

RODRIGO RODENBUSCH

**ANÁLISE DE SNPs EM GENES DE PIGMENTAÇÃO HUMANA EM INDIVÍDUOS  
COM ALTO OU BAIXO CONTEÚDO DE MELANINA**

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

Orientador

Profa. Dra. Clarice Sampaio Alho

Porto Alegre

2014

## AGRADECIMENTOS

Primeiramente, gostaria de agradecer à minha orientadora Prof<sup>a</sup>. Dr<sup>a</sup>. Clarice Sampaio Alho, que tive a oportunidade e alegria em ser seu primeiro aluno de Iniciação Científica na PUCRS, onde durante estes mais de 15 anos levei não só ensinamentos profissionais mas também ensinamentos para a minha vida pessoal.

Aos queridos colegas e amigos do Laboratório de Investigação de Paternidade, agradeço por toda a compreensão e apoio durante os momentos de ausência, momentos estes essenciais para a execução deste trabalho. Agradeço, principalmente ao colega André Gastaldo, pela ajuda nas correções da língua portuguesa.

Às bolsistas, e hoje mestrandas, Thayane Rispoli e Mayara Jorgens Prado, pela ajuda nas coletas e parceria nas viagens pelo Estado, a fim de conseguirmos todas as amostras necessárias para esta pesquisa. Aos colegas, do LabGen da PUCRS, Diego Gubert, Fernanda Sawitzki, Eduardo Capelesso e Bruna Schroeder Oliveira, por abraçarem esta causa da predição fenotípica e ajudarem em todas as etapas da coleta e validação da mesma. Agradeço também por toda ajuda com o levantamento populacional dos bancos de dados de frequência e conferência dos mesmos; ao Eduardo por passar horas organizando, normalizando e formatando as fotos dos olhos dos indivíduos, tarefa esta que foi muito árdua.

À minha família, pela base e apoio durante as horas de inquietude em busca de amostras, de resultados contundentes e na preparação dos seminários de acompanhamento. Pela vibração deles pelas minhas conquistas.

Agradeço do fundo do meu coração, à minha mãe Sirlei (*in memoriam*), pelo carinho incondicional, pelas incontáveis noites que passamos no hospital, onde eu, nervoso, preparando os seminários, tinha o seu apoio sempre dizendo que daria tudo certo. Embora ela não esteja mais neste plano conosco, sei que está muito orgulhosa do caçula dela.

Ao meu filho Vinícius Rodenbusch, agradeço pelo carinho e ao mesmo tempo me desculpo pela ausência em alguns momentos durante a execução deste trabalho. Filho, tu és uma das pedras fundamentais que sustentam a minha saúde afetiva.

Por fim, e não menos importante, à minha esposa, Jéssica Pegoraro Rodenbusch, agradeço pelo amor e compreensão ao longo desta caminhada, pelo incentivo e por nunca me deixar desistir nos momentos de maior aflição. Agradeço também a ela e a Deus pela chegada nos próximos meses da nossa filha Valentina, que será muito amada pelos seus pais e iluminará nossa família.

Aqueles não lembrados ou não citados me perdoem pelo esquecimento, pois apesar desta tese ser de responsabilidade e natureza individual, sem o apoio e incentivo de todos, não seria possível alcançar este objetivo.

## RESUMO

A compreensão da função e expressão dos genes envolvidos nos traços externamente visíveis (EVC; do inglês *externally visible characteristics*) têm sido amplamente utilizada em vários estudos de evolução humana e investigações forenses. Para este último propósito, vários esforços têm sido feitos para descobrir um modelo eficiente e fácil para a predição da cor da pele e dos olhos em seres humanos. A vantagem óbvia da predição de tais EVCs, através da utilização do DNA, é sua incorporação na rotina em laboratórios forenses, sendo assim aplicada em investigações policiais. Em nosso estudo, relacionamos o genótipo de oito SNPs em genes relacionados com a pigmentação (rs4778138 - OCA2; rs12913832 - HERC2; rs16891982 - SLC45A2; rs8045560 - MC1R; rs1426654 - SLC24A5; rs2733832 - TYRP1; rs1042602 - TYR; rs916977 - HERC2) com diferentes abordagens analíticas. Este painel de SNPs considerou as frequências alélicas obtidas de dados do HapMap e Alfred a partir de indivíduos com Alto Conteúdo de Melanina (HMC; do inglês *High Melanin Content*; a partir de populações africanas), e Baixo Conteúdo de Melanina (LMC; do inglês *Low Melanin Content*; a partir de populações europeias) e definiu os alelos H (para predizer os HMC) e alelos L (para predizer os LMC). A distribuição cumulativa dos alelos H e L nos dois grupos com características de pigmentação fenotipicamente distintas, dos 134 indivíduos da nossa população, mostrou que 82% dos indivíduos HMC (N = 61) tinham oito ou mais alelos H e 100% dos indivíduos de LMC (N = 73) tinham menos de oito alelo H, com o valor de precisão de 96,3%. Outras abordagens como AUC (do inglês; *Area Under the Receiver Operating Characteristic Curve*), cálculo de PGL (do inglês; *Pathway Genetic Load*) e GP (do inglês; *Genetic Probability*) foram realizadas. A análise AUC foi de 0,99 tanto para a predição fenotípica dos HMC quanto LMC; as análises PGL, para o painel com 8 SNPs, teve 93% de concordância genótipo-fenótipo nos HMC ou LMC; e a abordagem GP mostrou 91% de concordância para predição dos fenótipos HMC e LMC. Nossa tecnologia de genotipagem de alto rendimento, combinada com diferentes abordagens analíticas, chegou a uma precisão muito alta para predizer os fenótipos extremos de pigmentação humana. Acreditamos que esta técnica de fenotipagem forense pelo DNA (FDP; do inglês *forensic DNA phenotyping*), seria particularmente útil nos casos em que os perfis genéticos de locais de crime não fossem encontrados no bancos de dados de DNA ou para ajudar a classificar cadáveres degradados, esqueletos, ou vestígios biológicos de pessoas desaparecidas.

**PALAVRAS CHAVE:** Pigmentação Humana, DNA Forense, Predição Fenotípica.

## ABSTRACT

The understanding of gene function in the externally visible characteristics (EVC) expression has several uses in human population evolution studies or in forensic investigations. To this last, some effort has been done to discover an efficient and easy model for prediction of skin and eye color in humans. The obvious advantage of the prediction of such EVCs through the use of DNA is to be incorporated as routine in forensic labs and to be applied to police investigations. In our study we combined the genotyping of eight SNPs in pigment-related genes (*rs4778138 - OCA2*; *rs12913832 - HERC2*; *rs16891982 - SLC45A2*; *rs8045560 - MC1R*; *rs1426654 - SLC24A5*; *rs2733832 - TYRP1*; *rs1042602 - TYR*; *rs916977 - HERC2*) with different analytical approaches. Considering this SNP panel we evaluated allele frequencies from HAPMAP and ALFRED data obtained from subjects with High Melanin Content (HMC; from African populations), and Low Melanin Content (LMC; from European populations) and defined the alleles H (to predict HMC subjects) and alleles L (to predict LMC subjects). The cumulative distribution of alleles H and alleles L in two phenotypically different color groups of 134 South Brazilian subjects showed that 82% of HMC subjects (N = 61) had eight or more allele H and 100% of LMC subjects (N = 73) had less than eight allele H, with accuracy value of 96.3%. We performed other analyses using AUC (Area Under the Receiver Operating Characteristic Curve), PGL (Calculation of Pathway Genetic Load), and GP (Genetic Probability) approaches. The AUC was 0.99 in predicting both HMC and LMC phenotypes; PGL showed the eight SNPs panel had 93% of concordance between genotype and HMC or LMC phenotypes; and GP approach showed 91% of concordance between prediction and HMC or LMC phenotypes. Our high-throughput genotyping technology combined with different analytical approaches reached very high accuracy to predict the extreme phenotypes of human pigmentation. We believe this forensic DNA phenotyping (FDP) technique would be particularly useful in cases in which the genetic profiles of crime scenes were not found in the DNA data banks or to help classify degraded cadavers skeletons, or biological clues of dismissed people.

**KEYWORDS:** Human Pigmentation, Forensic DNA, Phenotypic Prediction

## LISTA DE ABREVIATURAS

EVC	Características externamente visíveis; do inglês <i>externally visible characteristics</i>
SNP	do inglês, <i>Single Nucleotide Polymorphism</i>
HMC	Alto conteúdo de melanina; do inglês <i>High Melanin Content</i>
LMC	Baixo conteúdo de melanina; do inglês <i>Low Melanin Content</i>
AUC	do inglês <i>Area Under the Receiver Operating Characteristic Curve</i>
PGL	do inglês <i>Pathway Genetic Load</i>
GP	do inglês <i>Genetic Probability</i>
FDP	fenotipagem forense pelo DNA; do inglês <i>Forensic DNA Phenotyping</i>
MC1R	do inglês, <i>melanocortin 1 receptor</i>
ASIP	do inglês, <i>Agouti signaling protein</i>
MATP	do inglês, <i>Membrane associated transporter protein</i>
SLC45A2	do inglês, <i>Solute Carrier Family 45, member 2</i>
SLC24A5	do inglês, <i>Solute Carrier Family 24, member 5</i>
TYR	do inglês, <i>Tyrosinase</i>
TYRP1	do inglês, <i>Tyrosinase-related protein 1</i>
OCA2	do inglês, <i>Oculocutaneous albinism type II</i>
HERC2	do inglês, <i>HECT domain and RCC1-like domain-containing protein 2</i>
VNTR	do inglês, <i>Variable Number of Tandem Repeat</i>
STR	do inglês, <i>Short of Tandem Repeat</i>
AIM	do inglês, <i>Ancestry Informative Marker</i>
PCR	do inglês, <i>Polymerase Chain Reaction</i>
CREB	do inglês, <i>cyclicAMP response-element binding protein</i>
MITF	do ingles, <i>Microphthalmia-associated Transcription Factor</i>
DQ	dopaquinona
UV	ultravioleta

## SUMÁRIO

<b>CAPÍTULO 1 - APRESENTAÇÃO DO TEMA E OBJETIVOS</b>	01
<b>1- REFERENCIAL TEÓRICO</b>	01
1.1 Importância da Predição de Fenótipo	02
1.2 Biologia da Pigmentação	02
1.2.1 Melanogênese	02
1.2.2 Pigmentação da Pele	04
1.2.3 Pigmentação do Olho	05
1.3 Pigmentação e Ancestralidade	06
1.4 Genes e SNPs Envolvidos na Determinação de Pigmentação de Pele e Olhos	08
1.4.1 Gene HERC2: rs916977 e rs12913832	09
1.4.2 Gene OCA2: rs4778138	10
1.4.3 Gene SLC24A5: rs1426654	11
1.4.4 Gene SLC45A2: rs16891982	12
1.4.5 Gene TYR: rs1042602	12
1.4.6 Gene TYRP1: rs2733832	13
1.4.7 Gene MC1R: rs8045560	13
1.5 Frequências alélicas dos SNPs em populações Africanas e Européias	13
<b>2- OBJETIVOS</b>	15
<b>CAPÍTULO 2 - ARTIGO CIENTÍFICO</b>	16
<b>CAPÍTULO 3 - CONSIDERAÇÕES FINAIS</b>	43
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	44
<b>ADENDOS</b>	50

## CAPÍTULO 1 - APRESENTAÇÃO DO TEMA E OBJETIVOS

### 1- REFERENCIAL TEÓRICO

Ao longo dos anos foram implantadas diferentes técnicas a fim de caracterizar e individualizar pessoas, como por exemplo os exames antropométricos, impressões digitais e estudos envolvendo marcadores genéticos e bioquímicos (STRs, VNTRs, polimorfismos dos sistemas HLA, ABO e de outras proteínas séricas) (Kobachuk, 2012). Os marcadores moleculares conhecidos como STRs (*Short tandem repeat*) são altamente polimórficos e amplamente utilizados na identificação humana em casos forenses há quase trinta anos (Jobling *et al.*, 2004). Na atualidade, o uso dos marcadores STR é muito vasto no sistema forense, sendo aplicado nos casos de identificação forense criminal e em testes de paternidade ou parentesco (Yoshida *et al.*, 2011). A confiabilidade elevada do resultado de uma análise de STR permite que ele seja usado como prova respeitável nos processos penal ou cível. Além disto, análises conjuntas de marcadores moleculares do tipo SNPs (*single nucleotide polymorphisms*) podem também auxiliar na identificação humana, pois permitem a obtenção de perfis genômicos únicos para cada indivíduo (Kobachuk, 2012). SNPs, que representam a classe mais abundante de polimorfismos humanos, podem servir tanto para a discriminação individual como para a predição de Características Externas Visíveis (EVCs). Sua abundância é a principal razão do grande interesse atual no campo forense; e o mapeamento do genoma humano tornou possível o desenvolvimento de mapas de genótipos e de haplótipos que podem caracterizar fenotipicamente um ser humano (Stacey *et al.*, 2002).

Entre as EVCs, a variação de pigmentação em humanos tem sido um bom alvo. Mutações polimórficas do tipo SNPs, podem determinar substituições de aminoácidos na proteína, alterando as propriedades funcionais da proteína traduzida e sendo expressa em fenótipos distintos (O'rahilly, 2009). Os traços fenotípicos mais promissores para a identificação forense são aqueles relacionados à pigmentação da pele, olhos e cabelos, por serem características muito marcantes, de fácil visualização e constituírem um dos fenótipos mais variáveis na população humana (Jablonski e Chaplin, 2000; Sturm *et al.*, 2001; Parra, 2007; Kayser e Schneider, 2009). O desenvolvimento de técnicas quantitativas para mensurar a pigmentação proporcionou uma maior objetividade na classificação da ampla variação de pigmentação existente, permitindo identificar de forma mais precisa os genes que de alguma forma exercem influência sobre essa característica (Jablonski, 2004).

Apesar de ser promissora a utilização de marcadores genéticos de traços fenotípicos na identificação humana para fins forenses, a descrição de genes e do mecanismo pelo qual eles influenciam a definição das características fenotípicas é de difícil elucidação, uma vez que essas características apresentam um padrão de herança complexo, por serem determinadas por múltiplos genes (poligenia) e por sofrerem forte influência do meio ambiente (Sturm e Larsson, 2009). Não obstante, estudos demonstram que a base genética da variação normal da pigmentação é passível de ser decifrada (Van Daal,

2008) se forem avaliadas as variantes genéticas adequadas (Liu *et al.*, 2009). Nós estamos, neste momento, tentando identificar quais são, portanto, elas.

### **1.1 – Importância da Predição de Fenótipo**

Estudos de associação genótipo-fenótipo sobre a cor do cabelo, a cor da íris e a pigmentação da pele têm identificado SNPs em genes diretamente envolvidos na síntese de pigmentos (Sturm, 2009; Walsh *et al.*, 2011a; Walsh *et al.*, 2011b; Walsh *et al.*, 2013). Todos eles sustentam que há uma necessidade grande por parte dos órgãos policiais de possuir uma base de dados com perfis genéticos que possam auxiliar na solução de crimes, excluindo ou inserindo indivíduos em processos criminais que buscam identificar suspeitos ou vítimas (Pulker *et al.*, 2007). Especialmente no caso de não se ter informações sobre o perfil genético de referência (para confronto de alelos de STRs), a amostra questionada pode ser estudada para fins da predição do fenótipo do indivíduo (Kayser *et al.*, 2011).

### **1.2 - Biologia da Pigmentação**

Acredita-se que existam mais de 120 genes de alguma forma envolvidos nas vias de pigmentação, os quais agem em diferentes estágios do processo de produção da melanina (Sturm, 2006; Branicki *et al.*, 2009). Tais genes agem em caráter quantitativo, influenciando cumulativamente nas diferentes tonalidades de coloração (Tully, 2007).

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

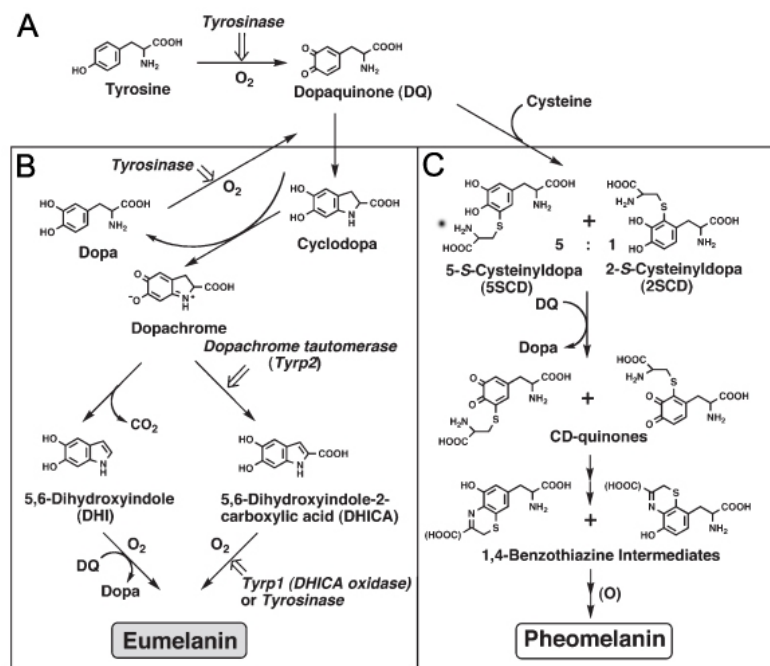
#### **1.2.1- Melanogênese**

O regulador 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-dihidroxifenilalanina (DOPA) como agente intermediário (Kondo e Hearing, 2011) (Figura 1A).



**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 1B).

**Feomelanogênese:** A feomelanogênese se dá pela adição de cisteína na oxidação da DQ para produzir dois isômeros principais de DOPAcisteinila (CD): a 5SCD (do inglês 5-S-cysteinyl-dopa) e a 2SCD ((do inglês 2-S-cysteinyl-dopa). 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 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 1C).



