the value of  $K$  calculates the delta  $K$  and then selects the appropriate  $K$  value [8, 9]. To estimate  
21  $K$  value the data was analyzed with twenty replicates for  $K=1$  to  $K=10$ , all runs were performed  
22 with 10,000 burn-in period and 100,000 Markov Chain Monte Carlo (MCMC) repeats after burn-  
23 in, ADMIXTURE MODEL, CORRELATED ALLELE [8]. Graphics were constructed using *Clumpp* [10]  
24 and *Distruct* [11] software.

#### 25 26 ***Factorial Discriminant Analysis (FDA).***

27  
28 An exploratory analysis discriminant function analysis is useful in determining whether  
29 a set of variables is effective in predicting category membership. In this study, discriminant  
30 function analysis (multiple; MANOVA) was used to know whether the SNPs set were able to  
31 correctly predict the phenotype. Cinderella-Like and Tiana-Like phenotypes was used as  
32 categorical dependent variables, and genotypes were used as multiple independent variables  
33 (predictor variables). Since Cinderella-Like and Tiana-Like subjects were known a priori, each  
34 genotype has a score per locus, and a score on a group measure. Then, the categories of the  
35 same type were distributed and were ascertained how the groups differ on the composite of  
36 dependent variables. A significant F test allows classification based on a linear combination of



1 predictor variables to eye color and to skin color. We included all SNP variables (16 alleles per  
2 person) and phenotype categories to eye and skin groups.

3

#### 4 **2- SUPPLEMENTAL FILE - REFERENCES**

5

6 1- Pathak MA. In memory of Thomas Bernhard Fitzpatrick. *J Invest Dermatol* 2004;122:20-1.

7 2- Otaka I, Kumagai K, Inagaki Y, Shimoyama M, Saegusa K, Hara T. (2002) *Am J Ophthalmol*  
8 133(1):140-142.

9 3- Metz CE (1978) Basic principles of ROC analysis. *Semin Nucl Med* 8(4):283-298

10 4- Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens ACJW, Kayser M (2009)  
11 Eye color and the prediction of complex phenotypes from genotypes. *Curr. Biol.*

12 19:R192–R193

13 5- Huebinger RM, Garner HR, Barber RC (2010) Pathway genetic load allows simultaneous  
14 evaluation of multiple genetic associations. *Burns* 36(6):787-792.

15 doi:10.1016/j.burns.2010.02.001

16 6- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using  
17 multilocus genotype data. *Genetics* 155:945–959.

18 7- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure: Extensions to  
19 linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.

20 8- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using  
21 the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620.

22 9- Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing  
23 STRUCTURE output and implementing the Evanno method. *Conservation Genetics*  
24 Resources 4(2): 359-361.





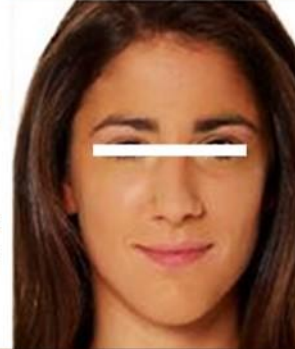
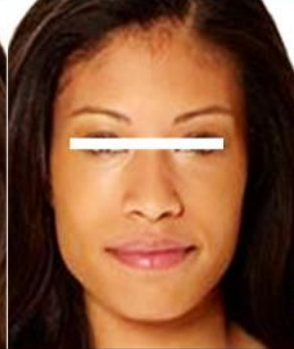






25 10- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program  
26 for dealing with label switching and multimodality in analysis of population structure.  
27 *Bioinformatics* 23:1801-1806.

28 11- Rosenberg NA (2004) Distruct: a program for the graphical display of population structure.  
29 *Molecular Ecology Notes* 4:137-138.

30

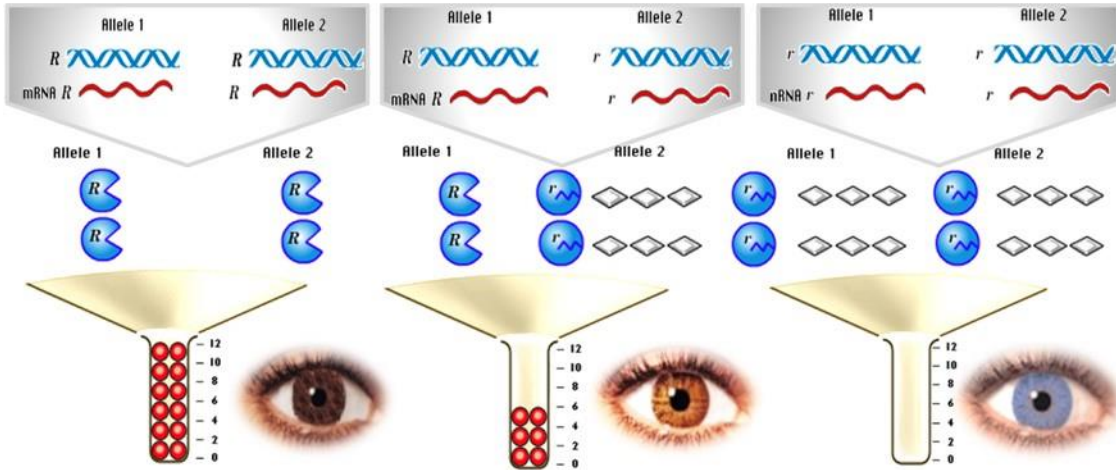
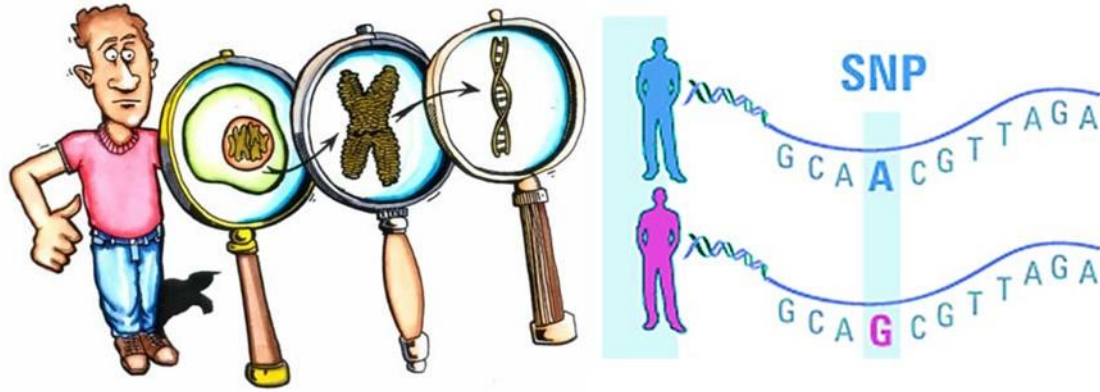
31

Forensic DNA Phenotyping Project – BRAZIL

			
<b>1- LIGHT</b> 1- White or very pale Always burns, never tans		2- Pale White with beige tint Always burns, sometimes tans	
<b>1- LIGHT</b>			
			
<b>2- MEDIUM</b> 3- Beige to light brown Sometimes burns, tans		4- Light to moderate brown Rarely burns, Always tans	
<b>2- INTERMED</b>			
			
<b>3- DARK</b> 5- Medium to dark brown Rarely burns, tans more		6- Dark brown to black Never burns	
<b>3- DARK</b>			

(Dr. Alho +55 51 33534334; csalho@puccs.br; The images are merely illustrative)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36



The images are merely illustrative



## FORENSIC DNA PHENOTYPING - BRAZIL | INSTRUMENTO DE DADOS

DATA DA COLETA: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

IDENTIFICAÇÃO:

NÚMERO			INICIAIS	

### DADOS PESSOAIS DO PARTICIPANTE

Nome: \_\_\_\_\_ Data de Nascimento: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

Contato (E-mail; telefones): \_\_\_\_\_

### DADOS ANTROPOMÉTRICOS / CARACTERÍSTICAS / INFORMAÇÕES

Altura: \_\_\_\_\_ Peso: \_\_\_\_\_ N° do sapato: \_\_\_\_\_

Ao usar a mão para escrever é: ( ) Destro ( ) Canhoto

A orelha é saliente? ( ) SIM ( ) NÃO Lóbulo orelha é solto? ( ) SIM ( ) NÃO

Está usando lente de contato? ( ) SIM ( ) NÃO Está bem bronzeado? ( ) SIM ( ) NÃO

Cabelo pintado/descolorido...? ( ) SIM ( ) NÃO ...Alisado/Encrespado? ( ) SIM ( ) NÃO

### DADOS EPIDEMIOLÓGICOS

Sexo: ( ) Masc ( ) Fem

Origem: ( ) Europeia ( ) Africana ( ) Indígena ( ) Asiática ( ) Árabe\* ( ) Outra: \_\_\_\_\_

Residente em: ( ) Sul ( ) Sudeste ( ) Centroeste ( ) Nordeste ( ) Norte ( ) Estado: \_\_\_\_\_

Natural de: ( ) Sul ( ) Sudeste ( ) Centroeste ( ) Nordeste ( ) Norte ( ) Estado: \_\_\_\_\_

(\*Árabes: Habitantes da Península Arábica, Oriente Médio, África setentrional).

