RODRIGO SECOLIN

APLICAÇÃO DE MODELOS ESTATÍSTICOS E DESENVOLVIMENTO DE ALGORITMOS PARA ESTUDOS GENÉTICOS DE DOENÇAS NEURO-PSIQUIÁTRICAS

CAMPINAS

Unicamp

2011

RODRIGO SECOLIN

APLICAÇÃO DE MODELOS ESTATÍSTICOS E DESENVOLVIMENTO DE ALGORITMOS PARA ESTUDO GENÉTICO DE DOENÇAS NEURO-PSIQUIÁTRICAS

Tese de Doutorado apresentada à Pós-Graduação da Faculdade de Ciências Médicas da Universidade Estadual de Campinas para obtenção do título de Doutor em Fisiopatologia Médica, área de concentração em Neurociências.

Orientador: Profa. Dra. Iscia Lopes-Cendes

CAMPINAS

Unicamp

2011

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DA FACULDADE DE CIÊNCIAS MÉDICAS DA UNICAMP

Bibliotecária: Rosana Evangelista Poderoso – CRB-8ª / 6652

Se24a

Secolin, Rodrigo

Aplicação de modelos estatísticos e desenvolvimento de algoritmos para estudo genético de doenças neuro-psiquiátricas / Rodrigo Secolin. -- Campinas, SP : [s.n.], 2011.

Orientador : Iscia Lopes-Cendes.

Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Ciências Médicas.

1. Bioinformática. 2. Genética humana. 3. Neurogenética. I. Lopes-Cendes, Iscia. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Título em Inglês: Statistical models application and algorithm development for genetic studies in neuropsychiatric diseases

Keywords:
- Bioinformatics
- Human genetics
- Neurogenetics

Titulação: Doutor em Fisiopatologia Médica
Área de Concentração: Neurociências

Banca examinadora:

Profª. Drª. Iscia Teresinha Lopes Cendes
Profº. Drº. Edi Lucia Sartorato
Profª. Drª. Monica Barbosa de Melo
Profª. Drª. Ana Lucia Brunialti Godard
Profº. Drº. Aguinaldo Luiz Simões

Data da defesa: 08.02.2011

Banca examinadora da tese de Doutorado

Rodrigo Secolin

Orientador(a) : Prof(a). Dr(a). Iscia Teresinha Lopes Cendes

Membros:

1. Prof(a). Dr(a). Iscia Teresinha Lopes Cendes
2. Prof(a). Dr(a). Edi Lucia Sartorato
3. Prof(a). Dr(a). Monica Barbosa de Melo
4. Prof(a). Dr(a). Aguinaldo Luiz Simões
5. Prof(a). Dr(a). Ana Lucia Brunialti Godard

Curso de pós-graduação em Fisiopatologia Médica da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

Data: 08/02/2011

DEDICATÓRIA

Dedico o meu trabalho a um ser fundamental na minha vida, que esteve sempre do meu lado, tanto nos momentos felizes e festivos quanto nos momentos tristes, estressantes e angustiantes. Um ser que fazia qualquer coisa por mim e, em troca, simplesmente queria a minha companhia, um carinho, deitar ao meu lado, dormir sentindo que estava perto de mim. E quando não estávamos perto um do outro durante minhas viagens, uma saudade mútua aparecia, somente sendo amenizada com um telefonema ou um vídeo pelo Skype. *Mas o suficiente para saber que estava tudo bem.*

Por isso, faço desta dedicatória uma homenagem ao meu grande amigo, filho, o qual um de seus nomes foi utilizado em um algoritmo desenvolvido durante meu Doutorado e imortalizado em revista científica internacional. E onde quer que ele esteja agora, está torcendo por mim, somente esperando o momento para estar na minha companhia novamente, pois era somente tudo o que ele sempre quis.

Tutri, você vai estar sempre comigo.

AGRADECIMENTOS

Inicialmente, eu quero agradecer à minha primeira família:

Minha namorada Ellen, minha mãe Edna, ao Vicente, minha querida prima Pamela, tia Cida, tio Valdir... tia Luzia, tio Dirceu, meus primos Milton, Alexandre Feitosa... Obrigado pelo carinho, amor e apoio nos momentos tristes e alegres; vocês sabem o quanto sou feliz de ter todos vocês ao meu lado

Ao pessoal do laboratório, a minha segunda família:

Fábio Augusto e Marilza, as primeiras pessoas que conheci no meu primeiro dia no laboratório há exatos 10 anos; obrigado pela ajuda e pela companhia em todos estes anos.

Luciana, Madá, Milena, Fábio Conte, Renato, Marina, Mariana, Cris, Alexandre, Oni, Anderson, Lidy, Stela, Ilária, Társis, Simoni, obrigado pela companhia na hora do almoço, pelas viagens, pela ajuda nos experimentos, pelos momentos super agradáveis.

Prof. Dra. Neide Ferreira, Marie-Jo, Iarinha, Bel, Maria do Carmo, Romenia, Patty, Vini, Danilo, Danizinha, Simone, Daniel, pessoas maravilhosas que passaram pelo laboratório e ainda mantenho contato; obrigado pelos momentos compartilhados. **Profas. Dras. Carmen, Cláudia, Vera Lúcia,** pela troca de experiência e pelo aprendizado, muito obrigado.

Sônia, Cláudia, Tereza, Jair, Henri, obrigado pela ajuda administrativa e pela companhia no departamento; sem vocês esta tese não teria sido realizada.

Aos funcionários da limpeza, da informática, do espaço *gourmet*, aos seguranças, muito obrigado por manter a Faculdade de Ciências Médicas um lugar ótimo para trabalhar e viver. **Aos médicos e profissionais** do HC da Unicamp, os quais sem a ajuda na parte clínica todos estes trabalhos não teriam sido realizados. **Aos pacientes e familiares** que, apesar do sofrimento que acompanha a enfermidade, se dispuseram em ajudar as nossas pesquisas.

Enfim, obrigado, **Profa. Dra. Iscia,** por você ter acreditado em mim desde o vestibular da UNICAMP em 2001, pelas palavras encorajadoras nos momentos de fracasso e pelas palavras de orgulho nos momentos de sucesso, pelas oportunidades a mim oferecidas, por me ensinar não só a fazer pesquisa científica, mas sim a ser um **cientista**, um sonho de infância que agora se torna realidade.

A todos vocês ... o meu sincero obrigado!

EPÍGRAFE

"A curiosidade é mais importante do que o conhecimento."

Albert Einstein

PREFÁCIO

Esta Tese foi elaborada no formato alternativo, conforme prescrito no art. 2º da informação CCPG/01/2008, com a inclusão de quatro artigos científicos publicados e um artigo científico em preparação para publicação. Estes artigos são resultado de mais de quatro anos de trabalho, em colaboração com vários outros profissionais. A partir das análises realizadas, pude colaborar de maneira significativa na resolução de importantes problemas da ordem de desenho experimental e processamento de dados, que culminaram com a elaboração da Tese.

Apesar do escopo desta Tese referir-se ao desenvolvimento de algoritmos e análises estatísticas utilizadas em técnicas de mapeamento genético, minha contribuição no trabalho descrito nestes cinco artigos também inclui: a aplicação dos termos de consentimento livre esclarecido aos pacientes avaliados no artigo 5, a organização dos dados fenotípicos nos artigos 1, 2, 3, 4 e 5; a extração de DNA e RNA de sangue periférico nos artigos 1, 3, 4 e 5; a obtenção dos genótipos de marcadores microssatélites por meio de técnicas de *Polymerase Chain Reaction* (PCR) no artigo 3; a genotipagem dos *Single Nucleotide Polymorfisms* (SNPs) por meio de PCR em tempo real nos artigos 4 e 5; e ensaios de expressão gênica no artigo 5.

Entretanto, a avaliação clínica e os respectivos exames dos indivíduos utilizados nestes artigos foram realizados por um médico especializado para cada patologia, garantindo ao máximo a homogeneidade fenotípica das amostras. Por fim, apesar de uma Tese ser tradicionalmente julgada tendo como base a demonstração de erudição pelo candidato e a elaboração de uma hipótese que deverá ser refutada, aceita ou ainda não resolvida, esta Tese teve como objetivo auxiliar na resolução de várias hipóteses, sendo cada uma delas discutida nos diferentes artigos incluídos, que são parte integral da Tese.

LISTA DE ABREVIATURAS

AFF2	*Fragile X mental retardation family, member 2*
AIC	*Akaike information criteria*
ANK3	*Ankyrin 3, node of Ranvier*
BPAD	*Bipolar affective disorder*
BPP	*Bilateral perisylvian polymicrogyria*
ELTM	Epilepsia do lobo temporal mesial
GNL3	*Guanine nucleotide binding protein-like 3 (nucleolar)*
INPI	Instituto Nacional de Propriedade Intelectual
ITGA9	*Integrin, alpha* 9
ITIH3	*Inter-alpha (globulin) inhibitor H3*
LRT	*Likelihood ratio test*
MLE	*Maximum likelihood estimation*
MTLE	*Mesial temporal lobe epilepsy*
NEK4	*Never in mitosis gene a-related kinase 4*
OMIM	*Online mendelian inheritance in man*
PCR	*Polymerase chain reaction*
PBRM1	*Polybromo 1*
PPBC	Polimicrogiria perisylviana bilateral congênita
SLITRK2	*Slit and Neurotrophic tyrosine kinase receptor-like family, member 4*
SLITRK4	*Slit and Neurotrophic tyrosine kinase receptor-like family, member 2*
SNPs	*Single nucleotide polymorphisms*
SRPX2	*Sushi-repeat-containing protein, X-linked 2*
TAB	Transtorno afetivo bipolar
TDT	*Transmission disequilibrium test*
TDTAE	*Transmission disequilibrium test allowing errors*
μ	média
σ^2	variância

LISTA DE TABELAS

LISTA DE FIGURAS

SUMÁRIO

RESUMO

Fatores genéticos têm sido descritos para diversas doenças do sistema nervoso central. Uma etapa importante na identificação de genes responsáveis por estas doenças são os estudos de mapeamento genético. Além disso, devido às novas tecnologias de aquisição de dados de genótipos dos indivíduos, é necessário o estudo e desenvolvimento de programas de processamento de grande quantidade de dados para as análises estatísticas. Os objetivos deste trabalho foram: 1) criar uma interface entre os equipamentos de aquisição de dados de genótipos e os programas estatísticos, por meio de programas de processamento de dados; 2) aplicar e avaliar os modelos estatísticos em amostras de famílias segregando três doenças neuro-psiquiátricas: epilepsia do lobo temporal mesial (ELTM), polimicrogiria perisylviana bilateral congênita (PPBC) e transtorno afetivo bipolar (TAB). A interface foi desenvolvida a partir de um algoritmo lógico, o qual adiciona a matriz dos dados dos genótipos provenientes dos equipamentos em uma matriz representativa dos dados das famílias. Este algoritmo, denominado JINGLEFIX, foi programado em linguagem de computador PERL e ambiente R e utilizado posteriormente nos estudos de mapeamento genético da ELTM, PPBC e TAB. Análise de segregação foi realizada em 148 famílias nucleares com ELTM, com um total de 698 indivíduos, visto que esta síndrome não possui padrão de herança conhecido. Uma família, segregando PPBC com um total de 15 indivíduos e um padrão conhecido de herança ligada ao X dominante, foi submetida à análise paramétrica de ligação por meio do pacote de programas LINKAGE, utilizando 18 marcadores microssátelites na região candidata Xq27-Xq28. Análise não paramétrica de ligação realizada em uma amostra de 74 famílias segregando TAB, totalizando 411 indivíduos, por meio do teste de transmissão de desequilíbrio de ligação (TDT), utilizando 21 *single nucleotide polymosphisms* (SNPs) para 21 regiões candidatas. A análise de segregação

revelou a presença de um gene de maior efeito com um padrão autossômico dominante, além da presença de genes de menor efeito influenciando no fenótipo da ELTM. O posterior mapeamento genômico da ELTM, utilizando os parâmetros definidos na análise de segregação, revelou ligação genética na região 18p11. A análise paramétrica de ligação genética levou ao mapeamento da região Xq27 para a família com PPBC, diferente da região candidata previamente descrita. Esta diferença pode ser explicada pelo tipo de amostra familiar utilizada pelos dois estudos. Em relação ao TAB, a análise não paramétrica identificou a região candidata 3p22. Posterior estudo de refinamento da região 3p21-3p22 utilizando 94 SNPs adicionais e estudo de expressão gênica identificou o gene *ITGA9* como possível gene de susceptibilidade para o TAB. Comparando o poder estatístico entre as análises, foi observado maior poder estatístico na análise paramétrica utilizando uma ou poucas famílias, com um número grande de indivíduos por família; enquanto que o poder estatístico foi maior nas análises não paramétricas utilizando múltiplas famílias de tamanhos moderados e estruturas variadas. Conclui-se que o algoritmo de processamento de dados e a adequada aplicação dos modelos estatísticos são fundamentais para sucesso do mapeamento genético das regiões e dos genes responsáveis pelas doenças neuro-psiquiátricas estudadas.

ABSTRACT

Genetic factors have been described for several central nervous system diseases. A main step for disease gene identification is genetic mapping study. In addition, due new genotype acquire technology, the development of genotype processing data software is required. The objectives of this work were: 1) to generate interface between genotype equipment and statistical software by processing data algorithm; 2) to apply and evaluate statistical models in family sample segregating three neurological diseases: mesial temporal lobe epilepsy (MTLE), bilateral perysilvian polymicrogyria (BPP) and bipolar affective disorder (BPAD). Data interface was developed from a logic algorithm, which adds a genotype matrix data from equipment to a family data matrix. This algorithm, called JINGLEFIX, was implemented in PERL computer language and R environment. In addition, this software was used in genetic mapping study for MTLE, BPP and BPAD. Segregation analysis was performed in 148 nuclear MTLE pedigrees, with a total of 698 individuals, since this syndrome has not known inheritance pattern. One BPP pedigree with known X-linked dominant pattern of inheritance, with a total of 15 individuals, was submitted to parametric linkage analysis by LINKAGE package, using 18 microsatellite markers on candidate region Xq27-Xq28. Non-parametric linkage analysis was performed from 74 BPAD families, with a total of 411 individuals, by transmission disequilibrium test (TDT) and using 21 single nucleotide polymorphisms (SNPs) for 21 candidate regions. Segregation analysis revealed a major effect gene with an autosomal dominant pattern of inheritance and minor gene effect, which could influence MTLE phenotype. Further whole genome analysis mapped the putative MTLE major gene on 18p11. Parametric linkage analysis mapped Xq27 locus for BPP, a different region compared to the Xq28 previous described. This difference could be explained to sample type used by the two studies. Non-parametric linkage for BPAD identified the candidate

region on 3p22. Further studies using 94 additional SNPs on 3p21-3p22 and gene expression analysis identified *ITGA9* as susceptibility gene for BPAD. A comparison of statistical power between statistical analyses showed a high statistical power for parametric linkage analysis from one or a few large families; whereas a high statistical power was observed for non-parametric linkage analysis using several moderate size families. The conclusion of this study is that data processing algorithm and adequate statistical model applying are fundamental tools for successful of genetic mapping of complex diseases.

1. INTRODUÇÃO

Os estudos de padrões de herança e mapeamento genético são uma etapa importante na identificação de genes responsáveis por diversas doenças (1). Neste contexto, o desenvolvimento de modelos estatísticos e programas computacionais tem auxiliado no processamento dos dados e na estimativa de parâmetros adequados para estas análises (2), principalmente para doenças complexas, incluindo doenças do sistema nervoso central.

1.1. Epilepsia do lobo temporal mesial

As epilepsias formam um grupo de doenças neurológicas crônicas, caracterizadas pela ocorrência de crises epilépticas recorrentes. Essas crises resultam de uma descarga neuronal anormal, com início em uma região generalizada ou restrita do cérebro (3). Entre as epilepsias, a epilepsia de lobo temporal mesial (ELTM), representando 40% dos casos, tem as crises geradas predominantemente pelo acometimento de estruturas mediais do lobo temporal, incluindo regiões hipocampais (4; 5). A presença de recorrência familiar (6) indica a evidência de fatores genéticos influenciando a ELTM, o que torna necessário análises que comprovem estas observações, além de definir o padrão de herança Mendeliano presente na ELTM.

1.2. Polimicrogiria perisylviana bilateral congênita

A polimicrogiria é uma malformação do desenvolvimento cortical, caracterizado por um grande número de pequenos giros cerebrais, separados por sulcos mais rasos. A polimicrogira perisylvina bilateral congênita (PPBC; OMIM #300388) é a forma mais comum de polimicrogiria, caracterizada pelo espessamento do córtex cerebral, ao redor da fissura de *Sylvius*. As características clínicas incluem paresia pseudobulbar, disartria e atraso ou dificuldade na fala (7; 8). Episódios de recorrência familiar têm sido descritos para a PPBC, sendo que tanto o padrão de herança autossômico recessivo quanto o padrão ligado ao X dominante tem sido observado (7-11). Estes achados indicam a necessidade de estudos genéticos para a identificação do gene responsável para PPBC.

1.3. Transtorno afetivo bipolar

O Transtorno Afetivo Bipolar (TAB; OMIM %611630) é uma doença psiquiátrica, clinicamente caracterizada por episódios de mania (TAB-I) ou hipomania (TAB-II) espaçados por períodos de depressão. A prevalência é de 0,8% a 2,6% na população. Em pacientes não tratados, a taxa de suicídio pode alcançar aproximadamente 20% (12). O TAB também tem sido associado a outras desordens psiquiátricas e outros fatores, tais como depressão, esquizofrenia, alcoolismo, abuso de drogas e suicídio. Apesar de que estudos familiares demonstram a contribuição de fatores genéticos para o TAB, vários trabalhos apresentam resultados de mapeamento genético conflituosos, mostrando a complexidade desta patologia do ponto de vista genético (12; 13).

1.4. Tipo de amostra

Quando há evidência de fatores genéticos relacionados a uma determinada doença, o primeiro passo é avaliar qual o tipo de amostra mais adequada para as análises. Neste caso, o tipo de amostra pode ser **familiar**, formada por indivíduos relacionados, ou **por indivíduos não relacionados**, formada por grupos de casos e de controles. Cada amostra apresenta suas características. Entretanto, entre os dois extremos há uma gradação de tipos de amostras com ambas as características descritas a seguir (Figura 1).

As doenças tipicamente monogênicas, com padrão de herança Mendeliano definido e com baixa prevalência na população possuem alta agregação familiar. Neste caso, o possível gene responsável pela doença apresenta um forte efeito genético no desenvolvimento da afecção e uma baixa frequência na população em geral (1). Entretanto, para doenças complexas comuns e sem um padrão de herança Mendeliano conhecido, Rish e Merikangas em 1996 (14; 15) propuseram uma teoria denominada *common disease-common variant*. De acordo com esta teoria, os fatores genéticos envolvidos em uma doença complexa incluem genes de susceptibilidade com baixo efeito genético e uma maior frequência gênica na população.

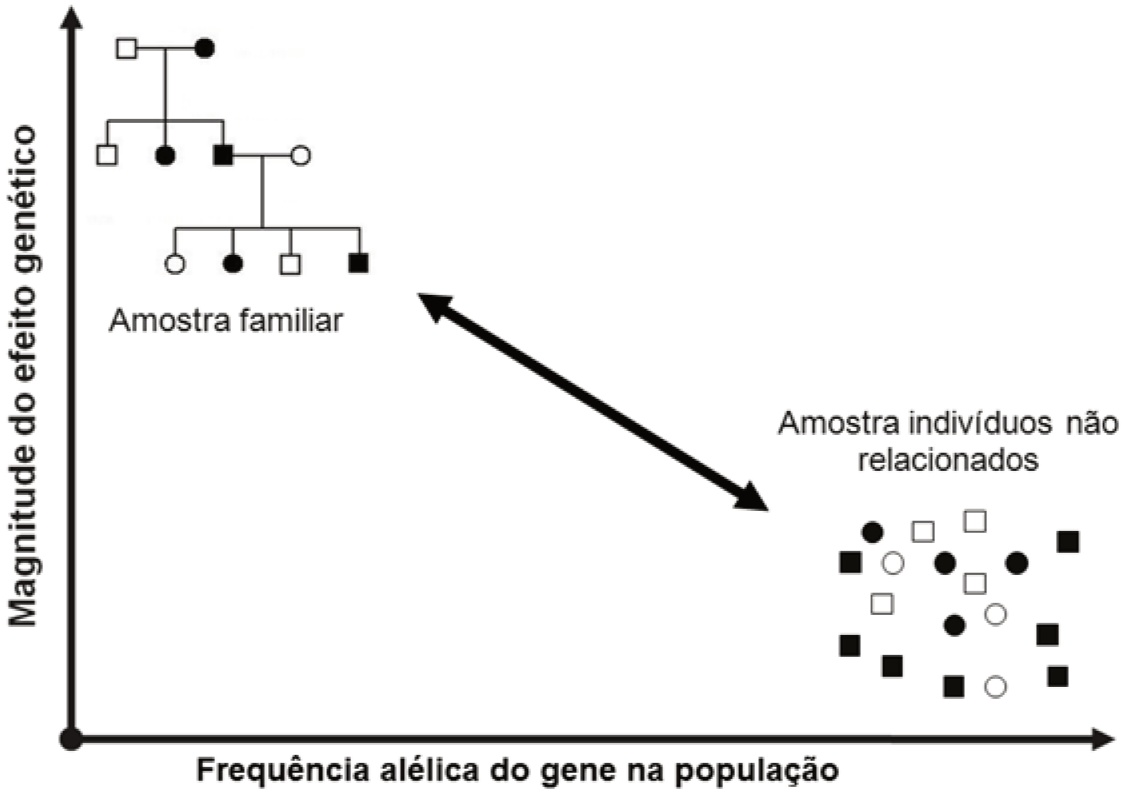


Figura 1. Relação entre frequência alélica do gene na população e seu efeito no fenótipo.

Observam-se diferenças entre os dois tipos de amostra. A amostra familiar apresenta limitações em estudos envolvendo doenças com idade de início avançada, visto a necessidade da obtenção de amostras de material nucléico dos pais ou até de avós. Este problema não é observado em uma amostra de indivíduos não relacionados, pois uma amostra incluindo casos e controles deve ser o mais homogênea possível, tanto em termos de idade quanto de gênero e, principalmente, de estrutura genética entre os grupos (1; 2).