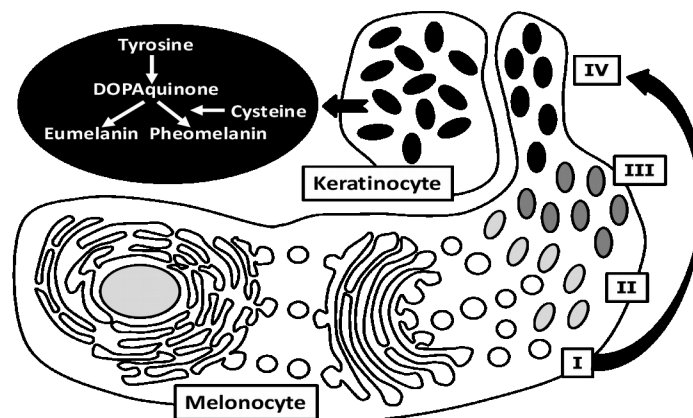
**Figura 1** – Via biossintética da produção de eumelanina/feomelanina a partir de tirosina. As atividades da TRY, TYRP1 e DCT estão envolvidas na produção da eumelanina, mas apenas TYR e cisteína são necessários para a produção de feomelanina a partir de DQ (Ito e Wakamatsu, 2010).

### 1.2.2- Pigmentação na Pele

A melanina é gradualmente produzida nos melanossomas que, quando maduros, são transportados ao longo das projeções dendríticas dos melanócitos até os queratinócitos adjacentes da pele (Sturm *et al.*,

2006). O número de melanócitos entre os diferentes tons de pele é constante, mas peles mais escuras apresentam melanossomas mais densos e individualmente dispersos, enquanto peles mais claras apresentam melanossomas menos densos e menores (Sturm *et al.*, 2012).

Os melanossomas desenvolvem gradualmente o pigmento enquanto amadurecem (Slominski *et al.*, 2004; Lin e Fisher, 2007; Ho *et al.*, 2011). Após sintetizado, o pigmento é passado aos queratinócitos circundantes da pele e do folículo piloso (Sturm, 2006) e, por fim, para as células do epitélio pigmentar da retina (Kondo e Hearing, 2011). O acúmulo de melanina e a distribuição do melanossomas variam quantitativamente entre indivíduos de diferentes grupos étnicos (Alaluf *et al.*, 2002; Sturm, 2006). Os melanossomas são divididos em quatro estágios de maturação (I a IV), determinados pela sua estrutura, quantidade, qualidade e arranjo da melanina produzida (Kushimoto *et al.*, 2001; Costin e Hearing, 2007) (Figura 2). O melanossoma no estágio I é uma organela esférica, com ausência da enzima tirosinase (enzima que promove a quebra da tirosina, formando os elementos precursores da eumelanina e feimelanina) e, ausência também, de componentes internos estruturais. No estágio II a organela apresenta uma estrutura ovoide, atividade de TYR e a expressão de uma proteína da matriz estrutural, conhecida como PMEL17, necessária para a produção da matriz fibrilar interna, e pela deposição mínima de melanina (Berson *et al.*, 2001; Sturm, 2006; Costin e Hearing, 2007; Kondo e Hearing, 2011). No estágio III, a síntese de melanina é iniciada e o pigmento é depositado sobre fibras internas. No estágio IV de desenvolvimento, o melanossoma apresenta formato elíptico ou elipsoide. Estes estádios de desenvolvimento referem à eumelanossomos (que sintetizam a eumelanina), e nos feomelanossomos (que sintetizam a feomelanina) o processo é similar com a diferença que durante o processo de maturação, há ausência de fibrilas (Costin e Hearing, 2007; Brenner e Berking, 2010).



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

### 1.2.3- Pigmentação no Olho

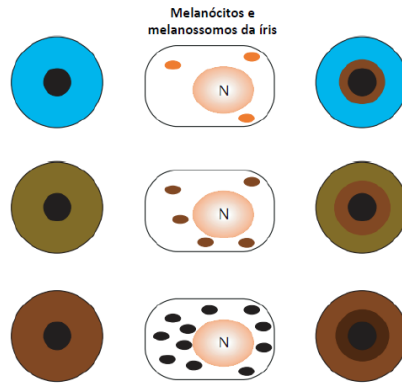
Os melanócitos dos olhos têm duas origens distintas, os que se encontram no epitélio pigmentoso da íris (IPE, do inglês *iris pigment epithelium*) que têm origem neuroectodérmica, e os melanócitos estromais, que têm a mesma origem embrionária dos melanócitos dérmicos (oriundos da crista neural), e

migram através do trato uveal durante o desenvolvimento. Nas íris de coloração castanha existe uma abundância de melanina na camada basal anterior e no estroma, enquanto que nas íris de coloração azul essas camadas contêm muito pouco de melanina (Sturm e Larsson, 2009).

A íris, região ocular composta por tecido conectivo e muscular, é responsável pelo controle da entrada da luz incidente através de sua abertura central denominada pupila, regulando a formação das imagens na retina. Essa região é subdividida em cinco camadas, sendo a mais externa denominada borda anterior, seguida do estroma, esfíncter e músculos dilatadores e, mais internamente, o epitélio posterior pigmentado (Sturm e Frudakis, 2004). Dessas, a borda anterior e o estroma são as principais camadas contribuintes para a variação da coloração da íris (Wilkerson *et al.*, 1996), uma vez que o epitélio posterior pigmentado apresenta a mesma pigmentação em todas as cores observadas, não contribuindo de forma relevante para a variação fenotípica final (Sturm e Frudakis, 2004). Íris de coloração marrom apresentam melanócitos com alto conteúdo de melanina na borda anterior e no estroma, enquanto em íris azuis, essas camadas possuem baixo teor de melanina, permitindo maior passagem da luz e consequente refletância de ondas curtas azuis pelas fibras de colágeno presentes. No estroma, os dendritos dos melanócitos são geralmente paralelos à superfície da íris, os quais tendem a se agrupar na borda anterior, e correspondem a aproximadamente 66% das células componentes dessas camadas, independentemente da coloração da íris (Wilkerson *et al.*, 1996).

Ao contrário do que ocorre na pele, onde a melanina é continuamente produzida e secretada, os melanossomos da íris não são secretados dos melanócitos, mas sim retidos e acumulados no citoplasma destes. As variações na coloração observadas são resultado das diferentes quantidades e tipos de melanossomos presentes nos melanócitos, sendo a proporção de eumelanina/feomelanina geralmente maior em colorações mais escuras, e colorações mais claras apresentando maior concentração do pigmento feomelanina (Sturm e Larsson, 2009).

Sabe-se que, apesar de todas as colorações de olhos possuírem um número similar de melanócitos, os olhos azuis contêm um mínimo de pigmentos e poucos melanossomos; os olhos verdes são resultado de moderados níveis de pigmentos, intensidade de melanina e grande quantidade de partículas melanossomais (Sturm e Frudakis, 2004) (Figura 3).



**Figura 3** - Pigmentação dos olhos. Diferenças entre os tipos e quantidades de melanossomos nos melanócitos dos olhos mais claros aos mais escuros (retirado de Sturm e Frudakis, 2004).

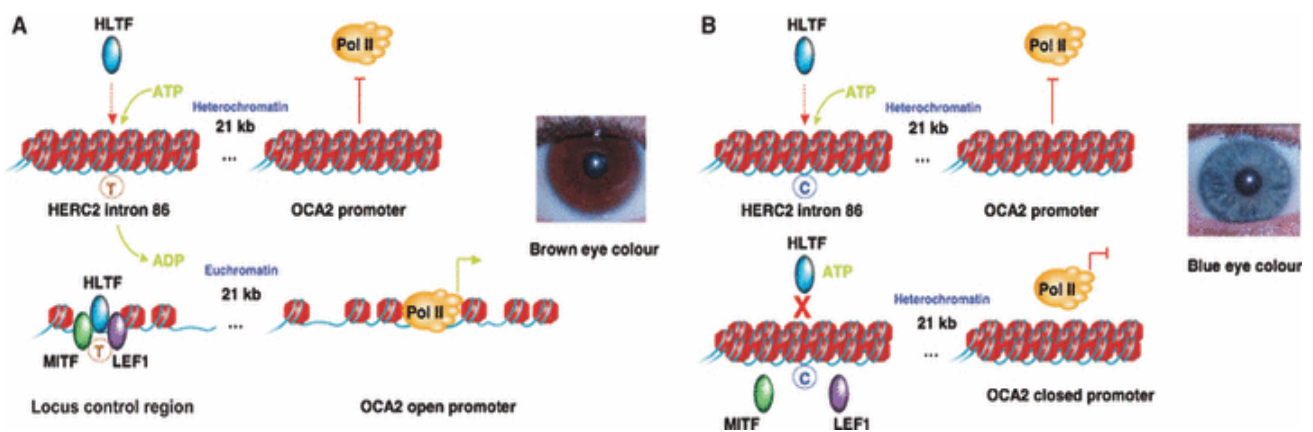
### 1.3 - Pigmentação e Ancestralidade

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

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

Diante da relação fortemente positiva entre pigmentação e origem étnica, poder-se-ia questionar se a presença de determinado alelo seria apenas um marcador de ancestralidade ou se, de fato, poderia ele estar diretamente envolvido na determinação fenotípica da pigmentação. Na tentativa de solucionar esta questão, muitos genes têm sido estudados, e os produtos proteicos e/ou a expressão de suas variantes alélicas têm sido associados à síntese ou não de melaninas. Porém, no presente momento é possível identificar apenas alguns SNPs em genes os quais, de fato, estão diretamente envolvidos na síntese de

pigmentos de pele e de olhos, independentemente da origem étnica da população em estudo (Kenny *et al.*, 2012). Por exemplo, como será citado a seguir, o SNP rs12913832 do gene *HERC2*, o qual se encontra a 21kb *upstream* do promotor de *OCA2*, regula a expressão de *OCA2* pelo dobramento da cromatina que permite o ancoramento dos fatores de transcrição de *OCA2*. O produto da expressão de *OCA2*, a proteína P, é essencial para a síntese de melanina. O alelo T (ou alelo A, considerando a fita complementar) do SNP rs12913832 ao permitir a abertura de cromatina, e o recrutamento dos fatores de transcrição de *OCA2*, induz à coloração mais escura, enquanto o alelo C (ou alelo G, considerando a fita complementar) mantém a cromatina mais fechada, tendo menos eficácia no recrutamento destes fatores, ocasionando uma coloração mais clara (Figura 4).



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

Mesmo se localizando no interior de segmentos gênicos chave para a produção de melanina, poucos os alelos foram, contudo, já reconhecidos como sendo os próprios responsáveis pelo efeito variante de síntese de pigmento. Por não ter sido ainda completamente elucidada a implicação efetiva no fenótipo de cada variante alélica, algumas delas podem estar associadas à pele/olho escuro ou à pele/olho claro apenas por estarem em desequilíbrio de ligação com uma variante, de fato, fenótipo-efetiva. De todas as formas, em geral, têm sido usados para explicar a coloração escura aqueles alelos cuja frequência está mais elevada nas populações africanas negras (alelos originais/ancestrais) e os alelos derivados, cuja frequência está mais elevada nas populações europeias, para explicar a coloração clara. O desafio na análise da maioria dos SNPs ainda é saber se o marcador é de fenótipo ou de ancestralidade. No intuito de focar nos marcadores de fenótipo, têm sido retirados dos estudos associativos e/ou de predição de fenótipo aqueles alelos derivados (isto é, mutações polimórficas mais recentes) cujas frequências estão mais aumentadas nas populações africanas, como é o caso do alelo derivado (mais recente) do SNP rs26722 do gene *SLC45A2* (regulador do pH que controla a via da síntese de melanina, no interior do melanócito). Algumas populações africanas de fenótipo escuro mostram a presença de um alelo derivado raro (alelo A; Lys272) deste SNP e, a despeito do seu fenótipo claro, populações caucasianas têm em frequência o alelo original deste SNP (alelo G; Glu272) similares às de aborígenes australianas escuros (frequência de G ~0,98) (Graf *et*

*al.*, 2005). Devido ao fato de o alelo original (G; Glu272) estar presente em grupos populacionais com diferentes fenótipos de pigmentação (caucasianos claros e aborígenes escuros), e ao fato de populações africanas de cor escura conterem ambas variantes (Lys272 e Glu272), Graf *et al.* (2005) concluíram que este polimorfismo não estaria associado com a determinação de pigmentação, mas sim que se trataria de uma variação associada à distribuição/isolamento geográfico. Assim, o alelo raro A surgiu como uma mutação recente apenas em africanos, e não interferiu na expressão da cor.

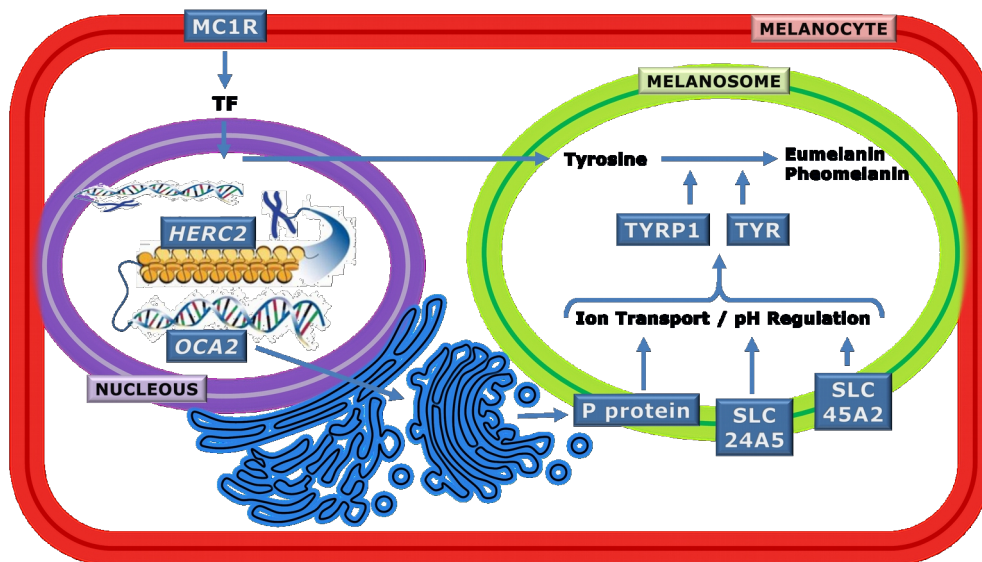
#### 1.4 - Genes e SNPs envolvidos na determinação de pigmentação de pele e olhos

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

Com base na literatura disponível, um estudo de revisão do nosso grupo selecionou oito SNPs em genes candidatos ao estudo de previsão de pigmentação em diversas populações, por estarem significativamente relacionados com a manifestação da pigmentação em seres humanos. A base do nosso estudo foi: Tully, 2007; Sulem *et al.*, 2007; Sturm *et al.*, 2008; Branicki *et al.*, 2009; Giardina *et al.*, 2011; Walsh *et al.*, 2011a; Walsh *et al.*, 2011b, Donnelly *et al.*, 2012; Allwood *et al.*, 2013; Liu *et al.*, 2013; Walsh *et al.*, 2013. No Quadro 1, estão apresentadas informações acerca dos genes e SNPs por nós selecionados e, na Figura 5, se pode visualizar a localização de cada produto gênico ou estrutura em questão. Adicionalmente, no ADENDO há informações detalhadas sobre cada SNP.

**Quadro 1. Alguns Genes e SNPs para predição de pigmentação em Humanos**

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	Leu 374 Phe	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



**Figura 5-** Indicação na estrutura celular dos genes (citados no quadro 1) e/ou seus produtos proteicos para predição de pigmentação em Humanos (Figura do autor)

#### 1.4.1- Gene *HERC2*: rs916977 e rs12913832

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

Visser *et al.* (2012) mostraram que o SNP rs12913832 de *HERC2*, que se encontra a 21kb *upstream* do promotor de *OCA2*, regula a expressão deste gene pelo dobramento da cromatina que gera um *loop* de longo alcance que acaba por promover a expressão de *OCA2*. Esta atividade é mediada pelos fatores de transcrição HLTF, LEF1 e MITF. Porém, diferentes versões alélicas de *HERC2* impedem o desenrolamento da cromatina, o que reduz a ligação dos fatores de transcrição de *OCA2*, reduzindo consequentemente sua taxa transcricional. O alelo T (ou alelo A, considerando a fita complementar) do SNP rs12913832 permite a abertura da cromatina e consegue recrutar os fatores de transcrição de *OCA2*, o que permite a produção de melanina e ocasiona uma coloração mais escura, enquanto o alelo C (ou alelo G, considerando a fita complementar) apresenta uma diminuição no recrutamento destes fatores, impedindo a síntese plena de pigmento e ocasionando uma coloração mais clara. Assim, via controle da expressão de *OCA2*, o gene *HERC2* tem o papel predominante na determinação de cor. Dados populacionais indicam que os SNPs rs916977 e rs12913832, localizados no íntron 12 e íntron 86, respectivamente, são melhores preditores de

cor de olho do que três SNPs já descritos no interior do íntron 1 do próprio gene *OCA2* (rs7495174, rs6497268 e rs11855019) (Branicki *et al.*, 2009). Estudos em outras populações também indicam o SNP rs12913832 como sendo o SNP com maior efeito na predição da cor de olhos (Allwood *et al.*, 2013).

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

#### **1.4.2- Gene *OCA2*: rs4778138**

O gene que, ao conter mutações deletérias, ocasiona o albinismo oculocutâneo tipo 2 (gene *OCA2*), codifica uma proteína com 12 domínios transmembrana denominada proteína P (Sturm *et al.*, 2008). Acredita-se que a proteína P esteja envolvida no transporte de ânions, na regulação do pH melanossomal (Visser *et al.*, 2012), e que esteja envolvida no tráfego de proteínas internas como a tirosinase (TYR) e as proteínas associadas a tirosinase 1 (TYRP1) (Valenzuela *et al.*, 2011). *OCA2* por si só, independente de *HERC2*, também desempenha um papel na variação de coloração comum em olhos, cabelo e pele. Três SNPs que se encontram no intron 1 de *OCA2* mostraram ter uma forte associação com a coloração dos olhos (Liu *et al.*, 2013). Além disto, o SNP rs1800407 encontrado no exón 13 do gene *OCA2* ocasiona uma troca do tipo Arg419Gln na sequência polipeptídica da proteína P (Donnelly *et al.*, 2012), esta troca aumenta a penetrância do fenótipo de olhos azuis associados ao SNP rs12913832 (Sturm *et al.*, 2008; Walsh, *et al.*, 2011a).

