

O genoma humano e as perspectivas para o estudo da esquizofrenia

ELIDA P. BENQUIQUE OJOPI¹
SHEILA PASSOS GREGORIO^{1,2}
PEDRO EDSON MOREIRA GUIMARÃES¹
CINTIA FRIDMAN¹
EMMANUEL DIAS NETO^{1*}

Resumo

O seqüenciamento de nosso genoma representa um passo essencial no entendimento da biologia humana e no planejamento racional de pesquisas biomédicas. Contudo, é importante notar que o seqüenciamento de um dado genoma é apenas uma parte de um complexo quebra-cabeças. A informação genética deve ser usada como um “mapa”, a partir do qual começamos a compreender a base das doenças e a importância da variação genética através da análise da complexidade e do comportamento das regiões reguladoras, genes e proteínas, funções gênicas e sistemas celulares. Apesar dos enormes esforços para identificar genes de susceptibilidade, os resultados de estudos de genética molecular de esquizofrenia até o momento têm sido modestos. O uso apropriado da genômica poderá ajudar imensamente na elucidação das causas da esquizofrenia, permitindo avaliar o papel de novos genes, das variações genéticas, das formas de *splicing* alternativo, das variações de expressão gênica e de vias metabólicas de interesse. A convergência de dados bioquímicos, de imagem, de neuroanatomia, farmacológicos, clínicos e genéticos permite prever que estamos muito próximos de uma melhor compreensão das bases biológicas da esquizofrenia. A disponibilidade desses avanços terá um enorme impacto na pesquisa desta doença.

Palavras-chave : Genoma, transcriptoma, neurogênese, esquizofrenia, polimorfismos.

Abstract

The sequencing of our genome represents an essential step in the comprehension of the human biology and in the rational planning of biomedical research. However, it is important to realize that the sequencing of a genome is only a piece of a complex puzzle. The genetic information must be used as a “map”, the starting-point to understand the basis of the diseases and the importance of the genetic variation through the analysis of the complexity and behavior of the regulatory regions, genes and proteins, gene functions and cellular systems. Despite the enormous efforts made towards the identification of susceptibility genes for schizophrenia, the results of molecular genetics studies have been modest. The appropriate use of genomics should add more information to help the elucidation of the causes of the schizophrenia, supporting areas such as gene discovery, determination of genetic variations, alternative splicing, analysis of differential gene expression and in the study of metabolic pathways of interest. The convergence of data from studies of biochemistry, imaging, neuroanatomy, pharmacology, clinical and genetics allows to foresee that we should be near to a better understanding of the biological basis of schizophrenia. The availability of these advances will have a dramatic impact on the research of this disease.

Key Words : Genome, transcriptome, neurogenesis, schizophrenia, polymorphisms.

Recebido: 17/11/2003 - Aceito: 26/11/2003

1 - Laboratório de Neurociências (LIM27), Instituto de Psiquiatria, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP).

2 - Instituto de Química da Universidade de São Paulo (IQ-USP).

* Pesquisador e vice-diretor do Laboratório de Neurociências (LIM27). *Endereço para correspondência*: Laboratório de Neurociências (LIM27), Instituto de Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Rua Dr. Ovídio Pires de Campos, s/n, Cerqueira César. 05403-010, São Paulo - SP, Brasil. Telefone: (11) 30697267; FAX: (11) 30624851; E-mail: emmanuel@usp.br

Introdução

As primeiras tentativas de compreensão dos fenômenos biológicos são provavelmente tão antigas quanto a linguagem e o pensamento consciente. Desde os primórdios temos tentado compreender nossas origens, buscando influenciar a dinâmica da vida, das doenças e da morte. Embora tenham ocorrido descobertas dramáticas nessa área, particularmente nos últimos 50 anos, nos últimos três anos, presenciamos um salto significativo em tecnologias da pesquisa biológica, principalmente no que se refere aos progressos na área de seqüenciamento genômico. A tecnologia de seqüenciamento automático de DNA utilizando raios laser (uma técnica que permite um acúmulo rápido de informação genética, gerando um rascunho genético de um dado organismo), tem modificado fundamentalmente o nosso conhecimento de biologia e de muitas áreas associadas. A implementação do seqüenciamento de DNA em larga escala permitiu a realização de um dos feitos científicos mais importantes da humanidade: o seqüenciamento do genoma humano.

O seqüenciamento de nosso genoma representa um passo essencial no entendimento da biologia humana e no planejamento racional de pesquisas biomédicas. A finalização do seqüenciamento de nosso material genético entusiasmou as buscas de novas abordagens para se determinar o risco associado a uma determinada doença ou o desenvolvimento de terapias individualizadas. Contudo, é importante notar que o seqüenciamento de um dado genoma é apenas uma parte de um complexo quebra-cabeça. A informação genética deve ser usada como um “mapa”, a partir do qual começamos a compreender a base das doenças e da variação genética através da análise da complexidade e comportamento das regiões reguladoras, genes e proteínas, funções gênicas e sistemas celulares. É crucial cruzarmos as informações genéticas com os dados clínicos, étnicos e informações ambientais, a fim de obtermos uma visão ampla do quadro. Com a informação genética disponível, podemos, por exemplo, criar bancos de dados de variações genéticas, buscando uma terapia personalizada, baseada no perfil genético dos pacientes. A medicina poderá saber por que um paciente tem suscetibilidade elevada a uma determinada doença, enquanto outro não. Nós poderemos também ser capazes de prever como uma doença vai progredir e de que forma poderíamos bloquear esse processo, utilizando drogas mais específicas ou outras intervenções. A era da

biociência vai mudar nossas vidas de forma fundamental - algumas mudanças sutis, outras mais claras - atualmente, temos apenas uma vaga idéia do que ainda está por vir (McLeod e Evans, 2001; Evans e Relling, 1999).