### DADOS BIOLÓGICOS - ORIGINAIS (NATURAIS)

OLHO COR: ( ) Azul ( ) Verde ( ) Mel ( ) Marron ( ) Preto

PELE COR: ( ) I-Leite ( ) II-Bege ( ) III-Cuia ( ) IV-Chocolate ( ) V-Café ( ) VI-Petróleo

CABELO COR: ( ) Loiro Plat ( ) Loiro Amar ( ) Loiro Esc ( ) Castanho Claro ( ) Cast Esc ( ) Negro

( ) Ruivo Claro ( ) Ruivo Esc

CABELO TIPO: ( ) Liso ( ) Ondulado ( ) Crespo ( ) Molinha

CABELO QTDE: ( ) Muito ( ) Médio ( ) Pouco CABELO OBSERV.: ( ) Calvície ( ) Grisalho

( ) Branco ( ) \_\_\_\_\_

### PARA HOMENS

QTD DE PELO EM FACE/PEITO/COSTAS/PERNAS/BRAÇOS: ( ) Muito ( ) Médio ( ) Pouco

### COLETADOR

Nome COLETADOR: \_\_\_\_\_

Matrícula PUCRS: \_\_\_\_\_

Como coletador, declaro que colhi e identifiquei corretamente todas as amostras de células da pessoa aqui indicada para fins de investigação de dados fenotípicos, tendo seguido rigorosamente os procedimentos indicados. Para tanto assino ao lado:

\_\_\_\_\_  
Assinatura COLETADOR

### OBSERVAÇÕES

(SE FOR O CASO: DESENHE O HEREDOGRAMA DA FAMÍLIA À QUAL O PARTICIPANTE PERTENCE, IDENTIFICANDO-O COM UMA SETA).



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

**MANUSCRITO | ARTIGO CIENTÍFICO 2 | FULL-LENGTH COMMUNICATION**

**NEW SET OF DNA MARKERS AND STATISTIC MODEL TO PREDICT SKIN AND EYE COLORS IN SOUTH BRAZILIAN SUBJECTS**

Fernanda Rosa Sawitzki<sup>1</sup>, Gustavo Adolfo Silva-Arias<sup>1</sup>, Deborah Soares Bispo Santos Silva<sup>1,2</sup>, Carlos Eduardo Ibaldo Gonçalves<sup>1,3</sup>, Maria da Graça Bicalho<sup>3</sup>, Marcelo Malaghini<sup>4</sup>, Eduardo Filipe Avila Silva<sup>1,5,6</sup>, Clarice Sampaio Alho<sup>1,6</sup>

<sup>1</sup>Laboratório de Genética Forense, PUCRS, Porto Alegre-RS, Brasil.

<sup>2</sup>Forensic Sciences Institute, North Carolina State University, USA.

<sup>3</sup>Laboratório de Imunogenética e Histocompatibilidade UFPR, Curitiba-PR, Brasil.

<sup>4</sup>Laboratório de Genética Molecular Forense, ICPR, Curitiba-PR, Brasil.

<sup>5</sup>Departamento de Polícia Federal, Porto Alegre-RS, Brasil.

<sup>6</sup>Instituto Nacional de Ciência e Tecnologia Forense, Porto Alegre-RS, Brasil.

**CORRESPONDENCE TO**

Clarice Sampaio Alho, Laboratório de Genética Forense, PUCRS. Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil. (+55) (51) 3320-3568. csalho@pucrs.br

**ABBREVIATIONS**

- AIC: Akaike information criterion
- DNA: deoxyribonucleic acid
- EVC: externally visible characteristics
- FDP: forensic DNA phenotyping
- HERC2: HECT domain and RCC1-like domain-containing protein 2
- HMC: high melanin content
- HSV: hue, saturation, value of brightness
- LMC: low melanin content
- MC1R: melanocortin 1 receptor
- OCA2: Oculocutaneous albinism type II
- PCR: polymerase chain reaction
- RGB: red, green, blue
- SAP: shrimp alkaline phosphatase
- SLC24A5: Solute Carrier Family 24, member 5
- SLC45A2: Solute Carrier Family 45, member 2
- SNP: single nucleotide polymorphism
- SW: snow white
- TYR: Tyrosinase
- TYRP1: Tyrosinase-related protein 1

1           **ABSTRACT**

2

3           Genes related to externally visible characteristics (EVC) present a great importance for

4 the prediction of skin and eye colors in humans. In this study, we tested the ability of a set of

5 SNPs in pigment-related genes to predict skin and eye colors in an admixed south Brazilian

6 population. We studied 436 South Brazilian subjects with different skin and eye colors to

7 construct a multinomial logistic regression model able to predict phenotype. The model was

8 tested in 40 random subjects to measure the efficiency of the prediction. The data presented

9 93% concordance between the predicted and the observed phenotype, showing the successful

10 phenotype prediction of the model in South Brazilian individuals. This is important since Brazil

11 has an ethnically mixed population in which genomic structure may favor epistatic and

12 pleiotropic effects not observed in populations that are more homogeneous.

13

14

15           **KEYWORDS**

16

17           Human Pigmentation; Phenotype; Eye color; Skin color; SNP; SNaPshot

18

19

20

21

## 1 INTRODUCTION

2  
3 Forensic DNA phenotyping (FDP) is a novel field within forensic genetics that aims to  
4 predict appearance traits of an individual based on a DNA sample. An efficient FDP technology  
5 would be especially useful in cases where the investigators have no leads to identify the  
6 individual [1-3]. Skin and eye colors are important externally visible characteristics (EVCs), since  
7 they are highly heritable genetic traits and are the most obvious and distinguishable EVCs to be  
8 used for human identification. Genotype-phenotype association studies have focused in  
9 detecting single nucleotide polymorphisms (SNPs) within genes directly or indirectly involved in  
10 the synthesis of pigments [1-4], which can be used for determining human pigmentation traits  
11 and predicting eye, hair, and/or skin colors.

12 In this work, SNPs were selected in the following human pigmentation associated genes:  
13 *HERC2* (rs12913832), *OCA2* (rs4778138), *SLC24A5* (rs1426654), *SLC45A2* (rs16891982), *TYRP1*  
14 (rs2733832), *TYR* (rs1042602), *SLC45A2* (rs26722), *HERC2* (rs916977) and *MC1R* (rs8045560);  
15 these SNPs were previously shown to significantly alter the melanin content in human  
16 melanocytes, supporting their functional role in pigmentation [5]. We aimed to investigate this  
17 set of SNPs and their efficiency to predict skin and eye colors for forensic purposes in individuals  
18 with different phenotypes from South Brazil.

## 21 MATERIALS AND METHODS

22  
23 Buccal swabs were collected from 436 South Brazilian population (ages ranging from 5  
24 to 70 years old) who, according to the latest census, consists of 83% Caucasian, 4% African  
25 descendants, 12% mulatto (term used in Brazil to describe an individual with mixed European  
26 and African lineage), 1% other ethnicities (<http://censo2010.ibge.gov.br>). This project was  
27 approved by the Research Ethics Committee (CEP) of the Pontifical Catholic University of Rio  
28 Grande do Sul (protocol #11-05722, Of.CEP-0295/12 and Of.CEP-1041/12). An informed written  
29 consent and a survey form on gender, age, ancestry, place of birth and residency was completed  
30 and signed by all participants.

31 The skin color of each participant was identified using the Fitzpatrick score (Types 1 to  
32 6) and combined into three skin categories: LIGHT SKIN (Types 1 and 2); MEDIUM SKIN (Types 3  
33 and 4); DARK SKIN (Types 5 and 6). Also, the amounts of red (R), green (G), and blue (B) pigments  
34 were measured in an inner and hairless portion of the right arm (below elbow) using a color  
35 analyzer equipment ACR-1023 (Instrutherm, Brazil). Photographs of participants' eyes were  
36 taken and the eye colors were identified as: light blue, dark blue, green, hazel, light brown, dark



1 brown, or black. Then they were combined and classified in three eye categories: LIGHT EYES  
2 (light blue and dark blue); INTERMEDIATE EYES (green, hazel and light brown); DARK EYES (dark  
3 brown and black). In addition, the amounts of RGB pigments were measured by analyzing each  
4 photo with the software COLORS [8] (phenotyping details are available on [9]).

5 Genomic DNA from buccal swabs was extracted using the NucleoSpin® Tissue kit  
6 (Macherey-Nagel Inc.). Amplification of SNPs was done by multiplex PCR in a final volume of 25  
7 µL containing 2-10 ng gDNA, 0.2 µM of each primer, and 1X Multiplex PCR Master Mix (Qiagen  
8 Inc, CA). Thermal cycler conditions consisted of 95°C/5min, 30 cycles of 94°C/30s, 57°C/90s and  
9 72°C/90s, and a final extension of 72°C/10min. PCR products were purified using USB® ExoSAP-  
10 IT® PCR Product Cleanup (Affymetrix®) according to the manufacturer's protocol. For SNP  
11 analysis, SNaPshot® Multiplex System ABI Prism (Applied Biosystems®) was used with eight  
12 SNaPshot primers. Reactions were performed in a final volume of 10 µL following the  
13 manufacturer's protocol. Multiplex single base extensions were carried out for 28 cycles of  
14 96°C/10s, 50°C/5s and 60°C/30s. SNaPshot products were treated with FastAP (Thermo  
15 Scientific™) according to the manufacturer's recommendation. 1µL of the labeled products  
16 were mixed with 8.5 µL of HiDi formamide and 0.5 µL of Genescan-120 LIZ size standard. They  
17 were then separated using an ABIPRISM® 3130xl Genetic Analyzer (Applied Biosystems) with  
18 POP-4 matrix in a capillary of 36cm. The final analysis was performed using GeneMapper ID  
19 software version 4.0 (Applied Biosystems®) (genotyping details are available on [9]).

20 Following Liu *et al.* [2], we performed multinomial logistic regression analysis and other  
21 additional statistic testes using IBM SPSS Statistics 17 [10] and the R 3.3.2 software [11].