A proteína P, relacionada ao transporte de sódio/sulfeto (família SLC13), pode estar diretamente implicada na regulação do pH melanossomal (Kondo e Hearing, 2011). Foi observado que a proteína P atuava na permuta de cátions  $\text{Na}^+/\text{H}^+$ , e/ou como um transportador de glutamato. Ambas as funções indicam que este produto do gene *OCA2* esteja envolvido no fornecimento de substrato para a TYR. A proteína P pode ainda estar envolvida no tráfego intracelular da enzima TYR, durante a maturação dos melanossomos (Eiberg *et al.*, 2008). O pH do melanossomo regula a atividade da TYR afetando, portanto, uma série de fatores na síntese da melanina, como a taxa de produção de melanina, a proporção de eumelanina/feomelanina e a maturação de melanossomos nos melanócitos e células de melanoma. Resultados experimentais indicaram que a produção de melanina é ideal em ambiente celular com pH próximo a pH-neutro (cerca de pH= 6.8). Paralelamente, a síntese da melanina é quase ausente em pH menor que 5.5. Em um pH menos ácido, por outro lado, a melanogênese pode prosseguir sem qualquer alteração na abundância de proteínas melanogênicas (Ancans *et al.*, 2001; Fuller *et al.*, 2001).

Outro fator envolvido ainda na regulação do pH é a enzima ATPase-vacuolar (V-ATPase), presente no interior do melanossomo, a qual atua na acidificação dos melanossomos. Em estudo realizado por Fuller e colaboradores foi relatado que a V-ATPase está em menor quantidade ou, até mesmo, ausente em melanossomos originados em pele negra (Fuller *et al.*, 2001).



### 1.4.3- Gene *SLC24A5*: rs1426654

Baseado no que se conhece do *NCKX5* (Na<sup>+</sup>/Ca<sup>2+</sup>/K<sup>+</sup> exchanger 5), a regulação do pH melanossomal pode ser o papel da proteína codificada pelo gene *SLC24A5* (Tully, 2007). O gene codifica uma proteína de 500 aminoácidos, que faz parte da família das proteínas envolvidas nas trocas de sódio/cálcio dependente de potássio, localiza-se próxima à membrana melanossomal e atua no transporte de membrana do melanossomo (Sturm, 2006; Norton *et al.*, 2007; Liu *et al.*, 2013). Lamason *et al.* (2005) descobriram que *NCKX5* está localizado em melanossomos ou seus precursores, e assumiram que ele poderia atuar acumulando Ca<sup>2+</sup> no melanossomo. O gene *SLC24A5* foi descoberto primeiramente em *zebra-fish*, no qual está relacionado ao fenótipo de uma listra clara significativa. O homólogo foi identificado e caracterizado em humanos. O gene humano mostrou que uma substituição de Ala111Thr, resultante do SNP rs1426654 (de Ala111-alelo G para Thr111-alelo A) tinha um efeito similar na pigmentação humana, isto é, conferia redução de melanina (fenótipo claro). De acordo com Ginger *et al.* (2008) o *SLC24A5* está associado com a rede trans-Golgi. Neste mesmo estudo, eles descobriram que o alelo G (Ala111) do SNP rs1426654, associado com pele mais escura, teve maior atividade de troca de íons em relação ao alelo A (Thr111), o qual está associado com pele mais clara. O alelo ancestral evolutivamente conservado (Ala111) do gene *SLC24A5* também foi encontrado em alta frequência em populações africanas e asiáticas, e que o alelo derivado (Thr111) atingiu a fixação em europeus. Especificamente, a frequência para a variante Thr111 varia entre 98,7 e 100% entre as amostras da população de origem europeia, enquanto Ala111 tem uma frequência de 93-100% em amostras de africanos, nativos americanos e população do Leste Asiático (Dimisianos *et al.*, 2008)

### 1.4.4- Gene *SLC45A2*: rs16891982

Originalmente identificado como um antígeno do melanoma humano (AIM1) e recentemente renomeado carregador de soluto família 45, membro 2, o gene *SLC45A2* contém sete éxons que se estendem a uma região de aproximadamente 40kb e codifica a proteína MATP de 530 aminoácidos que supõe-se ter 12 domínios transmembrana (Graf *et al.*, 2007). Mutações no gene *SLC45A2* causam uma forma de albinismo em humanos (OCA4) (Vierkötter *et al.*, 2012; Tully, 2007; Sturm, 2006). Graf *et al.* (2005) relataram dois SNPs em *SLC45A2* que não estavam associados a doenças, mas sim associados com a variação normal da pigmentação humana, e também a variações entre diferentes populações: Glu272Lys (rs26722) e Leu374Phe (rs16891982). A proteína MATP desempenha um papel crucial no processamento e tráfego intracelular da tirosinase (TYR), uma das enzimas cruciais e necessárias para síntese de melanina, assim como na via de outras proteínas melanossomais (Branicki *et al.*, 2009; Vierkötter *et al.*, 2012). O SNP rs16891982 gera o polimorfismo proteico Leu374Phe e regula a função de transporte da região transmembrana da proteína MATP. A variante Leu374 (alelo G) desempenha um papel importante no transporte de prótons, resultando em um pH ótimo intramelanossomal, o qual permite a atividade da tirosinase (TYR) e conseqüentemente a produção adequada da eumelanina (fenótipo pigmentado). A

variante Phe374 (alelo C) pode alterar o transporte, o pH e a síntese do pigmento (Graf *et al.*, 2005; Tully, 2007; Vierkötter *et al.*, 2012). Nakayama *et al.* (2006) referiram o alelo C (Phe374) como alelo “tipo-Caucasiano”, já que ele resultou na redução da função da MATP, por alterar o tráfego intracelular de elementos melanossomais, criando um ambiente de decréscimo na produção de melanina. Estudando o mesmo gene, Graf *et al.* (2005) observaram que o alelo Lys272 do gene *SLC45A2* mostrou um significativo aumento nas populações asiáticas e afro-americanas quando comparadas a caucasianos. Contudo, as populações aborígenes australianas mostraram frequências alélicas similares às das populações de caucasianos. Devido ao fato do fenótipo de pigmentação dos dois últimos grupos populacionais serem bem diferentes, os autores sugeriram que o polimorfismo rs26722 não estaria associado com a determinação de pigmentação, mas apenas seria uma variação associada à distribuição ou isolamento geográfico.

#### **1.4.5- Gene TYR: rs1042602**

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

#### **1.4.6- Gene TYRP1: rs2733832**

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

#### **1.4.7- Gene MC1R: rs8045560**

Até o momento, o gene *MC1R* é o mais bem caracterizado dentre os genes descritos que influenciam a variação normal de pigmentação em humanos (Gerstenblith *et al.*, 2007). Este gene apresenta um padrão

peculiar de distribuição dos seus alelos entre as populações humanas: populações africanas e outras caracterizadas pela pele escura apresentam as menores variações alélicas desse gene, enquanto em populações asiáticas e europeias, esse gene é altamente polimórfico (Makova e Norton, 2005). Estudos em diferentes populações têm demonstrado que a região codificadora do gene *MC1R* apresenta mais de 70 variantes já identificadas (Gerstenblith *et al.*, 2007), com impacto significativo no fenótipo de pigmentação (Harding *et al.*, 2000; Strum, 2006).

O gene *MC1R* transcreve para o receptor de melanocortina 1 (*MC1R*), uma proteína-G localizada na membrana de melanócitos (Valenzuela *et al.*, 2011) e diretamente responsável pela regulação da síntese de eumelanina/feomelanina (Makova *et al.*, 2005). A ligação do hormônio paracrino alfa-estimulante dos melanócitos ( $\alpha$ -MSH) ao gene *MC1R* causa uma cascata de sinalização de cAMP aumentando a produção de eumelanina (Makova *et al.*, 2005; Valenzuela *et al.*, 2011), enquanto que a ligação de seu antagonista, a proteína de sinalização agouti (ASIP), resulta em menos produção de eumelanina e no aumento da produção de feomelanina (Voisey *et al.*, 2006; Liu *et al.*, 2012).

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

O Quadro 2 apresenta um levantamento realizado com dados de indivíduos oriundos de populações Africanas e Europeias descritas nos bancos de dados ALFRED (The ALlele FREquency Database) e HAPMAP. O ALFRED (da U.S. National Science Foundation) foi projetado para tornar os dados de frequência de alelos em amostras de populações humanas prontamente disponíveis para utilização pelas comunidades; e o projeto internacional HAPMAP é resultado de uma parceria de cientistas e agências de financiamento do Canadá, China, Japão, Nigéria, Reino Unido e Estados Unidos, que tem por finalidade desenvolver e manter um recurso público para ajudar pesquisadores a encontrar genes e estes SNPs resultantes de estudos populacionais. Consultando estas duas bases de dados, compilamos resultados de todas as populações já estudadas até maio de 2013.

Neste levantamento não foram considerados dados de indivíduos oriundos de populações asiáticas devido a sua baixa representatividade na população do Rio Grande do Sul. Todos os dados foram plotados em uma planilha prévia para, finalmente, se obter o número total de indivíduos já genotipados e a frequência de cada alelo. Estes dados estão apresentados no Quadro 2.

**Quadro 2.** Genes, SNPs (rs) e Alelos associados de pigmentação em Humanos, segundo sua frequência nos conjuntos de populações Africanas e Europeias estudados nos projetos HAPMAP 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)

\* Ou alelo G, considerando a fita complementar

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

## 2- OBJETIVOS

### Objetivo Geral

Identificar e analisar os SNPs mais significativos para a previsão de pigmentação em humanos e medir suas frequências alélicas e genóticas em indivíduos com fenótipo de alto ou baixo conteúdo de melanina em pelo e olhos.

### Objetivos Específicos

- I. Selecionar, na literatura científica, SNPs significativos para a previsão de pigmentação em humanos.
- II. Identificar dois grupos de indivíduos fenotipicamente distintos: Grupo 1: indivíduos com muita pigmentação (olhos negros e pele escura); Grupo 2: indivíduos com pouca pigmentação (olhos azuis claro e pele muito clara).
- III. Genotipar os SNPs selecionados em genes de pigmentação.
- IV. Determinar as frequências alélicas e genóticas para cada SNP nos dois grupos de estudo.
- V. Comparar os grupos de acordo com a relação genótipo-fenótipo.

### **Analysis of eight SNPs in pigmentation-related genes in Southern Brazilian subjects with simultaneous high or low skin and eye melanin content.**

#### **ABSTRACT**

The understanding of gene function in the determination of externally visible characteristics (EVCs) has several uses in human population evolutionary studies and in forensic investigations. With regards to the latter, a fair amount of effort has been made to discover an easy and efficient model for the prediction of eye and skin color in humans. The obvious advantage of the prediction of such EVCs through the use of DNA is that it can be incorporated as routine procedure in forensic labs and be applied to police investigations. In the present study, we combined the genotypes of eight SNPs in pigmentation-related genes (rs4778138 - *OCA2*; rs12913832 - *HERC2*; rs16891982 - *SLC45A2*; rs8045560 - *MC1R*; rs1426654 - *SLC24A5*; rs2733832 - *TYRP1*; rs1042602 - *TYR*; rs916977 - *HERC2*) with multiple analytical approaches. Considering this SNP panel, we evaluated allele frequencies of subjects with High Melanin Content (HMC; from African populations) and Low Melanin Content (LMC; from European populations), as obtained from the HAPMAP and ALFRED databases, and defined the alleles H (to predict HMC subjects) and alleles L (to predict LMC subjects). The cumulative distribution of alleles H and alleles L in the two phenotypically different color groups comprised of a total of 134 southern Brazilian subjects showed that 82% of the HMC subjects (N = 61) had eight or more alleles H, and 100% of the LMC subjects (N = 73) had less than eight alleles H, with an accuracy value of 96.3%. We also performed analyses using the AUC (Area Under the Receiver Operating Characteristic Curve), PGL (Calculation of Pathway Genetic Load), and GP (Genetic Probability) approaches. The AUC for the prediction of both HMC and LMC phenotypes was 0.99; the PGL showed 93% concordance between genotype and HMC or LMC phenotypes across the eight SNPs panel; and the GP showed 91% concordance between predicted and observed HMC or LMC phenotypes. The combination of our high-throughput genotyping technology and multiple analytical approaches has achieved very high accuracy in the prediction of the extreme phenotypes for human pigmentation. We believe this forensic DNA phenotyping (FDP) technique can be particularly useful in instances where classic DNA profiling fails to generate a database match from crime scene DNA or where the need to classify skeletal or biological clues of dismissed people or degraded corpses arises.

## BACKGROUND

Different techniques, such as anthropometric examinations, fingerprinting and genetic studies, have been largely implemented in order to characterize and individualize people. For nearly twenty years, autosomal sets of short tandem repeat markers (STRs) and, more recently, single nucleotide polymorphisms (SNPs) have been used for human identification in forensic cases [1]. SNPs, which represent the most abundant class of human polymorphisms, may determine amino acid substitutions altering the functional properties of proteins and expressing distinct phenotypes [2]. SNP abundance has sparked the interest in forensics and the mapping of the human genome has made possible the development of sets of genotypes and haplotypes that can phenotypically characterize a human being [3]. Thus, SNP studies may serve for individual discrimination as well as for the prediction of externally visible characteristics (EVCs). Among the EVCs, the pigmentation of skin, eyes, and hair has become a good target for the prediction of phenotypes. Genotype-phenotype association studies regarding iris and skin color have been focusing on detecting SNPs within genes directly involved in the synthesis of pigments [4-16]. There are more than one hundred genes involved in the pigmentation process [17, 18] and they have a cumulative effect. However, for forensic purposes, despite the small number of known genes determining human pigmentary traits, it is possible to predict with high confidence the color of a person's eye, hair, and/or skin based on this small number of DNA polymorphisms [19]. Thus, it is fair to say that some genes have alleles which can definitely determine the extremes of melanin pigmentation patterns, i.e. , light and dark patterns.

In this work, we selected eight SNPs in seven genes that are linked to the melanin synthesis in order to investigate their significance in skin and eye color phenotypes for forensic purposes. SNPs were selected in the following human pigmentation associated genes: *HERC2*, *OCA2*, *SLC24A5*, *SLC45A2*, *TYR*, *TYRP1*, and *MC1R*; SNPs in the first three were shown to significantly alter the melanin content in human melanocytes, supporting their functional role in pigmentation [20].

***OCA2*, rs4778138:** The *OCA2* gene 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 [21-25]; changes on *OCA2* may interfere with melanin synthesis due to pH changes. The rs4778138 is a G>A SNP in the intron region of *OCA2* and it is strongly associated with eye color [19].

***HERC2*, rs12913832:** The Hect Domain and RCC1-like Domain 2 gene is located 10 Kb upstream of *OCA2* and it acts as a regulatory enhancer region of *OCA2*. The function of *HERC2* is still unknown. The rs12913832 is an A>G SNP in an intronic area of the *HERC2*, within the *OCA2* enhancer region. The derived allele G of this SNP is associated with the blue iris phenotype, being common in Europeans , particularly those of northwestern and eastern European descent [25]. The allele A (or allele T, if in the complementary DNA strand) allows for the opening of the chromatin and

recruitment of the *OCA2* transcription factors, which leads to a darker iris, while the allele G (or allele C, if in the complementary DNA strand) maintains the chromatin closed, being less effective in recruiting the aforementioned factors, which results in a lighter iris [20, 23].

***SLC45A2, rs16891982***: The solute carrier family 45 member 2 gene encodes the MATP protein [26], which plays a crucial role in the processing and intracellular trafficking of tyrosinase (TYR) and of other enzymes necessary for the synthesis of melanin [18, 28]. The rs16891982 is the 1122C>G SNP in exon 5 of the *SLC45A2* 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). The rs16891982 is associated with the normal variation of human skin pigmentation, since it regulates the MATP function; the Leu374 variant (allele C – complementary to G on codon TTG 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 [28-30]. A strong association between rs16891982 and eye color has been reported [19, 27].

***MC1R, rs8045560***: *MC1R* gene transcribes the melanocortin 1 receptor (alpha melanocyte stimulating hormone receptor), a G-protein located in the melanocyte membrane [24] which is directly responsible for the regulation of eumelanin / pheomelanin synthesis [19, 31, 32]. The *MC1R* gene is the best characterized among the genes that influence the normal variation of human skin pigmentation, and the ancestral allele C of rs8045560 (C>T SNP) is associated with darker skin populations [17, 31, 33].

***SLC24A5, rs1426654***: The solute carrier family 24, member 5, is responsible for the accumulating  $Ca^{+2}$  in the melanosome [34, 35]. The rs1426654 (G>A SNP in the coding region of *SLC24A5* 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 [25]. The ancestral allele (G allele; Ala111) is associated with darker skin [20, 36].

***TYRP1, rs2733832***: Located at 9p23, this gene encoding to tyrosinase-related protein 1, which is part of the enzyme tyrosinase in melanin production [4, 17, 19]. *TYRP1* polymorphisms are associated with coloring eyes in Europeans [37], and the presence of TYRP1 protein has been found to be 2.6 times higher in the skin of Africans and Indians (darker pigmentation) than in the skin of Mexicans, Chinese, and Europeans [38]. The ancestral allele C of rs2733832 (C>T SNP) is associated with darker skin populations.

***TYR, rs1042602***: *TYR* gene transcribes the tyrosinase, an important enzyme for the initial process of melanogenesis [19]. Mutations in the *TYR* gene are associated with vitiligo susceptibility [39] and oculocutaneous albinism type I [40], but polymorphic variations in the *TYR* are directly associated with



changes in the color of eye, hair, and skin [4, 41]. The rs1042602 (C>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 [42]. The ancestral allele C is associated with darker skin populations.

**HERC2, rs916977:** The derived allele G of rs916977 (A>G SNP) has been associated with blue eyes [4, 43].

In this paper, we have processed different analyses with these SNP variants and genotyped subjects from two southern Brazilian populations with either high or low melanin content using a multiplex system. We used this panel to predict color phenotype for forensic uses.

## MATERIAL AND METHODS

*Design, subjects, and approval.* This observational cohort study was conducted with data and DNA collected from adult subjects (over 18 years old) enrolled between March 1<sup>st</sup>, 2012 and October 31<sup>st</sup>, 2013. Subjects were selected based on their skin/eye phenotype characteristics. All subjects were from southern Brazil, an area that presents a distinctive genetic background: the majority of subjects were of European origin (Portuguese, Italian, Spanish, and German ancestry) and a smaller amount of individuals were of African origin [44]. A form containing information about gender, age, origin, place of birth and residence was filled out by all individuals. 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 for the study was obtained from all subjects.

