

Projeto Genoma Humano e suas implicações para a saúde humana: visão geral e contribuição brasileira para o projeto

Andrew J.G. Simpson; Otávia L. Caballero

A informação derivada do Projeto Genoma Humano, um esforço internacional para elucidar e caracterizar a seqüência completa de 3×10^9 pares de bases do genoma humano, vai revolucionar a prática da medicina no séc. XXI por prover instrumentos para determinar o componente genético de virtualmente todas as doenças. As conseqüências para a prática da medicina deverão ser profundas, tais como a geração de medidas que possam prever o risco, permitir o diagnóstico precoce e promover estratégias de tratamento mais efetivas. Neste artigo, os autores discutem as aplicações potenciais do conhecimento adquirido com o HGP, bem como a contribuição brasileira a esse projeto.



Andrew J.G. Simpson
Ph.D.; chefe do Laboratório de Genética do Câncer, Instituto Ludwig de Pesquisa sobre o Câncer, São Paulo

Otávia L. Caballero
Médica; Ph.D.; pesquisadora científica, Hospital do Câncer AC Camargo, São Paulo



Unitermos:

Projeto Genoma Humano, microarrays, genes

INTRODUÇÃO

O termo genoma refere-se ao conjunto completo de genes e cromossomos de um organismo. O mapeamento, seqüenciamento e análise dos genomas é chamado de genômica. A análise do genoma pode ser dividida em genômica estrutural e funcional. A genômica estrutural é a fase inicial que tem o objetivo da construção de mapas genéticos, físicos e de transcrição de um organismo. A genômica funcional, que se baseia na expressão gênica, usando as informações geradas pela genômica estrutural, leva à completa caracterização do padrão de expressão do conjunto completo dos genes, assim como à investigação sistemática das propriedades funcionais desse conjunto de genes.

A informação derivada do Projeto Genoma Humano

(Human Genome Project, HGP), um esforço internacional para elucidar e caracterizar a seqüência completa de 3×10^9 pares de bases do genoma humano, vai revolucionar a prática da medicina no séc. XXI por prover instrumentos para determinar o componente genético de virtualmente todas as doenças. A recente publicação da seqüência e análise iniciais do genoma humano, realizada simultaneamente por dois grupos, o Consórcio Internacional para o Seqüenciamento do Genoma Humano, liderado por Francis Collins, e pela companhia Celera, liderada por Craig Venter, marca, sem dúvida, um importante passo em direção à ciência pós-genômica(1,2). Embora muitas questões ainda permaneçam por ser respondidas, vários achados são de extremo interesse para a comunidade científica. Apenas 1,1 a 1,4% do genoma realmente codifica proteínas, sendo que 75% do DNA se situa entre os genes. Duplicações de segmentos, de tamanhos variáveis, são abundantes em todo o genoma e revelam uma história evolucionária complexa. Foram identificados entre 1,4 e 2,1 milhões de polimorfismos envolvendo um único nucleotídeo (single-nucleotide polymorphisms, SNPs). Dois genomas haplóides diferem em média 1 bp em cada 1.250. Menos de 1% de todos SNPs resultam em variações nas proteínas, mas a determinação dos SNPs que apresentam conseqüências funcionais permanece ainda por ser determinada. Segundo estimativas dos dois grupos, o número de genes do genoma humano varia entre 30 e 40 mil. Embora esse número seja apenas duas vezes maior que o de *Drosophila*, por exemplo, os genes humanos parecem ser muito mais complexos e apresen-

tam um número maior de produtos gerados a partir do mesmo gene.

As conseqüências para a prática da medicina deverão ser profundas, tais como a geração de medidas que possam prever o risco, permitir o diagnóstico precoce e promover estratégias de tratamento mais efetivas. A predição genética de riscos individuais de doença e resposta a drogas provavelmente estará ao alcance dos médicos na próxima década, bem como o desenvolvimento de drogas tendo como alvos vias moleculares alteradas em diferentes doenças. Como a pesquisa biomédica está rapidamente definindo os mecanismos moleculares dos efeitos farmacológicos, assim como os determinantes genéticos da patogênese das doenças e polimorfismos funcionalmente importantes em genes que governam o metabolismo das drogas, o planejamento e condução de testes clínicos e terapêuticos das doenças deverão sofrer mudanças radicais.

O estudo da variação genética humana, particularmente a distribuição dos polimorfismos em únicos nucleotídeos entre pacientes afetados e não-afetados por determinadas doenças, vai prover informações valiosas sobre a contribuição do componente genético do paciente a doenças como câncer, diabetes ou doenças mentais.