## 22 23 24 **RESULTS AND DISCUSSION**

25  
26 In a first analysis, a comparison was done between LIGHT SKIN / LIGHT EYES (N= 148)  
27 *versus* DARK SKIN / DARK EYES (N= 54) subjects and highly significant differences were found  
28 between allele frequencies of rs916977 (*HERC2*), rs12913832 (*HERC2*), rs4778138 (*OCA2*),  
29 rs1426654 (*SLC24A5*), rs16891982 (*SLC45A2*), rs2733832 (*TYRP1*), rs8045560 (*MC1R*) SNPs  
30 (Pearson's Test;  $p < 0.0005$  for all SNPs). It was noticed that SNP rs12913832 (*HERC2*) was in  
31 linkage disequilibrium with SNP rs916977 (*HERC2*). After further examination, SNPs that  
32 presented no relevant association were dropped from the SNP set and the following analyses  
33 were performed using the six most significant SNPs: rs12913832, rs4778138, rs1426654,  
34 rs16891982, rs2733832, and rs8045560.

35 Although all selected genes are involved in the melanogenesis process, it is still unclear  
36 whether their gene products present a tissue-specific expression, and if they have an influence

1 on skin or iris color or both. There was no significant differences between allele frequencies of  
 2 SNP rs12913832 ( $p=0.098$ ) among ( $N= 154$ ) DARK EYES subjects, independently of their skin  
 3 color. Based on this data, this SNP was considered a good eye color marker since all subjects  
 4 tested have DARK EYES but different skin tones. A comparison was also done between LIGHT  
 5 SKIN subjects ( $N= 248$ ) and there were no significant differences between allele frequencies of  
 6 SNPs rs16891982 ( $p=0.999$ ), rs8045560 ( $p=0.095$ ), and rs1426654 ( $p=0.999$ ). This last result  
 7 suggest that these three SNPs are good skin color markers, given that the subjects tested have  
 8 all LIGHT SKIN tone but different eye colors.

9 Next, a multinomial logistic regression was conducted and initially two models were  
 10 tested: 1- 'SNP ONLY', and 2- 'FULL', inclusive of 'SNP ONLY + GENDER' (TABLE 1). The test  
 11 accuracy (based on a 10-fold cross validation test) was higher for the 'FULL' model, which  
 12 included the six SNPs and the variable 'GENDER'.

13  
 14

15 **TABLE 1-** Multinomial logistic regression for skin and eye prediction; 'FULL' and 'SNP ONLY'  
 16 models, SNPs and GENDER variables are sorted by AIC values.

17

SKIN				EYE			
Model	Residual Deviance	AIC	InL	Model	Residual Deviance	AIC	InL
FULL	234.1125000	298.1125	117.0563	Full model	281.6249	345.6249	140.8124
SNPS ONLY	373.2593000	433.2593	186.6296	SNPs only	354.4310	414.4310	177.2155
rs12913832	467.4960644	479.4961	233.7480	rs1426654	428.4093	440.4093	214.2046
rs4778138	563.2855229	571.2855	281.6428	GENDER	751.1544	763.1544	375.5772
rs1426654	610.2770366	622.2770	305.1385	rs12913832	782.9468	794.9468	391.4734
GENDER	628.3360096	640.3360	314.1680	rs16891982	810.1636	818.1636	405.0818
rs2733832	669.4358247	681.4358	334.7179	rs4778138	858.3558	870.3558	429.1779
rs16891982	671.246956	683.2470	335.6235	rs8045560	858.9804	870.9804	429.4902
rs8045560	689.3647308	701.3647	344.6824	rs2733832	878.9304	890.9304	439.4652

18

19 AIC- Akaike information criterion; InL - log likelihood

20

21

22 Next, a 'STEPWISE' model, which included a selection of variables with the best AIC  
 23 (Akaike information criterion), was tested. This third model presented the best results for  
 24 phenotype DNA prediction when compared to 'FULL' and 'SNP ONLY' models. SNPs rs12913832,  
 25 rs4778138, rs1426654, rs2733832, and gender were selected as informative variables to predict  
 26 skin tone, and SNPs rs1426654, rs12913832, rs16891982, rs4778138, rs8045560, rs2733832,

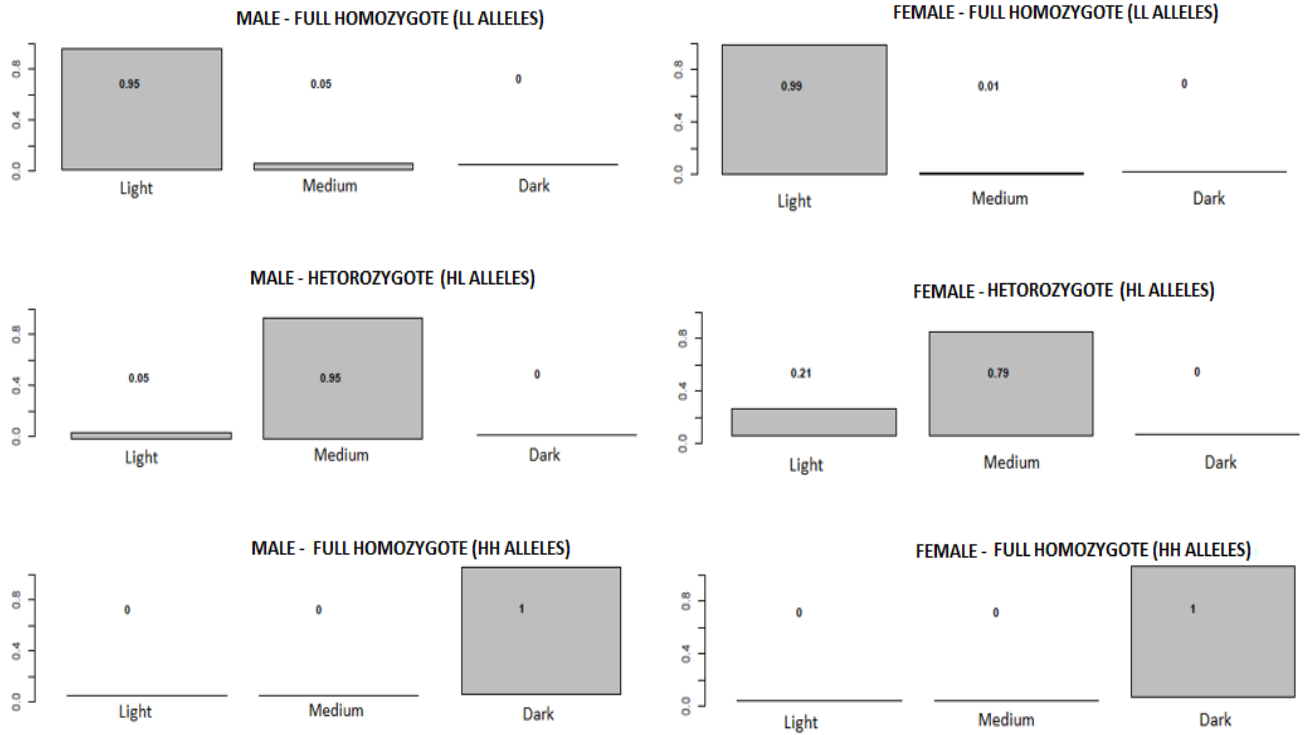
1 and gender were selected as informative variables to predict eye color. The selection of these  
2 variables by the 'STEPWISE' model was interesting given that the genotype-phenotype  
3 associations found here were also seen in previous studies conducted in ethnically different  
4 populations. Liu *et al.* [4] suggested that SNPs rs1426654 (*SLC24A5*) and rs16891982 (*SLC45A2*)  
5 are likely major predictors for skin color differences observed between continents. Our findings  
6 in South Brazilian subjects were consistent with Liu *et al.* since rs1426654 was the best ranked  
7 SNP for skin color prediction and the third best predictor for eye color. Similar findings were also  
8 described in a study by Edwards *et al.* [13] in a South Asian population. Also, in our findings,  
9 rs12913832 (*HERC2*) was the second best predictive variable for skin color and the best ranked  
10 SNP for predicting eye phenotype, reinforcing data presented by Walsh *et al.* [3] in individuals  
11 from Europe, Middle East and West Asia, and by other studies [4, 12, 13, 14].

12 When information on gender was included in the 'STEPWISE' model, it reduced AIC  
13 values, and so gender was selected as a good predictor of phenotype. Although it is still unclear  
14 how gender could have any effect on human eye color pigmentation, our findings are in  
15 accordance to those published by Pietroni *et al.* [13] and Póspiech *et al.* [14]. However, the  
16 inclusion of gender did not improve the accuracy in eye color prediction with the IrisPlex model  
17 proposed by Walsh *et al.* [3]. These data show how iris color variation seems to be much more  
18 complex than it appears.

19 Based on the genotype of six SNPs and gender (assigned by the 'STEPWISE' model), we  
20 calculated the joint probability of hypothetical subjects belonging to one of the skin (LIGHT,  
21 MEDIUM or DARK) and eye color categories (LIGHT, INTERMEDIATE or DARK). Figures 1 and 2  
22 illustrate the probabilities of a hypothetical set of genotypes and gender data belonging to  
23 subjects with different skin (FIGURE 1) and eye colors (FIGURE 2). Allele 'H' was used to call the  
24 alleles strongly associated with DARK SKIN / DARK EYES subjects (rs12913832=A; rs4778138=G;  
25 rs1426654=G; rs16891982=C; rs2733832=C; rs8045560=C), and allele 'L' was used for the alleles  
26 strongly associated with LIGHT SKIN / LIGHT EYES subjects (rs12913832=G; rs4778138=A;  
27 rs1426654=A; rs16891982=G; rs2733832=T; rs8045560=T).