*Skin Phenotyping.* The skin color of each participant was identified using the Fitzpatrick score, ranging from Type 1 to Type 6, where: 1- Highly sensitive skin, always burns, never tans; 2- Very sun-sensitive skin, burns easily, tans minimally; 3- Sun-sensitive skin, sometimes burns, slowly tans to light brown; 4- Minimally sun-sensitive skin, burns minimally, always tans to moderate brown; 5- Sun-insensitive skin, rarely burns, tans well; 6- Sun-insensitive skin, never burns, deeply pigmented. We also measured the skin amount of red (R), green (G), and blue (B) pigments in an inner and hairless portion (below elbow) of the right arm using the ACR-1023 (Instrutherm, São Paulo, Brazil), a battery-portable color analyzer equipment which employs a spectral analysis method to determine the color of the sample. Each R, G, and B values range from zero to 1023, where the minimum value (zero) represents the complete absence of color, and the maximum (1023), its complete presence. In this system, total white has a value of R=1023, G=1023, and B=1023, and total black has a value of R=0, G=0, and B=0. The RGB skin values of each person were measured three times and the average was recorded. The equipment was calibrated before each use by measuring a white plate. Potential subjects that reported having a recent or intense tan were excluded.

*Eye Phenotyping.* The eye color of each participant was classified from Type 1 to Type 6, where: 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 5500K (D55 illuminant), and a picture resolution of 4,000 X 3,000 pixels. To take the pictures, the camera was fixed on a tripod stand and

placed at a distance of  $6\text{cm} \pm 0.5\text{cm}$  from the subject's eye, having its focus locked onto the central area of the iris. The images were taken in a room with standardized artificial light conditions. The participants wearing contact lenses were required to remove them. We measured both eye amounts of R, G, and B pigments by analyzing each photo with the software COLOURS [45]. A triangular-shaped representative area of the iris covering approximately 2.500 pixels was selected in each picture and the R, G, and B values of each pixel were measured on a scale of zero to 255, where the minimum value (zero) represents the complete absence of a given color and the maximum (255) represents its complete presence. To calculate the R value, the software divides the total amount of red pigment present in each pixel of the selected area by the product of 255 times the total amount of pixels in the same area. For example (Figure 1, picture on the right):  $264002 / (255 \times 2542) = 0.41$  – as the total amount of red in the selected area is 264002 and the total amount of pixels in the same area is 2542. The same mathematical formula is used to G and B calculations. In this system, total black has a value of  $R=0$ ,  $G=0$ , and  $B=0$ , and total white has a value of  $R=1$ ,  $G=1$ , and  $B=1$ . The RGB eye values were measured on both eyes and the average was recorded. The primary RGB values were then converted into HSV (hue, saturation, value of brightness) system values using standard formulas. HSV is a cylindrical-coordinate representation of points in an RGB color model. This representation maps the values into a cylinder (color wheel), where: the angle around the central vertical axis corresponds to hue (H) and the distance from the axis corresponds to saturation (S). The height corresponds to a third value (V), the system's representation of the perceived luminance (brightness) in relation to the saturation. Potential subjects with *Heterochromia iridum* (excessively different colored irises) were excluded from the experiment.

*Genotyping.* Genomic DNA was extracted with a NucleoSpin® Blood kit (Macherey-Nagel Inc.). Amplification of regions flanking the SNPs were made by multiplex PCR in a volume of 25  $\mu\text{L}$  with 2 to 10 ng genomic DNA, 0.2  $\mu\text{M}$  of each primer (Table 1), and 1X of Qiagen Multiplex PCR Master Mix (© QIAGEN, Invitrogen). The amplification consisted of 95°C for 5 min, followed by 30 cycles of 30s at 94°C, 90s at 57°C, 90s at 72°C, and a final extension for 10 min at 72°C. After the amplification, the enzymatic purification of the PCR product using USB® ExoSAP-IT® PCR Product Cleanup (Affymetrix®) was carried out according to the user's manual recommendations. SNP analysis was performed by the SNaPshot® Multiplex System ABI Prism (Applied Biosystems®, São Paulo, Brazil), using eight SNaPshot primers (Table 1). Reactions were performed in a final volume of 10  $\mu\text{L}$ , containing 3.0  $\mu\text{L}$  of purified multiplex PCR product, 5.0  $\mu\text{L}$  of SNaPshot Multiplex Ready Reaction Mix, 1.0  $\mu\text{L}$  of pooled SNaPshot primers (each primer at 1  $\mu\text{M}$  final), and 1.0  $\mu\text{L}$  of sterile water. Multiplex single base extensions were carried out for 28 cycles according to the following program: 10 seconds at 96°C, 5 seconds at 50°C, and 30 seconds at 60°C. SNaPshot products were then treated at 37°C for 1 hour with 0.8  $\mu\text{L}$  of shrimp alkaline phosphatase (1.0 U/ $\mu\text{l}$ ) and 1.2  $\mu\text{L}$  of 10X SAP buffer reaction, added directly to 10  $\mu\text{L}$  of the SNaPshot product. After heat inactivation of shrimp alkaline phosphatase for 15 minutes at 75°C, 1 $\mu\text{L}$  of the labeled products were mixed with 9.5  $\mu\text{L}$  of HiDi formamide and 0.5  $\mu\text{L}$  of Genescan-120 LIZ size standard. They were then separated using an ABIPRISM® 3130xl Genetic Analyzer (Applied Biosystems) with POP-4 matrix and with respective run parameters: injection voltage of

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

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

*Area under the Receiver Operating Characteristic curve:* 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 0 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* [46]. Individual PGL scores potentially ranged from zero (homozygous for alleles L at all loci) to 16 (homozygous for alleles H at all loci). PGL scores weighted for color were created by multiplying the allele H count at each locus by its adjusted odds ratio (OR) – as determined from previously published HAPMAP and ALFRED frequencies. Weighted PGL scores for each locus were then computed to create a weighted PGL score for each individual. PGL and other continuous data were compared by the Mann-Whitney U test. Categorical data was compared using chi-square. The BioEstat 5.0 Statistical Package was used for all calculations.

*Cluster analysis using STRUCTURE Software:* We used the eight SNPs and the STRUCTURE software to estimate clusters among subjects. The STRUCTURE Software v2.3.4 (<http://pritch.bsd.uchicago.edu/structure.html>; 47) analyses the data plotting probabilistically the individuals in  $K$  populations by characterization of the allele frequencies set in each locus. The individuals may be assigned in one, two or more populations if their genotype shows an admixed pattern. Our data file was composed by: row of makers (SNP and gene data), column of labels of each individual, column of population data (1 for HMC; 2 for LMC), and the genotype data. The ADMIXTURE MODEL and CORRELATED ALLELE FREQUENCIES MODEL were chosen because the admixture between populations is a common characteristic of real genetic data, since subjects may have recent ancestors in more than one population [48]. To estimate the  $K$  value, the data was analyzed with twenty replicates for  $K=1$  to  $K=10$ , all runs were performed with 10,000 burn-in period and 100,000 Markov Chain Monte Carlo (MCMC) repeats after burn-in, ADMIXTURE MODEL, CORRELATED ALLELE. The  $K$  values were estimated using the  $\ln\Pr(X|K)$  values for each simulation. With these values, it is possible to calculate the probability for each  $K$  and estimate the best  $K$  [47]. The method used to estimate the value of  $K$  calculates the delta  $K$  and then selects the appropriate  $K$  value [49, 50]. Graphics were constructed using *Clumpp* [51] and *Distruct* [52] software.

## RESULTS

Table 2 presents the allele frequencies of the selected SNPs – based on population data available on the HAPMAP and ALFRED websites – obtained from African and European subjects (the database used to construct Table 2 is shown in a Supplementary Excel file). An association Chi-square test was performed to examine the relationship between allele frequencies in the aforementioned populations (Africans and Europeans). There was a highly significant association between one of the alleles with Africans and the other with Europeans for all the SNPs tested. We dubbed ‘allele H’ each allele strongly associated with subjects with High Melanin Content (HMC; from African populations), and ‘allele L’ each allele strongly associated with subjects with Low Melanin Content (LMC; from European populations). Considering the relationship between alleles and populations, we calculated (as explained in *Material and Methods*) the sensibility, specificity, predictive values, relative ratio, and odds ratio for each SNP present in table 1. The eight SNPs in pigmentation-related genes presented remarkable results in all of the parameters and they were considered robust markers for color prediction (Table 2).

We analyzed 134 southern Brazilian subjects belonging to two color categories: HMC – individuals with high melanin content (N = 61) – and LMC – individuals with low melanin content (N = 73). The phenotype data of these subjects is presented in Table 3. As expected by the previous selective sampling, the clustering toward the color was evident between these two groups: the HMC group presented only people with Types 5 and 6 of skin and eye classification, and with RGB values closer to zero (black); and the LMC group presented only people with Types 1 and 2 of skin and eye classification, and with RGB values closer to the highest values (white) (Table 3). However, the ancestral origin reported by the subjects was not clearly discriminated into the groups, especially in the HMC group, where 21% of the people assumed

they had European origin. We genotyped these 134 southern Brazilian subjects using the SNaPshot® Multiplex System with a set of eight primer pairs to analyze the eight target SNPs. Table 4 shows a highly significant association between 'alleles H' and HMC people, and 'alleles L' and LMC people. The other parameters (sensitivity, specificity, predictive values, relative ratio, odds ratio) were equally remarkable. Although not dealing with such a clear ethnic background in our sample, the allele frequencies evaluated confirmed the eight SNPs as robust markers for color prediction.

To observe the cumulative distribution of alleles H and alleles L in our two phenotypically different color groups, we plotted the frequency data from southern Brazilian subjects considering the total number of each allele H per person in the two groups, comprised of 61 HMC subjects and 73 LMC subjects, respectively. Figure 2 shows that LMC subjects had between zero and seven alleles H, while HMC subjects had between four and 14 alleles H. In our sample, no subject had 15 or 16 alleles H. The average of alleles H among the HMC subjects was 12.0 (SD=3.78), and 82% (56/61) of the HMC individuals had eight or more alleles H. The average of alleles H among the LMC subjects was 3.3 (SD=1.49), and 96% (70/73) of the LMC individuals had up to five alleles H, and 100% had seven or less alleles H. Some results are especially interesting because they present a divergence between the quantity of alleles H and the phenotype category: one HMC person had only four alleles H, one HMC person had five alleles H, two HMC people had six alleles H and one HMC person had seven alleles H. The individual allele analyses of these five subjects show that all loci were in some way involved in homo or heterozygosis, i.e., there was no evidence of any locus to explain the high skin/eye melanin content even with such few alleles H. We also analyzed one LMC person (HC120) with seven alleles H: this person was homozygote HH to rs8045560 (*MC1R*) and heterozygote to rs4778138 (*OCA2*), rs12913832 (*HERC2*), rs2733832 (*TYRP1*), rs1042602 (*TYR*), and rs916977 (*HERC2*). The analysis of these people could be important to identify which loci are more critical to color prediction – however, a conclusion about the matter was not possible with the present data because of the size and heterogeneity of the sample.

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). Thus, we changed the decision threshold from zero to 16 alleles H and plotted the "sensitivity" and "1 - specificity" values as "x" and "y" coordinate values of points 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" [53]. Based on the same assessment that Liu *et al* (2009) 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) [19]. Considering the cumulative distribution of alleles H (the total number of alleles H per person) in the two phenotypically

different color groups we obtained the very high value of AUC at 0.9908 in the prediction of both HMC and LMC phenotypes (SE= 0.0088; IC 95% = 0.99736-1.0).

For another cumulative analysis, we used the Pathway Genetic Load (PGL), which allows for a simultaneous evaluation of multiple loci with cumulative effect on the phenotype [46]. Individual unweighted PGL scores ranged from zero (homozygous for allele L at all eight *loci*) to 14, as previously shown in Figure 2. PGL weighted scores (PGL-WSs) 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 HAPMAP and ALFRED data (presented in Table 2). For example: to the following loci rs4778138 (*OCA*), rs12913832 (*HERC2*), rs16891982 (*SLC452A*), rs8045560 (*MC1R*), rs1426654 (*SLC24A5*), rs2733832 (*TYRP1*), rs1042602 (*TYR*), and rs916977 (*HERC2*), the complete genotype of subject HC123 was, respectively: LL – LL – LL – HL – LL – HL – HH – LL. Using respective OR values by locus, we obtained to HC123 subject this values: PGL-WS= (0.0 x 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) + (0.0 x 29.73) = 43.59. The PGL analysis in all of the 134 southern Brazilian subjects showed that the individual PGL-WSs in HMC population ranged from 65.29 to 870.92, with the median in 567.37 (mean= 584.63; SD=262.46); and the individual PGL-WSs in LMC population ranged from zero to 148.40, with the median in 38.31 (mean= 44.80; SD=30.93). The PGL-WS for the eight SNPs tested were significantly different between HMC and LMC groups (P<0.0001; Mann-Whitney Test), 95% (58/61) of the HCM subjects had the PGL-WSs higher than the maximum value among LMC subjects (148.40), and 88% (67/73) of the LCM subjects had the PGL-WSs lower than the minimum value among HMC subjects ( 65.29). Considering these two cut points (more than 148.40 to HMC, and less than 65.29 to LMC), the PGL analysis showed that the current eight SNPs genotyping system had 93% (125/134) concordance between genotype and HMC or LMC phenotype.

Based on the HAPMAP and ALFRED allele frequencies, we calculated each expected genotype frequency to HH, HL, and LL genotypes, by locus, in HMC and LMC groups (Table 5). Then, we examined the complete genotype of each South Brazilian subject and calculated the probability of the complete genotype belonging to expected HMC and LMC groups. For example: to the following loci rs4778138 (*OCA*), rs12913832 (*HERC2*), rs16891982 (*SLC452A*), rs8045560 (*MC1R*), rs1426654 (*SLC24A5*), rs2733832 (*TYRP1*), rs1042602 (*TYR*), and rs916977 (*HERC2*) the complete genotype of subject HC123 was, respectively: LL – LL – LL – HL – LL – HL – HH – LL. Using each genotype frequency by locus obtained from HMC population, we had the following values for subject HC123: 0.06554 - 0.00185 - 0.01416 - 0.15555 - 0.0756 - 0.15222 - 0.89492 - 0.01210 (see Table 5). The product of these values was  $3.33 \times 10^{-11}$ , which means the probability of this complete genotype is present in the HMC population. Then, using each genotype frequency by locus obtained from LMC population, we had the following values for subject HC123: 0.69890 - 0.50694 - 0.87049 - 0.49280 - 0.97812 - 0.48555 - 0.49280 - 0.61780 (see Table 5). The product of these last values was: 0.022, which means the probability of this complete genotype is present in the LMC population. In conclusion, comparing the two probabilities, the complete genotype of the HC123 subject has  $6.6 \times 10^8$  times higher chance of belonging to the LMC population than to the HML population. This finding was in perfect

agreement with the actual HC123 phenotype data obtained: a male with light skin color (Type 2), dark blue eyes (Type 2), high R-G-B values to skin (317-239-200) and eyes (0.385-0.365-0.530), and of European origin.

We applied this model to predict color in the HMC or LMC phenotype categories based on the genotype probability (GP) in all 134 South Brazilian subjects; we observed that 80% (49/61) of the genotypes of our HMC subjects had a higher chance of belonging to the HMC group than to the LMC group, with rates from  $1.95$  to  $8.27 \times 10^7$  (mean=  $7.92 \times 10^6$ ). Analyzing the genotypes of our LMC subjects, we noticed that 100% (73/73) had a higher chance of belonging to the LMC group, with rates from  $1.46 \times 10^4$  (in subject HC120, with seven alleles H) to  $7.77 \times 10^{12}$  (mean=  $2.02 \times 10^{11}$ ). This GP analysis using the current eight SNPs genotyping system showed 91% (122/134) concordance between predicted and observed HMC or LMC phenotype.

As a complementary approach, we used the eight SNPs and the STRUCTURE software to estimate clusters among subjects. The program assigned to each subject a value, which represented the probability of belonging to each of two phenotypic clusters. This value clustered the subjects into HMC or LMC groups. Figure 3 shows the dot plot of probability values in each phenotypic cluster; the set of SNPs separated the individuals into the two extreme HMC or LMC groups indicating a substructure across the total population.

## DISCUSSION

Forensic DNA phenotyping (FDP, a relatively young field within forensic genetics, aims to predict appearance traits of a sample donor from DNA left behind at a crime scene; an efficient FDP technology would be especially useful in cases where the Police have no clues (including conventional DNA profiling) as to who the crime scene sample donor might be [5]. As with other applications of DNA data in the forensic arena, FDP has to be backed up by the legislation as, for instance, is currently the case of forensic DNA-based eye color prediction in the Netherlands [19]. Skin and eye color are important externally visible traits (EVTs) in forensics, since they are highly heritable genetic traits and are the most obvious and distinguishing externally visible characteristics to be used in human identification.

Melanin is gradually produced in melanosomes inside melanocytes. Although the number of melanocytes in different skin colors is essentially constant, the distribution of melanosomes is not so, as darker skins have denser and individually dispersed melanosomes, while lighter skins have less dense melanosomes [54]. In the eyes, pigments are distributed in the epithelium and in the stromal cells of the iris; brown iris has an abundance of melanin while blue iris contains low levels of melanin allowing light transmittance and reflectance which results in a blue shortwave due to the presence of collagen fibers. Variations in both skin and eye color also result from the ratio of eumelanin/pheomelanin [55] and are strictly connected to human evolution as well. The presence of melanin is an important protection against ultraviolet radiation (UV): darker skin provides greater protection in environments with high UV radiation (low latitudes), while light skin facilitates the biosynthesis of vitamin D in regions where UV radiation is low

(high latitudes) [56]. Since pigmentation and ethnicity are correlated, it is important to understand if the presence of a particular allele in populations with high or low melanin content could be directly used in determining phenotypic pigmentation or if it is only a marker of ancestry.

It is still unclear whether a gene is expressed only in the skin or only in the iris, and for this reason we have selected subjects with HMC or LMC in both skin and eye phenotypes. Our results for the association study, sensibility, specificity or accuracy parameters, or the final analysis using different interpretation models would still have been the same even if only the eye color data had been considered. We have, nonetheless, preferred to work, in a first moment, with subjects with whole (skin and eye) HMC or LMC phenotypes, even though there are many Brazilian subjects with intermediary or ambiguous phenotypes, such as dark eye / light skin or light eye / dark skin. Further studies with the last subjects could select genes with preferential expression to skin or eye. In the same way, further studies with subjects with intermediate pigmentation phenotype should be mandatorily conducted. Although genetic markers have been informative to predict phenotypes in human pigmentation, this field of study still requires many efforts to be really applicable. Thus, we chose to design a panel that could at least detect the extreme phenotypes (subjects with high or low melanin content).