O desafio mais intrigante da biomedicina atual é a compreensão de doenças complexas que nos afetam. As doenças complexas apresentam origens multifatoriais e algumas vezes parecem até tratar-se de diferentes doenças classificadas erroneamente com o mesmo nome. Foram encontradas associações genéticas em vários cromossomos para todas as doenças complexas que nos afetam, e a finalização do seqüenciamento do genoma humano pode auxiliar o estudo de tais doenças. Muitas pesquisas têm concentrado seus esforços no estudo de regiões genômicas envolvidas, mas as técnicas empregadas até então na busca de genes-alvo eram muito laboriosas, dispendiosas e consumiam um tempo considerável. Após o seqüenciamento do genoma, temos a oportunidade de “olhar” cuidadosamente para essas regiões e apontar para uma coleção de alterações que podem elucidar a origem de características clínicas importantes em muitas dessas doenças.

Neste artigo discutiremos algumas das oportunidades de pesquisas criadas após o seqüenciamento do genoma humano, buscando vislumbrar novas possibilidades para a melhor compreensão da esquizofrenia.

O rascunho do genoma humano

“O genoma humano contém o código genético que preside o cerne de cada um dos 10 trilhões de células em cada ser humano; ele influencia profundamente os nossos corpos, comportamentos e nossas mentes. A finalização do seqüenciamento do genoma humano auxiliará o estudo das influências não-genéticas no desenvolvimento humano, revelará novas descobertas sobre as nossas origens e a nossa história, como uma espécie, e apontará para novas formas de se combater as doenças” (*Nature*, 2001). Com estas frases, os editores da revista *Nature* abriram a edição de um número histórico publicado em 15 de fevereiro de 2001, contendo os artigos que descrevem os achados derivados da análise do genoma humano. Esse número seminal deu origem à era da ciência pós-genômica.

Os artigos desse número da revista indicaram que o nosso genoma é formado por 2,9 a 3,2 bilhões de nucleotídeos. Destes, mais de 50% consistem de ele-

mentos repetitivos ou “parasitas” com função desconhecida. Apenas 1,1% a 1,4% de nosso genoma tem função conhecida, como codificador de proteínas. Cerca de 5% do nosso genoma consiste de grandes segmentos de DNA recentemente duplicados, incluindo vários genes. A iniciativa pública de seqüenciamento produziu uma lista de aproximadamente 22 mil genes humanos, enquanto a iniciativa privada encontrou cerca de 26 mil (IHGSC, 2001; Venter *et al.* 2001). Estes números podem ser comparados aos 6 mil genes de levedura, 13 mil de uma mosca (*Drosophila melanogaster*), 18 mil de um verme (*Caenorhabditis elegans*) e 26 mil de uma planta (*Arabidopsis thaliana*). Após o seqüenciamento completo, o número de genes preditos para a espécie humana foi mais baixo do que aqueles estimados previamente, os quais variam de 30 mil a 120 mil (Fields *et al.* 1994; Dunham *et al.*, 1999; Ewing e Green, 2000; Liang *et al.* 2000; Roest Crollius *et al.* 2000; Lander *et al.* 2001; Venter *et al.* 2001). A lista de genes preditos deve aumentar após o desenvolvimento de softwares mais adequados e sensíveis para a identificação gênica ou com o acúmulo de evidências experimentais futuras. A análise do conteúdo gênico não explica a nossa complexidade diante de outros organismos, e isso permanecerá como um desafio para o futuro.

A variação genética e sua relação com doenças humanas

Um dos mais importantes achados do Projeto Genoma Humano foi a determinação da diversidade de DNA existente entre diferentes indivíduos. Cada pessoa que já existiu no planeta - com exceção de gêmeos idênticos - possui um genoma único e, embora quaisquer dois genomas sejam ~99,9% idênticos, isso ainda deixa milhões de diferenças entre os 3,2 bilhões de pares de bases de nucleotídeos que compõem o genoma. O nível de identidade entre dois genomas é aproximadamente o mesmo, independentemente da origem étnica dos indivíduos. As diferenças genéticas que existem entre duas pessoas brancas de origem européia, ou entre um europeu branco e um asiático, é a mesma. Isso é válido para qualquer etnia e aniquila o conceito de raça.

São essas pequenas diferenças que resultam na variação herdada entre os indivíduos, incluindo suscetibilidade para determinadas doenças e as respostas a medicamentos. A maioria dessas diferenças aparece como substituições de um único nucleotídeo, ou SNPs (*Single Nucleotide Polymorphisms*). A alta fre-

quência dessas variações dentro da maioria, se não de todos os genes humanos (Lander and Weinberg, 2000) constitui a base da predisposição genética a doenças. Aproximadamente 90% de todas as variações são SNPs, e mais de 1,42 milhões de SNPs já foram identificados na época da primeira montagem disponível do genoma humano (*International SNP Map Working Group, 2001*). Esse número tem aumentado dramaticamente, e novas estratégias têm sido utilizadas com o objetivo de construir bibliotecas de todos os SNPs fenotipicamente importantes (Brookes, 1999; Cargill *et al.* 1999). Bancos de dados de polimorfismos são um grande benefício para os farmacogeneticistas e outros cientistas clínicos, particularmente aqueles que não possuem acesso às tecnologias genômicas e informáticas usadas na descoberta e avaliação de genes de interesse.

Enquanto os SNPs mais freqüentes (aqueles com uma freqüência alélica mínima de 40%) aparecem a cada 3.280 nucleotídeos do nosso genoma (com um número calculado de SNPs próximo a um milhão), alelos raros são mais comuns. Aproximadamente 11 milhões de bases de nosso genoma apresentam um pequeno índice de variação (alelos com freqüência mínima de 1%), mas ocorrem uma vez a cada 290 nt em nosso genoma. A combinação desses SNPs forma os genótipos. Os genótipos são únicos para cada indivíduo, e essa mesma combinação genética nunca se repete.

Um elemento-chave no quebra-cabeça das doenças genéticas é a compreensão da variação genética no genoma. O lado bom dos SNPs, especialmente nas doenças neuropsiquiátricas, consiste no seu baixo custo e relativa facilidade de estudo. Cada célula nucleada do nosso corpo possui o mesmo conteúdo de DNA. Este DNA é estável, independentemente da idade, uso de drogas, presença ou ausência de crises. Assim, não há necessidade de se manipular o tecido doente com o objetivo de analisar a presença de um SNP, e a simples coleta de um pequeno volume de sangue ou esfregaço bucal já permite a avaliação de centenas de polimorfismos.

