UNIVERSIDADE DE SÃO PAULO

FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

Programa de Pós-Graduação em Clínica Médica

Wagner Narciso de Campos

**SÍTIOS POLIMÓRFICOS DO GENE *HFE* EM ESTUDOS
POPULACIONAIS E EM DOENÇAS HEREDITÁRIAS OU
ADQUIRIDAS QUE CURSAM COM SOBRECARGA DE FERRO**

**RIBEIRÃO PRETO
2013**

WAGNER NARCISO DE CAMPOS

**SÍTIOS POLIMÓRFICOS DO GENE *HFE* EM ESTUDOS
POPULACIONAIS E EM DOENÇAS HEREDITÁRIAS OU
ADQUIRIDAS QUE CURSAM COM SOBRECARGA DE FERRO**

Tese apresentada ao Departamento de Clínica Médica da Faculdade de Medicina de Ribeirão Preto –FMRP/USP, para obtenção do título de Doutor em Ciências - Área de concentração: Clínica Médica.

Orientador: **Prof. Dr. Eduardo Antônio Donadi**

**RIBEIRÃO PRETO
2013**

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Campos, Wagner Narciso de

Sítios polimórficos do gene *HFE* em estudos populacionais e em doenças hereditárias ou adquiridas que cursam com sobrecarga de ferro. Ribeirão Preto, 2013.

Tese de doutorado, Faculdade de Medicina de Ribeirão Preto/USP – Departamento de Clínica Médica. Área de concentração: Clínica Médica.

Orientador: Donadi, Eduardo Antonio.

119 f. : il.; 30 cm

1. HFE. 2. Sobrecarga de ferro. 3. Hemocromatose. 4. Hepatite C. 5. Carcinoma Hepatocelular.

Wagner Narciso de Campos

SÍTIOS POLIMÓRFICOS DO GENE *HFE* EM ESTUDOS
POPULACIONAIS E EM DOENÇAS HEREDITÁRIAS OU
ADQUIRIDAS QUE CURSAM COM SOBRECARGA DE FERRO

Comissão Julgadora da
Tese de obtenção do grau de Doutor

Prof. Dr. Eduardo Antônio Donadi
Orientador/Presidente

1º. Examinador

2º. Examinador

3º. Examinador

4º. Examinador

Ribeirão Preto, _____ de _____ de 2013.

“(...) É mais ou menos em cada milhão de anos vai acontecer qualquer coisa no mundo e as pessoas vão mudar de aspecto.vão começar a ter duas cabeças e vão mandar as outras para qualquer lugar. Mas alguns não acham isso. Pensam que, (...) a única maneira de as pessoas não fazerem das outras serpentes é toda a gente aceitar toda a outra gente como filhos (...), filhos de uma única mãe, não importa o aspecto que têm ou aquilo que fazem.”

Alice Walker, “A cor púrpura”, 1982.

DEDICATÓRIA

*Dedico ao Alan Fontoura
pela sua paciência e
companheirismo constante.*

AGRADECIMENTOS

Ao **Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq**, pelo financiamento que possibilitou o desenvolvimento desse trabalho.

Ao **Prof. Dr. Eduardo Antônio Donadi**, que generosamente abriu seus laboratórios e ofereceu a preciosa oportunidade para meu desenvolvimento profissional com sua orientação, e teve o seu máximo esforço para que eu realizasse meu trabalho.

À **Dra. Juliana Doblas Massaro**, uma pesquisadora de excelência e amiga generosa, cuja ajuda integral durante todas as fases desse trabalho foi essencial para o desenvolvimento e manutenção da qualidade dessa pesquisa.

À **Dra. Ana de Lourdes Candolo Martinelli**, que ofereceu sua valiosa orientação clínica, acesso ao seu banco de amostra e se mostrou comprometida de forma generosa com a qualidade dessa pesquisa.

Ao **Prof. Dr. Eduardo Luiz Rachid Cançado** e à **Dra. Andréia Evangelista da Silva**, por ceder amostras de pacientes portadores de hemocromatose hereditária que foram importantes para um melhor desenvolvimento dessa pesquisa.

Às técnicas **Sandra Rodrigues** e **Flavia Tremeschin de Almeida**, que se mostram amigas e solícitas em todo momento facilitando o desenvolvimento, principalmente pelos seus impecáveis trabalhos no laboratório, de forma manter o ambiente organizado e acessível, além de prestar qualquer orientação técnica que eu necessitava de forma espontânea e amistosa.

À **Daiani Cristina Cilião Alves** e **Gustavo Martelli Palomino** pelo companheirismo e ajuda e orientação técnica em diversos momentos do desenvolvimento desse trabalho.

Finalmente ao **Alan Fontoura dos Santos**, que caminhou junto comigo nessa jornada com paciência e carinho e compreensão, enfrentados todas as dificuldades e obstáculos que por ventura tive que enfrentar para o desenvolvimento desse projeto.

Peço desculpa a aqueles que eu tenha me esquecido de agradecer, mas que se sintam profundamente reverenciado independente da extensão de sua ajuda.

RESUMO

Mutações no gene *HFE* têm sido apontadas como um dos principais fatores para o desenvolvimento de sobrecarga de ferro, especialmente os SNPs (*single nucleotide polymorphism*) clássicos C282Y e H63D. A sobrecarga de ferro pode ser primária, atribuída diretamente a mutações do gene *HFE* (hemocromatose hereditária - HH) ou adquirida em decorrência de fatores genéticos e de comorbidades, particularmente, infecção pelo vírus da hepatite C (HCV) e carcinoma hepatocelular (CHC). Desequilíbrio de ligação entre os SNPs principais do gene *HFE* com alelos dos genes *HLA-A* e *HLA-B* podem contribuir com alterações da função das células do sistema imunológico. Neste trabalho, sequenciamos a região codificadora do gene *HFE* (éxon 2, 3, 4 e 5) de pacientes que apresentam ou não sobrecarga de ferro primária ou adquirida. Assim, estudamos 130 pacientes com hepatite C, 60 pacientes com CHC, 14 pacientes com HH e 100 indivíduos saudáveis. Foram encontrados sete SNPs no gene *HFE* no sentido 5' para 3': i) H63D C>G no éxon 2 (rs1799945); ii) IVS2(+4)T>C no íntron 2 (rs2071303); iii) íntron 3 C>G (rs807209); iv) C282Y G>A no éxon 4 (rs1800562); v) íntron 4 G>A (rs2794717); vi) IVS4(-44) T>C (rs1800708); vii) no íntron 5 a deleção G>del, ainda não foi descrita na literatura. Foram realizadas as reconstruções de haplótipos da região codificadora e os haplótipos estendidos englobando o gene *HFE* e dez outros *loci* (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQB1*, *TNFA*, *TNFB*, *TNFC*, *TNFD* e *HLA-G*-14pb). Também, normatizamos a nomenclatura do gene de acordo com as regras do *The International ImMunoGeneTics Information System* - IMGT (<http://www.imgt.org/>). Diversas ferramentas foram utilizadas para avaliar a associação entre os sítios polimórficos do gene *HFE* com a sobrecarga de ferro, incluindo testes de associações avaliando alelos, genótipos, desequilíbrio de ligação, haplótipos e diplótipos, testes de diversidades e neutralidade. Finalmente, comparamos os SNPs do gene *HFE* encontrados na população saudável do presente trabalho com as diversas populações mundiais, disponíveis no sítio 1000genomes (<http://www.1000genomes.org/>). Os resultados demonstraram que a nossa população estava mais próxima das populações europeias e americanas, sendo parcialmente próxima das populações africanas e distante das populações asiáticas. O alelo C282Y G, contido no haplótipo da região codificadora *HFE**01:03, conferiu susceptibilidade a HH na população brasileira testada, concordando com os achados observados em outras populações mundiais. O diplótipo *HFE**01:02:01:01/*HFE**01:01:01:05 conferiu susceptibilidade para o desenvolvimento de sobrecarga de ferro em pacientes com CHC causado pelo HCV. Detectamos forte desequilíbrio entre os alelos H63D G e IVS2(+4) C no éxon 2 do gene *HFE*, e concomitantemente, desequilíbrio desses alelos com alelo *HLA-B**44, aparentemente, sem associação com sobrecarga de ferro, mas com aparente origem histórica. Finalmente, o teste de diversidade encontrou diferenças apenas na amostra de HH e o teste de neutralidade não encontrou pressões seletivas na população controle do presente trabalho. Concluindo, este é o primeiro trabalho, avaliando grande segmento da região codificadora do gene *HFE* em diversas amostras de pacientes apresentando ou não sobrecarga de ferro e em indivíduos controle.

SUMMARY

Mutations at the *HFE* gene have been identified as a major factor for the development of iron overload, especially the classical single nucleotide polymorphism (SNP) C282Y and H63D. Primary iron overload can be directly attributed to mutations in the *HFE* gene (hereditary hemochromatosis - HH), whereas secondary overload may be acquired due to genetic factors and comorbidities, particularly infection with hepatitis C virus (HCV) and hepatocellular carcinoma (HCC). Linkage disequilibrium between major SNPs at the *HFE* gene with *HLA-A* and *HLA-B* alleles may contribute to alterations in the function of immune cells. We sequenced the coding region of the *HFE* gene (exon 2, 3, 4 and 5) of patients exhibiting primary or secondary iron overload. Thus, we studied 130 patients with hepatitis C, 60 patients with HCC, 14 patients with HH and 100 healthy individuals. We observed seven SNPs in the *HFE* gene in 5' to 3' sequence: i) H63D C>G at exon 2 (rs1799945); ii) IVS2(+4)T>C at intron 2 (rs2071303); iii) intron 3 C>G (rs807209); iv) C282Y G>A at exon 4 (rs1800562); v) intron 4 G>A (rs2794717); vi) IVS4(-44) T>C (rs1800708); vii) at intron 5 we detected a yet undescribed G>deletion. We performed haplotype reconstruction of the coding region and extended haplotypes comprising the *HFE* gene and ten other loci (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQB1*, *TNFA*, *TNFB*, *TNFC*, *TNFD* and *HLA-G*14bp). In addition, we normalized the nomenclature of the *HFE* gene, according to the rules of the International ImMunoGeneTics Information System - IMGT (<http://www.imgt.org/>). Several tools were used to evaluate the association between *HFE* polymorphic sites with iron overload, including tests assessing associations of alleles, genotypes, linkage disequilibrium, haplotype and diplotypes, diversity and neutrality tests. Finally, we compare the SNPs at the *HFE* gene observed in the healthy population with those observed in diverse worldwide populations available at the 1000genomes site (<http://www.1000genomes.org/>) the results showed that our population was closer to American and European populations, partially close the populations of African and distant of Asian populations. The C282Y allele G, contained in the coding region *HFE* * 01:03 haplotype, conferred susceptibility to HH in the Brazilian population tested, agreeing with the findings observed in other worldwide populations. The *HFE**01:02:01:01/*HFE** 01:01:01:05 diplotype conferred susceptibility to the development of iron overload in patients with HCC caused by HCV. A strong linkage disequilibrium was detected between the H63D G and IVS2 (+4) C alleles in exon 2 of the *HFE* gene, and concomitantly, a linkage between these alleles with *HLA-B**44, apparently not associated with iron overload, but with apparent historical origin. Finally, the test of diversity revealed differences only in the HH sample and the neutrality test detected no selective pressures on the control population of this study. In conclusion, this is the first study evaluating a large segment of the coding region of the *HFE* gene in several samples of patients exhibiting or not iron overload and in control subjects.

FIGURAS

- Figura 1:** Estrutura do gene *HFE*, no cromossomo (6p21.3), identificada de acordo com o sítio Ensembl (<http://www.ensembl.org>). Gene id: ENSG00000010704. Os SNPs estudados neste trabalho estão indicados conforme identificação do NCBI SNP Database(<http://www.ncbi.nlm.nih.gov/snp>).14
- Figura 2:** *HFE* é um heterodímero, possuindo uma cadeia pesada e uma cadeia leve não covalentemente ligada a $\beta 2$ - μ globulina. A cadeia pesada é constituída por três domínios extracelulares $\alpha 1$, $\alpha 2$, $\alpha 3$, uma região transmembrana e uma cauda citoplasmática. Os locais das mutações C282Y, H63D e S65C, comumente encontradas no gene *HFE* de pacientes com hemocromatose hereditária são representados. Asteriscos indicam os resíduos Valina100 e Triptofano103, que interagem com o receptor de transferrina (SALTER-CID; PETERSON; YANG, 2000).15
- Figura 3:** O enterócito e as proteínas envolvidas na absorção do ferro. Dcytb - ferredoxina; DMT-1 - transportador de metal divalente-1; HCP-1 - proteína transportadora do heme-1; NU - núcleo; *HFE* - proteína da hemocromatose; TfR - receptor (figura de GROTO, 2010).17
- Figura 4:** Quatorze transcritos do gene *HFE* são descritos, sendo que dois não são traduzidos em proteínas. Esquemas e numeração da isoforma estão ilustrados conforme descrito no sítio ENSEMBL (www.ensembl.org).21
- Figura 5:** Foto ilustrativa do padrão de amplificação do éxon 3 em gel de agarose 1%.42
- Figura 6:** Esquema que indica o significado de cada algarismo na nomenclatura dos alelos do gene *HFE*.46
- Figura 7:** Cromatograma do SNP novo no íntron 5 identificada em um indivíduo da amostra da população controle. A deleção apresenta-se em heterozigose, portanto as bases posteriores se apresentam embaralhadas no sequenciamento normal (A). Isto não acontece com a fita clonada (B).55
- Figura 8:** Desequilíbrio de ligação entre os SNPs do gene *HFE*, imagem gerada no software HAPLOVIEW v.4.1 (BARRETT et al., 2005). Áreas em vermelho escuro forte desequilíbrio de ligação ($LOD \geq 2$; $D' = 1$); em vermelho claro indica desequilíbrio moderado ($LOD \geq 2$, $D' < 1$); azul indica fraco desequilíbrio ($LOD < 2$, $D' = 1$); e branco indica a ausência de desequilíbrio ($LOD < 2$, $D' < 1$). Valores de D' diferentes de 1,00 são representados em porcentagem dentro dos quadrados. LOD (Log do Odds); D' (correlação entre pares de SNPs).57
- Figura 9:** Esquema das formações Haplótípicas possíveis entre os SNPs H63D, IVS2(+4)T>C e C282Y, nas amostras de população Controle e Hemocromatose Hereditária. Em vermelho destaca-se a combinação alélica preferencial e exclusiva que ocorre entre os 3 SNPs (H63D G / IVS2(+4) C / C282Y G).59
- Figura 10:** Gráfico mostrando a soma de loci (frequência alélica e genotípica) semelhantes e diferentes ao da amostra de população Controle segundo o teste exato de Fisher (setas indicando as superpopulações por continente).69
- Figura 11:** Esquema do desequilíbrio de ligação entre o gene *HFE* e o HLA-B.74

TABELAS

Tabela 1. Comparação dos resíduos da fenda para ligação em peptídeos do HLA-A2 e HFE (LEBRÓN et al., 1998).....	18
Tabela 2. Frequência dos alelos do C282Y, H63D e S65D do gene HFE em diferentes populações - Modificado de Trifa et al. (2012); Leone et al. (2005) e Merryweather-Clarke et al. (1997).	24
Tabela 3. Quantidade de indivíduos, média de idade e desvio padrão, quantidade de homens e porcentagem por etnia.....	40
Tabela 4. Descrição das regiões analisadas, das sequências dos primers, da temperatura de melting, do tamanho em pb do fragmento amplificado e da sequência fasta da região (primer sem letras maiúsculas e grifado, região codificadora em letras maiúsculas e sequên sequência fasta da região (primer sem letras maiúsculas e grifado, região codificadora em letras maiúsculas e sequência dos aminoácidos acima da região codificadora). SNPs indicados de acordo com o código IUPAC.	42
Tabela 5. Frequências alélicas, genotípicas e desequilíbrio de Hardy-Weinberg (P) em amostra de população Controle e de pacientes com hepatite C (“com” e “sem” sobrecarga de ferro), carcinoma hepatocelular (com sobrecarga e diverso) e pacientes com hemocromatose hereditária. Ordem dos alelos: Selvagem/Mutado, valores sublinhados apontam combinações com testes exatos de Fisher significativos.	53
Tabela 6. Desequilíbrio de ligação, somente os pares com valores significativos entre os SNPs do gene HFE, em cada uma das amostras populacionais e no teste global. Grifados o par recorrente em todas as populações.	56
Tabela 7. Desequilíbrio de ligação entre o SNPs H63D, (IVS2(+4)T>C) e C282Y do gene HFE nas amostras de populações controle e hemocromatose hereditária.	58
Tabela 8. Heterozigose observada e esperada nas amostras de populações (realizados apenas nos loci que se apresentam polimórficos em cada amostra).	60
Tabela 9. Matriz de distância das amostras de populações baseada nos valores de FST (diagonal inferior: valores de P – diagonal superior: erro padrão de cada respectivo valor P.	61
Tabela 10. Identificação e constituição haplotípica, para cada SNP nas populações controle e de pacientes (em negrito as bases mutadas; com fundo cinza as regiões exônicas).	62
Tabela 11. Frequência haplotípica do gene HFE por população, teste exato de Fisher (em negrito: as frequências dos haplótipos discrepantes em comparações com outras amostras; grifado: combinações com valores significativos no teste exato de Fisher).	64
Tabela 12. Código de identificação das subpopulações que possuem seus dados depositados no site 1000genome (http://www.1000genomes.org/).	65
Tabela 13. Teste exato Fisher entre a população controle e as populações disponíveis no site 1000genomes (http://www.1000genomes.org/) calculados para os loci encontrados nesse trabalho, sombreado valores significativos.	67
Tabela 14. Desequilíbrio de Hardy-Weinberg (em negrito valores significativos).	70
Tabela 15. Teste de desequilíbrio de ligação entre loci do CPH (em negrito, valores significativos; sombreado de cinza, comparações com o gene HFE).	70
Tabela 16. Desequilíbrio de ligação SNPs clássicos do HFE (H63D e C282Y); com loci do CPH (em negrito P valor significativo). Desequilíbrio de ligação SNPs clássicos do HFE (H63D e C282Y); com loci do CPH (em negrito P valor significativo).	72
Tabela 17. Desequilíbrio de ligação entre os alelos do gene HLA-B e do SNP H63D do gene HFE.	73
Tabela 18. Desequilíbrio de ligação entre os alelos do gene HLA-B e do SNP (IVS2(+4)T>C) do gene HFE....	74
Tabela 19. Lista de sinónimas dos sete SNPs encontrados nos sites: NCBI (http://www.ncbi.nlm.nih.gov/snp/) e 1000genomes (http://www.1000genomes.org/).	107

SUMÁRIO

I. INTRODUÇÃO	14
I.1 O Gene <i>HFE</i>	14
I.1.1 Estrutura do Gene	14
I.1.2 Estrutura da Proteína <i>HFE</i>	15
I.1.3 Polimorfismos no Gene <i>HFE</i>	19
I.1.4 Isoformas da Proteína <i>HFE</i>	20
I.1.5 Estudos Populacionais e Nomenclatura dos Alelos do Gene <i>HFE</i> .	23
I.1.6 Desequilíbrio de Ligação entre os Genes <i>HFE</i> e <i>HLA-A</i> , <i>HLA-B</i>	26
I.2 Sobrecarga de Ferro	28
I.3 Hemocromatose Hereditária	28
I.4 Hepatite C	30
I.4.1 <i>HFE</i> e Hepatite C	32
I.5 Carcinoma Hepatocelular	32
I.5.1 Carcinoma Hepatocelular e <i>HFE</i>	34
II. HIPÓTESE	35
III. JUSTIFICATIVA	36
IV. OBJETIVOS	36
V. MATERIAL E MÉTODOS	37
V.1 Desenho do estudo e local de execução do projeto	37
V.2 Aspectos éticos e origem das amostras	37
V.3 Casuística	38
V.4 Aspectos clínicos e laboratoriais considerados importantes para estratificação dos pacientes	40
V.4.1 Sobrecarga de ferro	40
V.4.2 Coleta e processamento dos materiais biológicos para o estudo	41
V.5 Diversidade do Gene <i>HFE</i>	41
V.6 Clonagem do novo SNP no íntron 5	44
V.6.1 Purificação do Produto de PCR.	44
V.6.2 Reação de ligação do fragmento ao plasmídeo:	44
V.6.3 Plaqueamento das bactérias transformadas:	44
V.6.4 MINI PREP	45
V.7 Nomenclatura dos Alelos do gene <i>HFE</i>	45
VI. ANÁLISES ESTATÍSTICAS	46
VI.1 Estimativa das frequências alélicas e genotípicas	46
VI.2 Aderências ao equilíbrio de Hardy-Weinberg	47
VI.3 Desequilíbrio de Ligação	47
VI.4 Diversidade e Diferenciação Genética da Região Codificadora do gene <i>HFE</i>	48
VI.5 Teste Exato de Fisher	49
VI.6 Diversidade Haplotípica	49

VI.7 Teste exato de diferenciação populacional	50
VI.8 Testes de Neutralidade	50
VII. RESULTADOS	51
VII.1 Descrição das frequências alélicas, genotípicas, heterozigose, equilíbrio de Hardy-Weinberg e Teste Exato de Fisher.	51
VII.3 Desequilíbrio de Ligação	55
VII.4 Diversidade e Diferenciação Genética da Região Codificadora do gene <i>HFE</i>	59
VII.5 Dedução da Fase Meiótica dos Haplótipos no Gene <i>HFE</i>	61
VII.6 Teste de Neutralidade	62
VII.7 Descrição das frequências haplotípicas e teste Exato de Fisher para haplótipo e diplótipos do gene <i>HFE</i> .	63
VII.8 Teste exato de Fisher dos SNPs encontrados no trabalho comparado com diversas populações	65
VII.9 Equilíbrio de Hardy-Weinberg e Desequilíbrio de Ligação do Haplótipo Estendido	70
VIII. DISCUSSÃO	75
VIII.1 Frequências alélicas, genotípicas, equilíbrio de Hardy-Weinberg e heterozigose	75
VIII.2 Desequilíbrio de Ligação	78
VIII.3 Diversidade e Diferenciação Genética da Região Codificadora do gene <i>HFE</i>	86
VIII.4 Haplótipo de Fase Meiótica e Teste de Neutralidade	86
VIII.5 Descrição das frequências haplotípicas, teste exato de Fisher para haplótipo e diplótipos da região codificadora do gene <i>HFE</i> .	87
VIII.6 Teste exato de Fisher dos SNPs do gene <i>HFE</i> encontrados no trabalho, comparado com diversas populações	88
VIII.7 Hardy-Weinberg e Desequilíbrio de Ligação do Haplótipo Estendido	90
IX. CONCLUSÕES	93
X. REFERÊNCIAS BIBLIOGRÁFICAS	94
XI. APÊNDICE	107
XII. ANEXO	109

Figura 1: Estrutura do gene *HFE*, no cromossomo (6p21.3), identificada de acordo com o sítio Ensembl (<http://www.ensembl.org>). Gene id: ENSG0000010704. Os SNPs estudados neste trabalho estão indicados conforme identificação do NCBI SNP Database (<http://www.ncbi.nlm.nih.gov/snp>).

I. INTRODUÇÃO

I.1 O Gene *HFE*

I.1.1 Estrutura do Gene

O gene *HFE* (nomeado pela abreviação de *High Fe*, segundo o site HUGO Gene Nomenclature Committee – HGNC – <http://www.genenames.org>), inicialmente designado *HLA-H*, em 1996 (FEDER *et al.*, 1996), está localizado no braço curto do cromossomo 6 (6p21.3), na região telomérica do complexo principal de histocompatibilidade (CPH) estendido, fazendo parte dos genes *HLA* de classe I não-clássicos (HORTON *et al.*, 2004).

O gene possui 7952 pares de bases e contém sete éxons (Figura 1), sendo o sétimo não codificante, localizado dentro da região 3' não traduzida (3'NT). Os éxons 2, 3 e 4 codificam os domínios $\alpha 1$, $\alpha 2$ e $\alpha 3$ da molécula, respectivamente. O tamanho dos sete íntrons varia de 158 pares de bases (pb) (íntron4) até aproximadamente 3335 pares de base (íntron 1). Na região promotora, entre os 448 pb das regiões reguladoras putativas, observa-se uma região relativamente rica em GC (64% das bases de 1-78) e a ausência das caixas TATA e CAAT. Assim, o promotor do gene *HFE* humano tem características estruturais em comum com vários genes *housekeeping* e é consistente com o padrão de expressão do onipresente RNA mensageiro (RNAm). Além do aumento observado do conteúdo GC, 80-90 pb a montante do códon de iniciação, nenhuma homologia significativa da sequência reguladora putativa do gene *HFE* foi encontrada

entre as regiões promotoras de outros genes *HLA* de classe I clássicos (*HLA-A/B/C*). Uma caracterização mais precisa do promotor, pela análise funcional de gene repórter e a identificação de fatores de transcrição, pode fornecer pistas sobre o controle da expressão do gene *HFE* (FEDER *et al.*, 1996).

I.1.2 Estrutura da Proteína HFE

O gene *HFE* codifica uma molécula de 343 aminoácidos homóloga às proteínas *HLA* de classe I, apresentando cadeia α , constituída de três domínios extracelulares ($\alpha 1$, $\alpha 2$ e $\alpha 3$); sendo o domínio $\alpha 3$ ligado de forma não covalente à β_2 - μ globulina (codificada em outro gene, no cromossomo 15), além de uma região transmembrana e uma cauda citoplasmática (Figura 2). A molécula contém quatro resíduos de Cisteína altamente conservados que formam duas pontes dissulfeto intramoleculares, uma no domínio $\alpha 2$ (C225) e outra no $\alpha 3$ (C282) (DUPRADEAU *et al.*, 2003). Esses domínios são essenciais para a estrutura terciária da proteína, que gera a interação não covalente com a β_2 - μ globulina e por consequência sua exposição na superfície celular (POINTON *et al.*, 2000).

Figura 2: HFE é um heterodímero, possuindo uma cadeia pesada e uma cadeia leve não covalentemente ligada a β_2 - μ globulina. A cadeia pesada é constituída por três domínios extracelulares $\alpha 1$, $\alpha 2$, $\alpha 3$, uma região transmembrana e uma cauda citoplasmática. Os locais das mutações C282Y, H63D e S65C, comumente encontradas no gene *HFE* de pacientes com hemocromatose hereditária são representados. Asteriscos indicam os resíduos Valina100 e Triptofano103, que interagem com o receptor de transferrina (SALTER-CID; PETERSON; YANG, 2000).

A proteína codificada pelo gene *HFE* está fortemente relacionada com a regulação da absorção intestinal do ferro. A hepcidina, um hormônio peptídico que tem papel regulatório fundamental na homeostase do ferro, coordena o uso e o estoque do ferro com a sua aquisição (KRAUSE *et al.*, 2000; PARK *et al.*, 2001). Portanto, as alterações estruturais da hepcidina resultam em espectro de doenças de sobrecarga de ferro, incluindo as causas primárias, como a hemocromatose, e as secundárias, como a anemia da inflamação, outras anemias associadas à sobrecarga de

ferro, as hepatites virais, entre outras. A expressão da hepcidina é influenciada por estímulos sistêmicos tais como, reserva de ferro, nível da eritropoese, inflamação, hipóxia e estresse oxidativo. Esses estímulos controlam os níveis de hepcidina, atuando em proteínas existentes na superfície celular dos hepatócitos tais como: HFE, receptor 2 da transferrina (TfR2),

hemojuvelina, matriptase-2 e IL-6R (DARSHAN; ANDERSON, 2009). As proteínas de superfície ativam diversas cascatas de transdução de sinal, cujos elementos participantes são em grande parte ainda desconhecidos, de maneira a alterar (estimular ou reprimir) a transcrição do gene que codifica a hepcidina.

O papel da proteína HFE no metabolismo do ferro é ainda pouco conhecido. Sabe-se que a HFE interage com o receptor 1 da transferrina (TfR1) e especula-se que essa interação sirva para sequestrar HFE, impedindo-a, assim, de participar da cascata de sinalização conducente à estimulação da transcrição de gene da hepcidina (SCHMIDT *et al.*, 2008). A descoberta recente que a proteína HFE também pode interagir com o receptor de transferrina 2 (TfR2) indica que o complexo HFE-TfR2 pode estar envolvido na estimulação da hepcidina. Ainda, foi sugerido que a interação da holotransferrina com o receptor de transferrina 1 (TfR1) libera a HFE que fica disponível para interagir com o receptor de transferrina 2 (TfR2), o que facilitaria a transdução de sinal para a hepcidina (GAO *et al.*, 2009). Se algum dos elementos HFE ou receptor de transferrina 2 (TfR2) se encontrar mutado ou ausente, o complexo ficará incapaz de detectar o aumento da saturação da transferrina sérica, ocorrendo desregulação da homeostase do ferro.

O receptor de transferrina (TfR) tem papel central; sendo que ferro sérico ligado à transferrina (Tf) se acopla à superfície celular por meio do seu receptor (TfR) que, em seguida, sofre endocitose do complexo (TfR/Tf), e no ambiente ácido do endossomo, o ferro é liberado da transferrina (Tf) (RICHARDSON; PONKA, 1997; WEST *et al.*, 2001). O ferro livre atravessa a membrana endossomal e entra no *pool* de quelatos intracelular, aonde é utilizado para necessidades metabólicas celulares ou é incorporado e estocado dentro das proteínas de ferritina. A transferrina livre (Tf) de ferro (apo-Tf) remanescente se liga ao receptor de transferrina (TfR), em baixo pH, em vesículas ácidas ($\text{pH} \leq 6,5$) e o complexo TfR/apo-Tf é reciclado para a superfície da célula, onde apo-Tf se dissocia do complexo no pH mais alto do sangue ($\text{pH} \geq 7,4$) (DE SILVA; ASKWITH; KAPLAN, 1996; RICHARDSON; PONKA, 1997; WEST *et al.*, 2001) (Figura 3).

Figura 3: O enterócito e as proteínas envolvidas na absorção do ferro. Dcytb - ferredutase; DMT-1 - transportador de metal divalente-1; HCP-1 - proteína transportadora do heme-1; NU - núcleo; HFE - proteína da hemocromatose; TfR - receptor (figura de GROTO, 2010).

Estudos de cristalização demonstram homologia da molécula HFE com as moléculas HLA de classe I, pois ambas apresentam um superdomínio $\alpha 1$ - $\alpha 2$ (1-180 resíduos), formando uma plataforma que comporta oito fitas β antiparalelas sustentadas pelas duas α hélices, que geralmente são posicionadas no topo de moléculas do tipo imunoglobulina, por um domínio constante $\alpha 3$ (181-275 resíduos) e pela β_2 - μ globulina (BENNETT; LEBRÓN; BJORKMAN, 2000). Porém, a estrutura do HFE tem uma fenda muito estreita, que impede sua ligação com peptídeos, devido a uma translação da $\alpha 1$ hélice, trazendo-a ~ 4 Å mais próximo da hélice $\alpha 2$, que tem sua posição praticamente idêntica àquelas de uma molécula HLA de classe I clássica (LEBRÓN *et al.*, 1998; RAMMENSEE; FALK; RÖTZSCHKE, 1993).

Outra evidência de que a HFE não apresenta uma fenda de apresentação de peptídeo, vem do fato de que os *pockets* de interação com peptídeos se encontram “enterrados” na molécula, o que impede a ligação de peptídeos com HFE (LEBRÓN *et al.*, 1998). Embora

estreita, a HFE forma *pockets* análogos aos de uma molécula classe I clássica, porém posicionados de forma interna. Os *pockets* A e F, que se ligam ao N e C terminal dos peptídeos são resíduos mais conservados, enquanto os *pockets* intermediários, de B a E, que interagem com as cadeias laterais, contêm resíduos que mais variabilidade (LEBRÓN *et al.*, 1998; MATSUMURA *et al.*, 1992; SAPER; BJORKMAN; WILEY, 1991). Porém, quase metade dos resíduos HFE localizados em posições análogas aos *pockets* de A a F das moléculas classe I estão enterradas, impedindo que haja interação com peptídeo (Tabela 1) (LEBRÓN *et al.*, 1998). Os resíduos de Tirosina (7, 59, 159 e 171), formadores do *pocket* A e o Triptofano 167, resíduo da borda da fenda, nas moléculas HLA de classe I clássicas, interagem com o N terminal do peptídeo. Na molécula HFE apenas duas das Tirosinas são conservadas (HFE Tyr10 e Tyr160) e a Tirosina 10 é “enterrada”. Além disso, a cadeia lateral do HFE Glicina 168 (análogo ao resíduo 167 do CPH de classe I) se direciona para a fenda ocluindo o *pocket* A (LEBRÓN *et al.*, 1998).

Tabela 1. Comparação dos resíduos da fenda para ligação em peptídeos do HLA-A2 e HFE (LEBRÓN *et al.*, 1998)

HLA-A2	<i>Pocket</i>	HFE	<i>Pocket</i>	Interage c/ peptídeos
Metionina 5	A	Leucina 8	“Enterrado”	Não
Tirosina 7	A, B	Tirosina 10	“Enterrado”	Não
Fenilalanina 9	B, C	Fenilalanina 12	“Enterrado”	Não
Metionina 45	B	Valina 46	B	Não
Tirosina 59	A	Metionina 61	A	Não
Ác. Glutâmico 63	A, B	Leucina 65	A, B	Não
Lisina 66	A	Serina 68	A	Sim
Valina 67	B	Leucina 69	B	Sim
Histidina 70	B, C	Triptófano 72	B, C	Sim
Treonina 73	C	Metionina 75	C	Sim
Valina 76	Não localizado em <i>pocket</i>	Valina 78	Não localizado em <i>pocket</i>	Sim
Ác. Aspártico 77	F	Asparagina 79	“Enterrado”	Sim
Treonina 80	F	Treonina 82	F	Sim
Leucina 81	F	Isoleucina 83	“Enterrado”	Sim
Arginina 97	C, E	Valina 98	“Enterrado”	Não
Tirosina 99	A, B, C, D	Leucina 100	“Enterrado”	Não
Histidina 114	C, D, E	Triptófano 114	C, D, E	Não
Tirosina 116	C, F	Tirosina 116	“Enterrado”	Não
Tirosina 118	F	Tirosina 188	F	Não
Tirosina 123	F	Histidina 123	“Enterrado”	Não
Isoleucina 124	F	Leucina 124	“Enterrado”	Não
Treonina 143	F	Treonina 143	“Enterrado”	Não
Triptofano 147	E, F	Triptófano 147	E, F	Não
Valina 152	E	Arginina 153	E	Sim
Leucina 156	D, E	Asparagina 157	D, E	Não

HLA-A2	Pocket	HFE	Pocket	Interage c/ peptídeos
Tirosina 159	A, D	Tirosina 160	A, D	Não
Triptofano 167	A	Glutamina 168	A	Sim
Tirosina 171	A	Leucina 172	A	Não

Outro fator que evidencia que a fenda do HFE não é usada para apresentação de peptídeos é o fato de que a plataforma $\alpha 1$ - $\alpha 2$ e o domínio helicoidal da transferrina são locais de contato intramolecular primário entre as duas moléculas (BENNETT *et al.*, 2000). Os domínios $\alpha 3$ e o $\beta 2$ - μ globulina da HFE não entram em contato com o receptor de transferrina (TfR). (BENNETT *et al.*, 2000). Esse contato é importante para regulação do ferro, já que estudo de WEST *et al.* (2001) demonstrou que transferrina (Tf) e HFE competem por esse domínio, sugerindo que Tf e HFE se liguem no mesmo local no receptor de transferrina (TfR).

A hélice 1 do receptor de transferrina (TfR), que é mais curta do que as outras 2 hélices da estrutura, interage com a região C terminal da hélice $\alpha 1$ da proteína HFE, em arranjo paralelo. Essa hélice curta pareada tem um centro hidrofóbico constituído de Leucina 619 e Valina 622 que parecia com uma Valina 78 e um Triptofano 81 no HFE (BENNETT *et al.*, 2000).

I.1.3 Polimorfismos no Gene *HFE*

O gene *HFE* já tem várias mutações descritas, incluindo: i) éxon 2: V53M, V59M, H63H; ii) éxon3: Q127H, E168Q; iii) éxon 4: Q283P, P168X, , E168X e W169X), porém as três mais pesquisadas são: C282Y (éxon 4), H63D (éxon 2) e S65C (éxon 2) (ANDERSON *et al.*, 2004; FRANCHINI; VENERI, 2005; POINTON *et al.*, 2000).

A mutação C282Y caracteriza-se por apresentar transição de G para A no nucleotídeo 845 do gene, substituindo uma Tirosina por uma Cisteína no aminoácido 282 da proteína, que interrompe a ponte dissulfeto do domínio $\alpha 3$, prejudicando a interação com a $\beta 2$ -microglobulina, requerida para a sua expressão nas superfícies das células (FEDER *et al.*, 1996). Normalmente, o complexo HFE/ $\beta 2$ - μ globulina se liga ao receptor de transferrina (TfR) e estudos funcionais sugerem que a HFE interage com o TfR na superfície da célula e concorre com a transferrina para a ligação com TfR (ANDERSON *et al.*, 2004; DISTANTE *et al.*, 2004; GRIFFITHS, 2007). Cerca de 85% dos pacientes de hemocromatose hereditária são homozigotos para essa mutação e sua frequência na população caucasiana de europeus varia entre 5 a 10% (LUCOTTE, 2001).