Entretanto, diferenças na estrutura genética dos grupos podem não ser observadas por meio de características fenotípicas (16; 17). Neste caso, uma amostra populacional pode apresentar **estratificação genética**. Esta estratificação ocorre quando o grupo de pacientes ou controles apresenta um aumento na frequência alélica de um gene devido a fatores evolutivos, tais como deriva gênica, e não devido à possível associação deste gene com a doença em estudo, gerando resultados falso-positivos (18; 19). Por outro lado, um desenho

de estudo de associação genética baseada em famílias tem a vantagem de minimizar a estratificação da amostra, visto que esta análise considera associação dentro de cada família de modo independente (20; 21).

Devido ao grau de miscigenação da população brasileira (16; 17; 22), as análises por meio de amostras de famílias podem levar a resultados mais confiáveis em relação aos estudos utilizando indivíduos não relacionados em nosso meio. Neste caso, a definição dos padrões de herança Mendeliano em uma amostra familiar é de extrema importância, visto que esta informação é um dos principais fatores que influenciam os modelos estatísticos utilizados para o mapeamento genético do fenótipo em estudo.

1.5. Definição do padrão de herança: análise de segregação

O objetivo da análise de segregação é determinar a presença de padrões de herança genética de um determinado fenótipo a partir de uma amostra de dados obtidos de famílias onde tais fenótipos estão presentes.

Morton e MacLean (23) desenvolveram um modelo matemático o qual considera que um dado fenótipo é uma variável contínua (x), resultante do efeito de um gene principal (g), um componente de transmissão multifatorial (c) e um componente ambiental (e), como dado na seguinte fórmula:

$$x = g + c + e$$

Em relação ao gene principal, os autores pressupõem um par de alelos *A* e *a* com freqüências *p* e *q*, respectivamente. Os genótipos *AA*, *Aa* e *aa*, distribuem-se na população segundo a teoria de *Hardy-Weinberg* (24). Ambos os componentes de transmissão multifatorial e ambiental têm distribuição normal em torno das médias dos genótipos *AA*, *Aa* e *aa* (Figura 2). Portanto, a variância (V) total do fenótipo x é dada pela seguinte fórmula:

$$V = G + C + E$$

Quando o fenótipo apresenta um forte componente ambiental ou multifatorial, a coordenada da curva é mais larga. Caso o componente genético seja mais forte, a coordenada é mais estreita.

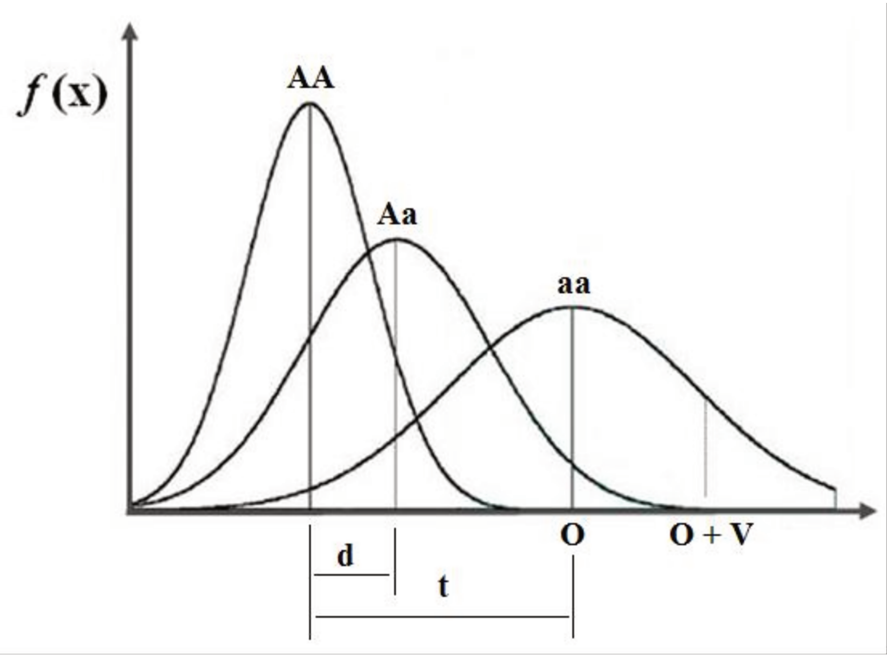


Figura 2. Curva de distribuição normal N (μ=O; σ^2=V) do fenótipo *x* em uma determinada amostra. d = dominância; t = deslocamento.

Estes valores são calculados por meio de uma função de probabilidades baseada em uma amostra de famílias nucleares, ou seja, pai, mãe e filhos. Esta função de probabilidade é dada pela seguinte fórmula:

$$p\left(\Psi_o \mid \Psi_p, S, s, \Theta\right) = \frac{p\left(S \mid \Psi_o, s\right) p\left(\Psi_o \mid \Psi_p, s, \Theta\right)}{p\left(S \mid \Psi_o, s, \Theta\right)}$$

onde Ψ_o = conjunto de fenótipos dos filhos; Ψ_p = conjunto de fenótipos dos pais; S = evento que significa pelo menos um indivíduo afetado na família; s = número de filhos; Θ = parâmetros utilizados na análise.

O conjunto de parâmetros Θ tem extrema importância na montagem dos modelos matemáticos de segregação. Lalouel & Morton (25) desenvolveram um modelo de análise de segregação baseados nos seguintes conjuntos de parâmetros: deslocamento (t), dominância (d), freqüência do alelo (q), herdabilidade (H) e probabilidades de transmissão (τ) do alelo *A*.

Como mostrado na Figura 2, o deslocamento (t) é a distância entre as médias dos dois genótipos homozigotos *AA* e *aa.* A dominância (d) é a posição da média do genótipo heterozigoto em relação às média do homozigoto *AA*. Caso o fenótipo apresente dominância completa, d = 1 enquanto que se houver codominância, d = 0,5. Caso exista dominância incompleta, 0,5 < d < 1. A herdabilidade (H) é dada pela razão C/V e reflete a transmissão multifatorial. As probabilidades de transmissão são parâmetros adicionais, os quais estimam desvios da transmissão Mendeliana do gene principal de pais para filhos. Estes parâmetros, indicados por τ_{AA}, τ_{Aa} e τ_{aa}, mostram respectivamente as probabilidades dos genótipos *AA*, *Aa* e *aa* de transmitirem o alelo *A*.

A partir destes parâmetros é elaborado um modelo misto, o qual servirá como hipótese nula (H0), além de modelos teste que refletem determinado padrão de herança, como mostrado na Tabela 1. Os parâmetros de cada modelo são estimados por meio de máxima verossimilhança (*Maximum Likelihood Estimation* – MLE) e, após esta estimativa, cada modelo teste é comparado ao modelo Misto por meio do teste da razão da máxima verossimilhança (*Likelihood Ratio Test* – LRT). Visto que -2*ln*(LRT) apresenta uma distribuição qui-quadrado com um grau de liberdade (χ^2), valores de p podem ser obtidos para estas comparações, sendo que valores de p < 0,05 **rejeitam** o modelo teste, enquanto que valores de p > 0,05 **não rejeitam** o modelo teste (25; 26)

Em situações onde é encontrado mais de um padrão de herança estatisticamente significativo, é utilizado o Critério de Informação de Akaike (*Akaike Information* Criteria - AIC), o qual é dado pela seguinte equação:

$$\left[-2\ln\left(LRT\right)\right]+2m$$

onde *m* = número de graus de liberdade de cada modelo. O modelo com menor valor de AIC indica o ajuste mais parcimonioso dos dados e, portanto, o mais provável (27).

Tabela 1. Modelos para análise de segregação.

Modelo	Parâmetro						
	d	**t**	**q**	**H**	τ_{AA}	τ_{Aa}	τ_{aa}
Misto (hipótese nula)	est.	est.	est.	est.	[1]	[0,5]	[0]
Esporádico	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Ausência de gene principal	[0]	[0]	[0]	est.	[1]	[0,5]	[0]
Multifatorial	est.	est.	est.	[0]	[1]	[0,5]	[0]
Autossômico dominante	[1]	est.	est.	est.	[1]	[0,5]	[0]
Autossômico codominante	[0,5]	est.	est.	est.	[1]	[0,5]	[0]
Autossômico recessivo	[0]	est.	est.	est.	[1]	[0,5]	[0]
Presença de transmissão Mendeliana	*	est.	est.	est.	[1]	[0,5]	[0]
Ausência de transmissão Mendeliana	*	est.	est.	est.	est.	est.	est.

est.= valores estimados; * = valor fixado dependendo do modelo de segregação autossômica encontrado; [ ] = valores fixados para cada modelo

Todos estes cálculos são realizados por meio do programa POINTER© (25) a partir de um arquivo contendo os dados das famílias em estudo e outro arquivo contendo as informações do fenótipo estudado. Entretanto, para uma boa estimativa dos parâmetros, é necessário um número grande de famílias nucleares (um mínimo de aproximadamente 90), além da adequada caracterização do fenótipo.

O padrão de herança definido pela análise de segregação, juntamente com os parâmetros adicionais estimados, são utilizados na escolha da análise de mapeamento genético mais adequado para o tipo de amostra avaliado. Há dois tipos de análises: as **paramétricas** e as **não paramétricas**, as quais serão discutidas adiante.

1.6. Análises paramétricas

Descrito primeiramente em 1931 por Bernstein (1), a **genética reversa** tem como princípio determinar a posição de um gene a partir de um marcador molecular conhecido por meio da frequência de recombinação meiótica (θ) entre eles, sendo que quanto menor a distância entre o gene e o marcador molecular, menor será θ. A ligação é detectada quando

ambos o gene e o marcador molecular são transmitidos de uma geração a outra de uma mesma família sem que ocorra recombinação meiótica entre eles. Portanto, uma vez detectada a ligação genética e conhecendo a posição cromossômica do marcador molecular, a posição cromossômica do gene de interesse pode ser inferida.

Em 1956, a partir das teorias de Bernstein, Morton (28) desenvolveu uma função matemática Z(θ) para detecção de ligação genética denominada *lod score*. Dados os genótipos G observados da amostra, esta função é calculada a partir da razão entre a probabilidade de ligação gênica, definida pela função $L(G \mid 0{,}00 \leq \theta < 0{,}5)$, e a probabilidade de não ligação, definida pela função $L(G \mid \theta = 0{,}5)$, visto que θ = 0,5 é definido como a freqüência de recombinação entre dois genes situados em cromossomos diferentes e, portanto, não ligados entre si. Este conceito se resume na seguinte fórmula:

$$Z(\theta) = \log_{10}\left[\frac{L(G \mid 0 \leq \theta < 0{,}5)}{L(G \mid \theta = 0{,}5)}\right]$$

De acordo com Morton (28), valores de $Z \geq 3$ indicam ligação, enquanto que valores de $Z \leq -2$ indicam recombinação aleatória, ou não ligação. Valores de $-2 < Z < 3$ indicam resultados não informativos.

Na prática, visto a complexidade de parâmetros adicionais que podem influenciar os cálculos de Z, tais como penetrância incompleta (1), Terwillinger e Ott (21), em 1994, desenvolveram um pacote de programas estatísticos, denominado LINKAGE®, para realização dos cálculos e estimação de Z.

Estes **estudos de ligação genética** são utilizados principalmente para mapeamento de regiões em doenças tipicamente monogênicas e com padrão de herança Mendeliano definido. Apesar de ser uma poderosa ferramenta para genética reversa, o cálculo de Z apresenta resultados informativos em situações em que a amostra estudada inclui somente uma ou poucas famílias, com um grande número de indivíduos avaliados em cada família. Em situações em que a amostra inclui um número grande de famílias com poucos indivíduos avaliados por família, um grande número de resultados não informativos é observado.

Para resolver este problema, Kruglyak et al. (29) desenvolveram um modelo estatístico de análise paramétrica robusto para famílias de tamanho moderado. Neste

modelo, o cálculo de Z utiliza a probabilidade da presença do fenótipo (Φ) condicionada a um vetor $v(x)$ que indica a transmissão do marcador molecular x de pais para filhos na família, como mostra a equação abaixo:

$$Z(\theta) = \log_{10}\left[\frac{L[\Phi \mid G, \theta = v(x)_{completo}]}{L[\Phi \mid G, \theta = v(x)_{uniforme}]}\right]$$

onde $v(x)_{completo}$ = a probabilidade de transmissão observada dos marcadores; $v(x)_{uniforme}$ = a probabilidade de transmissão esperada dos marcadores.

A estimação destas probabilidades é feita por meio do método de Cadeias de Markov Ocultas, visto que os genótipos, bem como a transmissão dos alelos, podem não ser diretamente observados em todos os indivíduos (30). Todos estes cálculos foram reunidos no pacote de programas GENEHUNTER® (29).

Apesar do alto poder estatístico apresentado por estes dois cálculos, o conhecimento do padrão de herança Mendeliano do fenótipo é a premissa principal destes estudos. No caso de doenças complexas, onde se espera a presença de genes de susceptibilidade e de menor efeito atuando no fenótipo, as análises não paramétricas de ligação genética são as mais adequadas para a clonagem posicional.

1.7. Análises não paramétricas

As análises não paramétricas de ligação tiveram início em 1990 com os trabalhos teóricos de Risch sobre análises de pares de irmãos afetados (31; 32). Em 1993, Spielman et al. (33) desenvolveram um método denominado Teste de Transmissão de Desequilíbrio de ligação (*Transmission Disequilibrium Test* – TDT), onde a probabilidade de transmissão de um alelo *A* de pais afetados para os filhos afetados é comparada com a probabilidade de não transmissão deste alelo *A* (Tabela 2).

Tabela 2. Relação entre a transmissão dos alelos *A* e *a* entre uma amostra familiar

	alelo *A*	**alelo *a***	**Total**
Transmitido	x	$n - x$	n
Não transmitido	y	$n - y$	n
Total	$x+y$	$2n-x-y$	$2n$

Visto que os dados dos genótipos dos pais podem ser desconhecidos ou não avaliados, Gordon et al. (34) propuseram um modelo de TDT baseado em LRT, o qual é robusto para este tipo de situação, como é mostrado na fórmula abaixo:

$$TDT = 2\ln \frac{L\left(G \mid \theta = \hat{R}_i, \hat{p}_i, \hat{E}\right)}{L\left(G \mid \theta = 1, \hat{\hat{p}}_i \hat{\hat{E}}\right)}$$

onde G = genótipos observados da amostra; R_i = risco relativo do genótipo i; p_i = frequência do genótipo i; E = categorização erros da amostra. Estes parâmetros são estimados via MLE e, visto que o valor de TDT tem uma distribuição χ^2, valores de p podem ser obtidos para avaliar a significância estatística da análise. Estes cálculos foram incluídos em um programa chamado TDTAE (TDT *allowing errors*) (34). Entretanto, um número baixo de famílias pode inflar os valores de TDT, gerando resultados falso-positivos (33; 34)

Devido a este problema, Dudbridge (20) elaborou um modelo onde a probabilidade da presença dos genótipos dos filhos fica condicionada à probabilidade dos genótipos dos pais em um modelo de regressão logística (26), como mostrado na seguinte fórmula:

$$P\left(c \mid f, m, y, z\right) = \frac{\exp\left[y\left(X_c \beta + X_{c,z}\gamma\right)\right]}{\sum_{c^* \in S(f,m)} \exp\left[y\left(X_{c^*}\beta + X_{c^*,z}\gamma\right)\right]}$$

onde c = genótipo do fílho; f = genótipo paterno; m = genótipo materno; y = fenótipo do filho; z = covariante do fenótipo do filho; X_c = vetor dos códigos dos possíveis genótipos do filho, ou seja, *AA*, *Aa* ou *aa*; β *e* γ = coeficientes de regressão (26).

As principais diferenças deste modelo em relação ao cálculo proposto por Gondor et al. (34) são o maior poder estatístico para um número menor de famílias e a possibilidade do uso de covariantes, devido à presença do coeficiente de regressão γ (26). Estes cálculos são realizados por meio do programa UNPHASED (20), sendo também denominadas de **estudos de associação baseadas em famílias**.

As análises não paramétricas também são utilizadas para estudos populacionais. Nestas análises, denominadas **estudos de associação genética populacional**, as frequências alélicas e genotípicas de um grupo de pacientes não relacionados (caso) são comparadas às frequências de um grupo de indivíduos não afetados com o fenótipo e não relacionados (controle). A partir destes dados, são montadas tabelas de frequências observadas e esperadas das amostras. Portanto, a distribuição de χ^2 é calculada pela seguinte fórmula:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

onde O_i = valores observados na célula i da tabela; E_i = valores esperados na célula i da tabela de esperados. Apesar de extremamente poderoso do ponto de vista estatístico, é necessário um número muito grande de indivíduos para evitar resultados tanto falso-positivos quanto falso-negativos. Além disso, o teste de associação alélica tem como consequência a perda de informação dos indivíduos heterozigotos, visto que possuem ambos os alelos, dificultando a categorização entre as classes.

Uma alternativa ao teste de χ^2 é a análise de associação genética utilizando um modelo de regressão logística, por meio da seguinte fórmula:

$$P(\Phi \mid \theta = \hat{\beta}, X_g, X_i) = \frac{e^{\hat{\beta}_0 + \hat{\beta}_g X_g + \sum_i \hat{\beta}_i X_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_g X_g + \sum_i \hat{\beta}_i X_i}}$$

onde Φ = fenótipo; β = conjunto de coeficientes de regressão; X_g = genótipos da amostra; X_i = covariantes observadas na amostra. Esta análise permite o uso de um número menor de amostra em relação ao teste de χ^2, além da inclusão de covariantes, as quais podem ser ambientais. Ambos os testes podem ser feitos em ambiente R (35)

1.8. Marcadores moleculares

Definida a amostra e o tipo de análise a ser utilizado, o próximo passo é a escolha do marcador molecular mais adequado para o estudo. Um mapa de marcadores moleculares funciona como um ponto de referência no genoma no processo de clonagem posicional (36). Os marcadores moleculares mais utilizados são as **regiões microssatélites** e os **SNPs** (*single nucleotide polymorphisms*).

As regiões microssatélites têm a principal característica de apresentar repetições de uma determinada sequência de nucleotídeos, denominada sequência em *tandem*. Por este motivo, estas regiões são altamente polimórficas (Figura 3). De acordo com o *National Center of Biotechnology Information*, há atualmente, no genoma humano, aproximadamente 7.000 regiões microssatélites (36; 37).

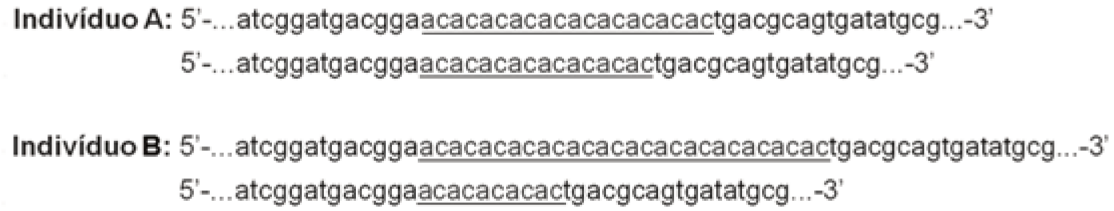


Figura 3. Exemplo de regiões microssatélites (sequência sublinhada).

Os SNPs são mutações de troca de um nucleotídeo na sequência de DNA (Figura 4). Apesar de raras exceções, visto que esta mutação de ponto envolve dois nucleotídeos, este tipo de marcador é bialélico (36). Com os estudos realizados pelo projeto HapMap já foram validados aproximadamente 19 milhões de SNPs pelo genoma humano (38; 39).

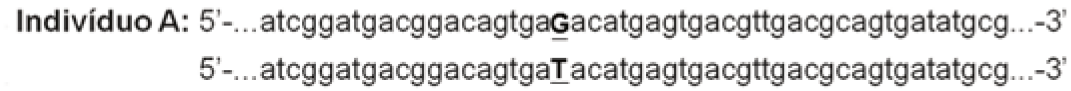


Figura 4. Exemplo de um SNP no indivíduo A (nucleotídeo sublinhado).

Um fator importante do ponto de vista estatístico é o número de marcadores moleculares a ser utilizado, o qual é condicionado ao grau de conhecimento prévio da fisiopatologia do fenótipo em estudo (2; 40).

A primeira situação é o caso de haver evidências de alguma via metabólica específica, por meio de ensaios bioquímicos ou ensaios indiretos com medicamentos, ou evidências de algum gene possivelmente relacionado com o fenótipo, descritos em trabalhos anteriores. Nesta situação os estudos a partir de **genes ou regiões candidatas** são os mais apropriados, pois um número pequeno de marcadores é utilizado (Tabela 3). A segunda situação é a total falta de evidências da fisiopatologia do fenótipo, ou evidências generalizadas demais, levando a uma confusa escolha dos genes/regiões candidatas. Nesta situação, o estudo de mapeamento em todo o genoma é mais adequado (2; 36; 40), como mostra a Tabela 3.

Tabela 3. Número de marcadores moleculares utilizados nas análises.

Análise	**Marcador**	**Distância média entre marcadores**	**Genes candidatos**	**Todo o genoma**
Paramétrica	Microssatélites	~10 Mb*	4 a 6 / gene	300 a 400
Não paramétrica	SNPs	~ 300 kb**	5 a 100 / gene	$1x10^4$ a $1x10^6$

*Mb = 1.000.000 pares de bases; **kb = 1.000 pares de bases.

A utilização de um número grande de marcadores moleculares gera um aumento no número de testes a ser computado, resultando em um viés estatístico denominado problema dos **múltiplos testes** (41), onde um aumento no número de testes infla a taxa de falso-positivos na amostra, como é exemplificado na Tabela 4. Por consequência, o limite dos valores de p deve tornar-se cada vez menor por meio de ajustes, tais como a correção

de Bonferroni e o *False Discovery Rate*, levando a um aumento do tamanho da amostra necessária, a qual pode chegar a um mínimo de 3000 indivíduos em estudos de associação populacional analisando todo o genoma (40; 41).

Tabela 4. Número de resultados falso-positivos por número de testes.

Número de testes	**Nível de significância α**					
	0,05	**0,01**	**0,001**	**0,0001**	**1×10^{-5}**	**1×10^{-6}**
1	0	0	0	0	0	0
10	0	0	0	0	0	0
20	1	0	0	0	0	0
100	5	1	0	0	0	0
1.000	50	10	1	0	0	0
10.000	500	100	10	1	0	0
100.000	5.000	1.000	100	10	1	0
1.000.000	50.000	10.000	1.000	100	10	1

O problema dos múltiplos testes é um dos fatores principais da inviabilidade de estudos de associação populacional (40). Apesar de afetar também as análises com amostras de famílias, o tamanho necessário da amostra familiar é consideravelmente menor em relação a uma amostra populacional, chegando a um mínimo de 400 indivíduos (42).