We have not analyzed the hair color because the melanin content in the hair is more strongly changed by intrinsic or extrinsic factors; hair color is also age-dependent, changing during late childhood, adolescence, adulthood, and old age, becoming darker, gray or white at any given stage [57]. These changes, added to the occurrence of baldness, make hair a not very useful trait to be useful for practical forensic DNA phenotyping applications.

Human skin and eye pigmentations are continuous traits and there is not a worldwide model to determine the limits of each mate category. However, in FDP studies, the use of categorized information about these continuous traits is inevitable. Thus, more accurate and more detailed pigmentation determinations may guarantee better conclusions in these FDP studies [16, 19]. To avoid any mistakes in the true color categorizations and to concentrate on high accuracy and high level of detail in color phenotype description, we have only worked with representatives of the two extremes of the continuous skin and eye color distribution. Considering only two extremely stratified groups of high (dark) and low (light) melanin content phenotypes and using reproducible methods to quantify the skin-eye pigmentation, we have obtained well-characterized information on both skin (using a surface *color analyzer*) and eye color (using digital photographs and a specially designed color quantifier software).

During the length of this study, important advances in FDP technology considering skin, eye, and/or hair color prediction were made. In 2009, Liu *et al.* constructed a model-building set using six SNPs from six genes (*HERC2* rs12913832, *OCA2* rs1800407, *SLC24A4* rs12896399, *SLC45A2* rs16891982, *TYR* rs1393350, and *IRF4* rs12203592) and tested it in a group of more than two thousand people from Rotterdam. They performed a multinomial logistic analysis for the interpretation of the six SNPs from six genes that revealed excellent accuracy for the prediction of eye color, with an overall area under the receiver characteristic operating curve (AUC) of 0.93 for brown, 0.91 for blue, and 0.72 for intermediate colored eyes. Nine



additional SNPs (from *TYRP1*, *OCA2*, *HERC2*, and *ASIP*) were also tested, but had only minimal additive effect on their sample. The remaining nine SNPs had no additive value to the predictive accuracy, though they were all significantly associated with eye color in the single-SNP analysis. The prediction accuracies presented by Liu *et al.* have improved considerably in comparison to their previous attempt, using three SNPs in *OCA2* and *HERC2* (e.g. AUC = 0.82 for brown eyes), as presented by Kayser *et al.* [54], or in comparison to another prediction analysis based on four SNPs in *OCA2*, *HERC2*, *SLC24A4*, and *TYR*, carried out applying a different methodology [4].

Mengel-From *et al.* [8] investigated in 395 unrelated, healthy Danish donors, the predictive values of various combinations of SNP alleles in the *HERC2*, *OCA2*, and *SLC45A2* (*MATP*) genes and compared the results to the eye colors as described by the individuals themselves. The highest predictive value of typing either the *HERC2* SNPs rs1129038 and/or rs12913832 (which are in strong linkage disequilibrium) was observed when eye color was split into two groups: light (blue, grey, and green) versus dark (brown and hazel). Sequence variations in rs11636232 and rs7170852 in *HERC2*, rs1800407 in *OCA2* and rs16891982 in *SLC45A2* showed further association with eye color in addition to the effect of *HERC2* rs1129038. In this work, the diplotype analysis of three sequence variations in *HERC2* and one sequence variation in *OCA2* showed the best discrimination between light and dark eye colors, with a 29.3 likelihood ratio.

In 2011, Walsh *et al.* [9, 10] presented a DNA tool, termed IrisPlex, for the accurate prediction of eye color performing a single multiplex assay based on SNaPshot chemistry and capillary electrophoresis. The researchers used a multiplex genotyping system based on the six eye-color informative SNPs: rs12913832, rs1800407, rs12896399, rs16891982, rs1393350 and rs12203592 from the *HERC2*, *OCA2*, *SLC24A4*, *SLC45A2* (*MATP*), *TYR* and *IRF4* genes, respectively. The IrisPlex eye color prediction interpretation was based on a multinomial, logistic regression model with three outcomes (blue, intermediate and brown). Their data revealed prevalence-adjusted prediction accuracies of over 90% for both blue and brown eye color in more than six thousand Dutch Europeans. They also presented a prediction model to correctly classify an individual's eye color via probability estimation based exclusively on DNA data and demonstrated the accuracy of the developed prediction test on 40 individuals from various geographic origins.

In the same year, Spichenok *et al.* [11] performed a TaqMan PCR assay studying and validating seven SNPs (7-plex-system), all located in and nearby genes known for their important role in pigmentation, on 554 samples donated from non-related individuals of various populations. Six SNPs were used to predict the eye color, and all seven were used to describe the skin color of the subjects. The outcome revealed that the markers rs12913832, rs1545397, rs16891982, rs1426654, rs885479, rs6119471, and rs12203592 can be applied to all populations to predict skin and eye colors of individuals, independently of their ethnic origin, with very low error-rates, even though the skin coloration varied between populations and demonstrated some level of complexity. In the following year, Pneuman *et al.* (2012) validated the eye and skin color predictors presented by Spichenok *et al.* on new sample sets and confirmed that the predictors can be applied to various populations, including African-American, South Asian (dark), East Asian (light), European,

and mixed populations. Prediction was made using classification trees for all samples and the average error-rate was calculated at approximately 3%, which corroborated the original publication of 7-plex-system [11]. The error-rates varied – they were lower for the predictors that describe the eye and skin color exclusively (non-brown or non-blue and non-white or non-dark, respectively) and higher for the predictor that describes individual eye colors (blue, brown, and intermediate/green) – due to uncertainties with the green eye color prediction.

More recently, Ruiz *et al.* [14] typed 37 SNPs in pigmentation-associated genes in order to investigate the genotype prediction of eye, skin, and hair color variation. The developed assays were subsequently used to evaluate the performance of eye color predictors in 416 subjects from six populations of north and south Europe through AUC analysis and a modified Bayesian likelihood-based classification approach (*Snipper* classifier). Using the online classifier provided two advantages to the researchers, namely the ability to submit incomplete SNP profiles (a common occurrence when typing challenging DNA) as well as handle physically linked SNPs showing independent effect (by allowing the user to input frequencies from SNP pairs). The study indicated that the SNP pair rs12913832 and rs1129038 encompasses the two SNPs most closely associated to eye color. The *Snipper* classifier was used to include the submission of data for these two SNPs.

Virtually at the same time, Allwood *et al.* [15] developed another model for prediction of eye color performing an assay with 19 SNPs studied in 101 unrelated individuals from the New Zealand population with different ancestral backgrounds. They constructed classification tree models with SNPs of known association with eye color phenotypes in three categories: blue versus non-blue, brown versus non-brown and intermediate versus non-intermediate. Four alternate models were capable of predicting eye color from the DNA genotypes of SNPs located within the *HERC2*, *OCA2*, *TYR*, and *SLC24A4* genes by using probability calculation and the classification trees. The final model exhibited high levels of accuracy for both blue (89%) and brown eye color (94%). The models using probability calculation and the classification trees were also assessed with a test set of 25 blind samples where phenotype was unknown, with blue and brown eye color predicted correctly. The authors avowed that classification trees offer an aesthetically simple and comprehensible model to predict blue and brown eye color.

Recently, Andersen *et al.* [16] focused their study on a new method for measuring the eye color phenotype, developing a continuous well-classified scale of different shades of eye color. They used the custom designed software Digital Iris Analysis Tool (DIAT) for automatic iris identification and creation of high resolution digital images. The correlations between the PIE-scores and the six IrisPlex SNPs (*HERC2* rs12913832, *OCA2* rs1800407, *SLC24A4* rs12896399, *TYR* rs1393350, *SLC45A2* rs16891982 and *IRF4* rs12203592) were analyzed in 570 individuals. Significant differences in the PIE-scores were observed in individuals typed as *HERC2* rs12913832 G (PIE=0.99) and rs12913832 GA (PIE=-0.71) or A (PIE=-0.87). After an adjustment for the effect of *HERC2* rs12913832, the quantitative PIE-scores were significantly associated with SNPs with minor effects (*OCA2* rs1800407, *SLC24A4* rs12896399, and *TYR* rs1393350) on the eye color. They found good concordance with the phenotype prediction from individuals typed as

*HERC2* rs12913832 when the multinomial, logistic regression IrisPlex [10] and the modified Bayesian likelihood-based *Snipper* classifier models were used [14].

In our study, we have combined the genotyping of eight SNPs (rs4778138 - *OCA2*; rs12913832 - *HERC2*; rs16891982 - *SLC45A2*; rs8045560 - *MC1R*; rs1426654 - *SLC24A5*; rs2733832 - *TYRP1*; rs1042602 - *TYR*; rs916977 - *HERC2*) with classic and new analytical approaches. Considering this SNP panel, we evaluated allele frequencies – as obtained from the HAPMAP and ALFRED databases – of populations with subjects with High Melanin Content (HMC; from African populations), and Low Melanin Content (LMC; from European populations) and defined the alleles H (to predict HMC subjects) and alleles L (to predict LMC subjects). The cumulative distribution of alleles H and alleles L in the two phenotypically different color groups comprised of a total of 134 Southern Brazilian subjects showed that 82% of the HMC subjects (N = 61) had eight or more alleles H, and 100% of the LMC subjects (N = 73) had less than eight alleles H, with an accuracy value of 96.3%. Considering the cumulative distribution of alleles H in the two phenotypically different color groups, we obtained the very high AUC value of 0.9908 for the prediction of both HMC and LMC phenotypes. PGL analysis, which allows simultaneous evaluation of multiple loci with cumulative effect on the phenotype, showed 93% concordance between genotype and HMC or LMC phenotypes across the eight SNPs panel. The last interpretation analysis, based on genotype probability (GP), showed 91% concordance between predicted and observed HMC or LMC phenotypes. As a complementary approach, the individuals were analyzed using STRUCTURE; the results confirmed our previous findings, showing two well-separated clusters.

In our interpretation analysis, we considered the cumulative effect of each allele on the whole set of SNPs. Even though the relationship between the two alleles in each SNP is unclear, markers in heterozygosis were considered co-dominant and were included in the analysis. We did not assume complete dominance with respect to any genes, based on previous assays with cultured human melanocytes where rs16891982 - *SLC45A2*, rs1426654 - *SLC24A5*, and rs4778138 - *OCA2* heterozygotes presented intermediate expression [20].

Since the epistasis, pleiotropic, expressivity, or penetrance effects among the genes are still uncertain, we also considered equal contribution for each allele, independently of SNP or gene. In previous studies, *HERC2* SNPs were recognized to have the most important effect on dark phenotypes. Subsequently, other approaches used dichotomy classification trees starting in the presence or absence of informative alleles in *HERC2* SNPs [12, 15] before continuing the evaluation with other SNPs. In our sample, 48.4% (15/61) and 50.8% (31/61) of HMC subjects had alleles L to rs12913832 and rs916977 *HERC2* SNPs, respectively. This could be explained by the different genetic background of our population. According to the ALFRED and HAPMAP databases, the global frequency of alleles L in HMC populations is 0.043 and 0.110 to rs12913832 and rs916977 *HERC2* SNPs, respectively. The use of dichotomy classification tree models considering first the rs12913832 and rs916977 *HERC2* SNPs in our population would cause a deviation from the better correct result.

Finally, we investigated a set of eight SNPs in previously confirmed pigmentation-related genes using a SNaPshot multiplex panel. Such high-throughput genotyping technology combined with multiple analytical approaches reached very high accuracy in the prediction of the extreme phenotypes for human pigmentation. We believe this forensic DNA phenotyping (FDP) technology could be particularly useful in instances where classic DNA profiling fails to generate a database match from crime scene DNA or where the need to classify skeletal or biological clues of dismissed people or degraded corpses arises.

## REFERENCES

1. Kobachuk LDG (2012) Estudo de frequências alélicas de dez locos STRs do cromossomo X na população do estado do Paraná e sua contribuição na identificação humana. Dissertation, Federal University of Paraná
2. O’Rahilly, S (2009) Human genetics illuminates the path to metabolic disease. *Nature* 462:307-314
3. Stacey GB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, Defelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–2229
4. Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, Jakobsdottir M, Steinberg S, Pálsson S, Jonasson F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediksdottir KR, Aben KK, Kiemenev LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K (2007) Genetic determinants of hair, eye and skin pigmentation in Europeans, *Nat Genet* 39:443–1452
5. Kayser M, Schneider P (2009) DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. *Forensic Science International* 3:154–161
6. Sturm RA (2009) Molecular genetics of human pigmentation diversity. *Hum Mol Genet Review* Issue 18:R9-R17
7. Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens ACJW, Kayser M (2009) Eye color and the prediction of complex phenotypes from genotypes. *Curr. Biol.* 19:R192–R193
8. Mengel-From J, Børsting C, Sanchez JJ, Eiberg H, Morling N (2010) Human eye colour and *HERC2*, *OCA2* and *MATP*. *Forensic Sci Int Genet* 4(5):323-328.
9. Walsh S, Lidenbrgh A, Zuniga SB, Sijen T, Knijff P, Kayser M, Ballantyne KN (2011a) Developmental validation of the IrisPlex system: Determination of blue and brown iris colour for forensic intelligence. *Forensic Sci Int Genet* 5 (5): 464-471.
10. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M (2011b) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 5(3):170-180.

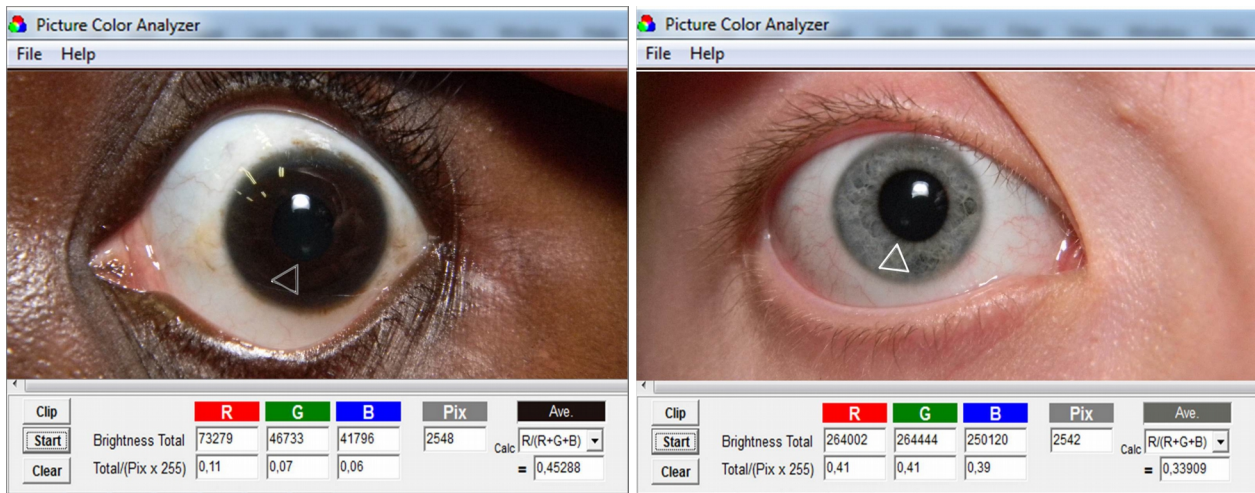
11. Spichenok O, Budimlija ZM, Mitchell AA, Jenny A, Kovacevic L, Marjanovic D, Caragine T, Prinz M, Wurmbach E. (2011) Prediction of eye and skin color in diverse populations using seven SNPs. *Forensic Sci Int Genet* 5(5):472-478.
12. Pneuman A, Budimlija ZM, Caragine T, Prinz M, Wurmbach E (2012) Verification of eye and skin color predictors in various populations. *Legal Medicine* 14(2):78-83.
13. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, Branicki W, Kayser M (2013) The HirisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci Int Genet* 7:98-115
14. Ruiz Y, Phillips C, Gomez-Tato A, Alvarez-Dios J, Casares de Cal M, Cruz R, Maroñas O, Söchtig J, Fondevila M, Rodriguez-Cid MJ, Carracedo A, Lareu MV (2013) Further development of forensic eye color predictive tests. *Forensic Sci Int Genet* 7(1):28-40.
15. Allwood JS, Harbison S (2013) SNP model development for the prediction of eye colour in New Zealand. *Forensic Sci Int Genet* 7(4):444-452.
16. Andersen JD, Johansen P, Harder S, Christoffersen SR, Delgado MC, Henriksen ST, Nielsen MM, Sørensen E, Ullum H, Hansen T, Dahl AL, Paulsen RR, Børsting C, Morling N (2013) Genetic analyses of the human eye colours using a novel objective method for eye colour classification. *Forensic Sci Int Genet* 7(5):508-515.
17. Sturm RA (2006) A golden age of human pigmentation genetics. *Trends Genet* 22:464-468
18. Branicki W, Brudnik U, Wojas-Pel A (2009) Interactions Between *HERC2*, *OCA2* and *MC1R* May Influence Human Pigmentation Phenotype. *Annals of Hum Genet* 73:160-170.
19. Liu F, Wenb B, Kayser M. (2013) Colorful DNA polymorphisms in humans. *Semin Cell Dev Biol* 24(6-7):562-575.
20. Cook AL, Chen W, Thurber AE, Smit DJ, Smith AG, Bladen TG, Brown DL, Duffy DL, Pastorino L, Bianchi-Scarra G, Leonard JH, Stow JL, Sturm RA. (2009) Analysis of Cultured Human Melanocytes Based on Polymorphisms within the *SLC45A2/MATP*, *SLC24A5/NCKX5*, and *OCA2/P* Loci. *J Invest Dermatol* 129(2):392-405.
21. Sturm RA, Duffy DL, Zhao ZZ, Leite FPN, Stark MS, Hayward NK, Martin NG, Montgomery GW (2008) A single SNP in an evolutionary conserved region within intron 86 of the *HERC2* gene determines human blue-brown eye color. *The American Journal of Human Genetics* 82:424-431.
22. Kondo T, Hearing VJ (2011) Update on the regulation of mammalian melanocyte function and skin pigmentation. *Expert Rev Dermatol* 6:97–108.
23. Visser M, Kayser M, Palstra RJ (2012) *HERC2* rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the *OCA2* promoter. *Genome Res* 22:446-455.
24. Valenzuela RK, Ito S, Wakamatsu K, Brilliant MH (2011) Prediction model validation: normal human pigmentation variation, *J Forensic Res* 2:139.