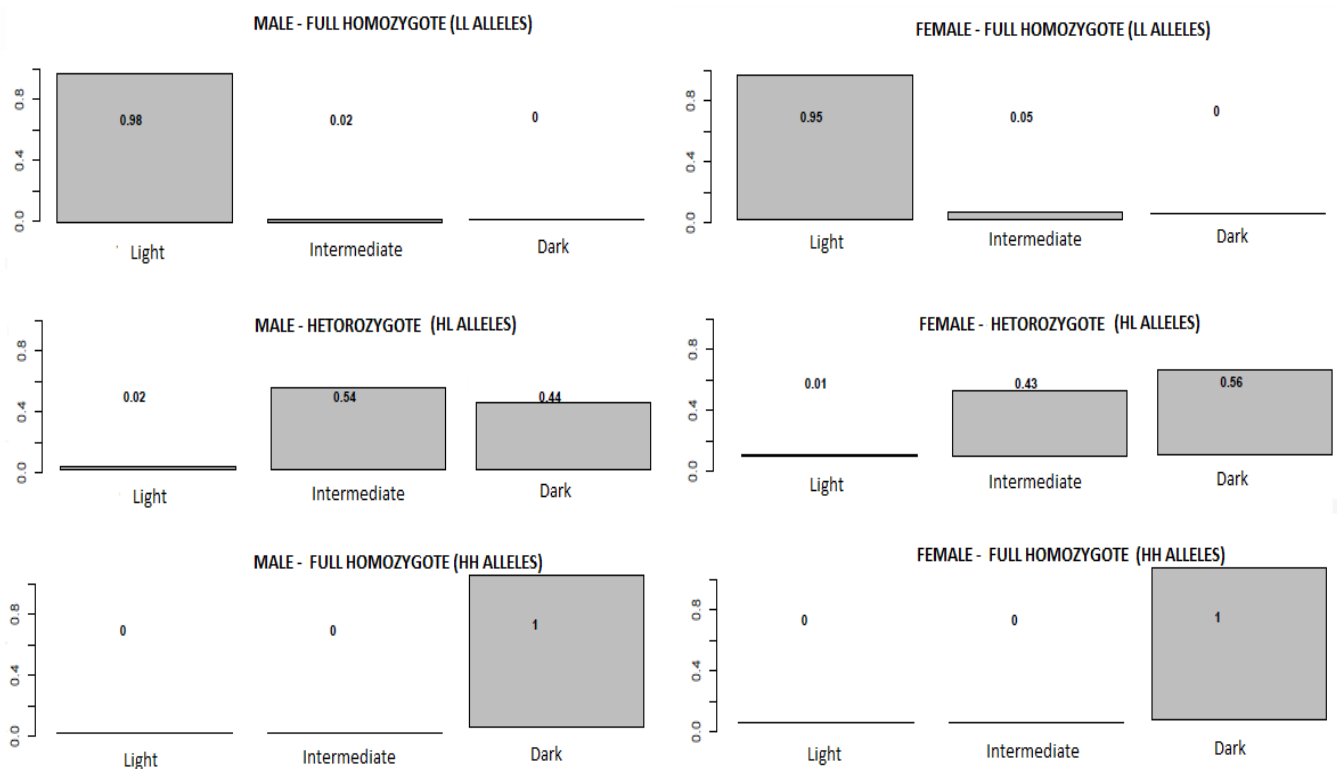
28

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25



**FIGURE 1**– Probability of a hypothetical set of genotypes and gender data to belong to subjects with LIGHT, MEDIUM or DARK SKIN tone according to the 'STEPWISE' multinomial logistic regression model (informative variables: rs12913832, rs4778138, rs1426654, rs2733832, and gender).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21



**FIGURE 2**– Probability of a hypothetical set of genotypes and gender data to belong to subjects with LIGHT, INTERMEDIATE or DARK EYE color according to the 'STEPWISE' multinomial logistic regression model (informative variables: rs1426654, rs12913832, rs16891982, rs4778138, rs8045560, rs2733832, and gender).

As shown in Figures 1 and 2, subjects that are homozygote for allele ‘H’ will always have DARK SKIN and DARK EYES, while subjects that are homozygote for allele ‘L’ will never have DARK SKIN or DARK EYES. In addition, heterozygote subjects will never have DARK SKIN and unlikely will have LIGHT EYES eyes.

Lastly, we performed a blind test with 40 random South Brazilian subjects to test the accuracy of the eye and skin colors prediction system based on genotype and gender data. Results showed a 93% (37/40) concordance between the predicted and observed phenotypes (TABLE 2). Three mismatches were found: one for the skin tone of one subject (PUC345: predicted= LIGHT SKIN; observed= MEDIUM SKIN), and two for the eye color of other subjects (PUC217: predicted= DARK EYES; observed= INTERMEDIATE EYES, and PUC224: predicted= LIGHT EYES; observed= INTERMEDIATE EYES). This model presented an acceptable error rate,

1 and it was able to predict skin and eye pigmentation, with high assertiveness, in subjects with  
 2 extreme phenotypes.

3

4 **TABLE 2-** Blind test: Data of predicted *versus* observed skin and eye pigmentation categories.

5

ID	GENDER	PROBABILITY OF SKIN TO BE			ACTUAL	PROBABILITY OF EYES TO BE			ACTUAL
		LIGHT	MEDIUM	DARK	SKIN TONE	LIGHT	INTERMEDIATE	DARK	EYE COLOR
PUC216	M	0.27	0.73	0	Medium	0.06	0.43	0.51	Dark Brown
PUC217	F	0.84	0.16	0	Light	0.02	0.11	<b>0.87*</b>	<b>Light Brown</b>
PUC219	F	0.50	0.50	0	Medium	0.05	0.77	0.18	Hazel
PUC222	M	0	0	1.00	Dark	0	0	1.00	Dark Brown
PUC223	M	0.98	0.02	0	Light	0.98	0.02	0	Blue
PUC224	F	0.99	0.01	0	Light	<b>0.89*</b>	0.11	0	<b>Green</b>
PUC225	F	0.99	0.01	0	Light	0.95	0.05	0	Blue
PUC227	M	0.93	0.07	0	Light	0.94	0.06	0	Blue
PUC252	M	0.98	0.02	0	Light	0.97	0.03	0	Blue
PUC253	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC262	M	0.93	0.07	0	Light	0.98	0.02	0	Blue
PUC280	F	0	1.00	0	Medium	0	1	0	Light Brown
PUC293	F	0.99	0.01	0	Light	0.96	0.04	0	Blue
PUC300	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC312	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC314	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC315	F	0.99	0.01	0	Light	0.95	0.05	0	Blue
PUC316	F	0.99	0.01	0	Light	0.95	0.05	0	Blue
PUC317	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC320	F	0	0.01	0.99	Dark	0	0	1.00	Dark Brown
PUC321	M	0.96	0.04	0	Light	0.94	0.06	0	Blue
PUC321	M	0.66	0.34	0	Light	0.03	0.61	0.36	Green
PUC322	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC326	M	0.95	0.05	0	Light	0.98	0.02	0	Blue
PUC327	M	0.93	0.07	0	Light	0.98	0.02	0	Blue
PUC330	M	0.68	0.32	0	Light	0.94	0.06	0	Blue
PUC331	M	0.94	0.06	0	Light	0.01	0.99	0	Green
PUC333	M	0.97	0.03	0	Light	0.98	0.02	0	Blue
PUC334	M	0	0	1.00	Dark	0	0	1.00	Dark Brown
PUC335	M	0	0	1.00	Dark	0	0	1.00	Dark Brown
PUC336	F	0.96	0.04	0	Light	0	0.06	0.94	Dark Brown
PUC337	F	0.99	0.01	0	Light	0.89	0.11	0	Blue
PUC338	M	0.93	0.07	0	Light	0.98	0.02	0	Blue
PUC339	F	0.95	0.05	0	Light	0.02	0.2	0.79	Dark Brown
PUC340	M	0.60	0.36	0.04	Light	0	0.04	0.96	Dark Brown
PUC341	M	0.93	0.07	0	Light	0.98	0.02	0	Blue
PUC342	F	0.97	0.02	0	Light	0	0.02	0.98	Dark Brown
PUC343	M	0.99	0.01	0	Light	0.95	0.05	0	Blue
PUC345	M	<b>0.65*</b>	0.14	0.21	<b>Medium</b>	0	0.09	0.91	Dark Brown
PUC346	F	0.88	0.12	0	Light	0.01	0.51	0.48	Green

6

7

\*Mismatches between predicted and observed phenotypes.

8

1           **CONCLUSION**

2  
3           Not rarely, global or foreign phenotype predictor systems fail in Brazilian samples since  
4 Brazil has an ethnically mixed population in which the genomic structure may favor epistatic and  
5 pleiotropic effects not observed in populations that are more homogeneous. However, the  
6 model presented in this study seems to be able to predict eye and skin colors in subjects from  
7 South Brazil. We believe this FDP system can be particularly useful in instances where classic  
8 DNA profiling fails to generate a database match, or to help classifying degraded corpses,  
9 skeletal, or other biological evidences for a missing person.

10  
11  
12           **ACKNOWLEDGMENTS**

13  
14           This work was supported by CAPES Brazil | Edital 25/2014 | Pró-Forenses, and CNPq  
15 Brasil | Edital INCT. Fernanda Rosa Sawitzki, Eduardo Filipe Avila Silva, Gustavo Adolfo Silva-  
16 Arias, Deborah Soares Bispo Santos Silva, and Carlos Eduardo Ibaldo Gonçalves were supported  
17 by CAPES. The authors would like to thank to Rodrigo Rodenbusch, Diego Wordell Gubert  
18 Mayara Jorgens Prado, and Pietra Graebin for collection and organization of phenotype data.