O impacto da genômica no estudo da esquizofrenia

É esperado que as informações genômicas tenham uma grande extensão de aplicações na medicina. Centenas de projetos estão usando informação genética na descoberta de drogas, farmacogenômica, busca de alvos

para terapia gênica, estudos de envelhecimento, planejamento de vetores virais, administração de doenças, biologia regenerativa, clonagem, diagnóstico, prognóstico, pesquisa clínica, reprodução etc.

A seqüência do genoma humano é uma peça importante em estudos de doenças complexas como hipertensão, câncer, doença de Alzheimer ou a esquizofrenia. Entretanto, o enfoque genético para investigação dessas doenças complexas não deve ser baseado somente em alguns SNPs ou na investigação da expressão diferencial de alguns poucos genes. A exploração completa da complexidade genômica deve incluir estudos de processamento alternativo dos genes (*alternative splicing*), genômica comparativa, interação proteína-proteína e haplótipos correspondentes a grande número de SNPs, estudados de maneira seriada em um grande grupo de pacientes.

A esquizofrenia (ESQ), assim como outras doenças complexas, parece ser causada por uma série de fatores, incluindo ambiente e genética. Estudos de gêmeos e de adoção têm revelado a importância do componente genético. Entretanto, é claro que a ESQ não é uma entidade única, mas reflete uma sintomatologia comum causada por diversas anormalidades genéticas distintas. Como a maioria dos cânceres, a ESQ deve precisar de mais de um “hit” genético para que haja a manifestação clínica (Sawa e Snyder, 2002). Os genes que devem estar envolvidos estão distribuídos em muitas regiões do nosso genoma e, de acordo com isso, *loci* genéticos que parecem conferir suscetibilidade para desenvolver ESQ têm sido mapeados em diversos cromossomos, incluindo 1q21-22, 6p25, 8p21, 10p14, 13q32, 18p11 e 22q11-13 (Berrettini, 2000; Brzustowics *et al.* 2000; Straub *et al.* 1995; Blouin *et al.* 1998; Ekelund *et al.* 2001). Embora a lista de genes candidatos continue crescendo, nenhum gene candidato altamente confiável para ESQ apareceu até o momento.

Estudos de ligação e associação têm sido utilizados na busca por variações de DNA que sejam responsáveis pelo desenvolvimento de certas doenças. Os primeiros estudos de ligação em ESQ, feitos há cerca de 15 anos (Sherrington *et al.* 1988; Kennedy *et al.* 1988), assim como estudos recentes (Hovatta *et al.* 1999; Kendler *et al.* 2000; Baron, 2001) têm sido marcados por descobertas controversas. Numerosas regiões genômicas têm sido identificadas como suspeitas de possuir genes que predisõem para a doença; entretanto, essas regiões supostamente de suscetibilidade variam de estudo para estudo e de famílias para famílias, sugerindo um alto grau de heterogeneidade

da doença. Após o seqüenciamento do genoma humano e sua interpretação com estudos que focam determinada função gênica, estamos vivendo um período em que muitas informações, de natureza diferente, podem ser analisadas simultaneamente.

A seleção racional de genes candidatos para estudos de associação genética em ESQ tem sido baseada em evidências farmacológicas, neuroquímicas e clínicas que apontam para receptores específicos, enzimas e outras moléculas que possam estar envolvidas na etiopatogênese da doença. Exemplos são os genes codificadores de receptores de dopamina e serotonina que apresentam alta afinidade por agentes antipsicóticos. Diversas variações gênicas nesses receptores têm sido estudadas, entretanto, os resultados têm-se mostrado contraditórios. Essa seleção pode agora ser aprimorada se forem incorporadas as seqüências genômicas que não estavam disponíveis anteriormente. Quando as análises baseadas em localização e função gênica são conjugadas, podemos obter uma lista de genes de interesse, que podem estar mapeados em regiões genômicas importantes para a ESQ (Tabela 1).

Um bom exemplo é a busca por expansões de repetições trinucleotídicas, que podem estar associadas ao intrigante fenômeno de antecipação, frequentemente sugerido para a ESQ (McInnis *et al.* 1999). Até o momento, muitos esforços têm sido feitos na busca por essas repetições, utilizando-se inúmeras técnicas, incluindo a RED (*Repeat Expansion Detection*; Schalling *et al.* 1993), mas até o momento, os resultados dessas buscas são controversos e não foram capazes de identificar uma repetição inequivocamente associada com antecipação na ESQ. A seqüência do genoma humano oferece agora um mapa onde todas as repetições podem ser localizadas, selecionadas de acordo com aspectos de interesse (localização cromossomal, proximidade a genes de interesse etc.), e estudadas por técnicas de amplificação que utilizam quantidades ínfimas de DNA como molde.

Outras formas de busca incluem a investigação de regiões envolvidas em rearranjos cromossômicos, como translocações e deleções. Um bom exemplo é uma deleção de ~3 Mb no braço longo do cromossomo 22 que causa a síndrome velocardiofacial (VCFS) que é frequentemente acompanhada de sintomas psicóticos muito similares à ESQ. Esse achado indica que os genes localizados na região deletada podem modular o risco de desenvolvimento da doença. Os genes ou polimorfismos dessa região, associados com ESQ, ainda não foram identificados, e a enorme quantidade

Tabela 1: Genes ou regiões genômicas possivelmente associadas com a esquizofrenia