A mutação H63D é originada pela transversão de C para G no nucleotídeo 187, determinando a substituição de Ácido aspártico por Histidina no aminoácido 63. A HFE com a substituição H63D interage com a β_2 -microglobulina e é expressa normalmente na superfície da célula. Evidências recentes sugerem que a sua função na modulação da interação da transferrina com o receptor (TfR) pode ser prejudicada (FEDER *et al.*, 1998). A frequência alélica da mutação H63D é de 16% na população europeia que apresenta expressão fenotípica branda para hemocromatose (LYON; FRANK, 2001). A substituição de Cisteína por Serina na posição 65 (S65C) também tem sido apontada na forma branda de hemocromatose hereditária (HANSON; IMPERATORE; BURKE, 2001; MURA; RAGUENES; FÉREC, 1999).

Embora a maioria dos pacientes com diagnóstico de hemocromatose hereditária seja homozigotos para a mutação C282Y no gene *HFE*, os pacientes com os demais genótipos (C282Y/H63D, H63D/H63D, H63D/WT, C282Y/WT, S65C/S65C, S65C/WT) podem ser considerados portadores de hemocromatose, desde que seja constatada a presença de sobrecarga de ferro (PIETRANGELO, 2004). Entre 4% e 21% dos pacientes clinicamente afetados não possuem essas duas mutações (ADAMS *et al.*, 2003; BARTON *et al.*, 1997; BEUTLER, *et al.*, 1996; BOROT *et al.*, 1997; CARDOSO, *et al.*, 1998; CARELLA *et al.*, 1997; FEDER *et al.*, 1996; JOUANOLLE *et al.*, 1997; MURA *et al.*, 1999; WORWOOD *et al.*, 1997).

I.1.4 Isoformas da Proteína HFE

É conhecido que a diversidade de isoformas das proteínas CPH é parcialmente alcançada por intermédio de edições alternativas do pré-RNAm (FUJII; ISHITANI; GERAGHTY, 1994; ISHITANI; GERAGHTY, 1992; KIRSZENBAUM *et al.*, 1994; MIKHAILOVA *et al.*, 2010). Isto não é diferente para o HFE, que também apresenta suas isoformas. Até o momento, já foram relatados 14 isoformas diferentes (Figura 4), porém dois deles não resultam em proteína (ENSEMBL-<http://www.ensembl.org>).

A maioria dos estudos do gene *HFE* foca suas análises nos SNPs C282Y (rs1800562), H63D (rs1799945), S65C (rs1800730) (FEDER *et al.*, 1996; HANSON *et al.*, 2001; MURA *et al.*, 1999), dos quais somente o primeiro altera a estrutura da proteína e sua associação com a β_2 - μ globulina, gerando assim, uma versão solúvel. Ou seja, esse SNP impede a formação do domínio transmembrana e a exposição exocelular, resultando em uma proteína que mantém sua localização difusa no citoplasma (MARTINS *et al.*, 2011).

Figura 4: Quatorze transcritos do gene *HFE* são descritos, sendo que dois não são traduzidos em proteínas. Esquemas e numeração da isoforma estão ilustrados conforme descrito no sítio ENSEMBL (www.ensembl.org).

Além do SNP C282Y, outros SNPs geradores de isoformas são relatados. O SNP IVS2(+4)T>C (rs2071303) está localizado na região de junção e afeta a escolha do local de edição alternativa na região 5' (5'ss) durante a montagem do *spliceosome*. Acredita-se que essa nova edição resultante da SNP IVS2(+4) C, produza proteínas com o íntron 5 ausente (CARMEL *et al.*, 2004; LUND; KJEMS, 2002; MIKHAILOVA *et al.*, 2010; NAGAI *et al.*, 2001; ROGOZIN; MILANESI, 1997). O SNP IVS4(-44)T>C (rs1800708), está localizado no íntron 4 que possui 158 pb e contem blocos de nucleotídeos de purina em grande quantidade. Esses blocos frequentemente ocorrem em locais de ligação de fatores de edição, sugerindo que o íntron 4 tenha alguma função regulatória (MCCULLOUGH; BERGET, 1997; MIKHAILOVA *et al.*, 2010; SOREK; AST, 2003; YEO; VAN NOSTRAND; LIANG, 2007). Galante *et al.* (2004) sugerem que o SNP IVS4(-44) T>C possa reter o íntron 4 em uma forma de edição alternativa. O RNAm com o inserto do íntron 4 pode gerar uma forma solúvel truncada da proteína HFE, isto devido a um código de parada cinco nucleotídeos a jusante do limite do éxon 4. Essa proteína conserva *reading frame* normal do gene e sua expressão é observada em tecidos hepáticos (JEFFREY *et al.*, 1999; MIKHAILOVA *et al.*, 2010).

Outro SNP formador de proteína solúvel é o IVS5(-47)G>A (rs1572982), localizado dentro do bloco de nucleotídeos -GCAAGATG(G/A)TG-, que é homólogo ao segmento de domínio de interação U4 no snRNA U6, -GCAAGGAUGA-. Além disso, esse bloco é complementar ao fragmento snRNA U2, -CAUCAGUUU, localizado a uma distância de seis nucleotídeos a jusante do segmento U2 – GUAGUA e que, geralmente, interage com o *branchpoint* nos íntrons (GUTHRIE; PATTERSON, 1988; MIKHAILOVA *et al.*, 2010; NAGAI *et al.*, 2001). O *branchpoint* consiste em uma Adenina de 18 a 40 nucleotídeos a jusante do sitio 3', utilizado para clivagem. Essa clivagem é realizada em três passos: i) o sitio 5' é clivado até o *branchpoint*, ii) o sitio 3' é clivado até o *branchpoint* e iii) os dois exóons são clivados (RUSKIN; ZAMORE; GREEN, 1988).

Martins *et al.* (2011) comentam que na proteína HFE, as principais variações nos processos de edição alternativa ocorrem nos éxons 2 e 4. A quantificação relativa dos transcritos das isoformas, produzidas a partir dessas variações, mostra expressão diferencial nos diversos tecidos, sendo que as isoformas curtas são tecido-específicas. Em linhas gerais, a ausência do éxon 2 produzirá uma proteína sem o domínio $\alpha 1$, que é essencial para interação com o receptor de transferrina 1 (ITfR1). Em contraste, a permanência do íntron 4 corresponderá a uma proteína com os três domínios α requeridos para a interação com TfR1 e

com o β_2 - μ globulina, no entanto, essa isoforma não contém a região transmembrana e a cauda citoplasmática.

A presença de edição alternativa e de proteínas solúveis pode ter importante papel em processos regulatórios. Algumas variantes de RNAm para *HFE*, ainda não estudados em nível proteico, sugerem que os peptídeos solúveis gerados a partir dessas variantes possam regular o transporte de ferro celular (JEFFREY *et al.*, 1999; LAHAM *et al.*, 2004; MARTINS *et al.*, 2011). Ensaios utilizando dímero HFE- β_2 - μ globulina artificial, mimetizando a proteína HFE solúvel, verificaram a efetiva redução na absorção de transferrina (Tf) pelas células (LAHAM *et al.*, 2004; MARTINS *et al.*, 2011). Porém, esse processo não está correlacionado com o receptor de transferrina (TFR) ou com a síntese de ferritina, contrastando com a indução normal do HFE. Isto sugere que as proteínas HFE solúveis possam ter função diferenciada ainda não desvendada (MARTINS *et al.*, 2011).

I.1.5 Estudos Populacionais e Nomenclatura dos Alelos do Gene *HFE*.

Apesar de o gene *HFE* ser estruturalmente análogo aos genes *HLA* de classe I, não apresenta sistema de nomenclatura de alelos padronizado, similar à nomenclatura dos outros genes do CPH. Provavelmente, a ausência da padronização da nomenclatura advém do fato de que a maioria dos trabalhos com *HFE* se limitem a tipificar os três SNPs principais (C282Y, H63D, e em menor frequência o S65C) por metodologias pontuais, enzimas de restrição ou discriminação alélica por PCR em tempo real, descrevendo sua frequência populacional e uma possível correlação com morbidades (hepatite, carcinoma hepatocelular, leucemia, porfiria, entre outras), negligenciando outros SNPs que contidos ao longo do gene e que apresentem possível relevância clínica (exemplos: DORAK; BURNETT; WORWOOD, 2005; ERHARDT *et al.*, 2003; MARTINELLI *et al.*, 2000; STUART *et al.*, 1998; TUNG *et al.*, 2003).

Também são comuns, em diversas populações do mundo, os estudos de *screening* para esses SNPs (geralmente feitos com doadores de sangue). Na Europa, a distribuição desses SNPs se apresenta da seguinte forma: O C282Y decresce do norte para o sul, com a menor frequência observada no Sudeste da Europa, o H63D apresenta frequência alta (de 10% a 20%) e o S65C apresenta frequência baixa (1% a 2%) na maioria das populações europeias (TRIFA *et al.*, 2012). A razão para o padrão de distribuição do SNP C282Y vem da hipotética origem Céltica ou Viking dessa mutação (LUCOTTE; DIETERLEN, 2003; SIMON *et al.*, 1980).

Levando em consideração a formação da população brasileira é consequência de cinco séculos de mistura entre povos advindos de três continentes, ou seja, de colonizadores europeus, principalmente portugueses, escravos africanos e ameríndios. Tendo que, primeiramente, chegaram 500.000 portugueses e quatro milhões de africanos que dividiram espaço na nova terra com 2,4 milhões de ameríndios então existentes (CURTIN, 1972; IBGE, 2000a). No período seguinte, até o presente, chegaram oficialmente no Brasil mais de cinco milhões de imigrantes portugueses, espanhóis, alemães, sírios, libaneses, japoneses e outros que não ultrapassaram 30% desse total (IBGE, 2000a; PARRA *et al.*, 2003). As migrações interregionais em massa durante o período colonial do Brasil promoveram a dispersão e a miscigenação desses três componentes ancestrais fundadores em diferentes proporções ao longo de um país que possui dimensões continentais. Em decorrência disso, a população brasileira atual é considerada trihíbrida, embora as proporções dos três grandes componentes ancestrais variem consideravelmente conforme a região geográfica. Diferenças de padrão na distribuição das frequências desses SNPs já foram observadas em amostras populacionais brasileiras (BUENO; DUCH; FIGUEIREDO, 2006; SANTOS *et al.*, 2010; TORRES *et al.*, 2008), o que reforça a importância de estudos mais abrangentes do gene *HFE* em diversos grupos de populações brasileiras (Tabela 2).

Tabela 2. Frequência dos alelos do C282Y, H63D e S65D do gene *HFE* em diferentes populações - Modificado de Trifa *et al.* (2012); Leone *et al.* (2005) e Merryweather-Clarke *et al.* (1997).

População	Tamanho da Amostra	Frequência de alelo (%)		
		C282Y	H63D	S65D
Europa				
Reino Unido ¹	368	6,0	12,1	-
Irlanda ¹	45	10	18,9	-
Islândia ¹	90	6,7	10,6	-
Suécia ^{13,17}	250	3,9	12,5	1,6
Noruega ¹	94	6,4	9,5	-
Finlândia ¹	38	0	11,8	-
Holanda ¹	39	2,6	29,5	-
Dinamarca ^{2*,18,19}	876	7	12,1	1,5
Suíça ^{2,}	502	6,1	12,4	-
França ^{2*,13,20}	10.395	7,1	16,7	2,5
Alemanha ^{2*}	425	3,8	13,2	-
Polônia ³	871	3,1	16,2	-
República Tcheca ⁴	481	3,4	14,9	1,25
Áustria ⁵	271	3,7	12,9	-
Hungria ^{2*}	1719	3,4	12,3	-
Eslovênia ⁶	200	4	14,5	0,5
Croácia ⁶	200	3,3	14,5	1,8
Sérvia e Montenegro ⁷	318	1,6	15,7	1,6
Bulgária ⁸	100	0	23	-
Espanha ^{2**}	1,194	3,2	20,6	-
Itália ^{9,13,21}	500	1,5	1,4	1,3

População	Tamanho da Amostra	Frequência de alelo (%)		
		C282Y	H63D	S65D
Grécia ¹	196	1,3	13,5	-
Romênia ¹⁰ (região da Moldávia)	200	1,75	13,25	-
Itália ¹	91	6,9	19	-
Grécia ¹	196	1,3	13,5	-
Espanha ¹	78	3,2	26,3	-
Turquia ¹	70	0	13,6	-
Portugal ^{11,12}	640	4,36	21,76	-
Oriente Médio e África				
Arábia Saudita ¹	118	0	8,5	-
Gâmbia ¹	39	0	1,3	-
Senegal ¹	130	0	0	-
Quênia ¹	78	0	1,3	-
Nigéria ¹	80	0	2,6	-
Zâmbia ¹	76	0	0,7	-
Ásia e Sub-contidente Indiano				
Índia ¹	109	0,2	8,4	-
Hong Kong ¹	72	0	2,8	-
Taiwan ¹ (aborígenes)	80	0	2,8	-
Indonésia ¹	90	0	2,8	-
Oceania				
Papua Nova Guiné ¹	139	0	0	-
Austrália ¹ (aborígenes)	93	0	0,2	-
América				
Mexicanos ¹	54	0	6,5	-
Jamaicanos ¹	90	1,1	2,2	-
Colombianos ¹	47	0	0	-
EUA ¹³	3.353	6,7	13,35	-
Ilhas Vancouver ¹ (aborígenes)	37	0	0	-
Equador ²²	100	0	3,5	4,0
Brasil	São Paulo ^{14,15}	148	1,4	10
		542	2,1	13,6
	Região Norte ¹⁶	400	0,6 ^{??}	9

¹(MERRYWEATHER-CLARKE *et al.*, 1997)– ²(MERRYWEATHER-CLARKE *et al.*, 2000)– ³(MOCZULSKI; GRZESZCZAK; GAWLIK, 2001)– ⁴(ČIMBUROVÁ *et al.*, 2005)– ⁵(DATZ *et al.*, 1997)– ⁶(RISTIĆ *et al.*, 2003)– ⁷(ŠARIĆ *et al.*, 2006)– ⁸(IVANOVA *et al.*, 2003)– ⁹(PIETRAPERTOSA *et al.*, 2003)– ¹⁰(VOICU *et al.*, 2009), ¹¹(LUCOTTE; DIETERLEN, 2003), ¹²(CARDOSO *et al.*, 2001)– ¹³(HANSON *et al.*, 2001)– ¹⁴(BUENO *et al.*, 2006)– ¹⁵(SANTOS *et al.*, 2010)– ¹⁶(TORRES *et al.*, 2008)– ¹⁷(HOLMSTRÖM *et al.*, 2002)– ¹⁸(MILMAN *et al.*, 2005)– ¹⁹(MILMAN *et al.*, 2004)– ²⁰(MURA *et al.*, 1999)– ²¹(MARIANI *et al.*, 2003)– ²²(LEONE *et al.*, 2005).

^{??} TORRES *et al.* (2008) usaram duas amostras de populações de 400 indivíduos cada, controles e pacientes com malária, para realização dos cálculos estatísticos. Nestas análises foram encontrados apenas 5 indivíduos heterozigotos para C282Y, no estado do Pará, e nenhum homozigoto para esse locus, porém não foi especificado em qual das amostras de populações.

*Dados provenientes de 4 estudos - ** Dados provenientes de 8 estudos.

I.1.6 Desequilíbrio de Ligação entre os Genes *HFE* e *HLA-A*, *HLA-B*

Apesar de o gene *HFE* estar localizado aproximadamente 4Mb do *HLA-A*, desequilíbrios de ligação são relatados entre o gene *HFE* e os genes *HLA-A* e *HLA-B*. Autores apontam desequilíbrio de ligação entre os SNPs C282Y e H63D com alelos dos loci *HLA-A* e *HLA-B*, principalmente C282Y, com o alelo *HLA-A**03 e H63D, com o alelo *HLA-A**29 (AJIOKA *et al.*, 1997; ALBERT; BAUR; MAYR, 1984; BARTON *et al.*, 2005a; CARDOSO *et al.*, 2002; CARDOSO; DE SOUSA, 2003; DE SOUSA, 1989; MCLAREN *et al.*, 2003; PORTO *et al.*, 1998; SALTER-CID *et al.*, 2000; SOUSA, 1992).

O desequilíbrio de ligação entre o SNP C282Y e os alelos *HLA-A**03 já foi observado em diversas populações (Irlanda, Britânia, Dinamarca, Suécia, Alemanha, Portugal, Itália, Alabama e Utah – EUA) (ALBERT *et al.*, 1984; BARTON; ACTON, 2002; BARTON *et al.*, 2005a; BRISSOT *et al.*, 1999; BULAJ *et al.*, 1996; CARDOSO *et al.*, 1998; CARELLA *et al.*, 1997; DATI *et al.*, 1996; GOTTSCHALK *et al.*, 1998; KNAU; ROSSNER; SCHMIDT, 1991; MAYR, 1977; MILMAN *et al.*, 1988; MURPHY *et al.*, 1998; NIELSEN *et al.*, 1998; OLSSON *et al.*, 1997; PORTO *et al.*, 1998; PORTO *et al.*, 1989; POWELL *et al.*, 1988; RITTER; SAFWENBERG; OLSSON, 1984; SIMON *et al.*, 1987a; SIMON *et al.*, 1988).

Na população portuguesa além da usual associação do SNP C282Y com o alelo *HLA-A**03, foram encontradas associações entre o SNP C282Y com os alelos *HLA-A**29, *HLA-B**07, *B**14 e *B**44 (BARTON; ACTON, 2002; CARVALHO, 1983; PORTO *et al.*, 1998). Porto *et al.* (1998) descreveram também uma associação entre o SNP H63D e o alelo *HLA-A**29.

Um dos haplótipos mais associados à hemocromatose hereditária é o C282Y *HLA-A**03-*B**07, cuja maior frequência está em países do noroeste europeu (principalmente na Irlanda, Bretanha, Dinamarca, Suécia e Alemanha), além da região central do Alabama, Austrália e Itália (em populações caucasianas) (BARTON *et al.*, 2005a). A origem e disseminação dessa associação são atribuídas aos Vikings e suas migrações (BARTON *et al.*, 2005a; CARDOSO; DE SOUSA, 2003). Estudo usando frequências do SNP C282Y e o haplótipo *HLA-A**03-*B**07 possibilitou, além da associação desse específico haplótipo com todas as populações europeias, a validação dos cálculos de estimativa do surgimento da mutação C282Y (DISTANTE *et al.*, 2004)

Em pacientes com hemocromatose hereditária do Alabama (EUA) e Itália foi reportado aumento da frequência absoluta do haplótipo C282Y *HLA-A**03-*B**35. No

Alabama, também foi reportado elevação na frequência absoluta dos haplótipos *HLA-A*01-B*60*, *A*02-B*39*, *A*02-B*62*, *A*03-B*13*, *A*03-B*15*, *A*03-B*27*, *A*03-B*44*, *A*03-B*47* e *A*03-B*57* (BARTON; ACTON, 2002).

Estudo realizado em linfócitos de indivíduos normais e em populações de pacientes homozigotos C282Y mostrou anormalidades no número e proporções de células CD4+, CD8+, CD8+ CD28+ e número de linfócitos totais, como também anormalidades do repertório dos receptores de células T do CD8+. Essas alterações são associadas ao desequilíbrio entre o SNP C282Y e o alelo *HLA-A*03* e *A*29*. Em pacientes com depósitos de ferro mais grave, foi observada associação entre o SNP C282Y em homozigose e o haplótipo ancestral *HLA-A*03-B*07* (CARDOSO; DE SOUSA, 2003; PIPERNO *et al.*, 1998). Barton *et al.* (2005b); apontam associação entre a contagem baixa de linfócitos e da gravidade da sobrecarga de ferro com a homozigose do SNP C282Y e com a presença do alelo *HLA-A*01* ou do alelo *HLA-B*08*. Já o alelo *HLA-B*14* foi associado com contagem de linfócito significativamente maior. Análises uni e multivariadas acerca do número de linfócitos totais do sangue periférico relatam que pacientes portadores do SNP C282Y e do haplótipo *HLA-A*01-B*08* apresentam contagens mais baixas em relação aos pacientes portadores do SNP C282Y e sem o haplótipo *HLA-A*01-B*08*. Em estudos familiares e em população controle de Portugal foi observada alta contagem de linfócito CD8+ no sangue periférico. Essa contagem foi encontrada em indivíduos que tinham tanto o SNP H63D quanto o alelo *HLA-A*29*. Também foi encontrada em população controle de Portugal, uma significativa correlação do alelo *HLA-A*01* e a alta quantidade de linfócitos CD8+ no sangue, e uma associação do *HLA-A*24* e a baixa quantidade de linfócitos CD8+ no sangue (BARTON *et al.*, 2005b; CARDOSO *et al.*, 2002; CRUZ *et al.*, 2004).

No Brasil, estudo feito com 15 homens com hemocromatose hereditária demonstrou que a maioria (53%) dos pacientes portava o SNP C282Y em homo ou heterozigose para o alelo *HLA-A*03*. Entretanto, três pacientes que não portavam o SNP C282Y, também eram positivos para o alelo *HLA-A*03*, sendo que um era homozigoto. Os três pacientes que eram homozigotos para C282Y, eram portadores do alelo *HLA-A*03* e dois eram homozigotos para ambos os SNPs. Outro dado destacado no trabalho é o fato de um dos pacientes que apresentou sintomas para hemocromatose hereditária de forma precoce (31 anos) era portador do haplótipo *HLA-A*03-C282Y G*, porém outro paciente também com sintomas precoce (20 anos) não portava nenhum dos alelos (BITTENCOURT *et al.*, 2002), entretanto um dos

autores informou posteriormente, ter diagnosticado o citado paciente com hemocromatose juvenil, devido a presença de mutações no gene *HAMP*.

Outro estudo com 35 pacientes com sobrecarga de ferro advinda de várias origens (HCV, consumo excessivo de álcool, anemia hemolítica crônica ou sem fatores secundários), encontrou o alelo *HLA-A*03* em 70% dos pacientes portadores do SNP C282Y em homozigose, e somente em 20 a 30% nos outros subgrupos de pacientes heterozigotos para C282Y. O trabalho destaca que essas frequências são as encontradas na população normal e que o grupo de alelos *HLA-A*03* se mostra claramente ligados ao SNP C282Y (CANÇADO *et al.*, 2006). O estudo da relação entre os SNPs do gene *HFE* e os alelos dos genes *HLA-A* e - *B* pode ser muito interessante para a população brasileira, já que a resistência de recombinação entre esses genes fazem o haplótipo *HLA-A-B-HFE* um excelente marcador para detectar migrações populacionais (DHILLON *et al.*, 2012; OLSSON *et al.*, 2008).

I.2 Sobrecarga de Ferro

A sobrecarga de ferro pode ser classificada em primária ou secundária. Na primária estão incluídas as alterações em genes de proteínas relacionadas à homeostase do ferro no organismo. É o que se observa nos pacientes com hemocromatose hereditária (HH), que apresentam aumento inapropriado da absorção intestinal de ferro, na maioria das vezes, associado à mutação no gene *HFE*. A sobrecarga de ferro secundária é observada em doenças congênitas ou adquiridas, que cursam com anemia hemolítica e/ou eritropoese ineficaz, e requerem múltiplas transfusões de hemácias. Isso ocorre em pacientes com doenças hematológicas, tais como: talassemia beta, anemia falciforme, síndrome mielodisplásica, anemia sideroblástica e anemia de Fanconi, e doenças hepáticas: hepatite C, esteatose hepática não-alcoólica, uso abusivo de bebidas alcoólicas (CAMASCHELLA, 2005; KUSHNER; PORTER; OLIVIERI, 2001; PIPERNO, 1998; SANTOS *et al.*, 2009).

I.3 Hemocromatose Hereditária

A hemocromatose hereditária é caracterizada pelo aumento na absorção de ferro e é frequente em caucasianos. O primeiro caso de hemocromatose hereditária foi relatado por Trousseau (1865), quando descreveu um paciente com a “tríade clínica” que consiste em: cirrose hepática, *diabetes mellitus* e hiperpigmentação da pele (ANDREWS, 1999; SANTOS *et al.*, 2009). Von Recklinghausen (1889), foi o primeiro a reconhecer como causa direta da doença o aumento progressivo do estoque de ferro, sendo o primeiro a utilizar a expressão

"hemocromatose". Posteriormente, a explicação sobre a natureza hereditária da doença foi realizada por SHELDON (1935). Os avanços no entendimento da transmissão genética e das bases moleculares desta doença ocorreram entre os anos 1970 e 1980 (SIMON *et al.*, 1975; SIMON *et al.*, 1977; SIMON *et al.*, 1980; SIMON *et al.*, 1987b; SIMON *et al.*, 1988), sendo que somente em 1996, Feder e colaboradores encontrou a primeira evidência de associação entre o gene *HFE* e a doença (FEDER *et al.*, 1996).

Pacientes com hemocromatose hereditária absorvem de duas a três vezes mais ferro da dieta do que pessoas normais. A maioria dos sintomas não se manifesta até a idade adulta, entretanto o aumento da saturação de transferrina sérica acontece na adolescência. O excesso de ferro é depositado nas células parenquimais do fígado, coração, pâncreas, glândula pituitária e glândula paratireoide. Os sintomas iniciais são inespecíficos, incluindo: fadiga, artralgia, disfunção erétil e aumento da pigmentação da pele. A progressão da doença inclui a tendência a desenvolver hepatomegalia, o que conduz à fibrose hepática, cirrose, carcinoma hepatocelular. Outra condição da progressão da doença é o depósito de ferro no coração, que pode estar associado com pericardite e arritmias. Este depósito causa cardiopatia, usualmente congestiva, mas pode ser também restritiva. A hemocromatose também foi associada a alguns tipos de endocrinopatias, além do *diabetes mellitus*, como hipogonadismo, hipotireoidismo e hipopituitarismo (ANDREWS, 1999).

Pacientes com hemocromatose são mais susceptíveis a infecções, em particular por *Vibrio vulnificus*, *Listeria monocytogenes*, *Yersinia enterocolitica*, *Salmonella enteritidis* (sorotipo typhimurium), *Klebsiella pneumoniae*, *Escherichia coli*, *Rhizopus arrhizus*, e algumas espécies de bolores. (ANDREWS, 1999).

A hemocromatose pode ser caracterizada pela deficiência ou baixa atividade da hepcidina (responsável pela redução na entrada de ferro na corrente sanguínea). A insuficiência de hepcidina pode ser oriunda da perda de função, por exemplo, do alelo *HAMP25* ou de mutações que prejudicam a interação da hepcidina com a proteína transmembrana de exportação de ferro (FPN). A maioria das hemocromatoses advêm da alteração em genes que regulam a hepcidina, incluindo o gene *HFE* (mutado em mais de 80% dos casos de hemocromatose), *TfR2* e *HJV* (CAMASCHELLA *et al.*, 2000; FEDER *et al.*, 1996; LOK *et al.*, 2009; PAPANIKOLAOU *et al.*, 2003; PIETRANGELO, 2010; SHAM *et al.*, 2005). A manifestação clínica da hemocromatose relacionada com o gene *HFE*, usualmente ocorre em pacientes de meia idade, com uma gama de anormalidades que vão de mudanças bioquímicas a danos severos em diversos órgãos (PIETRANGELO, 2004; 2010).

A biópsia do fígado é considerada o padrão ouro para quantificação do ferro. A sobrecarga é caracterizada pela concentração hepática que deve ser superior a 80 μmol por grama de fígado (peso seco), resultando em índice de ferro maior do que 1,9 mmol por quilograma por ano (o índice hepático é uma razão da concentração do ferro hepático pela idade do paciente em anos). A sobrecarga de ferro também pode ser avaliada pela flebotomia quantitativa até o ponto de depleção do ferro. O diagnóstico pode ser confirmado também por análises do gene *HFE*, homozigoto C282Y, conjuntamente com evidências de sobrecarga de ferro (ANDREWS, 1999).

Segundo Kowdley (2004) a correlação entre hemocromatose hereditária, carcinoma hepatocelular e hepatite C é controversa, enquanto alguns trabalhos como o de Fracanzani *et al.* (2001) apontam o aumento do risco, outros trabalhos (exemplo: MACSWEEN; SCOTT, 1973; PURTILO; GOTTLIEB, 2006) não encontraram correlação.

I.4 Hepatite C

Após a infecção aguda pelo vírus da Hepatite C (HCV), cerca de 50-80% dos casos evoluem para formas crônicas, sendo que, desses, 10-20% podem desenvolver cirrose fibrótica hepática, apresentando maior susceptibilidade para o desenvolvimento de câncer hepatocelular (CHC) (PAWLOTSKY, 2004). Fibrogênese é um processo dinâmico, caracterizado por síntese de constituintes da matriz extracelular (MEC), composta por uma mistura complexa de glicoproteínas (colágeno, elastina, fibronectina e laminina) e proteoglicanos, organizada em rede tridimensional (MARCELLIN; ASSELAH; BOYER, 2006).

A forma crônica da hepatite C é caracterizada por fibrose hepática progressiva, que geralmente tem evolução lenta, entretanto, em alguns casos, há evolução rápida para formas mais graves da doença (MCCAUGHAN; GEORGE, 2004). A fibrose hepática é entendida como resposta cicatricial do fígado à lesão decorrente do processo inflamatório crônico, em que o equilíbrio entre a produção e degradação da MEC favorece a produção. A inflamação crônica resulta em acúmulo de MEC e substituição de colágeno de baixa densidade (tipo IV) por colágeno de alta densidade (tipo I) no espaço subendotelial de Disse (ALBANIS; FRIEDMAN, 2005).

A formação da fibrose hepática envolve populações celulares heterogêneas. Além das células estreladas hepáticas (CEH) ativadas, com seu papel já bem estabelecido, os miofibroblastos periportais e perivenosos também participam do processo (RAMADORI;

SAILE; SAILE, 2002). As CEH são de origem mesenquimal, correspondendo a 5-8% das células hepáticas, e estão localizadas nos sinusoides do parênquima hepático. Essas células possuem 80% do estoque de retinóides do corpo humano, sendo responsáveis pela homeostase da MEC no fígado normal pela regulação do fluxo sanguíneo sinusoidal e secreção de mediadores parácrinos e autócrinos (FRIEDMAN, 2003). Devido à agressão hepática, as CEH são ativadas e se diferenciam modificando-se da forma de estrela, rica em vitamina A (forma quiescente), para células semelhantes aos miofibroblastos, pobres em retinóides, expressando alfa-actina de músculo liso (α -SMA) (forma ativada). As formas ativadas da CEH são importantes fontes de mediadores, componentes da MEC, proteases e inibidores de proteases que são fundamentais na formação da fibrose hepática (BEDOSSA; PARADIS, 2003; FRIEDMAN, 2003). A resposta imune inata na infecção pelo vírus da hepatite C é difícil de estudar *in vivo*, devido ao percentual pequeno de casos que apresentam manifestações clínicas da infecção (CHANG, 2004). A resposta imune humoral é importante mecanismo que pode conferir imunidade na infecção pelos vírus das hepatites (GORHAM, 2005). O anticorpo anti-HCV sérico pode ser detectado de sete a oito semanas após a infecção. : Diferentemente da infecção causada pelos vírus da hepatite A e B, a infecção pelo vírus C, mesmo após a cura, não confere imunidade persistente e o paciente permanece suscetível a outra infecção caso se exponha novamente ao HCV.

Células T CD4+ HCV específicas são detectadas no sangue periférico e no fígado de indivíduos infectados. A proliferação de linfócitos T CD4+ é detectada em um terço dos pacientes com HCV, sendo a resposta dirigida ao *core* do vírus, indicando que essas células sejam protetoras ou, pelo menos, marcadoras de melhora clínica. Células T CD8+ podem ser detectadas no fígado e no sangue periférico de pacientes cronicamente infectados pelo HCV, sendo a magnitude da atividade citotóxica inversamente proporcional aos títulos virais (CHANG, 2004).

Diversos fatores contribuem para o desenvolvimento e progressão da fibrose hepática em pacientes com hepatite C crônica, incluindo os epidemiológicos como idade do paciente na época da infecção maior do que 40 anos, sexo masculino, ingestão alcoólica maior do que 50 gramas de etanol ao dia, obesidade e coinfeção com HBV e HIV (AFDHAL, 2004; HU *et al.*, 2004; POYNARD *et al.*, 2001). Entre os fatores histopatológicos estão incluídos, esteatose hepática, espécies reativas de oxigênio, TGF-beta, apoptose de hepatócitos e fibrose hepática. Entre os fatores genéticos, cabe destacar sítios polimórficos de genes relacionados

com a resposta imune, incluindo *HLA-DR* (JEN-EING *et al.*, 2007), *CTLA-4*, *TLR-4*, *RANTES* e *HFE* (BATALLER; NORTH; BRENNER, 2003).

A relação entre a presença de mutações no gene *HFE* e a progressão da fibrose hepática em pacientes com hepatite C crônica ainda é controversa (BOIGE *et al.*, 2003). O risco aumentado de hepatite C crônica foi previamente relatado em pacientes com sobrecarga de ferro (TURLIN *et al.*, 1995), porém estudo mais recente relata que, exceto para homozigotos C282Y, as mutações do gene *HFE* não aumentam o risco de desenvolver carcinoma hepatocelular em pacientes com cirrose (CAUZA *et al.*, 2003). Recentemente, foi observada maior taxa de resposta virológica sustentada em portadores da mutação H63D no gene *HFE*. No entanto, os efeitos das mutações na estrutura e função da proteína HFE e/ou outros fatores genéticos que possam estar relacionados ou até mesmo fatores do hospedeiro devem ser ainda mais bem estudados (BONKOVSKY *et al.*, 2006).

I.4.1 *HFE* e Hepatite C

O perfil das mutações no gene *HFE* é significativamente diferente entre as várias populações. Até agora, a maioria dos estudos sobre a sobrecarga de ferro e fibrose em hepatite C foram realizados em países ocidentais. Thorburn *et al.* (2002) demonstraram que mutações C282Y e H63D no gene *HFE* não foram associadas com quaisquer características clínicas, bioquímicas, virológicas ou patológicas da sobrecarga de ferro intra-hepática, e, portanto, não têm papel na progressão da doença hepática na infecção pelo VHC. Tung *et al.* (2003) mostraram que a presença da mutação C282Y no gene *HFE* em doentes com doença hepática e infecção pelo VHC foram independentemente associados com elevados índices de ferro sérico, índice de ferro hepático e fibrose avançada. No Brasil, Martinelli *et al.* (1999) relataram que pacientes com infecção crônica pelo VHC que são portadores da mutação C282Y no gene *HFE* tendem a apresentar mais evidente acúmulo de ferro no corpo, maior grau de atividade necroinflamatória e fibrose no fígado e que, portanto, mutações no gene *HFE* podem ser fatores adicionais entre os implicados na determinação de pior prognóstico da doença hepática crônica, após infecção pelo VHC.

I.5 Carcinoma Hepatocelular

O carcinoma hepatocelular (CHC) é um tumor primário do fígado, que ocupa o quinto lugar em frequência entre as neoplasias, sendo o terceiro em causa de mortalidade dentre os

tumores malignos no mundo (PARKIN, 2001). Apresenta características epidemiológicas diversas, incluindo variações marcantes entre as diferentes regiões geográficas (< 80% dos casos estão localizados na África Subsaariana e Ásia oriental), predominância no sexo masculino e presença bem documentada dos fatores de risco precipitantes como a doença hepática crônica e a cirrose (aproximadamente 70-90% dos casos) (EL-SERAG; RUDOLPH, 2007).

As principais causas de cirrose nos portadores de CHC são: infecção crônica pelos vírus das hepatites B e C, doença hepática alcoólica, e possivelmente, a esteato-hepatite não alcoólica, estando menos frequentemente envolvidas, a hemocromatose hereditária, a deficiência de alfa1-antitripsina, a hepatite autoimune, e algumas porfirias) (EL-SERAG; RUDOLPH, 2007).

De acordo com as recomendações internacionais (BRUIX; SHERMAN, 2005), a triagem do CHC deve ser adotada naqueles pacientes com alto risco de desenvolver o tumor, ou seja, indivíduos portadores de cirrose hepática e com infecção crônica pelo vírus da hepatite B, realizando-se ultrassonografia de abdome semestralmente. Uma vez detectado um nódulo no fígado com cirrose, o diagnóstico de CHC deverá ser investigado. Estabelecido o diagnóstico de CHC, o prognóstico do paciente varia de acordo com o estágio tumoral (no momento do diagnóstico) e a terapêutica instituída. Pacientes com tumores em estádios precoces poderão receber tratamentos curativos, como ressecção cirúrgica ou transplante ortotópico de fígado, com até 70% de sobrevida em 5 anos (LLOVET; BRUIX, 2003). Entretanto, indivíduos com doença avançada e não tratados, apresentam taxas de sobrevida em 1 e 2 anos de 10-72% e 8-50%, respectivamente (LLOVET; BRUIX, 2003). Enquanto pacientes em estágio terminal sobrevivem menos de 6 meses (WILDI *et al.*, 2004)

Por ser doença com história natural ainda não completamente conhecida, o CHC tem despertado grande interesse para estudos genéticos em diferentes populações. Os principais marcadores são os sítios polimórficos observados nos genes *HLA* de classe II. El-Chennawi *et al.* (2008) estudaram pacientes egípcios portadores de CHC, relatando que os grupos de alelos *DRB1**04 e *DQB1**02 estão associados com susceptibilidade, enquanto o grupo de alelos *DQB1**06 com proteção contra o desenvolvimento do tumor. Estudo em chineses de Hong Kong em portadores de infecção crônica pelo vírus da hepatite B, apresentando ou não CHC, mostrou que os alelos *DRB1**15:01, *DQA1**01:02 e *DPB1**05:01 foram os mais frequentes nos pacientes com CHC (DONALDSON *et al.*, 2001).