25. Pulker H, Lareu MV, Phillips C, Carracedo A (2007) Finding genes that underlie physical traits of forensic interest using genetic tools. *Forensic Sci Int Genet* 1(2):100-104.
26. Graf J, Voisey J, Hughes I, Van Daal A (2007) Promoter polymorphisms in the *MATP (SLC24A5)* gene are associated with normal skin color variation. *Hum Mutat* 28:710-717.
27. Vierkötter A, Krämer U, Sugiri D, Morita A, Yamamoto A, Kaneko N, Matsui M, Krutmann J (2012) Development of lentigines in german and japanese women correlates with variants in the *SLC24A5*. *Gene Journal of Investigative Dermatology* 132:733–736.
28. Graf J, Hodgson R, Van Daal A (2005) Single nucleotide polymorphisms in the *MATP* gene are associated with normal human pigmentation variation. *Hum Mutat* 25:278-284.
29. Nakayama K, Soemantri A, Jin F, Dashnyam B, Ohtsuka R, Duanchang P, Isa MN, Settheetham-Ishida W, Harihara S, Ishida T (2006) Identification of novel functional variants of the melanocortin 1 receptor gene originated from Asians. *Human Genet* 119:322-330.
30. Tully G (2007) Genotype versus phenotype: Human pigmentation. *Forensic Sci Int Genet* 1(2):105-110.
31. Makova K, Norton H (2005) Worldwide polymorphism at the *MC1R* locus and normal pigmentation variation in humans. *Peptides* 26:1901-1908.
32. Voisey J, Gomez-Cabrera MC, Smit DJ, Leonard JH, Sturm RA, van Daal A (2006) A polymorphism in the agouti signalling protein (*ASIP*) is associated with decreased levels of mRNA. *Pigment Cell Res* 19:226–231.
33. Gerstenblith MR, Goldstein AM, Fagnoli MC, Peris K, Landi MT (2007) Comprehensive evaluation of allele frequency differences of *MC1R* variants across populations. *Human Mutat* 28: 495-505.
34. Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynek MJ, Mao X, Humphreville VR, Humbert JE, Sinha S, Moore JL, Jagadeeswaran P, Zhao W, Ning G, Makalowska I, McKeigue PM, O'donnell D, Kittles R, Parra EJ, Mangini NJ, Grunwald DJ, Shriver MD, Canfield VA, Cheng KC (2005) *SLC24A5*, a putative cation exchanger, affects pigmentation in zebrafish and humans, *Science* 310:1782–1786.
35. Ginger RS, Askew SE, Ogborne RM, Wilson S, Ferdinando D, Dadd T, Smith AM, Kazi S, Szerencsei RT, Winkfein RJ, Schnetkamp PP, Green MR (2008) *SLC24A5* encodes a trans-Golgi network protein with potassium-dependent sodium–calcium exchange activity that regulates human epidermal melanogenesis. *J Biol Chem* 283:5486–5495.
36. Parra EJ (2007) Human pigmentation variation: Evolution, genetic basis, and implications for public health, *Year book of physical anthropology*. 50:85-105.
37. Frudakis T, Thomas M, Gaskin Z, Venkateswarlu K, Chandra KS, Ginjupalli S, Gunturi S, Natrajan S, Ponnuswamy VK, Ponnuswamy KN (2003) Sequences associated with human iris pigmentation. *Genetics* 165:2071–2083.
38. Alaluf S, Atkins D, Barrett K, Carter N (2003) Ethnic variation in tyrosinase and *TYRP1* expression in photoexposed and photoprotected human skin, *Pigment Cell Res* 16 35-42.

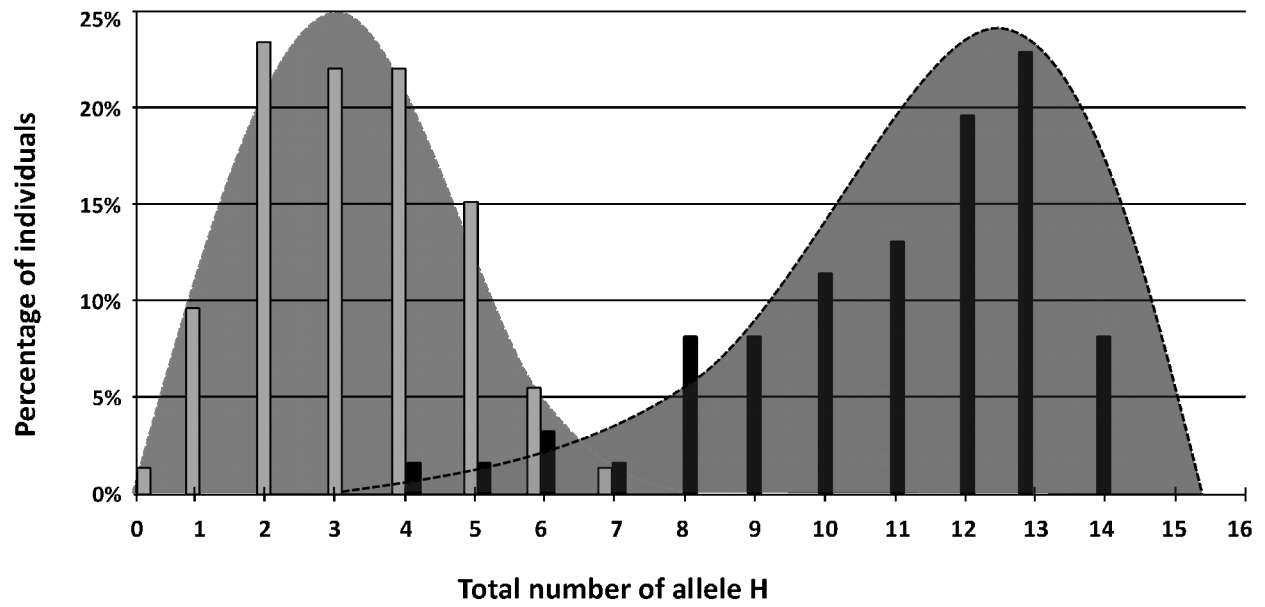
39. Jin Y, Birlea SA, Fain PR, Ferrara TM, Ben S, Riccardi SL, Cole JB, Gowan K, Holland PJ, Bennett DC, Luiten RM, Wolkerstorfer A, van der Veen JP, Hartmann A, Eichner S, Schuler G, van Geel N, Lambert J, Kemp EH, Gawkrödger DJ, Weetman AP, Taïeb A, Jouary T, Ezzedine K, Wallace MR, McCormack WT, Picardo M, Leone G, Overbeck A, Silverberg NB, Spritz RA (2012) Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. *Nature Genetics* 44:676–680
40. Norton HL, Kittles RA, Parra E, Mckeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, Mceovoy B, Shriver M (2007) Genetic evidence for the convergent evolution of light skin in Europeans and East Asian. *Mol Biol Evol* 24:710-22
41. Stokowski RP, Pant PVK, Dadd T, Fereday A, Hinds DA, Jarman C, Filsell W, Ginger RS, Green MR, van der Ouderaa FJ, Cox DR (2007) A genomewide association study of skin pigmentation in a south asian population, *The American Journal of Human Genetics* 81:1119–1132
42. Chaki M, Sengupta M, Mondal M, Bhattacharya A, Mallick S, Bhadra R, Ray K (2011) Molecular and functional studies of tyrosinase variants among indian oculocutaneous albinism type 1. *Patients Journal of Investigative Dermatology* 131:260–262
43. Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K, Vermeulen M, Arp P, Jhamai MM, van Ijcken WF, den Dunnen JT, Heath S, Zelenika D, Despriet DD, Klaver CC, Vingerling JR, de Jong PT, Hofman A, Aulchenko YS, Uitterlinden AG, Oostra BA, van Duijn CM (2008) Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 82(2):411-423. doi:10.1016/j.ajhg.2007.10.003
44. Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD (2003) Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA* 100(1):177-182
45. Otaka I, Kumagai K, Inagaki Y, Shimoyama M, Saegusa K, Hara T. (2002) Simple and inexpensive software designed for the evaluation of color. *Am J Ophthalmol* 133(1):140-142.
46. Huebinger RM, Garner HR, Barber RC (2010) Pathway genetic load allows simultaneous evaluation of multiple genetic associations. *Burns* 36(6):787-792.
47. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
48. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.
49. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620.
50. Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4(2): 359-361.

51. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.
52. Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* 4:137-138.
53. Metz CE (1978) Basic principles of ROC analysis. *Semin Nucl Med* 8(4):283-298
54. Sturm RA, Duffy DL (2012) Human pigmentation genes under environmental selection. *Genome Biology* 13:248. doi:10.1186/gb-2012-13-9-248
55. Sturm RA, Larsson M (2009) Genetics of human iris colour and patterns. *Pigment Cell Melanoma Res* 22:544-562
56. Soejima M, Koda Y (2007) Population differences of two coding SNPs in pigmentation-related genes *SLC24A5* and *SLC24A5*, *Int J Legal Med* 121:36-39
57. Rees JL (2003) Genetics of hair and skin color. *Annu Rev Genet* 37:67-90

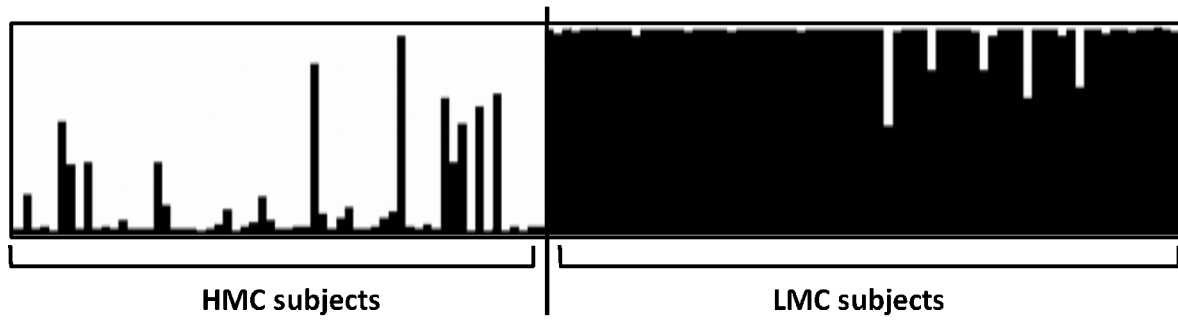




**FIGURE 1-** Layout of the COLOURS software (Otaka *et al.*, 2002). A triangular-shaped representative area of the iris covering approximately 2.500 pixels was selected in each picture. The R, G, and B values of each pixel were measured on a scale of 0 (zero) to 255 – where the minimum value represents the complete absence of a given color and the maximum represents its complete presence. To calculate the R value, COLOURS divides the total amount of red pigment present in each pixel of the selected area by the product of 255 times the total amount of pixels in the same area. Thus, in the picture on the right,  $264002 / (255 \times 2542) = 0.41$  – as the total amount of red in the selected area is 264002 and the total amount of pixels in the same area is 2542. The same mathematical formula is used for the calculation of G and B values.



**FIGURE 2-** Distribution of alleles H in Southern Brazilian subjects with High Melanin Content (gray bars; N=61) and Low Melanin Content (black bars; N=73).



**FIGURE 3-** STRUCTURE cluster plot for optimum  $K:3$ . Graph of the probability of each subject (each bar) belonging to two phenotype clusters: with high melanin content (in white) or with low melanin content (in black). Each vertical bar represents one subject (134 in total). The vertical black line separates the two different pre-categorized color groups, identified by the labels at the bottom.

**TABLE 1-** Primer sequences for conventional PCR and for SNaPShot multiplex system.

<b>RS (Gene; SNP)</b>	<b>Primer Sequence</b>
<b>rs916977 (HERC2; SNP A&gt;G)</b>	F-5'- ttgtttggcaaaactccaca-3' / R-5'- ttttcaattgcctgacat-3' (PCR Fragment: 200pb) SNaPShot (35pb): ggcaaaactccacagtggggatcagtttgagtaga <b>T/C</b>
<b>rs4778138 (OCA2; SNP G&gt;A)</b>	F-5'- ggattcaaaaagaaagtctcaagg-3' / R-5'- gctctccttttgataccagca-3' (PCR Fragment: 146bp) SNaPShot (42pb): tattgaactgaatgaaagtgaaaataacatatcaaaattg <b>G/A</b>
<b>rs12913832 (HERC2; SNP A&gt;G)</b>	F-5'- cagctccatcaatgtgtgca-3' / R-5'- ctgatgatgatagcgtgcaga-3' (PCR Fragment: 253pb) SNaPShot (56pb): ttcattgctctctgtgtctgatccaagaggcgaggccagtttcatttgagcattaa <b>A/G</b>
<b>rs16891982 (SLC452A; SNP C&gt;G)</b>	F-5'- ccaagttgtgctagaccagaaa-3' / R-5'- cctcaacagcctccaatctc-3' (PCR Fragment: 211pb) SNaPShot (64pb): aagacatccttaggagagagaaaagacttacaagaataaagtgaggaaaacacggagttgatgca <b>C/G</b>
<b>rs8045560 (MC1R; SNP C&gt;T)</b>	F-5'- aacgatgtttgtggtcagca-3' / R-5'- actcaaggcatctggaatgg-3' (PCR Fragment: 289pb) SNaPShot (73pb): caccacccttttccatggggatctgactcatctccaggaagatggttgggagataacccagctctgc <b>C/T</b>
<b>rs1426654 (SLC24A5; SNP G&gt;A)</b>	F-5'- gccttccctcaccctttcta-3' / R-5'- aggatggtgctaagccaat-3' (PCR Fragment: 425pb) SNaPShot (90pb): tagttgtaaagacatactttcactttattaggcataacaatcattcattatgttcagcccttgattgtctcaggatgttcaggc <b>G/A</b>
<b>rs2733832 (TYRP1; SNP C&gt;T)</b>	F-5'- atgacctgtctgtcgaagt-3' / R-5'- cttctgcctgcattttcaa-3' (PCR Fragment: 334pb) SNaPShot (98pb): ccaaatgatcctattttgtcctcctgcacaccttcacagatgcagctttgatgaatggctgaggagatacaatgctggaagacattttcatatgc <b>C/T</b>
<b>rs1042602 (TYR; SNP C&gt;A)</b>	F-5'- actgcaagtttgcttttgg-3' / R-5'- gcttcatgggcaaaatcaat-3' (PCR Fragment: 307pb) SNaPShot (106pb): ggccaaatgaaaatggatcaacaccatgttaacgacatcaatattatgacctttgtctggatgattattatgtgtcaatggatgactgcttggggat <b>C/A</b>

**TABLE 2-** Allele frequencies and other parameters of the selected SNPs based on population data from HAPMAP and ALFRED (the database used to create Table 1 is presented in a Supplementary Excel file).

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	13.38
	Allele L (T)	0.085 (134/1574)	0.560 (2203/3932)							
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	9.32	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)							

1- HMC (High Melanin Content) Populations: populations of Africans, as available on the HAPMAP and ALFRED websites. 2- LMC (Low Melanin Content) Populations: populations of Europeans, as available on the HAPMAP and ALFRED websites. 'Allele H': allele strongly associated with people with HMC (from African populations). 'Allele L': allele strongly associated with people with LMC (from European populations). 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.

**TABLE 3-** Phenotype data from the 134 previously selected Southern Brazilian subjects with high melanin content (HMC; N=61) and low melanin content (LMC; N=73).

		<b>HMC</b>	<b>LMC</b>
<b>TOTAL</b>	N (%)	61 (45.18)	73 (54.82)
<b>SEX</b>	Male [N (%)]	32 (52.46)	40 (54.80)
	Female [N (%)]	29 (47.54)	33 (45.20)
<b>AGE</b>	Years [Mean (SD)]	30.36 (8.74)	36.57 (13.65)
<b>SKIN TYPE*</b>	Type 1 [N (%)]	0	30 (41.09)
	Type 2 [N (%)]	0	43 (58.91)
	Type 5 [N (%)]	26 (42.62)	0
	Type 6 [N (%)]	35 (57.38)	0
<b>SKIN COLOR</b>	Red [Mean (SD)]	197.0 (51.9)	354.6 (52.5)
	Green [Mean (SD)]	141.8 (41.1)	282.2 (53.0)
	Blue [Mean (SD)]	111.8 (32.6)	248.0 (55.2)
<b>EYE TYPE</b>	Type 1 - Light blue [N (%)]	0	50 (68.49)
	Type 2 - Dark blue [N (%)]	0	23 (31.51)
	Type 5 - Dark brown [N (%)]	39 (63.94)	0
	Type 6 - Black [N (%)]	22 (36.06)	0
<b>EYE COLOR</b>	Red [Mean (SD)]	0.232 (0.075)	0.390 (0.049)
	Green [Mean (SD)]	0.128 (0.046)	0.376 (0.049)
	Blue [Mean (SD)]	0.168 (0.058)	0.541 (0.099)
	Hue [Mean (SD)]	325.9 (20.4)	224.6 (62.4)
	Saturation [Mean (SD)]	47.1 (9.0)	31.1 (9.2)
	Value of Brightness [Mean (SD)]	23.7 (8.5)	54.5 (9.1)
<b>AUTODECLARED ORIGIN</b>	European	13 (21.31)	66 (90.41)
	African	26 (37.70)	0
	Amerindian	2 (0.03)	0
	Mixed	10 (16.39)	4 (5.48)
	Unknown	10 (16.39)	3 (4.11)

\* See Type definitions on section 3 (Material and Methods).

**TABLE 4-** Allele frequencies and other parameters of the selected SNPs in 134 southern Brazilian subjects with high melanin content (N=61) and low melanin content (n=73).

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>SLC452A</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)	0.979	(143/146)							
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)	0.610	(89/146)							
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)	0.025	(4/146)	<0.0001	0.697	0.973	0.955	0.794	4.65	82.13
	Allele L (G)	0.303	(37/122)	0.975	(142/146)							

1- HMC SB People: Southern Brazilian subjects with High Melanin Content. 2- LMC SB People: Southern Brazilian subjects with Low Melanin Content. 'Allele H': allele strongly associated with people with HMC (from African populations). 'Allele L': allele strongly associated with people with LMC (from European populations). 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.

**TABLE 5-** Genotype frequencies of HH, HL and LL of the selected SNPs calculated from allele frequencies based on population data from HAPMAP and ALFRED (the database used to create Table 1 is presented in the Supplementary Excel file).