Localização cromossômica	Gene candidato	Descrição	Função gênica	Marcador	Referência
1q21-q22	KCNN3	Potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3	Proteína integral de membrana que forma um canal de cálcio dependente de voltagem com três outras subunidades de ligação à calmodulina	D1S1653-D1S1679	Brzustowicz <i>et al.</i> (2000)
1q23	RGS4	Regulator of G-protein signalling 4	Membros da família de reguladores de sinalização de proteína G, RGSs são moléculas regulatórias que agem como proteínas ativadoras de GTPases (GAPs)		Brzustowicz <i>et al.</i> (2000) Mirnic <i>et al.</i> (2001)
1q23.3	–	–	–	D1S196	Gurling <i>et al.</i> (2001)
2p13-p14	–	–	–	D2S358	Camp <i>et al.</i> (2001)
2q12-q13	–	–	–	D2S135	Moises <i>et al.</i> (1995)
2q37	–	–	–	D2S427	Paunio <i>et al.</i> (2001)
3q13.3	DRD3	Dopamine receptor D3	Este gene codifica o subtipo D3 do receptor de dopamina. Inibe adenilil ciclase através de proteínas G inibitórias e é expresso em regiões cerebrais envolvidas com emoção e funções cognitivas		Jonsson <i>et al.</i> (2003)
4p16.1-p15.3	DRD5	Dopamine receptor D5	Este gene codifica o subtipo D5 do receptor de dopamina, um receptor acoplado à proteína G que estimula a atividade de adenilil ciclase. É expresso em neurônios de regiões límbicas do cérebro	D4S615	Muir <i>et al.</i> (2001)
5p14.1-13.1	–	–	–	D5S111	Silverman <i>et al.</i> (1996)
5q32-q33	–	–	–	D5S422	Gurling <i>et al.</i> (2001)
6p21.3	NOTCH4	Notch homolog 4 (<i>Drosophila</i>)	Membros dessa família de proteínas transmembranas do tipo 1 possuem características estruturais comuns, incluindo um domínio extracelular que consiste de repetições semelhantes ao fator de crescimento epidermal (EGF), e um domínio intracelular que consiste de múltiplos tipos de domínios diferentes. NOTCH estão envolvidos com diversos processos de desenvolvimento através do controle de destino celular		Wassink <i>et al.</i> (2003) Skol <i>et al.</i> (2003)
6p23	–	–	–	D6S309-D6S1578	Lindholm <i>et al.</i> (1999)
6q25	–	–	–	D6S253	Lindholm <i>et al.</i> (2001)
8p21-p12	NRG1	Neuregulin 1	Neuregulin 1 (NRG1) foi originalmente identificada como uma glicoproteína de 44-kD que interage com o receptor tirosina quinase NEU/ERBB2, aumentando sua fosforilação em resíduos de tirosina. Pelas interações com receptores ERBB, as isoformas de NRG1 induzem o crescimento e diferenciação de células epiteliais, neuronais, gliais, e outros tipos		Stefansson <i>et al.</i> (2002)
11p15.5	DRD4	Dopamine receptor D4			Seeman <i>et al.</i> (1993) Lung <i>et al.</i> (2002)
11q23	DRD2	Dopamine receptor D2	Este gene codifica o subtipo D2 do receptor de dopamina, um receptor acoplado à proteína G que inibe a atividade de adenilil ciclase		Noble (2003)

continua

continuação

Localização cromossômica	Gene candidato	Descrição	Função gênica	Marcador	Referência
12q24	DAO	D-amino-acid oxidase	Este gene codifica a enzima peroxissomal D-aminoácido oxidase, com função biológica não conhecida; pode estar envolvido no balanço ácido-básico do rim		Chumakov <i>et al.</i> (2002)
13q12-22	–	–	–	D13S894	Camp <i>et al.</i> (2001)
13q31-q32	–	–	–	D13S793	Brzustowicz <i>et al.</i> (2000) Blouin <i>et al.</i> (1998)
13q32	–	–	–	D13S174	Blouin <i>et al.</i> (1998)
13q33.1	G72	Putative protein LG72	G72 e DAO podem interagir, resultando na ativação de DAO. DAO é expressa no cérebro, onde oxida D-serina, um potente ativador de um receptor de glutamato do tipo N-metil-D-aspartato		Chumakov <i>et al.</i> (2002)
15q14	CHRNA7	Cholinergic receptor, nicotinic, alpha polypeptide 7	Os receptores nicotínicos de acetilcolina (nAChRs) são membros da superfamília de canais de íons dependentes de ligantes que mediam a rápida transmissão de sinal das sinapses	D15S118	Freedman <i>et al.</i> (2001)
15q15	–	–	–	D15S1042- D15S659	Stober <i>et al.</i> (2000) Stober <i>et al.</i> (2002)
18p	GNAL	Guanine nucleotide binding protein (G protein), alpha activating activity polypeptide, olfactory type	A subunidade alfa de proteína G é componente de complexos heterotrimericos de proteína G. O heterotrímero sinaliza de receptores acoplados à proteína G para efetores intracelulares		Lara <i>et al.</i> (2001)
21q22.3	S100β	S100 calcium binding protein, beta (neural)	S100B é uma proteína de ligação a cálcio produzida e secretada por células gliais. Tem sido utilizada como marcador de danos cerebrais		Lara <i>et al.</i> (2001)
22q11.21	PRODH	Proline dehydrogenase (oxidase) 1	PRODH2 codifica a prolina desidrogenase, uma enzima mitocondrial que converte prolina em Δ ¹ -pirrolino-5-carboxilato. Está envolvida na transferência de potenciais redox através da membrana mitocondrial		Liu <i>et al.</i> (2002)
	DGCR6	DiGeorge syndrome critical region gene 6	O produto gênico é similar à proteína gonadal de <i>Drosophila</i> , que participa no desenvolvimento de células germinativas. Forma, em conjunto com a cadeia gama-1 de laminina humana, uma molécula de laminina, que se liga às células através de interações com um receptor. Tem funções relacionadas com ligação celular, migração e organização tecidual durante o desenvolvimento		Liu <i>et al.</i> (2002)
	COMT	Catechol-O-methyltransferase	A Catecol-O-metiltransferase cataliza a transferência de um metil S-adenosilmetionina para catecolaminas, incluindo dopamina, epinefrina, e norepinefrina. Esta Ometilação resulta em uma importante via de degradação de transmissores de catecolaminas	D22S315	Shifman <i>et al.</i> (2002)
22q11-q13	–	–	–	D22S278 D22S279- D22S276	Schwab & Wildenauer (1999) Jorgensen <i>et al.</i> (2002)

de de dados disponíveis atualmente oferece a possibilidade de identificação de blocos genômicos conservados nessa região, que é perdida, e que possivelmente contém seqüências de DNA fisiologicamente conservadas.