Em estudo sobre a expressão de oncogenes e genes supressores de tumor relacionados ao CHC e com as sobrecargas de ferro, o gene supressor de tumor *p53*, cuja expressão é aumentada em pacientes com CHC, encontra-se também mutado em 71% dos pacientes com hemocromatose hereditária (HH). Além disso, os pacientes homozigotos para a substituição C282Y do *HFE* apresentam também mutações no códon 220 de *p53*. Curiosamente, o promotor da hepcidina contém um elemento sensível ao *p53*, sendo que o *p53* pode ativar o promotor da hepcidina em células de hepatoma por via inflamatória dependente de IL-6. Porém, essa associação entre mutações de *p53* e genótipos *HFE* mutantes ainda não está bem esclarecida (VAUTIER *et al.*, 1999).

I.5.1 Carcinoma Hepatocelular e *HFE*

A hemocromatose hereditária é provavelmente a desordem genética na mais comum população caucasiana, com incidência de 1 em 200. O excesso de ferro absorvido pelo intestino, se não eliminado, pode acumular-se por anos e causar dano tecidual progressivo. A manifestação do dano tecidual é a cirrose, que pode levar ao desenvolvimento de câncer hepático (EDWARDS *et al.*, 1988; WILLIS *et al.*, 1997; WILLIS *et al.*, 2000).

Trabalhos têm demonstrado associação entre hemocromatose hereditária e carcinoma hepatocelular, tendo que a hemocromatose hereditária foi apontada como uma das principais complicações no tratamento do carcinoma. O risco de pacientes com hemocromatose hereditária desenvolverem carcinoma hepatocelular foi estimado em 200 vezes maior por alguns autores. (BRADBEAR *et al.*, 1985; NIEDERAU *et al.*, 1985; ; KOWDLEY, 2004). No entanto, há dúvidas se a associação de sobrecarga de ferro moderado, doenças hepáticas crônicas e mutações no gene *HFE* geram carcinoma hepatocelular (KOWDLEY, 2004).

O acúmulo de ferro no fígado leva a formação de oxigênio reativo, que causa stresse oxidativo, provocando mutações no DNA das células hepáticas e desenvolvimento de carcinoma hepatocelular (DEUGNIER, 2003; NAHON *et al.*, 2008). Além da idade e do sexo masculino serem fatores de risco bem conhecidos para o acúmulo de ferro, descobrir marcadores genéticos de risco para o câncer hepatocelular, como mutações no gene *HFE*, é de grande interesse para validar o papel carcinogênico do ferro no fígado e sugerir medidas preventivas (NAHON *et al.*, 2008).

Alguns estudos têm examinado se existem associações entre mutações no gene *HFE* e carcinoma hepatocelular, entre os pacientes com hepatite C crônica ou induzida por consumo de álcool. Muitos investigadores (BOIGE *et al.*, 2003; CAUZA *et al.*, 2003; FARGION *et*

et al., 2001; HELLERBRAND *et al.*, 2003; KOWDLEY, 2004) têm encontrado relativo aumento nas mutações clássicas no gene *HFE* em pacientes com câncer hepatocelular, independentemente de outros fatores de risco para o câncer hepático tais como: idade, sexo e duração da cirrose (KOWDLEY, 2004; RACCHI *et al.*, 1999).

Além do câncer hepatocelular, há controvérsia se pacientes com hemocromatose hereditária tem aumento no risco de câncer extra-hepático. Alguns estudos têm achado grande prevalência de câncer extra-hepático, enquanto outros não. Entretanto dados epidemiológicos sugerem relação positiva entre sobrecarga de ferro e câncer, dando suporte ao conceito que a sobrecarga de ferro pode ter efeito carcinogênico generalizado (AMMANN *et al.*, 1980; BOMFORD; WILLIAMS, 1976; BRADBEAR *et al.*, 1985; STEVENS *et al.*, 1988; ELMBERG *et al.*, 2003; FRACANZANI *et al.*, 2001; KOWDLEY, 2004; MALLORY; KOWDLEY, 2001; SELBY; FRIEDMAN, 2006).

A influência da sobrecarga de ferro e das mutações do gene *HFE* têm sido debatidas nos últimos anos. As duas mutações C282Y e H63D explicam a maioria dos casos de hemocromatose hereditária, condição caracterizada pela sobrecarga de ferro hepático que conduz a alta incidência de câncer. Ainda permanece obscura a questão se os pacientes com outras doenças hepáticas e heterozigotos para C282Y e H63D podem desenvolver sobrecarga hepática de ferro, o que conduziria ao aumento da lesão hepática e carcinomas (FEDER *et al.*, 1996; FARGION *et al.*, 2005; KEW, 2005; NAHON *et al.*, 2008).

II. HIPÓTESE

Considerando que os depósitos de ferro estejam relacionados com progressão para fibrose hepática e que o gene *HFE* esteja associado com a sobrecarga de ferro no fígado e que, poucos sítios polimórficos sejam estudados no gene *HFE*, aventamos as seguintes hipóteses: i) existem outros sítios na região codificadora do gene *HFE* ainda não estudados na literatura; ii) os diversos sítios polimórficos do gene *HFE* podem estar associados com susceptibilidade ou proteção ao desenvolvimento de sobrecarga de ferro em doenças hereditárias (hemocromatose) ou adquiridas (hepatite C e carcinoma hepatocelular) que podem cursar com sobrecarga de ferro no fígado.

III. JUSTIFICATIVA

Apenas dois sítios polimórficos (C282Y e H63D) na região codificadora do gene *HFE* são usualmente estudados em associação com doenças que cursam com depósitos de ferro no fígado. Considerando que a população brasileira seja uma das mais heterogêneas do mundo, sendo o resultado de cinco séculos de mistura entre indivíduos oriundos dos três continentes (africano, americano e europeu), a caracterização mais completa da região codificadora do gene *HFE* no Brasil permitiu a verificação da existência de outras variantes genéticas nesta população e sua associação com o controle do estado do ferro e possível progressão para fibrose hepática em doenças hereditárias (hemocromatose) ou adquiridas (hepatite C e carcinoma hepatocelular), associadas com sobrecarga de ferro. Esse estudo permitiu; i) conhecer o perfil da variabilidade genética da região codificadora do *HFE* na população brasileira, ii) padronização da nomenclatura do gene, ainda não disponível na literatura, iii) associação desses sítios polimórficos com doenças congênitas e adquiridas associadas com depósito crônico de ferro no fígado. Os resultados desse estudo impulsiona a realização de outros em outras populações brasileiras ou estrangeiras para viabilizar o entendimento acerca da participação desse gene na patogenia das doenças que evoluem com sobrecarga de ferro.

IV. OBJETIVOS

O objetivo geral consistiu na avaliação da diversidade genética da região codificadora (éxons 2, 3, 4 e 5) do gene *HFE* em indivíduos saudáveis e em pacientes que apresentam sobrecarga hepática de ferro hereditária (hemocromatose) ou adquirida, como na hepatite C e no carcinoma hepatocelular. Ainda, foram gerados haplótipos estendidos, englobando outros genes do CPH de classe I, II e III para estudos de desequilíbrio de ligação entre o *HFE* e os outros *loci* do CPH.

Os objetivos específicos incluíram:

- Identificar os SNPs recorrentes do gene *HFE* na população saudável e de pacientes e montar o perfil genético para essas amostras de populações;
- Avaliar o desequilíbrio de ligação entre os sítios polimórficos do gene *HFE*;
- Identificar associações entre sítios polimórficos do gene *HFE* com variáveis clínicas e laboratoriais (presença ou ausência de sobrecarga de ferro) entre os grupos de pacientes (hepatite C e carcinoma hepatocelular);

- Fazer associação entre sítios polimórficos do gene *HFE* e a sobrecarga primária mediante a comparação de portadores de hemocromatose hereditária e indivíduos saudáveis;
- Reconstrução de haplótipos dos SNPs do gene *HFE* em todas as amostras de populações desse estudo (controles, hemocromatose hereditária, hepatite C e câncer hepatocelular);
- Gerar nomenclatura para a variabilidade gene *HFE* baseando nas regras de nomenclatura dos genes do CPH e propor nomenclatura para os alelos advindos desses haplótipos;
- Fazer estudos de associação para haplótipos entre os grupos de pacientes com hepatite C e carcinoma hepatocelular e variáveis laboratoriais (presença ou ausência de sobrecarga de ferro);
- Fazer estudos de associação para haplótipos com a hemocromatose mediante a comparação com a população controle;
- Fazer estudo na população brasileira, identificando as principais variações presentes e comparando essas diferenças com diversas populações disponíveis no site 1000genome (<http://www.1000genomes.org/>).
- Montar o haplótipo estendido do *HFE* com outros genes do CPH, na população controle e fazer estudo de ligação entre os diversos *loci*;
- Investigar a existência de pressões seletivas influenciando a variabilidade genética no gene *HFE* na amostra de população Controle;

V. MATERIAL E MÉTODOS

V.1 Desenho do estudo e local de execução do projeto

O estudo é do tipo transversal, comparativo e descritivo, realizado a partir da análise de dados clínicos e laboratoriais dos pacientes atendidos nos Ambulatórios de Gastroenterologia e de Hepatites do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (HCFMRP-USP), sob a coordenação da Professora Ana de Lourdes C. Martinelli, que também participa do projeto.

V.2 Aspectos éticos e origem das amostras

O estudo foi aprovado pelo Comitê de Ética em Pesquisa do HCFMRP (Processos HCRP 4822/2011) e as amostras que foram usadas fazem parte de dois bancos de amostras. As amostras controles são provenientes do Banco de Amostras do Núcleo em Pesquisa em

Imunogenética (BAMPI) número processo HCFMRP 3530/2007 sob a coordenação do Professor Dr. Eduardo Antônio Donadi, e são parte integrante de projetos de rastreamento alélico, realizados pelo grupo, para diversos *loci*, o que possibilita uma análise de haplótipos na população brasileira. As amostras de pacientes fazem parte do banco de amostras número processo HCFMRP 3416/2003 sob a responsabilidade da Prof Dra Ana de Lôurdes Candolo Martinelli são amostras coletadas pelo ambulatório de Gastroenterologia para diversos fins. Também foram cedidas pelo Prof. Dr. Eduardo Luiz Rachid Cançado e a Dra, Andréia Evangelista da Silva, 8 amostras de pacientes com hemocromatose hereditária provenientes do projeto Cappesq sob nº 1065/05 sob o título “Hemocromatose Hereditária: Análise de Genótipo e Correlação com Aspectos Clínicos e Anátomo-Patológicos”.

Destes bancos foram selecionadas amostras que se enquadram em alguns critérios conforme item V.3, que isolam e determinam as características específicas requeridas para o presente estudo. Como se trata de amostras já coletadas e não houve nenhum contato direto ou indireto com estes indivíduos, e como as informações sobre os pacientes foram mantidas em sigilo, o estudo foi isento da aplicação do Termo de Consentimento Livre e Esclarecido.

V.3 Casuística

Hepatite C: Grupo composto por 130 pacientes (93 homens) com idades variando entre 19 à 69 anos (média de $42,60 \pm 10,98$), apresentando hepatopatia crônica pelo vírus C, sendo 71 com sobrecarga de ferro (SOB) e 59 sem SOB, acompanhados nos Ambulatórios de Doenças do Fígado e Ambulatório de Hepatites do Departamento de Clínica Médica do HCFMRP-USP e que preenchiam os critérios de inclusão: diagnóstico de hepatopatia crônica por infecção pelos vírus C da hepatite e ausência de ingestão alcoólica abusiva (60 g/dia). Os critérios de exclusão incluíram: não concordância em participar do estudo, infecção pelo vírus da imunodeficiência adquirida, existência de outras hepatopatias como hepatite B, hepatite autoimune, deficiência de alfa-1 antitripsina, doença de Wilson, cirrose biliar primária, colangite esclerosante primária e hepatite medicamentosa, e ainda, em situações nas quais possa ser observada sobrecarga de ferro como insuficiência renal crônica, hemoglobinopatias e porfiria cutânea tardia.

Carcinoma Hepatocelular: Grupo composto por 60 pacientes (43 homens), com idades variando de 14 a 78 anos (média de $57,00 \pm 14,00$), sendo que desses, 24 apresentavam sobrecarga de ferro e tinham como doença de base a hepatite C. Nos 36 restantes, as doenças de base eram diversas incluindo, hepatite criptogênica, hepatite B, esteato-hepatite não-

alcoólica e sem comorbidade. Para o diagnóstico do CHC, foram utilizadas as recomendações da *American Association for the Study of Liver Diseases - AASLD Practice Guideline* (GORHAN, 2005), obedecendo ao organograma já utilizado para os pacientes seguidos no Ambulatório de Lesões Focais do Fígado do Departamento de Clínica Médica do HCFMRP-USP e que considera o tamanho do nódulo hepático: i) lesões < 2 cm de diâmetro: cirrose com nódulo hepático hipervascular e *washout* na fase portal, em uma técnica dinâmica de imagem; ii) lesões entre 1-2 cm de diâmetro: cirrose com nódulo hepático hipervascular e *washout* na fase portal, em duas técnicas dinâmicas de imagem (tomografia computadorizada helicoidal trifásica, ressonância nuclear magnética ou arteriografia). Biópsia hepática com análise histopatológica compatível com CHC, independentemente do tamanho do nódulo.

Como o diagnóstico não depende de biópsia hepática, esses pacientes não tinham informação acerca da sobrecarga de ferro. Foram aceitos como sujeitos de pesquisa, pacientes com ou sem biópsia da lesão tumoral, realizada previamente para definição diagnóstica. Foram excluídos pacientes portadores de outras neoplasias malignas, portadores do HIV e imunossuprimidos.

Hemocromatose hereditária: Composto por 14 pacientes (9 homens) com idades variando de 32 a 81 anos (média de $55,35 \pm 15,16$), com diagnóstico de hemocromatose hereditária (HH) definida por evidências de sobrecarga de ferro no organismo (saturação de transferrina $\geq 45\%$ com ou sem elevação de ferritina sérica acima do valor normal ou presença depósitos de ferro em hepatócitos). Foram excluídos pacientes que sabidamente cursam com a sobrecarga de ferro, como história de múltiplas transfusões sanguíneas, hemoglobinopatias, doenças hepáticas crônicas de outras etiologias e insuficiência renal crônica em hemodiálise.

Grupo controle: Composto por 100 voluntários saudáveis (80 homens) com idades variando de 20 a 52 anos, provenientes de banco de sangue, com perfil hepático (função hepática e enzimas) normal e ausência de qualquer marcador sorológico para hepatites virais.

As características demográficas, clínicas e aquelas referentes à sobrecarga de ferro dos pacientes e as características demográficas dos controles estão ilustradas na Tabela 3.

Tabela 3. Número de indivíduos, média de idade e desvio padrão, gênero e porcentagem por etnia.

CASO CLÍNICO	NÚMERO DE INDIVÍDUOS	MÉDIA DE IDADE E DESVIO PADRÃO	QUANTIDADE E PORCENTAGEM DE HOMENS	PROPORÇÃO POR ETNIA
Controle	100	33,31 $\pm$ 8,18	80 (80%)	Branco: 79% Pardo: 11% Negro: 9% Asiático: 1%
Hemocromatose Hereditária	14	55,35 $\pm$ 15,16	9 (64%)	Branco: 86% Pardo: 7% Negro: 7%
Hepatite c/ Sobrecarga	71	42,92 $\pm$ 10,82	57 (80%)	Branco: 92% Pardo: 4% Negro: 4%
Hepatite s/ Sobrecarga	59	42,22 $\pm$ 11,24	36 (61%)	Branco: 95% Pardo: 3% Negro: 2%
Carcinoma Hepatocelular c/ Sobrecarga de Ferro c/ VHC	24	59,91 $\pm$ 8,61	18 (75%)	Branco: 84% Pardo: 8% Asiático: 8%
Hepatocarcinoma Diverso	36	55,05 $\pm$ 16,49	25 (69%)	Branco: 89% Pardo: 5% Negro: 3% Asiático: 3%

V.4 Aspectos clínicos e laboratoriais considerados importantes para estratificação dos pacientes

V.4.1 Sobrecarga de ferro

A sobrecarga de ferro foi classificada sob dois aspectos: 1. Semiquantificação dos depósitos no tecido hepático, segundo a classificação de Knodell *et al.* (1981), com algumas modificações; 2. Saturação de transferrina $\geq$ 45% com ou sem elevação de ferritina sérica.

Assim, foi considerado que o paciente apresentava sobrecarga de ferro na presença de saturação da transferrina > 45% e/ou presença de depósito de ferro em hepatócitos e, na ausência dessas duas condições, os pacientes foram considerados sem sobrecarga de ferro. Aqueles que não apresentaram essas informações foram considerados como grupo separado.

V.4.2 Coleta e processamento dos materiais biológicos para o estudo

Coleta do sangue: Para extração do DNA genômico, foram colhidos 10 mL de sangue venoso periférico de cada indivíduo em tubos Vacutainer (Beckton & Dickinson, Franklin Lakes, NJ), contendo EDTA K3 (0.054mL/tubo). O DNA foi extraído por *salting out*, utilizando o procedimento descrito por MILLER; DYKES; POLESKY (1988).

V.5 Diversidade do Gene *HFE*

A reação em cadeia da polimerase (PCR) foi utilizada para amplificação das regiões dos éxons 2, 3, 4 e 5 e, aproximadamente, 60pb dos íntrons que margeiam esses éxons, propiciando margem de distância entre os *primers* e a região a ser amplificada. Após confirmação da amplificação em gel de agarose (Figura 5), as amostras foram sequenciadas (ABI310 Genetic Analyzer Applied Biosystems, Foster City, CA). Os iniciadores foram desenhados utilizando o *Software* GENE RUNNER v.3.05 (Hastings Software, Inc., Hastings, NY) e o site UCSC Genome Bioinformatics Site (<http://genome.ucsc.edu/>) (Tabela 4). As reações A PCR para cada exon foi realizada em um volume final de 50µl contendo 1X *Buffer* (70 mM Tris-HCl pH 8.8, 20 mM (NH₄)₂SO₄, 3 mM MgCl₂, 1mM ditioneitol (DTT)), 0.2 mM de cada dNTP, 10 pmol de cada *primer*, 1.1 unidades de Taq DNA-polymerase Platinum (Invitrogen, Carlsbad, CA), e 200 ng de DNA genômico. O programa consistiu de um ciclo inicial de desnaturação a 94°C por 5 min, seguidos de 30 ciclos de 45 segundo à 94°C, 45 segundo na temperatura de anelamento (Tabela 4) e 1 minuto à 72°C, e uma extensão final à 72°C por 7 minutos. O produto de PCR foi sequenciado em um ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), de acordo com a sugestão do fabricante.

Figura 5: Foto ilustrativa do padrão de amplificação do éxon 3 em gel de agarose 1%.

Tabela 4. Descrição das regiões analisadas, das sequências dos primers, da temperatura de melting, do tamanho em pb do fragmento amplificado e da sequência fasta da região (primer indicado pelas letras maiúsculas e grifado, região codificadora em letras maiúsculas. SNPs indicados de acordo com o código IUPAC).

Regiões Amplificadas	Nomes e sequências dos <i>primers</i>	Temperatura de <i>melting</i>	No. de pares de bases do fragmento
Íntron 1 (parcial)	HFEex2F 5'- GCTGATGGTATGAGTTGATGC - 3'	60°C	444 pb
Éxon 2			
Íntron 2 (parcial)	HFEex2R 5'- CTCAGACTTCCAGCTGTTTC - 3'		
FASTA (Região éxon 2)			
4441	<u>GCTGATGGTATGAGTTGATGC</u> aggtgtgtgga		
4501	gcctcaacatcctgctccctcctactacacatgggtaaggcctgttgctctgtctccag		
27	S H S L H Y L F M G A S E Q D L G L S L		
4561	GTTCACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGACCTTGGTCTTTTCCT		
47	F E A L G Y V D D Q L F V F Y D H E S R		
4621	TGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTATGATG ^{s1} ATGAGAGTC		
67	R V E P R T P W V S S R I S S Q M W L Q		
4681	GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC		
87	L S Q S L K G W D H M F T V D F W T I M		
4741	AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA		
107	E N H N H S K E		
4801	TGGAAAATCACAACCACAGCAAGGgtay ² gtggagagggggcctcaccttcctgaggttgt		
4861	cagagctttttcatctttttcatgcacatcttgaag <u>GAAACAGCTGGAAGTCTGAG</u>		
Íntron 2 (parcial)	HFEex3F 5'-GCTTCCTGAGATCATTGGTCC -3'	59°C	485pb
Éxon 3			
Íntron 3 (parcial)	HFEex3R 5'-CAGAATTTGGAGAGGCACACAG -3'		
FASTA (Região éxon 3)			
4921	<u>GCTTCCTGAGATCATTGGTCC</u> ttggggatgggtgg		
115	S H		
4981	aaatagggacctatttccttttggttgacagtaacaaggctggggatttttccagAGTCCCA		
117	T L Q V I L G C E M Q E D N S T E G Y W		
5041	CACCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG		
137	K Y G Y D G Q D H L E F C P D T L D W R		

Regiões Amplificadas	Nomes e sequências dos <i>primers</i>	Temperatura de <i>melting</i>	No. de pares de bases do fragmento
5101	GAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAG		
157	A A E P R A W P T K L E W E R H K I R A		
5161	AGCAGCAGAACCCAGGGCCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC		
177	R Q N R A Y L E R D C P A Q L Q Q L L E		
5221	CAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA		
197	L G R G V L D Q Q V		
5281	GCTGGGGAGAGGTGTTTTGGACCAACAAGgtatggtggaaacacacttctgcccctatac		
5341	tctagtggcagagtggaggaggttgcagggcacggaatccctggttgaggtttcagaggt		
5401	ggctgagg <u>CTGTGTGCCTCTCCAAATTCTG</u>		
Íntron 3 (parcial)	HFEex4F 5'-GTATTTTCCTTCCTCCAACC - 3'	53°C	451pb
Éxon 4			
Íntron 4 (parcial)	HFEex4R 5'-CATAATTACCTCCTCAGGCAC-3'		
FASTA (Região éxon 4)			
6301	<u>GTATTTTCCTTCCTCCAACC</u> tatagaaggaagtgaaggttcagtccttcctggs ³ aag		
207		P P L V K	
6361	ggtaaacagatcccctctcctcatccttcctctttcctgtcaagTGCCTCCTTTGGTGAA		
212	V T H H V T S S V T T L R C R A L N Y Y		
6421	GGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGGCCTTGAATACTA		
232	P Q N I T M K W L K D K Q P M D A K E F		
6481	CCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTT		
252	E P K D V L P N G D G T Y Q G W I T L A		
6541	CGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGC		
272	V P P G E E Q R Y T C Q V E H P G L D Q		
6601	TGTACCCCTGGGGAAGAGCAGAGATATACGTR ⁴ CCAGGTGGAGCACCCAGGCCTGGATCA		
292	P L I V I W E		
6661	GCCCCTCATTGTGATCTGGGgtatgtgactgatgagagccaggagctgagaaaatctatt		
6721	gggggttgagagga <u>GTGCCTGAGGAGGTAATTATGG</u>		
Íntron 4 (parcial)	HFEex5F 5'-GTGCCTGAGGAGGTAATTATGG-3'	58°C	308bp
Éxon 5			
Íntron 5 (parcial)	HFEex5R 5'-CTTCCCATGGATGCCAGATC-3'		
FASTA (Região éxon 5)			
6721	<u>GTGCCTGAGGAGGTAATTATGG</u> cagtgagatgaggatctgctcttt		
6781	gttaggggr ⁵ tgggcy ⁶ gaggggtggcaatcaaaggctttaacttgctttttctgttttagAG		
299	P S P S G T L V I G V I S G I A V F V V		
6841	CCCTCACCGTCTGGCACCCCTAGTCATTGGAGTCATCAGTGAATTGCTGTTTTTGTCTGC		
319	I L F I G I L F I I L R K R Q G S R		
6901	ATCTTGTTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAgtagtag		
6961	gaacaagggg- ⁷ aagtctcttagtacctctgccccaggggcacagtgggaagaggggagag		
7021	gg <u>GATCTGGCATCCATGGGAAG</u>		

Identificação do SNP /alelo mais frequente segundo NCBI (<http://www.ncbi.nlm.nih.gov>) e ENSEMBL(<http://www.ensembl.org>): 1- rs1799945 (H63D) / C;2 - rs2071303 (IVS2(+4)T>C), / T;3- rs807209 / G; 4- rs1800562 (C282Y) / G; 5-rs2794717/ G; 6- rs1800708 (IVS4(-44)T>C) / T; 7- mutação no íntron 5 / G

V.6 Clonagem do novo SNP no íntron 5

Para validação do novo SNP encontrado no íntron 5 na amostra da população controle foi utilizado o protocolo abaixo:

V.6.1 Purificação do Produto de PCR.

A purificação do produto de PCR foi realizada, usando *kit* comercial (Jetx TM PCR DNA and Gel Band Purification Kit – GE Healthcare®, Buckinghamshire, UK).

V.6.2 Reação de ligação do fragmento ao plasmídeo:

Para o procedimento de ligação com o plasmídeo pGEM® - T easy Vector System (Promega, Madison, WI), foi preparada solução contendo 1,0µL do plasmídeo, 5,0µL de tampão do plasmídeo, 1,0µL de T₄ DNA ligase e 3,0µL do amplicon purificado, seguindo-se de incubação a 4°C *overnight*. Posteriormente, foram adicionados 10µL de suspensão de bactérias, agitando-se suavemente. Então, essa preparação foi incubada no gelo por 30 minutos, submetida a choque térmico a 42°C por 150 segundos e novamente incubada em gelo por mais 150 segundos. Em seguida, foram adicionados 800µL de meio de cultura líquido LB Broth (Fisher-Scientific, Hampton, New Hampshire, EUA), incubando-se a 37 °C por 1 hora, sob agitação leve.

V.6.3 Plaqueamento das bactérias transformadas:

As bactérias transformadas foram obtidas após centrifugação 3 minutos a 2.325 g, descartando-se de 300 a 750µL de sobrenadante e ressuspensando o botão celular com o restante do sobrenadante. Essa solução foi plaqueada em meio de cultura LB Ágar (Fisher-Scientific), ampicilina (Teuto, Anápolis, Brasil) (200 mg/mL), Xgal (Ludwig Biotec, Alvorada, Brasil) (40 mg/mL); incubou-se a 37°C *overnight* para o crescimento das colônias.

A partir dessas colônias, foram identificadas as colônias transformadas (brancas); as quais foram transferidas para tubos contendo 5 mL de meio líquido LB Broth (Fisher-Scientific), com ampicilina (Teuto) (200 mg/mL). As bactérias foram incubadas a 37°C com agitação leve na posição inclinada *overnight* para o crescimento. O crescimento das colônias foi verificado pela turvação dos meios. Dos meios cultivados foram reservados 2 mL, que foram centrifugados por 5 minutos de 2.325 a 5.232 g até obtenção de *pellet*, descartado o

sobrenadante e repetido o procedimento no mesmo tubo, de forma obter *pellet* com maior quantidade de bactérias.

V.6.4 MINI PREP

Para essa preparação foram adicionados sobre o *pellet* os tampões, de acordo com a especificação do fabricante do kit (Illustra™ Plasmid Prep Mini Spin Kit - GE Healthcare®).

V.7 Nomenclatura dos Alelos do gene *HFE*

Como citado na introdução, o gene *HFE* não possui nomenclatura de haplótipos padronizada, ou seja, os SNPs foram até agora nomeados separadamente. Nós utilizamos os softwares PHASE e ARLEQUIN para dedução da fase meiótica dos alelos e determinação dos haplótipos. A partir dessa inferência, os haplótipos foram nomeados seguindo as regras do site IMG/HLA Database (<http://hla.alleles.org/nomenclature/naming.html>), que são únicas para todos os genes HLA e foram padronizadas utilizando como referência o trabalho de MARSH *et al.* (2010). Assim:

1. Cada nome de alelo HLA refere-se ao grupo de alelos que forma o haplótipo. O alelo tem um número único que corresponde a até 4 grupos de 2 dígitos, separados por dois pontos. O comprimento da designação dos alelos é dependente da sequência do alelo de cada SNP, comparado ao alelo selvagem que está anotado na sequência referência (NCBI - <http://www.ncbi.nlm.nih.gov> e Ensembl - <http://www.ensembl.org>). Todos os alelos recebem pelo menos um nome de quatro dígitos, o que corresponde aos primeiros dois conjuntos de números. Nomes maiores são atribuídos apenas quando necessário.
2. Os dois primeiros dígitos antes do primeiro “dois pontos” descrevem o tipo, que na maioria das vezes corresponde ao antígeno sorológico que é carregado por um alotipo. O próximo grupo é usado para listar os subtipos, sendo que os números são atribuídos na ordem em que as sequências de DNA foram determinadas. Portanto, os alelos cujos números diferem nos dois primeiros conjuntos de dígitos devem diferir em uma ou mais substituições de nucleotídeos que alterem a sequência de aminoácidos da proteína. Os alelos que diferem apenas por substituições de nucleotídeos sinônimas (também chamadas de substituições silenciosas), isso dentro da sequência codificadora, são distinguidos pela utilização do terceiro conjunto de dígitos. E por fim, o quarto grupo de dígitos é usado para designar alterações em íntrons.

Apesar de as regras para o gene HLA sugerirem opcionalmente a utilização de sufixos para alelos nulos ou para alelos com expressões diferenciais, o presente trabalho não utilizará esses sufixos já que as técnicas utilizadas não permitiram esse nível de análise (Figura 6).

Figura 6: Esquema que indica o significado de cada algarismo na nomenclatura dos alelos do gene *HFE*.

VI. ANÁLISES ESTATÍSTICAS

VI.1 Estimativa das frequências alélicas e genotípicas

Os *loci* autossômicos do tipo SNPs apresentam alelos codominantes, o que permite inferir os genótipos a partir dos respectivos fenótipos. As frequências alélicas (x_i) e genotípicas (X_{ii}) de cada *locus* em cada população amostral foram estimadas por contagem direta, utilizando-se o *software* GENEPOP v.4.2 (RAYMOND e ROUSSET, 1995) (disponível em <http://wbiomed.curtin.edu.au/genepop>), segundo as equações:

$$x_i = \frac{2n_{ii} + \sum n_{ij}}{2n} \quad \text{e} \quad x_{ii} = \frac{n_{ii}}{n}$$

Em que:

x_i é a frequência do alelo i

x_{ii} é a frequência do genótipo ii ;

n_{ii} e n_{ij} correspondem ao número de homozigotos e heterozigotos observados para o alelo i , respectivamente;

n corresponde ao número de indivíduos analisados.

As frequências alélicas e genotípicas para cada *locus* foram estimadas por contagem direta, com uso do *software* GENEPOP v.4.2 (RAYMOND e ROUSSET, 1995). Para comparação das distribuições das frequências foi aplicado o teste exato de Fisher bicaudal (Siegel, 1975), comparando pacientes “com” e “sem” sobrecarga de ferro com diferentes comorbidades (Câncer Hepatocelular ou Hepatite C). Diferenças estatísticas foram consideradas significantes caso a probabilidade de igualdade fosse inferior a 5% ($p < 0,05$).

VI.2 Aderências ao equilíbrio de Hardy-Weinberg

Segundo o teorema de Hardy-Weinberg, as frequências genotípicas esperadas no equilíbrio podem ser estimadas a partir da expansão do seguinte binômio:

$$(x_i + x_j)^2 = x_i^2 + 2x_i x_j + x_j^2$$

Em que:

x_i^2 é a frequência esperada dos homozigotos do alelo i ;

$2x_i x_j$ é a frequência esperada do heterozigoto ij ;

x_j^2 é a frequência esperada dos homozigotos para o alelo j ;

A probabilidade de aderência às proporções teóricas do Equilíbrio de Hardy-Weinberg foi verificada pelo teste exato de Guo e Thompson (1992), utilizando-se o *software* GENEPOP v.4.2 (RAYMOND; ROUSSET, 1995). Foram realizados três testes baseados na hipótese nula de união aleatória dos gametas: teste exato de probabilidade, teste para detecção da deficiência e excesso de heterozigotos.

No teste exato de probabilidade (teste global), o valor de p corresponde à soma de probabilidades de todas as tabelas com probabilidade menor ou igual ao observado.

O segundo e o terceiro são testes mais sensíveis do que o de probabilidade e utilizam uma hipótese alternativa (H_1) de excesso ou de deficiência de heterozigotos, respectivamente.

VI.3 Desequilíbrio de Ligação

A análise de associações par-a-par entre *loci* foi realizada utilizando-se o programa GENEPOP v.4.2 (RAYMOND; ROUSSET, 1995) e ARLEQUIN v. 3.1 (EXCOFFIER; LAVAL; SCHNEIDER, 2005). Ambos os *softwares* baseiam-se na hipótese nula de que a distribuição genotípica em um *locus* é independente da distribuição em outro *locus*. Esta análise foi aplicada para verificar desvios do esperado pela regra de multiplicação entre pares de *loci* localizados em diferentes cromossomos. A palavra “ligação” neste caso não está relacionada com associação física entre alelos de *loci* de um mesmo cromossomo.

Para determinação de quais alelos estão efetivamente ligados e de qual é a intensidade dessa ligação, foram utilizadas duas estatísticas mais usuais para mensurar o desequilíbrio de ligação: r^2 e D' (CASTRO; PEREIRA, 2008; FLINT-GARCIA; THORNSBERRY, 2003).

A estatística r^2 , também descrita na literatura como D_2 , é a correlação ao quadrado entre a presença e ausência de alelos em diferentes *loci* (HILL; ROBERTSON, 1968):

$$r^2 = \frac{(D_{ab})^2}{\pi_A \pi_a \pi_B \pi_b}$$

A estatística D' é calculada da seguinte forma:

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A \pi_a \pi_B \pi_b)}, \text{ se } D_{ab} < 0$$

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A \pi_B \pi_a \pi_b)}, \text{ se } D_{ab} > 0$$

A estatística D' varia entre -1 e 1, enquanto o seu valor absoluto $|D'|$ varia de 0 a 1. Somente quando forem observados todos os quatro possíveis haplótipos (pab , pAb , paB e PAB) é que D' será inferior a 1. Nesse caso, consequentemente, um suposto evento de recombinação terá ocorrido entre os dois *loci*. Sempre que uma das quatro frequências haplotípicas for zero, $|D'|$ será igual a 1, o que ocorre com frequência, quando se trabalha com populações pequenas (WRAY; VISSCHER, 2008). As estatísticas r^2 e D' refletem diferentes aspectos de desequilíbrio de ligação e apresentam comportamentos diversos em condições distintas.

Na estatística r^2 , somente se encontra um valor igual a 1, ou seja, desequilíbrio de ligação completo, quando os dois *loci* em questão apresentam frequências alélicas idênticas. Por outro lado, na estatística D' , sempre que a combinação de um par de alelos não estiver presente na população em estudo, D' será igual a 1 (FLINT-GARCIA *et al.*, 2003; HEDRICK, 1987).

O desequilíbrio de ligação, baseado no D' e r^2 , foi calculado pelo ARLEQUIN v. 3.1 (EXCOFFIER *et al.*, 2005).

VI.4 Diversidade e Diferenciação Genética da Região Codificadora do gene *HFE*

O parâmetro F_{st} (θ_W) é uma das ferramentas utilizadas para se estimar a diversidade genética entre duas ou mais populações. Os cálculos, feitos por meio do *software* ARLEQUIN versão 3.1 (EXCOFFIER *et al.*, 2005), foram realizados utilizando-se método baseado na correlação entre alelos (WEIR; COCKERHAM, 1984). Embora os resultados não sejam exatamente os mesmos, o F_{ST} também pode ser definido em termos de probabilidade de identidade por descendência entre pares de alelos (NEI, 1987):

$$F_{ST} = \frac{(J_S - J_T)}{(1 - J_T)}$$

onde, considerando-se uma população total dividida em k sub populações S , sendo p_{ki} a frequência do alelo i na sub população k e p_i a frequência do alelo i na população total, temos:

$J_S = \frac{\sum_K J_K}{S} = \frac{\sum_K (\sum_I P_{KI}^2)}{S}$ correspondente à probabilidade média de que dois alelos amostrados ao acaso em uma mesma sub população sejam idênticos por descendência e $J_T = \sum_I P_i^2$ à probabilidade de que dois alelos amostrados ao acaso na população total sejam idênticos por descendência (NEI, 1987).