Além disso, a introdução da tecnologia de análise da expressão gênica usando-se "microarrays" facilitará a aplicação de abordagens experimentais nas quais alterações globais no nível de expressão de milhares de genes podem ser estudadas em um único experimento(3). Nos "microarrays" de c-DNAs, moléculas sin-

tetizadas a partir do RNA, que correspondem às moléculas que estão sendo transcritas na célula, aos milhares, são imobilizadas em superfícies de suportes sólidos. Como lâminas de vidro. Em seguida, c-DNAs marcados obtidos do paciente e referência normal são sintetizados e marcados com corantes fluorescentes diferentes e hibridizados no mesmo array. A relação entre a fluorescência das sondas hibridizadas detectadas por scanners identifica os genes com aumento ou diminuição de expressão em relação ao controle. Na próxima década, os "microarrays" de c-DNA, combinados com bancos de dados genômicos e sofisticadas técnicas de bioinformática, vão prover perfis específicos de cada processo patológico, bem como esclarecer prognóstico e orientar o tratamento.

Neste artigo, discutiremos as aplicações potenciais do conhecimento adquirido com o HGP, e a contribuição brasileira a esse projeto.

1. Farmacogenômica

A farmacogenômica envolve a aplicação de tecnologias como o sequenciamento de DNA, análise da expressão gênica e estatística em pesquisas e testes clínicos de drogas(4). Como muitas doenças podem resultar da alteração de uma rede de genes em diferentes vias, a farmacogenômica poderia identificar quais genes estariam envolvidos na determinação da resposta a determinada droga. Desse modo, a caracterização genética de populações de pacientes deverá ser parte integral do processo de descoberta e

desenvolvimento de drogas. Provavelmente, a seleção de drogas terapêuticas deverá ser substituída por seleção de pacientes nos quais uma determinada droga seria eficaz. Em nível clínico, enquanto os sintomas de uma determinada doença pareçam ser uniformes, variações individuais em vias poligênicas podem determinar que drogas adequadas para certos pacientes sejam tóxicas para outros. A abordagem farmacogenômica será importante tanto no desenvolvimento de novas drogas como também no resgate de drogas que, enquanto muito efetivas no tratamento de grupos de pacientes, claramente não beneficiam outros grupos que clinicamente não podiam ser diferenciados.

2. Terapia gênica

Terapia gênica se refere à transferência terapêutica de genes para células e tecidos in vivo. Suas potenciais aplicações clínicas são inúmeras e se baseiam na capacidade de alteração do "background" genético do paciente. Várias doenças genéticas humanas que resultam de uma lesão em um único gene têm sido propostas como candidatas à terapia gênica. Entretanto, o problema primordial na terapia gênica é o método de distribuição (transfecção) do gene, que até atualmente não tem sido efetivo(5). Os vetores mais comuns usados até o momento incluem adenovírus e retrovírus, mas o estágio de desenvolvimento dessa metodologia ainda não permite que essa técnica seja amplamente aplicada principalmente por causa da eficiência da transdução e da regulação da expressão gênica após a transdução.

3. Impacto dos avanços genéticos na clínica

Os avanços genéticos têm aumentado o nosso conhecimento sobre os mecanismos das doenças e o porquê de o curso clínico de algumas delas poder ser tão variável. No futuro, testes populacionais pré-sintomáticos para doenças de instalação tardia – como a doença de Alzheimer – podem generalizar-se e trazer benefícios importantes. A genotipagem pode se tornar parte das investigações de rotina para ajudar os clínicos a escolher o tratamento mais adequado a cada paciente. Entretanto, há alguns aspectos que precisam ser melhor entendidos. A relação entre seqüência de DNA e o fenótipo correspondente não é direta e clara. Em algumas condições, como diabetes e doenças cardiovasculares, onde fatores genéticos, sociais, biológicos e ambientais interagem, prever a importância clínica em um determinado paciente de diferentes mutações em genes de baixa penetrância pode ser muito difícil. Além disso, a determinação rigorosa dos custos e benefícios, tanto financeiros como psicossociais, dos testes genéticos é essencial, principalmente porque a informação gerada nos testes genéticos de triagem possui implicações sérias para os indivíduos e suas famílias(6).

