



**Universidade Federal do Pampa**

**UNIVERSIDADE FEDERAL DO PAMPA**

**CAMPUS URUGUAIANA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA**

**SAMARA MARQUES DOS REIS**

**CORRELAÇÃO ESTATÍSTICA ENTRE OS DADOS DE FREQUÊNCIAS  
GENÉTICAS E DADOS DE PREVALÊNCIA DE DOENÇA PODEM  
COMPLEMENTAR OS ESTUDOS DE CASO-CONTROLE PARA IDENTIFICAR  
LOCI SUSCEPTIBILIDADE EM ESTUDOS DE ASSOCIAÇÃO GENÉTICA**

**DISSERTAÇÃO DE MESTRADO**

**Uruguaiiana**

**2015**

**SAMARA MARQUES DOS REIS**

**CORRELAÇÃO ESTATÍSTICA ENTRE OS DADOS DE FREQUÊNCIAS  
GENÉTICAS E DADOS DE PREVALÊNCIA DE DOENÇA PODEM  
COMPLEMENTAR OS ESTUDOS DE CASO-CONTROLE PARA  
IDENTIFICAR LOCI SUSCEPTIBILIDADE EM ESTUDOS DE  
ASSOCIAÇÃO GENÉTICA**

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Pampa como requisito parcial para obtenção do Título de Mestre em Bioquímica.

Orientador: Prof. Dr. Andrés Delgado Cañedo

**Uruguaiiana**

**2015**

DOS REIS, Samara Marques

Correlação estatística entre os dados de frequências genéticas e dados de prevalência de doença podem complementar os estudos de caso-controle para identificar loci susceptibilidade em estudos de associação genética / Samara Marques Dos Reis.

50 folhas;

Dissertação (Mestrado) Universidade Federal do Pampa, 2015. Orientação: Prof. Dr. Andrés Delgado Cañedo.

1.Polimorfismo. 2.Depressão.

3.TPH2

I. Delgado-Cañedo, Andrés.

II. Doutor

**SAMARA MARQUES DOS REIS**

**CORRELAÇÃO ESTATÍSTICA ENTRE OS DADOS DE FREQUÊNCIAS  
GENÉTICAS E DADOS DE PREVALÊNCIA DE DOENÇA PODEM  
COMPLEMENTAR OS ESTUDOS DE CASO-CONTROLE PARA  
IDENTIFICAR LOCI SUSCEPTIBILIDADE EM ESTUDOS DE  
ASSOCIAÇÃO GENÉTICA**

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Pampa como requisito parcial para obtenção do Título de Mestre em Bioquímica.

Área de concentração: Química e Bioquímica dos Produtos Biologicamente Ativos.

Dissertação defendida e aprovada em: 04 de março de 2015.

Banca examinadora:

---

Prof. Dr. Andrés Delgado Cañedo

Orientador

Programa de Pós-Graduação em Bioquímica – UNIPAMPA

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Jacqueline da Costa Escobar Piccoli

Programa de Pós-Graduação em Bioquímica – UNIPAMPA

---

Prof. Dr. José Artur Bogo Chies

Programa de Pós-Graduação em Genética e Biologia Molecular – UFRGS

Dedico esta dissertação a minha amada filha  
Mariana e a meus queridos pais.

## **AGRADECIMENTOS**

Inicio meus agradecimentos por DEUS, já que ele colocou pessoas especiais, a meu lado, sem as quais, certamente, não teria conseguido alcançar meu objetivo.

Agradeço ao meu orientador, professor Andrés Cañedo, pela confiança, pela oportunidade de trabalhar ao seu lado e apoio para a elaboração desta dissertação. Obrigada pela atenção incondicional e amizade.

Aos meus professores que sempre acreditaram no meu potencial e incentivaram-me a seguir em frente nos estudos, em especial as minhas professoras da graduação Claudia Zamberlam e Vanessa Kirsten, que sempre foram meus exemplos de competência e profissionalismo.

Aos Professores da minha banca de qualificação de defesa da dissertação pela contribuição e ensinamentos.

Ao meu amado esposo Leandro, que sempre me incentivou e auxiliou nessa caminhada. Acredito no amor e é esse amor que nos impulsiona a continuar rumo aos nossos sonhos!

Aos meus colegas e amigos do Serviço de Aproveitamento do Hospital de Guarnição de Alegrete, Adilson, Fan, Martins, Lucas, Elda, Suzete, Doca, Sônia, Felipe, Barbosa, Silveira e César, que colaboraram para que a minha ausência não fosse sentida, auxiliaram-me desde o início deste curso, o meu muito Obrigada!

Ao Coronel Vilela, Coronel Otávio e o Coronel Burgarelli e a Márcia que permitiram que eu seguisse meus estudos e oportunizaram uma flexibilização nos meus horários de trabalho. Agradecer a atitudes tão humanas chega a ser pouco. Que eu possa ajudar ao próximo assim como eles me ajudaram.

Aos meus amigos, agradeço por todos os que me acompanharam antes e durante esse processo, estando presentes, sentindo a minha ausência em determinados momentos, mas acreditando que a amizade é superior a isso tudo.

À Letícia que me ajudou nessa dissertação, e a todos aqueles que, direta ou indiretamente, colaboraram para a realização deste trabalho.

Aos meus irmãos, Willian e João Vitor, meu agradecimento especial, pois, a seu modo, sempre se orgulharam de mim e confiaram em meu trabalho. Obrigada pela confiança!

Aos meus queridos pais, Paulo e Marliza, que comigo partilharam e preencheram, para a minha filha, a ausência da mãe, agradeço profundamente a possibilidade de realizar este trabalho. Com minha mãe, reparto a alegria e a satisfação do trabalho concluído.

Muito obrigada!!!

“A persistência é o caminho do êxito.”

Charles Chaplin

## RESUMO

### **CORRELAÇÃO ESTATÍSTICA ENTRE OS DADOS DE FREQUÊNCIAS GENÉTICAS E DADOS DE PREVALÊNCIA DE DOENÇA PODEM COMPLEMENTAR OS ESTUDOS DE CASO-CONTROLE PARA IDENTIFICAR LOCI SUSCEPTIBILIDADE EM ESTUDOS DE ASSOCIAÇÃO GENÉTICA**

Estudos de associação gene-doença mostraram uma relação entre TPH2 e a depressão em diferentes populações, estudos, no entanto, têm sido produzidos resultados contraditórios, sendo a Triptofano hidroxilase-2 (TPH2) uma enzima limitante da taxa na via sintética para a serotonina do cérebro, vários estudos relatam os polimorfismos da enzima TPH2. Dois grandes projetos, o HapMap e o 1000 genomas, organizaram a maioria dos polimorfismos a partir do estudo de várias populações disponibilizando estes dados. Este trabalho tem como objetivo desenvolver um método de estudo para obtenção de possíveis marcadores de predisposição a doença a partir da correlação entre os dados epidemiológicos e frequências populacionais de polimorfismos, baseado na hipótese de que se numa população existe maior frequência de uma determinada patologia determinada geneticamente, então as variantes envolvidas deveriam estar em maior frequência e vice-versa. O modelo usado foi o envolvimento de variantes do gene TPH2 na predisposição à depressão. Os dados obtidos com correlação positiva em um dos genótipos homocigotos e também no alelo deste homocigoto sugeriram a presença de 10 polimorfismos (14,49% do total) possivelmente envolvidos no desenvolvimento do processo depressivo. Estes dados foram comparados com dados da literatura envolvendo estudos do tipo caso controle. Nestes trabalhos foram estudados 20 dos 69 polimorfismos descritos para o gene TPH2. Com exceção de um único polimorfismo, todos os dados obtidos com a nossa estratégia apresentaram-se iguais aos dados da literatura, inclusive quanto ao alelo que determinaria predisposição à depressão quando demonstrada associação. Portanto, propomos esta estratégia como uma forma alternativa de se realizar estudos do tipo Genome-Wide Association sem a necessidade de estudos caso-controle, apenas usando dados epidemiológicos da doença, diminuindo o tempo e custo destes estudos.

Palavras-chave: Polimorfismo, TPH2 e depressão.