VI.5 Teste Exato de Fisher

O teste exato de Fisher foi usado com a finalidade de detectar possíveis variações nas frequências gênicas individuais, sendo testadas individualmente em subgrupos que compartilham de mesma característica ou de características que são comparáveis clinicamente, para detectar variações relacionadas com a presença ou ausência da sobrecarga de ferro. A confirmação do teste e subsequente calculo do *odds ratio* e intervalo de confiança foi feito utilizando o *software* PRISM® v.5 for Windows.

VI.6 Diversidade Haplotípica

A reconstrução dos haplótipos dessas regiões foi realizada, utilizando-se os algoritmos EM (EXCOFFIER; SLATKIN, 1995) ARLEQUIN v3.1 (EXCOFFIER *et al.*, 2005) e o *software* PHASE v.2 (STEPHENS; SMITH; DONNELLY, 2001). O primeiro usa um processo iterativo que visa à obtenção da estimativa de máxima verossimilhança das frequências haplotípicas, a partir de dados genotípicos multilocus, quando a fase gamética é desconhecida. Já que neste caso, uma contagem de gene simples não é viável, devido ao fato de que vários genótipos são possíveis em indivíduos heterozigotos para mais que um *locus*. O segundo executa um método estatístico Bayesiano para reconstrução dos haplótipos, onde a determinação da fase dos SNPs obtidos para cada indivíduo não leva em consideração nenhuma informação inicial além dos dados genotípicos.

A diversidade h49aplotípicas(h), equivalente à heterozigose em dados d49aplotípi foi estimada considerando a frequência dos haplótipos como definida por Nei (1987). No caso dos genomas h49aplotípi h representa a probabilidade de que dois haplótipos escolhidos aleatoriamente sejam diferentes na amostra. As fórmulas da diversidade h49aplotípicas(h) e seu desvio padrão (DP) são:

$$h = \frac{n}{n-1} \left(1 - \sum_{i=1}^k p_i^2 \right)$$

$$DP = \left\{ \frac{2}{n(n-1)} \{ 2(n-2) [\sum p_i^3 - (\sum p_i^2)^2] + \sum p_i^2 - (\sum p_i^2)^2 \} \right\}^{\frac{1}{2}}$$

Em que:

p_i é a frequência do haplótipo i ;

n é o número de cromossomos da amostra;

k é o número de haplótipos.

Os cálculos foram realizados com o *software* ARLEQUIN v.3.1 (EXCOFFIER *et al.*, 2005).

VI.7 Teste exato de diferenciação populacional

Essa análise testa a hipótese de distribuição aleatória de k (diferentes alelos) entre r (populações). O teste é análogo ao teste exato de Fisher em tabelas de contingência 2x2, estendido para tabelas de contingência $r \times k$. Diversos potenciais estados desta tabela de contingência, mantendo os mesmos totais marginais, são explorados via Cadeia de Markov, sendo então estimada a probabilidade (P) de se observar uma tabela menos ou igualmente provável do que a tabela original, isto sob-hipótese de panmixia. A fim de se assegurar que a Cadeia de Markov se iniciasse em posição aleatória, mil passos de desmemorização foram aplicados antes que os valores de P passassem a ser computados. Uma estimativa do erro padrão do valor de P é feita pela partição do número total de passos da Cadeia de Markov (30 mil) em B *batches*. O teste exato de diferenciação populacional foi realizado com o uso do *software* GENEPOP v.4.2 (RAYMOND; ROUSSET, 1995).

Esse procedimento compara cada *locus* em pares de populações, para determinar se existem diferenças nas frequências alélicas e genotípicas observadas, onde a hipótese nula testada é a de que a distribuição alélica é idêntica entre as populações (RAYMOND; ROUSSET, 1995).

VI.8 Testes de Neutralidade

Ewens-Watterson : Desvios das expectativas de evolução neutra podem ser avaliados pelo teste de neutralidade de Ewens-Watterson (EWENS, 1972; WATTERSON, 1978), por meio de um algoritmo *Markov-Chain Monte Carlo* de re-amostragem que permite a obtenção

de uma distribuição nula (SLATKIN, 1994), para o mesmo tamanho amostral e número de alelos. A homozigose observada assumindo-se o Equilíbrio de Hardy-Weinberg é então comparada à homozigose esperada sob a hipótese nula de neutralidade. Isto permite que se testem as hipóteses alternativas de seleção natural direcionada (homozigose observada maior do que a esperada) ou balanceadora (homozigose observada menor do que a esperada).

Tajima : O teste D de Tajima (1989) examina o relacionamento entre o número de sítios segregantes e a diversidade de nucleotídeos, comparando a diversidade estatística θ_W e π . Dentro do modelo neutro o θ_W e π esperados são iguais e consequentemente o valor esperado de D é zero, sob a neutralidade (NEI; KUMAR, 2000). Um valor positivo de D é evidência de uma vantagem do heterozigoto e um valor negativo é evidência para seleção de um alelo específico (STEPHENS *et al.*, 2001). A significância do D estatístico é testada gerando 99.999 amostragens aleatórias sob a hipótese de neutralidade seletiva e população em equilíbrio, usando um algoritmo de simulação de coalescência. Este teste foi feito usando o ARLEQUIN v.3.1 (EXCOFFIER *et al.*, 2005).

VII. RESULTADOS

VII.1 Descrição das frequências alélicas, genotípicas, heterozigose, equilíbrio de Hardy-Weinberg e Teste Exato de Fisher.

Nas amostras do presente estudo foram encontradas sete alterações na sequência do gene *HFE*, sendo que uma delas é uma deleção de uma base G, localizada no íntron 5 e descrita pela primeira vez neste trabalho. A deleção foi encontrada em um indivíduo da população controle, o qual foi clonado para sua validação (Tabela 4 e Figura 4 e 5), e encontra-se na posição 6:26093478, segundo numeração relativa do site Ensembl - <http://www.ensembl.org>. Essa mutação foi excluída dos cálculos de desequilíbrio de Hardy-Weinberg. As demais alterações já foram descritas em várias populações na literatura, com diversas nomenclaturas, conforme mostrado na Tabela 19 (apêndice 1), para esse trabalho foram adotadas as nomenclaturas mais frequentes na literatura para 4 SNPs mais importantes (H63D, C282Y, IVS2(+4)T>C e IVS4(−44)T>C) sendo os demais SNPs, para os quais não se encontrou uma nomenclatura comum frequente na literatura, optou-se em utilizar somente o número “rs”. Todas as amostras de indivíduos saudáveis ou pacientes com hepatite C (VHC) e/ou carcinoma hepatocelular (CHC) compartilharam o mesmo alelo mais frequente de cada um desses SNPs, exceto a amostra de pacientes com hemocromatose hereditária (HH) que

apresentou o alelo A do SNP C282Y como mais frequente (0,714), assim como o alelo T do SNP (IVS2(+4)T>C) rs2071303 com maior frequência quando comparadas com as demais amostras (0,857) (Tabela 5).

A amostra de CHC com sobrecarga de ferro e com hepatite C como doença de base (CHC C/ SOB) encontra-se em desequilíbrio de Hardy-Weinberg no SNP rs1800562 (C282Y). O teste global mostra que essa amostra ($P=0,043$) e o SNP rs807209 ($P=0,014$) possuem déficit de heterozigotos.

Foi aplicado o teste exato de Fisher, confrontando as amostras de populações de hepatite C, “com” e “sem” sobrecarga de ferro (VHC C/ SOB, VHC S/ SOB). Também foi realizado o teste exato de Fisher na amostra Câncer Hepatocelular com sobrecarga de ferro, confrontando-a com as amostras de populações de hepatite C (“com” e “sem” sobrecarga de ferro). Foi feito o mesmo teste confrontando novamente as amostras de populações de hepatite C (“com” e “sem” sobrecarga de ferro) com a de uma amostra geral Câncer Hepatocelular (CHC GR), constituído do agrupamento da amostra de população Câncer Hepatocelular com sobrecarga de ferro (CHC C/ SOB) e da amostra de população Câncer Hepatocelular de diversas outras etiologias (CHC DIV). Os resultados mostraram que não havia diferenças significativas entre essas populações.

Também foi aplicado o teste exato de Fisher entre a população controle (CTRL) deste trabalho e a amostras de populações de pacientes (hemocromatose hereditária, carcinoma hepatocelular com sobrecarga e carcinoma hepatocelular geral, hepatite C, “com” e “sem” sobrecarga de ferro). Nos resultados obtidos dos testes exatos de Fisher, a comparação entre a amostra de população Controle e amostra de população Hemocromatose Hereditária detectou diferenças significativas. O SNP (IVS2(+4)T>C) rs2071303 homozigoto TT ($P= 0,04$), foi estimado como sendo um genótipo de risco para desenvolver o fenótipo hemocromatose (*Odds Ratio* - OR = 3,91). O SNP clássico (C282Y) rs1800562 também apresentou diferenças significativas, o alelo G ($P< 0,0001$), foi detectado como um alelo protetivo (OR = 0,01), e por consequência o alelo A ($P< 0,001$), foi detectado como um alelo de risco (OR = 53,06). O homozigoto GG ($P= 0,007$) se mostrou, consequentemente, também como um genótipo protetivo (OR = 0,01), e o homozigoto AA ($P< 0,0001$), como um genótipo de risco (OR = 201,00).

Tabela 5. Frequências alélicas, genotípicas e desequilíbrio de Hardy-Weinberg (P) em amostra de população Controle e de pacientes com hepatite C (“com” e “sem” sobrecarga de ferro), carcinoma hepatocelular (com sobrecarga e diverso) e pacientes com hemocromatose hereditária. Ordem dos alelos: Selvagem/Mutado, valores sublinhados apontam combinações com testes exatos de Fisher significativos.

SNPs	ALELO	CONTROLE	HEPATITE C c/ SOBRECARGA de FERRO	HEPATITE C s/ SOBRECARGA de FERRO	CÂNCER HEPATOCELULAR c/ SOBRECARGA de VHC	CÂNCER HEPATOCELULAR DIVERSO	HEMOCROMATOSE HETERITÁRIA
(H63D) rs1799945	C	0,825	0,859	0,864	0,805	0,921	0,893
	G	0,175	0,141	0,136	0,195	0,079	0,107
	GG	0,030	0,042	0,000	0,024	0,000	0,000
	CG	0,290	0,197	0,271	0,341	0,158	0,214
	CC	0,680	0,761	0,729	0,634	0,842	0,786
Valor de <i>P</i>		1,000	0,130	0,580	1,000	1,000	1,000
(IVS2(+4)T>C) rs2071303	T	0,610	0,641	0,669	0,585	0,684	0,857
	C	0,390	0,359	0,331	0,415	0,316	0,143
	TT	<u>0,390¹</u>	0,408	0,424	0,390	0,526	<u>0,714¹</u>
	CC	0,170	0,127	0,085	0,220	0,158	0,000
	TC	0,440	0,465	0,492	0,390	0,316	0,286
Valor de <i>P</i>		0,528	1,000	0,556	0,212	0,295	1,000
rs807209	G	0,035	0,007	0,017	0,000	0,053	0,000
	C	0,965	0,993	0,983	1,000	0,947	1,000
	GG	0,010	0,000	0,000	0,000	0,000	0,000
	CG	0,050	0,014	0,034	0,000	0,105	0,000
	CC	0,940	0,986	0,966	1,000	0,895	1,000
Valor de <i>P</i>		0,103	-	1,000	-	1,000	-
(C282Y) rs1800562	G	<u>0,955²</u>	0,979	0,983	0,902	1,000	<u>0,286²</u>
	A	<u>0,045²</u>	0,021	0,017	0,098	0,000	<u>0,714²</u>
	GG	<u>0,910²</u>	0,958	0,966	0,854	1,000	<u>0,071²</u>
	AG	0,090	0,042	0,034	0,098	0,000	0,429
	AA	<u>0,000²</u>	0,000	0,000	0,049	0,000	<u>0,500²</u>
Valor de <i>P</i>		1,000	1,000	1,000	0,031	-	1,000

SNPs	ALELO	CONTROLE	HEPATITE C c/ SOBRECARGA de FERRO	HEPATITE C s/ SOBRECARGA de FERRO	CÂNCER HEPATOCELULAR c/ SOBRECARGA de VHC	CÂNCER HEPATOCELULAR DIVERSO	HEMOCROMATOSE HETERITÁRIA
rs2794717	G	1,000	1,000	0,992	0,988	1,000	1,000
	A	0,000	0,000	0,008	0,012	0,000	0,000
	GG	1,000	1,000	0,983	0,976	1,000	1,000
	GA	0,000	0,000	0,017	0,024	0,000	0,000
	AA	0,000	0,000	0,000	0,000	0,000	0,000
Valor de <i>P</i>		-	-	-	-	-	-
(IVS4(-44)T>C) rs1800708	T	0,925	0,880	0,907	0,817	0,842	1,000
	C	0,075	0,120	0,093	0,183	0,158	0,000
	CC	0,000	0,014	0,000	0,024	0,000	0,000
	TC	0,150	0,211	0,186	0,317	0,316	0,000
	TT	0,850	0,775	0,814	0,659	0,684	1,000
Valor de <i>P</i>		1,000	1,000	1,000	1,000	1,000	1,000
no íntron 5	G	0,995	1,000	1,000	1,000	1,000	1,000
	Del	0,005	0,000	0,000	0,000	0,000	0,000
	GG	0,990	1,000	1,000	1,000	1,000	1,000
	G del	0,010	0,000	0,000	0,000	0,000	0,000
	del del	0,000	0,000	0,000	0,000	0,000	0,000
Valor de <i>P</i>		-	-	-	-	-	-
Diversidade Média		0,1531	0,1393	0,1339	0,1878	0,1688	0,1250
(Desvio Padrão)		(±0,1123)	(±0,1053)	(±0,1026)	(±0,1321)	(±0,1220)	(±0,1006)

TESTE EXATO DE FISHER

¹SNP (IVS2(+4)T>C) rs2071303

HH X CRLT

TT (*P*=0,04; OR=3,91; IC= 1,14 à 13,34)²SNP (C282Y) rs1800562

HH X CRLT

G (*P*< 0,001; OR= 0,01; IC= 0,006 à 0,05)

A (*P*< 0,001; OR = 53,06; IC= 18,41 à 152,90)

GG (*P*< 0,001; OR = 0,007; IC= 0,0008 à 0,065)

AA (*P*< 0,001; OR = 201,00; IC= 10,44 à 3.871,00)*

*Devido a não ocorrência do genótipo AA na amostra de população Controle o *Odds Ratio* e o Intervalo de Confiança foi calculado utilizando o valor de *default* (0,5) do *software* PRISM® no caso de não ocorrência de evento.

Figura 7: Cromatograma do SNP novo no íntron 5 identificada em um indivíduo da amostra da população controle. A deleção apresenta-se em heterozigose, portanto as bases posteriores se apresentam embaralhadas no sequenciamento normal (A). Isto não acontece com a fita clonada (B).

VII.3 Desequilíbrio de Ligação

O teste de desequilíbrio de ligação entre os SNPs do gene *HFE* foi realizado pelo *software* GENEPOP v.4.2 (RAYMOND; ROUSSET, 1995) (Tabela 6) e foi realizado sem a deleção no íntron 5, pois esta variação foi encontrada em um único indivíduo da amostra da população controle. O desequilíbrio também foi calculado pelo *software* HAPLOVIEW v.4.1 (BARRETT *et al.*, 2005), no qual, apesar de ter sido incluído o SNP no íntron 5, para que o mesmo tenha representação gráfica, o *software* não o utilizou para os cálculos, devido sua baixa frequência (Figura 8).

Os resultados mais notórios foram o desequilíbrio de ligação entre os SNPs: i) (H63D) rs1799945 e (IVS2(+4)T>C) rs2071303 em quase todas as populações, analisadas separadamente, como também no teste global; ii) rs2071303 e (IVS4(-44)T>C) rs1800708, na amostra de CHC com sobrecarga de ferro e com hepatite C como doença de base, na amostra de câncer hepatocelular diverso e na amostra de hemocromatose hereditária, como também no teste global; e iii) (IVS2(+4)T>C) rs2071303 e (C282Y) rs1800562 na amostra de população Hemocromatose Hereditária.

O HAPLOVIEW, em concordância com o GENEPOP, também demonstrou forte desequilíbrio entre (H63D) rs1799945 e (IVS2(+4)T>C) rs2071303 (LOD = 95%), e também em menor intensidade (IVS2(+4)T>C) rs2071303 e (C282Y) rs1800562 (LOD = 77%). Outro

resultado em consonância com o GENEPOP é o forte desequilíbrio (IVS2(+4)T>C) rs2071303 e (IVS4(-44)T>C) rs1800708 (LOD= 90%). Um dado notório é a ausência de desequilíbrio de ligação entre os SNPs clássicos (H63D) rs1799945 e (C282Y) rs1800562, concordante com o teste global do GENEPOP.

Tabela 6. Desequilíbrio de ligação, somente os pares com valores significativos entre os SNPs do gene HFE, em cada uma das amostras populacionais e no teste global. Grifados o par recorrente em todas as populações.

SNP 1	SNP 2	Valor <i>P</i>
POPULAÇÃO CONTROLE		
<u>rs1799945 (H63D)</u>	<u>rs2071303 (IVS2(+4)T>C)</u>	< 0,001
rs2071303	rs1800708 (IVS4(-44)T>C)	< 0,001
HEPATITE C com SOBRECARGA		
<u>rs1799945 (H63D)</u>	<u>rs2071303 (IVS2(+4)T>C)</u>	< 0,001
rs2071303 (IVS2(+4)T>C)	rs1800708 (IVS4(-44) T→C)	< 0,001
HEPATITE sem SOBRECARGA		
<u>rs1799945 (H63D)</u>	<u>rs2071303 (IVS2(+4)T>C)</u>	< 0,001
rs2071303 (IVS2(+4)T>C)	rs1800708 (IVS4(-44)T>C)	< 0,001
CÂNCER HEPATOCELULAR com SOBRECARGA		
doença de base a HEPATITE C		
<u>rs1799945 (H63D)</u>	<u>rs2071303 (IVS2(+4)T>C)</u>	0,001
rs2071303 (IVS2(+4)T>C)	rs1800708 (IVS4(-44)T>C)	0,001
CÂNCER HEPATOCELULAR DIVERSO		
rs2071303 (IVS2(+4)T>C)	rs1800708 (IVS4(-44)T>C)	0,015
HEMOCROMATOSE HEREDITÁRIA		
<u>rs1799945 (H63D)</u>	<u>rs2071303 (IVS2(+4)T>C)</u>	0,010
rs1799945 (H63D)	rs1800562 (C282Y)	0,041
rs2071303 (IVS2(+4)T>C)	rs1800562 (C282Y)	0,034
TESTE GLOBAL de DESEQUILÍBRIO		
<u>rs1799945 (H63D)</u>	<u>rs2071303 (IVS2(+4)T>C)</u>	< 0,001
rs807209	rs1800562	< 0,001
rs1799945 (H63D)	rs2794717	< 0,001
rs2071303 (IVS2(+4)T>C)	rs2794717	< 0,001
rs807209	rs2794717	< 0,001
rs1800562 (C282Y)	rs2794717	< 0,001
rs2794717	rs1800708 (IVS4(-44)T>C)	< 0,001

Figura 8: Desequilíbrio de ligação entre os SNPs do gene *HFE*, imagem gerada no *software* HAPLOVIEW v.4.1 (BARRETT *et al.*, 2005). Áreas em vermelho escuro forte desequilíbrio de ligação ($\text{LOD} \geq 2$; $D' = 1$); em vermelho claro indica desequilíbrio moderado ($\text{LOD} \geq 2$, $D' < 1$); azul indica fraco desequilíbrio ($\text{LOD} < 2$, $D' = 1$); e branco indica a ausência de desequilíbrio ($\text{LOD} < 2$, $D' < 1$). Valores de D' diferentes de 1,00 são representados em porcentagem dentro dos quadrados. LOD (Log do Odds); D' (correlação entre pares de SNPs).

Considerando que: i) os SNPs H63D, (IVS2(+4)T>C) e C282Y são os únicos que se apresentaram em desequilíbrio na amostra de população HH; ii) o par de SNPs H63D e rs2071303 apresentaram-se em desequilíbrio em quase todas as populações (exceção câncer hepatocelular diverso) na análise par a par; iii) O resultados do *software* HAPLOVIEW confirmaram os desequilíbrios detectados pelo *software* GENEPOP; iv) os SNPs H63D e C282Y foram abundantemente relacionados com hemocromatose hereditária na literatura, foi realizada uma terceira análise de desequilíbrio de ligação, nas amostras de populações controle e de hemocromatose hereditária, utilizando uma tabela de contingência D' e r^2 (Tabela 7), para analisar quais alelos estariam ligados considerando esses três SNPs.

Os resultados demonstraram que tanto na amostra de população Controle, como na amostra Hemocromatose Hereditária existe um desequilíbrio de ligação entre os alelos selvagens dos SNPs H63D e (IVS2(+4)T>C) ($D'=1,000$ em ambas as amostras e $r^2=0,3318$ e $0,7200$, consecutivamente). Em ambos os casos, houve a completa ausência do recombinante mutado do SNP H63D, com o alelo selvagem do SNP (IVS2(+4)T>C). Outro resultado pertinente é a ligação tanto do alelo selvagem do H63D e do (IVS2(+4)T>C) com o alelo mutado do SNP C282Y ($D'=1,000$, em ambos os casos, e $r^2=0,3000$ e $0,4267$).

Tabela 7. Desequilíbrio de ligação entre o SNPs H63D, (IVS2(+4)T>C) e C282Y do gene HFE nas amostras de populações controle e hemocromatose hereditária.

CONTROLE E HEMOCROMATOSE HEREDITÁRIA										
		FREQUÊNCIA OBSERVADA	FREQUÊNCIA ESPERADA	VALOR DE DESIGUIBRIU NORMALIZADO (D')		VALOR DE RECOMBINAÇÃO NORMALIZADO (r^2)		VALOR DE Qui^2	VALOR DE P de Qui^2	
CONTROLE										
(H63D) rs1799945 x (IVS2(+4)T>C) rs2071303										
H63D		IVS2(+4)T>C								
	T	C	T	C	T	C				
	C	122	43	100,65	64,35	1,0000	-1,0000	0,3318	66,3559	0,0000
	G	0	35	21,35	13,65	-1,0000	1,0000	0,3318	66,3559	0,0000
(H63D) rs1799945 x (C282Y) rs1800562										
H63D		C282Y								
	G	A	G	A	G	A				
	C	156	9	157,57	7,42	-1,000	1,000	0,0100	1,9990	0.1574
	G	35	0	33,42	1,57	1,000	-1,000	0,0100	1,9990	0.1574
(IVS2(+4)T>C) rs2071303 x (C282Y) rs1800562										
IVS2(+4)T>C		C282Y								
	G	A	G	A	G	A				
	T	114	8	116,51	5,49	-0,7151	0,7151	0,0154	3,0811	0,0792
	C	77	1	74,49	3,51	0,7151	-0,7151	0,0154	3,0811	0,0792
HEMOCROMATOSE HEREDITÁRIA										
(H63D) rs1799945 x (IVS2(+4)T>C) rs2071303										
H63D		IVS2(+4)T>C								
	T	C	T	C	T	C				
	C	24	1	21,43	3,57	1,0000	-1,0000	0,7200	20,1600	0,0000
	G	0	3	2,57	0,43	-1,0000	1,0000	0,7200	20,1600	0,0000
(H63D) rs1799945 x (C282Y) rs1800562										
H63D		C282Y								
	G	A	G	A	G	A				
	C	5	20	7,14	17,86	-1,0000	1,0000	0,3000	8,4000	0,0038
	G	3	0	0,86	2,14	1,0000	-1,0000	0,3000	8,4000	0,0038
(IVS2(+4)T>C) rs2071303 x (C282Y) rs1800562										
IVS2(+4)T>C		C282Y								
	G	A	G	A	G	A				
	T	4	20	6,86	17,14	-1,0000	1,0000	0,4267	11,6667	0,0006
	C	4	0	1,14	2,86	1,0000	-1,0000	0,4267	11,6667	0,0006

Identificação do SNP /alelo mais frequente segundo NCBI (<http://www.ncbi.nlm.nih.gov>) e ENSEMBL (<http://www.ensembl.org>): rs1799945 (H63D) / C; rs2071303 (IVS2(+4)T>C) / T; rs1800562 (C282Y) / G.

Nessa análise de desequilíbrio observou-se um dado interessante que foi a ocorrência do haplótipo G-C-G do entre os SNPs H63D, IVS2(+4)T>C e C282Y, demonstrando que o alelo mutado do SNP H63D sempre ocorre com o mutado do SNP IVS2(+4)T>C e com o selvagem do SNP C282Y (Figura 9).

Figura 9: Esquema das formações Haplótípicas possíveis entre os SNPs H63D, IVS2(+4)T>C e C282Y, nas amostras de população Controle e Hemocromatose Hereditária. Em vermelho destaca-se a combinação alélica preferencial e exclusiva que ocorre entre os 3 SNPs (H63D G / IVS2(+4) C / C282Y G).

VII.4 Diversidade e Diferenciação Genética da Região Codificadora do gene *HFE*

As heterozigoses esperada e observada nas diversas amostras de populações não demonstraram diferenças significantes (Tabela 8), porém o teste global de diferenciação das amostras de populações conseguiu diferenciá-las ($P < 0,001$). Os cálculos de distância genética, entre os 6 pares de amostras de populações, baseados no valor de F_{ST} , demonstraram que apenas a amostra de população Hemocromatose Hereditária se diferencia das demais ($P = 0,00$) (Tabela 9).

Tabela 8. Heterozigose observada e esperada nas amostras de populações (realizados apenas nos loci que se apresentam polimórficos em cada amostra).

SNPs	HETEROZIGOSE OBSERVADA	HETEROZIGOSE ESPERADA
POPULAÇÃO CONTROLE		
rs1799945 (H63D)	0,29	0,29
rs2071303 (IVS2(+4)T>C)	0,44	0,47
rs807209	0,05	0,06
rs1800562 (C282Y)	0,09	0,08
rs1800708 (IVS4(-44)T>C)	0,15	0,13
no íntron 5	0,01	0,01
MÉDIA	0,14	0,15
DESVIO PADRAO	0,15	0,16
HEPATITE c/ SOBRECARGA de FERRO		
rs1799945 (H63D)	0,19	0,24
rs2071303 (IVS2(+4)T>C)	0,46	0,46
rs807209	0,01	0,01
rs1800562 (C282Y)	0,04	0,04
rs1800708 (IVS4(-44)T>C)	0,21	0,21
MÉDIA	0,13	0,13
DESVIO PADRAO	0,15	0,16
HEPATITE s/SOBRECARGA de FERRO		
rs1799945 (H63D)	0,27	0,23
rs2071303 (IVS2(+4)T>C)	0,49	0,44
rs807209	0,03	0,03
rs1800562 (C282Y)	0,03	0,03
rs2794717	0,01	0,01
rs1800708 (IVS4(-44)T>C)	0,18	0,17
MÉDIA	0,14	0,13
DESVIO PADRAO	0,16	0,15
CÂNCER HEPATOCELULARc/ VHC e SOBRECARGA de FERRO		
rs1799945 (H63D)	0,33	0,33
rs2071303 (IVS2(+4)T>C)	0,33	0,49
rs1800562 (C282Y)	0,00	0,15
rs2794717	0,04	0,04
rs1800708 (IVS4(-44)T>C)	0,25	0,28
MÉDIA	0,13	0,18
DESVIO PADRAO	0,14	0,17
CÂNCER HEPATOCELULAR DIVERSO		
rs1799945 (H63D)	0,29	0,25
rs2071303 (IVS2(+4)T>C)	0,40	0,46
rs807209	0,03	0,03
rs1800562 (C282Y)	0,14	0,13
rs1800708 (IVS4(-44)T>C)	0,33	0,28
MÉDIA	0,17	0,16
DESVIO PADRAO	0,15	0,16
HEMOCROMATOSE HEREDITÁRIA		
rs1799945 (H63D)	0,21	0,19
rs2071303 (IVS2(+4)T>C)	0,28	0,25
rs1800562 (C282Y)	0,42	0,42
MÉDIA	0,13	0,12
DESVIO PADRAO	0,16	0,15

Tabela 9. Matriz de distância das amostras de populações baseada nos valores de FST (diagonal inferior: valores de P – diagonal superior: erro padrão de cada respectivo valor P.

AMOSTRA	CONTROLE	HCV C/SOB.	HCV S/SOB.	CHC C/SOB	CHC DIV	HH
CONTROLE		-0,000	-0,000	-0,000	-0,001	0,322
HCV C/ SOB.	0,435		-0,005	-0,000	-0,007	0,352
HCV S/SOB	0,369	0,874		0,007	-0,004	0,355
CHC C/ VHC	0,441	0,422	0,209		-0,012	0,293
CHC DIV	0,415	0,741	0,489	0,823		0,296
HH	<0,0001	<0,0001	<0,0001	<0,0001	<0,0001	

VII.5 Dedução da Fase Meiótica dos Haplótipos no Gene *HFE*

Para estimar a fase meiótica dos haplótipos foram utilizados os *softwares* ARLEQUIN v3.1 e PHASE v.2. Foram feitas várias simulações na quais as amostras de populações foram inseridas como um único grupo e estruturadas em seus diversos subgrupos (controles, pacientes com ou sem sobrecarga de ferro, câncer hepatocelular, hepatite C, hemocromatose hereditária). Posteriormente, os haplótipos gerados em ambos os *softwares* foram comparados de forma a obter o haplótipo mais provável do indivíduo (Tabela 10). Apenas um único haplótipo em um indivíduo da amostra Câncer Hepatocelular Diverso, gerado pelo método EBL (o único considerado para esse estudo), na simulação gerada pelo *software* ARLEQUIN, se mostrou discordante. Todas as outras simulações foram idênticas, independente do *software* ou da estrutura populacional inserida. O haplótipo discordante apresentou baixa probabilidade de ocorrência (0,092, calculado pelo *software* PHASE) e foi exclusivo deste indivíduo, sendo assim, foi optado pelo outro haplótipo de maior probabilidade (0,908).

Tabela 10. Identificação e constituição haplotípica, para cada SNP nas populações controle e de pacientes (em negrito as bases mutadas; com fundo cinza as regiões exônicas).

Identificação do Haplótipo	Haplótipo						
	*1	*2	*3	*4	*5	*6	*7
<i>HFE</i> *01:01	C	T	C	G	G	T	G
<i>HFE</i> *01:01:01:01	C	C	C	G	G	T	G
<i>HFE</i> *01:01:01:02	C	T	G	G	G	T	G
<i>HFE</i> *01:01:01:03	C	T	C	G	A	T	G
<i>HFE</i> *01:01:01:04	C	T	C	G	G	C	G
<i>HFE</i> *01:01:01:05	C	C	C	G	G	C	G
<i>HFE</i> *01:01:01:06	C	C	C	G	G	C	del
<i>HFE</i> *01:02	G	T	C	G	G	T	G
<i>HFE</i> *01:02:01:01	G	C	C	G	G	T	G
<i>HFE</i> *01:02:01:05	G	C	C	G	G	C	G
<i>HFE</i> *01:03	C	T	C	A	G	T	G
<i>HFE</i> *01:03:01:01	C	C	C	A	G	T	G
<i>HFE</i> *01:04:01:01	G	C	C	A	G	T	G

Identificação do SNP/Alelo mais frequente segundo NCBI (<http://www.ncbi.nlm.nih.gov>) e ENSEMBL (<http://www.ensembl.org>): *1- rs1799945 (H63D) / C; *2 - rs2071303 (IVS2(+4)T>C) / T; *3- rs807209 / G; *4- rs1800562 (C282Y) / G; *5-rs2794717/ G; *6- rs1800708 (IVS4(-44)T>C) / T; *7- mutação no intron 5 / G

VII.6 Teste de Neutralidade

Para investigar se a seleção natural pode estar exercendo pressão significativa sobre as frequências alélicas e genotípicas, foi realizado na população controle o teste de neutralidade (Ewens-Watterson e Tajima), utilizando o *software* ARLEQUIN v3.1 (EXCOFFIER *et al.*, 2005). No teste Ewens-Watterson, considera-se que se encontrada homozigose observada menor do que a esperada, pode-se supor uma seleção, e caso contrário, em que encontrada homozigose observada maior do que a esperada, a hipótese é uma pressão de seleção balanceadora, desde que P seja significativo. Já o teste de neutralidade de D de Tajima leva em conta a diversidade da sequência, com P significativo. Os resultados desse trabalho não detectaram qualquer tipo de pressão de seleção na amostra de população Controle ($P < 0,05$).

VII.7 Descrição das frequências haplotípicas e teste Exato de Fisher para haplótipo e diplótipos do gene *HFE*.

Após a simulação dos haplótipos verificou-se que havia pronunciada discrepância na frequência dos haplótipos *HFE**01:01 e *HFE**01:03 na amostra de população Hemocromatose Hereditária. Devido a este fato, foi aplicado o teste exato de Fisher, considerando a frequência alélica, nas amostras de pacientes entre si: populações hepatite com sobrecarga de ferro, hepatite sem sobrecarga de ferro, câncer hepatocelular proveniente de hepatite C e câncer hepatocelular geral. Este teste foi realizado da mesma forma feita anteriormente com os SNPs. Os resultados mostraram que não havia diferenças significativas entre essas populações.

Esse mesmo teste foi aplicado entre a população controle e as amostras de população de pacientes (com hemocromatose hereditária, hepatocarcinoma e carcinoma hepatocelular geral e hepatite C com e sem sobrecarga de ferro).

Na análise: população Controle x Hemocromatose Hereditária dois haplótipos apresentaram diferenças significativas, o *HFE**01:01 como um alelo protetivo ($P = < 0,0001$ OR = 0,14) e o *HFE**01:03 como um alelo de risco ($P = < 0,0001$; OR = 60,00).

Nas análises de diplótipos, na comparação da amostra de população Hemocromatose Hereditária com a amostra de população Controle, os diplótipos *HFE**01:01 / *HFE**01:03 ($P = 0,03$; OR = 7,2) e *HFE**01:03 / *HFE**01:03 ($P < 0,001$; OR = 174,20) foram detectados como diplótipos de risco. Outro diplótipo detectado como sendo de risco foi o *HFE**01:02:01:01 / *HFE**01:01:01:05 ($P = 0,02$; OR = 14,14), encontrado na comparação entre as amostras de população de carcinoma com sobrecarga de ferro e controle.

Tabela 11. Frequência haplotípica do gene HFE por população, teste exato de Fisher (em negrito: as frequências dos haplótipos discrepantes em comparações com outras amostras; grifado: combinações com valores significativos no teste exato de Fisher).

ALELO	SEM SOBRECARGA de FERRO				POPULAÇÃO								DIVERSO					
					COM SOBRECARGA de FERRO													
							CÂNCER								CÂNCER		TOTAL	
	CONTROLE		HEPATITE C		HEPATITE C		HEPATOCELULAR de HEPATITE C		HEMOCROMATOSE HEREDITÁRIA		TOTAL		HEPATOCELULAR		POR ALELO			
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%		
HFE*01:01	107 ¹	0,54	74	0,63	86	0,61	23	0,48	4 ¹	0,14	113	0,52	40	0,56	334	0,55		
HFE*01:01:01:01	28	0,14	12	0,10	15	0,11	3	0,06	1	0,04	19	0,09	4	0,06	63	0,10		
HFE*01:01:01:02	7	0,04	2	0,02	1	0,01	0	0,00	0	0,00	1	0,00	2	0,03	12	0,02		
HFE*01:01:01:03	0	0,00	1	0,01	0	0,00	1	0,02	0	0,00	1	0,00	0	0,00	2	<0,01		
HFE*01:01:01:04	0	0,00	0	0,00	1	0,01	1	0,02	0	0,00	2	0,01	1	0,01	3	<0,01		
HFE*01:01:01:05	13 ²	0,07	11	0,09	16	0,11	7 ²	0,15	0	0,00	23	0,11	12	0,17	59	0,10		
HFE*01:01:01:06	1	0,01	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	1	<0,01		
HFE*01:02	0	0,00	1	0,01	0	0,00	0	0,00	0	0,00	0	0,00	1	0,01	2	<0,01		
HFE*01:02:01:01	34 ²	0,17	15	0,13	20	0,14	9 ²	0,19	3	0,11	32	0,15	8	0,11	89	0,14		
HFE*01:02:01:05	1	0,01	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	1	<0,01		
HFE*01:03	8 ¹	0,04	2	0,02	3	0,02	3	0,06	20 ¹	0,71	26	0,12	4	0,06	40	0,06		
HFE*01:03:01:01	1	0,01	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	1	<0,01		
HFE*01:04:01:01	0	0,00	0	0,00	0	0,00	1	0,02	0	0,00	1	0,00	0	0,00	1	<0,01		
TOTAL POR POP ³	200		118		142		48		28		218		72		608			
NO. DE INDIV ⁴	100		59		71		24		14		109		36		304			

TESTE EXATO DE FISHER

¹HH X CRLT

<i>HFE</i> *01:01	(<i>P</i> < 0,001; OR= 0,14; IC= 0,04 à 0,43)
<i>HFE</i> *01:03	(<i>P</i> < 0,001; OR= 60,00; IC=20,31 à 177,20)
<i>HFE</i> *01:01 / <i>HFE</i> *01:03	(<i>P</i> = 0,03; OR= 7,20; IC= 1,40 à 36,85)
<i>HFE</i> *01:03 / <i>HFE</i> *01:03	(<i>P</i> < 0,001; OR=174;20; IC= 8,92 à 400,00)*

²CHC C/ SOB X CRTL

<i>HFE</i> *01:02:01:01 / <i>HFE</i> *01:01:01:05	(<i>P</i> =0,02; OR=14,14; IC=1,40 à 142,80)
---	---

*Devido a não ocorrência do diplótipo *HFE**01:03 / *HFE**01:03 na amostra de população Controle o Odds Ratio e o Intervalo de Confiança foi calculado utilizando o valor de default (0,5) do software PRISM® no caso de não ocorrência de evento.