A identificação dos genótipos destes marcadores moleculares nos indivíduos é realizada por meio de equipamentos de aquisição de dados, tais como MegaBACE© 1000 (GE Healthcare, Buckinghamshine, UK), ABI® 7500 (Applied Biosystems, Foster City, CA, EUA), bem como chips de *microarray* (41). Além de gerar uma grande quantidade de informação, os dados provenientes destes equipamentos necessitam de um prévio

processamento para o uso adequado aos diversos programas estatísticos utilizados nas análises. Visto que não há programas para este tipo de tarefa, surge o primeiro problema: **como realizar a interface entre os dados dos equipamentos e os dados utilizados pelos cálculos estatísticos.**

Apesar das vantagens que uma amostra familiar apresenta em relação a uma amostra populacional, na prática são observadas várias famílias de tamanhos e estruturas familiares diferentes para doenças complexas neuro-psiquiátricas, levando a resultados espúrios somente devido ao desenho experimental empregado (2; 40). Neste momento, tem-se o segundo problema: **qual o melhor teste estatístico para uma amostra familiar variada.**

2. OBJETIVOS

2.1. OBJETIVO PRINCIPAL

Aplicar e avaliar os modelos estatísticos e desenvolver algoritmos para processamento de dados de genotipagem para as análises de mapeamento genético em famílias segregando doenças neuro-psiquiátricas.

2.1. OBJETIVOS ESPECÍFICOS

- Desenvolver algoritmos de interface entre os equipamentos de genotipagem e os cálculos estatísticos;
- Aplicar análises de segregação complexa nas famílias com ELTM;
- Realizar estudos paramétricos de ligação genética em uma família segregando PPBC;
- Realizar estudos não paramétricos de TDT em uma amostra de famílias segregando TAB

3. CAPÍTULOS

CAPÍTULO 1

ARTIGO 1

LINKGEN: A NEW ALGORITHM TO PROCESS DATA IN LINKAGE STUDIES

Rodrigo Secolin, Cristiane S. Rocha, Fábio R. Torres, Marilza L. Santos, Cláudia V. Maurer-Morelli, Neide F. Santos, Iscia Lopes-Cendes

Department of Medical Genetics, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil

Available online at www.sciencedirect.com
ScienceDirect
Genomics 91 (2008) 544–547
GENOMICS
www.elsevier.com/locate/ygeno

Short Communication

LINKGEN: A new algorithm to process data in genetic linkage studies

Rodrigo Secolin, Cristiane S. Rocha, Fábio R. Torres, Marilza L. Santos, Cláudia V. Maurer-Morelli, Neide F. Santos, Iscia Lopes-Cendes *

Department of Medical Genetics, Faculty of Medical Sciences, University of Campinas (UNICAMP), Tessália Vieira de Camargo, 126, Cidade Universitária Zeferino Vaz, Campinas SP, Brazil, CEP 13084-971

Received 24 July 2007; accepted 21 February 2008
Available online 2 April 2008

Abstract

Genetic linkage studies using whole genome scans are useful approaches for identifying genes related to human diseases. In general, these studies require genotyping of a large number of markers, which are used in statistical analysis. Recent technology has allowed easy genotyping of a large number of markers in less time; therefore, interface programs are required for manipulation of these large data sets. We present a new algorithm, which processes input data in LINKAGE format from data analyzed by automated genotyping systems. The algorithm was implemented in PERL script and R environment. Validation was performed with genotyped data from 127 individuals and 720 microsatellite markers of two whole genome scans. Our results showed a significant decrease in data processing time. In addition, this algorithm provides unbiased allele frequency estimation used for linkage analysis. LINKGEN is a freely available online tool and allows easier, faster, and reliable manipulation of large genotyping data sets.

Keywords: Linkage studies; Data handling; Bioinformatics

Introduction

Genetic linkage studies using whole genome scans are useful tools for finding human disease loci and genes [1,2]. Approximately 300–450 microsatellite markers or high-density SNP maps are required for these studies [3,4]. In parametric linkage studies, as well as meta-analysis genome scans, microsatellite markers are still the best choice for genotyping large single families or multiple small families [3,5,6]. Recently, high-throughput automated genotyping equipment has been used for analyzing these molecular markers, such as MegaBace (GE Heathcare, Chicago, IL) and ABI (Applied Biosystems, Foster City, CA) systems. However, data output format from most of this equipment is not compatible with input format required by several types of parametric and nonparametric genetic linkage software, such as LINKAGE, GENEHUNTER, MERLIN, and S.A.G.E. packages [7–10]. Although interface programs have been implemented for microarray-based SNP genotyping [11], there is no such tool available for processing of microsatellite genotype data. Furthermore, since not all individuals used in linkage studies are genotyped, the manual data processing of these large data sets leads to tedious and laborious work, which may increase the probability of clerical errors.

In this study we present a new algorithm, named LINKGEN, designed to generate an interface between most types of automated genotyping software and statistical packages for linkage analysis; in addition, it includes allele frequency estimates from pedigree data sets.

Results and discussion

LINKGEN achieves a faster and easier data processing for linkage analysis. Indeed, our results showed that all 720 microsatellite genotypes were correctly distributed across all 127 family members, using the LINKGEN algorithm in each validation study. Data manipulation can be carried out by the web; therefore, the software is able to operate in any operational system and for any type of analysis which uses genetic linkage statistical methods. In fact, several studies in our laboratory have used this algorithm [12–14], decreasing considerably the statistical processing time.

* Corresponding author. Fax: +55 19 3289 1818.
E-mail address: icendes@unicamp.br (I. Lopes-Cendes).

0888-7543/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2008.02.001

Since most output data formats from different automated genotyping equipment have at least a genotype call column, our software is able to accept output data from many genotyping systems currently in use, without a lot of preprocessing. In addition, most SNP genotyping equipment also includes genotype call columns, making LINKGEN also suitable for use in studies in which only a few SNP markers are genotyped.

Approximately 1500 disease genes have been identified and listed in the Online Mendelian Inheritance in Man (OMIM) using genetic linkage studies [2,15]. A large number of microsatellite or SNP markers are required for whole genome scans [2–4]. Since advanced genotyping techniques of microsatellite markers are now available, interface programs between genotyping and statistical software became an important tool for processing data in these studies, such as genome-wide scans and meta-analysis of genome scans [5,6]. In conclusion, this software allows construction of appropriate input files for most linkage analysis programs, providing an easier, faster, and reliable manipulation of large genotyping data sets in whole genome linkage analysis.

Materials and methods

Implementation

LINKGEN was implemented in PERL scripts and used two files, corresponding to two matrices defined as follows: a pedigree structure matrix $P=(p_{i,j})_{m\times n}$ built in PRE-LINKAGE format [3], and a marker genotype matrix $M=(m_{w,z})_{a\times b}$ from any genotyping system and any number of markers (Fig. 1). In P matrix, there are up to seven columns, as follows: $p_{i,1}$ = pedigree identification (ID); $p_{i,2}$=individual ID; $p_{i,3}$=father ID; $p_{i,4}$=mother ID; $p_{i,5}$=sex; $p_{i,6}$=affection status; and $p_{i,7}$=optional liability class. Here we included a last column $p_{i,g}$, with $g=7$ or 8, depending on the presence of a liability class column. This last column indicates whether the individuals have been genotyped ($p_{i,g}=1$) or not ($p_{i,g}=0$). In the allele marker matrix, we included two initial columns, where $m_{w,1}$=pedigree ID, and $m_{w,2}$=individual ID. Therefore, our algorithm combines these two matrices in a matrix $R=(r_{i,d})_{m\times[(n-1)+(z-2)]}$, where each i-th row follows the formula:

$$r_{i,1\ldots[(n-1)+(z-2)]}=\begin{cases}(p_{i,1}\ldots p_{i,g-1})\,U(m_{i,3}\ldots m_{i,z}) & \text{if } p_{i,1}=m_{w,1}\wedge p_{i,2}=m_{w,2}\wedge p_{i,g}=1\\ (p_{i,1}\ldots p_{i,g-1})\,U(m_{i,3}=0\ldots m_{i,z}=0) & \text{if } p_{i,1}=m_{w,1}\wedge p_{i,2}=m_{w,2}\wedge p_{i,g}=0\end{cases}.$$

Therefore, by this formula, each genotyped individual in pedigree data receives his/her corresponding group of allele markers, whereas each nongenotyped individual receives missing data code 0 (Fig. 1).

These files should have ***.fam** and ***.gen** extensions, regarding P and M matrices, respectively. Both files can be created in any text editor or calculus sheet processor software (i.e., Microsoft Word or Excel). In principle, the number of rows and columns is unlimited, depending on the file editor and computer configuration. As a result, a ***.pre** extension file in LINKAGE format is created.

Frequency estimates

Marker frequencies, denoted as p, are estimated based on sibship data according to the formula [16]:

$$p=\frac{\sum_{i=1}^{n}\sum_{j=1}^{k}X_{ij}(k_i+1)}{2\sum_{i=1}^{n}k_i(k_i+1)},$$

which, assuming that a marker can have a β allele, Xij is the number of β alelles carried by sibling j in family i (=0, 1, or 2); n is the number of families; and k_i denotes the number of siblings in i-th family. Considering Hardy-Weinberg equilibrium and absence of genotyping errors for markers studied, Broman [16] showed that this frequency estimation is unbiased and provides reliable results. These estimates were implemented in R environment and included to LINKGEN PERL scripts.

Validation

To validate the algorithm, we used genotyped data from two large whole genome scans performed in two Mendelian forms of partial epilepsy syndromes: familial mesial temporal lobe epilepsy (OMIM No. 608096) and autosomal dominant partial epilepsy with auditory features (OMIM No. 600512). These studies comprised 68 and 59 individuals and genotyping of 337 and 383 microsatellite markers, respectively. Genotyping was performed automatically using the MegaBACE 1000 (GE Heathcare) system and genotype call by the FRAGMENT PROFILE software. We used LINKAGE, GENEHUNTER, and S.A.G.E. packages for statistical analysis [7,8,10]. Pedigree structure matrix and genotyped allele matrix were built in Microsoft Excel software and saved as ***.fam** and ***.gen** files, respectively. Since genotypic data from MegaBace and ABI equipment do not have the same structure as ***.gen** file, this file is initially built with family and individual ID columns, followed by genotyped allele markers, which are copied and pasted from the genotype call column of FRAGMENT PROFILE output data of each marker file. Finally, these data are saved as ***.gen** file (Fig. 1).

Availability and requirements

Software home page: http://lgm.fcm.unicamp.br:9001/cgi-bin/linkgen/linkgen.cgi
Operating system(s): platform independent
Programming language: Perl (5.8.8) and R (2.5.1)
Other requirements: none
License: freely available

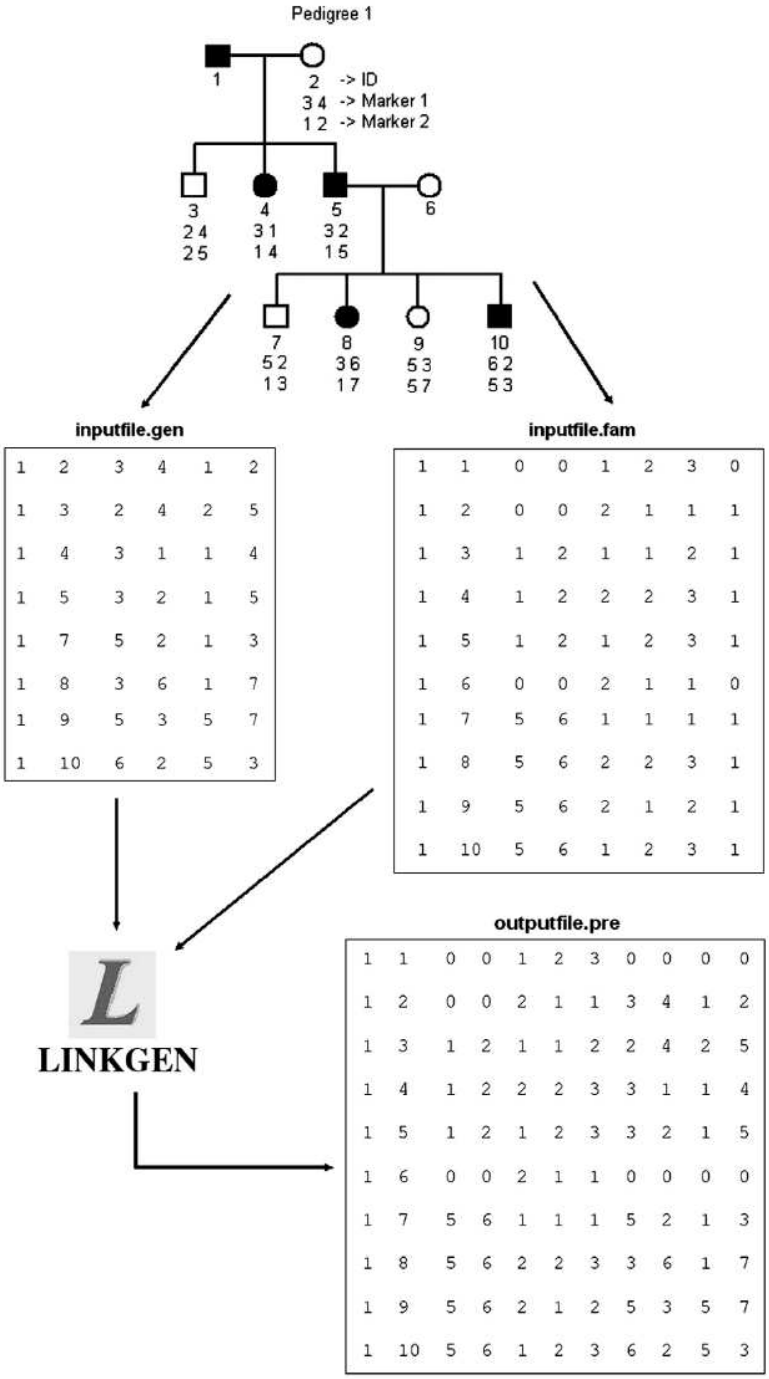


Fig. 1. Example of LINKGEN algorithm workflow. Squares and circles indicate males and females, respectively. White and black symbols represent unaffected and affected individuals, respectively. In this pedigree, individuals 1 and 6 are not genotyped.

Acknowledgments

This study was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), Brazil. Rodrigo Secolin, Cristiane S. Rocha, Fábio R. Torres, and Neide F. Santos received scholarships from FAPESP and Cláudia V. Maurer-Morelli received a fellowship from CNPq.

References

[1] J. Ott, Analysis of Human Genetic Linkage, Johns Hopkins University Press, MD, 1991.
[2] J.M. Vink, D.I. Boomsma, Gene finding strategies, Biol. Psychol. 61 (2002) 53–71.
[3] J.D. Terwillinger, J. Ott, Handbook of Human Genetic Linkage, Johns Hopkins University Press, MD, 1993.
[4] The International HapMap Consortium, a haplotype map of the human genome, Nature 437 (2005) 1299–1320.
[5] L.H. Wise, J.S. Lanchbury, C.M. Lewis, Meta-analysis of genome scans, Ann. Hum. Genet. 63 (1999) 263–272.
[6] C.J. Etzel, M. Liu, T.J. Costello, An updated meta-analysis approach for genetic linkage, BMC Genet. 6 (2005) S43.
[7] G.M. Lathrop, J.M. Lalouel, Easy calculations of lod scores and genetic risks on small computers, Am. J. Hum. Genet. 36 (1984) 460–465.
[8] L. Krgulyak, M.J. Daly, M.P. Reeve-Daly, E.S. Lander, Parametric and nonparametric linkage analysis: a unified multipoint approach, Am. J. Hum. Genet. 58 (1996) 1347–1363.
[9] G.R. Abecasis, S.S. Cherny, W.O. Cookson, L.R. Cardon, Merlin—rapid analysis of dense genetic maps using sparse gene flow trees, Nat. Genet. 30 (2002) 97–101.
[10] S.A.G.E. 5.1: Statistical Analysis for Genetic Epidemiology: http://darwin.cwru.edu/sage/.
[11] W. Liu, et al., Algorithms for large-scale genotyping microarrays, Bioinformatics 19 (2003) 2397–2403.
[12] C.V. Maurer-Morelli, et al., The SCN2A gene is not a likely candidate for familial temporal lobe epilepsy, Epilepsy Res. 71 (2006) 233–236.
[13] M.C. França Jr., et al., Prospective neuroimaging study in hereditary spastic paraplegia with thin corpus callosum, Mov. Disord. 22 (2007) 1556–1562.
[14] N.F. Santos, et al., A new candidate locus for bilateral perisylvian polymicrogyria mapped on chromosome Xq27, Am. J. Med. Genet. Part A, in press.
[15] Online Mendelian Inheritance in Man: http://www.ncbi.nlm.nih.gov.
[16] K.W. Broman, Estimation of allele frequencies with data on sibships, Genet. Epidemiol. 20 (2001) 307–315.

CAPÍTULO 2

ARTIGO 2

SEGREGATION ANALYSIS IN MESIAL TEMPORAL LOBE EPILEPSY WITH HIPPOCAMPAL ATROPHY

Rodrigo Secolin[1], Claudia Maurer-Morelli[1], Fernando Cendes[2], Iscia Lopes-Cendes[1]

Department of [1]Medical Genetics; and [2]Department of Neurology, Faculty of Medical Sciences, University of Campinas (UNICAMP), Campinas, Sao Paulo, Brazil

Epilepsia, 51(Suppl. 1):47–50, 2010
doi: 10.1111/j.1528-1167.2009.02445.x

Segregation analysis in mesial temporal lobe epilepsy with hippocampal atrophy

*Rodrigo Secolin, *Claudia Maurer-Morelli, †Fernando Cendes, and *Iscia Lopes-Cendes

*Department of Medical Genetics; and †Department of Neurology, Faculty of Medical Sciences, University of Campinas (UNICAMP), Campinas, Sao Paulo, Brazil

It is well known that among epilepsies with focal seizure onset, temporal lobe epilepsies (TLEs) are the most common forms (Gloor, 1991). Mesial temporal lobe epilepsy (MTLE) is one type of TLE frequently associated with mesial temporal sclerosis (MTS), a neuropathologic abnormality that can be diagnosed in vivo by high-resolution brain imaging as hippocampal atrophy (HA) and abnormal signal intensity (Berkovic et al., 1991; Cendes et al., 1993a). MTS, which is characterized by selective neuronal loss and gliosis in regions of hippocampus and hilus, has been associated with predisposing environmental factors, such as prolonged febrile seizures (FS) in childhood (Abou-Khalil et al., 1993; Cendes et al., 1993b).

Familial forms of TLE have been identified, suggesting that genetic factors could be involved in the determination of different types of TLE (Berkovic et al., 1996; Ottman et al., 1995; Gambardella et al., 2000). Familial recurrence of MTLE associated with HA has also been reported (Cendes et al., 1998; Fernandez et al., 1998; Kobayashi et al., 2001). Extensive phenotypic studies in large MTLE families showed a certain degree of variability in clinical and imaging findings, but most affected individuals have a benign course of the disease. In these families HA was observed in patients who had refractory seizures, requiring surgical treatment (Kobayashi et al., 2003a), as well as in individuals who had only a single partial seizure and, unexpectedly, in 34% of asymptomatic first-degree relatives of patients (Kobayashi et al., 2002, 2003b). These initial observations suggested that genetic factors may play a role in MTLE associated with HA. In order to further investigate this issue we performed complex segregation analysis in a sample of nuclear families of probands with MTLE and HA.

Address correspondence to Iscia Lopes-Cendes, MD, PhD, Department of Medical Genetics, FCM – UNICAMP, Tessália Vieira de Camargo, 126, Cidade Universitária Zeferino Vaz, 13084-971 Campinas, SP, Brazil. E-mail: icendes@unicamp.br

METHODS

Ascertainment of patients and data collection

Ascertainment of patients and clinical data collection are systematically performed in all probands with MTLE followed at the epilepsy clinic of our university hospital. The diagnosis of MTLE was based on clinical and electroencephalography (EEG) findings as defined by the International League against Epilepsy (ILAE) criteria (1989). Information on probands was collected regardless of family history and obtained from all patients who fulfilled the clinical and EEG criteria for MTLE. Subsequently, extended family history was obtained from the probands as well as available family members. HA was detected by volumetric magnetic resonance imaging (MRI) (Kobayashi et al., 2001, 2002). For calculation purposes the sample was divided into nuclear families, including the probands and their first-degree relatives. The ascertainment probability (π) was used for ascertainment correction (Lalouel & Morton, 1981; Ginsburg et al., 2003). All individuals evaluated provided written informed consent and this study was approved by the research ethics committee of our institution.

We considered as affected only individuals with MTLE associated with HA confirmed by MRI; unaffected individuals who did not present MTLE regardless of the presence of HA; and unknown individuals with MTLE but not evaluated with MRI.

Segregation analysis

Segregation analysis was performed under the mixed model implemented by the POINTER software (available at http://cedar.genetics.soton.ac.uk/pub/PROGRAMS/pointer, University of Southampton, U.K.) (Lalouel & Morton, 1981). The mixed model assumes a phenotype (x) with independent contribution of a major gene locus (g), a multifactorial component (c), and an environmental component (e). The overall phenotype is defined as $x = g + c + e$, and total variance is defined as $V = G + C + E$. The major locus has two alleles (A, a) and the

genotype frequencies follow the Hardy-Weinberg equilibrium.

Four parameters were estimated: dominance (d), or the relative position of heterozygote mean, where d = 1 indicates a dominant gene, whereas d = 0.5 indicates additive and d = 0 indicates a recessive gene; displacement (t) between the two extreme homozygote means; allele frequency (q); and heritability (H), which represents the proportion between the variance of multifactorial component and total variance (H = C/V).

Transmission probabilities (τ_1, τ_2, and τ_3) were analyzed for Mendelian pattern of transmission from parents to offspring. These parameters—τ_1, τ_2, and τ_3–denote probabilities of transmitting allele A for genotypes AA, Aa, and aa, respectively. Under Mendelian transmission, $\tau_1 = 1$, $\tau_2 = 0.5$, $\tau_3 = 0$ and no transmission, $\tau_1 = \tau_2 = \tau_3$.

Models were estimated by maximizing conditional likelihood (L) of nuclear family phenotypes. The difference between −2lnL under a general model with m parameters and −2lnL under a nested model with n parameters is χ^2 asymptotically distributed, with m−n degrees of freedom, where p-values <0.05 reject the analyzed model. Because these results can be biased for multiple tests, p-values were adjusted using false discovery rate (FDR) correction (Storey & Tibshirani, 2003) by *p.adjust* function in R environment (R Development Core Team, 2006). In addition, we used the Akaike Information Criterion (AIC) (Akaike, 1974), which is −2lnL plus twice the number of free parameters in the model. This comparison has the advantage that one model does not have to be a subset of the other one. The model with lowest AIC indicates most parsimonious fit of the observed data.