## **ABSTRACT**

### **STATISTICAL CORRELATION BETWEEN GENETIC FREQUENCIES DATA AND PREVALENCE OF DISEASE DATA COULD COMPLEMENT CASE-CONTROL ASSAYS FOR IDENTIFY SUSCEPTIBILITY LOCI IN GENOME-WIDE ASSOCIATION STUDIES**

Disease-gene association studies reported a relation between the TPH2 and depression in different populations, however some studies have produced contradictory results, being the tryptophan hydroxylase-2 (TPH2) a limiting enzyme in the rate of synthetic route of serotonin in the brain, many studies reported the polymorphisms of the TPH2 enzyme. Two big projects, HapMap and 1000 genomes, organized the major part of these polymorphisms from the study of several populations becoming these data available. This work is aimed to develop a system to obtain possible predisposition markers of a disease from the correlation between epidemiological data and population frequencies of polymorphism, based in the hypothesis that if in a population there is more frequency of a certain kind of pathology genetically determined, the variables involved should be more frequent and vice-versa. The model used was the involvement of variables of the TPH2 gene in the predisposition of depression. The data obtained with positive correlation in one of homozygous genotypes and in the allele of this homozygous suggested the presence of 10 polymorphisms (total 14,49%) possibly related to the development of depression. These data were compared to literature data involving case control studies. In these work were studied 20 from 69 polymorphisms described to the TPH2 gene. With the exception of only one polymorphism, all the data obtained through the strategy proposed in this work have been equals to the literature data, including the allele that is determinant to the predisposition of depression when it is demonstrated the association. Therefore, it is proposed this strategy as an alternative to realize this kind of Genome-Wide Association studies without the necessity of a case control study, only using the epidemiological data from the disease, decreasing the time and the cost of this study.

Key-words: polymorphs, TPH2 and depression

## LISTA DE TABELAS

### MANUSCRITO

Tabela 1. Estudos de caso-controle publicados sobre a associação entre o polimorfismo do TPH 2 e a depressão .....	31
Tabela 2- Correlação entre as frequências genéticas e os índices de depressão dos diferentes SNPs do gene TPH2 .....	33

## LISTA DE FIGURAS

Figura 1. Estrutura química da serotonina .....	18
Figura 2. Biossíntese da Serotonina .....	19

## MANUSCRITO

Figura 1. Fluxograma da estratégia utilizada para analisar os dados .....	42
---	----

## SUMÁRIO

1 INTRODUÇÃO.....	14
2 REFERENCIAL TEÓRICO.....	16
2.1 DEPRESSÃO .....	16
2.2 SEROTONINA.....	18
2.3 POLIMORFISMOS: CONCEITO E ORGANIZAÇÃO DE BANCO DE DADOS .....	20
2.4 POLIMORFISMOS DA ENZIMA TPH2 .....	21
3 JUSTIFICATIVA .....	22
4 OBJETIVOS.....	22
4.1 Objetivo Geral .....	22
4.2 Objetivos Específicos .....	22
5 MANUSCRITO.....	23
6 CONCLUSÃO.....	43
7 PERSPECTIVAS. ....	44
REFERÊNCIAS BIBLIOGRÁFICAS.....	45
ANEXOS.....	49

## APRESENTAÇÃO

A presente dissertação foi dividida em três partes principais. Na parte I encontram-se a INTRODUÇÃO, REFERENCIAL TEÓRICO, JUSTIFICATIVA e OBJETIVOS. Os resultados que fazem parte desta dissertação estão apresentados sob a forma de manuscrito, no item MANUSCRITO, que se encontra na parte II deste trabalho. A seção materiais e métodos, resultados, discussão e referências bibliográficas, encontra-se no manuscrito.

Na parte III, desta dissertação, encontram-se o item CONCLUSÃO, apresentando interpretações e comentários sobre o resultado mostrado no manuscrito deste trabalho, o item PERSPECTIVAS, onde está exposto o possível estudo para dar continuidade a este trabalho e o item REFERÊNCIAS BIBLIOGRÁFICAS refere-se somente às citações que aparecem nos itens introdução e revisão bibliográfica da parte I.

Devido ao grande tamanho das tabelas que contêm os dados dos polimorfismos nas diferentes populações, as mesmas são oferecidas em formato digital no CD que acompanha esta dissertação.

## PARTE I

### 1 INTRODUÇÃO

A depressão é uma doença que mais incapacita pessoas, a Organização Mundial de Saúde (OMS) divulgou um estudo mostrando que o Brasil é um dos países com a maior prevalência da doença nos últimos anos, apresentando este distúrbio em 10,8% da população (BROMETET et al., 2011).

Esta patologia é alvo de inúmeros estudos devido à complexidade dos fatores envolvidos; entretanto, está bem relacionada com alterações bioquímicas nos neurotransmissores. Uma das alterações mais estudadas é a diminuição do neurotransmissor serotonina. A diminuição no referido neurotransmissor provoca alguns sintomas característicos como: alterações de humor, ansiedade, tristeza vital, angústia e inibição da psicomotibilidade nos enfermos (IBANEZ et al., 2014).

A serotonina é sintetizada a partir do Triptofano, que é hidroxilado resultando no produto 5 – hidroxitriptofano que posteriormente é descarboxilado produzindo a serotonina (5 – hidroxitriptamina ou 5-HT) (RAVI et al., 2011).

A enzima triptofano hidroxilase 2 (TPH 2) é a principal enzima limitante na via da síntese da serotonina no sistema nervoso central realizando a hidroxilação acima descrita. Como tal, seu gene pode ser considerado um gene candidato ao estudo de uma grande variedade de fenômenos comportamentais, incluindo transtornos do humor, suicídio e traços de personalidade (ZHANG et al., 2011). Vários estudos relatam associação de polimorfismos deste gene com a depressão (SHEN et al., 2011; NAZREE et al., 2013; PEREIRA et al., 2009).

Existem duas grandes bases de dados disponíveis que permitem o estudo das variações genéticas, o Projeto 1000 genomas e o Projeto *HapMap*. Nestas bases de dados é possível obtermos frequências alélicas, genotípicas e haplótipos de polimorfismos em várias populações mundiais, com livre acesso.

A predisposição genética indica a suscetibilidade a determinadas doenças e distúrbios, mas essa suscetibilidade pode ser alterada ou modificada pelos fatores ambientais. Sendo assim, indivíduos podem apresentar a predisposição genética e não apresentar sintomas. Portanto, população com altos níveis de depressão poderiam apresentar altas frequências de alelos ou genótipos envolvidos na predisposição a esta doença.

Neste trabalho, usamos o teste de correlação de Pearson entre os dados alélicos e genotípicos disponíveis nas bases de dados acima descritas e a prevalência de depressão nas diferentes populações para testarmos a hipótese de que: populações com alta prevalência de depressão poderiam apresentar frequências alélicas e genotípicas aumentadas daqueles polimorfismos envolvidos com a patologia, comparando os dados obtidos com aqueles já publicados em estudos experimentais do tipo caso-controle.

## 2 REFERENCIAL TEÓRICO

### 2.1 DEPRESSÃO

Considerada um grande problema para a saúde pública, a depressão é uma patologia prevalente e altamente incapacitante para a saúde. Os sintomas são tristeza, sentimentos negativos sobre si mesmo, a satisfação reduzida, perda de vínculos emocionais, crises de choro, perda de resposta ao humor, perda de apetite, distúrbios do sono, perda de libido, cansaço e retardo (FLECK, 2009; FERRARI, et al., 2013; MASCELLA, 2013).

A Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde (CID-10), caracteriza a depressão em dois tipos:

#### *-F32- Episódios depressivos*

Nos episódios de cada um dos três graus de depressão: leve, moderado ou grave, o paciente apresenta uma diminuição do humor, diminuição da energia, diminuição da capacidade de concentração e fadiga acentuada.

Apresentam, também, distúrbios do sono, diminuição do apetite e autoestima, ideias de culpabilidade e ou de indignidade, alterações de humor ou, dependendo das circunstâncias, pode ser acompanhado de sintomas ditos “somáticos”, como: perda de interesse ou prazer, despertar matinal precoce, várias horas antes da hora habitual de despertar, agravamento matinal da depressão, lentidão psicomotora importante, agitação, perda de apetite, perda de peso e perda da libido.

Inclui: episódios isolados de (um) (uma):

- depressão:
- psicogênica
- reativa
- reação depressiva

Exclui: quando associados com transtornos de conduta em F91.- (F92.0) transtornos (de):

- adaptação (F43.2)
- depressivo recorrente

#### *F33 Transtorno depressivo recorrente*

Caracterizado pelo acontecimento repetido de episódios depressivos correspondentes à descrição de um episódio depressivo (F32) na falta de prévio de episódios independentes de



exaltação de humor e de aumento de energia. O transtorno pode, contudo, comportar breves episódios caracterizados por uma alteração (aumento) de humor e da atividade (hipomania), ocorrendo imediatamente a um episódio depressivo e por um tratamento antidepressivo.

As formas mais graves do transtorno depressivo recorrente (F33.2 e F33.3) apresentam numerosos pontos comuns com os conceitos anteriores da depressão maníaco-depressiva, melancolia, depressão vital e depressão endógena. O episódio inicial pode acontecer em qualquer idade, sendo que a duração pode variar de algumas semanas a alguns meses.