³Total de alelos por população

⁴Número de indivíduos por população

VII.8 Teste exato de Fisher dos SNPs encontrados no trabalho comparado com diversas populações

Com o intuito de comparar, pelo do teste Exato de Fisher, a frequência dos alelos e genótipos encontrados na amostra de população Controle, com amostra de populações mundiais, foi feito um levantamento no site 1000genome (<http://www.1000genomes.org/>)

O dados disponíveis no citado site se organizam em 4 superpopulações:

- AFR – Africanos
- AMR– Americanos de diversas origens
- ASN – Leste asiático
- EUR – Europeus

Essas superpopulações são estruturadas nas seguintes subpopulações conforme a Tabela 12.

Tabela 12. Código de identificação das subpopulações que possuem seus dados depositados no site 1000genome (<http://www.1000genomes.org/>).

Código da população	Descrição da População	Código da Superpopulação
MUD	População mundial	
CHB	Chineses Han em Bejing (Pequim), China	ASN
JPT	Japoneses em Tóquio, Japão	
CHS	Chineses Han do Sul, China	
CEU	Residentes de Utah com ancestralidade do norte e oeste da Europa	EUR
TSI	Toscanos na Itália	
FIN	Finlandeses na Finlândia	
GBR	Britânicos na Inglaterra e Escócia	
IBS	População Ibérica na Espanha	
YRI	Iorubá em Ibadan, Nigéria	AFR
LWK	Luhya em Webuye, Kenya	
ASW	Americanos de ancestralidade Africana no sudoeste dos EUA	AMR
MXL	Americanos de ancestralidade Mexicana proveniente de Los Angeles. EUA	
PUR	Porto Riquenhos de Porto Rico	
CLM	Colombianos de Medellin, Colômbia	

Os dados oriundos dessa análise foram tabulados com o intuito de verificar o quanto a população brasileira se distingue das populações mundiais disponíveis (Tabela 13) e gerar um gráfico com a soma dos *loci* semelhantes e dos *loci* diferentes (Figura 10).

No SNP rs1799945 (H63D) observou-se diferença significativa das seguintes para as seguintes populações: População Mundial (MUD), Africanos (AFR), Asiáticos (ASN), Americanos de ancestralidade Africana (ASW), Chineses Han em Bejing (CHB), Chineses Han do Sul (CHS), Japoneses em Tóquio (JPT), Luhya em Webuye (LKW), Porto Riquenhos (PUR) e Iorubá em Ibadan (YRI) ($P < 0,0001$ à $0,0451$).

O SNP rs2071303 (IVS2(+4)T>C) distinguiu a amostra de população Controle das seguintes populações: Asiáticos (ASN), Chineses Han em Bejing (CHB), Chineses Han do Sul (CHS), Japoneses em Tóquio (JPT) ($P < 0,0001$ à $0,0286$).

O SNP rs807209 foi capaz de distinguir, utilizando as frequências alélicas e genotípicas, as seguintes populações: Africanos (AFR), Asiáticos (ASN), Americanos de ancestralidade Africana (ASW) e Iorubá em Ibadan (YRI) ($P < 0,0001$ à $0,0050$). Esse mesmo SNP, assumindo somente a frequência alélica, separou as populações Chineses Han do Sul (CHS), Europeus (EUR), Finlandeses na Finlândia (FIN), População Ibérica na Espanha (IBS), Japoneses em Tóquio (JPT) e Luhya em Webuye (LWK) ($P < 0,0001$ à $0,0024$).

O SNP rs1800562 (C282Y) foi o que apresentou maior poder de separação, sendo pronunciadamente diferente das populações: População Mundial (MUD), Africanos (AFR), Asiáticos (ASN), Chineses Han em Bejing (CHB), Chineses Han do Sul (CHS), Japoneses em Tóquio (JPT), Luhya em Webuye (LWK) e Iorubá em Ibadan (YRI) ($P < 0,0001$ à $0,0040$). Além disso esse SNP se mostrou informativo, quando considerado apenas os dados genotípicos, nas comparações com as populações: Americanos (AMR), Americanos de ancestralidade Africana (ASW), Utah com ancestralidade do norte e oeste da Europa (CEU), Finlandeses na Finlândia (FIN), Porto Riquenhos (PUR), Toscanos na Itália (TSI) ($P < 0,0001$ à $0,0144$).

De forma oposta, o *locus* rs279417 apresenta poucas diferenças quando correlacionadas as populações, sendo que estas diferenças são estatisticamente significativas para os genótipos GG ou GA (Asiáticos (ASN), Chineses Han em Bejing (CHB), Chineses Han no Sul (CHS), Colombianos de Medellin (CLM), População Ibérica na Espanha (IBS), Japoneses em Tóquio (JPT), Luhya em Webuye (LWK), Americanos de ancestralidade Mexicana (MXL) e Iorubá em Ibadan (YRI) ($P < 0,0001$).

Finalmente, o *locus* rs1800708 (IVS4(-44)T>C) também se mostrou bem distinto nas diversas populações, sendo que o mesmo se diferenciou completamente das populações: População Mundial (MUD), Asiáticos (ASN), Chineses Han em Bejing (CHB), Chineses Han no Sul (CHS), Finlandeses na Finlândia (FIN) e Japoneses em Tóquio (JPT) ($P < 0,0001$ à $0,0209$). O SNP se diferenciou-se parcialmente pela frequência alélica as populações Americanas (AMR) e Americanos de ancestralidade Mexicana (MXL) ($P = 0,0076$ à $0,0236$). E também se diferenciou parcialmente pelas frequências genotípicas das populações: Utah com ancestralidade do norte e oeste da Europa (CEU), Colombianos de Medellin (CLM), População Ibérica na Espanha (IBS) ($P < 0,0001$ à $0,0038$).

Tabela 13. Teste exato Fisher entre a população controle e as populações disponíveis no site 1000genomes (<http://www.1000genomes.org/>) calculados para os loci encontrados nesse trabalho, sombreado valores significativos.

	MUD	AFR	AMR	ASN	ASW	CEU	CHB	CHS	CLM	EUR	FIN	GBR	IBS	JPT	LWK	MXL	PUR	TSI	YRI
rs1799945																			
C	0,0001	0,0000	0,0989	0,0000	0,0009	0,2497	0,0000	0,0000	0,6422	0,4463	0,1111	1,0000	0,6030	0,0000	0,0000	0,2146	0,0451	1,0000	0,0000
G	0,0001	0,0000	0,0989	0,0000	0,0009	0,2497	0,0000	0,0000	0,6422	0,4463	0,1111	1,0000	0,6030	0,0000	0,0000	0,2146	0,0451	1,0000	0,0000
CC	0,0001	0,0000	0,0865	0,0000	0,0011	0,3283	0,0000	0,0000	0,5933	0,5368	0,0745	0,7576	0,5459	0,0000	0,0000	0,2209	0,0378	1,0000	0,0000
CG	0,0004	0,0000	0,1038	0,0000	0,0053	0,5040	0,0000	0,0000	0,4669	0,6134	0,0924	0,5311	0,3558	0,0000	0,0000	0,2833	0,0500	0,8752	0,0000
GG	0,1444	0,0236	0,7023	0,0551	0,2896	0,6259	0,2463	0,6212	1,0000	0,7057	1,0000	0,6236	1,0000	0,2486	0,2463	1,0000	1,0000	0,7195	0,2491
rs2071303																			
C	0,0751	0,6652	0,7865	0,0000	1,0000	0,4497	0,0000	0,0000	0,3394	0,5652	0,6746	1,0000	0,6860	0,0000	0,5321	0,9084	0,5466	0,5337	0,7496
T	0,0751	0,6652	0,7865	0,0000	1,0000	0,4497	0,0000	0,0000	0,3394	0,5652	0,6746	1,0000	0,6860	0,0000	0,5321	0,9084	0,5466	0,5337	0,7496
CC	0,2562	0,6282	0,4882	0,0000	1,0000	0,4042	0,0001	0,0000	0,4934	0,2432	0,0906	1,0000	1,0000	0,0000	0,3095	0,8379	0,3538	0,3087	1,0000
CT	0,7533	1,0000	0,5344	0,9068	0,8716	0,8822	1,0000	1,0000	1,0000	0,3711	0,1138	1,0000	0,4007	0,7698	0,5691	0,6308	0,0286	0,5688	0,6571
TT	0,1180	0,8092	0,8987	0,0000	1,0000	0,6553	0,0000	0,0000	0,5085	1,0000	0,6565	1,0000	0,5632	0,0000	0,8850	0,7468	0,1116	1,0000	0,6559
rs807209																			
C	1,0000	0,0000	0,2685	0,0005	0,0050	0,0745	0,0678	0,0000	1,0000	0,0024	0,0000	0,1812	0,0000	0,0000	0,0004	0,3266	0,2674	0,0677	0,0002
G	1,0000	0,0000	0,2685	0,0005	0,0050	0,0745	0,0678	0,0000	1,0000	0,0024	0,0000	0,1812	0,0000	0,0000	0,0004	0,3266	0,2674	0,0677	0,0002
CC	0,8387	0,0000	0,5541	0,0015	0,0025	0,1265	0,1186	1,0000	1,0000	0,0073	1,0000	0,2853	1,0000	1,0000	0,0003	0,4795	0,4225	0,1185	0,0000
CG	0,6732	0,0000	0,7598	0,0051	0,0034	0,2204	0,2119	1,0000	0,7291	0,0223	1,0000	0,4500	1,0000	1,0000	0,0003	0,7040	0,4236	0,2117	0,0000
GG	0,2960	1,0000	0,3559	0,2591	1,0000	1,0000	1,0000	1,0000	1,0000	0,2088	1,0000	1,0000	1,0000	1,0000	0,6174	1,0000	1,0000	1,0000	1,0000
rs1800562																			
A	0,0363	0,0004	0,0564	0,0000	0,2173	0,8104	0,0036	0,0036	0,0968	1,0000	1,0000	0,6448	0,6296	0,0040	0,0036	0,0958	1,0000	0,6004	0,0041
G	0,0363	0,0004	0,0564	0,0000	0,2173	0,8104	0,0036	0,0036	0,0968	1,0000	1,0000	0,6448	0,6296	0,0040	0,0036	0,0958	1,0000	0,6004	0,0041
AA	1,0000	1,0000	0,0922	0,0000	0,1421	0,0007	0,0000	0,0000	0,3750	1,0000	0,0025	0,4709	0,0141	0,0000	0,0000	0,3976	0,0147	0,0136	0,0000
AG	0,0309	0,0000	0,0000	0,0000	0,0000	0,0000	0,0033	0,0032	0,0144	1,0000	0,0000	1,0000	0,5981	0,0036	0,0033	0,0119	0,0270	0,0032	0,0037
GG	0,0325	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0915	1,0000	0,0000	0,8093	0,6234	0,0000	0,0000	0,0528	1,0000	0,5929	0,0000

	MUD	AFR	AMR	ASN	ASW	CEU	CHB	CHS	CLM	EUR	FIN	GBR	IBS	JPT	LWK	MXL	PUR	TSI	YRI
rs2794717																			
A	1,0000	1,0000	0,5407	1,0000	0,3789	0,2104	1,0000	1,0000	1,0000	0,5898	0,4819	0,4709	1,0000	1,0000	1,0000	1,0000	0,1252	0,4949	1,0000
G	1,0000	1,0000	0,5407	1,0000	0,3789	0,2104	1,0000	1,0000	1,0000	0,5898	0,4819	0,4709	1,0000	1,0000	1,0000	1,0000	0,1252	0,4949	1,0000
AG	1,0000	1,0000	0,5399	0,0000	0,3789	0,2098	0,0000	0,0000	0,0000	0,5888	0,4819	0,4709	0,0000	0,0000	0,0000	0,0000	0,1244	0,4949	0,0000
GG	1,0000	1,0000	0,5399	0,0000	0,3789	0,2098	0,0000	0,0000	0,0000	0,5888	0,4819	0,4709	0,0000	0,0000	0,0000	0,0000	0,1244	0,4949	0,0000
AA	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
rs1800708																			
C	0,0000	0,5529	0,0076	0,0000	0,8265	0,7042	0,0000	0,0000	0,0377	0,3425	0,0020	1,0000	1,0000	0,0000	0,2955	0,0191	0,0732	0,6912	0,5799
T	0,0000	0,5529	0,0076	0,0000	0,8265	0,7042	0,0000	0,0000	0,0377	0,3425	0,0020	1,0000	1,0000	0,0000	0,2955	0,0191	0,0732	0,6912	0,5799
CC	0,0000	1,0000	0,0922	0,0000	0,3789	0,0000	0,0000	0,0000	0,3750	0,5888	0,1100	0,2204	0,0141	0,0000	0,4924	0,0236	0,3548	0,0002	0,0000
CT	0,0209	0,7501	0,0917	0,0000	0,4705	0,0000	0,0000	0,0000	0,0975	0,6538	0,0151	0,3838	0,0000	0,0000	0,4529	0,5259	0,1332	0,0000	0,0000
TT	0,0000	0,6363	0,0255	0,0000	0,6391	0,0000	0,0000	0,0000	0,0649	0,4645	0,0038	0,6750	0,0000	0,0000	0,3532	0,1080	0,0883	0,0000	0,0000

MUD - População mundial; **AFR** - Africanos; **AMR** - Americanos; **ASN** - Asiáticos; **ASW** - Americanos de ancestralidade Africana no sudoeste dos EUA; **CEU** - Residentes de Utah com ancestralidade do norte e oeste da Europa; **CHB** - Chineses Han em Bejing (Pequim), China; **CHS** - Chineses Han do Sul, China; **CLM** - Colombianos de Medellin, Colômbia; **EUR** - Europeus; **FIN** - Finlandeses na Finlândia; **GBR** - Britânicos na Inglaterra e Escócia; **IBS** - População Ibérica na Espanha; **JPT** - Japoneses em Tóquio, Japão; **LWK** - Luhya em Webuye, Kenya; **MXL** - Americanos de ancestralidade Mexicana proveniente de Los Angeles. EUA; **PUR** - Porto Riquenhos de Porto Rico; **TSI** - Toscanos na Itália; **YRI** - Iorubá em Ibadan, Nigéria.

MUD - População mundial; **AFR** - Africanos; **AMR** - Americanos; **ASN** - Asiáticos; **ASW** - Americanos de ancestralidade Africana no sudoeste dos EUA; **CEU** - Residentes de Utah com ancestralidade do norte e oeste da Europa; **CHB** - Chineses Han em Bejing (Pequim), China; **CHS** - Chineses Han do Sul, China; **CLM** - Colombianos de Medellin, Colômbia; **EUR** - Europeus; **FIN** - Finlandeses na Finlândia; **GBR** - Britânicos na Inglaterra e Escócia; **IBS** - População Ibérica na Espanha; **JPT** - Japoneses em Tóquio, Japão; **LWK** - Luhya em Webuye, Kenya; **MXL** - Americanos de ancestralidade Mexicana proveniente de Los Angeles, EUA; **PUR** - Porto Riquenhos de Porto Rico; **TSI** - Toscanos na Itália; **YRI** - Iorubá em Ibadan, Nigéria.

Figura 10: Gráfico mostrando a soma de *loci* (frequência alélica e genotípica) semelhantes e diferentes ao da amostra de população Controle segundo o teste exato de Fisher (setas indicando as superpopulações por continente).

VII.9 Equilíbrio de Hardy-Weinberg e Desequilíbrio de Ligação do Haplótipo Estendido

O equilíbrio de Hardy Weinberg (Tabela 14) e o desequilíbrio de ligação (Tabela 15) foram realizados usando dados de SNPs de outros genes do CPH, tipificados em trabalhos anteriores (CASTELLI, MENDES-JR *et al.*, 2009; CASTELLI, MENDES-JR *et al.*, 2011), que utilizaram as mesmas amostras controle do presente trabalho. Entre os dados, verificou-se desequilíbrio de Hardy-Weinberg no *locus HLA-DRB1* ($P=0,018$). A ausência de desequilíbrio entre o gene *HFE* e os *loci* do CPH foi um dado a ser destacado neste teste de desequilíbrio de ligação par a par. O desequilíbrio do *loci HLA-DRB1* foi determinado pelo excesso de heterozigotos detectado pelo teste global ($P=0,0001$).

Tabela 14. Desequilíbrio de Hardy-Weinberg (em negrito valores significativos).

LOCUS	VALOR P	DESVIO PADRÃO
<i>HLA-A</i>	0,40	0,03
<i>HLA-B</i>	0,32	0,03
<i>HLA-C</i>	0,55	0,02
<i>HLA-DRB1</i>	0,01	0,00
<i>HLA-DQB1</i>	0,12	0,00
<i>TNFA</i>	0,39	0,02
<i>TNFB</i>	0,41	0,01
<i>TNFC</i>	0,80	0,00
<i>TNFD</i>	0,52	0,01
<i>HLA-G14pb</i>	0,67	0,00
<i>HFE</i>	0,08	0,01

Tabela 15. Teste de desequilíbrio de ligação entre loci do CPH (em negrito, valores significativos; sombreado de cinza, comparações com o gene HFE).

LOCUS1	LOCUS 2	VALOR P	DESVIO PADRÃO
<i>HLA-A</i>	<i>HLA-B</i>	0,42	0,04
<i>HLA-A</i>	<i>HLA-C</i>	0,02	0,01
<i>HLA-A</i>	<i>HLA-DRB1</i>	0,00	0,00
<i>HLA-A</i>	<i>HLA-DQB1</i>	0,28	0,04
<i>HLA-A</i>	<i>TNFA</i>	0,37	0,04
<i>HLA-A</i>	<i>TNFB</i>	0,86	0,02
<i>HLA-A</i>	<i>TNFC</i>	0,68	0,02
<i>HLA-A</i>	<i>TNFD</i>	0,53	0,04
<i>HLA-B</i>	<i>HLA-C</i>	0,03	0,01
<i>HLA-B</i>	<i>HLA-DRB1</i>	0,20	0,04
<i>HLA-B</i>	<i>HLA-DQB1</i>	0,54	0,04
<i>HLA-B</i>	<i>TNFA</i>	0,02	0,01

<i>LOCUS 1</i>	<i>LOCUS 2</i>	<i>VALOR P</i>	<i>DESVIO PADRÃO</i>
<i>HLA-B</i>	<i>TNFB</i>	0,16	0,03
<i>HLA-B</i>	<i>TNFC</i>	0,26	0,02
<i>HLA-B</i>	<i>TNFD</i>	0,01	0,01
<i>HLA-C</i>	<i>HLA-DRB1</i>	0,25	0,04
<i>HLA-C</i>	<i>HLA-DQB1</i>	0,47	0,04
<i>HLA-C</i>	<i>TNFA</i>	0,25	0,04
<i>HLA-C</i>	<i>TNFB</i>	0,04	0,01
<i>HLA-C</i>	<i>TNFC</i>	0,01	0,00
<i>HLA-C</i>	<i>TNFD</i>	0,12	0,02
<i>HLA-DQB1</i>	<i>TNFA</i>	0,07	0,02
<i>HLA-DQB1</i>	<i>TNFB</i>	< 0,001	0,00
<i>HLA-DQB1</i>	<i>TNFC</i>	0,20	0,01
<i>HLA-DQB1</i>	<i>TNFD</i>	0,09	0,02
<i>HLA-DRB1</i>	<i>HLA-DQB1</i>	< 0,001	0,00
<i>HLA-DRB1</i>	<i>TNFA</i>	0,20	0,03
<i>HLA-DRB1</i>	<i>TNFB</i>	< 0,001	0,00
<i>HLA-DRB1</i>	<i>TNFC</i>	0,28	0,02
<i>HLA-DRB1</i>	<i>TNFD</i>	< 0,001	0,00
<i>TNFA</i>	<i>TNFB</i>	< 0,001	0,00
<i>TNFA</i>	<i>TNFC</i>	0,20	0,01
<i>TNFA</i>	<i>TNFD</i>	< 0,001	0,00
<i>TNFB</i>	<i>TNFC</i>	< 0,001	0,00
<i>TNFB</i>	<i>TNFD</i>	< 0,001	0,00
<i>TNFC</i>	<i>TNFD</i>	< 0,001	0,00
<i>HLA-A</i>	<i>HLA-G 14pb</i>	< 0,001	0,00
<i>HLA-B</i>	<i>HLA-G 14pb</i>	0,59	0,02
<i>HLA-C</i>	<i>HLA-G 14pb</i>	0,10	0,01
<i>HLA-DRB1</i>	<i>HLA-G 14pb</i>	0,86	0,01
<i>HLA-DQB1</i>	<i>HLA-G 14pb</i>	0,18	0,01
<i>TNFA</i>	<i>HLA-G 14pb</i>	0,71	0,01
<i>TNFB</i>	<i>HLA-G 14pb</i>	0,37	0,01
<i>TNFC</i>	<i>HLA-G 14pb</i>	0,70	0,00
<i>TNFD</i>	<i>HLA-G 14pb</i>	0,02	0,00
<i>HLA-A</i>	<i>HFE</i>	0,22	0,03
<i>HLA-B</i>	<i>HFE</i>	0,59	0,04
<i>HLA-C</i>	<i>HFE</i>	0,33	0,04
<i>HLA-DRB1</i>	<i>HFE</i>	0,11	0,02
<i>HLA-DQB1</i>	<i>HFE</i>	0,51	0,03
<i>TNFA</i>	<i>HFE</i>	0,22	0,03
<i>TNFB</i>	<i>HFE</i>	0,49	0,03
<i>TNFC</i>	<i>HFE</i>	0,65	0,01
<i>TNFD</i>	<i>HFE</i>	0,35	0,03
<i>HLA-G 14pb</i>	<i>HFE</i>	0,54	0,01

Partindo do resultado observado acima, um segundo teste, utilizando os SNPs clássicos e os demais dados do CPH, foi realizado para verificar se existia uma associação isolada entre eles (Tabela 16).

Tabela 16. Desequilíbrio de ligação SNPs clássicos do HFE (H63D e C282Y); com loci do CPH (em negrito P valor significativo).

<i>LOCUS</i>	<i>SNP</i>	<i>VALOR P</i>	<i>DESVIO PADRÃO</i>
<i>HLA-A</i>	H63D	0,09	0,01
<i>HLA-B</i>	H63D	0,03	0,00
<i>HLA-C</i>	H63D	0,15	0,01
<i>DRB1</i>	H63D	0,51	0,02
<i>DQB1</i>	H63D	0,31	0,01
<i>TNFA</i>	H63D	0,26	0,01
<i>TNFB</i>	H63D	0,90	0,00
<i>TNFC</i>	H63D	0,27	0,00
<i>TNFD</i>	H63D	0,67	0,01
<i>HLA-A</i>	C282Y	0,31	0,01
<i>HLA-B</i>	C282Y	0,57	0,01
<i>HLA-C</i>	C282Y	0,91	0,00
<i>DRB1</i>	C282Y	0,96	0,00
<i>DQB1</i>	C282Y	0,82	0,00
<i>TNFA</i>	C282Y	0,80	0,01
<i>TNFB</i>	C282Y	0,66	0,00
<i>TNFC</i>	C282Y	0,26	0,00
<i>TNFD</i>	C282Y	0,89	0,00
H63D	C282Y	0,79	0,00

O *locus HLA-B* e o SNP H63D apresentam-se em desequilíbrio de ligação ($P=0,0343$), quando calculada a hipótese nula, que é baseada na premissa de que a distribuição genotípica em um *locus* é independente da distribuição em outro *locus*. Mediante esse resultado foi realizado um segundo teste, que nos dá o desequilíbrio de ligação, usando as tabelas de contingência de desequilíbrio de ligação normalizado (D'), e o coeficiente de correlação (r^2). Isto para verificar quais os alelos estão em desequilíbrio e qual o comportamento e a intensidade desse desequilíbrio (Tabela 17). Esse teste indicou quatro conjuntos de alelos em desequilíbrio: *HLA-B*15* com H63D C, *HLA-B*37* com H63D G, *HLA-B*44* com H63D G e *HLA-B*56* com H63D G.

Tabela 17. Desequilíbrio de ligação entre os alelos do gene *HLA-B* e do SNP H63D do gene *HFE*.

<i>HLA-B</i>	FREQUÊNCIA OBSERVADA		FREQUÊNCIA ESPERADA		VALOR DE DESIQUILÍBRIO NORMALIZADO (D')		VALOR DE CORRELAÇÃO NORMALIZADO (r^2)	VALOR DE Qui^2	VALOR DE P de Qui^2
	<i>HFE</i> H63D								
	C	G	C	G	C	G			
07	16	2	14,83	3,17	0,3683	-0,3683	0,0029	0,5728	0,4492
08	12	1	10,71	2,29	0,5626	-0,5626	0,0047	0,9396	0,3324
13	3	1	3,30	0,70	-0,0899	0,0899	0,0008	0,1547	0,6941
14	10	0	8,24	1,76	1,0000	-1,0000	0,0113	2,2471	0,1339
15	17	0	14,01	2,99	1,0000	-1,0000	0,0199	3,9669	0,0464
18	10	0	8,24	1,76	1,0000	-1,0000	0,0113	2,2471	0,1339
27	2	2	3,30	0,70	-0,3933	0,3933	0,0149	2,9586	0,0854
35	17	4	17,31	3,69	-0,0177	0,0177	0,0002	0,0345	0,8526
37	1	2	2,47	0,53	-0,5955	0,5955	0,0254	5,0617	0,0245
38	3	0	2,47	0,53	1,0000	-1,0000	0,0033	0,6500	0,4201
39	4	2	4,94	1,06	-0,1911	0,1911	0,0053	1,0582	0,3036
40	3	0	2,47	0,53	1,0000	-1,0000	0,0033	0,6500	0,4201
41	3	0	2,47	0,53	1,0000	-1,0000	0,0033	0,6500	0,4201
42	1	0	0,82	0,18	1,0000	-1,0000	0,0011	0,2145	0,6433
44	8	11	15,66	3,34	-0,4891	0,4891	0,1183	23,5443	<0,00001
45	6	0	4,94	1,06	1,0000	-1,0000	0,0066	1,3203	0,2505
48	2	0	1,65	0,35	1,0000	-1,0000	0,0022	0,4312	0,5114
49	8	0	6,59	1,41	1,0000	-1,0000	0,0089	1,7788	0,1823
50	2	0	1,65	0,35	1,0000	-1,0000	0,0022	0,4312	0,5114
51	9	5	11,54	2,46	-0,2199	0,2199	0,0172	3,4137	0,0647
52	7	0	5,77	1,23	1,0000	-1,0000	0,0078	1,5484	0,2134
53	5	0	4,12	0,88	1,0000	-1,0000	0,0055	1,0946	0,2955
55	2	0	1,65	0,35	1,0000	-1,0000	0,0022	0,4312	0,5114
56	0	1	0,82	0,18	-1,0000	1,0000	0,0237	4,7094	0,0300
57	8	2	8,24	1,76	-0,0293	0,0293	0,0002	0,0423	0,8371
58	4	2	4,94	1,06	-0,1911	0,1911	0,0053	1,0582	0,3036
67	1	0	0,82	0,18	1,0000	-1,0000	0,0011	0,2145	0,6433

Identificação do SNP/Alelo mais frequente segundo NCBI (<http://www.ncbi.nlm.nih.gov>) e ENSEMBL (<http://www.ensembl.org>): rs1799945 (H63D) / C.

Devido ao fato do SNP rs2071303 estar a apenas 157bp do SNP H63D e de este SNP estar em desequilíbrio de ligação com o H63D em todas as nossas amostras, o definiu como um bom candidato para o teste de desequilíbrio entre os genes *HFE* e *HLA-B*. Os resultados desse novo teste também detectou o desequilíbrio com os alelos 37 e 44 do gene *HLA-B*, entre outros,. O desequilíbrio com o alelo 37, apesar de presente, apresenta valores com baixo poder estatístico ($D' = 0,5955$ e $r^2 = 0,02$). No alelo 44, o valor de D' (0,4823) ficou muito próximo ao do teste anterior, porém como o valor r^2 ficou muito mais baixo (0,0382) (Figura 11). Nesse caso também se verificou diferença pouco pronunciada entre a frequência observada e esperada (6,013,0; 11,59-7,41). Os outros desequilíbrios detectados (35-T, 49-T e 58-C) não coincidiram com os do teste anterior e não mostraram grandes valores de correlação, r^2 (0,0196 a 0,0266) (Tabela 18).

Tabela 18. Desequilíbrio de ligação entre os alelos do gene HLA-B e do SNP (IVS2(+4)T>C) do gene *HFE*.

	FREQUÊNCIA OBSERVADA	FREQUÊNCIA ESPERADA	VALOR DE DESIQUILÍBRIO NORMALIZADO (<i>D'</i>)		VALOR DE CORRELAÇÃO NORMALIZADO (<i>r</i> ²)		VALOR DE Qui ²	VALOR DE <i>P</i> de Qui ²	
<i>HLA-B</i>	<i>HFE</i> (IVS2(+4)T>C)								
	T	C	T	C	T	C			
07	13	5	10,98	7,02	0,2877	-0,2877	0,0052	1,0471	0,3062
08	8	5	7,93	5,07	0,0138	-0,0138	0,0000	0,0017	0,9672
13	4	0	2,44	1,56	1,0000	-1,0000	0,0130	2,6096	0,1062
14	4	6	6,10	3,90	-0,3443	0,3443	0,0098	1,9513	0,1624
15	11	6	10,37	6,63	0,0950	-0,0950	0,0005	0,1073	0,7433
18	8	2	6,10	3,90	0,4872	-0,4872	0,0080	1,5973	0,2063
27	1	3	2,44	1,56	-0,5902	0,5902	0,0111	2,2235	0,1359
35	17	4	12,81	8,19	0,5116	-0,5116	0,0196	3,9264	0,0475
37	0	3	1,83	1,17	-1,0000	1,0000	0,0238	4,7638	0,0291
38	2	1	1,83	1,17	0,1453	-0,1453	0,0002	0,0411	0,8393
39	1	5	3,66	2,34	-0,7268	0,7268	0,0256	5,1103	0,0238
40	3	0	1,83	1,17	1,0000	-1,0000	0,0097	1,9472	0,1629
41	2	1	1,83	1,17	0,1453	-0,1453	0,0002	0,0411	0,8393
42	1	0	0,61	0,39	1,0000	-1,0000	0,0032	0,6426	0,4228
44	6	13	11,59	7,41	-0,4823	0,4823	0,0382	7,6388	0,0057
45	3	3	3,66	2,34	-0,1803	0,1803	0,0016	0,3146	0,5749
48	2	0	1,22	0,78	1,0000	-1,0000	0,0065	1,2916	0,2558
49	8	0	4,88	3,12	1,0000	-1,0000	0,0266	5,3279	0,0210
50	2	0	1,22	0,78	1,0000	-1,0000	0,0065	1,2916	0,2558
51	10	4	8,54	5,46	0,2674	-0,2674	0,0034	0,6882	0,4068
52	3	4	4,27	2,73	-0,2974	0,2974	0,0050	1,0037	0,3164
53	5	0	3,05	1,95	1,0000	-1,0000	0,0164	3,2787	0,0702
55	0	2	1,22	0,78	-1,0000	1,0000	0,0158	3,1598	0,0755
56	1	0	0,61	0,39	1,0000	-1,0000	0,0032	0,6426	0,4228
57	4	6	6,10	3,90	-0,3443	0,3443	0,0098	1,9513	0,1624
58	1	5	3,66	2,34	-0,7268	0,7268	0,0256	5,1103	0,0238
67	1	0	0,61	0,39	1,0000	-1,0000	0,0032	0,6426	0,4228

Identificação do SNP/Alelo mais frequente segundo NCBI (<http://www.ncbi.nlm.nih.gov>) e ENSEMBL (<http://www.ensembl.org>): rs2071303 (IVS2(+4)T>C) / T.

Figura 11: Esquema do desequilíbrio de ligação entre o gene *HFE* e o *HLA-B*.

VIII. DISCUSSÃO

VIII.1 Frequências alélicas, genotípicas, equilíbrio de Hardy-Weinberg e heterozigose

No presente trabalho, foram usados dois tipos de denominações para os SNPs encontrados no gene *HFE*. O primeiro é o comumente utilizado nos trabalhos, onde o nome em si do SNP é usado para denominar indivíduos portadores do alelo mutado e/ou menos frequente na população, por exemplo: indivíduo portador do SNP C282Y, significa dizer que o indivíduo apresenta o alelo G (mutado e/ou menos frequente) para esse *locus*. A segunda denominação utilizada, para facilitar o entendimento dos resultados, principalmente nas análises de haplótipos e de desequilíbrio de ligação, foi a denominação do alelo em si. Utilizando o nome do SNP acompanhado da base nitrogenada característica do alelo, por exemplo: indivíduos portadores de alelo C282Y G. Em suma, dizer que um indivíduo é portador do SNP mutado C282Y é o mesmo que dizer que este indivíduo tem o alelo C282Y G, ao passo que o indivíduo não tem o SNP C282Y é o mesmo que dizer que ele tem o alelo C282Y C.

No presente trabalho foram encontrados seis SNPs e uma nova mutação no gene *HFE* no sentido 5' para 3': H63D no éxon 2 (rs1799945); IVS2(+4)T>C no íntron 2 (rs2071303); SNP no íntron 3 (rs807209); C282Y no éxon 4 (rs1800562); SNPs no íntron 4 (rs2794717) e IVS4(-44)T>C (rs1800708); no íntron 5 a deleção ainda não descrita. Apenas os SNPs clássicos H63D e C282Y apresentam literatura abundante para comparações.

A frequência alélica do H63D G nas amostras populacionais (pacientes e controles) deste estudo variou de 7,9% a 17,5% (Tabela 5), o que está de acordo com a literatura. Quando consideradas as populações disponíveis para comparação, as frequências mais altas foram encontradas em indivíduos saudáveis da Holanda e da península Ibérica (acima de 20%) (CARDOSO; DE SOUSA, 2003; PORTO; DE SOUSA, 2000). O fato de o alelo H63D G ocorrer em alta frequência na península Ibérica é muito relevante para estudos populacionais no Brasil, já que o Brasil foi colonizado por Portugal. Porém, apesar de a frequência do alelo H63D G ser concordante com a da literatura, não foram detectadas diferenças significativas entre as amostras de população controle e Hemocromatose Hereditária, tanto em relação às frequências alélicas, genotípicas e desequilíbrio de Hardy-Weinberg, como também no

teste exato de Fisher, contrariando autores como (BACON *et al.*, 1999; BARTON *et al.*, 1999; BARTON *et al.*, 1997; FEDER *et al.*, 1996), que apontam relação entre pacientes heterozigotos para C282Y A/ H63D G e o desenvolvimento da doença Hemocromatose Hereditária. Por exemplo; um estudo australiano com indivíduos gêmeos mostrou uma concordância principalmente em homens entre a sobrecarga de ferro e a presença do alelo H63D G (WHITFIELD *et al.*, 2000). Entretanto, em contraponto, (PHATAK *et al.*, 2002) comenta-se que frequência do homozigoto H63D/H63D e do heterozigoto composto C282Y/H63D é baixa e de baixa penetrância.

A homogeneidade da frequência do SNP H63D, inclusive na amostra Hemocromatose Hereditária, torna relevante a alteração da frequência do SNP adjacente (IVS2(+4)T>C), uma vez que o alelo IVS2(+4) C se mantém entre 31,6% a 42,5% em todas as nossas amostras de pacientes e controles e diminui para 14,3%, apenas na população com Hemocromatose Hereditária. Logo, ao se analisar o alelo T (selvagem) do mesmo SNP IVS2(+4) verifica-se aumento da sua frequência (85%) na amostra de Hemocromatose Hereditária se comparado com as outras amostras (58,5% a 68,4%). O teste exato de Fisher comparando a amostra população Controle e a amostra população Hemocromatose Hereditária para o IVS2(+4)T>C detectou diferenças significantes ($P=0,04$) entre essas amostras com relação ao genótipo TT, sendo indicado como genótipo de risco apenas para a hemocromatose (OR=3,91). No entanto, o intervalo de confiança é relativamente grande (IC= 1,14 a 13,34), devido às diferenças de tamanho das amostras.