Results

We evaluated a total of 698 individuals, distributed into 148 nuclear families, which were ascertained from 76 unrelated probands with MTLE ($\pi = 0.60$ for ascertainment correction). There were 95 nuclear families (64.2%) with only one affected individual and 53 (35.8%) with two or more affected individuals (Table 1). There were 123 affected, 543 unaffected, and 32 unknown individuals.

Table 1. Proportion of affected individuals per family

No. of individuals/nuclear family	No. of nuclear families (%)
1	95 (64.2)
2	22 (14.9)
3	15 (10.1)
4	4 (2.7)
5	7 (4.7)
6	5 (3.4)
Total	148 (100)

Among the 543 unaffected individuals, we found 15 individuals with HA; 91 individuals without HA; and 437 individuals who did not have MRI studies performed. Among the 123 affected individuals, there were 20 (12%) who reported an antecedent of febrile seizures.

As shown in Table 2, our results rejected the random model ($\chi^2_4 = 133.760$; $p < 0.001$) and the absence of a major a gene ($\chi^2_3 = 133.760$; $p < 0.001$). We could not reject the multifactorial model ($\chi^2_1 = 0.002$; $p = 0.999$), as well as recessive ($\chi^2_2 = 2.084$; $p = 0.407$), codominant ($\chi^2_2 = 0.112$; $p = 0.999$), and dominant ($\chi^2_2 = 0.002$; $p = 0.999$) inheritance. However, comparison of AIC values indicated that the autosomal dominant model (AIC = 315.096) is more parsimonious than recessive (AIC = 317.178) and codominant (AIC = 315.206) models. In addition, we did not reject a Mendelian transmission ($\chi^2_3 = 3.665$; $p = 0.450$) in comparison with non-Mendelian transmission ($\chi^2_3 = 30.159$; $p < 0.001$).

Discussion

Complex segregation analysis has been successfully used to evaluate the transmission of a trait from pedigree data in several diseases (Jarvik, 1998), showing that complex segregation analysis is a powerful and reliable tool even in the molecular genetics era. Complex segregation analysis can determine whether a Mendelian locus has an effect on a particular phenotype. In addition, it can test for the possible inheritance pattern and the magnitude of environmental and polygenic effects that could be influencing the final phenotype. All these inferred parameters can be subsequently used in linkage analysis, fine mapping, and other gene identification strategies related to the phenotype (Jarvik, 1998).

Several studies have suggested different inheritance patterns in various epilepsy syndromes. Direct observation of pedigrees with generalized epilepsy with febrile seizures plus, benign familial neonatal convulsions, autosomal dominant nocturnal frontal epilepsy, and familial partial epilepsy with variable foci is consistent with autosomal dominant inheritance with incomplete penetrance; whereas, most progressive myoclonus epilepsy syndromes present an autosomal recessive mode of inheritance (Callenbach et al., 2005). Moreover, Ottman et al. (1995) suggested an autosomal dominant inheritance in a single family with partial epilepsy and auditory auras based in a preliminary segregation analysis. Although Berkovic et al. (1996) have proposed an autosomal dominant inheritance in a familial form of TLE with no MRI abnormalities, to our knowledge there is no previous segregation analysis performed in MTLE with HA.

The relationship between MTLE and MTS has been recognized in classical histopathologic studies (Gloor, 1991; Blümcke et al., 1999) and more recently correlated with neuroimaging findings identified in vivo by high-

Epilepsia, 51(Suppl. 1):47–50, 2010
doi: 10.1111/j.1528-1167.2009.02445.x

Table 2. Segregation analysis performed for mesial temporal lobe epilepsy (MTLE) with hippocampal atrophy (HA). Values between square brackets were fixed for calculation purposes

Model	d	t	Q	H	τ_1	τ_2	τ_3	−2lnL	χ^2	d.f.	p*	Test	AIC
1. Mixed	1.000	1.736	0.286	0.064	[1]	[0.5]	[0]	311.094					319.094
2. Sporadic	[0]	[0]	[0]	[0]	[1]	[0.5]	[0]	444.854	133.760	4	0.000	2 × 1	444.854
3. Absence major gene	[0]	[0]	[0]	0.700	[1]	[0.5]	[0]	444.854	133.760	3	0.000	3 × 1	446.854
4. Multifactorial	1.002	1.736	0.286	[0]	[1]	[0.5]	[0]	311.096	0.002	1	0.999	4 × 1	317.096
5. Recessive (d = 0)	[0]	1.921	0.700	[0]	[1]	[0.5]	[0]	313.178	2.084	2	0.407	5 × 1	317.178
6. Codominant (d = 0,5)	[0.5]	2.841	0.287	[0]	[1]	[0.5]	[0]	311.206	0.112	2	0.999	6 × 1	315.206
7. Dominant (d = 1)	[1]	1.738	0.286	[0]	[1]	[0.5]	[0]	311.096	0.002	2	0.999	7 × 1	315.096
8. Mendelian	[1]	1.637	0.346	[0]	1.000	0.962	0.000	229.637	81.459	3	0.000	7 × 8	239.637
9. Non-Mendelian	[1]	1.808	0.299	[0]	[0.701]	[0.701]	[0.701]	444.854	215.217	3	0.000	9 × 8	448.854

d, dominance; t, displacement; q, allele frequency; H, heritability; τ_1, τ_2, τ_3, probabilities of transmitting allele A for genotypes AA, Aa and aa, respectively; L, likelihood; χ^2, chi-square; d.f., degrees of freedom; p*, corrected p-values; AIC, Akaike Information Criterion.

resolution MRI (Cendes et al., 1993a; Kobayashi et al., 2001), making HA a surrogate marker for MTS in patients with intractable MTLE. Until recently, only environmental risk factors were associated with the development of HA and MTLE, especially the occurrence of prolonged childhood FS (Cendes et al., 1993b). However, more recently, evidence suggesting the involvement of genetic factors predisposing to HA in MTLE was found by the study of a large cohort of families segregating MTLE (Kobayashi et al., 2001, 2002, 2003a). These previous clinical observations are supported by our results, since our complex segregation analysis strongly suggests that MTLE with HA could be influenced by a major gene inherited in an autosomal dominant pattern.

In addition, our results showed that a multifactorial effect could not be rejected, indicating that genes of minor effect could be acting as modifiers of the final phenotype. In fact, a certain degree of clinical variability is observed in patients with MTLE especially regarding seizures severity (Baulac et al., 2004). These genes of minor effect (or modifiers genes) could also explain the remarkable differences in disease severity observed even within families in which individuals with MTLE have good seizure control on antiepileptic medication; whereas, other affected family members have medically refractory seizures, needing surgical treatment (Kobayashi et al., 2001, 2003a). In addition, the still-complex relationship between the presence of MTS and the occurrence of seizures could be, at least in part, explained by genetic variants (i.e., sequence polymorphisms) in these genes of minor effect. In this context, it is interesting to note the presence of 15 individuals (first-degree relatives of probands with MTLE) who did not have seizures but show HA on volumetric MRI. These individuals, if found to carry the same major gene mutation as individuals with HA and MTLE, could be also carriers of genetic variants in genes of minor effect, which may protect them against seizures even in the presence of morphologic changes in the mesial temporal structures.

With the recent development of tools and strategies for disease gene mapping (Vink & Boomsma, 2002; The International HapMap Consortium, 2005), the identification of major genes, as well as of genes of minor effect, may be helpful in the better understanding of the mechanisms associated with the development of MTLE and HA and could help to clarify the complex relationship between MTS and the occurrence of seizures in patients with MTLE.

Acknowledgments

This study was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil). We are grateful to our patients and their families for their helpful cooperation and to Dr. Ricardo G.M. Ferreira for technical assistance with the POINTER software.

Disclosure

None of the authors has any conflict of interest to disclose.

References

Abou-Khalil B, Andermann E, Andermann F, Olivier A, Quesney LF. (1993) Temporal lobe epilepsy after prolonged febrile convulsions: excellent outcome after surgical treatment. *Epilepsia* 34:878–883.

Akaike H. (1974) A new look at the statistical model identification. *IEEE Trans Automat contr* 19:716–723.

Baulac S, Gourfinkel-An I, Nabbout R, Huberfeld G, Serratosa J, Leguern E, Baulac M. (2004) Fever, genes and epilepsy. *Lancet Neurol* 3:421–430.

Berkovic SF, Andermann F, Olivier A, Ethier R, Melanson D, Robitaille Y, Kuzniecky R, Peters T, Feindel W. (1991) Hippocampal sclerosis in temporal lobe epilepsy demonstrated by magnetic resonance imaging. *Ann Neurol* 29:175–182.

Berkovic SF, McIntosh A, Howell RA, Mitchell A, Sheffield LJ, Hopper JL. (1996) Familial temporal lobe epilepsy: a common disorder identified in twins. *Ann Neurol* 40:227–235.

Blümcke I, Beck H, Lie AA, Wiestler OD. (1999) Molecular neuropathology of human mesial temporal lobe epilepsy. *Epilepsy Res* 36:205–223.

Callenbach PMC, Maagdenberg AMJM, Frants RR, Brouwer OF. (2005) Clinical and genetic aspects of idiopathic epilepsies in childhood. *Neurology* 9:91–103.

Cendes F, Andermann F, Gloor P, Evans A, Jones-Gotman M, Watson C, Olivier MA, Peters T, Lopes-Cendes I. (1993a) MRI volumetric measurement of amygdala and hippocampus in temporal lobe epilepsy. *Neurology* 43:719–725.

Cendes F, Andermann F, Dubeau F, Gloor P, Evans A, Jones-Gotman M, Olivier A, Andermann E, Robitaille Y, Lopes-Cendes I. (1993b) Early childhood prolonged febrile convulsions, atrophy and sclerosis of mesial structures, and temporal lobe epilepsy: an MRI volumetric study. *Neurology* 43:1083–1087.

Cendes F, Lopes-Cendes I, Andermann E, Andermann F. (1998) Familial temporal lobe epilepsy: a clinically heterogeneous syndrome. *Neurology* 50:554–557.

Commission on Classification and Terminology of the International League against Epilepsy. (1989) Proposal for revised classification of epilepsies and epileptic syndromes. *Epilepsia* 30:389–399.

Fernandez G, Effenberger O, Vinz B, Steinlein O, Elger CE, Dohring W, Heinze HJ. (1998) Hippocampal malformation as a cause of familial febrile convulsions and subsequent hippocampal sclerosis. *Neurology* 50:909–917.

Gambardella A, Messina D, Le Piane E, Oliveri RL, Annesi G, Zappia M, Andermann E, Quattrone A, Aguglia U. (2000) Familial temporal lobe epilepsy autosomal dominant inheritance in a large pedigree from southern Italy. *Epilepsy Res* 38:127–132.

Ginsburg E, Malkin I, Elston RC. (2003) Sampling correction in pedigree analysis. *Stat Appl Genet Mol Biol* 2:1–22.

Gloor P. (1991) *Epilepsy surgery*. Raven Press, NY, USA.

Jarvik GP. (1998) Complex Segregation Analysis: uses and limitations. *Am J Hum Genet* 63:942–946.

Kobayashi E, Lopes-Cendes I, Guerreiro CAM, Sousa SC, Guerreiro MM, Cendes F. (2001) Seizure outcome and hipocampal atrophy in familial mesial temporal lobe epilepsy. *Neurology* 56:166–172.

Kobayashi E, Li ML, Lopes-Cendes I, Cendes F. (2002) MRI evidence of hippocampal sclerosis in asymptomatic first degree relatives of patients with familial mesial temporal lobe epilepsy. *Arch Neurol* 59:1891–1894.

Kobayashi E, D'Agostino MD, Lopes-Cendes I, Andermann E, Dubeau F, Guerreiro CA, Schenka AA, Queiroz LS, Olivier A, Cendes F, Andermann F. (2003a) Outcome of surgical treatment in familial mesial temporal lobe epilepsy. *Epilepsia* 44:1080–1084.

Kobayashi E, D'Agostino MD, Lopes-Cendes I, Berkovic SF, Li LM, Andermann E, Andermann F, Cendes F. (2003b) Hippocampal atrophy and T2 weighted signal changes in familial mesial temporal lobe epilepsy. *Neurology* 60:405–409.

Lalouel JM, Morton NE. (1981) Complex segregation analysis with pointers. *Hum Hered* 31:312–321.

Ottman R, Risch N, Hauser WA, Pedley TA, Lee JH, Barker-Cummings C, Lustenberger A, Nagle KJ, Lee KS, Scheuer ML, Neystat M, Susser M, Wilhelmsen KC. (1995) Localization of a gene for partial epilepsy to chromosome 10q. *Nat Genet* 10:56–60.

R Development Core Team. (2006) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Storey JD, Tibshirani R. (2003) Statistical significance for genomewide studies. *Proc Nat Acad Sci U S A* 100:9440–9445.

The International HapMap Consortium. (2005) A haplotype map of the human genome. *Nature* 437:1299–1320.

Vink JM, Boomsma DI. (2002) Gene finding strategies. *Biol Psychol* 61:53–71.

Epilepsia, 51(Suppl. 1):47–50, 2010
doi: 10.1111/j.1528-1167.2009.02445.x

CAPÍTULO 3

ARTIGO 3

A NEW CANDIDATE LOCUS FOR BILATERAL PERISYLVIAN POLYMICROGYRIA MAPPED ON CHROMOSOME Xq27

Neide F. Santos[1], Rodrigo Secolin[1], Iara L. Brandão-Almeida[1], Marilza S. Silva[1], Fábio R. Torres[1], Simone S. Tsuneda[1], Catarina A. Guimarães[2], Simone R.V. Hage[3], Fernando Cendes[2], Marilisa M. Guerreiro[2], Iscia Lopes-Cendes[1]

Departments of [1]Medical Genetics, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil; [2]Medical Neurology, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil; and [3] Speech Pathology, University of São Paulo (FOB), Bauru, SP, Brazil

American Journal of Medical Genetics Part A 146A:1151–1157 (2008)

A New Candidate Locus for Bilateral Perisylvian Polymicrogyria Mapped on Chromosome Xq27

Neide F. Santos,[1] Rodrigo Secolin,[1] Iara L. Brandão-Almeida,[1] Marilza S. Silva,[1] Fábio R. Torres,[1] Simone S. Tsuneda,[1] Catarina A. Guimarães,[2] Simone R.V. Hage,[3] Fernando Cendes,[2] Marilisa M. Guerreiro,[2] and Iscia Lopes-Cendes[1]*

[1]Department of Medical Genetics, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil
[2]Department of Medical Neurology, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil
[3]Department of Speech Pathology, University of São Paulo (FOB), Bauru, SP, Brazil

Received 16 April 2007; Accepted 17 December 2007

Polymicrogyria (PMG) is characterized by an excessive number of small and prominent brain gyri, separated by shallow sulci. Bilateral perisylvian polymicrogyria (BPP) is the most common form of PMG. Clinical signs include pseudobulbar paresis, mental retardation, and epilepsy. Familial forms of BPP have been described and a candidate locus was previously mapped to chromosome Xq28, distal do marker DXS8103. The objective of this study was to perform linkage analysis in one family segregating BPP. A total of 15 individuals, including 8 affected patients with BPP were evaluated. Family members were examined by a neurologist and subjected to magnetic resonance imaging scans. Individuals were genotyped for 18 microsatellite markers, flanking a 42.3 cM interval on ch Xq27-q28. Two-point and multipoint linkage analysis was performed using the LINKAGE package and haplotype reconstruction was performed by GENEHUNTER software. Our results showed a wide spectrum of clinical manifestations in affected individuals with BPP, ranging from normal to mild neurological abnormalities. Two-point linkage analysis yield a $Zmax = 2.06$ at $\theta = 0.00$ for markers DXS1205 and DXS1227. Multipoint lod-scores indicate a candidate interval of 13 cM between markers DSXS1205 and DXS8043, on ch Xq27.2-Xq27.3. These results point to a new locus for BPP in a more centromeric location than previously reported. © 2008 Wiley-Liss, Inc.

Key words: linkage analysis; X-linked inheritance; cortical malformation

How to cite this article: Santos NF, Secolin R, Brandão-Almeida IL, Silva MS, Torres FR, Tsuneda SS, Guimarães CA, Hage SRV, Cendes F, Guerreiro MM, Lopes-Cendes I. 2008. A new candidate locus for bilateral perisylvian polymicrogyria mapped on chromosome Xq27. Am J Med Genet Part A 146A:1151–1157.

INTRODUCTION

Polymicrogyria (PMG) is a cortical development malformation characterized by an excessive number of small and prominent brain gyri, separated by shallow sulci [Kuzniecky et al., 1993]. Bilateral perisylvian polymicrogyria (BPP; OMIM #300388) is the most common form of PMG, characterized by thickening of the cerebral cortex, around and on the depth of the Sylvian fissures. These abnormalities are often asymmetrical and can vary in extent among patients [Kuzniecky et al., 1993]. Clinical features include pseudobulbar paresis, causing restricted tongue movements, drooling, feeding problems, dysarthria and delay or difficult speech. Mental retardation and epilepsy may be present as well [Kuzniecky et al., 1993; Guerreiro et al., 2000; Montenegro et al., 2001].

Familial recurrence of BPP has been described and different patterns or inheritance have been proposed [Gropman et al., 1997; Yoshimura et al., 1998; Borgatti et al., 1999; Guerreiro et al., 2000; Villard et al., 2002]. A candidate locus was mapped on ch Xq28 [Villard et al., 2002] in five families with BPP. In addition, mutations in the *GPR56* gene were reported in patients with bilateral frontoparietal polymicrogyria and autosomal recessive inheritance [Piao et al., 2002, 2004] and recently, Roll et al. [2006]

N.F. Santos and R. Secolin contributed equally to this study.

Grant sponsor: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), SP, Brazil; Grant sponsor: Conselho Nacional de Pesquisa (CNPq), Brazil.

*Correspondence to: Iscia Lopes-Cendes, M.D., Ph.D., Department of Medical Genetics, FCM—UNICAMP, Tessália Vieira de Camargo, 126, Cidade Universitária Zeferino Vaz, Campinas, SP CEP 13084-971, Brazil. E-mail: icendes@unicamp.br

DOI 10.1002/ajmg.a.32270

found a mutation in the *SRPX2* gene (located on Xq22) in a male patient with BPP with severe seizures and mental retardation.

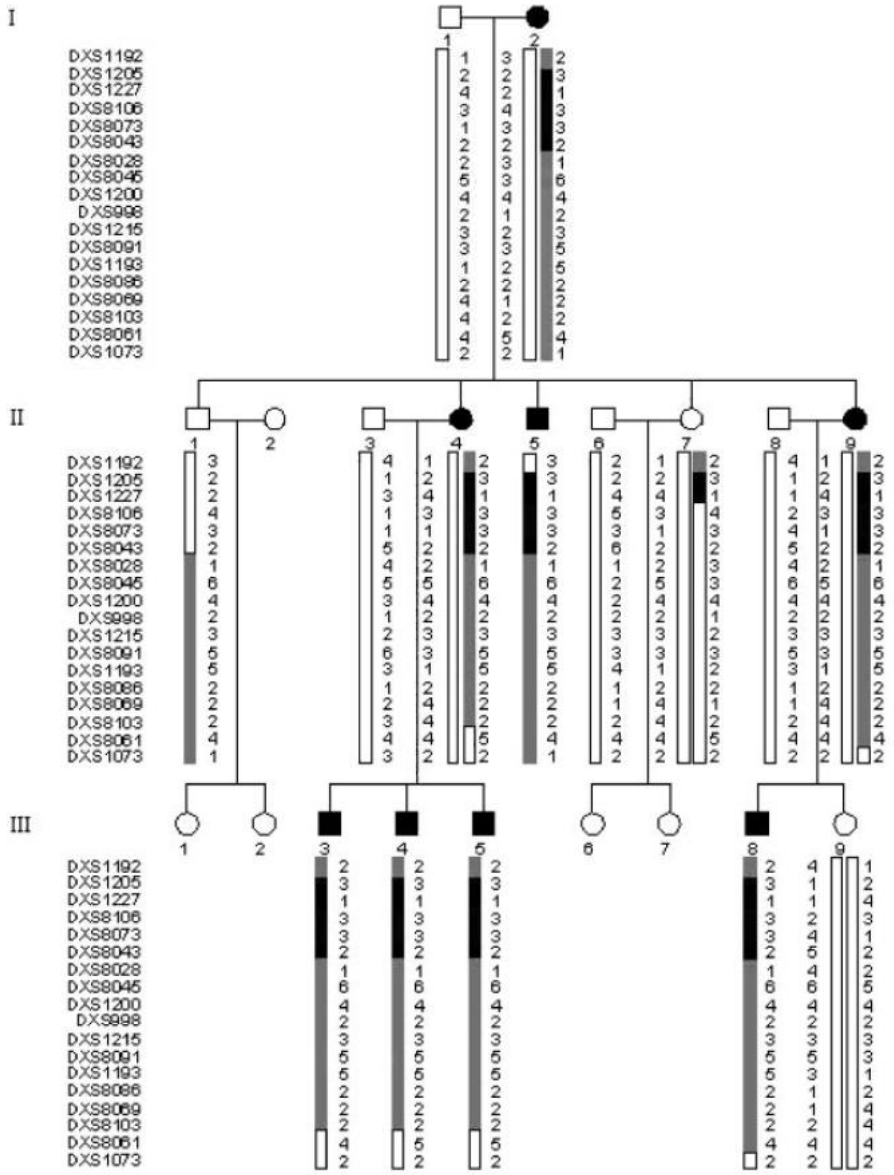
We identified a family with recurrence of BPP following an X-linked pattern of inheritance and report here on the results of clinical, neuroimaging, and genetic linkage studies

MATERIALS AND METHODS

We performed a detailed clinical investigation, imaging studies and linkage analysis in one family segregating BPP, including a total of 15 individuals. Among them, four were males married into the pedigree who were not included in the imaging studies but were clinically accessed and genotyped (Fig. 1). All 15 participating individuals gave informed consent and the study was approved by our Institution Research Ethics Committee.