O risco de ocorrência de um episódio maníaco pode acontecer, em caso de ocorrência de um episódio maníaco, o diagnóstico deve ser alterado pelo de transtorno afetivo bipolar (F31).

Inclui: episódios recorrentes de uma:

- depressão
- psicogênica
- reativa
- reação depressiva

Exclui: episódios depressivos recorrentes breves (F38.1).

Pesquisas epidemiológicas revelam que milhões de pessoas sofrem algum tipo de doença mental no mundo. Segundo a Organização Mundial da Saúde (2012), a depressão é um transtorno mental comum, que afeta mais de 350 milhões de pessoas de todas as idades, sendo a principal causa de incapacidade em todo o mundo, e um dos principais contribuintes para a carga global de doenças, afetando mais mulheres do que homens (MARAGNO et.al., 2006; BROMET et al., 2011; SILVA et al., 2014).

No Brasil, estima-se que há cerca de 13 milhões de pessoas afetadas pela depressão. De acordo com projeções, em 2030 poderá ser a doença deletéria mais prevalente no planeta, ficando à frente do câncer e de algumas doenças infecciosas (WHO, 2012).

Estudos revelam que a prevalência da depressão é mais acentuada em países de alta renda do que em países de baixa e média renda. Uma possível explicação para tal fato, é que a desigualdade de renda, promove uma ampla diversidade de condições, muitas vezes, levando o indivíduo a desenvolver a depressão (WILKINSON; PICKETT, 2006).

A depressão maior pode ter origem familiar, e sua herdabilidade resultaria principalmente de influências genéticas. Porém, a depressão é um distúrbio complexo com influências ambientais, psicossociais, bioquímicas e genéticas, que podem desencadear a patologia atuando de forma sinérgica (SULLIVAM et al 2000).

## 2.2 SEROTONINA

Entre as rotas bioquímicas mais estudadas na depressão, encontra-se aquela que afeta o neurotransmissor serotonina (5-hidroxitriptamina ou 5-HT). Esta molécula derivada do aminoácido essencial triptofano é conhecida por estar envolvida na emoção, aprendizagem e memória (DUMAN; VOLETI, 2012) e, a deficiência dela no cérebro pode ser um importante fator causal na ansiedade e depressão (Figura 1).

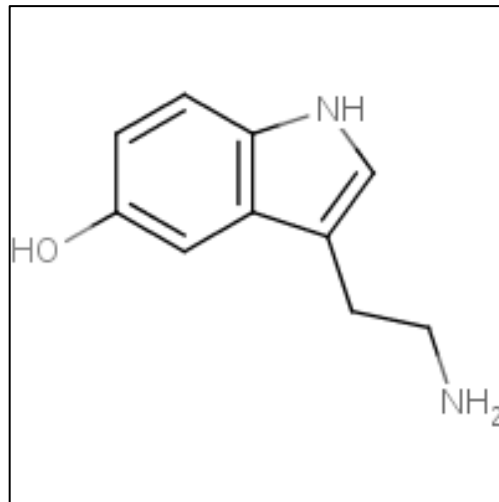


Figura 1- Estrutura química da serotonina

(Fonte: <http://chem.sis.nlm.nih.gov/chemidplus/rn/50-67-9>)

A 5HT é sintetizada e armazenada em diversos sítios do organismo (GUYTON; HALL, 1998). Sua fisiologia é complexa atuando como um neurotransmissor e neuromodulador (VEENSTRA et al; 2000). Age como inibidor das vias da dor, controle do humor e do sono (AMIREAULT; SIBON; CÔTE, 2013).

No sistema nervoso central, a transmissão serotoninérgica é regulada por meio da recaptura da serotonina pelo seu transportador (NI; WATTS, 2006). A primeira reação de síntese de 5HT é catalisada pela enzima triptofano hidroxilase que converte a L-triptofano em 5-hidroxitriptofano (5HT) (Figura 2) (RAVI et al., 2011).

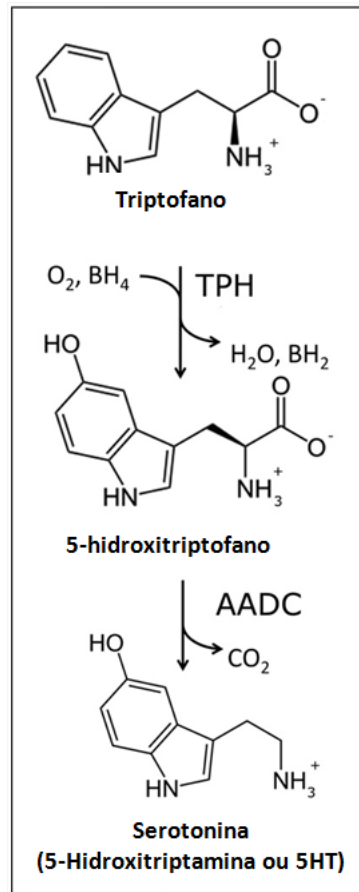


Figura 2- Biossíntese da Serotonina

(Adaptado de TORRENTE; GELENBERG; VRANA, 2012)

A triptofano hidroxilase (TPH) é a enzima limitante da velocidade na via de biossíntese da serotonina e desempenha uma função importante na sua regulação. Em 2003, apenas uma isoforma TPH tinha sido relatada para os vertebrados que se encontrava nos tecidos periféricos. Porém, mais tarde, foi descoberta uma nova forma de TPH específico do cérebro, chamada TPH2 (WALTHER et al., 2003).

Vários estudos demonstraram relação entre a depressão e a diminuição da captação do neurotransmissor serotonina. O Triptofano é um aminoácido aromático essencial, que não é produzido pelo organismo, e é o precursor da 5-hidroxitriptanina (Serotonina) (ROSSI; TIRAPEGUI, 2004).

Esse aminoácido é encontrado principalmente em alimentos de origem animal. Existem, evidências científicas de que a síntese de serotonina cerebral possa ser modulada dieteticamente através da oferta de macronutrientes (ROSSI; TIRAPEGUI, 2005).

O transporte de triptofano para o cérebro não se relaciona apenas com sua concentração sanguínea, mas também com sua concentração em relação à dos aminoácidos

neutros. Assim, concentrações dietéticas baixas de triptofano são associadas ao consumo elevado de aminoácidos que competem pelo transporte através da barreira hematoencefálica, reduzindo o conteúdo de 5-HT no cérebro e alterando certos comportamentos associados à serotonina (RAPPORT, 2007).

### **2.3 POLIMORFISMOS: CONCEITO E ORGANIZAÇÃO EM BANCO DE DADOS**

O conjunto de genes e outros aspectos da sequência do nosso DNA (ácido desoxirribonucleico) são praticamente idênticos entre os seres humanos, sendo que somente 0,5% dos genes dos indivíduos apresentam diferença. As variações na sequência de DNA são conhecidas como polimorfismos genéticos, que podem incluir adição, deleção ou alteração de bases em genes, essas alterações podem ocorrer em uma base ou mais. Ocorrem em sequências codificadoras e não-codificadoras, provocando alterações qualitativas e/ou quantitativas das proteínas. O Polimorfismo de nucleotídeo único (SNP) é uma das alterações mais frequentes na molécula de DNA e muito estudado em relação à causa ou predisposição a doenças (PIERCE, 2009).

Com objetivo de organização dos polimorfismos, criaram-se projetos para obter e organizar frequências alélicas, genótípicas e haplótipos de SNPs em diferentes populações. O projeto 1000 genomas é o primeiro projeto o qual foi sequenciado os genomas de um grande número de pessoas (pelo menos 1.000) para fornecer um recurso abrangente sobre a variação genética humana, como o de localizar a maioria das variantes genéticas que tem frequência de pelo menos 1% nas populações, descobrir genótipos e fornecer a informação exata sobre haplótipos de todos os polimorfismos do DNA humano em diversificadas populações. Ao mesmo tempo, permite comparar frequências e analisar os desequilíbrios de Hardy-Weinberg das populações incluídas no estudo (SIVA, 2008; 1000 Genomes Project Consortiu (2010); ALTSHULER, et al., 2005).

O projeto *HapMap* é um projeto criado a partir do sequenciamento do genoma humano e tem por objetivo guiar e analisar estudos de associação genética. É um banco de dados público onde estão armazenados os SNPs mais comuns do genoma humano. Sendo uma das aplicações mais importante dos dados do *HapMap* é tornar possível estudos do tipo *Genome-wide association* (GWA) (ALTSHULER et al., 2005).

## 2.4 POLIMORFISMOS DA ENZIMA TPH2

Vários genes têm sido investigados em estudos de associação com a depressão. Um gene candidato aos fenômenos comportamentais é o gene da enzima Triptofano hidroxilase 2 (TPH 2) por ser a principal enzima limitante na via da síntese da serotonina (CICHON et al., 2008). Como tal, é um gene candidato ao estudo de uma grande variedade de fenômenos comportamentais, incluindo transtornos do humor, suicídio e traços de personalidade (ZHANG et al., 2011).