Apesar dos enormes esforços para identificar genes de suscetibilidade, os resultados de estudos de genética molecular de ESQ até o momento têm sido modestos. Isso pode ser explicado por diversas razões. Devemos ressaltar que fatores ambientais possuem um enorme impacto no estabelecimento da doença. O impacto ambiental é tão forte quanto o apresentado pelo DNA, e pode distorcer dramaticamente os achados genéticos. O padrão de herança da ESQ não está completamente esclarecido e geralmente não segue um modelo de herança Mendeliano, sendo que a maioria dos casos é esporádica (Gottesman e Shields, 1976). Além disso, a variação clínica que encontramos na ESQ e as ferramentas diagnósticas disponíveis atualmente, baseadas em questionários e presença/ausência de uma série de sintomas, podem não ser tão adequadas e precisas como requerem os estudos de genética molecular.

Da mesma forma que são obtidos progressos em neuroquímica e tecnologias de imagem para a análise da ESQ, também podemos contar com novas descobertas derivadas de estudos genéticos. O uso apropriado da genômica poderá ajudar imensamente na elucidação das causas da ESQ, e os caminhos a serem seguidos devem focar uma série de aspectos, tais como:

- Descoberta gênica: regiões genômicas amplas, possivelmente envolvidas com a doença, já foram seqüenciadas. Entretanto, a busca por genes utilizando métodos computacionais não é suficiente, e um estudo detalhado de potenciais regiões codificadoras e mapeamento de dados é necessário para uma melhor análise do conteúdo no genoma humano e para a identificação de genes e regiões regulatórias potencialmente envolvidas. A genômica comparativa usando DNA de outros primatas deverá evidenciar regiões regulatórias importantes, cruciais para a regulação da expressão de nosso genoma.

- A compreensão de variações genéticas (SNPs) e do seu papel na gênese das doenças e na resposta aos quimioterápicos.

- Análise de isoformas resultantes de processamento gênico alternativo: estas novas formas gênicas já foram reveladas em cerca de 50% dos genes humanos. Este estudo em ESQ é uma área promissora.

- Análise de expressão gênica diferencial: estas análises, feitas com o uso de técnicas tais como “DNA

microarrays”, RDA e SAGE demonstraram ser poderosas para a análise de doenças complexas como o câncer e a hipertensão. A utilização dessas técnicas está em fase inicial no estudo da ESQ.

- Estudo de vias metabólicas de interesse: a seleção de genes pertencentes a vias metabólicas como as do metabolismo de fosfolípidos, neurogênese ou sinaptogênese, em conjunto com informações de mapeamento, permitirão a seleção de genes candidatos para estudos de associação, baseados em uma soma de critérios de seleção.

Um dos maiores problemas enfrentados pelas investigações biomédicas atuais, sendo também importante para a ESQ, é o fato de as populações apresentarem *backgrounds* étnicos diferentes, com freqüências alélicas distintas entre os diversos polimorfismos genéticos. A variação étnica também influencia estudos de haplótipos, visto que blocos genômicos podem variar dramaticamente entre esses grupos. Além disso, essas freqüências alélicas podem variar entre indivíduos do sexo masculino e feminino, assim como entre diferentes faixas etárias. Quando controles não são bem pareados por etnia, idade e sexo, há uma grande chance de a evidência de uma associação genética ser um falso-positivo. Por outro lado, é possível que uma verdadeira associação não seja válida e constante entre grupos étnicos distintos. Após a globalização da economia e o aumento do fluxo migratório que ocorreu por todo o planeta, a definição de um grupo étnico preciso torna-se cada vez mais difícil, se não artificial. Enquanto existe um imenso valor na identificação de associações para grupos étnicos, a definição de marcadores robustos de associação, válidos para todas as populações, deve ser considerada um dos mais importantes alvos na pesquisa bioquímica moderna.

Cruzando informações para selecionar genes candidatos

Em ESQ já é bem estabelecido que a desordem primária é causada por uma falha no desenvolvimento do sistema nervoso e não por um processo neurodegenerativo. Uma análise do genoma humano, usando o Banco de Ontologia dos Genes (Gene Ontology Bank - fevereiro de 2003 - <http://www.godatabase.org/cgi-bin/go.cgi>) indica a presença de 366 genes já relacionados à neurogênese (GO: 0007399). Adicionalmente, uma análise da literatura atual indica que pelo menos 27 regiões genômicas podem ser associadas à ESQ (Tabela).

Uma abordagem que tem sido utilizada em nosso laboratório é a combinação de vários dados, incluindo mapeamento e função gênica, com o objetivo de selecionar genes candidatos para proceder aos estudos de associação. Depois de avaliar a lista dos genes envolvidos com neurogênese e os dados das regiões genômicas envolvidas com ESQ, foi possível gerar uma lista com 100 genes de neurodesenvolvimento que estão localizados em regiões que foram identificadas como associadas à doença. Focando neste grupo de genes, fizemos buscas em bancos de dados a fim de identificar possíveis polimorfismos, tais como SNPs ou microssatélites de DNA, priorizando aquelas alterações que pudessem modificar a proteína a ser sintetizada. Após um passo de validação eletrônica, no qual a presença do polimorfismo foi avaliada em seqüências de bancos de dados, os polimorfismos confirmados seguiram para a validação experimental em amostras de DNA de 200 indivíduos (400 alelos). Todos os 18 possíveis polimorfismos que foram escolhidos para os passos de validação eletrônica foram confirmados por seqüências de bancos de dados e, para todos aqueles onde foi possível desenhar os iniciadores para os experimentos de validação experimental (no total de 11), a variação foi confirmada.

Juntamente com a busca por novos polimorfismos, nosso banco de DNA está em contínua expansão a fim de ampliarmos nossos grupos amostrais e produzir dados cada vez mais robustos. Adicionalmente, um software para identificação de polimorfismos foi desenvolvido em colaboração com a Empresa Brasileira Scylla Bioinformática (www.scylla.com.br), o qual também se encontra em contínuo aperfeiçoamento. Nosso grupo também desenvolveu um detalhado sistema para armazenamento dos dados clínicos e moleculares de todas as amostras, tanto pacientes e controles, incluindo dados referentes à extração de DNA, quantificação, qualidade e dados de estoque da amostra, a contínua inclusão de dados de genotipagem para cada amostra de DNA. Nosso objetivo é possuir uma coleção de DNA, em que a contínua inclusão dos dados de genotipagem possa ajudar-nos a construir um mapa fino das alterações e marcadores presentes em determinadas patologias.