Clinical and Language Evaluation

We systematically interviewed all patients and family members according to a standard detailed



questionnaire, emphasizing the history of problems with phonation and delayed speech, motor development and occurrence of prenatal events during the first two trimesters of pregnancy. All patients were examined by clinical neurologists. The neuropsychological assessment included the Wechsler Intelligence Scale for Children-III (WISC-III) or Wechsler Adult Intelligence Scale—Revised (WAIS-R) [Wechsler, 1981; Figueiredo, 2002]. Language evaluation was performed by a child speech therapist. Spontaneous language and free conversation were evaluated according to a semi-structured protocol that characterized the following aspects of language: phonologic, syntactic, semantic, pragmatic, and lexical [Guerreiro et al., 2002].

Magnetic Resonance Imaging Studies

Magnetic resonance imaging (MRI) scans were performed in a 2T scanner (Elscint Prestige®, Haifa, Israel), and included T1- and T2-weighted images in three orthogonal planes, as well as thin coronal T1 inversion recovery (IR) images. MRI visual analyses were performed using multiplanar reconstruction on OMNIPRO® workstation. According to MRI findings, individuals who clearly had BPP were classified as affected, whereas individuals with normal imaging studies were classified as unaffected.

Genotyping

Genomic DNA was obtained by direct extraction from lymphocytes of peripheric blood [Maniatis et al., 1989]. DNA samples were genotyped for 18 microsatellite markers: DXS1192, DXS1205, DXS1227, DXS8106, DXS8073, DXS8043, DXS8028, DXS8045, DXS1200, DXS998, DXS1215, DXS8091, DXS1193, DXS8086, DXS8069, DXS8103, DXS8061, and DXS1073, flanking a 42.3 cM interval on ch Xq27-q28. Microsatellite markers were selected from published data [Villard et al., 2002] and Généthon human genetic linkage map [Dib et al., 1996], available at the National Center of Biotechnology Information home page (http://www.ncbi.nlm.nih.gov).

PCR reactions were performed with 50 ng/μl of DNA, 5 μM of each oligonucleotide, 10× GeneAmp™ PCR Buffer II, 2.5 mM GeneAmp dNTP, 5 units/μl Taq™ DNA Polymerase and 25 mM $MgCl_2$ (Invitrogen, Carlsbad, CA). Sense oligonucleotides were fluorochrome labeled with FAM (6-carboxyfluorescein), VIC™ or NED (phosphoramidite fluorescein; Applied Biosystems, Foster City, CA). PCR products of the microsatellite markers labeled with fluorochromes were analyzed in the MegaBACE 1000™ automatic sequencer (GE Heathcare, Buckinghamshire, UK). Results were analyzed using Fragment Profiler Software™ (GE Heathcare).

Statistical Analysis

Data obtained from Fragment Profiler Software™ was processed to input files for linkage analysis by the LINKGEN program. This software was developed by our laboratory and is available online at http://lgm.fcm.unicamp.br:9001/cgi-bin/linkgen/linkgen.cgi for free use.

Two-point and multipoint genetic linkage analysis was calculated by the maximum likelihood method using the computer program MLINK© and LINKMAP© (version 5.2) (CEPH, University of Utah, and Columbia University 1990), respectively, from the LINKAGE© package [Lathrop and Lalouel, 1984; Terwillinger and Ott, 1993]. The GENEHUNTER© program was used for haplotype reconstruction [Kruglyak et al., 1996]. We assumed an X-linked dominant mode of inheritance with 0.8 penetrance. Allelic frequencies for each marker was calculated from unrelated individuals married into the pedigree

TABLE I. Demographic and Clinical Data on 15 Individuals Studied in One BPP Family

Pedigree number	Gender	Neurological examination	Language acquisition	PIQ/VIQ	MRI	Carrier of affected haplotype
I-1	Male	Normal	Normal	NE	NE	No
I-2	Female	Normal	Normal	102/83	BPP	Yes
II-1	Male	Normal	Normal	NE	Normal	No
II-3	Male	Normal	Normal	NE	NE	No
II-4	Female	Normal	Normal	93/80	BPP	Yes
II-5	Male	PBP	Delayed	79/NO	BPP	Yes
II-6	Male	Normal	Normal	NE	NE	No
II-7	Female	Normal	Normal	NE	Normal	No
II-8	Male	Normal	Normal	NE	NE	No
II-9	Female	Normal	Normal	105/91	BPP	Yes
III-3	Male	PBP	Delayed	74/57	BPP	Yes
III-4	Male	Normal	Delayed	95/97	BPP	Yes
III-5	Male	Normal	Delayed	97/93	BPP	Yes
III-8	Male	Normal	Delayed	107/97	BPP	Yes
III-9	Female	Normal	Normal	NE	Normal	No

PIQ, performance intelligence quotient; VIQ, verbal intelligence quotient; MRI, magnetic resonance imaging; PBP, pseudobulbar paresis; NE, not evaluated; NO, not obtained; BPP, bilateral perisylvian polymicrogyria.

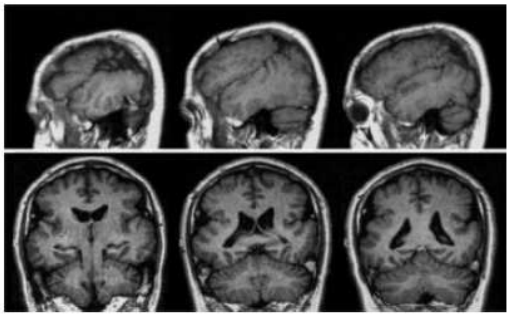


FIG. 2. MRI from Patient II-5 (male). T1-weighted sagittal and coronal images showing bilateral perisylvian polymicrogyria extending to both frontal and parietal regions.

and controls. Disease allele frequency was 0.0001 [Villard et al., 2002]. Lod score values (Z) equal or higher than 2.00 indicate genetic linkage between marker and disease [Terwillinger and Ott, 1993].

RESULTS

A summary of clinical information is given in Table I. We accessed a total of 15 individuals, including 5 females (Fig. 1). Clinical spectrum ranged from normal to mild neurological dysfunction, including pseudobulbar paresis, such as poor articulation and poor tongue movements. Overall, the MRIs showed a different pattern between males and females, with more pronounced and widespread polymicrogyric cortex in males (Figs. 2–5). Among the five affected males, two individuals (II-5, Fig. 2 and III-3, Fig. 3) had a more severe phenotype, with pseudobulbar paresis and delay in language acquisition (Table I). These individuals also had more severe MRI abnormalities with BPP extending over both frontal and parietal areas; whereas, the affected females (individuals I-2, Fig. 4; II-4, Fig. 5; and II-9) had the least severe MRI findings with asymmetrical and milder perisylvian polymicrogyria.

Significant two-point lod scores were observed for two markers, with a $Zmax = 2.06$ at $\theta = 0.00$ for markers DXS1205 and DXS1227 (Table II). Multipoint lod scores pointed to a candidate region of 13 cM between markers DXS1205 and DXS8043, located on cr Xq27.2-Xq27.3 (Fig. 6).

All eight patients with BPP on MRI had the affected haplotype (Table I, Fig. 1); however, findings in neurological and language evaluations were variable. Individuals I-1, II-3, II-6, and II-8 were not included in the MRI studies, but had no neurological complains or speech abnormalities (they were not formally tested by the speech therapist). Individuals

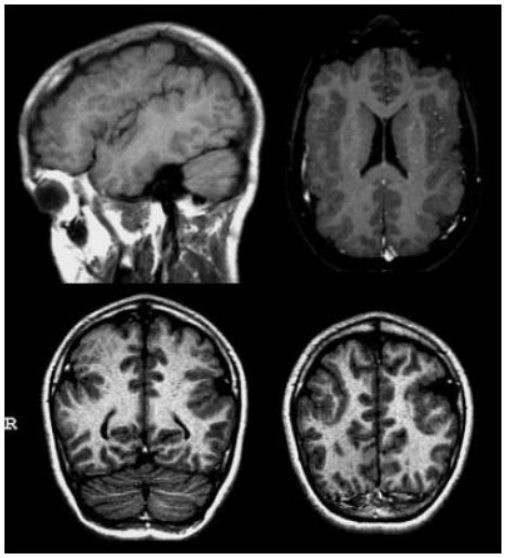


FIG. 3. MRI from Patient III-3 (male). T1-weighted sagittal (**top left**), axial (**top right**), and coronal (**bottom**) images showing bilateral perisylvian polymicrogyria extending to both frontal and parietal regions.

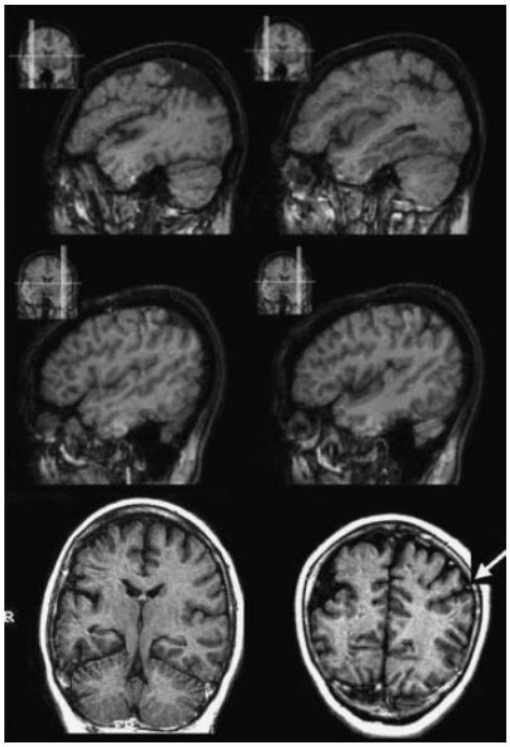


FIG. 4. MRI from Patient I-2 (female). T1-weighted sagittal and coronal images show asymmetrical bilateral perisylvian polymicrogyria more pronounced in right hemisphere (**top row**). The mild polymicrogyria on the left hemisphere is present in the posterior extension of the Sylvian fissure which is better appreciated in the coronal images (arrow).

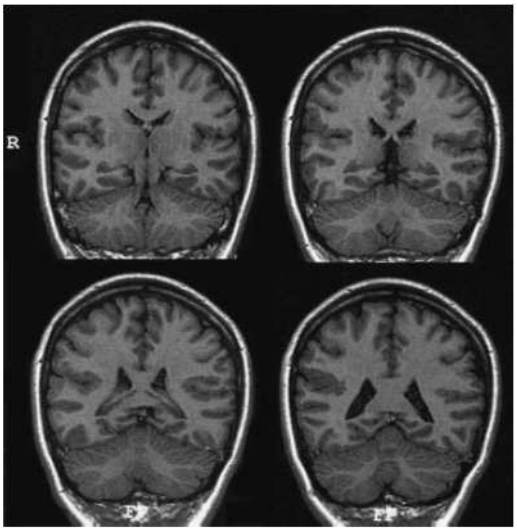


FIG. 5. MRI from Patient II-4 (female). Coronal T1 images showing mild bilateral asymmetrical perisylvian polymicrogyria (more evident on the right side) which spares frontal and parietal lobes.

II-1, III-9 had normal neurological exam, normal language acquisition normal MRI and do not carry the affected haplotype. Although individuals I-2, II-4, and II-9 had normal neurological exam and language acquisition, MRI clearly showed BPP and haplotype analysis revealed that they have the affected haplotype. Individuals II-5 and III-3 had pseudobulbar paresis and individuals III-4, III-5, and III-8 had normal neurological exams; however, all five had delayed language acquisition, MRI with BPP and the affected haplotype (Table I, Fig. 1). The only individual who carries the affected haplotype but does not have clinical or MRI abnormalities is II-7, a female who had normal development and has no neurological or speech abnormalities (Table I, Fig. 1). Unfortunately, we could not obtain a complete clinical, MRI or haplotype information on her two daughters (individuals III-6 and III-7), since the mother (individual II-7) did not agree to have the girls participate in the study; however, the mother informed us by a telephone interview that both girls (ages 19 and 8 years old) are presently asymptomatic and have had no developmental or speech problems.

DISCUSSION

Bilateral persylvian polymicrogyria is a brain abnormality currently classified under malformations due to abnormal cortical organization (including late neuronal migration) [Van Bogaert et al., 1998; Barkovich et al., 2005]. Genetic contribution to the development of BPP in some patients is supported by reports of familial cases, which suggests that gene mutations may cause this brain anomaly [Bartolomei et al., 1999; Borgatti et al., 1999; Caraballo et al., 2000; Guerreiro et al., 2000; Roll et al., 2006]. We identified one family segregating BPP, showing a possible X-linked dominant inheritance and in the present study we showed evidence for linkage to ch Xq27.2-Xq27.3 within a 13 cM candidate region.

In the family studied here, the matriarch individual I-2) who was normal on clinical evaluation had three affected children (II-4, II-5, and II-9) and four affected grandchildren (III-3, III-4, III-5, and III-8) with clinical and MRI abnormalities. Surprisingly, her MRI revealed BPP (Fig. 4). Furthermore, all eight individuals (clinically unaffected matriarch, three affected children and four affected grandchildren)

TABLE II. Two-Point Lod Scores for 18 Markers on Chromosome Xq27-28

	Recombinant fractions (θ)								
Markers	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
DXS1192	−4.64	0.61	0.76	0.79	0.75	0.68	0.59	0.47	0.33
DXS1205	2.06	1.89	1.71	1.52	1.32	1.10	0.88	0.65	0.42
DXS1227	2.06	1.89	1.71	1.52	1.32	1.10	0.88	0.65	0.42
DXS8106	1.03	0.93	0.82	0.71	0.60	0.47	0.35	0.22	0.11
DXS8073	1.33	1.23	1.12	1.01	0.90	0.77	0.64	0.50	0.34
DXS8043	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DXS8028	−0.64	0.89	1.0	0.98	0.91	0.80	0.67	0.52	0.36
DXS8045	−0.64	0.89	1.0	0.98	0.91	0.80	0.67	0.52	0.36
DXS1200	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DXS998	−1.97	−0.34	−0.12	−0.03	0.01	0.03	0.03	0.02	0.01
DXS1215	−1.97	−0.34	−0.12	−0.03	0.01	0.03	0.03	0.02	0.01
DXS8091	−0.64	0.89	1.0	0.98	0.91	0.80	0.67	0.52	0.36
DXS1193	−0.64	0.89	1.0	0.98	0.91	0.80	0.67	0.52	0.36
DXS8086	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DXS8069	−0.64	0.89	1.0	0.98	0.91	0.80	0.67	0.52	0.36
DXS8103	1.33	1.23	1.12	1.01	0.90	0.77	0.64	0.50	0.34
DXS8061	−14.04	−1.88	−1.07	−0.65	−0.38	−0.20	−0.08	−0.01	0.03
DXS1073	−5.97	−0.62	−0.36	−0.23	−0.14	−0.09	−0.05	−0.03	−0.01

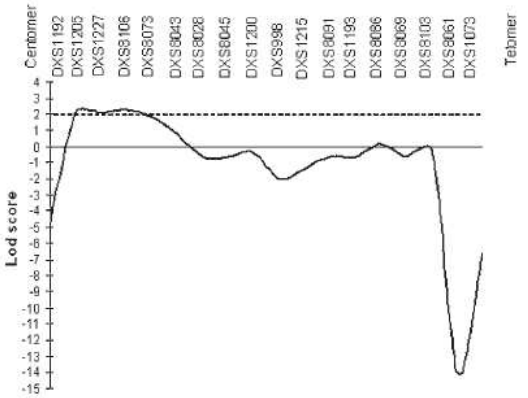


FIG. 6. Multipoint lod scores calculated for the 18 markers genotyped on chromosome Xq27-28. Dashed line indicates lod score threshold for linkage.

shared the same affected haplotype (Table I; Fig. 1). These findings point to the great importance of imaging studies in clinically unaffected individuals, especially females, belonging to families segregating BPP. In addition, we also found evidence for incomplete penetrance, since individual II-7 carries at least part of the affected haplotype (from markers DXS1205 to DXS1227) but had normal development, no neurological or speech abnormities and normal MRI. Incomplete penetrance has been previously reported in families with BPP and may result from the lack of sensitivity of clinically used imaging techniques to detect subtle brain abnormalities [Guerreiro et al., 2000; Jansen and Andermann, 2005].

Our linkage study focused initially at region Xq28 due to the reports of Villard et al. [2002] describing a candidate locus for BPP downstream to microsatellite marker DXS8103. However, the region of interest was extended to a more centromeric location, since the region telomeric to DXS8043 was excluded by haplotype analysis (Fig. 1). Our results point to a 13 cM candidate region for BPP in this family between markers DXS1205 and DXS8043 as defined by multipoint linkage analysis, on ch Xq27.2-q27.3. Our candidate locus is approximately 16 cM upstream marker DXS8103 and therefore, does not overlap the candidate region previously reported. The fact that two different candidate regions for BPP were found could be an evidence of locus heterogeneity and, since cortical development is a very complex molecular mechanism with the potential involvement of several different genes, we cannot exclude the possibility that the distal region on the long arm of ch X could harbor a cortical development genes cluster, with some of these genes involved in the etiology of BPP.

In conclusion, we mapped a candidate locus for BPP on ch Xq27.2-q27.3 which does not overlap with the candidate region previously mapped [Villard et al., 2002]. Our family demonstrated considerable variability of clinical expression with only males showing pseudobulbar paresis and low cognitive scores. We recommend detailed clinical evaluation, including examination by speech pathologists and MRI scans in all related family members for the precise phenotypic classification in linkage studies of BPP.

ACKNOWLEDGMENTS

We are grateful to the patients' family members for their helpful cooperation. NFS received a fellowship from CNPq. RS, ILBA, FRT, and SST received studentships from FAPESP.

REFERENCES

Barkovich AJ, Kuzniecky RI, Jackson GD, Guerrini R, Dobyns WB. 2005. A developmental and genetic classification for malformations of cortical development. Neurology 65:1873–1887.

Bartolomei F, Gavaret M, Dravet C, Guerrini R. 1999. Familial epilepsy with unilateral and bilateral malformations of cortical development. Epilepsia 40:47–51.

Borgatti R, Triulzi F, Zucca C, Piccinelli P, Balottin U, Carrozzo R, Guerrini R. 1999. Bilateral perisylvian polymicrogyria in three generations. Neurology 52:1910–1913.

Caraballo RH, Cersosimo RO, Mazza E, Fejerman N. 2000. Focal polymicrogyria in mother and son. Brain Dev 22:336–339.

Dib C, Fauré S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Selboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J. 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154.

Figueiredo VLM. 2002. WISC III: Wechsler Intelligence Scale for Children. Adapted in a Brazilian sample. São Paulo: Casa do Psicólogo.

Gropman AL, Barkovich AJ, Vezina LG, Conry JA, Dubovsky EC, Packer RJ. 1997. Pediatric congenital bilateral perisylvian syndrome: Clinical and MRI features in 12 patients. Neuropediatrics 28:198–203.

Guerreiro MM, Andermann E, Guerrini R, Dobyns WB, Kuzniecky R, Silver K, Van Bogaert P, Gillain C, David P, Ambrosetto G, Rosati A, Bartolomei F, Parmeggiani A, Paetau R, Salonen O, Ignatius J, Borgatti R, Zucca C, Bastos AC, Palmini A, Fernandes W, Montenegro MA, Cendes F, Andermann F. 2000. Familial perisylvian polymicrogyria: A new familial syndrome of cortical maldevelopment. Ann Neurol 48:39–48.

Guerreiro MM, Hage SRV, Guimarães CA, Abramides DV, Fernandes W, Pacheco PS, Piovesana AMSG, Montenegro MA, Cendes F. 2002. Developmental language disorder associated with polymicrogyria. Neurology 59:246–250.

Jansen A, Andermann E. 2005. Genetics of the polymicrogyria syndromes. J Med Genet 42:369–378.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. 1996. Parametric and nonparametric linkage analysis: A unified multipoint approach. Am J Hum Genet 58:1347–1363.

Kuzniecky R, Andermann F, Guerrini R. 1993. Congenital bilateral perisylvian syndrome: Study of 31 patients. The CBPS Multicenter Collaborative Study. Lancet 341:608–612.

Lathrop GM, Lalouel JM. 1984. Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36:460–465.

Maniatis T, Fritsch EF, Sambrook J. 1989. Molecular cloning: A laboratory manual. New York: Cold Spring Harbor Laboratory Press.

Montenegro MA, Guerreiro MM, Lopes-Cendes I, Cendes F. 2001. Bilaterial Posterior Parietal Polymicrogiria: A mild form of Congenital Bilateral Perisylvian Syndrome? Epilepsia 42:845–849.

Piao X, Basel-Vanagaite L, Straussberg R, Grant PE, Pugh EW, Doheny K, Doan B, Hong SE, Shugart YY, Walsh CA. 2002. An autosomal recessive form of bilateral frontoparietal polymicrogyria maps to chromosome 16p12.2-21. Am J Hum Genet 70:1028–1033.

Piao X, Hill RS, Bodell A, Chang BS, Basel-Vanagaite L, Straussberg R, Dobyns WB, Qasrawi B, Winter RM, Innes AM, Voit T, Ross ME, Michaud JL, Déscarie JC, Barkovich AJ, Walsh CA. 2004. G protein-coupled receptor-dependent development of human frontal cortex. Science 303:2033–2036.

Roll P, Rudolf G, Pereira S, Royer B, Scheffer IE, Massacrier A, Valenti MP, Roeckel-Trevisiol N, Jamali S, Beclin C, Seegmuller C, Metz-Lutz MN, Lemainque A, Delepine M, Caloustian C, de Saint Martin A, Bruneau N, Depetris D, Mattei MG, Flori E, Robaglia-Schlupp A, Levy N, Neubauer BA, Ravid R, Marescaux C, Berkovic SF, Hirsch E, Lathrop M, Cau P, Szepetowski P. 2006. SRPX2 mutations in disorders of language cortex and cognition. Hum Mol Genet 15:1195–1207.

Terwillinger JD, Ott J. 1993. Handbook of human genetic linkage. Maryland: Johns Hopkins.

Van Bogaert P, David P, Gillain CA, Wikler D, Damhaut P, Scalais E, Nuttin C, Wetzburger C, Szliwowski HB, Metens T, Goldman S. 1998. Perisylvian dysgenesis: Clinical, EEG, MRI and glucose metabolism features in 10 patients. Brain 121: 2229–2238.

Villard L, Nguyen K, Cardoso C, Martin CL, Weiss AM, Sifry-Platt M, Grix AW, Graham JM Jr, Winter RM, Leventer RJ, Dobyns WB. 2002. A locus for Bilateral Perisylvian Polymicrogyria maps to Xq28. Am J Hum Genet 70:1003–1008.

Wechsler D. 1981. Wechsler Adult Intelligence Scale—Revised. San Antonio: The Psychological Corporation.

Yoshimura K, Hamada F, Tomoda T, Wakiguchi H, Kurashige T. 1998. Focal pachypolymicrogyria in three siblings. Pediatr Neurol 18:435–438.