Diversos estudos vêm tentando elucidar a associação do polimorfismo da enzima TPH2 e a depressão (SHEN et al., 2011; NAZREE et al., 2013; Illi et al., 2009). Pesquisando o polimorfismo do gene TPH2 na população coreana, Serretti e colaboradores (2011), encontraram alelos significativamente associados com a depressão, eles observaram significância no Alelo G do SNP rs4570625 e no Alelo A do SNP rs10748185. Sugerindo uma associação expressiva entre esses polimorfismos e a depressão.

Também foi realizada uma investigação de variantes genéticas na população brasileira. Os autores investigaram a associação de oito SNPs com a depressão de início tardio em 84 pacientes ambulatoriais com depressão e 79 indivíduos do grupo controle, sugerindo uma associação entre o genótipo CT no SNP rs4565946 e a diminuição do risco de desenvolver depressão tardiamente. Para o SNP rs11179000 sugeriram que o homozigoto AA e alelo A aumentam o risco de depressão tardia (PEREIRA et al., 2011).

Tsai e colaboradores (2009) relataram uma associação entre o SNP rs17110747 e a depressão maior, observando que a associação é significativa entre o Alelo G e a depressão maior na população chinesa. Porém, nem todos os SNPs do gene TPH2 mostraram associação com a depressão.

### **3 JUSTIFICATIVA**

Levando-se em consideração que indivíduos com predisposição genética para qualquer doença podem não apresentar os sintomas, devido a não exposição a fatores ambientais, as populações com altos níveis de prevalência da doença em questão deveriam apresentar altas frequências de alelos ou genótipos envolvidos nesta predisposição e, por este motivo, estudos de correlação entre dados populacionais de polimorfismos genéticos e prevalência da doença a ser estudada poderiam levantar alvos interessantes a serem estudados.

Para testar esta hipótese, utilizamos, como modelo, a relação entre o gene TPH2 e a depressão baseado nas seguintes características: existem vários trabalhos mostrando a relação entre variantes do gene TPH2 e a depressão; existem dados epidemiológicos mundiais para a patologia; o gene TPH2 possui um grande número de SNPs e, sugere-se que alguns polimorfismos específicos contribuiriam para a predisposição à doença, embora a maioria não demonstre esta associação.

### **4 OBJETIVOS**

#### **4.1 Objetivo Geral**

- Desenvolver um método de estudo para obtenção de possíveis marcadores de predisposição à doença a partir de correlação entre os dados epidemiológicos e frequências populacionais de polimorfismos.

#### **4.2 Objetivos Específicos**

- Revisar em bases de dados genômicos os polimorfismos do gene TPH2 conhecidos;
- Revisar na literatura polimorfismos do gene TPH2 associados à depressão;
- Procurar na literatura e bases de dados as prevalências populacionais da depressão;
- Tabular as frequências alélicas e genótípicas dos diferentes polimorfismos do gene THP2 em populações descritas nos bancos de dados 1000 genomas e *HapMap*.
- Correlacionar a prevalência de depressão em diferentes populações mundiais com as frequências alélicas e genótípicas dos diferentes polimorfismos do gene THP2 em populações correlatas;
- Relacionar os dados obtidos com os dados da literatura.

## PARTE II

### 5 MANUSCRITO

**Statistical correlation between genetic frequencies data and prevalence of disease data could complement case-control assays for identify susceptibility loci in Genome-wide association studies.**

Samara Marques Dos Reis, Letícia Madeira Rosa, Andrés Delgado Cañedo

**Submetido a Plos One**



**Statistical correlation between genetic frequencies data and prevalence of disease data could complement case-control assays for identify susceptibility loci in Genome-wide association studies.**

Samara Marques Dos Reis<sup>1¶</sup>, Leticia Madeira Rosa<sup>1</sup>, Andrés Delgado-Cañedo<sup>1¶\*</sup>

<sup>1</sup>Laboratory of Gene Expression Control in Eukaryotes Universidade Federal do Pampa, São Gabriel, Rio Grande do Sul, Brazil

\*Correspondência autor

E-mail: andrescanedo@unipampa.edu.br

¶These authors contributed equally to this work.



## Abstract

It is estimated that the difference between two human beings, when it is compared their genome, do not go beyond 0,5%. This variation extends over the genome and includes, as we know the knowledge about genetic polymorphism. After the finalization of human genome sequencing it was expected understand the role of these polymorphisms in different situations and use it to contribute with the diagnoses and treatments. Two big projects, HapMap and 1000 genomes, organized the major part of these polymorphisms from the study of several populations becoming these data available. Nowadays, these data are used to control the study from Genome-Wide Association, where many polymorphisms are compared in case-control studies. In this work it is shown an alternative to this kind of study where genotypic and frequencies data from several population were analyzed through the Pearson correlation against epidemiological data of a certain kind of pathology, based in the hypothesis that if in a population there is more frequency of a certain kind of pathology genetically determined, the genetic variants involved should be more frequent and vice-versa. The model used was the involvement of polymorphisms of the TPH2 gene in the predisposition to depression. The data obtained with positive correlation in one of homozygous genotypes and in the allele of this homozygous suggested the presence of 10 polymorphisms possibly related to the development of depression. These data were compared to literature data involving case-control studies. In these works were studied 20 from 69 polymorphisms described to the TPH2 gene. With the exception of only one polymorphism, all the data obtained through the strategy proposed in this work have been equals to the literature data, including the allele that is determinant to the predisposition of depression, when it is demonstrated the association. Therefore, it is proposed this strategy as an alternative to carry out Genome-Wide Association studies without the necessity of a case-control study, only using the epidemiological data from the disease, decreasing the time and the cost of the study.

## Introduction

The detection and analysis of DNA polymorphisms have modified the studies about human genetic. The major impact of polymorphism studies is the possibility of identify genetic variants involved in diseases. Also, the polymorphism can be used to distinguish small differences either inside a population or between different populations [1].

Nowadays, population polymorphisms are organized in genetic databases, that are aimed to be a free database to the scientific community. Between these databases we can find the 1000 genome and HapMap projects that provide information about human genetic variants, including single nucleotide polymorphisms (SNPs) and structural variation; also, these projects offer linkage disequilibrium among the genetic variants [2-3].

Through the use of the tools offered by the project mentioned above and also other new databases, it is possible to obtain allelic, genotypic frequencies and haplotypes of different populations around the world.

There are several diseases with monogenic cause, but most diseases have a complex genetic component. Between them are included diseases of great impact in the world population as the depression, for example.

Depression is one of most common psychiatric disease. According to the World Health Organization [WHO], depression is the forth cause of incapacitation around the world, is one of the main contributor for the global burden of disease and affects more than 350 million people of different ages, being more frequent in women than in men [4].

The major depression has a multifaceted etiology with a familial component and it seems to occur as the interaction between genetic and environmental influences that would not act alone, but jointly [5].

Among the proteins related with the depressive process, the serotonin (5-HT or 5-hydroxytryptamine) is one of the most studied. It is a neurotransmitter that modulates the neural activity regulating a wide spectrum of neuropsychiatric processes [6]. It has been associated to the depression main symptoms such as alteration of the humor, appetite, sleep and in cognitive dysfunctions [7].

Serotonin is synthesized from tryptophan (an essential amino acid) through the tryptophan hydroxylase (TPH), this enzyme converts tryptophan into 5-hydroxytryptophan, that is converted to serotonin thereafter. The enzyme TPH2 is localized and stores in vesicles of the axons and dendrites [8].

TPH2 is a limiting in the serotonin synthesis pathway, controlling the synthesis rate in the brain and is considered a key factor for the maintenance of the normal level of serotonin in central nervous system. Experimental studies have suggested association between some SNPs of THP2 and depressive disorders [9-10], although the majority of the studied SNPs did not show this association [11, 12, 13].

Taking in account that: there are vast epidemiologic data about depression for different population around the world; TPH2 enzyme is related to depression; many TPH2 SNPs were studied regarding their association with de depressive process, reporting or not association; and the genotypic and allelic frequencies of TPH2 SNPs are available for several populations.

In the present work we analyzed all the known SNPs of TPH2 against epidemiologic depression data based in the follow hypothesis: “If a population has a high prevalence of a particular condition, genetically determined, then the genetic variation involved with this condition must also be high.” Ours data suggest that Pearson’s correlation test using genotypic frequencies and epidemiologic can be effective for obtain similar data to those determined experimentally.