4. Oncologia

Eventos genéticos críticos e específicos envolvendo oncogenes, genes supressores de tumor e enzimas de reparo do DNA têm sido caracterizados em muitos tumores. Essas alterações

moleculares, específicas das células tumorais, podem ser exploradas tanto para a detecção como para o tratamento do câncer. O conhecimento das características moleculares dos diferentes tipos de câncer tornará possível a identificação de certos pacientes que se beneficiariam de terapêuticas específicas para tais características. Atualmente, terapêuticas direcionadas para alterações moleculares específicas em vários tipos de tumor já estão sendo desenvolvidas e testadas. Por exemplo, células leucêmicas na maioria dos portadores de leucemia mielóide crônica (LMC) apresentam translocação entre os cromossomos 9 e 22, o que resulta na produção de uma enzima anormal (bcr-abl tirosina quinase) relacionada à etiologia da doença. STI571, uma droga inibidora específica da bcr-abl tirosina quinase foi desenvolvida como uma terapia específica que oferece uma opção de tratamento para pacientes com LMC(7). Em outro exemplo, o câncer de mama, em pacientes cujos tumores apresentam aumento de expressão do oncogene HER2, a terapia com Herceptina, anticorpo monoclonal anti-HER2, quando adicionada à quimioterapia em pacientes com câncer metastático aumenta a mediana da sobrevida em aproximadamente 25% em comparação com a quimioterapia isoladamente(8). Outros inibidores específicos do receptor HER2 se encontram em fase de teste. Nesse mesmo tipo de tumor, somente nas pacientes cujos tumores expressam receptores de estrogênio está indicado o tratamento com um modulador desse receptor, o Tamoxifen. Estudos recentes analisam a possibilidade de uma abordagem de classificação sistemática dos tumores baseada no

perfil de expressão gênica utilizando-se "micro-arrays". Os resultados preliminares desses trabalhos indicam que é possível a classificação dos tumores baseada no monitoramento da expressão de milhares de genes simultaneamente.

5. Contribuição brasileira para o Projeto Genoma Humano

Em 1997, a Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) decidiu que o Brasil não deveria perder as oportunidades científicas e econômicas que poderiam ser derivadas do seqüenciamento genômico, e que deveria ser capaz de produzir e analisar seus próprios dados, assim como usá-los para projetos científicos locais. Para o início, foi decidido que uma bactéria, preferencialmente de interesse para a economia local, deveria ter sua seqüência de DNA determinada. O microrganismo escolhido foi a *Xylella fastidiosa*, uma bactéria que infecta laranjeiras, fonte importante de renda para o estado. O principal objetivo desse primeiro projeto genoma foi levar a tecnologia a muitos laboratórios, que receberam seqüenciadores automáticos, reagentes e treinamento técnico. Trinta laboratórios foram selecionados para o projeto *Xylella* em todo o estado de São Paulo. O centro de bioinformática, localizado em Campinas, foi responsável pelo processamento dos dados gerados no projeto. O consórcio que surgiu do projeto *Xylella* foi denominado Rede ONSA (Organisation for Nucleotide Sequencing and Analysis)(9). Noventa por cento do genoma dessa bactéria foi seqüenciado em menos de um ano

(10). Esse mesmo consórcio está atualmente envolvido no seqüenciamento de outros organismos, como o *Xanthomonas axonopodis* pv. Citri, agente etiológico do cancro cítrico, cana-de-açúcar, entre outros.

Uma vez que o consórcio provou-se eficiente, outros projetos mais ambiciosos começaram a ser engendrados. Com o desenvolvimento de uma técnica de geração de fragmentos de cDNA preferencialmente da porção central dos genes, denominada Orestes (ORF expressed sequence tags)(11), no Instituto Ludwig de Pesquisa sobre o Câncer, um programa para descoberta gênica e compilação de seqüências codificadoras foi lançado em 1998 (Fapesp/LICR-HCGP) e iniciado em meados de 1999. Foi a primeira iniciativa da Rede ONSA na área altamente competitiva do seqüenciamento humano. As seqüências obtidas foram colocadas à disposição da comunidade científica através dos bancos de dados públicos. Enquanto o objetivo inicial era gerar 500.000 seqüências codificadoras derivadas de bibliotecas preparadas a partir de diferentes tecidos normal e tumoral, mais de um milhão foram geradas (1.111.949 seqüências), contribuição essa ao HGP somente superior a dos Estados Unidos. Com o reconhecimento de que o Fapesp/LICR-HCGP se transformou em um dos maiores contribuidores de seqüências dos bancos de dados públicos, a Fapesp e o Instituto Ludwig decidiram liderar e financiar uma iniciativa internacional para definir o transcriptoma humano, ou seja, a coleção de todos os genes expressos em uma célula, em parceria com o National Cancer Institute dos Estados Unidos, atual-

mente em progresso. Através da análise dos dados gerados pelo projeto genoma do câncer, a contribuição brasileira fica mais evidente. Duzentos e cinquenta mil seqüências geradas com a estratégia Orestes, que preferencialmente são geradas a partir da porção central das regiões codificadoras, foram analisadas. Dessas, 219 foram identificadas como sendo genes previamente não anotados que mapeavam no cromossomo 22. Apesar de, na época desse trabalho, as seqüências de Orestes representarem menos de 15% de todas as seqüências expressas depositadas nos bancos de dados públicos, com

elas conseguimos identificar 48 seqüências transcritas a partir do cromossomo 22 não definidas por outras seqüências(12).