<b>HMC Populations<sup>1</sup></b>	<b>HH</b>	<b>HL</b>	<b>LL</b>
<b>rs4778138 (OCA2; H=G, L=A)</b>	0.55503	0.38144	0.06554
<b>rs12913832 (HERC2; H=A, L=G)</b>	0.91585	0.08230	0.00185
<b>rs16891982 (SLC452A; H=C, L=G)</b>	0.77616	0.20968	0.01416
<b>rs8045560 (MC1R; H= C, L=T)</b>	0.83723	0.15555	0.00723
<b>rs1426654 (SLC24A5; H=G, L=A)</b>	0.52563	0.39875	0.07563
<b>rs2733832 (TYRP1; H=C, L=T)</b>	0.84089	0.15222	0.00689
<b>rs1042602 (TYR; H=C, L=A)</b>	0.89492	0.10217	0.00292
<b>rs916977 (HERC2; H=A, L=G)</b>	0.79210	0.19580	0.01210
<b>LMC Populations<sup>2</sup></b>	<b>HH</b>	<b>HL</b>	<b>LL</b>
<b>rs4778138 (OCA2; H=G, L=A)</b>	0.02690	0.27421	0.69890
<b>rs12913832 (HERC2; H=A, L=G)</b>	0.08123	0.40584	0.50694
<b>rs16891982 (SLC452A; H=C, L=G)</b>	0.00449	0.12502	0.87049
<b>rs8045560 (MC1R; H= C, L=T)</b>	0.19360	0.49280	0.31360
<b>rs1426654 (SLC24A5; H=G, L=A)</b>	0.00012	0.02176	0.97812
<b>rs2733832 (TYRP1; H=C, L=T)</b>	0.17223	0.48555	0.34223
<b>rs1042602 (TYR; H=C, L=A)</b>	0.49280	0.41839	0.08880
<b>rs916977 (HERC2; H=A, L=G)</b>	0.04580	0.33641	0.61780

1- HMC (High Melanin Content) Populations: populations of Africans, as available on the HAPMAP and ALFRED websites. 2- LMC (Low Melanin Content) Populations: populations of Europeans, as available on the HAPMAP and ALFRED websites. 'Allele H': allele strongly associated with people with HMC (from African populations). 'Allele L': allele strongly associated with people with LMC (from European populations).



### CAPÍTULO 3 - CONSIDERAÇÕES FINAIS

Recentemente, novas tecnologias começaram a ser aplicadas largamente na área em estudo. Podemos citar, por exemplo, a genotipagem de SNPs através de *chips* semicondutores. Estas tecnologias de genotipagem de SNPs em larga escala, associadas com a descoberta de novas mutações capazes de prever características externas visíveis, nos darão o norte para o desenvolvimento de novas pesquisas.

De qualquer forma, os resultados obtidos nesta tese, assim como a natureza geopolítica da região sul do Brasil, em especial a do Estado do Rio Grande do Sul, demonstraram que o nosso painel em SNaPSHOT, para genotipagem de oito SNPs, pode ser uma ferramenta para predição fenotípica com alto nível de acerto.

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 os mais de 120 genes envolvidos na pigmentação humana.

## REFERÊNCIAS BIBLIOGRÁFICAS

- A. Napolitano, C. Costantini, O. Crescenzi, G. Prota. Characterization of 1,4- benzothiazine intermediates in the oxidative conversion of 5-S-cysteinyldopa to pheomelanins, *Tetrahedron Lett* 35 (1994) 6365–6368.
- A. Napolitano, P. Di Donato, G. Prota, E.J. Land. Transient quinonimines and 1,4-benzothiazines of pheomelanogenesis: new pulse radiolytic and spectrophotometric evidence, *Free Radic Biol Med.* 27 (1999) 521–528.
- A. Napolitano, P. Di Donato, G. Prota. New regulatory mechanisms in the biosynthesis of pheomelanins: rearrangement vs. redox exchange reaction routes of a transient 2H-1,4-benzothiazine-o-quinonimine intermediate, *Biochim Biophys Acta.* 1475 (2000) 47–54.
- A. Palumbo, M. d'Ischia, G. Prota. Tyrosinase-promoted oxidation of 5,6-dihydroxyindole-2-carboxylic acid to melanin. Isolation and characterization of oligomer intermediates, *Tetrahedron*, 43 (1987) 4203–4206.
- A. Palumbo, F. Solano, G. Misuraca, P. Aroca, J.C. Garcia-Borron, J.A. Lozano, G. Prota. Comparative action of dopachrome tautomerase and metal ions on the rearrangement of dopachrome, *Biochim Biophys Acta.* 1115 (1991) 1–5.
- A. Slominski, D. J. Tobin, S. Shibahara, J. Wortsman. Melanin pigmentation in mammalian skin and Its hormonal regulation. *American Physiological Reviews* 84(2004) 1155–1228.
- A. Van Daal, A. The genetic basis of human pigmentation, *Forensic Science International* 1(2008). 541–543.
- A. Vierkötter, U. Krämer, D. Sugiri, A. Morita, A. Yamamoto, N. Kaneko, M. Matsui, J. Krutmann. Development of lentigines in german and japanese women correlates with variants in the *SLC24A5*, *Gene Journal of Investigative Dermatology* 132 (2012) 733–736.
- C.L. Wilkerson, N.A. Syed, M.R. Fisher, N.L. Robinson, H.L. Wallow, D.M. Albert. Melanocytes and iris color. Light microscopic findings, *Archives of Ophthalmology* 114(1996)437–442.
- E. Giardina, A. Spinella, G. Novelli. Past, present and future of forensic DNA typing, *Nanomedicine* 6 (2011) 257–270.
- E.E. Kenny, N.J. Timpson, M.Sikora, M.C. Yee, A. Moreno-Estrada, C. Eng, S. Huntsman, E.G. Burchard, M. Stoneking, C.D. Bustamante, S. Myles. Melanesian blond hair is caused by an amino acid change in *TYRP1*, *Science* 336 (2012) 554.
- E.J. Land, S. Ito, K. Wakamatsu, P.A. Riley. Rate constants for the first two chemical steps of Eumelanogenesis, *Pigment Cell Res.* 16 (2003) 487–493.
- E.J. Parra. Human pigmentation variation: Evolution, genetic basis, and implications for public health, *Year book of physical anthropology.* 50 (2007) 85-105.
- F. Liu, K. van Duijn, J.R. Vingerling, A. Hofman, A.G. Uitterlinden, A.C.J.W. Janssens, M. Kayser. Eye color and the prediction of complex phenotypes from genotypes, *Curr. Biol.* 19 (2009) R192–R193.
- F. Liu, B. Wenb, M. Kayser. Colorful DNA polymorphisms in humans, *Semin Cell Dev Biol* (2013).

- G.B. Stacey, S.F. Schaffner, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. Defelice, A. Lochner, M. Faggart, S.N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E.S. Lander, M.J. Daly, D. Altshuler. The structure of haplotype blocks in the human genome, *Science* 296 (2002) 2225–2229.
- G. Dimisianos, I. Stefanaki, V. Nicolaou, V. Sypsa, C. Antoniou, M. Poulou, O. Papadopoulos, H. Gogas, E. Kanavakis, E. Nicolaidou, A. D. Katsambas, A. J. Stratigos. A study of a single variant allele (rs1426654) of the pigmentation-related gene *SLC24A5* in Greek subjects, *Experimental Dermatology* 18 (2009)175–7.
- G. E Costin, V. J Hearing. Human skin pigmentation: melanocytes modulate skin color in response to stress, *The Journal of the Federation of American Societies for Experimental Biology* 21 (2007) 976-994.
- G. Greco, K. Wakamatsu, L. Panzella, S. Ito, A. Napolitano, M. D'Ischia. Isomeric cysteinyl dopas provide a (photo)degradable bulk component and a robust structural element in red human hair pheomelanin, *Pigment Cell Melanoma Res.* 22 (2009) 319–327.
- G. Tully. Genotype versus phenotype: Human pigmentation, *Forensic Sci Int. Genet.* 1 (2007) 100-104.
- H. Ho, R. Kapadia, S. Al-Tahan, S.; et al. WIPI1 Coordinates Melanogenic Gene Transcription and Melanosome Formation via TORC1 Inhibition, *The Journal of Biological Chemistry* 286, n. 14. (2011) 12509-12523.
- H.L. Norton, R.A Kittles, E. Parra, P. Mckeigue, X. Mao, K. Cheng, V. A. Canfield, D.G Bradley, B. Mceovoy, M. Shriver. Genetic evidence for the convergent evolution of light skin in Europeans and East Asian, *Mol Biol Evol* 24(2007) 710-22.
- H. Pulker, M.V. Lareu, C. Phillips, A. Carracedo. Finding genes that underlie physical traits of forensic interest using genetic tools, *Forensic Sci Int Genet.* 1 (2007) 100-104.
- INTERNATIONAL FEDERATION OF PIGMENT CELL SOCIETIES – IFPCS 2009: Disponível em: <http://www.espcr.org/micemut/> - Acessado em abril de 2013.
- J. Berson, D. C. Harper, D. Tenza, G. Raposo, M.S. Marks. Pmel17 initiates premelanosome morphogenesis within multivesicular bodies, *Molecular Biology of the cell* 12(2001) 3451-3464.
- J. Graf, J. Voisey, I. Hughes, A. Van Daal. Promoter polymorphisms in the *MATP (SLC24A5)* gene are associated with normal skin color variation, *Hum. Mutat.* 28 (2007) 710-717.
- J. Graf, R. Hodgson, A. Van Daal. Single nucleotide polymorphisms in the *MATP* gene are associated with normal human pigmentation variation, *Hum. Mutat.* 25 (2005) 278-284.
- J.L. Rees. The genetics of Sun sensitivity in humans, *Am. J. Hum. Genet.* 75 (2004) 739-751.
- J.R. Landis, G.G Koch. The measurement of observer agreement for categorical data, *Biometrics* 33(1977)159-174.
- J.S. Allwood, S.A. Harbison. SNP model development for the prediction of eye colour in New Zealand, *Forensic Science International: Genetics* 7 (2013) 444–452.

- J. Voisey, M.C. Gomez-Cabrera, D.J. Smit, J.H. Leonard, R.A. Sturm, A. van Daal. A polymorphism in the agouti signalling protein (ASIP) is associated with decreased levels of mRNA Pigment Cell Res. 19 (2006) 226–231.
- J.Y Lin, D.E Fisher. Melanocyte biology and skin pigmentation, Nature 445 (2007) 843-850.
- K. Makova, H. Norton. Worldwide polymorphism at the *MC1R* locus and normal pigmentation variation in humans, Peptides 26 (2005) 1901-1908.
- K. Nakayama, A. Soemantri, F.J. Bumbein Dashnyam, R. Ohtsuka, P. Duanchang, M. N. Isa, W. Settheetham-Ishida, S. Harihara, T. Ishida. Identification of novel functional variants of the melanocortin 1 receptor gene originated from asians, Human Genet 119(2006) 322-30.
- K. Wakamatsu, K. Ohtara, S. Ito. Chemical analysis of late stages of pheomelanogenesis: conversion of dihydrobenzothiazine to a benzothiazole structure, Pigment Cell Melanoma Res. 22 (2009) 474–486.
- K. Yoshida, K. Yayama, A. Hatanaka, K. Tamaki. Efficacy of extended kinship analyses utilizing commercial STR kit in establishing personal identification, Leg. Med. 13 (2011) 12–15.
- L.D.G. Kobachuk. Estudo de frequências alélicas de dez locos STRs do cromossomo X na população do estado do Paraná e sua contribuição na identificação humana, (2012).
- L. Excoffier, H.E.L. Lischer. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol Ecol Resour 10 (2010) 564-67.
- M.A. Jobling, P. Gill. Encoded evidence: DNA in forensic analysis, Nature Rev. Genet. 5 (2004) 739–751.
- M. Brenner, M. Berking. Grundlagen der Hautpigmentierung: Biochemie und regulation der melaninsynthese, Der Hautarzt 61(2010) 554-560.
- M. Chaki, M. Sengupta, M. Mondal, A. Bhattacharya, S. Mallick, R. Bhadra, K. Ray. Molecular and functional studies of tyrosinase variants among indian oculocutaneous albinism type 1, Patients Journal of Investigative Dermatology 131 (2011) 260–262.
- M. Kayser, P. Schneider. DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations, Forensic Science International 3(2009)154–161.
- M. Kayser, P. Knijff. Improving human forensics through advances in genetics, genomics and molecular biology, Nature Reviews 12 (2011) 179-192.
- M.P. Donnelly, P. Paschou, E. Grigorenko, D. Gurwitz, C. Barta, R.B. Lu, O.V. Zhukova, J.J. Kim, M. Siniscalco, M. New, H. Li, S.L.B. Kajuna, V.G. Manolopoulos, W.C. Speed, A.J. Pakstis, J.R. Kidd, K.K. Kidd. A global view of the *OCA2-HERC2* region and pigmentation, Hum Genet 131 (2012) 683-696.
- M.R, Gerstenblith, A.M, Goldstein, M.C, Fargnoli, K. Peris, M.T Landi. Comprehensive evaluation of allele frequency differences of *MC1R* variants across populations, Human Mutat 28(2007) 495-505.
- M. Soejima, Y. Koda. Population differences of two coding SNPs in pigmentation-related genes *SLC24A5* and *SLC24A5*, Int J Legal Med 121 (2007) 36-39.

- M. Visser, M. Kayser, R.J. Palstra. *HERC2* rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the *OCA2* promoter *Genome Res.* 22 (2012) 446-455.
- N. Jablonski, G. Chaplin. The evolution of human skin coloration, *Journal of Human Evolution*, v. 39(2000) 57-106.
- N. Jablonski. The Evolution of Human Skin and Skin Color, *Annual Review of Anthropology* v. 33 (2004) 585–623.
- O’Rahilly, S. Human genetics illuminates the path to metabolic disease, *Nature* 462(2009) 307-14.
- P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, K.P. Magnusson, A. Manolescu, A. Karason, A. Palsson, G. Thorleifsson, M. Jakobsdottir, S. Steinberg, S. Palsson, F. Jonasson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K.R. Benediksdottir, K.K. Aben, L.A. Kiemenev, J.H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir, K. Stefansson. Genetic determinants of hair, eye and skin pigmentation in Europeans, *Nat. Genet.* 39 (2007) 1443–1452.
- R.A Sturm, R. Teasdale, N. Box. Human pigmentation genes: Identification, Structure and Consequences of Polymorphic Variation, *Gene* 277(2001) 49–62.
- R.A Sturm, T.N Frudakis. Eye colour: portals into pigmentation genes and ancestry, *Trends in Genetics* 20(2004) 327–332.
- R.A. Sturm. A golden age of human pigmentation genetics, *Trends Genet.* 22 (2006) 464-468.
- R.A. Sturm, D.L. Duffy, Z.Z. Zhao, F.P.N. Leite, M.S. Stark, N.K. Hayward, N.G. Martin, G.W. Montgomery, A single SNP in an evolutionary conserved region within intron 86 of the *HERC2* gene determines human blue-brown eye color. *The American Journal of Human Genetics* 82 (2008) 424-431.
- R.A. Sturm. Molecular genetics of human pigmentation diversity, *Hum. Mol. Genet. Review Issue* 18 (2009) R9-R17.
- R.A. Sturm, M. Larsson. Genetics of human iris colour and patterns, *Pigment Cell Melanoma Res.* 22 (2009) 544-562.
- R.A. Sturm, D.L. Duffy. Human pigmentation genes under environmental selection, *Genome Biology* (2012) 13:248. doi:10.1186/gb-2012-13-9-248
- R.K. Valenzuela, S. Ito, K. Wakamatsu, M.H. Brilliant. Prediction model validation: normal human pigmentation variation, *J. Forensic Res.* 2 (2011). doi:10.4172/2157-7145.1000139
- R.L. Lamason, M.A. Mohideen, J.R. Mest, A.C. Wrong, H.L. Norton, M.C. Aros, M.J. Juryneec, X. Mao, V.R. Humphreville, J.E. Humbert, S. Sinha, J.L. Moore, P. Jagadeeswaran, W. Zhao, G. Ning, I. Makalowska, P.M. McKeigue, D. O’Donnell, R. Kittles, E.J. Parra, N.J. Mangini, D.J. Grunwald, M.D. Sriver, V.A. Canfield, K.C. Cheng. *SLC24A5*, a putative cation exchanger, affects pigmentation in zebrafish and humans, *Science* 310(2005) 1782–1786.
- R. M. Harding, E. Healy, A. J. Ray, N. S. Ellis, N. Flanagan, C. Todd, C. Dixon, A. Sajantila, I. J. Jackson, M. A. Birch-Machin, J. L. Rees. Evidence for Variable Selective Pressures at *MC1R*, *Am J Human Genet* 66(2000)1351-61.