## **Materials and Methods**

### **Search of TPH2 polymorphisms and their population frequencies**

All the SNPs for TPH2 gene was searched in the National Center for Biotechnology Information (NCBI) “gene” database (<http://www.ncbi.nlm.nih.gov/gene/121278>). Subsequently, each of the polymorphisms was used for population data research in Databases from the 1000genomes project [2] and the International HapMap Project [14].

Sample size and allelic and genotypic frequencies of all the available populations were tabulated in Excel datasheets.

### **Data sources, searches and eligibility criteria**

We performed an electronic search in four databases (PubMed, Medline, Google Scholar and SNPedia) from the inception of each database to July 2014. We searched for the following keywords: “TPH2”, “polymorphism” and “depression” at same time. A second search was done using each SNP name in combination with the keyword “depression”.

The articles retrieved were cured in an independent manner by two researchers. It was considered articles that presented data from case-control studies, relating TPH2 gene polymorphisms and depression were considered eligible for this study. Studies without genotypic frequencies were excluded.

For each article selected we recorded the first author, publication year, ethnic group, population size and allelic and genotypic frequencies.

### **Statistical analysis**

Both allelic and genotype frequencies of each population studied by HapMap and 1000 genomes projects were subjected to Pearson's correlation test against the population rates of

depression inferred for these populations, obtained from previously published articles [15-16]. Values  $<0.4$  were considered as without correlation and values  $\geq 0.4$  were considered as with correlation, according to suggestions of Taylor [17] and Dancey [18].

## Results

### Review of TPH2 Polymorphisms

The search for polymorphisms in the gene which codify the TPH2 enzyme has found 69 results, all of them were SNPs (March 2014). 67 SNPs (97%) were described at 1000 genome database and 57 SNPs (83%) at HapMap project database S1 Table.

Regarding the location, 7 of 69 SNPs mapped in the potential promoter region (-1000 to -1), 52 SNPs mapped in intronic regions, 9 SNPs mapped in exons and 1 SNP mapped upstream to the transcribed region. 20 of 69 SNPs (30%) have already been studied in depressive patients S2 Table.

### Systematic review

The search in databases found 228 articles that contained the keywords TPH2 polymorphism and depression in the text. 221 of 228 articles were thereafter excluded because did not studied depression but others related syndromes such as panic syndrome, schizophrenia, bipolar and unipolar disorder, among others; or because they studied TPH2 polymorphism in depressive patients but they did not present genotypic frequencies or case-control study.

Finally, 7 articles were considered usable for our work. They were realized in 7 populations (Brazilian, Malaysian, Chinese, Korean, European, Asiatic and Finnish). The study workflow is showed in Fig. 1.

**Fig.1. Flow chart of the strategy used to analyze the data.**

From the 20 SNPs already studied in depressive patients only 4 SNPs (rs4570625, rs10748185, rs11179000 and rs17110747) presented statistically significant results. The SNPs rs1386494, rs7305115 and rs17110747 were studied in two publications and rs17110747 was studied in one publication. In the case of rs17110747 one article did not present significance between case and control frequencies.

Data of all polymorphism studied are presented in Table 1.

**Table 1. Published case-control studies of the association between TPH2 polymorphism and depression.**

SNP	Population	Sample Case/Control	Resultados dos artigos	Authors	resultados da Correlação de Pearson
rs4131347	European/Asian	324/130	Not significant	Mann et al., 2007	Significant Allele G
rs4448731	Brazilian	84/79	Not Significant	Pereira et al., 2011	Significant Allele C
rs4570625	Korean	145/170	Significant Allele G	Serretti et al., 2011	Significant Allele G
rs10748185	Korean	145/170	Significant Allele A	Serretti et al., 2011	Not Significant
rs4565946	Brazilian	84/79	Not Significant	Pereira et al., 2011	Significant Allele T
rs11179000	Brazilian	84/79	Significant Allele A	Pereira et al., 2011	Significant Allele A
rs7955501	Brazilian	84/79	Not Significant	Pereira et al., 2011	Not Significant
rs1386495	Malaysian	265/332	Not Significant	Nazree et al., 2013	Not Significant
rs1386494	Malaysian (Malay, Chinese and Indian)	265/332	Not Significant	Nazree et al., 2013	Not Significant
	Finnish	86/395	Not Significant	Lli A. et al., 2009	Not Significant
rs2171363	Chinese	508/463	Not Significant	Tsai et al., 2009	Not Significant
rs1386492	Chinese	508/463	Not Significant	Tsai et al., 2009	Not Significant
rs7305115	Malaysian	265/332	Not Significant	Nazree et al., 2013	Not Significant
	Chinese	368/371	Not Significant	Shen et al., 2011	Not Significant
rs11179027	Korean	145/170	Not Significant	Serretti et al., 2011	Not Significant
rs10506645	Brazilian	84/79	Not Significant	Pereira et al., 2011	Significante Allele C
rs4760820	Brazilian	84/79	Not Significant	Pereira et al., 2011	Significante Allele G

rs1386498	Korean	145/170	Not significant	Serretti et al., 2011	Not Significant
rs1487275	Brazilian	84/79	Not Significant	Pereira et al., 2011	Not Significant
rs4469933	Korean	145/170	Not significant	Serretti et al., 2011	Significante Allele T
rs4290270	Chinese	368/371	Not Significant	Shen et al., 2011	Not Significant
rs17110747	Korean	145/170	Not Significant	Serretti et al., 2011	Significante Allele G
	Chinese	508/463	Significante Allele G	Tsai et al., 2009	



## **Correlation between TPH2 polymorphisms and world rate of depression**

To test our hypothesis we used the Pearson's correlation, using allelic and genotypic frequencies of TPH2 SNPs, obtained from HapMap and 1000 genome project, and the depression rate previously published for the region of each population studied by HapMap and 1000 genomes project. The tests suggested correlation ( $r > |0.4|$ ), either positive or negative, in 32 of 66 SNP analyzed, regardless of whether the correlation occurred in allelic frequencies, genotypic frequencies or both. Considering those data with positive correlation, 25 SNPs presented positive correlation only in allelic frequencies, 01 SNPs presented positive correlation only in genotypic frequencies and 10 SNPs presented positive correlation both in allelic and in homozygote genotypic frequencies (Table 2).

**Table 2 – Correlation between genetic frequencies and depression rate of different SNPs of the TPH2 gene.**

SNP	Allele	(r) Allele 1	Allele	(r) Allele2	Genotype	(r) Homo 1	Genotype	(r) Hetero2	Genotype	(r) Homo 2
<b>rs4131347</b>	C	-0.54	G	0.54	CC	-0.63	CG	-0.21	GG	0.47
rs4131348	C	-0.25	T	-0.25	CC	-0.53	CT	-0.16	TT	0.23
rs4448731	C	0.46	T	-0.46	CC	0.25	CT	0.27	TT	-0.37
<b>rs6582071</b>	A	-0.69	G	0.69	AA	-0.64	AG	-0.50	GG	0.63
rs7963803	A	-0.29	C	0.29	AA	0.33	AC	-0.36	CC	0.33
<b>rs4570625</b>	G	0.53	T	-0.53	GG	0.51	GT	-0.43	TT	-0.45
<b>rs11178997</b>	A	-0.53	T	0.53	AA	-0.26	AT	-0.54	TT	0.53
rs11178998	A	0.33	G	-0.33	AA	0.32	AG	-0.32	GG	0.13
<b>rs4641527</b>	G	0.61	T	-0.61	GG	0.47	GT	-0.19	TT	-0.62
rs4341581	G	-0.06	T	0.06	GG	-0.78	GT	-0.13	TT	0.08
rs34115267	C	0.27	G	-0.27	CC	0.11	CG	-0.11	GG	0.00
rs7954758	A	-0.22	G	0.22	AA	-0.13	AG	0.05	GG	0.26
rs10748185	A	0.33	G	-0.33	AA	0.31	AG	0.06	GG	-0.26
rs10784941	A	-0.39	G	0.39	AA	-0.27	AG	0.22	GG	0.18
rs4565946	C	-0.41	T	0.41	CC	-0.30	CT	0.23	TT	0.23
<b>rs11179000</b>	A	0.63	T	-0.63	AA	0.54	AT	-0.38	TT	-0.58
rs11179001	A	0.01	G	-0.01	AA	0.11	AG	0.02	GG	-0.02
rs2129575	G	0.23	T	-0.23	GG	0.22	GT	-0.15	TT	-0.19
rs1386488	A	0.07	C	-0.07	AA	0.00	AC	0.10	CC	-0.35
rs1843809	G	-0.27	T	0.27	GG	-0.37	GT	-0.14	TT	0.21
rs7955501	A	-0.26	G	0.26	AA	-0.23	AG	-0.03	GG	0.11
rs10879346	C	-0.46	T	0.46	CC	-0.42	CT	0.03	TT	0.32
rs1386495	A	0.15	G	-0.15	AA	0.17	AG	-0.08	GG	-0.31
rs1386494	C	0.06	T	-0.06	CC	0.02	CT	0.00	TT	-0.19
rs6582072	A	-0.11	G	0.11	AA	-0.32	AG	-0.03	GG	0.07
rs1386493	A	-0.15	G	0.15	AA	-0.15	AG	-0.12	GG	0.13
rs2171363	C	-0.29	T	0.29	CC	-0.37	CT	0.41	TT	0.13
rs1386492	C	-0.12	T	0.12	CC	-0.11	CT	-0.19	TT	0.05
rs1386491	C	0.07	G	-0.07	CC	-0.37	CG	0.18	GG	-0.11
rs7963720	C	-0.42	T	0.42	CC	-0.51	CT	0.28	TT	0.32
rs17110563	C	-0.05	T	0.05	CC	0.22	CT	0.13	TT	0.13
rs17110566	A	-0.26	G	0.26	AA	-0.46	AG	-0.10	GG	0.15
rs4760815	A	0.15	T	-0.15	AA	0.16	AT	-0.04	TT	-0.14
rs4760816	C	-0.06	T	0.06	CC	-0.11	CT	0.11	TT	0.06
rs7305115	A	-0.05	G	0.05	AA	-0.07	AG	0.03	GG	0.03
rs6582078	G	-0.06	T	0.06	GG	0.01	GT	0.02	TT	0.03
rs11179027	C	-0.06	G	0.06	CC	-0.05	CG	-0.05	GG	0.05
rs4760750	A	-0.15	C	0.15	AA	-0.12	AC	-0.12	CC	0.15
rs1023990	A	0.56	G	-0.56	AA	0.39	AG	-0.14	GG	-0.51