19  
20  
21           **CONFLICT OF INTEREST**

22  
23           The authors declare no conflict of interest.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

1                   **SUPPLEMENTAL FILE**

2

3

4                   The following tables contain all the models tested to choose the best

5 one to predict eye and skin phenotype. First we present data from skin analysis and then from

6 eye analysis. The “full model” considered an additive model of all SNPs and gender while the

7 “SNP only “excluded gender. Finally, the stepwise model was constructed with variable

8 selection based on AIC, which compares full, reduced and single variables model fit by using

9 forwards, backwards, and both-way variable selection schemes. In addition, we performed an

10 eye analysis with the stepwise model including the rs4778138. When cross-analyzing HMC x

11 SW and LMC x SW it was possible to notice this SNP could also predict the eye color. This

12 analysis led to a 0.88 accuracy result, higher than the standard step-wise accuracy index of

13 0.86, and higher than “full model”. The remaining SNPs had no additive value to the predictive

14 accuracy.

15

16

17



1  
2

**SUPPLEMENTAL FILE - SKIN STATISCS**

<b>FULL MODEL</b>																
<b>Call:</b>																
multinom(formula = pele_ref ~ GEN + RS4778138 + RS12913832 + RS16891982 + RS8045560 + RS1426654 + RS2733832 + RS1042602, data = data_matrix)																
Reference category: "3-Dark"																
<b>Coefficients:</b>																
	(Intercept)	GENDER-M	RS4778138C	RS4778138C	RS12913832GA	RS12913832	RS16891982	RS16891982	RS8045560C	RS8045560T	RS1426654C	RS1426654C	RS2733832C	RS2733832T	RS1042602C	RS1042602CC
1-Light	28,46688	-2,0823793	-16,47403	-17,53528	1,956854	29,22061	3,986956	4,093621	2,81969	2,481316	-2,5369083	-54,0451	0,9261111	28,05842	-13,25351	-14,19725
2-Medium	27,73051	-0,3421549	-17,17381	-17,6529	2,12525	26,95703	3,011503	1,19067	3,266257	2,520276	-0,4761373	-17,73518	1,8364191	29,05716	-13,08093	-14,05777
<b>Std. Errors:</b>																
	(Intercept)	GENDER-M	RS4778138C	RS4778138C	RS12913832GA	RS12913832	RS16891982	RS16891982	RS8045560C	RS8045560T	RS1426654C	RS1426654C	RS2733832C	RS2733832T	RS1042602C	RS1042602CC
1-Light	1,505262	1,105034	1,965047	2,15E+00	1,402274	0,3267151	2,155774	2,135984	1,596742	1,770191	1,44E+00	1,26E-13	1,155228	0,2830033	1,835806	1,435158
2-Medium	1,337999	1,067128	1,940133	2,17E+00	1,35243	0,3267152	1,84309	1,84068	1,590167	1,765392	1,47E+00	1,55E+00	1,146389	0,2830033	1,801784	1,374211
<b>p-value:</b>																
	(Intercept)	GENDER-M	RS4778138C	RS4778138C	RS12913832GA	RS12913832	RS16891982	RS16891982	RS8045560C	RS8045560T	RS1426654C	RS1426654C	RS2733832C	RS2733832T	RS1042602C	RS1042602CC
1-Light	0	0,0595045	0	2,22E-16	1,63E-01	0	0,0643955	0,055301	0,0774124	0,1609982	0,0781523	0	4,23E-01	0	5,22E-13	0
2-Medium	0	0,7484895	0	4,44E-16	1,16E-01	0	0,1022709	0,5177201	0,0399724	0,1534066	0,7451977	0	1,09E-01	0	3,87E-13	0
<b>Residual Deviance</b>	234,1125															
<b>AIC</b>	298,1125															
<b>InL</b>	117,05626															
		<b>Predictive statistics</b>														
		<b>Accuracy</b> 0,86 (0.73 - 0.98)														
			<b>Sensitivity</b>	<b>sd</b>	<b>Specificity</b>	<b>sd</b>	<b>Precision</b>	<b>sd</b>	<b>Prevalence</b>	<b>sd</b>						
		3-Dark	0,63	0,43	0,99	0,01	0,50	0,55	0,06	0,10						
		1-Light	0,95	0,07	0,47	0,42	0,89	0,11	0,80	0,13						
		2-Medium	0,50	0,41	0,95	0,06	0,60	0,37	0,14	0,12						

3  
4

SNPs ONLY MODEL (excluding gender)															
<b>Call:</b>															
multinom(formula = pele_ref ~ RS4778138 + RS12913832 + RS16891982 + RS8045560 + RS1426654 + RS2733832 + RS1042602, data = data_matrix)															
Reference category: "3-Dark"															
<b>Coefficients:</b>															
	(Intercept	RS4778138	RS4778138	RS1291383	RS1291383	RS1689198	RS1689198	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832	RS1042602	RS1042602CC
1-Light	3,464737	-0,25494	-1,59659	0,743814	0,670584	1,097925	2,283772	1,078601	3,260428	-4,04358	-21,0026	-0,04629	18,7959	-1,61546	-2,39991
2-Medium	4,299191	-0,68662	-1,79171	0,990263	-0,94617	0,517393	-0,11993	0,51822	2,668322	-2,17946	-4,41787	0,167128	19,46307	-1,47206	-2,17943
<b>Std. Errors:</b>															
	(Intercept	RS4778138	RS4778138	RS1291383	RS1291383	RS1689198	RS1689198	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832	RS1042602	RS1042602CC
1-Light	2,323405	0,784456	9,58E-01	0,743499	1,214819	1,461596	1,428375	0,64258	1,230853	0,863466	9,77E-06	0,639075	0,223519	1,538363	1,457653
2-Medium	2,043452	0,748193	8,73E-01	0,698085	1,209563	0,911906	0,907624	0,607134	1,177534	0,860712	1,05E+00	0,593936	0,223519	1,494698	1,402516
<b>p-value:</b>															
	(Intercept	RS4778138	RS4778138	RS1291383	RS1291383	RS1689198	RS1689198	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832	RS1042602	RS1042602CC
1-Light	0,1359	0,74519	0,095739	3,17E-01	5,81E-01	0,452543	0,109852	0,09324	0,008075	2,83E-06	0,00E+00	9,42E-01	0	0,293665	0,099677
2-Medium	0,035388	0,35877	0,040184	1,56E-01	4,34E-01	0,570459	0,894875	0,393354	0,02345	1,13E-02	2,40E-05	7,78E-01	0	0,324694	0,120198
<b>Residual D</b>	373,2593														
<b>AIC:</b>	433,2593														
<b>InL</b>	186,6296														
		<b>Predictive statistics</b> (based on a 10-fold cross validation test)													
		<b>Accuracy</b> 0,81 (0.72 - 0.90)													
		<b>Sensitivity</b>		<b>sd</b>	<b>Specificity</b>		<b>sd</b>	<b>Precision</b>		<b>sd</b>	<b>Prevalence</b>		<b>sd</b>		
		1-Light	0,97	0,04	0,94	0,05	0,91	0,09	0,355705	0,230016					
		2-Medium	0,29	0,25	0,92	0,06	0,45	0,31	0,185833	0,119674					
		3-Dark	0,86	0,11	0,75	0,25	0,79	0,17	0,458462	0,228779					

VARIABLE IMPORTANCE EXPLORATORY ANALYSES							
Variable importance (explanatory contribution to full model)				Comparison of full, reduced and single variables model fit			
				Model	Residual Deviance	AIC	InL
RS1426654GG	71,78			Full model	234,1125	<b>298,1125</b>	117,0563
RS2733832TT	57,12			SNPs only	373,2593	433,2593	186,6296
RS12913832GG	56,18			RS1426654	467,4960644	479,4961	233,748
RS4778138GG	35,19			GENERO	563,2855229	571,2855	281,6428
RS4778138GA	33,65			RS12913832	610,2770366	622,277	305,1385
RS1042602CC	28,25502			RS16891982	628,3360096	640,336	314,168
RS1042602CA	26,33			RS4778138	669,4358247	681,4358	334,7179
RS16891982GC	7,00			RS8045560	671,246956	683,247	335,6235
RS8045560CT	6,09			RS2733832	689,3647308	701,3647	344,6824
RS16891982GG	5,28			RS1042602	700,8324539	712,8325	350,4162
RS8045560TT	5,00						
RS12913832GA	4,08						
RS1426654GA	3,01						
RS2733832CT	2,76						
GENDER-M	2,42						
<b>Variable selection for reduced models based on AIC</b>							
Backwards selection							
pele_ref ~ GENERO + RS4778138 + RS12913832 + RS16891982 + RS8045560 + RS1426654 + RS2733832							
Forwards selection							
pele_ref ~ RS1426654 + RS12913832 + RS16891982 + GENERO + RS4778138 + RS2733832 + RS8045560							
Bothway selection							
pele_ref ~ RS1426654 + RS12913832 + RS16891982 + GENERO + RS4778138 + RS2733832 + RS8045560							