CAPÍTULO 4

ARTIGO 4

FAMILY-BASED ASSOCIATION STUDY FOR BIPOLAR AFFECTIVE DISORDER

Rodrigo Secolin[1], Cláudio E.M. Banzato[2], Maria C.M. Oliveira[2], Maria F.R. Bittar[1], Marilza L. Santos[1], Paulo Dalgalarrondo[2], Iscia Lopes-Cendes[1]

Departments of [1]Medical Genetics and [2]Medical Psychology and Psychiatry,
University of Campinas (UNICAMP), Campinas, São Paulo, Brazil

Family-based association study for bipolar affective disorder

Rodrigo Secolin[a], Cláudio E.M. Banzato[b], Maria C.M. Oliveira[b], Maria F.R. Bittar[a], Marilza L. Santos[a], Paulo Dalgalarrondo[b] and Iscia Lopes-Cendes[a]

In this study we aimed to evaluate 21 candidate loci for bipolar affective disorder (BPAD) identified earlier in a large genome-wide association study. We evaluated 74 pedigrees with BPAD, with a total of 411 individuals, including 96 patients who fulfilled clinical criteria for BPAD according to *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition classification. Family-based association analysis was performed using the UNPHASED software. We identified a single nucleotide polymorphism (rs9834970) localized on chromosome 3p22.3, showing statistically significant association with BPAD after the Bonferroni correction for multiple comparisons ($P_{corrected} = 0.0025$) with an odds ratio = 2.64 (95% confidence interval: 1.30–5.35). Single nucleotide polymorphism rs9834970 is located in an intergenic region and is not known to be associated to regulatory genomic sequences. *Psychiatr Genet* 20:126–129 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Psychiatric Genetics 2010, 20:126–129

Keywords: genetics, mood disorders, pedigree disequilibrium test, single nucleotide polymorphisms

Departments of [a]Medical Genetics and [b]Medical Psychology and Psychiatry, University of Campinas-UNICAMP, Campinas, Sao Paulo, Brazil

Correspondence to Iscia Lopes-Cendes, MD, PhD, Department of Medical Genetics, FCM – UNICAMP, Tessália Vieira de Camargo, 126, Cidade Universitária Zeferino Vaz, Campinas SP, CEP 13084-971, Brazil
Tel: +55 19 3521 8907; fax: +55 19 3289 1818; e-mail: icendes@unicamp.br

Received 26 May 2009 Revised 19 November 2009
Accepted 24 November 2009

Introduction

Family, twin and adoption studies clearly showed the contribution of heritable factors in bipolar affective disorder [BPAD (OMIM # 125480)] (Mcguffin and Katz, 1989; Tsuang and Faraone, 1990; Maier *et al.*, 1993; Spence *et al.*, 1995; Alda, 1999; Kato *et al.*, 2005). Indeed, many candidate loci and candidate genes have been reported in BPAD, but few of these studies have been replicated in different population (Kato, 2007; Kaneva *et al.*, 2008; McAuley *et al.*, 2008; Serretti and Mandelli, 2008; Shi *et al.*, 2008; Barnett and Smoller, 2009). The Wellcome Trust Case Control Consortium (2007) reported 21 candidate loci with strong or moderate evidence of association to BPAD in the British population. In an attempt to replicate these results in an ethnic diverse population, we evaluate these same candidate loci using a family-based association approach.

Methods

Ascertainment of patients

We ascertained unrelated patients with BPAD in three medical centers: a general teaching hospital at the University of Campinas, a psychiatric hospital (Bairral Institute of Psychiatry) and a community mental health center (Paulinia Public Health Care System), all located in São Paulo, Brazil. Detailed family histories were obtained from all probands and available family members, and pedigrees were constructed for all patients included in the study. Each proband and all available relatives were interviewed by the same psychiatrist (MCMO) using the Mini International Neuropsychiatric Interview Plus (Sheehan *et al.*, 2000). This study was approved by the research ethics committee of our institution and all patients and family members signed a consent form before entering the study.

We considered affected all individuals with bipolar I or bipolar II disorder according to *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV) criteria (American Psychiatry Association, 1994). Among 411 individuals distributed in 74 families there were 96 patients with bipolar I disorder. Interestingly, we did not find probands or any family member with bipolar II disorder. We considered unaffected, individuals (307) without BPAD, which included 280 individuals without any psychiatric disorder, one with schizophrenia, one individual with schizoaffective disorder and 25 with depressive disorder. In addition, eight individuals were not evaluated and therefore classified as unknown.

Genotyping

Genomic DNA was obtained by direct extraction from lymphocytes of peripheral blood according to standard procedures (Maniatis *et al.*, 1989). DNA samples were quantified using NanoVue V1.7.2 Spectrophotometer (GE Healthcare, Chicago, Illinois, USA).

We chose 21 single nucleotide polymorphisms (SNPs), 20 autosomal and one on the X chromosome, with strong and moderate evidence of association with BPAD according to the supplementary information published

0955-8829 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins

DOI: 10.1097/YPG.0b013e32833a2050

by The Wellcome Trust Case Control Consortium (2007). SNPs were genotyped by real-time PCR using the ABI 7500 TaqMan System (Applied Biosystems, Foster City, California, USA) and fluorescence signals were detected by the 7500 System Sequence Detection Software (Applied Biosystems).

Statistical analysis

SNP genotypes obtained from the System Sequence Detection Software were processed to input files for statistical analysis using the JINGLEFIX software, which also estimated minor allele frequencies (threshold = 0.05) from processed data (Secolin *et al.*, 2008). Mendelian inconsistencies and Hardy–Weinberg equilibrium (*P* value threshold = 0.01) were evaluated by PEDCHECK and HAPLOVIEW software, respectively (O'Connell and Weeks, 1998; Barrett *et al.*, 2005). We used the UNPHASED program for family-based association analysis (Dudbridge, 2008) and Bonferroni correction to account for multiple testing. To verify statistical power of our sample, we used the TDT POWER CALCULATOR program, as this software enables estimation from pedigrees with different family structures (Chen and Deng, 2001). The parameters used in the TDT POWER CALCULATOR were: dominant pattern of inheritance with incomplete penetrance, marker frequency = 0.10, allele frequency = 0.25, and a Bonferroni adjusted statistical significance level $\alpha = 0.0024$.

Results

Descriptive statistics

Family investigation revealed 191 males (46.5%), 220 females (53.5%) with a mean age of 46.3 ± 16.4 years. Among patients, 34 (35.4%) had both parents genotyped, whereas 42 patients (43.7%) had at least one parent genotyped. SNP average genotype call rate was 99% and average genotyping success rate of sample patients was also 99%. Statistical power for our family sample was higher than 80% for the detection of association.

Statistical analysis

As shown in Table 1, SNPs rs2609653 and rs10134944 presented minor allele frequency of less than 0.05 and SNP rs10134944 was not in Hardy–Weinberg equilibrium ($P < 0.01$); therefore, these SNPs were dropped from further analysis. We found significant association signals for SNPs rs2989476 ($P = 0.0031$), rs7570682 ($P = 0.0407$), rs9834970 ($P = 0.00013$) and rs11622475 ($P = 0.0153$). However, only SNP rs9834970 maintained statistical significance for association after Bonferroni correction ($P_{corrected} = 0.0025$), with an odds ratio = 2.64 (95% confidence interval: 1.30–5.35). In addition, we found that allele C of SNP rs9834970 was over transmitted from parents to patients (15 : 1).

Discussion

In 2007, The Wellcome Trust Case Control Consortium (2007) published 21 candidate loci with strong or moderate evidence of association to BPAD in the British population. This initial large-scale association study was followed by a meta-analysis study showing several additional important chromosomal regions associated with BPAD (Baum *et al.*, 2008). Subsequently, strong signals for association were also found in myosin 5B (*MYO5B*), tetraspanin-8 (*TSPAN8*), epidermal growth factor receptor (*EGFR*), ankyrin G (*ANK3*) and α 1C subunit of L-type voltage-gated channel (*CACNA1C*) genes, again in the British population (Ferreira *et al.*, 2008; Sklar *et al.*, 2008). Among these

Table 1 Family-based association results for 21 SNPs in bipolar affective disorder

					P value		Allele effect estimation		
dbSNP	Locus	MAF	HWE	χ^2	Nominal	Corrected	Allele	OR	95% CI
rs2989476	1p31.3	0.447	1.000	8.708	0.0031	0.0589	G	2.15	1.15–4.03
rs4027132	2p25.1	0.381	0.352	1.828	0.1764	1.0000	G	1.15	0.64–2.06
rs7570682	2q12.1	0.207	1.000	4.186	0.0407	0.7733	G	1.51	0.79–2.89
rs1375144	2q14.1	0.342	0.025	2.318	0.1279	1.0000	A	1.68	0.95–2.97
rs11888446	2q31.3	0.161	0.417	0.000	1.0000	1.0000	C	0.86	0.44–1.68
rs4673905	2q33.1	0.230	0.224	3.397	0.0653	1.0000	G	0.89	0.47–1.68
rs2953145	2q37.3	0.170	0.048	0.402	0.5257	1.0000	G	0.88	0.46–1.69
rs4276227	3p23	0.361	0.578	1.699	0.1924	1.0000	C	1.09	0.59–2.02
rs9834970	3p22.3	0.382	0.436	14.700	0.00013	0.0025	C	2.64	1.30–5.35
rs683395	3q27.1	0.093	1.000	0.091	0.7629	1.0000	C	1.86	0.82–4.21
rs6458307	6p21.1	0.404	0.895	0.040	0.8415	1.0000	C	1.23	0.70–2.16
rs6901299	6q22.31	0.101	0.048	0.505	0.4772	1.0000	A	1.48	0.67–3.25
rs1405318	7p21.3	0.105	0.958	1.328	0.2491	1.0000	G	1.44	0.58–3.52
rs2609653[a]	8p12	0.038	0.034	–	–	–	–	–	–
rs10982256	9q32	0.481	0.968	1.096	0.2952	1.0000	C	0.97	0.56–1.68
rs10134944[a]	14q22.3	0.022	1.4e-6	–	–	–	–	–	–
rs11622475	14q32.33	0.167	0.649	5.884	0.0153	0.2907	C	1.53	0.84–2.81
rs420259	16p12.1	0.291	0.298	0.000	1.0000	1.0000	C	0.77	0.40–1.29
rs1344484	16q12.2	0.423	0.063	2.379	0.1230	1.0000	C	1.11	0.62–1.98
rs3761218	20p13	0.398	1.000	0.617	0.4319	1.0000	C	1.13	0.67–1.91
rs975687	Xq23	0.095	0.032	1.386	0.2390	1.0000	T	5.94	0.49–71.51

χ^2, chi square; CI, confidence interval; HWE, Hardy–Weinberg equilibrium; MAF, minor allele frequency; OR, odds ratio; SNP, single nucleotide polymorphisms.
[a]These SNPs were not evaluated due to MAF <0.05 and HWE status.

genes, association with *ANK3* and BPAD has been replicated in further studies (Scott *et al.*, 2009; Smith *et al.*, 2009). In addition, Scott *et al.* (2009) also found a strong signal for association on chromosome 3p21. Linkage studies have also been performed in BPAD and identified several different loci (Fullerton *et al.*, 2008; Kaneva *et al.*, 2008; McAuley *et al.*, 2008; Ross *et al.*, 2008). However, many of the association and linkage studies performed to date on BPAD show different and even contradictory results (Kato, 2007; Barnett and Smoller, 2009). Differences in phenotype definition, sample ancestry and study design are key elements to explain these apparent discrepancies. Therefore, replication studies in different samples and using different mapping strategies are of utmost importance to confirm findings reported earlier.

In an attempt to replicate the results reported by The Wellcome Trust Case Control Consortium (2007) in an ethnic diverse population, we evaluated the same 21 candidate loci identified by the original study using a family-based association approach. Our family sample was found to have significant power to detect association. However, there were important differences in the way patients were ascertained in the two studies. The Wellcome Trust Case Control Consortium (2007) used the research diagnostic criteria (Spitzer *et al.*, 1978) for phenotype definition and performed a case–control study in a population-based sample. In our study we used DSM-IV diagnostic criteria (American Psychiatry Association, 1994) and a family-based sample. In fact, Barnett and Smoller (2009) suggested that different specific susceptibility genes for BPAD could be identified depending on sampling – family based versus population based. In addition, unlike The Wellcome Trust Case Control Consortium (2007), we did not include schizoaffective disorder as a bipolar subtype, as we found only one individual with this phenotype. One advantage of our study design is the absence of problems with sample stratification, as family-based association studies consider the association within but not between families, making it unnecessary for the use of genomic controls (Dudbridge, 2008). By contrast, one limitation of our study is that as association was achieved in a sample with familial aggregation of cases, the significance of this association for patients without familial recurrence of the disease is unclear.

Except for SNP rs9834970, we did not find significant association for any other SNP analyzed, including SNP rs420259 which presented the highest association signal according to The Wellcome Trust Case Control Consortium (2007). Furthermore, despite the fact that finding four SNPs with significant signals is more than expected by chance after performing 21 tests, this amount of analyses can inflate the false-positive rate; therefore, a statistical correction for multiple comparisons was performed in our study.

Although SNP rs9834970 is located in an intergenic region and is not known to be associated with regulatory genomic sequences, we found four candidate genes located nearby which are related to nerve growth factors (*ITAG9* gene) (Staniszewska *et al.*, 2008) and voltage-gated sodium channels (*SCN5A*, *SCN10A* and *SCN11A* genes).

In conclusion, we showed that SNP rs9834970, localized on chromosome 3p22.3, is associated with the disease phenotype in BAPD pedigrees. Our results strengthened earlier findings of association between this SNP and BPAD reported by The Wellcome Trust Case Control Consortium (2007). Additional SNPs should be genotyped in the candidate region to estimate haplotype association and to identify the putative susceptibility variant associated with BPAD in this group of families.

Acknowledgements

The authors are grateful to all patients and family members for their helpful cooperation. This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP-grant # 08/55827-9), SP, Brazil. Rodrigo Secolin and Maria C.M. Oliveira received studentships from FAPESP (grants # 06/57332-1 and # 02/06831-7, respectively).

References

Alda M (1999). Bipolar disorder: from families to genes. *Can J Psychiatry* **42**:378–387.

American Psychiatry Association (1994). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Press.

Barnett JH, Smoller JW (2009). The genetics of bipolar disorder. *Neuroscience* **164**:331–343.

Barrett JC, Fry B, Maller J, Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**:263–265.

Baum AE, Hamshere M, Green E, Cichon S, Rietsche M, Noethen MM, *et al.* (2008). Meta-analysis of two genome-wide association studies of bipolar disorder reveals important points of agreement. *Mol Psychiatry* **13**:466–467.

Chen WM, Deng HW (2001). A general and accurate approach for computing the statistical power of the transmission disequilibrium test for complex disease genes. *Genetic Epidemiol* **21**:53–67.

Dudbridge F (2008). Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* **66**:87–98.

Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, *et al.* (2008). Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet* **40**:1056–1058.

Fullerton JM, Liu Z, Badenhop RF, Scimone A, Blair IP, Van Herten M, *et al.* (2008). Genome screen of 15 Australian bipolar affective disorder pedigrees supports previously identified loci for bipolar susceptibility genes. *Psychiatr Genet* **18**:156–161.

Kaneva R, Milanova V, Angelicheva D, Macgregor S, Kostov C, Vladimirova R, *et al.* (2008). Bipolar disorder in the Bulgarian gypsies: genetic heterogeneity in a young founder population. *Am J Med Genet B Neuropsychiatr Genet* **150B**:191–201.

Kato T (2007). Molecular genetics of bipolar disorder and depression. *Psychiatry Clin Neurosci* **61**:3–19.

Kato T, Kuratomi G, Kato N (2005). Genetics of bipolar disorder. *Drugs Today (Barc)* **41**:335–344.

Maier W, Lichtermann D, Minges J, Hallmayer J, Heun R, Benkert O, *et al.* (1993). Continuity and discontinuity of affective disorder and schizophrenia: results of a controlled family study. *Arch Gen Psychiatry* **50**:871–883.

Maniatis T, Fritsch EF, Sambrook J (1989). *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.

McAuley EZ, Blair IP, Liu Z, Fullerton JM, Scimone A, Van Herten M, *et al.* (2008). A genome screen of 35 bipolar affective disorder pedigrees provides

significant evidence for a susceptibility locus on chromosome 15q25–26. *Mol Psychiatry* **14**:492–500.
Mcguffin P, Katz R (1989). The genetics of depression and manic-depressive disorder. *Br J Psychiatry* **155**:294–304.
O'Connell JR, Weeks DE (1998). PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* **63**: 259–266.
Ross J, Berrettini W, Coryell W, Gershon ES, Badner JA, Kelsoe JR, *et al.* (2008). Genome-wide parametric linkage analyses of 644 bipolar pedigrees suggest susceptibility loci at chromosomes 16 and 20. *Psychiatr Genet* **18**:191–198.
Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, *et al.* (2009). Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proc Natl Acad Sci U S A* **106**:7501–7506.
Secolin R, Rocha CS, Torres FR, Santos ML, Maurer-Morelli CV, Santos NF, *et al.* (2008). LINKGEN: a new algorithm to process data in genetic linkage studies. *Genomics* **91**:544–547.
Serretti A, Mandelli L (2008). The genetics of bipolar disorder: genome hot regions, genes, new potential candidates and future directions. *Mol Psychiatry* **13**:742–771.
Sheehan D, Lecrubier Y, Harnett-Sheehan K, Janavs J, Weiller E, Bonora I, *et al.* (2000). *Mini International neuropsychiatric interview plus*. Florida: University of South Florida.
Shi J, Badner JA, Liu C (2008). PDLIM5 and susceptibility to bipolar disorder: a family-based association study and meta-analysis. *Psychiatr Genet* **18**:116–121.
Sklar P, Smoller JW, Fan J, Ferreira MAR, Perlis RH, Chambert K, *et al.* (2008). Whole-genome association study of bipolar disorder. *Mol Psychiatry* **13**:558–569.
Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, *et al.* (2009). Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* **14**:755–763.
Spence MA, Flodman P, Sadovnikck AD, Bailey-Wilson JE, Ameli H, Remick RA (1995). Bipolar disorder: evidence for a major locus. *Am J. Med Genet* **60**:370–376.
Spitzer RL, Endicott J, Robins E (1978). Research diagnostic criteria: rationale and reliability. *Arch Gen Psychiatry* **35**:773–782.
Staniszewska I, Sariyer IK, Lecht S, Brown MC, Walsh EM, Tuszynski GP, *et al.* (2008). Integrin alpha9 beta1 is a receptor for nerve growth factor and other neurotrophins. *J Cell Sci* **121**:504–513.
The Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* **447**:661–678.
Tsuang MT, Faraone SV (1990). *The genetics of mood disorders*. Baltimore: Johns Hopkins University Press.

CAPÍTULO 5

ARTIGO 5

INTEGRIN, ALPHA 9 GENE AS SUSCEPTIBILITY GENE FOR BIPOLAR AFFECTIVE DISORDER

Rodrigo Secolin[1], Cláudio E.M. Banzato[2], Lucas Pilli[2], Marilza L. Santos[1], Paulo Dalgalarrondo[2], Iscia Lopes-Cendes[1]

Departments of [1]Medical Genetics and [2]Medical Psychology and Psychiatry,
University of Campinas (UNICAMP), Campinas, São Paulo, Brazil

Integrin, alpha 9 gene as susceptibility gene for bipolar affective disorder

Rodrigo Secolin BSc[1], Cláudio E.M. Banzato MD, PhD[2], Lucas Pilli MD[2], Marilza L. Santos BSc[1], Paulo Dalgalarrondo MD, PhD[2], Iscia Lopes-Cendes MD, PhD[1]

Departments of Medical Genetics[1] and Medical Psychology and Psychiatry[2], University of Campinas - UNICAMP, Campinas, SP, Brazil

Main text word count: 640

Abstract word count: 99

Running title: Family based association study for BAPD

Key words: Genetics, SNPs, Mood disorders, microRNA

Correspondence to: Iscia Lopes-Cendes MD, PhD
Department of Medical Genetics, FCM – UNICAMP
Tessália Vieira de Camargo, 126
Cidade Universitária Zeferino Vaz
Campinas SP, Brazil, CEP 13083 -887
TEL: +55 19 3521 8907
FAX: +55 19 3289 1818
e-mail: icendes@unicamp.br

ABSTRACT

In the present study we aimed to identify a putative susceptibility gene from 3p22 candidate region, previously described for bipolar affective disorder (BPAD). We evaluated 74 pedigrees with BPAD, with a total of 411 individuals, including 96 patients who fulfilled clinical criteria for BPAD according to DSM-IV classification. Among 94 SNPs evaluated by Transmition Disequilibrium Test (TDT), we identified one SNP (rs166508), localized within integrin alpha 9 (*ITGA9*) gene. We observed an increased expression level of *ITGA9* mRNA for mutant AA genotype patients by gene expression analysis. We suggest microRNA upregulation could be under *ITGA9* susceptibility of BPAD phenotype.

INTRODUCTION

Bipolar affective disorder (BPAD [OMIM # 125480]) is a severe and potentially disabling psychiatric illness, with a lifetime prevalence of 0.8-2.6 % in the population. The most characteristic clinical features of this disease are episodes of mania or hypomania interspersed with periods of depression. In untreated patients, the suicide rate can reach approximately 20%[1,2]. BPAD have been also associated to other psychiatric disorders, such as depression and schizophrenia [2-4].

Family, twin and adoption studies clearly demonstrated the contribution of heritable factors BPAD[5-8]. Indeed, many candidate loci and candidate genes have been reported in BPAD, but few of these studies have been replicated in different population [2,8]. The Wellcome Trust Case Control Consortium[9] reported 21 candidate loci with strong or moderate evidence of association to BPAD in the British population. In a previous family based approach, we found a significant signal on 3p22 region[10]. In order further our studies, we aim to refine the 3p22 region in order to identify a putative susceptibility gene for BPAD.

RESULTS

Both SNP average genotype call rate and average genotyping success rate of sample was 90%. Statistical power for our family sample was higher than 80% for the detection of association. SNPs rs3772105, rs35107818, rs11928905 and rs35479964 showed MAF < 0.05 (Figure 1a); SNPs rs2290528 and rs2685112 were not in Hardy-Weinberg equilibrium (Figure 1b); therefore, these SNPs were drop from further analysis. In addition, there was not LD between adjacent SNPs.