rs10506645	C	0.46	T	-0.46	CC	0.27	CT	0.02	TT	-0.48
rs4641528	C	-0.50	T	0.50	CC	-0.51	CT	0.14	TT	0.36
rs1386497	A	-0.05	C	0.05	AA	-0.05	AC	0.15	CC	-0.20
rs12229394	A	-0.32	G	0.32	AA	-0.38	AG	0.00	GG	0.21
rs4760820	C	-0.45	G	0.45	CC	-0.42	CG	0.45	GG	0.09
rs1352250	A	-0.57	G	0.57	AA	-0.60	AG	0.15	GG	0.38
rs1386498	A	-0.39	G	0.39	AA	-0.48	AG	0.25	GG	0.18
rs1487278	C	0.00	T	0.00	CC	-0.33	CT	0.14	TT	0.00
rs11179039	C	0.43	T	-0.43	CC	0.29	CT	0.04	TT	-0.45
rs1473473	A	-0.27	T	0.27	AA	-0.30	AT	-0.14	TT	0.26
rs1487276	C	0.22	T	-0.22	CC	0.17	CT	-0.08	TT	-0.45
rs10879352	C	0.46	T	-0.46	CC	0.32	CT	0.03	TT	-0.50
rs9325202	A	-0.40	G	0.40	AA	-0.49	AG	0.26	GG	0.26
rs17110690	A	-0.02	G	0.02	AA	-0.10	AG	0.08	GG	0.00
rs10879354	C	-0.46	T	0.46	CC	-0.48	CT	0.24	TT	0.30
rs11179050	A	0.43	G	-0.43	AA	0.32	AG	0.01	GG	-0.45
rs1487275	A	0.21	C	-0.21	AA	0.25	AC	-0.08	CC	0.06
rs1386486	A	-0.31	G	0.31	AA	-0.17	AG	0.01	GG	0.34
rs1386483	C	0.38	T	-0.38	CC	0.35	CT	0.11	TT	-0.28
<b>rs1386482</b>	G	0.55	T	-0.55	GG	0.46	GT	0.05	TT	-0.46
rs10879357	A	-0.09	G	0.09	AA	0.01	AG	-0.14	GG	0.09
rs4469933	C	-0.46	T	0.46	CC	-0.34	CT	-0.17	TT	0.36
<b>rs11615016</b>	A	-0.42	G	0.42	AA	-0.38	AG	0.36	GG	0.59
rs4290270	A	-0.35	T	0.35	AA	-0.36	AT	-0.13	TT	0.33
<b>rs1487280</b>	C	-0.41	T	0.41	CC	-0.41	CT	-0.11	TT	0.41
<b>rs17110747</b>	A	-0.46	G	0.46	AA	-0.36	AG	0.48	GG	0.51
rs1872824	A	-0.39	G	0.39	AA	-0.40	AG	0.27	GG	0.32

(r)- Pearson product-moment *correlation coefficient*.

Bold SNPs correspond to those that presented positive correlation both in allelic and in genotypic frequencies.

## **Comparison of the correlation results and data previously published**

20 of 66 TPH2 SNPs included in our study were previously analyzed by others author using case-control studies. When we considered significant only those SNPs that presented correlation positive both in allelic and in genotypic frequencies (genotypes and alleles that predispose people to some disease), our data strongly agree with the data previously published. We had only one exception with rs4131347, located in the promoter region. For this SNP our data suggested positive correlation in the GG homozygote and G allele, while Man and co-workers [11] did not find correlation in a case-control study (data compiled in Table 2).

## **Discussion**

The development of the project of Human Genome allowed innovations in medicine leading to an easily improving in the determination of diseases, an specific diagnosis and more effective treatments [19]. Thus, the genetic variation of individuals should be considered in a very near future, thereby facilitating accurate early diagnosis and treatment of diseases.

The need for organization of the genetic variability data for future studies, led to the creation of databases such as the 1000 genomes and the HapMap project that aim to organize a catalog of human genetic variation to characterize diverse populations, data are freely available to the public through the databases [19-20]. These initiatives provide allele and genotype frequencies and the haplotype for several SNPs in different world populations.

Our hypothesis for this study is that: if a population has a high prevalence of a given disease, genetically determined, then the genetic component positively correlated should also be present in high frequency, and vice versa.

Thus, correlating epidemiological data of a conjunct of different populations with their allelic and/or genotypic frequencies could reveal potential locus to be studied for obtaining genetic factors related to the disease in question. This strategy could enable the execution of Genome-wide association tests without the prior need for case-control studies and the use of data from non-selected sample of individuals, reducing the cost and work to obtain and validate the data.

In this context, we analyzed epidemiological data of depression in different populations correlating them by the Pearson test with TPH2 SNPs data available in the databases, since this gene variants had been related to depression, this gene has a high number of SNPs and epidemiological data of depression around the world were available. At the same time, we compare the data obtained by our method with experimental data from literature involving case-control studies.

From the correlation between prevalence data for worldwide depression and SNPs data for the TPH2 gene in different populations, considering SNPs that presented positive correlated data in some of the homozygotes and one of the alleles, our results suggest the presence of 10 TPH2 SNPs (rs4131347, rs6582071, rs4570625, rs11178997, rs4641527, rs11179000, rs4760820, rs1386482, rs11615016 and rs17110747) correlated with depression, four of these mapping in the potential promoter region.

Among the 10 SNPs selected risk in our study, four have already been studied in the literature, showing significant relationship with depression, three SNPs showed correlation with the same alleles determined by our method (rs4570625, rs11179000, rs17110747) [13-22-23]. At the same time, the results obtained by studying SNPs rs7955501 [22], rs1386495 [12], rs1386494 [12], rs1386492 [13], rs7305115 [12], rs11179027 [23], rs1487275 [22], rs4290270 [24] did not show significant difference either in allele and genotype, in agreement

with our results that did not show correlation. Thus, the high trend between our data and the literature gives strength to our hypothesis.

There was only a polymorphism (rs10748185), where the A allele was significant in a survey in Korean [23], disagreeing with our results, which had no correlation between this SNP and depression. But there's another population study data agree with ours, where rs17110747 was no correlation in allele G. In addition, a meta-analysis of work of the TPH2 gene polymorphism also suggest the ratio of SNPs rs4565946 and rs17110747 and with major depression these results agree with the data obtained in this study [24].

Also, we highlight that there are some TPH2 polymorphism that show genotype positive correlation regarding to the depression that have not been studied and should be validated for adding data to our results.

## Conclusions

In conclusion, the data from this study suggest that it is possible to obtain genetic markers associated with specific diseases if data from allelic and genotypic frequencies are statistically analyzed by Pearson correlation test against epidemiological data of these diseases in the populations studied. This type of procedure could be applied to other diseases or other genes potentially involved in the development of the depression.

Furthermore, it would be interesting that databases that gather information on gene variants in different populations provide a tool to insert epidemiological data for potential targets for future studies, approaching each day over the promises of human genome sequencing. At the same time, to give more strength to the analysis, new populations should be included in the databases.