stepwise model																			
<b>Call:</b>																			
multinom(formula = pele_ref ~ GEN + RS4778138 + RS12913832 + RS16891982 + RS8045560 + RS1426654 + RS2733832, data = data_matrix)																			
Reference category: "3-Dark"																			
<b>Coefficients:</b>																			
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832G	RS16891982GC	RS16891982G	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832TT					
1-Light	3,006502	-2,2114228	-4,296571	-5,693507	2,262895	25,35494	3,949701	3,844804	2,979065	2,998314	-2,66203	-36,9364	0,764586	14,96028					
2-Medium	2,361023	-0,4971882	-4,971532	-5,801038	2,430956	23,11041	3,011271	0,99828	3,426377	3,029403	-0,61026	-6,29636	1,681179	15,93536					
<b>exp(B)</b>																			
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832G	RS16891982GC	RS16891982G	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832TT					
1-Light	2,02E+01	0,1095447	0,013615161	0,003367763	9,61E+00	1,03E+11	51,91982	4,67E+01	19,66941	2,01E+01	0,069807	9,09E-17	2,148105	3141725					
2-Medium	1,06E+01	0,6082385	0,006932516	0,003024413	1,14E+01	1,09E+10	20,3132	2,71E+00	30,76496	2,07E+01	0,543207	1,84E-03	5,371888	8329918					
<b>Std.Errors:</b>																			
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832G	RS16891982GC	RS16891982G	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832TT					
1-Light	2,346838	1,098369	2,45E+00	2,44E+00	1,348422	0,3231973	2,080278	1,944611	1,535254	1,740834	1,449795	7,16E-09	1,068485	402,4728					
2-Medium	2,111235	1,061722	2,43E+00	2,42E+00	1,305899	0,3231973	1,765367	1,645623	1,512606	1,722914	1,464409	2,04E+00	1,043111	402,4726					
<b>p-value:</b>																			
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832G	RS16891982GC	RS16891982G	RS8045560	RS8045560	RS1426654	RS1426654	RS2733832	RS2733832TT					
1-Light	2,00E-01	0,04407572	0,07988726	1,99E-02	9,33E-02	0	5,76E-02	4,80E-02	0,052327	0,085007	0,066337	0	0,474251	9,70E-01					
2-Medium	2,63E-01	0,63958081	0,04071972	0,01649223	6,27E-02	0	0,08805507	5,44E-01	0,0235	0,078696	0,676875	0,002024	0,107028	9,68E-01					
<b>Residual Deviance:</b>	237,8685																		
<b>AIC:</b>	293,8685																		
<b>InL</b>	118,93425																		
		<b>Predictive statistics</b> (based on a 10-fold cross validation test)																	
		<b>Accuracy</b> 0,89 (0.77 - 1.00)																	
				<b>Sensitivity</b>		<b>sd</b>		<b>Specificity</b>		<b>sd</b>		<b>Precision</b>		<b>sd</b>		<b>Prevalence</b>		<b>sd</b>	
		3-Dark		0,88		0,25		1,00		0,01		0,80		0,45		0,07		0,11	
		1-Light		0,98		0,04		0,54		0,36		0,89		0,12		0,77		0,15	
		2-Medium		0,56		0,38		0,98		0,04		0,84		0,27		0,16		0,14	

1  
2  
3  
4  
5

<b>Comparison of full, reduced and single variable models</b>					
<b>Model</b>	<b>Residual Deviance</b>	<b>AIC</b>	<b>Δ AIC</b>	<b>InL</b>	<b>LRT</b>
stepwise models	237,8685	293,8685		118,93425	0,4400374
Full model	234,1125	298,1125	4,244	117,05626	
Dominant + other SNPs + gender	259,1147	307,1147	13,2462	129,557343	0,00155325
stepwise models (without Gender)	259,6213	311,6213	17,7528	129,810672	0,00027477
SNPs only	255,886	315,886	22,0175	127,943011	1,8704E-05
RS1426654	351,0581572	363,0582	69,18966	175,529079	1,6833E-13
RS16891982	458,3850486	470,385	176,5165	229,192524	1,8412E-33
Only loci with dominant alleles	452,1565	472,1565	178,288	226,07825	3,2291E-34
RS12913832	464,70099	476,701	182,8325	232,350495	1,0888E-34
RS4778138	521,0075765	533,0076	239,1391	260,503788	8,6927E-46
RS8045560	522,2678404	534,2678	240,3993	261,13392	4,8771E-46
RS2733832	534,9485091	546,9485	253,08	267,474255	1,4368E-48
RS1042602	550,6062961	562,6063	268,7378	275,303148	1,047E-51
GENERO	562,8665205	570,8665	276,998	281,43326	4,5621E-53
NULL	575,3048	579,3048	285,4363	287,652386	1,80E-54

6  
7  
8  
9  
10  
11  
12  
13

1  
2

**SUPPLEMENTAL FILE - EYE STATISTICS**

FULL MODEL																					
Call:																					
multinom(formula = olho_ref ~ GEN + RS4778138 + RS12913832 + RS16891982 + RS8045560 + RS1426654 + RS2733832 + RS1042602, data = data_matrix)																					
Reference category: "3-Dark"																					
<b>Coefficients:</b>																					
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS16891982GC	RS16891982GG	RS8045560CT	RS8045560TT	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT	RS1042602CA	RS1042602CC					
1- Light	-44,050542	1,2030344	0,1329261	-15,86803	19,660247	66,55267	21,4129233	20,5179407	0,2387867	0,5024152	-0,9599069	-75,90235	0,2311341	2,716526	-0,34303848	-0,6140441					
2- intermed	-3,659494	0,4667996	-0,6576409	-1,36924	2,150376	43,28438	0,7317219	0,4925415	-0,5120872	-0,3286113	0,3646818	-15,21951	1,4293365	2,929863	0,05620956	-0,2649666					
<b>Std.Errors:</b>																					
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS16891982GC	RS16891982GG	RS8045560CT	RS8045560TT	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT	RS1042602CA	RS1042602CC					
1- Light	0,6869263	0,6646491	0,905561	2,81E-08	0,5809099	0,3893675	0,6969682	0,5323684	0,888059	0,9857193	1,4975037	3,81E-16	0,9882309	1,0383201	0,9681855	0,9922432					
2- intermed	1,3560049	0,3864455	0,4003014	8,86E-01	0,4675935	0,3893675	1,1217676	1,0768523	0,4668883	0,5409442	0,4796347	1,36E-06	0,5211123	0,6081605	0,5387764	0,5650525					
<b>p-value:</b>																					
	(Intercept)	GENDER-M	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS16891982GC	RS16891982GG	RS8045560CT	RS8045560TT	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT	RS1042602CA	RS1042602CC					
1- Light	0	0,07029122	0,8832988	0	0,00E+00	0	0	0	0,7880174	0,6102659	0,5215196	0	0,815072901	8,89E-03	0,723106	0,5360189					
2- intermed	0,00696041	0,22707371	0,100411	0,1224264	4,25E-06	0	0,5142118	0,6473907	0,272725	0,5435342	0,4470559	0	0,006090719	1,45E-06	0,9169089	0,639124					
	<b>Residual De</b>	281,6249																			
	<b>AIC</b>	345,6249																			
	<b>lnL</b>	140,812436																			
				<b>Predictive statistics</b>																	
				<b>Accuracy</b> 0,85 (0,77 - 0,92)																	
				<b>Sensitivity</b>		<b>sd</b>		<b>Specificity</b>		<b>sd</b>		<b>Precision</b>		<b>sd</b>		<b>Prevalence</b>		<b>sd</b>			
				1- Light		0,97		0,04		0,89		0,16		0,92		0,06		0,45		0,33	
				2- intermediate		0,39		0,30		0,93		0,08		0,48		0,36		0,16		0,12	
				3- Dark		0,86		0,15		0,84		0,18		0,79		0,17		0,38		0,29	

3



<b>VARIABLE IMPORTANCE EXPLORATORY ANALYSES</b>							
<b>Variable importance (explanatory contribution to full model)</b>				<b>Comparison of full, reduced and single variables model fit</b>			
				<b>Model</b>	<b>Residual Deviance</b>	<b>AIC</b>	<b>InL</b>
RS12913832GG	109,84			Full model	281,6249	<b>345,6249</b>	140,812436
RS1426654GG	91,12			SNPs only	354,431	414,431	177,215482
RS16891982GC	22,14			RS12913832	428,4093	440,4093	214,204639
RS12913832GA	21,81			RS4778138	751,1544	763,1544	375,577224
RS16891982GG	21,01			RS1426654	782,9468	794,9468	391,473385
RS4778138GG	17,24			GENDER	810,1636	818,1636	405,081778
RS2733832TT	5,65			RS2733832	858,3558	870,3558	429,177902
GENDER-M	1,67			RS16891982	858,9804	870,9804	429,490198
RS2733832CT	1,66			RS8045560	878,9304	890,9304	439,465205
RS1426654GA	1,32			RS1042602	894,8409	906,8409	447,420473
RS1042602CC	0,88						
RS8045560TT	0,83						
RS4778138GA	0,79						
RS8045560CT	0,75						
RS1042602CA	0,40						
<b>Variable selection for reduced models based on AIC</b>							
Backwards selection							
olho_ref ~ GEN + RS12913832 + RS1426654 + RS2733832							
Forwards selection							
olho_ref ~ RS12913832 + RS2733832 + RS1426654 + GEN							
Bothway selection							
olho_ref ~ RS12913832 + RS2733832 + RS1426654 + GEN							