O alelo C282Y A apresentou frequência alélica de 2,1% na amostra de população Controle, variando de 1,7% a 9,8% nas amostras de populações de pacientes (CHC com sobrecarga de ferro, CHC diverso e HCV com ou sem sobrecarga de ferro) e apresentou alta frequência na amostra de população Hemocromatose Hereditária (71,4%), sendo que, somente um indivíduo dessa amostra não apresentou mutação para esse SNP (Tabela 5). A alta frequência alélica do SNP C282Y A na amostra de população Hemocromatose Hereditária também é extensamente documentada na literatura, sendo este SNP apontado como o principal sítio polimórfico associado com hemocromatose hereditária, especialmente em populações europeias (PIETRANGELO, 2004).

No teste de Hardy-Weinberg, a amostra Câncer Hepatocelular com Sobrecarga de Ferro, se apresentou em desequilíbrio no SNP C282Y. O teste global feito tanto por amostra de população, quanto por *loci*, mostrou que a amostra Câncer Hepatocelular

com Sobrecarga de Ferro ($P=0,043$) e o SNP rs807209 ($P=0,014$) possuem déficit de heterozigotos. Isto nos leva a deduzir que o déficit de heterozigotos encontrado nessa amostra de população Carcinoma com Sobrecarga seja resultado do déficit encontrado no SNP rs807209, e que o desequilíbrio com o C282Y não é determinado por déficit ou excesso de heterozigotos.

As frequências dos alelos H63D G e C282Y A na nossa amostra de população Controle (Tabela 2) foram similares às aquelas relatadas em outras populações do Sudeste do Brasil (especificamente da cidade de São Paulo), isto é a frequência do o alelo C282Y A variou de 1,4 a 2,1% e a o alelo H63D G, de 10 a 13,6% (SANTOS *et al.*, 2010).

O perfil de frequência do SNP C282Y na região Sudeste do Brasil é próximo ao de populações europeias contrastando com o de populações de outras regiões do Brasil, nas quais o alelo C282Y A é raro ou ausente. Por exemplo, na região Norte do Brasil ocorre ausência do alelo C282Y A em três estados (Amapá, Acre, Rondônia) (TORRES *et al.*, 2008). Em nosso estudo, o SNP C282Y foi associado com susceptibilidade apenas à hemocromatose hereditária. O alelo selvagem G e o genótipo GG foram identificados como protetivos ($OR=0,01$ e $0,007$, respectivamente), enquanto que o alelo mutado A e o genótipo AA foram indicados como fatores de risco ($OR= 53,06$ e $201,00$, respectivamente) (Tabela 5). É notório o intervalo de confiança muito dilatado, principalmente para o genótipo AA ($IC=10,44$ a $3.871,00$). Essa dilatação se deve a grande diferença de tamanhos amostrais (Controle, $n = 100$ e Hemocromatose Hereditária, $n = 14$) e, principalmente, pela não ocorrência do genótipo AA na população Controle. Apesar de esses dois fatores tornarem o valor de *Odds Ratio* impreciso, o mesmo não deixa de ser relevante, já que a amostra de população com o n maior (população Controle) não apresenta o homozigoto mutado (AA), enquanto o mesmo genótipo se apresenta abundantemente na amostra de população com o n menor (Hemocromatose Hereditária). No intuito de superar o problema dos intervalos de confiança dilatados, foi feita uma busca para obtenção de outras amostras de indivíduos com hemocromatose hereditária, porém só obtivemos mais oito indivíduos, o que não expandiu expressivamente o n dessa população. Deve-se ressaltar que a dificuldade no aumento do n na amostra Hemocromatose Hereditária advém do fato da raridade dessa doença.

VIII.2 Desequilíbrio de Ligação

Depois de observada diferença significativa no SNP IVS2(+4)T>C entre as amostras de população Controle e Hemocromatose Hereditária, foram realizados testes de desequilíbrio de ligação, que demonstraram desequilíbrio entre os SNPs H63D / IVS2(+4)T>C, com exceção da amostra de população Câncer Hepatocelular Diverso e entre os SNPs IVS2(+4)T>C e IVS4(-44)T>C, com exceção da amostra Hemocromatose Hereditária. A amostra de população Hemocromatose Hereditária estava em desequilíbrio para os SNPs clássicos H63D e C282Y.

O diagrama gerado pelo *software* HAPLOVIEW (Figura 8) detectou forte desequilíbrio entre os SNPs H63D e IVS2(+4)T>C (LOD=95) e entre os SNPs IVS2(+4)T>C e IVS4(-44)T>C (LOD=90), desequilíbrio moderado entre os SNPs IVS2(+4)T>C e C282Y (LOD=77) e ausência de desequilíbrio entre os SNPs clássicos H63D e C282Y (LOD=53). Dados semelhantes foram obtidos por Dhillon *et al.* (2012) em estudo realizado no Norte da Índia, sendo encontrado desequilíbrio completo entre H63D, IVS2(+4)T>C e IVS4(-44)T>C. No trabalho de Dhillon *et al.*, (2007), não foram realizados testes de desequilíbrio entre o SNP C282Y e os outros SNPs (H63D, IVS2(+4)T>C e IVS4(-44)T>C) devido a ausência do SNP C282Y, ausência é extensamente documentada na população indiana. (KAUR *et al.*, 2003; THAKUR *et al.*, 2004; GAREWAL *et al.*, 2005; DHILLON *et al.*, 2007).

Esses dados motivaram a realização de estudo mais aprofundado do desequilíbrio para compreender quais alelos estariam ligados e qual seria o padrão dessas ligações. Para essa finalidade, foi gerada uma tabela de contingência D' e r^2 (Tabela 7), utilizando o *software* ARLEQUIN. A tabela mostrou que o alelo mutado H63D G está ligado completamente ao alelo selvagem C282Y G, tanto na amostra de população hemocromatose hereditária como na amostra de população Controle, e que não há ocorrência do alelo G (mutado) do SNP H63D com o alelo A (mutado) do SNP C282Y. Esses resultados estão de acordo com os haplótipos encontrados no Canadá e USA, onde os alelos C282Y A e H63G também não foram encontrados nos mesmos cromossomos (FEDER *et al.*, 1996; BARTON *et al.*, 1997; BARTON *et al.*, 1999). Por outro lado, contrastam com os resultados de BEST; HARRIS; SPRIGGS (2001) que apontaram que o desequilíbrio de ligação entre esses dois SNPs é observado em *cis*, ao passo que nosso trabalho foi observado em *trans*. A ausência do haplótipo mutado-

mutado (G-A) também ocorreu na amostra de população Controle, porém o desequilíbrio de ligação não foi significativo.

Na amostra de população Hemocromatose Hereditária, 40% dos pacientes eram heterozigotos para o SNP C282Y, e desses, apenas 1/3 era portador do alelo H63D G em heterozigose com o alelo C282Y A. Trabalho desenvolvido no Brasil por Bittencourt *et al.* (2002) encontrou resultados semelhantes, tendo que, por volta da metade dos 15 pacientes com hemocromatose hereditária eram homozigotos para o alelo C282Y A e nenhum paciente apresentava a heterozigose C282Y A/H63D G.

Esses resultados contrastam com os encontrados nas populações do Norte da Europa, onde 74% a 100% dos pacientes com hemocromatose hereditária, que carregam apenas uma cópia do SNP C282Y, têm ambos os alelos, C282Y A e H63D G, em heterozigose (AGUILAR MARTINEZ *et al.*, 1997; BEUTLER *et al.*, 1996; CARDOSO; DE SOUSA, 2003; CARDOSO *et al.*, 1998; FEDER *et al.*, 1996; MERRYWEATHER-CLARKE *et al.*, 2000; MIEDZYBRODZKA *et al.*, 1999; MILMAN *et al.*, 1988; PORTO; DE SOUSA, 2000; SÁNCHEZ *et al.*, 1998; WORWOOD *et al.*, 1997).

Um comportamento notório foi o alelo C do SNP IVS2(+4), que se apresenta em desequilíbrio de ligação com o alelo H63D G em ambas as amostras de população (Controle e Hemocromatose Hereditária). No entanto, na amostra de população Hemocromatose Hereditária o mesmo alelo C do IVS2(+4) também se apresentou em desequilíbrio com o alelo C282Y G. Além disso, o alelo G (mutado) do SNP H63D, nunca se apresentou ligado ao alelo T (selvagem) do SNP IVS2(+4) e com o alelo A (mutado) do SNP C282Y, mostrando a inexistência do haplótipo (G-T-A) (Figura 9). Esse resultado foi semelhante ao encontrado por De Lucas *et al.* (2005), que ao genotipar 23 indivíduos para o SNP IVS2(+4)T>C, não encontraram a variação intrônica IVS2(+4) C. Uma vez que esses indivíduos eram sabidamente homozigotos para o alelo C282Y A, a presença do SNP C282Y A poderia excluir a presença do SNP IVS2(+4) C no mesmo cromossomo. Em conjunto, esses achados reforçam a inexistência ou baixa frequência do haplótipo recombinante G-T-A.

Um estudo do gene *HFE* em populações controles advindas da região que abrange a Eurásia Central até o Leste Europeu genotipou três SNPs (IVS2(+4)T>C, IVS4(-44)T>C e IVS5(-47)A>G) em indivíduos previamente tipificados como homozigotos para o alelo H63D G, relatando desequilíbrio de ligação entre o alelo H63D G com os outros dois SNPs, definindo três haplótipos (CTA, TTG e TTA)

(KHUSAINOVA *et al.*, 2013). Estudo feito na Califórnia (EUA) encontrou o haplótipo “CTA” ligado ao alelo H63D G em todos os pacientes com hemocromatose hereditária (BEUTLER, ERNEST e WEST, 1997). Esse haplótipo foi fortemente associado com cromossomos europeus carregando o alelo H63D G (DHILLON *et al.*, 2012). Indivíduos do Norte da Índia que portavam o alelo H63D G eram todos portadores desse haplótipo “europeu”. Dhillon *et al.* (2012) sugerem que esse haplótipo tenha sido introduzido na população indiana pela miscigenação com europeus. A ligação entre os alelos H63D G e IVS2(+4) C encontrada no presente trabalho, provavelmente, reforça a origem europeia desse haplótipo.

De Lucas *et al.* (2005) sugerem que o SNP IVS2(+4)T>C não influencia o desenvolvimento da hemocromatose hereditária, mesmo tendo encontrado diferenças estatísticas significantes ($P=0,009$) na frequência do alelo IVS2(+4) C entre grupo controle (34,8%) e em pacientes heterozigos para o SNP C282Y (16,2%). O autor aponta que a presença do alelo C282Y A exclui a presença do alelo IVS2(+4) C no mesmo haplótipo, logo, análises baseadas na frequência alélica simples do SNP IVS2(+4)T>C não são adequadas, uma vez que ocorre alta frequência do alelo C282Y A em amostras de pacientes com hemocromatose hereditária e, consequentemente, ocorre baixa frequência de IVS2(+4) C nessa amostra. Apesar de não estar claro no trabalho, o autor sugere recalcular a frequência do alelo IVS2(+4) C, provavelmente utilizando uma razão entre a contagem do alelo IVS2(+4) C e um tamanho amostral, que foi calculado pela subtração do tamanho total da amostra (“*n*”) da quantidade de vezes em que o alelo C282Y A aparece na referida amostra. No trabalho de De Lucas *et al.* (2005), isso resultou em uma nova frequência (32,4%), ou seja, o dobro da frequência observada anteriormente. Isto ocorre porque a amostra de pacientes com hemocromatose hereditária do autor, foi inteiramente heterozigota para o alelo C282Y A. O autor comenta, que essa nova frequência não apresenta diferenças estatísticas entre os grupos controle e pacientes com hemocromatose hereditária.

Apesar do presente trabalho não ter encontrado diferenças significativas entre as amostras Controle e Hemocromatose Hereditária, tanto para o alelo IVS2(+4) C quanto para o genótipo IVS2(+4) CC, foram encontradas diferenças significativas ($P=0,04$) para o genótipo IVS2(+4) TT, provavelmente também efeito do desequilíbrio entre os alelos IVS2(+4) C e C282Y A.

Sendo assim, mesmo não tendo sido observadas diferenças significativas, foi testado o conceito de De Lucas *et al.* (2005). Conceito, no qual se subentende-se que a

presença de um dado alelo (C282Y A) de um SNP que está em desequilíbrio com um segundo alelo (IVS2(+4) C) de outro SNP, exclui a possibilidade da presença desse segundo alelo (IVS2(+4) C). Sendo assim, a presença do primeiro alelo (C282Y A) deve ser descontado nos cálculos de frequência e teste exato de Fisher para o segundo alelo (IVS2(+4) C). Seguindo esses preceitos, recalculamos a frequência do alelo IVS2(+4) C e os genótipos IVS2(+4) CC e IVS2(+4) CT. Porém, a amostra Hemocromatose Hereditária não tem a característica homogênea como a do autor, que é de todos os pacientes serem heterozigotos para o alelo C282Y A. Sendo assim, para uma adaptação viável do conceito, foram criadas quatro expressões matemáticas para o recálculo das frequências, ponderando pelo desequilíbrio entre os dois alelos i de dois SNPs (SNP₁ e SNP₂) e a porcentagem de recombinação dos haplótipos desses dois alelos, conforme mostrado abaixo:

Para frequência alélica:

$$x_{iSNP_1} = \frac{2 n_{ii\ SNP_1} + n_{ij\ SNP_1}}{2n - (\%_{recb} \cdot (2 n_{ii\ SNP_2} + n_{ij\ SNP_2}))}$$

Para frequência genotípica das combinações homozigotas:

$$x_{iiSNP_1} = \frac{n_{ii\ SNP_1}}{n - (\%_{recb} \cdot (n_{ii\ SNP_2} + n_{ij\ SNP_2}))}$$

Para frequência genotípica das combinações heterozigotas:

$$x_{ij} = \frac{n_{ij\ SNP_1}}{n - (\%_{recb} \cdot (n_{ii\ SNP_2}))}$$

Onde:

$$\%_{recb} = \frac{p_{i\ desq} - p_{i\ recb}}{p_{i\ desq} + p_{i\ recb}}$$

x_{iSNP_1} é a frequência do alelo i em desequilíbrio do SNP₁

x_{iiSNP_1} é a frequência do genótipo **ii** homozigoto em desequilíbrio do SNP₁

x_{ijSNP_1} é a frequência do genótipo **ij** heterozigotos em desequilíbrio do SNP₁

n_{iiSNP_1} e n_{ijSNP_1} correspondem ao número de homozigotos e heterozigotos no SNP₁ observados para o alelo **i** em desequilíbrio, respectivamente;

n_{iiSNP_2} e n_{ijSNP_2} correspondem ao número de homozigotos e heterozigotos no SNP₂ observados para o alelo **i** em desequilíbrio, respectivamente;

n corresponde ao número de indivíduos analisados.

$\%_{recb}$ é uma porcentagem entre os **p_{i desq}** (haplótipos que apresentam a combinação do desequilíbrio entre os alelos “**i**” dos SNP₁ e SNP₂) subtraído do haplótipos que se apresentem recombinados entre os dois alelos “**i**”.

Dado, o IVS2(+4) C como n_{iSNP_1} e C282Y A como n_{iSNP_2} e, considerando a $\%_{recb}$ no presente trabalho como 1 (100%), já que não foram observados haplótipos recombinantes (IVS2(+4) C/C282Y A) nas presentes amostras, obtivemos as seguintes frequências: na população Controle o alelo IVS2(+4) C = 17,8%; o genótipo IVS2(+4) CC = 42,8% e o genótipo IVS2(+4) CT = 9%; na população Hemocromatose Hereditária o genótipo IVS2(+4) CC = 0%; o genótipo CT e alelo C um aumento expressivo: IVS2(+4) CT = 50% e IVS2(+4) C = 50%.

Utilizando o mesmo conceito, foram redefinidos dois novos tamanhos amostrais “**n**” para as amostras de Controle e Hemocromatose Hereditária, calculado pelo denominador da fração das expressões matemáticas $n_{SNP_1} = 2n - \left(\%_{recb} \cdot (2 n_{iiSNP_2} + n_{ijSNP_2}) \right)$. Este novo tamanho amostral foi utilizado para realizar o cálculo do teste exato de Fisher, considerando o alelo IVS2(+4) C, $n_{SNP_1} = n - \left(\%_{recb} \cdot (n_{iiSNP_2} + n_{ijSNP_2}) \right)$, para realizar o mesmo teste para os genótipos IVS2(+4) CC e, $n_{SNP_1} = n - \left(\%_{recb} \cdot (n_{iiSNP_2}) \right)$, para testar o genótipo IVS2(+4) CT. Os resultados dos testes para o alelo C e para o genótipo CC e CT não foram significativos ($P = 0,46$ e $1,00$ e $0,09$, consecutivamente), corroborando a provável neutralidade do SNP IVS2(+4)T>C.

Dhillon *et al.* (2012) também relatam que o SNP IVS2(+4)T>C é um polimorfismo neutro com relação à sobrecarga de ferro. Estudo recente demonstrou que o surgimento e o tempo de ocorrência evolutiva concomitante dos SNPs clássicos H63D

e C282Y é considerável, devido à presença dos haplótipos recombinantes, H63D C / IVS2(+4) C / C282Y G e H63D G / IVS2(+4) T / C282Y A (CURCIO *et al.*, 2008), sendo que esse último haplótipo não foi encontrado nas amostras do presente trabalho. Porém, análises posteriores de reconstrução de haplótipos detectaram dois indivíduos recombinantes (H63D G/IVS2(+4) T/C282Y G), correspondente ao haplótipo *HFE**01:02, um indivíduo na amostra de população Hepatite C sem sobrecarga e outro na amostra de população Cancer Hepatocelular Diverso.

Apesar na literatura relatar a neutralidade do SNP IVS2(+4)T>C para o fenótipo hemocromatose herediária, também o descreve como responsável pela mudança no padrão de edição alternativa da proteína HFE (MIKHAILOVA *et al.*, 2010), já que a variação intrônica IVS2(+4) C está localizada perto do sítio de junção, resultando em proteína com o exon 2 ausente. Além disso, estudo genético avaliando indivíduo do sexo feminino e seu pai, que apresentava uma sobrecarga de ferro moderada, encontrou que ambos portavam os SNPs (S65C e IVS2(+4)T>C), provavelmente localizados no mesmo cromossomo e ambos não portavam os SNPs clássicos (C282Y e H63D). Ainda, esse estudo identificou em ambos indivíduos duas moléculas da proteína HFE, uma completa e outra com ausência dos aminoácidos transcritos do éxon 2. O autor relata que provavelmente o alelo SNP S65C T (mutado) combinado com o alelo IVS2(+4) C possa ser responsável por isso. A função da proteína HFE sem o éxon 2, ainda não é clara, porém o autor sugere que essa proteína solúvel possa contribuir para o desenvolvimento da sobrecarga de ferro (FLOREANI *et al.*, 2005).

No presente trabalho, não foi encontrada a variação S65C (também localizada no éxon 2), o que está de acordo com dois outros trabalhos conduzidos na população brasileira, avaliando pacientes com sobrecarga de ferro coletados na cidade de São Paulo (CANÇADO *et al.*, 2007; CANÇADO *et al.*, 2006). Outro estudo conduzido também na cidade de São Paulo, com 148 indivíduos saudáveis, encontrou o SNP S65C em uma frequência de 1% (BUENO *et al.*, 2006). Um trabalho com maior *n* (542 indivíduos saudáveis) encontrou o SNP S65C com frequência de 0,6%, confirmando a raridade do SNP S65C na população brasileira (SANTOS *et al.*, 2010). Trabalhos têm apontado o SNP S65C como raro em diversas populações saudáveis ao redor do mundo, sendo que em populações no Norte da Europa sua frequência varia de 1,5% a 1,6% (HANSON *et al.*, 2001; HOLMSTRÖM *et al.*, 2002; MERRYWEATHER-CLARKE *et al.*, 2000; MILMAN *et al.*, 2005; MILMAN *et al.*, 2004), na Itália 1,3% (PIETRAPEROSA *et al.*, 2003), na França em 2,5% (HANSON *et al.*, 2001; MURA

et al., 1999) e no Leste Europeu variando entre 0,5 a 1,8% (ČIMBUROVÁ *et al.*, 2005; RISTIĆ *et al.*, 2003; ŠARIĆ *et al.*, 2006). Na Espanha, esse SNP é encontrado apenas nos pacientes com sobrecarga de ferro, apresentando baixa frequência (3%) (DE DIEGO; JOSÉ MURGA; MARTÍNEZ-CASTRO, 2004). Um achado notório é o relativo aumento na frequência (4%) desse SNP, tanto na população saudável, como em pacientes com hemocromatose hereditária do Equador e, concomitantemente, a ausência do SNP C282Y nas nessas mesmas populações (LEONE *et al.*, 2005) (Tabela 2).

Resumindo, apesar de que o genótipo TT do IVS2(+4)T>C possa indicar susceptibilidade ao desenvolvimento da hemocromatose hereditária, uma vez que apresenta forte desequilíbrio entre os alelos H63D G e IVS2(+4) C, esses achados podem gerar fator de confusão sobre a participação desse SNP na susceptibilidade à doença, pois: i) a nossa amostra de população Hemocromatose Hereditária é reduzida, com apenas 14 indivíduos; ii) o alelo apontado como gerador da proteína HFE solúvel é o IVS2(+4) C e não o IVS2(+4) T que compõe o genótipo de risco na comparação entre população Controle e com Hemocromatose Hereditária; iii) o forte desequilíbrio de ligação entre os alelos IVS2(+4) T e C282Y G pode gerar diferenças significativas entre as amostras, já que a presença do alelo C282Y G exclui a presença do alelo IVS2(+4) C, sem que isso tenha uma correlação direta com o fenótipo hemocromatose hereditária, conforme argumenta De Lucas *et al.* (2005), mesmo recalculando as frequências e o teste exato de Fisher baseado nos conceitos desse trabalho.

O SNP IVS4(-44)T>C, encontrado no presente trabalho, também é apontado por, possivelmente, alterar a edição alternativa da proteína HFE. Esse SNP retém o íntron 4, que contém um código de término de transcrição e, por consequência, produz uma proteína curta e solúvel que pode ter função na regulação do ferro. A expressão dessa proteína já foi observada no fígado, porém sua função ainda não está bem determinada (MIKHAILOVA *et al.*, 2010).

Um ponto importante a ser destacado é a relevância da estratificação das amostras de pacientes com comorbidades por etnia e sexo, já que a literatura demonstra consideráveis diferenças nas frequências dos SNPs clássicos em diferentes grupos étnicos. BUENO *et al.* (2006) apontam que, quando estes são estratificados por grupos étnicos, o SNP C282Y, por exemplo, foi encontrado apenas na população Caucasiana, e que H63D tem uma diferença significativa quando comparado entre a população Caucasiana (8,6%) e Africana (2,4%). Já o SNP S65C foi encontrado em baixa frequência em ambas as populações de origem Caucasiana (0,6%) e Africana (0,3%)

(BUENO *et al.*, 2006). Estudo de Santos *et al.*, (2010) não encontrou diferenças significativas entre as etnias para o SNP C282Y, o que o autor justifica pelo fato de ter sido utilizado a autodeclaração para classificação étnica. Isto pode ter limitado o trabalho, dada a alta taxa de miscigenação e a difícil classificação étnica. Entretanto, diferenças étnicas significativas foram encontradas por Agostinho *et al.*, (1999), trabalhando com uma população brasileira de Caucasianos e Miscigenados (origem Afro-descendente e Caucasiana) oriundos da cidade de Campinas (região Sudeste), Afro-descendentes, oriundos do estado da Bahia (região Nordeste) e Ameríndios (Índios Parakanã pertencentes a tribo Tupi). A frequência alélica para C282Y foi 1,4% para Caucasianos, 1,1% para Afro-descendentes, 1,1% para mestiços e notórios 0% para índios Parakanã. O SNP H63D apresentou uma frequência de 16,3% para Caucasianos, 7,5% para Afro-descendentes e 9,8 para mestiços e novamente 0% para Ameríndios.

Em relação ao gênero, Santos *et al.* (2010) não encontraram diferença de frequência para o SNP C282Y, porém, encontraram diferenças de fenótipo mórbido relacionado ao gênero, isto é, mulheres parecem ser menos susceptíveis a alterações relacionadas com a presença do SNP C282Y, o que pode estar correlacionado com a perda de sangue no período menstrual como um fator protetivo.

Em nosso trabalho, não foi possível realizar estratificações, devido à inerente dificuldade de identificar a etnia dos indivíduos, além da constituição majoritariamente masculina (74%) e de brancos (87%) das amostras, sendo que homens brancos constituem 63% dos indivíduos da população, o que inviabilizaria qualquer análise estatística consistente. Em adição, a diversidade média entre as amostras de populações se mostraram muito próximas, variando de 0,13 a 0,18. Essa homogeneidade nas diversidades médias demonstra que não existe estruturação interpopulacional, o que se justifica pela busca aleatória de indivíduos para formação das amostras estudadas, ou seja, não houve qualquer parâmetro de etnia ou localização geográfica interferindo nessa escolha. Entretanto, por motivos óbvios de acesso, as amostras em sua maioria são provenientes de indivíduos advindos da região de Ribeirão Preto e que pela sua história de colonização são predominantemente brancos. Em última análise, pode se supor que as morbidades em si não estejam direcionando um fluxo de alelos entre as populações e alterando a composição e consequentemente a estrutura genética.

Finalmente, cabe salientar que o presente trabalho faz a primeira descrição da deleção de uma base G, encontrada no íntron 5, encontrada em apenas um indivíduo saudável. Mutações ocasionais em um indivíduo ou em seus ascendentes ou

descendentes podem ocorrer, sem que isso represente frequência polimórfica. No entanto, deve ser enfatizada e descrita, pois pode ser encontrada em outras populações com diferentes frequências.

VIII.3 Diversidade e Diferenciação Genética da Região Codificadora do gene *HFE*

Já que foi observada homogeneidade da diversidade média nas amostras de populações do presente trabalho, foi realizada análise de diversidade e diferenciação da região codificadora. Essa análise não encontrou grandes diferenças entre a heterozigose esperada e observada. Porém, o teste global de diferenciação conseguiu diferenciar as amostras de população ($P < 0,001$) e a matriz de distância genética baseada no F_{ST} demonstrou que a amostra da população Hemocromatose Hereditária se diferencia das populações nos sete *loci* (Tabela 9). Esses resultados são concordantes com as diferenças significativas encontradas, em relação às frequências alélicas e genótípicas, tanto como no teste exato de Fisher. Esse resultado nos permite supor que os SNPs clássicos C282Y e o SNP IVS2(+4)T>C, pelo intenso desequilíbrio de ligação, diferenciam a amostra Hemocromatose Hereditária e podem alterar a composição dessa população como um todo, arrastando os outros alelos. Outra hipótese seria o fato do pequeno n da amostra Hemocromatose Hereditária estar potencializando a deriva amostral.

VIII.4 Haplótipo de Fase Meiótica e Teste de Neutralidade

A partir da dedução das fases meióticas das amostras desse trabalho foi aplicado na amostra população Controle os testes de neutralidade de Tajima e Ewens-Watterson. Os resultados não demonstraram qualquer tipo de pressão seletiva nas amostras, indicando que nenhum alelo ou genótipo está sendo preferencialmente selecionado na população saudável, o que até o presente momento demonstra que esse *locus* está seguindo um modelo neutro de transmissão. Isso indica que não está ocorrendo e/ou não ocorreu, recentemente, eventos populacionais, pelo menos no que tange o presente gene testado, incluindo seleção preferencial de alelos, demasiada expansão ou diminuição da população, isolamento geográfico, grandes migrações ou cruzamentos preferenciais.

Esses resultados são congruentes com uma população panmítica, na qual um SNP ou combinação de SNPs não geram mortalidade e/ou dificuldade de reprodução, de forma a eliminar, dificultar ou promover a transmissão de alelos e, por consequência, gerar encontros genotípicos preferenciais. Outra explicação para os testes não serem significativos, pode estar na insuficiência de variações obtidas, sendo assim, seria necessária a ampliação das sequências a serem pesquisadas no gene *HFE*.

VIII.5 Descrição das frequências haplotípicas, teste exato de Fisher para haplótipo e diplótipos da região codificadora do gene *HFE*.

Nas Tabela 10 e 11 estão descritos os 13 haplótipos, como também suas frequências. Cabe salientar que ainda não existe nomenclatura padronizada para o gene *HFE*. Assim, à semelhança do que já está descrito para os outros genes clássicos e não clássicos de histocompatibilidade, sugerimos uma nomenclatura, cuja normatização ainda depende de reconhecimento pelo The International ImMunoGeneTics Information System - IMGT (<http://www.imgt.org>). Já enviamos essa normatização e o IMGT, que está analisando nosso pedido, mediante o aceite será submetido o trabalho para publicação (manuscrito no ANEXO XII).

O haplótipo *HFE**01:01 foi o mais frequente nas amostras de populações do presente estudo (de 48% a 63%), exceto na amostra de população Hemocromatose Hereditária (14%), que teve como haplótipo mais frequente o *HFE**01:03 (71%). Esse resultado faz perfeita analogia com os dados obtidos com o SNP C282Y em análises anteriores neste trabalho, já que a única diferença entre a sequência do *HFE**01:01 (padrão) e à *HFE**01:03 é que a última contém o alelo mutado do SNP C282Y.

Seguindo a mesma analogia, o teste exato de Fisher que comparou a amostra de população Hemocromatose Hereditária e a amostra de população Controle, verificando diferenças entre os haplótipos *HFE**01:01 e *HFE**01:03, identificou o *HFE**01:01 como um haplótipo protetivo (OR=0,14) e *HFE**01:03 como um haplótipo de risco (OR=60,00). Novamente, devido ao *n* pequeno na amostra de população Hemocromatose Hereditária o haplótipo *HFE**01:03 teve seu intervalo de confiança dilatado (IC= 20,31 a 177,20).

Os diplótipos gerados da combinação desses dois haplótipos *HFE**01:01 e *HFE**01:03, na comparação entre as amostras população Controle e Hemocromatose Hereditária, também apresentou diferença significativa (*P*=0,03). Os diplótipos *HFE**01:01 / *HFE**01:03 (OR=7,20) e *HFE**01:03 / *HFE**01:03 (OR=174,20) se

mostraram associados com risco para o desenvolvimento de hemocromatose hereditária. Novamente, deve-se verificar que, apesar de um presumível maior risco para o diplótipo *HFE**01:03 / *HFE**01:03, já que o mesmo se apresenta em dupla dose, deve-se considerar que a não recorrência desse diplótipo na população controle gera o valor de *Odds Ratio* e o intervalo de confiança dilatados.

Uma exceção ao aplicar analogias entre os diplótipos e genótipo é o genótipo TT do SNP IVS2(+4)T>C, já que os haplótipos *HFE**01:01:01:01; *01:01:01:05; *01:01:01:06; *01:02:01:01; *01:02:01:05; *01:03:01:01 e *01:04:01:01 são portadores do alelo T, e diplótipos oriundos da combinação desses haplótipos não apresentaram mudanças significativas de frequência e nem diferenças significativas no teste exato de Fisher (incluindo a comparação de amostras de população Controle e de Hemocromatose Hereditária). O provável motivo para não detecção desses haplótipos como sendo de risco, provavelmente deve-se ao fato da ocorrência do alelo mutado desse SNP IVS2(+4)T>C em sete haplótipos, o que pode “dissolver” a diferença entre as amostras, fazendo com que os testes percam seu poder de detecção.

Em contraponto, foi detectado um diplótipo (*HFE**01:02:01:01 / *HFE**01:01:01:05) de risco (OR=14,14) na comparação entre as amostras de população de câncer hepatocelular com sobrecarga de ferro e controle, resultado esse sem análogos nas análises de SNPs. Esse achado indica que a susceptibilidade à sobrecarga de ferro pode ser atribuída à composição geral da proteína e não devido a um sítio em particular.

Em resumo, esses resultados demonstram que apesar de, em linhas gerais, se esperar que padrões de associações encontradas em SNPs se mantenham em haplótipos, e vice-versa, essa relação pode não ser direta, sendo necessário que sempre se façam análises em vários níveis (SNP, haplótipo, haplótipo estendido) para encontrar padrões associadas ao fenótipo de interesse.

VIII.6 Teste exato de Fisher dos SNPs do gene *HFE* encontrados no trabalho, comparado com diversas populações

Além das comparações com estudos específicos avaliando os diversos sítios polimórficos do gene *HFE*, foram também realizadas comparações com populações geradas a partir do agrupamento dos indivíduos de ancestral comum dos continentes: africano, americano, asiático e europeu e, uma última com o agrupamento geral, chamada de população mundial.

A comparação entre as amostras do presente estudo e as 14 populações mundiais registradas no sítio 1000genomes (<http://www.1000genomes.org>) demonstrou que a amostra de população Controle do presente trabalho se aproxima muito das populações americanas por motivo óbvio, já que a amostra de população Controle do trabalho é proveniente do continente americano. Porém, a amostra de população Controle também se assemelha muito com as populações europeias, o que se justifica principalmente pela composição da população controle que é majoritariamente composta por brancos (80%). No entanto, surpreende o fato que, entre todas as populações testadas, a amostra de população Controle deste trabalho mais se assemelhe à população britânica, já que o Brasil não tem relatos históricos de colonização ou de grandes migrações advindas dessa região. Outro dado interessante é que quando comparamos o quanto as populações são semelhantes ou diferentes em relação à população controle do presente trabalho, verificamos que as populações africanas apresentam aproximadamente a mesma quantidade de SNPs com características semelhantes e diferentes. Isso demonstra a forte influência da miscigenação da população brasileira com africanos, mesmo em uma população majoritariamente branca. Um dado a ser destacado, é que as semelhanças aumentam ao se comparar a amostra de população Controle com a população de Americanos de ancestralidade Africana no Sudoeste dos EUA, provavelmente, essa maior proximidade com a população de americanos com ancestralidade africana advém do fato dessa população ser mais miscigenada com caucasianos do que as populações advindas da África. Já com relação às outras populações africanas, e foi observado que a quantidade de comparações “semelhantes” e “diferentes” se equilibra perfeitamente, quando confrontado a amostra de população Controle com a de Luhya em Webuye, Kenya, e se diferencia mais da população de Iorubá em Ibadan, Nigéria, sendo esse um resultado discrepante. Já que, chefes políticos e mercadores da África Centro-Ocidental (hoje região ocupada por Angola), forneceram a maior parte dos escravos utilizados em toda a América portuguesa. No século XVIII, o comércio do Rio de Janeiro, Recife e São Paulo era suprido por escravos que vinham da costa leste africana (oceano Índico), particularmente Moçambique. No comércio baiano, a partir de meados do século XVII, e até o fim do tráfico, os escravos eram oriundos da região do Golfo de Benin (sudoeste da atual Nigéria), tendo em mente que, o golfo do Benin compreende o litoral que vai da costa do Marfim até a Nigéria. (IBGE, 2000b)

Finalmente, verifica-se um maior distanciamento com as populações asiáticas, o que é esperado já que apesar de ter havido migrações para o Brasil de populações

advindas dessas regiões, não foi em volume suficiente para influenciar a composição genética brasileira, principalmente quando lembramos que a população controle usada na comparação é predominantemente branca.

Em linhas gerais, pode-se verificar que a amostra de população Controle se aproxima muito das amostras de população européias (ou de ancestralidade européia) e americanas, com maior semelhança com a população de Britânicos na Inglaterra e Escócia. Com relação às populações africanas (ou de ancestralidade africana), mostrou-se equilibrado, tendo proporcionalmente SNPs com “semelhanças” e “diferenças”. E por fim, a amostra de população asiática se comportou de forma muito discrepante (Asiáticos; Chineses Han em Bejing; Chineses Han do Sul; Japonenses em Tóquio). No presente trabalho, não se conseguiu explicar os resultados discrepantes encontrados nas análises de comparação entre a população Controle e a população britânica, do Kenya e da Nigéria. Provavelmente, para melhor análise desses dados são necessários outros testes, seja pela montagem dos haplótipos desses *loci*, e/ou pelo aumento da quantidade de *loci* analisados.

VIII.7 Hardy-Weinberg e Desequilíbrio de Ligação do Haplótipo Estendido

A amostra de população Controle foi anteriormente genotipada para dez *loci* do CPH (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQB1*, *TNFA*, *TNFB*, *TNFC*, *TNFD*, *HLA-G* 14pb). Esses dados nos permitiram fazer testes de desequilíbrio de ligação com o gene *HFE*, para verificar se o tão descrito desequilíbrio (exemplo: AJIOKA *et al.*, 1997; CARDOSO *et al.*, 2002; PORTO *et al.*, 1998) entre *HFE* e *HLA-A* e *HLA-B* também seria encontrado nesta amostra.

O primeiro teste foi de desequilíbrio de Hardy-Weinberg utilizando o *software* GENEPOP, que encontrou desequilíbrio significativo apenas no *locus HLA-DRB1* ($P=0,01$) (Tabela 14). Tal desequilíbrio foi determinado por excesso de heterozigoto determinado pelo teste global.