## References

1. Barreiro LB. Laval G. Quach H. Patin E. Quintana-Murc L. (2008) Natural selection has driven population differentiation in modern humans. *Nature Genetics* 40: 340-345.
2. McVean GA. Altshuler DM. Durbin RM. Abecasis GR. Bentley DT. et al. (2012) An integrated map of genetic variation from 1.092 human genomes. *Nature* 491: 56-65.
3. Daniel F. Carr. Ana Alfirevic and Munir Pirmohamed (2014) Pharmacogenomics: Current State-of-the-Art. *Genes* 5: 430-443.
4. World Health Organization. Depression. Fact sheet N°369 October 2012.
5. Sullivan PF. Neale MC. Kendler KS (2000) Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry* 157: 1552–1562.
6. Berger M. Gray JA. Roth BL. The expanded biology of serotonin (2009). *Annu Rev Med* 60:355–66.
7. Meltzer H. (1998) Serotonergic dysfunction in depression. *Br J Psychiatry Suppl.* 8: 25–31.
8. Ravi V. Jessica J. Tommy L. Nikolaos V. (2011) Tryptophan Transport in Human Fibroblast Cells—A Functional Characterization. *International Journal of Tryptophan Research.* 4: 19–27.
9. Gao J. Pan Z. Jiao F. Li F. Zhao G. et al. Evangelou (2012) - TPH2 Gene Polymorphisms and Major Depression – A Meta-Analysis. *Plos One* 7: 1-5.
10. Goenjian AK. Bailey JN. Walling DP. Steinberg AM. Schmidt D. et al. (2012) Association of TPH1, TPH2, and 5HTTLPR with PTSD and depressive symptoms. *J Affect Disord.* 140 (3): 244-52.

11. Mann JJ. Currier D. Murphy L. Huang Y. Galfalvy H. et al. (2008) No association between a TPH2 promoter polymorphism and mood disorders or monoamine turnover. *J Affect Disord.* 106(1-2): 117–121.
12. Nazree NE. Loke AC. Zainal NZ. Mohamed Z. (2013) Lack of association between TPH2 gene polymorphisms with major depressive disorder in multiethnic Malaysian population. *Asia Pac Psychiatry.*
13. Tsai SJ. Hong CJ. Liou YJ. Yu YW. Chen TJ. et al. (2009) Tryptophan hydroxylase 2 gene is associated with major depression and antidepressant treatment response. *Progress in Neuropsychopharmacol Biol Psychiatry.* 33(4): 637-41.
14. The International HapMap Project. *Nature.* 2003; 426 (6968): 789-96.
15. Bromberg E. Andrade LH. Hwang I. Sampson NA. Alonso J. et al. (2011) Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine.* 9:90
16. Ferrari AJ. Charlson FJ. Norman RE. Patten SB. Freedman G. et al. (2013) Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Med* 10(11): 1-12.
17. Taylor R. (1990) Interpretation of the Correlation Coefficient: A Basic Review. *JDMS* 1:35-39.
18. Dancey C. REIDY J. (2006) *Estatística Sem Matemática para Psicologia: Usando SPSS para Windows.* Porto Alegre. Artmed.
19. Corrêa MV. (2002). O Admirável Projeto Genoma Humano. *PHYSIS: Rev. Saúde Coletiva* 12(2): 277-299.
20. Clarke L. Xiangqun Zheng-Bradley X. Smith R. Kulesha E. Xiao C et al. (2012) The 1000 Genomes Project: Data Management and Community. *Access. Nat Methods* 9(5): 459–462.



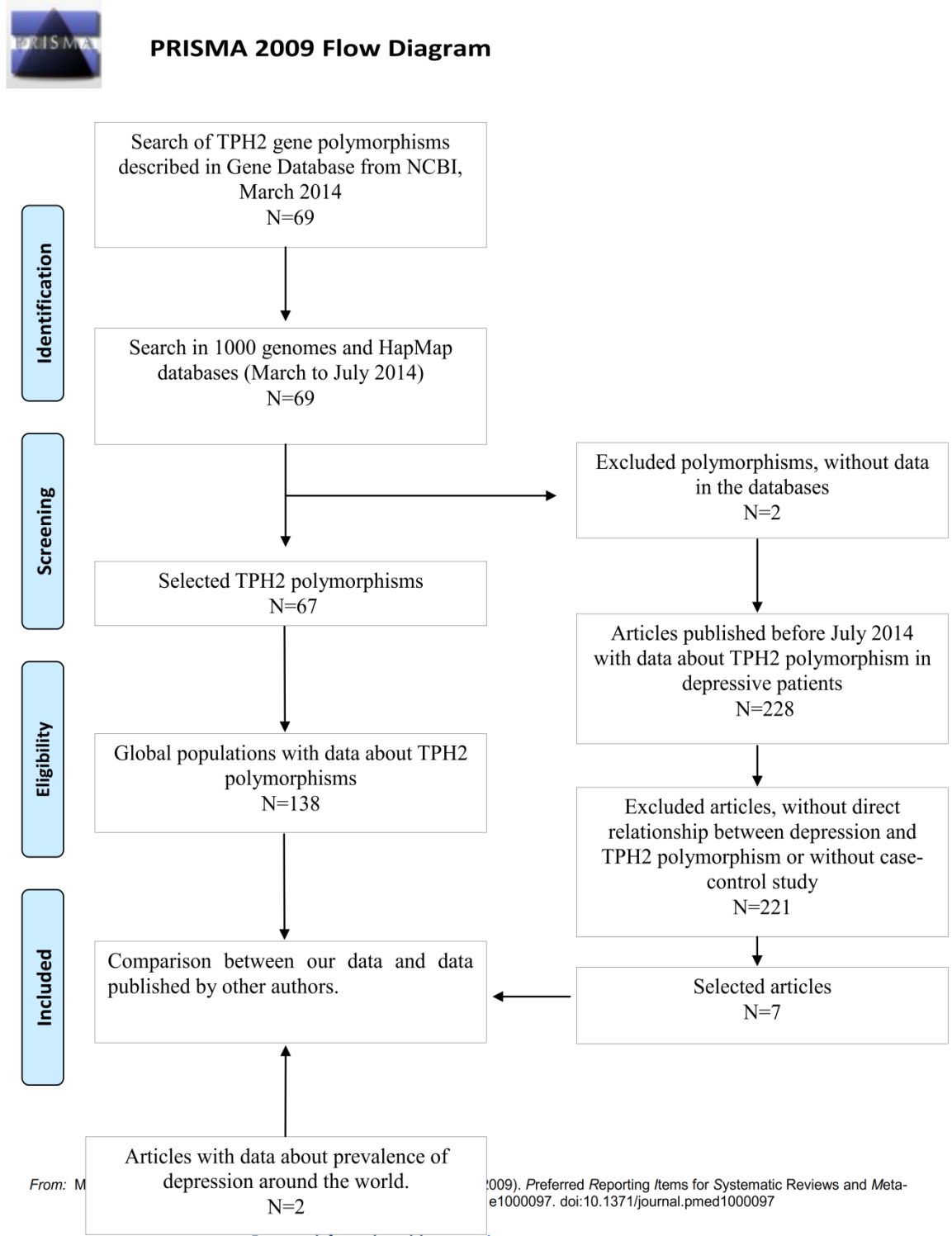
21. D. Beck J. Beiswanger C. Coppock D. et al. (2005) A haplotype map of the human genome. *Nature*. 437(7063): 1299-320.
22. Pereira P de A. Romano-Silva MA. Bicalho MA. De Marco L. Correa H. et al. (2011) Association Between Tryptophan Hydroxylase-2 Gene and Late-Onset Depression. *Am J Geriatr Psychiatry*. 19(9): 825-9.
23. Serretti A. Chiesa A. Porcelli S. Han C. Patkar AA. et al. (2011) Influence of TPH2 variants on diagnosis and response to treatment in patients with major depression, bipolar disorder and schizophrenia. *Psychiatry Research* 189: 26–32.
24. Shen X. Wu Y. Qian M. Wang X. Hou Z. et al. (2011) Tryptophan hydroxylase 2 gene is associated with major depressive disorder in a female Chinese population. *Journal of Affective Disorders* 133: 619–624.
25. Gao J. Pan Z. Jiao Z. Li F. Zhao G. et al. (2012) TPH2 Gene Polymorphisms and Major Depression – A Meta-Analysis. *Plos One* 7: e36721.

## Supporting Information

**S1 Table. Allelic and genotypic frequencies and sample size from all the TPH2 SNPs available in HapMap and 1000 genomes database.** This table present all the data used in this work, including the depression rate assign for each population studied by HapMap or 1000 genome projects.