STEPWISE+RS47+GEN															
<b>Call:</b>															
multinom(formula = olho_ref ~ GEN + RS4778138 + RS12913832 + RS1426654 + RS2733832, data = data_matrix)															
Reference category: "3-Dark"															
<b>Coefficients:</b>															
	(Intercept)	GENEROM	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT					
1- Light	-21,698317	1,1842942	0,1663211	-9,523992	17,851354	63,52622	-0,8207795	-44,37783	0,3246238	2,649689					
2- intermed	-3,486471	0,4706484	-0,686395	-1,494969	2,097553	42,11093	0,3652895	-22,87992	1,4436892	2,850271					
<b>exp(B)</b>															
	(Intercept)	GENEROM	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT					
1- Light	3,77E-10	3,268379	1,1809523	7,31E-05	5,66E+07	3,88E+27	0,4400885	5,33E-20	1,38351	14,14964					
2- intermed	3,06E-02	1,601032	0,5033875	2,24E-01	8,15E+00	1,94E+18	1,4409311	1,16E-10	4,236296	17,29246					
<b>Std.Errors:</b>															
	(Intercept)	GENEROM	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT					
1- Light	0,6049319	0,6567022	8,99E-01	2,43E+02	0,5658324	0,334737	1,4712845	9,59E-07	0,9802039	1,0256557					
2- intermed	0,6604365	0,3807215	3,95E-01	8,84E-01	0,4586102	0,334737	0,4410811	2,21E-06	0,5127402	0,5937434					
<b>p-value:</b>															
	(Intercept)	GENEROM	RS4778138GA	RS4778138GG	RS12913832GA	RS12913832GG	RS1426654GA	RS1426654GG	RS2733832CT	RS2733832TT					
1- Light	0,00E+00	0,07132604	0,8532803	9,69E-01	0,00E+00	0	5,77E-01	0,00E+00	0,740508595	9,78E-03					
2- intermed	1,30E-07	0,21638378	0,0821586	0,09095968	4,79E-06	0	0,4075751	0,00E+00	0,004868092	1,58E-06					
<b>Residual De</b>	285,8046														
<b>AIC:</b>	325,8046														
<b>InL</b>	142,902322														
							<b>Predictive statistics</b> (based on a 10-fold cross validation test)								
							<b>Accuracy</b> 0,88 (0.81 - 0.94)								
							<b>Sensitivity</b>	<b>sd</b>	<b>Specificity</b>	<b>sd</b>	<b>Precision</b>	<b>sd</b>	<b>Prevalence</b>	<b>sd</b>	
							1- Light	0,90	0,15	0,87	0,16	0,85	0,13	0,42	0,30
							2- intermediate	0,97	0,05	0,95	0,06	0,93	0,06	0,39	0,31
							3- Dark	0,58	0,28	0,94	0,06	0,69	0,25	0,19	0,13

1  
2  
3  
4  
5

<b>Comparison of full, reduced and single variable models</b>					
<b>Model</b>	<b>Residual Deviance</b>	<b>AIC</b>	<b>Δ AIC</b>	<b>InL</b>	<b>LRT</b>
stepwise models	291,6852	323,6852		145,8426	0,8635
Stepwise+ rs47	289,3334	325,3334	1,6	144,666694	0,903952
Stepwise+ rs47+ gen	285,8046	325,8046	2,1	142,902322	0,97997
Dominant + other SNPs + gender	283,149	335,149	11,5	141,574523	9,58E-01
SNPs only	285,1573	345,1573	21,5	142,578664	0,171
Full model	281,6249	345,6249	21,9	140,812436	
RS12913832	343,7162	355,7162	32,0	171,858106	8,72E-05
RS4778138	668,7237	680,7237	357,0	334,361847	< 2.2e-16
RS1426654	721,5364	733,5364	409,9	360,768212	< 2.2e-16
RS2733832	764,616	776,616	452,9	382,308002	< 2.2e-16
RS16891982	766,6773	778,6773	455,0	383,338655	< 2.2e-16
Only loci with dominant alleles	776,4851	792,4851	468,8	388,242558	1,95E-89
RS8045560	793,6942	805,6942	482,0	396,847095	< 2.2e-16
GENERO	806,8082	814,8082	491,1	403,40411	< 2.2e-16
RS1042602	804,0464	816,0464	492,4	402,023178	< 2.2e-16
NULL	815,1954	819,1954	495,5	407,597721	< 2.2e-16

6  
7  
8  
9

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

## REFERENCES

- [1] Kayser, M., *Forensic Sci. Int. Genet.* 2015, 18, 33-48.
- [2] Liu, F., van Duijn, K., Vingerling, J. R., Hofman, A., Uitterlinden, A. G., Janssens, A. C. J. W., Kayser, M., *Curr. Biol.* 2009, 19, R192–R193.
- [3] Walsh, S., Lidenbrgh, A., Zuniga, S. B., Sijen, T., Knijff, P., Kayser, M., Ballantyne, K. N., *Forensic Sci. Int. Genet.* 2011, 5 464-471.
- [4] Liu, F., Wenb, B., Kayser, M. *Semin. Cell. Dev. Biol.* 2013, 24, 562-575.
- [5] Cook, A. L., Chen, W., Thurber, A. E., Smit, D. J., Smith, A. G., Bladen, T. G., Brown, D. L., Duffy, D. L., Pastorino, L., Bianchi-Scarra, G., Leonard, J. H., Stow, J. L., Sturm, R. A., *J. Invest. Dermatol.* 129, 2009, 392-405.
- [6] Parra, F. C., Amado, R. C., Lambertucci, J. R., Rocha, J., Antunes, C. M., Pena, S. D., *Proc. Natl. Acad. Sci. USA.* 2003, 100, 177-182.
- [7] Fitzpatrick, T. B., *Arch. Dermatol.* 1988, 124, 869–871.
- [8] Otaka, I., Kumagai, K., Inagaki, Y., Shimoyama, M., Saegusa, K., Hara, T., *Am. J. Ophthalmol.* 2002, 133, 140-142
- [9] Sawitzki F.R, Rodenbusch R., Gubert D.W, Prado M.J, Silva D.S.B.S and Alho C.S. Analysis of Eight SNPs in Pigment-Related Genes in South Brazilian Subjects with High or Low Skin and Eye Melanin Content. *SM J Forensic Res. Criminol.*, 2017, 1(2), 1008.
- [10] Venables, W. N., Ripley, B. D. (2002) *Modern Applied Statistics with S.* Fourth Edition. Springer, New York. ISBN 0-387-95457-0
- [11] R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [12] Edwards, M., Cha, D., Krithika, S., Johnson, M., Cook, G., Parra, E. J., *Pigment Cell. Melanoma Res.* 2016, 29, 141-162.
- [13] Pietroni, C., Andersen, J. D., Johansen, P., Andersen, M. M., Harder, S., Paulsen, R. R., Børsting, C., Morling, N. *Forensic Science International: Genetics.* 2014, 11, 1-6.
- [14] Pospiech, E., Wojas-Pelc, A., Walsh, S., Liu, F., Maeda, H, Ishikawa, T, Skowron, M, Kayser, M, Branicki, W., *Forensic Sci. Int. Genet.* 2014, 11, 64-72.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18

**CAPÍTULO 3**  
**CONSIDERAÇÕES FINAIS E ANEXOS**

### 3- CONSIDERAÇÕES FINAIS

Recentemente, novas tecnologias começaram a ser aplicadas largamente na presente área em estudo. Podemos citar, por exemplo, a genotipagem de SNPs através de sequenciamento de nova geração, com *chips* condutores/reatores. Estas novas tecnologias de genotipagem de SNPs em larga escala, associadas com a descoberta de novas mutações capazes de prever características externas visíveis, serão capazes de nortear o desenvolvimento de novas pesquisas e conseqüentemente a obtenção de painéis capazes de atender as mais diferentes populações.

Os resultados obtidos nos nossos estudos e apresentados nesta tese, demonstraram que o nosso painel em SNaPShot, para genotipagem de seis SNPs, pode ser uma ferramenta para predição fenotípica de cor de olho e cor de pele, devido ao nível de concordância entre o genótipo e o fenótipo encontrado na população do sul do Brasil.

Fazendo uso do progresso observável neste campo, novos estudos que possam prever características fenotípicas devem ser desenvolvidos, a fim de esclarecer as mais complexas associações e interações entre as dezenas de milhares de variações genéticas conhecidamente envolvidos na pigmentação humana.

Através dos presentes dados e de estudos futuros poderá se estabelecer painéis que possam prever as CEV de forma assertiva e confiável, ajudando, dessa forma, na solução de crimes, desaparecimentos, entre outros casos da rotina forense.

**FORENSIC DNA PHENOTYPING - BRAZIL | POP COLETA****1- Abordar um possível participante da pesquisa, cumprimentá-lo e apresentar o trabalho.**

1- Abordar o participante: -"Bom dia/Boa tarde/Boa noite, somos da PUCRS, estamos realizando uma pesquisa sobre os genes que determinam características do corpo humano, você gostaria de participar?" / "Serão três passos simples e não demora muito tempo, faremos o seguinte: anotaremos seus dados; fotografaremos seu rosto e passaremos um cotonete na sua bochecha para coletar células da boca".

2- Se a pessoa discordar: Agradecer e se despedir. Se a pessoa concordar: Prosseguir.

**2- Passar o TCLE para que o participante da pesquisa conheça e assine.**

1- Instruir o participante: -"Assine, por favor, e fique com uma cópia deste TCLE. Nele há explicações sobre a pesquisa, e constam os telefones de contato caso você queira: saber mais sobre a pesquisa; sair a qualquer momento da pesquisa; receber os resultados da pesquisa". "Permita também, se quiser, que seus dados sejam guardados para serem usados em pesquisas futuras".

**3- Guardar o TCLE assinado e entregar uma cópia do TCLE para o participante.****4- Preencher completa e corretamente o INSTRUMENTO DE COLETA de dados.**

O PESQUISADOR faz as perguntas ao participante. O PESQUISADOR preenche o instrumento.

1- O código de identificação deve ser preenchido usando quatro números e duas letras; as letras se referem às iniciais do primeiro nome e do último sobrenome do participante.

2- Durante o preenchimento, o PESQUISADOR questiona o participante quanto aos seus dados, identifica e anota as categorias de tipo/cor de olho, pele e cabelo. Há um espaço para observações que sejam pertinentes.

**5- Fotografar face e olhos do participante da pesquisa.**

1- Posicionar a pessoa em ambiente iluminado, com fundo liso (p.ex. uma parede) ou afastado de interferentes de fundo.

2- Ajustar a câmara para as seguintes posições: SEM flash; resolução máxima; foco na face da pessoa a ser fotografada.

3- Pedir que a pessoa se mantenha serena (sem expressões faciais) e NÃO se mova enquanto as fotos são tomadas.