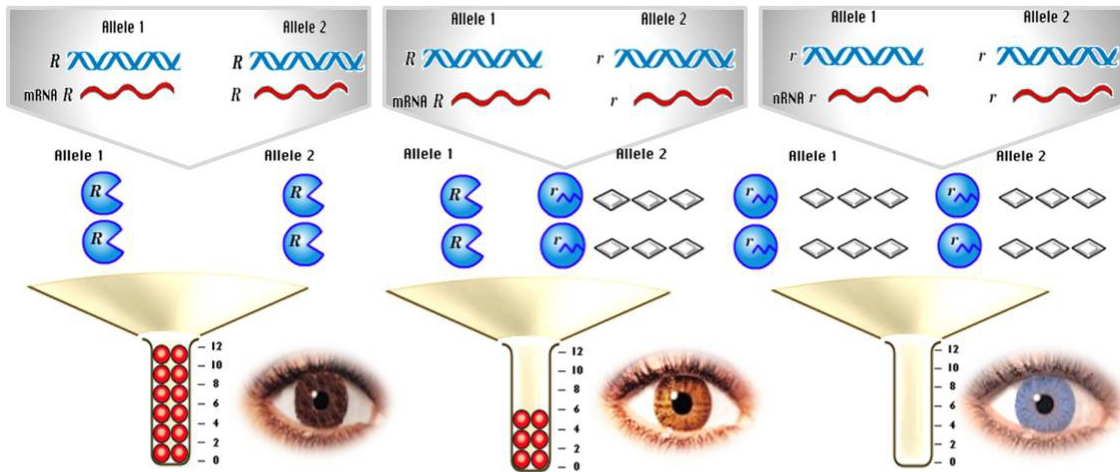
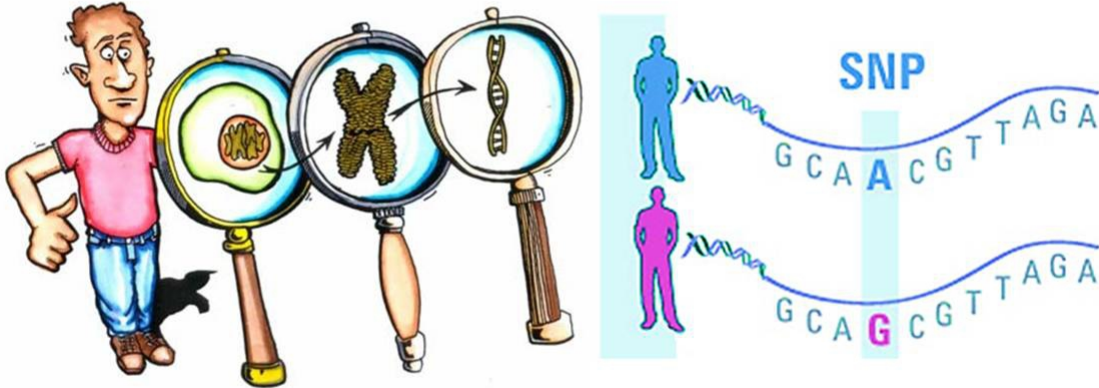
- R.P. Stokowski, P.V.K. Pant, T. Dadd, A. Fereday, D.A. Hinds, C. Jarman, W. Filsell, R.S. Ginger, M.R. Green, F.J. van der Ouderaa, D.R. Cox. A genomewide association study of skin pigmentation in a south asian population, *The American Journal of Human Genetics*. 81 (2007) 1119–1132.
- R.S. Ginger, S. E. Askew, R.M. Ogborne, R.M Wilson, D. Ferdinando, T. Dadd , A.M Smith, S. Kazi, R.T Szerencsei, R.J Winkfein, P.P Schnetkamp, M.R Green. SLC24A5 encodes a trans-Golgi network protein with potassium-dependent sodium–calcium exchange activity that regulates human epidermal melanogenesis, *J Biol Chem* 283(2008) 5486–5495.
- S. Alaluf, D. Atkins, K. Barrett, M. Blout, N. Carter, A. Heath. Ethnic variation in melanin content and composition in photoexposed and photoprotected human skin, *Pigment Cell Res* 15(2002) 112-118.
- S. Alaluf, K. Barrett, M. Blount, N. Carter. Ethnic variation in tyrosinase and *TYRP1* expression in photoexposed and photoprotected human skin, *Pigment. Cell. Res.* 16 (2003) 35-42.
- S. Ito, K. Wakamatsu, Chemistry of Melanins. In: J.J. Nordlund, R.E. Boissy, V.J. Hearing, R.A. King, W.S. Oetting, J.P. Ortonne (Eds.). *The Pigmentary System. Physiology and Pathophysiology*, Blackwell Publishing; Oxford, 2006, pp. 282-310.
- S. Walsh, A. Lidenbrgh, S.B. Zuniga, T. Sijen, P. Knijff, M. Kayser, K.N. Ballantyne. Developmental validation of the IrisPlex system: Determination of blue and brown iris colour for forensic intelligence, *Forensic Sci Int. Genet.* 5 (2011a) 464-471.
- S. Walsh, F. Liu, K.N. Ballantyne, M. van Oven, O. Lao, M. Kayser. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information, *Forensic Sci Int. Genet.* (2011b). doi:10.1016/j.fsigen.2010.02.004
- S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser. The HirisPlex system for simultaneous prediction of hair and eye colour from DNA, *Forensic Sci Int. Genet.* 7 (2013) 98-115.
- T. Frudakis, M. Thomas, Z. Gaskin, K. Venkateswarlu, K.S. Chandra, S. Ginjupalli, S. Gunturi, S. Natrajan, V.K. Ponnuswamy, K.N. Ponnuswamy. Sequences associated with human iris pigmentation, *Genetics* 165 (2003) 2071–2083.
- T. Kondo, V.J. Hearing. Update on the regulation of mammalian melanocyte function and skin pigmentation, *Expert Rev Dermatol.* 6 (2011) 97–108.
- T. Kushimoto, V. Basrur, J. Valencia, J. Matsunaga, W.D. Vieira, V.J. Ferrans, J. Muller, E. Appella, V.J. Hearing. A model for melanosome biogenesis based on the purification and analysis of early melanosomes. *Proceedings of the National Academy of Sciences of the United States of America.* 98(2001) 10698-10703.
- W. Branicki, U. Brudnik, A. Wojas-Pel. Interactions Between *HERC2*, *OCA2* and *MC1R* May Influence Human Pigmentation Phenotype, *Annals of Hum Genet.* 73 (2009) 160-170.
- Y. Jin, S.A. Birlea, P.R Fain, T.M. Ferrara, S. Ben, S.L. Riccardi, J.B. Cole, K. Gowan, P.J. Holland, D.C. Bennett, R.M. Luiten, A. Wolkerstorfer, J.P.W. van der Veen, A. Hartmann, S. Eichner, G. Schuler, N. van Geel, J. Lambert, E.H. Kemp, D.J. Gawkrödger, A.P. Weetman, A. Taïeb, T. Jouary, K. Ezzedine, M.R.

Wallace, W.T. McCormack, M. Picardo, G. Leone, A. Overbeck, N.B. Silverberg, R.A. Spritz. Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo, *Nature Genetics* 44 (2012) 676–680.

MATERIAL INFORMATIVO APRESENTADO AOS PARTICIPANTES DA PESQUISA



PROJETO: ANÁLISE DE SNPs EM GENES DE  
PIGMENTAÇÃO HUMANA EM INDIVÍDUOS  
FENOTIPICAMENTE DISTINTOS  
LABORATÓRIO DE GENÉTICA HUMANA





## PASSO 1: CUMPRIMENTO, APRESENTAÇÃO E EXPLICAÇÕES



## PASSO 2: TCLE



## PASSO 3: DADOS

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

INFORMAÇÃO NÚMERO: COE [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]

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

GRUPOS EPIDEMIOLÓGICOS  
Sexo:  M  F  
Etnia:  Asiática  Indígena  Portuguesa  Brasileira  
 Espanhola  Europeia  Desconhecida  Outra  
Religião:  Católica  Protestante  Outra  
Escolaridade:  < 8 anos  8 a 11 anos  12 a 14 anos  15 a 17 anos  > 17 anos

GRUPOS BIOLÓGICOS  
COR DOS OLHOS:  Azul Claro  Azul Escuro  Verde  Amarelo  Verde Amarelado  Marrom Claro  Marrom Escuro  Preto

COR DA PELE:  Tipo I  Tipo II  Tipo III  Tipo IV  Tipo V  Tipo VI

RESERVAÇÕES / COMPLEMENTAÇÕES

LOCAL E ASSINATURA DE QUEM PREENCHEU O INSTRUMENTO

## PASSO 5: REGISTRO DA COR DA PELE



## PASSO 4: FOTO DOS OLHOS



## PASSO 6: COLETA DE CÉLULAS



## PASSO 7: OBTENÇÃO E ANÁLISE DE DNA



## FINAL: COMPARAÇÃO ENTRE DNA, COR DA PELE E COR DOS OLHOS.

## PROCEDIMENTO OPERACIONAL PADRÃO PARA COLETA DE DADOS

### PESQUISA: MARCADORES GENÉTICOS PARA PIGMENTAÇÃO HUMANA POP COLETA

#### 1- Abordar o sujeito da pesquisa, cumprimentá-lo e apresentar o trabalho.

PESQUISADOR: “Bom dia/Boa tarde/Boa noite, somos do LGHM da PUCRS, estamos realizando uma pesquisa sobre os genes que determinam a coloração da pele e dos olhos, o senhor(a) gostaria de participar?”

PESQUISADOR: “São três passos simples e não demora muito tempo, faremos o seguinte: registraremos a cor da pele do seu braço com um medidor de cor; fotografaremos seus olhos e passaremos um cotonete na sua bochecha para coletar células”.

SE O SUJEITO REJEITAR: Agradecer e se despedir.

SE O SUJEITO ACEITAR: Prosseguir.

#### 2- Passar o TCLE para que o sujeito da pesquisa assine e leve cópia.

PESQUISADOR: “Assine, por favor, e fique com uma cópia deste TCLE. Neles há explicações sobre a pesquisa, e constam os telefones de contato caso o senhor queira: 1- saber mais sobre a pesquisa; 2- sair a qualquer momento da pesquisa; 3- receber os resultados da pesquisa”.

#### 3- Guardar o TCLE

#### 4- Preencher completa e corretamente o Instrumento de coleta de dados do projeto.

O PESQUISADOR pergunta ao SUJEITO e preenche o Instrumento.

Lembrar-se de perguntar e anotar sobre:

1- Usa lentes de contato?

2- Usa autobronzeador?

3- Tomou banho de sol recentemente [está bronzeado]?

4- Comeu algo a pouco tempo?

#### 5- Fotografar os dois olhos do SUJEITO.

INSTRUÇÕES:

1- Posicionar o SUJEITO em ambiente com luz artificial.

2- Ajustar a câmara para as seguintes posições: Com flash; modo macro, zoom macro máximo, sensibilidade ISO 80, modo imagem 12 megapixels com resolução de 4000X3000 qualidade máxima, apoiada em tripé.

3- Aproximar a câmara fotográfica até 6±0,5cm de distância entre o olho e a objetiva. Olho direito primeiro.

4- Disparar a foto usando flash atenuado por material translúcido.

5- Anotar a numeração, a data e o horário das fotos no Instrumento de coleta.

#### 6- Medir a coloração do SUJEITO.

INSTRUÇÕES:

1- Posicionar o SUJEITO em ambiente com luz artificial.

2- Realizar a calibração do equipamento com o auxílio do cartão de calibração fornecido com o 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 3(três) BIPS e identificar R, G e B com o valor 1023, confirmando o processo de calibragem.

3- Encostar o sensor do instrumento na parte interna do braço a 5±0,5cm abaixo 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 dados 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.

#### 7- Coletar células da mucosa oral do SUJEITO com dois swabs.

INSTRUÇÕES:

1- Usar luvas.

2- Posicionar o SUJEITO com a boca aberta.

3- Mostrar ao SUJEITO que o swab é estéril e está fechado.

4- Abrir o swab rasgando a ponta do envelope de papel.

5- Não tirar o swab do envelope, apenas amarrotá-lo para liberar a ponta do swab.

6- Fazer esfregaço do lado direito interno da boca do SUJEITO por dez segundos, girando o swab.

7- Transferir a ponta do swab para dentro do tubo de eppendorff.

8- Com a tesoura esterilizada com hipoclorito, cortar a haste do swab, por cima do papel, tomando cuidado para que não caia nenhum vestígio do papel no eppendorff.

9- Fechar e identificar o eppendorff com o número de coleta e as iniciais do sujeito.

10- Repetir o procedimento do lado esquerdo interno da boca.

#### 8- Encerrar o procedimento.

PESQUISADOR: “Muito obrigado(a) pela sua valiosa participação. Esteja à vontade para fazer qualquer pergunta agora ou quando quiser.”

## TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado para participar de uma pesquisa da PUCRS. O título da pesquisa é **“ANÁLISE DE SNPs EM GENES DE PIGMENTAÇÃO HUMANA EM INDIVÍDUOS GENETICAMENTE DISTINTOS”**, que tem por objetivo estudar partes do material genético que possam estar relacionadas à cor da pele e à cor dos olhos das pessoas. Para isso 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.

Este estudo não trará benefícios para você diretamente, nesse momento. Mas, haverá benefícios para a sociedade se conseguirmos encontrar relações entre o material genético e a cor da pele e olhos da pessoa. Isso poderá ajudar a identificar pessoas desaparecidas, por exemplo.

Sinta-se à vontade para fazer qualquer pergunta sobre o estudo e esclarecer suas dúvidas. Os responsáveis pela pesquisa garantidamente lhe responderão. Caso você tenha vontade de ser retirado da pesquisa, diga isso a qualquer momento, sem necessidade de 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 a você, sempre que você quiser. A dignidade e a privacidade dos participantes desta pesquisa, em relação a qualquer dado, serão mantidas. Os dados dos participantes da pesquisa 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. No estudo dos dados, a identidade do participante será desconhecida tanto pela equipe do trabalho como por qualquer outra pessoa externa.

A partir do mesmo material poderão ser futuramente feitos outros estudos. Para isso, você poderá permitir que sua amostra seja usada em novos estudos. Mesmo que esta permissão seja dada agora, estudos futuros só serão realizados se forem aprovados por um Comitê de Ética em Pesquisa.

Eu, \_\_\_\_\_ (*preencher com nome do participante da pesquisa*) fui informado dos objetivos da pesquisa acima 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 modificar minha decisão se assim eu o desejar. A equipe da Dra. Clarice S. Alho certificou-me de que todos os dados referentes ao participante desta pesquisa serão confidenciais, bem como sua rotina de vida não será modificada em razão desta pesquisa e que, ainda, terei a liberdade de retirar meu consentimento de participação na pesquisa se assim o desejar. Assim, como sujeito de pesquisa, dou consentimento para participar deste estudo. Eu também permito que a amostra de material genético seja armazenada na universidade, sob responsabilidade dos Pesquisadores Coordenadores, para estudos futuros. Minha permissão vale desde que os novos estudos sejam sempre devidamente aprovados por um Comitê de Ética em Pesquisa e estejam sob responsabilidade dos mesmos Pesquisadores Coordenadores. Fui informado também que, nesta situação, caso alguma informação considerada importante seja identificada no material genético analisado em decorrência da pesquisa, os pesquisadores poderão buscar a mim para oferecer acesso a tal informação.

Caso eu tenha novas perguntas sobre este estudo, posso telefonar a Dra. Clarice S. Alho nos telefones (51) 91136334 ou (51) 3320 3545 ramal 4534, ou para esclarecer sobre os meus direitos como participante deste estudo, ou se penso que fui prejudicado pela minha participação, posso chamar também os responsáveis pelo Comitê de Ética em Pesquisa – CEP-PUCRS no telefone (51) 3320 3000 ramal 3345. As ligações para os celulares poderão ser realizadas a cobrar. Declaro também que recebi cópia do presente Termo de Consentimento.

**Assinatura do Participante** Nome do Participante \_\_\_\_\_/\_\_\_\_\_/20\_\_\_\_\_

**Assinatura do Pesquisador** Nome do Pesquisador \_\_\_\_\_/\_\_\_\_\_/20\_\_\_\_\_

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

**Assinatura da Testemunha** Nome da Testemunha \_\_\_\_\_/\_\_\_\_\_/20\_\_\_\_\_

**INSTRUMENTO DE COLETA DE DADOS**

**PESQUISA: MARCADORES GENÉTICOS PARA PIGMENTAÇÃO HUMANA**

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

Indivíduo NÚMERO:    COR

número			iniciais		

**DADOS PESSOAIS**

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

Nome da Mãe: \_\_\_\_\_

Contato: \_\_\_\_\_

**DADOS EPIDEMIOLÓGICOS**

Sexo:            ( ) M            ( ) F  
 Origem:        ( ) Alemã      ( ) Italiana    ( ) Africana    ( ) Indígena    ( ) Portuguesa ( ) Espanhola  
                   ( ) Brasileira ( ) Árabe        ( ) Desconhecida ( ) Outra: \_\_\_\_\_  
 Residente em:            ( ) RS            ( ) SC            ( ) PR            ( ) Outro: \_\_\_\_\_  
 Natural de:                ( ) RS            ( ) SC            ( ) PR            ( ) Outro: \_\_\_\_\_

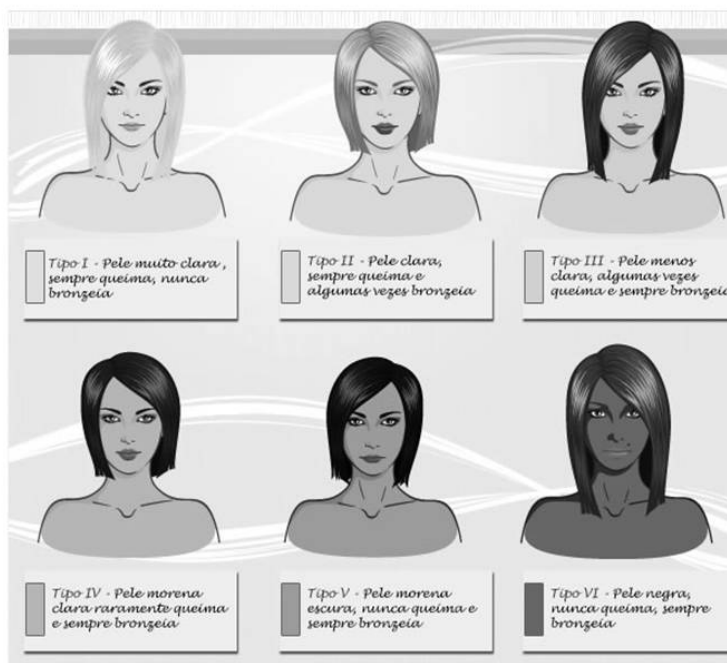
**DADOS BIOLÓGICOS**

**COR OLHO:**    ( ) Azul Claro  
                   ( ) Azul Escuro  
                   ( ) Verde  
                   ( ) Mel (Amarelado)  
                   ( ) Marrom Claro  
                   ( ) Marrom Escuro  
                   ( ) Preto

**Número da Foto:** \_\_\_\_\_

**COR PELE:**    ( ) Tipo I  
                   ( ) Tipo II  
                   ( ) Tipo III  
                   ( ) Tipo IV  
                   ( ) Tipo V  
                   ( ) Tipo VI

**Medidor:**    R: \_\_\_\_ R: \_\_\_\_ R: \_\_\_\_ Média: \_\_\_\_  
                   G: \_\_\_\_ G: \_\_\_\_ G: \_\_\_\_ Média: \_\_\_\_  
                   B: \_\_\_\_ B: \_\_\_\_ B: \_\_\_\_ Média: \_\_\_\_



**OBSERVAÇÕES / COMPLEMENTAÇÕES**


**NOME E ASSINATURA DE QUEM PREENCHEU O INSTRUMENTO**

--