**S2 Table. TPH2 SNPs location and publications.** This table present the location of each SNP within of the genome and of the TPH2 gene. Also, those SNPs that were already published in case-control studies are highlighted in bold text and the external links for the NCBI abstracts are available in the last column.

**Fig.1. Flow chart of the strategy used to analyze the data.**



### PARTE III

## 6 CONCLUSÃO

Os dados do presente estudo sugerem que seria possível de se obterem marcadores gênicos associados a doenças específicas, se dados de frequências gênicas e genotípicas fossem analisadas estatisticamente pelo método de correlação de Pearson contra dados epidemiológicos destas doenças nas populações estudadas. Este tipo de procedimento poderia ser aplicado em outras doenças ou outros genes potencialmente envolvidos no desenvolvimento de depressão.

Além disso, seria importante que nas bases de dados *HapMap*, 1000 genomas e outros bancos de dados de variantes populacionais existisse uma ferramenta para inserir dados epidemiológicos a serem estudados e obter SNPs alvos para futuros estudos, aproximando-nos, a cada dia mais, da promessa do sequenciamento do genoma humano. Ao mesmo tempo, para dar mais forças à análise, novas populações deveriam ser incluídas nas bases de dados acima descritas.

Também, observa-se que existem alguns polimorfismos do gene TPH2 com correlação genotípica positiva com respeito à depressão que ainda não foram estudados e que deveriam ser validados para acrescentar dados a este estudo.

## **7 PERSPECTIVAS**

Levando em consideração a alta taxa de congruência entre os dados obtidos, neste trabalho, e os dados publicados na literatura, este trabalho é o início para a abertura de um leque de futuros trabalhos.

Assim, daremos continuidade ao estudo dos polimorfismos genéticos do gene do TPH2 estudando, em ensaios do tipo caso-controle, aqueles SNPs que obtiveram uma correlação genotípica positiva com respeito à depressão para poder validar os dados desse estudo; por outro lado, outros genes envolvidos com a depressão serão testados com o mesmo sistema aqui proposto.

## REFERÊNCIAS BIBLIOGRÁFICAS

ALTSHULER, David et al. A haplotype map of the human genome. **Nature**, v.437, n.7063, p.1299-320, out. 2005.

AMIREAULT, Pascal et al. Life without Peripheral Serotonin: Insights from TryptophanHydroxylase 1 Knockout Mice Reveal the Existence of Paracrine/ Autocrine Serotonergic Networks. **ACS Chemical Neurosci**, v. 4, n. 1, p. 64-71, jan. 2013.

BROMETET, Evelyn et al. Cross-national epidemiology of DSM-IV major depressive episode. **BMC Medicine**, v.9, p. 90-106, 2011.

CHICHON, Sven, et al. Brain-specific tryptophan hydroxylase 2 (TPH2): a functional Pro206Ser substitution and variation in the 5-region are associated with bipolar affective disorder. **Human Molecular genetics**, v. 17, p.87-97, 2008.

DUMAN, Ronald S., VOLETI, Bhavya Signaling pathways underlying the pathophysiology and treatment of depression: novel mechanisms for rapid-acting agentes. **Trends in Neurosciences**, v.35, n.1, p. 47-56, jan. 2012.

FERRARI, Alize J. et al. Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. **Plos Medicine**, v.10, n.11, p. e1001547, nov. 2013.

FLECK, Marcelo P. Temas atuais los Depressão. **Revista Brasileira de Psiquiatria**, São Paulo, v. 31, supl. 1, mai. 2009.

GOLDSTEIN, D B., CAVALLERI G. L. Understanding human diversity. **Nature**, v. 437, p.1241-1242, 2005.

GUYTON, Arthur C., HALL John E. **Fisiologia Humana e Mecanismos das doenças**. 6. ed. Rio de Janeiro: Ganabara Koogan, 1998. p. 440.

IBANEZ, Grazielle et al. Adesão e Dificuldades Relacionadas Ao Tratamento medicamentoso em pacientes com Depressão. **Revista Brasileira de Enfermagem**, Brasília, v. 67, n. 4, ago. 2014.

ILLI, Ari et al. 5-HTR1A, 5-HTR2A, 5-HTR6, TPH1 and TPH2 polymorphisms and major depression. **NeuroReport**, v.20, n.12, p.1125-8, ago. 2009.

MARAGNO, Luciana et al. Prevalência de Transtornos mentais Comuns em Populações atendidas pelo Programa Saúde da Família (QUALIS) no Município de São Paulo, Brasil. **Caderno de Saúde Pública**, Rio de Janeiro, v. 22, n. 8, ago. 2006.

MASCELLA, Vivian. Depressão: Causas e tratamento. **Estudos de Psicologia**, Campinas, v.30, n. 2, jun. 2013.

NAZREE, NE, et al. Lack of association between TPH2 gene polymorphisms with major depressive disorder in multiethnic Malaysian population. **Asia Pac Psychiatry**, 2013.

NATIONAL INSTITUTE OF MENTAL HEALTH. **Depression**. 2011. Disponível: [http://www.nimh.nih.gov/health/publications/depression/depression-booklet\\_34625.pdf](http://www.nimh.nih.gov/health/publications/depression/depression-booklet_34625.pdf). Acesso em: 27 set. 2014.

NI, Wei; WATTS Stephanie W. 5-hydroxytryptamine in the cardiovascular system: focus on the serotonin transporter (SERT). **Clinical and Experimental Pharmacology and Physiology**, v. 33, n.7, p. 575-83, jun. 2006.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-10)**. 10<sup>a</sup> rev. São Paulo: Universidade de São Paulo; 1997.

PEREIRA, Patrícia de Araujo et al. Association Between Tryptophan Hydroxylase-2 Gene and Late-Onset Depression. **The American Journal Geriatric Psychiatry**, v. 19, n. 9, p. 825-9, 2011.

PIERCE, Benjamin. **Genética - Um Enfoque Conceitual**. Guanabara Koogan, 2011.

RAPPORT, Maurice M., GREEN, Arda Aldren, PAGE, Irvine H. Serum vasoconstrictor (serotonin). IV. Isolation and characterization. **Journal of Biological Chemistry**, v.176, p. 1243-1251, 1948.

RAVI, Vumma et al. Tryptophan Transport in Human Fibroblast Cells—A Functional Characterization. **International Journal of Tryptophan Research**, v.4, p. 19–27, 2011.

ROSSI L; TIRAPEGUI J. Implicações do sistema serotoninérgico no exercício físico. **Arquivo Brasileiro de Endocrinologia e Metabolismo**. 48. 2004.

ROSSIL, TIRAPEGUI J. Serotonina e a Neuromodulação Alimentar. **Revista Nutrição em Pauta**. 72, 36-40. 2005.

SERRETTI, Alessandro et al. Influence of TPH2 variants on diagnosis and response to treatment in patients with major depression, bipolar disorder and schizophrenia. **Psychiatry Research**, v.189, p.26–32, 2011.

SHEN, Xinhua et al. Tryptophan hydroxylase 2 gene is associated with major depressive disorder in a female Chinese population. **Journal of Affective Disorders**, v.133, p. 619–624, 2011.

SILVA, Marcus T. et al. A prevalência de morbidade depressão entre adultos brasileiros.: Uma revisão sistemática e meta-análise **Revista Brasileira de Psiquiatria**, São Paulo, v. 36, n. 3, set. 2014.

SIVA, Nayanah. 1000 Genomes Project. **Nature Biotechnology**, v.26, n. 3, p. 256, 2008.

SULLIVAN, Patrick F. et al. Genetic epidemiology of major depression: review and meta-analysis. **American Journal of Psychiatry**, v. 157, p. 1552–1562, 2000.

TSAI, Shih-Jen et al. Tryptophan hydroxylase 2 gene is associated with major depression and antidepressant treatment response. **Progress in Neuro-Psychopharmacology & Biological Psychiatry**, v. 33, n. 4, p. 637-41, 2009.

TORRENTE, Mariana P., GELEMBERG, Alan J., VRANA, Kent A. Boosting serotonin in the brain: is it time to revamp the treatment of depression? **Journal of Psychopharmacology**, v.26, n.5, p. 629-635, mai., 2012.

The International *HapMap* Project. **Nature**. v.426, n.6968, p. 789-96, 2003.



## ANEXOS

**S1 Table. Allelic and genotypic frequencies and sample size from all the TPH2 SNPs available in HapMap and 1000 genomes database.** This table present all the data used in this work, including the depression rate assign for each population studied by HapMap or 1000 genome projects.

**S2 Table. TPH2 SNPs location and publications.** This table present the location of each SNP within of the genome and of the TPH2 gene. Also, those SNPs that were already published in case-control studies are highlighted in bold text and the external links for the NCBI abstracts are available in the last column.