Um grande grupo de amostras de pacientes no banco garante que os fenótipos clínicos mais diversos estejam bem representados, permitindo a estratificação dos pacientes sem a perda de poder estatístico durante as análises realizadas. A estratificação será importante para cobrir diferentes formas da doença, bem como para dados relevantes em estudos de farmacogenômica. A maior vantagem de possuímos um grande número de genes para serem avaliados nas mesmas amostras de

pacientes é que podemos avaliar o efeito cumulativo ou mesmo sinérgico de vários genes ou variações de menor efeito. Novos genes podem facilmente ser adicionados a este sistema, utilizando informações derivadas de *microarrays* de cDNA ou de dados oriundos de SAGE que possam sugerir a expressão alterada de alguns genes.

Perspectivas

A convergência de dados bioquímicos, de imagem, neuroanatômicos, psicofarmacológicos, clínicos e genéticos permite prever que estaremos muito próximos de uma melhor compreensão das bases biológicas da ESQ. O que podemos esperar no campo da genética nos próximos 15 a 20 anos? Embora todas as possibilidades estejam um pouco distantes para serem discutidas aqui, alguns pontos parecem ser relativamente certos:

- Uma lista satisfatória dos produtos dos genes humanos poderá fornecer um vasto conjunto de medicamentos potenciais similares a proteínas humanas (tal como insulina, interferons e hormônios de crescimento);
- Os futuros cadastros de dados médicos deverão incluir uma grande quantidade de dados genéticos e, eventualmente, o genoma completo da pessoa, associado a uma longa lista de polimorfismos que possam ser usados para prever resposta a determinadas drogas, a substâncias químicas e ambientais, juntamente com a predisposição do indivíduo;
- A compreensão das bases genéticas de doenças complexas irá permitir o desenvolvimento de estratégias de prevenção para impedir o estabelecimento da patologia e o desenvolvimento de drogas para um tratamento mais eficiente;
- Empresas de farmacogenômica irão desenvolver, utilizando a individualidade genética, uma medicina personalizada na qual drogas serão geradas para atender às necessidades específicas dos indivíduos.

A disponibilidade destes avanços terá um enorme impacto na pesquisa da ESQ. A fim de avaliar os mais diferentes aspectos desta doença, é extremamente desejável a criação de um grupo internacional de pesquisadores que possam contribuir com amostras e dados, incluindo aí os mais diversos aspectos relacionados à doença. Uma base de dados unificada, envolvendo as diferentes disciplinas já mencionadas e associada a um banco de DNA de indivíduos de diferentes etnias, irá permitir um importante salto na pesquisa em ESQ.

Agradecimentos

Os autores agradecem aos Drs. Juliana Yacubian, Hildeberto Tavares e Paulo Sallet pelas amostras e dados clínicos, e ao Prof. Wagner Farid Gattaz pelo contínuo apoio ao grupo de genética do LIM27. O Laboratório de Neurociências (LIM27) recebe suporte financeiro da Associação Beneficente Alzira Denise Hertzog Silva (ABADHS), Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) e Conselho Nacional de Pesquisas (CNPq). Os autores reconhecem o importante apoio da empresa Applied Biosystems do Brasil.