Significant signal was found for SNP rs166508, even when adjusted by Bonferroni correction for multiple tests (*Pcorrected* = 0.0187; Figure 1c). Analysis of allele transmission for rs166508 revealed that 80% of total found A allele was transmitted to affected individuals, whereas only 33% of total found allele G was transmitted to affected individuals (Figure 2).

This SNP is located on intron 15 within integrin alpha 9 gene (*ITGA9*), located on 3p21.3 according the National Center of Biotechnology Information (GeneID: 3680; http://www.ncbi.nlm.nih.gov/gene/3680). As shown in Figure 3, expression analysis of *ITGA9* gene revealed a statistically significant increased expression among individuals who carry the mutant AA genotype compared with GG and GA genotypes (Kruskal-Wallis P = 0.0339).

DISCUSSION

Subsequent to the first GWAS performed by The Wellcome Trust Case Control Consortium [9], other GWAS, replication and meta-analysis studies have been strength the association between 3p21-3p22 region and BPAD[11-13]. Thereby, we further our previous findings, refining the 3p22 region nearby SNP rs9834970 [10].

According to TDT results, SNP rs166508 is associated to our BPAD family sample, with an increase of allele A transmission compared to allele G (Figure 3). This SNP is located within intron 15 of the integrin alpha 9 (*ITGA9*) gene (Figure 4). *ITGA9* encodes a subunit of the alpha 9 integrin, an integral membrane glycoprotein which is a receptor for a nerve growth factor, neurotrophin 3 and brain-derived neurotrophic factor [14].

Interesting, expression analysis showed an increase in *ITGA9* level in the blood for mutant AA patients. Despite we did not evaluate human brain tissue, we can suggest that *ITGA9* level in this tissue could be increased in our AA patients, since we observed similarity in expression level for both healthy human blood and brain tissue [15]. However, further *ITGA9* expression studies in account to brain tissues in BPAD patients are necessary to confirm these previous results.

In conclusion, we demonstrated that SNP rs166508, localized within *ITGA9* gene, is associated with the disease phenotype in BAPD pedigrees. We observed that rs166508 is four base pairs downstream a microRNA regulation site (hsa-mir-4271). In fact, several studies have associated microRNA regulation and neuropsychiatric diseases [16,17]. Since hsa-mir-2471 has already implicated in human neural precursors regulation [18], we suggest that rs166508 could be in linkage disequilibrium with a variant within the increased hsa-mir-

4271 site and this variant could generate an upregulation in *ITGA9* level in BPAD patients. However, sequencing analysis of hsa-mir-4271 site is needed to confirm our hypothesis.

METHODS

Ascertainment of subjects and SNP genotyping

We used clinical information and genomic DNA from our previous BPAD familial sample [10]. This sample consisted of 411 individuals distributed in 74 families with 96 bipolar I disorder individuals, classified according DSM-IV criteria [19]. Further study pilot gene expression analysis was conducted from RNA sample from nine unrelated BPAD affected individuals within this family data. This study was approved by the Research Ethics Committee of our institution and all patients and family members signed a consent form prior to entering the study.

SNP Genotyping

We choose 94 SNPs using SNPBROWSER© 4.0 (Applied Biosystems, Foster City, California, USA) from the four populations available in phase II HapMap [20]. The parameters used were: minimum allele frequency (MAF) > 0.05; and linkage disequilibrium $r^2 < 0.8$. SNPs were genotyped by Real Time PCR using SNPlex System® (Applied Biosystems, Foster City, California, USA).

Gene expression analysis

Since *ITGA9* gene is expressed in blood [15], we evaluate differential expression among the three SNP genotypes from RNA was obtained by direct extraction from lymphocytes of peripheral blood according to Ficoll-Plaque Plus (GE Heathcare, Chicago, Illinois, USA) and Trizol (Invitrogen, Carlsbad, Califórnia, USA) procedure. RNA samples were quantified using NanoVue V1.7.2 Spectophotometer (GE Heathcare, Chicago, Illinois, USA).

Expression analysis was performed in triplicate by real-time PCR using the ABI 7500 TaqMan System using human *GAPDH* as endogenous control (probe ID: Hs03929097_g1; Applied Biosystems, Foster City, California, USA) and a *ITGA9* assay (probe ID: Hs00174408_m1; Applied Biosystems, Foster City, California, USA).

Fluorescence signals were detected by the 7500 System Sequence Detection Software (Applied Biosystems, Foster City, California, USA).

Statistical analysis

We used the JINGLEFIX academic software for preprocessing data and MAF estimation from SNPlex data [21]. Mendelian inconsistencies and Hardy-Weinberg Equilibrium (HWE; pvalue threshold = 0.01) were evaluated by PEDCHECK and HAPLOVIEW software, respectively [22,23]. We used the TDTae program for family-based association analysis [24] and Bonferroni correction to account for multiple testing. In order to verify statistical power of our sample, we used the TDT POWER CALCULATOR program, since this software enables estimation from pedigrees with different family structures [22]. The parameters used in the TDT POWER CALCULATOR were: dominant pattern of inheritance with incomplete penetrance [6]; marker frequency = 0.10; allele frequency = 0.25; and a Bonferroni adjusted statistical significance level $\alpha = 0.00053$. Kruskal-Wallis rank sum test was used in order to evaluate differential expression by R environment [25].

ACKNOWLEDGMENTS

We are grateful to all patients and family members for their helpful cooperation. This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP-grant # 08/55827-9), SP, Brazil. Rodrigo Secolin studentship from FAPESP (grants # 06/57332-1).

REFERENCES

1. Alda, M. Bipolar disorder: from families to genes. *Canadian journal of psychiatry* **42**, 378-87(1997).

2. Kato, T. Molecular genetics of bipolar disorder and depression. *Psychiatry and clinical neurosciences* **61**, 3-19(2007).

3. Maier, W. et al. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. *Archives of general psychiatry* **50**, 871-83(1993).

4. Goodwin, F.K. & Jamison, K.R. *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression, 2nd Edition*. 1262 (Oxford University Press, USA: 2007).

5. McGuffin, P. & Katz, R. The genetics of depression and manic-depressive disorder. *The British journal of psychiatry : the journal of mental science* **155**, 294-304(1989).

6. Spence, M.A. et al. Bipolar disorder: evidence for a major locus. *American journal of medical genetics* **60**, 370-6(1995).

7. Kato, T., Kuratomi, G. & Kato, N. Genetics of bipolar disorder. *Drugs of today* **41**, 335-44(2005).

8. Barnett, J.H. & Smoller, J.W. The genetics of bipolar disorder. *Neuroscience* 1-13(2009).

9. Consortium, T.W.T.C.C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78(2007).

10. Secolin, R. et al. Family-based association study for bipolar affective disorder. *Psychiatric genetics* **20**, 126-9(2010).

11. McMahon, F.J. et al. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nature genetics* **42**, 128-131(2010).

12. Sklar, P. et al. Whole-genome association study of bipolar disorder. *Molecular psychiatry* **13**, 558-69(2008).

13. Scott, L.J. et al. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7501-6(2009).

14. Staniszewska, I. et al. Integrin alpha9 beta1 is a receptor for nerve growth factor and other neurotrophins. *Journal of cell science* **121**, 504-13(2008).

15. Su, A.I. et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 6062-7(2004).

16. Miller, B.H. & Wahlestedt, C. MicroRNA dysregulation in psychiatric disease. *Brain research* **1338**, 89-99(2010).

17. Xu, B., Karayiorgou, M. & Gogos, J.A. MicroRNAs in psychiatric and neurodevelopmental disorders. *Brain research* **1338**, 78-88(2010).

18. Goff, L.A. et al. Ago2 immunoprecipitation identifies predicted microRNAs in human embryonic stem cells and neural precursors. *PloS one* **4**, e7192(2009).

19. Association, A.P. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV*. 886(American Psychiatric Association: 1994).

20. Consortium, T.I.H. A haplotype map of the human genome. *Nature* **437**, 1299-320(2005).

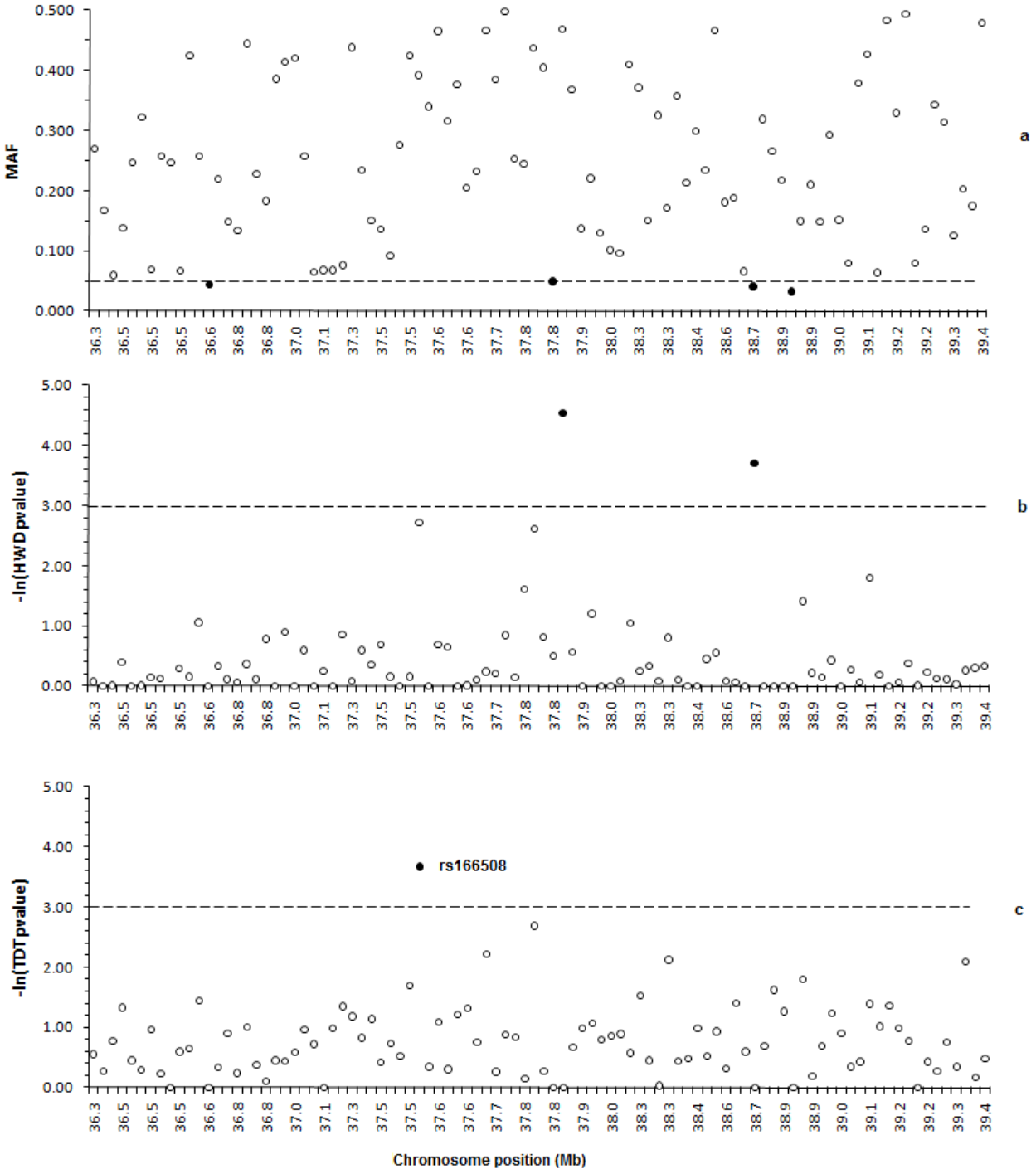
21. Secolin, R. et al. LINKGEN: a new algorithm to process data in genetic linkage studies. *Genomics* **91**, 544-7(2008).

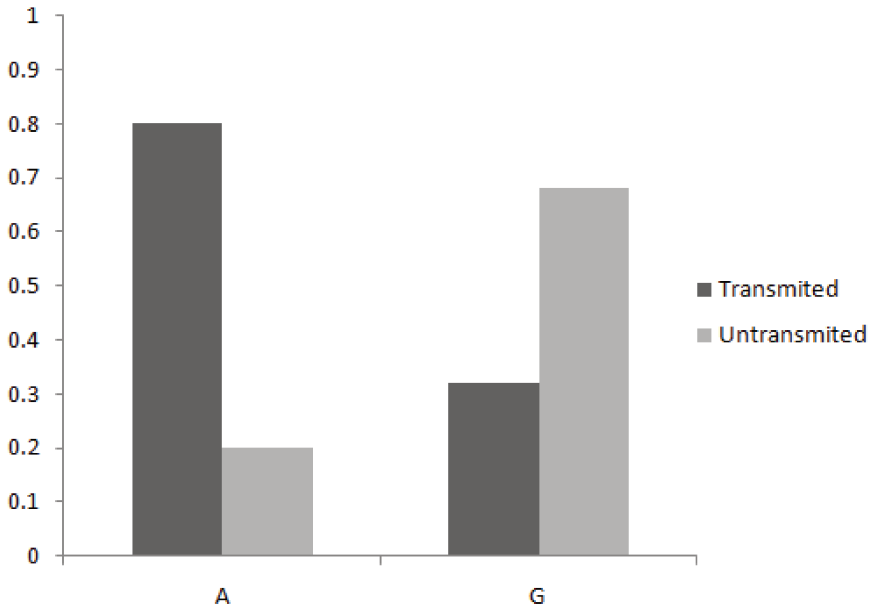
22. Chen, W.M. & Deng, H.W. A general and accurate approach for computing the statistical power of the transmission disequilibrium test for complex disease genes. *Genetic epidemiology* **21**, 53-67(2001).

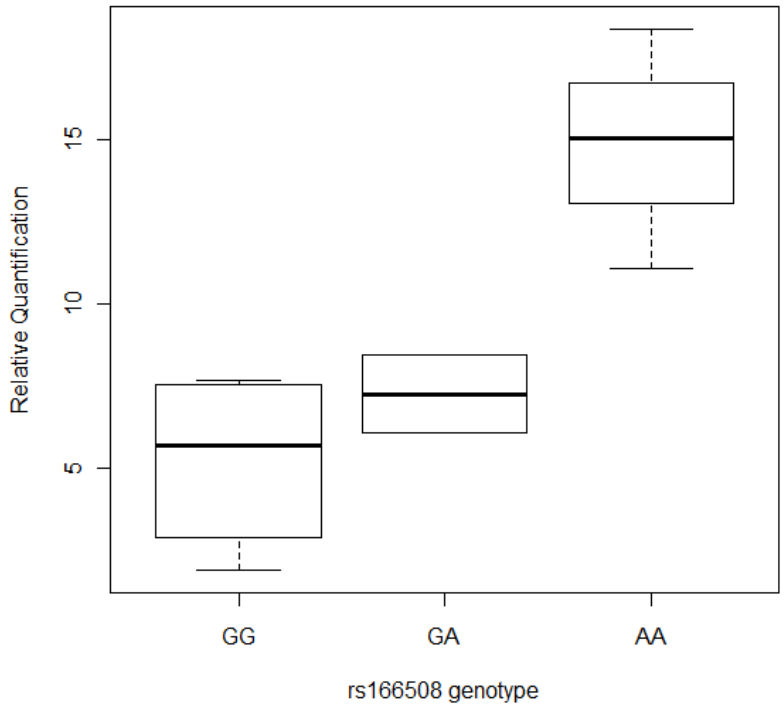
23. Barrett, J.C. et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-5(2005).

24. Gordon, D. et al. A transmission disequilibrium test for general pedigrees that is robust to the presence of random genotyping errors and any number of untyped parents. *Human Heredity* **12**, 752-761(2006).

25. R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing. (2007).at <http://www.r-project.org>


MAF
0.500
0.400
0.300
0.200
0.100
0.000
a
-ln(HWDpvalue)
5.00
4.00
3.00
2.00
1.00
0.00
b
-ln(TDTpvalue)
rs166508
c
36.3
36.5
36.5
36.5
36.6
36.8
36.8
37.0
37.1
37.3
37.5
37.5
37.6
37.6
37.7
37.8
37.8
37.9
38.0
38.3
38.3
38.4
38.6
38.7
38.9
38.9
39.0
39.1
39.2
39.2
39.3
39.4
Chromosome position (Mb)


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
A
G
Transmited
Untransmited


Relative Quantification
5
10
15
GG
GA
AA
rs166508 genotype

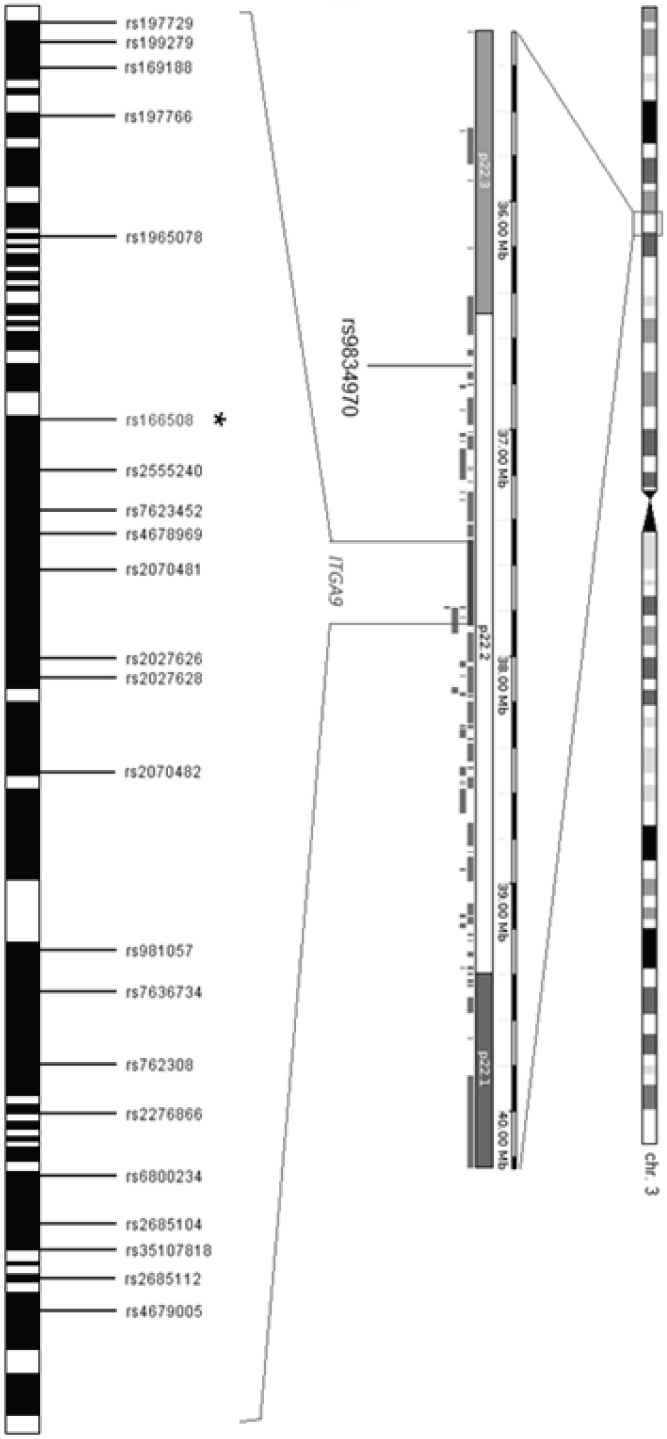

rs197729
rs199279
rs169188
rs197766
rs1965078
rs166508 *
rs2555240
rs7623452
rs4678969
rs2070481
rs2027626
rs2027628
rs2070482
rs981057
rs7636734
rs762308
rs2276866
rs6800234
rs2685104
rs35107818
rs2685112
rs4679005
rs9834970
ITGA9
p22.3
p22.2
p22.1
36.00 Mb
37.00 Mb
38.00 Mb
39.00 Mb
40.00 Mb
chr. 3

Figure 1. Quality control and TDT results from 3p22 region: a) Minor allele frequency (MAF) showing three SNPs with MAF < 0.05; b) Hardy-Weinberg Disequilibrium (HWD) showing two SNPs with altered distribution; c) TDT *P* results, showing SNPs rs166508 as statistically significant association.

Figure 2. Proportion of transmitted and untransmitted rs166508 alleles within BPAD families.

Figure 3. Boxplot of relative quantification analysis from the three different genotypes of unrelated BPAD individuals.

Figure 4. Location SNP rs166508 (asterisk) among SNPs evaluated within *ITGA9* gene (exons = white; introns = black).

4. DISCUSSÃO

A escolha do tipo de amostra, do modelo estatístico e dos marcadores moleculares é de extrema importância para o correto mapeamento genético. Neste contexto, foram levantados dois problemas: 1) visto a grande quantidade de dados gerados para as análises estatísticas e a ausência de programas de processamento deste tipo de dados, **como realizar a interface entre os dados dos equipamentos e os dados utilizados pelos cálculos estatísticos**; 2) visto que, apesar das amostras familiares serem mais adequadas em nosso meio em relação a uma amostra populacional, a heterogeneidade das amostras familiares gera dúvidas quanto **ao melhor modelo estatístico a ser utilizado**.

O algoritmo LINKGEN (Capítulo 1) foi desenvolvido com o objetivo inicial de processar os dados de genotipagem de marcadores microssatélites em amostras familiares, sendo utilizado em mais três estudos de mapeamento genético (43-45). Além disso, este programa também se mostrou útil para processamento dos dados de genotipagem de SNPs, como mostrado no artigo do Capítulo 4. Renomeado para JINGLEFIX, este programa está sendo registrado no Instituto Nacional de Propriedade Intelectual (INPI pedido nº 0000270805133430).

Visto que, entre as três doenças estudadas, somente a ELTM não possui um padrão de herança definido (9; 11; 46), foi realizada uma análise de segregação complexa para definição destes parâmetros, como descrito no Capítulo 2. Os resultados apontam para a presença de um gene de maior efeito, segregando em um padrão autossômico dominante, além da presença de genes de menor efeito influenciando o fenótipo da ELTM. Estes parâmetros foram utilizados em posteriores estudos de mapeamento em todo o genoma para ELTM, onde foi encontrada ligação genética na região 18p11 (47).

Entretanto, além do padrão de herança, a estrutura e o tamanho das famílias também influenciam os modelos estatísticos de mapeamento genético. A fim de comparar o poder estatístico das análises paramétricas e não paramétricas em diferentes tipos de amostras, foram realizados cálculos de poder estatístico utilizando um teste paramétrico (LINKAGE) e um não paramétrico (TDT), em cada uma das amostras segregando PPBC (Capítulo 3) e TAB (Capítulo 4), para a utilização de até 100 marcadores moleculares. A análise paramétrica se mostrou mais robusta do que a análise não paramétrica para a família

segregando PPBC. De modo contrário, a análise não paramétrica se mostrou mais robusta para as 74 famílias segregando TDT (Figura 5).