Como conclusão, em pouco tempo o Brasil evoluiu da ausência absoluta de experiência em seqüenciamento genômico para se transformar em um dos países que mais produziu seqüências no mundo. A experiência acumulada com os projetos genoma no Brasil com certeza deverá ser extremamente importante para colocar o país também na pesquisa de ponta na era pós-genômica.

RESUME

Proyecto Genoma Humano y sus implicaciones para la salud humana: visión general y contribución brasileña para el proyecto

La información derivada del Proyecto Genoma Humano (Human Genome Project, HGP), un esfuerzo internacional para elucidar y caracterizar la secuencia completa de 3x10⁹ pares de bases del genoma humano, va a revolucionar la práctica de la medicina en el siglo 21 por proveer instrumentos para determinar el componente genético de virtualmente todas las enfermedades. Las consecuencias para la práctica de la medicina deberán ser profundas, tales como la generación de medidas que puedan prever el riesgo, permitir el diagnóstico precoz y promover estrategias de tratamiento más efectivas. En ese artículo, los autores discuten las aplicaciones potenciales del conocimiento adquirido con el HGP, bien como la contribución brasileña a ese proyecto.

 ABSTRACT

Human Genome Project and its implications to human health: general view and brazilian contribution to the project

The information derived from the Human Genome Project (HGP), an international endeavor to elucidate and characterize the complete sequence of 3x10⁹ base pairs of the human genome, will revolutionize the practice of medicine in the 21st century by providing tools to determine the genetic component of virtually all illnesses. The consequences of this for the practice of medicine should be profound, such as generating measures capable of predicting risk, enabling early diagnosis and promoting strategies for more effective treatment. In this article, the authors discuss the potential applications of the knowledge acquired by the HGP, as well as the Brazilian contribution to this project.

 BIBLIOGRAFIA

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* 2001;291:1304-51.
2. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860- 921.
3. Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, et al. A concise guide to cDNA microarray analysis. *Biotechniques* 2000 ;29:548-50,552-4,556.
4. Emilien G, Ponchon M, Caldas C, Isacson O, Maloteaux JM Impact of genomics on drug discovery and clinical medicine. *QJM* 2000 ;93:391-423.
5. Kay MA, Glorioso JC, Naldini L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med* 2001 ;7:33-40.
6. Lander ES. Scientific commentary: the scientific foundations and medical and social prospects of the Human Genome Project. *J Law Med Ethics* 1998;26:184-8, 178.
7. Vigneri P, Wang JY Induction of apoptosis in chronic myelogenous leukemia cells through nuclear entrapment of BCR-ABL tyrosine kinase. *Nat Med* 2001;7:228-34.
8. D, Gutheil JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, et al. First-line, single-agent Herceptin(R) (trastuzumab) in metastatic breast cancer: a preliminary report. *Eur J Cancer* 2001;37 (1 Suppl):25-9.
9. Simpson AJ, Perez JF . ONSA, the Sao Paulo

Virtual Genomics Institute. Organization for Nucleotide Sequencing and Analysis. *Nat Biotechnol* 1998 Sep;16(9):795-6.

10. Simpson AJ, Reinach FC, Arruda P, Abreu FA, Acencio M, Alvarenga R, et al. The genome sequence of the plant pathogen *Xylella fastidiosa*: the *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. *CL Nature* 2000;406:151-7.

11. Dias Neto E, Garcia Correa R, Verjovski-Almeida S, Briones MR, Nagai MA, da Silva W, et al. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc Natl Acad Sci USA* 2000;97:3491-6.

12. de Souza SJ, Camargo AA, Briones MR, Costa FF, Nagai MA, Verjovski-Almeida S, et al. Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags. *Proc Natl Acad Sci USA* 2000;97:12690-3.

ENDEREÇO PARA CORRESPONDÊNCIA

Andrew J.G. Simpson
Instituto Ludwig de Pesquisa sobre o Câncer
Rua Prof. Antônio Prudente 109 – 4º andar
CEP: 01509-010
São Paulo - SP - Brasil
E-mail: asimpson@nodel.com.br