O teste par a par entre os *loci*, não detectou desequilíbrio com o gene *HFE* (Tabela 15), o que se pode explicar pelo fato do gene *HFE* estar em região mais telomérica no cromossomo 6, e muito distante dos outros *loci* testados.

Quando comparados os SNPs C282Y e H63D, verificamos desequilíbrio significativo ($P=0,03$) entre o SNP H63D e o *loci HLA-B* e, sendo assim, para verificar qual alelo faria parte dessa combinação realizamos os cálculos de D' e r^2 e encontramos

quatro combinações de alelos em desequilíbrio significativo entre *HLA-B*/H63 (15/C; 37/G; 44/G e 56/G) (Tabela 17).

Dessas quatro combinações, duas (*HLA-B**15 com H63D C; *HLA-B**56 com H63D G) apresentaram desequilíbrio com $D'=1$, e isso se deve ao fato de o *HLA-B* ser um gene multialélico, e do alelo G do SNP H63D ser um alelo raro, o que aumenta muito as chances de nem todos os haplótipos possíveis serem representados. Outro dado importante é a fraca intensidade dessa ligação, já que o coeficiente de recombinação varia entre 0,0199 a 0,1183. O coeficiente de recombinação da combinação *HLA-B**44 com H63D G, Qui^2 (23,54) e r^2 (0,11), mostra-se relativamente maior quando maior se comparado com outros pares significativos (3,96-5,06 e 0,01-0,02, respectivamente), o que se justifica pela diferença um pouco mais pronunciada entre as frequências observadas e esperadas na combinação desses dois alelos (observada: 8-11; esperada: 15,66-3,34). Porém, mesmo para este caso, os valores de desequilíbrio são considerados baixos. Muito provavelmente, esse desequilíbrio tenha uma origem histórica, ou seja, esses alelos surgiram em momentos próximos, e pensando assim, conclui-se que esses alelos não estão sofrendo pressão de seleção e que, portanto, esse valor tende a se diluir de geração em geração.

Devido ao fato do SNP IVS2(+4)T>C ter apresentado expressivo desequilíbrio de ligação com o SNP H63D, e ambos ficarem em uma distância relativamente pequena (157 pb), foi feito um segundo teste com esse SNP e o *loci HLA-B*, no intuito de confirmar o possível desequilíbrio entre esses dois *loci*, *HFE* e *HLA-B*. Os resultados obtidos nesse novo teste foi muito similar ao anterior. Foram encontrados cinco pares de alelos dos genes *HLA-B* / IVS2(+4)T>C com desequilíbrios significativos (35/T; 37/C; 44/C; 49/T e 58/C), sendo que os únicos pares coincidentes com os pares do teste anterior foram os alelos 37 e 44, porém nenhum dos dois se mostrou expressivo ($D'=1,00$ e 0,48; $r^2=0,02$ e 0,03) (Tabela 18).

Um dado importante a ser destacado, é a coerência dos desequilíbrios de ligação entre os alelos dos SNPS H63D, IVS2(+4)T>C e o *loci HLA-B*. Em testes anteriores, identificamos o desequilíbrio entre os alelos H63D G e IVS2(+4) C (Tabela 7 e Figura 9), neste teste houve desequilíbrio entre os alelos H63D G e *HLA-B**44 e IVS2(+4) C e *HLA-B**44 (Tabela 17). Tendo esse resultado, foi deduzido um haplótipo de desequilíbrio entre os dois SNPs, H63D e IVS2(+4) e o *loci HLA-B* (G-C-44) (Figura 11).

Um estudo realizado na Itália, em população saudável, demonstrou aumento na frequência de indivíduos homozigotos para IVS2(+4) C portadores do haplótipo *HLA-A*29-B*44*. No mesmo trabalho foi feita uma segunda análise para se determinar se os dois alelos dos SNPs H63D G e IVS2(+4) C, e os alelos dos *loci HLA-A*29* e *B*44* constituíam um único haplótipo (*HLA-A*29/B*44-H63D G-IVS2(+4) C*). Para esta análise foi utilizada uma amostra de indivíduos homozigotos para o SNP H63D G. Ao serem observados os resultados verificou-se que todos eram homozigotos para IVS2(+4) C. Entretanto, ao realizar outra análise, com indivíduos homozigotos para IVS2(+4) C, somente 65% eram portadores do SNP H63D G. Correlacionando essas análises às observações de frequências, De Lucas *et al.* (2005) deduziram que, tendo que o IVS2(+4) C se mostrou mais frequente que este seja mais antigo que o H63D G.

Na população portuguesa (PORTO *et al.*, 1998), não foi encontrada associação entre alelos do *loci HLA-B* e alelos do gene *HFE* em pacientes com hemocromatose hereditária, porém no mesmo estudo foram encontradas associações em grupo de pacientes com sobrecarga de ferro advinda de doenças hepáticas virais, distúrbios lipídicos, alcoolismo crônico. Esse estudo apontou desequilíbrio de ligação entre os alelos *HLA-B*44 / C282Y A*. Trabalhos realizados na Suécia, Grã Bretanha, Utah e Alabama também apontam aumento na frequência absoluta do *HLA-A*03-B*44* em pacientes com hemocromatose hereditária, portadores do alelo C282Y A (BARTON; ACTON, 2002; LESAGE *et al.*, 1983; SIMON *et al.*, 1987b).

Em linhas gerais, no presente trabalho não se detectou desequilíbrio entre os alelos do gene *HLA-A* e SNPs do gene *HFE* (em especial C282Y), contrariando dois trabalhos feitos na população brasileira anteriormente (BITTENCOURT *et al.*, 2002; CANÇADO *et al.*, 2006), mas detectou desequilíbrio entre os alelos H63D G e *HLA-B*44*, provavelmente de origem histórica e não por pressão de seleção. Corrobora essa hipótese, além da fraca intensidade do desequilíbrio, o fato de que o haplótipo *HLA-A*29-B*44-DRB1*07* é um dos haplótipos mais difundidos na Europa, comum no Oeste e Norte da África e o segundo mais comum haplótipo em Portugal (ARNAIZ-VILLENA *et al.*, 1995; ARNAIZ-VILLENA *et al.*, 1997). Barton *et al.* (2005a) também relatam que o haplótipo *HLA-A*02-B*44* é muito comum em caucasianos do Alabama. Sendo assim, a chance de ocorrência da combinação haplotípica *HLA-B*44 / H63D G* deve-se possivelmente ao fato de que o alelo G tenha surgido em indivíduos portadores do alelo *HLA-B*44*, dado que este alelo é muito frequente e, provavelmente, de origem mais antiga. Entretanto, já que não havia dados disponíveis dos *loci* do CPH

para as populações de pacientes do presente trabalho, não foi possível fazer os testes de desequilíbrio de ligação nessas populações. Porém, baseado nos dados descritos acima, sugerimos que talvez haja desequilíbrio até mais pronunciado nas amostras de populações de pacientes, ligado diretamente aos fenótipos das morbidades (especialmente com hemocromatose hereditária), mas que, no entanto, não pode ser verificado.

IX. CONCLUSÕES

No presente trabalho, foi possível determinar o alelo C282Y A como relacionado com o desenvolvimento da HH, e analogamente, o haplótipo que contém essa mutação, *HFE**01:03. O diplótipo *HFE**01:02:01:01/*HFE**01:01:01:05 também foi identificado como de risco em pacientes com CHC com sobrecarga de ferro.

Com base nos seis SNPs e uma nova deleção encontrada, normatizamos a nomenclatura dos haplótipos para o gene *HFE*, tendo como modelo a nomenclatura oficial dos alelos dos genes *HLA*.

Estudo de desequilíbrio de ligação demonstrou forte ligação entre os alelos H63D G e IVS2(+4) C, nas amostras das populações com HH e Controle, entretanto essa ligação não se mostrou associada com o fenótipo da doença.

Comparações entre a amostra de população Controle e várias populações disponíveis do site 1000genomes, demonstraram que a nossa população estava mais próxima das populações europeias e americanas, sendo parcialmente próxima das populações africanas e distante das populações asiáticas.

Estudo de diversidade não apontou grandes diferenças entre as diversas populações do presente estudo, mas mostrou uma relativa diferenciação da amostra da população HH, atribuída pelo forte desequilíbrio dos SNPs IVS2(+4)T>C, H63D e C282Y ou ao pequeno tamanho amostral.

O teste de neutralidade não identificou nenhuma pressão de seleção na amostra Controle, ou seja, a amostra se mostrou como uma típica população panmítica.

O estudo do haplótipo estendido, entre o gene *HFE* e dez *loci* do CPH, identificou desequilíbrio entre os alelos do gene *HFE* (H63D G, IVS2(+44) C) e o *locus* *HLA-B**44, sendo que pela intensidade desse desequilíbrio, parece ser de origem histórica.

X. REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, P. C. *et al.* Clinical and family studies in genetic hemochromatosis: microsatellite and HFE studies in five atypical families. **Hepatology**, v. 26, n. 4, p. 986-990, 2003.
- AFDHAL, N. H. The natural history of hepatitis C. **Seminars in liver disease**, 2004. Copyright© 2004 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. p.3-8.
- AGOSTINHO, M. F. *et al.* Mutation analysis of the HFE gene in Brazilian populations. **Blood Cells, Molecules, and Diseases**, v. 25, n. 6, p. 324-327, 1999.
- AGUILAR MARTINEZ, P. *et al.* Compound heterozygotes for hemochromatosis gene mutations: may they help to understand the pathophysiology of the disease? **Blood Cells, Molecules, and Diseases**, v. 23, n. 2, p. 269-276, 1997.
- AJIOKA, R. S. *et al.* Haplotype analysis of hemochromatosis: evaluation of different linkage-disequilibrium approaches and evolution of disease chromosomes. **The American Journal of Human Genetics**, v. 60, n. 6, p. 1439-1447, 1997.
- ALBANIS, E.; FRIEDMAN, S. Antifibrotic agents for liver disease. **American Journal of transplantation**, v. 6, n. 1, p. 12-19, 2005.
- ALBERT, E. D.; BAUR, M.; MAYR, W. Histocompatibility testing 1984. **Class II Antigen Reports. Berlin, Springer-Verlag**, 1984.
- AMMANN, R. W. *et al.* High incidence of extrahepatic carcinomas in idiopathic hemochromatosis. **Scandinavian journal of gastroenterology**, v. 15, n. 6, p. 733-736, 1980.
- ANDERSON, G. J. *et al.* HFE gene and hemochromatosis. **Journal of gastroenterology and hepatology**, v. 19, n. 6, p. 712-712, 2004.
- ANDREWS, N. C. Disorders of Iron Metabolism. **New England Journal of Medicine**, v. 341, n. 26, p. 1986-1995, 1999.
- ARNAIZ-VILLENA, A. *et al.* HLA allele and haplotype frequencies in Algerians: Relatedness to Spaniards and Basques. **Human immunology**, v. 43, n. 4, p. 259-268, 1995.
- ARNAIZ-VILLENA, A. *et al.* Relatedness among Basques, Portuguese, Spaniards, and Algerians studied by HLA allelic frequencies and haplotypes. **Immunogenetics**, v. 47, n. 1, p. 37-43, 1997.
- BACON, B. R. *et al.* Molecular medicine and hemochromatosis: at the crossroads. **Gastroenterology**, v. 116, n. 1, p. 193-207, 1999.
- BARRETT, J. *et al.* Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v. 21, n. 2, p. 263-265, 2005.
- BARTON, J. C.; ACTON, R. HLA-A and -B alleles and haplotypes in hemochromatosis probands with HFE C282Y homozygosity in central Alabama. **BMC Medical Genetics**, v. 3, n. 1, p. 9, 2002.
- BARTON, J. C. *et al.* Two novel missense mutations of the HFE gene (I105T and G93R) and identification of the S65C mutation in Alabama hemochromatosis probands. **Blood cells, molecules & diseases**, v. 25, n. 3, p. 147-155, 1999.
- BARTON, J. C. *et al.* Genetic and clinical description of hemochromatosis probands and heterozygotes: evidence that multiple genes linked to the major histocompatibility complex are responsible for hemochromatosis. **Blood Cells, Molecules, and Diseases**, v. 23, n. 1, p. 135-145, 1997.

- BARTON, J. C. *et al.* HLA haplotype A*03-B*07 in hemochromatosis probands with HFE C282Y homozygosity: frequency disparity in men and women and lack of association with severity of iron overload. **Blood Cells, Molecules, and Diseases**, v. 34, n. 1, p. 38-47, 2005a.
- BARTON, J. C. *et al.* Total blood lymphocyte counts in hemochromatosis probands with HFE C282Y homozygosity : relationship to severity of iron overload and HLA-A and -B alleles and haplotypes. v. 9, p. 1-9, 2005b.
- BATALLER, R.; NORTH, K. E.; BRENNER, D. A. Genetic polymorphisms and the progression of liver fibrosis: a critical appraisal. **Hepatology**, v. 37, n. 3, p. 493-503, 2003.
- BEDOSSA, P.; PARADIS, V. Liver extracellular matrix in health and disease. **The Journal of pathology**, v. 200, n. 4, p. 504-515, 2003.
- BENNETT, M. J.; LEBRÓN, J. A.; BJORKMAN, P. J. Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor. **Nature**, v. 403, n. 6765, p. 46-53, 2000.
- BEST, L.; HARRIS, P.; SPRIGGS, E. Hemochromatosis mutations C282Y and H63D in 'cis' phase. **Clinical genetics**, v. 60, n. 1, p. 68-72, 2001.
- BEUTLER, E. *et al.* Mutation analysis in hereditary hemochromatosis. **Blood Cells, Molecules, and Diseases**, v. 22, n. 2, p. 187-194, 1996.
- BEUTLER, E.; WEST, C. New Diallelic Markers in the HLA Region of Chromosome 6. **Blood Cells, Molecules, and Diseases**, v. 23, n. 2, p. 219-229, 8// 1997.
- BITTENCOURT, P. *et al.* Analysis of HLA-A antigens and C282Y and H63D mutations of the HFE gene in Brazilian patients with hemochromatosis. **Brazilian journal of medical and biological research**, v. 35, n. 3, p. 329-335, 2002.
- BOIGE, V. *et al.* Lack of association between HFE gene mutations and hepatocellular carcinoma in patients with cirrhosis. **Gut**, v. 52, n. 8, p. 1178-1181, 2003.
- BOMFORD, A.; WILLIAMS, R. Long term results of venesection therapy in idiopathic haemochromatosis. **QJM**, v. 45, n. 4, p. 611-623, 1976.
- BONKOVSKY, H. L. *et al.* Roles of Iron and HFE Mutations on Severity and Response to Therapy During Retreatment of Advanced Chronic Hepatitis C. **Gastroenterology**, v. 131, n. 5, p. 1440-1451, 2006.
- BOROT, N. *et al.* Mutations in the MHC class I-like candidate gene for hemochromatosis in French patients. **Immunogenetics**, v. 45, n. 5, p. 320-324, 1997.
- BRADBPEAR, R. A. *et al.* Cohort study of internal malignancy in genetic hemochromatosis and other chronic nonalcoholic liver diseases. **Journal of the National Cancer Institute**, v. 75, n. 1, p. 81-84, 1985.
- BRISOT, P. *et al.* A genotypic study of 217 unrelated probands diagnosed as "genetic hemochromatosis" on "classical" phenotypic criteria. **J Hepatol**, v. 30, n. 4, p. 588-93, Apr 1999.
- BRUIX, J.; SHERMAN, M. Management of hepatocellular carcinoma. **Hepatology**, v. 42, n. 5, p. 1208-1236, 2005.
- BUENO, S.; DUCH, C. R.; FIGUEIREDO, M. S. Mutations in the HFE gene (C282Y, H63D, S65C) in a Brazilian population. **Revista Brasileira de Hematologia e Hemoterapia**, v. 28, n. 4, p. 293-295, 2006.
- BULAJ, Z. J. *et al.* Clinical and biochemical abnormalities in people heterozygous for hemochromatosis. **N Engl J Med**, v. 335, n. 24, p. 1799-805, Dec 12 1996.

CAMASCHELLA, C. Understanding iron homeostasis through genetic analysis of hemochromatosis and related disorders. **Blood**, v. 106, n. 12, p. 3710-3717, 2005.

CAMASCHELLA, C. *et al.* The gene TFR2 is mutated in a new type of haemochromatosis mapping to 7q22. **Nature genetics**, v. 25, n. 1, p. 14-15, 2000.

CANÇADO, R. D. *et al.* Estudo das mutações C282Y, H63D e S65C do gene HFE em doentes brasileiros com sobrecarga de ferro. **Rev. bras. hematol. hemoter**, v. 29, n. 4, p. 351-60, 2007.

CANÇADO, R. D. *et al.* Analysis of HFE gene mutations and HLA-A alleles in Brazilian patients with iron overload. **Sao Paulo Med J**, v. 124, n. 2, p. 55-60, 2006.

CARDOSO, C. S.; DE SOUSA, M. HFE, the MHC and hemochromatosis: paradigm for an extended function for MHC class I. **Tissue Antigens**, v. 61, n. 4, p. 263-275, Apr 2003.

CARDOSO, C. S. *et al.* Co-selection of the H63D mutation and the HLA-A29 allele: a new paradigm of linkage disequilibrium? **Immunogenetics-New York Then Heidelberg-**, v. 53, n. 12, p. 1002-1008, 2002.

CARDOSO, C. S. *et al.* Comparative study of the two more frequent HFE mutations (C282Y and H63D): significant different allelic frequencies between the North and South of Portugal. **European journal of human genetics: EJHG**, v. 9, n. 11, p. 843, 2001.

CARDOSO, E. M. *et al.* HFE mutations in patients with hereditary haemochromatosis in Sweden. **J Intern Med**, v. 243, n. 3, p. 203-8, Mar 1998.

CARELLA, M. *et al.* Mutation analysis of the HLA-H gene in Italian hemochromatosis patients. **American journal of human genetics**, v. 60, n. 4, p. 828, 1997.

CARMEL, I. *et al.* Comparative analysis detects dependencies among the 5' splice-site positions. **Rna**, v. 10, n. 5, p. 828-840, 2004.

CARVALHO, A. S. HLA-A, B and C markers in the Portuguese population. **Tissue Antigens**, v. 21, n. 1, p. 39-44, 1983.

CASTELLI, E. C. *et al.* The genetic structure of 3' untranslated region of the HLA-G gene: polymorphisms and haplotypes. **Genes and immunity**, v. 11, n. 2, p. 134-141, 2009.

CASTELLI, E. C. *et al.* A comprehensive study of polymorphic sites along the HLA-G gene: implication for gene regulation and evolution. **Molecular biology and evolution**, v. 28, n. 11, p. 3069-3086, 2011.

CASTRO, C. M.; PEREIRA, A. D. S. Mapeamento por associação em plantas: conceitos básicos e perspectivas de uso no melhoramento de batata. **Embrapa Clima Temperado. Documentos**, v. 248, 2008.

CAUZA, E. *et al.* Mutations of the HFE gene in patients with hepatocellular carcinoma. **The American journal of gastroenterology**, v. 98, n. 2, p. 442-447, 2003.

CHANG, K. HCV: Immune response and its modifiers in clinical development. In: WRIGHT; ROCKY (Ed.). **Liver disease: from bench to bedside: Postgraduate course. American Association for the Study of Liver Diseases, Boston** 2004.

ČIMBUROVÁ, M. *et al.* S65C and other mutations in the haemochromatosis gene in the Czech population. **Folia Biologica (Praha)**, v. 51, p. 172-176, 2005.

CRUZ, E. *et al.* Involvement of the major histocompatibility complex region in the genetic regulation of circulating CD8 T-cell numbers in humans. **Tissue Antigens**, v. 64, n. 1, p. 25-34, Jul 2004.

CURCIO, M. *et al.* Haplotype analysis of the H63D, IVS2+ 4t/c, and C282Y polymorphisms of the HFE gene reveals rare events of intragenic recombination. **European journal of haematology**, v. 80, n. 4, p. 341-345, 2008.

CURTIN, P. D. **The Atlantic slave trade: a census**. University of Wisconsin Press, 1972. ISBN 0299054047.

DARSHAN, D.; ANDERSON, G. J. Interacting signals in the control of hepcidin expression. **Biometals**, v. 22, n. 1, p. 77-87, 2009.

DATI, F. *et al.* Consensus of a group of professional societies and diagnostic companies on guidelines for interim reference ranges for 14 proteins in serum based on the standardization against the IFCC/BCR/CAP Reference Material (CRM 470). International Federation of Clinical Chemistry. Community Bureau of Reference of the Commission of the European Communities. College of American Pathologists. **Eur J Clin Chem Clin Biochem**, v. 34, n. 6, p. 517-20, Jun 1996.

DATZ, C. *et al.* Predominance of the HLA-H Cys282Tyr mutation in Austrian patients with genetic haemochromatosis. **Journal of hepatology**, v. 27, n. 5, p. 773, 1997.

DE DIEGO, C.; JOSÉ MURGA, M; MARTÍNEZ-CASTRO, P. Frequency of HFE H63D, S65C, and C282Y mutations in patients with iron overload and controls from Toledo, Spain. **Genetic testing**, v. 8, n. 3, p. 263-267, 2004.

DE LUCAS, A. P. *et al.* Is the IVS2+ 4T> C variant of the HFE gene a splicing mutation or a polymorphism? A study in the Spanish population. **Genetics in Medicine**, v. 7, n. 3, p. 212, 2005.

DE SILVA, D. M.; ASKWITH, C. C.; KAPLAN, J. Molecular mechanisms of iron uptake in eukaryotes. **Physiological reviews**, v. 76, n. 1, p. 31-47, 1996.

DE SOUSA, M. **Iron in immunity, cancer, and inflammation**. John Wiley & Sons Inc, 1989. ISBN 0471921505.

DEUGNIER, Y. M. Iron and liver cancer. **Alcohol**, v. 30, n. 2, p. 145-150, 2003.

DHILLON, B. K. *et al.* Frequency of primary iron overload and HFE gene mutations (C282Y, H63D and S65C) in chronic liver disease patients in north India. **World Journal of Gastroenterology**, v. 13, n. 21, p. 2956, 2007.

DHILLON, B. K. *et al.* H63D mutation in HFE gene is common in Indians and is associated with the European haplotype. **Journal of Genetics**, v. 91, n. 2, p. 229, 2012.

DISTANTE, S. *et al.* The origin and spread of the HFE-C282Y haemochromatosis mutation. **Human genetics**, v. 115, n. 4, p. 269-279, 2004.

DONALDSON, P. *et al.* HLA class II alleles in Chinese patients with hepatocellular carcinoma. **Liver**, v. 21, n. 2, p. 143-148, 2001.

DORAK, M. T.; BURNETT, A. K; WORWOOD, M. HFE gene mutations in susceptibility to childhood leukemia: HuGE review. **Genetics in Medicine**, v. 7, n. 3, p. 159-168, 2005.

DUPRADEAU, F.-Y. *et al.* Phenotypic expression of the C282Y/Q283P compound heterozygosity in< i> HFE</i> and molecular modeling of the Q283P mutation effect. **Blood Cells, Molecules, and Diseases**, v. 30, n. 3, p. 231-237, 2003.

EDWARDS, C. Q. *et al.* Prevalence of hemochromatosis among 11,065 presumably healthy blood donors. **The New England journal of medicine**, v. 318, n. 21, p. 1355, 1988.

EL-CHENNAWI, F. A. *et al.* HLA-class II alleles in Egyptian patients with hepatocellular carcinoma. **Immunological Investigations**, v. 37, n. 7, p. 661-674, 2008.

- EL-SERAG, H. B.; RUDOLPH, K. L. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. **Gastroenterology**, v. 132, n. 7, p. 2557-2576, 2007.
- ELMBERG, M. *et al.* Cancer risk in patients with hereditary hemochromatosis and in their first-degree relatives. **Gastroenterology**, v. 125, n. 6, p. 1733-1741, 2003.
- ERHARDT, A. *et al.* HFE mutations and chronic hepatitis C: H63D and C282Y heterozygosity are independent risk factors for liver fibrosis and cirrhosis. **Journal of hepatology**, v. 38, n. 3, p. 335, 2003.
- EWENS, W. J. The sampling theory of selectively neutral alleles. **Theoretical population biology**, v. 3, n. 1, p. 87-112, 1972.
- EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. **Evolutionary bioinformatics online**, v. 1, p. 47, 2005.
- EXCOFFIER, L.; SLATKIN, M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. **Molecular biology and evolution**, v. 12, n. 5, p. 921-927, 1995.
- FARGION, S. *et al.* Survival and prognostic factors in 212 Italian patients with genetic hemochromatosis. **Hepatology**, v. 15, n. 4, p. 655-659, 2005.
- FARGION, S. *et al.* Mutations in the HFE gene and their interaction with exogenous risk factors in hepatocellular carcinoma. **Blood cells, molecules & diseases**, v. 27, n. 2, p. 505, 2001.
- FEDER, J. N. *et al.* A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. **Nature genetics**, v. 13, n. 4, p. 399-408, 1996.
- FEDER, J. N. *et al.* The hemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding. **Proceedings of the National Academy of Sciences**, v. 95, n. 4, p. 1472-1477, 1998.
- FLINT-GARCIA, S. A.; THORNSBERRY, J. M. Structure of Linkage Disequilibrium in Plants*. **Annual Review of Plant Biology**, v. 54, n. 1, p. 357-374, 2003.
- FLOREANI, A. *et al.* Intron 2 [IVS2, T-C +4] HFE gene mutation associated with S65C causes alternative RNA splicing and is responsible for iron overload. **Hepatology Research**, v. 33, n. 1, p. 57-60, 9// 2005.
- FRACANZANI, A. L. *et al.* Increased cancer risk in a cohort of 230 patients with hereditary hemochromatosis in comparison to matched control patients with non-iron-related chronic liver disease. **Hepatology**, v. 33, n. 3, p. 647-651, 2001.
- FRANCHINI, M.; VENERI, D. Recent advances in hereditary hemochromatosis. **Annals of hematology**, v. 84, n. 6, p. 347-352, 2005.
- FRIEDMAN, S. L. Liver fibrosis—from bench to bedside. **Journal of hepatology**, v. 38, n. supplement 1, p. S38-S53, 2003.
- FUJII, T.; ISHITANI, A.; GERAGHTY, D. E. A soluble form of the HLA-G antigen is encoded by a messenger ribonucleic acid containing intron 4. **J Immunol**, v. 153, n. 12, p. 5516-24, Dec 15 1994.
- GALANTE, P. A. F. *et al.* Detection and evaluation of intron retention events in the human transcriptome. **Rna**, v. 10, n. 5, p. 757-765, 2004.
- GAO, J. *et al.* Interaction of the hereditary hemochromatosis protein, HFE, with transferrin receptor 2 is required for transferrin-induced hepcidin expression. **Cell metabolism**, v. 9, n. 3, p. 217, 2009.
- GAREWAL, G. *et al.* Prevalence of the H63D mutation of the HFE in north India: its presence does not cause iron overload in beta thalassemia trait. **European journal of haematology**, v. 74, n. 4, p. 333-336, 2005.

- GORHAM, J. D. Acute and chronic liver disease: immunologic mechanisms and therapy. **AASLD, Postgraduate Course**. In: VIERLING;PETER, *et al* (Ed.), 2005. p.11-12: 15-21.
- GOTTSCHALK, R. *et al*. HFE codon 63/282 (H63D/C282Y) dimorphism in German patients with genetic hemochromatosis. **Tissue Antigens**, v. 51, n. 3, p. 270-5, Mar 1998.
- GRIFFITHS, W. J. H. Review article: the genetic basis of haemochromatosis. **Alimentary pharmacology & therapeutics**, v. 26, n. 3, p. 331-342, 2007.
- GROTTO, H. Z. Fisiologia e metabolismo do ferro. **Rev. Bras. Hematol. Hemoter**, v. 32, n. Supl 2, p. 8-17, 2010.
- GUO, S. W.; THOMPSON, E. A. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. **Biometrics**, p. 361-372, 1992.
- GUTHRIE, C.; PATTERSON, B. Spliceosomal snRNAs. **Annual review of genetics**, v. 22, n. 1, p. 387-419, 1988.
- HANSON, E. H.; IMPERATORE, G.; BURKE, W. HFE gene and hereditary hemochromatosis: a HuGE review. **American journal of epidemiology**, v. 154, n. 3, p. 193-206, 2001.
- HEDRICK, P. W. Gametic disequilibrium measures: proceed with caution. **Genetics**, v. 117, n. 2, p. 331-341, 1987.
- HELLERBRAND, C. *et al*. HFE C282Y heterozygosity in hepatocellular carcinoma: evidence for an increased prevalence. **Clinical gastroenterology and hepatology: the official clinical practice journal of the American Gastroenterological Association**, v. 1, n. 4, p. 279, 2003.
- HILL, W.; ROBERTSON, A. Linkage disequilibrium in finite populations. **TAG Theoretical and Applied Genetics**, v. 38, n. 6, p. 226-231, 1968.
- HOLMSTRÖM, P. *et al*. Mild iron overload in patients carrying the HFE S65C gene mutation: a retrospective study in patients with suspected iron overload and healthy controls. **Gut**, v. 51, n. 5, p. 723-730, 2002.
- HORTON, R. *et al*. Gene map of the extended human MHC. **Nature Reviews Genetics**, v. 5, n. 12, p. 889-899, 2004.
- HU, K. Q. *et al*. Overweight and obesity, hepatic steatosis, and progression of chronic hepatitis C: a retrospective study on a large cohort of patients in the United States. **Journal of hepatology**, v. 40, n. 1, p. 147-154, 2004.
- IBGE - Censo Demográfico de 2000a - Instituto Brasileiro de Geografia e Estatística. **Acessado em 06/03/2001**,
- IBGE– Brasil: 500 Anos de Povoamento, Instituto Brasileiro de Geografia Estatística (ed) IBGE, Rio de Janeiro (2000b).
- ISHITANI, A.; GERAGHTY, D. E. Alternative splicing of HLA-G transcripts yields proteins with primary structures resembling both class I and class II antigens. **Proc Natl Acad Sci U S A**, v. 89, n. 9, p. 3947-51, May 1 1992.
- IVANOVA, A. *et al*. C282Y and H63D mutations in the HFE gene are not associated with porphyria cutanea tarda in Bulgaria. **Hepatology**, v. 30, n. 6, p. 1531-1532, 2003.
- JEFFREY, G. P. *et al*. Alternate splicing produces a soluble form of the hereditary hemochromatosis protein HFE. **Blood Cells, Molecules, and Diseases**, v. 25, n. 1, p. 61-67, 1999.

- JEN-EING, J. *et al.* Tumor necrosis factor- α 308.2 polymorphism is associated with advanced hepatic fibrosis and higher risk for hepatocellular carcinoma. **Neoplasia (New York, NY)**, v. 9, n. 11, p. 987, 2007.
- JOUANOLLE, A. M. *et al.* A candidate gene for hemochromatosis: frequency of the C282Y and H63D mutations. **Human genetics**, v. 100, n. 5, p. 544-547, 1997.
- KAUR, G. *et al.* Distribution of C282Y and H63D mutations in the HFE gene in healthy Asian Indians and patients with thalassaemia major. **National medical journal of India**, v. 16, n. 6, p. 309-310, 2003.
- KEW, M. D. Pathogenesis of hepatocellular carcinoma in hereditary hemochromatosis: occurrence in noncirrhotic patients. **Hepatology**, v. 11, n. 6, p. 1086-1087, 2005.
- KHUSAINOVA, R. I. *et al.* Analysis of H63D mutation in hemochromatosis (HFE) gene in populations of Central Eurasia. **Russian Journal of Genetics**, v. 49, n. 2, p. 237-245, 2013/02/01 2013.
- KIRSZENBAUM, M. *et al.* An alternatively spliced form of HLA-G mRNA in human trophoblasts and evidence for the presence of HLA-G transcript in adult lymphocytes. **Proc Natl Acad Sci U S A**, v. 91, n. 10, p. 4209-13, May 10 1994.
- KNAU, B.; ROSSNER, M.; SCHMIDT, U. [Idiopathic hemochromatosis: association with antigens of the HLA system]. **Med Klin (Munich)**, v. 86, n. 10, p. 503-7, Oct 15 1991.
- KNODELL, R. G. *et al.* Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. **Hepatology**, v. 1, n. 5, p. 431-435, 1981.
- KOWDLEY, K. V. Iron, hemochromatosis, and hepatocellular carcinoma. **Gastroenterology**, v. 127, n. 5, p. S79-S86, 2004.
- KRAUSE, A. *et al.* LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. **FEBS letters**, v. 480, n. 2, p. 147-150, 2000.
- KUSHNER, J. P.; PORTER, J. P.; OLIVIERI, N. F. Secondary iron overload. **ASH Education Program Book**, v. 2001, n. 1, p. 47-61, 2001.
- LAHAM, N. *et al.* Transferrins receptor association and endosomal localization of soluble HFE are not sufficient for regulation of cellular iron homeostasis. **Journal of cellular biochemistry**, v. 91, n. 6, p. 1130-1145, 2004.
- LEBRÓN, J. A. *et al.* Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. **Cell**, v. 93, n. 1, p. 111, 1998.
- LEONE, P. E. *et al.* Analysis of HFE gene mutations (C282Y, H63D, and S65C) in the Ecuadorian population. **Annals of hematology**, v. 84, n. 2, p. 103-105, 2005.
- LESAGE, G. D. *et al.* Hemochromatosis: genetic or alcohol-induced? **Gastroenterology**, v. 84, n. 6, p. 1471-7, Jun 1983.
- LLOVET, J. M.; BRUIX, J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: chemoembolization improves survival. **Hepatology**, v. 37, n. 2, p. 429-442, 2003.
- LOK, C. Y. *et al.* Iron overload in the Asian community. **Blood**, v. 114, n. 1, p. 20-25, 2009.
- LUCOTTE, G. Frequency analysis and allele map in favor of the Celtic origin of the C282Y mutation of hemochromatosis. **Blood Cells, Molecules, and Diseases**, v. 27, n. 2, p. 549-556, 2001.
- LUCOTTE, G.; DIETERLEN, F. A European allele map of the C282Y mutation of hemochromatosis: Celtic versus Viking origin of the mutation? **Blood Cells, Molecules, and Diseases**, v. 31, n. 2, p. 262-267, 2003.

- LUND, M.; KJEMS, J. Defining a 5'splice site by functional selection in the presence and absence of U1 snRNA 5'end. **Rna**, v. 8, n. 2, p. 166-179, 2002.
- LYON, E; FRANK, E. L. Hereditary hemochromatosis since discovery of the HFE gene. **Clinical chemistry**, v. 47, n. 7, p. 1147-1156, 2001.
- MACSWEEN, R.; SCOTT, A. Hepatic cirrhosis: a clinico-pathological review of 520 cases. **Journal of clinical pathology**, v. 26, n. 12, p. 936-942, 1973.
- MALLORY, M. A.; KOWDLEY, K. V. Hereditary hemochromatosis and cancer risk: more fuel to the fire? **Gastroenterology**, v. 121, n. 5, p. 1253, 2001.
- MARCELLIN, P.; ASSELAH, T.; BOYER, N. Fibrosis and disease progression in hepatitis C. **Hepatology**, v. 36, n. S1, p. S47-S56, 2006.
- MARIANI, R. *et al.* Prevalence of HFE mutations in upper Northern Italy: study of 1132 unrelated blood donors. **Digestive and liver disease**, v. 35, n. 7, p. 479-481, 2003.
- MARSH, S. G. E. *et al.* Nomenclature for factors of the HLA system, 2010. **Tissue antigens**, v. 75, n. 4, p. 291-455, 2010.
- MARTINELLI, A. L. C. *et al.* Are haemochromatosis mutations related to the severity of liver disease in hepatitis C virus infection? **Acta haematologica**, v. 102, n. 3, p. 152-156, 1999.
- MARTINELLI, A. L. C. *et al.* Porphyria cutanea tarda in Brazilian patients: association with hemochromatosis C282Y mutation and hepatitis C virus infection. **The American journal of gastroenterology**, v. 95, n. 12, p. 3516-3521, 2000.
- MARTINS, R. *et al.* Differential HFE gene expression is regulated by alternative splicing in human tissues. **PloS one**, v. 6, n. 3, p. e17542, 2011.
- MATSUMURA, M. *et al.* Emerging principles for the recognition of peptide antigens by MHC class I molecules. **Science (New York, NY)**, v. 257, n. 5072, p. 927, 1992.
- MAYR, W. R. Berechnung der vaterschaftsausschlusschance und der Vaterschaftswahrscheinlichkeit im HLA system. **International Journal of Legal Medicine**, v. 79, n. 2, p. 87-95, 1977.
- MCCAUGHAN, G. W.; GEORGE, J. Fibrosis progression in chronic hepatitis C virus infection. **Gut**, v. 53, n. 3, p. 318-321, 2004.
- MCCULLOUGH, A. J.; BERGET, S. M. G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. **Molecular and cellular biology**, v. 17, n. 8, p. 4562-4571, 1997.
- MCLAREN, C. E. *et al.* Hemochromatosis and Iron Overload Screening (HEIRS) study design for an evaluation of 100,000 primary care-based adults. **The American journal of the medical sciences**, v. 325, n. 2, p. 53-62, 2003.
- MERRYWEATHER-CLARKE, A. T. *et al.* Geography of HFE C282Y and H63D mutations. **Genetic testing**, v. 4, n. 2, p. 183-198, 2000.
- MERRYWEATHER-CLARKE, A. T. *et al.* Global prevalence of putative haemochromatosis mutations. **Journal of Medical Genetics**, v. 34, n. 4, p. 275-278, 1997.
- MIEDZYBRODZKA, Z. *et al.* Haemochromatosis mutations in North-East Scotland. **British journal of haematology**, v. 106, n. 2, p. 385-387, 1999.
- MIKHAILOVA, S. V. *et al.* The ethnospecific distribution of the HFE haplotypes for IVS2 (+ 4) t/c, IVS4 (- 44) t/c, and IVS5 (- 47) g/a in populations of Russia and possible effects of these single-

nucleotide polymorphisms in splicing. **Genetic Testing and Molecular Biomarkers**, v. 14, n. 4, p. 461-469, 2010.