Referências bibliográficas

- BARON, M. - Genetics of schizophrenia and the new millennium: progress and pitfalls. *Am J Hum Genet* 68: 299-312, 2001.
- BERRETTINI, W.H. - Are schizophrenic and bipolar disorders related? A review of family and molecular studies. *Biol Psychiatry* 48: 531-8, 2000.
- BLOUIN, J.L.; DOMBROSKI, B.A.; NATH, S.K. ET AL - Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nature Genet* 20:70-3, 1998.
- BROOKES, A.J. - The essence of SNPs. *Gene* 234:177-86, 1999.
- BROOKES, A.J.; LEHVASLAIHO, H.; SIEGFRIED, M.; BOEHM, J.G.; YUAN, Y.P.; SARKAR, C.M.; BORK, P.; ORTIGAO, F. - HGBASE: a database of SNPs and other variations in and around human genes. *Nucleic Acids Res* 28: 356-60, 2000.
- BRZUSTOWICZ, L.M.; HODGKINSON, K.A.; CHOW, E.W.; HONER, W.G.; BASSETT, A.S. - Location of a major susceptibility locus for familiar schizophrenia on chromosome 1q21-q22. *Science* 288: 678-82, 2000.
- BUETOW, K.H.; EDMONSON, M.N.; CASSIDY, A.B. - Reliable identification of large numbers of candidate SNPs from public EST data. *Nature Genet* 21: 323-5, 1999.
- CAMP, N.J.; NEUHAUSEN, S.L.; TIOBECH, J. ET AL - Genomewide multipoint linkage analysis of seven extended Palauan pedigrees with schizophrenia, by a Markov-Chain Monte Carlo method. *Am J Hum Genet* 69: 1278-89, 2001.
- CARGILL, M.; ALTSHULER, D.; IRELAND, J.; SKLAR, P.; ARDLIE, K.; PATIL, N.; SHAW, N.; LANE, C.R.; LIM, E.P.; ET AL. - Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genet* 22: 231-8, 1999.
- CHUMAKOV, I.; BLUMENFELD, M.; GUERASSIMENKO, O.; CAVAREC, L.; PALICIO, M.; ABDERRAHIM, H.; ET AL. - Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia. *Proc Natl Acad Sci* 99: 13675-80, 2002.
- DUNHAM, I.; SHIMIZU, N.; ROE, B.A.; CHISSOE, S.; HUNT, A.R.; COLLINS, J.E.; BRUSKIEWICH, R.; BEARE J. ET AL. - The DNA sequence of human chromosome 22. *Nature* 402: 489-95, 1999.
- EKELUND, J.; HOVATTA, I.; PARKER, A.; PAUNIO, T.; VARILLO, T.; MARTIN R.; SUHONEN J.; ELLONEN P.; ET AL. - Chromosome 1 loci in Finnish schizophrenia families. *Hum Mol Genet* 10: 1611-7, 2001.
- EVANS, W.E.; RELING, M.V. - Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* 286: 487-91, 1999.
- EWING, B.; GREEN, P. - Analysis of expressed sequence tags indicates 35,000 human genes. *Nature Genet* 25: 232-4, 2000.
- FIELDS, C.; ADAMS, M.D.; WHITE, O.; VENTER, J.C. - How many genes are in the human genome? *Nature Genet* 7: 345-6, 1994.
- FREEDMAN, R.; LEONARD, S.; GAULT, J.M.; HOPKINS, J.; CLONINGER, C.R.; KAUFMANN, C.A.; TSUANG, M.T.; FARONE, S.V.; MALASPINA, D.; SVRAKIC, D.M.; SANDERS, A.; GEJMAN, P. - Linkage disequilibrium for schizophrenia at the chromosome 15q13-14 locus of the alpha7-nicotinic acetylcholine receptor subunit gene (CHRNA7). *Am J Med Genet* 105: 20-2, 2001.
- GOTTESMAN, I.I.; SHIELDS, J. - A critical review of recent adoption, twin, and family studies of schizophrenia: behavioral genetics perspectives. *Schizophr Bull* 2: 360-401, 1976.
- GURLING, H.M.; KALSI, G.; BRYNJOLFSSON, J.; SIGMUNDSSON, T.; SHERRINGTON, R.; MANKOO, B.S.; READ, T.; MURPHY, P.; BLAVERI, E.; MCQUILLIN, A.; PETURSSON, H.; CURTIS, D. - Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21-22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3-24 and 20q12.1-11.23. *Am J Hum Genet* 68: 661-73, 2001.
- HOVATTA, I.; VARILLO, T.; SUVISAARI, J.; TERWILLIGER, J.D.; OLLIKAINEN, V.; ARAJARVI, R.; JUVONEN, H.; KOKKO-SAHIN, M.L.; VAISANEN, L.; MANNILA, H.; LONNQVIST, J.; PELTONEN, L. - A genomewide screen for schizophrenia genes in an isolated Finnish subpopulation, suggesting multiple susceptibility loci. *Am J Hum Genet* 65:1114-24, 1999.
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM - Initial sequencing and analysis of the human genome. *Nature* 409: 860-921, 2001.
- INTERNATIONAL SNP MAP WORKING GROUP - A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409: 928-33, 2001.
- JONSSON, E.G.; FLYCKT, L.; BURGERT, E.; CROCO, M.A.; FORSLUND, K.; MATTILA-EVENDEN, M.; RYLANDER, G.; ASBERG, M.; NIMGAONKAR, V.L.; EDMAN, G.; BJERKENSTEDT, L.; WIESEL, F.A.; SEDVALL, G.C. - Dopamine D3 receptor gene Ser9Gly variant and schizophrenia: association study and meta-analysis. *Psychiatr Genet* 13: 1-12, 2003.
- JORGENSEN, T.H.; BORGLUM, A.D.; MORS, O.; WANG, A.G.; PINAUD, M.; FLINT, T.J.; DAHL, H.A.; VANG, M.; KRUSE, T.A.; EWALD, H. - Search for common haplotypes on chromosome 22q in patients with schizophrenia or bipolar disorder from the Faroe Islands. *Am J Med Genet* 114: 245-52, 2002.
- KENDLER, K.S.; MYERS, J.M.; O'NEILL, F.A.; MARTIN, R.; MURPHY, B.; MACLEAN, C.J.; WALSH, D.; STRAUB, R.E. - Clinical features of schizophrenia and linkage to chromosomes 5q, 6p, 8p, and 10p in the Irish Study of High-Density Schizophrenia Families. *Am J Psychiatry* 157: 402-8, 2003.
- KENNEDY, J.L.; GIUFFRÀ, L.A.; MOISES, H.W.; CAVALLI-SFORZA, L.L.; PAKSTIS, A.J.; KIDD, J.R.; CASTIGLIONE, C.M.; SJOGREN, B.; WETTERBERG, L.; KIDD, K.K. - Evidence against linkage of schizophrenia to markers on chromosome 5 in a northern Swedish pedigree. *Nature* 336: 167-70, 1988.
- LANDER, E.S.; WEINBERG, R.A. - Genomics: journey to the center of biology. *Science* 287: 1777-82, 2000.
- LANDER, E.S.; LINTON, L.M.; BIRREN, B.; NUSBAUM, C.; ZODY, M.C.; BALDWIN, J.; DEVON, K.; DEWAR, K.; DOYLE, M.; FITZHUGH, W. ET AL - Initial sequencing and analysis of the human genome. *Nature* 409: 860-921, 2001.
- LARA, D.R.; GAMA, C.S.; BELMONTE-DE-ABREU, P.; PORTELA, L.V.; GONCALVES, C.A.; FONSECA, M.; HAUCK, S.; SOUZA, D.O. - Increased serum S100B protein in schizophrenia: a study in medication-free patients. *J Psychiatr Res* 35: 11-4, 2001.
- LIANG, F.; HOLT, I.; PERTEA, G.; KARAMYCHEVA, S.; SALZBERG, S.L.; QUACKENBUSH, J. - Gene index analysis of the human genome estimates approximately 120,000 genes. *Nature Genet* 25: 239-40, 2000.
- LINDHOLM, E.; EKHMOLM, B.; BALCIUNIENE, J.; JOHANSSON, G.; CASTENSSON, A.; KOISTI, M.; NYLANDER, P.O.; PETERSSON, U.; ADOLFSSON, R.; JAZIN, E. - Linkage analysis of a large Swedish kindred provides further support for a susceptibility locus for schizophrenia on chromosome 6p23. *Am J Med Genet* 88: 369-77, 1999.
- LINDHOLM, E.; EKHMOLM, B.; SHAW, S.; JALONEN, P.; JOHANSSON, G.; PETERSSON, U.; SHERRINGTON, R.; ADOLFSSON, R.; JAZIN, E. - A schizophrenia susceptibility locus at 6q25, in one of the world's largest reported pedigrees. *Am J Hum Genet* 69: 96-105, 2001.
- LIU, H.; HEATH, S.C.; SOBIN, C.; ROOS, J.L.; GALKE, B.L.; BLUNDELL, M.L.; LENANE, M.; ROBERTSON, B. ET AL. - Genetic variation at the 22q11 PRODH2/DGCR6 locus presents an unusual pattern and increases susceptibility to schizophrenia. *Proc Natl Acad Sci USA* 99: 3717-22, 2002.
- LUNG, F.W.; TZENG, D.S.; SHU, B.C. - Ethnic heterogeneity in allele variation in the DRD4 gene in schizophrenia. *Schizophr Res* 57: 239-45, 2002.