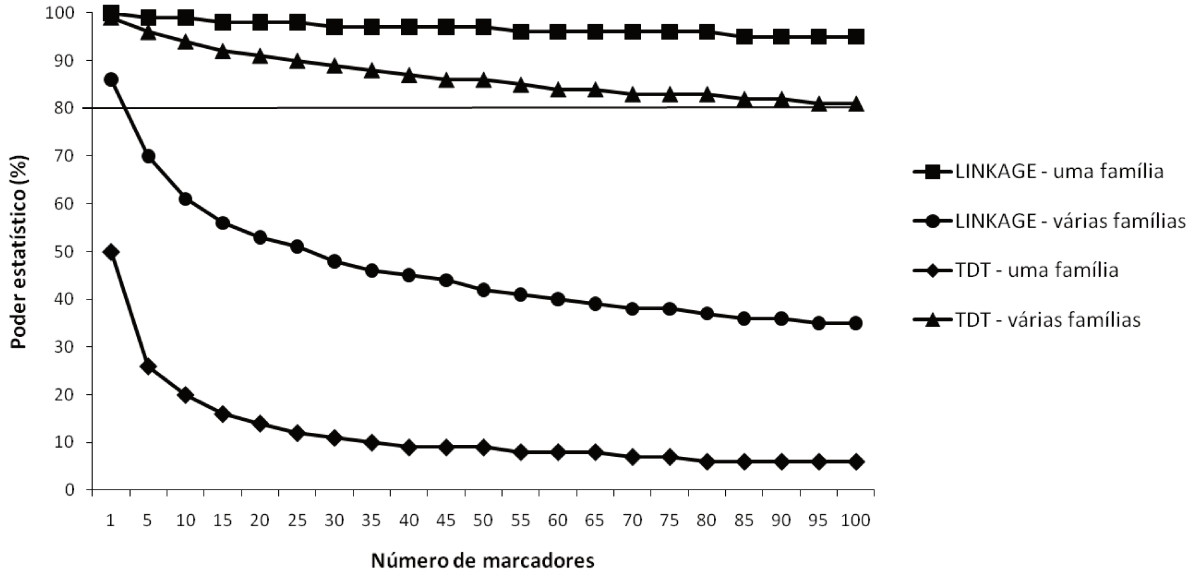


Figura 5. Comparação entre as análises paramétricas (LINKAGE) e não paramétricas (TDT) para as duas amostras estudadas.

Em relação aos achados descritos no Capítulo 3, Santos et al. encontraram ligação genética na região Xq27, mais centromérica da região Xq28 descrita por Villard et al. (11). Vink et al. (2) sugerem que as diferenças nos resultados das análises estatísticas podem ser explicadas pela estrutura da amostra familiar. De fato, diferenças entre as amostras podem ser observadas: Santos et al. (Capítulo 3) avaliaram somente uma família com 15 indivíduos; enquanto que Villard et al. (11) avaliaram cinco famílias com uma média de cinco indivíduos por família. Entretanto, as cinco famílias não se apresentam informativas se avaliadas independentemente. Atualmente, estudos com alguns genes candidatos na região Xq27 estão sendo realizados em nosso laboratório, tais como genes relacionados com modulação de neuritos (*SLITRK2* e *SLITRK4*, *SRPX2*, *AFF2*), visto que nenhum gene foi identificado na região Xq27-Xq28 para PPBC.

Apesar de menor poder estatístico em relação aos testes paramétricos (2), o sucesso do mapeamento genético utilizando análises não paramétricas para um número

maior de famílias de tamanhos e estruturas variadas é evidente no Capítulo 4, onde foi encontrada associação entre a região 3p22 e TAB e posterior identificação do SNP rs166508, localizado no intron 15 do gene *ITGA9* (Capítulo V).

Entretanto, diferentes genes foram identificados na mesma região, tais como os genes *ANK3, GNL3*, *NEK4*, *ITIH3* (48) e o gene *PBRM1* (49). Estes diferentes resultados podem ser explicados novamente pelo tipo de amostra estudada (2), pois incluem tanto estudos de meta-análise quanto estudos populacionais e familiares. Barnett e Smoller (50) reforçam esta hipótese, sugerindo que casos familiares segregando TAB podem apresentar genes de susceptibilidade diferentes dos casos não familiares. Portanto, o gene *ITGA9* continua como um gene de susceptibilidade para TAB em nossa amostra familiar.

O SNP rs166508 está a quatro pares de bases no sentido 3' de um provável sitio de ligação do microRNA hsa-mir-2471, o qual está relacionado com precursores neurais. Visto que indivíduos afetados com TAB portadores do genótipo AA apresentam aumento de expressão do gene *ITGA9*, o SNP rs166508 pode estar em desequilíbrio de ligação com uma possível variante no sítio de ligação do hsa-mir-2471. Portanto, esta variação pode levar a não ligação do microRNA na região e consequente diminuição na regulação da expressão de *ITGA9*.

Entretanto, análise da sequência do provável sítio de ligação utilizando os indivíduos avaliados no estudo de expressão, juntamente com seus familiares, não revelou alteração na sequência entre os indivíduos portadores dos genótipos GG, GA e AA do SNP RS 166508. Além disso, a análise dos 10 SNPs não sinônimos presentes na região codificante do gene *ITGA9* também não revelou associação com o TAB. Portanto, até o momento, não foi encontrado um substrato que justifique o aumento de expressão do gene *ITGA9* para os indivíduos portadores do genótipo AA para o SNP rs166508. A análise da sequência dos mRNA a partir de amostras de cultura de linfoblastos dos mesmos indivíduos utilizados no estudo de expressão, além da análise das regiões promotora e UTR em toda a amostra estão sendo realizadas a fim de identificar o substrato que justifique o aumento de expressão do gene *ITGA9* em pacientes com TAB.

5. CONCLUSÕES

Artigo 1:

O programa JINGLEFIX é apto para uma adequada interface de dados entre os diversos equipamentos de genotipagem (MegaBACE® 1000, ABI® 7500 e chips de *microarray*) e os programas estatísticos mais utilizados para mapeamento genético.

Artigo 2:

A análise de segregação complexa mostrou a presença de um gene principal, com padrão de herança autossômico dominante, além da presença de genes de menor efeito, atuando no fenótipo da ELTM. Este gene foi posteriormente mapeado na região 18p11.

Artigo 3:

A análise paramétrica de ligação genética em uma família segregando PPBC identificou a região Xq27 como região candidata a possuir o possível gene responsável por PPBC nesta família.

Artigo 4:

A análise não paramétrica de ligação mapeou a região 3p22 para a amostra de famílias segregando TAB.

Artigo 5:

O refinamento da região 3p22 identificou o gene *ITGA9* como possível gene de susceptibilidade para o TAB.

Conclusão geral:

Tanto o algoritmo de processamento de dados quanto à adequada aplicação dos modelos estatísticos foram fundamentais para o mapeamento genético das regiões e dos genes responsáveis pelas doenças neuro-psiquiátricas estudadas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

1. Ott J. Analysis of human genetic linkage. Rev. ed. Maryland: John Hopkins; 1991.

2. Vink JM, Boomsma DI. Gene finding strategies. Biological psychology. 2002 Oct ;61(1-2):53-71.

3. Laidlaw J, Richens A. Textbook of Epilepsy. Churchill Livingstone; 1993.

4. Berkovic SF, Andermann F, Olivier A, Ethier R, Melanson D, Robitaille Y, et al. Hippocampal sclerosis in temporal lobe epilepsy demonstrated by magnetic resonance imaging. Annals of neurology. 1991 Feb ;29(2):175-82.

5. Cendes F, Andermann F, Gloor P, Evans A, Jones-Gotman M, Watson C, et al. MRI volumetric measurement of amygdala and hippocampus in temporal lobe epilepsy. Neurology. 1993 Apr ;43(4):719-25.

6. Kobayashi E, Lopes-Cendes I, Guerreiro CA, Sousa SC, Guerreiro MM, Cendes F. Seizure outcome and hippocampal atrophy in familial mesial temporal lobe epilepsy. Neurology. 2001 Jan ;56(2):166-72.

7. Piao X, Basel-Vanagaite L, Straussberg R, Grant PE, Pugh EW, Doheny K, et al. An autosomal recessive form of bilateral frontoparietal polymicrogyria maps to chromosome 16q12.2-21. American journal of human genetics. 2002 Apr ;70(4):1028-33.

8. Kuzniecky R, Andermann F, Guerrini R. Congenital bilateral perisylvian syndrome: study of 31 patients. The CBPS Multicenter Collaborative Study. Lancet. 1993 Mar ;341(8845):608-12.

9. Guerreiro MM, Andermann E, Guerrini R, Dobyns WB, Kuzniecky R, Silver K, et al. Familial perisylvian polymicrogyria: a new familial syndrome of cortical maldevelopment. Annals of neurology. 2000 Jul ;48(1):39-48.

10. Gropman AL, Barkovich AJ, Vezina LG, Conry JA, Dubovsky EC, Packer RJ. Pediatric congenital bilateral perisylvian syndrome: clinical and MRI features in 12 patients. Neuropediatrics. 1997 Aug ;28(4):198-203.

11. Villard L, Nguyen K, Cardoso C, Martin CL, Weiss AM, Sifry-Platt M, et al. A locus for bilateral perisylvian polymicrogyria maps to Xq28. American journal of human genetics. 2002 Apr ;70(4):1003-8.

12. Kato T. Molecular genetics of bipolar disorder and depression. [Internet]. Psychiatry and clinical neurosciences. 2007 Mar ;61(1):3-19.Available from: http://www.ncbi.nlm.nih.gov/pubmed/17239033

13. Alda M. Bipolar disorder: from families to genes. Canadian journal of psychiatry. Revue canadienne de psychiatrie. 1997 May ;42(4):378-87.

14. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996 Sep ;273(5281):1516-7.

15. McClellan J, King M-C. Genetic heterogeneity in human disease. Cell. 2010 Apr ;141(2):210-7.

16. Ferreira LB, Mendes-Junior CT, Wiezel CEV, Luizon MR, Simões AL. Genomic ancestry of a sample population from the state of São Paulo, Brazil. American journal of human biology. 2008 ;18(5):702-5.

17. Lambertucci R, Rocha J, Antunes CM, Parra FC, Amado RC. Color and genomic ancestry in Brazilians ´. PNAS. 2003 ;100(1):177-182.

18. Law B, Buckleton JS, Triggs CM, Weir BS. Effects of population structure and admixture on exact tests for association between Loci. Genetics. 2003 May ;164(1):381-7.

19. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. Human molecular genetics. 2008 Oct ;17(R2):R143-50.

20. Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. Human heredity. 2008 Jan ;66(2):87-98.

21. Terwilliger JD, Ott J. Handbook of Human Genetic Linkage. The Johns Hopkins University Press; 1994.

22. Muniz YCN, Ferreira LB, Mendes-Junior CT, Wiezel CEV, Simões AL. Genomic ancestry in urban Afro-Brazilians. Annals of human biology. 2008 ;35(1):104-11.

23. Morton NE, MacLean CJ. Analysis of family resemblance. 3. Complex segregation of quantitative traits. American journal of human genetics. 1974 Jul ;26(4):489-503.

24. Futuyma DJ. Evolutionary Biology. Sinauer Associates; 1997.

25. Lalouel JM, Morton NE. Complex segregation analysis with pointers. Human heredity. 1981 Jan ;31(5):312-21.

26. Casella G, Berger RL. Statistical Inference. Duxbury Press; 2001.

27. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974 ;19(6):716-723.

28. Morton NE. The detection and estimation of linkage between the genes for elliptocytosis and the Rh blood type. American journal of human genetics. 1956 Jun ;8(2):80-96.

29. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. American journal of human genetics. 1996 Jun ;58(6):1347-63.

30. Lander ES, Green P. Construction of multilocus genetic linkage maps in humans. Proceedings of the National Academy of Sciences of the United States of America. 1987 Apr ;84(8):2363-7.

31. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. American journal of human genetics. 1990 Feb ;46(2):222-8.

32. Risch N. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. American journal of human genetics. 1990 Feb ;46(2):229-41.

33. Spielman RS, McGinnis RE, Ewens WJ. Transmission Test for Linkage Disequilibrium: American Journal of Human Genetics. 1993 ;52(3):506-516.

34. Gordon D, Haynes C, Johnnidis C, Patel SB, Bowcock AM, Ott J. A transmission disequilibrium test for general pedigrees that is robust to the presence of random genotyping errors and any number of untyped parents. Human Heredity. 2006 ;12(9):752-761.

35. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Internet]. 2007 ;Available from: http://www.r-project.org

36. Strachan T, Read A. Human Molecular Genetics, Third Edition. Garland Science/Taylor & Francis Group; 2003.

37. National Center of Biotechnology Information. Homo sapiens Genome: Statistics Build 37 version 1 [Internet]. 2010 ;Available from: http://www.ncbi.nlm.nih.gov/mapview/stats/BuildStats.cgi?taxid=9606&build=37&ver=1

38. Consortium TIH. A haplotype map of the human genome. Nature. 2005 ;437(7063):1299-320.

39. Frazer K a, Ballinger DG, Cox DR, Hinds D a, Stuve LL, Gibbs R a, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007 ;449(7164):851-61.

40. Cui Y, Li G, Li S, Wu R. Designs for linkage analysis and association studies of complex diseases. Methods in molecular biology. 2010 Jan ;620219-42.

41. Draghici S. Data Analysis Tools for DNA Microarrays. Chapman and Hall/CRC; 2003.

42. Secolin R, Banzato CEM, Oliveira MCM, Bittar MFR, Santos ML, Dalgalarrondo P, et al. Family-based association study for bipolar affective disorder. Psychiatric genetics. 2010 Jun ;20(3):126-9.

43. França MC, D'Abreu A, Maurer-Morelli CV, Seccolin R, Appenzeller S, Alessio A, et al. Prospective neuroimaging study in hereditary spastic paraplegia with thin corpus callosum. Movement disorders. 2007 Aug ;22(11):1556-62.

44. Maurer-Morelli CV, Marchesini RB, Secolin R, Santos NF, Kobayashi E, Cendes F, et al. Linkage study of voltage-gated potassium channels in familial mesial temporal lobe epilepsy. Arquivos de Neuropsiquiatria. 2007 ;65(1):20-23.

45. Maurer-Morelli CV, Secolin R, Marchesini RB, Santos NF, Kobayashi E, Cendes F, et al. THE SCN2A gene is not a likely candidate for familial mesial temporal lobe epilepsy. Epilepsy research. 2006 Oct ;71(2-3):233-6.

46. Spence MA, Flodman PL, Sadovnick AD, Bailey-Wilson JE, Ameli H, Remick RA. Bipolar disorder: evidence for a major locus. American journal of medical genetics. 1995 Oct ;60(5):370-6.

47. Maurer-Morelli CV, Secolin R, Domingues RR, Marchesini RB, Santos NF, Kobayashi E, et al. Identification of the locus for familial mesial temporal lobe epilepsy with hippocampal atrophy and search for candidate genes. Supplement of Neurology. 2007 ;68(S1):A338-A338.

48. Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, et al. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. Proceedings of the National Academy of Sciences of the United States of America. 2009 May ;106(18):7501-6.

49. McMahon FJ, Akula N, Schulze TG, Muglia P, Tozzi F, Detera-Wadleigh SD, et al. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. [Internet]. Nature genetics. 2010 ;42(2):128-131.Available from: http://www.ncbi.nlm.nih.gov/pubmed/20081856

50. Barnett JH, Smoller JW. The genetics of bipolar disorder. Neuroscience. 2009 Apr ;1-13.

7. ANEXOS

ANEXO 1 – Licença de uso de *Copyright* para o artigo 1

ELSEVIER LICENSE
TERMS AND CONDITIONS

Dec 03, 2010

This is a License Agreement between Rodrigo Secolin ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Rodrigo Secolin
Customer address	R. Tessalia Vieira de Camargo, 126 Campinas, other 13083887
License number	2561360187187
License date	Dec 03, 2010
Licensed content publisher	Elsevier
Licensed content publication	Genomics
Licensed content title	LINKGEN: A new algorithm to process data in genetic linkage studies
Licensed content author	Rodrigo Secolin, Cristiane S. Rocha, Fábio R. Torres, Marilza L. Santos, Cláudia V. Maurer-Morelli, Neide F. Santos, Iscia Lopes-Cendes
Licensed content date	June 2008
Licensed content volume number	91
Licensed content issue number	6
Number of pages	4
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Application of statistical models and algorithm development for genetic studies of neuropsychiatric diseases
Expected completion date	Feb 2011
Estimated size (number of pages)	107

Elsevier VAT number GB 494 6272 12

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily

granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Website**: The following terms and conditions apply to electronic reserve and author websites:
Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:
This license was made in connection with a course,
This permission is granted for 1 year only. You may obtain a license for future website posting,
All content posted to the web site must maintain the copyright information line on the bottom of each image,
A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com , and
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and
he permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version,
A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx , As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. **Author website** for books with the following additional clauses:
Authors are permitted to place a brief summary of their work online only.
A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com

All content posted to the web site must maintain the copyright information line on the bottom of each image
You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. or for books to the Elsevier homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. **Other Conditions**:

v1.6

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10893952.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2561360187187 and noting the reason for cancellation.

Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

ANEXO 2 – Licença de uso de *Copyright* para o artigo 2

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Dec 03, 2010

This is a License Agreement between Rodrigo Secolin ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2561380915186
License date	Dec 03, 2010
Licensed content publisher	John Wiley and Sons
Licensed content publication	Epilepsia
Licensed content title	Segregation analysis in mesial temporal lobe epilepsy with hippocampal atrophy
Licensed content author	Rodrigo Secolin,Claudia Maurer-Morelli,Fernando Cendes,Iscia Lopes-Cendes
Licensed content date	Feb 1, 2010
Start page	47
End page	50
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one if its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing

process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). Any electronic posting of the Materials is limited to one year from the date permission is granted and is on the condition that a link is placed to the journal homepage on Wiley's online journals publication platform at www.interscience.wiley.com. The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher and on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. No part of the Materials may be copied, modified, adapted, translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. WILEY DOES NOT MAKE ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this

Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in a writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

13. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

14. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

15. This Agreement shall be governed by and construed in accordance with the laws of England and you agree to submit to the exclusive jurisdiction of the English courts.

16. Other Terms and Conditions:

BY CLICKING ON THE "I ACCEPT" BUTTON, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

V1.2

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10893990.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2561380915186 and noting the reason for cancellation.

Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

ANEXO 3 – Licença de uso de *Copyright* para o artigo 3

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Dec 03, 2010

This is a License Agreement between Rodrigo Secolin ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2561381309119
License date	Dec 03, 2010
Licensed content publisher	John Wiley and Sons
Licensed content publication	American Journal of Medical Genetics Part A
Licensed content title	A new candidate locus for bilateral perisylvian polymicrogyria mapped on chromosome Xq27
Licensed content author	Neide F. Santos,Rodrigo Secolin,Iara L. Brandão-Almeida,Marilza S. Silva,Fábio R. Torres,Simone S. Tsuneda,Catarina A. Guimarães,Simone R.V. Hage,Fernando Cendes,Marilisa M. Guerreiro,Iscia Lopes-Cendes
Licensed content date	May 1, 2008
Start page	1151
End page	1157
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one if its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). Any electronic posting of the Materials is limited to one year from the date permission is granted and is on the condition that a link is placed to the journal homepage on Wiley's online journals publication platform at www.interscience.wiley.com. The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher and on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. No part of the Materials may be copied, modified, adapted, translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. WILEY DOES NOT MAKE ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in a writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

13. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

14. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

15. This Agreement shall be governed by and construed in accordance with the laws of England and you agree to submit to the exclusive jurisdiction of the English courts.

16. Other Terms and Conditions:

BY CLICKING ON THE "I ACCEPT" BUTTON, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

V1.2

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10894000.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2561381309119 and noting the reason for cancellation.

Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

ANEXO 4 – Licença de uso de *Copyright* para o artigo 4

WOLTERS KLUWER HEALTH LICENSE TERMS AND CONDITIONS

Dec 03, 2010

This is a License Agreement between Rodrigo Secolin ("You") and Wolters Kluwer Health ("Wolters Kluwer Health") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Wolters Kluwer Health, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2561381144011
License date	Dec 03, 2010
Licensed content publisher	Wolters Kluwer Health
Licensed content publication	Psychiatric Genetics
Licensed content title	Family-based association study for bipolar affective disorder
Licensed content author	Rodrigo Secolin, Cláudio Banzato, Maria Oliveira, et al
Licensed content date	Jan 1, 2010
Volume Number	20
Issue Number	3
Type of Use	Dissertation/Thesis
Requestor type	Individual
Title of your thesis / dissertation	Application of statistical models and algorithm development for genetic studies of neuropsychiatric diseases
Expected completion date	Feb 2011
Estimated size(pages)	107
Billing Type	Invoice
Billing Address	R. Tessalia Vieira de Camargo, 126 Universidade Estadual de Campinas Campinas, other 13083887 Brazil
Customer reference info	
Total	0.00 USD
Terms and Conditions	

Terms and Conditions

1. A credit line will be prominently placed and include: for books - the author(s), title of book, editor, copyright holder, year of publication; For journals - the author(s), title of article, title of journal, volume number, issue number and inclusive pages.
2. The requestor warrants that the material shall not be used in any manner which may be considered derogatory to the title, content, or authors of the material, or to Wolters Kluwer/Lippincott, Williams & Wilkins.

3. Permission is granted for one time use only as specified in your correspondence. Rights herein do not apply to future reproductions, editions, revisions, or other derivative works. Once term has expired, permission to renew must be made in writing.
4. Permission granted is non-exclusive, and is valid throughout the world in the English language and the languages specified in your original request.
5. Wolters Kluwer Health/ Lippincott, Williams & Wilkins, cannot supply the requestor with the original artwork or a "clean copy."
6. The requestor agrees to secure written permission from the author (for book material only).
7. Permission is valid if the borrowed material is original to a LWW imprint (Lippincott-Raven Publishers, Williams & Wilkins, Lea & Febiger, Harwal, Igaku-Shoin, Rapid Science, Little Brown & Company, Harper & Row Medical, American Journal of Nursing Co, and Urban & Schwarzenberg - English Language).
8. If you opt not to use the material requested above, please notify Rightslink within 90 days of the original invoice date.
9. Other Terms and Conditions:

v1.0

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10893994.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2561381144011 and noting the reason for cancellation.

Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.