MILLER, S.; DYKES, D.; POLESKY, H. A simple salting out procedure for extracting DNA from human nucleated cells. **Nucleic acids research**, v. 16, n. 3, p. 1215, 1988.

MILMAN, N. *et al.* Frequency of the hemochromatosis HFE mutations C282Y, H63D, and S65C in blood donors in the Faroe Islands. **Annals of hematology**, v. 84, n. 3, p. 146-149, 2005.

MILMAN, N. *et al.* HLA determinants in 70 Danish patients with idiopathic haemochromatosis. **Clinical genetics**, v. 33, n. 4, p. 286-292, 1988.

MILMAN, N. *et al.* Frequency of the C282Y and H63D mutations of the hemochromatosis gene (HFE) in 2501 ethnic Danes. **Annals of hematology**, v. 83, n. 10, p. 654-657, 2004.

MOCZULSKI, D. K.; GRZESZCZAK, W.; GAWLIK, B. Role of hemochromatosis C282Y and H63D mutations in HFE gene in development of type 2 diabetes and diabetic nephropathy. **Diabetes Care**, v. 24, n. 7, p. 1187-1191, 2001.

MURA, C.; RAGUENES, O.; FÉREC, C. HFE mutations analysis in 711 hemochromatosis probands: evidence for S65C implication in mild form of hemochromatosis. **Blood**, v. 93, n. 8, p. 2502-2505, 1999.

MURPHY, S. *et al.* High incidence of the Cys 282 Tyr mutation in the HFE gene in the Irish population-implications for haemochromatosis. **Tissue Antigens**, v. 52, n. 5, p. 484-8, Nov 1998.

NAGAI, K. *et al.* Structure and assembly of the spliceosomal snRNPs. **Biochemical Society Transactions**, v. 29, n. 2, p. 15-26, 2001.

NAHON, P. *et al.* Liver iron, HFE gene mutations, and hepatocellular carcinoma occurrence in patients with cirrhosis. **Gastroenterology**, v. 134, n. 1, p. 102-110, 2008.

NEI, M. **Molecular evolutionary genetics**. Columbia University Press, 1987. ISBN 0231063210.

NEI, M.; KUMAR, S. **Molecular evolution and phylogenetics**. Oxford University Press, USA, 2000. ISBN 0199881227.

NIEDERAU, C. *et al.* Survival and causes of death in cirrhotic and in noncirrhotic patients with primary hemochromatosis. **New England Journal of Medicine**, v. 313, n. 20, p. 1256-1262, 1985.

NIELSEN, P. *et al.* Prevalence of the C282Y and H63D mutations in the HFE gene in patients with hereditary haemochromatosis and in control subjects from Northern Germany. **Br J Haematol**, v. 103, n. 3, p. 842-5, Dec 1998.

OLSSON, K. S. *et al.* HLA haplotype map of river valley populations with hemochromatosis traced through five centuries in Central Sweden. **European journal of haematology**, v. 81, n. 1, p. 36-46, 2008.

OLSSON, K. S. *et al.* The ancestral haplotype in patients with genetic hemochromatosis from Central and Western Sweden. International Symposium: Iron in Biology and Medicine, St. Malo, 1997.

PAPANIKOLAOU, G. *et al.* Mutations in HFE2 cause iron overload in chromosome 1q-linked juvenile hemochromatosis. **Nature genetics**, v. 36, n. 1, p. 77-82, 2003.

PARK, C. H. *et al.* Hepsidin, a urinary antimicrobial peptide synthesized in the liver. **Journal of Biological Chemistry**, v. 276, n. 11, p. 7806-7810, 2001.

PARKIN, D. M. Global cancer statistics in the year 2000. **The lancet oncology**, v. 2, n. 9, p. 533-543, 2001.

- PARRA, F. C. *et al.* Color and genomic ancestry in Brazilians. **Proceedings of the National Academy of Sciences**, v. 100, n. 1, p. 177-182, 2003.
- PAWLOTSKY, J. M. Pathophysiology of hepatitis C virus infection and related liver disease. **Trends in microbiology**, v. 12, n. 2, p. 96-102, 2004.
- PHATAK, P. D. *et al.* Prevalence and penetrance of HFE mutations in 4865 unselected primary care patients. **Blood Cells, Molecules, and Diseases**, v. 29, n. 1, p. 41-47, 2002.
- PIETRANGELO, A. Hereditary hemochromatosis—a new look at an old disease. **New England Journal of Medicine**, v. 350, n. 23, p. 2383-2397, 2004.
- PIETRANGELO, A. Hereditary Hemochromatosis: Pathogenesis, Diagnosis, and Treatment. **Gastroenterology**, v. 139, n. 2, p. 393-408.e2, 8// 2010.
- PIETRAPERIOSA, A. *et al.* HFE gene mutations in an Apulian population: allele frequencies. **European journal of epidemiology**, v. 18, n. 7, p. 685-690, 2003.
- PIPERNO, A. Classification and diagnosis of iron overload. **Haematologica**, v. 83, n. 5, p. 447-455, 1998.
- PIPERNO, A. *et al.* Heterogeneity of hemochromatosis in Italy. **Gastroenterology**, v. 114, n. 5, p. 996-1002, 1998.
- POINTON, J. J. *et al.* Uncommon mutations and polymorphisms in the hemochromatosis gene. **Genetic testing**, v. 4, n. 2, p. 151-161, 2000.
- PORTO, G. *et al.* Major histocompatibility complex class I associations in iron overload: evidence for a new link between the HFE H63D mutation, HLA-A29, and non-classical forms of hemochromatosis. **Immunogenetics**, v. 47, n. 5, p. 404-410, 1998.
- PORTO, G. *et al.* Idiopathic haemochromatosis in north Portugal: association with haplotype A3B7. **J Clin Pathol**, v. 42, n. 6, p. 667-8, Jun 1989.
- PORTO, G.; DE SOUSA, M. Variation of hemochromatosis prevalence and genotype in national groups. **Hemochromatosis. Genetics, Pathophysiology, Diagnosis, and Treatment**, p. 51-62, 2000.
- POWELL, L. W. *et al.* Is all genetic (hereditary) hemochromatosis HLA-associated. **Ann N Y Acad Sci**, v. 526, p. 23-33, 1988.
- POYNARD, T. *et al.* Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis C. **Journal of hepatology**, v. 34, n. 5, p. 730-739, 2001.
- PURTILO, D. T.; GOTTLIEB, L. S. Cirrhosis and hepatoma occurring at Boston city hospital (1917-1968). **Cancer**, v. 32, n. 2, p. 458-462, 2006.
- RACCHI, O. *et al.* Mutations of the HFE Gene and the Risk of Hepatocellular Carcinoma. **Blood Cells, Molecules, and Diseases**, v. 25, n. 6, p. 350-353, 1999.
- RAMADORI, G.; SAILE, B.; SAILE, B. Mesenchymal cells in the liver—one cell type or two? **Liver**, v. 22, n. 4, p. 283-294, 2002.
- RAMMENSEE, H. G.; FALK, K.; RÖTZSCHKE, O. Peptides naturally presented by MHC class I molecules. **Annual review of immunology**, v. 11, n. 1, p. 213-244, 1993.
- RAYMOND, M.; ROUSSET, F. An exact test for population differentiation. **Evolution**, p. 1280-1283, 1995.
- RICHARDSON, D. R.; PONKA, P. The molecular mechanisms of the metabolism and transport of iron in normal and neoplastic cells. **Biochimica et Biophysica acta**, v. 1331, n. 1, p. 1, 1997.

- RISTIĆ, S. *et al.* Hemochromatosis gene mutations in the Croatian and Slovenian populations. **Clinical genetics**, v. 64, n. 5, p. 444-446, 2003.
- RITTER, B.; SAFWENBERG, J.; OLSSON, K. S. HLA as a marker of the hemochromatosis gene in Sweden. **Hum Genet**, v. 68, n. 1, p. 62-6, 1984.
- ROGOZIN, I. B.; MILANESI, L. Analysis of donor splice sites in different eukaryotic organisms. **Journal of molecular evolution**, v. 45, n. 1, p. 50-59, 1997.
- RUSKIN, B.; ZAMORE, P. D.; GREEN, M. R. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. **Cell**, v. 52, n. 2, p. 207-219, 1/29/ 1988.
- SALTER-CID, L.; PETERSON, P. A.; YANG, Y. The major histocompatibility complex-encoded HFE in iron homeostasis and immune function. **Immunologic research**, v. 22, n. 1, p. 43-59, 2000.
- SÁNCHEZ, M. *et al.* Prevalence of the Cys282Tyr and His63Asp HFE gene mutations in Spanish patients with hereditary hemochromatosis and in controls. **Journal of hepatology**, v. 29, n. 5, p. 725, 1998.
- SANTOS, P. C. J. L. *et al.* Alterações moleculares associadas à hemocromatose hereditária. **Revista Brasileira de Hematologia e Hemoterapia**, v. 31, n. 3, p. 192-202, 2009.
- SANTOS, P. C. J. L. *et al.* HFE gene mutations and iron status of Brazilian blood donors. **Brazilian Journal of Medical and Biological Research**, v. 43, n. 1, p. 107-114, 2010.
- SAPER, M.; BJORKMAN, P.; WILEY, D. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 resolution. **Journal of molecular biology**, v. 219, n. 2, p. 277-319, 1991.
- ŠARIĆ, M. *et al.* Frequency of the hemochromatosis gene mutations in the population of Serbia and Montenegro. **Clinical genetics**, v. 70, n. 2, p. 170-172, 2006.
- SCHMIDT, P. J. *et al.* Transferrin receptor modulates Hfe-dependent regulation of hepcidin expression. **Cell metabolism**, v. 7, n. 3, p. 205, 2008.
- SELBY, J. V.; FRIEDMAN, G. D. Epidemiologic evidence of an association between body iron stores and risk of cancer. **International journal of cancer**, v. 41, n. 5, p. 677-682, 2006.
- SHAM, R. L. *et al.* Autosomal dominant hereditary hemochromatosis associated with a novel ferroportin mutation and unique clinical features. **Blood Cells, Molecules, and Diseases**, v. 34, n. 2, p. 157-161, 2005.
- SHELDON, J. H. **Haemochromatosis**. Oxford University Press, Humphrey Milford, publisher to the University, 1935.
- SIMON, M. *et al.* The genetics of hemochromatosis. **Progress in medical genetics**, v. 4, p. 135, 1980.
- SIMON, M. *et al.* Idiopathic hemochromatosis. Demonstration of recessive transmission and early detection by family HLA typing. **The New England journal of medicine**, v. 297, n. 19, p. 1017, 1977.
- SIMON, M. *et al.* A study of 609 HLA haplotypes marking for the hemochromatosis gene: (1) mapping of the gene near the HLA-A locus and characters required to define a heterozygous population and (2) hypothesis concerning the underlying cause of hemochromatosis-HLA association. **Am J Hum Genet**, v. 41, n. 2, p. 89-105, Aug 1987a.
- SIMON, M. *et al.* A study of 609 HLA haplotypes marking for the hemochromatosis gene:(1) mapping of the gene near the HLA-A locus and characters required to define a heterozygous population and (2) hypothesis concerning the underlying cause of hemochromatosis-HLA association. **American journal of human genetics**, v. 41, n. 2, p. 89, 1987b.

- SIMON, M. *et al.* Letter: Idiopathic hemochromatosis associated with HL-A 3 tissular antigen. **La Nouvelle presse médicale**, v. 4, n. 19, p. 1432, 1975.
- SIMON, M. *et al.* Genetics of hemochromatosis: HLA association and mode of inheritance. **Ann N Y Acad Sci**, v. 526, p. 11-22, 1988.
- SLATKIN, M. An exact test for neutrality based on the Ewens sampling distribution. **Genetical Research**, v. 64, n. 1, p. 71-74, 1994.
- SOREK, R.; AST, G. Intronic sequences flanking alternatively spliced exons are conserved between human and mouse. **Genome research**, v. 13, n. 7, p. 1631-1637, 2003.
- SOUSA, M. D. T lymphocytes and iron overload: novel correlations of possible significance to the biology of the immunological system. **Memórias do Instituto Oswaldo Cruz**, v. 87, p. 23-29, 1992.
- STEPHENS, M.; SMITH, N. J. e DONNELLY, P. A new statistical method for haplotype reconstruction from population data. **The American Journal of Human Genetics**, v. 68, n. 4, p. 978-989, 2001.
- STEVENS, R. G. *et al.* Body iron stores and the risk of cancer. **New England Journal of Medicine**, v. 319, n. 16, p. 1047-1052, 1988.
- STUART, K. A. *et al.* The C282Y mutation in the haemochromatosis gene (HFE) and hepatitis C virus infection are independent cofactors for porphyria cutanea tarda in Australian patients. **Journal of hepatology**, v. 28, n. 3, p. 404, 1998.
- TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, n. 3, p. 585-595, 1989.
- THAKUR, V. *et al.* Absence of hemochromatosis associated Cys282Tyr HFE gene mutation and low frequency of hemochromatosis phenotype in nonalcoholic chronic liver disease patients in India. **Journal of gastroenterology and hepatology**, v. 19, n. 1, p. 86-90, 2004.
- THORBURN, D. *et al.* The role of iron and haemochromatosis gene mutations in the progression of liver disease in chronic hepatitis C. **Gut**, v. 50, n. 2, p. 248-252, 2002.
- TORRES, F. *et al.* Frequency of the HFE C282Y and H63D polymorphisms in Brazilian malaria patients and blood donors from the Amazon region. **Genet Mol Res**, v. 7, n. 1, p. 60-4, 2008.
- TRIFA, A. P. *et al.* HFE Gene C282Y, H63D and S65C Mutations Frequency in the Transylvania Region, Romania. **Journal of gastrointestinal and liver diseases: JGLD**, v. 21, n. 2, p. 177, 2012.
- TROUSSEAU, A. Glycosurie, diabete sucre. **Anonymous clinique medicale de l'Hotel-Dieu de Paris**, v. 2, p. 663, 1865.
- TUNG, B. Y. *et al.* Hepatitis C, iron status, and disease severity: relationship with HFE mutations. **Gastroenterology**, v. 124, n. 2, p. 318, 2003.
- TURLIN, B. *et al.* Increased liver iron stores in patients with hepatocellular carcinoma developed on a noncirrhotic liver. **Hepatology**, v. 22, n. 2, p. 446-450, 1995.
- VAUTIER, G. *et al.* p53 mutations in British patients with hepatocellular carcinoma: clustering in genetic hemochromatosis. **Gastroenterology**, v. 117, n. 1, p. 154-160, 1999.
- VOICU, P. M. *et al.* Prevalence of HFE (hemochromatosis) gene mutations C282Y and H63D in a Romanian population. **Blood cells, molecules & diseases**, v. 42, n. 1, p. 14, 2009.
- WATTERSON, G. The homozygosity test of neutrality. **Genetics**, v. 88, n. 2, p. 405-417, 1978.
- WEIR, B.; COCKERHAM, C. C. Estimating F-statistics for the analysis of population structure. **Evolution**, p. 1358-1370, 1984.

- WEST, A. P. *et al.* Mutational analysis of the transferrin receptor reveals overlapping HFE and transferrin binding sites. **Journal of molecular biology**, v. 313, n. 2, p. 385-397, 2001.
- WHITFIELD, J. *et al.* Effects of HFE C282Y and H63D Polymorphisms and Polygenic Background on Iron Stores in a Large Community Sample of Twins. **The American Journal of Human Genetics**, v. 66, n. 4, p. 1246-1258, 2000.
- WILDI, S. *et al.* Critical evaluation of the different staging systems for hepatocellular carcinoma. **British journal of surgery**, v. 91, n. 4, p. 400-408, 2004.
- WILLIS, G. *et al.* A high prevalence of HLA-H 845A mutations in hemochromatosis patients and the normal population in eastern England. **Blood Cells, Molecules, and Diseases**, v. 23, n. 2, p. 288-291, 1997.
- WILLIS, G. *et al.* Incidence of liver disease in people with HFE mutations. **Gut**, v. 46, n. 3, p. 401-404, March 1, 2000 2000.
- WORWOOD, M. *et al.* The HLA A1-B8 haplotype extends 6 Mb beyond HLA-A: associations between HLA-A, B, F and 15 microsatellite markers. **Tissue Antigens**, v. 50, n. 5, p. 521-6, Nov 1997.
- WRAY, N. R.; VISSCHER, P. M. Statistical genetics: gene mapping through linkage and association. In: NEALE; FERREIRA, *et al* (Ed.): Taylor & Francis New York, 2008. p.87-112. ISBN 0415410401.
- YEO, G. W.; VAN NOSTRAND, E. L.; LIANG, T. Y. Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. **PLoS genetics**, v. 3, n. 5, p. e85, 2007.

XI. APÊNDICE

Tabela 19. Lista de sinónimas dos sete SNPs encontrados nos sites: NCBI (<http://www.ncbi.nlm.nih.gov/snp/>) e 1000genomes (<http://www.1000genomes.org/>)

Nome do SNPs	Sinónimos
rs1799945 (H63D)	CM960827
	NC_000006.11:g.26091179C< G
	NG_008720.1:g.8671C< G
	NM_000410.3:c.187C< G
	NM_139003.2:c.187C< G
	NM_139004.2:c.187C< G
	NM_139006.2:c.187C< G
	NM_139007.2:c.77-363C< G
	NM_139008.2:c.77-363C< G
	NM_139009.2:c.118C< G
	NM_139010.2:c.77-1734C< G
	NM_139011.2:c.77-2168C< G
	NP_000401.1:p.His63Asp
	NP_620572.1:p.His63Asp
	NP_620573.1:p.His63Asp
	NP_620575.1:p.His63Asp
	NP_620578.1:p.His40Asp
	NT_007592.15:g.26031179C< G
rs2071303 (IVS2(+4) C)	rs56939457
	rs56518267
	rs60682736
	rs17515611
	CS991415
	NC_000006.11:g.26091336T< C
	NG_008720.1:g.8828T< C
	NM_000410.3:c.340+4T< C
	NM_139003.2:c.340+4T< C
	NM_139004.2:c.340+4T< C
	NM_139006.2:c.340+4T< C
	NM_139007.2:c.77-206T< C
	NM_139008.2:c.77-206T< C
	NM_139009.2:c.271+4T< C
	NM_139010.2:c.77-1577T< C
	NM_139011.2:c.77-2011T< C
	NT_007592.15:g.26031336T< C
rs807209	NC_000006.11:g.26092865C< G
	NG_008720.1:g.10357C< G
	NM_000410.3:c.617-48C< G
	NM_139003.2:c.341-90C< G
	NM_139004.2:c.341-48C< G
	NM_139006.2:c.617-90C< G
	NM_139007.2:c.353-48C< G
	NM_139008.2:c.353-90C< G
	NM_139009.2:c.548-48C< G
	NM_139010.2:c.77-48C< G
	NM_139011.2:c.77-482C< G
	NT_007592.15:g.26032865C< G

Nome do SNPs	Sinônimos
rs1800562	rs58044250
	rs111535158
	rs17530654
	rs4134660
	CM960828
	CM004391
	NC_000006.11:g.26093141G< A
	NG_008720.1:g.10633G< A
	NM_000410.3:c.845G< A
	NM_139003.2:c.527G< A
	NM_139004.2:c.569G< A
	NM_139006.2:c.803G< A
	NM_139007.2:c.581G< A
	NM_139008.2:c.539G< A
	NM_139009.2:c.776G< A
	NM_139010.2:c.305G< A
	NM_139011.2:c.77-206G< A
	NP_000401.1:p.Cys282Tyr
	NP_620572.1:p.Cys176Tyr
	NP_620573.1:p.Cys190Tyr
	NP_620575.1:p.Cys268Tyr
	NP_620576.1:p.Cys194Tyr
	NP_620577.1:p.Cys180Tyr
	NP_620578.1:p.Cys259Tyr
	NP_620579.1:p.Cys102Tyr
	NT_007592.15:g.26033141G< A
rs2794717	NC_000006.11:g.26093297G< A
	NG_008720.1:g.10789G< A
	NM_000410.3:c.893-50G< A
	NM_139003.2:c.575-50G< A
	NM_139004.2:c.617-50G< A
	NM_139006.2:c.851-50G< A
	NM_139007.2:c.629-50G< A
	NM_139008.2:c.587-50G< A
	NM_139009.2:c.824-50G< A
	NM_139010.2:c.353-50G< A
	NM_139011.2:c.77-50G< A
(IVS4(-44)T>C) rs1800708	NT_007592.15:g.26033297G< A
	NC_000006.11:g.26093303T< C
	NG_008720.1:g.10795T< C
	NM_000410.3:c.893-44T< C
	NM_139003.2:c.575-44T< C
	NM_139004.2:c.617-44T< C
	NM_139006.2:c.851-44T< C
	NM_139007.2:c.629-44T< C
	NM_139008.2:c.587-44T< C
	NM_139009.2:c.824-44T< C
	NM_139010.2:c.353-44T< C
	NM_139011.2:c.77-44T< C
	NT_007592.15:g.26033303T< C

XII. ANEXO

Genetic structure of *HFE* coding region: SNPs, haplotypes and allele nomenclature standardization.

Campos, W.N.¹, Massaro, J.D.¹; Martinelli, A.L.C.²; Mendes-Junior, C.T.³; Donadi, E.A.¹

¹ Division of Clinical Immunology, Department of Medicine, School of Medicine of Ribeirão Preto, University of São Paulo, Brazil

² Division of Gastroenterology, Department of Medicine, School of Medicine of Ribeirão Preto, University of São Paulo, Brazil

³ Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 14040-901, Ribeirão Preto-SP, Brazil.

Key words: HFE; haplotype; nomenclature; Brazilian population; new deletion

Correspondence

Wagner Narciso de Campos
Division of Clinical Immunology
Department of Medicine
School of Medicine of Ribeirão Preto
University of São Paulo
Av. Bandeirantes, 3900
14049-900, Ribeirão Preto
SP, Brazil
Tel: +55 16 3602 3373
Fax: +55 16 3633 6695
e-mail: wncampos@yahoo.com.br

Abstract

The coding region of the *HFE* gene (exons 2, 3, 4 and 5) was sequenced in 304 individuals, consisting of healthy individuals and patients with hereditary hemochromatosis, hepatitis C and hepatocellular carcinoma in the Brazilian population. Considering all samples, seven variation sites were detected: H63D C>G in exon 2 (rs1799945), IVS2 (+4) T>C in intron 2 (rs2071303), intron 3 C>G (rs807209), C282Y G>A in exon 4 (rs1800562), intron 4 G>A (rs2794717), IVS4 (-44) T>C (rs1800708), and a new deletion G>del in intron 5. The variations observed were used for the

inference of 13 haplotypes (*HFE* alleles), which were named according to the IMGT nomenclature rules for HLA class I non-classical available at IMGT website.

HFE gene (named for the abbreviation of High Fe, according to the HUGO Gene Nomenclature Committee - HGNC - <http://www.genenames.org>), initially designated as *HLA-H* (1) is located at the chromosomal region 6p21.3, in the telomeric region of the extended major histocompatibility complex (MHC) and was classified as a similar *HLA* class I non-classical gene (2).

The *HFE* gene has 7.952 Kb and is composed of seven exons, with the seventh exon corresponding to the 3' untranslated region (3'UTR). The introns sizes range from 158 base pairs (bp) (intron 4) to approximately 3335 bp (intron 1) (1). This gene translates 343 amino acids in the protein molecule, which presents α chain consisting of three extracellular domains, a transmembrane region and a cytoplasmic tail. The exon 1 contains only the signal peptide (3), exons 2, 3 and 4 encode the domains $\alpha 1$, $\alpha 2$ and $\alpha 3$, respectively, exon 5 encodes the transmembrane domain and finally exon 6 contains an in frame stop codon and result in a shortened cytoplasmic tail (3). The $\alpha 3$ domain is linked by a noncovalent bind to β_2 - μ globulin (encoded by another gene on chromosome 15). The *HFE* molecule contains four highly conserved cysteine residues which forms two intramolecular disulfide bridges, one in the $\alpha 2$ (C225) and the other in the $\alpha 3$ (C282) domains (4). These domains ($\alpha 2$ and $\alpha 3$) are essential for the tertiary structure of the protein and its interaction with β_2 - μ globulin, enabling *HFE* exposure on the cell surface (5). The protein encoded by the *HFE* gene is strongly related to the regulation of intestinal absorption of iron (6, 7).

The *HFE* gene has several SNPs and mutations previously described throughout exon 2 (V53M, V59M, H63D and S65C), exon 3 (Q127H and E168Q) and exon 4 (C282Y, Q283P, P168X, E168X and W169X), but only three SNPs are frequently studied: C282Y, H63D and S65C (5, 8, 9). However, there is no comprehensive work that verifies the genetic diversity and haplotype structure of the entire *HFE* coding region and present a standardized nomenclature system shared by others structurally analogous MHC genes. Admixed populations from Brazil are particularly suitable to achieving these goals, since Brazilians represent one of the most heterogeneous populations in the world as the result of five centuries of variable degrees of interethnic crosses of people from three continents: European colonizers mainly represented by the Portuguese, Africans brought by them as slaves, and autochthonous Amerindians (10).

In order to identify recurrent SNPs in the *HFE* gene on Brazilian population we sequenced its coding region (exon 2, 3, 4, 5) and intron boundaries, in a population

sample of 130 healthy individuals and 184 patients of various morbidities (14 subjects with hereditary hemochromatosis, 130 with hepatitis C and 60 with hepatocellular cancer). These 304 individuals are from the city of Ribeirão Preto region, in the state of São Paulo, southeastern Brazil. Exons 2, 3, 4 and 5 were chosen since these exons contain the principal domains of HFE protein (exons 2, 3 and 4), and the transmembrane domain (exon 5).

DNA samples were extracted from peripheral blood by a salting-out procedure (11). *HFE* alleles were defined by nucleotide sequence variations at exons 2, 3, 4 and 5. Each exon was individually amplified by Polymerase Chain Reaction (PCR) using the sets of primers describe at Table 1. Primers were designed using the GENNERUNNER v. 3.05 software (Hastings Software, Inc., Hastings, NY), as well as tools from the UCSC Genome Bioinformatics Site (<http://genome.ucsc.edu>). PCR for each exon was performed in a final volume of 50 µl containing 1X PCR buffer (70 mM Tris-HCl pH 8.8, 20 mM (NH₄)₂SO₄, 3 mM MgCl₂, 1mM dithiothreitol (DTT)), 0.2 mM of each dNTP, 10 pmol of each primer, 1.1 units of Taq DNA-polymerase Platinum (Invitrogen, Carlsbad, CA), and 200 ng of genomic DNA. The initial denaturation cycle was carried out at 94°C for 5 minutes, followed by 30 cycles of 45 seconds at 94°C, 45 seconds at annealing temperature (Table 1) and 1 minute at 72°C, and by a final extension step at 72°C for 7 minutes. PCR products were directly sequenced in a ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), using the same primers designed for amplification (Table 1), according to manufacturer's suggestions.

The SNPs observed were used to assemble the genetic profile of these samples and reconstruct their *HFE* haplotypes by computational procedures. The haplotype inference was performed using the and ELB (alpha value: 0.01; epsilon value: 0.01; gamma value: 0; sampling interval: 500; burn-in steps: 100,000; recombination steps: 0) algorithms as well as the PHASE method [number of iterations: 1000 (iterations in the final run increased 10 times); thinning interval: 1; burn-in value: 100; δ value for each locus: 1; number of independent runs: 10], implemented at the ARLEQUIN v3.1 (12), and PHASE v.2. (13) softwares, respectively. For haplotype estimates several simulations in which the population samples were either inputted as a single group or in their various structured subgroups (controls, patients with or without iron overload, hepatocellular cancer, hepatitis C, hereditary hemochromatosis) were performed in order to obtained consensual and reliable haplotypes for each sample. Subsequently, based on the rules of the MHC genes official nomenclature available at IMGT / HLA

Database (<http://hla.alleles.org/nomenclature/naming.html>), which are unique for all HLA genes, we established a standardized nomenclature for *HFE* alleles (Figure 1) using as reference the work of Marsh 2010 (13). The rules are described below:

1. Each *HFE* allele name has a unique number that corresponds to up to four groups of two digits, separated by colons. The length of allele assignment is dependent on the sequence of the allele regarding each SNP, compared to the reference allele from the reference sequence (NCBI - NG_008720.1 and Ensembl - ENSG00000010704). All alleles receive at least two groups of digits. More groups of digits are assigned only when necessary.
2. The first two digits before the first "colon" describe the type, which in most cases corresponds to the serological antigen that is carried by an allotype. The next group is used to list subtypes, of which the numbers are assigned in the order in which the DNA sequences were determined. Therefore, alleles which the numbers differ in the first two sets of digits must differ by one or more non-synonymous nucleotide substitutions that alter the amino acid sequence of the protein. The alleles that differ only by synonymous nucleotide substitutions (called silent substitutions) within the coding sequence, are distinguished by using the third set of digits. Finally, the fourth group of digits is used to designate alleles that differ due to changes in introns.

Six SNPs and a new mutation were observed in the *HFE* gene: i) H63D C>G in exon 2 (rs1799945), ii) IVS2 (+4) T>C in intron 2 (rs2071303), iii) the intron 3 C>G (rs807209), iv) C282Y G>A in exon 4 (rs1800562), v) intron 4 G>A (rs2794717), vi) IVS4 (-44) T>C (rs1800708), vii) in intron 5 a new deletion G>del. The sequence including this new mutation was cloned and sequenced in order to validate it, as can be seen in the chromatogram (Figure 2). The identification of SNPs was performed using the NCBI (<http://www.ncbi.nlm.nih.gov>) and Ensembl (<http://www.ensembl.org>) databases.

The haplotype inference analysis resulted in 13 haplotypes as described in Table 2. A single individual with hepatocellular cancer showed a discordant result by the ELB method in only one simulation by ARLEQUIN software v.3.1 (11). All other simulations provided identical results, regardless of the method or population structure

inserted. The discordant haplotype showed low probability of occurrence (0.092), when recalculated by PHASE v.2 software (12). Given that, the r haplotype inferred by all PHASE and EM runs and by almost all ELB runs was assigned for that sample.

This work represents a more comprehensive evaluation of the variability in the *HFE* coding region in the Brazilian population, describing the most frequent *HFE* alleles. It provides a relevant contribution for population genetics and case-control association studies, since haplotypes instead of isolated SNPs may be related to human diseases. Finally, this paper proposes a standardized nomenclature for these alleles, so that the alleles derived from future haplotype studies may be easily identified, handled and compared.

Acknowledgments

CNPq-Brazil - National Council for Scientific and Technological Development for the financial support. Prof. Dr. Eduardo Luiz Rachid Cançado and Dra. Andréia Evangelista da Silva, for giving samples of patients with hereditary hemochromatosis who were important to a better development of this research.

References

1. Feder JN, Gnirke A, Thomas W, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nature genetics*. 1996;**13**:399-408.
2. Horton R, Wilming L, Rand V, et al. Gene map of the extended human MHC. *Nature Reviews Genetics*. 2004;**5**:889-99.
3. Rhodes D, Trowsdale J. Alternate splice variants of the hemochromatosis gene HFE. *Immunogenetics*. 1999;**49**:357-9.
4. Dupradeau F-Y, Mura C, Jacolot S, et al. Phenotypic expression of the C282Y/Q283P compound heterozygosity in *HFE* and molecular modeling of the Q283P mutation effect. *Blood Cells, Molecules, and Diseases*. 2003;**30**:231-7.
5. Merryweather-Clarke AT, Pointon JJ, Jouanolle AM, Rochette J, Robson KJ. Geography of HFE C282Y and H63D mutations. *Genet Test*. 2000;**4**:183-98.
6. Krause A, Neitz S, Mägert HJ, et al. LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. *FEBS letters*. 2000;**480**:147-50.
7. Park CH, Valore EV, Waring AJ, Ganz T. Hepcidin, a urinary antimicrobial peptide synthesized in the liver. *Journal of Biological Chemistry*. 2001;**276**:7806-10.

8. Franchini M, Veneri D. Recent advances in hereditary hemochromatosis. *Annals of hematology*. 2005;**84**:347-52.
9. Anderson GJ, Ramm G, Subramaniam V, Powell L. HFE gene and hemochromatosis. *Journal of gastroenterology and hepatology*. 2004;**19**:712-.
10. Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. Color and genomic ancestry in Brazilians. *Proceedings of the National Academy of Sciences*. 2003;**100**:177-82.
11. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic acids research*. 1988;**16**:1215.
12. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*. 2005;**1**:47.
13. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics*. 2001;**68**:978-89.

TABLES

Table 1. Description of *HFE* gene target regions, the sequences of the primers with annealing temperature, and the amplicon length.

Amplified Region	Names and primers sequences	Annealing temperature	Amplicon length
Intron 1 (partial)	HFEex2F: 5'-GCTGATGGTATGAGTTGATGC-3'	60°C	444 bp
Exon 2	HFEex2R: 5'-CTCAGACTTCCAGCTGTTTC-3'		
Intron 2 (partial)			
Intron 2 (partial)	HFEex3F: 5'-GCTTCCTGAGATCATTTGGTCC-3'	59°C	485 bp
Exon 3	HFEex3R: 5'-CAGAATTTGGAGAGGCACACAG-3'		
Intron 3 (partial)			
Intron 3 (partial)	HFEex4F: 5'-GTATTCCTTCCTCCAACC-3'	53°C	451 bp
Exon 4	HFEex4R: 5'-CATAATTACCTCCTCAGGCAC-3'		
Intron 4 (partial)			
Intron 4 (partial)	HFEex5F: 5'-GTGCCTGAGGAGGTAATTATGG-3'	58°C	308 bp
Exon 5	3'		
Intron 5 (partial)	HFEex5R: 5'-CTTCCCATGGATGCCAGATC-3'		

Table 2. Haplotype identification and SNP alleles constitution. The mutated bases are in boldface and variation sites at exonic regions are indicated with gray background.

Haplotype identification	Haplotype composition							Exchanged aminoacids	Absolute allele frequency	Relative allele frequency
	+3511 ^a	+3668 ^b	+5197 ^c	+5473 ^d	+5629 ^e	+5635 ^f	+5811 ^g			
<i>HFE</i> *01:01	C	T	C	G	G	T	G		334	0.5493
<i>HFE</i> *01:01:01:01	C	C	C	G	G	T	G		63	0.1036
<i>HFE</i> *01:01:01:02	C	T	G	G	G	T	G		12	0.1974
<i>HFE</i> *01:01:01:03	C	T	C	G	A	T	G		2	0.0329
<i>HFE</i> *01:01:01:04	C	T	C	G	G	C	G		3	0.0493
<i>HFE</i> *01:01:01:05	C	C	C	G	G	C	G		59	0.9704
<i>HFE</i> *01:01:01:06	C	C	C	G	G	C	del		1	0.0164
<i>HFE</i> *01:02	G	T	C	G	G	T	G	H63D	2	0.0329
<i>HFE</i> *01:02:01:01	G	C	C	G	G	T	G	H63D	89	0.1464
<i>HFE</i> *01:02:01:05	G	C	C	G	G	C	G	H63D	1	0.0164
<i>HFE</i> *01:03	C	T	C	A	G	T	G	C282Y	40	0.6579
<i>HFE</i> *01:03:01:01	C	C	C	A	G	T	G	C282Y	1	0.0164
<i>HFE</i> *01:04:01:01	G	C	C	A	G	T	G	H63D/C282Y	1	0.0164

Identification of SNPs and most frequent allele according to NCBI (<http://www.ncbi.nlm.nih.gov>) and Ensembl (<http://www.ensembl.org>). Nucleotide locations were assigned considering the first transcribed base as +1: a - rs1799945 (H63D); b - rs2071303 (IVS2 (+4) T>C); c - rs807209; d - rs1800562 (C282Y); e - rs2794717; f - rs1800708 (IVS4 (-44) T>C); g - deletion in intron 5.

FIGURES

Figure 1. Description of the meaning of each element in the nomenclature of the *HFE* gene alleles.

Figure 2. Chromatogram indicating a new indel at in intron 5, identified in an individual from the healthy population sample. The deletion is present in heterozygosis, causing a shift in the reading frame and shuffling the bases in downstream sequencing (A). This is not observed at the cloned strand (B).