- MARSH, S.; KWOK, P.; MCLEOD, H.L. - SNP Databases and pharmacogenetics: Great Start, but a Long Way to Go. *Human Mut* 20: 174-9, 2002.
- MCINNIS, M.G.; MCMAHON, F.J.; CROW T.; ROSS C.A.; DELISI L.E. - Anticipation in schizophrenia: a review and reconsideration. *Am J Med Genet* 88: 686-93, 1999.
- MCLEOD, H.L.; EVANS, W.E. - Pharmacogenomics: unlocking the human genome for better drug therapy. *Ann Rev Pharmacol Tox* 41: 101-21, 2001.
- MIRNICS, K.; MIDDLETON, F.A.; STANWOOD, G.D.; LEWIS, D.A.; LEVITT, P. - Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia. *Mol Psychiatry* 6: 293-301, 2001.
- MOISES, H.W.; YANG, L.; KRISTBJARNARSON, H.; WIESE, C.; BYERLEY, W.; MACCIARDI, F.; AROLT, V.; BLACKWOOD, D.; LIU, X.; SJOGREN, B.; ET AL - An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nature Genet* 11: 321-4, 1995.
- MUIR, W.J.; THOMSON, M.L.; MCKEON, P.; MYNETT-JOHNSON, L.; WHITTON, C.; EVANS, K.L.; PORTEOUS, D.J.; BLACKWOOD, D.H. - Markers close to the dopamine D5 receptor gene (DRD5) show significant association with schizophrenia but not bipolar disorder. *Am J Med Genet* 105: 152-8, 2001.
- NATURE - Human genomes, public and private. *Nature* 409: 745, 2001.
- NOBLE, E.P. - D2 dopamine receptor gene in psychiatric and neurologic disorders and its phenotypes. *Am J Med Genet* 116: 103-25, 2003.
- PAUNIO, T.; EKELUND, J.; VARILLO, T. ET AL - Genome-wide scan in a nationwide study sample of schizophrenia families in Finland reveals susceptibility loci on chromosomes 2q and 5q. *Hum Mol Genet* 10: 3037-48, 2001.
- ROEST-CROLLIUS, H.; JAILLON, O.; BERNOT, A.; DASILVA, C.; BOUNEAU, L.; FISCHER, C.; FIZAMES, C.; WINCKER, P.; BROTTIER, P.; QUETIER, F.; SAURIN, W.; WEISSENBACH, J. - Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence. *Nature Genet* 25: 235-8, 2000.
- SAWA, A.; SNYDER, S.H. - Schizophrenia: diverse approaches to a complex disease. *Science* 296: 692-5, 2002.
- SCHALLING, M.; HUDSON, T.J.; BUETOW, K.H.; HOUSMAN, D.E. - Direct detection of novel expanded trinucleotide repeats in the human genome. *Nature Genet* 4: 135-9, 1993.
- SCHWAB, S.G.; WILDENAUER, D.B. - Chromosome 22 workshop report. *Am J Med Genet* 88: 276-8, 1999.
- SEEMAN, P.; GUAN, H.C.; VAN TOL, H.H.M. - Dopamine D4 receptors elevated in schizophrenia. *Nature* 365: 441-5, 1993.
- SHERRINGTON, R.; BRYNJOLFSSON, J.; PETURSSON, H.; POTTER, M.; DUDLESTON, K.; BARRACLOUGH, B.; ET AL. - Localization of a susceptibility locus for schizophrenia on chromosome 5. *Nature* 336:164-7, 1988.
- SHIFMAN, S.; BRONSTEIN, M.; STERNFELD, M.; PISANTE-SHALOM, A.; LEV-LEHMAN, E.; WEIZMAN, A. ET AL. - A highly significant association between a COMT haplotype and schizophrenia. *Am J Hum Genet* 71: 1296-302, 2002.
- SILVERMAN, J.M.; GREENBERG, D.A.; ALTSTIEL, L.D.; SIEVER, L.J.; MOHS, R.C.; SMITH, C.J.; ET AL. - Evidence of a locus for schizophrenia and related disorders on the short arm of chromosome 5 in a large pedigree. *Am J Med Genet* 67: 162-71, 1996.
- STEFANSSON, H., SIGURDSSON E.; STEINTHORSDOTTIR V.; BJORNSDOTTIR, S.; SIGMUNDSSON, T.; GHOSH, S.; BRYNJOLFSSON, J.; GUNNARSDOTTIR, S.; IVARSSON, O.; CHOU, T.T.; HJALTASON, O.; ET AL. - Neuregulin 1 and susceptibility to schizophrenia. *Am J Hum Genet* 71: 877-92, 2002.
- STOBER, G.; SAAR K.; RUSCHENDORF, F.; MEYER, J.; NURNBERG, G.; JATZKE, S. ET AL - Splitting schizophrenia: periodic catatonia-susceptibility locus on chromosome 15q15. *Am J Hum Genet* 67: 1201-7, 2000.
- STOBER, G.; SEELow, D.; RUSCHENDORF, F.; EKICI, A.; BECKMANN, H.; REIS, A. - Periodic catatonia: confirmation of linkage to chromosome 15 and further evidence for genetic heterogeneity. *Hum Genet* 111: 323-30, 2002.
- STRAUB, R.E.; MACLEAN, C.J.; O'NEILL, F.A.; BURKE, J.; MURPHY, B.; DUKE, F.; SHINKWIN, R.; WEBB, B.T.; ZHANG, J.; WALSH, D. ET AL - A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. *Nature Genet* 11: 287, 1995.
- VENTER, J.C.; ADAMS, M.D.; MYERS, E.W. ET AL - The sequence of the human genome. *Science* 291: 1304-51, 2001.
- WASSINK, T.H.; NOPOULOS, P.; PIETILA, J.; CROWE, R.R.; ANDREASEN, N.C. - NOTCH4 and the frontal lobe in schizophrenia. *Am J Med Genet* 118: 1-7, 2003.