4- Disparar 10-12 fotos de uma distância de  $\pm 50$ cm, circundando em torno da cabeça do participante, de orelha a orelha.

5- Aproximar a câmara fotográfica  $\pm 15$ cm de distância entre o olho do participante e a objetiva, fotografar cada um dos olhos.

6- Por fim, fotografar o participante de frente, segurando o INSTRUMENTO DE COLETA de forma que apareça o seu número.

**6- Coletar células da mucosa oral do participante.**

1- Perguntar ao participante: -"Você comeu algo há pouco tempo?" Anotar a resposta no INSTRUMENTO DE COLETA.

2- Separar DOIS envelopes com swabs, e anotar o número de identificação do participante em cada um deles.

3- Mostrar ao participante que o swab é estéril e que está fechado. Insistir que não há risco de contaminação para a pessoa.

4- Ensinar ao participante como ele deverá proceder com o swab na bochecha, demonstrando:

"Você deve fazer um esfregão do lado interno da boca por dez segundos, lentamente, sempre girando o swab. Há dois swabs, um será usado no lado direito da bochecha e o outro do lado esquerdo. Um de cada vez".

5- Abrir o primeiro swab rasgando um lado do envelope; não retirar o swab totalmente do envelope, apenas liberar a ponta.

6- Entregar ao participante para que ele proceda a coleta de células de um dos lados da boca.

7- Recolher e voltar a fechar o envelope protegendo muito bem o material coletado.

8- Repetir o procedimento do outro lado interno da boca.

**7- Encerrar o procedimento.**

1- Agradecer ao participante: -"Muito obrigado(a) pela sua valiosa participação. Esteja à vontade para fazer qualquer pergunta agora ou quando quiser".

**No caso de medir a coloração da pele do participante.****INSTRUÇÕES:**

1- Posicionar o participante em ambiente iluminado.

2- Realizar a calibração do equipamento com o auxílio do cartão de calibração presente no estojo do instrumento: encostar o sensor firmemente no cartão; pressionar os botões FUNCTION e CAL ao mesmo tempo; aguardar aproximadamente 50 segundos até o aparelho gerar o som de três BIPS; confirmar que R, G e B estão com o valor 1023.

3- Encostar o sensor do instrumento na parte interna do braço a  $5 \pm 0,5$ cm acima da dobra do cotovelo.

4- Apertar o gatilho e aguardar os valores aparecerem no visor do aparelho, realizando a medição.

5- Anotar a numeração dos valores RGB que aparecerem no visor; pressionar o botão RGB/HSL, anotar os valores de HSL que aparecerem no visor.

6- Repetir duas vezes mais os passos de 3 a 5, e anotar as três medições.

## TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO - TCLE

Você (ou a pessoa pela qual você é responsável) está sendo convidado para participar de uma pesquisa da PUCRS. O título da pesquisa é 'FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS DE SNPs EM GENES QUE REGULAM A PIGMENTAÇÃO DE OLHOS PARA FINS FORENSES: ESTUDO EM TRIOS FAMILIAIS (MÃE-PAI-CRIANÇA)', que tem por objetivo estudar partes do DNA (o DNA é o que define as características de cada pessoa) que possam estar relacionadas à cor da pele e à cor dos olhos das pessoas. As partes do DNA que vamos estudar serão aquelas responsáveis pela produção dos pigmentos, tanto na pele como nos olhos. Imagina-se que se o DNA induz a produção de mais pigmentos, a pele e os olhos das pessoas ficam mais escuros. Nós queremos avaliar se isso está mesmo associado. Você não terá custo algum em participar deste estudo.

Nós precisamos: 1- registrar a cor da pele do seu braço com um medidor; 2- fotografar cada um de seus olhos (não o rosto); 3- coletar células da sua bochecha com um cotonete esterilizado. Talvez a coleta da saliva lhe cause um desconforto momentâneo, mas nós qualificamos que sua participação nesta pesquisa lhe oferece um risco mínimo, isto é, o mesmo risco existente em atividades rotineiras como conversar, tomar banho, caminhar, etc. Este estudo não trará benefícios diretos para você (ou à pessoa pela qual você é responsável), nesse momento. Mas, esta pesquisa poderá trazer benefícios para a sociedade e para a ciência. Sinta-se à vontade para fazer qualquer pergunta sobre o estudo e/ou para esclarecer suas dúvidas, nós vamos sempre lhe responder. A qualquer momento diga se você tiver vontade de sair da pesquisa, sem necessidade de se explicar. Você não será prejudicado se resolver tomar essa decisão; apenas comunique aos pesquisadores se quiser abandonar a pesquisa e nós respeitaremos a sua decisão.

Toda e qualquer informação atualizada sobre este estudo estará à disposição de todos os que participam da pesquisa. Os resultados deste estudo poderão ser comunicados sempre que você quiser. Sua dignidade e privacidade, em relação a qualquer dado, serão mantidas sempre. Seus dados estarão protegidos pelos pesquisadores responsáveis contra qualquer tipo de uso que não os aqui previstos. Além disto, os dados receberão um número para garantir o anonimato dos participantes da pesquisa, ou seja, no estudo dos dados, a identidade será desconhecida tanto pela equipe do trabalho como por qualquer outra pessoa externa.

A partir das mesmas amostras poderão ser futuramente estudados outras partes do DNA. Para isso, você poderá permitir que sua amostra seja usada em novos estudos (através de re consentimento por meio de um TCLE específico referente ao novo projeto de pesquisa), que serão avaliados novamente pelo Comitê de Ética e Pesquisa da PUCRS e/ou Comissão Nacional de Ética em Pesquisa-CONEP. Por isso, entendemos que é importante armazenar o material que estamos coletando hoje.

Eu, \_\_\_\_\_ (*preencher com nome do participante ou responsável*) fui informado dos objetivos da pesquisa de maneira clara e objetiva. Recebi informação a respeito dos procedimentos do estudo e esclareci minhas dúvidas. Sei que em qualquer momento poderei solicitar novas informações e/ou modificar minha decisão se eu desejar. A equipe da Dra. Clarice S. Alho certificou-me de que todos os meus dados serão confidenciais, bem como minha rotina de vida não será modificada em razão desta pesquisa. Assim, como participante de pesquisa, aceito participar deste estudo.

Caso eu tenha novas perguntas ou queira esclarecer algo sobre os meus direitos, ou se penso que fui prejudicado pela minha participação neste estudo, eu posso procurar a Dra. Clarice S. Alho nos telefones (51) 9113-6334 ou (51) 3320-3545, ou na Av. Ipiranga, 6681, Prédio 12; ou contatar o Comitê de Ética em Pesquisa da PUCRS (CEP-PUCRS) no telefone (51) 3320-345 ou na Av. Ipiranga, 6690, Prédio 60, Sala 314, em Porto Alegre/RS, de segunda à sexta-feira em horário comercial (a ligação para celular pode ser realizada a cobrar). O CEP-PUCRS é grupo de assessores das áreas biológicas, saúde, exatas, sociais e humanas, e tem por objetivo avaliar os aspectos éticos e metodológicos das pesquisas envolvendo os seres humanos.

**Acredito ter sido suficientemente informado a respeito das informações que li, ou que foram lidas para mim, descrevendo o estudo e concordo voluntariamente em participar desta pesquisa.**

Nome e assinatura do participante ou do responsável	Local e data
Nome e assinatura do Pesquisador ou pessoa por ele delegada	Local e data

Gostaríamos de pedir seu consentimento para que o material que você forneceu para estudos seja devidamente guardado e reservado para ser usado em outras pesquisas futuramente.

Eu também permito que meu material seja armazenado na PUCRS: para a realização de futuros trabalhos, consinto em ser consultado através de meus dados de contato, que constarão neste mesmo TCLE. Minha permissão vale desde que os novos estudos sejam sempre devidamente aprovados por um CEP. Fui informado também que caso alguma informação considerada importante seja identificada no meu DNA em decorrência da pesquisa, os pesquisadores poderão buscar contato comigo para oferecer acesso a tal informação.

Declaro que recebi uma via do presente TCLE (conforme a Resolução CNS 466/2012 item IV.3.f, IV.5.d., este TCLE foi elaborado em DUAS VIAS igualmente válidas, assinadas e rubricadas em todas as suas páginas, sendo que uma será arquivada pelo pesquisador responsável e outra será fornecida ao participante de pesquisa).

Nome e assinatura do participante ou do responsável	Local e data
---	--------------

Na impossibilidade de autonomia para a leitura deste Termo de Consentimento Livre e Esclarecido, o mesmo foi lido para \_\_\_\_\_ (preencher com nome do participante ou responsável da pesquisa) pelo pesquisador \_\_\_\_\_ (preencher com o nome do pesquisador) enquanto eu estava presente como testemunha.

Nome e assinatura da Testemunha	Local e data
---------------------------------	--------------

**DADOS DE CONTATO:**

Nome completo: \_\_\_\_\_

Nome da mãe: \_\_\_\_\_

Data de nascimento: \_\_\_\_/\_\_\_\_/\_\_\_\_

Telefone de contato 1: \_\_\_\_\_

Telefone de contato 2: \_\_\_\_\_

Endereço para contato (físico ou eletrônico): \_\_\_\_\_

Nome e assinatura do Pesquisador ou pessoa por ele delegada	Local e data
---	--------------





Pontifícia Universidade Católica do Rio Grande do Sul  
Pró-Reitoria Acadêmica  
Av. Ipiranga, 6681 - Prédio 1 - 3º andar  
Porto Alegre - RS - Brasil  
Fone: (51) 3320-3500 - Fax: (51) 3339-1564  
E-mail: [proacad@pucrs.br](mailto:proacad@pucrs.br)  
Site: [www.pucrs.br/proacad](http://www.pucrs.br/proacad)